

Dal catalogo McGraw-Hill:

Chiacchio, *Tecnologie informatiche per l'automazione*
Chiaverini, Caccavale, Villani, Sciavicco, *Fondamenti di sistemi dinamici*
Di Febbraro, Giua, *Sistemi ad eventi discreti*
Finzi, Pirovano, Volta, *Gestione della qualità dell'aria. Modelli di simulazione e previsione*
Magnani, *Tecnologia dei sistemi di controllo*
Sciavicco, Siciliano, *Robotica industriale*

**Paolo Bolzern
Riccardo Scattolini
Nicola Schiavoni**

Fondamenti di controlli automatici

Seconda edizione

Al lettore

La realizzazione di un libro comporta costi variabili (carta, stampa, legatura) e costi fissi, cioè indipendenti dal numero di copie stampate (traduzione, preparazione degli originali, redazione, composizione, impaginazione). I fotocopiatori possono contenere il prezzo perché, oltre a non pagare i diritti d'autore, non hanno costi fissi. Ogni fotocopia, d'altra parte, riducendo il numero di copie vendute dall'editore, aumenta l'incidenza dei costi fissi a copia e costringe l'editore ad aumentare il prezzo; questo, naturalmente, fornisce un ulteriore incentivo a fotocopiare. Se questo circolo vizioso non verrà spezzato, arriveremo al punto in cui gli editori non avranno più convenienza economica a realizzare libri di testo per l'università.

In quel momento non ci saranno più neppure fotocopie.

L'editore

McGraw-Hill

Milano • New York • San Francisco • Washington D.C. • Auckland
Bogotá • Lisboa • London • Madrid • Mexico City • Montreal
New Delhi • San Juan • Singapore • Sydney • Tokyo • Toronto

*a Paola (P. B.)
ad Anna, Elena e Carla (R. S.)
a Stefania, Mariateresa, Lorenzo e Stefano (N. S.)*

Copyright © 2004, 1998 The McGraw-Hill Companies, srl
Publishing Group Italia
via Ripamonti, 89 - 20139 Milano

McGraw-Hill
A Division of The McGraw-Hill Companies



I diritti di traduzione, di riproduzione, di memorizzazione elettronica e di adattamento totale e parziale con qualsiasi mezzo (compresi i microfilm e le copie fotostatiche) sono riservati per tutti i paesi.

Nomi e marchi citati nel testo sono generalmente depositati o registrati dalle rispettive case produttrici.

Editor: Paolo Roncoroni
Produzione: Donatella Giuliani
Realizzazione editoriale: CompoMat s.a.s., Configni (RI)
Grafica di copertina: G&G
Stampa: Arti Grafiche Battaia, Zibido San Giacomo (Mi)

ISBN 88-386-6099-9
Printed in Italy
3456789AGBLIL9875

Prefazione	XVII
------------	------

Problemi e sistemi di controllo	1
1.1 Introduzione	1
1.2 Problemi di controllo	2
1.2.1 Definizioni ed elementi costitutivi	2
1.2.2 Alcuni esempi	3
1.3 Sistemi di controllo	4
1.3.1 Definizioni fondamentali	4
1.3.2 Specifiche di progetto	5
1.3.3 Controllo in anello aperto e controllo in anello chiuso	6
1.3.4 Aspetti realizzativi	8
1.4 Ruolo della modellistica matematica	11
1.4.1 Riformulazione dei problemi di controllo	11
1.4.2 Problemi di sintesi	12
1.4.3 Problemi di analisi	13
1.5 Sistemi di controllo in anello chiuso	14
1.5.1 Confronto con i sistemi di controllo in anello aperto	14
1.5.2 Classi di controllori in anello chiuso	19
1.6 Controllo, supervisione e automazione	21
1.7 Conclusioni	22
Esercizi	23
Sistemi dinamici a tempo continuo	27
2.1 Introduzione	27
2.2 Concetti fondamentali	27
2.2.1 Variabili di ingresso, stato e uscita	27
2.2.2 Rappresentazione di stato	29
2.2.3 Esempi	30
2.2.4 Commenti sul concetto di stato	33
2.3 Classificazione	34
2.4 Ritardo di tempo e sistemi a parametri distribuiti	37
2.5 Equilibrio	38
2.6 Stabilità	41
2.6.1 Stabilità dell'equilibrio	41

2.6.2 Stabilità del movimento	45
2.7 Conclusioni	46
Esercizi	47
3. Sistemi lineari e stazionari a tempo continuo	
3.1 Introduzione	49
3.2 Movimento	49
3.2.1 Formula di Lagrange	50
3.2.2 Movimento libero e movimento forzato	50
3.2.3 Princípio di sovrapposizione degli effetti	51
3.2.4 Rappresentazioni equivalenti	52
3.2.5 Autovalori e modi	54
3.2.6 Risposta all'impulso e movimento forzato	58
3.3 Equilibrio	60
3.4 Stabilità	61
3.4.1 Stabilità del sistema	62
3.4.2 Stabilità e movimento libero	62
3.4.3 Stabilità e autovalori	63
3.4.4 Stabilità e polinomio caratteristico	64
3.4.5 Stabilità e parametri incerti	67
3.4.6 Proprietà dei sistemi asintoticamente stabili	69
3.5 Linearizzazione e stabilità dell'equilibrio di sistemi non lineari	70
3.5.1 Linearizzazione	70
3.5.2 Stabilità dell'equilibrio	72
3.6 Raggiungibilità, osservabilità e scomposizione canonica	73
3.6.1 Esempi introduttivi	73
3.6.2 Raggiungibilità	76
3.6.3 Osservabilità	79
3.6.4 Scomposizione canonica e forma minima	81
3.7 Conclusioni	85
Esercizi	85
4. Funzione di trasferimento	
4.1 Introduzione	87
4.2 Definizione della funzione di trasferimento	87
4.2.1 Definizione e interpretazioni	87
4.2.2 Struttura della funzione di trasferimento	88
4.2.3 Equazioni differenziali e funzione di trasferimento	90
4.2.4 Cancellazioni e stabilità	91
4.2.5 Cancellazioni, raggiungibilità e osservabilità	93
4.2.6 Ritardo di tempo	93
4.3 Rappresentazioni e parametri della funzione di trasferimento	94
4.3.1 Guadagno	95
4.3.2 Derivatore ideale	96
4.3.3 Integratore	96
4.3.4 Costanti di tempo	96
4.3.5 Pulsazione naturale e smorzamento	96
4.4 Risposta allo scalino	98
4.4.1 Valore iniziale e finale	98
4.4.2 Caratteristiche della risposta allo scalino	99
4.4.3 Sistemi del primo ordine	99
4.4.4 Sistemi del secondo ordine	100

4.4.5 Sistemi con ritardo di tempo	107
4.4.6 Sistemi di ordine superiore al secondo	107
4.5 Realizzazione	110
4.5.1 Forma canonica di raggiungibilità	110
4.5.2 Forma canonica di osservabilità	111
4.5.3 Relazioni tra diverse rappresentazioni dei sistemi lineari SISO	111
4.6 Conclusioni	112
Esercizi	112
5. Schemi a blocchi	
5.1 Introduzione	115
5.2 Componenti di uno schema a blocchi	115
5.3 Regole di elaborazione	118
5.3.1 Sistemi in serie	118
5.3.2 Sistemi in parallelo	118
5.3.3 Sistemi in retroazione	119
5.3.4 Riduzione di schemi a blocchi	120
5.3.5 Cancellazioni	121
5.4 Stabilità dei sistemi interconnessi	122
5.4.1 Stabilità di sistemi in serie	122
5.4.2 Stabilità di sistemi in parallelo	123
5.4.3 Stabilità di sistemi retroazionati	124
5.4.4 Stabilità e schemi equivalenti	126
5.5 Raggiungibilità e osservabilità dei sistemi interconnessi	126
5.6 Conclusioni	128
Esercizi	128
6. Risposta in frequenza	
6.1 Introduzione	131
6.2 Risposta alla sinusoide	131
6.2.1 Calcolo dell'uscita sinusoidale	132
6.2.2 Risposta in frequenza: definizione e proprietà	134
6.3 Risposta a segnali dotati di serie o trasformata di Fourier	135
6.3.1 Segnali sviluppabili in serie di Fourier	135
6.3.2 Segnali dotati di trasformata di Fourier	137
6.4 Complementi	139
6.4.1 Risposta esponenziale	139
6.4.2 Il caso di sistemi instabili	139
6.5 Identificazione sperimentale della risposta in frequenza	141
6.6 Diagrammi cartesiani o di Bode	142
6.6.1 Diagramma del modulo	142
6.6.2 Diagramma della fase	148
6.6.3 Ritardo di tempo	151
6.6.4 Sistemi a sfasamento minimo	152
6.7 Diagrammi polari	153
6.8 Azione filtrante dei sistemi dinamici	158
6.8.1 Filtri passa-basso	158
6.8.2 Filtri passa-alto	161

6.9	Approssimazione a poli dominanti	163
6.10	Conclusioni	164
	Esercizi	165
7/	Sistemi dinamici a tempo discreto	167
7.1	Introduzione	167
7.2	Generalità	168
7.2.1	Variabili di ingresso, stato e uscita, e rappresentazione di stato	168
7.2.2	Esempi	169
7.2.3	Classificazione	171
7.2.4	Equilibrio	172
7.3	Stabilità	173
7.3.1	Stabilità dell'equilibrio	173
7.3.2	Stabilità del movimento	174
7.4	Movimento ed equilibrio dei sistemi lineari e stazionari	175
7.4.1	Calcolo del movimento	175
7.4.2	Movimento libero e movimento forzato	175
7.4.3	Princípio di sovrapposizione degli effetti	176
7.4.4	Rappresentazioni equivalenti	177
7.4.5	Autovalori e modi	178
7.4.6	Risposta all'impulso e movimento forzato	181
7.4.7	Equilibrio	182
7.5	Stabilità dei sistemi lineari e stazionari	183
7.5.1	Stabilità del sistema	183
7.5.2	Stabilità e movimento libero	184
7.5.3	Stabilità e autovalori	184
7.5.4	Stabilità e polinomio caratteristico	185
7.5.5	Stabilità e parametri incerti	188
7.5.6	Proprietà dei sistemi asintoticamente stabili	189
7.6	Linearizzazione e stabilità dell'equilibrio di sistemi non lineari	190
7.6.1	Linearizzazione	190
7.6.2	Stabilità dell'equilibrio	192
7.7	Raggiungibilità, osservabilità e scomposizione canonica dei sistemi lineari e stazionari	193
7.7.1	Raggiungibilità	193
7.7.2	Osservabilità	195
7.7.3	Scomposizione canonica e forma minima	197
7.8	Conclusioni	199
	Esercizi	200
8	Analisi in frequenza dei sistemi a tempo discreto	203
8.1	Introduzione	203
8.2	Funzione di trasferimento	203
8.2.1	Definizione e interpretazioni	203
8.2.2	Struttura della funzione di trasferimento	204
8.2.3	Cancellazioni e stabilità	205
8.2.4	Ritardo di tempo	206
8.2.5	Rappresentazioni e parametri della funzione di trasferimento	207
8.3	Risposta allo scalino	209

8.3.1	Valore iniziale e finale	209
8.3.2	Andamento del transitorio	210
8.3.3	Sistemi del primo ordine	210
8.3.4	Sistemi del secondo ordine	211
8.3.5	Sistemi FIR	213
8.3.6	Poli dominanti	214
8.3.7	Modelli approssimanti FIR	215
8.4	Realizzazione	216
8.5	Schemi a blocchi	217
8.6	Risposta in frequenza	218
8.6.1	Calcolo dell'uscita sinusoidale	218
8.6.2	Risposta in frequenza: definizione e proprietà	220
8.6.3	Segnali dotati di sviluppo di Fourier	220
8.6.4	Segnali dotati di trasformata di Fourier	221
8.7	Complementi	222
8.7.1	Risposta esponenziale	222
8.7.2	Il caso di sistemi instabili	222
8.8	Diagrammi di Bode e polari	224
8.9	Conclusioni	227
	Esercizi	227
9	Sistemi di controllo a tempo continuo: stabilità	229
9.1	Introduzione	229
9.2	Controllo nell'intorno di un equilibrio	230
9.3	Schema generale di controllo in retroazione	233
9.4	Requisiti di un sistema di controllo	234
9.4.1	Stabilità	234
9.4.2	Prestazioni	235
9.5	Stabilità in condizioni nominali	238
9.5.1	Diagramma di Nyquist	238
9.5.2	Criterio di Nyquist	239
9.5.3	Estensioni del criterio di Nyquist	242
9.6	Stabilità in condizioni perturbate	245
9.6.1	Margine di stabilità vettoriale	246
9.6.2	Margine di guadagno	247
9.6.3	Margine di fase	248
9.6.4	Significatività del margine di guadagno e di fase	249
9.6.5	Criterio di Bode	252
9.7	Complementi sulla stabilità in condizioni perturbate	254
9.7.1	Criteri di stabilità robusta	254
9.7.2	Legami tra indicatori di robustezza	256
9.8	Conclusioni	258
	Esercizi	258
10	Sistemi di controllo a tempo continuo: prestazioni	261
10.1	Introduzione	261
10.2	Funzioni di sensitività e limiti alle prestazioni	262
10.3	Analisi della funzione di sensitività complementare	263
10.3.1	Analisi statica	264
10.3.2	Poli e zeri	265
10.3.3	Risposta in frequenza	265
10.3.4	Smorzamento e margine di fase	268

10.3.5 Valutazione esatta della banda passante	270	13.4.1 Osservatore banale	342
10.3.6 Risposta allo scalino	272	13.4.2 Osservatore asintotico con dinamica arbitraria	342
10.3.7 Effetto di un ritardo di tempo	273	13.5 Assegnamento degli autovalori con stato non misurabile	344
10.4 Analisi della funzione di sensitività	274	13.6 Interpretazione in termini di funzioni di trasferimento	346
10.4.1 Analisi statica	274	13.7 Conclusioni	349
10.4.2 Poli e zeri	276	Esercizi	349
10.4.3 Risposta in frequenza	276		
10.4.4 Altri limiti alle prestazioni	278		
10.5 Analisi della funzione di sensitività del controllo	279		
10.5.1 Analisi statica	280		
10.5.2 Poli e zeri	280		
10.5.3 Risposta in frequenza	281		
10.6 Prestazioni in condizioni perturbate	282		
10.6.1 Regolazione robusta a zero dell'errore	282	14.1 Introduzione	351
10.6.2 Reiezione robusta di disturbi sinusoidali	282	14.2 Modello dei regolatori PID	351
10.6.3 Attenuazione robusta di disturbi a banda limitata	283	14.3 Realizzazione dei regolatori PID	354
10.6.4 Sensitività rispetto a incertezze parametriche	284	14.3.1 Limitazione dell'azione derivativa	355
10.7 Conclusioni	286	14.3.2 Desaturazione dell'azione integrale	357
Esercizi	286	14.3.3 Inserimento "morbido" della regolazione automatica	362
11 Sintesi dei sistemi di controllo a tempo continuo	289	14.4 Metodi di taratura automatica	364
11.1 Introduzione	289	14.4.1 Metodi in anello chiuso	364
11.2 Requisiti e specifiche	290	14.4.2 Metodi in anello aperto	369
11.2.1 Requisiti principali	290	14.5 Conclusioni	376
11.2.2 Rappresentazione grafica dei vincoli	291	Esercizi	376
11.2.3 Altri requisiti	293		
11.3 Approcci alla sintesi	293		
11.4 Esempi di progetto	295	15 Schemi di controllo avanzati	379
11.5 Principali reti stabilizzatrici	308	15.1 Introduzione	379
11.5.1 Rete anticipatrice	309	15.2 Regolatori in anello aperto	379
11.5.2 Rete ritardatrice	310	15.2.1 Prefiltraggio del segnale di riferimento	380
11.5.3 Rete a sella	311	15.2.2 Compensazione del segnale di riferimento	383
11.6 Conclusioni	312	15.2.3 Schemi di controllo a due gradi di libertà	384
Esercizi	313	15.2.4 Compensazione dei disturbi misurabili	385
12 Luogo delle radici	315	15.3 Predittore di Smith	386
12.1 Introduzione	315	15.3.1 Schema a predittore di Smith	387
12.2 Definizione e proprietà	315	15.3.2 Approssimanti di Padé	388
12.2.1 Caratterizzazione del luogo	316	15.3.3 Impiego delle approssimanti di Padé nel predittore	
12.2.2 Regole di tracciamento	319	di Smith	389
12.3 Uso del luogo delle radici nell'analisi	323	15.4 Regolatori in cascata	390
12.4 Uso del luogo delle radici nella sintesi	327	15.5 Controllo di sistemi instabili	393
12.5 Conclusioni	333	15.6 Regolatori di disaccoppiamento	394
Esercizi	333	15.6.1 Disaccoppiamento di sistemi triangolari	395
13 Assegnamento degli autovalori	335	15.6.2 Disaccoppiamento di sistemi generici	395
13.1 Introduzione	335	15.6.3 Disaccoppiamento "in avanti"	396
13.2 Retroazione statica dall'uscita o dallo stato	335	15.6.4 Disaccoppiamento "all'indietro"	398
13.3 Assegnamento degli autovalori con stato misurabile	337	15.7 Controllo decentralizzato	399
13.3.1 Sistema in forma canonica	339	15.7.1 Matrice dei guadagni relativi	400
13.3.2 Sistema non in forma canonica	340	15.8 Conclusioni	405
13.4 Osservatore dello stato	341	Esercizi	405
16 Sistemi di controllo non lineari	407		
16.1 Introduzione	407		
16.2 Considerazioni preliminari	408		
16.2.1 Schemi a blocchi di sistemi non lineari	408		
16.2.2 Sistema canonico	409		
16.3 Stabilità assoluta	409		
16.3.1 Generalità	409		
16.3.2 Una condizione necessaria	411		

16.3.3 Una condizione sufficiente	411
16.3.4 Esempi di applicazione	412
16.4 Oscillazioni permanenti	416
16.4.1 Generalità	416
16.4.2 Funzione descrittiva	416
16.4.3 Metodo della funzione descrittiva	420
16.4.4 Stabilità delle oscillazioni	423
16.4.5 Esempi di applicazione	424
16.4.6 Taratura automatica di un PID	427
16.5 Conclusioni	429
Esercizi	429
17 Sistemi di controllo digitale: analisi e sintesi a tempo continuo	431
17.1 Introduzione	431
17.2 Schemi di controllo digitale	432
17.2.1 Campionatore	432
17.2.2 Mantenitore	434
17.2.3 Regolatore digitale	435
17.2.4 Temporizzazione	435
17.3 Campionamento	437
17.3.1 Trasformata di un segnale campionato	437
17.3.2 Aliasing	438
17.3.3 Teorema del campionamento	440
17.3.4 Filtri anti-aliasing	441
17.4 Mantenitore di ordine zero	441
17.5 Analisi a tempo continuo dei sistemi di controllo ibridi	443
17.6 Scelta del periodo di campionamento	446
17.7 Discretizzazione di un regolatore a tempo continuo	448
17.7.1 Metodo della trasformazione bilineare	449
17.7.2 Esempio di sintesi	453
17.8 Problemi realizzativi	455
17.8.1 Quantizzazione	455
17.8.2 Desaturazione dell'azione integrale	455
17.9 Conclusioni	457
Esercizi	458
18 Sistemi di controllo digitale: analisi e sintesi a tempo discreto	459
18.1 Introduzione	459
18.2 Sistema a segnali campionati	459
18.2.1 Calcolo della funzione di trasferimento	460
18.2.2 Presenza di ritardi di tempo	461
18.2.3 Autovalori del sistema a segnali campionati	463
18.2.4 Determinazione diretta della funzione di trasferimento	466
18.2.5 Zeri del sistema a segnali campionati	467
18.2.6 Guadagno e tipo del sistema a segnali campionati	469
18.2.7 Esempio riassuntivo	469
18.3 Altri metodi di discretizzazione del regolatore	470
18.3.1 Tenuta e campionamento	470
18.3.2 Trasformazione diretta di poli e zeri	471
18.4 Analisi di sistemi retroazionati a tempo discreto	472
18.4.1 Stabilità	473
18.4.2 Criterio di Nyquist	473

18.4.3 Luogo delle radici	475
18.4.4 Funzioni di sensitività	477
18.4.5 Analisi statica	478
18.4.6 Analisi dinamica	478
18.5 Sintesi diretta a tempo discreto	479
18.5.1 Requisiti e specifiche	480
18.5.2 Sintesi mediante il luogo delle radici	482
18.5.3 Assegnamento del modello	483
18.5.4 Assegnamento dei poli	487
18.6 Conclusioni	492
Esercizi	493
A Matrici	495
A.1 Introduzione	495
A.2 Definizioni e operazioni fondamentali	495
A.2.1 Generalità	495
A.2.2 Operazioni su una matrice	496
A.2.3 Operazioni tra matrici	498
A.3 Autovalori e autovettori	499
A.3.1 Generalità	499
A.3.2 Forma diagonale e forma di Jordan	500
A.4 Potenza	501
A.5 Esponenziale	503
A.6 Limiti, derivate e integrali	505
B Segnali a tempo continuo	507
B.1 Introduzione	507
B.2 Impulso e altri segnali canonici	507
B.3 Trasformata di Laplace	509
B.3.1 Generalità	509
B.3.2 Proprietà principali	511
B.3.3 Sviluppo di Heaviside e antitrasformazione di trasformate razionali	515
B.4 Serie di Fourier	520
B.4.1 Forma esponenziale	520
B.4.2 Forma trigonometrica	521
B.4.3 Proprietà principali	525
B.5 Trasformata di Fourier	525
B.5.1 Forma esponenziale	525
B.5.2 Forma trigonometrica	526
B.5.3 Proprietà principali	528
B.5.4 Relazioni con la trasformata di Laplace	530
C Segnali a tempo discreto	533
C.1 Introduzione	533
C.2 Impulso e altri segnali canonici discreti	533
C.3 Trasformata Zeta	534
C.3.1 Generalità	534
C.3.2 Proprietà principali	536
C.3.3 Sviluppo di Heaviside, lunga divisione e antitrasformazione di trasformate razionali	538
C.4 Sviluppo di Fourier discreto	541

C.4.1 Forma esponenziale	541
C.4.2 Forma trigonometrica	542
C.5 Trasformata di Fourier discreta	543
C.5.1 Forma esponenziale	543
C.5.2 Forma trigonometrica	544
C.5.3 Relazioni con la trasformata Zeta	546
Riferimenti bibliografici	547
Indice analitico	553

Prefazione

Questo libro presenta gli elementi di base della teoria dei sistemi e del controllo. La trattazione inizia con una discussione delle problematiche fondamentali del controllo automatico, tendente anche a motivare l'importanza dei modelli matematici. Quindi, si introducono i principali concetti e risultati sui sistemi dinamici a tempo continuo e discreto, con particolare riguardo a quelli lineari e invarianti nel tempo, per i quali si utilizzano sia rappresentazioni nel dominio del tempo sia rappresentazioni nel dominio della frequenza. Successivamente, si illustrano i metodi per il controllo dei sistemi monovariabili a tempo continuo e, dopo avere toccato alcuni aspetti del controllo multivariabile e di quello non lineare, il volume si chiude considerando il controllo digitale. Malgrado tale sommario faccia apparire del tutto tradizionale la struttura del lavoro, noi riteniamo che esso si caratterizzi per diversi aspetti.

- La dinamica dei sistemi occupa uno spazio piuttosto ampio, vista la sua importanza intrinseca anche in applicazioni diverse dal controllo.
- La teoria del controllo classica è presentata seguendo una linea che risente dei recenti approcci allo studio della robustezza.
- Sono trattati in dettaglio i controllori industriali e i corrispondenti metodi di taratura automatica.
- Si integra la sintesi ad assegnamento degli autovalori con i classici metodi di progetto nel dominio delle frequenze.
- Dopo aver illustrato la teoria per il singolo anello di regolazione, si prendono in considerazione anche sistemi multivariabili e schemi più articolati, con elementi di controllo in anello aperto o ad anelli annidati.
- Sono introdotte alcune nozioni sul controllo non lineare, che permettono di chiarire gli effetti provocati dalla presenza di saturazioni o di controllori a relè.
- Il controllo digitale è descritto nei suoi aspetti teorici e applicativi, così da metterne in luce potenzialità e difficoltà.

Fin dalla prima edizione, risalente al 1998, abbiamo cercato di integrare argomenti metodologici e tecnici, sforzandoci di contenere gli sviluppi matematici e rinunciando talvolta al rigore formale per privilegiare le valenze applicative dei concetti e degli algoritmi forniti dalla teoria, anche mediante un gran numero di esemplificazioni. In particolare, i sistemi dinamici sono presentati in ambito puramente algebrico, e non geometrico, senza che ciò impedisca di arrivare a introdurre i primi risultati sulle proprietà strutturali. Come prerequisito, al lettore è richiesta

solo la conoscenza delle operazioni con i numeri complessi e di qualche nozione sulle equazioni differenziali, mentre quanto occorre di algebra delle matrici e di rappresentazione di segnali nel dominio delle frequenze è riportato nelle appendici. Sinteticamente, abbiamo voluto misurarsi con la sfida di presentare le basi dell'automatica in maniera semplice, ma aperta, cioè tale da consentire lo studio successivo di teorie e tecniche più avanzate.

In questo senso, abbiamo sempre ritenuto che il libro potesse essere di interesse per coloro che già operavano professionalmente nel settore del controllo e desiderassero allargare le loro conoscenze, magari di tipo prevalentemente empirico. È chiaro però che, primariamente, abbiamo pensato a questo libro quale manuale di riferimento per i corsi universitari di base nell'ambito dell'automatica, cercando di riversare in esso tutta la nostra esperienza pluriennale nella didattica relativa a questi temi.

Negli anni, il nostro lavoro ha ricevuto significativi apprezzamenti da parte di colleghi, che hanno utilizzato il testo a supporto dei corsi tenuti, uniti a preziosi suggerimenti. Perciò, nello stendere la nuova edizione, abbiamo ritenuto di conservare l'impostazione originaria, ma anche di effettuare un profondo lavoro di revisione, che ci ha suggerito a volte di modificare la presentazione degli argomenti e la loro distribuzione nei capitoli, a volte di eliminare alcune parti, a volte di introdurre argomenti che prima non erano trattati. Tra questi ultimi, ci pare importante segnalare la tecnica di sintesi del controllore ad assegnamento degli autovalori.

Abbiamo costantemente perseguito l'obiettivo di rendere più flessibile l'utilizzazione del testo. Sappiamo che le esigenze dei docenti possono essere molto differenziate, in dipendenza, per esempio, dal numero di crediti formativi dei loro corsi e dalla tipologia degli allievi, il cui percorso di studio può prevedere una specializzazione nel settore dell'automatica oppure no. Per esemplificare, suggeriamo allora due utilizzazioni tipiche, rese possibili dalla struttura logica con cui sono connessi i 18 capitoli e le 3 appendici (Figura P.1). In ogni caso queste ultime (che riguardano le matrici, i segnali a tempo continuo e quelli a tempo discreto) potranno essere trattate in misura dipendente dalla cultura preecessa degli allievi e dall'impostazione, più formale o più pragmatica, che il docente intende dare al suo insegnamento.

Per allievi che non abbiano alcuna conoscenza della teoria dei sistemi, dopo il Capitolo 1, dedicato a introdurre i problemi e i sistemi di controllo, i Capitoli 2-6 sui sistemi dinamici a tempo continuo saranno indispensabili, anche se non necessariamente al completo. Successivamente, si passerà ai Capitoli 9-12 sui metodi classici di analisi e sintesi dei sistemi retroazionati monovariabili e al Capitolo 14, sui regolatori PID, da svolgere almeno in parte. A scelta del docente, potranno poi essere trattati argomenti selezionati sui sistemi a tempo discreto (Capitoli 7 e 8) e sul controllo digitale (Capitoli 17 e 18), sull'assegnamento degli autovalori (Capitolo 13), sulle strutture avanzate di controllo (Capitolo 15) o sui sistemi di controllo non lineari (Capitolo 16). Invece, per allievi cui la teoria dei sistemi sia stata presentata in precedenti corsi, dopo il Capitolo 1, i Capitoli 2-8 saranno svolti in modo rapido concentrando l'attenzione sui temi più rilevanti per il controllo e i Capitoli 9-18 potranno essere trattati in misura estesa.

Un'altra novità di rilievo della nuova edizione è costituita dall'introduzione di un cospicuo numero di esercizi al termine di ogni capitolo. Essi non sono solo tesi alla verifica della comprensione degli argomenti trattati e della capacità acquisita di applicare i metodi presentati, ma intendono spesso stimolare il lettore alla riflessione sulla materia e proporgli qualche estensione di quanto esposto.

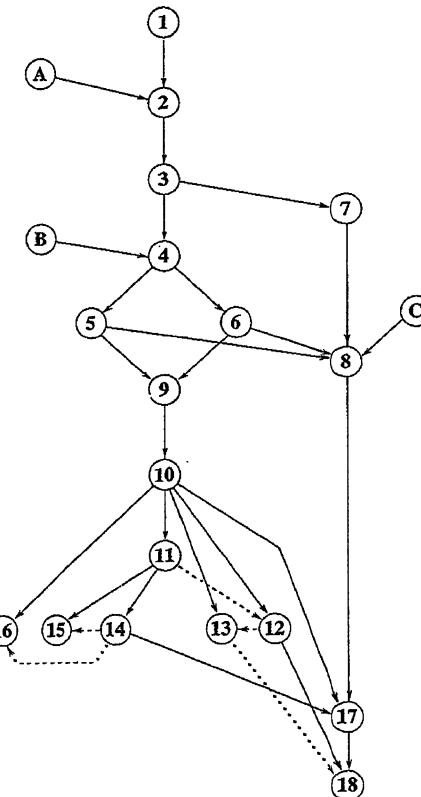


Figura P.1
Schema logico del libro.
Numeri e lettere fanno riferimento rispettivamente a Capitoli e Appendici. Le frecce tratteggiate rappresentano connessioni più deboli di quelle indicate dalle frecce continue.

Il libro si conclude con un'aggiornata sezione di riferimenti bibliografici, nella quale suggeriamo alcune letture utili per l'approfondimento degli argomenti presentati e per un'introduzione a tematiche più avanzate.

Nel testo non si fa riferimento esplicito allo studio dei sistemi dinamici e di controllo con l'ausilio di strumenti software per la simulazione e il "computer aided design", anche se essi rivestono ovviamente una grande importanza quando si affrontano problemi pratici. Pensiamo, infatti, che, a causa della rapida obsolescenza cui sono soggetti i prodotti hardware e software, qualsiasi indicazione specifica diventerebbe in breve tempo inadeguata. Diamo tuttavia per scontato che negli insegnamenti universitari per cui questo testo sarà adottato gli strumenti più attuali vengano presentati e ampiamente utilizzati, magari nelle esercitazioni.

In fine, desideriamo ringraziare colleghi e studenti che, con le loro osservazioni sulla precedente edizione, ci hanno aiutato a revisionare il testo, nonché la casa editrice per il supporto fornito e, in particolare, per la nuova veste grafica che ha voluto dare al libro, augurandoci che essa contribuisca a migliorarne la leggibilità.

Milano, aprile 2004

Paolo Bolzern, Riccardo Scattolini, Nicola Schiavoni

1.1 Introduzione

Nella conduzione di un gran numero di apparati della più svariata natura, si riconosce la necessità di intervenire dall'esterno in modo che essi si comportino nel modo desiderato. Per esempio, affinché un autoveicolo si muova su un dato percorso seguendo un prefissato profilo di velocità, è indispensabile agire in maniera continua sugli organi che determinano la direzione del movimento (sterzo) e la coppia applicata alle ruote motrici (acceleratore, cambio e freni).

Quindi, questo tipo di problematica si incontra frequentemente già nella vita di ogni giorno; di maggiore interesse però è la sua larghissima diffusione in tutta l'ingegneria. Con riferimento al settore civile, per climatizzare un edificio bisogna scegliere le caratteristiche dell'aria entrante nei singoli locali (portata, temperatura, umidità) in modo che le condizioni ambientali siano quelle volute. Invece, nel settore industriale, si pensi al caso di una caldaia destinata alla generazione di vapore per il quale siano di interesse la portata e la pressione, che dipendono dalla portata d'acqua addotta e dalla potenza termica ceduta dalle fiamme dei bruciatori; inoltre è necessario che il livello dell'acqua nella cavità ove essa viene trasformata in vapore sia tenuto costante per evitare pericolosi fenomeni di svuotamento.

Eigenze dello stesso genere si possono riconoscere anche in ambiti differenti da quelli dell'ingegneria tradizionale, come nella gestione di complessi fenomeni di carattere economico, sociologico o ambientale. È noto a tutti, per esempio, che le autorità di governo di un paese che si propongano di ridurre il deficit finanziario possono perseguire tale scopo agendo su leve di vario tipo, quali l'inasprimento della tassazione e il contenimento delle spese correnti.

Per determinare le azioni da compiere in situazioni come quelle delineate, si è condotti ad affrontare i cosiddetti problemi di controllo, che trovano soluzione nei sistemi di controllo. Questo capitolo ha l'obiettivo di introdurre allo studio di tali problemi e sistemi, e in particolare presenta:

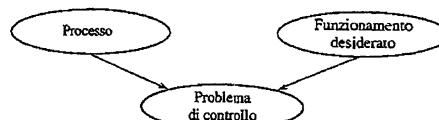
- gli elementi costitutivi dei problemi di controllo;
- le caratteristiche salienti dei sistemi di controllo, con un cenno agli aspetti realizzativi;
- le speciali potenzialità dei sistemi di controllo retroazionati;
- il ruolo svolto dalla modellistica matematica nell'ambito dell'analisi e del progetto.

1.2 Problemi di controllo

1.2.1 Definizioni ed elementi costitutivi

Processo e variabili principali I *problemi di controllo* consistono nell'imporre un *funzionamento desiderato* a un *processo assegnato* (Figura 1.1). Con il termine *processo*, o *sistema sotto controllo*, si fa quindi riferimento all'oggetto sul quale il problema è posto: un impianto, un'apparecchiatura, una macchina o anche un fenomeno di natura fisica o di altro genere. Invece il *funzionamento desiderato* è

Figura 1.1
Elementi di un problema di controllo.



espresso dalla richiesta che l'andamento nel tempo di alcune variabili del processo coincida con quello di altre variabili preassegnate. Le prime, quindi, rappresentano le grandezze di interesse e, nell'insieme, costituiscono la *variabile controllata*; le seconde, in forma vettoriale, costituiscono invece l'*andamento desiderato della variabile controllata*, detto altimamente *segnaile di riferimento*. Quest'ultimo, qualora sia costante nel tempo, con terminologia inglese si dice anche *set-point*. Pertanto, l'obiettivo ideale di un problema di controllo si può sintetizzare nella formula

$$\text{variabile controllata} = \text{segnaile di riferimento} \quad (1.1)$$

per tutto l'intervallo di tempo in cui il funzionamento del processo è di interesse.

Per perseguire l'obiettivo (1.1) occorre evidentemente avere la possibilità di condizionare la variabile controllata. A tale scopo, si suppone di poter agire sul processo manipolando altre variabili scalari, che nell'insieme formano la *variabile di controllo*, i cui andamenti temporali si possano scegliere a piacimento, eventualmente nel rispetto di qualche vincolo. Queste variabili sono quindi assegnabili, in maniera relativamente arbitraria, da chi effettua il controllo.

A volte, quando il segnale di riferimento è costante, si usa dire che ci si trova di fronte a un *problema di regolazione*. Per semplicità, in questa trattazione non si farà distinzione tra i termini "controllo" e "regolazione" che saranno considerati sinonimi.

Incertezza In realtà le *variabili controllate* non dipendono solamente dalle *variabili di controllo*: esistono altre grandezze che non sono manipolabili e tuttavia hanno un'influenza sul comportamento del processo. Tipicamente esse non sono note con precisione e, per loro tramite, si può descrivere l'*incertezza* che è sempre presente in un processo da controllare. Per cominciare, quasi mai si può conoscere perfettamente la situazione in cui si trova il processo nell'istante in cui si inizia a esercitare l'azione di controllo. Inoltre, alcuni *parametri* interni al processo possono avere un valore differente dal *valore nominale*, o previsto, e almeno parzialmente ignoto. Analogamente, si può verificare la presenza di variabili che influiscono sul processo dall'esterno come quelle di controllo, ma che, a differenza di queste ultime, non si possono scegliere liberamente. L'esistenza di queste variabili, raccolte nel vettore chiamato *disturbo*, complica in maniera significativa i problemi di controllo, soprattutto per il fatto che esse usualmente differiscono, in modo ignoto, rispetto ai loro *andamenti nominali*.

Quando tutti i parametri e i disturbi assumono i loro valori e andamenti nominali, si dice che ci si trova in *condizioni nominali*; altrimenti si parla di *condizioni perturbate*.

Il tempo Tutte le variabili che intervengono in un problema di controllo sono funzioni del tempo. Esso può essere di tipo *continuo*, cioè descritto da una variabile reale che si indica con la lettera t , oppure *discreto*, cioè descritto da una variabile intera, che si denota con la lettera k .

1.2.2 Alcuni esempi

Per illustrare i concetti introdotti, è opportuno presentare un certo numero di esempi concreti.

Esempio 1.1 Si supponga che l'autoveicolo cui si è accennato all'inizio del capitolo debba percorrere una strada pianeggiante, lungo una traiettoria e con una velocità assegnate. Allora le variabili controllate sono quattro: due individuano la posizione in un piano orizzontale e due la velocità. A esse corrispondono, come segnali di riferimento, quattro funzioni che specificano la posizione e la velocità desiderate in ogni istante di tempo. Invece, le variabili di controllo sono quelle che individuano la posizione del volante, quella dei pedali dell'acceleratore e del freno, e quella del cambio.

È evidente che le variabili di controllo, sottoposte tra l'altro a ovvie limitazioni, non determinano completamente il moto del veicolo, che dipende anche, per esempio, dalla posizione e dalla velocità iniziali. Tra i parametri incerti che, a parità di azione prodotta dalle variabili di controllo, modificano l'andamento delle variabili controllate, si possono annoverare la massa del veicolo, l'efficienza del motore e dell'apparato frenante, lo stato degli pneumatici e dell'asfalto. Infine, un disturbo è certamente costituito dalla presenza di vento, la cui velocità è incerta in modulo, direzione e verso.

Esempio 1.2 L'impianto centralizzato di climatizzazione di un edificio, già citato nel paragrafo precedente, deve essere realizzato in modo che le temperature nei singoli ambienti (variabili controllate) siano pari a quelle desiderate (segnali di riferimento). A questo scopo si utilizzano come variabili di controllo le portate d'aria inviate ai locali, modificabili agendo su saracinesche. Evidentemente, le portate non potranno essere né negative, né superiori a un valore massimo noto.

Le temperature interne, però, dipendono anche da molti altri fattori, tra i quali la temperatura esterna e l'insolazione (disturbi), nonché i coefficienti di scambio termico (parametri). Inoltre, la posizione delle saracinesche non determina completamente la potenza termica addotta. Infatti, la portata e la temperatura dell'aria condizionante dipendono dal funzionamento delle pompe e degli scambiatori di calore a monile e, ragionevolmente, non si possono ritenere perfettamente note: in altri termini, costituiscono ulteriori fonti di incertezza.

Esempio 1.3 Si assuma che nella gestione di un impianto industriale vi sia interesse affinché il livello dell'acqua in un serbatoio (variabile controllata) sia pari a un certo valore desiderato (segnaile di riferimento). Ciò è richiesto, per esempio, per la cavità di un generatore di vapore nella quale risiedono l'acqua e il vapore stesso. La pompa che alimenta il serbatoio è azionata elettricamente e la polenza impiegata, certamente soggetta a qualche limitazione, si può assumere come variabile di controllo. Tuttavia il rendimento della pompa, che determina la portata afferente al serbatoio, non può essere perfettamente noto; esso è invece un parametro incerto. Ovviamente, poi, è l'utenza a decidere quale sia la portata uscente, che quindi gioca il ruolo di disturbo.

Nei tre casi precedenti le variabili erano funzioni del tempo continuo; l'esempio seguente mostra un caso in cui invece è più naturale assumere che le variabili in gioco dipendano da un tempo discreto.

Esempio 1.4 Un'azienda manifatturiera deve decidere settimanalmente (tempo discreto) la tipologia e la quantità dei semilavorati da acquistare (variabili di controllo) in modo che i livelli delle scorte presenti in magazzino (variabili controllate) abbiano nel tempo degli andamenti (segnali di riferimento) scelti in modo che possano essere effettuate le lavorazioni previste senza che vi siano eccessivi immobilizzati di capitale. Naturalmente ciò viene fatto sulla base di ipotesi di utilizzo (valori nominali di variabili incerte) che si possono rivelare errate perché, per esempio, l'azienda potrebbe ricevere ordini imprevisti da evadere con urgenza oppure, a causa di scioperi, il lavoro potrebbe essere temporaneamente sospeso.

Questi esempi, nella loro semplicità, dimostrano quanto diversificati siano i problemi di controllo che si possono incontrare. In ambito industriale, innumerevoli sono i casi in cui siamo di interesse variabili quali temperature e livelli, come accade per i reattori chimici nei quali si svolgono reazioni esotermiche o endotermiche, i fornì e tutte le apparecchiature che prevedono la circolazione di fluidi. Spesso sono controllate anche variabili quali portate, pressioni, concentrazioni e umidità. Invece, negli apparati elettronici (convertitori, amplificatori, generatori di segnali, modulatori e demodulatori), ci si trova di fronte a obiettivi posti su correnti e tensioni. Frequentissime, poi, sono le situazioni in cui le variabili controllate siano di tipo meccanico, come posizione e velocità. Basta pensare a processi costituiti da veicoli di ogni tipo (terrestri, marini, aerei e spaziali), ad antenne che debbano seguire satelliti o a telescopi utilizzati per osservare astri in movimento relativo rispetto alla Terra e, più in generale, a motori di varia natura per i quali le posizioni e le velocità debbano essere tenute sotto controllo. In questo contesto, una menzione particolare meritano le applicazioni robotiche.

Esempio 1.5 Un robot costituito da diversi bracci (*link*) connessi tra loro deve essere utilizzato per lucidare la superficie di un pezzo meccanico. Allora, è necessario che l'estremità dell'ultimo braccio (*end effector*), su cui è montato l'utensile, effettui un percorso prefissato esercitando una forza opportuna sul componente da lucidare. Le variabili controllate sono dunque quelle che individuano la posizione dell'utensile e la forza da esso applicata. Le variabili di controllo sono invece le grandezze elettriche (tensioni o correnti) relative ai motori che muovono i bracci. Tra le variabili incerte si possono annoverare le resistenze e le induttanze dei motori, le coppie d'attrito, le elasticità dei giunti e la rugosità della superficie da lucidare.

Per chiudere, conviene notare che anche le dimensioni fisiche dei sistemi sotto controllo possono essere molto varie. Si pensi, da una parte, al problema di tenere costante la velocità angolare con cui un lettore fa ruotare un compact disc durante l'ascolto e, dall'altra, a quello di controllare l'ampiezza e la frequenza della tensione sinusoidale in tutti i punti della rete elettrica di un'intera nazione.

1.3 Sistemi di controllo

1.3.1 Definizioni fondamentali

La determinazione dell'andamento della variabile di controllo viene compiuta da un organo detto *controllore*, o *regolatore*. Il complesso costituito dal processo e dal controllore è denominato *sistema di controllo* (Figura 1.2). Può accadere che il processo e il controllore siano strettamente connessi tra loro, cioè che il processo sia dotato di un meccanismo di autoregolazione così intrinsecamente legato al processo stesso da rendere praticamente indistinguibili i due elementi del sistema di controllo. Si parla allora di controllori e sistemi di controllo *naturali*, mentre in caso contrario si dice che essi sono *artificiali*. Questi ultimi si possono classificare in *manuali* e *automatici*, a seconda che l'azione di controllo sia esercitata dall'uomo o da un dispositivo progettato appositamente.

Figura 1.2
Elementi di un sistema di controllo.



Esempio 1.6 All'interno del corpo umano esistono molti sistemi di controllo naturali; tra questi si possono citare quelli che provvedono a "risolvere" i seguenti problemi di controllo:

- tenere costante la pressione arteriosa;
- tenere costante la concentrazione di glucosio nel sangue;
- fare in modo che il ritmo del battito cardiaco sia il più adeguato alle condizioni di sforzo cui l'organismo è sottoposto;
- tenere costante la temperatura corporea;
- consentire movimenti complessi delle varie parti del corpo senza perdita dell'equilibrio da parte dell'individuo.

A seguito di malattie, i controllori naturali possono produrre dei "malfunzionamenti". In questo caso essi possono essere sostituiti da opportune terapie, che svolgono il ruolo di controllori artificiali. Per esempio, tramite l'assunzione di farmaci è possibile ridurre la pressione arteriosa quando tende ad assumere valori troppo elevati.

Esempio 1.7 (Seguito dell'Esempio 1.1) Un autoveicolo è normalmente condotto da un pilota umano, che ne costituisce un controllore manuale. Tuttavia, si trovano ormai in commercio dispositivi che consentono il controllo automatico della velocità (*cruise control*) senza alcun intervento umano, se non quello necessario per impostare la velocità desiderata. Inoltre, sono in fase avanzata di studio i sistemi per il controllo automatico della traiettoria, da utilizzare su percorsi autostradali attrezzati.

Si osservi che sistemi di controllo automatico del moto sono già da tempo diffusissimi sugli aeromobili (piloti automatici). Essi agiscono sulla potenza erogata dai motori e sulla posizione delle superfici mobili (timone e alettoni) per seguire una determinata rotta con una velocità prefissata.

La trattazione che segue è dedicata ai sistemi di controllo automatici o, come si dice comunemente, ai *controlli automatici*. Quanto verrà detto con riferimento ai controllori artificiali è largamente utilizzabile anche per lo studio dei sistemi di controllo naturali.

1.3.2 Specifiche di progetto

L'obiettivo di progetto espresso dalla relazione (1.1) prevede la perfetta identità tra la variabile controllata e il segnale di riferimento. Esso è già stato indicato come ideale in quanto per molte ragioni è, di fatto, irraggiungibile: per rendersi conto di ciò si faccia riferimento all'esperienza comune relativa agli esempi del paragrafo precedente.

Nella pratica applicativa, fortunatamente, si può ritenere che un problema di controllo sia stato risolto convenientemente anche se la (1.1) è soddisfatta solo in via approssimata. Naturalmente le approssimazioni accettabili devono essere precise caso per caso. Però, in termini generali, si può dire che la relazione

$$\text{variabile controllata} \simeq \text{segnale di riferimento}$$

viene tradotta nell'imporre che l'*errore* del sistema di controllo definito come

$$\text{errore} = \text{segnale di riferimento} - \text{variabile controllata} \quad (1.2)$$

soddisfi un insieme di *requisiti*, o *specifiche*, che esprimono la necessità che esso risulti "accettabilmente piccolo" in tutte le condizioni di funzionamento di interesse. Queste, a loro volta, corrispondono ai valori che possono assumere i parametri del processo e agli andamenti prevedibili dei disturbi, oltre che alla situazione in cui si trova il processo all'inizio dell'intervallo di tempo di controllo.

Per quanto riguarda i parametri e i disturbi, è molto utile conoscere i rispettivi valori e andamenti nominali, nonché le classi cui appartengono i valori e gli andamenti veri. Inoltre, di solito un unico controllore deve essere in grado di imporre

al processo comportamenti di volta in volta differenti, cioè, nel momento in cui si progetta il controllore, il segnale di riferimento non è del tutto noto, ma di esso si conosce soltanto una classe funzionale di appartenenza.

Esempio 1.8 (Seguito dell'Esempio 1.2) Pensando al periodo estivo, si può assumere che le temperature desiderate nei diversi ambienti siano funzioni costanti a tratti i cui valori non scendano mai al di sotto di 24°C e non salgano mai oltre 28°C . Per quanto riguarda i disturbi, la temperatura esterna potrebbe essere descritta anch'essa con riferimento alle sue caratteristiche temporali, considerandola una funzione continua con valori appartenenti un intervallo noto; il suo andamento nominale potrebbe essere costituito da una sinusoida di ampiezza nota e periodo di 24 ore. Informazioni analoghe potrebbero essere disponibili a proposito della potenza termica trasmessa dal sole per irraggiamento. Per i coefficienti di scambio termico delle pareti, la conoscenza delle caratteristiche dei materiali potrebbe consentire almeno di considerare noti i rispettivi intervalli di appartenenza.

Chiaramente, non si può pensare che, durante il funzionamento, la temperatura controllata coincida sempre con quella desiderata: di solito può bastare che l'errore sia piccolo. Per esempio, ci si può accontentare del fatto che l'errore sia nullo, se la temperatura desiderata è costante, e che, di fronte a sue brusche variazioni, esso ritorni a zero in tempi brevi, indipendentemente dagli andamenti dei disturbi e dai valori dei parametri.

All'obiettivo di rendere "piccolo" l'errore si aggiunge di solito una richiesta di *moderazione del controllo*, che è motivata sia dalla tipica presenza di vincoli sul valore massimo, o minimo, che può assumere la variabile di controllo sia dall'ovvia necessità di evitare che il processo riceva sollecitazioni eccessive. Pertanto, anche per le variabili di controllo potranno essere definite opportune specifiche. È chiaro che l'ottenimento di un buon comportamento dell'errore è in contrasto, fortunatamente solo parziale, con l'obiettivo di evitare una sollecitazione eccessiva del controllo. Si rivedano, a questo proposito, gli esempi trattati in precedenza.

1.3.3 Controllo in anello aperto e controllo in anello chiuso

Le informazioni in possesso del progettista del controllore sulle variabili in gioco, che sono state oggetto del paragrafo precedente, non devono essere confuse con le informazioni di cui dispone il regolatore durante il funzionamento.

Ogni controllore, per poter agire sul processo in maniera opportuna, deve necessariamente avere delle informazioni sul segnale di riferimento. Molto spesso l'azione di controllo dipende dal passato e dal presente dell'andamento desiderato della variabile controllata, ma non dal futuro; cioè, all'istante $t = \bar{t}$ è noto al regolatore l'andamento effettivo del segnale di riferimento per $t \leq \bar{t}$, se il tempo è continuo. Se invece il tempo è discreto, all'istante $k = \bar{k}$ il controllore conosce l'andamento effettivo del segnale di riferimento per $k \leq \bar{k}$. È inoltre possibile, anche se non indispensabile, che il controllore conosca passato e presente del disturbo, o di qualche sua componente. Quando il controllore possiede informazioni solo sul segnale di riferimento ed eventualmente sul disturbo, esso si dice in *anello aperto*, o ad azione diretta (in inglese *feedforward*), e così si denomina anche il sistema di controllo corrispondente. Questa struttura di controllo è illustrata dallo schema di Figura 1.3.

Figura 1.3
Sistema di controllo in anello aperto.

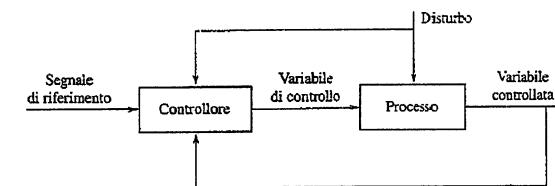
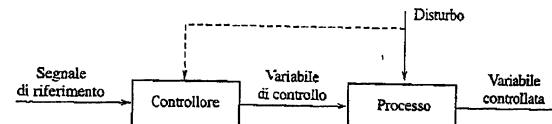


Figura 1.4
Sistema di controllo in anello chiuso.

In aggiunta a quanto sopra, accade spesso che la variabile controllata sia misurabile e disponibile al regolatore. Allora l'azione di controllo impressa al processo in $t = \bar{t}$ (o $k = \bar{k}$) dipende anche dall'andamento della variabile controllata per $t \leq \bar{t}$ (o $k \leq \bar{k}$) e il controllore si dice in *anello chiuso*, o in *retroazione* (in inglese *feedback*); analogamente si parla di sistema di controllo in *anello chiuso*, o *retroazionato*. Un tale sistema è rappresentato nella Figura 1.4. Se il disturbo è misurabile e la variabile di controllo, in anello aperto o chiuso, ne dipende, si usa dire che il controllore effettua una *compensazione del disturbo*. La compensazione costituisce comunque un'azione in anello aperto: le dizioni "anello chiuso" e "retroazione" vengono infatti riservate per indicare la dipendenza della variabile di controllo da quella controllata, che evidentemente dipende dalla prima. Questi termini sono anche usati con riferimento al caso in cui la variabile generata dal controllore sia fatta dipendere, invece che dalla variabile controllata, da un'altra variabile misurabile del processo, a sua volta influenzata dalla variabile di controllo (Figura 1.5). Per esempio, nell'ambito dell'industria chimica le variabili controllate sono spesso le concentrazioni dei prodotti, che sono misurabili solo con apparecchiature complesse e costose. Si usa perciò effettuare un controllo in anello chiuso retroazionando temperature, che si possono misurare in maniera più semplice ed economica e portano con sé informazioni utili per il controllo delle concentrazioni. Si osservi che è piuttosto rara l'eventualità che tutte le componenti del disturbo siano effettivamente misurate. Analogamente, non si considera qui l'ipotesi di misurabilità dei parametri, in quanto di solito irrealistica.

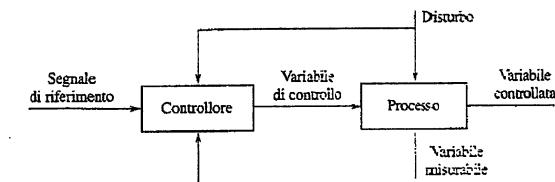
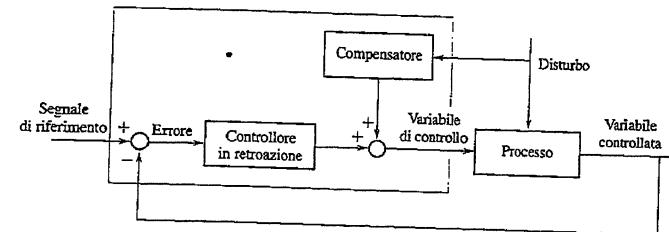


Figura 1.5
Sistema di controllo in anello chiuso con variabili controllata e misurata diverse.

Una situazione tipica di controllo in retroazione con compensazione si ha quando la variabile di controllo è costituita dalla somma di due addendi: uno di essi dipende solo dal disturbo ed è generato da un *compensatore* (pertanto è in anello aperto); l'altro, piuttosto che dipendere separatamente dal segnale di riferimento e dalla variabile controllata, è una funzione dell'errore (1.2) che viene generata da un controllore in retroazione. Come mostrato nella Figura 1.6, l'intero controllore è dunque costituito da due elementi di regolazione e due nodi sommatori. Tuttavia, con lieve abuso di linguaggio, il termine controllore sarà spesso riferito al solo compensatore o al solo controllore in retroazione, dando per scontata la presenza dei nodi sommatori.

Figura 1.6
Sistema di controllo sull'errore con compensazione.



Esempio 1.9 (Seguito degli Esempi 1.2 e 1.8) Realizzare in anello aperto il sistema di controllo della temperatura di un edificio significa assegnare la portata d'aria da addurre in ogni istante di tempo in funzione della temperatura desiderata ed eventualmente di quella esterna (compensazione). Nel far ciò ci si deve riferire a opportune ipotesi sui parametri e i disturbi non misurati: è prassi comune assumere che essi siano tutti pari ai loro valori o andamenti nominali. Se durante il funzionamento queste ipotesi non fossero verificate, gli obiettivi del controllo non sarebbero raggiunti.

Lo stesso sistema di controllo, realizzato in anello chiuso, fa dipendere la portata d'aria anche dalla temperatura controllata. Il controllore ha allora l'informazione necessaria per accettare continuamente le deviazioni dal comportamento nominale del processo e operare congruentemente. Anche le conseguenze di eventi imprevisti (come l'apertura di una finestra) saranno rilevate e potranno essere contrastate.

La struttura di controllo in anello chiuso è evidentemente più potente di quella in anello aperto, che del resto ne rappresenta un caso particolare. In termini generali, è chiaro che quanto maggiori sono le informazioni sul processo a disposizione del controllore, tanto migliori sono le prestazioni ottenibili. Più avanti questo argomento sarà ripreso per mostrare che alcuni risultati di grande importanza pratica, connessi con il contenimento degli effetti dell'incertezza presente sul processo, sono ottenibili solo mediante controllori in anello chiuso, soprattutto quelli in cui si misura proprio la variabile controllata (Figura 1.4). Solo essi infatti, mediane la misura di variabili dipendenti dalle grandezze incerte, possono superare gli inconvenienti procurati dalla non perfetta conoscenza del processo. Invece i controllori in anello aperto sono "ciechi", nel senso che non hanno modo di rilevare le conseguenze del fatto che il processo si trovi in condizioni perturbate, invece che nominali. Tuttavia, proprio la circostanza che in un sistema di controllo retroazionato la variabile di controllo dipende da un'altra variabile, che a sua volta dipende da essa stessa (la struttura informativa "si mangia la coda"), può essere fonte di problemi, per cui i sistemi di controllo in anello chiuso richiedono particolare attenzione nel progetto.

1.3.4 Aspetti realizzativi

Tecnologia del controllore Un regolatore deve elaborare l'informazione disponibile per generare un'appropriata azione di controllo sul processo. Per costruire un controllore automatico si possono utilizzare varie tecnologie, di cui è possibile tracciare una breve cronistoria. Dopo le prime realizzazioni di tipo meccanico in uso fino agli inizi del XX secolo, si sono largamente diffusi i controllori idraulici e pneumatici, nei quali le variabili in gioco sono associate alle portate e alle pressioni di fluidi, quali olio o aria compressa, circolanti all'interno delle apparecchiature. In tal modo è stato possibile realizzare controllori di uso molto generale e dotati di discrete caratteristiche di flessibilità, cioè capaci di svolgere funzioni abbastanza complesse. Essi però risultano costosi e a volte di notevoli ingombri

e peso. Questi controllori sono stati in seguito affiancati prima da controllori eletromecanici o elettrici e poi, riducendo le potenze dei segnali e aumentandone le frequenze, da controllori elettronici. In questo modo si sono potuti realizzare apparati più piccoli, leggeri, economici e di facile manutenzione. Inoltre, si è ottenuto un forte incremento di flessibilità, cosicché tramite l'elettronica è stato possibile realizzare funzioni sempre più avanzate e risolvere problemi di controllo via via più complessi. Tutto ciò si applica già ai primi controllori elettronici, di tipo analogico, ma risulta particolarmente vero per quelli ancora più moderni realizzati con tecnologia digitale, ormai decisamente prevalenti su tutti gli altri. Si tratta, in questo caso, di microprocessori o calcolatori appositamente predisposti, nei quali gli algoritmi di controllo implementati dipendono solo dalla natura di un programma di calcolo e quindi possono essere facilmente impostati, magari sfruttando interfacce appropriate. Di fatto, i controllori digitali hanno veramente rivoluzionato lo scenario, consentendo di superare le limitazioni realizzative poste dai controllori di altro tipo: è nato così il controllo digitale. Non si può però tacere il fatto che l'uso di controllori digitali, nei quali le variabili in gioco sono sequenze di numeri i cui valori sono soggetti alle approssimazioni dovute alle codifiche usate, comporta non piccole difficoltà in sede di analisi e di progetto, anche per la necessità di considerare la presenza nell'anello di apparecchiature che effettuano le necessarie conversioni delle variabili dal mondo del tempo continuo, tipico per i processi, a quello digitale e viceversa.

Esempio 1.10 (Seguito dell'Esempio 1.6) L'ipertensione arteriosa si può curare misurando di tanto in tanto la pressione (per esempio ogni 8 ore) e scegliendo, in funzione del suo valore, la quantità di medicinale, spesso multipla di una quantità base (una pillola), da assumere fino alla misurazione successiva. Si realizza così un controllo digitale, seppure manuale, su un processo governato da variabili a tempo continuo.

Strumentazione di processo Affinché un processo possa essere connesso a un controllore, è indispensabile che esso sia corredata da un'adeguata strumentazione che lo predispone al controllo. Pertanto, gli schemi delle Figure 1.3-1.6 devono essere interpretati come configurazioni valide in linea di principio. Invece, nella pratica, a monte del processo vi è un attuatore per ogni variabile di controllo e ogni variabile misurata richiede la presenza di un trasduttore. Allora, per esempio, lo schema a controllo sull'errore con compensazione di Figura 1.6 diviene quello di Figura 1.7. In esso compaiono anche disturbi sull'attuatore e sui trasduttori.

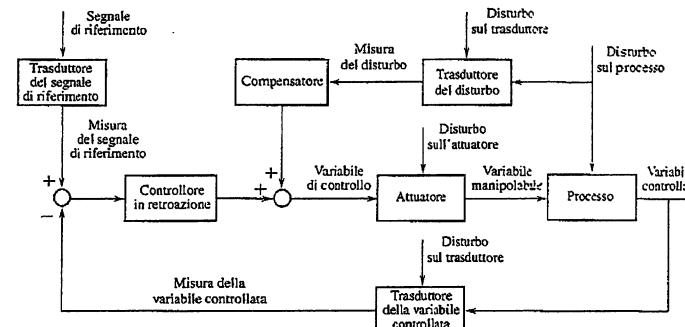


Figura 1.7
Sistema di controllo in anello chiuso con compensazione, completo di strumentazione.

tori: quello sull'attuatore potrebbe essere compensato come quello sul processo, mentre non è realistico pensare di poter misurare i disturbi sui trasduttori.

→ I trasduttori della variabile controllata e del disturbo effettuano la misura delle rispettive grandezze fisiche tramite opportuni sensori, e quindi elaborano le informazioni rendendole compatibili con la tecnologia del controllore, cui poi le trasmettono, eventualmente a distanza. Per esempio, se si adotta un controllore elettronico, tutte le variabili di processo (come temperature, pressioni o velocità) vengono trasformate in correnti elettriche oppure tensioni. Per coerenza la stessa operazione deve essere effettuata, almeno in via concettuale, sull'andamento desiderato della variabile controllata.

→ L'attuatore ha la funzione di convertire le variabili prodotte dal controllore nelle variabili manipolabili, che sono relative al processo. Inoltre, imprime un'amplificazione alle variabili di controllo, necessaria perché esse possano effettivamente influire sul processo: infatti, per motivi costruttivi, il controllore funziona di solito a bassa potenza. Per esempio, utilizzando un apparato elettronico per controllare un robot, le variabili di controllo potrebbero essere tensioni con associata una bassa potenza e l'attuatore potrebbe essere costituito da amplificatori e motori elettrici che generano le coppie (variabili manipolabili) necessarie per ottenere i movimenti desiderati.

Nel caso si adottino controllori digitali, all'interno della strumentazione occorre considerare anche i convertitori cui si è accennato in precedenza.

È opportuno aggiungere che, in ogni caso, i controllori e la strumentazione di processo utilizzati in ambito industriale sono soggetti a stringenti normative, tese almeno a imporre che tutte le variabili di collegamento tra i componenti assumano valori all'interno di intervalli prefissati. Il loro rispetto consente di interconnettere senza problemi apparecchiature di produttori diversi.

Esempio 1.11 (Seguito degli Esempi 1.1 e 1.7) Il pilota di un autoveicolo riunisce in sé le funzioni dei trasduttori e del controllore (manuale). Egli innanzitutto è capace, tramile la vista e con l'aiuto del lachimetro, di rilevare la posizione e la velocità (trasduzione). Quindi, per confronto con la posizione e la velocità desiderate, sa elaborare i dati raccolti così da decidere istante per istante quale debba essere la più opportuna posizione per il volante, per i pedali dell'acceleratore e del freno, e per la leva del cambio (controllo). Tramite i meccanismi di attuazione queste grandezze determinano infine le variabili manipolabili: posizione delle ruote, quantità di miscela inviata ai cilindri, posizione delle ganasce dei freni, marcia innestata.

Invece, un sistema di controllo automatico della traiettoria è capace di misurare mediante opportuni sensori, per esempio ottici, le variabili di posizione e velocità e le rappresenta con segnali elettrici digitali (trasduzione). Successivamente essa determina le variabili di controllo, che sono altri segnali elettrici digitali (controllo). Si osservi che il volante, i pedali e la leva del cambio non sono più utilizzati. Infine, le variabili di controllo, tramite l'attuazione, agiscono sulle variabili manipolabili. Ciò può richiedere l'uso di motori elettrici e pompe idrauliche.

Esempio 1.12 In un serbatoio affluiscono due soluzioni liquide con concentrazioni diverse e parzialmente ignote di soluto in uno stesso solvente (Figura 1.8). Esse si mescolano producendo una soluzione di concentrazione ancora diversa che viene usata dall'utenza secondo le sue necessità. Si desidera che il livello e la concentrazione della soluzione nel serbatoio siano costanti e pari a valori prefissati.

Per affrontare questo problema si può pensare di utilizzare un controllore in anello chiuso di tipo elettronico che svolga anche un'azione di compensazione della portata di fluido uscente. Occorrono allora tre trasduttori che misurino il livello di fluido nel serbatoio e la concentrazione della soluzione (variabili controllate), nonché la portata di fluido all'uscita (disturbo). Questi trasduttori devono fornire variabili elettriche che assumano valori all'interno di opportuni intervalli, così da essere intellegibili dal controllore. Su di esso andranno poi impostati i valori di altre variabili elettriche scalate in modo tale da rappresentare il segnale di riferimento (trasduttore del segnale di riferimento). Il controllore genererà, poi, due variabili elettriche che, opportunamente amplificate, agiranno su due motori e quindi su due valvole (attuatori), così da determinare le portate di ingresso dei due liquidi.

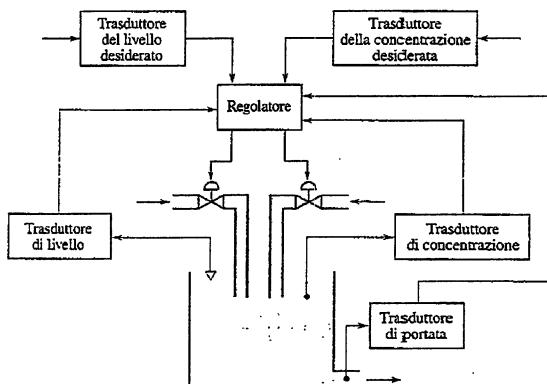


Figura 1.8
Sistema idraulico
dell'Esempio 1.12.

1.4 Ruolo della modellistica matematica

In base alla discussione e agli esempi dei paragrafi precedenti, risulta chiaro che i contesti nei quali i problemi di controllo si possono porre sono svariatissimi, sia nell'ingegneria tradizionale, sia al di fuori di essa, così come diverse possono essere le tecnologie di realizzazione dei controllori. Perciò è naturale chiedersi se e come sia possibile trattare in maniera unitaria una casistica così differenziata, che spazia dalla fisica, alla chimica, all'elettronica, fino all'economia e addirittura alle scienze sociali. In altri termini è lecito domandarsi se si possa istituire e su che cosa si possa fondare una *scienza del controllo*. Lo strumento chiave per rispondere in maniera positiva a tali quesiti sta nell'uso estensivo dei *modelli matematici*, le cui caratteristiche generali sono studiate dalla *teoria dei sistemi*.

1.4.1 Riformulazione dei problemi di controllo

Per affrontare in maniera razionale un problema di controllo è molto conveniente darne innanzitutto una riformulazione in termini puramente matematici. Si noti a questo proposito che nel Paragrafo 1.3.2 si è già prospettata l'opportunità di conoscere gli insiemi cui appartengono i parametri e le classi funzionali cui appartengono i disturbi e i segnali di riferimento. Analogamente, tutte le specifiche di progetto (sull'errore e sul controllo) saranno espresse in termini formali. Tutto questo però non basta. È necessario anche disporre di una descrizione matematica degli elementi che compaiono nel sistema di controllo. Per il processo, i trasduttori e gli attuatori bisogna quindi ricavare opportuni modelli matematici, costituiti da insiemi di relazioni che li descrivono: essi, sostanzialmente, sono sistemi di equazioni algebriche, differenziali, o alle differenze, definite sulle variabili di interconnessione dei singoli componenti.

Procedendo in questa maniera, i problemi di controllo, posti originariamente "nel mondo della realtà", vengono quindi riformulati "nel mondo della matematica", dove è facile riconoscere profonde similitudini formali tra problemi corrispondenti a realtà molto diverse tra loro. Si può allora istituire una *teoria del*

controllo, detta anche *automatica*, che serve ad affrontare in termini matematici astratti tutti questi problemi, per pervenire alla determinazione delle *leggi di controllo*, cioè dei modelli dei controllori. Successivamente, si procederà alla realizzazione dei controllori, cioè alla scelta e messa a punto dei componenti, per esempio elettronici, che svolgono proprio le funzioni previste dai modelli.

Esempio 1.13 Un serbatoio cilindrico di area di base A è alimentato con una portata d'acqua q ; detta h la quota del pelo libero, per la legge di conservazione della massa si può scrivere il modello

$$A\dot{h}(t) = q(t)$$

Si assume ora che si voglia portare il livello al valore costante \bar{h} agendo sulla portata. Trascurando il trascrittore e l'attuatore, si può pensare di adottare un controllore in anello chiuso descritto da

$$q(t) = K_1(\bar{h} - h(t))$$

con $K_1 > 0$. Con questa scelta, quando il livello è inferiore a quello desiderato la portata è positiva, mentre è negativa nel caso opposto; inoltre, il valore assoluto di $q(t)$ è proporzionale alla distanza dall'obiettivo. Allora il sistema di controllo è retto dall'equazione

$$A\dot{h}(t) = K_1(\bar{h} - h(t))$$

che, detto h_{t_0} il livello al tempo t_0 , ha come soluzione

$$h(t) = e^{-(K_1/A)(t-t_0)}h_{t_0} + \left(1 - e^{-(K_1/A)(t-t_0)}\right)\bar{h} \quad t \geq t_0$$

Pertanto, al crescere di t il livello tende esponenzialmente al valore desiderato.

Si abbia ora invece un condensatore di capacità C attraversato da una corrente i ; detta v la tensione ai morsetti, il suo comportamento è retto dall'equazione

$$C\dot{v}(t) = i(t)$$

Se si vuole caricarlo a una tensione \bar{v} , si può assegnare la corrente secondo la formula

$$i(t) = K_2(\bar{v} - v(t))$$

cosicché, detta v_{t_0} la tensione al tempo t_0 , si ottiene

$$v(t) = e^{-(K_2/C)(t-t_0)}v_{t_0} + \left(1 - e^{-(K_2/C)(t-t_0)}\right)\bar{v} \quad t \geq t_0$$

Di nuovo, per $K_2 > 0$, la tensione tende esponenzialmente al valore desiderato.

Si noti come i due problemi, fisicamente diversi, ammettano formulazioni matematiche e soluzioni sostanzialmente identiche.

1.4.2 Problemi di sintesi

In termini estremamente concisi si può dire che la soluzione di un problema di controllo consta di tre passi:

1. la riformulazione matematica del problema di controllo;
2. la determinazione del modello matematico del controllore;
3. la realizzazione del controllore.

Il secondo passo, denominato *problema di sintesi*, o *di progetto*, prescinde quindi in larga misura dalla "realità fisica" del problema, che viene invece confinata nel primo e nel terzo (Figura 1.9). Si osservi però che, se è vero che quanto sopra rende possibile l'esistenza di una scienza del controllo unitaria, bisogna pure sottolineare che un certo grado di interconnessione fra i tre passi continua a susseguirsi, in quanto, per esempio, a ogni fase del secondo passo è sempre opportuno

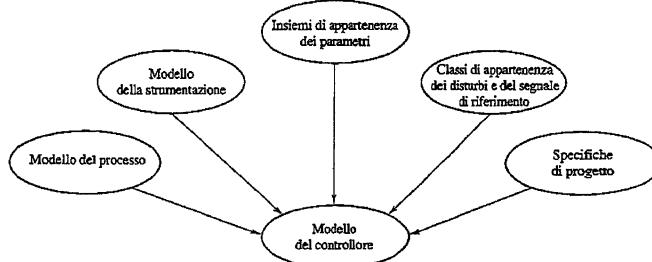


Figura 1.9
Problema di sintesi.

tenere presente il significato reale dei modelli astratti con cui si opera, nonché le questioni realizzative.

Questo testo tratta, quasi esclusivamente, la determinazione del modello matematico del controllore e pertanto, salvo nei casi in cui sia utile fare il contrario, per motivi di semplicità si scriverà "problema di controllo" per indicare la riformulazione matematica del problema "fisico", "controllore" per denotare il modello del componente, e così via.

1.4.3 Problemi di analisi

Invece, un *problema di analisi* nasce quando si vogliono accettare le prestazioni di un sistema di controllo completamente specificato anche per quanto riguarda il modello del controllore (Figura 1.10). Ciò è praticamente indispensabile al termine di un progetto, prima di procedere alla realizzazione e all'installazione sull'impianto, per verificare che siano soddisfatte tutte le esigenze delle quali magari non si è potuto tenere conto esplicitamente. È pure molto utile durante le fasi intermedie di un progetto svolto mediante tecniche di tipo *trial-and-error*, cioè per successivi tentativi guidati proprio dai risultati dell'applicazione di metodi di pura analisi.

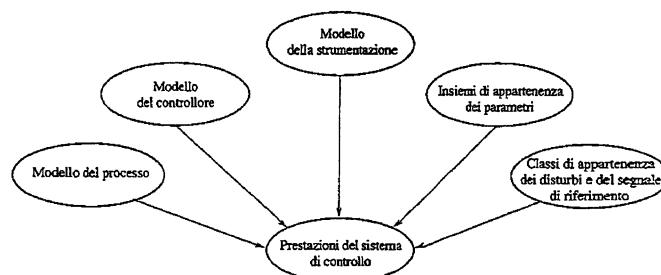


Figura 1.10
Problema di analisi.

Simulazione Malgrado non ci sia lo spazio per discuterne nel seguito, tra le tecniche di analisi merita una segnalazione particolare la *simulazione*, che consiste nello studiare l'evoluzione temporale delle variabili del sistema di controllo risolvendo le equazioni che governano il suo comportamento, ottenute aggregando i modelli matematici dei singoli componenti. Questa operazione non può essere fatta manualmente se non in casi particolarmente semplici; invece, viene effettua-

ta con l'ausilio di calcolatori digitali, capaci di affrontare il problema sfruttando le potenti tecniche del calcolo numerico: si parla perciò di *simulazione digitale*.

Vale la pena di osservare che in passato si ricorreva alla *simulazione analogica*. Essa consisteva nel realizzare un circuito elettronico il cui modello, pur di stabilire opportune corrispondenze tra le variabili, era identico a quello del sistema di controllo da studiare (si riveda a questo proposito l'Esempio 1.13). Dall'andamento delle variabili elettriche si risaliva quindi a quello delle variabili di interesse reale.

Gli elaboratori digitali odierni consentono di utilizzare nella simulazione modelli matematici molto dettagliati, e quindi complessi. Altre tecniche di analisi sono utilizzabili solo su modelli semplificati, mentre modelli ancora più semplici sono richiesti di solito dai metodi di sintesi. È quindi normale l'uso di più modelli della stessa realtà nelle varie fasi dello studio di un sistema di controllo.

1.5 Sistemi di controllo in anello chiuso

La struttura in anello chiuso svolge un ruolo cruciale nella soluzione dei problemi di controllo. Perciò, è opportuno introdurre subito, partendo da esemplificazioni, le sue potenzialità e le ragioni delle difficoltà connesse con il progetto dei controllori, per passare poi a descrivere le caratteristiche di funzionamento dei più comuni di essi.

1.5.1 Confronto con i sistemi di controllo in anello aperto

Conviene iniziare confrontando le prestazioni ottenibili dai sistemi di controllo in anello chiuso con quelle ottenibili dai sistemi di controllo in anello aperto. A questo scopo si tratterà prima il *comportamento statico*, o *di regime*, corrispondente al caso in cui tutte le variabili in gioco siano costanti, e poi il *comportamento dinamico*. Da ciò saranno dedotte alcune importanti conseguenze di cui tenere conto nel progetto del controllore.

Comportamento statico Si consideri l'esempio seguente nel quale la strumentazione sarà trascurata per motivi di semplicità.

Esempio 1.14 Un corpo di massa $M > 0$ è vincolato a un riferimento fisso mediante una molla di costante elastica $k > 0$ e si muove lungo una guida rettilinea caratterizzata da un coefficiente di attrito viscoso (proporzionale alla velocità) $h > 0$; esso è soggetto a una forza motrice F_m e a un'altra forza esterna F_e costante (Figura 1.11). Si vuole controllare la sua posizioni, rappresentata dallo spostamento s rispetto alla posizione di riposo della molla (variabile controllata), agendo sulla

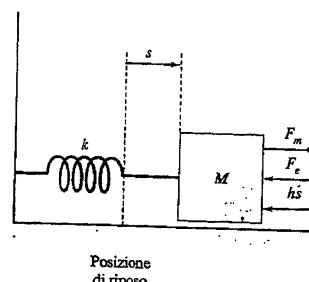


Figura 1.11
Processo dell'Esempio 1.14.

forza motrice (variabile di controllo). Le costanti M , k e h sono parametri di cui sono noti solo i valori nominali \bar{M} , \bar{k} e \bar{h} . Anche la forza esterna non è perfettamente nota: essa è un disturbo di valore nominale \bar{F}_e .

Quando tutte le variabili sono costanti, l'equilibrio delle forze impone la relazione

$$F_m = ks + F_e \quad (1.3)$$

Se l'obiettivo consiste nell'ottenere un certo valore \bar{s} costante di s , si può pensare di assegnare alla forza motrice il valore

$$F_m = F_{ma} = \bar{k}\bar{s} + \bar{F}_e \quad (1.4)$$

Con questa scelta l'obiettivo è palesemente raggiunto quando il processo è in condizioni nominali. L'Equazione (1.4) descrive un controllore in anello aperto, in quanto la variabile di controllo dipende solo dal valore desiderato \bar{s} della variabile controllata (oltre che dal valore nominale di k e F_e); non vi è invece dipendenza da s .

Le prestazioni di questo controllore si degradano però immediatamente non appena si considerino condizioni perturbate. Infatti, sostituendo la (1.4) nella (1.3) si trova un errore

$$e = \bar{s} - s \quad (1.5)$$

pari a

$$e = e_a = \bar{s} - \frac{1}{k}(\bar{k}\bar{s} + \bar{F}_e - F_e) = \frac{k - \bar{k}}{k}\bar{s} + \frac{F_e - \bar{F}_e}{k}$$

cioè costituito da due addendi proporzionali alle incertezze $k - \bar{k}$ e $F_e - \bar{F}_e$.

Supponendo di poter misurare la posizione s , si può invece utilizzare un controllore in anello chiuso il quale aggiunga a F_{ma} un contributo tanto maggiore quanto maggiore è l'errore; in altri termini si può sostituire l'Equazione (1.4) con la relazione

$$F_m = F_{mc} = \bar{k}\bar{s} + \bar{F}_e + \mu e \quad , \quad \mu > 0 \quad (1.6)$$

ottenendo per la (1.5)

$$e = e_c = \frac{k - \bar{k}}{\mu + k}\bar{s} + \frac{F_e - \bar{F}_e}{\mu + k} = \frac{k}{\mu + k}e_a$$

È chiaro così che il controllore in anello chiuso (1.6) consente di ottenere un errore minore in modulo di quello prodotto dal controllore in anello aperto (1.4). Inoltre, se sono noti almeno i valori massimi di $|k - \bar{k}|$ e $|F_e - \bar{F}_e|$, il modulo di e_c può essere reso piccolo a piacere pur di scegliere μ sufficientemente elevato.

Nel caso limite in cui neppure i valori nominali \bar{k} e \bar{F}_e siano ben conosciuti, mentre in anello aperto è praticamente impossibile ottenere un valore soddisfacente dell'errore, in anello chiuso si può ricorrere al controllore

$$F_m = F'_{mc} = \mu e \quad , \quad \mu > 0 \quad (1.7)$$

al quale corrisponde l'errore

$$e = e_c = \frac{k}{\mu + k}\bar{s} + \frac{F_e}{\mu + k}$$

il cui modulo non è nullo nemmeno in condizioni nominali, ma diminuisce in ogni caso all'aumentare di μ .

I risultati trovati nell'esempio hanno una valenza assolutamente generale. Di regola, non è necessario ricorrere a un sistema di controllo in anello chiuso se si può supporre che tutti i parametri e i disturbi abbiano valori e andamenti nominali noti: in questo caso si possono ottenere buone prestazioni statiche con un controllore in anello aperto, che è più semplice perché non richiede la misura della variabile controllata. Però, nelle applicazioni è di fatto inevitabile assumere che le condizioni del sistema di controllo siano perturbate e, allora, la superiorità delle prestazioni statiche raggiungibili con strutture in anello chiuso appare palese: in altre parole i sistemi retroazionati possono avere una *precisione statica* migliore di quella dei sistemi in anello aperto.

Comportamento dinamico È necessario chiedersi se sia lecito studiare i sistemi di controllo ipotizzando che tutte le variabili siano costanti o se invece, così facendo, non si perdano di vista aspetti essenziali per l'analisi e il progetto. In altri termini, bisogna capire se e come le variabili tendono verso i valori previsti dall'analisi statica, quando inizialmente si trovino ad assumere valori differenti.

Esempio 1.15 (Seguito dell'Esempio 1.14) Un modello del processo in esame valido anche in transitorio si può ricavare semplicemente in base alla legge fondamentale della dinamica; si ha così

$$M\ddot{s}(t) = -h\dot{s}(t) - ks(t) + F_m(t) - F_e \quad (1.8)$$

Applicando al processo (1.8) il controllore in anello aperto (1.4) e quello in anello chiuso (1.6), si ottengono rispettivamente le equazioni

$$M\ddot{s}(t) = -h\dot{s}(t) - ks(t) + \bar{k}\ddot{s} - F_e + \bar{F}_e \quad (1.9)$$

e

$$M\ddot{s}(t) = -h\dot{s}(t) - (k + \mu)s(t) + (\bar{k} + \mu)\ddot{s} - F_e + \bar{F}_e \quad (1.10)$$

la seconda delle quali mostra come l'effetto prodotto dall'introduzione del termine proporzionale all'errore (anello chiuso) sia equivalente a un aumento della costante elastica della molla. Queste equazioni confermano che in condizioni statiche (velocità e accelerazione nulle, $\ddot{s}(t) = \dot{s}(t) = 0$) l'errore è nominalmente nullo. Per determinare l'evoluzione temporale della variabile controllata $s(t)$, conseguente alle condizioni iniziali $s(0) = s_0$, $\dot{s}(0) = v_0$, bisogna risolvere le Equazioni (1.9) e (1.10). La teoria delle equazioni differenziali lineari insegnano che le loro soluzioni generali dipendono dalle soluzioni di opportune equazioni algebriche associate, dette equazioni caratteristiche, che, nei casi (1.9) e (1.10), risultano essere

$$M\lambda^2 + h\lambda + k = 0 \quad (1.11)$$

e

$$M\lambda^2 + h\lambda + k + \mu = 0 \quad (1.12)$$

Dette λ_1 e λ_2 le radici di queste equazioni (assunte distinte), le soluzioni delle Equazioni (1.9) e (1.10) assumono la forma

$$s(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + C_3 \quad (1.13)$$

dove le costanti C_1 , C_2 e C_3 dipendono da s_0 e v_0 , oltre che da \ddot{s} (in particolare $C_3 = \ddot{s}$ se il sistema di controllo è in condizioni nominali). Perciò, a parte la costante C_3 , la variabile controllata è la somma di due esponenziali. Quando gli scalari λ_1 e λ_2 sono complessi, essi sono anche coniugati e, posto $\lambda_1 = \sigma + j\omega$, la (1.13) si può riscrivere nella forma

$$s(t) = C_4 e^{\sigma t} \cos(\omega t + \phi) + C_5 \quad (1.14)$$

dove C_4 , C_5 e C_6 sono altre costanti (con $C_6 = \ddot{s}$ se il sistema di controllo è in condizioni nominali).

L'Equazione (1.11) è la stessa che si scriverebbe per il processo non controllato (1.8). Pertanto l'inserzione del controllore in anello aperto non ha modificato la dinamica del sistema. Invece, la dinamica del sistema di controllo in anello chiuso dipende dal parametro di progetto μ . È facile verificare che per valori positivi ed elevati di μ , che sono quelli di maggior interesse per ridurre la sensitività dell'errore statico alle variazioni parametriche e ai disturbi, la soluzione della (1.12) è proprio del tipo (1.14) con $\sigma < 0$ fisso e ω crescente con μ .

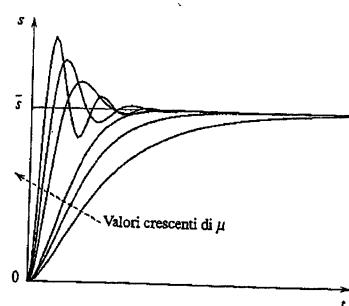


Figura 1.12
Trasitori della posizione del processo massa-molla nominale in condizioni iniziali nulle sotto l'azione del controllore (1.6).

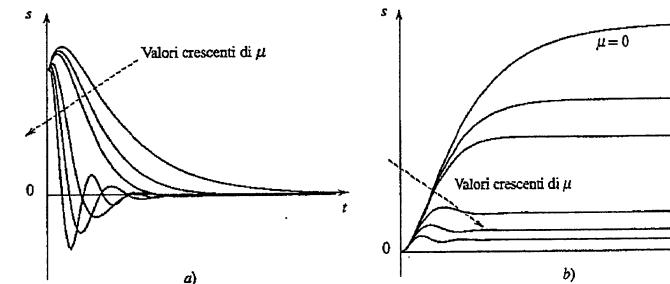


Figura 1.13
Transitori della posizione del processo massa-molla sotto l'azione del controllore (1.6) con $r = 0$: a) in condizioni nominali con $s_0 \neq 0$ e $v_0 \neq 0$; b) in condizioni iniziali nulle con $F_e \neq \bar{F}_e$.

La Figura 1.12 mostra alcuni possibili transitori relativi al caso $s_0 = v_0 = 0$, tracciati in condizioni nominali sotto l'ipotesi che le soluzioni della (1.11), relativa al processo privo di controllore in anello chiuso, siano reali, cioè, come si dice nella meccanica, esso sia sovraromorzato. Si può osservare che all'aumentare di μ corrisponde dapprima un incremento della velocità con cui la posizione s tende al suo valore finale e poi l'insorgere di sovraclonazioni di ampiezza sempre maggiore, perché l'azione di controllo diventa "troppo" sensibile all'errore. Considerazioni analoghe si potrebbero fare se si studiasse il sistema di controllo con diverse condizioni iniziali o in condizioni perturbate. Per esempio, la Figura 1.13 riporta i transitori relativi al caso in cui il sistema di controllo sia in condizioni nominali con $s_0 \neq 0$ e $v_0 \neq 0$, nonché a quello in cui F_e sia una costante diversa da \bar{F}_e e le condizioni iniziali siano nulle; per semplicità si è assunto $\ddot{s} = 0$. Si noti in particolare nella Figura 1.13b come il valore finale dell'errore diminuisca all'aumentare di μ . È anche interessante osservare che, al crescere di μ , la forza motrice assume ampiezze sempre più elevate e tende a oscillare. Si veda a questo proposito la Figura 1.14, nella quale le curve riportate riguardano i più bassi valori di μ considerati nella Figura 1.12. Almeno dal punto di vista qualitativo, i risultati validi in condizioni nominali possono essere ritenuti significativi anche in condizioni perturbate, se i parametri M , h e k hanno solo "piccole" variazioni, perché le radici delle Equazioni (1.11) e (1.12) sono funzioni continue dei coefficienti delle potenze di λ .

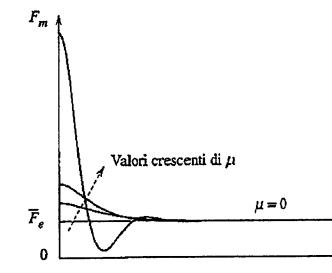


Figura 1.14
Transitori della forza motrice del processo massa-molla nominale in condizioni iniziali nulle sotto l'azione del controllore (1.6).

Vale la pena di notare che, qualora a μ si assegnasse un valore negativo elevato, una delle due radici dell'Equazione (1.12) diventerebbe positiva e di conseguenza l'esponentiale corrispondente che appare nell'Equazione (1.13) sarebbe crescente: in questo caso la posizione $s(t)$ della massa non tenderebbe a \ddot{s} , ma divergerebbe verso l'infinito.

È facile verificare che, se al controllore (1.6) si sostituisse quello di Equazione (1.7), il comportamento dinamico del sistema di controllo sarebbe ancora descritto da equazioni della forma (1.13), (1.14) (con differenti valori delle costanti C_i , ma non di λ_1 , λ_2 , σ e ω) e quindi non cambierebbe qualitativamente: infatti, la modifica delle caratteristiche dinamiche del processo è dovuta al termine μe che appare in entrambe le equazioni.

Anche in questo caso i risultati dell'esempio hanno carattere generale. Nello studiare un sistema di controllo, non è mai sufficiente riferirsi al suo comportamento di regime. Invece occorre sempre esaminare con cura anche le caratteristiche dei transitori per accettare che le variabili mostriano la tendenza a portarsi sui valori individuati nell'analisi statica: a questo proposito si parlerà di proprietà di stabilità. Il processo può goderne oppure no; le proprietà di stabilità comunque si

possono acquisire, o perdere, in conseguenza dell'applicazione di un controllore in anello chiuso. Bisogna poi verificare che i transitori delle variabili controllate siano sufficientemente veloci e privi di sovraelongazioni troppo ampie, cioè che il sistema di controllo abbia una buona *precisione dinamica*. Se, oltre a ciò, le azioni di controllo sono moderate, si dice che i sistemi di controllo sono dotati di buone *prestazioni dinamiche*.

Un controllore in anello aperto non ha sostanziale influenza sul comportamento dinamico complessivo, che resta strettamente dipendente da quello del solo processo. Invece, con un controllore in anello chiuso adeguatamente progettato, si può tarare, almeno in qualche misura, il comportamento dinamico del sistema di controllo, riducendone la sensibilità ai parametri, ai disturbi e alle incertezze sulle condizioni iniziali.

Modelli matematici Dovendo fare riferimento anche allo studio dei transitori, per l'analisi e la sintesi dei sistemi di controllo è indispensabile disporre di modelli *dinamici*, e non solo *statici* (o *non dinamici*), di tutti gli elementi del sistema di controllo, in particolare del processo. I primi sono sostanzialmente costituiti da equazioni differenziali (come la (1.8)), per il tempo continuo, o alle differenze, per il tempo discreto, come si è già accennato al Paragrafo 1.4.1. I secondi, invece, sono semplici equazioni algebriche (come la (1.3)).

Progetto dei sistemi di controllo in anello chiuso Nel Paragrafo 1.3.2 si era affermato che, in un progetto, occorre imporre il soddisfacimento di opportuni requisiti posti sull'errore del sistema di controllo e sulla variabile di controllo. Si può ora precisare che le specifiche devono riguardare, separatamente, sia il comportamento in condizioni di regime, cioè al termine dei transitori, sia le caratteristiche del comportamento durante i transitori stessi.

Si deve notare che i requisiti di tipo statico spesso contrastano parzialmente con i requisiti di tipo dinamico, e quelli sull'errore con quelli sul controllo. Negli Esempi 1.14 e 1.15, lo scalare μ deve essere il più elevato possibile affinché il modulo dell'errore statico sia piccolo anche in condizioni perturbate, mentre il punto di vista dinamico consiglia di attribuirgli valori non troppo elevati per evitare l'instaurarsi di sovraelongazioni ampie durante il transitorio. Inoltre, per valori di μ soddisfacenti dal punto di vista della variabile controllata, la variabile di controllo assume già ampiezze elevate.

Le difficoltà nel progetto nascono sia dalle forti potenzialità dei controllori in anello chiuso nel determinare le proprietà dei corrispondenti sistemi di controllo, con il paradossale pericolo quindi di peggiorare quelle originarie, sia dalla necessità di tenere conto di specifiche statiche e dinamiche le quali impongono che ogni controllore ben progettato realizzzi un buon compromesso tra esigenze opposte.

Uso di elementi di controllo in anello aperto Tutta la discussione precedente era volta a mostrare che certe indispensabili caratteristiche dei sistemi di controllo possono essere ottenute esclusivamente mediante strutture in anello chiuso. Però, ciò non significa che elementi in anello aperto non siano utili: essi, anzi, consentono spesso di migliorare le prestazioni di un sistema di controllo in anello chiuso. Per esempio, se un disturbo è misurabile, l'inserzione di un compensatore permette di ottenere un sistema di controllo più pronto a rispondere alle sue variazioni. Infatti, facendo riferimento allo schema di Figura 1.6, il compensatore agisce al sopravvenire della causa di errore, mentre il controllore in retroazione opera in base agli effetti sulla variabile controllata.

Esempio 1.16 (Seguito degli Esempi 1.14 e 1.15) Se la forza esterna F_e è misurabile, diventa naturale sostituire ai controllori (1.4) e (1.6) quelli descritti dalle equazioni

$$F_m = F_{ma} = \bar{k}\bar{s} + F_e \quad (1.15)$$

$$F_m = F_{mc} = \bar{k}\bar{s} + F_e + \mu e, \quad \mu > 0 \quad (1.16)$$

Si può verificare facilmente, sostituendo le (1.15) e (1.16) nell'Equazione (1.8), che questi controllori fanno sì che il disturbo F_e non abbia alcun effetto sulla posizione s .

1.5.2 Classi di controllori in anello chiuso

Le leggi di funzionamento dei controllori in anello chiuso sono di varia natura: questo paragrafo è dedicato a introdurre le più tipiche. Naturalmente, qui non c'è l'obiettivo di spiegare la logica di progetto dei controllori, che sarà argomento di capitoli successivi, ma solo di illustrare mediante esempi alcuni risultati ottenibili.

Controllori statici o non dinamici I regolatori più semplici sono quelli *statici*, o *non dinamici*, che fanno dipendere il valore istantaneo della variabile di controllo esclusivamente dai valori allo stesso istante del segnale di riferimento, della variabile controllata ed eventualmente dei disturbi misurabili. Per esempio, i controllori (1.6) e (1.7) appartengono a questa classe.

Controllori dinamici Maggiori sono le potenzialità di quei regolatori, detti *dinamici*, nei quali il valore della variabile di controllo in un certo istante dipende anche dai valori passati del segnale di riferimento, della variabile controllata e dei disturbi misurabili. Comunemente, ma non sempre, in questo caso il legame tra le variabili in gioco è espresso mediante equazioni differenziali, o alle differenze.

Esempio 1.17 (Seguito degli Esempi 1.14-1.16) Si prenda come riferimento il semplice controllore (1.7) e lo si sostituisca con

$$F_m(t) = F_{mc}(t) = \mu e(t) + v \int_0^t e(\tau) d\tau, \quad \mu > 0, v > 0 \quad (1.17)$$

il quale fa dipendere la forza motrice da tutto l'andamento dell'errore tra l'istante iniziale $t = 0$ e l'istante corrente. Sostituendo la (1.17) nella (1.8), che descrive il processo, si ottiene

$$M\ddot{e}(t) = -h\dot{s}(t) - ks(t) + \mu e(t) + v \int_0^t e(\tau) d\tau - F_e$$

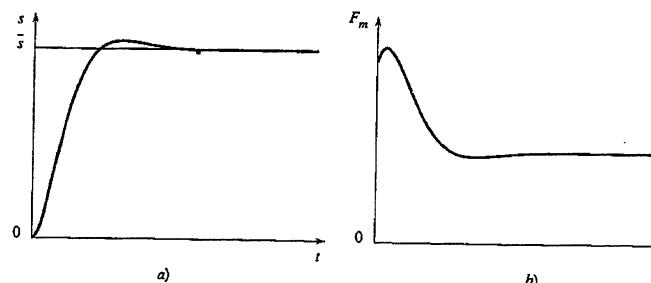
Se si deriva questa equazione rispetto al tempo, ricordando che \dot{s} e F_e sono costanti, per l'Equazione (1.5) che definisce l'errore, si può anche scrivere

$$M\ddot{e}(t) = -h\ddot{e}(t) - (k + \mu)\dot{e}(t) - ve(t) \quad (1.18)$$

e da questa relazione si deduce che se l'errore è costante ($\ddot{e}(t) = \dot{e}(t) = \dot{e}(t) = 0$) esso è anche nullo. Si osservi che questo fatto è assolutamente indipendente dal valore assunto dai parametri M , h , k e dal disturbo F_e e, pertanto, l'errore a regime è nullo in condizioni sia nominali sia perturbate. Inoltre non si è assunto che F_e fosse misurabile. Dall'Equazione (1.18) si conclude immediatamente che l'importante risultato raggiunto è frutto dell'inserzione nel controllore dell'azione integrale ($v \neq 0$), il cui peso nell'Equazione (1.17) è influente (v si può scegliere piccolo a piacere).

Rimane ora da verificare che l'errore, inizialmente non nullo insieme alle prime sue due derivate, tende a portarsi verso lo zero, almeno quando il processo è in condizioni nominali. Si rinuncia qui a effettuare una dimostrazione della verità di questo fatto, limitandosi ad affermare che in realtà è possibile scegliere gli scalari μ e v non solo in modo da raggiungere il risultato desiderato, ma addirittura così da fornire al sistema di controllo transitori veloci e ben smorzati, che rimangono tali anche al sopravvenire di piccole perturbazioni dei parametri. La Figura 1.15 mostra in a) un

Figura 1.15
Processo massa-molla nominale sotto l'azione del controllore (1.17) in condizioni iniziali nulle:
a) transitorio della posizione;
b) transitorio della forza motrice.

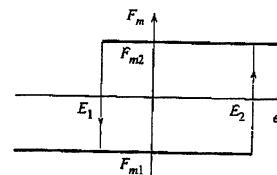


possibile transitorio della posizione s ottenuto per adeguati valori di μ e ν , e in b) il corrispondente andamento della forza motrice. Si nota così che è possibile ottenere prestazioni soddisfacenti a un tempo sia sotto il profilo statico sia sotto quello dinamico.

Controllori a relè Una classe di controllori dinamici molto utilizzati è costituita dai controllori a relè. Essi generano una variabile di controllo che assume solamente un numero finito di valori, determinati in sede di progetto, e commuta dall'uno all'altro di essi quando l'errore supera determinate soglie. Rimandando i dettagli al Capitolo 16, ci si limita qui a considerare un esempio.

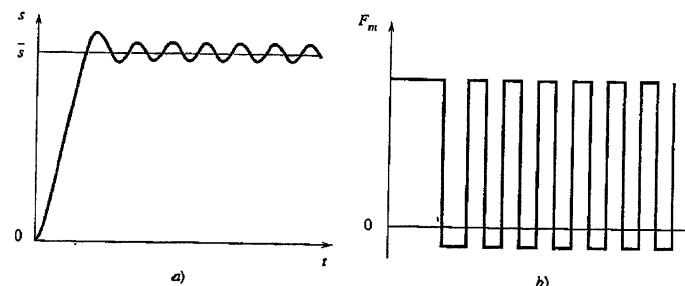
Esempio 1.18 (Seguito degli Esempi 1.14-1.17) Per controllare il processo massa-molla di Equazione (1.8) si utilizzi il relè con isteresi il cui comportamento generico è descritto dalla caratteristica di Figura 1.16. Esso assegna alla forza motrice F_m due soli valori, detti F_{m1} e F_{m2} . La commu-

Figura 1.16
Relè con isteresi.



tazione dal primo al secondo di essi avviene quando l'errore e raggiunge il valore E_2 , mentre la commutazione di F_m a F_{m1} si verifica per $e = E_1$. Il comportamento del sistema in anello chiuso si può ricavare risolvendo l'Equazione (1.8) per assegnate condizioni iniziali e ricordando che la forza motrice è determinata come indicato nella Figura 1.16. L'andamento della posizione s del processo nominale con condizioni iniziali nulle, per opportuni valori degli scalari E_1 , E_2 , F_{m1} e F_{m2} è riportato in a) nella Figura 1.17; in b) è invece rappresentata la forza motrice corrispondente.

Figura 1.17
Processo massa-molla nominale con controllore a relè: a) transitorio della posizione;
b) transitorio della forza motrice.



Si può verificare che i transitori che si ottengono per differenti condizioni iniziali sono del tutto simili a quelli di queste figure.

In generale, al termine di un transitorio, un relè tende a fornire alle variabili del sistema di controllo un andamento di tipo oscillatorio. Se esso è progettato accuratamente, almeno in condizioni nominali la variabile controllata oscilla attorno al suo andamento desiderato, che è tipicamente costante, e l'ampiezza dell'oscillazione è modesta. In senso stretto quindi l'errore a regime non è nullo, né costante, come invece può capitare con controllori di altro genere (si rivedano gli Esempi 1.14-1.17). Però esso può avere moduli così piccoli da risultare accettabili in molti casi, soprattutto in problemi nei quali, più che l'ottenimento di prestazioni di particolare rilievo, conta la semplicità e l'economicità del controllore. Di fatto, molti sistemi di controllo di livello (Esempio 1.3) e temperatura poco sofisticati utilizzano proprio controllori a relè.

1.6 Controllo, supervisione e automazione

I sistemi di controllo sono spesso correddati da una sorta di controllore di secondo livello che effettua operazioni di *supervisione*. Pertanto, partendo dallo schema in anello chiuso di Figura 1.4, dove è stato eliminato il disturbo per motivi di semplicità, si ottiene quanto rappresentato in Figura 1.18.

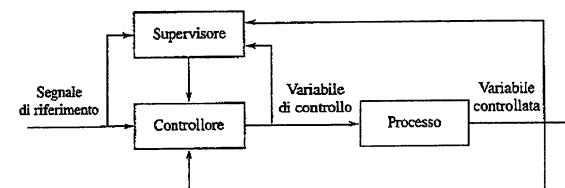


Figura 1.18
Sistema di controllo in anello chiuso con supervisione.

Parecchi e di vario genere sono i compiti svolti dal supervisore, realizzato sempre con tecnologia (elettronica) digitale. Tra questi vale la pena di segnalare innanzitutto l'elaborazione dei dati raccolti sulla variabile di controllo e su quella controllata al fine di aggiornare continuamente il modello del processo. Su questa base esso può di tanto in tanto modificare, o se si preferisce "riprogettare", il controllore di primo livello in modo automatico: si ottiene così un *controllore adattativo*. Un'altra funzione importante del supervisore consiste nella *diagnosica*, cioè nel rilevare la presenza di guasti al processo, al controllore o alla strumentazione, e gestire le situazioni di emergenza.

In realtà, almeno una parte di questi compiti piuttosto avanzati è svolta comunemente con qualche contributo da parte di operatori, ai quali il supervisore si limita a presentare i dati raccolti, magari preelaborati in modo che essi possano prendere più facilmente le decisioni del caso. Il colloquio operatore-supervisore avviene mediante le cosiddette *interfacce uomo-macchina*, poste nelle *sale controllo* e costituite da schermi, allarmi, tastiere e quant'altro è necessario affinché l'operatore possa rendersi conto con la massima facilità e tempestività di quale sia la situazione del sistema di controllo, prendere le migliori decisioni del caso e comunicarle alle apparecchiature di controllo.

In un impianto industriale di grandi dimensioni i sistemi di controllo possono essere centinaia. Singoli gruppi di essi possono essere affidati a singoli supervisori i quali a loro volta sono gestiti da ulteriori elementi di coordinamento, connessi a

un'unica sala controllo. Si ottiene così un *sistema di controllo a molti livelli*, in cui i livelli più elevati hanno anche la funzione di scegliere i segnali di riferimento per i livelli inferiori, su basi che, oltre che tecniche, sono anche di natura economica globale.

Inoltre, all'interno di tali architetture, sistemi di controllo del genere finora considerato coesistono con altri sistemi automatici che presiedono alla corretta sequenzializzazione di alcune operazioni. Si pensi, per esempio, a una cella di produzione automatizzata per la quale vi è la necessità di coordinare le operazioni svolte dalle macchine utensili e dai sistemi di movimentazione dei pezzi da lavorare, in modo da garantire la massima produttività, adattare il funzionamento complessivo alle diverse lavorazioni da effettuare e gestire eventuali situazioni anomale, quali guasti alle macchine, blocco nell'alimentazione dei pezzi o riempimento dell'unità di raccolta in uscita. Nei problemi di questo tipo l'obiettivo non è quello di far seguire ad alcune variabili un andamento predeterminato, quanto piuttosto quello di gestire in maniera corretta una sequenza di *eventi*. Il processo da controllare viene chiamato *a eventi discreti* e i dispositivi che ne governano il funzionamento sono detti in questo caso *controllori logici*. Per l'analisi e la progettazione di questi sistemi di controllo occorre fare ricorso a formalizzazioni matematiche diverse da quella considerata nel Paragrafo 1.4 e basate su strumenti quali la *logica booleana*, gli *automi a stati finiti* e le *reti di Petri*.

I sistemi di controllo complessi, che vanno ben oltre l'imporre a un processo la semplice condizione (1.1) da cui questa trattazione ha preso lo spunto, sono meglio detti *sistemi di automazione*. Malgrado la loro importanza, essi non saranno ulteriormente considerati.

Esempio 1.19 (Seguito degli Esempi 1.1, 1.7 e 1.11) In un moderno autoveicolo esistono, tra gli altri, i sistemi di controllo automatico della frenatura (servofreno e sistemi antibloccaggio), della sterzata (servosterzo), della trazione (cambio automatico e sistemi antisillamento) e della temperatura all'interno dell'abitacolo. Il pilota allora agisce quale controllore manuale di livello superiore (o supervisore) che assegna gli andamenti desiderati delle variabili controllate agendo, rispettivamente, sul pedale del freno, sul volante, sull'acceleratore e sulla manopola del condizionatore.

1.7 Conclusioni

Un problema di controllo si presenta tutte le volte che si voglia imprimere un andamento desiderato a una certa variabile agendo su un'altra grandezza manipolabile che ne determina il comportamento. Quando si affida questo intervento a un dispositivo appositamente progettato si parla di controllo automatico. Si è visto in questo capitolo che le strategie utilizzabili per esercitare tale azione di controllo sono sostanzialmente di due tipi diversi (in anello aperto o in anello chiuso) a seconda della natura dell'informazione disponibile al regolatore. In presenza di incertezza, dovuta a deviazioni dei disturbi dai loro andamenti nominali o all'imprecisa conoscenza di alcuni parametri del processo, solo le strutture in anello chiuso sono in grado, almeno potenzialmente, di fornire al sistema di controllo le caratteristiche di comportamento (statico e dinamico) richieste. È stato inoltre sottolineato il ruolo svolto in questo ambito dalla modellistica matematica, che consente di affrontare con tecniche identiche problemi provenienti da aree applicative differenti. Oltre a fare un breve cenno sulla strumentazione e sulle principali tecnologie costruttive dei controllori, si è messo in evidenza come i sistemi di controllo qui considerati siano spesso solo componenti elementari di sistemi di automazione ben più complessi e articolati.

I Capitoli 2-8 saranno dedicati ai modelli matematici. Per quanto si è illustrato, la conoscenza delle loro principali proprietà e dei relativi metodi di studio è un requisito preliminare per la presentazione dei metodi di analisi e sintesi dei sistemi di controllo, che costituisce il tema dei Capitoli 9-18.

Esercizi

Esercizio 1.1 Nel III secolo A.C., lo scienziato alessandrino Ctesibio inventò l'orologio ad acqua riportato nella Figura 1.19. Se la portata d'acqua q è costante nel tempo, allora il galleggiante A sale nel recipiente cilindrico B e la quota raggiunta dall'indicatore C è proporzionale al tempo trascorso, che è possibile leggere sulla scala graduata D . Per tenere costante la portata q , occorre tenere costante il livello d'acqua h nel piccolo serbatoio cilindrico E , alimentato dalla portata p , dipendente dalla pressione nella condotta di alimentazione e dall'ampiezza dell'orifizio di uscita. A ciò provvede il galleggiante F , che, almeno in prima approssimazione, modifica la portata in modo proporzionale alla variazione di livello.

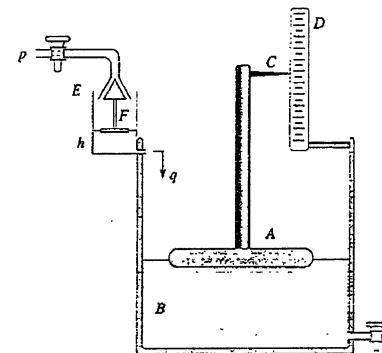


Figura 1.19
Orologio di Ctesibio.

Si rifletta sul funzionamento dell'apparecchiatura, individuando le principali variabili del sistema di controllo del livello del serbatoio E e scrivendone un modello.

Esercizio 1.2 Facendo riferimento all'esperienza quotidiana, si individuino sistemi di controllo (di temperatura, livello, posizione, velocità, o altro ancora). Indicando se siano in anello aperto o chiuso e riconoscendone le variabili controllate, quelle di controllo e i disturbi.

Esercizio 1.3 Si consideri un sistema sotto controllo descritto in condizioni statiche dalla relazione

$$y = ku + hd \quad (1.19)$$

tra la variabile di controllo u , il disturbo d e la variabile controllata y . Per esempio, se le costanti k e h sono positive, l'Equazione (1.19) può rappresentare la dipendenza della velocità di rotazione di un motore elettrico a corrente continua dalla tensione d'armatura e dalla coppia resistente (si veda l'Esempio 2.6). Detto w il valore desiderato di y , allo scopo di fare in modo che risulti

$$y = w$$

si può innanzitutto fare riferimento ai valori nominali \bar{k} , \bar{h} e \bar{d} di k , h e d , e utilizzare un controllore in anello aperto della forma

$$u = \bar{u} \quad (1.20)$$

Si assegna a \bar{u} il valore più opportuno, e quindi si descrivano le conseguenze prodotte sul sistema di controllo (1.19), (1.20) da perturbazioni di k , h e d .

Alternativamente, si può adottare il controllore in anello chiuso descritto da

$$u = \mu(w - y) \quad (1.21)$$

Si studi il comportamento del sistema di controllo (1.19), (1.21) al variare del parametro μ , in condizioni sia nominali sia perturbate.

Si analizzino anche le migliori prestazioni ottenibili quando il disturbo è misurabile e il controllore (1.21) può essere sostituito da

$$u = \mu(w - y) + \beta d$$

con μ e β costanti da scegliere in modo opportuno.

Esercizio 1.4 Secondo un semplice modello economico, il consumo C delle famiglie di una nazione dipende dall'investimento I delle imprese secondo la legge

$$\dot{C}(t) = aC(t) + bI(t)$$

in cui a e b sono due coefficienti positivi con valori nominali \bar{a} e \bar{b} . Interpretando C come variabile controllata I come variabile di controllo, si determini il livello di investimento \bar{I} tale per cui risulti

$$C = \bar{C}$$

in condizioni nominali. Quindi, si studino gli effetti prodotti dalle variazioni di a e b sul sistema di controllo in anello aperto in cui $I = \bar{I}$.

Successivamente, si valuti l'effetto, in transitorio e a regime, dei controllori in anello chiuso descritti da

$$I(t) = \mu(\bar{C} - C(t))$$

e

$$\dot{I}(t) = \mu(\bar{C} - C(t))$$

al variare del parametro di progetto μ , considerando valori sia nominali sia perturbati di a e b .

Esercizio 1.5 Si consideri nuovamente l'Esempio 1.17 e in particolare l'Equazione (1.18), con $M = 1$, $h = 1$, $k = 1$. Si verifichi che, se si assume $\mu = 0.25$ e $v = 0.5$, l'errore tende asintoticamente a zero qualunque sia il valore iniziale suo e delle sue prime due derivate, mentre ciò non accade per $\mu = 1$ e $v = 8$.

(Suggerimento: analogamente a quanto già si è fatto nell'Esempio 1.15, si scriva l'equazione caratteristica associata all'Equazione (1.18), quindi si determinino le sue radici, notando che una di esse è $\lambda = -0.5$, per $\mu = 0.25$ e $v = 0.5$, e $\lambda = -2$, per $\mu = 1$ e $v = 8$.)

Esercizio 1.6 Si applichi al processo di Equazione (1.8) il controllore descritto da

$$F_m(t) = F_{mc}(t) = \tilde{k}\ddot{s} + \tilde{F}_e + \mu e(t) + \eta\dot{e}(t), \quad \mu > 0, \quad \eta > 0$$

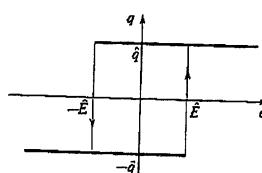
e si ripeta l'analisi del comportamento statico e dinamico del sistema complessivo, seguendo le linee sviluppate negli Esempi 1.14, 1.15 e 1.17.

Esercizio 1.7 Si applichi il controllore a relè con isteresi riportato nella Figura 1.20, in cui

$$e = \bar{h} - h(t)$$

al serbatoio dell'Esempio 1.13, e si studi l'evoluzione del livello del fluido per una generica condizione iniziale $h(t_0) = h_0$, al variare dei parametri \hat{E} e \hat{q} . Si esamini anche l'effetto di un'incertezza sull'area di base A .

Figura 1.20
Relè dell'Esercizio 1.7.



Esercizio 1.8 Si consideri il processo a tempo discreto, con variabile controllata y e variabile di controllo u , descritto dalla relazione

$$y(k+1) = ay(k) + bu(k)$$

dove a e b sono due parametri reali. Quindi si applichi ad esso il controllore

$$u(k) = \mu(w - y(k))$$

dove w è il valore desiderato di y , assunto costante, e μ è un parametro oggetto di scelta. Si valutino le caratteristiche statiche e dinamiche del sistema di controllo così ottenuto, seguendo le

linee sviluppate negli Esempi 1.14, 1.15 e 1.17. Successivamente si affronti lo stesso problema per il controllore

$$u(k+1) = u(k) + \mu(w - y(k))$$

Esercizio 1.9 I componenti usati di solito nel progetto degli amplificatori hanno un "guadagno" elevato, ma fortemente incerto. Accanto a essi sono fortunatamente a disposizione anche componenti capaci di attenuare i segnali in maniera ben nota. Allora, è prassi comune realizzare gli amplificatori sfruttando le caratteristiche della struttura retroazionata di Figura 1.21, dove, limitandosi a un'analisi di tipo statico, il blocco A posto lungo la linea di andata è descritto dalla relazione

$$y = \mu u \quad (1.22)$$

con $\mu \gg 1$, mentre il blocco in retroazione R è un attenuatore in cui

$$v = \alpha y \quad (1.23)$$

con $0 < \alpha \ll 1$. Si calcoli il guadagno y/w del sistema retroazionato e si verifichi che, purché sia $\alpha\mu \gg 1$, esso risulta largamente indipendente dal valore incerto di μ .

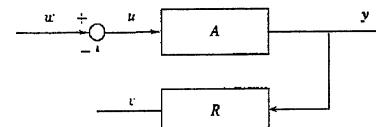


Figura 1.21
Schema in anello chiuso dell'Esercizio 1.9.

Esercizio 1.10 Si rifletta sul sistema di automazione, a molti livelli anche nelle sue forme più semplici, di un ascensore. L'utente indica il piano desiderato, scegliendolo all'interno di un numero finito di valori. A seguito di ciò, si verifica una sequenza di eventi, costituiti dalla chiusura delle porte, dal trasferimento della cabina al piano desiderato e dalla riapertura delle porte. Ognuno di questi eventi, poi, richiede un controllo specifico. Per esempio, il movimento della cabina deve essere effettuato facendole seguire a una fase di accelerazione, una fase di crociera e una di decelerazione; ciò è garantito da adeguati comandi al motore primario. Anche i movimenti delle porte sono effettuati inviando opportuni segnali di controllo ai relativi motori.

2.1 Introduzione

Come si è osservato nel capitolo precedente, la possibilità di affrontare in modo rigoroso i problemi di controllo è strettamente legata alla disponibilità di modelli matematici dei componenti dei sistemi di controllo: processi, trasduttori, attuatori e regolatori. In particolare, è poi indispensabile che questi modelli siano capaci di descrivere il comportamento degli oggetti considerati anche quando le variabili in gioco non siano tutte costanti nel tempo.

In questo capitolo vengono introdotte le nozioni di base sui sistemi dinamici, che sono i modelli matematici correntemente utilizzati nelle applicazioni, e le loro proprietà principali.

In dettaglio, si trattano i temi seguenti:

- gli elementi costitutivi dei sistemi dinamici;
- i criteri di classificazione basati sulle caratteristiche delle equazioni che li descrivono;
- uno speciale sistema dinamico detto ritardo di tempo;
- la particolare modalità di funzionamento chiamata equilibrio;
- la fondamentale proprietà di stabilità.

2.2 Concetti fondamentali

2.2.1 Variabili di ingresso, stato e uscita

Un sistema dinamico a tempo continuo (Figura 2.1) costituisce un modello matematico di un oggetto fisico il quale interagisce con il mondo circostante tramite due vettori di variabili dipendenti dal tempo, assunto reale e indicato con il simbolo t .

Le une, dette *variabili di ingresso*, rappresentano le azioni che vengono compiute sull'oggetto in esame da agenti esterni che ne influenzano il comportamento. Le altre, dette *variabili di uscita*, rappresentano quanto del comportamento dell'oggetto stesso è, per qualche ragione, di interesse. In altri termini, tra le variabili

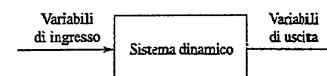


Figura 2.1
Sistema dinamico.

di ingresso e quelle di uscita vi è un rapporto di causa ed effetto; l'evoluzione delle seconde descrive il modo in cui l'oggetto che si sta modellando risponde alle sollecitazioni impresse con le prime.

Esempio 2.1 Si consideri il circuito elettrico nella Figura 2.2. Se si assume come variabile di ingresso la tensione v_G ai morsetti del generatore e come variabile di uscita la tensione v_{R_1} ai morsetti del primo resistore, per la legge delle tensioni si ha

$$v_{R_1}(t) = v_G(t) - v_{R_2}(t)$$

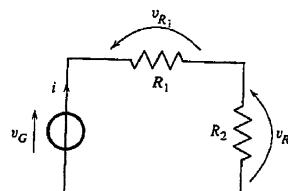
D'altra parte la legge del resistore conduce a scrivere

$$v_{R_1}(t) = R_1 i(t)$$

$$v_{R_2}(t) = R_2 i(t)$$

dove i è la corrente nella maglia. Eliminando da queste equazioni le variabili v_{R_2} e i , si trova in definitiva

$$v_{R_1}(t) = \frac{R_1}{R_1 + R_2} v_G(t) \quad (2.1)$$



La situazione incontrata nell'esempio precedente è del tutto particolare in quanto l'Equazione (2.1) implica che la conoscenza del valore assunto in un certo istante di tempo della variabile di ingresso basta a determinare il valore assunto nello stesso istante di tempo della variabile di uscita. Nella grande maggioranza dei casi di interesse invece non è così, e la conoscenza del valore in un certo istante delle variabili di ingresso non è sufficiente a individuare il valore allo stesso istante delle variabili di uscita, anzi spesso non è nemmeno necessaria.

Esempio 2.2 Si consideri il circuito elettrico nella Figura 2.3. Se si assume come variabile di ingresso la tensione v_G ai morsetti del generatore e come variabile di uscita la tensione v_R ai morsetti del resistore, per la legge delle tensioni si ha

$$v_R(t) = v_G(t) - v_C(t) \quad (2.2)$$

Il valore al tempo t dell'uscita v_R , quindi, non dipende solo dal valore dell'ingresso v_G allo stesso istante di tempo, ma anche dalla situazione interna del circuito. Quest'ultima è caratterizzata dall'energia elettrica accumulata nel condensatore, che è funzione della tensione v_C ai morsetti.

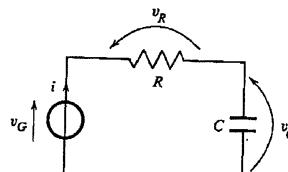


Figura 2.2
Circuito elettrico dell'Esempio 2.1.

Figura 2.3
Circuito elettrico dell'Esempio 2.2.

Generalizzando quanto risulta dall'esempio appena trattato, si conclude allora che per descrivere compiutamente l'oggetto da modellare è spesso necessario introdurre un terzo vettore di variabili, chiamate *variabili di stato*, che descrivono la sua situazione interna.

2.2.2 Rappresentazione di stato

Introdotte le variabili necessarie, si può ora passare a descrivere la struttura delle equazioni che le mettono in relazione o, in altri termini, si può dare la definizione di sistema dinamico a tempo continuo. Se si indicano rispettivamente con $u \in R^n$, $x \in R^p$ e $y \in R^q$ i vettori delle variabili di ingresso, stato e uscita, e con f e g due funzioni vettoriali, un *sistema dinamico a tempo continuo* è costituito dalle equazioni

$$\dot{x}(t) = f(x(t), u(t), t) \quad (2.3)$$

$$y(t) = g(x(t), u(t), t) \quad (2.4)$$

e può essere rappresentato come nella Figura 2.4. Per brevità si parla anche semplicemente di *sistema*; se invece si vuole sottolineare il fatto che la variabile tempo, da cui dipendono tutte le altre variabili utilizzate, è reale, si dice più propriamente che il sistema dinamico è *a tempo continuo*, o *continuo nel tempo*. L'Equazione (2.3) costituisce l'*equazione di stato*, la (2.4) va sotto il nome di *trasformazione d'uscita*, mentre il numero n delle variabili di stato si dice *ordine* del sistema.

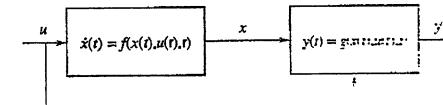


Figura 2.4
Sistema dinamico a tempo continuo: equazione di stato e trasformazione d'uscita.

Esempio 2.3 (Seguito dell'Esempio 2.2) Per il circuito elettrico di Figura 2.3 le leggi del condensatore e del resistore portano a scrivere

$$\begin{aligned} C \dot{v}_C(t) &= i(t) \\ v_R(t) &= R i(t) \end{aligned}$$

dove i è la corrente nella maglia. Allora, ponendo $u = v_G$, $x = v_C$, $y = v_R$ e ricordando la (2.2), si ottiene

$$\begin{aligned} \dot{x}(t) &= \frac{1}{RC} (u(t) - x(t)) \\ y(t) &= u(t) - x(t) \end{aligned} \quad (2.5) \quad (2.6)$$

cioè un sistema dinamico di ordine $n = 1$.

Il legame tra l'ingresso e l'uscita risulta quindi scisso in due parti e descritto mediante due equazioni vettoriali:

- un'equazione differenziale (l'*equazione di stato*), che mette in relazione con l'ingresso le variabili che descrivono la situazione interna del sistema;
- un'equazione algebrica (la *trasformazione d'uscita*), che consente di determinare l'uscita a uno specifico istante di tempo sulla base della conoscenza di tale situazione e dell'ingresso allo stesso istante di tempo.

L'Equazione (2.3) definisce l'evoluzione dello stato $x(t)$ per $t > t_0$, in corrispondenza di ogni terza costituita da un istante iniziale t_0 , una funzione di ingresso $u(t)$, $t \geq t_0$, e una condizione iniziale $x(t_0) = x_0$. La funzione $x(t)$, $t \geq t_0$, si dice *movimento dello stato* del sistema. L'Equazione (2.4) permette invece di determinare l'evoluzione dell'uscita $y(t)$ per $t \geq t_0$, in corrispondenza della funzione di ingresso $u(t)$ e dell'andamento dello stato $x(t)$ per $t \geq t_0$. La funzione $y(t)$, $t \geq t_0$, va sotto il nome di *movimento dell'uscita*.

Esempio 2.4 (Seguito degli Esempi 2.2 e 2.3) Per semplice sostituzione nelle equazioni che lo descrivono, si può verificare che i movimenti dello stato e dell'uscita del sistema (2.5), (2.6) che si ottengono in corrispondenza di $t_0 = 0$, $u(t) = U \sin(\omega t)$, $t \geq 0$, e $x(0) = x_0$ sono

$$x(t) = e^{-t/RC} x_0 + \frac{U \omega R C}{1 + \omega^2 R^2 C^2} e^{-t/RC} + \frac{U}{\sqrt{1 + \omega^2 R^2 C^2}} \sin(\omega t - \gamma) \quad (2.7)$$

$$y(t) = -e^{-t/RC} x_0 - \frac{U \omega R C}{1 + \omega^2 R^2 C^2} e^{-t/RC} + \frac{U \omega R C}{\sqrt{1 + \omega^2 R^2 C^2}} \cos(\omega t - \gamma) \quad (2.8)$$

dove

$$\gamma = \arctan(\omega R C) \quad (2.9)$$

La possibilità di determinare i movimenti dei sistemi dinamici è legata alla capacità di risolvere, analiticamente o numericamente, le equazioni differenziali di stato (2.3), in generale non lineari e a coefficienti variabili nel tempo, che li definiscono. La soluzione di tali equazioni è di solito decisamente complessa. Nel caso particolare in cui esse siano invece lineari e a coefficienti costanti, esiste una formula risolutiva esplicita molto espressiva, che sarà illustrata nel capitolo seguente.

Infine, conviene notare che si usa dire che le Equazioni (2.3), (2.4) costituiscono una *rappresentazione di stato, o ingresso-stato-uscita, o interna* di un sistema dinamico a tempo continuo.

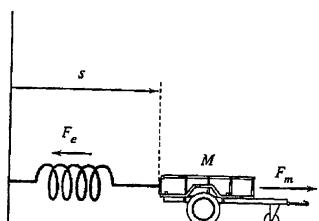
2.2.3 Esempi

Gli esempi seguenti mostrano come effettivamente i sistemi dinamici permettano di rappresentare in maniera unitaria un gran numero di oggetti reali facenti capo a settori parecchio differenti tra loro, quali l'elettrotecnica, la meccanica e la termodinamica.

Esempio 2.5 Il carrello nella Figura 2.5 si muove di moto rettilineo ed è soggetto alla forza motrice F_m e alla forza elastica F_e , dipendente dall'allungamento s della molla secondo la relazione

$$F_e(s(t), t) = k(t)s(t) = k_{t_0} e^{-\alpha(t-t_0)} s(t), \quad t \geq t_0 \quad (2.10)$$

Figura 2.5
Carrello dell'Esempio 2.5.



dove $k_{t_0} > 0$ e $\alpha > 0$ in quanto si suppone che la costante elastica k diminuisca esponenzialmente nel tempo. Assumendo come variabile di uscita l'energia totale E_T del carrello di massa M , considerato puntiforme, si ottiene

$$E_T(t) = \frac{1}{2} (k(t)s^2(t) + M\dot{s}^2(t))$$

In base alla legge fondamentale della meccanica si ha poi

$$M\ddot{s}(t) = F_m(t) - F_e(t)$$

Perciò, ponendo $u = F_m$, $x_1 = s$, $x_2 = \dot{s}$, $y = E_T$, per $t \geq t_0$ risulta

$$\dot{x}_1(t) = x_2(t) \quad (2.11)$$

$$\dot{x}_2(t) = \frac{1}{M} (u(t) - k_{t_0} e^{-\alpha(t-t_0)} x_1(t)) \quad (2.12)$$

$$y(t) = \frac{1}{2} (k_{t_0} e^{-\alpha(t-t_0)} x_1^2(t) + M x_2^2(t)) \quad (2.13)$$

Si noti che l'energia totale (uscita) non dipende esplicitamente dalla forza motrice che, quale azione di agenti esterni sull'oggetto in esame, rappresenta la variabile di ingresso. L'energia totale dipende invece dall'allungamento e dalla velocità, che descrivono la situazione interna del carrello. Si osservi pure che, perfino nel caso in cui posizione e velocità assumessero gli stessi valori in due diversi istanti di tempo, i corrispondenti valori dell'energia sarebbero comunque differenti per la dipendenza dal tempo della costante elastica della molla.

Esempio 2.6 Un motore elettrico a corrente continua con eccitazione indipendente può essere schematizzato come nella Figura 2.6, dove v e i rappresentano rispettivamente la tensione e la corrente nel circuito d'armatura collocato nel rotore, R ed L sono i valori di resistenza e induttanza dello stesso circuito ed e_f è la cosiddetta forza elettromotrice, che corrisponde alla tensione generata sul circuito di armatura per effetto dell'interazione con il circuito di eccitazione, disposto sullo statora. Se quest'ultimo circuito è alimentato da una tensione di eccitazione costante, si può ritenere che la forza elettromotrice sia proporzionale alla velocità di rotazione w , ovvero $e_f = kw$, dove $k > 0$ è un'opportuna costante. D'altra parte, in assenza di perdite energetiche nel motore, la coppia generata risulta $C_m = ki$. Questa coppia motrice, insieme alla coppia resistente C_r dovuta al carico meccanico e all'eventuale coppia di attrito C_a che si assumrà proporzionale alla velocità (ovvero $C_a = hw$, dove $h \geq 0$ è il coefficiente di attrito), sono responsabili della rotazione dell'albero motore, di momento di inerzia J .

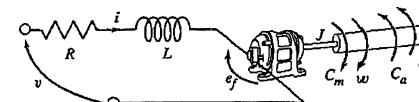


Figura 2.6
Motore elettrico
dell'Esempio 2.6.

Utilizzando le convenzioni di segno mostrate nella Figura 2.6, la "parte elettrica" del motore è quindi descritta dalle equazioni

$$v(t) - e_f(t) = Ri(t) + L\dot{i}(t)$$

$$e_f(t) = kw(t)$$

mentre la "parte meccanica" è descritta da

$$J\ddot{\omega}(t) = C_m(t) - C_r(t) - hw(t)$$

$$C_m(t) = ki(t)$$

Scegliendo $x_1 = y_1 = i$ e $x_2 = y_2 = w$ come variabili di stato e di uscita e denotando con $u_1 = v$ e $u_2 = C_r$ le due variabili di ingresso, il sistema è rappresentabile mediante le equazioni

$$\dot{x}_1(t) = -\frac{R}{L} x_1(t) - \frac{k}{L} x_2(t) + \frac{1}{L} u_1(t) \quad (2.14)$$

$$\dot{x}_2(t) = \frac{k}{J} x_1(t) - \frac{h}{J} x_2(t) - \frac{1}{J} u_2(t) \quad (2.15)$$

$$y_1(t) = x_1(t) \quad (2.16)$$

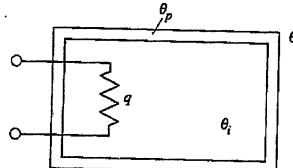
$$y_2(t) = x_2(t) \quad (2.17)$$

È quindi un sistema di ordine $n = 2$, con due variabili di ingresso e due di uscita.

Esempio 2.7 Per il forno rappresentato schematicamente nella Figura 2.7 la temperatura interna θ_i rappresenta la variabile di uscita, mentre la temperatura esterna θ_e e la potenza q fornita all'ambiente sono le variabili di ingresso. Detti C_f la capacità termica del forno e k_{ip} il coefficiente di scambio tra l'interno e l'esterno del forno, per il principio di conservazione dell'energia si ha

$$C_f \dot{\theta}_i(t) = k_{ip}(\theta_e(t) - \theta_i(t)) + q(t)$$

Figura 2.7
Forno dell'Esempio 2.7.



Il processo è quindi descritto da un sistema dinamico del primo ordine con variabile di stato θ_i . Un modello più dettagliato si può scrivere se si trattano in maniera separata gli accumuli energetici nel fluido all'interno del forno e nella parete. Denotando con C_i la capacità termica del fluido, con C_p quella della parete a temperatura θ_p , con k_{ip} il coefficiente di scambio tra fluido e parete e con k_{pe} quello tra parete ed esterno si ha allora

$$C_i \dot{\theta}_i(t) = k_{ip}(\theta_p(t) - \theta_i(t)) + q(t)$$

$$C_p \dot{\theta}_p(t) = k_{pe}(\theta_e(t) - \theta_p(t)) - k_{ip}(\theta_p(t) - \theta_i(t))$$

Ponendo $u_1 = q$, $u_2 = \theta_e$, $x_1 = \theta_i$, $x_2 = \theta_p$ e $y = \theta_i$, si ottengono le equazioni

$$\dot{x}_1(t) = \frac{k_{ip}}{C_i}(x_2(t) - x_1(t)) + \frac{1}{C_i}u_1(t) \quad (2.18)$$

$$\dot{x}_2(t) = \frac{k_{pe}}{C_p}(u_2(t) - x_2(t)) - \frac{k_{ip}}{C_p}(x_2(t) - x_1(t)) \quad (2.19)$$

$$y(t) = x_1(t) \quad (2.20)$$

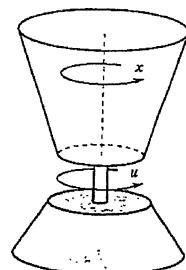
che costituiscono un sistema dinamico di ordine $n = 2$.

Esempio 2.8 Per la centrifuga rappresentata schematicamente nella Figura 2.8, J indica il momento d'inerzia rispetto all'asse di rotazione e $k > 0$ la costante caratteristica di una coppia d'attrito proporzionale al quadrato della velocità angolare di rotazione. Tale velocità costituisce la variabile di uscita del sistema, mentre la variabile di ingresso u è la coppia motrice. Ipotizzando che J non dipenda dalla velocità angolare e assumendo quest'ultima anche come variabile di stato, per le leggi della dinamica si ha

$$\ddot{x}(t) = -\frac{k}{J}x^2(t)\operatorname{sgn}(x(t)) + \frac{1}{J}u(t) \quad (2.21)$$

$$y(t) = x(t) \quad (2.22)$$

Figura 2.8
Centrifuga dell'Esempio 2.8.



dove

$$\operatorname{sgn}(x) = \begin{cases} \frac{x}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Se la coppia motrice è nulla, cioè $u(t) = 0$, il movimento conseguente a una condizione iniziale $x(0) = x_0$ è dato da

$$y(t) = x(t) = \frac{Jx_0}{J + kx_0 \operatorname{sgn}(x_0)} t \quad (2.23)$$

Se invece la coppia motrice è $u(t) = M > 0$, ponendo $h = \sqrt{M/k}$, il movimento conseguente a una condizione iniziale $x(0) = x_0$ è

$$y(t) = x(t) = \begin{cases} h(x_0 + h) + (x_0 - h)e^{-(2kh/J)t} & , x_0 \geq 0 \\ h(x_0 + h) - (x_0 - h)e^{-(2kh/J)t} & , x_0 < 0 \\ h \tan\left(\frac{kh}{J}t + \arctan\left(\frac{x_0}{h}\right)\right), & t \leq t_1 \\ h \frac{1 - e^{-(2kh/J)(t-t_1)}}{1 + e^{-(2kh/J)(t-t_1)}}, & t > t_1 \end{cases} \quad (2.24)$$

dove

$$t_1 = -\frac{J}{kh} \arctan\left(\frac{x_0}{h}\right) \quad (2.25)$$

Questi risultati sono facilmente verificabili per sostituzione.

Come mostrato dai casi appena presentati, alle variabili di ingresso e di uscita non sono necessariamente associati i concetti fisici di afferenza ed efferenza di massa, energia o altro, ma solo quelli logici di causa ed effetto. Come ulteriore esempio di quanto affermato, nella descrizione analitica dell'andamento temporale del livello del liquido contenuto in un serbatoio alimentato per mezzo di una pompa (Figura 2.9), il livello è senz'altro da considerare quale variabile di uscita, mentre la portata della pompa è da annoverare tra le variabili di ingresso, in quanto contribuisce a determinare i valori successivamente assunti dal livello stesso, indipendentemente dal fatto che essa sia positiva o negativa, cioè corrisponda ad afferenza o efferenza di liquido.

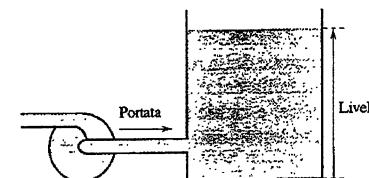


Figura 2.9
Serbatoio alimentato da una pompa.

2.2.4 Commenti sul concetto di stato

Vale la pena riflettere brevemente sul problema della scelta delle variabili e sulla formulazione delle equazioni di stato. Le variabili di stato rendono conto di quanto è necessario conoscere della situazione interna, o, se si preferisce, della storia passata del sistema, per poter calcolare l'uscita. Nei sistemi fisici la situazione interna è tipicamente determinata da accumuli di energia, quantità di moto o massa e perciò può essere opportuno scegliere come variabili di stato quelle variabili da cui questi accumuli dipendono. Pertanto, nei modelli dei circuiti elettrici si possono scegliere come variabili di stato le tensioni ai morsetti dei condensatori,

come nell'Esempio 2.3, perché da queste dipende l'energia elettrica accumulata in essi, e le correnti negli induttori, come nell'Esempio 2.6, perché da esse dipende l'energia magnetica accumulata al loro interno. Invece nei modelli dei sistemi meccanici può convenire scegliere come variabili di stato posizioni e velocità dei vari elementi, come fatto negli Esempi 2.5, ancora 2.6 e 2.8, perché legate ad accumuli di energia potenziale e cinetica o di quantità di moto. Infine, nei sistemi termici, come quello dell'Esempio 2.7, le temperature sono adatte a essere assunte come variabili di stato, perché è da esse che dipendono le energie termiche immagazzinate. Le leggi fondamentali della fisica (elettromagnetismo, meccanica, termodinamica, idraulica o altro) esprimono poi delle relazioni tra variabili di ingresso, variabili di stato e loro derivate che, opportunamente elaborate, conducono alle equazioni di stato. Anche se i criteri di scelta delle variabili di stato sopra suggeriti andrebbero meglio precisati, in ogni caso essi costituiscono un'indicazione molto utile nelle applicazioni pratiche.

È opportuno aggiungere pure che le variabili e le equazioni di stato di un sistema dinamico non sono definite in maniera unica. Nel Paragrafo 3.2.4, seppure per una classe particolare di sistemi dinamici, si mostrerà che esistono addirittura infinite possibili scelte delle variabili di stato: alcune, come quelle effettuate sulla base delle considerazioni appena svolte, potranno essere più comode per una prima scrittura delle equazioni del sistema; altre invece potranno essere più adatte a uno studio analitico delle sue proprietà. Queste ultime scelte potranno portare addirittura a definire lo stato non in campo reale, come si è finora ipotizzato, bensì in campo complesso, ponendo $x \in C^n$. Tutte le scelte, comunque, saranno tali da stabilire la stessa relazione tra l'ingresso e l'uscita del sistema.

Va anche osservato che neppure il numero delle variabili di stato, cioè l'ordine del sistema, è fissato a priori; in altre parole, esistono più sistemi dinamici associabili a un unico oggetto fisico da modellizzare. Quest'ultima molteplicità è legata alla maggiore o minore accuratezza con cui si decide di descrivere i fenomeni in gioco. Questo concetto è già presente nell'Esempio 2.7. A sua ulteriore illustrazione si consideri ancora il carrello dell'Esempio 2.5. Nel calcolo dell'energia totale lo si è considerato come un corpo rigido, trascurando quindi i contributi energetici associati ai movimenti relativi tra le singole parti, la cui descrizione avrebbe richiesto l'introduzione di altre variabili di stato. In termini generali, il modello matematico di un oggetto fisico deve essere individuato, fra tutti quelli possibili, evitando ridondanze e stabilendo un adeguato compromesso tra semplicità e precisione nella descrizione della realtà, in relazione al particolare problema in considerazione. La scelta della complessità del modello più opportuna è operazione tutt'altro che semplice, e richiede sempre una buona dimestichezza con il settore applicativo di riferimento, unita alla conoscenza delle tecniche di studio dei sistemi dinamici qui trattate.

2.3 Classificazione

I sistemi dinamici descritti dalle Equazioni (2.3), (2.4) possono essere classificati in vari modi sulla base delle proprietà delle funzioni f e g .

Sistemi monovariabili e multivariabili (SISO e MIMO) Si dicono *monovariabili* (o *SISO*, dall'inglese *Single-Input-Single-Output*) i sistemi dotati di una sola variabile di ingresso e di una sola variabile di uscita, cioè quelli per cui $m = p = 1$; si dicono *multivariabili* (o *MIMO*, dall'inglese *Multiple-Input-Multiple-Output*) gli altri.

Il circuito elettrico dell'Esempio 2.3 è un sistema monovariabile, così come il carrello dell'Esempio 2.5 e la centrifuga dell'Esempio 2.8; invece il motore dell'Esempio 2.6 e il forno dell'Esempio 2.7 sono sistemi multivariabili perché dotati di due variabili di ingresso (il primo di essi ha anche due variabili di uscita).

Sistemi propri, strettamente propri e non dinamici Se la funzione g non dipende dall'ingresso, cioè se la trasformazione d'uscita (2.4) si può scrivere nella forma

$$y(t) = g(x(t), t)$$

allora il sistema si dice *strettamente proprio*, o anche *puramente dinamico*, perché l'uscita dipende dall'ingresso non direttamente, ma solo attraverso lo stato. Invece in generale il sistema si dice *proprio*.

Come mostra la trasformazione d'uscita (2.6), il circuito elettrico dell'Esempio 2.3 è un sistema proprio, ma non strettamente; il carrello dell'Esempio 2.5, il motore dell'Esempio 2.6, il forno dell'Esempio 2.7 e la centrifuga dell'Esempio 2.8 sono invece sistemi strettamente propri.

Si osservi pure che un caso particolare di sistema proprio è quello di sistema *non dinamico*, o *statico*, per descrivere il cui comportamento non è necessario introdurre alcuna variabile di stato. In questo caso il legame ingresso-uscita è istantaneo e descritto dalla sola equazione

$$y(t) = g(u(t), t)$$

Di questo tipo è il circuito elettrico dell'Esempio 2.1. Il sistema è non dinamico perché i resistori dissipano energia, ma non ne accumulano.

Sistemi invarianti e varianti nel tempo Si parla di sistema *invariante nel tempo*, o anche *stazionario*, nel caso in cui le funzioni f e g non dipendano esplicitamente dal tempo, cioè risult

$$\dot{x}(t) = f(x(t), u(t)) \quad (2.26)$$

$$y(t) = g(x(t), u(t)) \quad (2.27)$$

mentre, se anche una sola delle funzioni f e g dipende esplicitamente da t , il sistema si dice *variante nel tempo*.

Tutti gli esempi introdotti finora sono relativi a sistemi stazionari, tranne il carrello dell'Esempio 2.5, che è variante nel tempo a causa della supposta dipendenza dal tempo della costante elastica.

È importante osservare che il movimento (dello stato e dell'uscita) di un sistema stazionario, cioè la sua risposta a una qualunque sollecitazione, costituita da uno stato iniziale e da una funzione di ingresso, non dipende dall'istante di applicazione della sollecitazione stessa. In altre parole, la stessa sollecitazione inviata in due istanti diversi t_1 e t_2 al medesimo sistema produce due movimenti che differiscono solo per una traslazione temporale pari alla differenza tra t_2 e t_1 . Pertanto, nello studio di questi sistemi è possibile assumere un qualunque valore del tempo come istante iniziale: la scelta più tipica è $t_0 = 0$.

Sistemi lineari e non lineari Quando le funzioni f e g sono lineari in x e u , cioè quando sia $\dot{x}(t)$ sia $y(t)$ sono combinazioni lineari delle varie componenti

dei vettori $x(t)$ e $u(t)$, allora il sistema (2.3), (2.4) si può scrivere nella forma

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

$$y(t) = C(t)x(t) + D(t)u(t)$$

dove le matrici $A(t) \in R^{n \times n}$, $B(t) \in R^{n \times m}$, $C(t) \in R^{p \times n}$ e $D(t) \in R^{p \times m}$ sono in generale funzioni del tempo. In questo caso il sistema dinamico si dice *lineare*, mentre, se le equazioni del sistema non assumono la forma sopra riportata, si parla di sistema *non lineare*. La matrice $A(t)$ viene detta *matrice della dinamica*; i suoi elementi e quelli delle matrici $B(t)$ e $C(t)$ assumono per la verità valori complessi nei casi in cui lo stato sia definito in campo complesso.

Il circuito elettrico dell'Esempio 2.3 è lineare e, in particolare, risulta $A = -1/RC$, $B = 1/RC$, $C = -1$ e $D = 1$. Il motore dell'Esempio 2.6 è lineare con

$$A = \begin{bmatrix} -\frac{R}{L} & -\frac{k}{L} \\ \frac{k}{J} & -\frac{h}{J} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & -\frac{1}{J} \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Il forno dell'Esempio 2.7 è anch'esso lineare e per il modello (2.18)-(2.20) si ha

$$A = \begin{bmatrix} -\frac{k_{ip}}{C_i} & \frac{k_{ip}}{C_i} \\ \frac{k_{ip}}{C_p} & -\frac{k_{ip} + k_{pe}}{C_p} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{C_i} & 0 \\ 0 & \frac{k_{pe}}{C_p} \end{bmatrix}, \quad C = [1 \ 0], \quad D = [0 \ 0]$$

Il carrello dell'Esempio 2.5 è non lineare, perché le equazioni di stato (2.11), (2.12) sono lineari, ma la trasformazione d'uscita (2.13) è quadratica nelle variabili di stato. La centrifuga dell'Esempio 2.8 è non lineare, perché l'equazione di stato (2.21) è quadratica e contiene la funzione $\text{sgn}(\alpha)$.

Un caso di particolare interesse La linearità è una proprietà di grande importanza per un sistema, specie quando è associata all'invarianza nel tempo, come accade per i sistemi degli Esempi 2.3, 2.6 e 2.7. Un sistema che gode di queste due proprietà è descritto dalle equazioni

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (2.28)$$

$$y(t) = Cx(t) + Du(t) \quad (2.29)$$

dove A , B , C e D sono matrici costanti. In questo caso il sistema è retto da un'equazione di stato e una trasformazione d'uscita lineari e a coefficienti costanti, ed è definito solo dagli elementi che compaiono nelle matrici in gioco. Queste particolarità fanno sì che lo studio dei sistemi lineari e invarianti nel tempo possa essere molto approfondito e semplice. A essi sono dedicati gli interi Capitoli 3-6. Inoltre, tutte le volte che sia possibile, si cerca di fornire al sistema in esame le due proprietà suddette mediante un procedimento di linearizzazione e di "congelamento" di eventuali parametri lentamente variabili nel tempo. La linearizzazione sarà trattata nel Paragrafo 3.5.1. Per quanto riguarda invece la possibilità di ritenere approssimativamente costanti i parametri che non lo sono, si consideri l'esempio seguente.

Esempio 2.9 (Seguito dell'Esempio 2.5) Si supponga di voler studiare il comportamento del carrello di Equazioni (2.11)-(2.13) in un intervallo di tempo $[t_0, t_f]$ piccolo rispetto alla costante di decadimento $1/\alpha$ della costante elastica della molla, per esempio $t_f - t_0 = 0.1/\alpha$. Allora, nell'intervallo di tempo di interesse i valori assunti dall'esponenziale nelle (2.12), (2.13) variano nell'intervallo $[\approx 0.9, 1]$ e hanno un valor medio di ≈ 0.95 . Il modello variante nel tempo (2.11)-(2.13) potrebbe quindi essere sostituito dal modello stazionario approssimato

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= \frac{1}{M}(u(t) - 0.95k_{t_0}x_1(t)) \\ y(t) &= \frac{1}{2}(0.95k_{t_0}x_1^2(t) + Mx_2^2(t)) \end{aligned}$$

2.4 Ritardo di tempo e sistemi a parametri distribuiti

In realtà, non tutti i sistemi dinamici sono descritti dalle Equazioni (2.3), (2.4). Tra tutti gli altri, che possono richiedere formulazioni matematiche anche molto complesse, vale la pena di segnalare uno largamente utilizzato come semplice modello di fenomeni di trasporto. Si tratta del cosiddetto *ritardo di tempo*: un sistema SISO la cui uscita replica l'ingresso ritardandolo di un tempo $\tau > 0$, come illustrato nella Figura 2.10. Esso è descritto dalla semplice relazione

$$y(t) = u(t - \tau) \quad (2.30)$$

che mette l'uscita in diretta relazione con l'ingresso.

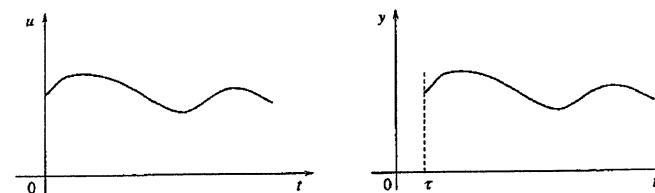


Figura 2.10
Ritardo di tempo.

Esempio 2.10 Il nastro trasportatore di Figura 2.11 è lungo L e si muove a velocità costante V . La portata eferente di materiale q_e , presa come uscita, dipende dalla portata aferente q_a , che è l'ingresso, secondo l'Equazione (2.30) dove $\tau = L/V$. Si osservi però che la portata eferente $q_e(t)$, $0 \leq t < \tau$, non dipende da $q_a(t)$, $t \geq 0$, ma dalla densità di materiale che si trova in ogni punto del nastro al tempo $t = 0$. Questo fatto non è descritto dall'Equazione (2.30).

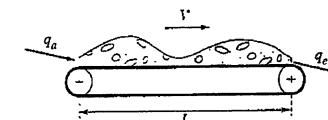


Figura 2.11
Nastro trasportatore
dell'Esempio 2.10.

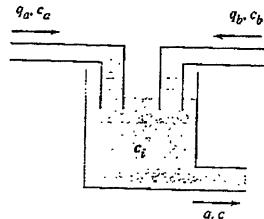
In termini generali, assumendo di conoscere l'ingresso per $t \geq 0$, l'Equazione (2.30) non determina l'uscita per $0 \leq t < \tau$, ma solo per $t \geq \tau$. La situazione interna al tempo $t = 0$, non è definita come nei sistemi (2.3), (2.4) assegnando un numero finito di scalari reali, cioè le n componenti del vettore di stato $x(0)$;

è definita invece da un'intera funzione di un'ulteriore variabile. Nell'esempio appena trattato, quest'ultima è una coordinata spaziale che corre da 0 a L lungo il nastro trasportatore, mentre come stato si può assumere la densità di materiale lungo il nastro. Se si volesse pervenire a una descrizione più completa del sistema di cui all'Equazione (2.30), si potrebbe allora constatare la necessità di introdurre un'equazione di stato costituita da un'equazione differenziale a derivate parziali, anziché totali come la (2.3). Sulla base di essa sarebbe agevole verificare la stazionarietà e la linearità del sistema, seppure in un senso allargato rispetto alle definizioni fornite. A ogni modo, fortunatamente la semplice Equazione (2.30) risulta sufficiente a descrivere quanto occorre in molte applicazioni di tipo ingegneristico.

I sistemi nei quali lo stato al tempo t è costituito da un'intera funzione di una o più variabili ed è retto da un'equazione di stato differenziale a derivate parziali si dicono *a dimensione infinita*, o *a parametri distribuiti*, in contrapposizione ai sistemi dinamici descritti da equazioni come le (2.3), (2.4), in cui lo stato al tempo t è definito da un numero finito n di scalari reali ed è retto un'equazione di stato differenziale a derivate totali, che si dicono *a dimensione finita*, o anche *a parametri concentrati*.

Esempio 2.11 Il serbatoio nella Figura 2.12 è alimentato da una portata q_a di liquido nel quale è disciolta una sostanza con concentrazione c_a e anche da una portata q_b della stessa soluzione, ma con concentrazione c_b ; q è la portata eferente. Per studiare la dipendenza da q_a , q_b , c_a e c_b della concentrazione c del fluido efferente, occorre assumere come variabile di stato almeno la concentrazione c_i della soluzione all'interno del serbatoio. Tuttavia, essa è esclusivamente funzione del tempo solo se si può assumere che la soluzione sia perfettamente mescolata. Altrimenti c_i è funzione anche del punto in cui la si valuta, cioè dipende dalle coordinate spaziali, e l'equazione differenziale che ne descrive il legame con gli ingressi è a derivate parziali.

Figura 2.12
Serbatoio dell'Esempio 2.11.



Osservazioni analoghe si possono fare a proposito delle temperature che intervengono nell'Esempio 2.7.

2.5 Equilibrio

Per i sistemi stazionari di Equazioni (2.26), (2.27) soggetti a ingressi costanti $u(t) = \bar{u}$, sono di particolare importanza quei movimenti dello stato e dell'uscita che risultano anch'essi costanti nel tempo. Questi movimenti sono detti, rispettivamente, *stati e uscite di equilibrio*. Gli stati di equilibrio devono soddisfare l'equazione $\dot{x}(t) = 0$, cioè sono le soluzioni \bar{x} costanti nel tempo dell'equazione

$$f(\bar{x}, \bar{u}) = 0 \quad (2.31)$$

Quest'equazione può avere una sola soluzione, può averne tante, oppure nessuna. In ogni caso, a ciascuna di esse corrisponde un'uscita di equilibrio \bar{y} calcolabile

mediante la relazione

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

Uno stato di equilibrio è dunque uno stato in cui un sistema, sollecitato da un ingresso costante, permane indefinitamente se vi si trova in un qualunque istante di tempo. In altri termini, se al tempo \bar{t} lo stato del sistema è $x(\bar{t}) = \bar{x}$, allora risulta $x(t) = \bar{x}$ anche per ogni $t > \bar{t}$; quindi, l'uscita rimane costante al valore di equilibrio \bar{y} .

Esempio 2.12 (Seguito dell'Esempio 2.3) Se $u(t) = \bar{u} = U$, lo stato di equilibrio ottenuto annullando $\dot{x} = \bar{x} = U$, cui corrisponde l'uscita di equilibrio $\bar{y} = 0$. L'ovvia interpretazione di questo risultato è che, quando tutte le variabili in gioco nel circuito sono costanti, il condensatore è un circuito aperto e pertanto la corrente nella maglia e la tensione ai morsetti del resistore sono nulle, mentre la tensione ai morsetti del condensatore coincide con quella ai morsetti del generatore.

Esempio 2.13 (Seguito dell'Esempio 2.8) Se la coppia motrice è nulla, cioè $u(t) = 0$, all'equilibrio si ha $\bar{y} = \bar{x} = 0$.

Se invece la coppia motrice è $u(t) = M > 0$, all'equilibrio risulta

$$-\frac{k}{J}\bar{x}^2 \operatorname{sgn}(\bar{x}) + \frac{M}{J} = 0$$

cioè

$$\bar{y} = \bar{x} = \sqrt{\frac{M}{k}} = h$$

Esempio 2.14 Si consideri il sistema dinamico con equazione di stato

$$\dot{x}(t) = x^2(t) + x(t) + u(t)$$

Esso ha due stati di equilibrio per $u(t) = \bar{u} < 1/4$, ne ha uno (doppio) per $u(t) = \bar{u} = 1/4$, mentre non ne ha nessuno per $u(t) = \bar{u} > 1/4$.

Esempio 2.15 Il sistema dinamico con equazioni di stato

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= u(t) \end{aligned}$$

può descrivere il moto rettilineo di un corpo rigido di massa unitaria, con posizione x_1 e velocità x_2 , soggetto a una forza u . Quando $u(t) = \bar{u} \neq 0$, non vi sono stati di equilibrio perché $\dot{x}_2(t) \neq 0$. Invece quando $u(t) = \bar{u} = 0$, all'equilibrio risulta

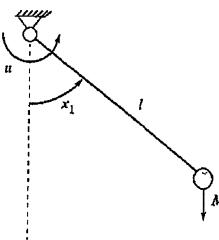
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \text{qualsiasi} \\ 0 \end{bmatrix} \quad (2.32)$$

Gli stati di equilibrio sono dunque infiniti e di interpretazione immediata dal punto di vista della meccanica.

Esempio 2.16 Il pendolo di Figura 2.13 è costituito da un corpo puntiforme di massa M collegato mediante un'asta rigida di lunghezza l e massa trascurabile a una cerniera; esso si muove in un piano verticale. L'ingresso u è costituito da una coppia motrice applicata e l'uscita y dall'energia totale riferita al livello della cerniera. Indicando con x_1 la posizione angolare rispetto alla verticale e con x_2 la velocità angolare, e supponendo che nella cerniera agisca una coppia d'attrito proporzionale alla velocità angolare secondo un coefficiente d'attrito $k > 0$, in base alle leggi della dinamica le equazioni che reggono il comportamento del sistema, che risulta stazionario, sono

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -\frac{g}{l} \sin(x_1(t)) - \frac{k}{Ml^2}x_2(t) + \frac{1}{Ml^2}u(t) \\ y(t) &= \frac{1}{2}Ml^2x_2^2(t) - Mgl \cos(x_1(t)) \end{aligned}$$

Figura 2.13
Pendolo dell'Esempio 2.16.



dove g è l'accelerazione di gravità. Per

$$u(t) = \ddot{u} = Mg/l \quad (2.33)$$

all'equilibrio si ha

$$0 = \ddot{x}_2$$

$$0 = -\frac{g}{l} \sin(\bar{x}_1) - \frac{k}{Ml^2} \bar{x}_2 + \frac{g}{l}$$

$$\ddot{y} = \frac{1}{2} Ml^2 \ddot{x}_2^2 - Mg/l \cos(\bar{x}_1)$$

da cui

$$\ddot{x} = \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} \pi/2 + 2\pi i \\ 0 \end{bmatrix} \quad i \text{ intero} \quad (2.34)$$

$$\ddot{y} = 0 \quad (2.35)$$

Esistono, quindi, infiniti stati di equilibrio, tutti dall'ovvia identica interpretazione fisica: i momenti della coppia u e di quella prodotta dalla forza di gravità si egualano e il pendolo rimane fermo in posizione orizzontale a destra rispetto alla cerniera. Se invece

$$u(t) = \ddot{u} = 0 \quad (2.36)$$

all'equilibrio si ha

$$0 = \ddot{x}_2$$

$$0 = -\frac{g}{l} \sin(\bar{x}_1) - \frac{k}{Ml^2} \bar{x}_2$$

$$\ddot{y} = \frac{1}{2} Ml^2 \ddot{x}_2^2 - Mg/l \cos(\bar{x}_1)$$

da cui

$$\ddot{x} = \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} \pi i \\ 0 \end{bmatrix} \quad i \text{ intero} \quad (2.37)$$

$$\ddot{y} = (-1)^{i+1} Mg/l \quad i \text{ intero} \quad (2.38)$$

In questa situazione esistono stati di equilibrio di due tipi diversi: il pendolo è comunque fermo, ma per valori pari di i è orientato verso il basso, mentre per valori dispari di i è orientato verso l'alto.

Per chiudere, si osservi che il concetto di uscita di equilibrio si può estendere anche al ritardo di tempo (2.30), ove si conclude che risulta

$$\ddot{y} = \ddot{u}$$

Nel caso dell'Esempio 2.10 il concetto di equilibrio implica che la densità di materiale lungo il nastro ("stato" del sistema) sia costante nel tempo. Ragionando in termini intuitivi, si può poi comprendere come essa risulti identica in ogni punto del nastro e coerente con le portate afferente ed efferente, costanti e uguali tra loro.

2.6 Stabilità

In questo paragrafo si introduce la nozione di *stabilità*, proposta alla fine del XIX secolo dal matematico russo A.M. Liapunov, che considera le conseguenze sul movimento di un sistema di un'incertezza sul valore iniziale del suo stato, nell'ipotesi che gli ingressi siano fissi e noti. In particolare, essa sostanzialmente richiede che "piccole" perturbazioni dello stato iniziale rispetto a un valore di riferimento provochino solo "piccole" perturbazioni del movimento dello stato, eventualmente destinate ad annullarsi su tempi lunghi.

La stabilità risulta proprietà di grande interesse, perché lo stato iniziale di un sistema di solito non è né noto, né misurabile. Inoltre, può accadere che sul sistema intervengano perturbazioni di breve durata, anche se di intensità elevata, e quando ciò accade è molto difficile studiare il comportamento del sistema durante l'azione delle perturbazioni stesse, mentre è relativamente facile farlo a partire dall'istante in cui esse scompaiono se si sanno trattare incertezze sullo stato, che, in quell'istante, sarà ignoto.

Per quanto affermato, è chiaro che la trasformazione d'uscita del sistema non svolge alcun ruolo in questo contesto e quindi non sarà considerata. Inoltre, la trattazione che segue è riferita ai soli sistemi stazionari di equazione di stato (2.26) con istante iniziale $t_0 = 0$.

2.6.1 Stabilità dell'equilibrio

Per un sistema dinamico invariante nel tempo, si considerino un ingresso costante $u(t) = \ddot{u}$, $t \geq 0$, e un corrispondente stato di equilibrio \ddot{x} , detto *nominales*. Si consideri anche un movimento dello stato $x(t)$, detto *perturbato*, generato a partire ancora da \ddot{u} , ma da uno stato iniziale x_0 (in generale diverso da \ddot{x}). Si può allora dare la seguente definizione.

Definizione 2.1 Uno stato di equilibrio \ddot{x} si dice *stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \ddot{x}\| \leq \delta$$

risulti

$$\|x(t) - \ddot{x}\| \leq \varepsilon$$

per tutti $t \geq 0$.

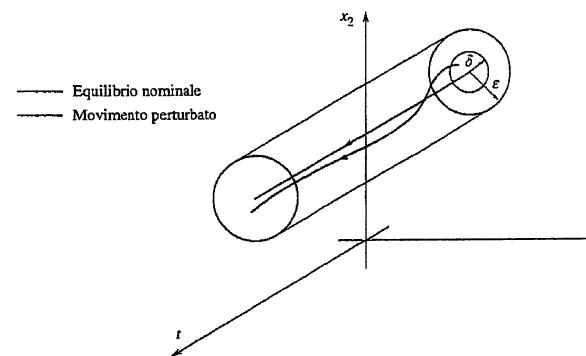
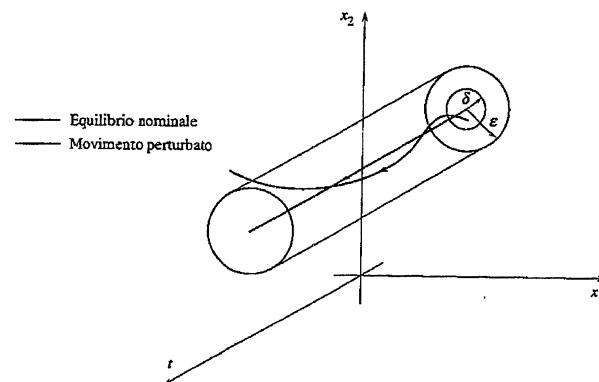


Figura 2.14
Equilibrio stabile.

La proprietà di stabilità dell'equilibrio richiede quindi che il movimento perturbato rimanga "vicino" all'equilibrio nominale, come illustrato nella Figura 2.14 per un sistema di ordine 2. Più precisamente, scelta arbitrariamente piccola la massima distanza accettabile in un qualunque istante di tempo tra il movimento perturbato e l'equilibrio nominale, quest'ultimo è stabile se la condizione su tale distanza è rispettata, pur di prendere lo stato iniziale del movimento perturbato sufficientemente prossimo all'equilibrio nominale. Se ciò non accade l'equilibrio si dice instabile come specificato qui di seguito e illustrato nella Figura 2.15.

Definizione 2.2 Uno stato di equilibrio \bar{x} si dice *instabile* se non è stabile. ■

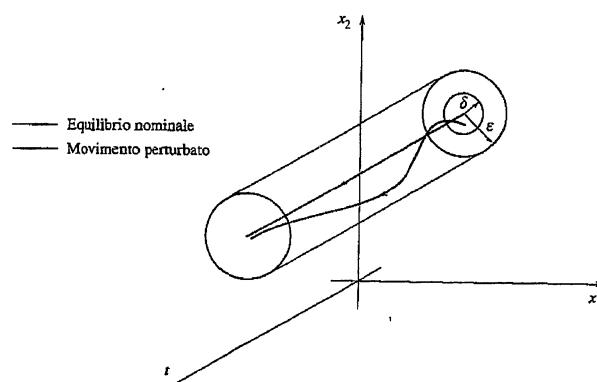
Figura 2.15
Equilibrio instabile.



L'instabilità di uno stato di equilibrio implica perciò che esistono perturbazioni arbitrariamente piccole dello stato iniziale che provocano l'allontanamento dello stato del sistema dall'equilibrio stesso.

La proprietà di stabilità può essere rafforzata richiedendo anche che il movimento perturbato tenda all'equilibrio nominale per $t \rightarrow \infty$. Ciò è specificato dalla seguente definizione di stabilità asintotica, che costituisce la proprietà di maggiore importanza, concettuale e nella pratica (Figura 2.16).

Figura 2.16
Equilibrio asintoticamente stabile.



Definizione 2.3 Uno stato di equilibrio \bar{x} si dice *asintoticamente stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \bar{x}\| \leq \delta$$

risulti

$$\|x(t) - \bar{x}\| \leq \varepsilon$$

per tutti i $t \geq 0$, e inoltre

$$\lim_{t \rightarrow \infty} \|x(t) - \bar{x}\| = 0$$

Esempio 2.17 (Seguito degli Esempi 2.2, 2.3 e 2.12) Come già detto, per $u(t) = U$ all'equilibrio si ha $\bar{x} = U$. Dall'equazione di stato (2.5) si deduce che il movimento perturbato relativo allo stato iniziale $x(0) = x_0$ è

$$x(t) = e^{-t/RC} x_0 + (1 - e^{-t/RC}) U$$

Pertanto

$$x(t) - \bar{x} = e^{-t/RC} (x_0 - \bar{x})$$

e quindi, se si desidera che risulti

$$|x(t) - \bar{x}| \leq \varepsilon$$

basta prendere $\delta = \varepsilon$, visto che per $t \geq 0$

$$|e^{-t/RC} (x_0 - \bar{x})| \leq |x_0 - \bar{x}|$$

Lo stato di equilibrio \bar{x} è pertanto stabile. Inoltre si può anche affermare che esso è asintoticamente stabile perché

$$\lim_{t \rightarrow \infty} e^{-t/RC} = 0$$

Esempio 2.18 (Seguito degli Esempi 2.8 e 2.13) Si è già visto che, per $u(t) = M > 0$, l'unico stato di equilibrio della centrifuga è $\bar{x} = h$. Per determinare le proprietà di stabilità di questo stato di equilibrio, invece di studiare il movimento (2.24), (2.25), conviene osservare la Figura 2.17 in cui è riportato l'andamento di \dot{x} in funzione di x . Si nota che la derivata dello stato è positiva per tutti i valori di x minori di h , mentre è negativa per tutti i valori di x maggiori di h . Ciò significa che, qualunque sia lo stato iniziale del sistema, la distanza tra lo stato del sistema e lo stato di equilibrio diminuisce all'aumentare del tempo e tende asintoticamente a zero. Pertanto $\bar{x} = h$ è uno stato di equilibrio asintoticamente stabile.

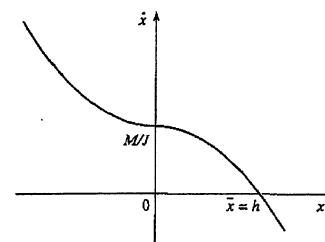
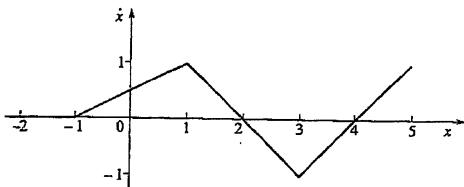


Figura 2.17
Relazione tra stato e sua derivata per la centrifuga degli Esempi 2.8, 2.13 e 2.18.

Regione di attrazione Gli stati di equilibrio considerati nei due ultimi esempi non solo godono della proprietà di asintotica stabilità, ma in più sono, come si dice, *globalmente stabili*: i movimenti perturbati generati da un qualunque stato iniziale, "vicino" allo stato di equilibrio nominale o "lontano" da esso, convergono tutti al medesimo equilibrio nominale. In altri termini, l'ampiezza della perturbazione è irrilevante. Ciò non costituisce certo il caso generale. In corrispondenza di un certo ingresso fissato, il sistema potrebbe essere dotato di più stati di equilibrio con differenti proprietà di stabilità. Per ognuno di quelli asintoticamente stabili esisterà poi un insieme di stati iniziali, che si dicono costituire la *regione di attrazione* dello stato di equilibrio, che generano movimenti perturbati convergenti asintoticamente allo stato di equilibrio stesso.

Esempio 2.19 Gli stati di equilibrio del sistema dinamico del primo ordine la cui relazione tra \dot{x} e x è riportata nella Figura 2.18 sono $\bar{x} \leq -1$, $\bar{x} = 2$ e $\bar{x} = 4$. Ragionando come fatto nell'Esempio 2.18 sulla base del segno di \dot{x} , si nota che lo stato del sistema tende a crescere quando è compreso tra -1 e 2 o è maggiore di 4 , mentre tende a diminuire quando è compreso tra 2 e 4 . Pertanto gli stati di equilibrio $\bar{x} < -1$ sono stabili, non asintoticamente, perché ogni piccola perturbazione dello stato iniziale porta comunque in uno stato di equilibrio diverso. Lo stato $\bar{x} = -1$ è instabile in quanto non è possibile trovare un intorno proprio e simmetrico rispetto a esso di stati iniziali che generino movimenti che non si allontanino da \bar{x} stesso. La stessa considerazione vale per $\bar{x} = 4$. Invece, $\bar{x} = 2$ è uno stato di equilibrio asintoticamente stabile con regione di attrazione costituita dall'intervallo aperto $(-1, 4)$, perché tutti i movimenti generati da stati iniziali all'interno di questo intervallo convergono asintoticamente proprio a $\bar{x} = 2$.

Figura 2.18
Relazione tra stato e sua derivata per il sistema dell'Esempio 2.19.



Determinazione delle proprietà di stabilità Lo studio della stabilità di uno stato di equilibrio è in generale un problema non banale. A volte possono essere d'aiuto considerazioni di natura fisica, come nel caso seguente,

Esempio 2.20 (Seguito dell'Esempio 2.16) Ragionando in maniera puramente intuitiva sugli stati di equilibrio (2.37), si comprende che quelli caratterizzati da valori dispari di i , cioè quelli in cui il pendolo è fermo e orientato verso l'alto, sono instabili, in quanto per poco che si perturbi la posizione o la velocità, il pendolo tende ad allontanarsi dalla condizione di equilibrio di partenza. D'altra parte gli stati di equilibrio caratterizzati da valori pari di i , cioè quelli in cui il pendolo è fermo e orientato verso il basso, sono asintoticamente stabili, in quanto, perturbando leggermente la posizione e la velocità si instaurano movimenti che, a causa dell'attrito presente, tendono allo stato di equilibrio considerato. Naturalmente il fatto stesso che esistano più stati di equilibrio impedisce che quelli asintoticamente stabili lo siano globalmente.

Analizzare rigorosamente questo esempio, a questo punto, è però impossibile, in quanto non sono ancora stati introdotti gli strumenti matematici necessari. Infatti l'applicazione diretta delle Definizioni 2.1-2.3 è quasi sempre una via impraticabile, perché richiede il calcolo dei movimenti perturbati, cosa di solito parecchio complessa, mentre la tecnica consistente nello studiare l'andamento di \dot{x} in funzione di x (Esempi 2.18 e 2.19) è applicabile solo a sistemi del primo ordine. La

determinazione, o anche solo la stima, delle regioni di attrazione è poi ancora meno semplice. Fortunatamente, per classi importanti di sistemi dinamici esistono metodi generali di analisi che saranno introdotti nel seguito.

2.6.2 Stabilità del movimento

Le considerazioni sulla stabilità dell'equilibrio appena svolte possono essere generalizzate sostituendo all'equilibrio nominale di riferimento un generico movimento nominale. In proposito, qui ci si limita a mostrare la nuova, più generale, forma presa dalle Definizioni 2.1-2.3.

Per un sistema dinamico invariante nel tempo si considerino un ingresso $\bar{u}(t)$, $t \geq 0$, uno stato iniziale \bar{x}_0 e il movimento dello stato $\bar{x}(t)$, detto *nominale*, da essi prodotto. Si consideri anche un secondo movimento dello stato $x(t)$, detto *perturbato*, generato a partire ancora da $\bar{u}(t)$, ma da uno stato iniziale x_0 (in generale diverso da \bar{x}_0). Si può allora dare la seguente definizione.

Definizione 2.4 Un movimento $\bar{x}(t)$ si dice *stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \bar{x}_0\| \leq \delta$$

risulti

$$\|x(t) - \bar{x}(t)\| \leq \varepsilon$$

per tutti $t \geq 0$.

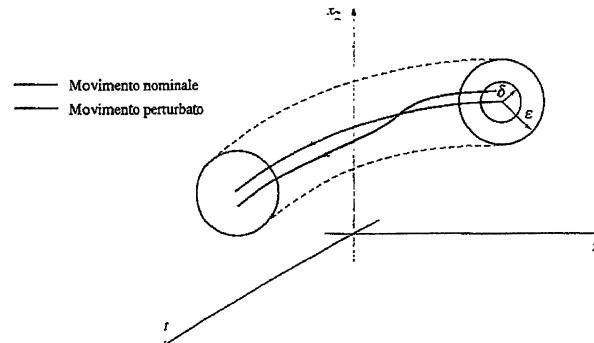


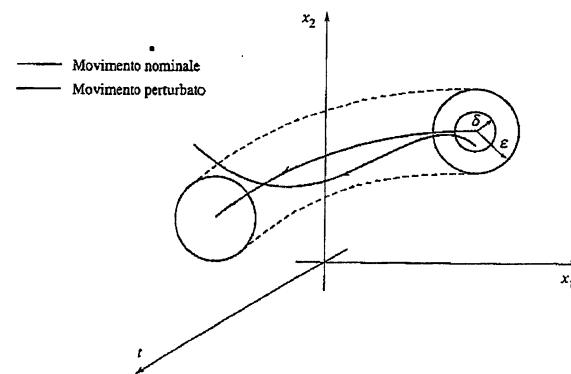
Figura 2.19
Movimento stabile.

La proprietà di stabilità del movimento richiede quindi che il movimento perturbato sia "vicino" al movimento nominale (Figura 2.19). Più precisamente, scelta arbitrariamente piccola la massima distanza accettabile in un qualunque istante di tempo tra il movimento perturbato e quello nominale, quest'ultimo è stabile se la condizione su tale distanza è rispettata, pur di prendere lo stato iniziale del movimento perturbato sufficientemente prossimo a quello del movimento nominale. Se ciò non accade il movimento si dice instabile come specificato qui di seguito (Figura 2.20).

Definizione 2.5 Un movimento $\bar{x}(t)$ si dice *instabile* se non è stabile.

La proprietà di stabilità può essere rafforzata richiedendo anche che il movimento perturbato tenda al movimento nominale per $t \rightarrow \infty$. Ciò è specificato dalla seguente definizione di stabilità asintotica (Figura 2.21).

Figura 2.20
Movimento instabile.



Definizione 2.6 Un movimento $\tilde{x}(t)$ si dice *asintoticamente stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \tilde{x}_0\| \leq \delta$$

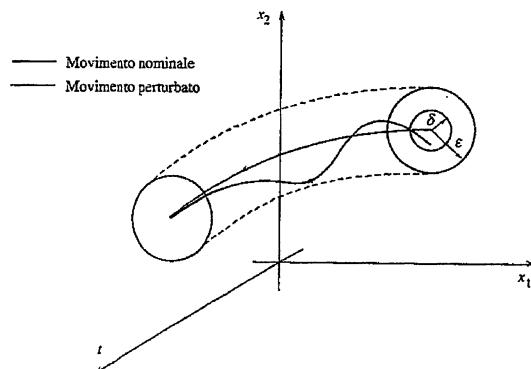
risulti

$$\|x(t) - \tilde{x}(t)\| \leq \varepsilon$$

per tutti $t \geq 0$, e inoltre

$$\lim_{t \rightarrow \infty} \|x(t) - \tilde{x}(t)\| = 0$$

Figura 2.21
Movimento
asintoticamente stabile.



2.7 Conclusioni

In questo capitolo sono stati introdotti i sistemi dinamici a tempo continuo. Questa classe di modelli matematici ben si presta a descrivere un'ampia varietà di

fenomeni per i quali sostanzialmente i legami tra le diverse variabili sono espressi nella forma di equazioni differenziali. Dopo avere illustrato il significato delle variabili di ingresso, di stato e di uscita, sono state introdotte l'equazione di stato e la trasformazione d'uscita. Quindi si sono discussi vari criteri di classificazione e il concetto di equilibrio. Infine si è dedicata attenzione alla fondamentale nozione di stabilità. I Capitoli 3-6 saranno dedicati allo studio dettagliato dei sistemi che, oltre a essere lineari, godono anche della proprietà di invarianza nel tempo. Questi sistemi infatti, grazie alle particolari proprietà che li contraddistinguono, sono quelli di maggiore importanza nella teoria del controllo e in gran parte delle applicazioni ingegneristiche.

Esercizi

Esercizio 2.1 Si consideri il sistema

$$\dot{x}_1(t) = \alpha x_2(t) \quad (2.39)$$

$$\dot{x}_2(t) = \beta x_1(t) + \gamma u(t) \quad (2.40)$$

$$y(t) = \delta x_1(t) \quad (2.41)$$

Si determinino un circuito elettrico e un sistema meccanico massa-molla che si lascino descrivere dalle equazioni sopra riportate per opportuni valori dei parametri α, β, γ e δ .

(*Suggerimento:* si osservi che le Equazioni (2.39)-(2.41) sono equivalenti all'unica equazione differenziale del secondo ordine $\ddot{y}(t) - \alpha\beta y(t) = \gamma\delta u(t)$).

Esercizio 2.2 Si scriva il sistema dinamico che descrive un circuito elettrico costituito da un generatore di tensione, un resistore, un condensatore e un induttore in serie, assumendo come variabile di ingresso la tensione impressa dal generatore e come variabile di uscita la corrente nella maglia. Come variabili di stato si prendano dapprima la tensione ai morsetti del condensatore e la corrente; quindi si mostri che, in alternativa, si possono assumere come variabili di stato anche la corrente e la sua derivata.

Infine si determini un sistema meccanico analogo del circuito elettrico, cioè descritto da equazioni della stessa forma.

Esercizio 2.3 Si voglia scrivere il modello di un circuito elettrico nel quale compaiono due condensatori in parallelo. Si rifletta sull'eventuale opportunità di assumere tra le variabili di stato le due tensioni ai loro morsetti. Si consideri poi il problema che si pone, in maniera del tutto analoga, se al circuito appartengono due induttori in serie.

Esercizio 2.4 Si determinino gli stati e le uscite di equilibrio del sistema

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

$$\dot{x}_2(t) = \cos(u(t)x_1(t))$$

$$y(t) = (x_1(t) + x_2(t))^2$$

per $u(t) = \bar{u} = 3$.

Esercizio 2.5 Si calcoli il movimento dello stato del sistema dell'Esempio 2.15 per $x_1(0) = x_{10}$, $x_2(0) = x_{20}$ e $u(t) = \sin(\omega t)$.

Esercizio 2.6 Si determinino le proprietà di stabilità degli stati di equilibrio del sistema dell'Esempio 2.14 al variare di $\bar{u} \leq 1/4$.

Esercizio 2.7 Si consideri un sistema con equazione di stato

$$\dot{x}(t) = u(t)x^3(t)$$

Per $u(t) = \bar{u} = 1$ e per $u(t) = \bar{u} = -1$, si determinino gli stati di equilibrio e le corrispondenti proprietà di stabilità.

Esercizio 2.8 Si consideri un sistema con equazione di stato

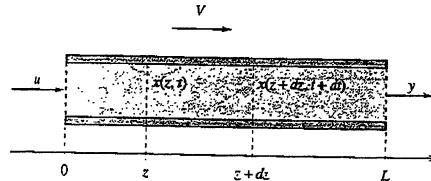
$$\dot{x}(t) = |x(t)| + u(t)$$

Si calcoli il movimento dello stato corrispondente a $\bar{u} = 0$ e $x(0) = x_0$. Poi, per $u(t) = \bar{u}$, si determinino gli stati di equilibrio e le corrispondenti proprietà di stabilità.

Esercizio 2.9 Nella meccanica a volte si usa dire che una pallina (massa puntiforme) ferma in un piano orizzontale è in una posizione di "equilibrio indifferente". Si analizzi la situazione indicata alla luce delle Definizioni 2.1-2.3, facendo riferimento al sistema dell'Esempio 2.15 con $\ddot{u} = 0$.

Esercizio 2.10 La Figura 2.22 mostra una condotta cilindrica, di area di base S e lunghezza L , attraversata da una soluzione incompressibile a velocità V ; l'ingresso u è la concentrazione di soluto all'inizio della condotta, l'uscita y quella al termine. Si considerino due sezioni a distanza z e $z + dz$ dall'origine della condotta e due istanti di tempo t e $t + dt$.

Figura 2.22
Condotta percorsa
da una soluzione
incompressibile
dell'Esercizio 2.10.



Per la sua legge di conservazione, la massa totale di soluto nel volume compreso tra le sezioni z e $z + dz$ al tempo $t + dt$ deve risultare pari a quella presente al tempo t , più quella entrata attraverso la sezione z , meno quella uscita dalla sezione $z + dz$. Allora, detta $x(z, t)$ la concentrazione di soluto alla sezione z e al tempo t , si può scrivere

$$x(z, t + dt) S dz = x(z, t) S dz + x(z, t) SV dt - x(z + dz, t) SV dt$$

Da qui, dividendo per $S dz dt$ e facendo tendere a zero dz e dt , si ottiene

$$\frac{\partial x(z, t)}{\partial t} = -V \frac{\partial x(z, t)}{\partial z} \quad (2.42)$$

Inoltre l'uscita è data da

$$y(t) = x(L, t) \quad (2.43)$$

Si verifichi che la soluzione dell'"equazione di stato" (2.42), integrata con condizione iniziale $x(z, 0) = x_0(z)$ e condizione al contorno $x(0, t) = u(t)$, è costituita dalla funzione

$$x(z, t) = \begin{cases} x_0(z - Vt) & 0 \leq t < z/V \\ u(t - z/V) & t \geq z/V \end{cases}$$

Da qui e dalla "trasformazione d'uscita" (2.43), si ricavi la (2.30) per $t \geq \tau = L/V$. Si determini pure l'andamento dell'uscita per $0 \leq t < \tau$.

Sistemi lineari e stazionari a tempo continuo

3

3.1 Introduzione

Questo capitolo e i tre successivi sono dedicati alla presentazione degli elementi fondamentali della teoria dei sistemi dinamici a tempo continuo che godono anche delle proprietà di linearità e stazionarietà. Essi infatti, come anticipato in parte al Paragrafo 2.3, rivestono un'importanza del tutto particolare in quanto sono utilizzati in molti casi per descrivere in maniera approssimata anche il comportamento di sistemi non lineari e varianti nel tempo. D'altra parte, le proprietà che contraddistinguono questi sistemi sono così forti da rendere estremamente consigliabile, tutte le volte in cui ciò sia possibile, il riferirsi a essi nell'affrontare problemi di qualunque natura.

Gli argomenti trattati in questo capitolo sono i seguenti:

- le espressioni esplicite dei movimenti dello stato e dell'uscita, con l'introduzione del concetto di modo e del fondamentale principio di sovrapposizione degli effetti;
- le espressioni esplicite degli stati e delle uscite di equilibrio;
- la formulazione di semplici condizioni necessarie e sufficienti di stabilità asintotica;
- l'applicazione dei risultati allo studio approssimato del movimento e all'analisi della stabilità dell'equilibrio di sistemi non lineari, effettuati mediante la procedura della linearizzazione;
- due interessanti nuove proprietà strutturali dei sistemi dinamici, che conducono alla definizione di una scomposizione in parti dei sistemi stessi.

3.2 Movimento

Si consideri il sistema a tempo continuo lineare, invariante nel tempo e proprio descritto da

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3.1)$$

$$y(t) = Cx(t) + Du(t) \quad (3.2)$$

dove $u \in R^m$, $x \in R^n$ e $y \in R^p$, mentre le matrici A , B , C e D sono reali, costanti e di dimensioni opportune. Si assuma che l'istante iniziale sia t_0 , anche se, per quanto affermato al Paragrafo 2.3, lo studio si potrebbe effettuare con riferimento al caso $t_0 = 0$ senza perdere la generalità.

3.2.1 Formula di Lagrange

È facile verificare per semplice sostituzione che il movimento dello stato corrispondente all'ingresso $u(t)$, definito per $t \geq t_0$, e allo stato iniziale $x(t_0) = x_{t_0}$ è dato dalla seguente espressione, detta *formula di Lagrange*:

$$x(t) = e^{A(t-t_0)}x_{t_0} + \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau, \quad t \geq t_0 \quad (3.3)$$

Il corrispondente movimento dell'uscita è

$$y(t) = Ce^{A(t-t_0)}x_{t_0} + C \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t), \quad t \geq t_0 \quad (3.4)$$

Per la definizione dell'esponenziale di matrice e^{At} si rimanda al Paragrafo A.5. Dalle Equazioni (3.3), (3.4) si possono trarre importanti indicazioni sulle caratteristiche dei movimenti dei sistemi corrispondenti.

3.2.2 Movimento libero e movimento forzato

Nei movimenti (3.3), (3.4) dello stato e dell'uscita del sistema (3.1), (3.2) si può individuare un contributo dipendente solo dallo stato iniziale e uno dipendente solo dall'ingresso, dai quali il movimento complessivo si ottiene per semplice somma.

Il contributo al movimento dello stato e dell'uscita funzione solo dello stato iniziale, cioè quello che si avrebbe se, a pari stato iniziale, l'ingresso fosse nullo, si chiama *movimento libero*, ed è dato da

$$x_f(t) = e^{A(t-t_0)}x_{t_0} \quad (3.5)$$

$$y_f(t) = Ce^{A(t-t_0)}x_{t_0} \quad (3.6)$$

mentre il contributo funzione solo dell'ingresso, cioè quello che si avrebbe se, a pari ingresso, lo stato iniziale fosse nullo, si chiama *movimento forzato*, ed è dato da

$$x_f(t) = \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau \quad (3.7)$$

$$y_f(t) = C \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t) \quad (3.8)$$

Esempio 3.1 (Seguito degli Esempi 2.2-2.4) Il sistema è lineare e quindi nelle Equazioni (2.7)-(2.9) si individuano i movimenti liberi e quelli forzati: i primi dipendono solo da $x(0) = x_0$ e si ottengono ponendo $U = 0$, cioè $u(t) = 0$; i secondi dipendono solo da u e si ottengono ponendo $x_0 = 0$; sia gli uni sia gli altri sono lineari nelle rispettive cause, cioè, per esempio, raddoppiando al raddoppiare di x_0 e U .

Esempio 3.2 (Seguito dell'Esempio 2.8) Nel movimento della centrifuga (2.24), (2.25) ottenuto con ingresso non nullo non è individuabile in alcun modo il contributo (2.23) che si ha per ingresso nullo. Infatti il sistema (2.21), (2.22) non è lineare e per esso non ha alcun senso parlare di movimento libero e movimento forzato.

3.2.3 Principio di sovrapposizione degli effetti

Per il sistema (3.1), (3.2), si supponga che in corrispondenza dell'ingresso u' e dello stato iniziale x'_{t_0} si ottengano i movimenti dello stato e dell'uscita x' e y' , mentre in corrispondenza dell'ingresso u'' e dello stato iniziale x''_{t_0} si ottengano i movimenti dello stato e dell'uscita x'' e y'' . Allora, sono soddisfatte le relazioni

$$x'(t) = e^{A(t-t_0)}x'_{t_0} + \int_{t_0}^t e^{A(t-\tau)}Bu'(\tau)d\tau, \quad t \geq t_0$$

$$y'(t) = Ce^{A(t-t_0)}x'_{t_0} + C \int_{t_0}^t e^{A(t-\tau)}Bu'(\tau)d\tau + Du'(t), \quad t \geq t_0$$

$$x''(t) = e^{A(t-t_0)}x''_{t_0} + \int_{t_0}^t e^{A(t-\tau)}Bu''(\tau)d\tau, \quad t \geq t_0$$

$$y''(t) = Ce^{A(t-t_0)}x''_{t_0} + C \int_{t_0}^t e^{A(t-\tau)}Bu''(\tau)d\tau + Du''(t), \quad t \geq t_0$$

Si consideri ora una terza situazione in cui l'ingresso u''' e lo stato iniziale x'''_{t_0} siano ambedue costituiti dalla stessa combinazione lineare degli ingressi e degli stati iniziali già considerati, cioè si assuma che esistano due scalari qualunque α e β tali che

$$u'''(t) = \alpha u'(t) + \beta u''(t) \quad (3.9)$$

$$x'''_{t_0} = \alpha x'_{t_0} + \beta x''_{t_0} \quad (3.10)$$

Per sostituzione si può verificare che anche i movimenti dello stato e dell'uscita x''' e y''' prodotti dalla coppia (3.9), (3.10) sono dati dalla medesima combinazione: infatti i movimenti

$$x'''(t) = \alpha x'(t) + \beta x''(t), \quad t \geq t_0$$

e

$$y'''(t) = \alpha y'(t) + \beta y''(t), \quad t \geq t_0$$

soddisfano identicamente le Equazioni (3.3), (3.4). Quanto visto sopra costituisce la dimostrazione del seguente risultato.

Teorema 3.1 Considerato il sistema lineare e stazionario (3.1), (3.2) con istante iniziale t_0 , siano x' e y' i movimenti dello stato e dell'uscita generati dall'ingresso u' e dallo stato iniziale x'_{t_0} , e x'' e y'' i movimenti dello stato e dell'uscita generati dall'ingresso u'' e dallo stato iniziale x''_{t_0} . Allora, per ogni coppia di scalari α e β , i movimenti dello stato x''' e dell'uscita y''' generati dall'ingresso

$$u'''(t) = \alpha u'(t) + \beta u''(t)$$

e dallo stato iniziale

sono

$$x'''(t) = \alpha x'(t) + \beta x''(t)$$

$$y'''(t) = \alpha y'(t) + \beta y''(t)$$

Pertanto, in particolare, il comportamento dei sistemi in oggetto per "grandi perturbazioni" dello stato iniziale e dell'ingresso differisce da quello relativo a "piccole perturbazioni" solo per un fattore di scala. Si ricordi a questo proposito quanto notato alla fine dell'Esempio 3.1.

Il Teorema 3.1, introdotto qui per sistemi descritti dalle Equazioni (3.1), (3.2), va sotto il nome di *principio di sovrapposizione degli effetti* ed è di grande importanza in quanto consente di calcolare il movimento generato da più cause, cioè stati iniziali e ingressi, semplicemente come somma pesata dei singoli effetti provocati dalle cause suddette. In realtà, si può dimostrare che questo risultato vale indipendentemente dall'ipotesi di stazionarietà del sistema cui lo si applica, mentre è legato in maniera indissolubile all'ipotesi di linearità, tanto che la classe dei sistemi lineari potrebbe essere definita come costituita da tutti e soli quei sistemi per i quali è possibile parlare di effetto di una variabile (stato iniziale o ingresso) su un'altra (stato o uscita).

3.2.4 Rappresentazioni equivalenti

Nel Paragrafo 2.2.4 si è accennato al fatto che la scelta delle variabili di stato da utilizzare per descrivere un oggetto fisico mediante un sistema dinamico non è unica. Con riferimento al sistema (3.1), (3.2) è facile essere più concreti. Si consideri una matrice costante $T \in \mathbb{R}^{n \times n}$ non singolare e, mediante un *cambio di variabili*, si definisca un nuovo vettore di stato \hat{x} come

$$\hat{x}(t) = Tx(t) \quad (3.11)$$

Grazie al fatto che T è invertibile, la corrispondenza tra x e \hat{x} risulta biunivoca; infatti

$$x(t) = T^{-1}\hat{x}(t) \quad (3.12)$$

Allora, sostituendo l'Equazione (3.12) nelle (3.1), (3.2), dopo aver moltiplicato la (3.1) a sinistra per T , si ottiene

$$\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t) \quad (3.13)$$

$$y(t) = \hat{C}\hat{x}(t) + \hat{D}u(t) \quad (3.14)$$

con

$$\hat{A} = TAT^{-1}, \quad \hat{B} = TB, \quad \hat{C} = CT^{-1}, \quad \hat{D} = D \quad (3.15)$$

Questo sistema dinamico è *equivalente* a quello descritto dalle Equazioni (3.1), (3.2), nel senso che, per un ingresso $u(t)$, $t \geq t_0$, e due stati iniziali x_{t_0} e \hat{x}_{t_0} legati dalla condizione $\hat{x}_{t_0} = Tx_{t_0}$, i movimenti dello stato dei sistemi (3.1), (3.2)

e (3.13), (3.14) sono effettivamente legati dalla relazione (3.11), cioè $\hat{x}(t) = Tx(t)$, $t > t_0$, e i movimenti dell'uscita sono identici. Lo stesso risultato vale per le singole componenti libere e forzate. Tutto ciò si può verificare per semplice sostituzione. Pertanto le due quadruple di matrici (A, B, C, D) e $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ sono semplicemente due maniere differenti di descrivere un medesimo oggetto fisico.

Esempio 3.3 Il sistema meccanico di Figura 3.1 è costituito da due carrelli di massa M_1 ed M_2 collegati da una molla di rigidezza k e soggetti rispettivamente alle forze u_1 e u_2 . Indicando con x_1 e x_2 le loro posizioni misurate rispetto a un riferimento fisso e con x_3 e x_4 le corrispondenti velocità, detta y la posizione del primo carrello, si ha

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k}{M_1} & \frac{k}{M_1} & 0 & 0 \\ \frac{k}{M_2} & -\frac{k}{M_2} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{M_1} & 0 \\ 0 & \frac{1}{M_2} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

$$y(t) = [1 \ 0 \ 0 \ 0] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix}$$

dove si è assunto che l'allungamento della molla sia pari a $x_2 - x_1$. In alternativa, la posizione e la velocità assolute del secondo carrello possono essere sostituite da quelle relative alle coordinate del primo; si può cioè porre

$$\hat{x}_1(t) = x_1(t)$$

$$\hat{x}_2(t) = x_2(t) - x_1(t)$$

$$\hat{x}_3(t) = x_3(t)$$

$$\hat{x}_4(t) = x_4(t) - x_3(t)$$

cui corrisponde la matrice di trasformazione

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

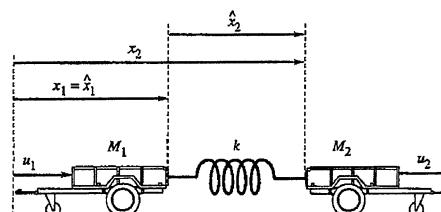


Figura 3.1
Sistema meccanico
dell'Esempio 3.3.

Le nuove equazioni del sistema risultano allora

$$\begin{bmatrix} \dot{\hat{x}}_1(t) \\ \dot{\hat{x}}_2(t) \\ \dot{\hat{x}}_3(t) \\ \dot{\hat{x}}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{k}{M_1} & 0 & 0 \\ 0 & -\frac{k(M_1 + M_2)}{M_1 M_2} & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \\ \hat{x}_4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{M_1} & 0 \\ -\frac{1}{M_1} & \frac{1}{M_2} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

$$y(t) = [1 \ 0 \ 0 \ 0] \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \\ \hat{x}_4(t) \end{bmatrix}$$

Se è necessario, la matrice T e il vettore di stato \hat{x} possono anche essere definiti in campo complesso senza che ciò crei alcun particolare problema, a parte la sicura perdita di significato fisico da parte delle nuove variabili di stato. Si noti in particolare che le matrici A e \hat{A} sono simili e, di conseguenza, i loro autovalori sono identici. Come si mostrerà tra breve utilizzando i risultati dell'Appendice A, questi ultimi contribuiscono in maniera così notevole a determinare le caratteristiche dei movimenti dei sistemi (3.1), (3.2), o (3.13), (3.14), che vengono spesso chiamati addirittura *autovalori del sistema*.

3.2.5 Autovalori e modi

Le formule (3.5), (3.6) permettono di calcolare il movimento libero del sistema (3.1), (3.2). Nel caso il suo ordine sia $n = 1$, e quindi $A = a$ è scalare, x_i e y_i dipendono dalla ben nota funzione esponenziale e^{at} . Per studiare il movimento libero nel caso generale $n \neq 1$ bisogna invece analizzare in profondità la struttura dell'esponenziale di matrice e^{At} , utilizzando i risultati del Paragrafo A.5. Conviene trattare separatamente i casi in cui A è diagonalizzabile oppure no, riferendosi per semplicità al caso $t_0 = 0$.

Matrice della dinamica diagonalizzabile In molti casi, come per esempio tutte le volte che gli autovalori s_i , $i = 1, 2, \dots, n$, della matrice A sono tra loro distinti, è possibile scegliere la matrice $T = T_D$ in modo che la matrice della dinamica \hat{A} assuma la *forma diagonale* \hat{A}_D , cioè risulti

$$\hat{A} = \hat{A}_D = \text{diag}\{s_1, s_2, \dots, s_n\}$$

Quando ciò accade il movimento libero dello stato del sistema (3.13)-(3.15) risulta semplicemente

$$\begin{aligned} \hat{x}_l(t) &= e^{\hat{A}_D t} \hat{x}_0 = \sum_{k=0}^{\infty} \frac{(\hat{A}_D t)^k}{k!} \hat{x}_0 = \\ &= \text{diag} \left\{ \sum_{k=0}^{\infty} \frac{(s_1 t)^k}{k!}, \sum_{k=0}^{\infty} \frac{(s_2 t)^k}{k!}, \dots, \sum_{k=0}^{\infty} \frac{(s_n t)^k}{k!} \right\} \hat{x}_0 = \\ &= \text{diag} \{e^{s_1 t}, e^{s_2 t}, \dots, e^{s_n t}\} \hat{x}_0 \end{aligned}$$

Per le (3.11), (3.12), i corrispondenti movimenti liberi dello stato e dell'uscita del sistema (3.1), (3.2) sono

$$x_l(t) = T_D^{-1} \hat{x}_l(t) = T_D^{-1} \text{diag} \{e^{s_1 t}, e^{s_2 t}, \dots, e^{s_n t}\} T_D x_0 \quad (3.16)$$

$$y_l(t) = C T_D^{-1} \text{diag} \{e^{s_1 t}, e^{s_2 t}, \dots, e^{s_n t}\} T_D x_0 \quad (3.17)$$

Essi pertanto sono combinazioni lineari, con coefficienti dipendenti dagli elementi di x_0 , T_D e C , di termini esponenziali $e^{s_i t}$, $i = 1, 2, \dots, n$, che sono detti *modi*. È molto importante notare che le coppie $s_i = \sigma_i + j\omega_i$, $\bar{s}_i = \sigma_i - j\omega_i$ di autovalori complessi coniugati di A generano termini che, sommati, danno luogo in x_l e y_l a un unico termine reale del tipo $e^{\sigma_i t} \sin(\omega_i t + \varphi_i)$ dove φ_i è una fase opportuna. Con un lieve abuso di linguaggio, anche questi termini saranno denominati modi.

Esempio 3.4 Si consideri il sistema (3.1), (3.2) con

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad C = [1 \ 1]$$

Gli autovalori sono $s_{1,2} = 1 \pm j$. Prendendo

$$T_D^{-1} = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}$$

si ha

$$\hat{A} = T_D A T_D^{-1} = \begin{bmatrix} 1+j & 0 \\ 0 & 1-j \end{bmatrix}$$

da cui, per la (3.16),

$$\begin{aligned} x_l(t) &= 0.5 \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \begin{bmatrix} e^{(1+j)t} & 0 \\ 0 & e^{(1-j)t} \end{bmatrix} \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix} x_0 = \\ &= e^t \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} x_0 \end{aligned}$$

mentre, per la (3.17),

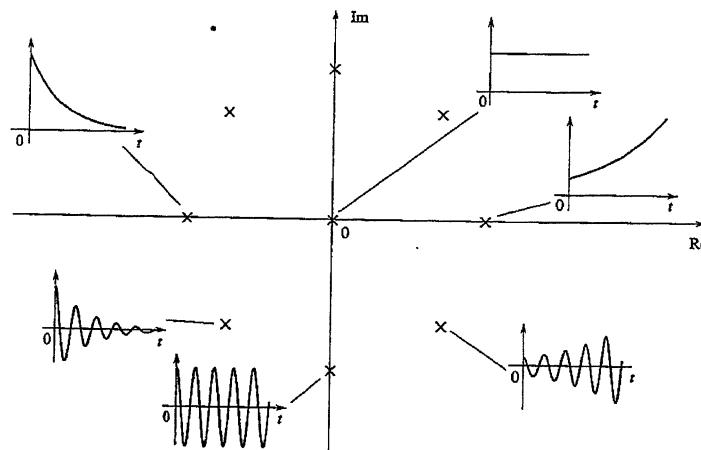
$$y_l(t) = e^t \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} x_0 = \sqrt{2} e^t [\cos(t + \pi/4) \ \cos(t - \pi/4)] x_0$$

La Figura 3.2 riporta un quadro sintetico dell'andamento temporale qualitativo dei modi in dipendenza della posizione nel piano complesso dei corrispondenti autovalori, rappresentati con crocette.

Matrice della dinamica non diagonalizzabile Quando la matrice A possiede autovalori multipli, cioè quando non tutti gli autovalori s_i , $i = 1, 2, \dots, n$, sono distinti tra loro, può darsi che risulti impossibile renderla diagonale mediante un'opportuna trasformazione. In questo caso esiste comunque una matrice di trasformazione $T = T_J$ capace di trasformare A nella cosiddetta *forma di Jordan* \hat{A}_J . Come descritto nel Paragrafo A.3.2, la matrice \hat{A}_J ha una struttura "quasi diagonale", cioè i suoi soli elementi non nulli sono quelli sulla diagonale principale, coincidenti con gli autovalori, insieme ad alcuni, di valore unitario, posti sulla sopradiagonale. In base ai risultati del Paragrafo A.5, i modi che costituiscono i movimenti liberi dello stato e dell'uscita sono allora della forma

$$t^{\eta-1} e^{s_i t} \quad (3.18)$$

Figura 3.2
Modi dei sistemi con autovalori distinti.

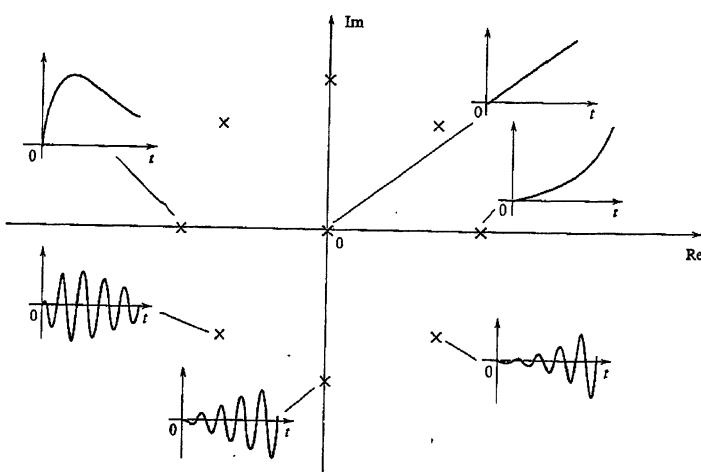


se s_i è reale e

$$t^{\eta-1} e^{\sigma_i t} \sin(\omega_i t + \varphi_i) \quad (3.19)$$

se $s_i = \sigma_i + j\omega_i$ è complesso, dove η è un qualunque intero compreso tra 1 e la massima dimensione dei miniblocchi di Jordan associati a s_i , e φ_i una fase opportuna. La Figura 3.3 riporta l'andamento temporale qualitativo dei modi relativi ad autovalori doppi che si affiancano a quelli della Figura 3.2.

Figura 3.3
Modi dei sistemi con autovalori doppi.



Esempio 3.5 Si prenda in esame il corpo di massa M , vincolato a un riferimento fisso mediante una molla di costante elastica $k \geq 0$, in moto su una guida rettilinea caratterizzata da un coefficiente di attrito viscoso rispetto al corpo $h \geq 0$, rappresentato nella Figura 3.4 e simile a quello già considerato negli Esempi 1.14-1.18. Assumendo come variabile di uscita y lo spostamento del corpo rispetto alla posizione di riposo della molla e come variabile di ingresso u la forza esterna applicata al corpo, la legge fondamentale della dinamica consente di scrivere l'equazione

$$M\ddot{y}(t) = -ky(t) - h\dot{y}(t) + u(t) \quad (3.20)$$

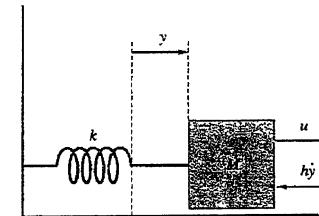


Figura 3.4
Sistema massa-molla
dell'Esempio 3.5.

Posizione
di riposo

Scegliendo come variabili di stato la stessa uscita, cioè la posizione del corpo, e la sua derivata, cioè la velocità, vale a dire ponendo $x_1 = y$ e $x_2 = \dot{y}$, è possibile sostituire all'Equazione (3.20) le equazioni

$$\dot{x}_1(t) = x_2(t) \quad (3.21)$$

$$\dot{x}_2(t) = -\frac{k}{M}x_1(t) - \frac{h}{M}x_2(t) + \frac{1}{M}u(t) \quad (3.22)$$

$$y(t) = x_1(t) \quad (3.23)$$

Il sistema si può pertanto porre nella forma (3.1), (3.2), con $m = p = 1, n = 2$ e

$$A = \begin{bmatrix} 0 & 1 \\ -k/M & -h/M \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1/M \end{bmatrix}, \quad C = [1 \ 0], \quad D = 0 \quad (3.24)$$

mentre i suoi autovalori sono

$$s_1 = -\frac{h}{2M} + \sqrt{\frac{h^2}{4M^2} - \frac{k}{M}}, \quad s_2 = -\frac{h}{2M} - \sqrt{\frac{h^2}{4M^2} - \frac{k}{M}} \quad (3.25)$$

Si inizi con il considerare il caso $h^2 \neq 4Mk$, in cui gli autovalori sono distinti ed è facile verificare che ponendo

$$T_D = \frac{1}{s_2 - s_1} \begin{bmatrix} s_2 & -1 \\ -s_1 & 1 \end{bmatrix} \quad (3.26)$$

risulta

$$\hat{A}_D = T_D A T_D^{-1} = \text{diag}\{s_1, s_2\}$$

e quindi, per le (3.16), (3.17),

$$x_1(t) = \frac{1}{s_2 - s_1} \begin{bmatrix} (s_2 x_{01} - x_{02}) e^{s_1 t} - (s_1 x_{01} - x_{02}) e^{s_2 t} \\ (s_1 s_2 x_{01} - s_1 x_{02}) e^{s_1 t} - (s_1 s_2 x_{01} - s_2 x_{02}) e^{s_2 t} \end{bmatrix} \quad (3.27)$$

$$y_1(t) = \frac{1}{s_2 - s_1} (s_2 x_{01} - x_{02}) e^{s_1 t} - (s_1 x_{01} - x_{02}) e^{s_2 t} \quad (3.28)$$

Queste formule sono espansive solo se $h^2 > 4Mk$, cioè se gli autovalori sono reali. Se invece $h^2 < 4Mk$, gli autovalori sono complessi coniugati e ponendo

$$\sigma = -\frac{h}{2M}, \quad \omega = \sqrt{\frac{k}{M} - \frac{h^2}{4M^2}}$$

si possono scrivere nella forma $s_{1,2} = \sigma \pm j\omega$. Sostituendo queste espressioni nelle (3.27), (3.28), per

$$\delta = \sqrt{\sigma^2 + \omega^2} \quad , \quad \gamma = \arcsin\left(\frac{\omega}{\delta}\right) = \arccos\left(\frac{\sigma}{\delta}\right)$$

si ottiene

$$\begin{aligned} x_f(t) &= e^{\sigma t} \begin{bmatrix} -\frac{\delta}{\omega} \sin(\omega t - \gamma) x_{01} + \frac{1}{\omega} \sin(\omega t) x_{02} \\ -\frac{\delta^2}{\omega} \sin(\omega t) x_{01} + \frac{\delta}{\omega} \sin(\omega t + \gamma) x_{02} \end{bmatrix} \\ y_f(t) &= e^{\sigma t} \left(-\frac{\delta}{\omega} \sin(\omega t - \gamma) x_{01} + \frac{1}{\omega} \sin(\omega t) x_{02} \right) \end{aligned} \quad (3.29)$$

D'altra parte, se $h^2 = 4Mk$, gli autovalori sono coincidenti e valgono $s_1 = s_2 = s_0 = -h/2M$. Allora ponendo

$$T_J = \begin{bmatrix} 0 & -1/s_0^2 \\ 1 & -1/s_0 \end{bmatrix} \quad (3.30)$$

risulta

$$\hat{A}_J = T_J A T_J^{-1} = \begin{bmatrix} s_0 & 1 \\ 0 & s_0 \end{bmatrix} \quad (3.31)$$

e quindi

$$\begin{aligned} x_f(t) &= e^{s_0 t} \begin{bmatrix} x_{01} - (s_0 x_{01} - x_{02})t \\ x_{02} - s_0(s_0 x_{01} - x_{02})t \end{bmatrix} \\ y_f(t) &= e^{s_0 t} (x_{01} - (s_0 x_{01} - x_{02})t) \end{aligned} \quad (3.32)$$

3.2.6 Risposta all'impulso e movimento forzato

Si supponga che il sistema (3.1), (3.2) abbia un solo ingresso e sia sollecitato con un impulso unitario applicato in $t_0 = 0$: si assuma cioè $m = 1$ e $u(t) = \text{imp}(t)$. Denotando con g_x e g_y i particolari movimenti forzati dello stato e dell'uscita che si ottengono dalle (3.7), (3.8), per la proprietà (B.3) della funzione impulso risulta

$$g_x(t) = e^{At} B \quad (3.33)$$

$$g_y(t) = C e^{At} B + D \text{imp}(t) \quad (3.34)$$

Questi movimenti sono detti *risposta all'impulso dello stato e dell'uscita*. Si noti che, per $t > 0$, essi coincidono con i movimenti liberi prodotti dallo stato iniziale $x(0) = B$ e come tali sono costituiti da combinazioni dei modi del sistema.

L'estensione di queste definizioni al caso $m > 1$ è molto naturale. In generale le (3.33), (3.34) definiscono due matrici di funzioni di dimensioni $n \times m$ e $p \times m$ e il generico elemento di posizione (i, j) di g_x , o di g_y , rappresenta la risposta della variabile di stato i -esima, o dell'uscita i -esima, a un impulso unitario inviato all'ingresso j -esimo (avendo posto nulle le altre variabili di ingresso).

Qualunque sia il valore di m , le risposte all'impulso rivestono un'importanza del tutto speciale, in quanto la loro sola conoscenza è sufficiente per calcolare i movimenti forzati dello stato e dell'uscita generati da un ingresso qualunque (applicato in $t_0 = 0$). Ricordando ancora la proprietà (B.3) della funzione impulso, è infatti immediato verificare che questi movimenti sono ottenibili come prodotto di convoluzione delle funzioni (3.33), (3.34) con la funzione di ingresso; visto che

g_x , g_y e u sono nulle per tempi negativi, si ha cioè

$$\begin{aligned} g_x(t) * u(t) &= \int_{-\infty}^{+\infty} g_x(t - \tau) u(\tau) d\tau = \int_0^t g_x(t - \tau) u(\tau) d\tau = \\ &= \int_0^t e^{A(t-\tau)} B u(\tau) d\tau = x_f(t) \\ g_y(t) * u(t) &= \int_{-\infty}^{+\infty} g_y(t - \tau) u(\tau) d\tau = \int_0^t g_y(t - \tau) u(\tau) d\tau = \\ &= \int_0^t (C e^{A(t-\tau)} B + D \text{imp}(t - \tau)) u(\tau) d\tau = \\ &= C \int_0^t e^{A(t-\tau)} B u(\tau) d\tau + D u(t) = y_f(t) \end{aligned} \quad (3.35)$$

D'altra parte, supponendo per semplicità che la matrice A sia diagonalizzabile, per le (3.12), (3.15) i movimenti forzati (3.7), (3.8) del sistema (3.1), (3.2) risultano allora

$$\begin{aligned} x_f(t) &= T_D^{-1} \hat{x}_f(t) = T_D^{-1} \int_0^t e^{\hat{A}_D(t-\tau)} \hat{B} u(\tau) d\tau = \\ &= T_D^{-1} \int_0^t \text{diag}\{e^{s_1(t-\tau)}, e^{s_2(t-\tau)}, \dots, e^{s_n(t-\tau)}\} T_D B u(\tau) d\tau \end{aligned} \quad (3.36)$$

$$y_f(t) = C T_D^{-1} \int_0^t \text{diag}\{e^{s_1(t-\tau)}, e^{s_2(t-\tau)}, \dots, e^{s_n(t-\tau)}\} T_D B u(\tau) d\tau + D u(t) \quad (3.37)$$

mentre è ovvia la generalizzazione al caso in cui A non sia diagonalizzabile, basata sulla sua forma di Jordan \hat{A}_J .

Esempio 3.6 (Seguito dell'Esempio 3.5) Assumendo $u(t) = \text{sca}(t)$, nel caso $s_1 \neq s_2$ con $k \neq 0$, le Equazioni (3.36), (3.37) per la (3.26) fanno sì che risulti

$$\begin{aligned} x_f(t) &= \begin{bmatrix} 1 & 1 \\ s_1 & s_2 \end{bmatrix} \int_0^t \text{diag}\{e^{s_1(t-\tau)}, e^{s_2(t-\tau)}\} d\tau \begin{bmatrix} \frac{1}{M(s_2 - s_1)} \\ \frac{1}{M(s_2 - s_1)} \end{bmatrix} = \\ &= \frac{1}{M(s_1 - s_2)} \begin{bmatrix} \frac{e^{s_1 t} - 1}{s_1} - \frac{e^{s_2 t} - 1}{s_2} \\ \frac{e^{s_1 t} - 1}{s_1} - \frac{e^{s_2 t} - 1}{s_2} \end{bmatrix} \\ y_f(t) &= \frac{1}{M(s_1 - s_2)} \left(\frac{e^{s_1 t} - 1}{s_1} - \frac{e^{s_2 t} - 1}{s_2} \right) \end{aligned}$$

Anche in questo caso se gli autovalori sono complessi coniugati conviene riscrivere le espressioni dei movimenti in modo che risultino chiari che x_f e y_f assumono valori reali. Precisamente si ha

$$\begin{aligned} x_f(t) &= \frac{1}{M} \begin{bmatrix} \frac{1}{\delta^2} \left(1 + \frac{\delta}{\omega} e^{\sigma t} \sin(\omega t - \gamma) \right) \\ \frac{1}{\omega} e^{\sigma t} \sin(\omega t) \end{bmatrix} \\ y_f(t) &= \frac{1}{M\delta^2} \left(1 + \frac{\delta}{\omega} e^{\sigma t} \sin(\omega t - \gamma) \right) \end{aligned}$$

Se invece $s_1 \neq s_2$ con $k = 0$, allora risulta $s_1 = 0$ e si ricava

$$\begin{aligned} x_f(t) &= \left[\begin{array}{cc} 1 & 1 \\ 0 & s_2 \end{array} \right] \int_0^t \text{diag} \left\{ 1, e^{s_2(t-\tau)} \right\} d\tau \left[\begin{array}{c} -\frac{1}{Ms_2} \\ \frac{1}{Ms_2} \end{array} \right] = \frac{1}{Ms_2} \left[\begin{array}{c} e^{s_2 t} - 1 \\ \frac{s_2}{e^{s_2 t} - 1} - t \end{array} \right] \\ y_f(t) &= \frac{1}{Ms_2} \left(\frac{e^{s_2 t} - 1}{s_2} - t \right) \end{aligned}$$

Se poi $s_1 = s_2 = s_0 \neq 0$, allora, ricordando le (3.30), (3.31), si può verificare che risulta

$$x_f(t) = \frac{1}{M} \left[\begin{array}{c} \frac{1}{s_0^2} - \frac{e^{s_0 t}}{s_0^2} + \frac{t e^{s_0 t}}{s_0} \\ t e^{s_0 t} \end{array} \right], \quad y_f(t) = \frac{1}{M} \left(\frac{1}{s_0^2} - \frac{e^{s_0 t}}{s_0^2} + \frac{t e^{s_0 t}}{s_0} \right)$$

e, se infine $s_1 = s_2 = 0$,

$$x_f(t) = \frac{1}{M} \left[\begin{array}{c} 0.5t^2 \\ t \end{array} \right], \quad y_f(t) = \frac{1}{2M} t^2$$

3.3 Equilibrio

Si vogliono ora analizzare le caratteristiche del sistema (3.1), (3.2) per quanto riguarda le condizioni di equilibrio. Assumendo costante e pari a \bar{u} l'ingresso del sistema, sulla base dei risultati del Paragrafo 2.5 si può affermare che gli stati di equilibrio \bar{x} sono le soluzioni dell'equazione

$$A\bar{x} + B\bar{u} = 0 \quad (3.38)$$

A ogni stato di equilibrio corrisponde l'uscita di equilibrio

$$\bar{y} = C\bar{x} + D\bar{u} \quad (3.39)$$

Notevole è il caso in cui A è invertibile, cioè $\det(A) \neq 0$, o, equivalentemente, A non ha autovalori nulli: in questa circostanza, infatti, l'Equazione (3.38) ammette una e una sola soluzione, per cui lo stato di equilibrio è unico e risulta

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

$$\bar{y} = (-CA^{-1}B + D)\bar{u}$$

La matrice $D - CA^{-1}B \in R^{p \times m}$ rappresenta il *guadagno statico* del sistema. Per i sistemi SISO esso costituisce il rapporto tra l'uscita e l'ingresso quando tutte le variabili del sistema stesso, compreso lo stato, sono costanti.

Esempio 3.7 (Seguito dell'Esempio 3.5) Per il sistema (3.21)-(3.23) le Equazioni (3.38)-(3.39) diventano

$$\begin{bmatrix} 0 & 1 \\ -k/M & -h/M \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1/M \end{bmatrix} \bar{u} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\bar{y} = [0 \ 1] \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}$$

da cui si trova

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 1/k \\ 0 \end{bmatrix} \bar{u}, \quad \bar{y} = \frac{1}{k} \bar{u}$$

Esempio 3.8 (Seguito dell'Esempio 2.6) Per il motore elettrico a corrente continua descritto dalle Equazioni (2.14)-(2.17) si desideri calcolare lo stato di equilibrio corrispondente all'applicazione di una tensione costante di alimentazione \bar{v} e di una coppia resistente costante \bar{C}_r . Risolvendo il sistema di equazioni che si ottiene dalle (2.14)-(2.15) ponendo $u_1(t) = \bar{v}$, $u_2(t) = \bar{C}_r$ e annullando le derivate delle variabili di stato, si ricava facilmente

$$\bar{x}_1 = \frac{\bar{v} + k\bar{C}_r}{k^2 + Rh} \quad (3.40)$$

$$\bar{x}_2 = \frac{k\bar{v} - R\bar{C}_r}{k^2 + Rh} \quad (3.41)$$

Queste espressioni rappresentano dunque i valori di equilibrio rispettivamente della corrente di armatura e della velocità di rotazione.

Quando invece $\det(A) = 0$, l'Equazione (3.38) può ammettere infinite soluzioni o anche non ammetterne alcuna, anche in dipendenza dal valore di \bar{u} , a seconda che il vettore $B\bar{u}$ sia tra quelli generabili moltiplicando a destra A per un opportuno vettore \bar{x} oppure no. Vi sono allora infiniti stati di equilibrio, oppure nessuno, e la nozione di guadagno statico perde comunque senso.

Esempio 3.9 Si consideri il sistema

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} x(t) + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} u(t) \\ y(t) &= [5 \ 6] x(t) \end{aligned}$$

Se \bar{u} è tale che $\bar{u}_1 \neq -2\bar{u}_2$, non vi sono stati di equilibrio perché $\dot{x} \neq 0$. Se invece $\bar{u}_1 = -2\bar{u}_2$, allora vi sono infiniti stati e uscite di equilibrio descritti da

$$\bar{x} = \begin{bmatrix} 2(\bar{u}_2 - \alpha) \\ \alpha \end{bmatrix}, \quad \bar{y} = 2(5\bar{u}_2 - 2\alpha)$$

essendo α uno scalare arbitrario.

3.4 Stabilità

Per un sistema lineare stazionario la determinazione delle proprietà di stabilità si può effettuare in modo semplice, perché esse di fatto sono relative all'intero sistema e dipendono dal solo movimento libero. Per di più anche il calcolo del movimento libero può essere evitato. Infatti, è possibile dare condizioni necessarie e sufficienti di stabilità asintotica e condizioni solo sufficienti di instabilità riferite esclusivamente agli autovalori del sistema. In termini di soli autovalori non si possono quindi individuare le caratteristiche di stabilità proprio in tutti i casi, ma i risultati disponibili sono comunque utilissimi, soprattutto in vista del fatto che la proprietà di stabilità asintotica è quella di gran lunga più interessante nelle applicazioni. Nel seguito del paragrafo si vedrà anche come perfino il calcolo degli autovalori sia sostituibile con procedure più semplici.

3.4.1 Stabilità del sistema

Si supponga di voler studiare la stabilità del movimento $\tilde{x}(t)$ prodotto nel sistema lineare di equazione di stato (3.1) da $\tilde{u}(t)$, $t \geq 0$, e $\tilde{x}(0) = \tilde{x}_0$; in particolare questo movimento può essere anche uno stato di equilibrio. Applicando il principio di sovrapposizione degli effetti (Teorema 3.1) con $u'(t) = u''(t) = \tilde{u}(t)$, $x'_0 = \tilde{x}_0$, $x''_0 = \tilde{x}_0 + \delta x_0$, $\alpha = -1$ e $\beta = 1$, si trova che la differenza $\delta x(t) = x(t) - \tilde{x}(t)$ tra movimento perturbato e movimento nominale è retta dall'equazione

$$\delta\dot{x}(t) = A\delta x(t), \quad \delta x(0) = \delta x_0 \quad (3.42)$$

Pertanto, per le Definizioni 2.4-2.6 (o 2.1-2.3 se ci si riferisce a un equilibrio), il movimento \tilde{x} è stabile se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti i δx_0 per cui

$$\|\delta x_0\| \leq \delta \quad (3.43)$$

risulti

$$\|\delta x(t)\| \leq \varepsilon, \quad t \geq 0 \quad (3.44)$$

Il movimento (o l'equilibrio) \tilde{x} è instabile se ciò non accade, ed è invece asintoticamente stabile se, oltre a essere stabile, verifica pure la condizione

$$\lim_{t \rightarrow \infty} \|\delta x(t)\| = 0 \quad (3.45)$$

Si osservi subito che nella riformulazione appena data delle Definizioni 2.4-2.6 (o 2.1-2.3) concernenti la stabilità, \tilde{x} non appare più in modo esplicito, o in altri termini l'analisi della stabilità di un qualunque movimento \tilde{x} , e in particolare anche di un qualunque stato di equilibrio, porta comunque allo studio delle soluzioni dell'Equazione (3.42) al variare dello stato iniziale come specificato dalla condizione (3.43). Si può quindi enunciare il seguente fondamentale risultato.

Teorema 3.2 Un movimento (o uno stato di equilibrio) di un sistema lineare stazionario è stabile, asintoticamente stabile o instabile se e solo se tutti i movimenti (o gli stati di equilibrio) del sistema sono rispettivamente stabili, asintoticamente stabili o instabili. ■

Come conseguenza di questo teorema, per i sistemi considerati ha senso parlare di stabilità, stabilità asintotica o instabilità del sistema, invece che di singoli movimenti (o stati di equilibrio), intendendo con questo il sussistere delle proprietà di stabilità, stabilità asintotica o instabilità di tutti i movimenti (o gli stati di equilibrio) del sistema stesso secondo le Definizioni 2.4-2.6 (o 2.1-2.3).

3.4.2 Stabilità e movimento libero

Al di là dei simboli utilizzati, le soluzioni dell'Equazione (3.42) non sono altro che i movimenti liberi dello stato del sistema di equazione di stato (3.1). La stabilità allora sussiste quando, fissata arbitrariamente una norma massima ε per questi movimenti, la condizione (3.44) è verificata almeno per stati iniziali di norma abbastanza piccola, come specificato dalla condizione (3.43). Però, visto che, moltiplicando per un qualunque coefficiente lo stato iniziale, il movimento libero rimane moltiplicato per lo stesso coefficiente, ciò che conta è esclusivamente che il movimento libero sia limitato, cioè non vada all'infinito, per tutti i $t \geq 0$ e per tutti gli stati iniziali. In questo caso, e solo in questo caso, infatti, pur di prendere δ abbastanza piccolo, la condizione (3.44) è verificata. Se questo non accade invece il sistema è instabile. Infine, perché il sistema sia asintoticamente

stabile la condizione (3.45) impone che, sempre per linearità, tutti i movimenti liberi si annullino asintoticamente; essi devono anche essere limitati, ma, vista la struttura dei modi costituenti il movimento libero (Paragrafo 3.2.5), quest'ultima condizione è implicata da quella concernente l'azzeramento asintotico. Quanto qui osservato costituisce la dimostrazione del seguente importantissimo risultato.

Teorema 3.3 Un sistema lineare stazionario è stabile se e solo se tutti i movimenti liberi dello stato sono limitati: è asintoticamente stabile se e solo se tutti i movimenti liberi dello stato tendono a zero per $t \rightarrow \infty$; è instabile se e solo se almeno un movimento libero dello stato non è limitato. ■

Le proprietà di stabilità dipendono dunque esclusivamente dalle caratteristiche della matrice della dinamica A . Inoltre, per i sistemi che qui si stanno considerando la proprietà di stabilità asintotica implica quella di stabilità globale di tutti i movimenti, perché, indipendentemente dal valore della norma di δx_0 , sotto l'ipotesi di asintotica stabilità tutti i movimenti liberi tendono a zero e quindi tutti i movimenti perturbati tendono al movimento nominale, qualunque esso sia.

Esempio 3.10 (Seguito degli Esempi 2.2-2.4 e 2.17) Dall'Equazione (2.7) si ricava che il movimento libero dello stato è

$$x_l(t) = e^{-t/R} C x(0)$$

che, qualunque sia $x(0)$, è limitato per $t \geq 0$ e si annulla per $t \rightarrow \infty$. Il Teorema 3.3 conferma così i risultati ottenuti nell'Esempio 2.17.

Si prenda ancora in esame il ritardo di tempo (2.30). Poiché questo sistema non è descritto dalle Equazioni (3.1) e (3.2), in senso stretto non sono applicabili a esso né le Definizioni 2.4-2.6, né i Teoremi 3.2 e 3.3. Si è già affermato però nel Paragrafo 2.4 che esso è lineare, e pertanto gli si possono applicare il principio di sovrapposizione degli effetti e i concetti di movimento libero e forzato. Generalizzando i risultati dell'Esempio 2.10 e sulla base dei commenti a esso successivi, si può concludere che il movimento libero dello stato del ritardo è limitato per $t > 0$ e nullo per $t \geq \tau$ qualunque sia la condizione iniziale, in quanto è chiaro che la densità di materiale sul nastro trasportatore per $t \geq \tau$ non dipende dalla densità iniziale. Perciò, assumendo di poter applicare comunque il Teorema 3.3, si può ben dire che il ritardo di tempo (2.30) è asintoticamente stabile, sia pure in un senso allargato.

3.4.3 Stabilità e autovalori

L'analisi del movimento libero condotta nel paragrafo precedente ha mostrato che esso è costituito da una combinazione lineare di modi del tipo (3.18), (3.19). Si può verificare che, al variare dello stato iniziale, tutti i modi compaiono nel movimento libero dello stato e pertanto, ricordando il Teorema 3.3, per accettare che un sistema sia asintoticamente stabile occorre e basta verificare che i singoli modi tendano a zero per t tendente all'infinito. Si arriva allora immediatamente alla formulazione della seguente condizione necessaria e sufficiente.

Teorema 3.4 Il sistema lineare e stazionario (3.1), (3.2) è asintoticamente stabile se e solo se tutti i suoi autovalori hanno parte reale negativa. ■

Alla luce di questo teorema si usa dire che i punti del piano complesso con parte reale negativa costituiscono la *regione di asintotica stabilità*. Analogamente, è facile pervenire alla condizione sufficiente di instabilità riportata qui di seguito.

Teorema 3.5 Il sistema lineare e stazionario (3.1), (3.2) è instabile se almeno uno dei suoi autovalori ha parte reale positiva. ■

Confrontando le condizioni espresse dai teoremi precedenti, si conclude che l'unico caso indecidibile dalla sola analisi degli autovalori si verifica quando il sistema possiede autovalori con parte reale nulla insieme ad altri eventuali con parte reale negativa. In questo caso, esclusa la stabilità asintotica, si potrebbe avere sia la stabilità sia l'instabilità. Quest'ultima si ha se e solo se, tra gli autovalori con parte reale nulla, ce n'è almeno uno cui corrisponda almeno un miniblocco di Jordan di dimensioni maggiori di 1, perché in questo caso, e solo in questo caso, nel movimento libero dello stato compare un modo del tipo t^k oppure $t^k \sin(\omega_i t + \varphi_i)$, con $k > 0$, illimitato per $t \rightarrow \infty$. Si noti che questi termini sono sicuramente assenti quando non vi sono autovalori multipli sull'asse immaginario: in questo caso allora la condizione sufficiente del Teorema 3.5 diventa anche necessaria, e il sistema risulta stabile se e solo se tutti i suoi autovalori hanno parte reale negativa o nulla.

Esempio 3.11 (Seguito degli Esempi 2.2-2.4, 2.17 e 3.10) Il Teorema 3.4 consente di confermare immediatamente la conclusione già raggiunta che il circuito elettrico è asintoticamente stabile: il suo unico autovalore $-1/RC$ è infatti reale e negativo.

Esempio 3.12 (Seguito dell'Esempio 3.5) Gli autovalori (3.25) del sistema hanno parte reale negativa quando h è positivo, cioè l'attrito è effettivamente presente. In questo caso pertanto il sistema è asintoticamente stabile. Se invece h è nullo, gli autovalori hanno parte reale nulla e occorre considerare due diversi casi in dipendenza del valore di k : quando $k \neq 0$, gli autovalori sono distinti e il sistema è stabile; quando invece $k = 0$, gli autovalori sono entrambi nulli e non è possibile trarre conclusioni definitive sulla base dei Teoremi 3.4 e 3.5. Le Equazioni (3.27), (3.29), (3.32) che riportano le formule del movimento libero corrispondente alle varie combinazioni di parametri permettono di verificare i risultati sopra riportati. La (3.32) poi consente di concludere che per $h = k = 0$ il sistema è instabile: infatti la forma di Jordan (3.31) di A è costituita da un unico miniblocco 2×2 corrispondente all'autovalore nullo.

3.4.4 Stabilità e polinomio caratteristico

Concentrandosi ora sulla proprietà di stabilità asintotica, si osservi che, per il Teorema 3.4, l'accertamento della sua presenza non richiede più la soluzione di un'equazione differenziale per il calcolo del movimento libero, ma solo la soluzione di un'equazione algebrica per il calcolo degli autovalori. Si è così ottenuta una notevole semplificazione. D'altra parte l'*equazione caratteristica*

$$\varphi(s) = 0 \quad (3.46)$$

ottenuta ugagliando a zero il *polinomio caratteristico*

$$\varphi(s) = \det(sI - A) = s^n + p_1 s^{n-1} + p_2 s^{n-2} + \dots + p_{n-1} s + p_n \quad (3.47)$$

è un'equazione polinomiale di grado n e per essa la soluzione esplicita non è banale già per $n > 2$, mentre la formula risolutiva non esiste addirittura per $n > 4$. Di fatto quindi per risolvere l'Equazione (3.46) bisogna usare metodi numerici iterativi con il supporto di adeguati strumenti di calcolo. Pur dando per scontato che questi mezzi siano disponibili a chi intraprende l'analisi di un sistema, la loro utilizzazione risulta per certi versi scomoda, come sarà più chiaro nel seguito, e comunque fondamentalmente inutile. Esistono infatti dei risultati che, senza

richiedere la soluzione dell'equazione caratteristica, consentono di dire se tutti gli n autovalori s_i hanno parte reale negativa oppure no, dando così luogo, per il Teorema 3.4, a criteri di stabilità asintotica tanto semplici da poter essere applicati senza l'ausilio di particolari supporti di calcolo anche a sistemi di ordine elevato. Per la presentazione di tali risultati conviene riferirsi al polinomio caratteristico

$$\varphi(s) = \varphi_0 s^n + \varphi_1 s^{n-1} + \varphi_2 s^{n-2} + \dots + \varphi_{n-1} s + \varphi_n \quad \varphi_0 \neq 0 \quad (3.48)$$

scritto in una forma un poco più generale della (3.47).

Osservando che

$$\varphi(s) = \varphi_0 \prod_{i=1}^n (s - s_i) \quad (3.49)$$

e ricordando le proprietà della traccia e del determinante (Paragrafo A.3), si inizi col notare che

$$\varphi_1/\varphi_0 = -\text{tr}(A) = -\sum_{i=1}^n s_i, \quad \varphi_n/\varphi_0 = \det(-A) = (-1)^n \prod_{i=1}^n s_i$$

Quando il sistema è asintoticamente stabile risulta $\sum_{i=1}^n s_i < 0$; quindi $\text{tr}(A) < 0$ e perciò $\varphi_1/\varphi_0 > 0$. Risulta anche $(-1)^n \prod_{i=1}^n s_i > 0$ e quindi $\det(-A) > 0$ da cui $\varphi_n/\varphi_0 > 0$. Pertanto le condizioni

$$\varphi_1/\varphi_0 > 0, \quad \varphi_n/\varphi_0 > 0$$

sono necessarie per la stabilità asintotica. Più in generale, imponendo $\text{Re}(s_i) < 0$, $i = 1, 2, \dots, n$, nell'Equazione (3.49), risulta dimostrata la seguente condizione necessaria.

Teorema 3.6 Se il sistema (3.1), (3.2) è asintoticamente stabile, allora i coefficienti φ_i , $i = 0, 1, \dots, n$, del polinomio caratteristico hanno tutti lo stesso segno. ■

Si può poi verificare che questa condizione è anche sufficiente per sistemi di ordine $n = 1$ e $n = 2$.

Esempio 3.13 Il Teorema 3.6 consente di concludere circa la mancanza di stabilità asintotica per i sistemi dotati di polinomi caratteristici come

$$\varphi(s) = s^3 + 3s^2 - s - 3$$

oppure

$$\varphi(s) = s^4 + 5s^2 + 4$$

cui corrispondono rispettivamente gli autovalori $-3, -1, +1$ e $\pm j, \mp j$. La condizione non è anche sufficiente per $n > 2$: per esempio le radici dell'equazione

$$\varphi(s) = s^3 + s^2 + s + 1 = 0 \quad (3.50)$$

sono -1 e $\pm j$.

Criterio di Routh Per formulare una nuova condizione di stabilità asintotica che sia sufficiente, oltre che necessaria, occorre per prima cosa definire la cosiddetta *tabella di Routh*, costruita a partire dai coefficienti del polinomio caratteristico. Essa possiede $n + 1$ righe e ha una struttura triangolare in quanto ogni due righe, con l'esclusione della prima se n è pari, il numero di elementi diminuisce di uno:

φ_0	φ_2	φ_4
φ_1	φ_3	φ_5
.
.
h_1	h_2	h_3
k_1	k_2	k_3
l_1	l_2	l_3
.
.

Le prime due righe contengono i coefficienti del polinomio caratteristico: in ordine, la prima contiene $\varphi_0, \varphi_2, \varphi_4, \dots$, la seconda $\varphi_1, \varphi_3, \varphi_5, \dots$, fino all'esaurimento. Successivamente ogni riga si costruisce sulla base degli elementi delle due righe precedenti e, avendo indicato con h_i, k_i e l_i gli elementi di tre generiche righe consecutive, si ha

$$l_i = -\frac{1}{k_1} \det \left(\begin{bmatrix} h_i & h_{i+1} \\ k_i & k_{i+1} \end{bmatrix} \right) = h_{i+1} - \frac{h_i k_{i+1}}{k_1} \quad (3.51)$$

Gli elementi necessari per applicare la (3.51) e non definiti nelle righe precedenti quella corrente devono essere considerati nulli. Inoltre, se risulta $k_1 = 0$, la (3.51) non è applicabile e allora si dirà, per convenzione, che la tabella di Routh non è ben definita.

Si può ora enunciare la seguente condizione necessaria e sufficiente di asintotica stabilità, di cui qui non si fornirà la dimostrazione, che va sotto il nome di *criterio di Routh*.

Teorema 3.7 Il sistema (3.1), (3.2) è asintoticamente stabile se e solo se la tabella di Routh relativa al suo polinomio caratteristico è ben definita e tutti gli elementi della sua prima colonna hanno lo stesso segno.

Esempio 3.14 La tabella relativa al polinomio

$$\varphi(s) = s^5 + 15s^4 + 85s^3 + 225s^2 + 274s + 120 \quad (3.52)$$

è

1	85	274
15	225	120
70	266	
168	120	
216		
120		

Pertanto, tutte le sue radici hanno parte reale negativa; infatti il polinomio si può scrivere nella forma

$$\varphi(s) = (s + 1)(s + 2)(s + 3)(s + 4)(s + 5)$$

Esempio 3.15 (Seguito dell'Esempio 3.13) La tabella di Routh corrispondente al polinomio (3.50) non è ben definita, perché nella terza riga compare uno zero in prima colonna. Come già visto, tra le sue radici ve ne sono due complesse coniugate con parte reale nulla.

Si noti che, per il Teorema 3.7, la costruzione dell'intera tabella può anche non essere necessaria e inoltre la divisione per k_1 del determinante nell'Equazione (3.51) non è indispensabile per la validità del risultato. Quest'ultima tuttavia, con un'analisi supplementare della tabella di Routh, permetterebbe di determinare il numero di autovalori con parte reale positiva e nulla posseduti da sistemi non asintoticamente stabili.

3.4.5 Stabilità e parametri incerti

A volte può accadere che un sistema dinamico sia definito a meno del valore di qualche parametro e si desideri determinare la *regione di asintotica stabilità nell'insieme dei parametri*, cioè si voglia stabilire quali siano i valori di questi parametri che fanno sì che il sistema stesso sia asintoticamente stabile. Una situazione di questo genere si verifica, per esempio, nei problemi di analisi della stabilità di sistemi parzialmente incerti e anche in quelli di sintesi consistenti nello stabilizzare un sistema mediante un controllore dotato di struttura prefissata e definito a meno del valore di alcuni parametri di progetto. Allora, può essere comodo fare riferimento al criterio di Routh per individuare i valori dei parametri per cui le condizioni necessarie e sufficienti di stabilità asintotica sono verificate.

Esempio 3.16 Si abbia un sistema dotato del polinomio caratteristico

$$\varphi(s) = s^3 + (2 + \beta)s^2 + (1 + 2\beta)s + \alpha + \beta$$

La tabella di Routh corrispondente è

1	1 + 2\beta
2 + \beta	\alpha + \beta
2(\beta + 1)^2 - \alpha	
2 + \beta	
\alpha + \beta	

Pertanto, il sistema è asintoticamente stabile per tutte le coppie di valori α, β che soddisfano le disequazioni

$$\beta > -2, \quad 2(\beta + 1)^2 > \alpha, \quad \beta > -\alpha$$

cioè per tutte le coppie cui corrispondono punti nella regione in colore chiaro della Figura 3.5. Per confronto si noti che la condizione necessaria del Teorema 3.6 è verificata quando sono verificate le diseguaglianze

$$\beta > -2, \quad \beta > -\frac{1}{2}, \quad \beta > -\alpha$$

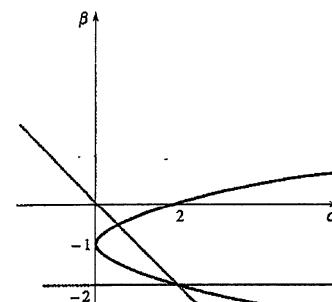
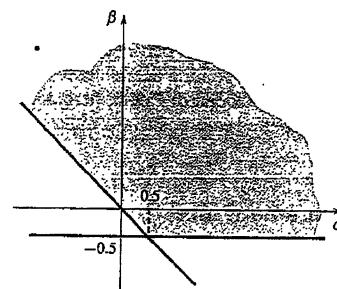


Figura 3.5
Regione di asintotica stabilità nell'insieme dei parametri per l'Esempio 3.16.

Figura 3.6
Condizioni necessarie di asintotica stabilità per l'Esempio 3.16.



cui corrisponde nell'insieme dei parametri la regione in colore chiaro della Figura 3.6, ovviamente contenente la precedente.

Vale la pena di rilevare la semplicità del metodo adottato per la soluzione del problema in esame, soprattutto in confronto alle difficoltà che avrebbe comportato la determinazione degli autovalori in funzione dei parametri liberi.

Criterio di Kharitonov Quando i coefficienti del polinomio caratteristico sono essi stessi i parametri incerti o liberi, l'analisi di stabilità si può effettuare in maniera molto semplice con elaborazioni puramente numeriche piuttosto che simboliche.

Innanzitutto si descriva sinteticamente il polinomio caratteristico (3.48) nella forma

$$\varphi(s) = \{\varphi_0, \varphi_1, \dots, \varphi_n\}$$

Si assume poi che ogni parametro φ_i appartenga a un intervallo noto

$$\varphi_i^- \leq \varphi_i \leq \varphi_i^+ \quad (3.53)$$

e si definiscano i quattro polinomi

$$\varphi_a(s) : \{\varphi_0^+, \varphi_1^+, \varphi_2^-, \varphi_3^-, \varphi_4^+, \varphi_5^+, \varphi_6^-, \dots\}$$

$$\varphi_b(s) : \{\varphi_0^-, \varphi_1^-, \varphi_2^+, \varphi_3^+, \varphi_4^-, \varphi_5^-, \varphi_6^+, \dots\}$$

$$\varphi_c(s) : \{\varphi_0^+, \varphi_1^-, \varphi_2^-, \varphi_3^+, \varphi_4^+, \varphi_5^-, \varphi_6^-, \dots\}$$

$$\varphi_d(s) : \{\varphi_0^-, \varphi_1^+, \varphi_2^+, \varphi_3^-, \varphi_4^-, \varphi_5^+, \varphi_6^+, \dots\}$$

Vale allora la seguente condizione necessaria e sufficiente, che va sotto il nome di *criterio di Kharitonov*.

Teorema 3.8 Le radici del polinomio (3.48) hanno parte reale negativa per tutti i valori dei coefficienti φ_i , purché appartenenti agli intervalli (3.53), se e solo se hanno parte reale negativa tutte le radici dei polinomi φ_a , φ_b , φ_c e φ_d . ■

Si noti la potenza di questo risultato che, sebbene in un caso molto particolare, riduce l'analisi di stabilità di un sistema incerto a quella di quattro sistemi perfettamente noti, qualunque sia l'ordine del sistema stesso, cioè indipendentemente dal numero dei parametri ignoti. Per studiare i singoli polinomi φ_a , φ_b , φ_c e φ_d si può poi usare un qualunque criterio, per esempio quello di Routh.

Esempio 3.17 (Seguito dell'Esempio 3.14) Si supponga che i coefficienti del polinomo caratteristico (3.52) rappresentino i valori nominali dei coefficienti φ_i del polinomio (3.48) e che questi ultimi possano variare entro un intervallo di $\pm 10\%$ attorno ai valori nominali stessi. In altri termini si consideri il polinomio (3.48) con $n = 5$ e

$$\begin{aligned} 0.9 &\leq \varphi_0 \leq 1.1 \\ 13.5 &\leq \varphi_1 \leq 16.5 \\ 76.5 &\leq \varphi_2 \leq 93.5 \\ 202.5 &\leq \varphi_3 \leq 247.5 \\ 246.6 &\leq \varphi_4 \leq 301.4 \\ 108 &\leq \varphi_5 \leq 132 \end{aligned}$$

Come si può accettare facilmente col criterio di Routh, le radici dei polinomi

$$\varphi_a(s) = 1.1s^5 + 16.5s^4 + 76.5s^3 + 202.5s^2 + 301.4s + 132$$

$$\varphi_b(s) = 0.9s^5 + 13.5s^4 + 93.5s^3 + 247.5s^2 + 246.6s + 108$$

$$\varphi_c(s) = 1.1s^5 + 13.5s^4 + 76.5s^3 + 247.5s^2 + 301.4s + 108$$

$$\varphi_d(s) = 0.9s^5 + 16.5s^4 + 93.5s^3 + 202.5s^2 + 246.6s + 132$$

hanno tutte parte reale negativa. Allora, per il criterio di Kharitonov, le radici del polinomio (3.48) hanno parte reale negativa per tutti i valori assunti dai parametri φ_i negli intervalli sopra specificati.

3.4.6 Proprietà dei sistemi asintoticamente stabili

I sistemi asintoticamente stabili sono quelli di maggiore interesse per le applicazioni. In genere, non è accettabile che esistano movimenti liberi illimitati, né che ve ne siano di limitati, ma non asintoticamente nulli. Tra l'altro, la stabilità non asintotica è contraddistinta da condizioni molto critiche, quali la presenza di autovalori sull'asse immaginario, circa le quali raramente può esserci certezza quando si ha a che fare con modelli, necessariamente approssimati, di una determinata realtà fisica. Per di più, la proprietà di stabilità asintotica comporta alcune conseguenze veramente notevoli sui sistemi che ne godono.

- Innanzitutto, in un sistema asintoticamente stabile, il movimento che si ha per $t \rightarrow \infty$ è indipendente dallo stato iniziale, perché il movimento asintotico coincide con quello forzato, visto che il movimento libero tende ad annullarsi.
- Inoltre, la risposta all'impulso, sia dello stato sia dell'uscita, tende asintoticamente a zero, perché essa, come si è osservato al Paragrafo 3.2.6, per $t > 0$ coincide con un movimento libero.
- Analogamente, la risposta a un qualunque ingresso di durata limitata nel tempo tende a zero in modo asintotico. Infatti, detti \tilde{t} l'istante in cui l'ingresso diventa definitivamente nullo e \tilde{x} il corrispondente valore dello stato, i movimenti del sistema negli istanti di tempo successivi sono quelli liberi generati da \tilde{x} ; si ha cioè

$$x(t) = e^{A(t-\tilde{t})}\tilde{x}, \quad t > \tilde{t}$$

$$y(t) = Ce^{A(t-\tilde{t})}\tilde{x}, \quad t > \tilde{t}$$

Visto poi che $\det(A) \neq 0$, perché un sistema asintoticamente stabile non può avere autovalori nulli, lo stato e l'uscita di equilibrio conseguenti a un qualunque ingresso $u(t) = \bar{u}$ sono unici (si veda il Paragrafo 3.3). A essi tendono i movimenti generati da $u(t) = \bar{u}$ sca(t), qualunque sia lo stato iniziale.

- Infine, un sistema asintoticamente stabile gode anche della proprietà di *stabilità esterna* (o *stabilità BIBO*, dall'inglese *Bounded-Input-Bounded-Output*), cioè produce un movimento forzato dell'uscita limitato in corrispondenza di ogni ingresso limitato. Ricordando le Equazioni (3.3), (3.4), tale proprietà si dimostra facilmente se A è diagonalizzabile sfruttando la (3.37). La dimostrazione si estende al caso di matrici non diagonalizzabili e la proprietà vale anche per il ritardo di tempo (2.30). Viceversa, si può anche verificare che, sotto blande ipotesi, i sistemi stabili esternamente sono anche asintoticamente stabili. Si veda a questo proposito il Teorema 3.17.

3.5 Linearizzazione e stabilità dell'equilibrio di sistemi non lineari

Come anticipato al Paragrafo 2.3, si mostra ora come i risultati presentati nei paragrafi precedenti possano essere sfruttati anche per lo studio di sistemi non lineari e stazionari. Si consideri perciò un sistema non lineare, in generale MIMO, invariante nel tempo e proprio, descritto da

$$\dot{x}(t) = f(x(t), u(t)) \quad (3.54)$$

$$y(t) = g(x(t), u(t)) \quad (3.55)$$

soggetto all'ingresso costante $u(t) = \bar{u}$. Si faccia poi riferimento a un suo stato di equilibrio \bar{x} e alla corrispondente uscita di equilibrio \bar{y} , detti *nominali*, che soddisfano le identità

$$0 = f(\bar{x}, \bar{u}) \quad (3.56)$$

$$\bar{y} = g(\bar{x}, \bar{u}) \quad (3.57)$$

3.5.1 Linearizzazione

Il procedimento della *linearizzazione* consiste nel descrivere il comportamento di un sistema non lineare attorno all'equilibrio nominale mediante un particolare sistema lineare. Quest'ultimo sistema costituisce solo un'approssimazione del sistema originario, ma è estremamente utile per affrontare molti problemi specifici, perché a esso sono applicabili i potenti metodi di analisi e sintesi disponibili per i sistemi lineari.

Se si introducono le variazioni $\delta u(t)$, $\delta x(t)$ e $\delta y(t)$ delle variabili di ingresso, stato e uscita rispetto a \bar{u} , \bar{x} e \bar{y} , nonché dello stato iniziale δx_{t_0} ancora rispetto a \bar{x} , cioè se si pone

$$u(t) = \bar{u} + \delta u(t)$$

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

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

$$x_{t_0} = \bar{x} + \delta x_{t_0}$$

le equazioni (3.54), (3.55) diventano

$$\dot{\bar{x}} + \delta \dot{x}(t) = f(\bar{x} + \delta x(t), \bar{u} + \delta u(t)) \quad (3.58)$$

$$\dot{\bar{y}} + \delta \dot{y}(t) = g(\bar{x} + \delta x(t), \bar{u} + \delta u(t)) \quad (3.59)$$

con la condizione iniziale

$$\bar{x} + \delta x(t_0) = \bar{x} + \delta x_{t_0} \quad (3.60)$$

Supponendo poi che le funzioni f e g siano sufficientemente regolari, esse possono essere sviluppate in serie di Taylor rispetto a x e u in $x = \bar{x}$ e $u = \bar{u}$. Sostituendo questo sviluppo, arrestato ai termini del primo ordine, nelle (3.58), (3.59) e ricordando che \bar{x} non dipende dal tempo, si ottiene

$$\begin{aligned} \dot{\bar{x}}(t) &= f(\bar{x}, \bar{u}) + \left. \frac{\partial f(x, u)}{\partial x} \right|_{x=\bar{x}, u=\bar{u}} \delta x(t) + \left. \frac{\partial f(x, u)}{\partial u} \right|_{x=\bar{x}, u=\bar{u}} \delta u(t) \\ \dot{\bar{y}} + \delta \dot{y}(t) &= g(\bar{x}, \bar{u}) + \left. \frac{\partial g(x, u)}{\partial x} \right|_{x=\bar{x}, u=\bar{u}} \delta x(t) + \left. \frac{\partial g(x, u)}{\partial u} \right|_{x=\bar{x}, u=\bar{u}} \delta u(t) \end{aligned}$$

Da queste ultime equazioni e dalle (3.56), (3.57), (3.60) si ha infine

$$\delta \dot{x}(t) = A \delta x(t) + B \delta u(t) \quad , \quad \delta x(t_0) = \delta x_{t_0} \quad (3.61)$$

$$\delta \dot{y}(t) = C \delta x(t) + D \delta u(t) \quad (3.62)$$

dove

$$A = \left. \frac{\partial f(x, u)}{\partial x} \right|_{x=\bar{x}, u=\bar{u}} \quad (3.63)$$

$$B = \left. \frac{\partial f(x, u)}{\partial u} \right|_{x=\bar{x}, u=\bar{u}} \quad (3.64)$$

$$C = \left. \frac{\partial g(x, u)}{\partial x} \right|_{x=\bar{x}, u=\bar{u}} \quad (3.65)$$

$$D = \left. \frac{\partial g(x, u)}{\partial u} \right|_{x=\bar{x}, u=\bar{u}} \quad (3.66)$$

Si è così ricavato un sistema lineare e stazionario che in senso stretto lega le variazioni prime delle variabili in gioco e che si chiama *sistema linearizzato*. Esso è utilizzabile per descrivere in maniera approssimata il comportamento del sistema (3.54), (3.55) attorno al particolare equilibrio considerato nel caso in cui le variazioni delle funzioni di ingresso e dello stato iniziale $\delta u(t)$ e δx_{t_0} , nonché le variazioni $\delta x(t)$ e $\delta y(t)$ da esse provocate, siano sufficientemente piccole in norma.

Esempio 3.18 (Seguito dell'Esempio 2.16) I sistemi linearizzati attorno alla condizione di equilibrio (2.33)-(2.35) sono descritti dalle Equazioni (3.61)-(3.66) con

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -k/Ml^2 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 0 \\ 1/Ml^2 \end{bmatrix} \quad , \quad C = [Mgl \quad 0] \quad , \quad D = 0 \quad (3.67)$$

Invece, per l'equilibrio (2.36)-(2.38) vi sono due differenti situazioni da esaminare, corrispondenti a valori dell'intero i nelle Equazioni (2.37) e (2.38) pari o dispari, cioè a pendolo orientato verso il basso o verso l'alto.

Nel primo caso il sistema linearizzato (3.61)-(3.66) ha come matrice della dinamica

$$A = \begin{bmatrix} 0 & 1 \\ -g/l & -k/Ml^2 \end{bmatrix} \quad (3.68)$$

Invece nel secondo caso risulta

$$A = \begin{bmatrix} 0 & 1 \\ g/l & -k/Ml^2 \end{bmatrix} \quad (3.69)$$

In entrambi i casi si ha

$$B = \begin{bmatrix} 0 \\ 1/Ml^2 \end{bmatrix}, \quad C = [0 \ 0], \quad D = 0$$

Il fatto che si ottenga $C = 0$ e $D = 0$ significa che, con riferimento alle condizioni di equilibrio (2.36)-(2.38), è nulla la variazione prima dell'energia totale del sistema rispetto a perturbazioni delle altre variabili.

3.5.2 Stabilità dell'equilibrio

Ora si mostrerà in che modo l'analisi della stabilità del sistema linearizzato (3.61)-(3.66) consenta in molti casi di determinare le proprietà di stabilità dello stato di equilibrio \bar{x} del sistema non lineare originario (3.54), (3.55). Si noti che i risultati saranno esatti, malgrado essi si ottengano usando un modello approssimato. Questo fatto in realtà non dovrebbe meravigliare in quanto, per definizione, le proprietà di stabilità sono locali, o "in piccolo", e quindi è ragionevole che esse possano essere studiate facendo uso di modelli linearizzati, almeno in situazioni non critiche.

Teorema 3.9 Lo stato di equilibrio \bar{x} , relativo all'ingresso \bar{u} , del sistema non lineare (3.54)-(3.55) è asintoticamente stabile se tutti gli autovalori del sistema linearizzato corrispondente (3.61)-(3.66) hanno parte reale negativa. ■

Visto che questo teorema richiede la stabilità asintotica del sistema linearizzato, la determinazione degli autovalori può anche essere evitata e sostituita dallo studio del polinomio caratteristico, secondo quanto indicato al Teorema 3.7.

Teorema 3.10 Lo stato di equilibrio \bar{x} , relativo all'ingresso \bar{u} , del sistema non lineare (3.54)-(3.55) è instabile se almeno uno degli autovalori del sistema linearizzato corrispondente (3.61)-(3.66) ha parte reale positiva. ■

Pertanto, mentre il Teorema 3.10 è l'esatta controparte del Teorema 3.5, il Teorema 3.9 rappresenta una "estensione" della sola parte sufficiente del Teorema 3.4. Se quindi il sistema (3.61)-(3.66) ha alcuni autovalori con parte reale negativa e altri con parte reale nulla, niente si può dire circa la stabilità di \bar{x} esaminando solo il sistema linearizzato: essa dipende dai termini successivi al primo dello sviluppo in serie di Taylor di $f(x, u)$ attorno a \bar{x}, \bar{u} . Per affrontare tali situazioni sono disponibili altre tecniche che qui non sono trattate.

Esempio 3.19 (Seguito degli Esempi 2.16, 2.20 e 3.18) Per quanto riguarda gli stati di equilibrio (2.34) i Teoremi 3.9 e 3.10 non consentono di trarre alcuna conclusione circa le proprietà di stabilità, in quanto la matrice A della dinamica del sistema linearizzato riportata nella (3.67) ha un autovalore nullo e uno reale negativo pari a $-k/Ml^2$. Per quanto concerne invece gli stati di equilibrio (2.37) l'applicazione dei Teoremi 3.9 e 3.10 conferma i risultati trovati su basi puramente intuitive nell'Esempio 2.20. Infatti, per valori pari di i la matrice A è riportata nella (3.68) e ha autovalori reali o complessi, ma comunque sempre con parte reale negativa, mentre per valori dispari di i essa è riportata nella (3.69) e ha un autovalore con parte reale positiva.

3.6 Raggiungibilità, osservabilità e scomposizione canonica

In questo paragrafo saranno presentate due nuove proprietà di un sistema dinamico: una riguarda la possibilità di far assumere al suo stato un prefissato valore agendo sull'ingresso, l'altra invece la possibilità di determinare il valore dello stato iniziale a partire dalla conoscenza del movimento della sua uscita. Sulla base di queste proprietà si può effettuare una "scomposizione in parti" del sistema stesso. La scomposizione è tale da consentire di concludere che, per il calcolo del movimento forzato dell'uscita, è sufficiente fare riferimento a una sola di queste "parti", cioè a meno variabili di stato di quante ne occorrono per descrivere completamente il funzionamento del sistema originario.

3.6.1 Esempi introduttivi

Per introdurre le proprietà indicate conviene iniziare a considerare alcuni esempi.

Esempio 3.20 Nel circuito elettrico di Figura 3.7 la variabile di ingresso u è la tensione ai morsetti del generatore, mentre la variabile di uscita y è la tensione ai morsetti di uno dei due condensatori. Poiché il circuito contiene due elementi in grado di accumulare energia, è ragionevole ritenere che il sistema dinamico corrispondente si possa descrivere con un sistema del secondo ordine. Di fatto, prendendo come variabili di stato x_1 e x_2 le tensioni ai morsetti dei due condensatori, si può scrivere

$$\dot{x}_1(t) = -\frac{1}{RC}(x_1(t) + x_2(t) - u(t)) \quad (3.70)$$

$$\dot{x}_2(t) = -\frac{1}{RC}(x_1(t) + x_2(t) - u(t)) \quad (3.71)$$

$$y(t) = x_1(t) \quad (3.72)$$

Come spiegato nei Paragrafi 2.2.4 e 3.2.4, le variabili di stato di un sistema dinamico non sono definite in maniera unica. La libertà di scelta può essere sfruttata, tra l'altro, per sostituire a x_1 e x_2 le due nuove variabili \dot{x}_1 e \dot{x}_2 somma e differenza delle precedenti:

$$\dot{x}_1(t) = x_1(t) + x_2(t) \quad (3.73)$$

$$\dot{x}_2(t) = x_1(t) - x_2(t) \quad (3.74)$$

Si osservi che \dot{x}_1 e \dot{x}_2 possono effettivamente essere prese come variabili di stato, in quanto le Equazioni (3.73), (3.74) stabiliscono una relazione biumivoca tra esse e le variabili utilizzate in precedenza. Il sistema (3.70)-(3.72) è allora equivalente al sistema

$$\dot{\dot{x}}_1(t) = -\frac{2}{RC}(\dot{x}_1(t) - u(t)) \quad (3.75)$$

$$\dot{x}_2(t) = 0 \quad (3.76)$$

$$y(t) = \frac{1}{2}(\dot{x}_1(t) + \dot{x}_2(t)) \quad (3.77)$$

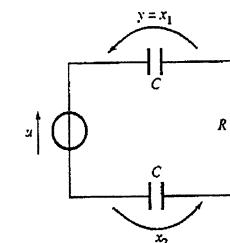
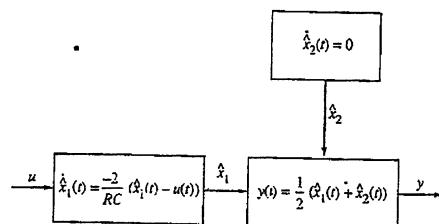


Figura 3.7
Circuito elettrico
dell'Esempio 3.20.

Figura 3.8
Scomposizione in parti del sistema dell'Esempio 3.20.



Le relazioni di dipendenza tra le variabili u , \dot{x}_1 , \dot{x}_2 e y espresse dalle Equazioni (3.75)-(3.77) sono rappresentate schematicamente nella Figura 3.8. In modo specifico è immediato notare che la differenza \dot{x}_2 tra le tensioni ai morsetti dei condensatori non dipende in alcun modo dall'ingresso u (Equazione (3.76)), mentre è facile concludere (Equazione (3.75)) che si può scegliere l'ingresso u in modo che, a partire da stato nullo, la somma delle tensioni \dot{x}_1 assuma in un arbitrario istante di tempo $t > 0$ un qualunque valore \dot{x}_{1fin} . Infatti la soluzione della (3.75) per $\dot{x}_1(0) = 0$ è

$$\dot{x}_1(t) = \frac{2}{RC} \int_0^t e^{-(2/RC)(t-\tau)} u(\tau) d\tau \quad (3.78)$$

Assumendo per esempio che l'ingresso sia del tipo

$$\tilde{u}(t) = k e^{-(2/RC)t}$$

si ottiene

$$\dot{x}_1(t) = \frac{2}{RC} e^{-(2/RC)t} k t$$

e allora, affinché risulti $\dot{x}_1(t) = \dot{x}_{1fin}$, basta porre

$$k = \frac{RC}{2} \dot{x}_{1fin} \frac{e^{(2/RC)t}}{t}$$

Per concludere l'analisi si noti che, come mostra la Figura 3.8, la variabile di stato \dot{x}_2 non è coinvolta nel calcolo del movimento forzato dell'uscita, perché quest'ultimo, in vista dell'Equazione (3.77), risulta la metà di quello della prima variabile di stato fornito dall'Equazione (3.78).

Esempio 3.21 La Figura 3.9 rappresenta un corpo di massa M in moto su una guida rettilinea caratterizzata da un coefficiente d'attrito viscoso h . Scegliendo come variabile di ingresso u la forza esterna applicata al corpo, come variabili di stato x_1 e x_2 la velocità e la posizione del corpo e come variabile di uscita ancora la velocità, si può scrivere

$$\dot{x}_1(t) = -\frac{h}{M}x_1(t) + \frac{1}{M}u(t) \quad (3.79)$$

$$\dot{x}_2(t) = x_1(t) \quad (3.80)$$

$$y(t) = x_1(t) \quad (3.81)$$

Le relazioni di dipendenza tra le variabili u , x_1 , x_2 e y espresse dalle Equazioni (3.79)-(3.81) sono rappresentate schematicamente nella Figura 3.10. In particolare, si può notare che l'esame di un qualunque transitorio dell'uscita non permette in alcun modo di ricavare informazioni circa la posizione x_2 all'istante iniziale in quanto la velocità y non dipende da quest'ultima variabile: in

Figura 3.9
Corpo dell'Esempio 3.21.

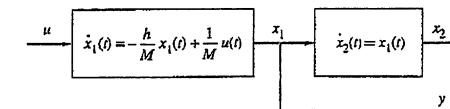
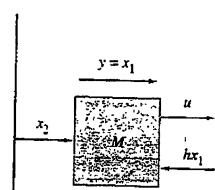


Figura 3.10
Scomposizione in parti del sistema dell'Esempio 3.21.

realità il calcolo del movimento dell'uscita coinvolge solo le Equazioni (3.79) e (3.81) nelle quali x_2 non appare. Viceversa, la conoscenza dell'uscita y consente per l'Equazione (3.81) di determinare banalmente il valore iniziale della variabile di stato x_1 . Si osservi che se, contrariamente a quanto fatto sopra, si prendesse come variabile di uscita la posizione, se cioè si ponesse

$$y(t) = x_2(t) \quad (3.82)$$

la determinazione dello stato iniziale a partire dalla conoscenza del movimento dell'uscita sarebbe possibile in quanto

$$\begin{aligned} x_1(0) &= \dot{x}_2(0) = \dot{y}(0) \\ x_2(0) &= y(0) \end{aligned}$$

Il sistema (3.79), (3.80), (3.82) è rappresentato schematicamente nella Figura 3.11.

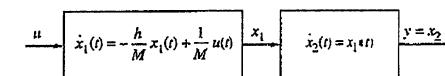


Figura 3.11
Scomposizione in parti del sistema modificato dell'Esempio 3.21.

Esempio 3.22 I due serbatoi della Figura 3.12 sono cilindrici con area di base S_1 e S_2 . Essi sono alimentati da un'unica pompa la cui portata volumetrica si suddivide in parti uguali tra loro. La variabile di uscita y è il volume totale di fluido presente nei serbatoi al di sopra della quota assunta come zero nella misurazione dei livelli x_1 e x_2 . Supponendo che non si verifichino fenomeni di trabocchamento o svuotamento, il funzionamento dei serbatoi è retto dalle equazioni

$$\dot{x}_1(t) = \frac{1}{2S_1}u(t) \quad (3.83)$$

$$\dot{x}_2(t) = \frac{1}{2S_2}u(t) \quad (3.84)$$

$$y(t) = S_1x_1(t) + S_2x_2(t) \quad (3.85)$$

Se ora si scelgono come nuove variabili di stato la somma e la differenza dei volumi di fluido nei due serbatoi, se cioè si pone

$$\dot{x}_1(t) = S_1x_1(t) + S_2x_2(t) \quad (3.86)$$

$$\dot{x}_2(t) = S_1x_1(t) - S_2x_2(t) \quad (3.87)$$

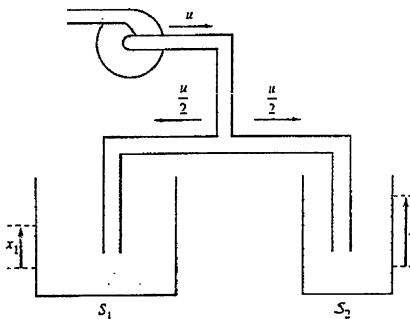
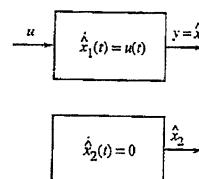


Figura 3.12
Serbatoi dell'Esempio 3.22.

Figura 3.13
Scomposizione in parti del sistema dell'Esempio 3.22.



il sistema (3.83)-(3.85) prende la forma

$$\dot{x}_1(t) = u(t) \quad (3.88)$$

$$\dot{x}_2(t) = 0 \quad (3.89)$$

$$y(t) = \dot{x}_1(t) \quad (3.90)$$

Queste equazioni sono rappresentate nella Figura 3.13 e hanno un'interpretazione molto semplice: la portata u influenza esclusivamente sul volume totale $\dot{x}_1 = y$ di fluido nei serbatoi, mentre la differenza dei volumi parziali \dot{x}_2 è indipendente dall'ingresso e non contribuisce in alcun modo a determinare l'uscita. Il calcolo del movimento forzato coinvolge dunque solo le Equazioni (3.88), (3.90).

Gli esempi precedenti hanno mostrato come mediante opportuni cambi di variabili sia possibile in generale mettere in evidenza "parti" di un sistema dinamico che non sono influenzate dalla variabile di ingresso oppure non influenzano la variabile di uscita. Per formalizzare quanto visto sopra ci si riferirà al sistema dinamico di ordine n con m ingressi e p uscite

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3.91)$$

$$y(t) = Cx(t) + Du(t) \quad (3.92)$$

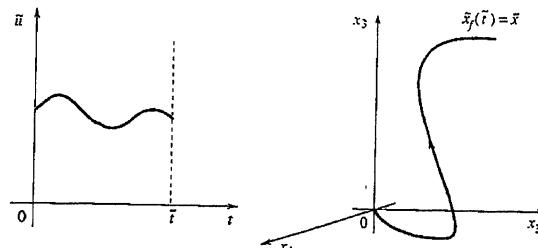
3.6.2 Raggiungibilità

Si consideri la proprietà del movimento forzato dello stato espressa dalla seguente definizione e simbolicamente illustrata nella Figura 3.14.

Definizione 3.1 Uno stato \tilde{x} del sistema (3.91), (3.92) si dice *raggiungibile* se esistono un istante di tempo finito $\tilde{t} > 0$ e un ingresso \tilde{u} , definito tra 0 e \tilde{t} , tali che, detto $\tilde{x}_f(\tilde{t})$, $0 \leq t \leq \tilde{t}$, il movimento forzato dello stato generato da \tilde{u} , risulti $\tilde{x}_f(\tilde{t}) = \tilde{x}$.

Un sistema i cui stati siano tutti raggiungibili si dice *completamente raggiungibile*. ■

Figura 3.14
Raggiungibilità.



In altri termini, quindi, un particolare vettore \tilde{x} costituisce uno stato raggiungibile se è possibile, con un'opportuna scelta dell'ingresso, trasferirvi dall'origine lo stato del sistema in un tempo finito arbitrario. Si osservi che la trasformazione d'uscita (3.92) non gioca in realtà alcun ruolo circa la proprietà di raggiungibilità. Quest'ultima dipende esclusivamente dall'Equazione (3.91), cosicché a volte essa viene attribuita addirittura alla coppia (A, B) .

Per accertare se un dato sistema gode della proprietà di completa raggiungibilità si può applicare il successivo Teorema 3.11 che fa riferimento alla cosiddetta *matrice di raggiungibilità*, definita come

$$M_r = [B \ AB \ A^2B \ \cdots \ A^{n-1}B] \in R^{n \times mn}$$

Teorema 3.11 Il sistema (3.91), (3.92) è completamente raggiungibile, ovvero la coppia (A, B) è completamente raggiungibile, se e solo se il rango della matrice di raggiungibilità è pari a n , cioè

$$\rho(M_r) = n \quad (3.93)$$

Se il sistema ha un solo ingresso, cioè $m = 1$, la matrice M_r è quadrata e la condizione necessaria e sufficiente (3.93) è equivalente a $\det(M_r) \neq 0$.

Nel caso in cui un sistema non sia completamente raggiungibile, si può isolare la sua "parte" dotata della proprietà di raggiungibilità, così come specificato nel teorema seguente dove $n_r = \rho(M_r)$.

Teorema 3.12 Mediante un opportuno, non unico, cambio di variabili di stato

$$\hat{x}(t) = T_r x(t)$$

l'equazione di stato (3.91) può essere posta nella forma

$$\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t) \quad (3.94)$$

dove

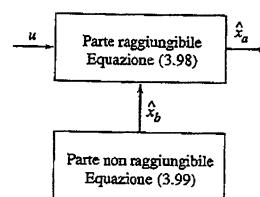
$$\hat{A} = \begin{bmatrix} \hat{A}_a & \hat{A}_{ab} \\ 0 & \hat{A}_b \end{bmatrix}, \quad \hat{A}_a \in R^{n_r \times n_r} \quad (3.95)$$

$$\hat{B} = \begin{bmatrix} \hat{B}_a \\ 0 \end{bmatrix}, \quad \hat{B}_a \in R^{n_r \times m} \quad (3.96)$$

$$\rho([\hat{B}_a \ \hat{A}_a \hat{B}_a \ \hat{A}_a^2 \hat{B}_a \ \cdots \ \hat{A}_a^{n_r-1} \hat{B}_a]) = n_r \quad (3.97)$$

Per costruire la matrice T_r si selezionino innanzitutto n_r colonne linearmente indipendenti in M_r . Ognuna di tali colonne rappresenta uno stato raggiungibile ed esse, nel complesso, descrivono gli stati raggiungibili, nel senso che sono raggiungibili tutti e soli gli stati ottenibili combinando linearmente queste colonne. Esse, eventualmente moltiplicate per costanti non nulle, vengono poste in T_r^{-1} dove precedono altre $n - n_r$ colonne arbitrarie, ma tali che risulti $\det(T_r^{-1}) \neq 0$. La procedura ha sempre buon esito in virtù della (3.97).

Figura 3.15
Parte raggiungibile e parte non raggiungibile.



Se si partitiona il vettore \hat{x} come

$$\hat{x} = \begin{bmatrix} \hat{x}_a \\ \hat{x}_b \end{bmatrix}, \quad \hat{x}_a \in R^{n_a}$$

l'Equazione (3.94), grazie alle Equazioni (3.95), (3.96), si può scrivere nella forma

$$\dot{\hat{x}}_a(t) = \hat{A}_a \hat{x}_a(t) + \hat{A}_{ab} \hat{x}_b(t) + \hat{B}_a u(t) \quad (3.98)$$

$$\dot{\hat{x}}_b(t) = \hat{A}_b \hat{x}_b(t) \quad (3.99)$$

La struttura di queste equazioni mostra con chiarezza che tutti i movimenti forzati producono $\dot{\hat{x}}_b(t) = 0, t > 0$, in quanto u non influisce né direttamente, né tramite \hat{x}_a sulla derivata di \hat{x}_b e, d'altra parte, un'opportuna scelta di u consente per la (3.97) di far assumere un qualunque valore al vettore \hat{x}_a a un istante di tempo finito. Si dirà pertanto che l'Equazione (3.98) costituisce la *parte raggiungibile* del sistema (3.91), (3.92) e l'Equazione (3.99) la *parte non raggiungibile*. La Figura 3.15 illustra sinteticamente la struttura delle equazioni.

La matrice \hat{A} è triangolare a blocchi e perciò i suoi autovalori sono quelli dei blocchi sulla diagonale \hat{A}_a e \hat{A}_b : si usa dire che gli autovalori di \hat{A}_a sono quelli della parte raggiungibile e gli autovalori di \hat{A}_b quelli della parte non raggiungibile.

Esempio 3.23 (Seguito dell'Esempio 3.20) Per le Equazioni (3.70), (3.71) risulta

$$M_r = [B \ AB] = \begin{bmatrix} 1/RC & -2/(RC)^2 \\ 1/RC & -2/(RC)^2 \end{bmatrix}$$

Questa matrice ha determinante nullo, la qual cosa conferma, per il Teorema 3.11, la non completa raggiungibilità del sistema dinamico. Per separare la parte raggiungibile da quella non raggiungibile si può scegliere

$$T_r^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (3.100)$$

la cui prima colonna coincide con una qualunque delle colonne di M_r moltiplicata per una costante. Il fatto che gli elementi della prima colonna della matrice di trasformazione (3.100) coincidano è connesso al fatto che, come già spiegato, gli stati raggiungibili sono caratterizzati dall'ugualanza delle componenti x_1 e x_2 . Si ottengono così le Equazioni (3.75), (3.76), da cui si riconosce che l'autovalore della parte raggiungibile è $-2/RC$ mentre quello della parte non raggiungibile è 0.

Esempio 3.24 (Seguito dell'Esempio 3.22) Dalle Equazioni (3.83), (3.84) si trova

$$M_r = [B \ AB] = \begin{bmatrix} 1/2S_1 & 0 \\ 1/2S_2 & 0 \end{bmatrix}$$

Gli stati raggiungibili del sistema (3.83)-(3.85) sono quelli corrispondenti a pari volumi di fluido nei due serbatoi e come matrice di trasformazione si può scegliere

$$T_r^{-1} = \begin{bmatrix} 1/2S_1 & 1/2S_1 \\ 1/2S_2 & -1/2S_2 \end{bmatrix}$$

che corrisponde al cambio di variabili (3.86), (3.87). Dalle Equazioni (3.88), (3.89) si riconosce infine che sia l'autovalore della parte raggiungibile sia quello della parte non raggiungibile sono nulli.

Per concludere, si noti che la proprietà di raggiungibilità di uno stato \tilde{x} coincide, per la classe di sistemi dinamici qui considerata, con quella di *controllabilità*, consistente nella possibilità di portare lo stato del sistema da $x(0) = \tilde{x}$ a $x = 0$ in un tempo arbitrario finito mediante un'opportuna scelta dell'ingresso.

3.6.3 Osservabilità

Si può ora passare a considerare una seconda proprietà del sistema (3.91), (3.92), questa volta relativa al movimento libero dell'uscita. Essa è simbolicamente illustrata nella Figura 3.16 ed è definita come segue.

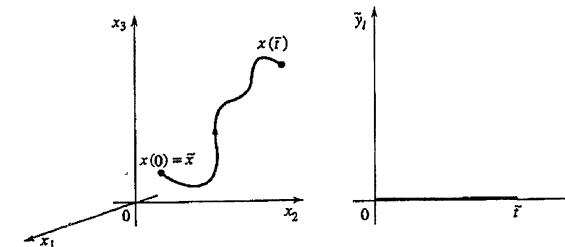


Figura 3.16
Non osservabilità.

Definizione 3.2 Uno stato $\tilde{x} \neq 0$ del sistema (3.91), (3.92) si dice *non osservabile* se, qualunque sia $\tilde{t} > 0$ finito, detto $\tilde{y}_l(t), t \geq 0$, il movimento libero dell'uscita generato da \tilde{x} , risulta $\tilde{y}_l(t) = 0, 0 \leq t \leq \tilde{t}$.

Un sistema privo di stati non osservabili si dice *completamente osservabile*. ■

In altri termini, un particolare vettore \tilde{x} costituisce uno stato non osservabile se l'esame di un tratto di qualunque durata del movimento libero dell'uscita da esso generata non consente di distinguerlo dal vettore $x = 0$, che ovviamente genera un movimento libero dell'uscita identicamente nullo. Si osservi che le matrici B e D delle Equazioni (3.91), (3.92) non hanno alcun ruolo circa la proprietà di osservabilità; quest'ultima viene attribuita addirittura alla coppia (A, C) .

Per accettare se un dato sistema gode della proprietà di completa osservabilità, si può utilizzare il Teorema 3.13 che fa riferimento alla cosiddetta *matrice di osservabilità*, definita come

$$M_o = [C' \ A'C' \ A^2C' \ \dots \ A^{n-1}C'] \in R^{n \times p n}$$

dove l'apice indica l'operazione di trasposizione.

Teorema 3.13 Il sistema (3.91), (3.92) è completamente osservabile, ovvero la coppia (A, C) è completamente osservabile, se e solo se il rango della matrice di osservabilità è pari a n , cioè

$$\rho(M_o) = n \quad (3.101)$$

Se il sistema ha una sola uscita, cioè $p = 1$, la matrice M_o è quadrata e la condizione necessaria e sufficiente (3.101) è equivalente a $\det(M_o) \neq 0$.

Nel caso in cui un sistema non sia completamente osservabile, si può isolare la sua "parte" dotata della proprietà di non osservabilità, così come specificato nel teorema seguente, dove $n_o = \rho(M_o)$.

Teorema 3.14 Mediante un opportuno, non unico, cambio di variabili di stato

$$\dot{\hat{x}}(t) = T_o x(t)$$

il sistema (3.91), (3.92) con ingresso nullo può essere posto nella forma

$$\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) \quad (3.102)$$

$$y(t) = \hat{C}\hat{x}(t) \quad (3.103)$$

dove

$$\hat{A} = \begin{bmatrix} \hat{A}_a & 0 \\ \hat{A}_{ba} & \hat{A}_b \end{bmatrix}, \quad \hat{A}_a \in R^{n_o \times n_o} \quad (3.104)$$

$$\hat{C} = [\hat{C}_a \ 0], \quad \hat{C}_a \in R^{p \times n_o} \quad (3.105)$$

$$\rho([\hat{C}'_a \hat{A}'_a \hat{C}'_a \hat{A}'^2 \hat{C}'_a \dots \hat{A}'^{n_o-1} \hat{C}'_a]) = n_o \quad (3.106)$$

Per costruire la matrice T_o si selezionino innanzitutto $n - n_o$ colonne linearmente indipendenti ζ_i tali che

$$M'_o \zeta_i = 0$$

Ognuno di questi vettori rappresenta uno stato non osservabile ed essi nel complesso descrivono gli stati non osservabili, nel senso che sono non osservabili tutti e soli gli stati ottenuti combinando linearmente questi vettori. Essi, moltiplicati eventualmente per costanti non nulle, vengono posti come colonne in T_o^{-1} preceduti da altre n_o colonne arbitrarie, ma tali che risulti $\det(T_o^{-1}) \neq 0$. La procedura ha sempre buon esito in virtù della (3.106).

Se si partitiona il vettore \hat{x} come

$$\hat{x} = \begin{bmatrix} \hat{x}_a \\ \hat{x}_b \end{bmatrix}, \quad \hat{x}_a \in R^{n_o}$$

le Equazioni (3.102), (3.103), grazie alle Equazioni (3.104), (3.105), si possono scrivere nella forma

$$\dot{\hat{x}}_a(t) = \hat{A}_a \hat{x}_a(t) \quad (3.107)$$

$$\dot{\hat{x}}_b(t) = \hat{A}_{ba} \hat{x}_a(t) + \hat{A}_b \hat{x}_b(t) \quad (3.108)$$

$$y(t) = \hat{C}_a \hat{x}_a(t) \quad (3.109)$$

La struttura di queste equazioni mostra con chiarezza che tutti i movimenti liberi dell'uscita generati da stati iniziali con $\hat{x}_a(0) = 0$ sono identicamente nulli, in quanto $\dot{\hat{x}}_b$ non influenza né direttamente, né tramite \hat{x}_a su y , e, d'altra parte, l'esame di un qualunque transitorio di y consente per la (3.106) di determinare $\hat{x}_a(0)$. Si dirà pertanto che l'Equazione (3.107) costituisce la *parte osservabile* del sistema (3.91), (3.92) e l'Equazione (3.108) la *parte non osservabile*. La Figura 3.17 illustra sinteticamente la struttura delle equazioni.

Anche qui è importante notare che gli autovalori della matrice \hat{A} sono quelli di \hat{A}_a , detti autovalori della parte osservabile, e quelli di \hat{A}_b , detti autovalori della parte non osservabile.

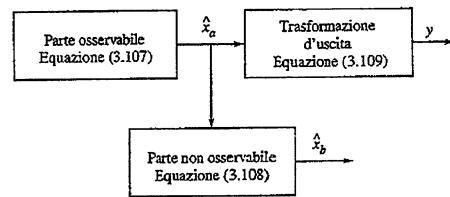


Figura 3.17
Parte osservabile e parte non osservabile.

Esempio 3.25 (Seguito dell'Esempio 3.21) Dalle Equazioni (3.79)-(3.81) risulta

$$M_o = [C' \ A'C'] = \begin{bmatrix} 1 & -h/M \\ 0 & 0 \end{bmatrix}$$

e infatti la discussione dell'Esempio 3.21 ha illustrato come siano non osservabili tutti e soli gli stati con $x_1 = 0$ e $x_2 \neq 0$. Il sistema è già nella forma indicata dal Teorema 3.13, per cui $T_o = I$, mentre gli autovalori della parte osservabile e di quella non osservabile sono rispettivamente $-h/M$ e 0. Il sistema (3.79), (3.80), (3.82) è invece completamente osservabile perché

$$M_o = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Esempio 3.26 (Seguito dell'Esempio 3.22) Dalle Equazioni (3.83)-(3.85) si ha

$$M_o = [C' \ A'C'] = \begin{bmatrix} S_1 & 0 \\ S_2 & 0 \end{bmatrix}$$

Questa matrice ha determinante nullo e pertanto, per il Teorema 3.13, il sistema dinamico corrispondente non è completamente osservabile; è infatti già noto che un'analisi dell'uscita non consente di distinguere dallo stato nullo tutti e soli gli stati corrispondenti a volumi di fluido non nulli nei due serbatoi, ma con somma nulla. Il cambio di variabili che separa la parte non osservabile da quella osservabile può essere quello specificato dalle Equazioni (3.86), (3.87), e l'esame delle Equazioni (3.88), (3.89) consente di affermare che sia l'autovalore della parte osservabile sia quello della parte non osservabile sono nulli.

Per concludere, si noti che la proprietà di non osservabilità di uno stato \tilde{x} coincide, per la classe di sistemi dinamici qui considerata, con quella di *non ricostruibilità*, consistente nell'impossibilità di distinguere lo stato finale (anziché iniziale) \tilde{x} di un sistema da quello nullo mediante l'analisi di un transitorio libero dell'uscita di qualunque durata.

3.6.4 Scomposizione canonica e forma minima

Un sistema dinamico può essere sia non completamente raggiungibile sia non completamente osservabile. In questo caso, come è facile intuire, è possibile scomporre il sistema stesso come specificato nel teorema seguente.

Teorema 3.15 Mediante un opportuno, non unico, cambio di variabili di stato

$$\dot{\hat{x}}(t) = T_K x(t)$$

il sistema (3.91), (3.92) può essere posto nella forma

$$\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t) \quad (3.110)$$

$$y(t) = \hat{C}\hat{x}(t) + Du(t) \quad (3.111)$$

dove

$$\hat{x} = \begin{bmatrix} \hat{x}_a \\ \hat{x}_b \\ \hat{x}_c \\ \hat{x}_d \end{bmatrix} \quad (3.112)$$

$$\hat{A} = \begin{bmatrix} \hat{A}_a & \hat{A}_{ab} & \hat{A}_{ac} & \hat{A}_{ad} \\ 0 & \hat{A}_b & 0 & \hat{A}_{bd} \\ 0 & 0 & \hat{A}_c & \hat{A}_{cd} \\ 0 & 0 & 0 & \hat{A}_d \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \hat{B}_a \\ \hat{B}_b \\ 0 \\ 0 \end{bmatrix}, \quad \hat{C} = [0 \ \hat{C}_b \ 0 \ \hat{C}_d] \quad (3.113)$$

e le dimensioni delle sottomatrici in cui sono stati partizionati \hat{x} , \hat{A} , \hat{B} e \hat{C} sono congruenti tra loro. Inoltre il sottosistema

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

è completamente raggiungibile e osservabile; il sottosistema

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}}_a(t) \\ \dot{\hat{x}}_b(t) \end{bmatrix} &= \begin{bmatrix} \hat{A}_a & \hat{A}_{ab} \\ 0 & \hat{A}_b \end{bmatrix} \begin{bmatrix} \hat{x}_a(t) \\ \hat{x}_b(t) \end{bmatrix} + \begin{bmatrix} \hat{B}_a \\ \hat{B}_b \end{bmatrix} u(t) \\ y(t) &= \hat{C}_b \hat{x}_b(t) + D u(t) \end{aligned}$$

ha ordine n_r ed è completamente raggiungibile; il sottosistema

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}}_b(t) \\ \dot{\hat{x}}_d(t) \end{bmatrix} &= \begin{bmatrix} \hat{A}_b & \hat{A}_{bd} \\ 0 & \hat{A}_d \end{bmatrix} \begin{bmatrix} \hat{x}_b(t) \\ \hat{x}_d(t) \end{bmatrix} + \begin{bmatrix} \hat{B}_b \\ 0 \end{bmatrix} u(t) \\ y(t) &= \hat{C}_b \hat{x}_b(t) + \hat{C}_d \hat{x}_d(t) + D u(t) \end{aligned}$$

ha ordine n_o ed è completamente osservabile.

La matrice T_K si determina con un algoritmo che, in sostanza, consiste nell'effettuare prima una scomposizione basata, per esempio, sulla proprietà di raggiungibilità e poi una seconda scomposizione connessa alla proprietà di non osservabilità.

Le Equazioni (3.110), (3.111), grazie alle (3.112), (3.113), si possono scrivere nella forma

$$\dot{\hat{x}}_a(t) = \hat{A}_a \hat{x}_a(t) + \hat{A}_{ab} \hat{x}_b(t) + \hat{A}_{ac} \hat{x}_c(t) + \hat{A}_{ad} \hat{x}_d(t) + \hat{B}_a u(t) \quad (3.114)$$

$$\dot{\hat{x}}_b(t) = \hat{A}_b \hat{x}_b(t) + \hat{A}_{bd} \hat{x}_d(t) + \hat{B}_b u(t) \quad (3.115)$$

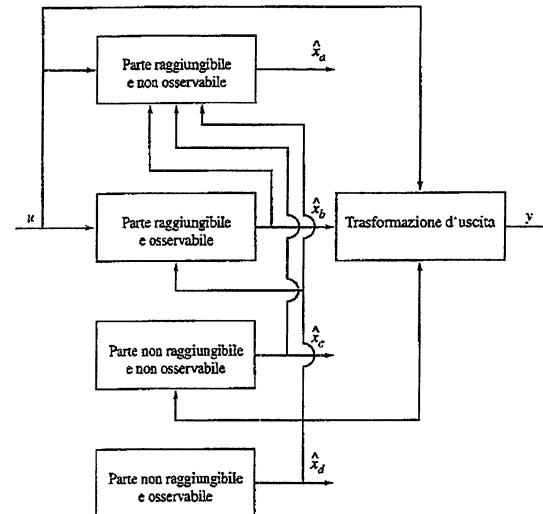
$$\dot{\hat{x}}_c(t) = \hat{A}_c \hat{x}_c(t) + \hat{A}_{cd} \hat{x}_d(t) \quad (3.116)$$

$$\dot{\hat{x}}_d(t) = \hat{A}_d \hat{x}_d(t) \quad (3.117)$$

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

e, alla luce del Teorema 3.15, si usa dire che l'Equazione (3.114) costituisce la parte raggiungibile e non osservabile del sistema, la (3.115) la parte raggiungibile e osservabile, la (3.116) la parte non raggiungibile e non osservabile e la (3.117)

Figura 3.18
Scomposizione canonica.



la parte non raggiungibile e osservabile. Questa scomposizione viene detta *canonica*, o *di Kalman*, dal nome dell'autore del Teorema 3.15, ed è rappresentata sinteticamente nella Figura 3.18.

Gli autovalori della matrice \hat{A} , che è triangolare a blocchi, sono la riunione di quelli delle matrici \hat{A}_a , \hat{A}_b , \hat{A}_c e \hat{A}_d . Sono detti autovalori della parte raggiungibile e non osservabile, della parte raggiungibile e osservabile e così via. Analogamente, con ovvio significato dei termini, si parla di stabilità, stabilità asintotica e instabilità delle varie parti in cui il sistema è scomposto.

Esempio 3.27 (Seguito degli Esempi 3.20-3.26) È ovvio che non sempre in un sistema dinamico sono presenti tutte le quattro parti sopra nominate.

- Il sistema (3.70)-(3.72) è completamente osservabile, per cui nella sua scomposizione canonica mancheranno le parti non osservabili: essa è costituita dalle Equazioni (3.75)-(3.77).
- Il sistema (3.79)-(3.81) manca invece della parte non raggiungibile: la sua scomposizione canonica è data, per esempio, dalle Equazioni (3.79)-(3.81) stesse.
- Il sistema (3.79), (3.80), (3.82) è completamente raggiungibile e osservabile.
- Il sistema (3.83)-(3.85) infine ha solo la parte raggiungibile e osservabile e quella non raggiungibile e non osservabile: la corrispondente scomposizione canonica è quella riportata nelle Equazioni (3.88)-(3.90).

Un esame della Figura 3.18 mostra che, delle quattro parti in cui è scomposto il sistema, l'unica affetta dalla variabile di ingresso e che influenza la variabile di uscita è la parte raggiungibile e osservabile. Ciò è confermato da un'analisi delle Equazioni (3.110)-(3.113). Infatti, grazie alla triangolarità a blocchi di \hat{A} , si ha

$$e^{\hat{A}t} = \begin{bmatrix} e^{\hat{A}_a t} & ? & ? & ? \\ 0 & e^{\hat{A}_b t} & ? & ? \\ 0 & 0 & e^{\hat{A}_c t} & ? \\ 0 & 0 & 0 & e^{\hat{A}_d t} \end{bmatrix}$$

dove col simbolo “?” si sono indicati termini che non interessa precisare. Perciò

$$\hat{C}e^{\hat{A}t}\hat{B} = \hat{C}_b e^{\hat{A}_b t} \hat{B}_b$$

mentre per l'Equazione (3.15) si ha

$$\hat{C}e^{\hat{A}t}\hat{B} = CT_K^{-1}e^{T_K\hat{A}T_K^{-1}}T_K B = CT_K^{-1}T_K e^{A_t}T_K^{-1}T_K B = Ce^{At}B$$

e quindi

$$Ce^{At}B = \hat{C}_b e^{\hat{A}_b t} \hat{B}_b \quad (3.118)$$

L'Equazione (3.118) consente di affermare che la risposta all'impulso dell'uscita (3.34) di un sistema dipende esclusivamente dalla sua parte raggiungibile e osservabile e non dalle altre tre parti in cui il sistema si può scomporre. Allora, ricordando l'Equazione (3.35), si può enunciare il seguente risultato.

Teorema 3.16 Il movimento forzato dell'uscita del sistema (3.91), (3.92), tra tutte le quattro parti della scomposizione canonica di quest'ultimo, dipende solo dalla parte raggiungibile e osservabile ed è dato da

$$\begin{aligned} y_f(t) &= g_y(t) * u(t) = \int_0^t g_y(t-\tau) u(\tau) d\tau = \\ &= \hat{C}_b \int_0^t e^{\hat{A}_b(t-\tau)} \hat{B}_b u(\tau) d\tau + Du(t) \end{aligned} \quad (3.119)$$

Esempio 3.28 (Seguito degli Esempi 3.20, 3.23 e 3.27) Il sistema retto dalle Equazioni (3.70)-(3.72) ha come risposta all'impulso dell'uscita la funzione

$$g_y(t) = \frac{1}{RC} e^{-(2/RC)t}$$

facilmente calcolabile sia sulla base delle Equazioni (3.70)-(3.72) stesse, sia anche sulla base delle Equazioni (3.73)-(3.77). Il movimento forzato dell'uscita prodotto dall'ingresso $u(t) = \text{sca}(t)$, per l'Equazione (3.35), o anche per la (3.119), è

$$y_f(t) = g_y(t) * u(t) = \frac{1}{RC} \int_0^t e^{-(2/RC)(t-\tau)} d\tau = \frac{1}{2} (1 - e^{-(2/RC)t})$$

Un sistema raggiungibile e osservabile si dice essere *in forma minima*, in quanto non è possibile adoperare un numero di variabili di stato minore del suo ordine per descrivere la relazione tra ingresso e uscita che esso stabilisce. Alla luce del Teorema 3.16, la risposta all'impulso dell'uscita costituisce un modello *ingresso-uscita* per i sistemi dinamici in forma minima, nel senso che la sola conoscenza di g_y consente di calcolare direttamente il movimento forzato dell'uscita tramite l'Equazione (3.119), senza la necessità di utilizzare, o anche solo definire, le variabili di stato.

In altri termini le parti non raggiungibili e/o non osservabili di un sistema sono sovrabbondanti se esso deve essere utilizzato esclusivamente per lo studio del movimento forzato dell'uscita. Queste parti sono, insomma, il prodotto di un “errore di modellizzazione” nel quale si può incorrere in casi complessi. Possono anche nascere quando un sistema è il risultato dell’interconnessione di più sottosistemi; si veda a questo proposito il Paragrafo 5.5. Naturalmente, le parti non

raggiungibili e/o non osservabili non sono affatto sovrabbondanti nello studio del movimento libero dell'uscita e/o del movimento dello stato, come, per esempio, nell'analisi della stabilità.

Riguardo ai legami tra la stabilità asintotica e la stabilità esterna vale infine la proprietà seguente.

Teorema 3.17 Si assuma che il sistema (3.91), (3.92) sia in forma minima. Allora esso è esternamente stabile se e solo se è asintoticamente stabile. ■

3.7 Conclusioni

I sistemi dinamici a tempo continuo lineari e stazionari costituiscono una famiglia di modelli per i quali è possibile formulare una ricca serie di semplici e potenti risultati teorici. In particolare, in questo capitolo è stata presentata la formula esplicita per il calcolo del movimento, libero e forzato, e si è messo in evidenza che il movimento libero è costituito da una combinazione lineare dei modi del sistema, strettamente connessi con gli autovalori. Dopo aver osservato che in un sistema lineare e stazionario le proprietà di stabilità di qualsiasi movimento dipendono solo dalle caratteristiche del movimento libero, si è arrivati a concludere che la stabilità asintotica dipende dal segno della parte reale degli autovalori. In questo modo, un problema di natura differenziale, quale quello dell'analisi di stabilità, è stato ricondotto a uno studio puramente algebrico sulle proprietà del polinomio caratteristico. Gli strumenti matematici, più o meno classici, sull'algebra dei polinomi hanno poi consentito di formulare efficaci criteri di stabilità, anche nel caso di sistemi soggetti a incertezze.

Successivamente si è mostrato che i metodi di analisi introdotti possono essere utilizzati anche per lo studio dei sistemi non lineari e stazionari, grazie alla procedura della linearizzazione, e si è visto come i risultati sulla stabilità dei sistemi lineari permettano di studiare anche la stabilità degli stati di equilibrio dei sistemi non lineari.

L'ultima parte del capitolo è stata dedicata a una concisa illustrazione delle proprietà di raggiungibilità e di osservabilità, che permettono di introdurre la scomposizione canonica di un sistema e di discutere la proprietà di minimalità di una rappresentazione di stato. Ai fini di una descrizione del legame ingresso-uscita, ciò che conta è soltanto la parte raggiungibile e osservabile del sistema.

Lo studio finora condotto ha fatto riferimento a una formulazione dei sistemi dinamici lineari e stazionari che si dice nel dominio del tempo, perché tutte le variabili in gioco sono espresse in maniera assolutamente naturale come funzioni della variabile reale tempo. Nei Capitoli 4-6 si mostrerà che esiste anche una rappresentazione alternativa che fa riferimento alle trasformate di Laplace e di Fourier delle variabili del sistema: essa sarà detta nel dominio della variabile complessa, o nel dominio della frequenza.

Esercizi

Esercizio 3.1 Si consideri un sistema lineare e stazionario con matrice della dinamica

$$A = \begin{bmatrix} -9 & 4 & 2 \\ -1 & -1 & 0 \\ -17 & 10 & 3 \end{bmatrix}$$

Si determinino i suoi modi e quindi il movimento libero dello stato conseguente a un qualunque stato iniziale.

Esercizio 3.2 Si costruisca un sistema, di ordine minimo possibile, dotato almeno dei modi e^{-t} , t , e^{jt} . Quindi si valutino le sue proprietà di stabilità.

Esercizio 3.3 Si consideri il sistema di equazione di stato

$$\dot{x}_1(t) = x_2(t) \quad (3.120)$$

$$\dot{x}_2(t) = -x_1(t) \quad (3.121)$$

e si valutino le sue proprietà di stabilità. Quindi si verifichi il risultato trovato analizzando le traiettorie percorse dallo stato del sistema nel piano di ascissa x_1 e ordinata x_2 .

(Suggerimento: si dividia la (3.120) per la (3.121) e si integri l'equazione così trovata.)

Esercizio 3.4 Si calcoli la risposta all'impulso dell'uscita del sistema

$$\dot{x}_1(t) = -x_1(t) + u(t)$$

$$\dot{x}_2(t) = -2x_2(t) + 5u(t)$$

$$y(t) = x_1(t) + 3x_2(t)$$

e si verifichi che è asintoticamente nulla, essendo il sistema asintoticamente stabile. Quindi, in base a essa, si calcoli la risposta dell'uscita del sistema all'ingresso $u(t) = r$.

Si calcoli poi il guadagno statico del sistema.

Esercizio 3.5 Si individuino i valori di α e β che rendono asintoticamente stabile un sistema di polinomio caratteristico

$$\varphi(s) = s^3 + \alpha s^2 + \beta s + 1$$

e si determini la regione a loro corrispondente nel piano di ascissa α e ordinata β .

Esercizio 3.6 Si consideri il polinomio caratteristico

$$\varphi(s) = \alpha s^3 + \beta s^2 + s + \gamma$$

e si dica se è vero che tutte le sue radici hanno parte reale negativa qualunque valore assumano α , β e γ negli intervalli seguenti:

$$1 \leq \alpha \leq 2$$

$$2 \leq \beta \leq 4$$

$$4 \leq \gamma \leq 6$$

Esercizio 3.7 Si consideri il sistema non lineare

$$\dot{x}(t) = x^2(t) - u(t)x(t) - 2u(t)$$

$$y(t) = x^3(t) + u^3(t)$$

con ingresso $u(t) = \bar{u} = 1$. Si determinino gli stati di equilibrio, con le loro proprietà di stabilità, e le corrispondenti uscite. Per i sistemi linearizzati attorno a tali equilibri, si calcolino quindi le risposte $y(t)$ all'ingresso $\delta u(t) = 0.1 \sin(0.25t)$.

Esercizio 3.8 In analogia con la teoria presentata nel Paragrafo 3.5.1, si definisca il procedimento di linearizzazione di un sistema non lineare e variante nel tempo attorno a un suo generico movimento. Si rifletta sull'eventualità che la tecnica di verifica della stabilità del Paragrafo 3.5.2 possa essere estesa a tale situazione.

Esercizio 3.9 Si consideri il sistema

$$\dot{x}_1(t) = -2x_1(t) + 3x_2(t) + \alpha u(t)$$

$$\dot{x}_2(t) = -2x_1(t) + \beta u(t)$$

$$y(t) = \gamma x_1(t) + \delta x_2(t) + \varepsilon u(t)$$

Si individuino i valori dei parametri α , β , γ , δ e ε per i quali esso non è completamente raggiungibile e successivamente quelli per cui esso non è completamente osservabile.

Esercizio 3.10 Si consideri il sistema

$$\dot{x}_1(t) = -x_1(t) + x_2(t) + 2u(t)$$

$$\dot{x}_2(t) = -x_2(t)$$

$$\dot{x}_3(t) = -2x_3(t) + 4u(t)$$

$$y(t) = x_1(t) + 2x_2(t) + 3x_3(t)$$

che non è in forma minima. Si determini tale forma e si verifichi che la risposta all'impulso dell'uscita del sistema originario coincide con quella della sua forma minima.

Funzione di trasferimento

4.1 Introduzione

In questo capitolo sarà introdotta una nuova rappresentazione dei sistemi dinamici a tempo continuo lineari e stazionari. Essa è chiamata funzione di trasferimento e mette in relazione tra loro le trasformate di Laplace (Paragrafo B.3) delle variabili di ingresso e di uscita. Si usa dire che mediante questa funzione i sistemi dinamici sono descritti nel dominio della variabile complessa, anche se, per ragioni che saranno chiarite nel Capitolo 6, è più comune parlare di rappresentazione nel dominio della frequenza.

Nel capitolo saranno trattati i seguenti argomenti:

- la definizione della funzione di trasferimento e l'analisi delle sue principali proprietà;
- la rappresentazione e i parametri della funzione di trasferimento;
- le caratteristiche della risposta allo scalino di sistemi SISO del primo e del secondo ordine;
- i legami tra diverse rappresentazioni dei sistemi dinamici lineari e stazionari: le variabili di stato, la risposta all'impulso, la funzione di trasferimento.

4.2 Definizione della funzione di trasferimento

4.2.1 Definizione e interpretazioni

Si consideri il sistema con n variabili di stato, m variabili di ingresso e p variabili di uscita

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (4.1)$$

$$y(t) = Cx(t) + Du(t) \quad (4.2)$$

e si indichino con $U(s)$, $X(s)$ e $Y(s)$, funzioni della variabile complessa s , le trasformate di Laplace di $u(t)$, $x(t)$ e $y(t)$. Applicando la trasformazione di Laplace ad ambo i membri delle Equazioni (4.1), (4.2) e ricordando le sue proprietà, si ottiene

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

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

da cui risulta

$$X(s) = (sI - A)^{-1}BU(s) + (sI - A)^{-1}x(0) \quad (4.3)$$

$$Y(s) = (C(sI - A)^{-1}B + D)U(s) + C(sI - A)^{-1}x(0) \quad (4.4)$$

Le Equazioni (4.3), (4.4) forniscono le trasformate di Laplace dei movimenti dello stato e dell'uscita. Al loro interno è immediato individuare le componenti libera $(sI - A)^{-1}x(0)$ e forzata $(sI - A)^{-1}BU(s)$ del movimento dello stato e le corrispondenti componenti $C(sI - A)^{-1}x(0)$ e $(C(sI - A)^{-1}B + D)U(s)$ del movimento dell'uscita. Si ha così un'ulteriore conferma della validità del principio di sovrapposizione degli effetti presentato al Paragrafo 3.2.3.

La matrice $p \times m$

$$G(s) = C(sI - A)^{-1}B + D \quad (4.5)$$

che appare nella (4.4) viene detta *funzione di trasferimento* e, moltiplicata a destra per la trasformata di Laplace dell'ingresso u , fornisce la trasformata di Laplace dell'uscita y corrispondente a stato iniziale nullo, cioè dell'uscita forzata. Per condizioni iniziali nulle il sistema (4.1), (4.2) può quindi essere descritto con la *rappresentazione ingresso-uscita*

$$Y(s) = G(s)U(s) \quad (4.6)$$

Nota la funzione di trasferimento $G(s)$ di un sistema e nota la trasformata di Laplace $U(s)$ dell'ingresso, è possibile calcolare, mediante antitrasformazione della (4.6), il movimento forzato y_f dell'uscita che coincide con il movimento y nell'ipotesi di stato iniziale nullo.

La funzione di trasferimento è quindi una *rappresentazione esterna* del sistema, in contrasto con la rappresentazione interna espressa dalla forma in variabili di stato (4.1), (4.2), introdotta e analizzata nei Capitoli 2 e 3.

Nel caso particolare di sistemi SISO, dalla (4.6) segue che la funzione di trasferimento può essere vista come il rapporto tra le trasformate di Laplace dell'uscita forzata e dell'ingresso che l'ha prodotta.

Sempre nel caso di sistemi SISO di ordine n , si consideri il particolare ingresso $u(t) = \text{imp}(t)$. Allora, dato che $U(s) = 1$, dalla (4.6) risulta $Y(s) = G(s)$, cioè la funzione di trasferimento si può interpretare come trasformata di Laplace della risposta all'impulso g_y (Equazione (3.34)) del sistema (4.1), (4.2). Questa interpretazione della funzione di trasferimento è immediatamente generalizzabile al caso di sistemi MIMO.

4.2.2 Struttura della funzione di trasferimento

Se il sistema (4.1), (4.2) è SISO, $G(s) = C(sI - A)^{-1}B + D$ è una funzione razionale in s , data dal rapporto di due polinomi, per la quale valgono le osservazioni riportate nel seguente.

- Nel caso particolare di sistemi non dinamici descritti dalla relazione $y(t) = Du(t)$, $G(s)$ è indipendente da s e risulta pari a D .
- Per costruzione, la matrice $(sI - A)^{-1}$ ha $n \times n$ elementi, tutti funzioni razionali in s , con denominatore di grado n che coincide con il polinomio caratteristico di A , e con numeratore al più di grado $n - 1$. Poiché la moltiplicazione a sinistra di $(sI - A)^{-1}$ per la matrice C e la sua moltiplicazione a destra per la

matrice B corrispondono a combinare linearmente i suoi elementi, si conclude che $C(sI - A)^{-1}B$, che coincide con $G(s)$ se il sistema è strettamente proprio ($D = 0$), è una funzione razionale in s con polinomio a denominatore di grado n e polinomio a numeratore al più di grado $n - 1$.

- Se il sistema è proprio ($D \neq 0$), la somma di una funzione razionale $C(sI - A)^{-1}B$ con denominatore di grado n con una costante D produce una funzione razionale con polinomi a numeratore e denominatore ambedue di grado n .
- Nel calcolo di $G(s)$ può avvenire che i polinomi a numeratore e denominatore abbiano una o più radici in comune. In tal caso, dopo aver effettuato la cancellazione di questi fattori comuni, $G(s)$ è una funzione razionale con polinomio a denominatore di grado $v < n$ e polinomio a numeratore di grado al più $v - 1$ per $D = 0$, o v per $D \neq 0$. La funzione di trasferimento risultante è detta in *forma minima*.

Per quanto detto, in generale risulta

$$G(s) = \frac{N_G(s)}{D_G(s)} = \frac{\beta_v s^v + \beta_{v-1}s^{v-1} + \dots + \beta_1s + \beta_0}{\alpha_v s^v + \alpha_{v-1}s^{v-1} + \dots + \alpha_1s + \alpha_0} \quad (4.7)$$

dove $v \leq n$, $\beta_v = 0$ se il sistema è strettamente proprio ($D = 0$) e, senza perdita di generalità, si può assumere $\alpha_v = 1$. La differenza tra il grado del denominatore $D_G(s)$ e quello del numeratore $N_G(s)$ è chiamata *grado relativo* di $G(s)$.

In apparente contrasto con quanto ora esposto, talvolta può essere utile considerare funzioni di trasferimento con grado del numeratore strettamente maggiore di quello del denominatore. I sistemi dinamici, puramente fitizi, corrispondenti a tali funzioni di trasferimento sono chiamati *impropri*. Il loro impiego è comunque giustificato dal fatto che è sempre possibile determinare un sistema proprio, o strettamente proprio, che ne approssimi le caratteristiche di interesse nello specifico problema di volta in volta trattato.

Poli e zeri Con riferimento a sistemi SISO descritti dalla (4.7), si può utilizzare la terminologia introdotta nel Paragrafo B.3 e dire che \hat{s} è uno *zero* di $G(s)$ se ne annulla il numeratore, cioè se $N_G(\hat{s}) = 0$, mentre è un *polo* di $G(s)$ se ne annulla il denominatore, cioè se $D_G(\hat{s}) = 0$. Per quanto precedentemente detto, i poli sono anche radici dell'equazione $\det(sI - A) = 0$, cioè sono autovalori del sistema. Sia gli zeri sia i poli sono reali o complessi coniugati a coppie, in quanto sono reali i coefficienti α_i e β_i nella (4.7). Poli e zeri, nel loro insieme, sono detti *singularità*.

Nel caso più generale di sistemi MIMO, \hat{s} è un polo di $G(s)$ se annulla il denominatore di almeno una delle funzioni razionali che compongono la funzione di trasferimento; mentre la definizione di zero risulta più complessa e non viene qui fornita.

Esempio 4.1 (Seguito dell'Esempio 2.6) Le equazioni del motore elettrico (2.14)-(2.17) possono essere poste nella forma (4.1), (4.2) dove

$$A = \begin{bmatrix} -R/L & -k/L \\ k/J & -h/J \end{bmatrix}, \quad B = \begin{bmatrix} 1/L & 0 \\ 0 & -1/J \end{bmatrix}, \quad C = I, \quad D = 0$$

La funzione di trasferimento del sistema è

$$G(s) = \frac{1}{\varphi(s)} \begin{bmatrix} \frac{s + k/J}{L} & \frac{k}{LJ} \\ \frac{k}{LJ} & \frac{-s + R/L}{J} \end{bmatrix} \quad (4.8)$$

con

$$\varphi(s) = s^2 + \left(\frac{R}{L} + \frac{h}{J}\right)s + \frac{k^2 + Rh}{LJ}$$

Si noti che tutti gli elementi di $G(s)$ sono funzioni razionali con denominatore coincidente con il polinomio caratteristico di A .

Esempio 4.2 (Seguito dell'Esempio 3.5) La funzione di trasferimento del sistema (3.21)-(3.23) è

$$G(s) = [1 \ 0] \begin{bmatrix} s & -1 \\ k/M & s + h/M \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1/M \end{bmatrix} = \frac{1/M}{s^2 + sh/M + k/M} \quad (4.9)$$

Questa funzione non ha zeri, mentre ha due poli che coincidono con gli autovalori. Il grado relativo è due.

In casi semplici come quello in esame, posto $x(0) = 0$, la funzione di trasferimento si può anche ottenere trasformando con Laplace le equazioni del sistema e quindi eliminando progressivamente le trasformate delle variabili di stato nella determinazione del legame tra $U(s)$ e $Y(s)$. Nel caso specifico, dalle (3.21)-(3.23) risulta

$$\begin{aligned} sX_1(s) &= X_2(s) \\ sX_2(s) &= -\frac{k}{M}X_1(s) - \frac{h}{M}X_2(s) + \frac{1}{M}U(s) \\ Y(s) &= X_1(s) \end{aligned}$$

e, per sostituzione,

$$s^2Y(s) = -\frac{k}{M}Y(s) - \frac{h}{M}sY(s) + \frac{1}{M}U(s)$$

Da quest'ultima equazione si ricava direttamente la funzione di trasferimento (4.9). Una terza possibilità per la valutazione di $G(s)$ si ha ponendo $y(0) = \dot{y}(0) = 0$ e trasformando la (3.20); si ottiene così

$$Ms^2Y(s) = -kY(s) - hsY(s) + U(s)$$

da cui segue direttamente la (4.9).

4.2.3 Equazioni differenziali e funzione di trasferimento

L'Esempio 4.2 mostra come, limitandosi a considerare sistemi SISO descritti dall'equazione differenziale di ordine n

$$\frac{d^n y(t)}{dt^n} + \alpha_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + \alpha_1 \frac{dy(t)}{dt} + \alpha_0 y(t) = \beta_0 u(t) \quad (4.10)$$

per ricavare la funzione di trasferimento non sia necessario passare attraverso una rappresentazione in variabili di stato, ma si possa più semplicemente ipotizzare

$$y(0) = 0, \quad \frac{d^i y(0)}{dt^i} = 0, \quad i = 1, 2, \dots, n-1$$

e quindi applicare direttamente l'operatore di trasformazione di Laplace alla (4.10) per ottenere

$$s^n Y(s) + \alpha_{n-1} s^{n-1} Y(s) + \dots + \alpha_1 s Y(s) + \alpha_0 Y(s) = \beta_0 U(s)$$

da cui

$$\frac{Y(s)}{U(s)} = G(s) = \frac{\beta_0}{s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_1 s + \alpha_0}$$

Nel caso generale in cui il sistema sia rappresentato dall'equazione differenziale

$$\begin{aligned} \frac{d^n y(t)}{dt^n} + \alpha_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + \alpha_1 \frac{dy(t)}{dt} + \alpha_0 y(t) &= \\ \beta_n \frac{d^n u(t)}{dt^n} + \beta_{n-1} \frac{d^{n-1} u(t)}{dt^{n-1}} + \dots + \beta_1 \frac{du(t)}{dt} + \beta_0 u(t) & \quad (4.11) \end{aligned}$$

un procedimento analogo consiste nel porre

$$u(0) = 0, \quad \frac{d^i u(0)}{dt^i} = 0, \quad i = 1, 2, \dots, n-1 \quad (4.12)$$

$$y(0) = 0, \quad \frac{d^i y(0)}{dt^i} = 0, \quad i = 1, 2, \dots, n-1 \quad (4.13)$$

Il numero di vincoli imposti dalle (4.12), (4.13) è certamente sovrabbondante, tuttavia, applicando ancora la trasformazione di Laplace alla (4.11), per le (4.12), (4.13) si ottiene immediatamente

$$\begin{aligned} s^n Y(s) + \alpha_{n-1} s^{n-1} Y(s) + \dots + \alpha_1 s Y(s) + \alpha_0 Y(s) &= \\ \beta_n s^n U(s) + \beta_{n-1} s^{n-1} U(s) + \dots + \beta_1 s U(s) + \beta_0 U(s) & \end{aligned}$$

da cui

$$\frac{Y(s)}{U(s)} = G(s) = \frac{\beta_n s^n + \beta_{n-1} s^{n-1} + \dots + \beta_1 s + \beta_0}{s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_1 s + \alpha_0}$$

4.2.4 Cancellazioni e stabilità

Nel calcolo della funzione di trasferimento $G(s)$ data dalla (4.7), la cancellazione di radici in comune tra i polinomi a numeratore e denominatore fa sì che il numero ν dei poli possa essere inferiore a quello n degli autovalori. Poiché $G(s)$ è una rappresentazione esterna, descrive cioè il legame tra l'ingresso e l'uscita del sistema, si può intuitivamente ipotizzare che gli autovalori che non coincidono con i poli di $G(s)$ siano associati a parti "nascoste" del sistema che non influenzano tale legame, come mostrato nel seguito mediante semplici esempi.

Esempio 4.3 (Seguito dell'Esempio 3.20) Il circuito elettrico rappresentato nella Figura 3.7 è descritto dal sistema lineare e invariante (4.1), (4.2) in cui

$$A = \begin{bmatrix} -\frac{1}{RC} & -\frac{1}{RC} \\ \frac{1}{RC} & -\frac{1}{RC} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{RC} \\ \frac{1}{RC} \end{bmatrix}, \quad C = [1 \ 0], \quad D = 0$$

con autovalori $s = 0$ e $s = -\frac{2}{RC}$. Dalla definizione (4.5), la funzione di trasferimento di questo sistema è

$$G(s) = [1 \ 0] \begin{bmatrix} s + \frac{1}{RC} & \frac{1}{RC} \\ \frac{1}{RC} & s + \frac{1}{RC} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{RC} \\ \frac{1}{RC} \end{bmatrix} = \frac{1/RC}{s + 2/RC} \quad (4.14)$$

che mostra come l'autovalore in $s = 0$ non sia un polo del sistema, ma appartenga alla sua parte "nascosta". Si noti che dalla semplice analisi della funzione di trasferimento (4.14) si potrebbe erroneamente concludere che il sistema è del primo ordine ($\nu = 1$) e asintoticamente stabile, in quanto l'unico polo è reale negativo poiché i parametri R e C sono positivi.

Esempio 4.4 (Seguito dell'Esempio 3.21) Il sistema meccanico rappresentato nella Figura 3.9 è descritto dalle matrici

$$A = \begin{bmatrix} -\frac{h}{M} & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{M} \\ 0 \end{bmatrix}, \quad C = [1 \ 0], \quad D = 0$$

e dalla funzione di trasferimento

$$G(s) = [1 \ 0] \begin{bmatrix} s + \frac{h}{M} & 0 \\ -1 & s \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{M} \\ 0 \end{bmatrix} = \frac{1/M}{s + h/M} \quad (4.15)$$

Anche in questo caso, soltanto l'autovalore $s = -\frac{h}{M}$ è polo della funzione di trasferimento, mentre l'autovalore $s = 0$ è stato cancellato da un'analogia radice del polinomio a numeratore. Un'immediata giustificazione di questo risultato si può ricavare dall'esame della struttura delle matrici A , B , C . L'uscita del sistema, infatti, coincide con la prima variabile di stato x_1 , il cui movimento non dipende da x_2 . È chiaro quindi che il legame tra l'ingresso e l'uscita rappresentato da $G(s)$ non può descrivere la dinamica "nascosta" associata a x_2 .

Esempio 4.5 (Seguito dell'Esempio 3.22) Il sistema idraulico rappresentato nella Figura 3.12 è descritto dal modello (3.83)-(3.85) ha due autovalori coincidenti in $s = 0$ e funzione di trasferimento

$$G(s) = [S_1 \ S_2] \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2S_1} \\ \frac{1}{2S_2} \end{bmatrix} = \frac{1}{s} \quad (4.16)$$

Quindi uno dei due autovalori nulli non appare come polo della funzione di trasferimento.

Gli esempi precedenti mostrano che le eventuali parti nascoste, non rappresentate dalla funzione di trasferimento, possono essere associate anche ad autovalori con parte reale non negativa. Per valutare se un sistema è asintoticamente stabile a partire dalla conoscenza della sua funzione di trasferimento è quindi necessario riferirsi a uno dei seguenti casi:

- quando nel calcolo della funzione di trasferimento $G(s)$ non avvengono semplificazioni, il suo denominatore coincide con il polinomio caratteristico, l'insieme dei poli coincide con quello degli autovalori e la conoscenza dei poli è sufficiente per accettare la proprietà di stabilità del sistema;
- se vi sono cancellazioni, la semplice conoscenza dei poli di $G(s)$ non consente di trarre alcuna conclusione sulla stabilità del sistema; potrebbe infatti accadere che uno o più autovalori non appartenenti all'insieme dei poli non abbiano parte reale negativa (si vedano gli Esempi 4.3, 4.4, 4.5 in cui gli autovalori in questione sono nulli).

La presenza di cancellazioni, in genere, si può considerare dovuta a una non corretta modellizzazione del sistema, almeno per quanto riguarda la sua descrizione esterna. Pertanto, nel seguito si ipotizzerà sempre che, qualora un sistema sia descritto in termini di funzione di trasferimento, l'insieme dei suoi poli coincida con quello dei suoi autovalori, cioè che non siano avvenute cancellazioni. Per questo motivo si supporrà sempre che il polinomio $D_G(s)$ al denominatore di $G(s)$ coincida con il polinomio caratteristico e il suo grado sarà chiamato *ordine del sistema*.

4.2.5 Cancellazioni, raggiungibilità e osservabilità

I risultati appena visti, legati alla presenza di eventuali parti "nascoste" del sistema possono essere interpretati in base alle proprietà di raggiungibilità e osservabilità presentate nel Paragrafo 3.6. Senza ledere la generalità, si può infatti pensare che, eventualmente dopo un opportuno cambio di variabili, il sistema (4.1), (4.2) sia rappresentato nella sua scomposizione canonica specificata nel Teorema 3.15. Al Paragrafo 3.6.4 è stato dimostrato che la risposta all'impulso dell'uscita di un sistema dipende esclusivamente dalla sua parte raggiungibile e osservabile e quindi (Teorema 3.16) che anche il movimento forzato dipende solo da questa parte. Pertanto, visto che $G(s) = \mathcal{L}[g_y(t)]$, anche la funzione di trasferimento dipende dalla sola parte raggiungibile e osservabile. Ritornando a considerare il caso SISO, si può allora affermare che:

- i poli di $G(s)$ sono anche autovalori della parte raggiungibile e osservabile del sistema (4.1), (4.2);
- il grado del polinomio $D_G(s)$, cioè il numero dei poli, coincide con l'ordine del sottosistema costituente la parte raggiungibile e osservabile del sistema (4.1), (4.2).

In definitiva, è possibile enunciare il seguente risultato.

Teorema 4.1 I poli del sistema (4.1), (4.2), assunto SISO, coincidono con gli autovalori della parte raggiungibile e osservabile del sistema stesso, compresa la molteplicità. ■

In base al risultato precedente si può infine concludere che gli eventuali autovalori del sistema (4.1), (4.2) che non sono poli della funzione di trasferimento $G(s)$, sono necessariamente autovalori della parte non raggiungibile e/o non osservabile del sistema stesso. Essi non sono presenti tra i poli di $G(s)$ per effetto di cancellazioni.

Esempio 4.6 (Seguito degli Esempi 3.20, 4.3) Dall'esame della scomposizione in parti del sistema dopo un opportuno cambio di variabili, rappresentata nella Figura 3.8, si può osservare che l'autovalore $s = 0$, che non compare come polo della funzione di trasferimento $G(s)$ data dalla (4.14), è associato alla parte non raggiungibile con variabile di stato \hat{x}_2 .

Esempio 4.7 (Seguito degli Esempi 3.21, 4.4) Le relazioni di dipendenza tra le variabili rappresentate nella Figura 3.10 mostrano che il legame tra la variabile di ingresso e quella di uscita non dipende dalla parte non osservabile del sistema con variabile di stato x_2 e autovalore $s = 0$. La funzione di trasferimento (4.15) non ha quindi questo autovalore come suo polo.

Esempio 4.8 (Seguito degli Esempi 3.22, 4.5) La scomposizione illustrata nella Figura 3.13 mostra come vi sia una parte non raggiungibile e non osservabile associata alla variabile di stato \hat{x}_2 e con l'autovalore $s = 0$, che non è polo della funzione di trasferimento (4.16).

4.2.6 Ritardo di tempo

Il ritardo di tempo descritto dalla relazione ingresso-uscita

$$y(t) = u(t - \tau) \quad (4.17)$$

è un sistema lineare e stazionario. Trasformando con Laplace ambo i membri dell'Equazione (4.17) (Appendice B.3.2), si ha $Y(s) = e^{-\tau s} U(s)$, per cui la funzione di trasferimento è

$$G(s) = e^{-\tau s} \quad (4.18)$$

Si osservi che la (4.18) non è una funzione razionale, pertanto in questo caso non si usa parlare di poli o di zeri.

Per tener conto di molti casi pratici è anche utile considerare sistemi SISO lineari e stazionari in cui l'ingresso agisce sullo stato e sull'uscita con un ritardo τ e descritti da

$$v(t) = u(t - \tau) \quad (4.19)$$

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (4.20)$$

$$y(t) = Cx(t) + Du(t) \quad (4.21)$$

o in cui la misura dell'uscita è disponibile con un ritardo τ , cioè

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (4.22)$$

$$z(t) = Cx(t) + Du(t) \quad (4.23)$$

$$y(t + \tau) = z(t) \quad (4.24)$$

Per mezzo della trasformazione di Laplace, dalle (4.19)-(4.24) si determina in entrambi i casi la funzione di trasferimento

$$G(s) = e^{-\tau s} G'(s), \quad G'(s) = C(sI - A)^{-1} B + D$$

che è quindi data dal prodotto del termine $e^{-\tau s}$, che rappresenta il ritardo sull'ingresso o sull'uscita, con la funzione $G'(s)$, razionale in s , a cui si applicano tutte le considerazioni precedenti sui sistemi con funzione di trasferimento razionale.

4.3 Rappresentazioni e parametri della funzione di trasferimento

Molto spesso è conveniente rappresentare la funzione di trasferimento di un sistema SISO in una delle due seguenti forme fattorizzate

$$G(s) = \frac{\rho \prod_i (s + z_i) \prod_i (s^2 + 2\zeta_i \alpha_{ni} s + \alpha_{ni}^2)}{s^g \prod_i (s + p_i) \prod_i (s^2 + 2\xi_i \omega_{ni} s + \omega_{ni}^2)} \quad (4.25)$$

$$G(s) = \frac{\mu \prod_i (1 + \tau_i s) \prod_i (1 + 2\zeta_i s/\alpha_{ni} + s^2/\alpha_{ni}^2)}{s^g \prod_i (1 + T_i s) \prod_i (1 + 2\xi_i s/\omega_{ni} + s^2/\omega_{ni}^2)} \quad (4.26)$$

Malgrado le Equazioni (4.25), (4.26) comprendano il caso in cui non tutte le singolarità siano reali, in esse non appaiono coefficienti complessi. Infatti, ogni coppia di singolarità complesse coniugate dà luogo a un termine di secondo grado. I parametri che compaiono nelle Equazioni (4.25), (4.26) vengono così chiamati:

- lo scalare ρ è detto *costante di trasferimento*;
- l'intero g , che può assumere valori positivi e negativi o essere nullo, è detto *tipo*;
- gli scalari $z_i \neq 0$ e $p_i \neq 0$ sono gli *zeri* e i *poli* reali non nulli cambiati di segno;
- gli scalari $\alpha_{ni} > 0$ e $\omega_{ni} > 0$ sono le *pulsazioni naturali* delle coppie di zeri e poli complessi coniugati;
- gli scalari ζ_i e ξ_i , in modulo minori di 1, sono gli *smorzamenti* degli zeri e dei poli complessi coniugati;
- lo scalare μ è detto *guadagno*;
- gli scalari $\tau_i \neq 0$ e $T_i \neq 0$ sono le *costanti di tempo*.

Con semplici calcoli si può facilmente determinare il legame tra i parametri delle diverse rappresentazioni (4.25), (4.26). In particolare risulta

$$\mu = \frac{\rho \prod_i z_i \prod_i \alpha_{ni}^2}{\prod_i p_i \prod_i \omega_{ni}^2}, \quad \rho = \frac{\mu \prod_i \tau_i \prod_i \omega_{ni}^2}{\prod_i T_i \prod_i \alpha_{ni}^2}, \quad \tau_i = \frac{1}{z_i}, \quad T_i = \frac{1}{p_i} \quad (4.27)$$

4.3.1 Guadagno

Sia dato un sistema descritto dalla funzione di trasferimento $G(s)$, posta nella forma (4.26), asintoticamente stabile ($g \leq 0$, $T_i > 0$ e $\xi_i > 0$). Si consideri inoltre il caso particolare $g = 0$. Supponendo che il sistema sia sottoposto a un ingresso costante \bar{u} , modellizzabile come un segnale a scalino con trasformata di Laplace $U(s) = \bar{u}/s$, l'uscita tende a un valore di regime \bar{y} determinabile mediante il teorema del valore finale (Paragrafo B.3.2) come

$$\begin{aligned} \bar{y} &= \lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} s G(s) \frac{\bar{u}}{s} = \lim_{s \rightarrow 0} s (C(sI - A)^{-1} B + D) \frac{\bar{u}}{s} = \\ &= G(0) \bar{u} = (-CA^{-1}B + D) \bar{u} \end{aligned} \quad (4.28)$$

e, per la (4.26), come

$$\bar{y} = \lim_{s \rightarrow 0} s \frac{\mu \prod_i (1 + \tau_i s) \prod_i (1 + 2\zeta_i s/\alpha_{ni} + s^2/\alpha_{ni}^2) \bar{u}}{\prod_i (1 + T_i s) \prod_i (1 + 2\xi_i s/\omega_{ni} + s^2/\omega_{ni}^2) s} = \mu \bar{u} \quad (4.29)$$

Confrontando le (4.28), (4.29), risulta

$$\mu = G(0) = -CA^{-1}B + D \quad (4.30)$$

La costante $\mu = \bar{y}/\bar{u}$ rappresenta il rapporto tra il valore di regime dell'uscita del sistema e il valore dell'ingresso costante che l'ha prodotta. Questa osservazione motiva la denominazione di *guadagno* attribuita a μ , che coincide perciò, in questo caso, con il *guadagno statico* introdotto al Paragrafo 3.3 relativamente all'analisi dell'equilibrio.

Al guadagno si può attribuire il significato ora detto di rapporto tra uscita e ingresso costanti anche nel caso di sistemi non asintoticamente stabili, purché di tipo $g = 0$, in quanto, dato un equilibrio caratterizzato dai valori costanti \bar{u} e \bar{y} di ingresso e uscita, come mostrato nel Paragrafo 3.3 risulta

$$\bar{y} = \mu \bar{u}$$

dove ancora μ è definito dalla (4.30).

Infine, nel caso in cui $g \neq 0$, il coefficiente μ viene ancora chiamato *guadagno*, o *guadagno generalizzato*, e il suo valore è dato da

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

In tal caso però μ non rappresenta più il guadagno statico del sistema.

4.3.2 Derivatore ideale

Si consideri un sistema descritto da una funzione di trasferimento di tipo $g < 0$ (il sistema, cioè, possiede uno o più zeri nell'origine). Nell'ipotesi di asintotica stabilità, si può verificare mediante il teorema del valore finale che il valore di regime dell'uscita a fronte di uno scalino d'ingresso di ampiezza arbitraria è nullo, cioè $\bar{y} = 0$. Il sistema esercita dunque sull'ingresso un'azione derivativa semplice se $g = -1$, o multipla se $g < -1$. Per questo motivo il sistema improprio con funzione di trasferimento $G(s) = s$ è comunemente chiamato *derivatore ideale*, definizione di facile interpretazione ricordando che l'operatore s è associato, in termini di trasformate di Laplace, all'operazione di derivazione nel tempo.

4.3.3 Integratore

Per valutare il comportamento di un sistema con poli nell'origine, cioè con $g > 0$, si consideri il caso semplice in cui $G(s) = 1/s$. Risulta allora $D = 0$, $C(sI - A)^{-1}B = 1/s$, cioè $A = 0$, e $CB = 1$, quindi per esempio $B = 1$, $C = 1$. Il sistema in variabili di stato può dunque essere rappresentato nella forma

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

L'uscita è quindi l'integrale dell'ingresso, e per questo motivo un sistema con $G(s) = 1/s$ viene comunemente chiamato *integratore*. Più in generale, un sistema con funzione di trasferimento $G(s) = 1/s^g$ esercita sull'ingresso del sistema un'azione integrale semplice se $g = 1$, o multipla se $g > 1$. Anche questo risultato può essere facilmente interpretato in termini di proprietà della trasformata di Laplace.

4.3.4 Costanti di tempo

Per la loro definizione (4.27), le costanti di tempo coincidono con il reciproco cambiato di segno dei poli e degli zeri non nulli. L'effetto delle costanti di tempo al denominatore sulla risposta forzata del sistema, e in particolare sulla risposta all'impulso, può quindi essere immediatamente dedotto dall'analisi del Paragrafo 3.2.5. Per esempio, nel caso di una funzione di trasferimento con soltanto poli reali $-p_i = -1/T_i$ di molteplicità uno, si può verificare mediante la procedura di anttrasformazione di Heaviside che la risposta all'impulso è una combinazione lineare dei modi e^{-t/T_i} . Se il sistema è asintoticamente stabile ($p_i > 0$), i modi si esauriscono tanto più velocemente quanto maggiore è il valore di p_i , cioè quanto più i poli sono lontani dall'origine del piano complesso. Se il polo $-p_i$ ha molteplicità $v_i > 1$, l'uscita è costituita da una combinazione lineare di termini dati da e^{-t/T_i} moltiplicati per t^j , $j = 0, 1, \dots, v_i - 1$. È possibile allora concludere che le costanti di tempo al denominatore sono legate alla velocità con cui si esauriscono i transitori del sistema: maggiore è il valore di una costante di tempo, più lentamente si esaurirà il contributo al movimento dell'uscita dato dal polo corrispondente.

4.3.5 Pulsazione naturale e smorzamento

Per comprendere il significato degli scalari α_{ni} , ω_{ni} , ξ_i e ξ_i che compaiono nelle (4.25), (4.26) si consideri una coppia di poli complessi coniugati (considera-

zioni analoghe possono essere fatte per gli zeri) $a \pm jb$, definiti come le radici dell'equazione $s^2 + 2\xi\omega_n s + \omega_n^2 = 0$. Risulta

$$a = -\xi\omega_n \quad b = \omega_n\sqrt{1-\xi^2}$$

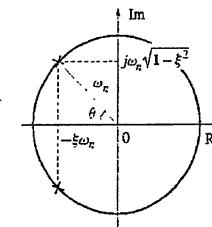


Figura 4.1
Parametri caratteristici dei poli complessi coniugati.

Il modulo dei poli coincide quindi con la pulsazione naturale ω_n , mentre lo smorzamento ξ è il coseno dell'angolo θ compreso tra la coniungente i poli con l'origine e il semiasse reale negativo (si veda la Figura 4.1). Pertanto, per un valore fissato di ω_n , al variare di ξ da -1 a $+1$ i poli si spostano su una circonferenza di raggio ω_n centrata nell'origine. In particolare, si possono distinguere i casi riportati in Tabella 4.1

$\xi = 0$	poli immaginari puri in $s = \pm j\omega_n$
$\xi > 0$	poli a parte reale negativa
$\xi < 0$	poli a parte reale positiva
$\xi = 1$	poli reali coincidenti in $s = -\omega_n$
$\xi = -1$	poli reali coincidenti in $s = \omega_n$

Tabella 4.1
Posizione dei poli in funzione dello smorzamento.

Se il sistema ha soltanto una coppia di poli complessi coniugati

$$-\xi\omega_n \pm j\omega_n\sqrt{1-\xi^2}$$

e guadagno unitario, è possibile verificare mediante la procedura di anttrasformazione di Heaviside che la risposta all'impulso è

$$y(t) = \frac{\omega_n}{\sqrt{1-\xi^2}} e^{-\xi\omega_n t} \sin(\omega_n t \sqrt{1-\xi^2})$$

Essa è perciò costituita da un termine sinusoidale, con pulsazione data dalla parte immaginaria dei poli, moltiplicato per un'esponenziale che, se il sistema è asintoticamente stabile, decresce tanto più velocemente quanto più è elevato il valore di $\xi\omega_n > 0$, cioè del modulo della parte reale dei poli.

Se la coppia di poli ha molteplicità $v_i > 1$, l'uscita è costituita da una combinazione lineare di termini analoghi al precedente moltiplicati per t^j , $j = 0, 1, \dots, v_i - 1$.

4.4 Risposta allo scalino

A partire dalla conoscenza della funzione di trasferimento, saranno ora esaminate in dettaglio le caratteristiche della risposta allo scalino di sistemi SISO asintoticamente stabili. Studiare il movimento dell'uscita in risposta a uno scalino è importante, perché permette di capire come si comporta il sistema nel passaggio da una condizione di equilibrio a un'altra a seguito di un'improvvisa commutazione del valore dell'ingresso.

Ci si concentrerà in particolare sui sistemi del primo e secondo ordine, con l'obiettivo di stabilire una relazione diretta tra i parametri della funzione di trasferimento e le principali caratteristiche della risposta allo scalino. Come si vedrà, tale analisi risulta valida, in certa misura, anche nello studio di sistemi di ordine superiore al secondo. Inoltre, i risultati di questo paragrafo possono essere utilmente impiegati per risolvere un problema di tipo diverso, cioè quello di dedurre una funzione di trasferimento approssimante del primo o del secondo ordine a partire dalla risposta sperimentale di un sistema soggetto a un ingresso a scalino.

Lo scalino d'ingresso sarà assunto di ampiezza unitaria, in quanto, per linearità, la risposta a uno scalino di ampiezza \bar{u} si ricava da quella allo scalino unitario moltiplicandola per \bar{u} .

Prima di intraprendere lo studio delle risposte è opportuno fare alcune considerazioni di validità generale.

4.4.1 Valore iniziale e finale

Nel caso di sistemi asintoticamente stabili, il calcolo del valore di regime y_∞ (o \bar{y}) della variabile di uscita è già stato effettuato nei Paragrafi 4.3.1 e 4.3.2 per mezzo del teorema del valore finale, giungendo alla conclusione che y_∞ è nullo in presenza di eventuali azioni derivate ($g < 0$), altrimenti è pari al guadagno μ .

Analogamente, per un sistema descritto dalla funzione di trasferimento razionale

$$G(s) = \frac{\beta_m s^m + \beta_{m-1} s^{m-1} + \dots + \beta_0}{\alpha_n s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_0}$$

dove $m \leq n$, il valore iniziale della risposta allo scalino può essere determinato mediante il teorema del valore iniziale, che fornisce

$$y(0) = \lim_{s \rightarrow \infty} s \frac{\beta_m s^m + \beta_{m-1} s^{m-1} + \dots + \beta_0}{\alpha_n s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_0} \frac{1}{s} = \begin{cases} 0, & m < n \\ \frac{\beta_m}{\alpha_n}, & m = n \end{cases}$$

Ricordando le regole di trasformazione della derivata di una funzione, il teorema può anche essere applicato per dedurre il valore iniziale delle derivate successive dell'uscita nell'istante iniziale. Per esempio, se $m < n$, $y(0) = 0$ e risulta

$$\begin{aligned} \dot{y}(0) &= \lim_{s \rightarrow \infty} s(sY(s) - y(0)) = \\ &= \lim_{s \rightarrow \infty} s^2 \frac{\beta_m s^m + \beta_{m-1} s^{m-1} + \dots + \beta_0}{\alpha_n s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_0} \frac{1}{s} = \begin{cases} 0, & m < n-1 \\ \frac{\beta_m}{\alpha_n}, & m = n-1 \end{cases} \end{aligned}$$

Iterando il procedimento è possibile concludere che per $m < n$ sono nulle le prime $n-m-1$ derivate di y in $t = 0$.

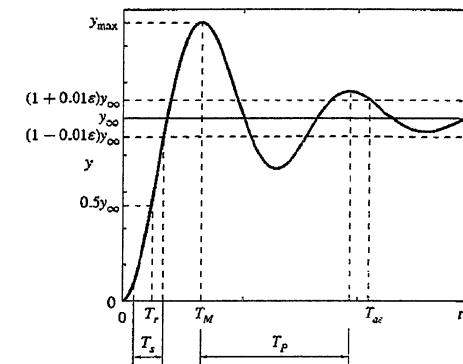


Figura 4.2
Parametri caratteristici della risposta allo scalino (con $\varepsilon = 10$).

4.4.2 Caratteristiche della risposta allo scalino

Per qualificare la risposta dei sistemi asintoticamente stabili, assumendo senza perdita di generalità che il guadagno μ sia positivo, nel seguito si farà riferimento ai parametri sottoelencati, il cui significato è illustrato nella Figura 4.2:

- *valore di regime y_∞* : valore dell'uscita a transitorio esaurito; per quanto detto è pari a μ ($g = 0$) o a zero ($g < 0$);
- *valore massimo y_{\max}* : massimo valore assunto dall'uscita;
- *sovraelongazione massima percentuale S%*: ampiezza, in percentuale, della sovraelongazione massima rispetto al valore di regime, cioè

$$S\% = 100 \cdot \frac{y_{\max} - y_\infty}{y_\infty}$$

- *tempo di massima sovraelongazione T_M* : primo istante in cui $y = y_{\max}$;
- *tempo di salita T_s* : tempo richiesto perché l'uscita passi per la prima volta dal 10% al 90% del suo valore di regime;
- *tempo di ritardo T_r* : tempo necessario perché l'uscita raggiunga la prima volta il valore $0.5y_\infty$;
- *tempo di assottigliamento T_{ae}* : tempo necessario perché il modulo della differenza tra l'uscita e il valore di regime y_∞ rimanga definitivamente al di sotto di $\varepsilon\%$, cioè l'uscita sia nell'intervallo $(1 - 0.01\varepsilon)y_\infty, (1 + 0.01\varepsilon)y_\infty$; per esempio con T_{ae} si indicherà il tempo necessario perché l'uscita entri definitivamente nella fascia di ampiezza $\pm 0.01y_\infty$ attorno al valore di regime y_∞ . Si dirà inoltre che T_{ae} rappresenta il tempo di assottigliamento al $(100 - \varepsilon)\%$;
- *periodo dell'oscillazione T_P* : distanza temporale tra i primi due massimi dell'uscita.

4.4.3 Sistemi del primo ordine

Si consideri un sistema del primo ordine con funzione di trasferimento

$$G(s) = \frac{\mu}{1 + Ts} \quad (4.31)$$

L'andamento dell'uscita nel tempo, determinato antitrasformando la funzione $Y(s) = G(s)/s$, è

$$y(t) = \mu (1 - e^{-t/T}) \quad , \quad t \geq 0 \quad (4.32)$$

Dalla (4.32) segue che $y(0) = 0$, $dy(0)/dt = \mu/T$ e, purché il sistema sia asintoticamente stabile ($T > 0$), $y_\infty = \mu$. È facile verificare inoltre che la tangente a y in $t = 0$ raggiunge il valore asintotico y_∞ dopo un tempo pari a T .

La (4.32) mostra anche che la risposta è di tipo esponenziale; per $T > 0$ risulta pertanto $y_{\max} = y_\infty$ e $S\% = 0$. Sempre dalla (4.32) è possibile calcolare esplicitamente il tempo di assestamento

$$T_{ae} = T \ln \frac{1}{0.01\epsilon} = -T \ln 0.01\epsilon \quad (4.33)$$

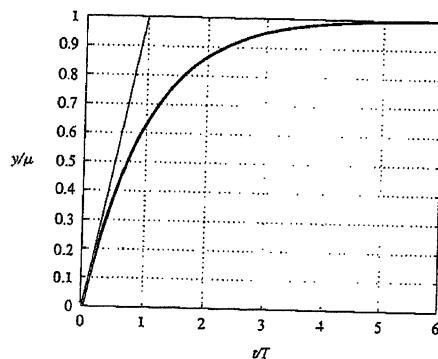
e derivare gli altri parametri caratteristici della risposta allo scalino, riportati nella Tabella 4.2.

Tabella 4.2
Parametri caratteristici
della risposta allo scalino
del sistema (4.31).

y_∞	T_s	T_r	T_{a5}	T_{a1}
μ	$\simeq 2.2T$	$\simeq 0.7T$	$\simeq 3T$	$\simeq 4.6T$

Come atteso, la velocità di risposta del sistema, sempre per $T > 0$, dipende dalla costante di tempo T ; a una diminuzione di T , cioè a uno spostamento verso sinistra del polo, corrisponde una diminuzione del tempo di salita, del tempo di ritardo e del tempo di assestamento. In particolare, il transitorio si può ritenere in pratica esaurito per $t \simeq 4 \div 5T$. L'andamento dell'uscita normalizzata y/μ è riportato nella Figura 4.3 in funzione del tempo normalizzato t/T . È infine possibile verificare che risulta $y(T) \simeq 0.63y_\infty$.

Figura 4.3
Risposta allo scalino del
sistema (4.31) e tangente
nell'origine.



4.4.4 Sistemi del secondo ordine

Sistemi con solo poli reali

I casi: poli distinti. La risposta allo scalino di un sistema del secondo ordine con funzione di trasferimento

$$G(s) = \frac{\mu}{(1+T_1s)(1+T_2s)} \quad , \quad T_1 > T_2 \quad (4.34)$$

è data da

$$y(t) = \mu \left(1 - \frac{T_1}{T_1 - T_2} e^{-t/T_1} + \frac{T_2}{T_1 - T_2} e^{-t/T_2} \right) \quad , \quad t \geq 0 \quad (4.35)$$

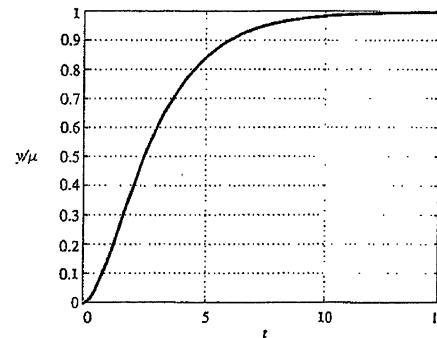


Figura 4.4
Risposta allo scalino del
sistema (4.34) per $T_1 = 2$,
 $T_2 = 1$.

La (4.35) mostra che, in condizioni di stabilità asintotica ($T_1 > T_2 > 0$), il sistema è tanto più veloce quanto più le costanti di tempo T_1 e T_2 sono piccole, però T_{ae} , T_s e T_r sono funzioni non semplici di T_1 e T_2 . Si può inoltre verificare che la risposta non presenta alcuna sovraelongazione. A titolo di esempio, la risposta del sistema in un caso particolare è riportata nella Figura 4.4.

Dall'espressione (4.35) è anche possibile verificare che se $T_1 \gg T_2$, l'esponenziale più "lenta", con costante di tempo T_1 , domina la forma della risposta e, per t non troppo piccolo ($t \simeq (4 \div 5)T_2$), risulta

$$y(t) \simeq \mu \left(1 - e^{-t/T_1} \right) \quad , \quad t \geq 0 \quad (4.36)$$

cioè il sistema diventa "di fatto" del primo ordine.

Il caso: poli coincidenti. Nel caso in cui la funzione di trasferimento sia

$$G(s) = \frac{\mu}{(1+Ts)^2} \quad (4.37)$$

la risposta allo scalino ha la seguente espressione analitica

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

L'andamento nel tempo di $y(t)$ è qualitativamente non diverso da quello relativo alla (4.35) se $T \simeq T_1 \simeq T_2$. Tuttavia in questo caso per $T > 0$ è possibile valutare esplicitamente i valori caratteristici riportati nella Tabella 4.3.

y_∞	T_s	T_r	T_{a5}	T_{a1}
μ	$\simeq 3.36T$	$\simeq 1.68T$	$\simeq 4.74T$	$\simeq 6.64T$

Tabella 4.3
Parametri caratteristici
della risposta allo scalino
del sistema (4.37).

Sistemi con poli reali e uno zero Si considerino sistemi con funzione di trasferimento

$$G(s) = \frac{\mu(1+\tau s)}{(1+T_1s)(1+T_2s)} \quad , \quad T_1 \neq \tau \quad , \quad T_2 \neq \tau \quad (4.38)$$

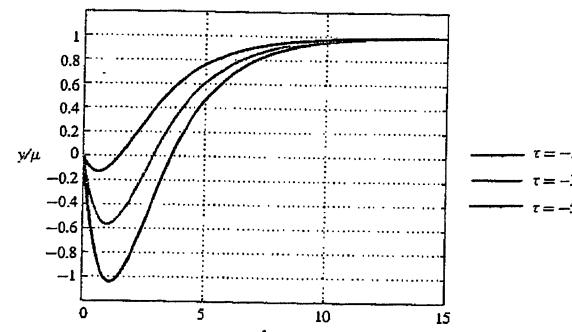
a cui corrisponde la risposta allo scalino

$$y(t) = \mu \left(1 - \frac{T_1 - \tau}{T_1 - T_2} e^{-t/T_1} + \frac{T_2 - \tau}{T_1 - T_2} e^{-t/T_2} \right) \quad , \quad t \geq 0 \quad (4.39)$$

I parametri caratteristici della (4.39) sono funzioni non semplici di τ , T_1 e T_2 . Supponendo $T_1 > T_2 > 0$, per valutare convenientemente l'andamento nel tempo dell'uscita in funzione della posizione relativa dello zero rispetto ai poli è opportuno distinguere i casi seguenti.

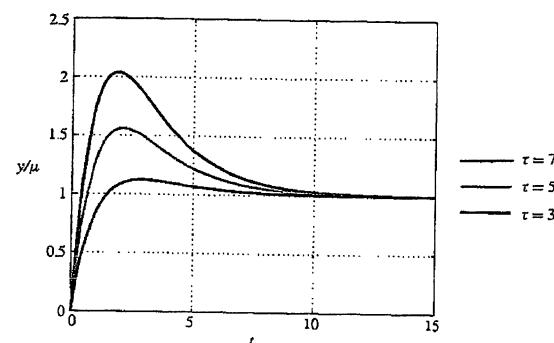
I caso: $\tau < 0$. Come mostrato nella Figura 4.5, dove sono riportati gli andamenti di y per T_1 e T_2 fissate e per diversi valori di τ , la risposta presenta una *sottelongazione iniziale* o, come spesso si dice, una *risposta inversa*, che è tanto più pronunciata quanto più lo zero $-1/\tau$ si avvicina all'origine del piano complesso.

Figura 4.5
Risposta allo scalino del sistema (4.38), con $T_1 = 2$, $T_2 = 1$, per diversi valori di $\tau < 0$.



II caso: $\tau > T_1 > T_2$. La risposta presenta una *sovraelongazione* tanto più marcata quanto più lo zero negativo è vicino all'origine del piano complesso rispetto alla posizione dei poli. Nella Figura 4.6 sono riportati gli andamenti di y per valori fissati di T_1 e T_2 e al variare di τ . Anche nel caso di sistemi di ordine più elevato la presenza di una sovraelongazione (ma non di un andamento oscillante!) nella risposta allo scalino è indice della presenza di uno zero reale negativo di modulo minore di quello dei poli.

Figura 4.6
Risposta allo scalino del sistema (4.38), con $T_1 = 2$, $T_2 = 1$, per diversi valori di $\tau > 0$.

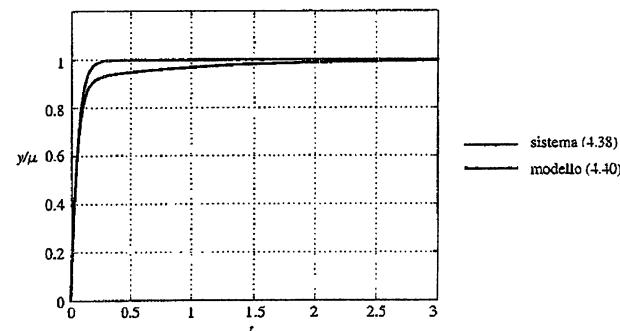


III caso: $\tau \simeq T_1 \gg T_2$. Dalla (4.39) con semplici calcoli si può verificare che l'andamento dell'uscita può essere così approssimato

$$y(t) \simeq \mu (1 - e^{-t/T_2}) \quad , \quad t \geq 0 \quad (4.40)$$

e il sistema può essere considerato "di fatto" del primo ordine. Tuttavia, come mostrato nella Figura 4.7, la risposta è lievemente diversa da quella data dalla

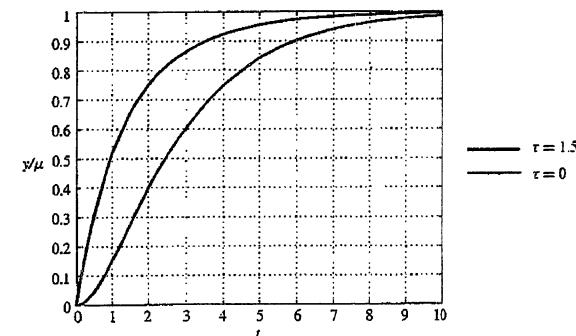
Figura 4.7
Risposta allo scalino del sistema (4.38) per $T_1 = 1$, $T_2 = 0.05$ e $\tau = 0.92$ e andamento della (4.40) con $T_2 = 0.05$.



(4.40) a causa della coppia polo-zero trascurata nell'approssimazione. Tali singolarità generano comunque un transitorio, di piccola entità e ignorato nella (4.40), che produce la lenta deriva di y verso y_∞ .

IV caso: $T_1 > \tau > T_2$. La presenza dello zero tende a velocizzare la risposta rispetto al caso $\tau = 0$. Un esempio di ciò è riportato nella Figura 4.8 dove, assumendo $T_1 = 2$, $T_2 = 1$, sono poste a confronto le risposte per $\tau = 0$ e $\tau = 1.5$.

Figura 4.8
Risposta allo scalino del sistema (4.38), con $T_1 = 2$, $T_2 = 1$, e due diversi valori di τ .



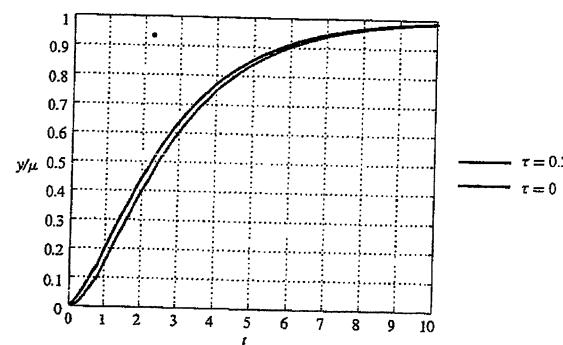
Si noti anche che, se $\tau \simeq T_2$, con ragionamenti analoghi a quelli del caso precedente si può ipotizzare che la risposta sia approssimabile con quella di un sistema del primo ordine con costante di tempo T_1

$$y(t) \simeq \mu (1 - e^{-t/T_1}) \quad , \quad t \geq 0$$

L'effetto di lenta deriva dell'uscita verso il valore di regime in questo caso non si verifica, perché il polo trascurato in $-1/T_2$ produce un transitorio che si annulla più rapidamente di quello del polo in $-1/T_1$.

V caso: $T_1 > T_2 > \tau > 0$. Al diminuire del valore di τ , cioè per uno zero sempre più lontano dall'origine del piano complesso, la risposta (4.39) tende a quella di un sistema con gli stessi poli, ma senza lo zero, data dalla (4.35). Per esempio nella Figura 4.9 sono riportati gli andamenti di y per valori di T_1 e T_2 fissati, con $\tau = 0$ e $\tau = 0.2$.

Figura 4.9
Risposta allo scalino del sistema (4.38) con $T_1 = 2$, $T_2 = 1$, e due diversi valori di τ .



Sistemi con poli complessi coniugati Si consideri un sistema con funzione di trasferimento

$$G(s) = \frac{\mu \omega_n^2}{s^2 + 2\xi \omega_n s + \omega_n^2} \quad (4.41)$$

dove $\omega_n > 0$ e ξ , $|\xi| < 1$, rappresentano la pulsazione naturale e lo smorzamento dei poli. Dai risultati del Paragrafo B.3.3, si deduce che l'espressione analitica della risposta allo scalino è

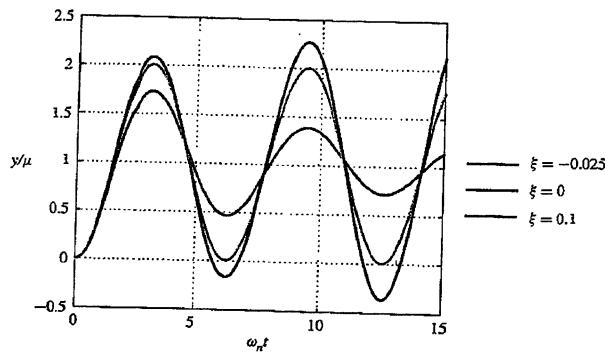
$$y(t) = \mu \left(1 - \frac{1}{\sqrt{1-\xi^2}} e^{-\xi \omega_n t} \sin \left(\omega_n t \sqrt{1-\xi^2} + \arccos(\xi) \right) \right), \quad t \geq 0 \quad (4.42)$$

ed è quindi costituita da un termine sinusoidale moltiplicato per un'esponenziale. In particolare, per $\xi > 0$ l'esponenziale è decrescente e la risposta tende asintoticamente al valore di regime μ , per $\xi < 0$ il sistema è instabile e l'uscita diverge, mentre per $\xi = 0$ il sistema è stabile, non asintoticamente, e la (4.42) assume la forma

$$y(t) = \mu (1 - \cos(\omega_n t)), \quad t \geq 0$$

Nella Figura 4.10 è riportato l'andamento dell'uscita normalizzata y/μ in funzione del tempo normalizzato $\omega_n t$ nei casi di asintotica stabilità ($\xi = 0.1$), stabilità ($\xi = 0$) e instabilità ($\xi = -0.025$).

Figura 4.10
Risposta allo scalino del sistema (4.41) per diversi valori di ξ .



Particolare interesse merita, come di consueto, lo studio della (4.42) nel caso di stabilità asintotica ($0 < \xi < 1$). Uguagliando a zero la derivata della (4.42) rispetto al tempo, si possono determinare facilmente gli istanti di stazionarietà della funzione, che risultano

$$\bar{t}_k = \frac{k\pi}{\omega_n \sqrt{1-\xi^2}}, \quad k \text{ intero positivo.} \quad (4.43)$$

Essi sono istanti di massimo per k dispari e di minimo per k pari. I corrispondenti valori dell'uscita sono

$$y(\bar{t}_k) = \mu \left(1 - (-1)^k e^{-\xi k \pi / \sqrt{1-\xi^2}} \right) \quad (4.44)$$

Le (4.43) e (4.44) portano immediatamente a concludere che risulta

$$y_{\max} = \mu \left(1 + e^{-\xi \pi / \sqrt{1-\xi^2}} \right) \quad (4.45)$$

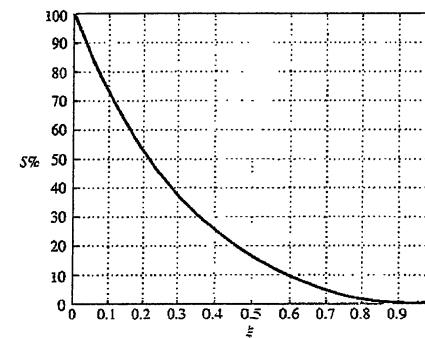
$$T_M = \frac{\pi}{\omega_n \sqrt{1-\xi^2}} \quad (4.46)$$

$$T_P = \frac{2\pi}{\omega_n \sqrt{1-\xi^2}} \quad (4.47)$$

$$S\% = 100 e^{-\xi \pi / \sqrt{1-\xi^2}} \quad (4.47)$$

La sovraelongazione massima percentuale dipende quindi esclusivamente dallo smorzamento ξ , come mostra la Figura 4.11, è una funzione monotona decrescente di ξ .

Figura 4.11
Andamento di $S\%$ al variare di ξ nella risposta allo scalino del sistema (4.41).



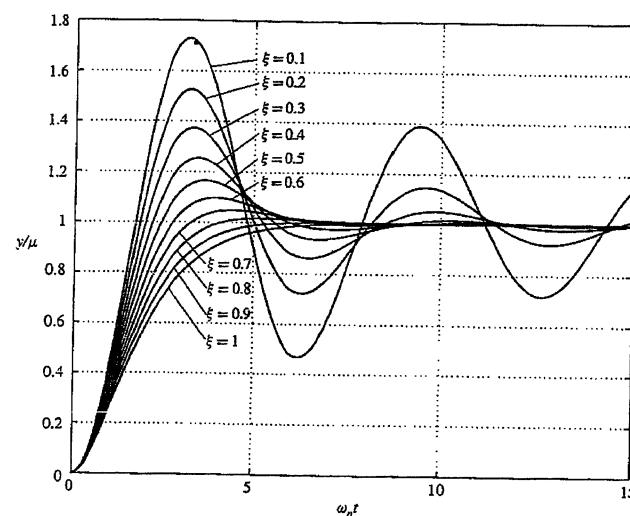
I transitori di y/μ in funzione di $\omega_n t$ relativi a vari valori dello smorzamento sono riportati nella Figura 4.12; dalla figura è anche evidente che il tempo di ritardo T_r e il tempo di salita T_s dipendono anch'essi da ξ . Per quanto riguarda il tempo di assestamento T_{ae} , se non è agevole ricavare la relazione esatta che lo lega a ω_n e ξ , è però molto facile fornire un'approssimazione. Si osservi infatti che i massimi e i minimi di y/μ dati dalla (4.44) giacciono rispettivamente sulle funzioni

$$y_M(t) = 1 + e^{-\xi \omega_n t}$$

e

$$y_m(t) = 1 - e^{-\xi \omega_n t}$$

Figura 4.12
Risposta allo scalino del sistema (4.41) per diversi valori di $\xi > 0$.



Si consideri allora l'istante \bar{t}_e in corrispondenza del quale le funzioni $y_M(t)$ e $y_m(t)$ entrano nella fascia $[(1 - 0.01\epsilon), (1 + 0.01\epsilon)]$. Per determinare il valore di \bar{t}_e è quindi sufficiente imporre

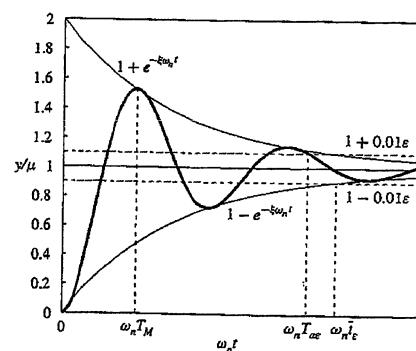
$$e^{-\xi\omega_n\bar{t}_e} = 0.01\epsilon$$

da cui

$$\bar{t}_e = -\frac{1}{\xi\omega_n} \ln 0.01\epsilon$$

Dalla Figura 4.13 si deduce che \bar{t}_e rappresenta una buona approssimazione di T_{ae} . Si noti che la stima così ottenuta dipende da $\xi\omega_n$, cioè solo dal modulo della parte reale degli autovalori. In particolare risulta $T_{a1} \approx 4.6/\xi\omega_n$, identico a quello di un sistema del primo ordine con costante di tempo $T = 1/\xi\omega_n$.

Figura 4.13
Risposta allo scalino del sistema (4.41) e parametri caratteristici.



y_∞	$S\%$	T_M	T_P	stima di T_{ae}
μ	$100e^{-\xi\pi/\sqrt{1-\xi^2}}$	$\frac{\pi}{\omega_n\sqrt{1-\xi^2}}$	$\frac{2\pi}{\omega_n\sqrt{1-\xi^2}}$	$-\frac{1}{\xi\omega_n} \ln 0.01\epsilon$

Le caratteristiche principali della risposta allo scalino sono riassunte nella Tabella 4.4.

4.4.5 Sistemi con ritardo di tempo

Nel Paragrafo 4.2.6 si è mostrato che un sistema lineare e stazionario in cui sia presente un ritardo τ nell'applicazione dell'ingresso o nella rilevazione dell'uscita è caratterizzato dalla funzione di trasferimento

$$G(s) = e^{-\tau s} G'(s)$$

dove $G'(s)$ è razionale in s . La risposta allo scalino di questo sistema è nulla per $t < \tau$ e il teorema del valore iniziale si può ancora impiegare per determinare l'uscita e le sue derivate all'istante $t = \tau$. Se $G'(s)$ è del primo o del secondo ordine, le caratteristiche della risposta per $t > \tau$ sono immediatamente deducibili dall'analisi condotta in precedenza e applicata a $G'(s)$. Infatti, la risposta è quella di $G'(s)$ traslata di τ ; per esempio, i valori dei parametri caratteristici T_M , T_P , T_{ae} precedentemente ricavati nei vari casi risultano tutti aumentati di τ .

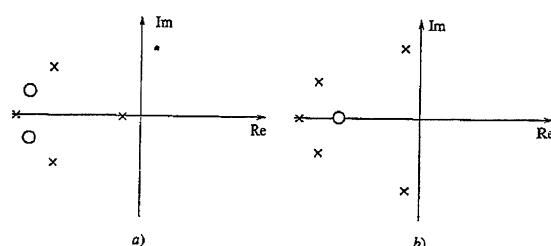
4.4.6 Sistemi di ordine superiore al secondo

Nel Paragrafo 4.4.4 si è analizzata la risposta allo scalino di un sistema asintoticamente stabile con due poli reali e uno zero e si è mostrato che se la costante di tempo dello zero è simile a quella di uno dei due poli, la risposta è assimilabile a quella di un sistema del primo ordine con uguale guadagno e con il polo rimanente dopo l'eliminazione delle singolarità prossime tra loro. In altri termini, benché la cancellazione della coppia polo/zero non sia in senso stretto lecita, l'aver eliminato queste singolarità produce effetti trascurabili sulla risposta allo scalino. Queste considerazioni possono essere estese anche al caso di sistemi di ordine superiore. Dato un sistema rappresentato dalla funzione di trasferimento $G(s)$, qualora vi siano coppie costituite da un polo e uno zero di $G(s)$ vicini tra loro nel piano complesso e con parte reale negativa, è possibile forzare la cancellazione delle stesse mantenendo invariati tutti gli altri parametri, compreso il guadagno, per ottenere un modello approssimato di ordine ridotto, ma con caratteristiche simili a quelle del sistema di partenza, per esempio per quanto riguarda l'andamento della risposta allo scalino.

Poli dominanti Data una funzione di trasferimento $G(s)$ di un sistema asintoticamente stabile, una volta cancellate eventuali coppie di poli e zeri prossimi tra loro nel piano complesso, vengono chiamati *poli dominanti* i poli, reali o complessi, nettamente più vicini all'asse immaginario rispetto agli altri poli, come esemplificato nella Figura 4.14. La risposta allo scalino di un sistema con poli dominanti può essere approssimata con quella di un sistema la cui funzione di trasferimento $G_a(s)$ possiede soltanto questi poli e un guadagno pari a quello del sistema di partenza. Pertanto, se un sistema ha solo uno o due poli dominanti, l'andamento qualitativo della sua risposta allo scalino può essere determinato a partire

Tabella 4.4
Parametri caratteristici della risposta allo scalino del sistema (4.41).

Figura 4.14
a) polo dominante reale;
b) poli dominanti complessi coniugati.



dai risultati precedentemente illustrati. Naturalmente nel fare l'approssimazione è opportuno tenere conto anche di eventuali zeri che hanno distanza dall'asse immaginario comparabile, o addirittura inferiore, a quella dei poli dominanti.

Esempio 4.9 (Seguito degli Esempi 2.6, 4.1) Si considerino, nelle rispettive unità di misura, i seguenti valori dei parametri caratteristici del motore a corrente continua: $R = 0.46$, $L = 1$, $h = 0.0008$, $k = 0.25$, $J = 0.012$. Dalla (4.8), la funzione di trasferimento tra la tensione nel circuito d'armatura v e la velocità di rotazione w è

$$G(s) = \frac{3.98}{(1 + 0.0856s)(1 + 0.0022s^2)} \quad (4.48)$$

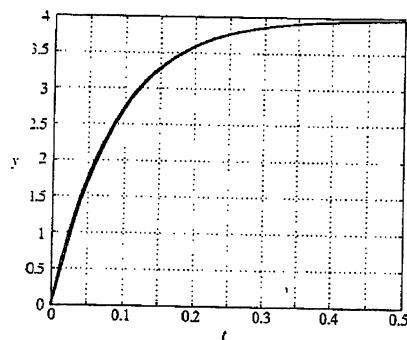
in cui una costante di tempo è nettamente maggiore dell'altra. Il polo dominante è quindi in $s = -1/0.0856 = -11.68$ e l'approssimazione a polo dominante è

$$G_a(s) = \frac{3.98}{1 + 0.0856s} \quad (4.49)$$

Le risposte allo scalino associate ai sistemi (4.48) e (4.49) sono riportate nella Figura 4.15, che mostra come in fase di modellizzazione sia possibile trascurare la costante di tempo elettrica $L/R = 0.0022$ ottenendo comunque un modello accurato.

Questo esempio mostra come spesso la presenza di uno o più poli dominanti sia dovuta al fatto che nel sistema in esame vi è un fenomeno che influenza maggiormente rispetto ad altri nella caratterizzazione del comportamento dinamico del sistema stesso. Nel caso specifico, la dinamica relativa alla rotazione dell'albero motore ha certamente un rilievo maggiore rispetto alla dinamica elettrica, che si esaurisce molto più velocemente della prima.

Figura 4.15
Risposta allo scalino del sistema (4.48) (linea nera) e del modello (4.49) (linea in colore).



Esempio 4.10 Si consideri il sistema

$$G(s) = \frac{I + \tau s}{(1 + 0.1s)(1 + 0.02s + 0.002s^2)(1 + 0.1s + s^2)} \quad (4.50)$$

con uno zero in $s = -1/\tau$ e con poli in $s = -10$, $s = -5 \pm j21.8$ e $s = -0.05 \pm j0.998$. Questi ultimi sono i poli dominanti del sistema e, quando $\tau = 0.5$, un'approssimante della (4.50) ricavata secondo le regole precedentemente esposte è

$$G_a(s) = \frac{1}{1 + 0.1s + s^2} \quad (4.51)$$

Le risposte allo scalino relative alle (4.50) e (4.51), riportate nella Figura 4.16, mostrano come le considerazioni fatte nell'esempio precedente siano valide anche nel caso in esame.

Se invece nella (4.50) si pone $\tau = 20$, il corrispondente zero in $s = -0.05$ non può essere trascurato nella determinazione della funzione di trasferimento approssimante, che risulta

$$G_a(s) = \frac{I + 20s}{1 + 0.1s + s^2} \quad (4.52)$$

Nella Figura 4.17 sono riportate le risposte allo scalino delle funzioni di trasferimento (4.50), (4.51) e (4.52), dalla figura è evidente l'opportunità di includere lo zero del sistema nella funzione di trasferimento approssimante.

Figura 4.16
Risposta allo scalino del sistema (4.50) (linea in nero) e del modello (4.51) (linea in colore).

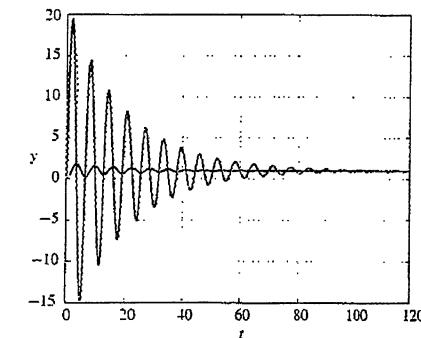
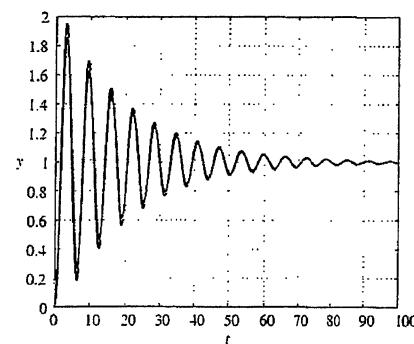


Figura 4.17
Risposta allo scalino del sistema (4.50) (linea nera), del modello (4.51) (linea in colore continua) e del modello (4.52) (linea in colore tratteggiata).

4.5 Realizzazione

Il problema della *realizzazione* consiste nel ricavare un modello in variabili di stato equivalente al modello assegnato tramite la funzione di trasferimento $G(s)$. Questo problema è decisamente molto complesso nel caso dei sistemi MIMO; ci si limiterà quindi a considerare sistemi SISO descritti da

$$G(s) = \frac{N_G(s)}{D_G(s)} = \frac{\beta_n s^n + \beta_{n-1} s^{n-1} + \dots + \beta_1 s + \beta_0}{s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_1 s + \alpha_0} \quad (4.53)$$

Poiché la rappresentazione di un sistema in termini di variabili di stato non è univoca, si veda il Paragrafo 3.2.4, il problema della determinazione di una realizzazione di ordine n ammette infinite soluzioni, caratterizzate da diverse quadruple di matrici A , B , C e D della rappresentazione di stato, legate tra loro dalle trasformazioni (3.15). Tra le soluzioni possibili, assumono grande interesse quelle in cui le matrici A , B , C e D hanno una struttura particolare. Conviene cioè utilizzare speciali *forme canoniche* per le matrici in gioco, caratterizzate dall'avere tutti gli elementi indipendenti dai valori α_i e β_i della (4.53), salvo $2n+1$ termini. Nel seguito vengono presentate due forme canoniche di largo impiego.

4.5.1 Forma canonica di raggiungibilità

Si osservi innanzitutto che la (4.53) può sempre essere scritta come

$$G(s) = \frac{\hat{\beta}_{n-1} s^{n-1} + \hat{\beta}_{n-2} s^{n-2} + \dots + \hat{\beta}_1 s + \hat{\beta}_0 + \hat{\beta}_n}{s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_1 s + \alpha_0} \quad (4.54)$$

dove

$$\hat{\beta}_n = \beta_n \quad \hat{\beta}_i = \beta_i - \alpha_i \hat{\beta}_n, \quad i = 0, \dots, n-1 \quad (4.55)$$

La *forma canonica di raggiungibilità*, o *di controllo*, è data dalle matrici

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-1} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (4.56)$$

$$C = [\hat{\beta}_0 \ \hat{\beta}_1 \ \hat{\beta}_2 \ \dots \ \hat{\beta}_{n-1}], \quad D = \hat{\beta}_n \quad (4.57)$$

Si noti che le (4.56), (4.57) hanno $2n+1$ parametri liberi, mentre gli altri elementi sono strutturalmente 0 o 1: a questo riguardo si può affermare che $2n+1$ rappresenta, ovviamente, il minimo numero di parametri necessario per descrivere la (4.53) in termini di variabili di stato.

Il nome di questa forma canonica è dovuto al fatto che, indipendentemente dal valore dei parametri α_i , il sistema che essa definisce è sempre completamente raggiungibile, come facilmente verificabile per mezzo del Teorema 3.11. La proprietà di osservabilità scaturisce invece dall'ipotesi che i polinomi $N_G(s)$ e $D_G(s)$ nella (4.53) siano primi tra loro.

4.5.2 Forma canonica di osservabilità

Un modo alternativo di rappresentare in termini di variabili di stato la (4.54) è costituito dalla *forma canonica di osservabilità*, o *di ricostruzione*, definita da

$$A = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \dots & 0 & -\alpha_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\alpha_{n-1} \end{bmatrix}, \quad B = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_{n-1} \end{bmatrix} \quad (4.58)$$

$$C = [0 \ 0 \ 0 \ \dots \ 0 \ 1], \quad D = \hat{\beta}_n \quad (4.59)$$

Anche le (4.58), (4.59) sono caratterizzate da $2n+1$ parametri liberi e dai rimanenti elementi fissati strutturalmente a 0 o 1.

Il sistema (4.58), (4.59) è sempre completamente osservabile, come facilmente verificabile applicando il Teorema 3.13, mentre la proprietà di raggiungibilità è dovuta all'ipotesi che i polinomi $N_G(s)$ e $D_G(s)$ nella (4.53) siano primi tra loro.

Esempio 4.11 (Seguito dell'Esempio 4.2) La funzione di trasferimento del carrello è

$$G(s) = \frac{1/M}{s^2 + sh/M + k/M}$$

la cui realizzazione in forma canonica di raggiungibilità è descritta da

$$A = \begin{bmatrix} 0 & 1 \\ -k/M & -h/M \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1/M \ 0], \quad D = 0 \quad (4.60)$$

mentre quella in forma canonica di osservabilità è data da

$$A = \begin{bmatrix} 0 & -k/M \\ 1 & -h/M \end{bmatrix}, \quad B = \begin{bmatrix} 1/M \\ 0 \end{bmatrix}, \quad C = [0 \ 1], \quad D = 0 \quad (4.61)$$

Nessuna di queste due forme coincide con la rappresentazione in variabili di stato precedentemente definita tramite le (3.21)-(3.23). La matrice di trasformazione dalle (4.60) alle (3.24) è

$$T = \begin{bmatrix} 1/M & 0 \\ 0 & 1/M \end{bmatrix}$$

mentre il passaggio dalle (4.61) alle (3.24) è garantito dalla trasformazione

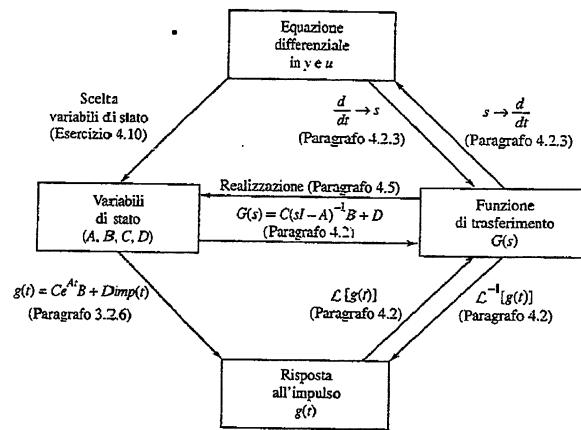
$$T = \begin{bmatrix} 0 & 1 \\ 1 & -h/M \end{bmatrix}$$

4.5.3 Relazioni tra diverse rappresentazioni dei sistemi lineari SISO

A conclusione di questo capitolo è opportuno ricordare che le rappresentazioni sin qui introdotte per i sistemi dinamici lineari, stazionari, SISO e con funzione di trasferimento razionale sono:

- la rappresentazione mediante variabili di stato;
- la risposta all'impulso;
- la funzione di trasferimento;
- la descrizione mediante un'unica equazione differenziale.

Figura 4.18
Rappresentazioni dei sistemi dinamici e relazioni corrispondenti.



Le principali relazioni che intercorrono tra queste rappresentazioni sono riassunte nella Figura 4.18, che evidenzia ancora come sia possibile trasformare il sistema da una forma in un'altra qualunque, anche se non sempre per via diretta. Si osservi a questo riguardo che sarebbe comunque possibile completare la Figura 4.18 con i percorsi diretti mancanti mediante un'analisi più dettagliata di quella che si è ritenuto opportuno presentare in questo testo.

4.6 Conclusioni

In questo capitolo è stata introdotta la funzione di trasferimento, che costituisce la rappresentazione nel dominio della variabile complessa dei sistemi lineari e stazionari. Essa esprime il legame tra le trasformate di Laplace dell'ingresso e dell'uscita e, per sistemi SISO senza ritardi, consiste in una funzione razionale. Sono stati discussi i legami tra la funzione di trasferimento e i modelli descritti nel dominio del tempo e, per sistemi del primo e del secondo ordine, sono state analizzate le caratteristiche dell'uscita a fronte di ingressi a scalino. In particolare, si è mostrato come il valore di alcuni parametri caratteristici della risposta allo scalino, quali il tempo di salita, il tempo di assestamento e la sovraetensione massima percentuale, in molti casi possa essere determinato direttamente in base alla posizione dei poli e degli zeri del sistema. Sono state infine ricordate le relazioni che intercorrono tra le diverse rappresentazioni dei sistemi lineari e stazionari sin qui introdotte. Avere presenti tali relazioni risulta di grande utilità pratica, poiché consente di fare ricorso alla rappresentazione che meglio si adatta alla soluzione dei problemi specifici di volta in volta considerati.

Esercizi

Esercizio 4.1 Si mostri che la rappresentazione di un sistema in termini di funzione di trasferimento è unica, a differenza di quanto visto nel Paragrafo 3.2.4 relativamente alla rappresentazione di stato.

Esercizio 4.2 Si verifichi, ricordando la trasformata di Laplace del prodotto di convoluzione, che $G(s)$ può essere interpretata come la trasformata della risposta all'impulso del sistema (4.1), (4.2).

Esercizio 4.3 Si determini la funzione di trasferimento del sistema descritto da

$$\begin{aligned}\dot{x}(t) &= x(t - \tau) + u(t), \quad \tau > 0 \\ y(t) &= x(t)\end{aligned}$$

Esercizio 4.4 Si rappresenti la funzione di trasferimento

$$G(s) = \frac{s + 10}{s^2 + 2s + 4}$$

nella forma fattorizzata (4.26) e la funzione di trasferimento

$$G(s) = \frac{1 - 0.1s}{1 + 7s + 10s^2}$$

nella forma fattorizzata (4.25).

Esercizio 4.5 Sfruttando le proprietà della trasformazione di Laplace, si determini la risposta all'impulso, alla rampa e alla parabola a partire dalla risposta allo scalino di un sistema descritto dalla funzione di trasferimento $G(s)$.

Esercizio 4.6 Si determini la funzione di trasferimento di un sistema la cui risposta all'impulso è data dal grafico riportato nella Figura 4.3. Si valuti qualitativamente la risposta allo scalino di questo sistema.

Esercizio 4.7 Si determini l'andamento della risposta allo scalino del sistema con funzione di trasferimento

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

supposto asintoticamente stabile. Si valuti quindi il corrispondente tempo di assestamento.

Esercizio 4.8 Si determini l'andamento nel tempo della risposta allo scalino del sistema con funzione di trasferimento

$$G(s) = \frac{\mu\omega_n^2(1 + \tau s)}{s^2 + 2\xi\omega_n s + \omega_n^2}$$

Esercizio 4.9 Si verifichi che le (4.55), (4.56), (4.57) rappresentano una realizzazione della (4.53).

Esercizio 4.10 Si verifichi che il sistema descritto dall'equazione differenziale

$$\frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = \beta_0 u(t)$$

può essere descritto in variabili di stato in forma canonica di raggiungibilità se si pone

$$x_i(t) = \frac{1}{\beta_0} \frac{d^{i-1} y(t)}{dt^{i-1}}, \quad i = 1, \dots, n$$

5.1 Introduzione

Nello studio dei sistemi dinamici lineari e stazionari costituiti da più sottosistemi variamente collegati tra loro risulta spesso conveniente l'uso di una rappresentazione grafica basata sui cosiddetti "schemi a blocchi". Questo tipo di descrizione, oltre a mettere in luce con chiarezza le interazioni tra i diversi sottosistemi, rende agevole il calcolo della funzione di trasferimento tra una data variabile di ingresso e una data variabile di uscita del sistema complessivo.

In questo capitolo si presenteranno:

- i componenti base di uno schema a blocchi;
- le regole per elaborare uno schema a blocchi senza modificare le relazioni ingresso-uscita del sistema;
- lo studio di tre tipiche configurazioni derivanti dalla connessione in serie, in parallelo e in retroazione di due sottosistemi;
- il legame tra le proprietà di stabilità dei singoli componenti di uno schema a blocchi e la stabilità dell'intero sistema;
- le conseguenze di eventuali cancellazioni di poli e zeri causate dalla connessione tra due o più sottosistemi.

5.2 Componenti di uno schema a blocchi

Nella simbologia degli schemi a blocchi una variabile è rappresentata da una *frecce* etichettata con il nome della variabile stessa e un sistema è rappresentato da un *blocco*, cioè da un rettangolo al cui interno è indicata la funzione di trasferimento, dotato di una freccia entrante e una uscente che corrispondono rispettivamente all'ingresso e all'uscita. Pertanto un sistema con ingresso u , uscita y e funzione di trasferimento $G(s)$ viene rappresentato come mostrato nella Figura 5.1. Negli schemi a blocchi compaiono anche due altri elementi: il *nodo sommatore* e il *punto di diramazione*. Il nodo sommatore è costituito da un cerchio con una freccia uscente e alcune frecce entrate, ognuna caratterizzata da un segno (+ oppure -): esso indica che la variabile associata alla freccia uscente è la somma algebrica,

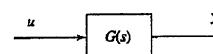
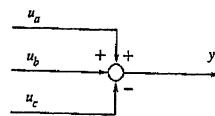


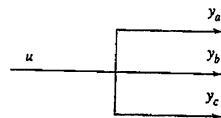
Figura 5.1
Rappresentazione di un sistema dinamico.

Figura 5.2
Nodo sommatore.



secondo quanto specificato dai segni, delle variabili associate alle frecce entranti. Per esempio la Figura 5.2 rappresenta la relazione $y(t) = u_a(t) + u_b(t) - u_c(t)$. Invece, il punto di diramazione si adopera quando si vuole descrivere il fatto che due o più variabili, rappresentate da frecce uscenti, costituiscono la replica dell'unica variabile a monte del punto. Così, il fatto che le variabili y_a , y_b e y_c costituiscano una replica della variabile u , cioè $y_a(t) = y_b(t) = y_c(t) = u(t)$, si rappresenta come nella Figura 5.3.

Figura 5.3
Punto di diramazione.



D'ora in poi si assumerà che tutte le variabili che appaiono in uno schema a blocchi siano scalari e, coerentemente, che tutti i sottosistemi rappresentati siano SISO. Ciò non esclude che sia possibile descrivere in questo modo anche sistemi MIMO. Per esempio si consideri il sistema con due ingressi e due uscite, cioè con

$$u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}, \quad y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

definito dalla relazione $Y(s) = G(s)U(s)$ con

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

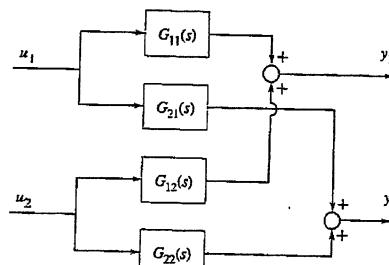
Poiché il legame tra le variabili è descritto in modo equivalente dalle equazioni

$$Y_1(s) = G_{11}(s)U_1(s) + G_{12}(s)U_2(s)$$

$$Y_2(s) = G_{21}(s)U_1(s) + G_{22}(s)U_2(s)$$

il sistema si può rappresentare mediante lo schema a blocchi di Figura 5.4, in cui tutte le variabili che compaiono sono scalari.

Figura 5.4
Schema a blocchi di un sistema MIMO.



Esempio 5.1 Si consideri il sistema idraulico descritto nella Figura 5.5. Il primo serbatoio è alimentato attraverso l'elettrovalvola E_1 , comandata dal controllore C in anello chiuso in modo da mantenere il livello h_2 nel secondo serbatoio prossimo al valore desiderato h_2^0 .

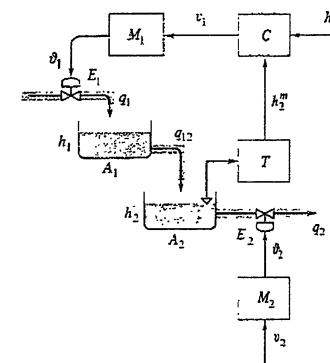


Figura 5.5
Sistema idraulico dell'Esempio 5.1.

A tale scopo il controllore utilizza la misura h_2^m fornita dal trasduttore T e opera in modo puramente statico in base all'errore rilevato $e^m = h_2^0 - h_2^m$. Si suppone inoltre che la portata q_{12} tra il primo e il secondo serbatoio sia proporzionale al livello h_1 , cioè $q_{12} = kh_1$, e che la portata q_2 in uscita dal secondo serbatoio dipenda solo dall'apertura ϑ_2 della valvola di scarico E_2 . Le due valvole sono azionate da motori elettrici, denotati con M_1 e M_2 e comandati dalle tensioni v_1 e v_2 , rispettivamente. Indicando con A_1 e A_2 le sezioni dei due serbatoi, e supponendo che tutti i componenti siano descrivibili mediante modelli lineari e stazionari, è facile verificare che il sistema di Figura 5.5 può essere rappresentato dallo schema a blocchi di Figura 5.6, dove

$$G_1(s) = \frac{1}{1 + sA_1/k} \quad , \quad G_2(s) = \frac{1}{A_2 s}$$

mentre $M_1(s)$, $M_2(s)$ e $T(s)$ sono le funzioni di trasferimento dei due motori e del trasduttore, la costante β è il guadagno del controllore, e le costanti a_i , $i = 1, 2$, descrivono il legame, supposto lineare e istantaneo, tra l'apertura delle valvole ϑ_i e le portate erogate q_i .

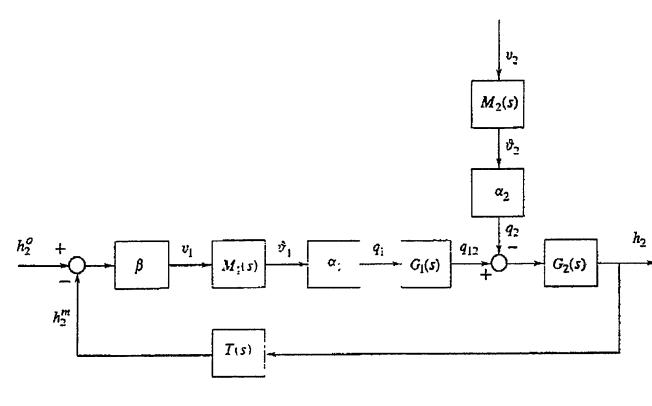


Figura 5.6
Schema a blocchi del sistema idraulico dell'Esempio 5.1.

5.3 Regole di elaborazione

È spesso utile saper calcolare, a partire da uno schema a blocchi, la funzione di trasferimento tra una variabile di ingresso e una determinata uscita. A tale riguardo è conveniente cominciare a trattare i casi più semplici di due soli blocchi collegati in serie, in parallelo e in retroazione.

5.3.1 Sistemi in serie

Due sistemi descritti dalle equazioni

$$Y_a(s) = G_a(s)U_a(s) \quad (5.1)$$

$$Y_b(s) = G_b(s)U_b(s) \quad (5.2)$$

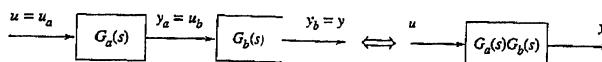
si dicono interconnessi *in serie* (o *in cascata*) quando l'uscita y_a del primo coincide con l'ingresso u_b del secondo (Figura 5.7). Considerando $u(t) = u_a(t)$ e $y(t) = y_b(t)$ come l'ingresso e l'uscita del sistema complessivo, è immediato ricavare che

$$Y(s) = G_b(s)Y_a(s) = G_b(s)G_a(s)U(s)$$

Dunque la funzione di trasferimento complessiva risulta

$$G(s) = \frac{Y(s)}{U(s)} = G_a(s)G_b(s) \quad (5.3)$$

ovvero è pari al prodotto delle funzioni di trasferimento dei singoli sottosistemi. La regola espressa dall'Equazione (5.3) è ovviamente generalizzabile al caso in cui i sistemi connessi in serie siano in numero maggiore di due.



5.3.2 Sistemi in parallelo

I sistemi (5.1), (5.2) si dicono connessi *in parallelo* (Figura 5.8) se hanno lo stesso ingresso, che si assume anche come ingresso del sistema complessivo, mentre le loro uscite si sommano per generare l'uscita del sistema complessivo. Poiché è immediato verificare che

$$Y(s) = Y_a(s) + Y_b(s) = G_a(s)U_a(s) + G_b(s)U_b(s) = (G_a(s) + G_b(s))U(s)$$

la funzione di trasferimento complessiva risulta

$$G(s) = \frac{Y(s)}{U(s)} = G_a(s) + G_b(s) \quad (5.4)$$

ovvero è pari alla somma delle funzioni di trasferimento dei singoli sottosistemi. È ovvio che, nel caso in cui i segni associati alle variabili y_a e y_b entranti nel

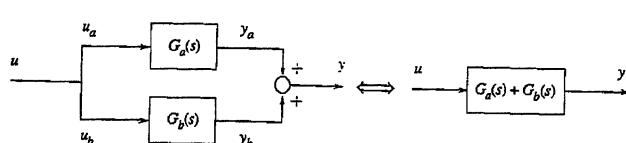


Figura 5.7
Sistemi in serie.

Figura 5.8
Sistemi in parallelo.

nodo sommatore di Figura 5.8 non siano entrambi positivi, la formula (5.4) va opportunamente modificata considerando la somma algebrica delle funzioni di trasferimento invece della semplice somma aritmetica. Altrettanto banale è la generalizzazione del risultato trovato al caso in cui i sistemi connessi in parallelo siano più di due.

5.3.3 Sistemi in retroazione

Quando i sistemi (5.1), (5.2) sono collegati come nella Figura 5.9 essi si dicono connessi *in retroazione* e costituiscono un sistema *in anello chiuso o retroazionato*. Si ricava

$$Y(s) = G_a(s)(U(s) - Y_b(s)) = G_a(s)(U(s) - G_b(s)Y(s))$$

La funzione di trasferimento complessiva pertanto risulta

$$G(s) = \frac{Y(s)}{U(s)} = \frac{G_a(s)}{1 + G_a(s)G_b(s)} \quad (5.5)$$

ovvero è pari al rapporto tra la funzione di trasferimento del sottosistema che appare lungo la linea di andata tra u e y e la somma tra 1 e la cosiddetta *funzione di trasferimento d'anello*

$$L(s) = G_a(s)G_b(s)$$

del sistema, definita come prodotto delle funzioni di trasferimento dei due sottosistemi presenti lungo l'anello di Figura 5.9.

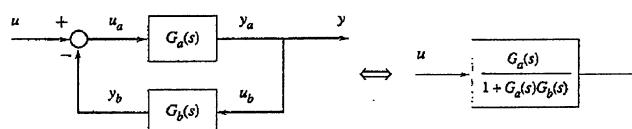


Figura 5.9
Sistemi in retroazione negativa.

La configurazione descritta nella Figura 5.9, di grande importanza nelle applicazioni, viene indicata con il termine di *retroazione negativa*, e quest'ultimo aggettivo fa riferimento al fatto che la variabile y_b entra nel nodo sommatore di Figura 5.9 con il segno meno. Nel caso della connessione in *retroazione positiva*, rappresentata nella Figura 5.10, è facile verificare che la funzione di trasferimento del sistema è

$$G(s) = \frac{Y(s)}{U(s)} = \frac{G_a(s)}{1 - G_a(s)G_b(s)}$$

pari al rapporto tra la funzione di trasferimento del sottosistema in andata tra u e y e la differenza tra 1 e la funzione di trasferimento d'anello.

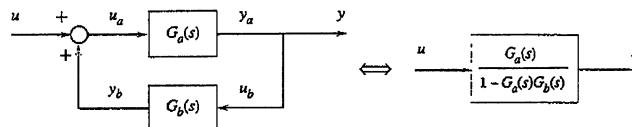


Figura 5.10
Sistemi in retroazione positiva.

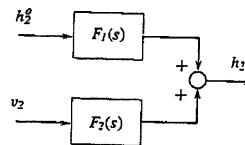
5.3.4 Riduzione di schemi a blocchi

Le regole di calcolo delle funzioni di trasferimento di sistemi costituiti da sottosistemi connessi in serie, parallelo e retroazione possono essere utilizzate più volte in sequenza per elaborare schemi a blocchi complessi, che presentano anche più d'una variabile di ingresso. Si ricordi che, grazie alla linearità, l'effetto di ogni ingresso su un'uscita è indipendente dalla presenza degli altri eventuali ingressi, e può essere perciò valutato ponendo questi ultimi tutti a zero. Un'illustrazione di questo procedimento è fornita nell'esempio seguente.

Esempio 5.2 (Seguito dell'Esempio 5.1) Assumendo come uscita il livello h_2 , lo schema di Figura 5.6 si deve poter ricondurre a quello di Figura 5.11, visto che gli unici ingressi indipendenti sono il segnale di riferimento h_2^0 e la tensione v_2 applicata al motore M_2 . Si ha cioè, con ovvio significato dei simboli,

$$H_2(s) = F_1(s)H_2^0(s) + F_2(s)V_2(s)$$

Figura 5.11
Schema equivalente ridotto del sistema dell'Esempio 5.2.



Grazie al principio di sovrapposizione degli effetti, per calcolare $F_1(s)$ si può supporre nullo l'ingresso v_2 e ignorare pertanto nello schema a blocchi di Figura 5.6 la presenza dei blocchi $M_2(s)$ e α_2 . Osservando a questo punto che sulla linea di andata tra h_2^0 e h_2 c'è una sequenza di cinque blocchi in serie, lo schema può essere rielaborato come nella Figura 5.12, dove

$$G_3(s) = \beta\alpha_1 M_1(s) G_1(s) G_2(s)$$

Ci si è così ricondotti a un semplice schema a retroazione negativa, dal quale è immediato ricavare che

$$F_1(s) = \frac{G_3(s)}{1 + G_3(s)T(s)}$$

Per calcolare invece $F_2(s)$ si suppone nullo l'ingresso h_2^0 e, con operazioni del tutto analoghe a quelle effettuate per il calcolo di $F_1(s)$, si ottiene

$$F_2(s) = \frac{-\alpha_2 M_2(s) G_2(s)}{1 + G_3(s)T(s)}$$

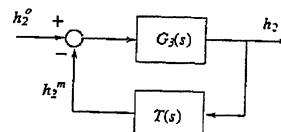


Figura 5.12
Elaborazione dello schema a blocchi dell'Esempio 5.2 per il calcolo di $F_1(s)$.

Vale la pena di sottolineare che l'elaborazione dello schema a blocchi di un sistema corrisponde semplicemente a una manipolazione delle equazioni algebriche che descrivono le relazioni tra le trasformate di Laplace delle variabili in gioco.

Oltre alle elaborazioni finora presentate, costituite dalla sostituzione di sottosistemi in serie, parallelo e retroazione con i loro equivalenti, esistono altre operazioni sugli schemi a blocchi che conservano inalterate le relazioni ingresso-uscita. Tali operazioni comprendono:

- lo spostamento di una variabile a valle di un blocco, moltiplicandola per la funzione di trasferimento del blocco (vedi Figura 5.13);

- lo spostamento di una variabile a monte di un blocco, dividendola per la funzione di trasferimento del blocco (vedi Figura 5.14);

o equivalentemente

- lo spostamento di un blocco a monte di un nodo sommatore, a patto di moltiplicare per la funzione di trasferimento del blocco tutte le altre variabili entrate nel nodo (vedi ancora Figura 5.13);
- lo spostamento di un blocco a valle di un nodo sommatore, a patto di dividere per la funzione di trasferimento del blocco tutte le altre variabili entrate nel nodo (vedi ancora Figura 5.14).

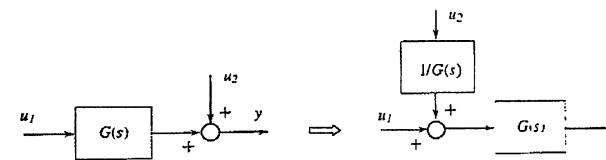


Figura 5.14
Schema equivalente:
spostamento di una
variabile a monte di un
blocco o,
equivalentemente,
spostamento di un blocco
a valle di un nodo
sommatore.

Ai fini della semplificazione di uno schema a blocchi è sempre lecito utilizzare tali regole, anche se a volte possono dar luogo a risultati apparentemente incerti. Per esempio, se la funzione $G(s)$ in Figura 5.14 possiede più poli che zeri, il blocco descritto da $1/G(s)$ non può rappresentare un sistema dinamico proprio. Inoltre, operazioni come quelle mostrate nelle Figure 5.13 e 5.14 fanno aumentare fintizialmente il numero dei poli del sistema complessivo.

Si osservi infine che le regole di elaborazione finora illustrate valgono anche nel caso di schemi in cui appaiono funzioni di trasferimento non razionali, associate per esempio a sistemi lineari e stazionari che contengono un ritardo di tempo.

5.3.5 Cancellazioni

Quando si connettono in uno schema a blocchi un certo numero di sottosistemi ci si aspetta che l'ordine del sistema complessivo (cioè il numero di variabili di stato necessario per descriverlo) sia uguale alla somma degli ordini dei singoli sottosistemi. Può tuttavia accadere che, nel calcolo di una determinata funzione di trasferimento, si pervenga a una funzione razionale con un numero di poli inferiore rispetto alla somma dei numeri di poli dei singoli sottosistemi. In tal caso è evidente che, per effetto della connessione, si sono generate delle cancellazioni, che corrispondono a parti "nascoste" del sistema. L'esempio seguente mostra una di queste situazioni.

Figura 5.13
Schema equivalente:
spostamento di una
variabile a valle di un
blocco o,
equivalentemente,
spostamento di un blocco
a monte di un nodo
sommatore.

Esempio 5.3 Nello schema di Figura 5.15 sia

$$G_1(s) = \frac{s+1}{s+2}, \quad G_2(s) = \frac{1}{s}$$

Il sistema complessivo è allora di ordine 2. Tuttavia, applicando le regole di elaborazione prima illustrate si ricava che la funzione di trasferimento tra u e y è data da

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

e pertanto possiede un solo polo. Si conclude che la connessione di due sistemi del primo ordine ha dato luogo a un sistema che contiene una parte nascosta, poiché agli effetti del legame ingresso-uscita si comporta ancora come un sistema del primo ordine.

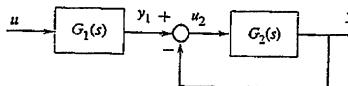


Figura 5.15
Schema a blocchi analizzato nell'Esempio 5.3.

La possibile presenza di parti nascoste dovute alla connessione di sottosistemi sarà discussa in modo più rigoroso nel Paragrafo 5.5.

5.4 Stabilità dei sistemi interconnessi

In questo paragrafo si affronterà il problema della determinazione di condizioni di stabilità asintotica per un sistema costituito da più blocchi interconnessi tra loro. In particolare si discuterà in quali situazioni la stabilità asintotica dei singoli sottosistemi rappresenta una condizione sufficiente e/o necessaria per la stabilità asintotica del sistema complessivo. Si supporrà che i sottosistemi che compongono lo schema siano descritti da funzioni razionali con numeratore e denominatore primi tra loro.

5.4.1 Stabilità di sistemi in serie

Si consideri il sistema di Figura 5.7, ottenuto dalla connessione in serie di due sottosistemi aventi funzioni di trasferimento $G_a(s) = N_a(s)/D_a(s)$ e $G_b(s) = N_b(s)/D_b(s)$. Si è visto al Paragrafo 5.3.1 che la funzione di trasferimento del sistema di Figura 5.7 è data da

$$G(s) = G_a(s)G_b(s) = \frac{N_a(s)N_b(s)}{D_a(s)D_b(s)} \quad (5.6)$$

Quindi, se non intervengono cancellazioni, il denominatore di $G(s)$ è dato dal prodotto dei denominatori di $G_a(s)$ e $G_b(s)$ e i poli di $G(s)$ sono la riunione dei poli dei due sottosistemi. In particolare se i poli di $G_a(s)$ e $G_b(s)$ hanno parte reale negativa, lo stesso vale per i poli di $G(s)$. Si conclude quindi che il sistema di Figura 5.7 è asintoticamente stabile se e solo se lo sono singolarmente i due sottosistemi.

Se invece esistono fattori comuni tra i polinomi N_a e D_b , oppure tra N_b e D_a , il sistema complessivo ha un numero di poli minore della somma dei numeri di poli dei sottosistemi che lo compongono, e perciò la connessione in serie genera una parte nascosta corrispondente a una cancellazione. Se tale cancellazione

riguarda poli con parte reale positiva o nulla, questa parte nascosta non è asintoticamente stabile, e non lo è quindi il sistema complessivo, anche se ciò non viene messo in evidenza dall'analisi dei poli della funzione di trasferimento $G(s)$. Prima di trarre erronee conclusioni sulla stabilità del sistema in base all'esame dei poli della funzione di trasferimento complessiva occorre pertanto assicurarsi che non siano avvenute cancellazioni critiche.

Esempio 5.4 Si consideri la connessione in serie dei sistemi del primo ordine con funzioni di trasferimento

$$G_a(s) = \frac{1}{s-\alpha}, \quad G_b(s) = \frac{s-\alpha}{s+1}$$

dove $\alpha \neq -1$. Ovviamente il primo sistema è asintoticamente stabile per $\alpha < 0$, mentre il secondo lo è per ogni valore di α . Connessione i due sistemi in serie si ottiene un nuovo sistema che, avendo funzione di trasferimento

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

contiene in ogni caso una parte nascosta, associata al polo in $s = \alpha$. Se ci si limitasse a considerare il segno dell'unico polo della funzione di trasferimento $G(s)$ si potrebbe erroneamente concludere che il sistema è asintoticamente stabile indipendentemente da α . Esso invece è asintoticamente stabile se e solo se $\alpha < 0$.

In definitiva, si può affermare che:

- la connessione in serie di sottosistemi asintoticamente stabili genera sempre un sistema asintoticamente stabile;
- la presenza di un sottosistema non asintoticamente stabile in un collegamento in serie rende non asintoticamente stabile il sistema complessivo.

5.4.2 Stabilità di sistemi in parallelo

Si consideri il sistema di Figura 5.8, ottenuto dalla connessione in parallelo di due sottosistemi con funzioni di trasferimento $G_a(s) = N_a(s)/D_a(s)$ e $G_b(s) = N_b(s)/D_b(s)$. Per quanto ricavato al Paragrafo 5.3.2, la funzione di trasferimento del sistema di Figura 5.8 è data da

$$G(s) = G_a(s) + G_b(s) = \frac{N_a(s)D_b(s) + N_b(s)D_a(s)}{D_a(s)D_b(s)} \quad (5.7)$$

Quindi, anche in questo caso, se non intervengono cancellazioni, il denominatore di $G(s)$ è dato dal prodotto dei denominatori di $G_a(s)$ e $G_b(s)$ e i poli di $G(s)$ sono la riunione dei poli dei due sottosistemi. In particolare se i poli di $G_a(s)$ e $G_b(s)$ hanno parte reale negativa, lo stesso vale per i poli di $G(s)$. Si conclude quindi che il sistema di Figura 5.8 è asintoticamente stabile se e solo se lo sono singolarmente i due sottosistemi.

Può invece accadere che nel calcolo di $G(s)$ avvengano delle cancellazioni che, come si è visto, corrispondono a parti nascoste del sistema complessivo. Se tali cancellazioni riguardano poli con parte reale positiva o nulla, queste parti nascoste non sono asintoticamente stabili, e non lo è quindi il sistema complessivo. anche se ciò non viene messo in evidenza dall'analisi dei poli della funzione di trasferimento $G(s)$. Per giudicare la stabilità del sistema complessivo sulla base dei poli di $G(s)$ occorre quindi la stessa cautela segnalata in precedenza riguardo alla presenza di cancellazioni critiche.

Esempio 5.5 Si consideri la connessione in parallelo dei sistemi con funzioni di trasferimento

$$G_a(s) = \frac{1}{s - \alpha} \quad , \quad G_b(s) = \frac{\beta}{(s - \alpha)(s + 1)}$$

Entrambi i sottosistemi sono asintoticamente stabili per $\alpha < 0$. La funzione di trasferimento del sistema complessivo è data da

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

e presenta solo due poli in α e -1 , anziché i tre che ci si sarebbe potuto aspettare. Se poi addirittura fosse $\beta = -1 - \alpha$ risulterebbe

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

e l'analisi del denominatore potrebbe condurre a un'errata conclusione sulla stabilità del sistema.

In definitiva, si può affermare che:

- la connessione in parallelo di sottosistemi asintoticamente stabili genera sempre un sistema asintoticamente stabile;
- la presenza di un sottosistema non asintoticamente stabile in un collegamento in parallelo rende non asintoticamente stabile il sistema complessivo.

5.4.3 Stabilità di sistemi retroazionati

Si passi ora a considerare la stabilità del sistema retroazionato di Figura 5.9, supponendo che sia $G_a(s) = N_a(s)/D_a(s)$ e $G_b(s) = N_b(s)/D_b(s)$. Utilizzando la formula (5.5), la funzione di trasferimento del sistema di Figura 5.9 è data da

$$G(s) = \frac{G_a(s)}{1 + G_a(s)G_b(s)} = \frac{N_a(s)D_b(s)}{D_a(s)D_b(s) + N_a(s)N_b(s)} \quad (5.8)$$

Si osserva immediatamente che i poli del sistema retroazionato (ovvero le radici del denominatore) non dipendono solo dai poli dei sottosistemi (radici di D_a e D_b).

Supponendo che non vi siano cancellazioni nella (5.8), i poli del sistema complessivo (detti anche poli in anello chiuso) sono le radici dell'equazione

$$D_L(s) + N_L(s) = 0 \quad (5.9)$$

o, se si preferisce, dell'equazione

$$1 + L(s) = 0 \quad (5.10)$$

dove si sono indicati con N_L e D_L i polinomi a numeratore e denominatore della funzione d'anello $L(s) = G_a(s)G_b(s)$. L'Equazione (5.9), o anche l'Equazione (5.10), viene detta *equazione caratteristica del sistema retroazionato*.

Ovviamente il sistema di Figura 5.9 è asintoticamente stabile se e solo se tutte le radici dell'Equazione (5.9), o dell'Equazione (5.10), hanno parte reale negativa, ma questa proprietà non ha legami diretti con la stabilità dei singoli sottosistemi.

Si noti che nel calcolo della funzione d'anello $L(s)$ possono avvenire delle cancellazioni polo-zero che, come negli altri casi, corrispondono a parti nascoste del sistema complessivo (pur di scegliere a seconda dei casi $y = y_a$ oppure y_b come variabile di uscita). Se tali cancellazioni riguardano poli con parte reale positiva o nulla, queste parti nascoste non sono asintoticamente stabili, e non lo è

quindi il sistema complessivo, anche se ciò non viene messo in evidenza dall'analisi delle radici dell'Equazione (5.9). Perciò prima di giudicare la stabilità del sistema in base all'Equazione (5.9) occorre assicurarsi che nel calcolo di $L(s)$ non siano avvenute cancellazioni critiche.

Esempio 5.6 Si consideri il sistema retroazionato di Figura 5.9 con

$$G_a(s) = \frac{1}{s - \alpha} \quad , \quad G_b(s) = \frac{s - \alpha}{s - \beta} \quad , \quad \alpha \neq \beta$$

Nel calcolo della funzione d'anello $L(s) = G_a(s)G_b(s) = 1/(s - \beta)$ avviene una cancellazione polo-zero. Se $\alpha \geq 0$ questa è una cancellazione critica e il sistema complessivo non è asintoticamente stabile. Se invece $\alpha < 0$ la stabilità del sistema retroazionato può essere giudicata in base all'equazione caratteristica

$$D_L(s) + N_L(s) = s - \beta + 1 = 0$$

che ha come unica radice $s = \beta - 1$. In conclusione il sistema è asintoticamente stabile per $\alpha < 0$ e $\beta < 1$. Si osservi in particolare che il sistema retroazionato può essere asintoticamente stabile anche quando il sistema descritto da $G_b(s)$ non lo è. Ciò avviene per valori di β compresi tra 0 e 1.

Quanto detto finora si può facilmente riformulare a proposito dei sistemi retroazionati positivamente come quello mostrato nella Figura 5.10, con l'unica accortezza di sostituire alle (5.9), (5.10) le equazioni

$$D_L(s) - N_L(s) = 0 \quad (5.11)$$

$$1 - L(s) = 0 \quad (5.12)$$

Allora, a meno di cancellazioni critiche, la proprietà di stabilità asintotica è presente se e solo se le radici dell'Equazione (5.11), o dell'Equazione (5.12), hanno parte reale negativa.

Determinare la posizione dei poli di un sistema retroazionato, o anche solo accettarne le proprietà di stabilità, è dunque cosa non semplice in quanto i poli di un sistema retroazionato sono in generale diversi dai poli dei sottosistemi che lo compongono. Questa circostanza, d'altra parte, rende particolarmente interessante l'uso di schemi retroazionati al fine di modificare la dinamica di un sistema. Il collegamento in retroazione può infatti servire a collocare in posizioni opportune i poli del sistema o anche a stabilizzare un sistema instabile, spostandone i poli all'interno della regione di stabilità. Si noti che è impossibile ottenere gli stessi risultati limitandosi a collegamenti in serie e in parallelo che, come visto in precedenza, non alterano la posizione dei poli.

Naturalmente, ai potenziali vantaggi offerti dalla retroazione si affiancano altrettanto evidenti pericoli. Una scelta malaccorta dello schema di retroazione può infatti generare instabilità anche a partire da singoli componenti asintoticamente stabili.

Riassumendo:

- la connessione in retroazione di sistemi asintoticamente stabili può generare un sistema non asintoticamente stabile;
- un sistema retroazionato può essere asintoticamente stabile anche se alcuni dei sottosistemi non sono asintoticamente stabili.

Una discussione approfondita sui metodi che consentono di analizzare le proprietà di stabilità di un sistema retroazionato è rimandata ai Capitoli 9 e 12.

5.4.4 Stabilità e schemi equivalenti

Quando si elabora uno schema a blocchi usando le regole di spostamento di variabili o spostamento di blocchi discusse al termine del Paragrafo 5.3.4 occorre ricordarsi che, anche se tali regole conservano inalterate le relazioni ingresso-uscita, esse possono condurre a conclusioni errate nei riguardi della stabilità.

Si supponga, per esempio, che la funzione di trasferimento $G(s)$ di Figura 5.14 corrisponda a un sistema asintoticamente stabile, ma con uno zero a parte reale positiva. Esaminando lo schema di destra si potrebbe erroneamente dedurre che il sistema complessivo è instabile, visto che nel calcolo della funzione di trasferimento tra u e y avviene una cancellazione critica. In realtà, la parte nascosta instabile è stata fintiziamente introdotta nel passaggio dallo schema a blocchi originario a quello equivalente. Anche se i due schemi sono equivalenti ai fini del calcolo delle funzioni di trasferimento, essi non lo sono affatto per quanto riguarda l'analisi della stabilità. Quest'ultima può essere correttamente giudicata solo sullo schema originario.

5.5 Raggiungibilità e osservabilità dei sistemi interconnessi

Una questione di un certo interesse riguarda le proprietà di raggiungibilità e di osservabilità di un sistema descritto attraverso uno schema a blocchi. In particolare ci si può chiedere sotto quali condizioni l'interconnessione di sistemi in forma minima, cioè completamente raggiungibili e osservabili, produce un sistema complessivo in forma minima. Anche in questo caso, per semplicità, ci si limiterà a trattare il caso di sistemi SISO.

Come si ricorderà, la presenza di parti non raggiungibili e/o non osservabili (cioè di parti nascoste) viene messa in luce dal fatto che il grado del denominatore della funzione di trasferimento complessiva è inferiore all'ordine del sistema, cioè al numero di variabili di stato. La perdita delle proprietà di raggiungibilità e osservabilità è dunque strettamente connessa al fenomeno delle cancellazioni discusso nel Paragrafo 5.3.5, come evidenziato dal seguente esempio.

Esempio 5.7 (Seguito dell'Esempio 5.3). Si considerino le seguenti realizzazioni (minime) dei sottosistemi descritti da $G_1(s)$ e $G_2(s)$:

$$\dot{x}_1(t) = -2x_1(t) - u(t)$$

$$y_1(t) = x_1(t) + u(t)$$

e

$$\dot{x}_2(t) = u_2(t)$$

$$y(t) = x_2(t)$$

Osservando che dallo schema di Figura 5.15 risulta $u_2(t) = y_1(t) - y(t)$, il sistema complessivo è descritto dalle seguenti equazioni:

$$\dot{x}_1(t) = -2x_1(t) - u(t)$$

$$\dot{x}_2(t) = x_1(t) - x_2(t) + u(t)$$

$$y(t) = x_2(t)$$

Si può facilmente verificare che quest'ultimo sistema non è completamente raggiungibile. Infatti la sua matrice di raggiungibilità vale

$$M_r = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}$$

e dunque ha rango pari a 1. La matrice di osservabilità è invece data da

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

e pertanto il sistema è completamente osservabile. Dal punto di vista del legame ingresso-uscita, esso è quindi equivalente a un sistema del primo ordine, come evidenziato dalla funzione di trasferimento, calcolata in precedenza,

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

Saranno ora discusse più in dettaglio le condizioni sotto le quali la connessione in serie, parallelo o in retroazione di due sottosistemi può generare sistemi in forma non minima.

Sistemi in serie Nel Paragrafo 5.4.1 si è visto che, quando si collegano in serie due sottosistemi con funzioni di trasferimento $G_a(s) = N_a(s)/D_a(s)$ e $G_b(s) = N_b(s)/D_b(s)$, possono verificarsi delle cancellazioni che coinvolgono fattori comuni tra i polinomi N_a e D_b , oppure tra i polinomi N_b e D_a . Si tratta di capire se le parti nascoste così generate sono non raggiungibili o non osservabili. A tale proposito si può dimostrare che:

- quando uno zero di $G_a(s)$ cancella un polo di $G_b(s)$ si genera una parte non raggiungibile e osservabile;
- quando un polo di $G_a(s)$ cancella uno zero di $G_b(s)$ si genera una parte raggiungibile e non osservabile.

Sistemi in parallelo Come si è visto nel Paragrafo 5.4.2, quando si collegano in parallelo due sottosistemi con funzioni di trasferimento $G_a(s) = N_a(s)/D_a(s)$ e $G_b(s) = N_b(s)/D_b(s)$ possono verificarsi delle cancellazioni nel calcolo della funzione di trasferimento complessiva. Alcuni fattori del polinomio $N_a D_b + N_b D_a$ che compare al numeratore nella (5.7) possono infatti semplificarsi con fattori dei polinomi D_a o D_b . In realtà è immediato rendersi conto che tali semplificazioni si verificano solo se D_a e D_b hanno fattori in comune. Più precisamente si può dimostrare che:

- quando $G_a(s)$ e $G_b(s)$ hanno in comune dei poli si genera una parte contemporaneamente non raggiungibile e non osservabile.

Sistemi in retroazione Anche nel caso di collegamento in retroazione di due sottosistemi può verificarsi una perdita di raggiungibilità o di osservabilità. Con riferimento agli schemi di Figura 5.9 e Figura 5.10, e fattorizzando le due funzioni di trasferimento come $G_a(s) = N_a(s)/D_a(s)$ e $G_b(s) = N_b(s)/D_b(s)$, avviene una cancellazione nella funzione di trasferimento tra u e y quando i polinomi N_a e D_b possiedono fattori comuni. In particolare, si dimostra che:

- quando uno zero di $G_a(s)$ coincide con un polo di $G_b(s)$ si genera una parte contemporaneamente non raggiungibile e non osservabile.

Si noti invece che quando uno zero di $G_b(s)$ coincide con un polo di $G_a(s)$ il sistema con ingresso u e uscita y rimane completamente raggiungibile e osservabile (infatti il denominatore della (5.8) non perde di grado), ma il corrispondente polo

di $G_a(s)$ non viene spostato dalla retroazione. In effetti un'eventuale radice comune tra i polinomi N_b e D_a può essere raccolta a fattore comune al denominatore della (5.8). Questo è ciò che avviene nel caso dell'Esempio 5.6 prima discusso.

Naturalmente, in un sistema retroazionato più complicato, che contenga un numero maggiore di blocchi, possono verificarsi cancellazioni che coinvolgono le sole funzioni di trasferimento sulla linea di andata oppure soltanto quelle sulla linea di retroazione. In tal caso si applicano, a seconda di come sono collegati i blocchi, le conclusioni viste a proposito di sistemi in serie o in parallelo. Un aspetto importante è che le parti non raggiungibili e/o non osservabili così generate non vengono alterate dalla presenza della retroazione.

5.6 Conclusioni

In questo capitolo è stato introdotto il formalismo grafico degli schemi a blocchi, particolarmente utile per descrivere e analizzare il comportamento di sistemi anche di elevata complessità, ma costituiti dall'interconnessione di componenti elementari. Con riferimento a sistemi SISO sono state ricavate le regole per calcolare le funzioni di trasferimento di sistemi formati da blocchi in serie, in parallelo e in retroazione. Particolare attenzione è stata dedicata al problema dell'analisi di stabilità, arrivando a mettere in luce la criticità di tale analisi nel caso dei sistemi retroazionati. Per questi sistemi infatti la stabilità dei singoli sottosistemi non rappresenta né una condizione necessaria né una condizione sufficiente per la stabilità del sistema complessivo. Quest'ultimo argomento sarà ampiamente ripreso nel Capitolo 9.

Esercizi

Esercizio 5.1 Disegnare lo schema a blocchi del sistema idraulico considerato nell'Esempio 5.1 supponendo che il sistema di controllo sia progettato per regolare il livello del primo serbatoio, anziché quello del secondo. Si calcolino poi le nuove funzioni di trasferimento tra gli ingressi h_1^0 (andamento desiderato del livello h_1) e v_1 e l'uscita h_2 .

Esercizio 5.2 Verificare che la (5.5) rappresenta effettivamente la funzione di trasferimento di un sistema dinamico proprio solo se vale la relazione

$$\lim_{s \rightarrow \infty} G_a(s)G_b(s) = \lim_{s \rightarrow \infty} L(s) \neq -1$$

In caso contrario si usa dire che il sistema contiene un *anello algebrico non ben definito*.

Esercizio 5.3 Utilizzando le regole per lo spostamento di blocchi a monte o a valle di un nodo, modificare lo schema di Figura 5.6 in modo che sulla linea di retroazione non compaia alcun blocco (*retroazione unitaria*) e che l'effetto del disturbo v_2 sia spostato a valle del blocco $G_2(s)$.

Esercizio 5.4 Dimostrare, utilizzando le rappresentazioni di stato, che gli autovalori del sistema di Figura 5.7 sono la riunione degli autovalori dei due sottosistemi. Questo risulta conferma che, anche se un autovalore instabile di un sottosistema non appare tra i poli della funzione di trasferimento complessiva per effetto di una cancellazione, il sistema rimane comunque instabile a causa di tale autovalore.

Ripetere poi l'esercizio nel caso della connessione in parallelo di Figura 5.8.

Esercizio 5.5 Con riferimento allo schema a retroazione negativa di Figura 5.9 si calcolino i poli in anello chiuso con

$$G_a(s) = \frac{1}{s^2 + 3s - 4}, \quad G_b(s) = 6$$

e si verifichi che la retroazione consente in questo caso di stabilizzare un sistema instabile.

Esercizio 5.6 Con riferimento allo schema a retroazione negativa di Figura 5.9 si calcolino i poli in anello chiuso con

$$G_a(s) = \frac{9}{(1+s)^2}, \quad G_b(s) = \frac{3}{1+s}$$

e si verifichi che in questo caso il sistema retroazionato è instabile anche se i due sottosistemi che lo compongono sono asintoticamente stabili.

Esercizio 5.7 Calcolare le funzioni di trasferimento tra gli ingressi w e d e l'uscita y per il sistema retroazionato di Figura 5.16. Si noti che solo nella funzione di trasferimento tra d e y compare il polo instabile $s = 1$. Interpretare tale risultato in termini di presenza di parti nascoste.

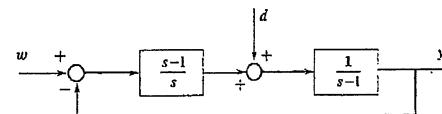


Figura 5.16
Sistema retroazionato
dell'Esercizio 5.7.

Esercizio 5.8 Dopo aver ricavato le rappresentazioni di stato delle funzioni di trasferimento $G_a(s)$ e $G_b(s)$ dell'Esempio 5.4, verificare che il sistema ottenuto dalla loro connessione in serie risulta non completamente osservabile (se $G_a(s)$ è a monte di $G_b(s)$) oppure non completamente raggiungibile (se $G_a(s)$ è a valle di $G_b(s)$).

Esercizio 5.9 Mostrare che, collegando in parallelo due integratori, il sistema risultante è ancora un integratore, con guadagno pari alla somma dei guadagni. Verificare inoltre mediante una rappresentazione di stato che il sistema complessivo non è né completamente raggiungibile né completamente osservabile.

Esercizio 5.10 Si riprenda l'Esempio 5.6, considerando però y_b come variabile di uscita. Si verifichi che $s = \alpha$ è l'autovalore di una parte raggiungibile ma non osservabile.

6.1 Introduzione

L'analisi nel dominio della frequenza dei sistemi dinamici lineari e stazionari costituisce uno degli strumenti più potenti, versatili ed espressivi per lo studio di alcune loro importanti proprietà, a integrazione delle tecniche nel dominio del tempo descritte nel Capitolo 3. L'analisi in frequenza di un sistema si basa in prima istanza sullo studio del suo comportamento quando esso viene sollecitato da un ingresso di tipo sinusoidale, ma, grazie al principio di sovrapposizione degli effetti, può essere estesa a classi di segnali ben più ampie. Si potranno infatti prendere in considerazione tutte quelle funzioni per cui è possibile effettuare una scomposizione armonica, che permette di rappresentare il segnale come una combinazione lineare di un numero finito o infinito di componenti sinusoidali. Ricadono in questa famiglia le funzioni periodiche sviluppabili in serie di Fourier e le funzioni dotate di trasformata di Fourier.

In particolare l'attenzione sarà rivolta ai seguenti argomenti:

- il calcolo della risposta a un ingresso sinusoidale e l'enunciazione del teorema fondamentale della risposta in frequenza;
- l'estensione al caso di ingressi esprimibili come combinazioni lineari di sinusoidi;
- l'introduzione del fondamentale concetto di risposta in frequenza di un sistema e dei suoi legami con la funzione di trasferimento;
- alcuni cenni sul problema della rilevazione sperimentale della risposta in frequenza;
- lo studio delle diverse rappresentazioni grafiche della risposta in frequenza;
- la classificazione dei sistemi dinamici in relazione alle loro caratteristiche di azione filtrante, con una breve discussione su alcune importanti tipologie di filtri;
- l'interpretazione del concetto di poli dominanti nel dominio della frequenza.

6.2 Risposta alla sinusoide

Si vuole qui studiare il movimento di un sistema lineare e stazionario quando viene sollecitato da un ingresso di tipo sinusoidale. Per il calcolo del movimento si potrebbe naturalmente utilizzare la formula di Lagrange, presentata nel Paragrafo 3.2.1, specializzandola al caso in cui $u(t)$ è una sinusode. Si vedrà però dall'analisi che segue che si possono trarre interessanti conclusioni sul comportamento del sistema anche a partire dalla sua funzione di trasferimento.

6.2.1 Calcolo dell'uscita sinusoidale

Si consideri il sistema SISO di ordine n descritto da

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (6.1)$$

$$y(t) = Cx(t) + Du(t) \quad (6.2)$$

e la funzione di trasferimento associata

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

Si voglia determinare la risposta del sistema (6.1), (6.2) a un ingresso sinusoidale del tipo

$$u(t) = U \sin(\omega_0 t), \quad t \geq 0$$

sotto l'ipotesi che il sistema sia asintoticamente stabile. Ricordando la trasformata di Laplace della sinusoide, si ricava

$$Y(s) = G(s) \frac{U\omega}{s^2 + \omega^2} \quad (6.3)$$

e l'uscita $y(t)$ può essere ottenuta antitrasformando tale espressione mediante lo sviluppo di Heaviside. Per esempio, se $G(s)$ possiede solo poli reali distinti, si ottiene

$$y(t) = \mathcal{L}^{-1} \left[\underbrace{\sum_{i=1}^n \frac{P_i}{s + p_i}}_{Y_1(s)} + \underbrace{\frac{Q}{s - j\omega} + \frac{\bar{Q}}{s + j\omega}}_{Y_2(s)} \right] = y_1(t) + y_2(t) \quad (6.4)$$

con opportuni valori delle costanti P_i e Q , mentre \bar{Q} è il complesso coniugato di Q . In particolare, il numero complesso Q è il residuo associato al polo in $j\omega$, cioè, ricordando la (6.3),

$$Q = [(s - j\omega) Y(s)]_{s=j\omega} = G(j\omega) \frac{U\omega}{j\omega + j\omega} = G(j\omega) \frac{U}{2j} \quad (6.5)$$

Osservando poi che per una funzione razionale $G(s)$ vale l'identità $G(-j\omega) = \tilde{G}(j\omega)$, risulta anche

$$\bar{Q} = -\tilde{G}(j\omega) \frac{U}{2j} \quad (6.6)$$

Grazie all'ipotesi di asintotica stabilità, la componente $y_1(t)$, che è una combinazione di esponenziali decrescenti, tende asintoticamente a zero per $t \rightarrow \infty$ e pertanto l'uscita $y(t)$ tende asintoticamente a $y_2(t)$. Si osservi ora che

$$y_2(t) = \mathcal{L}^{-1} \left[\frac{Q}{s - j\omega} + \frac{\bar{Q}}{s + j\omega} \right] = Q e^{j\omega t} + \bar{Q} e^{-j\omega t}$$

e, in base alle (6.5), (6.6) e all'applicazione di semplici regole sui numeri complessi, si ottiene

$$\begin{aligned} y_2(t) &= G(j\omega) \frac{U}{2j} e^{j\omega t} - \tilde{G}(j\omega) \frac{U}{2j} e^{-j\omega t} = \\ &= \frac{U}{2j} [(G(j\omega) - \tilde{G}(j\omega)) \cos \omega t + j(G(j\omega) + \tilde{G}(j\omega)) \sin \omega t] = \\ &= \frac{U}{2j} [2j \operatorname{Im} G(j\omega) \cos \omega t + 2j \operatorname{Re} G(j\omega) \sin \omega t] = \\ &= U [|G(j\omega)| \sin(\arg G(j\omega)) \cos \omega t + |G(j\omega)| \cos(\arg G(j\omega)) \sin \omega t] = \\ &= |G(j\omega)| U \sin(\omega t + \arg G(j\omega)) \end{aligned}$$

In definitiva, l'uscita $y(t)$ converge verso una sinusoide che ha la stessa pulsazione della sinusoide in ingresso ed è caratterizzata dall'ampiezza $Y = |G(j\omega)| U$ e dallo sfasamento $\vartheta = \arg G(j\omega)$. Si noti tra l'altro che ciò avviene per qualunque condizione iniziale, visto che l'effetto sull'uscita di un arbitrario stato iniziale $x(0)$ tende comunque ad annullarsi in virtù della proprietà di asintotica stabilità.

È quindi possibile formulare il seguente importante risultato, che viene spesso indicato come *teorema fondamentale della risposta in frequenza*.

Teorema 6.1 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(s)$ l'ingresso sinusoidale

$$u(t) = U \sin(\omega_0 t)$$

l'uscita a transitorio esaurito assume l'andamento

$$\tilde{y}(t) = |G(j\omega_0)| U \sin(\omega_0 t + \arg G(j\omega_0))$$

indipendentemente dallo stato iniziale. ■

In base a tale risultato, la conoscenza di modulo e fase della funzione $G(j\omega)$ al variare di ω permette di determinare il comportamento (asintotico) del sistema nei riguardi di ingressi sinusoidali di diversa pulsazione.

Si noti inoltre che, in risposta a un ingresso sinusoidale di pulsazione ω_0 , l'uscita a regime può risultare nulla solo se la funzione di trasferimento del sistema possiede una coppia di zeri puramente immaginari in $\pm j\omega_0$, così da far risultare $G(j\omega_0) = 0$.

Un'ultima osservazione riguarda la velocità con cui l'uscita si assesta sull'andamento asintotico. Ovviamente, la durata del transitorio che si osserva prima che y possa ritenersi a tutti gli effetti pratici coincidente con l'uscita asintotica \tilde{y} dipende dalla dinamica propria del sistema, e può essere valutata in base al tempo di assestamento T_{ae} introdotto al Paragrafo 4.4.

Esempio 6.1 (Seguito dell'Esempio 2.4) Con riferimento al circuito elettrico della Figura 2.3, si è già calcolato il movimento dell'uscita conseguente alla condizione iniziale $x(0) = x_0$ e all'ingresso $u(t) = U \sin(\omega_0 t)$, $t \geq 0$. Ricordando le (2.8)-(2.9) esso è dato da

$$\begin{aligned} y(t) &= -e^{-t/RC} x_0 - \frac{U\omega_0 RC}{1 + \omega_0^2 R^2 C^2} e^{-t/RC} + \\ &+ \frac{U\omega_0 RC}{\sqrt{1 + \omega_0^2 R^2 C^2}} \cos(\omega_0 t - \arctan(\omega_0 RC)) \end{aligned} \quad (6.7)$$

Al crescere del tempo, i primi due termini tendono asintoticamente a zero e l'uscita converge verso l'andamento periodico

$$\tilde{y}(t) = \frac{U\omega_0 RC}{\sqrt{1 + \omega_0^2 R^2 C^2}} \cos(\omega_0 t - \arctan(\omega_0 RC)) \quad (6.8)$$

Questo risultato è coerente con la teoria sviluppata dato che la funzione di trasferimento del sistema è

$$G(s) = \frac{RCs}{1 + RCs} \quad (6.9)$$

e pertanto

$$|G(j\omega_0)| = \frac{\omega_0 RC}{\sqrt{1 + \omega_0^2 R^2 C^2}}, \quad \arg G(j\omega_0) = \frac{\pi}{2} - \arctan(\omega_0 RC)$$

Uno studio dettagliato della funzione $|G(j\omega)|$ al variare di ω permette inoltre di affermare che, nei confronti di ingressi sinusoidali, il sistema tende ad attenuare sensibilmente le sinusoide a "bassa"

pulsazione (fino a bloccare completamente quelle a pulsazione $\omega = 0$, cioè gli ingressi costanti), mentre lascia passare praticamente inalterate le sinusoidi a pulsazione "elevata" (si noti che per $\omega \rightarrow \infty$ risulta $|G(j\omega)| \rightarrow 1$ e $\arg G(j\omega) \rightarrow 0$).

6.2.2 Risposta in frequenza: definizione e proprietà

La funzione complessa

$$G(j\omega) = C(j\omega I - A)^{-1}B + D$$

definita per valori della variabile reale ω non negativi viene chiamata *risposta in frequenza* associata al sistema (6.1), (6.2). Formalmente essa coincide con la restrizione della funzione di trasferimento $G(s)$ ai punti appartenenti al semiasse immaginario positivo. Per motivi che saranno più chiari nel seguito è opportuno estendere tale definizione anche a sistemi non asintoticamente stabili, per i quali il risultato enunciato nel Teorema 6.1 non è valido nella stessa forma. Se però $G(s)$ possiede dei poli sull'asse immaginario, i corrispondenti valori della pulsazione ω vanno esclusi nella definizione della risposta in frequenza per evitare che essa assuma valore infinito.

Si osservi poi che, come già ricordato, risulta

$$G(-j\omega) = \bar{G}(j\omega) \quad (6.10)$$

dove la sopralineatura indica l'operazione di coniugazione. Quindi la conoscenza della risposta in frequenza permette, se necessario, il calcolo di $G(j\omega)$ anche per valori negativi di ω .

Il concetto di risposta in frequenza, che è stato qui introdotto per sistemi SISO e di ordine n finito, può essere in realtà esteso anche a sistemi multivariabili e a sistemi a dimensione infinita, purché lineari e stazionari. In generale, per un sistema con funzione di trasferimento $G(s)$, si chiamerà risposta in frequenza la funzione $G(j\omega)$ definita per tutti i valori non negativi di ω per cui la definizione è ben posta. Per esempio, la risposta in frequenza associata al ritardo di tempo, descritto dalla relazione

$$y(t) = u(t - \tau)$$

e con funzione di trasferimento $G(s) = e^{-\tau s}$, è data da

$$G(j\omega) = e^{-j\omega\tau} \quad , \quad \omega \geq 0 \quad (6.11)$$

Si noti che anche per questo sistema vale un risultato identico a quello del Teorema 6.1. Infatti applicando l'ingresso

$$u(t) = U \sin(\omega_0 t) \quad , \quad t \geq 0$$

al ritardo di tempo, l'uscita vale

$$y(t) = U \sin(\omega_0 t - \omega_0 \tau) \quad , \quad t \geq \tau$$

e inoltre in base alla (6.11) risulta $|G(j\omega_0)| = 1$ e $\arg G(j\omega_0) = -\omega_0 \tau$. Dunque, anche in questo caso, il modulo e l'argomento della risposta in frequenza rappresentano rispettivamente il fattore di amplificazione e lo sfasamento che subisce la sinusode di ingresso. Oltre tutto qui si osserva una perfetta coincidenza tra l'uscita y e il suo andamento asintotico già a partire dall'istante τ .

6.3 Risposta a segnali dotati di serie o trasformata di Fourier

Il risultato enunciato nel Teorema 6.1 è alla base di importanti generalizzazioni relative a segnali di ingresso esprimibili come combinazioni lineari di funzioni sinusoidali. In effetti, come discusso nell'Appendice B, esistono importanti classi di segnali che possono essere rappresentati come combinazioni lineari di un insieme (finito o infinito, numerabile o no) di componenti sinusoidali, dette anche armoniche. I risultati discussi nel paragrafo precedente possono essere estesi facilmente anche a questo tipo di segnali osservando che, grazie al principio di sovrapposizione degli effetti, in un sistema lineare e stazionario l'effetto di una singola armonica in ingresso può essere calcolato indipendentemente dalla presenza delle altre componenti. Ciò verrà illustrato nel seguito con riferimento alla classe dei segnali periodici dotati di sviluppo in serie di Fourier e alla classe dei segnali dotati di trasformata di Fourier.

6.3.1 Segnali sviluppabili in serie di Fourier

Si supponga di applicare al sistema (6.1), (6.2) un ingresso u periodico di periodo $T > 0$, cioè tale per cui $u(t + T) = u(t)$, $\forall t$. Come è noto (si veda il Paragrafo B.4), sotto ipotesi molto blande un segnale di questo tipo è rappresentabile attraverso il suo sviluppo in serie di Fourier come

$$u(t) = \sum_{n=-\infty}^{+\infty} U_n e^{jn\omega_0 t} \quad , \quad \omega_0 = 2\pi/T \quad (6.12)$$

I coefficienti complessi U_n costituiscono lo spettro del segnale u e rappresentano il contributo delle varie armoniche alla formazione del segnale stesso.

In base all'analisi sviluppata nel precedente paragrafo e sfruttando il principio di sovrapposizione degli effetti non è difficile dimostrare il seguente risultato.

Teorema 6.2 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(s)$ l'ingresso periodico (6.12), l'uscita a transitorio esaurito assume l'andamento

$$\bar{y}(t) = \sum_{n=-\infty}^{+\infty} Y_n e^{jn\omega_0 t}$$

con

$$Y_n = G(jn\omega_0)U_n \quad (6.13)$$

indipendentemente dallo stato iniziale.

Quindi un ingresso periodico genera un'uscita periodica, con lo stesso periodo dell'ingresso, il cui spettro $\{Y_n\}$ è legato allo spettro $\{U_n\}$ dell'ingresso attraverso la relazione (6.13), dove $\omega_0 = 2\pi/T$ è la pulsazione fondamentale. In altre parole, la n -esima armonica presente nell'ingresso subisce un'amplificazione pari al fattore $|G(jn\omega_0)|$ e uno sfasamento angolare pari ad $\arg G(jn\omega_0)$.

In particolare un'armonica di pulsazione $\bar{\omega} > 0$ può essere presente nel movimento periodico dell'uscita solo se è presente nell'ingresso, cioè se $\bar{\omega} = n\omega_0$, per qualche n intero, e se $U_n \neq 0$. Se però la funzione di trasferimento $G(s)$ possiede una coppia di zeri immaginari in $s = \pm j\bar{\omega}$, allora $G(j\bar{\omega}) = 0$ e tale armonica viene "bloccata" dal sistema.

Esempio 6.2 (Seguito dell'Esempio 6.1) Si alimenti il circuito mostrato nella Figura 2.3 con un ingresso periodico a onda quadra definito da

$$\tilde{u}(t) = \begin{cases} 0 & -T/2 \leq t < -T/4 \\ V & -T/4 \leq t < T/4 \\ 0 & T/4 \leq t < T/2 \end{cases}$$

e rappresentabile mediante la (6.12) ponendo (si veda l'Esempio B.14)

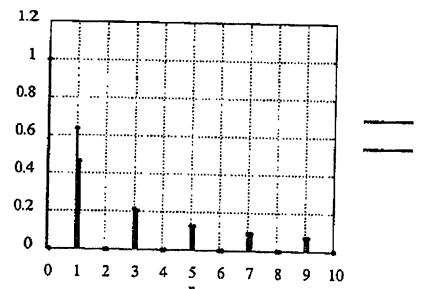
$$U_0 = \frac{V}{2} \quad , \quad U_n = \frac{V \sin(n\pi/2)}{n\pi/2} \quad , \quad n = \dots, -2, -1, 1, 2, \dots$$

Grazie al Teorema 6.2 e ricordando la (6.9), si può affermare che la corrispondente uscita periodica \tilde{y} ha uno spettro pari a

$$Y_0 = 0 \quad , \quad Y_n = \frac{jn\omega_0 RC}{1 + jn\omega_0 RC} \frac{V \sin(n\pi/2)}{n\pi/2} \quad , \quad n = \dots, -2, -1, 1, 2, \dots$$

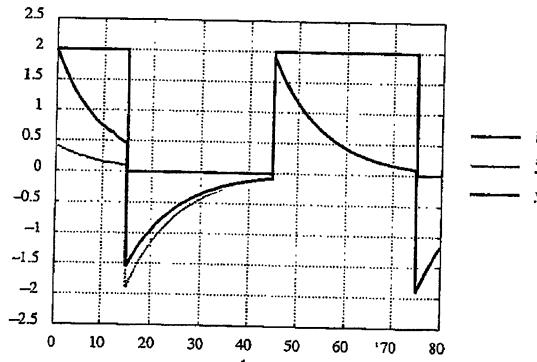
dove $\omega_0 = 2\pi/T$. In particolare si osservi che il sistema "blocca" completamente la componente costante del segnale in ingresso, corrispondente all'armonica "zero", e che l'uscita, così come l'ingresso, non contiene altre armoniche di indice pari poiché risulta $Y_n = 0$ per $n = \pm 2, \pm 4, \dots$. Un confronto tra gli spettri di ampiezza dell'ingresso e dell'uscita quando $RC = 10$, $V = 2$ e $T = 60$ è riportato nella Figura 6.1. Come si nota, tra le armoniche di indice dispari solo la prima subisce

Figura 6.1
Spettro dell'ingresso \tilde{u}
e dell'uscita \tilde{y}
nell'Esempio 6.2.



una significativa attenuazione. La Figura 6.2 mostra l'andamento temporale di \tilde{y} ed evidenzia come l'uscita y del sistema a partire da stato iniziale nullo converga asintoticamente verso tale andamento periodico. Il relativo transitorio può ritenersi praticamente esaurito dopo un intervallo pari a circa cinque costanti di tempo, e cioè $T_{el} \approx 5RC = 50$.

Figura 6.2
Andamento dell'ingresso \tilde{u} ,
dell'uscita asintotica \tilde{y}
e dell'uscita y a partire da
stato iniziale nullo
nell'Esempio 6.2.



6.3.2 Segnali dotati di trasformata di Fourier

Si consideri ora il caso di un ingresso u dotato di trasformata di Fourier $U(j\omega) = \mathcal{F}[u(t)]$ (si veda il Paragrafo B.5). È ben noto che la trasformata di Fourier costituisce una rappresentazione in frequenza del segnale, nel senso che si può scrivere

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} U(j\omega) e^{j\omega t} d\omega \quad (6.14)$$

e quindi la funzione u è scomponibile in un'infinità non numerabile di armoniche, con pulsazioni che coprono l'intero asse reale, ognuna moltiplicata per il coefficiente $U(j\omega)$.

La presenza nella (6.14) dell'integrale, al posto della sommatoria che compare nella (6.12), non modifica concettualmente l'analisi svolta nel caso precedente. In virtù della proprietà di linearità dell'operazione di integrazione, si può infatti fare ricorso ancora al principio di sovrapposizione degli effetti per giungere alla seguente conclusione.

Teorema 6.3 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(s)$ l'ingresso (6.14), l'uscita a transitorio esaurito assume l'andamento

$$\tilde{y}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Y(j\omega) e^{j\omega t} d\omega$$

con

$$Y(j\omega) = G(j\omega)U(j\omega) \quad (6.15)$$

indipendentemente dallo stato iniziale. ■

Quindi, anche in questo caso, il movimento asintotico dell'uscita non contiene armoniche che non siano presenti nel segnale di ingresso, e le armoniche dell'ingresso sono modulate secondo la (6.15). Alcune di esse possono risultare completamente cancellate dal sistema se $G(s)$ possiede zeri sull'asse immaginario.

Particolarmente interessante è il caso in cui l'ingresso è nullo prima dell'istante $t = 0$. In tal caso si dimostra facilmente che la (6.15) rappresenta la trasformata di Fourier della componente forzata dell'uscita, cioè quella che si ottiene partendo da $x(0) = 0$.

Tutto ciò consente di attribuire una nuova interpretazione alla risposta in frequenza $G(j\omega)$. Per sistemi asintoticamente stabili la formula (6.15) mette in luce che $G(j\omega)$ rappresenta il rapporto tra gli spettri dell'uscita asintotica e dell'ingresso per tutti i valori di ω per cui non sia nullo lo spettro dell'ingresso $U(j\omega)$. Se l'ingresso è nullo prima dell'istante $t = 0$, allora $G(j\omega)$ rappresenta il rapporto tra gli spettri dell'uscita e dell'ingresso a partire da condizioni iniziali nulle.

Esempio 6.3 (Seguito degli Esempi 3.5 e 4.2) Il sistema meccanico mostrato nella Figura 3.4 è asintoticamente stabile ed è descritto dalla funzione di trasferimento

$$G(s) = \frac{1}{Ms^2 + hs + k}$$

Partendo da condizioni di quiete si solleciti il sistema con il seguente ingresso a forma di impulso rettangolare

$$u(t) = \begin{cases} 1 & 0 \leq t < L \\ 0 & \text{altrimenti} \end{cases}$$

la cui trasformata di Fourier (si veda l'Esempio B.20) è

$$U(j\omega) = e^{-j\omega L/2} L \frac{\sin(\omega L/2)}{\omega L/2}$$

In particolare lo spettro di ampiezza di u è

$$|U(j\omega)| = L \frac{|\sin(\omega L/2)|}{\omega L/2}$$

In accordo con la precedente teoria, lo spettro dell'uscita è dato da

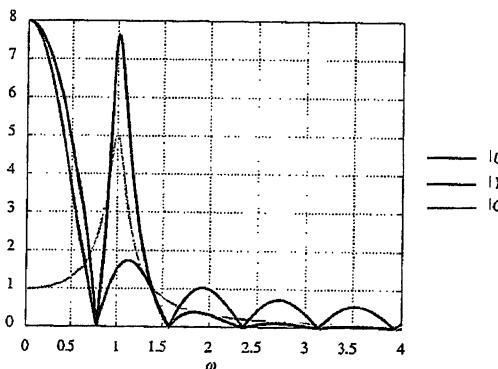
$$Y(j\omega) = G(j\omega)U(j\omega) = \frac{1}{k - M\omega^2 + jh\omega} e^{-j\omega L/2} L \frac{\sin(\omega L/2)}{\omega L/2}$$

e dunque lo spettro di ampiezza è

$$|Y(j\omega)| = \frac{1}{\sqrt{(k - M\omega^2)^2 + h^2\omega^2}} L \frac{|\sin(\omega L/2)|}{\omega L/2}$$

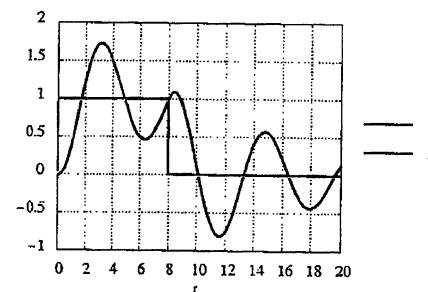
Nella Figura 6.3 sono mostrati gli andamenti di $|U(j\omega)|$, $|G(j\omega)|$ e $|Y(j\omega)|$ per $k = 1$, $M = 1$, $h = 0.2$, $L = 8$. Come si nota, il sistema tende ad amplificare sensibilmente le componenti

Figura 6.3
Spettro d'ampiezza
dell'ingresso e dell'uscita e
modulo della risposta in
frequenza nell'Esempio
6.3.



armoniche dell'ingresso a pulsazioni vicine alla pulsazione naturale $\omega_n = \sqrt{k/M} = 1$ associata ai poli. Questo effetto, legato al fenomeno della *risonanza*, è dovuto alla presenza di un coefficiente di attrito h relativamente basso, che induce un debole smorzamento ($\xi = h/2\sqrt{kM} = 0.1$) e può essere apprezzato anche osservando l'andamento temporale dell'uscita y , mostrato nella Figura 6.4, in cui è evidente una forte componente sinusoidale con periodo prossimo a 2π .

Figura 6.4
Andamento dell'ingresso e
dell'uscita nell'Esempio 6.3.



6.4 Complementi

I risultati appena visti sulla risposta in frequenza possono essere estesi in varie direzioni. Si possono per esempio applicare gli stessi ragionamenti al caso della risposta a un ingresso esponenziale. Inoltre è possibile trarre qualche conclusione anche nel caso in cui il sistema in esame non sia asintoticamente stabile.

6.4.1 Risposta esponenziale

Si voglia calcolare il movimento del sistema (6.1), (6.2) in risposta a un ingresso esponenziale del tipo

$$u(t) = U e^{\lambda t}, \quad t \geq 0$$

supponendo che λ non coincida con alcun autovalore di A e, per il momento, che il sistema sia asintoticamente stabile.

Con considerazioni simili a quelle del Paragrafo 6.2.1, non è difficile dimostrare che vale il risultato seguente.

Teorema 6.4 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(s)$ l'ingresso esponenziale

$$u(t) = U e^{\lambda t}$$

l'uscita a transitorio esaurito assume l'andamento

$$\hat{y}(t) = G(\lambda) U e^{\lambda t} \quad (6.16)$$

indipendentemente dallo stato iniziale. ■

Naturalmente, la proprietà appena enunciata è davvero significativa solo quando l'andamento asintotico (6.16) non converge a zero, cioè quando la costante λ è positiva o nulla, oppure quando la velocità con cui $e^{\lambda t}$ tende a zero è inferiore a quella con cui si esaurisce il transitorio dovuto ai poli del sistema. Per essere più precisi, il risultato per un generico valore di λ andrebbe enunciato affermando che per $t \rightarrow \infty$ la differenza tra l'uscita $y(t)$ e la funzione $\hat{y}(t)$ tende a zero per qualunque stato iniziale.

Il Teorema 6.4 consente di illustrare una fondamentale proprietà degli zeri di un sistema dinamico. A questo proposito si ricordi che si chiama zero di un sistema SISO un valore complesso s che rende nulla la funzione di trasferimento $G(s)$ (Paragrafo 4.2.2). Alla luce di quanto appena visto, applicando l'ingresso $u(t) = U e^{\lambda t}$ con λ coincidente con uno zero del sistema, l'uscita tende asintoticamente ad annullarsi per qualsiasi valore dello stato iniziale.

Questa importante proprietà degli zeri viene di solito chiamata *proprietà bloccante*. Con opportune precisazioni, che però esulano dagli scopi della presente trattazione, essa potrebbe essere estesa anche al caso dei sistemi multivariabili.

6.4.2 Il caso di sistemi instabili

Quando il sistema in esame non è asintoticamente stabile, i risultati dei paragrafi precedenti non sono più validi. Per esempio, se si applica un ingresso sinusoidale a un sistema instabile non è più vero in generale che l'uscita tenda a una sinusoida, visto che il termine $y_1(t)$ nella formula (6.4) del movimento forzato non converge necessariamente a zero. Ciononostante si può dimostrare che, con un'opportuna

scelta dello stato iniziale, il termine $y_1(t)$ può essere esattamente controbilanciato dalla componente libera del movimento.

Per discutere più in dettaglio la questione, si consideri dapprima il caso della risposta esponenziale del sistema (6.1), (6.2), rimuovendo l'ipotesi di asintotica stabilità. Ci si può chiedere se, alimentando il sistema con l'ingresso

$$u(t) = U e^{\lambda t}, \quad t \geq 0 \quad (6.17)$$

esiste un valore di $x(0)$ tale che il movimento dello stato sia anch'esso esponenziale del tipo $x(t) = x(0)e^{\lambda t}$. Usando la (6.1) si ricava che deve essere

$$\lambda x(0)e^{\lambda t} = Ax(0)e^{\lambda t} + BUe^{\lambda t}$$

ovvero

$$(\lambda I - A)x(0) = BU$$

Dunque il valore cercato di $x(0)$ esiste ed è unico se e solo se λ non è un autovalore di A . Sotto tale ipotesi il movimento dello stato associato all'ingresso (6.17) è allo stato iniziale

$$x(0) = (\lambda I - A)^{-1}BU$$

è dato da

$$x(t) = (\lambda I - A)^{-1}BUe^{\lambda t}, \quad t \geq 0$$

In corrispondenza l'uscita vale

$$y(t) = [C(\lambda I - A)^{-1}B + D]Ue^{\lambda t} = G(\lambda)Ue^{\lambda t}, \quad t \geq 0$$

e quindi presenta un andamento esponenziale dello stesso tipo di quello dell'ingresso, ma moltiplicato per il fattore $G(\lambda)$.

In definitiva, quando manca l'ipotesi di asintotica stabilità, solo scegliendo opportunamente lo stato iniziale il movimento dell'uscita coincide con la funzione (6.16).

Ad analoghe conclusioni si perviene quando si considera il caso della risposta alla sinusoide. Basta infatti osservare che $\sin(\omega t) = \text{Im}(e^{j\omega t})$ e che i risultati precedenti continuano formalmente a valere anche quando la costante λ dell'Equazione (6.17) assume valori complessi. Grazie alla linearità si possono poi ricavare risultati simili anche nel caso di segnali in ingresso dotati di serie o trasformata di Fourier. Valgono dunque i seguenti teoremi.

Teorema 6.5 Si supponga che il sistema (6.1), (6.2) con funzione di trasferimento $G(s)$ non abbia autovalori in $\pm j\omega_0$ e si applichi a esso l'ingresso

$$u(t) = U \sin(\omega_0 t), \quad t \geq 0$$

Allora esiste uno stato iniziale $x(0)$ per cui l'uscita è sinusoidale e vale

$$y(t) = |G(j\omega_0)|U \sin(\omega_0 t + \arg G(j\omega_0)) \quad t \geq 0$$

Teorema 6.6 Si supponga che il sistema (6.1), (6.2) con funzione di trasferimento $G(s)$ non abbia autovalori in $\pm jn\omega_0$ (con n intero) e si applichi a esso l'ingresso

$$u(t) = \sum_{n=-\infty}^{+\infty} U_n e^{jn\omega_0 t}, \quad t \geq 0$$

Allora esiste uno stato iniziale $x(0)$ per cui l'uscita è periodica e vale

$$y(t) = \sum_{n=-\infty}^{+\infty} G(jn\omega_0)U_n e^{jn\omega_0 t}, \quad t \geq 0$$

Teorema 6.7 Si supponga che il sistema (6.1), (6.2) con funzione di trasferimento $G(s)$ non abbia autovalori sull'asse immaginario e si applichi a esso l'ingresso

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} U(j\omega)e^{j\omega t} d\omega, \quad t \geq 0$$

Allora esiste uno stato iniziale $x(0)$ per cui l'uscita è dotata di trasformata di Fourier e vale

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G(j\omega)U(j\omega)e^{j\omega t} d\omega, \quad t \geq 0$$

6.5 Identificazione sperimentale della risposta in frequenza

Poiché la risposta in frequenza costituisce una rappresentazione completa e particolarmente utile di un sistema dinamico, è interessante porsi il problema della sua determinazione a partire dalla rilevazione di misure sperimentali. I risultati discussi nei precedenti paragrafi suggeriscono tecniche di identificazione della risposta in frequenza ampiamente diffuse nelle applicazioni pratiche.

Se ci si limita a considerare sistemi fisici con una variabile di ingresso e una variabile di uscita e si ritiene che essi siano, almeno in prima approssimazione, descrivibili attraverso un sistema lineare, stazionario e asintoticamente stabile, sulla base del Teorema 6.1 e dei commenti che lo seguono, il valore della risposta in frequenza in $\omega = \omega_0$ può essere determinato attraverso il seguente esperimento. A partire da una condizione di equilibrio, si sollecita il sistema con un ingresso sinusoidale $u(t) = U \sin(\omega_0 t)$ e si attende il tempo necessario perché il sistema si possa considerare a regime. Quindi si misurano l'ampiezza e la fase della sinusoide di uscita. Il modulo della risposta in frequenza in $\omega = \omega_0$ risulta pari al rapporto tra le ampiezze della sinusoide di uscita e di quella di ingresso, mentre la sua fase è pari alla differenza tra le loro fasi. Ripetendo questa operazione applicando funzioni di ingresso con diverse pulsazioni, si ottengono così i valori della risposta in frequenza per tutti i valori di pulsazione di interesse. Da qui, se necessario, con tecniche di interpolazione si può determinare un'espressione analitica della risposta in frequenza, e quindi della funzione di trasferimento.

L'identificazione sperimentale della risposta in frequenza può essere compiuta più velocemente se si utilizzano segnali di ingresso più "ricchi" di armoniche e si fa riferimento ai Teoremi 6.2 o 6.3. Impiegando segnali di questo tipo si possono determinare in un singolo esperimento più valori della risposta in frequenza. Per far ciò occorre solamente separare le armoniche che compongono il segnale di uscita (mediante un calcolo numerico dello spettro o attraverso l'uso di opportuni filtri) e riportarle a quelle presenti nell'ingresso.

6.6 Diagrammi cartesiani o di Bode

Per i sistemi SISO, i *diagrammi cartesiani*, o *diagrammi di Bode*, sono probabilmente la forma più usata per rappresentare graficamente la risposta in frequenza $G(j\omega)$ associata alla funzione di trasferimento $G(s)$. Essi sono costituiti da una coppia di curve che rappresentano in funzione della pulsazione ω il modulo e la fase di $G(j\omega)$. Le due curve sono dette *diagramma di Bode del modulo* e *diagramma di Bode della fase*.

Nel tracciamento dei diagrammi è comodo utilizzare una scala logaritmica in base dieci per l'ascissa, dove è riportata la pulsazione ω , eventualmente normalizzata rispetto a un valore di riferimento. Così facendo, la distanza tra i due punti che rappresentano le due generiche pulsazioni ω_1 e $\omega_2 > \omega_1$ è proporzionale, secondo un opportuno fattore di scala, al rapporto ω_2/ω_1 , cioè alla differenza tra i logaritmi di ω_2 e ω_1 , anziché alla differenza $\omega_2 - \omega_1$, come accadrebbe in una scala lineare. Quindi, considerate quattro pulsazioni $\omega_1, \omega_2, \omega_3, \omega_4$ tali che $\omega_1/\omega_2 = \omega_3/\omega_4$, la distanza tra i punti che rappresentano ω_3 e ω_4 è uguale a quella tra i punti che rappresentano ω_1 e ω_2 . In particolare, si chiama *decade* un intervallo tra due pulsazioni che sono tra loro in un rapporto pari a dieci. Da quanto sopra specificato segue che la pulsazione nulla non compare al finito sull'asse.

Nel tracciamento dei diagrammi cartesiani è conveniente assumere che la funzione $G(s)$ sia assegnata nella forma fattorizzata (4.26), qui riportata per comodità

$$G(s) = \frac{\mu \prod_i (1 + \tau_i s) \prod_i (1 + 2\xi_i s / \alpha_{ni} + s^2 / \alpha_{ni}^2)}{s^g \prod_i (1 + T_i s) \prod_i (1 + 2\xi_i s / \omega_{ni} + s^2 / \omega_{ni}^2)}$$

a cui corrisponde la risposta in frequenza

$$G(j\omega) = \frac{\mu \prod_i (1 + j\omega\tau_i) \prod_i (1 + 2j\xi_i \omega / \alpha_{ni} - \omega^2 / \alpha_{ni}^2)}{(j\omega)^g \prod_i (1 + j\omega T_i) \prod_i (1 + 2j\xi_i \omega / \omega_{ni} - \omega^2 / \omega_{ni}^2)} \quad (6.18)$$

6.6.1 Diagramma del modulo

Nel diagramma del modulo l'asse delle ordinate riporta in scala lineare il valore del modulo della risposta in frequenza espresso in *decibel*, o in dB. Convenzionalmente, il valore in decibel di una quantità positiva x è dato da $20 \log x$, dove il logaritmo è in base dieci. Pertanto

$$|G(j\omega)|_{dB} = 20 \log |G(j\omega)|$$

e valori positivi, negativi e nulli di $|G(j\omega)|_{dB}$ corrispondono a valori di $|G(j\omega)|$ maggiori, minori e pari a uno.

Per le scelte effettuate il modulo in dB della (6.18) è

$$\begin{aligned} |G(j\omega)|_{dB} = & 20 \log |\mu| - 20g \log |j\omega| + \sum_i 20 \log |1 + j\omega\tau_i| + \\ & + \sum_i 20 \log |1 + 2j\xi_i \omega / \alpha_{ni} - \omega^2 / \alpha_{ni}^2| - \sum_i 20 \log |1 + j\omega T_i| + \\ & - \sum_i 20 \log |1 + 2j\xi_i \omega / \omega_{ni} - \omega^2 / \omega_{ni}^2| \end{aligned} \quad (6.19)$$

e quindi il tracciamento del diagramma può essere compiuto considerando dapprima separatamente i termini che compaiono nella (6.19) e successivamente sommando i relativi contributi. Inoltre, osservando che per un qualunque numero complesso $s \neq 0$ risulta $|1/s|_{dB} = -|s|_{dB}$, i contributi al diagramma di Bode

del modulo dei fattori di $G(j\omega)$ corrispondenti agli zeri di $G(s)$ si ricavano immediatamente con un semplice cambio di segno a partire dai contributi dei poli. Per un'analisi completa è sufficiente quindi considerare il modulo della risposta in frequenza associata ai termini

$$G_a(s) = \mu \quad (6.20)$$

$$G_b(s) = \frac{1}{s} \quad (6.21)$$

$$G_c(s) = \frac{1}{1 + Ts} \quad (6.22)$$

$$G_d(s) = \frac{1}{1 + 2\xi s / \omega_n + s^2 / \omega_n^2} \quad (6.23)$$

Nel seguito si mostreranno le principali caratteristiche del diagramma di Bode del modulo relativo alle funzioni (6.20)-(6.23) e per le stesse si introdurranno dei diagrammi approssimati, detti *diagrammi asintotici*. Questi ultimi sono di semplissimo tracciamento e il loro impiego consente nella maggioranza dei casi di determinare, senza l'ausilio di strumenti di calcolo e spesso con un accettabile livello di approssimazione, un andamento qualitativo del diagramma esatto associato a funzioni $G(s)$ di struttura anche molto complessa. In ogni caso è possibile valutare l'errore che si commette nell'approssimazione e quindi, se necessario, apportare le opportune correzioni.

Diagramma del modulo di $G_a(j\omega)$ Dalla definizione (6.20) risulta

$$|G_a(j\omega)|_{dB} = 20 \log |\mu|$$

a cui corrisponde un grafico costituito da una retta parallela all'asse delle ω con ordinata positiva, negativa o nulla a seconda che il modulo di μ sia maggiore di uno, minore di uno o unitario.

Diagramma del modulo di $G_b(j\omega)$ In questo caso, dalla (6.21) segue che

$$|G_b(j\omega)|_{dB} = 20 \log \left| \frac{1}{j\omega} \right| = -20 \log \omega \quad (6.24)$$

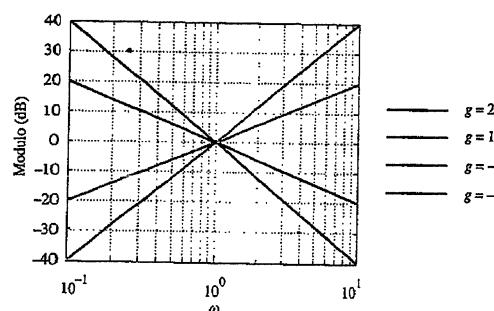
Il diagramma è una retta, in quanto sia la scala delle ascisse sia la dipendenza da ω sono di tipo logaritmico. È allora sufficiente conoscere il diagramma in due punti per poterlo tracciare completamente; a questo riguardo si osservi che $|G_b(j1)|_{dB} = 0$ e $|G_b(j10)|_{dB} = -20$. Convenzionalmente si usa indicare come unitaria la pendenza di 20 dB/decade e si usa dire che la retta ha pendenza -1 o che "perde" 20 dB/decade.

Più in generale, il diagramma di Bode del modulo associato a $G(s) = 1/s^g$ è una retta di ordinata 0 in $\omega = 1$ e con pendenza $-g$, che "perde" quindi $20g$ dB/decade per $g > 0$ (azioni integrali) e "guadagna" $20|g|$ dB/decade per $g < 0$ (azioni derivative). Il diagramma relativo ai casi $g = \pm 1$ e $g = \pm 2$ è riportato nella Figura 6.5.

Diagramma del modulo di $G_c(j\omega)$ La risposta in frequenza associata alla (6.22) ha come modulo

$$|G_c(j\omega)|_{dB} = 20 \log \left| \frac{1}{1 + j\omega T} \right| = -20 \log \sqrt{1 + \omega^2 T^2} \quad (6.25)$$

Figura 6.5
Diagramma di Bode di $|G_b(j\omega)|$.

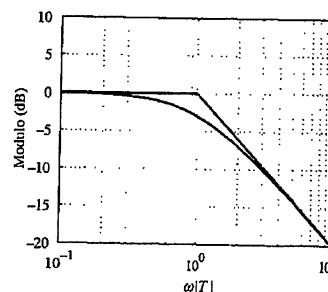


il cui grafico è riportato nella Figura 6.6. Notando che risulta

$$|G_c(j\omega)|_{\text{dB}} \simeq \begin{cases} -20 \log 1 = 0, & \omega \ll 1/|T| \\ -20 \log \omega |T|, & \omega \gg 1/|T| \end{cases}$$

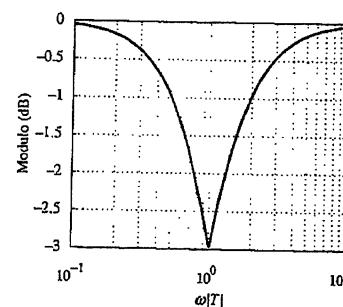
si può rappresentare in modo approssimato la curva di Equazione (6.25) con il diagramma asintotico, anch'esso riportato nella Figura 6.6, costituito per $\omega \leq 1/|T|$ dalla semiretta di pendenza 0 e ordinata nulla e per $\omega > 1/|T|$ dalla semiretta di pendenza -1 e ordinata nulla in $\omega = 1/|T|$. Questo diagramma asintotico è tangente a quello esatto per $\omega \rightarrow 0$ e $\omega \rightarrow \infty$.

Figura 6.6
Diagramma di Bode esatto e asintotico di $|G_c(j\omega)|$.



Il diagramma dell'errore $E_c(\omega)$ che si commette sostituendo il diagramma asintotico a quello esatto è riportato nella Figura 6.7. Esso assume il modulo massimo pari a $20 \log \sqrt{2} \simeq -3$ dB in $\omega = 1/|T|$.

Figura 6.7
Diagramma dell'errore $E_c(\omega)$.



Come già affermato, il diagramma esatto e quello asintotico associati al termine $G(s) = 1 + Ts$, corrispondente a uno zero reale, si ottengono direttamente come immagine speculare rispetto all'asse delle pulsazioni di quelli relativi a $G_c(s)$. Si osservi da ultimo che l'andamento di questi diagrammi è indipendente dal segno della costante di tempo T .

Diagramma del modulo di $G_d(j\omega)$ In questo caso, dalla (6.23) si ha

$$\begin{aligned} |G_d(j\omega)|_{\text{dB}} &= 20 \log \left| \frac{1}{1 + 2j\xi\omega/\omega_n - \omega^2/\omega_n^2} \right| = \\ &= -20 \log \sqrt{(1 - \omega^2/\omega_n^2)^2 + 4\xi^2\omega^2/\omega_n^2} \end{aligned} \quad (6.26)$$

Il grafico di questa funzione dipende ovviamente dal modulo, ma non dal segno, dello smorzamento ξ ed è riportato per diversi valori di questo parametro nella Figura 6.8. In particolare, un semplice studio di funzione relativo alla (6.26) mostra che il grafico ha un massimo se $|\xi| < 1/\sqrt{2} \simeq 0.707$. Questo massimo, chiamato *picco di risonanza*, è in corrispondenza della *pulsazione di risonanza*

$$\omega_r = \omega_n \sqrt{1 - 2\xi^2} \quad (6.27)$$

e risulta

$$|G_d(j\omega_r)| = \frac{1}{2|\xi| \sqrt{1 - \xi^2}} \quad (6.28)$$

L'andamento del picco di risonanza (6.28) in funzione di $|\xi|$ è riportato nella Figura 6.9.

È anche facile verificare che

$$|G_d(j\omega_n)| = \frac{1}{2|\xi|} \quad (6.29)$$

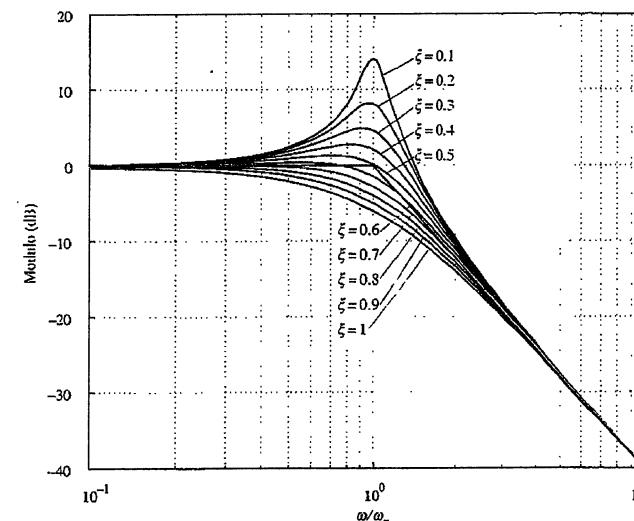
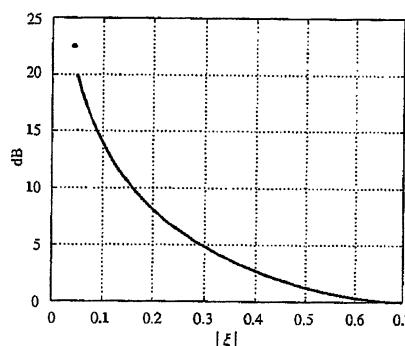


Figura 6.8
Diagramma di Bode esatto e asintotico di $|G_d(j\omega)|$.

Figura 6.9
Andamento del picco di risonanza in funzione di $|\xi|$.



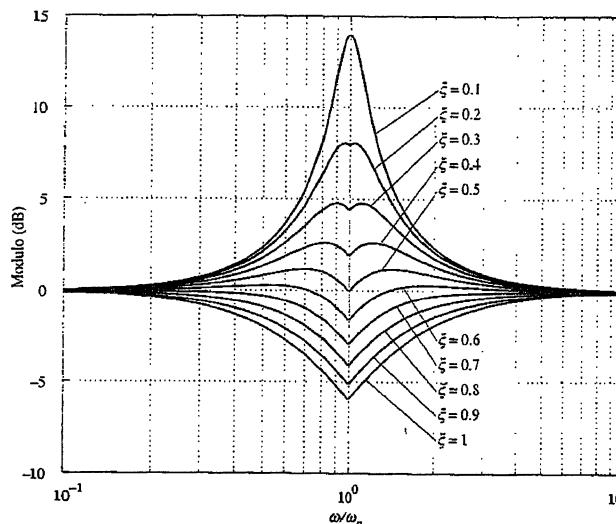
Pertanto per $|\xi| = 0.5$ risulta $|G_d(j\omega_n)|_{dB} = 0$, mentre per $\xi = 0$ dalle (6.27), (6.28) segue che $\omega_r = \omega_n$ e il picco di risonanza è infinito, coerentemente con il fatto che in questo caso la funzione ha una coppia di poli in $s = \pm j\omega_n$.

Per quanto riguarda il diagramma asintotico, si noti che dalla (6.26) si ha

$$|G_d(j\omega)|_{dB} \approx \begin{cases} -20 \log 1 = 0, & \omega \ll \omega_n \\ -40 \log(\omega/\omega_n), & \omega \gg \omega_n \end{cases}$$

Ciò suggerisce di rappresentare in modo approssimato la curva con il diagramma asintotico costituito per $\omega \leq \omega_n$ dalla semiretta di pendenza e ordinata nulle, e per $\omega > \omega_n$ dalla semiretta di pendenza -2 e ordinata nulla per $\omega = \omega_n$. Questo diagramma asintotico coincide quindi con quello relativo a una coppia di poli reali in $s = -\omega_n$ ed è tangente a quello effettivo per $\omega \rightarrow 0$ e $\omega \rightarrow \infty$. Il diagramma asintotico è anch'esso riportato nella Figura 6.8. L'errore $E_d(\omega)$ che si commette sostituendo il diagramma asintotico a quello esatto è riportato nella Figura 6.10; esso può assumere valori anche molto elevati per $\omega \simeq \omega_n$ e bassi valori di smorzamento.

Figura 6.10
Diagramma dell'errore $E_d(\omega)$.



Anche in questo caso i diagrammi esatto e asintotico associati a $G(s) = 1 + 2\xi s/\omega_n + s^2/\omega_n^2$ si possono ottenere come immagine speculare rispetto all'asse delle pulsazioni di quelli associati a $G_d(s)$. In particolare, se la coppia di zeri considerata è puramente immaginaria, cioè $\xi = 0$, il diagramma tende a $-\infty$, coerentemente con il fatto che la funzione ha una coppia di zeri in $s = \pm j\omega_n$ e quindi per $\omega = \omega_n$ il suo modulo è uguale a zero.

Tracciamento del diagramma asintotico del modulo. Il diagramma asintotico di una qualsiasi risposta in frequenza si ottiene sommando come specificato nell'Equazione (6.19) i diagrammi associati ai singoli fattori della (6.18). Si comprende però facilmente che questo diagramma, che è costituito da una spezzata i cui vari tratti hanno pendenza intera, può essere tracciato direttamente senza ricavare prima i diagrammi degli addendi. Infatti, a pulsazioni minori di tutti i termini $1/|\tau_i|$, $1/|T_i|$, α_{ni} e ω_{ni} , gli unici fattori di $G(j\omega)$ che lo determinano sono μ e $1/(j\omega)^g$. All'aumentare della pulsazione la pendenza viene poi modificata dalla presenza degli altri poli e zeri.

Si riconosce pertanto che per il tracciamento si possono applicare le seguenti regole:

- il tratto iniziale del diagramma ha pendenza $-g$ e in $\omega = 1$ esso, o il suo prolungamento, assume il valore $20 \log |\mu|$;
- in corrispondenza dei valori di ω uguali a $1/|\tau_i|$ o $1/|T_i|$ la pendenza aumenta o diminuisce di un numero di unità pari alla molteplicità dello zero o del polo corrispondenti;
- in corrispondenza dei valori di ω uguali alle pulsazioni naturali α_{ni} o ω_{ni} la pendenza aumenta o diminuisce di un numero di unità pari al doppio della molteplicità delle coppie di zeri o poli corrispondenti.

Il diagramma esatto si può poi ottenere sommando i diagrammi esatti dei singoli fattori di $G(s)$, oppure sommando al diagramma asintotico tutti i diagrammi degli errori relativi alle singolarità di $G(s)$ fuori dall'origine.

Esempio 6.4 Si consideri la funzione di trasferimento

$$G(s) = \frac{100s(1 - 0.1s)}{(1 + 16s + 100s^2)(1 + 0.01s)} \quad (6.30)$$

I diagrammi asintotico ed esatto del modulo di $G(j\omega)$ sono presentati nella Figura 6.11. Poiché la funzione ha guadagno 100 e uno zero nell'origine, il diagramma asintotico è bassa pulsazione

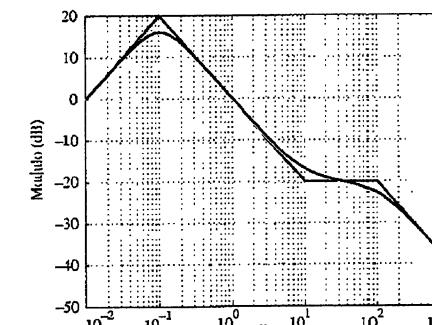


Figura 6.11
Diagramma di Bode del modulo della funzione (6.30).

ha pendenza +1 e il prolungamento del primo tratto in $\omega = 1$ assume il valore -40. Quindi alla pulsazione $\omega = 0.1$, corrispondente alla pulsazione naturale della coppia di poli complessi con smorzamento 0.8, la pendenza diminuisce di due e assume il valore -1. In $\omega = 10$, per la presenza dello zero, la pendenza aumenta di uno, e infine in $\omega = 100$, per la presenza del polo, la pendenza diminuisce di uno. Si noti che, poiché la minima distanza tra le singolarità del sistema è pari a una decade, il diagramma asintotico costituisce una buona approssimazione di quello effettivo.

Per concludere, si osservi che la pendenza assunta dal diagramma asintotico del modulo per $\omega \rightarrow \infty$ è sempre pari al grado relativo, con il segno cambiato, della corrispondente funzione di trasferimento, cioè alla differenza tra il grado del polinomio a numeratore e il grado di quello a denominatore. Pertanto tale pendenza è negativa per sistemi strettamente propri e nulla per sistemi propri, ma non strettamente.

6.6.2 Diagramma della fase

Nel diagramma della fase l'asse delle ordinate riporta in scala lineare, tarata in gradi o radienti, il valore di $\arg G(j\omega)$; la scelta qui effettuata è quella di esprimere questo valore in gradi. Dalla (6.18), ricordando che per un qualunque numero complesso $s \neq 0$ risulta $\arg(1/s) = -\arg(s)$, segue che

$$\begin{aligned} \arg G(j\omega) &= \arg \mu - g \arg(j\omega) + \sum_i \arg(1 + j\omega\tau_i) + \\ &\quad + \sum_i \arg(1 + 2j\xi_i \omega / \alpha_{ni} - \omega^2 / \alpha_{ni}^2) - \sum_i \arg(1 + j\omega T_i) + \\ &\quad - \sum_i \arg(1 + 2j\xi_i \omega / \omega_{hi} - \omega^2 / \omega_{ni}^2) \end{aligned} \quad (6.31)$$

e quindi, anche in questo caso, il tracciamento del diagramma può essere effettuato considerando dapprima separatamente i termini che compaiono nella (6.31) e successivamente sommando i singoli contributi. Per questo motivo nel seguito verranno analizzati i diagrammi della fase associati alle funzioni di trasferimento elementari definite nelle (6.20)-(6.23); per alcuni di essi verrà inoltre introdotto un corrispondente diagramma asintotico.

La convenzione che si adotterà per il calcolo dell'argomento dei singoli contributi della (6.31) è quella di assegnare il valore di minimo modulo tra tutti i possibili, che comporta di considerare valori in modulo non superiori a 180° . Naturalmente il valore $\arg G(j\omega)$ risultante dalla somma dei singoli contributi della (6.31) potrà invece avere modulo maggiore di 180° .

Diagramma della fase di $G_a(j\omega)$ In base alla definizione (6.20) si ha

$$\arg G_a(j\omega) = \arg \mu = \begin{cases} 0^\circ, & \mu > 0 \\ -180^\circ, & \mu < 0 \end{cases} \quad (6.32)$$

a cui corrisponde una retta parallela all'asse delle pulsazioni con ordinata 0° o -180° . Si noti che la scelta di assegnare lo sfasamento di -180° piuttosto che $+180^\circ$, quando $\mu < 0$ è puramente convenzionale.

Diagramma della fase di $G_b(j\omega)$ Dalla (6.21) risulta

$$\arg G_b(j\omega) = \arg \left(\frac{1}{j\omega} \right) = -90^\circ \quad (6.33)$$

Poiché la fase è negativa, si usa dire che il polo nell'origine produce un *ritardo di fase, o ritarda*.

Più in generale, il diagramma di Bode della fase associato a $G(s) = 1/s^g$ è una retta parallela all'asse delle pulsazioni e di ordinata $-g90^\circ$. Nel caso di azioni derivate ($g < 0$) questo contributo è positivo e pertanto si usa dire che zeri nell'origine producono un *anticipo di fase, o anticipano*.

Diagramma della fase di $G_c(j\omega)$ La risposta in frequenza associata a $G_c(s)$, definita tramite la (6.22), ha come fase

$$\arg G_c(j\omega) = -\arg(1 + j\omega T) = -\arctan(\omega T)$$

Il diagramma corrispondente è riportato per $T > 0$ nella Figura 6.12, mentre per $T < 0$ è speculare rispetto all'asse delle pulsazioni. La fase è negativa se il polo $s = -1/T$ è negativo, mentre è positiva se il polo è positivo; nel primo caso il polo "ritarda", nel secondo "anticipa". Risulta inoltre $\arg G_c(j/T) = \pm 45^\circ$.

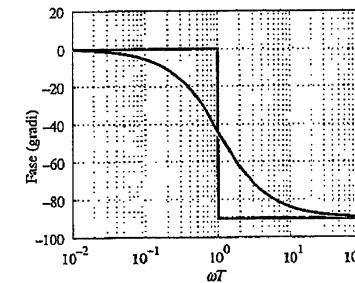


Figura 6.12
Diagramma di Bode esatto e asintotico di $\arg G_c(j\omega)$. $T > 0$.

Si noti poi che

$$\arg G_c(j\omega) \simeq \begin{cases} -\arg(1) = 0^\circ, & \omega \ll 1/|T| \\ -\arg(j\omega T) = \begin{cases} -90^\circ, & T > 0 \\ +90^\circ, & T < 0 \end{cases}, & \omega \gg 1/|T| \end{cases}$$

e quindi un possibile diagramma asintotico associato alla fase di $G_c(j\omega)$ è costituito per $\omega < 1/|T|$ dalla semiretta orizzontale di ordinata nulla e per $\omega > 1/|T|$ dalla semiretta orizzontale di ordinata -90° , se $T > 0$, o $+90^\circ$, se $T < 0$. Il diagramma asintotico è tangente a quello esatto per $\omega \rightarrow 0$ e $\omega \rightarrow \infty$, ma anche per valori di ω molto diversi da $1/|T|$ fornisce un'approssimazione piuttosto scadente del diagramma esatto, come mostrato nella Figura 6.12 per $T > 0$.

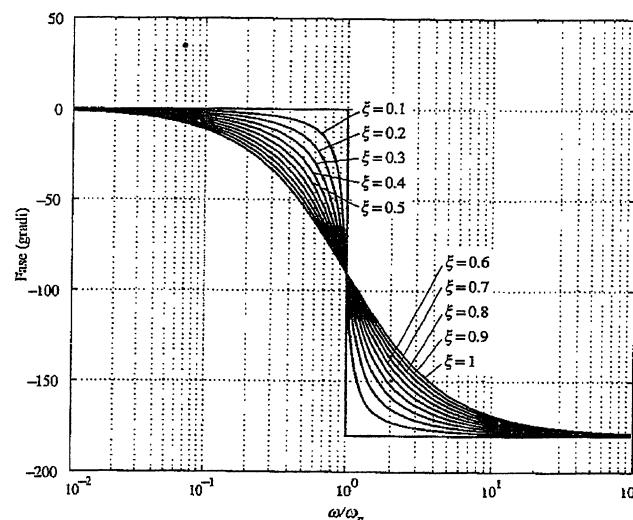
È anche possibile impiegare altri tipi di diagramma asintotico (per esempio lineari a tratti) che forniscono una rappresentazione più precisa e il cui tracciamento risulta lievemente più complesso. D'altra parte, per l'uso che ne verrà fatto in seguito, l'approssimazione introdotta è più che sufficiente.

Il diagramma associato al termine $G(s) = 1 + Ts$, corrispondente a uno zero reale, è simmetrico rispetto a quello di $G_c(s)$. Perciò, uno zero negativo "anticipa" e uno zero positivo "ritarda".

Diagramma della fase di $G_d(j\omega)$ In questo caso, dalla definizione (6.23) si ha

$$\arg G_d(j\omega) = -\arg \left(1 + \frac{2j\xi\omega}{\omega_n} - \frac{\omega^2}{\omega_n^2} \right)$$

Figura 6.13
Diagramma di Bode esatto e asintotico di $\arg G_d(j\omega)$.
 $\xi > 0$.



Il grafico di questa funzione, riportato nella Figura 6.13 per valori di $\xi > 0$, dipende ovviamente dal valore, in modulo e segno, dello smorzamento ξ . Si noti in particolare che per $\xi > 0$ (poli a parte reale negativa) risulta $\arg G_d(j\omega_n) = -90^\circ$, cioè i poli danno un contributo di ritardo, mentre per $\xi < 0$ risulta $\arg G_d(j\omega_n) = +90^\circ$, cioè i poli danno un contributo di anticipo. Per $\xi = 0$ si ha

$$G_d(j\omega) = \frac{1}{1 - \omega^2/\omega_n^2}$$

e quindi $G_d(j\omega)$ è un numero reale positivo (con sfasamento nullo) per $\omega < \omega_n$ e negativo (con sfasamento -180° per convenzione) per $\omega > \omega_n$.

Anche per la fase di $G_d(j\omega)$ si può definire un diagramma asintotico, riportato sempre nella Figura 6.13, costituito per $\omega \leq \omega_n$ dalla semiretta orizzontale di ordinata nulla e per $\omega > \omega_n$ dalla semiretta orizzontale di ordinata -180° se $\xi \geq 0$ e $+180^\circ$ se $\xi < 0$. Si osservi che il diagramma asintotico coincide quindi con quello relativo a una coppia di poli puramente immaginari in $s = \pm j\omega_n$ ed è tangente a quello esatto per $\omega \rightarrow 0$ e $\omega \rightarrow \infty$; tuttavia, per valori di ω diversi da ω_n e per smorzamenti ξ elevati, l'approssimazione che si ottiene usando il diagramma asintotico è scadente.

Per quanto riguarda i diagrammi esatto e asintotico associati al termine $G(s) = 1 + 2\xi s/\omega_n + s^2/\omega_n^2$, essi si ottengono come immagine speculare rispetto all'asse delle pulsazioni di quelli associati a $G_d(j\omega)$. Pertanto i diagrammi asintotico ed effettivo di zeri con $\xi \geq 0$ e $\xi < 0$ si ottengono ribaltando rispetto all'asse delle pulsazioni quelli di corrispondenti poli con $\xi \geq 0$ e $\xi < 0$. In conclusione, zeri con $\xi \geq 0$ danno anticipo di fase, mentre zeri con $\xi < 0$ producono ritardo di fase.

Tracciamento del diagramma asintotico della fase Il diagramma asintotico della fase di una qualsiasi risposta in frequenza è quello che si ottiene sommando, come specificato nell'Equazione (6.31), i diagrammi associati ai singoli fattori della (6.18). Si comprende però facilmente che questo diagramma, che risulta

costante a tratti, può essere ottenuto direttamente senza il tracciamento preliminare dei diagrammi degli addendi. Infatti, a pulsazioni minori di tutti i termini $1/|\tau_i|$, $1/|T_i|$, α_{ni} e ω_{ni} , gli unici fattori di $G(j\omega)$ che lo determinano sono μ e $1/(j\omega)^\xi$. I cambi di valore avvengono poi in corrispondenza degli altri poli e zeri del sistema.

Le regole per il tracciamento possono essere allora così riassunte:

- il tratto iniziale del diagramma ha ordinata $\arg \mu - g90^\circ$;
- in corrispondenza delle pulsazioni pari a $1/|\tau_i|$ e $1/|T_i|$ l'ordinata aumenta ($\tau_i > 0$ e $T_i < 0$) o diminuisce ($\tau_i < 0$ e $T_i > 0$) di 90° ;
- in corrispondenza delle pulsazioni naturali α_{ni} o ω_{ni} l'ordinata aumenta ($\xi_i \geq 0$ e $\xi_i < 0$) o diminuisce ($\xi_i < 0$ e $\xi_i \geq 0$) di 180° .

Il diagramma esatto si può infine ottenere sommando i diagrammi esatti dei singoli fattori di $G(s)$.

Esempio 6.5 (Seguito dell'Esempio 6.4). I diagrammi asintotico ed esatto della fase della (6.30) sono presentati nella Figura 6.14. Poiché la funzione ha guadago 100 e uno zero nell'origine, il diagramma asintotico a bassa pulsazione ha ordinata $+90^\circ$. Quindi alla pulsazione $\omega = 0.1$, corrispondente alla pulsazione naturale della coppia di poli complessi coniugati, l'ordinata diminuisce di 180° e assume il valore -90° . In $\omega = 10$, per la presenza dello zero positivo, l'ordinata diminuisce ancora di 90° , e infine in $\omega = 100$, per la presenza del polo, l'ordinata diminuisce di ulteriori 90° .

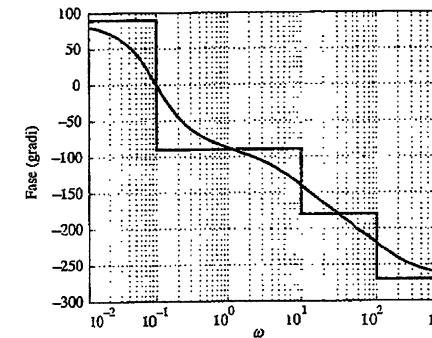


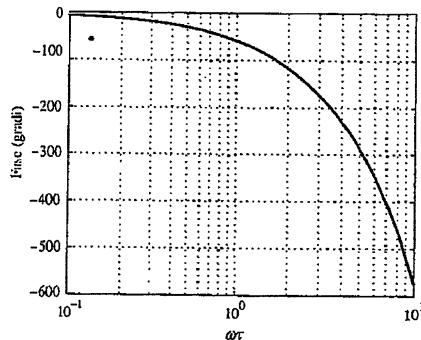
Figura 6.14
Diagramma di Bode della fase della funzione (6.30).

6.6.3 Ritardo di tempo

È ovviamente possibile tracciare i diagrammi di Bode della risposta in frequenza associata alla funzione di trasferimento del ritardo $G(s) = e^{-\tau s}$, $\tau > 0$. In particolare, risulta $|G(j\omega)| = 1$, $\forall \omega$, e quindi $|G(j\omega)|_{dB} = 0$, $\forall \omega$, mentre la fase espressa in gradi è $\arg G(j\omega) = -\omega\tau 180/\pi$, per cui, per esempio, $\arg G(j\tau) = -180/\pi \simeq -57^\circ$. Il diagramma del modulo è quindi una retta orizzontale di ordinata nulla, mentre il diagramma della fase ha l'andamento riportato nella Figura 6.15 in funzione della pulsazione normalizzata $\omega\tau$. Si osservi che in questo caso la fase ha un andamento lineare in ω e quindi l'uso di una scala logaritmica genera una curva con andamento esponenziale.

Per quanto detto, se si ha una funzione di trasferimento del tipo $G(s) = e^{-\tau s} G'(s)$, dove $G'(s)$ è razionale, il diagramma di Bode del modulo della risposta in frequenza è identico a quello associato a $G'(s)$, mentre quello della fase si ottiene sommando il contributo $-\omega\tau 180/\pi$ a quello della fase associato a $G'(s)$.

Figura 6.15
Diagramma di Bode della fase del ritardo di tempo.



6.6.4 Sistemi a sfasamento minimo

In generale, il diagramma di Bode della fase di una risposta in frequenza $G(j\omega)$ non si può dedurre dalla sola conoscenza del diagramma di $|G(j\omega)|_{dB}$. Infatti, anche ipotizzando di conoscere il segno del guadagno della funzione, nel diagramma del modulo non sono contenute le informazioni relative al segno delle parti reali delle singolarità di $G(s)$, che a loro volta determinano i segni dei vari contributi alla fase. Un caso di particolare interesse si verifica qualora sia noto che il guadagno è positivo, che tutti i poli e gli zeri di $G(s)$ hanno parte reale negativa o nulla e che $G(s)$ non contiene ritardi di tempo. In questo caso $G(s)$ viene detta a *sfasamento minimo*, o a *fase minima*, ed è immediato ricavare il diagramma asintotico della fase da quello asintotico del modulo. Infatti, a ogni cambio di pendenza in diminuzione di quest'ultimo corrisponde la presenza di un contributo angolare di ritardo dovuto a poli, mentre ogni cambio di pendenza in aumento comporta la presenza di un contributo angolare di anticipo dovuto a zeri. Ciò tra l'altro implica che quando il diagramma asintotico del modulo ha pendenza k , il diagramma asintotico della fase assume il valore $k90^\circ$.

In realtà, per i sistemi a sfasamento minimo è addirittura possibile ricavare il diagramma esatto della fase a partire da quello esatto del modulo, o viceversa. Le formule che legano queste due funzioni, qui non riportate perché non saranno utilizzate in seguito, vengono chiamate *formule di Bode*.

Sfasatore puro Una generica funzione di trasferimento $G(s)$ può sempre essere vista come il prodotto di una funzione di trasferimento a fase minima $G_{fm}(s)$, con modulo della risposta in frequenza identico a quello di $G(s)$, e di una funzione di trasferimento $G_{fnm}(s)$ la cui risposta in frequenza ha modulo unitario a tutte le pulsazioni. Il sistema $G_{fnm}(s)$ è detto *sfasatore puro*. Infatti, si può sempre fattorizzare $G(s)$ come

$$G(s) = \frac{N^-(s)N^+(s)}{D^-(s)D^+(s)} \quad (6.34)$$

dove $N^-(s)$ e $D^-(s)$ sono polinomi in s , primi tra loro, le cui radici hanno parte reale negativa o nulla e $N^+(s)$ e $D^+(s)$ sono polinomi in s , primi tra loro, le cui radici hanno parte reale positiva e tali che $N^+(0) = D^+(0)$. Evidentemente la (6.34) può essere scritta come

$$G(s) = G_{fm}(s)G_{fnm}(s) \quad (6.35)$$

dove

$$G_{fm}(s) = \frac{N^-(s)N^+(-s)}{D^-(s)D^+(-s)}$$

è a sfasamento minimo e tale che $|G(j\omega)| = |G_{fm}(j\omega)|$, $\forall \omega$, mentre la parte a fase non minima

$$G_{fnm}(s) = \frac{N^+(s)D^+(-s)}{N^+(-s)D^+(s)}$$

è tale che $|G_{fnm}(j\omega)| = 1$, $\forall \omega$. Nel caso in cui la funzione $G(s)$ contenga un ritardo di tempo, questo dovrà essere ovviamente incluso nel termine $G_{fnm}(s)$.

6.7 Diagrammi polari

Un'altra rappresentazione della risposta in frequenza è quella che si ottiene tracciando nel piano complesso il relativo *diagramma polare*, cioè il luogo dei punti $G(j\omega)$ con $\omega \geq 0$. Il diagramma polare può quindi essere interpretato come l'immagine attraverso $G(s)$ del semiasse immaginario non negativo ed è costituito da una curva giacente nel piano complesso e punteggiata in ω . Su questa curva è prassi porre delle frecce che indicano il verso di percorrenza al crescere della pulsazione da zero all'infinito.

Il tracciamento preciso del diagramma polare richiede lo studio di una funzione complessa, oppure la valutazione puntuale del numero complesso $G(j\omega)$ per un numero sufficientemente elevato di valori di ω . Tale determinazione è certamente laboriosa se compiuta manualmente, mentre non presenta difficoltà se si dispone di una macchina di calcolo, anche di modesta potenzialità. Tuttavia, l'andamento qualitativo del diagramma polare può essere determinato nella maggior parte dei casi in modo semplice anche senza l'ausilio di mezzi di calcolo. Il resto di questo paragrafo è dedicato all'illustrazione, condotta essenzialmente mediante esempi, dei metodi che si possono seguire per ottenere in modo qualitativo il diagramma cercato. In particolare si mostrerà come, data una funzione $G(s)$, il diagramma polare della risposta in frequenza associata possa essere ricavato o per mezzo di un'opportuna costruzione grafica nel piano complesso o direttamente dai diagrammi di Bode relativi a $G(j\omega)$. Questi ultimi non sono altro che una diversa rappresentazione di $G(j\omega)$ e pertanto hanno un contenuto informativo equivalente a quello del diagramma polare.

Per determinare l'andamento del diagramma polare di $G(j\omega)$ con il primo dei due metodi, si scriva la funzione di trasferimento nella forma fattorizzata

$$G(s) = \frac{\rho \prod_{i=1}^q (s + z_i) \prod_{i=1}^r (s^2 + 2\xi_i \alpha_{ni} s + \alpha_{ni}^2)}{s^g \prod_{i=1}^k (s + p_i) \prod_{i=1}^h (s^2 + 2\xi_i \omega_{ni} s + \omega_{ni}^2)}$$

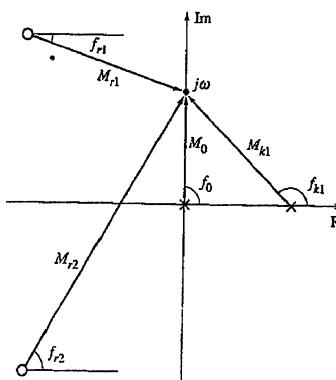
la cui risposta in frequenza è

$$G(j\omega) = \frac{\rho \prod_{i=1}^q (j\omega + z_i) \prod_{i=1}^r (-\omega^2 + 2j\xi_i \alpha_{ni} \omega + \alpha_{ni}^2)}{(j\omega)^g \prod_{i=1}^k (j\omega + p_i) \prod_{i=1}^h (-\omega^2 + 2j\xi_i \omega_{ni} \omega + \omega_{ni}^2)}$$

È allora immediato verificare che risulta

$$G(j\omega) = M e^{j\phi} \quad (6.36)$$

Figura 6.16
Valutazione di $G(j\omega)$ per via grafica.



dove

$$M = \frac{|\rho| \prod_{i=1}^q M_{qi} \prod_{i=1}^{2r} M_{ri}}{M_o^s \prod_{i=1}^k M_{ki} \prod_{i=1}^{2h} M_{hi}}$$

e

$$f = f_\rho + \sum_{i=1}^q f_{qi} + \sum_{t=1}^{2r} f_{ti} - g f_o - \sum_{i=1}^k f_{ki} - \sum_{i=1}^{2h} f_{hi}$$

In queste formule M_{qi} , M_{ri} , M_o , M_{ki} , M_{hi} sono i moduli e f_{qi} , f_{ri} , f_o , f_{ki} , f_{hi} sono le fasi dei vettori spiccati dalle corrispondenti singolarità di $G(s)$ e con il secondo estremo in $j\omega$ (Figura 6.16), mentre f_ρ è l'argomento della costante di trasferimento ρ . Analizzando al crescere di ω da zero all'infinito il modulo e la fase dei singoli vettori si può quindi determinare l'andamento di $G(j\omega)$.

Nei casi più complicati la procedura descritta può risultare difficoltosa e inoltre è critica la valutazione dell'andamento del diagramma polare quando $G(s)$ ha poli sull'asse immaginario. Per questi motivi, nel seguito, i diagrammi polari di alcune funzioni di trasferimento significative verranno direttamente ricavati dai corrispondenti diagrammi di Bode, mostrando soltanto in un caso come sia possibile pervenire ai medesimi risultati con il metodo appena illustrato e lasciando al lettore questa verifica negli altri esempi trattati.

Esempio 6.6 Si consideri la funzione di trasferimento

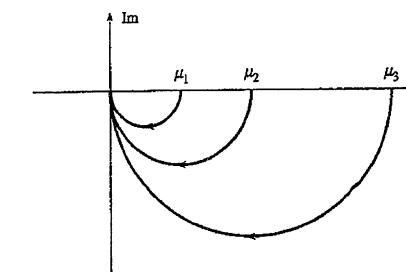
$$G(s) = \frac{\mu}{1+Ts} \quad (6.37)$$

in cui $\mu > 0$ e $T > 0$. Il diagramma di Bode del modulo di $G(s)/\mu$ è riportato nella Figura 6.6, mentre il corrispondente diagramma della fase è mostrato nella Figura 6.12. Da questi diagrammi, è evidente che all'aumentare di ω da zero all'infinito, il modulo e la fase della risposta in frequenza associata decrescono monotonicamente da μ a 0 e da 0° a -90° rispettivamente. Si supponga ora di aumentare il valore di μ mantenendo costante T : per ogni ω la fase di $G(j\omega)$ rimane inalterata, mentre il modulo aumenta proporzionalmente. Il diagramma polare per tre diversi valori del guadagno μ_1 , $\mu_2 = 2\mu_1$ e $\mu_3 = 4\mu_1$ è riportato nella Figura 6.17. Si noti che al variare di T cambia la punteggiatura delle curve in ω .

Volendo giungere alle medesime conclusioni per mezzo della costruzione grafica precedentemente introdotta, si riscriva la (6.37) nella forma

$$G(s) = \frac{\rho}{s+p} \quad (6.38)$$

Figura 6.17
Diagramma polare associato alla (6.37) ($\mu > 0$, $T > 0$) per $\mu = \mu_1$, $\mu = \mu_2 = 2\mu_1$ e $\mu = \mu_3 = 4\mu_1$.



dove $\rho = \mu/T$ e $p = 1/T$. È quindi necessario valutare il modulo M_{ki} e l'angolo f_{ki} del vettore rappresentato nella Figura 6.18 al variare del punto corrente sul semiasse immaginario non negativo. In particolare risulta

$$|G(j\omega)| = \frac{\rho}{M_{ki}} \quad , \quad \arg G(j\omega) = -f_{ki}$$

Dalla costruzione grafica di Figura 6.18 è immediato verificare che per $\omega = 0$ si ha $M_{ki} = p$ e $f_{ki} = 0$, quindi $|G(j0)| = \rho/p = \mu$ e $\arg G(j0) = 0^\circ$; questo punto potrebbe essere dedotto direttamente dalla (6.37) o dalla (6.38) valutando $G(0)$. All'aumentare di ω , sia il modulo M_{ki} sia l'argomento f_{ki} crescono monotonicamente, così che $|G(j\omega)|$ e $\arg G(j\omega)$ sono decrescenti. Per $\omega \rightarrow \infty$ si ha $M_{ki} \rightarrow \infty$ e $f_{ki} \rightarrow 90^\circ$; corrispondentemente il modulo di $G(j\omega)$ tende a zero e la fase a -90° , in coerenza con il diagramma precedentemente ricavato e riportato nella Figura 6.17.

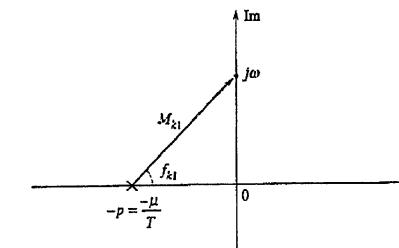


Figura 6.18
Costruzione grafica per il tracciamento del diagramma polare associato alla (6.37) con $\mu > 0$ e $T > 0$.

Sempre con riferimento alla (6.37), nel caso in cui $\mu < 0$ e $T > 0$, il modulo di $G(j\omega)$ è identico a quello dell'esempio precedente, mentre, in accordo con la convenzione introdotta nella (6.32), lo sfasamento iniziale è di -180° . All'aumentare di ω la fase diminuisce da -180° a -270° . Il corrispondente diagramma polare è riportato nella Figura 6.19. Se invece si ha $\mu < 0$ e $T < 0$, a pari modulo la fase passa da -180° per $\omega = 0$ a -90° per $\omega \rightarrow \infty$ (si veda ancora la Figura 6.19). Se infine risulta $\mu > 0$ e $T < 0$, sempre a pari modulo, la fase passa da 0° per $\omega = 0$ a 90° per $\omega \rightarrow \infty$. Anche questo diagramma è riportato nella Figura 6.19.

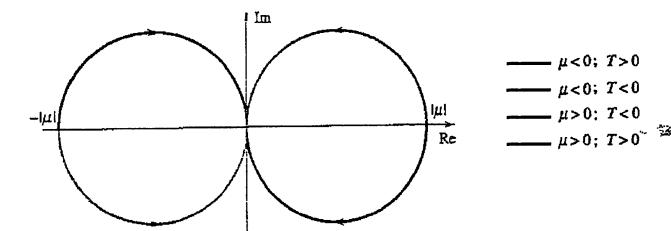


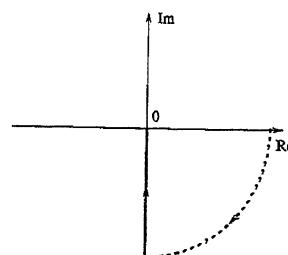
Figura 6.19
Diagramma polare associato alla (6.37) per diverse combinazioni di segno di μ e T .

Esempio 6.7 Il diagramma polare della risposta in frequenza dell'integratore

$$G(s) = \frac{1}{s} \quad (6.39)$$

coincide con il semiasse immaginario negativo (Figura 6.20). Infatti, come si deduce immediatamente dai corrispondenti diagrammi di Bode (si vedano le (6.24) e (6.33)), il modulo decresce all'aumentare di ω , mentre la fase si mantiene costante e pari a -90° .

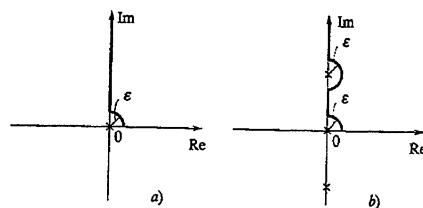
Figura 6.20
Diagramma polare
associato alla (6.39).



Il medesimo risultato potrebbe anche essere dedotto per continuità dall'analisi del diagramma polare associato alla (6.37) con $\mu > 0$ e $T > 0$. A questo riguardo, si consideri ancora la forma (6.38) e si supponga $\mu = T$, cioè $\rho = 1$. Al crescere di T il guadagno μ della funzione aumenta e il polo si sposta sempre più vicino all'origine del piano complesso. Per $T \rightarrow \infty$ la (6.38) tende alla funzione di trasferimento dell'integratore (6.39) ed è facile verificare che i diagrammi di Bode delle Figure 6.6 (con $|G(j\omega)/\mu|_{dB}$ sull'asse delle ordinate) e 6.12 tendono a coincidere con quelli dell'integratore (si veda la Figura 6.5 e si osservi che per $T \rightarrow \infty$ la fase riportata nella Figura 6.12 tende a -90° per $\omega \rightarrow 0$). Pertanto il diagramma di Figura 6.20 si può anche ricavare come limite per $\mu = T \rightarrow \infty$ del diagramma associato alla (6.37) e riportato nella Figura 6.17. In questa interpretazione, per $\omega = 0$ il diagramma compie un quarto di circonferenza di raggio infinito e fase che passa da 0° a -90° , mentre per $\omega > 0$ e crescente, la fase è costante a -90° e il modulo decresce.

Nell'esempio precedente si è mostrato come il problema del tracciamento del diagramma polare quando $s = 0$ è un polo di $G(s)$ possa essere risolto analizzando il diagramma di una funzione con polo in $s = -p$, $p > 0$, e studiando il caso limite $p \rightarrow 0$. In alternativa, volendo procedere mediante la costruzione grafica precedentemente illustrata, è possibile pervenire ai medesimi risultati pur di definire il diagramma polare associato alla funzione $G(s) = 1/s$ come l'immagine attraverso $G(s)$ dei punti del semiasse immaginario positivo $s = j\omega$, per $\omega > 0$, e dei punti $\varepsilon e^{j\theta}$, $\theta \in [0, \pi/2]$, con ε infinitesimo (si veda la Figura 6.21a). In quest'ottica è come se il polo tendendo all'origine "deformasse" l'asse immaginario nell'intorno dell'origine stessa. Seguendo la procedura grafica di costruzione e considerando il luogo dei punti definito nella Figura 6.21a, il diagramma polare che si determina per $\omega = 0$ compie un quarto di circonferenza di raggio infinito e fase che passa da 0° a -90° , in analogia con l'interpretazione fornita nell'esempio.

Figura 6.21
Luogo dei punti modificato
per il tracciamento del
diagramma polare.



Coerentemente con quanto detto e per ragioni che saranno meglio chiarite nel seguito, è opportuno ridefinire il diagramma polare, quando $G(s)$ possiede poli sull'asse immaginario, come l'immagine attraverso $G(s)$ dei punti del semiasse immaginario positivo (esclusi i poli di $G(s)$), dei punti $\varepsilon e^{j\theta}$, $\theta \in [0, \pi/2]$ e ε infinitesimo, se $s = 0$ è un polo di $G(s)$, e dei punti $j\bar{\omega} + \varepsilon e^{j\theta}$, $\theta \in (-\pi/2, \pi/2)$ e ε infinitesimo, se $j\bar{\omega}$ è un polo di $G(s)$. In questo modo, mediante la modifica del luogo dei punti di cui si traccia l'immagine per mezzo di semicirconferenze (o di un quarto di circonferenza) di raggio infinitesimo e poste nel semipiano destro, come mostrato nella Figura 6.21b, il diagramma polare associato a $G(s)$ risulta comunque connesso.

Esempio 6.8 Il diagramma polare associato alla funzione di trasferimento

$$G(s) = \frac{\mu}{1 + 2\xi s/\omega_n + s^2/\omega_n^2} \quad (6.40)$$

con $\xi > 0$ e $\mu > 0$, può essere determinato direttamente dal diagramma del modulo riportato nella Figura 6.8, in cui sull'asse delle ordinate è da leggersi $|G(j\omega)/\mu|_{dB}$, e da quello della fase di Figura 6.13. Per diversi valori di ξ , fissati μ e ω_n , il diagramma polare della (6.40) è riportato nella Figura 6.22. Come già discusso riguardo ai diagrammi di Bode corrispondenti, tutte le curve di Figura 6.22 attraversano il semiasse immaginario negativo per $\omega = \omega_n$; inoltre il modulo non è monotono decrescente per $\xi < 1/\sqrt{2}$.

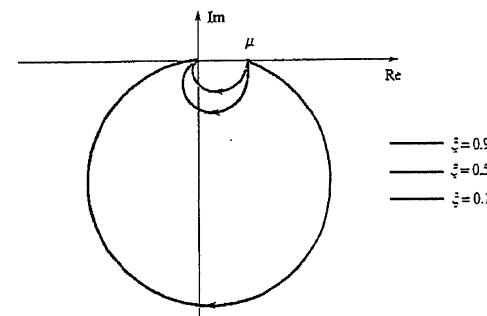


Figura 6.22
Diagramma polare
associato alla (6.40), $\mu > 0$,
per diversi valori di ξ .

Dalle Figure 6.8, 6.13 e 6.22 si può anche dedurre che per $\xi = 0$, cioè quando i poli complessi coniugati sono sull'asse immaginario, il modulo cresce con fase nulla per $\omega \rightarrow \omega_n$, per $\omega = \omega_n$ il modulo è infinito e la fase passa da 0° a -180° . In corrispondenza, il diagramma compie una semicirconferenza di raggio infinito nel semipiano immaginario negativo (si veda in particolare la Figura 6.13), mentre per $\omega \rightarrow \infty$ la fase è costante e pari a -180° e il modulo tende a zero. Il diagramma polare è riportato nella Figura 6.23. Si lascia al lettore il compito di determinare tale diagramma per via grafica considerando l'immagine attraverso $G(s)$ del semiasse immaginario positivo modificato come nella Figura 6.21.

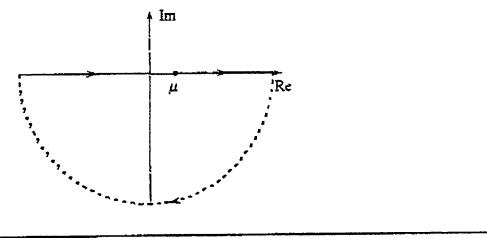


Figura 6.23
Diagramma polare
associato alla (6.40), $\mu > 0$,
per $\xi = 0$.

Esempio 6.9 Il diagramma polare associato al ritardo di tempo

$$G(s) = e^{-\tau s}$$

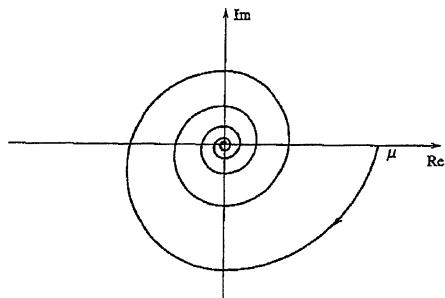
è costituito da una circonferenza di raggio unitario percorsa un numero infinito di volte in senso orario a partire dal semiasse reale positivo. Infatti, il modulo è costante e pari a uno per ogni ω , mentre la fase è decrescente con ω .

Esempio 6.10 L'andamento qualitativo del diagramma polare della funzione

$$G(s) = \frac{\mu}{1+Ts} e^{-\tau s} \quad \mu > 0, T > 0 \quad (6.41)$$

è riportato nella Figura 6.24 e si ottiene osservando che il modulo della (6.41) è identico a quello della (6.37), mentre lo sfasamento differisce per il contributo $-\omega\tau/2\pi$ dovuto al ritardo. Pertanto, al crescere di ω il diagramma compie un numero sempre maggiore di giri in senso orario attorno all'origine del piano complesso. Questo comportamento è tipico del diagramma polare delle funzioni di trasferimento con il termine di ritardo. Infatti, data la generica funzione di trasferimento $G(s) = G'(s)e^{-\tau s}$, con $G'(s)$ razionale, per ω sufficientemente elevata il contributo di fase di $G'(s)$ è comunque costante e la fase decresce per la presenza del ritardo.

Figura 6.24
Diagramma polare
associato alla (6.41).



6.8 Azione filtrante dei sistemi dinamici

I risultati del Paragrafo 6.3 mostrano come un sistema dinamico lineare e stazionario agisce sul segnale di ingresso, scomposto nelle sue armoniche, per produrre il segnale d'uscita. In particolare, si è già sottolineato come il sistema non può generare armoniche in uscita assenti nello spettro del segnale di ingresso, ma può unicamente amplificare o attenuare e sfasare quelle presenti. In altri termini, il sistema si comporta nei confronti del segnale di ingresso come un *filtro* che ne modella lo spettro in accordo con la propria risposta in frequenza. Da questo punto di vista è utile classificare i sistemi dinamici in base alle caratteristiche della loro azione filtrante. Nel seguito del paragrafo saranno illustrate due importanti classi di filtri, facendo riferimento esclusivamente a sistemi asintoticamente stabili.

6.8.1 Filtri passa-basso

Particolare importanza rivestono quei sistemi che lasciano passare inalterate, o al più amplificate di un valore costante, unicamente le armoniche del segnale di ingresso con pulsazione inferiore o uguale a un dato valore $\bar{\omega}$, e che eliminano le armoniche con pulsazione maggiore di $\bar{\omega}$. Questi sistemi vengono chiamati *filtri*

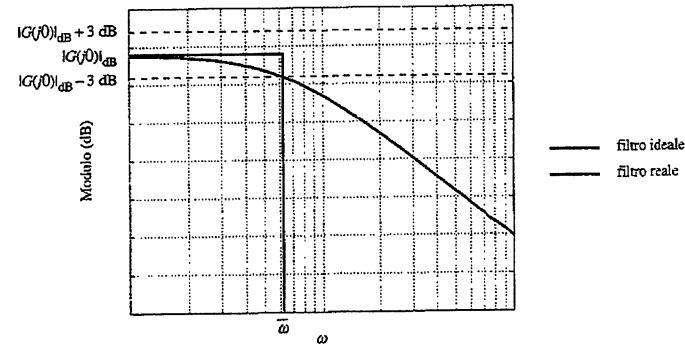


Figura 6.25
Diagramma di Bode del
modulo di un filtro ideale
e di un filtro reale
passa-basso.

ideali passa-basso; il diagramma di Bode del modulo a essi associato è costante sino a $\bar{\omega}$ e vale $-\infty$ dB per $\omega > \bar{\omega}$ (Figura 6.25), mentre il corrispondente diagramma della fase è nullo, almeno fino a $\bar{\omega}$.

La realizzazione di un filtro ideale passa-basso è di fatto impossibile e quindi è interessante definire i *filteri reali passa-basso* come quei sistemi caratterizzati da una risposta in frequenza $G(j\omega)$ con modulo circa costante a bassa frequenza e decrescente per $\omega > \bar{\omega}$. In particolare, si usa definire come filtro passa-basso un sistema la cui risposta in frequenza soddisfa le relazioni

$$\left\{ \begin{array}{l} \frac{1}{\sqrt{2}} \leq \frac{|G(j\omega)|}{|G(j0)|} \leq \sqrt{2}, \quad \omega \leq \bar{\omega} \\ \frac{|G(j\omega)|}{|G(j0)|} < \frac{1}{\sqrt{2}}, \quad \omega > \bar{\omega} \end{array} \right. \quad (6.42)$$

Come illustrato dalla Figura 6.25, il diagramma di Bode del modulo di un filtro reale passa-basso è quindi compreso nell'intervallo $[-3 \text{ dB}, 3 \text{ dB}]$ attorno al valore $|G(j0)|$ per $\omega \leq \bar{\omega}$, mentre assume valori inferiori a $|G(j0)|_{dB} - 3 \text{ dB}$ per $\omega > \bar{\omega}$. L'intervallo di pulsazioni $[0, \bar{\omega}]$ viene chiamato *banda passante* del filtro e $\bar{\omega}$ è l'estremo superiore della banda passante. La banda passante quindi rappresenta un indice della capacità del sistema di attenuare armoniche di pulsazione elevata e di far passare, con limitata variazione del modulo, le armoniche di pulsazione minore o uguale a $\bar{\omega}$.

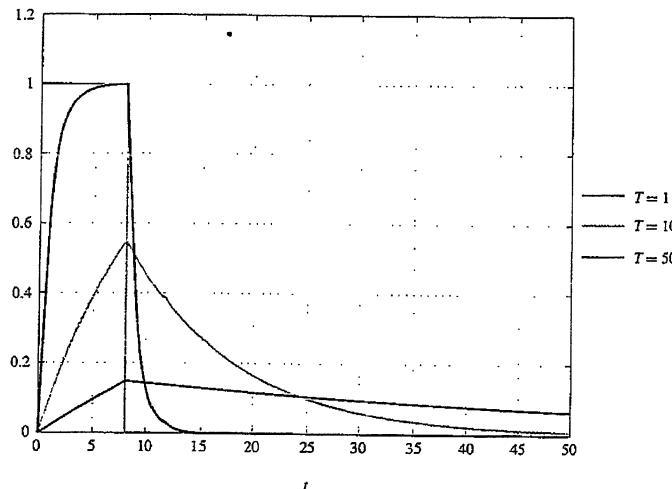
Si noti inoltre che nella definizione fornita tramite la (6.42) lo sfasamento introdotto da $G(s)$ non ha alcun ruolo, mentre in realtà il contributo di fase del filtro può essere significativo e di norma non va trascurato. Infine è ovvio osservare che la (6.42) implica che $G(s)$ sia di tipo zero, cioè non contenga azioni derivate o integrali. In tal caso $G(j0)$ coincide con il guadagno μ .

Esempio 6.11 Un sistema del primo ordine con funzione di trasferimento

$$G(s) = \frac{\mu}{1+Ts} \quad \mu > 0, T > 0 \quad (6.43)$$

può essere interpretato come un filtro passa-basso (si veda la Figura 6.6). Dalla figura è anche immediato verificare che la banda passante del filtro è $[0, 1/T]$ poiché il modulo decresce uniformemente e assume il valore $\mu/\sqrt{2}$ in corrispondenza della pulsazione del polo $\omega = 1/T$.

Figura 6.26
Filtraggio passa-basso di un impulso rettangolare attraverso il sistema (6.43), con $\mu = 1$ e diversi valori di T .



Nel Paragrafo 4.4.3 si è mostrato che la risposta allo scalino del sistema (6.43) giunge tanto più rapidamente al valore di regime quanto più è piccolo il valore della costante di tempo T , cioè quanto più è ampia la banda passante. Ciò è facilmente interpretabile alla luce dei risultati precedenti: all'allargamento della banda passante aumenta il numero di armoniche di ingresso che si ritrovano (quasi) inalterate in uscita. Il segnale di ingresso viene quindi riprodotto sempre più fedelmente. Al contrario, se T cresce, il sistema fa passare solo armoniche con bassa pulsazione e il segnale di uscita corrispondente è caratterizzato da variazioni lente. Queste considerazioni giustificano il fatto che all'ampiezza della banda passante sia usualmente associato il concetto di "velocità" del sistema.

A titolo di esempio, si confrontino nella Figura 6.26 le diverse risposte del sistema (6.43), con $\mu = 1$ e diversi valori di T , all'ingresso a forma di impulso rettangolare

$$u(t) = \begin{cases} 1 & , 0 \leq t < 8 \\ 0 & , \text{altrove} \end{cases} \quad (6.44)$$

Si era già visto nell'Esempio 6.3 che lo spettro di questo ingresso è prevalentemente concentrato "a bassa frequenza", ma contiene componenti armoniche significative fino a pulsazioni intorno a $\omega = 4$ (si veda la Figura 6.3). Quando $T = 1$, il sistema lascia passare la maggior parte delle componenti del segnale e la forma dell'uscita non differisce molto dalla forma dell'impulso rettangolare in ingresso. Se invece $T = 10$ o addirittura $T = 50$, il filtro ha una banda molto più stretta e le distorsioni e l'attenuazione che subisce il segnale sono molto più accentuate.

Filtri di Butterworth Nel campo dell'elaborazione dei segnali, sono numerose le applicazioni che richiedono l'uso di filtri passa-basso per depurare un segnale dalle componenti ad alta frequenza. Un tipico esempio riguarda i cosiddetti filtri anti-aliasing, di cui si parlerà nel Capitolo 17, dedicato all'analisi dei sistemi di controllo digitale. Rimandando il lettore a testi specialistici per maggiori dettagli sui criteri di progetto di tali filtri, ci si limita qui a parlare dei *filtri di Butterworth*, con funzione di trasferimento

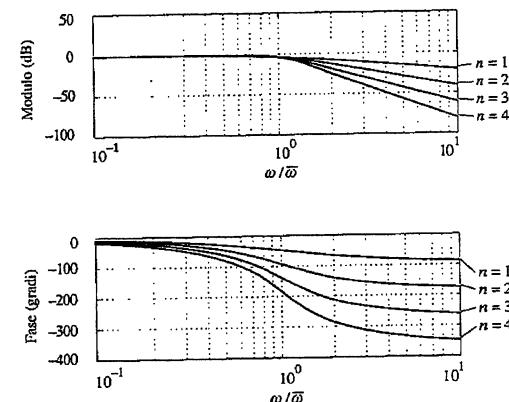
$$G(s) = \frac{1}{B_n(s)}$$

Tabella 6.1
Polinomi di Butterworth.

Ordine n	$B_n(s')$
1	$(s' + 1)$
2	$(s'^2 + 1.414s' + 1)$
3	$(s'^2 + s' + 1)(s' + 1)$
4	$(s'^2 + 0.765s' + 1)(s'^2 + 1.848s' + 1)$

dove il polinomio $B_n(s)$ è univocamente definito una volta fissato l'ordine n del filtro e la banda passante $[0, \bar{\omega}]$ desiderata. In particolare, definendo la variabile complessa normalizzata $s' = s/\bar{\omega}$, i polinomi $B_n(s)$, per $n = 1, 2, 3, 4$, si ricavano dalla Tabella 6.1. Come si può verificare, il filtro di Butterworth di ordine n è caratterizzato da una collocazione equispaziata degli n poli lungo la semicirconferenza di raggio $\bar{\omega}$ centrata nell'origine e appartenente al semipiano sinistro.

Figura 6.27
Diagrammi di Bode dei filtri di Butterworth di ordine 1, 2, 3, 4.

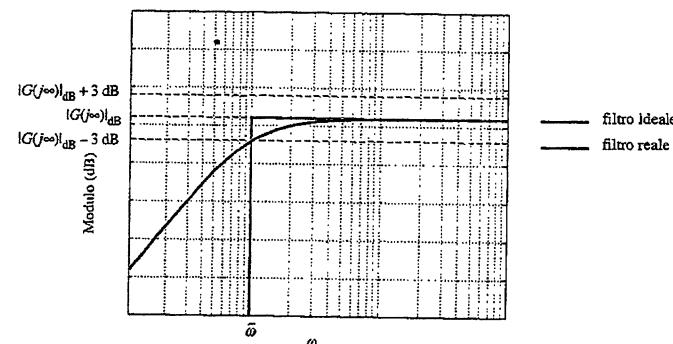


I corrispondenti diagrammi di Bode di modulo e fase sono riportati, nelle pulsazioni normalizzate $\omega/\bar{\omega}$, nella Figura 6.27. Si osservi che l'azione filtrante è tanto più vicina a quella di un filtro ideale, almeno per quanto riguarda il modulo, quanto più è elevato n . Per approfondimenti sulle proprietà dei filtri di Butterworth e sullo studio di altri filtri passa-basso, si rimanda ai testi specializzati.

6.8.2 Filtri passa-alto

Analogamente a quanto fatto per i filtri passa-basso è possibile definire i *filtri ideal passa-alto* come quei sistemi che lasciano passare inalterate, o al più amplificate di una quantità costante, unicamente le armoniche del segnale di ingresso con pulsazione maggiore o uguale a $\bar{\omega}$, e che eliminano le armoniche con pulsazione minore di $\bar{\omega}$. Il loro diagramma di Bode del modulo è costante per $\omega \geq \bar{\omega}$ e

Figura 6.28
Diagramma di Bode del modulo di un filtro ideale e di un filtro reale passa-alto.



vale $-\infty$ dB per $\omega < \bar{\omega}$ (Figura 6.28), mentre il corrispondente diagramma della fase è nullo, almeno per $\omega \geq \bar{\omega}$.

Si definiscono poi *filtri reali passa-alto* quei sistemi caratterizzati da una risposta in frequenza $G(j\omega)$ il cui modulo soddisfa le seguenti relazioni

$$\begin{cases} \frac{|G(j\omega)|}{|G(j\infty)|} < \frac{1}{\sqrt{2}}, & \omega < \bar{\omega} \\ \frac{1}{\sqrt{2}} \leq \frac{|G(j\omega)|}{|G(j\infty)|} \leq \sqrt{2}, & \omega \geq \bar{\omega} \end{cases} \quad (6.45)$$

Come mostrato nella Figura 6.28, il diagramma di Bode del modulo di un filtro reale passa-alto è quindi compreso nell'intervallo $[-3 \text{ dB}, 3 \text{ dB}]$ attorno al valore $|G(j\infty)|_{\text{dB}}$ per $\omega \geq \bar{\omega}$, mentre assume valori inferiori a $|G(j\infty)|_{\text{dB}} - 3 \text{ dB}$ per $\omega < \bar{\omega}$. L'insieme di pulsazioni $[\bar{\omega}, \infty)$ viene chiamato *banda passante* del filtro e $\bar{\omega}$ è l'estremo inferiore della banda passante.

La banda passante quindi rappresenta un indice della capacità del sistema di attenuare le armoniche di pulsazione inferiore a $\bar{\omega}$ e di far passare, con limitata variazione del modulo, le armoniche di pulsazione contenuta nella banda passante.

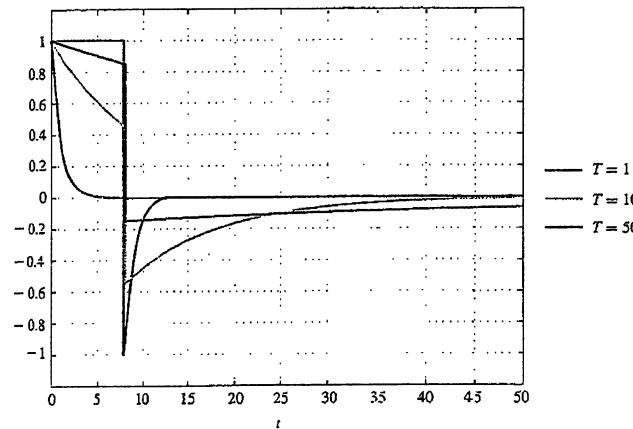
Si noti che solo sistemi non strettamente propri possono avere le caratteristiche di un filtro passa-alto, dato che la (6.45) implica necessariamente $G(j\infty) \neq 0$.

Esempio 6.12 Tracciando il diagramma di Bode del modulo associato alla funzione di trasferimento

$$G(s) = \frac{\mu s}{1 + Ts}, \quad T > 0 \quad (6.46)$$

si nota che tale sistema può essere considerato un filtro reale passa-alto con banda passante $[1/T, \infty)$.

Se ne può valutare empiricamente il comportamento calcolando l'uscita corrispondente all'ingresso a impulso rettangolare (6.44), già utilizzato nell'esempio precedente. Nella Figura 6.29 sono riportate le uscite del sistema (6.46) con tre diversi valori di T e ponendo $\mu = T$. Solo le componenti "veloci" del segnale in ingresso passano inalterate in uscita e, al diminuire di T , cioè via via



che la banda passante si concentra sempre più ad alta frequenza, l'uscita tende a risentire solo delle discontinuità del segnale di ingresso agli istanti $t = 0$ e $t = 8$.

6.9 Approssimazione a poli dominanti

Quando si è parlato nel Paragrafo 4.4.6 del concetto di poli dominanti, si è spiegato che l'approssimazione a poli dominanti di una data $G(s)$ di ordine elevato mira a conservare la dinamica "lenta" del sistema, trascurando quella "veloce". Ci si aspetta quindi che la risposta in frequenza dell'approssimazione a poli dominanti non si discosti molto da quella del sistema originario per bassi valori della pulsazione ω . Come si usa dire, l'approssimazione a poli dominanti rappresenta un modello "in bassa frequenza" del sistema. Ciò sarà evidente dai diagrammi di Bode mostrati negli esempi che seguono. D'altra parte si ricordi che, nel tracciamento dei diagrammi asintotici, poli e zeri con costanti di tempo "piccole" o con pulsazioni naturali "elevate" danno contributi trascurabili all'andamento dei diagrammi per bassi valori della pulsazione.

Esempio 6.13 (Seguito dell'Esempio 4.9). Con riferimento al modello del motore elettrico già considerato, si è visto come la funzione di trasferimento

$$G(s) = \frac{3.98}{(1 + 0.0856s)(1 + 0.0022s)} \quad (6.47)$$

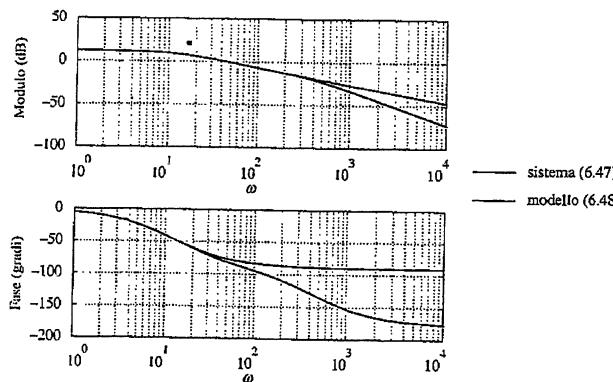
che descrive il legame tra la tensione d'armatura e la velocità di rotazione possa essere ragionevolmente approssimata dalla funzione

$$G_a(s) = \frac{3.98}{1 + 0.0856s} \quad (6.48)$$

in cui viene conservato il polo dominante. Dal confronto tra i diagrammi di Bode riportati nella Figura 6.30 si nota in effetti che la risposta in frequenza associata alla (6.48) costituisce una buona approssimazione della risposta in frequenza "vera" solo per bassi valori di ω .

Figura 6.29
Filtraggio passa-alto di un impulso rettangolare attraverso il sistema (6.46), con $\mu = T$ e diversi valori di T .

Figura 6.30
Diagrammi di Bode del sistema (6.47) e del modello (6.48).



Esempio 6.14 (Seguito dell'Esempio 4.10). Nei diagrammi di Bode di Figura 6.31 viene confrontata la risposta in frequenza associata alla funzione di trasferimento

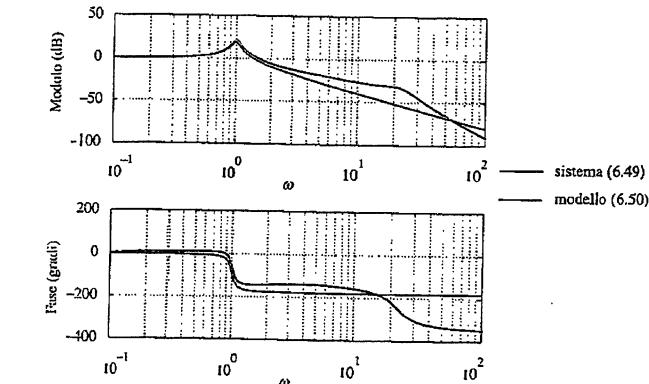
$$G(s) = \frac{1 + 0.5s}{(1 + 0.1s)(1 + 0.02s + 0.002s^2)(1 + 0.1s + s^2)} \quad (6.49)$$

con quella dell'approssimante a poli dominanti

$$G_a(s) = \frac{1}{1 + 0.1s + s^2} \quad (6.50)$$

Anche in questo caso si osserva, come previsto, che le maggiori discrepanze si verificano "in alta frequenza".

Figura 6.31
Diagrammi di Bode del sistema (6.49) e del modello (6.50).



6.10 Conclusioni

In questo capitolo è stata introdotta l'importante nozione di risposta in frequenza per i sistemi dinamici lineari e stazionari. Questa funzione può essere ricavata dalla funzione di trasferimento valutandola nei punti del semiasse immaginario positivo e caratterizza il comportamento del sistema quando questo è sollecitato

da ingressi sinusoidali o esprimibili come combinazione lineare di componenti sinusoidali. Sotto opportune ipotesi, si è visto come il modulo e la fase della risposta in frequenza rappresentino rispettivamente l'amplificazione e lo sfasamento che subiscono le diverse componenti armoniche del segnale di ingresso nel loro passaggio attraverso il sistema per la formazione del segnale di uscita. Si è così in grado di classificare i sistemi in base alla loro azione filtrante, individuando in particolare le importanti classi dei filtri passa-basso e passa-alto. Gran parte del capitolo è stata dedicata a due importanti rappresentazioni grafiche della risposta in frequenza (i diagrammi di Bode e i diagrammi polari), presentando in dettaglio i metodi per il loro tracciamento, anche solo qualitativo. Nel seguito questi diagrammi verranno largamente impiegati nello studio e nella sintesi dei sistemi retroazionati.

Esercizi

Esercizio 6.1 Calcolare mediante lo sviluppo di Heaviside la risposta del sistema con funzione di trasferimento

$$G(s) = \frac{20}{1 + 4s}$$

all'ingresso $u(t) = 8 \cos(5t)$, a partire da condizioni iniziali nulle. Verificare poi che tale risposta ha un andamento asintotico coerente con quello previsto dal teorema fondamentale della risposta in frequenza.

Esercizio 6.2 Dimostrare il Teorema 6.4.

Esercizio 6.3 Si consideri il sistema descritto dalla funzione di trasferimento

$$G(s) = \frac{60}{s^2 + 0.3s + 9}$$

Supponendo che l'ingresso sia $u(t) = \sin(\omega t)$, valutare l'andamento asintotico dell'uscita nei tre casi $\omega = 0.3$, $\omega = 3$ e $\omega = 30$, interpretando i risultati in termini del fenomeno della risonanza.

Esercizio 6.4 Si calcolino modulo e fase della risposta in frequenza associata alla funzione di trasferimento

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

e si disegnino i relativi grafici in funzione della pulsazione ω , prima in scala lineare e poi in scala semilogaritmica (diagrammi di Bode).

Esercizio 6.5 Con riferimento al modello del motore elettrico già considerato negli Esempi 2.6 e 3.8, si analizzi l'effetto di perturbazioni di tipo esponenziale della coppia resistente nell'intorno della condizione di equilibrio ricavata nell'Esempio 3.8 (si ponga cioè $\delta C_r = e^{j\omega t}$). In particolare si verifichi che quando $\lambda = -R/L$ la variazione della velocità angolare può rimanere nulla, anche in presenza della perturbazione sulla coppia resistente, purché il valore iniziale della corrente sia pari a $i(0) = (h\bar{v} + k\bar{C}_r) / (k^2 + Rh) + 1/k$.

Esercizio 6.6 Attraverso la ricerca del punto di massimo della funzione (6.26), verificare la correttezza delle formule (6.27) e (6.28) per il calcolo della pulsazione e del picco di risonanza per un sistema del secondo ordine descritto dalla funzione di trasferimento (6.23).

Esercizio 6.7 Si verifichi che il sistema con funzione di trasferimento

$$G(s) = \frac{(1-2s)(1-0.8s+s^2)}{(1+2s)(1+0.8s+s^2)} e^{-5s}$$

è uno sfasatore puro, nel senso che il diagramma di Bode del modulo è costante al variare della pulsazione.

Esercizio 6.8 Una maniera alternativa per rappresentare graficamente una risposta in frequenza $G(j\omega)$ consiste nel tracciare il grafico del modulo $|G(j\omega)|_{dB}$ in funzione della fase $\arg G(j\omega)$. La curva risultante, parametrizzata in ω , viene chiamata *diagramma di Nichols*. Come esempio, si tracci il diagramma di Nichols associato alla funzione $G_b(s) = 1/s$ e quello associato alla funzione $G_c(s) = 1/(1+Ts)$.

Esercizio 6.9 Tracciare il diagramma polare della risposta in frequenza associata a

$$G(s) = \frac{10}{(1 + 0.25s^2)(1 + 0.1s)}$$

e verificare che esso è costituito da due rami che si adagiano su un asintoto obliqua e da una semicirconferenza di raggio infinito che li collega.

Esercizio 6.10 Verificare che il sistema descritto da

$$G(s) = \frac{s^2 + 1}{s^2 + 1.414s + 1}$$

può essere usato per depurare un segnale in ingresso dalla componente armonica a pulsazione $\omega = 1$, lasciando praticamente inalterate le altre componenti dello spettro. Un filtro di questo genere viene di solito chiamato *notch filter* o *filtro a spillo*.

Sistemi dinamici a tempo discreto

7

7.1 Introduzione

Questo capitolo tratta una nuova classe di sistemi dinamici, caratterizzati dal fatto che la variabile tempo, da cui dipendono tutte le variabili che intervengono nella descrizione del sistema stesso, è intera invece che reale. Essi, per distinguerli da quelli a tempo continuo, sono detti sistemi a tempo discreto, o discreti nel tempo.

È bene precisare che nell'ingegneria più tradizionale non è molto comune trovare esempi di processi che si lascino descrivere naturalmente con sistemi a tempo discreto. Modelli di questo genere sono invece spesso utilizzati nella descrizione di fenomeni in ambito economico, gestionale, ambientale o sociologico, per affrontare problemi sia di controllo sia di previsione. Inoltre, nell'ingegneria classica, questi modelli hanno larga applicazione nel settore dell'elaborazione e trasmissione dei segnali e nello studio dei problemi di controllo digitale, nei quali un processo descritto da un sistema a tempo continuo è controllato da un'apparecchiatura digitale. Infatti i sistemi a tempo discreto costituiscono buoni modelli per il calcolatore-controllore e per il sistema dinamico sotto controllo, corredato della strumentazione necessaria per la connessione con la macchina digitale. Si vedano a questo proposito i Capitoli 17 e 18.

In questo capitolo, e nel successivo, ci si propone di introdurre le principali tecniche di studio disponibili per i sistemi a tempo discreto, le cui proprietà sono molto simili a quelle discusse nei Capitoli 2-6 per i sistemi a tempo continuo. Per questa ragione la trattazione sarà piuttosto sintetica.

In dettaglio, si presentano i temi seguenti:

- gli elementi costitutivi dei sistemi dinamici a tempo discreto;
- i criteri di classificazione basati sulle caratteristiche delle equazioni che li descrivono;
- la particolare modalità di funzionamento detta equilibrio;
- la fondamentale proprietà di stabilità.

Un'attenzione particolare viene dedicata ai sistemi lineari e stazionari, spesso utilizzati per descrivere in modo approssimato anche il comportamento di sistemi non lineari e varianti nel tempo. In proposito, si trattano gli argomenti sottoelencati:

- le espressioni esplicative dei movimenti dello stato e dell'uscita, con l'introduzione del concetto di modo e del fondamentale principio di sovrapposizione degli effetti;

- le espressioni esplicite degli stati e delle uscite di equilibrio;
- la formulazione di semplici condizioni necessarie e sufficienti di stabilità asintotica;
- l'applicazione dei risultati allo studio approssimato del movimento e all'analisi della stabilità dell'equilibrio di sistemi non lineari, effettuati mediante la procedura della linearizzazione;
- due interessanti proprietà strutturali, che conducono alla definizione di una scomposizione in parti dei sistemi dinamici.

7.2 Generalità

7.2.1 Variabili di ingresso, stato e uscita, e rappresentazione di stato

Un sistema dinamico a tempo discreto costituisce un modello matematico di un oggetto che interagisce con il mondo circostante tramite due vettori, u e y , di variabili dipendenti dal tempo, assunto intero e indicato con il simbolo k . In particolare, $u \in R^m$ è il vettore delle *variabili di ingresso*, che rappresentano le azioni che vengono compiute sull'oggetto in esame da agenti esterni che ne influenzano il comportamento; $y \in R^p$ è invece il vettore delle *variabili di uscita*, che rappresentano quanto del comportamento dell'oggetto in esame è, per qualche ragione, di interesse.

Occorre poi introdurre anche il vettore $x \in R^n$ delle *variabili di stato*, che descrivono la situazione interna dell'oggetto da modellare quanto basta per permettere il calcolo delle variabili di uscita, noti che siano il valore di quelle di ingresso e l'istante di tempo. Allora, un *sistema dinamico a tempo discreto di ordine n* è costituito dalla coppia di equazioni

$$x(k+1) = f(x(k), u(k), k) \quad (7.1)$$

$$y(k) = g(x(k), u(k), k) \quad (7.2)$$

dette rispettivamente *equazione di stato* e *trasformazione d'uscita*, dove f e g sono due opportune funzioni vettoriali. Esso è anche detto semplicemente *sistema* quando non possano sussistere ambiguità.

Il legame tra l'ingresso e l'uscita risulta quindi scisso in due parti e descritto mediante due equazioni vettoriali:

- un'equazione alle differenze (l'*equazione di stato*), che mette in relazione con l'ingresso le variabili che descrivono la situazione interna del sistema;
- un'equazione algebrica (la *trasformazione d'uscita*), che consente di determinare l'uscita a uno specifico istante di tempo sulla base della conoscenza di tale situazione e dell'ingresso allo stesso istante di tempo.

L'Equazione (7.1) definisce l'evoluzione dello stato $x(k)$ per $k > k_0$, in corrispondenza di ogni terna costituita da un istante iniziale k_0 , una funzione di ingresso $u(k)$, $k \geq k_0$, e una condizione iniziale $x(k_0) = x_{k_0}$. Infatti, si trova facilmente

$$\begin{aligned} x(k_0 + 1) &= f(x(k_0), u(k_0), k_0) = f(x_{k_0}, u(k_0), k_0) \\ x(k_0 + 2) &= f(x(k_0 + 1), u(k_0 + 1), k_0 + 1) = \\ &= f(f(x_{k_0}, u(k_0), k_0), u(k_0 + 1), k_0 + 1) \end{aligned}$$

e così via. La funzione $x(k)$, $k \geq k_0$, si dice *movimento dello stato* del sistema. L'Equazione (7.2) permette invece di determinare l'evoluzione dell'uscita $y(k)$

per $k \geq k_0$, in corrispondenza della funzione di ingresso $u(k)$ e dell'andamento dello stato $x(k)$ per $k \geq k_0$. La funzione $y(k)$, $k \geq k_0$, va sotto il nome di *movimento dell'uscita*.

Le Equazioni (7.1), (7.2) costituiscono una *rappresentazione di stato*, o *ingresso-stato-uscita*, o *interna* di un sistema dinamico a tempo discreto.

Le considerazioni fatte al Paragrafo 2.2.4, circa la scelta della natura e del numero delle variabili di stato, si applicano ugualmente ai sistemi che si stanno trattando qui. Anche in questo caso esse possono essere spesso selezionate come quelle quantità che determinano gli accumuli all'interno dei sistemi, rendendo così conto della storia passata. La scelta, comunque, non è mai unica e a volte, per una maggiore comodità nell'analisi, può essere opportuno definire lo stato in campo complesso invece che reale.

7.2.2 Esempi

Si mostrano, qui di seguito, alcuni esempi di sistemi dinamici a tempo discreto.

Esempio 7.1 Si voglia descrivere l'evoluzione temporale, mese per mese, del guadagno g di un'azienda che vende un certo bene che essa stessa produce in quantità $q \geq 0$. Se k indica il generico mese e si denota con $s \geq 0$ la scorta all'inizio del mese, un semplice bilancio consente di scrivere

$$s(k+1) = s(k) + q(k) - v(k) \quad (7.3)$$

dove $v \geq 0$ rappresenta la quantità venduta nel mese. Si può assumere che quest'ultima dipenda dal prezzo $p \geq 0$ secondo la relazione

$$v(k) = \alpha(k) - \beta(k)p(k) \quad (7.4)$$

dove α e β sono due parametri non negativi che descrivono l'atteggiamento dei consumatori rispetto al prezzo; le vendite cioè contengono una quota indipendente dal prezzo e un'altra che diminuisce all'aumentare del prezzo; a prezzo costante, le due quote sono funzioni del periodo considerato (si può pensare che α e β dipendano dalla congiuntura economica). L'azienda, operando in regime di monopolio, può fissare ogni mese il prezzo del bene secondo le proprie esigenze; essa lo innalzerà quanto più scarsa è la scorta, rispetto a una quantità $s^0 > 0$ oggetto di scelta, e quanto più riesce a vendere. Detti γ e δ due parametri positivi dipendenti dal tempo, si avrà allora

$$p(k+1) = p(k) + \gamma(k)[s^0(k) - s(k)] + \delta(k)v(k) \quad (7.5)$$

Se si trascura il costo di mantenimento della scorta, il guadagno mensile, differenza tra ricavi e spese di produzione, è

$$g(k) = p(k)v(k) - \zeta(k)q(k) \quad (7.6)$$

dove anche il parametro ζ , che rappresenta il costo unitario di produzione, è positivo. Se ora si pone

$$u_1(k) = q(k) \quad , \quad u_2(k) = \alpha(k) \quad , \quad u_3(k) = s^0(k)$$

$$x_1(k) = s(k) \quad , \quad x_2(k) = p(k) \quad , \quad y(k) = g(k)$$

eliminando v dalle Equazioni (7.3)-(7.6) si ottiene il sistema

$$x_1(k+1) = x_1(k) + \beta(k)x_2(k) + u_1(k) - u_2(k) \quad (7.7)$$

$$\begin{aligned} x_2(k+1) &= -\gamma(k)x_1(k) - [1 - \delta(k)\beta(k)]x_2(k) + \\ &+ \delta(k)u_2(k) - \gamma(k)u_3(k) \end{aligned} \quad (7.8)$$

$$y(k) = -\beta(k)x_2^2(k) - \zeta(k)u_1(k) + x_2(k)u_3(k) \quad (7.9)$$

Se invece come variabile di uscita si prende il prezzo p , la trasformazione d'uscita diviene semplicemente

$$y(k) = x_2(k) \quad (7.10)$$

In ogni caso, $n = 2$.

Esempio 7.2 Per studiare l'evoluzione del numero di prede $x_1 \geq 0$ e di predatori $x_2 \geq 0$ che convivono in uno stesso ecosistema al variare delle prede immesse $u \geq 0$, si può fare riferimento a un modello proposto da A.J. Lotka e V. Volterra negli anni '20 del secolo scorso. Denotando con k il periodo temporale di riferimento, una versione a tempo discreto di tale modello si può scrivere come

$$\begin{aligned}x_1(k+1) &= x_1(k) + \alpha(1 - \beta x_1(k))x_1(k) - \gamma x_1(k)x_2(k) + u(k) \\x_2(k+1) &= x_2(k) - \delta x_2(k) + \zeta x_1(k)x_2(k)\end{aligned}$$

dove tutti i parametri sono positivi. La prima equazione afferma che, in assenza di predatori ($x_2(k) = 0$), il numero di prede al tempo $k+1$ è pari a quello nel periodo precedente aumentato della quantità di prede immesse e di una quantità dipendente in modo non lineare dal numero delle prede presenti. Quest'ultima tiene conto della natalità attraverso il termine $\alpha x_1(k)$ e anche della limitazione delle risorse nutritive a disposizione tramite il termine $\alpha \beta x_1(k)^2$. Inoltre, la presenza dei predatori ($x_2(k) > 0$) fa sì che le prede diminuiscano in maniera proporzionale sia al numero di predatori stessi sia a quello di prede. La seconda equazione descrive invece l'evoluzione del numero di predatori. In assenza di prede ($x_1(k) = 0$) i predatori tendono naturalmente a diminuire per mancanza di nutrimento, mentre alla presenza di prede ($x_1(k) > 0$) corrisponde un contributo che tende a far aumentare i predatori, perché questi hanno a disposizione più sostanze nutritive. Assumendo che la variabile di interesse sia costituita dal numero di predatori, si potrà

$$y(k) = x_2(k)$$

ottenendo così un sistema di ordine $n = 2$.

Esempio 7.3 Molti algoritmi matematici si possono descrivere come sistemi dinamici a tempo discreto. Per esempio, si consideri il seguente classico algoritmo di stima. Si supponga che due variabili $\alpha(k)$ e $\beta(k)$ siano tra loro proporzionali secondo un coefficiente μ , cioè

$$\beta(k) = \mu\alpha(k)$$

e si assuma di poter fissare ad arbitrio il valore di α . Per determinare μ , che è costante, ma ignoto, si può pensare di impostare un singolo valore non nullo ad α , misurare il corrispondente valore di β ed eseguire il rapporto dei due numeri. Una situazione più realistica si ha però quando si assume che nel misurare β si commetta un errore ε , differente da una prova all'altra; cioè

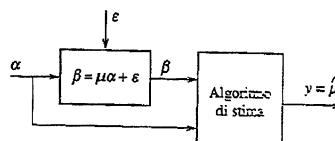
$$\beta(k) = \mu\alpha(k) + \varepsilon(k) \quad (7.11)$$

In tal caso è ragionevole basare la stima di μ su una serie di misure effettuate con diversi valori di α , aggiornando ricorsivamente la stima dopo ogni prova. Un semplice algoritmo utilizzabile per la stima ricorsiva di μ può essere descritto dalle relazioni

$$\hat{\mu}(k+1) = \hat{\mu}(k) + \eta\alpha(k)(\beta(k) - \hat{\mu}(k)\alpha(k)) \quad (7.12)$$

dove $\hat{\mu}(k)$ rappresenta la stima $\hat{\mu}$ di μ disponibile al tempo k e il parametro η è positivo. La quantità $\beta(k) - \hat{\mu}(k)\alpha(k)$ rappresenta l'errore che si commette al passo k nell'approssimare μ con $\hat{\mu}(k)$ e la logica dell'algoritmo consiste nell'aggiornare a ogni passo la stima di μ in dipendenza di tale quantità: per esempio, se l'errore è positivo con α positivo, μ è sottostimato per cui conviene aumentare $\hat{\mu}$. L'insieme delle relazioni (7.11), (7.12) può essere interpretato come un sistema dinamico in cui α ed ε costituiscono le variabili di ingresso e $\hat{\mu}$ l'uscita (Figura 7.1). Precisamente, ponendo

$$u_1(k) = \alpha(k), \quad u_2(k) = \varepsilon(k), \quad y(k) = x(k) = \hat{\mu}(k)$$



si ottiene il sistema, di ordine $n = 1$,

$$\begin{aligned}x(k+1) &= x(k) \left(1 - \eta u_1^2(k)\right) + \mu \eta u_1^2(k) - \eta u_1(k) u_2(k) \\y(k) &= x(k)\end{aligned}$$

Figura 7.1
Stima ricorsiva
di un parametro.

7.2.3 Classificazione

I sistemi dinamici descritti dalle Equazioni (7.1), (7.2) possono essere classificati in vari modi sulla base delle proprietà delle funzioni f e g .

Sistemi monovariabili e multivariabili (SISO e MIMO) Si dice *monovariabile* (o *SISO*) un sistema dotato di una sola variabile di ingresso e di una sola variabile di uscita, cioè un sistema per cui $m = p = 1$: in caso contrario esso si dice *multivariabile* (o *MIMO*).

I sistemi degli Esempi 7.1 e 7.3 sono multivariabili, perché hanno rispettivamente tre e due ingressi; invece quello dell'Esempio 7.2 è monovariabile, perché ha un ingresso e un'uscita.

Sistemi propri, strettamente propri e non dinamici Se la funzione g non dipende dall'ingresso, cioè se la trasformazione d'uscita (7.2) si può scrivere nella forma

$$y(k) = g(x(k), k)$$

allora il sistema si dice *strettamente proprio*, o anche *puramente dinamico*, perché l'uscita dipende dall'ingresso non direttamente, ma solo attraverso lo stato. Viceversa, in generale il sistema si dice *proprio*.

I sistemi degli Esempi 7.1-7.3 sono strettamente propri, tranne quello descritto dalle Equazioni (7.7)-(7.9), che è proprio.

Si osservi pure che un caso particolare di sistema proprio è quello di sistema *non dinamico, o statico*, per descrivere il cui comportamento non è necessario introdurre alcuna variabile di stato. Il legame ingresso-uscita è pertanto istantaneo e descritto dalla sola equazione

$$y(k) = g(u(k), k)$$

Sistemi invarianti e varianti nel tempo Si parla di sistema *invariante nel tempo*, o anche *stazionario*, nel caso in cui le funzioni f e g non dipendano esplicitamente dal tempo, cioè risultino

$$x(k+1) = f(x(k), u(k))$$

$$y(k) = g(x(k), u(k))$$

mentre, se anche una sola delle funzioni f e g dipende esplicitamente dal tempo, il sistema si dice *variante nel tempo*.

I sistemi degli Esempi 7.2 e 7.3 sono stazionari, mentre quelli dell'Esempio 7.1 sono varianti nel tempo, a meno che i parametri β , γ , δ e ζ siano indipendenti da k .

La risposta di un sistema stazionario a una qualunque sollecitazione, costituita da uno stato iniziale e da una funzione di ingresso, non dipende dall'istante di applicazione della sollecitazione stessa. In altre parole, la stessa sollecitazione inviata in due istanti diversi k_1 e k_2 al medesimo sistema produce due risposte che differiscono solo per una traslazione temporale pari alla differenza tra k_2 e k_1 . Pertanto, nello studio di questi sistemi, è possibile assumere qualunque valore del tempo come istante iniziale: la scelta più adottata nell'ambito della teoria dei sistemi e del controllo, e quindi anche in questo testo, è $k_0 = 0$.

Sistemi lineari e non lineari Quando le funzioni f e g sono lineari in x e u , cioè quando $x(k+1)$ e $y(k)$ sono combinazioni lineari delle varie componenti dei vettori $x(k)$ e $u(k)$, allora il sistema (7.1), (7.2) si può scrivere nella forma

$$x(k+1) = A(k)x(k) + B(k)u(k) \quad (7.13)$$

$$y(k) = C(k)x(k) + D(k)u(k) \quad (7.14)$$

dove le matrici $A \in R^{n \times n}$, $B \in R^{n \times m}$, $C \in R^{p \times n}$ e $D \in R^{p \times m}$ sono in generale funzioni del tempo. In questo caso il sistema dinamico si dice *lineare*; altrimenti si parla di sistema *non lineare*. La matrice A viene detta *matrice della dinamica*; i suoi elementi, come pure quelli delle matrici B e C , assumono valori complessi nei casi in cui lo stato sia definito in campo complesso.

Tra i sistemi degli Esempi 7.1-7.3, l'unico lineare è quello descritto dalle Equazioni (7.7), (7.8), (7.10) (a patto di trascurare i vincoli sul segno delle variabili di ingresso e di stato); esso può essere descritto dalle matrici

$$A(k) = \begin{bmatrix} 1 & \beta(k) \\ -\gamma(k) & 1 - \delta(k)\beta(k) \end{bmatrix}, \quad B(k) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & \delta(k) & \gamma(k) \end{bmatrix}$$

$$C(k) = [0 \ 1], \quad D(k) = [0 \ 0 \ 0]$$

Un caso di particolare interesse Un sistema che gode della proprietà di linearità unitamente a quella di invarianza nel tempo è descritto dalle equazioni

$$x(k+1) = Ax(k) + Bu(k) \quad (7.15)$$

$$y(k) = Cx(k) + Du(k) \quad (7.16)$$

dove A , B , C e D sono matrici costanti. La forma speciale delle Equazioni (7.15), (7.16) fa sì che lo studio dei sistemi lineari e invarianti nel tempo sia particolarmente semplice. Per questa ragione, tutte le volte che ciò sia possibile, si cerca di fornire al sistema le due proprietà suddette mediante un procedimento di linearizzazione e di "congelamento" di eventuali parametri lentamente variabili nel tempo. La linearizzazione sarà trattata al Paragrafo 7.6.1.

Sistemi a parametri concentrati e a parametri distribuiti I sistemi finora presi in esame hanno la caratteristica di essere, come si dice, *a dimensione finita*, o anche *a parametri concentrati*. Ciò significa che, per descrivere la loro situazione interna al tempo k , è sufficiente conoscere un numero finito n di scalari reali, cioè le n componenti del vettore di stato x . Esistono anche casi in cui invece lo stato del sistema al tempo k è costituito da un'intera funzione di una o più variabili. Si parla allora di sistemi *a dimensione infinita*, o *a parametri distribuiti*.

7.2.4 Equilibrio

Per i sistemi stazionari di Equazioni (7.15), (7.16) soggetti a ingressi costanti $u(k) = \bar{u}$, i movimenti dello stato e dell'uscita che risultano anch'essi costanti nel tempo sono detti, rispettivamente, *stati e uscite di equilibrio*. Gli stati di equilibrio devono evidentemente soddisfare l'equazione $x(k+1) = x(k)$, cioè sono le soluzioni \bar{x} costanti nel tempo, se ne esistono, dell'equazione

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

mentre a ognuna di queste soluzioni corrisponde un'uscita di equilibrio \bar{y} , calcolabile mediante la relazione

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

Esempio 7.4 (Seguito dell'Esempio 7.2) Gli stati di equilibrio del modello preda-predatore per $u(k) = \bar{u} = 0$ soddisfano le equazioni

$$\begin{aligned} \bar{x}_1 &= \bar{x}_1 + \alpha(1 - \beta\bar{x}_1)\bar{x}_1 - \gamma\bar{x}_1\bar{x}_2 \\ \bar{x}_2 &= \bar{x}_2 - \delta\bar{x}_2 + \zeta\bar{x}_1\bar{x}_2 \end{aligned}$$

cioè sono

$$\bar{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} 1/\beta \\ 0 \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} \delta/\zeta \\ \alpha(1 - \beta\delta/\zeta)/\gamma \end{bmatrix}$$

7.3 Stabilità

Le definizioni delle varie proprietà di stabilità relative ai sistemi a tempo discreto sono del tutto simili a quelle introdotte per i sistemi a tempo continuo nel Paragrafo 2.6. Analogamente a quanto fatto in quella sede, ci si riferisce, quindi, a sistemi stazionari con $k_0 = 0$, iniziando con il trattare la stabilità dell'equilibrio.

7.3.1 Stabilità dell'equilibrio

Per un sistema dinamico invariante nel tempo, si considerino un ingresso costante $u(k) = \bar{u}$, $k \geq 0$, e un corrispondente stato di equilibrio \bar{x} , detto *nominale*. Si consideri anche un movimento dello stato $x(k)$, detto *perturbato*, generato a partire ancora da \bar{u} , ma da uno stato iniziale x_0 (in generale diverso da \bar{x}). Si possono allora dare le seguenti definizioni.

Definizione 7.1 Uno stato di equilibrio \bar{x} si dice *stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \bar{x}\| \leq \delta$$

risulti

$$\|x(k) - \bar{x}\| \leq \varepsilon$$

per tutti $i k \geq 0$. ■

Definizione 7.2 Uno stato di equilibrio \bar{x} si dice *instabile* se non è stabile. ■

Definizione 7.3 Uno stato di equilibrio \bar{x} si dice *asintoticamente stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \bar{x}\| \leq \delta$$

risulti

$$\|x(k) - \bar{x}\| \leq \varepsilon$$

per tutti $i k \geq 0$, e inoltre

$$\lim_{k \rightarrow \infty} \|x(k) - \bar{x}\| = 0$$

La proprietà di stabilità dell'equilibrio richiede quindi che il movimento perturbato rimanga "vicino" all'equilibrio nominale. Più precisamente, scelta arbitrariamente piccola la massima distanza accettabile in un qualunque istante di tempo tra il movimento perturbato e l'equilibrio nominale, quest'ultimo è stabile se la condizione su tale distanza è rispettata, pur di prendere lo stato iniziale del movimento perturbato sufficientemente prossimo all'equilibrio nominale. Se ciò non accade l'equilibrio si dice instabile. L'instabilità di uno stato di equilibrio implica perciò che esistono perturbazioni arbitrariamente piccole dello stato iniziale che provocano l'allontanamento dello stato del sistema dall'equilibrio stesso. Infine, la proprietà di stabilità può essere rafforzata nella proprietà di stabilità asintotica, richiedendo anche che il movimento perturbato tenda all'equilibrio nominale per $k \rightarrow \infty$.

È possibile che uno stato di equilibrio asintoticamente stabile sia anche *globalmente stabile*; cioè può accadere che la Definizione 7.3 sia applicabile con qualunque valore di $\delta > 0$, se si preferisce, che i movimenti perturbati generati da un qualunque stato iniziale, "vicino" all'equilibrio nominale o "lontano" da esso, convergano tutti all'equilibrio nominale stesso. Questo però non è certo il caso generale. Infatti, il sistema potrebbe essere dotato di più stati di equilibrio con differenti proprietà di stabilità. Per ognuno di quelli asintoticamente stabili esisterà poi un insieme di stati iniziali, che si dicono costituire la *regione di attrazione* dello stato di equilibrio, che generano movimenti perturbati convergenti asintoticamente a esso.

L'applicazione diretta delle Definizioni 7.1-7.3 per accettare le proprietà di stabilità di un sistema è quasi sempre una via impraticabile, perché richiede il calcolo dei movimenti perturbati. La determinazione, o anche solo la stima, delle regioni di attrazione è poi ancora meno semplice. Fortunatamente, per classi importanti di sistemi dinamici esistono metodi generali di analisi che saranno introdotti nel seguito.

7.3.2 Stabilità del movimento

Le considerazioni sulla stabilità dell'equilibrio appena svolte, e in particolare la proprietà di stabilità asintotica, possono essere generalizzate sostituendo all'equilibrio nominale di riferimento un generico movimento nominale. In proposito, qui ci si limita a mostrare la nuova, più generale, forma presa dalle Definizioni 7.1-7.3.

Per un sistema dinamico invariante nel tempo si considerino un ingresso $\tilde{u}(k)$, $k \geq 0$, uno stato iniziale \tilde{x}_0 e il movimento dello stato $\tilde{x}(k)$, detto *nominale*, da essi prodotto. Si consideri anche un secondo movimento dello stato $x(k)$, detto *perturbato*, generato a partire ancora da $\tilde{u}(k)$, ma da un nuovo stato iniziale x_0 (in generale diverso da \tilde{x}_0). Si possono allora dare le seguenti definizioni.

Definizione 7.4 Un movimento $\tilde{x}(k)$ si dice *stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \tilde{x}_0\| \leq \delta$$

risulti

$$\|x(k) - \tilde{x}(k)\| \leq \varepsilon$$

per tutti $i \geq 0$. ■

Definizione 7.5 Un movimento $\tilde{x}(k)$ si dice *instabile* se non è stabile. ■

Definizione 7.6 Un movimento $\tilde{x}(k)$ si dice *asintoticamente stabile* se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti gli stati iniziali x_0 che soddisfano la relazione

$$\|x_0 - \tilde{x}_0\| \leq \delta$$

risulti

$$\|x(k) - \tilde{x}(k)\| \leq \varepsilon$$

per tutti $i \geq 0$, e inoltre

$$\lim_{k \rightarrow \infty} \|x(k) - \tilde{x}(k)\| = 0$$

■

7.4 Movimento ed equilibrio dei sistemi lineari e stazionari

Ci si soffermerà ora, per alcuni paragrafi, sui sistemi a tempo discreto lineari e invarianti nel tempo, per i quali molte delle nozioni introdotte in precedenza possono essere ulteriormente elaborate, così da renderle più specifiche e di facile utilizzazione.

Si consideri perciò il sistema dinamico descritto da

$$x(k+1) = Ax(k) + Bu(k) \quad (7.17)$$

$$y(k) = Cx(k) + Du(k) \quad (7.18)$$

dove $u \in R^m$, $x \in R^n$ e $y \in R^p$, mentre le matrici A , B , C e D sono reali, costanti e di dimensioni opportune.

7.4.1 Calcolo del movimento

È facile verificare per semplice sostituzione che il movimento dello stato corrispondente all'ingresso $u(k)$, definito per $k \geq k_0$, e allo stato iniziale $x(k_0) = x_{k_0}$ è dato da

$$x(k) = A^{k-k_0}x_{k_0} + \sum_{i=k_0}^{k-1} A^{k-i-1}Bu(i) \quad (7.19)$$

Il corrispondente movimento dell'uscita è

$$y(k) = CA^{k-k_0}x_{k_0} + C \sum_{i=k_0}^{k-1} A^{k-i-1}Bu(i) + Du(k) \quad (7.20)$$

Dalle Equazioni (7.19), (7.20) si possono trarre importanti indicazioni sulle caratteristiche dei movimenti dei sistemi corrispondenti.

7.4.2 Movimento libero e movimento forzato

Nei movimenti (7.19), (7.20) dello stato e dell'uscita del sistema (7.17), (7.18) si può individuare un contributo dipendente solo dallo stato iniziale e uno dipendente solo dall'ingresso, dai quali il movimento complessivo si ottiene per semplice somma.

Il contributo al movimento dello stato e dell'uscita funziona solo dello stato iniziale, cioè quello che si avrebbe se, a pari stato iniziale, l'ingresso fosse nullo, si chiama *movimento libero*, ed è dato da

$$x_l(k) = A^{k-k_0} x_{k_0} \quad (7.21)$$

$$y_l(k) = C A^{k-k_0} x_{k_0} \quad (7.22)$$

mentre il contributo funziona solo dell'ingresso, cioè quello che si avrebbe se, a pari ingresso, lo stato iniziale fosse nullo, si chiama *movimento forzato*, ed è dato da

$$x_f(k) = \sum_{i=k_0}^{k-1} A^{k-i-1} B u(i) \quad (7.23)$$

$$y_f(k) = C \sum_{i=k_0}^{k-1} A^{k-i-1} B u(i) + D u(k) \quad (7.24)$$

7.4.3 Principio di sovrapposizione degli effetti

Per il sistema (7.17), (7.18), si supponga che in corrispondenza dell'ingresso u' e dello stato iniziale x'_{k_0} si ottengano i movimenti dello stato e dell'uscita x' e y' , mentre in corrispondenza dell'ingresso u'' e dello stato iniziale x''_{k_0} si ottengano i movimenti dello stato e dell'uscita x'' e y'' . Allora, sono soddisfatte le relazioni

$$x'(k) = A^{k-k_0} x'_{k_0} + \sum_{i=k_0}^{k-1} A^{k-i-1} B u'(i) \quad , \quad k \geq k_0$$

$$y'(k) = C A^{k-k_0} x'_{k_0} + C \sum_{i=k_0}^{k-1} A^{k-i-1} B u'(i) + D u'(k) \quad , \quad k \geq k_0$$

$$x''(k) = A^{k-k_0} x''_{k_0} + \sum_{i=k_0}^{k-1} A^{k-i-1} B u''(i) \quad , \quad k \geq k_0$$

$$y''(k) = C A^{k-k_0} x''_{k_0} + C \sum_{i=k_0}^{k-1} A^{k-i-1} B u''(i) + D u''(k) \quad , \quad k \geq k_0$$

Si consideri ora una terza situazione in cui l'ingresso u''' e lo stato iniziale x'''_{k_0} siano ambedue costituiti dalla stessa combinazione lineare degli ingressi e degli stati iniziali già considerati, cioè si assuma che esistano due scalari qualunque α e β tali che

$$u'''(k) = \alpha u'(k) + \beta u''(k) \quad (7.25)$$

$$x'''_{k_0} = \alpha x'_{k_0} + \beta x''_{k_0} \quad (7.26)$$

Per sostituzione si può verificare che anche i movimenti dello stato e dell'uscita x''' e y''' prodotti dalla coppia (7.25), (7.26) sono dati dalla medesima combinazione: infatti i movimenti

$$x'''(k) = \alpha x'(k) + \beta x''(k) \quad , \quad k \geq k_0$$

e

$$y'''(k) = \alpha y'(k) + \beta y''(k) \quad , \quad k \geq k_0$$

soddisfano identicamente le Equazioni (7.19), (7.20). Quanto visto sopra costituisce la dimostrazione del seguente risultato.

Teorema 7.1 Considerato il sistema lineare (7.17), (7.18) con istante iniziale k_0 , siano x' e y' i movimenti dello stato e dell'uscita generati dall'ingresso u' e dallo stato iniziale x'_{k_0} , e x'' e y'' i movimenti dello stato e dell'uscita generati dall'ingresso u'' e dallo stato iniziale x''_{k_0} . Allora, per ogni coppia di scalari α e β , i movimenti dello stato x''' e dell'uscita y''' generati dall'ingresso

$$u'''(k) = \alpha u'(k) + \beta u''(k)$$

e dallo stato iniziale

$$x'''_{k_0} = \alpha x'_{k_0} + \beta x''_{k_0}$$

sono

$$x'''(k) = \alpha x'(k) + \beta x''(k)$$

$$y'''(k) = \alpha y'(k) + \beta y''(k) \quad \blacksquare$$

Pertanto, in particolare, il comportamento dei sistemi in oggetto per "grandi perturbazioni" dello stato iniziale e dell'ingresso differisce da quello relativo a "piccole perturbazioni" solo per un fattore di scala.

Il Teorema 7.1, introdotto qui per sistemi descritti dalle Equazioni (7.17), (7.18), va sotto il nome di *principio di sovrapposizione degli effetti* ed è di grande importanza in quanto consente di calcolare il movimento generato da più cause, cioè stati iniziali e ingressi, semplicemente come somma pesata dei singoli effetti provocati dalle cause suddette. In realtà, si può dimostrare che questo risultato vale indipendentemente dall'ipotesi di stazionarietà del sistema cui lo si applica, mentre è legato in maniera indissolubile all'ipotesi di linearità, tanto che la classe dei sistemi lineari potrebbe essere definita come costituita da tutti e soli quei sistemi per i quali è possibile parlare di effetto di una variabile (stato iniziale o ingresso) su un'altra (stato o uscita).

7.4.4 Rappresentazioni equivalenti

Si consideri ora una matrice costante $T \in \mathbb{R}^{n \times n}$ non singolare e, mediante un *cambio di variabili*, si definisca un nuovo vettore di stato \hat{x} come

$$\hat{x}(k) = T x(k) \quad (7.27)$$

Allora, per semplice sostituzione, si ottiene

$$\hat{x}(k+1) = \hat{A} \hat{x}(k) + \hat{B} u(k) \quad (7.28)$$

$$y(k) = \hat{C} \hat{x}(k) + \hat{D} u(k) \quad (7.29)$$

con

$$\hat{A} = T A T^{-1} \quad , \quad \hat{B} = T B \quad , \quad \hat{C} = C T^{-1} \quad , \quad \hat{D} = D \quad (7.30)$$

Questo sistema dinamico è *equivalente* a quello descritto dalle equazioni (7.17), (7.18), nel senso che, per un ingresso $u(k)$, $k \geq k_0$, e due stati iniziali x_{k_0} e \hat{x}_{k_0} legati dalla condizione $\hat{x}_{k_0} = T x_{k_0}$, i movimenti dello stato dei sistemi (7.17), (7.18) e (7.28), (7.29) sono effettivamente legati dalla relazione (7.27),

cioè $\hat{x}(k) = Tx(k)$, $k > k_0$, e i movimenti dell'uscita sono identici. Lo stesso risultato vale per le singole componenti libere e forzate. Pertanto le due quadruple di matrici (A, B, C, D) e $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ sono semplicemente due maniere differenti di descrivere una medesima realtà. Se è necessario, la matrice T e il vettore di stato \hat{x} possono anche essere definiti in campo complesso senza che ciò crei alcun particolare problema, a parte la sicura perdita di significato fisico da parte delle nuove variabili di stato. Si noti in particolare che le matrici A e \hat{A} sono simili e, di conseguenza, i loro autovalori sono identici. Come si mostrerà tra breve utilizzando i risultati dell'Appendice A, questi ultimi contribuiscono in maniera così notevole a determinare le caratteristiche dei movimenti dei sistemi (7.17), (7.18), o (7.28), (7.29), che vengono spesso chiamati addirittura *autovalori del sistema*.

7.4.5 Autovalori e modi

Le formule (7.21), (7.22) permettono di calcolare il movimento libero del sistema (7.17), (7.18). Nel caso in cui il suo ordine sia $n = 1$, e quindi $A = a$ è scalare, x_l e y_l dipendono dalla ben nota funzione a^k . Per studiare il movimento libero nel caso generale $n \neq 1$ bisogna invece analizzare in profondità la struttura della potenza di matrice A^k , utilizzando i risultati del Paragrafo A.4. Conviene trattare separatamente i casi in cui A è diagonalizzabile oppure no, riferendosi per semplicità al caso $k_0 = 0$.

Matrice della dinamica diagonalizzabile In molti casi, come per esempio tutte le volte che gli autovalori z_i , $i = 1, 2, \dots, n$, sono tra loro distinti, è possibile scegliere la matrice $T = T_D$ in modo che la matrice della dinamica \hat{A} assuma la forma diagonale \hat{A}_D , cioè risulti

$$\hat{A} = \hat{A}_D = \text{diag}\{z_1, z_2, \dots, z_n\}$$

Quando ciò accade il movimento libero dello stato del sistema (7.28), (7.29) risulta semplicemente

$$\hat{x}_l(k) = \hat{A}_D^k \hat{x}_0 = \text{diag}\{z_1^k, z_2^k, \dots, z_n^k\} \hat{x}_0$$

In corrispondenza, per la (7.27), i movimenti liberi dello stato e dell'uscita del sistema (7.17), (7.18) sono

$$x_l(k) = T_D^{-1} \hat{x}_l(k) = T_D^{-1} \text{diag}\{z_1^k, z_2^k, \dots, z_n^k\} T_D x_0 \quad (7.31)$$

$$y_l(k) = C T_D^{-1} \text{diag}\{z_1^k, z_2^k, \dots, z_n^k\} T_D x_0 \quad (7.32)$$

Essi pertanto sono combinazioni lineari, con coefficienti dipendenti dagli elementi di x_0 , T_D e C , di termini z_i^k , $i = 1, 2, \dots, n$, che sono detti *modi*. È molto importante notare che le coppie $z_i = \rho_i e^{j\theta_i}$, $\bar{z}_i = \rho_i e^{-j\theta_i}$ di autovalori complessi coniugati di A generano termini che, sommati, danno luogo in x_l e y_l a un unico termine reale del tipo $\rho_i^k \sin(\theta_i k + \phi_i)$ dove ϕ_i è una fase opportuna. Con un lieve abuso di linguaggio, anche questi termini saranno denominati modi.

Esempio 7.5 Si consideri il sistema (7.17), (7.18) con

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad C = [1 \ 1]$$

Gli autovalori sono $z_{1,2} = 1 \pm j = \sqrt{2}e^{\pm j\pi/4}$. Prendendo

$$T_D^{-1} = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}$$

si ha

$$\hat{A} = T_D A T_D^{-1} = \begin{bmatrix} 1+j & 0 \\ 0 & 1-j \end{bmatrix}$$

da cui, per la (7.31),

$$\begin{aligned} x_l(k) &= 0.5 \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \begin{bmatrix} (\sqrt{2}e^{j\pi/4})^k & 0 \\ 0 & (\sqrt{2}e^{-j\pi/4})^k \end{bmatrix} \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix} x_0 = \\ &= 0.5\sqrt{2^k} \begin{bmatrix} e^{jk\pi/4} + e^{-jk\pi/4} & -je^{jk\pi/4} + je^{-jk\pi/4} \\ je^{jk\pi/4} - je^{-jk\pi/4} & e^{jk\pi/4} + e^{-jk\pi/4} \end{bmatrix} x_0 = \\ &= \sqrt{2^k} \begin{bmatrix} \cos(k\pi/4) & \sin(k\pi/4) \\ -\sin(k\pi/4) & \cos(k\pi/4) \end{bmatrix} x_0 \end{aligned}$$

mentre, per la (7.32),

$$\begin{aligned} y_l(k) &= \sqrt{2^k} \begin{bmatrix} \cos(k\pi/4) & \sin(k\pi/4) \\ -\sin(k\pi/4) & \cos(k\pi/4) \end{bmatrix} x_0 = \\ &= \sqrt{2^{k+1}} [\cos((k+1)\pi/4) \quad \cos((k-1)\pi/4)] x_0 \end{aligned}$$

Le Figure 7.2-7.4 riportano un quadro sintetico dell'andamento temporale qualitativo dei modi in dipendenza della posizione nel piano complesso dei corrispondenti autovalori, rappresentati con crocette. Per chiarezza, i punti corrispondenti ai valori assunti da ogni modo, indicati con asterischi, sono collegati tra loro da una linea sottile. Si osservi in particolare che autovalori reali negativi producono modi di tipo oscillatorio, mentre un autovalore nell'origine del piano complesso genera un modo che è nullo per $k \geq 1$; comportamenti di questo tipo non hanno riscontro nei sistemi a tempo continuo, dove autovalori reali generano modi monotoni e l'eventuale annullamento di un modo è sempre asintotico.

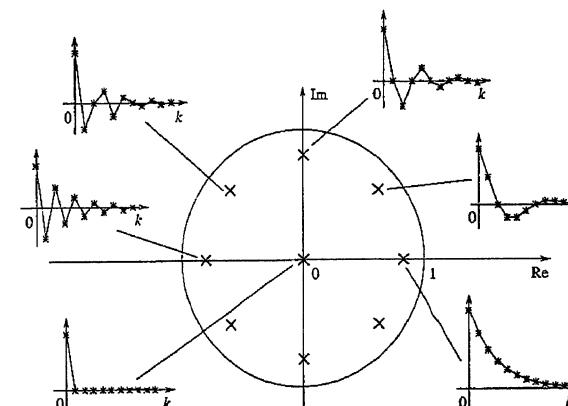


Figura 7.2
Modi dei sistemi con autovalori distinti e modulo minore di 1.

Figura 7.3
Modi dei sistemi con autovalori distinti e modulo uguale a 1.

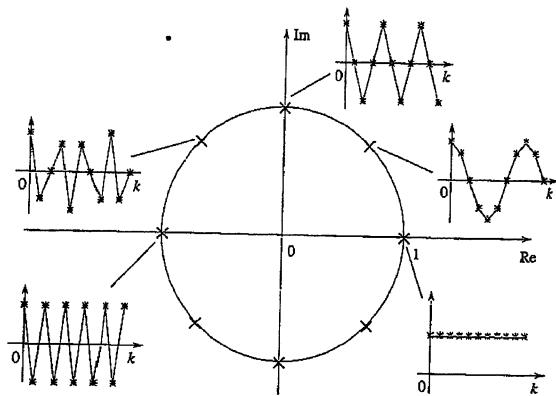
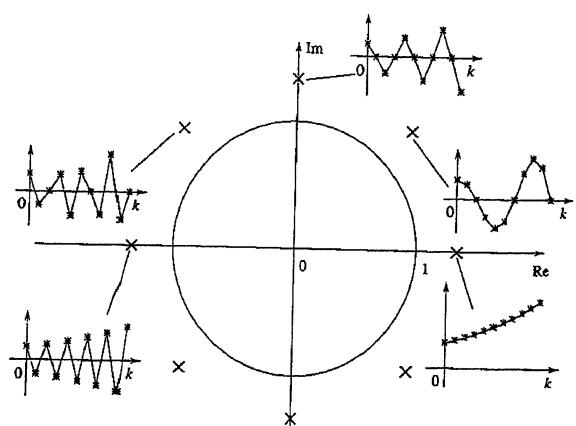


Figura 7.4
Modi dei sistemi con autovalori distinti e modulo maggiore di 1.



Matrice della dinamica non diagonalizzabile Quando la matrice A possiede autovalori multipli, cioè quando non tutti gli autovalori z_i , $i = 1, 2, \dots, n$, sono distinti tra loro, può darsi che risulti impossibile renderla diagonale mediante un'opportuna trasformazione. In questo caso però esiste comunque una matrice di trasformazione $T = T_J$ capace di trasformare A nella cosiddetta *forma di Jordan* \hat{A}_J . In base ai risultati del Paragrafo A.4, i modi che costituiscono i movimenti liberi dello stato e dell'uscita sono allora nulli per $k < \eta - 1$, mentre per $k \geq \eta - 1$, e assumendo convenzionalmente $0^0 = 1$, sono della forma

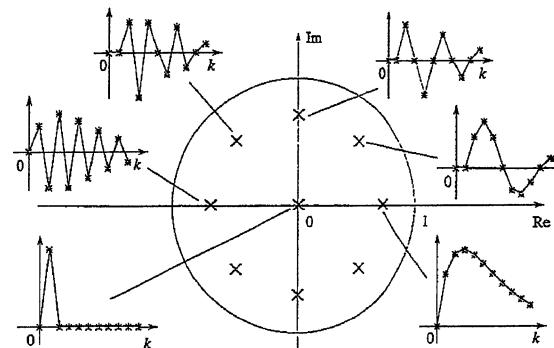
$$k^{\eta-1} z_i^{k-\eta+1} \quad (7.33)$$

se z_i è reale e

$$k^{\eta-1} \rho_i^{k-\eta+1} \sin(\theta_i(k-\eta+1) + \varphi_i) \quad (7.34)$$

se $z_i = \rho_i e^{j\theta_i}$ è complesso, dove η è un qualunque intero compreso tra uno e la massima dimensione $\hat{\eta}_i$ dei miniblocchi di Jordan associati a z_i , e φ_i una fase opportuna. Si osservi che i modi corrispondenti agli autovalori nell'origine del piano

Figura 7.5
Modi dei sistemi con autovalori doppi e modulo minore di 1.



complesso sono nulli almeno dall'istante di tempo $k = \hat{\eta}_i$ in avanti. La Figura 7.5 riporta l'andamento temporale qualitativo dei modi relativi ad autovalori doppi e modulo minore di 1 che si affiancano a quelli della Figura 7.2.

7.4.6 Risposta all'impulso e movimento forzato

Si supponga che il sistema (7.17), (7.18) abbia un solo ingresso e sia sollecitato con un impulso unitario applicato in $k_0 = 0$: si assuma cioè $m = 1$ e $u(k) = \text{imp}^*(k)$. Denotando con g_x e g_y i particolari movimenti forzati dello stato e dell'uscita che si ottengono dalle (7.23), (7.24), per la definizione della funzione impulso discreto del Paragrafo C.2, risulta

$$g_x(k) = \begin{cases} 0 & k=0 \\ A^{k-1}B & k>0 \end{cases} \quad (7.35)$$

$$g_y(k) = \begin{cases} D & k=0 \\ CA^{k-1}B & k>0 \end{cases} \quad (7.36)$$

Questi movimenti sono detti *risposta all'impulso dello stato e dell'uscita*. Si noti che, per $k > 0$, essi coincidono coi movimenti liberi prodotti dallo stato iniziale $x(0) = B$ ritardati di un passo. Le risposte all'impulso dei sistemi che hanno solo autovalori in $z = 0$ si annullano in un tempo finito, che è pari alla massima tra le dimensioni dei miniblocchi di Jordan e quindi non maggiore dell'ordine n , cioè $g_y(k) = 0$, $k \geq n$. In virtù di questa proprietà essi sono detti *sistemi FIR* (dall'inglese *Finite-Impulse-Response*).

L'estensione di queste definizioni al caso $m > 1$ è molto naturale. In generale le (7.35), (7.36) definiscono due matrici di funzioni di dimensioni $n \times m$ e $p \times m$ e il generico elemento di posizione (i, j) di g_x o di g_y , rappresenta la risposta della variabile di stato i -esima, o dell'uscita i -esima, a un impulso unitario inviato all'ingresso j -esimo (avendo posto nulle le altre variabili di ingresso).

Qualunque sia il valore di m , le risposte all'impulso rivestono un'importanza del tutto speciale in quanto la loro sola conoscenza è sufficiente per calcolare i movimenti forzati dello stato e dell'uscita generati da un ingresso qualunque (applicato in $k_0 = 0$). Ricordando ancora la definizione dell'impulso, è infatti immediato verificare che questi movimenti sono ottenibili come prodotto di con-

voluzione delle funzioni (7.35), (7.36) con la funzione di ingresso; visto che g_x , g_y e u sono nulle per tempi negativi, si ha cioè

$$\begin{aligned} g_x(k) * u(k) &= \sum_{i=-\infty}^{+\infty} g_x(k-i) u(i) = \sum_{i=0}^k g_x(k-i) u(i) = \\ &= \sum_{i=0}^{k-1} A^{k-i-1} B u(i) = x_f(k) \\ g_y(k) * u(k) &= \sum_{i=-\infty}^{+\infty} g_y(k-i) u(i) = \sum_{i=0}^k g_y(k-i) u(i) = \\ &= \sum_{i=0}^{k-1} C A^{k-i-1} B u(i) + D u(k) = y_f(k) \end{aligned} \quad (7.37)$$

D'altra parte, supponendo per semplicità che la matrice A sia diagonalizzabile, per le (7.27), (7.30) i movimenti forzati (7.23), (7.24) del sistema (7.17), (7.18) risultano

$$\begin{aligned} x_f(k) &= T_D^{-1} \hat{x}_f(k) = T_D^{-1} \sum_{i=0}^{k-1} \hat{A}^{k-i-1} \hat{B} u(i) = \\ &= T_D^{-1} \sum_{i=0}^{k-1} \text{diag}\{z_1^{k-i-1}, z_2^{k-i-1}, \dots, z_n^{k-i-1}\} T_D B u(i) \\ y_f(k) &= C T_D^{-1} \sum_{i=0}^{k-1} \text{diag}\{z_1^{k-i-1}, z_2^{k-i-1}, \dots, z_n^{k-i-1}\} T_D B u(i) + D u(k) \end{aligned}$$

mentre è ovvia la generalizzazione al caso in cui A non sia diagonalizzabile.

In particolare, la risposta $h_y(k)$ dell'uscita di un sistema SISO allo scalino si può calcolare ponendo $u(k) = \text{sca}^*(k)$ nella (7.37). Ricordando la definizione dello scalino discreto del Paragrafo C.2 si ha

$$h_y(k) = \sum_{i=0}^{k-1} C A^{k-i-1} B + D = \sum_{i=0}^k g_y(k-i) = \sum_{i=0}^k g_y(i) \quad (7.38)$$

È interessante notare che, viceversa, dall'Equazione (7.38) si può ottenere la risposta all'impulso in funzione della risposta allo scalino. Infatti, ponendo $h_y(-1) = 0$ per convenzione, risulta

$$g_y(k) = h_y(k) - h_y(k-1) \quad , \quad k \geq 0 \quad (7.39)$$

7.4.7 Equilibrio

Si vogliono ora analizzare le caratteristiche del sistema (7.17), (7.18) per quanto riguarda le condizioni di equilibrio. Assumendo costante e pari a \bar{u} l'ingresso del sistema, sulla base dei risultati del Paragrafo 7.2.4 si può affermare che gli stati di equilibrio \bar{x} sono le soluzioni dell'equazione

$$\bar{x} = A\bar{x} + B\bar{u} \quad (7.40)$$

A ogni stato di equilibrio corrisponde l'uscita di equilibrio

$$\bar{y} = C\bar{x} + D\bar{u}$$

Notevole è il caso in cui $I - A$ è invertibile, cioè $\det(I - A) \neq 0$, o, equivalentemente, A non ha autovalori pari a 1: in questa circostanza, infatti, l'Equazione (7.40) ammette una e una sola soluzione, per cui lo stato di equilibrio è unico e risulta

$$\begin{aligned} \bar{x} &= (I - A)^{-1} B\bar{u} \\ \bar{y} &= [C(I - A)^{-1} B + D]\bar{u} \end{aligned}$$

La matrice $C(I - A)^{-1} B + D \in R^{p \times m}$ rappresenta il *guadagno statico* del sistema. Per i sistemi SISO esso costituisce il rapporto tra l'uscita e l'ingresso quando tutte le variabili del sistema stesso, compreso lo stato, sono costanti.

Quando invece $\det(I - A) = 0$, l'Equazione (7.40) può ammettere infinite soluzioni o anche non ammetterne alcuna, anche in dipendenza dal valore di \bar{u} , a seconda che il vettore $B\bar{u}$ sia tra quelli generabili moltiplicando a destra $I - A$ per un opportuno vettore \bar{x} oppure no. Vi sono allora infiniti stati di equilibrio, oppure nessuno, e la nozione di guadagno statico perde comunque senso.

7.5 Stabilità dei sistemi lineari e stazionari

7.5.1 Stabilità del sistema

Si voglia studiare la stabilità del movimento $\bar{x}(k)$ prodotto nel sistema lineare di equazione di stato (7.17) da $\bar{u}(k)$, $k \geq 0$, e $\bar{x}(0) = \bar{x}_0$; in particolare questo movimento può essere anche uno stato di equilibrio. Applicando il principio di sovrapposizione degli effetti (Teorema 7.1) con $u'(k) = u''(k) = \bar{u}(k)$, $x'_0 = \bar{x}_0$, $x''_0 = \bar{x}_0 + \delta x_0$, $\alpha = -1$ e $\beta = 1$, si trova che la differenza $\delta x(k) = x(k) - \bar{x}(k)$ tra movimento perturbato e movimento nominale è retta dall'equazione

$$\delta x(k+1) = A\delta x(k) \quad , \quad \delta x(0) = \delta x_0 \quad (7.41)$$

Pertanto, per le Definizioni 7.4-7.6 (o 7.1-7.3 se ci si riferisce a un equilibrio), il movimento \bar{x} è stabile se, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che per tutti i δx_0 per cui

$$\|\delta x_0\| \leq \delta \quad (7.42)$$

risulti

$$\|\delta x(k)\| \leq \varepsilon \quad , \quad k \geq 0$$

Il movimento (o l'equilibrio) \bar{x} è instabile se ciò non accade, ed è asintoticamente stabile se, oltre a essere stabile, verifica la condizione

$$\lim_{k \rightarrow \infty} \|\delta x(k)\| = 0$$

Nella riformulazione appena data delle Definizioni 7.4-7.6 (o 7.1-7.3) concernenti la stabilità, \bar{x} non appare più in modo esplicito. In altri termini, l'analisi della stabilità di un qualunque movimento \bar{x} , e in particolare anche di un qualunque stato di equilibrio, porta in ogni caso allo studio delle soluzioni dell'Equazione (7.41). Si al variare della condizione iniziale come specificato dalla condizione (7.42). Si può quindi enunciare il seguente fondamentale risultato.

Teorema 7.2 Un movimento (o uno stato di equilibrio) di un sistema lineare stazionario è stabile, asintoticamente stabile o instabile se e solo se tutti i movimenti (o gli stati di equilibrio) del sistema sono rispettivamente stabili, asintoticamente stabili o instabili. ■

Come conseguenza di questo teorema, ha quindi senso per i sistemi considerati parlare di stabilità, stabilità asintotica o instabilità del sistema, invece che di singoli movimenti (o stati di equilibrio), intendendo con questo il sussistere delle proprietà di stabilità, stabilità asintotica o instabilità di tutti i movimenti (o gli stati di equilibrio) del sistema stesso secondo le Definizioni 7.4-7.6 (o 7.1-7.3).

7.5.2 Stabilità e movimento libero

Basandosi sulla (7.41) e ragionando in maniera del tutto simile a quanto fatto per i sistemi continui nel Paragrafo 3.4.2, si può dimostrare il seguente importantissimo risultato.

Teorema 7.3 Un sistema lineare stazionario è stabile se e solo se tutti i movimenti liberi dello stato sono limitati; è asintoticamente stabile se e solo se tutti i movimenti liberi dello stato tendono a zero per $k \rightarrow \infty$; è instabile se e solo se almeno un movimento libero dello stato non è limitato. ■

Le proprietà di stabilità dipendono dunque esclusivamente dalle caratteristiche della matrice della dinamica A . Inoltre, per i sistemi che qui si stanno considerando, la proprietà di stabilità asintotica implica quella di stabilità globale di tutti i movimenti, perché, indipendentemente dal valore della norma di δx_0 , sotto l'ipotesi di asintotica stabilità tutti i movimenti liberi tendono a zero e quindi tutti i movimenti perturbati tendono al movimento nominale, qualunque esso sia.

7.5.3 Stabilità e autovalori

Al variare dello stato iniziale, tutti i modi (7.33), (7.34) compaiono nel movimento libero dello stato e pertanto, ricordando il Teorema 7.3, per accettare che un sistema sia asintoticamente stabile occorre e basta verificare che i singoli modi tendano a zero per k tendente all'infinito. Si arriva allora immediatamente alla formulazione della seguente condizione necessaria e sufficiente.

Teorema 7.4 Il sistema lineare e stazionario (7.17), (7.18) è asintoticamente stabile se e solo se tutti i suoi autovalori hanno modulo minore di 1. ■

Alla luce di questo teorema si usa dire che i punti del piano complesso con modulo minore di 1, o interni al cerchio di raggio unitario centrato nell'origine, costituiscono la *regione di asintotica stabilità*. Analogamente, è facile pervenire alla condizione sufficiente di instabilità riportata qui di seguito.

Teorema 7.5 Il sistema lineare e stazionario (7.17), (7.18) è instabile se almeno uno dei suoi autovalori ha modulo maggiore di 1. ■

Confrontando le condizioni espresse dai teoremi precedenti, si conclude che l'unico caso indecidibile dalla sola analisi degli autovalori si verifica quando il sistema possiede autovalori con modulo unitario insieme ad altri eventuali con modulo minore di 1. In questo caso, esclusa la stabilità asintotica, si potrebbe avere sia la stabilità sia l'instabilità. Quest'ultima si ha se e solo se, tra gli autovalori con modulo unitario, ce n'è almeno uno cui corrisponda almeno un miniblocco di Jordan di dimensioni maggiori di 1, perché in questo caso, e solo in questo caso, nel movimento libero dello stato compare un modo del tipo k^h oppure $k^h \sin(\theta_i k + \varphi_i)$, con $h > 0$, illimitato per $k \rightarrow \infty$. Si noti che questi termini sono sicuramente assenti quando non vi sono autovalori multipli sulla circonferenza di raggio unitario centrata nell'origine: in questo caso allora la condizione sufficiente del Teorema

7.5 diventa anche necessaria, mentre il sistema risulta stabile se e solo se tutti i suoi autovalori hanno modulo non maggiore di 1.

Esempio 7.6 Si consideri il sistema con matrice della dinamica

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

Il movimento libero della prima variabile di stato prodotto dallo stato iniziale $x_1(0) = 1, x_2(0) = 0$ assume per $k \geq 0$, i valori 1, 1, 2, 3, 5, 8, ... Essi sono i cosiddetti numeri di Fibonacci, che tendono all'infinito con k : il sistema è dunque instabile. Infatti i suoi autovalori sono $(1 \pm \sqrt{5})/2$ e uno di essi ha modulo maggiore di 1.

In sintesi, il ruolo giocato per i sistemi a tempo continuo dal semipiano sinistro del piano complesso, dall'asse immaginario e dal semipiano destro viene svolto per i sistemi a tempo discreto rispettivamente dal cerchio di raggio unitario centrato nell'origine, dalla circonferenza sua frontiera e dalla regione all'esterno di essa.

7.5.4 Stabilità e polinomio caratteristico

Concentrandosi ora sulla proprietà di stabilità asintotica, si osservi che, per il Teorema 7.4, l'accertamento della sua presenza richiede la soluzione di un'equazione algebrica per il calcolo degli autovalori. L'*equazione caratteristica*

$$\varphi(z) = 0$$

ottenuta uguagliando a zero il *polinomio caratteristico*

$$\varphi(z) = \det(zI - A) = z^n + p_1 z^{n-1} + p_2 z^{n-2} + \dots + p_{n-1} z + p_n \quad (7.43)$$

è un'equazione polinomiale per risolvere la quale bisogna in generale usare metodi numerici iterativi. Esistono però alcuni risultati che, senza richiedere la soluzione dell'equazione caratteristica, consentono di dire se tutti gli n autovalori z_i hanno modulo minore di 1 oppure no, dando così luogo, per il Teorema 7.4, a criteri di stabilità asintotica così semplici da poter essere applicati senza l'ausilio di particolari supporti di calcolo anche a sistemi di ordine elevato.

Facendo riferimento al polinomio caratteristico

$$\varphi(z) = \varphi_0 z^n + \varphi_1 z^{n-1} + \varphi_2 z^{n-2} + \dots + \varphi_{n-1} z + \varphi_n, \quad \varphi_0 \neq 0 \quad (7.44)$$

scritto in una forma un poco più generale della (7.43), si possono innanzitutto formulare le seguenti condizioni solo necessarie, o solo sufficienti, di cui si omettono le dimostrazioni.

Teorema 7.6 Condizione necessaria affinché il sistema (7.17), (7.18) sia asintoticamente stabile è che il modulo di φ_n/φ_0 risulti minore di 1. ■

Teorema 7.7 Condizione necessaria affinché il sistema (7.17), (7.18) sia asintoticamente stabile è che il polinomio caratteristico valutato in $z = 1$ sia concorde con φ_0 , cioè

$$\varphi_0 \varphi(1) = \varphi_0 \sum_{i=0}^n \varphi_i > 0$$

Teorema 7.8 Condizione necessaria affinché il sistema (7.17), (7.18) sia asintoticamente stabile è che il polinomio caratteristico valutato in $z = -1$ sia concorde con φ_0 , se n è pari, e discorde da φ_0 , se n è dispari, cioè

$$(-1)^n \varphi_0 \varphi(-1) > 0$$

Teorema 7.9 Condizione sufficiente affinché il sistema (7.17), (7.18) sia asintoticamente stabile è che la sequenza dei coefficienti φ_i , $i = 0, 1, \dots, n$, sia decrescente e positiva, cioè

$$\varphi_0 > \varphi_1 > \varphi_2 > \dots > \varphi_{n-1} > \varphi_n > 0$$

Teorema 7.10 Condizione sufficiente affinché il sistema (7.17), (7.18) sia asintoticamente stabile è che la somma dei moduli dei coefficienti φ_i , $i = 1, 2, \dots, n$, sia minore del modulo di φ_0 , cioè

$$\sum_{i=1}^n |\varphi_i| < |\varphi_0|$$

Esempio 7.7 I Teoremi 7.6 e 7.8 consentono di concludere immediatamente circa la mancanza di stabilità asintotica per un sistema dotato del polinomio caratteristico

$$\varphi(z) = z^3 + 4z^2 - 0.81z - 3.24$$

cui corrispondono le radici $0.9, -0.9$ e -4 .

I Teoremi 7.9 e 7.10 permettono di accettare la stabilità asintotica per un sistema con polinomio caratteristico

$$\varphi(z) = z^3 + 0.6z^2 + 0.11z + 0.006$$

cui corrispondono le radici $-0.1, -0.2$ e -0.3 .

Invece, applicando i criteri appena introdotti, nulla si può dire per un polinomio caratteristico come

$$\varphi(z) = 8z^3 + 12z^2 + 6z + 1 \quad (7.45)$$

cui corrispondono tre radici coincidenti in $z = -0.5$.

Trasformazione bilineare Una classica tecnica per ricavare condizioni necessarie e sufficienti di stabilità asintotica per sistemi a tempo discreto consiste nell'effettuare un cambio di variabile $z = \psi(s)$ nel polinomio caratteristico $\varphi(z)$ in modo tale che la nuova equazione caratteristica $\varphi(\psi(s)) = 0$ sia ancora di tipo polinomiale e le sue radici abbiano parte reale negativa in tutti e soli i casi in cui le radici del polinomio caratteristico originario $\varphi(z)$ hanno modulo minore di 1. Una volta che ci si sia ricondotti al caso dei sistemi a tempo continuo, si può applicare un qualunque criterio di asintotica stabilità valido per quest'ultima classe di sistemi.

Un cambio di variabile adatto allo scopo sopra indicato è definito dalla *trasformazione bilineare*

$$z = \frac{1+s}{1-s} \quad (7.46)$$

cui corrisponde la trasformazione inversa

$$s = \frac{z-1}{z+1} \quad (7.47)$$

Ricordando la (7.44), l'equazione caratteristica diventa allora

$$\varphi_0 \left(\frac{1+s}{1-s} \right)^n + \varphi_1 \left(\frac{1+s}{1-s} \right)^{n-1} + \varphi_2 \left(\frac{1+s}{1-s} \right)^{n-2} + \dots + \varphi_{n-1} \left(\frac{1+s}{1-s} \right) + \varphi_n = 0$$

cioè, in forma polinomiale,

$$\begin{aligned} \varphi_0 (1+s)^n + \varphi_1 (1+s)^{n-1} (1-s) + \varphi_2 (1+s)^{n-2} (1-s)^2 + \\ + \dots + \varphi_{n-1} (1+s) (1-s)^{n-1} + \varphi_n (1-s)^n = 0 \end{aligned} \quad (7.48)$$

Si osservi che il coefficiente di s^n in tale equazione è

$$\sum_{i=0}^n (-1)^i \varphi_i = (-1)^n \varphi(-1)$$

Esso può risultare nullo, e in tal caso il sistema corrispondente al polinomio caratteristico in esame non è asintoticamente stabile perché tra gli autovalori c'è anche $z = -1$ (l'abbassamento di grado è legato al fatto che la trasformazione (7.47) non è definita per $z = -1$). Quando ciò accade, il sistema con polinomio caratteristico (7.44) non è asintoticamente stabile. Assumendo viceversa che il grado dell'equazione polinomiale (7.48) sia n , le n radici dell'Equazione (7.44) hanno modulo minore di 1, se e solo se le n radici dell'Equazione (7.48) hanno parte reale negativa, proprietà che può essere accertata, per esempio, con il criterio di Routh.

Esempio 7.8 (Seguito dell'Esempio 7.7) Effettuando la sostituzione (7.46) nel polinomio (7.45) e azzerando, si ottiene

$$\begin{aligned} 8 \left(\frac{1+s}{1-s} \right)^3 + 12 \left(\frac{1+s}{1-s} \right)^2 + 6 \left(\frac{1+s}{1-s} \right) + 1 &= 0 \\ 8(1+s)^3 + 12(1+s)^2(1-s) + 6(1+s)(1-s)^2 + (1-s)^3 &= 0 \\ s^3 + 9s^2 + 27s + 27 &= 0 \end{aligned}$$

Dall'ultima equazione, per il criterio di Routh (Teorema 3.7), si ha la conferma che tutte le radici del polinomio (7.45) hanno modulo minore di 1; infatti la tabella di Routh risulta

1	27
9	27
24	
27	

e tutti gli elementi della prima colonna sono positivi.

Criterio di Jury Per i sistemi discreti esistono anche delle condizioni necessarie e sufficienti di stabilità asintotica di tipo diretto, che cioè non richiedono la preventiva riformulazione del problema in quello dell'analisi di stabilità di un sistema a tempo continuo. Per introdurre la più comune tra queste, occorre per prima cosa definire la cosiddetta *tabella di Jury*, costruita a partire dai coefficienti del polinomio caratteristico. Essa possiede $n+1$ righe e ha una struttura triangolare, perché in ogni riga il numero di elementi diminuisce di uno:

φ_0	φ_1	φ_2	\dots	\dots	\dots	φ_{n-1}	φ_n
h_1	h_2	h_3	\dots	\dots	\dots	h_{n-1}	h_n
l_1	l_2	l_3	\dots	\dots	\dots	l_{n-1}	
\vdots	\vdots	\vdots	\ddots	\ddots	\ddots	\vdots	
\vdots	\vdots	\vdots	\ddots	\ddots	\ddots	\vdots	

La prima riga contiene i coefficienti del polinomio caratteristico ordinati secondo potenze decrescenti di z ; ogni delle righe successive si costruisce sulla base degli elementi della riga precedente. Avendo indicato con h_i gli elementi di una generica riga, di v elementi, e con l_i gli elementi della riga successiva, di $v - 1$ elementi, si ha

$$l_i = \frac{1}{h_1} \det \begin{pmatrix} h_1 & h_{v-i+1} \\ h_v & h_i \end{pmatrix} = h_i - \frac{h_1 h_{v-i+1}}{h_1} \quad (7.49)$$

Se risulta $h_1 = 0$, la (7.49) non è applicabile e allora si dirà, per convenzione, che la tabella di Jury non è ben definita.

Si può ora enunciare la seguente condizione necessaria e sufficiente di asintotica stabilità, di cui qui non si fornirà la dimostrazione, che costituisce una delle possibili varianti del *criterio di Jury*.

Teorema 7.11 Il sistema (7.17), (7.18) è asintoticamente stabile se e solo se la tabella di Jury relativa al suo polinomio caratteristico è ben definita e tutti gli elementi della sua prima colonna hanno lo stesso segno. ■

Si noti, quindi, che la costruzione dell'intera tabella può anche non essere necessaria, e inoltre la divisione per h_1 del determinante nell'Equazione (7.49) non è indispensabile per la validità del teorema.

Esempio 7.9 Utilizzando i Teoremi 7.6-7.10 non si può trarre alcuna conclusione circa la collocazione nel piano complesso delle radici del polinomio

$$\varphi(z) = 8z^3 + 6z^2 + 2z + 3$$

La tabella di Jury corrispondente è

$$\begin{array}{cccc} 8 & 6 & 2 & 3 \\ \simeq 6.87 & 5.25 & 0.25 & \\ \simeq 6.86 & \simeq 5.06 & & \\ \simeq 3.13 & & & \end{array}$$

Pertanto, tutte le radici di φ hanno modulo minore di 1, e infatti esso si può scrivere nella forma

$$\varphi(z) \simeq 8(z + 0.92)(z - 0.085 + j0.63)(z + 0.085 + j0.63)$$

7.5.5 Stabilità e parametri incerti

Le tecniche del paragrafo precedente consentono anche di trattare un polinomio caratteristico affatto da parametri incerti per determinare la *regione di asintotica stabilità nell'insieme dei parametri*, costituita dai valori dei parametri che fanno sì che il sistema corrispondente sia asintoticamente stabile. In sostanza, si tratta di imporre che sia soddisfatto un opportuno insieme di disequazioni, che costituiscono condizioni necessarie e sufficienti di stabilità asintotica. Quando si adotta il metodo della trasformazione bilineare è possibile che i calcoli, seppure concettualmente banali, diventino complicati anche per piccoli valori di n e con pochi parametri liberi. Può essere allora più comodo riferirsi al criterio di Jury.

Esempio 7.10 Si abbia un sistema dotato del polinomio caratteristico

$$\varphi(z) = z^2 + \alpha z + \beta$$

La tabella di Jury corrispondente è

$$\begin{array}{ccc} 1 & \alpha & \beta \\ t - \beta^2 & \alpha(1 - \beta) & \\ \frac{t - \beta}{t + \beta} [(1 + \beta)^2 - \alpha^2] & & \end{array}$$

Pertanto, il sistema è asintoticamente stabile per tutte le coppie di valori α, β che soddisfano le disequazioni

$$t - \beta^2 > 0 \quad , \quad \frac{t - \beta}{t + \beta} [(1 + \beta)^2 - \alpha^2] > 0$$

cioè

$$\beta > -\alpha - 1 \quad , \quad \beta > \alpha - 1 \quad , \quad \beta < 1$$

o, in altri termini, per tutte le coppie cui corrispondono punti nella regione in colore chiaro della Figura 7.6.

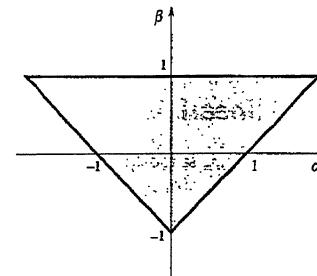


Figura 7.6
Regione di asintotica
stabilità nell'insieme dei
parametri per l'Esempio
7.10.

Vale la pena di osservare che per i sistemi a tempo discreto non esiste un risultato comparabile quanto a potenza e semplicità con il criterio di Kharitonov, valido per i sistemi a tempo continuo (Paragrafo 3.4.5).

7.5.6 Proprietà dei sistemi asintoticamente stabili

I sistemi asintoticamente stabili sono quelli di maggiore interesse per le applicazioni. In genere, non è accettabile che esistano movimenti liberi illimitati, né che ve ne siano di limitati, ma non asintoticamente nulli. Tra l'altro, la stabilità non asintotica è contraddistinta da condizioni molto critiche, quali la presenza di autovalori sulla circonferenza unitaria centrata nell'origine, circa le quali raramente può esserci certezza quando si ha a che fare con modelli, necessariamente approssimati, di una determinata realtà. Per di più, la proprietà di stabilità asintotica comporta alcune conseguenze veramente notevoli sui sistemi che ne godono.

- Innanzitutto, in un sistema asintoticamente stabile, il movimento che si ha per $k \rightarrow \infty$ è indipendente dallo stato iniziale, perché il movimento asintotico coincide con quello forzato, visto che il movimento libero tende ad annullarsi.
- Inoltre, la risposta all'impulso, sia dello stato sia dell'uscita, tende asintoticamente a zero, perché essa, come si è osservato al Paragrafo 7.4.6, per $k > 0$ coincide con un movimento libero ritardato di un passo.

- Analogamente, la risposta a un qualunque ingresso di durata limitata nel tempo tende a zero in modo asintotico. Infatti, detti \tilde{k} l'istante in cui l'ingresso diventa definitivamente nullo e \tilde{x} il corrispondente valore dello stato, i movimenti del sistema negli istanti di tempo successivi sono quelli liberi generati da \tilde{x} ; si ha cioè

$$x(k) = A^{k-\tilde{k}} \tilde{x} \quad , \quad k > \tilde{k}$$

$$y(k) = C A^{k-\tilde{k}} \tilde{x} \quad , \quad k > \tilde{k}$$

- Visto che $\det(I - A) \neq 0$, perché un sistema asintoticamente stabile non può avere autovalori in $\lambda = 1$, lo stato e l'uscita di equilibrio conseguenti a un qualunque ingresso $u(k) = \bar{u}$ sono unici (si veda il Paragrafo 7.4.7). A essi tendono i movimenti generati da $u(k) = \bar{u}$ $\text{sca}^*(k)$, qualunque sia lo stato iniziale.
- Infine, un sistema asintoticamente stabile gode anche della proprietà di *stabilità esterna* (o *stabilità BIBO*), cioè produce un movimento forzato dell'uscita limitato in corrispondenza di ogni ingresso limitato. Viceversa, si può anche verificare che, sotto blande ipotesi, i sistemi stabili esternamente sono anche asintoticamente stabili. Si veda a questo proposito il Teorema 7.20.

7.6 Linearizzazione e stabilità dell'equilibrio di sistemi non lineari

Come anticipato al Paragrafo 7.2.3, si mostra ora come i risultati presentati nei paragrafi precedenti possano essere sfruttati anche per lo studio di sistemi non lineari e stazionari. Si consideri perciò un sistema non lineare, in generale MIMO, descritto da

$$x(k+1) = f(x(k), u(k)) \quad (7.50)$$

$$y(k) = g(x(k), u(k)) \quad (7.51)$$

soggetto all'ingresso costante $u(t) = \bar{u}$. Si faccia poi riferimento a un suo stato di equilibrio \tilde{x} e alla corrispondente uscita di equilibrio \bar{y} , detti *nominali*, che soddisfano le identità

$$\tilde{x} = f(\tilde{x}, \bar{u}) \quad (7.52)$$

$$\bar{y} = g(\tilde{x}, \bar{u}) \quad (7.53)$$

7.6.1 Linearizzazione

Il procedimento della *linearizzazione* consiste nel descrivere il comportamento di un sistema non lineare attorno all'equilibrio nominale mediante un particolare sistema lineare. Quest'ultimo sistema costituisce solo un'approssimazione del sistema originario, ma è estremamente utile per affrontare molti problemi specifici, perché a esso sono applicabili i potenti metodi di analisi e sintesi disponibili per i sistemi lineari.

Se si introducono le variazioni $\delta u(k)$, $\delta x(k)$ e $\delta y(k)$ delle variabili di ingresso, stato e uscita rispetto a \bar{u} , \tilde{x} e \bar{y} , nonché dello stato iniziale δx_{k_0} ancora rispetto a \tilde{x} , cioè se si pone

$$u(k) = \bar{u} + \delta u(k)$$

$$x(k) = \tilde{x} + \delta x(k)$$

$$y(k) = \bar{y} + \delta y(k)$$

$$x_{k_0} = \tilde{x} + \delta x_{k_0}$$

le equazioni (7.50), (7.51) diventano

$$\tilde{x} + \delta x(k+1) = f(\tilde{x} + \delta x(k), \bar{u} + \delta u(k))$$

$$\bar{y} + \delta y(k) = g(\tilde{x} + \delta x(k), \bar{u} + \delta u(k))$$

con la condizione iniziale

$$\tilde{x} + \delta x(k_0) = \tilde{x} + \delta x_{k_0} \quad (7.54)$$

Se poi si suppone che le funzioni f e g siano sufficientemente regolari, esse possono essere sviluppate in serie di Taylor rispetto a x e u in $x = \tilde{x}$ e $u = \bar{u}$, arrestandosi ai termini del primo ordine; si ottiene allora

$$\tilde{x} + \delta x(k+1) = f(\tilde{x}, \bar{u}) + \left. \frac{\partial f(x, u)}{\partial x} \right|_{x=\tilde{x}, u=\bar{u}} \delta x(k) + \left. \frac{\partial f(x, u)}{\partial u} \right|_{x=\tilde{x}, u=\bar{u}} \delta u(k)$$

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

Da queste ultime equazioni e dalle (7.52)-(7.54) si ha infine

$$\delta x(k+1) = A \delta x(k) - B \delta u(k) \quad , \quad \delta x(k_0) = \delta x_{k_0} \quad (7.55)$$

$$\delta y(k) = C \delta x(k) - D \delta u(k) \quad (7.56)$$

dove

$$A = \left. \frac{\partial f(x, u)}{\partial x} \right|_{x=\tilde{x}, u=\bar{u}} \quad (7.57)$$

$$B = \left. \frac{\partial f(x, u)}{\partial u} \right|_{x=\tilde{x}, u=\bar{u}} \quad (7.58)$$

$$C = \left. \frac{\partial g(x, u)}{\partial x} \right|_{x=\tilde{x}, u=\bar{u}} \quad (7.59)$$

$$D = \left. \frac{\partial g(x, u)}{\partial u} \right|_{x=\tilde{x}, u=\bar{u}} \quad (7.60)$$

Si è così ottenuto un sistema lineare che in senso stretto lega le variazioni prime delle variabili in gioco e che si chiama *sistema linearizzato*. Esso è utilizzabile per descrivere in maniera approssimata il comportamento del sistema (7.50), (7.51) attorno al particolare equilibrio considerato nel caso in cui le variazioni delle funzioni di ingresso e dello stato iniziale $\delta u(k)$ e δx_{k_0} , nonché le variazioni $\delta x(k)$ e $\delta y(k)$ da esse provocate, siano sufficientemente piccole in norma.

Esempio 7.11 (Seguito dell'Esempio 7.3) Assumendo che ε e α siano costanti, ponendo cioè $u_1(k) = \tilde{u}$, $u_2(k) = \tilde{s}$, all'equilibrio si ha

$$\bar{x} = \tilde{x} \left(1 - \eta\tilde{\alpha}^2\right) + \mu\eta\tilde{\alpha}^2 + \eta\tilde{\alpha}\tilde{\varepsilon}$$

Pertanto, assumendo $\tilde{\alpha} \neq 0$, si trova

$$\bar{x} = \mu + \tilde{\varepsilon}/\tilde{\alpha} \quad (7.61)$$

cioè il valore di equilibrio della stima x è pari al valore esatto del parametro ignoto μ quando $\tilde{\varepsilon} = 0$ (si usa dire che la stima è non polarizzata). Inoltre il sistema linearizzato risulta

$$\delta x(k+1) = (1 - \eta\tilde{\alpha}^2)\delta x(k) - \eta\tilde{\varepsilon}\delta u_1(k) + \eta\tilde{\alpha}\delta u_2(k) \quad (7.62)$$

$$\delta y(k) = \delta x(k) \quad (7.63)$$

7.6.2 Stabilità dell'equilibrio

Ora si mostrerà in che modo l'analisi della stabilità del sistema linearizzato (7.55)-(7.60) consenta in molti casi di determinare le proprietà di stabilità dello stato di equilibrio \tilde{x} del sistema non lineare originario (7.50), (7.51). Si noti che i risultati saranno esatti, maigrado essi si ottengano usando un modello approssimato. Questo fatto in realtà non dovrebbe meravigliare in quanto, per definizione, le proprietà di stabilità sono locali, o "in piccolo", e quindi è ragionevole che esse possano essere studiate facendo uso di modelli linearizzati, almeno in situazioni non critiche.

Teorema 7.12 Lo stato di equilibrio \tilde{x} , relativo all'ingresso \tilde{u} , del sistema non lineare (7.50), (7.51) è asintoticamente stabile se tutti gli autovalori del sistema linearizzato corrispondente (7.55)-(7.60) hanno modulo minore di 1. ■

Visto che questo teorema richiede la stabilità asintotica del sistema linearizzato, la determinazione degli autovalori può anche essere evitata e sostituita dallo studio del polinomio caratteristico, secondo quanto indicato dai risultati del Paragrafo 7.5.4.

Teorema 7.13 Lo stato di equilibrio \tilde{x} , relativo all'ingresso \tilde{u} , del sistema non lineare (7.50), (7.51) è instabile se almeno uno degli autovalori del sistema linearizzato corrispondente (7.55)-(7.60) ha modulo maggiore di 1. ■

Pertanto, mentre il Teorema 7.13 è l'esatta controparte del Teorema 7.5, il Teorema 7.12 rappresenta una "estensione" della sola parte sufficiente del Teorema 7.4. Se quindi il sistema (7.55)-(7.60) ha alcuni autovalori con modulo minore di 1 e altri con modulo unitario, niente si può dire circa la stabilità di \tilde{x} esaminando solo il sistema linearizzato: essa dipende dai termini successivi al primo dello sviluppo in serie di Taylor di $f(x, u)$ attorno a \tilde{x}, \tilde{u} . Per affrontare tali situazioni sono disponibili altre tecniche che qui non sono trattate.

Esempio 7.12 (Seguito degli Esempi 7.3 e 7.11) Nello studiare le proprietà di convergenza dell'algoritmo di stima, è importante poter garantire che, almeno quando α ed ε sono costanti, la stima converga al valore di equilibrio (7.61). A questo proposito, si può fare riferimento alle Equazioni (7.62), (7.63) del sistema linearizzato. La matrice della dinamica è costituita in questo caso dallo scalare $1 - \eta\tilde{\alpha}^2$, il cui modulo è minore di 1 purché risulti

$$0 < \eta\tilde{\alpha}^2 < 2 \quad (7.64)$$

Quindi, si noti che, per garantire la convergenza, la costante η non può essere troppo elevata.

Anche assumendo $\eta < 2/\tilde{\alpha}^2$, il risultato appena trovato garantisce la convergenza locale: in altri termini, la convergenza è assicurata se si dispone di una stima iniziale di μ prossima a quella di equilibrio. In realtà si può dimostrare che, nelle ipotesi fatte, la convergenza è globale, cioè la regione di attrazione dello stato di equilibrio è costituita dall'intero insieme di stato. Infatti, quando α ed ε sono costanti l'equazione di stato del sistema non lineare diventa

$$x(k+1) = \left(1 - \eta\tilde{\alpha}^2\right)x(k) + \mu\eta\tilde{\alpha}^2 + \eta\tilde{\alpha}\tilde{\varepsilon}$$

e assume la forma di un'equazione di stato di un sistema lineare e stazionario, con ingresso costante. Quindi, qualunque sia lo stato iniziale $x(0)$, la stima converge al valore di equilibrio $\tilde{x} = \mu + \tilde{\varepsilon}/\tilde{\alpha}$. purché valga la (7.64). Si osservi che tale proprietà, valida in questo caso, non è vera in generale.

7.7 Raggiungibilità, osservabilità e scomposizione canonica dei sistemi lineari e stazionari

In questo paragrafo saranno presentate per i sistemi a tempo discreto le proprietà di raggiungibilità e osservabilità, nonché la scomposizione canonica, già introdotte per i sistemi a tempo continuo nel Paragrafo 3.6. Saranno qui approfondite alcune questioni che là non erano state trattate in dettaglio per motivi di semplicità. Si farà riferimento al sistema dinamico di ordine n con m ingressi e p uscite

$$x(k+1) = Ax(k) + Bu(k) \quad (7.65)$$

$$y(k) = Cx(k) + Du(k) \quad (7.66)$$

7.7.1 Raggiungibilità

Si consideri la seguente proprietà del movimento forzato dello stato.

Definizione 7.7 Uno stato \tilde{x} del sistema (7.65), (7.66) si dice *raggiungibile* se esistono un istante di tempo finito $\tilde{k} > 0$ e un ingresso \tilde{u} , definito tra 0 e $\tilde{k}-1$, tali che, detto $\tilde{x}_f(k)$, $0 \leq k \leq \tilde{k}$, il movimento forzato dello stato generato da \tilde{u} , risulti $\tilde{x}_f(\tilde{k}) = \tilde{x}$.

Un sistema i cui stati siano tutti raggiungibili si dice *completamente raggiungibile*. ■

In altri termini, quindi, un particolare vettore \tilde{x} costituisce uno stato raggiungibile se è possibile, con un'opportuna scelta dell'ingresso, trasferirvi dall'origine lo stato del sistema in un tempo finito arbitrario. Si osservi che la trasformazione d'uscita (7.66) non gioca in realtà alcun ruolo in questo contesto. La raggiungibilità dipende esclusivamente dall'Equazione (7.65), cosicché a volte essa viene attribuita addirittura alla coppia (A, B) .

Per accettare le proprietà di raggiungibilità del sistema (7.65), (7.66) si può ragionare come segue. Innanzitutto, gli stati raggiungibili al tempo $k=1$ sono quelli che si possono descrivere come combinazioni lineari delle colonne di B ; infatti

$$x_f(1) = Bu(0) = \sum_{i=1}^m b_i u_i(0)$$

dove b_i rappresenta la colonna i -esima di B . Analogamente, al tempo $k=2$ sono raggiungibili gli stati che risultino combinazioni lineari delle colonne di B

e di quelle di AB , in quanto $x_f(2) = Bu(1) + ABu(0)$. Al generico tempo k , allora, sono raggiungibili gli stati ottenibili come combinazioni lineari delle colonne della matrice

$$[B \ AB \ A^2B \ \dots \ A^{k-1}B]$$

Si osservi ora che, per il teorema di Cayley-Hamilton del Paragrafo A.3.1, risulta

$$A^n = -(\alpha_1 A^{n-1} + \alpha_2 A^{n-2} + \dots + \alpha_{n-1} A + \alpha_n I)$$

dove gli scalari α_i sono i coefficienti del polinomio caratteristico di A . Perciò le colonne della matrice $A^n B$ sono combinazioni lineari delle colonne delle matrici $A^i B$, $i = 0, 1, \dots, n-1$. Quindi, se uno stato è raggiungibile, lo è al massimo in n passi. Quanto sopra costituisce la dimostrazione dei risultati seguenti, che fanno riferimento alla cosiddetta *matrice di raggiungibilità*, definita come

$$M_r = [B \ AB \ A^2B \ \dots \ A^{n-1}B] \in R^{n \times mn}$$

Teorema 7.14 Il sistema (7.65), (7.66) è completamente raggiungibile, ovvero la coppia (A, B) è completamente raggiungibile, se e solo se il rango della matrice di raggiungibilità è pari a n , cioè

$$\rho(M_r) = n \quad (7.67)$$

Se il sistema ha un solo ingresso, cioè $m = 1$, la matrice M_r è quadrata e la condizione necessaria e sufficiente (7.67) è equivalente a $\det(M_r) \neq 0$.

Nel caso in cui un sistema non sia completamente raggiungibile, si può isolare la sua "parte" dotata della proprietà di raggiungibilità, così come specificato nel teorema seguente dove $n_r = \rho(M_r)$.

Teorema 7.15 Mediante un opportuno, non unico, cambio di variabili di stato

$$\hat{x}(k) = T_r x(k)$$

l'equazione di stato (7.65) può essere posta nella forma

$$\hat{x}(k+1) = \hat{A}\hat{x}(k) + \hat{B}u(k) \quad (7.68)$$

dove

$$\hat{A} = \begin{bmatrix} \hat{A}_a & \hat{A}_{ab} \\ 0 & \hat{A}_b \end{bmatrix}, \quad \hat{A}_a \in R^{n_r \times n_r} \quad (7.69)$$

$$\hat{B} = \begin{bmatrix} \hat{B}_a \\ 0 \end{bmatrix}, \quad \hat{B}_a \in R^{n_r \times m} \quad (7.70)$$

$$\rho([\hat{B}_a \ \hat{A}_a \hat{B}_a \ \hat{A}_a^2 \hat{B}_a \ \dots \ \hat{A}_a^{n_r-1} \hat{B}_a]) = n_r \quad (7.71)$$

Per costruire la matrice T_r si selezionino innanzitutto n_r colonne linearmente indipendenti in M_r . Ognuna di tali colonne rappresenta uno stato raggiungibile ed esse nel complesso descrivono gli stati raggiungibili, nel senso che, come già dimostrato, sono raggiungibili tutti e soli gli stati ottenibili combinando linearmente queste colonne. Esse, eventualmente moltiplicate per costanti non nulle, vengono

poste in T_r^{-1} dove precedono altre $n - n_r$ colonne arbitrarie, ma tali che risult $\det(T_r^{-1}) \neq 0$. La procedura ha sempre buon esito in virtù della (7.71).

Se si partiziona il vettore \hat{x} come

$$\hat{x} = \begin{bmatrix} \hat{x}_a \\ \hat{x}_b \end{bmatrix}, \quad \hat{x}_a \in R^{n_r}$$

l'Equazione (7.68), grazie alle Equazioni (7.69), (7.70), si può scrivere nella forma

$$\hat{x}_a(k+1) = \hat{A}_a \hat{x}_a(k) + \hat{A}_{ab} \hat{x}_b(k) + \hat{B}_a u(k) \quad (7.72)$$

$$\hat{x}_b(k+1) = \hat{A}_b \hat{x}_b(k) \quad (7.73)$$

La struttura di queste equazioni mostra con chiarezza che tutti i movimenti forzati producono $\hat{x}_b(k) = 0$, $k > 0$, in quanto u non influisce né direttamente né tramite \hat{x}_a sulla variabile \hat{x}_b , e, d'altra parte, un'opportuna scelta di u consente per la (7.71) di far assumere un qualunque valore al vettore \hat{x}_a a un istante di tempo finito maggiore o uguale a n_r . Si dirà pertanto che l'Equazione (7.72) costituisce la *parte raggiungibile* del sistema (7.65), (7.66) e l'Equazione (7.73) la *parte non raggiungibile*. La matrice \hat{A} ha una struttura triangolare a blocchi e perciò i suoi autovalori sono quelli dei blocchi sulla diagonale \hat{A}_a e \hat{A}_b : si usa dire che gli autovalori di \hat{A}_a sono quelli della parte raggiungibile e gli autovalori di \hat{A}_b quelli della parte non raggiungibile.

Per concludere, la proprietà di raggiungibilità di uno stato \bar{x} implica quella di *controllabilità*, consistente nella possibilità di portare lo stato del sistema da $x(0) = \bar{x}$ a $x = 0$ in un tempo finito mediante un'opportuna scelta dell'ingresso. Tuttavia, per la classe di sistemi dinamici qui considerata, le due proprietà non coincidono. Per esempio, il sistema descritto da $x(k+1) = 0$ di ordine n qualunque, è completamente controllabile, ma non ha alcuno stato raggiungibile, tranne quello nullo. Si noti la differenza rispetto al caso dei sistemi a tempo continuo (Paragrafo 3.6.2).

7.7.2 Osservabilità

Si può ora passare a considerare una seconda proprietà del sistema (7.65), (7.66), questa volta relativa al movimento libero dell'uscita.

Definizione 7.8 Uno stato $\bar{x} \neq 0$ del sistema (7.65), (7.66) si dice *non osservabile* se, qualunque sia $\bar{k} > 0$ finito, detto $\bar{y}_l(k)$, $k \geq 0$, il movimento libero dell'uscita generato da \bar{x} , risulta $\bar{y}_l(k) = 0$, $0 \leq k \leq \bar{k}$.

Un sistema privo di stati non osservabili si dice *completamente osservabile*. ■

In altri termini, un particolare vettore \bar{x} costituisce uno stato non osservabile se l'esame di un tratto di qualunque durata del movimento libero dell'uscita da esso generata non consente di distinguerlo dal vettore $x = 0$, che ovviamente genera un movimento libero dell'uscita identicamente nullo. Si osservi che le matrici B e D delle Equazioni (7.65), (7.66) non hanno alcun ruolo in questo contesto; l'osservabilità viene attribuita addirittura alla coppia (A, C) .

Per accettare le proprietà di osservabilità del sistema (7.65), (7.66) si può ragionare come segue. Innanzitutto, gli stati non osservabili devono produrre un'uscita nulla al tempo $k = 0$, cioè devono soddisfare l'equazione $Cx(0) = 0$.

Analogamente, al tempo $k = 1$ deve essere $y_1(1) = CAx(0) = 0$. Considerando l'uscita fino al generico tempo k , la non osservabilità implica che

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^k \end{bmatrix} x(0) = 0$$

Si osservi ora che, per il teorema di Cayley-Hamilton del Paragrafo A.3.1, le righe della matrice CA^n sono combinazioni lineari delle righe delle matrici CA^i , $i = 0, 1, \dots, n-1$. Quindi, se uno stato iniziale genera un'uscita non nulla, questa deve assumere per la prima volta un valore diverso da zero in un istante di tempo $k < n$. Quanto sopra costituisce la dimostrazione dei risultati seguenti, che fanno riferimento alla cosiddetta *matrice di osservabilità*, definita come

$$M_o = [C' A'C' A'^2C' \cdots A'^{n-1}C'] \in R^{n \times pn}$$

Teorema 7.16 Il sistema (7.65), (7.66) è completamente osservabile, ovvero la coppia (A, C) è completamente osservabile, se e solo se il rango della matrice di osservabilità è pari a n , cioè

$$\rho(M_o) = n \quad (7.74)$$

■

Se il sistema ha una sola uscita, cioè $p = 1$, la matrice M_o è quadrata e la condizione necessaria e sufficiente (7.74) è equivalente a $\det(M_o) \neq 0$.

Nel caso in cui un sistema non sia completamente osservabile, si può isolare la sua "parte" dotata della proprietà di non osservabilità, così come specificato nel teorema seguente, dove $n_o = \rho(M_o)$.

Teorema 7.17 Mediante un opportuno, non unico, cambio di variabili di stato

$$\hat{x}(k) = T_o x(k)$$

il sistema (7.65), (7.66) con ingresso nullo può essere posto nella forma

$$\hat{x}(k+1) = \hat{A}\hat{x}(k) \quad (7.75)$$

$$y(k) = \hat{C}\hat{x}(k) \quad (7.76)$$

dove

$$\hat{A} = \begin{bmatrix} \hat{A}_a & 0 \\ \hat{A}_{ba} & \hat{A}_b \end{bmatrix}, \quad \hat{A}_a \in R^{n_o \times n_o} \quad (7.77)$$

$$\hat{C} = [\hat{C}_a \ 0], \quad \hat{C}_a \in R^{p \times n_o} \quad (7.78)$$

$$\rho([\hat{C}'_a \hat{A}'_a \hat{C}'_a \hat{A}'^2_a \hat{C}'_a \cdots \hat{A}'^{n_o-1}_a \hat{C}'_a]) = n_o \quad (7.79)$$

■

Per costruire la matrice T_o si selezionino innanzitutto $n - n_o$ vettori colonna linearmente indipendenti ζ_i tali che

$$M'_o \zeta_i = 0$$

Ognuno di questi vettori rappresenta uno stato non osservabile ed essi nel complesso descrivono gli stati non osservabili, nel senso che sono non osservabili tutti e soli gli stati ottenuti combinando linearmente questi vettori. Essi, moltiplicati eventualmente per costanti non nulle, vengono posti come colonne in T_o^{-1} preceduti da altre n_o colonne arbitrarie, ma tali che risulti $\det(T_o^{-1}) \neq 0$. La procedura ha sempre buon esito in virtù della (7.79).

Se si partitiona il vettore \hat{x} come

$$\hat{x} = \begin{bmatrix} \hat{x}_a \\ \hat{x}_b \end{bmatrix}, \quad \hat{x}_a \in R^{n_o}$$

le Equazioni (7.75), (7.76), grazie alle Equazioni (7.77), (7.78), si possono scrivere nella forma

$$\hat{x}_a(k+1) = \hat{A}_a \hat{x}_a(k) \quad (7.80)$$

$$\hat{x}_b(k+1) = \hat{A}_{ba} \hat{x}_a(k) + \hat{A}_b \hat{x}_b(k) \quad (7.81)$$

$$y(k) = \hat{C}_a \hat{x}_a(k)$$

La struttura di queste equazioni mostra con chiarezza che tutti i movimenti liberi dell'uscita generati da stati iniziali con $\hat{x}_a(0) = 0$ sono identicamente nulli, in quanto \hat{x}_b non influisce né direttamente, né tramite \hat{x}_a su y e, d'altra parte, l'esame di un qualunque transitorio di y consente per la (7.79) di determinare $\hat{x}_a(0)$. Si dirà pertanto che l'Equazione (7.80) costituisce la *parte osservabile* del sistema (7.65), (7.66) e l'Equazione (7.81) la *parte non osservabile*. Anche qui è importante notare che gli autovalori della matrice \hat{A} sono quelli di \hat{A}_a , detti autovalori della parte osservabile, e quelli di \hat{A}_b , detti autovalori della parte non osservabile.

Per concludere, la proprietà di non osservabilità di uno stato \tilde{x} implica quella di *non ricostruibilità*, consistente nell'impossibilità di distinguere lo stato finale (anziché iniziale) \tilde{x} di un sistema da quello nullo mediante l'analisi di un transitorio libero dell'uscita di qualunque durata. Tuttavia, per la classe di sistemi dinamici qui considerata, le due proprietà non coincidono, a differenza di quanto accade per i sistemi a tempo continuo (si veda il Paragrafo 3.6.3).

7.7.3 Scomposizione canonica e forma minima

Un sistema dinamico che sia insieme non completamente raggiungibile e non completamente osservabile può essere scomposto in quattro parti come specificato nel teorema seguente.

Teorema 7.18 Mediante un opportuno, non unico, cambio di variabili di stato

$$\hat{x}(k) = T_k x(k)$$

il sistema (7.65), (7.66) può essere posto nella forma

$$\hat{x}(k+1) = \hat{A}\hat{x}(k) + \hat{B}u(k) \quad (7.82)$$

$$y(k) = \hat{C}\hat{x}(k) + Du(k) \quad (7.83)$$

dove

$$\hat{x} = \begin{bmatrix} \hat{x}_a \\ \hat{x}_b \\ \hat{x}_c \\ \hat{x}_d \end{bmatrix} \quad (7.84)$$

$$\hat{A} = \begin{bmatrix} \hat{A}_a & \hat{A}_{ab} & \hat{A}_{ac} & \hat{A}_{ad} \\ 0 & \hat{A}_b & 0 & \hat{A}_{bd} \\ 0 & 0 & \hat{A}_c & \hat{A}_{cd} \\ 0 & 0 & 0 & \hat{A}_d \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \hat{B}_a \\ \hat{B}_b \\ 0 \\ 0 \end{bmatrix}, \quad \hat{C} = [0 \ \hat{C}_b \ 0 \ \hat{C}_d] \quad (7.85)$$

e le dimensioni delle sottomatrici in cui sono stati partizionati \hat{x} , \hat{A} , \hat{B} e \hat{C} sono congruenti tra loro. Inoltre il sottosistema

$$\begin{aligned} \hat{x}_b(k+1) &= \hat{A}_b \hat{x}_b(k) + \hat{B}_b u(k) \\ y(k) &= \hat{C}_b \hat{x}_b(k) + D u(k) \end{aligned}$$

è completamente raggiungibile e osservabile; il sottosistema

$$\begin{aligned} \begin{bmatrix} \hat{x}_a(k+1) \\ \hat{x}_b(k+1) \end{bmatrix} &= \begin{bmatrix} \hat{A}_a & \hat{A}_{ab} \\ 0 & \hat{A}_b \end{bmatrix} \begin{bmatrix} \hat{x}_a(k) \\ \hat{x}_b(k) \end{bmatrix} + \begin{bmatrix} \hat{B}_a \\ \hat{B}_b \end{bmatrix} u(k) \\ y(k) &= \hat{C}_b \hat{x}_b(k) + D u(k) \end{aligned}$$

ha ordine n_r ed è completamente raggiungibile; il sottosistema

$$\begin{aligned} \begin{bmatrix} \hat{x}_b(k+1) \\ \hat{x}_d(k+1) \end{bmatrix} &= \begin{bmatrix} \hat{A}_b & \hat{A}_{bd} \\ 0 & \hat{A}_d \end{bmatrix} \begin{bmatrix} \hat{x}_b(k) \\ \hat{x}_d(k) \end{bmatrix} + \begin{bmatrix} \hat{B}_b \\ 0 \end{bmatrix} u(k) \\ y(k) &= \hat{C}_b \hat{x}_b(k) + \hat{C}_d \hat{x}_d(k) + D u(k) \end{aligned}$$

ha ordine n_o ed è completamente osservabile. ■

La matrice T_K si determina con un algoritmo che, in sostanza, consiste nell'effettuare prima una scomposizione basata, per esempio, sulla proprietà di raggiungibilità e poi una seconda scomposizione connessa alla proprietà di non osservabilità.

Le Equazioni (7.82), (7.83), grazie alle (7.84), (7.85), si possono scrivere nella forma

$$\hat{x}_a(k+1) = \hat{A}_a \hat{x}_a(k) + \hat{A}_{ab} \hat{x}_b(k) + \hat{A}_{ac} \hat{x}_c(k) + \hat{A}_{ad} \hat{x}_d(k) + \hat{B}_a u(k) \quad (7.86)$$

$$\hat{x}_b(k+1) = \hat{A}_b \hat{x}_b(k) + \hat{A}_{bd} \hat{x}_d(k) + \hat{B}_b u(k) \quad (7.87)$$

$$\hat{x}_c(k+1) = \hat{A}_c \hat{x}_c(k) + \hat{A}_{cd} \hat{x}_d(k) \quad (7.88)$$

$$\hat{x}_d(k+1) = \hat{A}_d \hat{x}_d(k) \quad (7.89)$$

$$y(k) = \hat{C}_b \hat{x}_b(k) + \hat{C}_d \hat{x}_d(k) + D u(k)$$

e, sulla base di quanto affermato nel Teorema 7.18, si usa dire che l'Equazione (7.86) costituisce la parte raggiungibile e non osservabile del sistema, la (7.87) la parte raggiungibile e osservabile, la (7.88) la parte non raggiungibile e non osservabile e la (7.89) la parte non raggiungibile e osservabile. Nell'insieme, la scomposizione di un sistema in sottosistemi secondo quanto indicato dal Teorema 7.18 viene detta *scomposizione canonica*, o di *Kalman*, dal nome dell'autore del teorema stesso. Gli autovalori della matrice \hat{A} , che è triangolare a blocchi,

sono la riunione di quelli delle matrici \hat{A}_a , \hat{A}_b , \hat{A}_c e \hat{A}_d . Essi sono detti autovalori della parte raggiungibile e non osservabile, della parte raggiungibile e osservabile e così via. Analogamente, con ovvio significato dei termini, si parla di stabilità, stabilità asintotica e instabilità delle varie parti in cui il sistema è scomposto.

Delle quattro parti in cui è scomposto il sistema, l'unica affetta dalla variabile di ingresso e che influenza la variabile di uscita è la parte raggiungibile e osservabile; per la struttura delle matrici \hat{A} , \hat{B} e \hat{C} risulta poi

$$CA^k B = \hat{C}_b \hat{A}_b^k \hat{B}_b \quad (7.90)$$

L'Equazione (7.90) consente di affermare che la risposta all'impulso dell'uscita (7.36) di un sistema dipende esclusivamente dalla sua parte raggiungibile e osservabile e non dalle altre tre parti in cui il sistema si può scomporre. Allora, ricordando l'Equazione (7.37), si può enunciare il seguente risultato.

Teorema 7.19 Il movimento forzato dell'uscita del sistema (7.65), (7.66), tra tutte le quattro parti della scomposizione canonica di quest'ultimo, dipende solo dalla parte raggiungibile e osservabile ed è dato da

$$\begin{aligned} y_f(k) &= g_y(k) * u(k) = \sum_{i=0}^k g_y(k-i) u(i) = \\ &= \sum_{i=0}^{k-1} \hat{C}_b \hat{A}_b^{k-i-1} \hat{B}_b u(i) + D u(k) \quad (7.91) \end{aligned}$$

Un sistema completamente raggiungibile e osservabile si dice essere *in forma minima*, in quanto non è possibile adoperare un numero di variabili di stato minore del suo ordine per descrivere la relazione tra ingresso e uscita che esso stabilisce. Alla luce del Teorema 7.19, la risposta all'impulso dell'uscita costituisce un modello *ingresso-uscita* per i sistemi dinamici in forma minima, nel senso che la sola conoscenza di g_y consente di calcolare direttamente il movimento forzato dell'uscita tramite l'Equazione (7.91), senza la necessità di utilizzare, o anche solo definire, le variabili di stato.

Riguardo ai legami tra la stabilità asintotica e la stabilità esterna vale infine la proprietà seguente.

Teorema 7.20 Si assuma che il sistema (7.65), (7.66) sia in forma minima. Allora esso è esternamente stabile se e solo se è asintoticamente stabile. ■

7.8 Conclusioni

In questo capitolo sono state introdotte le principali caratteristiche e proprietà dei sistemi a tempo discreto, che sono molto simili a quelle dei sistemi a tempo continuo, motivate e discusse ampiamente al Capitolo 2. Quindi, si è dedicata un'attenzione specifica ai sistemi lineari e stazionari, che sono quelli di maggiore importanza nella teoria del controllo e in tutte le applicazioni in genere, e si sono trovati risultati strettamente correlati a quelli presentati nel Capitolo 3.

Per i sistemi lineari e stazionari, lo studio del movimento ha messo in evidenza che la componente libera è formata da una combinazione lineare dei modi, associati agli autovalori della matrice della dinamica. Inoltre, si è visto che la

stabilità di un generico movimento o equilibrio dipende solo dalle proprietà del movimento libero, e quindi dagli autovalori. Si sono poi formulati altri criteri di stabilità, che non richiedono il calcolo esplicito degli autovalori, ma sono basati sull'esame dei coefficienti del polinomio caratteristico. Questi risultati sono stati successivamente utilizzati per analizzare i sistemi non lineari nell'intorno di loro condizioni di equilibrio. Infine, anche per i sistemi a tempo discreto sono state introdotte le proprietà di raggiungibilità e di osservabilità e la scomposizione canonica basata su di esse.

La trattazione ha fatto riferimento a una descrizione dei sistemi dinamici nel dominio del tempo. Il Capitolo 8 sarà invece dedicato all'analisi dei sistemi lineari e stazionari effettuata tramite le trasformate Zeta e di Fourier, cioè nel dominio della variabile complessa, o della frequenza.

Esercizi

Esercizio 7.1 Si consideri il sistema

$$\begin{aligned}x(k+1) &= x^2(k) + u(k) \\y(k) &= u(k)x(k)\end{aligned}$$

Si calcoli il movimento corrispondente a $u(k) = \bar{u}$ e $x(0) = x_0$, almeno per $k = 1, 2, 3, 4$.

Esercizio 7.2 Si consideri un sistema a tempo discreto lineare, stazionario e soggetto a un ingresso fissato. Si dica se è possibile che, al variare dello stato iniziale, si ottengano due movimenti dello stato inizialmente diversi e poi identici, da un istante di tempo (finito) in avanti. Si costruisca un esempio e si rifletta sulla possibilità che il fenomeno indicato possa verificarsi in un sistema a tempo continuo.

(*Suggerimento:* si osservi che, quando $\det(A) \neq 0$, si può pensare di far funzionare il sistema "all'indietro", ricavando in maniera univoca $x(k)$ da $x(k+1)$, mentre per $\det(A) = 0$ questo non è possibile.)

Esercizio 7.3 Si consideri un sistema lineare e stazionario con matrice della dinamica

$$A = \begin{bmatrix} 1.1 & -1.2 \\ 0.8 & 1.1 \end{bmatrix}$$

Si determinino i suoi modi e quindi il movimento libero dello stato conseguente a un qualunque stato iniziale.

Esercizio 7.4 Si costruisca un sistema, di ordine minimo possibile, dotato almeno dei modi $(-1)^k, k, j^k$. Quindi si valutino le sue proprietà di stabilità.

Esercizio 7.5 Si calcoli la risposta all'impulso dell'uscita del sistema

$$\begin{aligned}x_1(k+1) &= 0.1x_1(k) + u(k) \\x_2(k+1) &= 0.2x_2(k) + 5u(k) \\y(k) &= x_1(k) + 3x_2(k)\end{aligned}$$

Si verifichi che essa è asintoticamente nulla, essendo il sistema asintoticamente stabile. Quindi, in base a essa, si calcoli la risposta dell'uscita del sistema all'ingresso $u(k) = k$.

Si calcoli poi il guadagno statico del sistema.

Esercizio 7.6 Si dimostri il Teorema 7.6.

Esercizio 7.7 Si verifichi che la trasformazione bilineare

$$z = \frac{1+s}{1-s}$$

trasforma il semipiano sinistro del piano complesso S nel cerchio di raggio unitario centrato nell'origine del piano complesso Z .

Esercizio 7.8 Si consideri il polinomio caratteristico

$$\varphi(z) = 20z^2 + \alpha z + \beta$$

e si dica se è vero che tutte le sue radici hanno modulo minore di uno qualunque valore assumano α e β negli intervalli seguenti:

$$\begin{aligned}5 &\leq \alpha \leq 20 \\1 &\leq \beta \leq 2\end{aligned}$$

Esercizio 7.9 Si verifichi la correttezza dell'Equazione (7.90).

Esercizio 7.10 Si determinino gli stati e le uscite di equilibrio del sistema

$$\begin{aligned}x_1(k+1) &= x_1(k) + (1-x_1(k))(1+x_2(k)) + u(k) \\x_2(k+1) &= x_1(k) + (1+x_1(k))(1-x_2(k)) - u(k) \\y(k) &= (x_1(k) + x_2(k))^3\end{aligned}$$

per $u(k) = \bar{u} = 0$, e si analizzi la stabilità dei primi. Quindi, si determinino i sistemi linearizzati attorno a essi.

8.1 Introduzione

In questo capitolo l'analisi in frequenza precedentemente introdotta nei Capitoli 4-6 per i sistemi a tempo continuo sarà estesa ai sistemi a tempo discreto. In particolare, per i sistemi lineari e stazionari sarà definita la funzione di trasferimento, che esprime il legame tra le trasformate Zeta dell'ingresso e dell'uscita. Nel capitolo saranno trattati i seguenti argomenti:

- la definizione della funzione di trasferimento e l'analisi delle sue principali proprietà;
- la rappresentazione e i parametri della funzione di trasferimento;
- le caratteristiche della risposta allo scalino di sistemi SISO del primo e del secondo ordine;
- la derivazione dalla risposta allo scalino di approssimanti con la struttura dei sistemi FIR;
- la realizzazione nel tempo di sistemi descritti dalla funzione di trasferimento;
- la rappresentazione con schemi a blocchi di sistemi interconnessi e le relative regole di elaborazione;
- il calcolo della risposta a un ingresso sinusoidale, la definizione di risposta in frequenza e dei suoi legami con la funzione di trasferimento, l'enunciazione del teorema fondamentale della risposta in frequenza e le estensioni al caso di ingressi esprimibili come combinazioni lineari di sinusoidi;
- lo studio delle diverse rappresentazioni grafiche della risposta in frequenza.

8.2 Funzione di trasferimento

8.2.1 Definizione e interpretazioni

Si consideri il sistema con n variabili di stato, m variabili di ingresso e p variabili di uscita

$$x(k+1) = Ax(k) + Bu(k) \quad (8.1)$$

$$y(k) = Cx(k) + Du(k) \quad (8.2)$$

e si indichino con $U(z)$, $X(z)$ e $Y(z)$ le trasformate Zeta (si veda il Paragrafo C.3) di $u(k)$, $x(k)$ e $y(k)$. Applicando la trasformata Zeta alle Equazioni (8.1), (8.2) e

ricordando le sue proprietà, si ottiene

$$zX(z) - zx(0) = AX(z) + BU(z) \quad (8.3)$$

$$Y(z) = CX(z) + DU(z) \quad (8.4)$$

da cui risulta

$$X(z) = (zI - A)^{-1}BU(z) + z(zI - A)^{-1}x(0) \quad (8.5)$$

$$Y(z) = (C(zI - A)^{-1}B + D)U(z) + zC(zI - A)^{-1}x(0) \quad (8.6)$$

Le Equazioni (8.5), (8.6) forniscono le trasformate Zeta dei movimenti dello stato e dell'uscita. In particolare, esse evidenziano le componenti libera $z(zI - A)^{-1}x(0)$ e forzata $(zI - A)^{-1}BU(z)$ del movimento dello stato, e le corrispondenti componenti $zC(zI - A)^{-1}x(0)$ e $(C(zI - A)^{-1}B + D)U(z)$ del movimento dell'uscita.

La matrice $p \times m$

$$G(z) = C(zI - A)^{-1}B + D \quad (8.7)$$

che appare nella (8.6) è detta *funzione di trasferimento* e, moltiplicata a destra per la trasformata Zeta dell'ingresso u , fornisce la trasformata Zeta dell'uscita y corrispondente a stato iniziale nullo, cioè dell'uscita forzata. Nell'ipotesi di condizioni iniziali nulle, il sistema (8.1), (8.2) può quindi essere descritto con la *rappresentazione ingresso-uscita*

$$Y(z) = G(z)U(z) \quad (8.8)$$

Nei sistemi SISO, la funzione di trasferimento (8.8) è scalare ed è data dal rapporto tra la trasformata Zeta dell'uscita forzata e quella dell'ingresso che l'ha prodotta. Inoltre, dato l'ingresso $u(k) = \text{imp}^*(k)$, con $U(z) = 1$, dalla (8.8) risulta $Y(z) = G(z)$. Pertanto, la funzione di trasferimento può essere definita in modo alternativo come la trasformata Zeta della risposta all'impulso g_y (Equazione (7.36)) del sistema (8.1), (8.2). Questa interpretazione della funzione di trasferimento è immediatamente generalizzabile al caso di sistemi MIMO.

8.2.2 Struttura della funzione di trasferimento

La definizione (8.7) di funzione di trasferimento per sistemi a tempo discreto è identica alla (4.5) relativa ai sistemi a tempo continuo pur di sostituire nella (4.5) la variabile complessa z alla variabile complessa s . Pertanto, nel caso di sistemi SISO, con considerazioni identiche a quelle sviluppate nel Paragrafo 4.2.2, si possono fare le seguenti osservazioni:

- nel caso particolare di sistemi non dinamici, descritti dalla relazione $y(k) = Du(k)$, $G(z)$ è indipendente da z e risulta pari a D ;
- per costruzione, $C(zI - A)^{-1}B$ è una funzione razionale in z con denominatore di grado n e numeratore al più di grado $n - 1$. Se il sistema (8.1), (8.2) è strettamente proprio ($D = 0$), $G(z)$ ha quindi la struttura suddetta;
- se il sistema è proprio ($D \neq 0$), la somma della funzione razionale $C(zI - A)^{-1}B$ con denominatore di grado n con una costante produce una funzione razionale con numeratore e denominatore ambedue di grado n ;

- se il numeratore e il denominatore della funzione di trasferimento hanno radici in comune, è necessario effettuare la cancellazione dei fattori comuni, ottenendo in conclusione una funzione razionale con denominatore di grado $v < n$ e numeratore di grado al più $v - 1$, per $D = 0$, o v , per $D \neq 0$. La funzione di trasferimento risultante è detta in *forma minima*.

Per quanto detto, in generale risulta

$$G(z) = \frac{N_G(z)}{D_G(z)} = \frac{\beta_v z^v + \beta_{v-1} z^{v-1} + \dots + \beta_1 z + \beta_0}{\alpha_v z^v + \alpha_{v-1} z^{v-1} + \dots + \alpha_1 z + \alpha_0} \quad (8.9)$$

dove $v \leq n$, $\beta_v = 0$ se il sistema è strettamente proprio e, senza perdita di generalità, si può assumere $\alpha_v = 1$. La differenza tra il grado del denominatore e quello del numeratore di $G(z)$ è detta *grado relativo*.

Poli e zeri Con riferimento a sistemi SISO descritti dalla (8.9), \hat{z} è uno zero di $G(z)$ se annulla il suo numeratore, cioè se $N_G(\hat{z}) = 0$, mentre \hat{z} è un polo di $G(z)$ se annulla il suo denominatore, cioè se $D_G(\hat{z}) = 0$. Sia gli zeri sia i poli sono reali o complessi coniugati a coppie, in quanto sono reali i coefficienti α_i e β_i nella (8.9). Poli e zeri, nel loro insieme, sono detti *singolarità*.

Poiché per costruzione le radici di $D_G(z) = 0$ sono anche radici di $\det(zI - A) = 0$, si può affermare che i poli sono anche autovalori, mentre un autovalore può non essere un polo qualora esso corrisponda a uno dei fattori cancellati tra numeratore e denominatore nella costruzione di $G(z)$. Inoltre, in assenza di cancellazioni, la molteplicità di un polo coincide con quella del corrispondente autovalore.

Nel caso più generale di sistemi MIMO, \hat{z} è un polo di $G(z)$ se annulla il denominatore di almeno una delle funzioni razionali che compongono la funzione di trasferimento, mentre la definizione di zero risulta più complicata e non viene qui fornita.

8.2.3 Cancellazioni e stabilità

Quando nel calcolo della funzione di trasferimento $G(z)$, data dalla (8.9), il numero v dei poli risulta inferiore all'ordine n del sistema, gli autovalori che non coincidono con i poli di $G(z)$ sono associati a parti "nascoste" del sistema che non influenzano il legame tra l'ingresso e l'uscita. Queste parti nascoste possono essere associate anche ad autovalori con modulo maggiore o uguale a uno. Per valutare se un sistema è asintoticamente stabile a partire dalla conoscenza della sua funzione di trasferimento, è quindi necessario riferirsi a uno dei casi seguenti:

- quando nel calcolo di $G(z)$ non avvengono semplificazioni, il suo denominatore coincide con il polinomio caratteristico, l'insieme dei poli coincide con quello degli autovalori e la conoscenza dei poli è sufficiente per accettare la proprietà di stabilità del sistema;
- se vi sono cancellazioni, la semplice conoscenza dei poli di $G(z)$ non consente di trarre alcuna conclusione sulla stabilità del sistema: potrebbe infatti succedere che uno o più autovalori non appartenenti all'insieme dei poli non abbiano modulo minore di uno.

In generale, la presenza di cancellazioni si può considerare dovuta a una non corretta modellizzazione del sistema, almeno per quanto riguarda la sua descrizione

esterna. Pertanto, nel seguito si ipotizzerà sempre che, qualora un sistema sia descritto in termini di funzione di trasferimento, l'insieme dei suoi poli coincida con quello dei suoi autovalori, cioè che non siano avvenute cancellazioni. Per questo motivo, se non altrimenti specificato, si supporrà che il polinomio $D_G(z)$ al denominatore di $G(z)$ coincida con il polinomio caratteristico e il suo grado sarà chiamato ordine del sistema.

Cancellazioni, raggiungibilità e osservabilità Senza perdere la generalità si può pensare che, eventualmente dopo un opportuno cambio di variabili, il sistema (8.1), (8.2) sia rappresentato nella sua scomposizione canonica specificata nel Teorema 7.18. Poiché la risposta all'impulso di un sistema dipende esclusivamente dalla sua parte raggiungibile e osservabile (Paragrafo 7.7.3) e $G(z) = \mathcal{Z}[g_y(k)]$, anche la funzione di trasferimento dipende solo da questa parte. Con riferimento al caso SISO, si può dire allora che:

- i poli di $G(z)$ sono anche gli autovalori della parte raggiungibile e osservabile del sistema (8.1), (8.2);
- il grado del polinomio $D_G(z)$, cioè il numero dei poli, coincide con l'ordine del sottosistema che costituisce la parte raggiungibile e osservabile del sistema (8.1), (8.2).

In definitiva, è possibile enunciare il seguente risultato.

Teorema 8.1 I poli del sistema (8.1), (8.2), assunto SISO, coincidono con gli autovalori della parte raggiungibile e osservabile del sistema stesso, compresa la molteplicità. ■

In base a tale risultato, si può infine concludere che gli eventuali autovalori del sistema (8.1), (8.2), che non sono poli della funzione di trasferimento $G(z)$, sono necessariamente autovalori della parte non raggiungibile e/o non osservabile del sistema stesso. Essi non sono presenti tra i poli di $G(z)$ per effetto di cancellazioni.

8.2.4 Ritardo di tempo

Nei sistemi a tempo discreto il *ritardo di tempo*, descritto dalla relazione ingresso-uscita

$$y(k) = u(k-h) \quad (8.10)$$

con h intero positivo, è un sistema dinamico a dimensione finita, lineare e stazionario. Infatti, applicando a entrambi i membri dell'Equazione (8.10) la trasformata Zeta e ricordando le sue proprietà, si ricava

$$Y(z) = z^{-h}U(z)$$

Pertanto, la funzione di trasferimento associata al ritardo di tempo h è

$$G(z) = z^{-h}$$

ed è razionale in z , con h poli nell'origine del piano complesso.

Si considerino ora sistemi SISO, lineari e stazionari, in cui l'ingresso agisce sullo stato e sull'uscita con un ritardo h e quindi descritti da

$$v(k) = u(k-h) \quad (8.11)$$

$$x(k+1) = Ax(k) + Bu(k) \quad (8.12)$$

$$y(k) = Cx(k) + Du(k) \quad (8.13)$$

o in cui la misura dell'uscita è disponibile con un ritardo h , cioè

$$x(k+1) = Ax(k) + Bu(k) \quad (8.14)$$

$$w(k) = Cx(k) + Du(k) \quad (8.15)$$

$$y(k+h) = w(k) \quad (8.16)$$

Applicando ancora la trasformata Zeta alle (8.11)-(8.13), o alle (8.14)-(8.16), si determina in entrambi i casi la funzione di trasferimento

$$G(z) = z^{-h}G'(z) \quad , \quad G'(z) = C(zI - A)^{-1}B + D$$

data dal prodotto del termine z^{-h} , che rappresenta il ritardo sull'ingresso o sull'uscita, con la funzione $G'(z)$, anch'essa razionale in z .

8.2.5 Rappresentazioni e parametri della funzione di trasferimento

Nel seguito spesso sarà conveniente rappresentare la funzione di trasferimento di un sistema SISO, o i suoi elementi se si considerano sistemi MIMO, nella seguente forma fattorizzata:

$$G(z) = \frac{\gamma \prod_i (z - z_i) \prod_j (z^2 - 2\varphi_i \cos(\xi_i)z + \varphi_i^2)}{(z-1)^g \prod_i (z - p_i) \prod_j (z^2 - 2\rho_i \cos(\theta_i)z + \rho_i^2)} \quad (8.17)$$

Malgrado l'Equazione (8.17) comprenda il caso in cui non tutte le singolarità siano reali, in essa non appaiono coefficienti complessi. Infatti, ogni coppia di singolarità complesse coniugate dà luogo a un termine di secondo grado.

I parametri che compaiono nella (8.17) vengono così chiamati:

- lo scalare γ è detto *costante di trasferimento*;
- l'intero g , che può assumere valori positivi e negativi o essere nullo, è detto *tipo*;
- gli scalari $z_i \neq 1$ e $p_i \neq 1$ sono gli *zeri* e i *poli* reali non unitari;
- gli scalari $\varphi_i > 0$ e $\rho_i > 0$ sono i *moduli* delle coppie di zeri e poli complessi coniugati e gli scalari ξ_i e θ_i i corrispondenti *argomenti*.

Guadagno Sia dato un sistema asintoticamente stabile e descritto dalla funzione di trasferimento $G(z)$, posta nella forma (8.17); per quanto in precedenza affermato, i poli hanno modulo minore di 1, cioè $g \leq 0$, $|p_i| < 1$ e $\rho_i < 1$. Si consideri inoltre il caso particolare $g = 0$. Supponendo che il sistema sia sottoposto a un ingresso costante \bar{u} , modellizzabile come un segnale a scalino con trasformata Zeta $U(z) = \bar{u}z/(z-1)$, l'uscita tende a un valore di regime \bar{y} determinabile mediante il teorema del valore finale come

$$\bar{y} = \lim_{k \rightarrow \infty} y(k) = \lim_{z \rightarrow 1} (z-1)G(z) \frac{\bar{u}z}{z-1} = G(1)\bar{u} = (C(I-A)^{-1}B + D)\bar{u} \quad (8.18)$$

Dalla (8.18) segue che il rapporto $\mu = \bar{y}/\bar{u}$ tra il valore di regime dell'uscita del sistema e il valore dell'ingresso costante che l'ha prodotta, definito come *guadagno statico* nel Paragrafo 7.4.7, è

$$\mu = G(1) = C(I-A)^{-1}B + D \quad (8.19)$$

Per quanto esposto, la costante μ definita dall'Equazione (8.19) prende il nome di *guadagno*.

Al guadagno si può comunque attribuire il significato di rapporto tra uscita e ingresso costanti anche nel caso di sistemi non asintoticamente stabili, purché di tipo $g = 0$, in quanto, dato un equilibrio caratterizzato dai valori costanti \bar{u} e \bar{y} di ingresso e uscita, risulta $\bar{y} = \mu\bar{u}$ (si veda il Paragrafo 7.4.7), dove ancora μ è definito dall'Equazione (8.19).

Infine, nel caso in cui $g \neq 0$, il coefficiente μ viene ancora chiamato guadagno, o *guadagno generalizzato*, e il suo valore è dato da

$$\mu = \lim_{z \rightarrow 1} (z - 1)^g G(z)$$

Derivatore a tempo discreto Si consideri il sistema di tipo $g = -1$ con funzione di trasferimento $G(z) = (z - 1)/z$. Risulta allora $A = 0$, $CB = -1$, $D = 1$, per cui una possibile rappresentazione di stato è

$$\begin{aligned} x(k+1) &= u(k) \\ y(k) &= -x(k) + u(k) \end{aligned}$$

che corrisponde all'equazione alle differenze

$$y(k) = u(k) - u(k-1) \quad (8.20)$$

I valori dell'uscita sono quindi ottenuti come differenza di due valori consecutivi di u . Poiché il sistema è asintoticamente stabile, è possibile applicare il teorema del valore finale per verificare che il valore di regime dell'uscita a fronte di uno scalino d'ingresso di ampiezza arbitraria è nullo, cioè $\bar{y} = 0$, come peraltro evidente dalla (8.20). Il sistema esercita quindi sull'ingresso del sistema un'azione di *derivazione a tempo discreto*. Più in generale, un sistema di tipo $g < 0$ esercita un'azione di derivazione dell'ingresso semplice se $g = -1$, o multipla se $g < -1$.

Integratore a tempo discreto Si consideri il caso in cui $G(z) = z/(z - 1)$. Risulta allora $A = 1$, $CB = 1$, e $D = 1$. Il sistema in variabili di stato può dunque essere rappresentato nella forma

$$\begin{aligned} x(k+1) &= x(k) + u(k) \\ y(k) &= x(k) + u(k) \end{aligned}$$

che corrisponde all'equazione alle differenze

$$y(k) = y(k-1) + u(k)$$

L'uscita è quindi la somma dei valori dell'ingresso; per tale motivo questo sistema viene comunemente chiamato *sommatore*, o *integratore a tempo discreto*. Più in generale, un sistema di tipo $g > 0$ esercita sull'ingresso del sistema un'azione integrale semplice se $g = 1$, o multipla se $g > 1$.

Esempio 8.1 (Seguito dell'Esempio (7.1)). Si ipotizzi che, nel modello (7.7), (7.8), (7.10) che descrive l'evoluzione temporale del prezzo di un prodotto, i parametri β , γ e δ siano costanti e assumano i seguenti valori $\beta = 0.5$, $\gamma = 0.5$, $\delta = 2$. Le matrici corrispondenti sono

$$A = \begin{bmatrix} 1 & 0.5 \\ -0.5 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0.5 \end{bmatrix}, \quad C = [0 \ 1], \quad D = [0 \ 0 \ 0]$$

Applicando la definizione (8.7), la funzione di trasferimento risultante è

$$G(z) = [G_1(z) \ G_2(z) \ G_3(z)]$$

$$G_1(z) = \frac{-0.5}{(z - 0.5)^2}, \quad G_2(z) = \frac{2z - 1.5}{(z - 0.5)^2}, \quad G_3(z) = \frac{0.5(z - 1)}{(z - 0.5)^2}$$

Il sistema è dunque asintoticamente stabile con una coppia di autovalori coincidenti in $z = 0.5$, che sono anche poli delle funzioni di trasferimento tra ogni singolo ingresso e l'uscita. Il valore $G_1(1) = -2$ mostra che a una variazione positiva a scalino di u_1 (quantità prodotta) corrisponde asintoticamente una riduzione del prezzo. Al contrario, poiché $G_2(1) = 2$, una variazione positiva a scalino di u_2 (quota venduta indipendentemente dal prezzo), produce un aumento del prezzo. Infine, per la presenza dell'azione di derivazione nella funzione di trasferimento $G_3(z)$, asintoticamente il prezzo non dipende da una variazione a scalino di u_3 (livello ottimale delle scorte).

8.3 Risposta allo scalino

Si analizza ora l'andamento della risposta a un ingresso a scalino, di ampiezza unitaria e con trasformata $U(z) = z/(z - 1)$, di un sistema SISO asintoticamente stabile con funzione di trasferimento

$$G(z) = \frac{\beta_m z^m + \beta_{m-1} z^{m-1} + \dots + \beta_0}{z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_0}, \quad m \leq n$$

8.3.1 Valore iniziale e finale

Si è già osservato che il valore di regime y_∞ dell'uscita è nullo in presenza di eventuali azioni derivate ($g < 0$) e altrimenti è pari al guadagno $\mu = G(1)$.

Il valore $y(0)$ può invece essere determinato mediante il teorema del valore iniziale

$$y(0) = \lim_{z \rightarrow \infty} \frac{\beta_m z^m + \beta_{m-1} z^{m-1} + \dots + \beta_0}{z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_0} \frac{z}{z - 1} = \begin{cases} 0, & m < n \\ \beta_m, & m = n \end{cases}$$

Ricordando che

$$\mathcal{Z}[y(k+i)] = z^i \left[Y(z) - \sum_{h=0}^{i-1} y(h)z^{-h} \right]$$

(si veda il Paragrafo C.3.2), il teorema può anche essere applicato iterativamente per dedurre il valore dell'uscita negli istanti successivi a quello iniziale. Per esempio, se $m < n$ e quindi $y(0) = 0$, risulta

$$y(1) = \lim_{z \rightarrow \infty} \frac{\beta_m z^m + \beta_{m-1} z^{m-1} + \dots + \beta_0}{z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_0} \frac{z^2}{z - 1} = \begin{cases} 0, & m < n - 1 \\ \beta_m, & m = n - 1 \end{cases}$$

Iterando il procedimento, è possibile concludere che per $m < n$ sono nulli i valori $y(0)$, $y(1)$, ..., $y(n-m-1)$. Allo stesso risultato si può pervenire mediante il metodo di antitrasformazione della lunga divisione di $Y(z)$, o anche dall'analisi dei valori $g_y(k)$ della risposta all'impulso di $G(z)$, ricordando che il k -esimo valore $h_y(k)$ della risposta allo scalino del sistema è, in base alla (7.38),

$$h_y(k) = \sum_{i=0}^k g_y(i)$$

Il numero $n - m$ dei primi valori nulli della risposta allo scalino, cioè dei primi coefficienti nulli della risposta all'impulso, viene chiamato *tempo di latenza*. Per quanto detto, esso coincide con il grado relativo di $G(z)$.

8.3.2 Andamento del transitorio

Si consideri ora la funzione di trasferimento $G(z)$ espressa nella forma

$$G(z) = \frac{N_G(z)}{D_G(z)} = \frac{N_G(z)}{\prod_{i=1}^s (z - p_i) \prod_{i=1}^h (z^2 - 2\rho_i \cos(\theta_i)z + \rho_i^2)}$$

in cui si suppone per semplicità che i poli siano distinti. Applicando lo sviluppo di Heaviside (Paragrafo C.3.3) per l'anttrasformazione di $Y(z) = zG(z)/(z-1)$, si ottiene

$$y(k) = \frac{N_G(1)}{D_G(1)} + \sum_{i=1}^s P_i p_i^k + \sum_{i=1}^h 2|Q_i| \rho_i^k \cos(\theta_i k + \arg Q_i), \quad k \geq 0 \quad (8.21)$$

dove

$$P_i = \frac{N_G(p_i)}{(p_i - 1) \left. \frac{dD_G(z)}{dz} \right|_{z=p_i}}, \quad Q_i = \frac{N_G(\rho_i e^{j\theta_i})}{(\rho_i e^{j\theta_i} - 1) \left. \frac{dD_G(z)}{dz} \right|_{z=\rho_i e^{j\theta_i}}} \quad (8.22)$$

L'Equazione (8.21) evidenzia il contributo fornito dagli elementi di $G(z)$ alla formazione dell'uscita. In particolare, quando $g = 0$ il primo termine rappresenta il guadagno del sistema; la sommatoria che compare al secondo termine è relativa al contributo dei poli reali, mentre la sommatoria al terzo termine è relativa ai poli complessi coniugati. Per l'ipotesi di asintotica stabilità, si ha $|p_i| < 1$, $\forall i$, e $\rho_i < 1$, $\forall i$; pertanto i termini dovuti sia ai poli reali sia a quelli complessi coniugati tendono a zero per $k \rightarrow \infty$ e si esauriscono tanto più velocemente quanto più sono piccoli i valori di $|p_i|$ e ρ_i .

Nel seguito si studierà la risposta allo scalino di alcuni semplici sistemi del primo e del secondo ordine, nonché quella dei sistemi FIR. Si estenderà inoltre il concetto di poli dominanti (si veda il Paragrafo 4.4.6) ai sistemi a tempo discreto e si mostrerà come ricavare dalla risposta allo scalino di un sistema non noto un'approssimante con la struttura di un sistema FIR.

8.3.3 Sistemi del primo ordine

Si consideri un sistema del primo ordine con funzione di trasferimento

$$G(z) = \frac{\mu(1-p)}{z-p}, \quad p \neq 0 \quad (8.23)$$

È immediato verificare che la risposta allo scalino del sistema è

$$y(k) = \mu(1-p^k), \quad k \geq 0 \quad (8.24)$$

Dalla (8.24) segue che $y(0) = 0$ e, poiché il sistema è asintoticamente stabile, $y_\infty = G(1) = \mu$. L'Equazione (8.24) mostra anche che se $p > 0$ la risposta è monotona, mentre se $p < 0$ l'uscita è oscillante. In ogni caso l'uscita giunge tanto più velocemente a regime quanto minore è $|p|$. L'andamento della risposta allo scalino per diversi valori di p è riportato nelle Figure 8.1 e 8.2 (in queste e nelle successive figure i raccordi rettilinei tra i valori della risposta sono disegnati solo per migliorare la leggibilità).

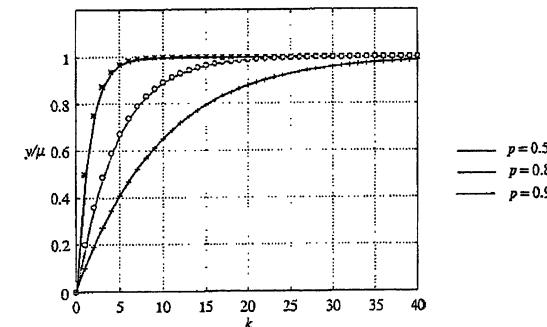


Figura 8.1
Andamento della risposta allo scalino normalizzata y/μ del sistema (8.23) per diversi valori di $0 < p < 1$.

Sempre dall'Equazione (8.24) è possibile valutare il primo istante k_{as} per cui l'uscita entra nella fascia $[(1 - 0.01\epsilon)y_\infty, (1 + 0.01\epsilon)y_\infty]$. Il valore di k_{as} è infatti il primo intero positivo per cui

$$k_{as} \geq \frac{\ln 0.01\epsilon}{\ln |p|}$$

e può essere considerato il *tempo di assestamento* al $(100 - \epsilon)\%$ del sistema. In Tabella 8.1 sono riportati i tempi di assestamento k_{a10} e k_{a1} per diversi valori di $|p|$.

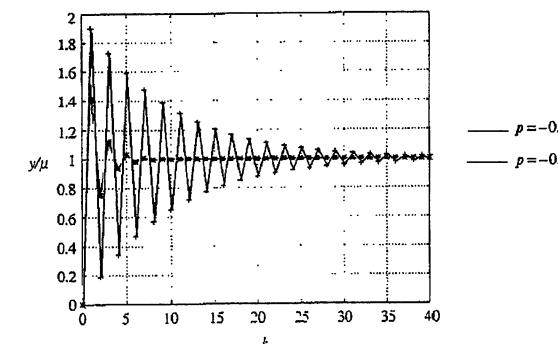


Figura 8.2
Andamento della risposta allo scalino normalizzata y/μ del sistema (8.23) per diversi valori di $-1 < p < 0$.

$ p $	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.98	0.99
k_{a10}	1	2	2	3	4	5	7	11	22	45	114	230
k_{a1}	2	3	4	6	7	10	13	21	44	90	228	459

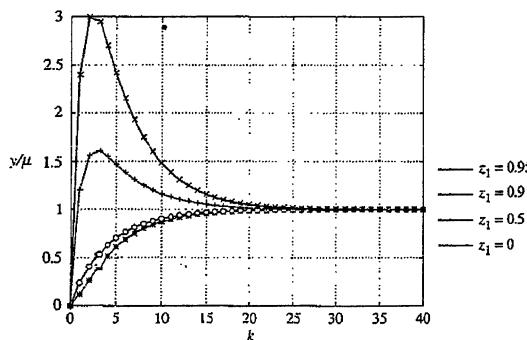
Tabella 8.1
Tempo di assestamento al 90% e al 99% della risposta allo scalino del sistema (8.23).

8.3.4 Sistemi del secondo ordine

Sistemi con poli reali Si consideri ora il sistema con due poli reali distinti, p_1 e p_2 , uno zero z_1 e descritto da

$$G(z) = \frac{\mu(1-p_1)(1-p_2)}{(1-z_1)} \frac{z-z_1}{(z-p_1)(z-p_2)} \quad (8.25)$$

Figura 8.3
Andamento della risposta allo scalino normalizzata y/μ del sistema (8.25) con $p_1 = 0.4$, $p_2 = 0.8$ e $0 \leq z_1 < 1$.



in cui μ è il guadagno. Mediante l'Equazione (8.21), è possibile determinare l'espressione analitica della risposta allo scalino, che risulta

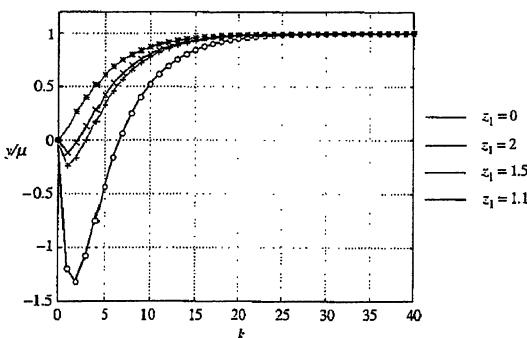
$$y(k) = \mu + P_1 p_1^k + P_2 p_2^k, \quad k \geq 0$$

dove P_1 e P_2 , ricavabili dalla (8.22), dipendono anche dal valore dello zero z_1 . Per quanto riguarda il contributo dei due poli, si possono estendere a questo caso le considerazioni relative ai sistemi del primo ordine. Per comprendere invece il contributo dello zero alla risposta allo scalino, nelle Figure 8.3 e 8.4 è riportato l'andamento dell'uscita per valori fissati e positivi di p_1 e p_2 al variare di z_1 ; in particolare vengono considerati i casi distinti: $0 \leq z_1 < 1$ e $z_1 > 1$. Dall'analisi delle Figure 8.3 e 8.4 si può concludere quanto segue:

- se $p_2 > p_1$ e $z_1 < p_2$, la risposta del sistema è tanto più veloce quanto più aumenta z_1 ;
- per $z_1 > p_2$, si manifesta una sovraelongazione tanto più pronunciata quanto più lo zero è vicino al punto 1;
- quando $z_1 > 1$, l'uscita presenta una sottoelongazione tanto più marcata quanto più lo zero è vicino al punto 1.

Infine, si può verificare che la presenza di uno zero $z_1 < 0$ non altera significativamente la risposta del sistema rispetto al caso in cui $z_1 = 0$, indipendentemente dal fatto che lo zero sia interno o esterno alla circonferenza di raggio unitario centrata nell'origine del piano complesso.

Figura 8.4
Andamento della risposta allo scalino normalizzata y/μ del sistema (8.25) con $p_1 = 0.4$, $p_2 = 0.8$ e $z_1 > 1$.



Sistemi con poli complessi coniugati L'espressione della risposta allo scalino del sistema con funzione di trasferimento

$$G(z) = \frac{\mu(1 - 2\rho \cos(\theta) + \rho^2)}{z^2 - 2\rho \cos(\theta)z + \rho^2} \quad (8.26)$$

è data da

$$y(k) = \mu + 2|\mathcal{Q}| \rho^k \cos(k\theta + \arg \mathcal{Q}), \quad k \geq 0 \quad (8.27)$$

dove il valore di \mathcal{Q} si può determinare applicando la (8.22). Dall'analisi della (8.27) si conclude che l'andamento dell'uscita è oscillante e il numero N di campioni per periodo di oscillazione dipende dall'angolo θ . Infatti, considerando valori di θ che siano divisori interi di 360° e imponendo $\cos(k\theta) = \cos((k+N)\theta)$, risulta $N = 360^\circ/\theta$. Per esempio, quando $\theta = 30^\circ$, la risposta allo scalino ha 12 campioni per periodo. L'andamento dell'uscita normalizzata y/μ per $\rho = 0.9$, $\theta = 30^\circ$ è riportato nella Figura 8.5.

Sempre dalla (8.27) si può osservare che le oscillazioni dell'uscita sono smorzate ($\rho < 1$) e modulate dal termine decrescente ρ^k ; pertanto il tempo di assestamento della risposta è tanto maggiore quanto più il modulo dei poli è prossimo a 1.

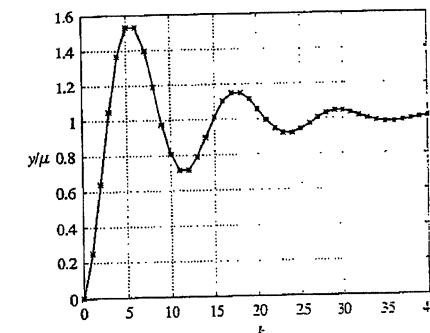


Figura 8.5
Andamento della risposta allo scalino normalizzata y/μ del sistema (8.26) con $\rho = 0.9$, $\theta = 30^\circ$.

8.3.5 Sistemi FIR

Nel Paragrafo 7.4.6 sono stati definiti i sistemi FIR come quei sistemi che hanno tutti gli autovalori in $z = 0$. La loro funzione di trasferimento è quindi

$$G(z) = \frac{\beta_m z^m + \beta_{m-1} z^{m-1} + \dots + \beta_0}{z^n} = \sum_{i=0}^m \beta_i z^{i-n} \quad (8.28)$$

con $m \leq n$. Dalla (8.28) si conclude direttamente che i valori $g_y(k)$, $k = 0, 1, \dots$ della risposta all'impulso del sistema sono

$$g_y(k) = \beta_{n-k} \quad (8.29)$$

avendo posto $\beta_i = 0$ per $i < 0$ e $i > m$. Inoltre, in base alla (7.38), sempre a partire dalla conoscenza dei coefficienti β_i , è possibile dedurre i valori $h_y(k)$,

$k = 0, 1, \dots$, della risposta allo scalino del sistema, in quanto risulta

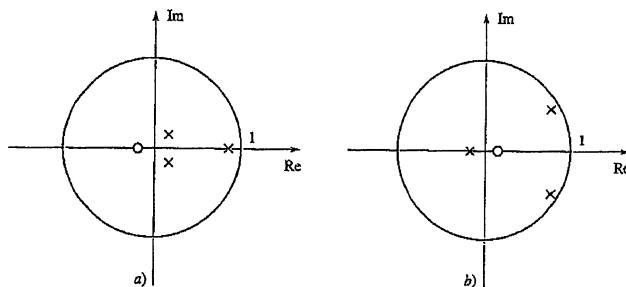
$$h_y(k) = \sum_{i=0}^k \beta_{n-i}$$

sempre con $\beta_i = 0$ per $i < 0$ e $i > m$. Quindi, la risposta allo scalino di un sistema FIR, così come la sua risposta all'impulso, si esaurisce in tempo finito e l'uscita si assesta sul valore del guadagno $\mu = G(1) = \sum_{i=0}^m \beta_i$ dopo un numero di passi pari all'ordine n .

8.3.6 Poli dominanti

Data la funzione di trasferimento $G(z)$ di un sistema SISO asintoticamente stabile, nell'ipotesi che eventuali coppie di poli e zeri prossime tra loro nel piano complesso siano state preventivamente cancellate, i poli dominanti sono i poli (reali o complessi) con modulo minore di 1, ma "decisamente" maggiore del modulo degli altri poli del sistema. Essi pertanto sono i poli di $G(z)$ più vicini alla circonferenza di raggio unitario e centrata nell'origine del piano complesso, come per esempio mostrato nella Figura 8.6.

Figura 8.6
Esempi di poli dominanti:
a) polo dominante reale;
b) poli dominanti complessi coniugati.



La risposta allo scalino di un sistema con poli dominanti è simile a quella di un sistema approssimante che possiede questi poli, un guadagno pari a quello del sistema di partenza e il numero di poli o zeri nell'origine necessario per far sì che il grado relativo, cioè la differenza tra l'ordine n del denominatore e quello m del numeratore, sia uguale a quello del sistema di partenza. Per quanto discusso in precedenza, questi poli o zeri nell'origine sono necessari per imporre che siano uguali i tempi di latenza del sistema originario e dell'approssimante, cioè che siano nulli i primi $n - m$ valori della risposta allo scalino dei due sistemi.

Esempio 8.2 Il sistema con funzione di trasferimento

$$G(z) = \frac{0.861(z - 0.5)}{(z + 0.1)(z - 0.3)(z^2 - 1.9 \cos(45^\circ)z + 0.9025)} \quad (8.30)$$

ha una coppia di poli dominanti complessi coniugati di modulo 0.95. Il corrispondente modello approssimante è

$$G_a(z) = \frac{0.56}{z(z^2 - 1.9 \cos(45^\circ)z + 0.9025)} \quad (8.31)$$

Le risposte allo scalino dei sistemi (8.30) e (8.31) sono riportate nella Figura 8.7.

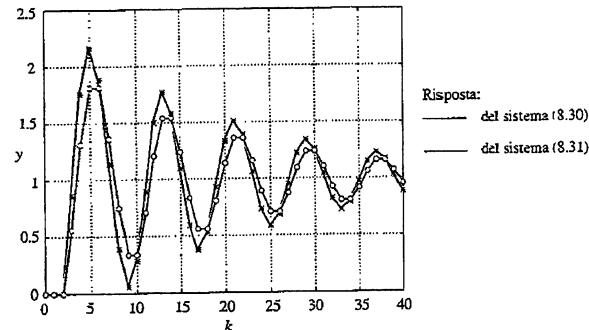


Figura 8.7
Risposta del sistema (8.30)
e del modello
approssimante (8.31).

8.3.7 Modelli approssimanti FIR

A partire dalla conoscenza della risposta all'impulso o allo scalino di un sistema incognito $G(z)$, è molto semplice ricavare un modello approssimante FIR di ordine opportuno. Come discusso nel Paragrafo 7.5.6, i valori $g_y(k)$ della risposta all'impulso di un sistema asintoticamente stabile tendono a zero per $k \rightarrow \infty$. Pertanto, se si suppone che per M sufficientemente elevato sia

$$g_y(M+i) = 0 \quad , \quad i > 0 \quad (8.32)$$

ricordando le (8.28), (8.29), la funzione di trasferimento $G(z)$ può essere scritta nella forma approssimata

$$G_a(z) = \sum_{k=0}^M g_y(k)z^{-k} = \frac{g_y(0)z^M + g_y(1)z^{M-1} + \dots + g_y(M)}{z^M} \quad (8.33)$$

che corrisponde alla funzione di trasferimento di un sistema FIR. Si ricordi ora che, per la (7.39), se $h_y(k)$ è il k -esimo valore della risposta allo scalino del sistema, il valore della risposta all'impulso $g_y(k)$ è

$$g_y(k) = h_y(k) - h_y(k-1) \quad , \quad k \geq 0 \quad (8.34)$$

In base a queste considerazioni, si può quindi concludere che la determinazione del modello approssimato di tipo FIR (8.33) a partire dalla conoscenza della risposta allo scalino è immediata sfruttando la (8.34). La funzione di trasferimento approssimato $G_a(z)$ così ottenuta deve essere moltiplicata per una costante in modo che risulti $G(1) = G_a(1)$, così che il guadagno del modello sia uguale a quello del sistema, nonostante l'approssimazione introdotta dall'Equazione (8.32).

Si noti che l'ordine M del modello approssimante può essere elevato, certamente superiore a quello del sistema di partenza. Pertanto, l'interesse per la procedura descritta risiede nella possibilità di ricavare facilmente un modello approssimato dalla sola conoscenza della risposta allo scalino del sistema. Al contrario, il metodo non deve essere in alcun caso interpretato come una tecnica di riduzione dell'ordine.

Esempio 8.3 I primi valori della risposta allo scalino del sistema

$$G(z) = \frac{0.5}{z - 0.5} \quad (8.35)$$

sono riportati nella Tabella 8.2; da questa e per mezzo della 8.34 si possono ricavare i valori della risposta all'impulso riportati nella Tabella 8.3.

Tabella 8.2
Valori della risposta allo scalino del sistema (8.35).

k	0	1	2	3	4	5	6	7	8	9	10
$h_y(k)$	0	.5	.75	.875	.938	.969	.984	.992	.996	.998	.999

Tabella 8.3
Valori della risposta all'impulso del sistema (8.35).

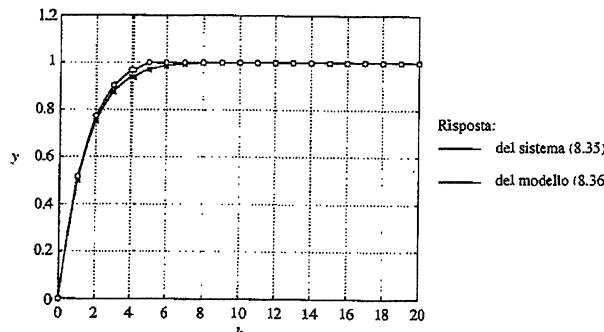
k	0	1	2	3	4	5	6	7	8	9	10
$g_y(k)$	0	.5	.25	.125	.063	.031	.015	.008	.003	.002	.001

Ipotizzando di poter trascurare i valori di $g_y(k)$ per $k > M = 6$, il modello approssimato risultante è

$$G_a(z) = 1.016 \frac{0.5z^5 + 0.25z^4 + 0.125z^3 + 0.063z^2 + 0.031z + 0.015}{z^6} \quad (8.36)$$

Le risposte allo scalino dei sistemi (8.35) e (8.36) sono riportate nella Figura 8.8.

Figura 8.8
Risposta del sistema (8.35) e del modello FIR (8.36).



8.4 Realizzazione

Dato un sistema SISO in forma minima e descritto dalla funzione di trasferimento

$$G(z) = \frac{N_G(z)}{D_G(z)} = \frac{\beta_v z^v + \beta_{v-1} z^{v-1} + \dots + \beta_1 z + \beta_0}{z^v + \alpha_{v-1} z^{v-1} + \dots + \alpha_1 z + \alpha_0} \quad (8.37)$$

il problema della *realizzazione*, cioè della determinazione di una rappresentazione di stato, può essere risolto facendo ricorso alle stesse forme canoniche introdotte nel Paragrafo 4.5 per i sistemi a tempo continuo. In particolare, posto $n = v$, la (8.37) può sempre essere scritta come

$$G(z) = \frac{\hat{\beta}_{n-1} z^{n-1} + \hat{\beta}_{n-2} z^{n-2} + \dots + \hat{\beta}_1 z + \hat{\beta}_0}{z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_1 z + \alpha_0} + \hat{\beta}_n \quad (8.38)$$

dove

$$\hat{\beta}_n = \beta_n, \quad \hat{\beta}_i = \beta_i - \alpha_i \hat{\beta}_n, \quad i = 0, \dots, n-1$$

Le matrici (A, B, C, D) del sistema (8.1), (8.2), espresso in forma canonica di raggiungibilità, sono ancora definite dalle (4.56), (4.57); inoltre il sistema risultante è sempre completamente raggiungibile, mentre la proprietà di osservabilità deriva dall'ipotesi che il sistema sia in forma minima.

Le matrici (A, B, C, D) del sistema in forma canonica di osservabilità sono definite tramite le (4.58), (4.59); inoltre il sistema è sempre completamente osservabile, mentre la proprietà di raggiungibilità scaturisce dall'ipotesi di minimalità del sistema stesso.

Anche per i sistemi a tempo discreto si può quindi formulare il seguente risultato.

Teorema 8.2 Il problema della realizzazione minima della funzione di trasferimento (8.37) ammette sempre soluzione, cioè esistono infiniti sistemi dinamici (8.1), (8.2) di ordine $n = v$ che hanno come funzione di trasferimento la (8.37). ■

Rappresentazione ingresso-uscita nel tempo A partire dalla funzione di trasferimento (8.37) è possibile ricavare direttamente una rappresentazione nel tempo di tipo ingresso-uscita sfruttando le proprietà della trasformata Zeta. A questo proposito, si consideri il sistema

$$Y(z) = \frac{\beta_v z^v + \beta_{v-1} z^{v-1} + \dots + \beta_1 z + \beta_0}{z^v + \alpha_{v-1} z^{v-1} + \dots + \alpha_1 z + \alpha_0} U(z) \quad (8.39)$$

e si moltiplich e divida il termine a destra per z^{-v} , ottenendo

$$Y(z) = \frac{\beta_v + \beta_{v-1} z^{-1} + \dots + \beta_1 z^{-v+1} + \beta_0 z^{-v}}{1 + \alpha_{v-1} z^{-1} + \dots + \alpha_1 z^{-v+1} + \alpha_0 z^{-v}} U(z)$$

da cui

$$Y(z) + \alpha_{v-1} z^{-1} Y(z) + \dots + \alpha_1 z^{-v+1} Y(z) + \alpha_0 z^{-v} Y(z) = \\ = \beta_v U(z) + \beta_{v-1} z^{-1} U(z) + \dots + \beta_1 z^{-v+1} U(z) + \beta_0 z^{-v} U(z) \quad (8.40)$$

Per le proprietà della trasformata Zeta, risulta $z^{-i} Y(z) = Z[y(k-i)]$ e $z^{-j} U(z) = Z[u(k-j)]$; pertanto l'Equazione (8.40) può essere riscritta nella forma

$$y(k) = -\alpha_{v-1} y(k-1) - \dots - \alpha_1 y(k-v+1) - \alpha_0 y(k-v) + \\ + \beta_v u(k) + \beta_{v-1} u(k-1) + \dots + \beta_1 u(k-v+1) + \beta_0 u(k-v) \quad (8.41)$$

La (8.41) è un'equazione alle differenze, direttamente ricavabile dalla funzione di trasferimento, che può essere utilizzata per la determinazione del movimento formato a partire dalla conoscenza di $G(z)$ senza fare ricorso a una rappresentazione di stato.

8.5 Schemi a blocchi

Nello studio dei sistemi a tempo discreto costituiti da più sottosistemi SISO variamente interconnessi, è spesso conveniente ricorrere alla rappresentazione grafica basata sugli schemi a blocchi, introdotta nel Capitolo 5 per i sistemi a tempo continuo. I simboli impiegati (frecce, blocchi, nodi sommatori e punti di diramazione) hanno lo stesso significato visto nel Paragrafo 5.2, così come non cambiano le regole di elaborazione:

Figura 8.9
Sistemi in serie.

- dati i due sistemi in serie di Figura 8.9 e le rispettive funzioni di trasferimento $G_a(z)$ e $G_b(z)$, la funzione di trasferimento risultante è $G(z) = G_a(z)G_b(z)$;

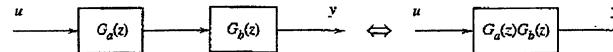
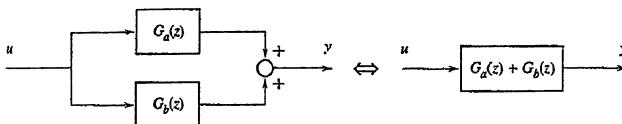


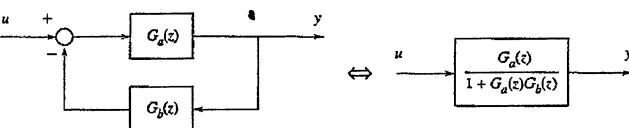
Figura 8.10
Sistemi in parallelo.

- dati i due sistemi in parallelo di Figura 8.10, le cui funzioni di trasferimento sono $G_a(z)$ e $G_b(z)$, il sistema complessivo è descritto da $G(z) = G_a(z) + G_b(z)$;



- il sistema di Figura 8.11 con retroazione negativa, funzione di trasferimento in linea di andata $G_a(z)$ e funzione di linea di retroazione $G_b(z)$, è descritto da

$$G(z) = \frac{Y(z)}{U(z)} = \frac{G_a(z)}{1 + G_a(z)G_b(z)}$$



Con le opportune modifiche è infine possibile estendere al caso in esame tutte le considerazioni riportate nei Paragrafi 5.3 e 5.4 sulle regole di elaborazione degli schemi a blocchi e sulla stabilità dei sistemi interconnessi. Poiché tali estensioni non comportano alcuna difficoltà concettuale, esse vengono lasciate al lettore.

8.6 Risposta in frequenza

In questo paragrafo è introdotta la nozione di *risposta in frequenza*, che consente di determinare il comportamento del sistema a fronte di ingressi sinusoidali o esprimibili come combinazione lineare di sinusoidi.

8.6.1 Calcolo dell'uscita sinusoidale

Si consideri il sistema SISO di ordine n descritto da

$$x(k+1) = Ax(k) + Bu(k) \quad (8.42)$$

$$y(k) = Cx(k) + Du(k) \quad (8.43)$$

e la funzione di trasferimento associata

$$G(z) = C(zI - A)^{-1}B + D$$

Supponendo che il sistema (8.42), (8.43) sia asintoticamente stabile, si voglia determinare la sua risposta a un ingresso sinusoidale del tipo

$$u(k) = U \sin(\theta k), \quad k \geq 0$$

Ricordando la trasformata Zeta della sinusode, risulta

$$Y(z) = G(z) \frac{z \sin(\theta) U}{(z - e^{j\theta})(z - e^{-j\theta})}$$

e l'uscita $y(k)$ può essere ottenuta antitrasformando tale espressione mediante lo sviluppo di Heaviside. Per esempio, se $G(z)$ possiede solo poli reali distinti, si ottiene

$$y(k) = Z^{-1} \left[\underbrace{\sum_{i=1}^n \frac{P_i}{z + p_i}}_{Y_1(z)} + \underbrace{\frac{Q}{z - e^{j\theta}} + \frac{\bar{Q}}{z - e^{-j\theta}}}_{Y_2(z)} \right] = y_1(k) + y_2(k) \quad (8.44)$$

dove P_i sono opportune costanti,

$$Q = G(e^{j\theta}) \frac{e^{j\theta} U}{2j} \quad (8.45)$$

e \bar{Q} è il complesso coniugato di Q . Per l'ipotesi di asintotica stabilità, $y_1(k)$ tende asintoticamente a zero per $k \rightarrow \infty$; inoltre con semplici passaggi si ottiene

$$y_2(k) = |G(e^{j\theta})| U \sin(\theta k + \arg G(e^{j\theta})) \quad (8.46)$$

In definitiva, l'uscita converge verso una sinusode che ha la stessa pulsazione della sinusode d'ingresso, ampiezza $Y = |G(e^{j\theta})| U$ e sfasamento $\varphi = \arg G(e^{j\theta})$. Ciò avviene per qualunque condizione iniziale, dato che l'effetto sull'uscita di un arbitrario stato iniziale $x(0)$ tende comunque ad annullarsi per la proprietà di asintotica stabilità.

È quindi possibile formulare il seguente *teorema fondamentale della risposta in frequenza*.

Teorema 8.3 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(z)$ l'ingresso sinusoidale

$$u(k) = U \sin(\theta_0 k)$$

l'uscita a transitorio esaurito assume l'andamento

$$\bar{y}(k) = |G(e^{j\theta_0})| U \sin(\theta_0 k + \arg G(e^{j\theta_0}))$$

indipendentemente dallo stato iniziale. ■

In base a questo risultato, in risposta a un ingresso sinusoidale di pulsazione θ_0 , l'uscita a regime può risultare nulla solo se $|G(e^{j\theta_0})| = 0$, cioè solo se la funzione di trasferimento possiede una coppia di zeri complessi coniugati in $e^{\pm j\theta_0}$.

La velocità con cui l'uscita si assesta sull'andamento asintotico dipende dalla dinamica propria del sistema e può essere valutata in base al tempo di assestamento introdotto al Paragrafo 8.3.3 per i sistemi del primo ordine.

8.6.2 Risposta in frequenza: definizione e proprietà

La funzione complessa

$$G(e^{j\theta}) = C(e^{j\theta}I - A)^{-1}B + D$$

definita per valori di $\theta \in [0, \pi]$ tali che $e^{j\theta}$ non sia un polo di $G(z)$, viene chiamata *risposta in frequenza* associata al sistema (8.1), (8.2). Formalmente essa coincide con la restrizione della funzione di trasferimento $G(z)$ alla semicirconferenza definita dai punti del piano complesso a parte immaginaria positiva o nulla e modulo unitario, escluso al più un numero finito di punti corrispondenti agli eventuali poli di $G(z)$ su questa semicirconferenza. D'altra parte, per la teoria delle funzioni analitiche, che comprendono come sottoinsieme le funzioni razionali di variabile complessa, a una risposta in frequenza $G(e^{j\theta})$, $\theta \in [0, \pi]$, corrisponde una singola funzione di trasferimento $G(z)$. La corrispondenza tra le due funzioni è dunque bivoca e l'informazione in esse contenuta riguardo al sistema (8.1), (8.2) è identica.

Inoltre, è facile verificare che $G(e^{j\theta})$ è una funzione periodica di periodo 2π e risulta

$$G(e^{-j\theta}) = \bar{G}(e^{j\theta}) \quad (8.47)$$

Quindi la conoscenza della risposta in frequenza permette di calcolare $G(e^{j\theta})$ anche per $\theta \in (-\pi, 0)$.

Il concetto di risposta in frequenza, qui introdotto per sistemi SISO, può essere esteso anche a sistemi multivariabili purché lineari e stazionari.

8.6.3 Segnali dotati di sviluppo di Fourier

Grazie al principio di sovrapposizione degli effetti, è possibile estendere il risultato enunciato nel Teorema 8.3 anche al caso in cui la funzione di ingresso appartenga alla classe dei segnali periodici dotati di sviluppo di Fourier, o alla classe dei segnali non periodici dotati di trasformata di Fourier (si vedano i Paragrafi C.4 e C.5).

Si supponga di applicare al sistema (8.1), (8.2) un ingresso u periodico di periodo N , con N intero positivo, cioè tale per cui $u(k+N) = u(k)$, $\forall k$. Un segnale di questo tipo (Paragrafo C.4) è rappresentabile attraverso il suo sviluppo di Fourier come

$$u(k) = \sum_N U_n e^{jn\theta_0 k}, \quad \theta_0 = 2\pi/N \quad (8.48)$$

dove il simbolo \sum rappresenta la somma estesa a un qualunque intervallo di valori di n di lunghezza N . I coefficienti complessi U_n costituiscono lo spettro del segnale u e rappresentano il contributo delle varie armoniche alla formazione del segnale stesso. In base all'analisi precedente, è possibile dimostrare il seguente risultato.

Teorema 8.4 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(z)$ l'ingresso periodico (8.48), l'uscita a transitorio esaurito assume l'andamento

$$\tilde{y}(k) = \sum_N Y_n e^{jn\theta_0 k}$$

con

$$Y_n = G(e^{jn\theta_0})U_n \quad (8.49)$$

indipendentemente dallo stato iniziale. ■

Quindi, un ingresso periodico genera un'uscita periodica, con lo stesso periodo dell'ingresso, il cui spettro $\{Y_n\}$ è legato allo spettro $\{U_n\}$ dell'ingresso attraverso la relazione (8.49), dove $\theta_0 = 2\pi/N$ è la pulsazione fondamentale. In altre parole, la n -esima armonica presente nell'ingresso subisce un'amplificazione pari al fattore $|G(e^{jn\theta_0})|$ e uno sfasamento angolare pari ad $\arg G(e^{jn\theta_0})$. Per la proprietà di "simmetria" (8.47) e per la sua periodicità, la conoscenza della risposta in frequenza $G(e^{j\theta})$ consente dunque di calcolare l'uscita periodica (asintotica) associata a un generico ingresso di periodo N .

In particolare, un'armonica di pulsazione $\bar{\theta} > 0$ può essere presente nel movimento periodico dell'uscita solo se è presente nell'ingresso, cioè se $\bar{\theta} = n\theta_0$, per qualche n intero, e se $U_n \neq 0$. D'altra parte, se $G(z)$ ha una coppia di zeri $z = e^{\pm j\bar{\theta}}$, allora tale armonica viene "bloccata" dal sistema.

8.6.4 Segnali dotati di trasformata di Fourier

Si consideri ora il caso di un ingresso u dotato di trasformata di Fourier $U(e^{j\theta}) = \mathcal{F}[u(k)]$ (si veda il Paragrafo C.5). È ben noto che la trasformata di Fourier costituisce una rappresentazione in frequenza del segnale, nel senso che si può scrivere

$$u(k) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} U(e^{j\theta}) e^{j\theta k} d\theta \quad (8.50)$$

e quindi la funzione u è scomposta in un'infinità non numerabile di armoniche, con pulsazioni che coprono l'intero intervallo $[-\pi, \pi]$, ognuna moltiplicata per il coefficiente $U(e^{j\theta})$.

La presenza nell'Equazione (8.50) dell'integrale, al posto della sommatoria che compare nella (8.48), non modifica concettualmente l'analisi svolta nel caso precedente. In virtù della proprietà di linearità dell'operazione di integrazione, si può infatti fare ricorso ancora al principio di sovrapposizione degli effetti per giungere alla conclusione riportata nel seguente teorema.

Teorema 8.5 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(z)$ l'ingresso (8.50), a transitorio esaurito l'uscita si assesta sull'andamento

$$\tilde{y}(k) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} Y(e^{j\theta}) e^{j\theta k} d\theta$$

con

$$Y(e^{j\theta}) = G(e^{j\theta})U(e^{j\theta}) \quad (8.51)$$

indipendentemente dallo stato iniziale. ■

Anche a proposito di questo risultato si possono fare commenti analoghi a quelli che seguono il Teorema 8.4. In particolare, il movimento asintotico dell'uscita non contiene armoniche che non siano presenti nel segnale di ingresso, ma alcune di queste possono risultare completamente cancellate dal sistema se esso possiede zeri di modulo unitario.

Particolaramente interessante è il caso di sistemi alimentati da un ingresso nullo prima dell'istante $k = 0$ (come discusso nel Paragrafo C.5.3 un segnale di questo tipo è anche dotato della trasformata Zeta). In questo caso la componente forzata dell'uscita è anch'essa nulla per $k < 0$ e coincide con il movimento asintotico \tilde{y} per $k \geq 0$, dato che l'effetto di un arbitrario stato iniziale $x(k_0)$ è nullo in

ogni istante finito quando $k_0 \rightarrow -\infty$, in virtù dell'asintotica stabilità del sistema. Quindi la (8.51) rappresenta la trasformata di Fourier della componente forzata dell'uscita, cioè quella che si ottiene partendo dallo stato iniziale $x(0) = 0$.

Tutto ciò consente di dare una nuova interpretazione alla risposta in frequenza $G(e^{j\theta})$. Per sistemi asintoticamente stabili, la formula (8.51) mette in luce che $G(e^{j\theta})$ rappresenta il rapporto tra gli spettri dell'uscita asintotica e dell'ingresso per tutti i valori di θ per cui non sia nullo lo spettro dell'ingresso $U(e^{j\theta})$. Se l'ingresso è nullo prima dell'istante $k = 0$, allora $G(e^{j\theta})$ rappresenta il rapporto tra gli spettri dell'uscita e dell'ingresso a partire da stato iniziale nullo.

8.7 Complementi

I risultati appena introdotti sulla risposta in frequenza possono essere estesi in varie direzioni. Per esempio, si possono applicare gli stessi ragionamenti al caso della risposta a un ingresso che, per analogia con il caso a tempo continuo, sarà detto di tipo "esponenziale", ed è possibile trarre qualche conclusione anche quando il sistema in esame non è asintoticamente stabile.

8.7.1 Risposta esponenziale

Si voglia calcolare il movimento di un sistema SISO, asintoticamente stabile e con funzione di trasferimento $G(z)$, in risposta a un ingresso del tipo

$$u(k) = U\lambda^k, \quad k \geq 0 \quad (8.52)$$

supponendo che λ non coincida con alcun polo del sistema.

Con considerazioni analoghe a quelle del Paragrafo 8.6, si può dimostrare che vale il risultato seguente.

Teorema 8.6 Se si applica a un sistema lineare asintoticamente stabile con funzione di trasferimento $G(z)$ l'ingresso

$$u(k) = U\lambda^k$$

l'uscita a transitorio esaurito assume l'andamento

$$\hat{y}(k) = G(\lambda)U\lambda^k \quad (8.53)$$

indipendentemente dallo stato iniziale. ■

Il risultato precedente consente di illustrare una fondamentale proprietà degli zeri di un sistema dinamico. Infatti, alla luce di quanto appena visto, applicando l'ingresso (8.52) con λ coincidente con uno zero del sistema, l'uscita tende asintoticamente ad annullarsi per qualsiasi valore dello stato iniziale. Questa importante proprietà degli zeri viene di solito chiamata *proprietà bloccante*.

8.7.2 Il caso di sistemi instabili

Quando il sistema non è asintoticamente stabile, i risultati dei paragrafi precedenti non sono più validi. Per esempio, se si applica un ingresso sinusoidale a un sistema instabile non è più vero in generale che l'uscita tende a una sinusode, visto che il termine $y_1(k)$ nella formula (8.44) del movimento forzato non converge a zero. Ciononostante si può dimostrare che, con un'opportuna scelta dello stato iniziale,

il termine $y_1(k)$ può essere esattamente controbilanciato dalla componente libera di un particolare movimento.

Per discutere più in dettaglio la questione, ci si può chiedere se, alimentando con l'ingresso (8.52) il sistema (8.1), (8.2), supposto non asintoticamente stabile, esiste un valore di $x(0)$ tale che il movimento dello stato sia del tipo $x(k) = x(0)\lambda^k$. Usando l'Equazione (8.1) deve essere

$$x(0)\lambda^{k+1} = Ax(0)\lambda^k + BU\lambda^k$$

ovvero

$$(\lambda I - A)x(0) = BU$$

Dunque il valore cercato di $x(0)$ esiste ed è unico se e solo se λ non è un autovalore di A e, sotto tale ipotesi, il movimento dello stato associato all'ingresso (8.52) è allo stato iniziale

$$x(0) = (\lambda I - A)^{-1}BU$$

è dato da

$$x(k) = (\lambda I - A)^{-1}BU\lambda^k, \quad k \geq 0 \quad (8.54)$$

In corrispondenza, l'uscita y è

$$y(k) = [C(\lambda I - A)^{-1}B + D]U\lambda^k = G(\lambda)U\lambda^k, \quad k \geq 0 \quad (8.55)$$

e quindi ha un andamento dello stesso tipo di quello dell'ingresso, ma moltiplicato per il fattore $G(\lambda)$.

In definitiva, quando manca l'ipotesi di asintotica stabilità, solo scegliendo opportunamente lo stato iniziale il movimento dell'uscita coincide con la funzione (8.53).

Ad analoghe conclusioni si perviene quando si considera il caso della risposta alla sinusode. Basta infatti osservare che $\sin(\theta_0 k) = \text{Im}(e^{j\theta_0 k})$ e che i risultati precedenti continuano formalmente a valere anche qualora nella (8.52) si ponga $\lambda = e^{j\theta}$. Grazie alla linearità si possono poi ricavare risultati simili anche nel caso di segnali in ingresso dotati di serie o trasformata di Fourier. Valgono dunque i seguenti teoremi.

Teorema 8.7 Si supponga che il sistema (8.1), (8.2) con funzione di trasferimento $G(z)$ non abbia autovalori in $e^{-j\pi\theta_0}$ e si applichi a esso l'ingresso

$$u(k) = U \sin(\theta_0 k), \quad k \geq 0$$

Allora esiste uno stato iniziale $x(0)$ per cui l'uscita è sinusoidale e vale

$$y(k) = |G(e^{j\theta_0})|U \sin(\theta_0 k + \arg G(e^{j\theta_0})) \quad k \geq 0$$

Teorema 8.8 Si supponga che il sistema (8.1), (8.2) con funzione di trasferimento $G(z)$ non abbia autovalori in $e^{jn\theta_0}$ (con n intero) e si applichi a esso l'ingresso

$$u(k) = \sum_n U_n e^{jn\theta_0 k}, \quad k \geq 0$$

Allora esiste uno stato iniziale $x(0)$ per cui l'uscita è dotata di sviluppo di Fourier e vale

$$y(k) = \sum_n G(e^{jn\theta_0})U_n e^{jn\theta_0 k}, \quad k \geq 0$$

Teorema 8.9 Si supponga che il sistema (8.1), (8.2) con funzione di trasferimento $G(z)$ non abbia autovalori di modulo unitario e si applichi a esso l'ingresso

$$u(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} U(e^{j\theta}) e^{j\theta k} d\theta \quad , \quad k \geq 0$$

Allora esiste uno stato iniziale $x(0)$ per cui l'uscita è dotata di trasformata di Fourier e vale

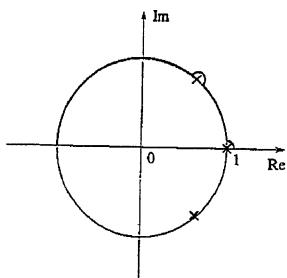
$$y(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\theta}) U(e^{j\theta}) e^{j\theta k} d\theta \quad , \quad k \geq 0$$

8.8 Diagrammi di Bode e polari

La risposta in frequenza $G(e^{j\theta})$, $\theta \in [0, \pi]$, può essere rappresentata graficamente tramite i *diagrammi di Bode* di modulo e fase a essa associati. Tuttavia, a differenza del caso dei sistemi a tempo continuo, questa rappresentazione non è molto utilizzata poiché nel tracciamento dei diagrammi non è facile determinare, neppure in modo qualitativo, il singolo contributo dei poli e degli zeri. Per questa ragione non è possibile definire diagrammi approssimati, o asintotici, e, anche in casi semplici, per il tracciamento è richiesto l'uso di adeguati strumenti di calcolo.

Più utilizzato, nell'analisi delle proprietà di stabilità dei sistemi retroazionati (si veda il Paragrafo 18.4) è il *diagramma polare* associato a $G(e^{j\theta})$, cioè il luogo dei punti $G(e^{j\theta})$ con $\theta \in [0, \pi]$ e tale che $e^{j\theta}$ non sia polo di $G(z)$. Questo luogo, costituito da una curva giacente nel piano complesso e punteggiata in θ , rappresenta l'immagine attraverso $G(z)$ della semicirconferenza di raggio unitario, centrata nell'origine e a parte immaginaria positiva o nulla. Se $G(z)$ ha dei poli su questa circonferenza, la definizione del diagramma polare va modificata considerando delle indentature infinitesime e rivolte verso l'esterno della circonferenza che consentono di aggirare le singolarità in questione, come mostrato nella Figura 8.12.

Figura 8.12
Luogo dei punti modificato per il tracciamento del diagramma polare.



Il tracciamento preciso del diagramma polare richiede la valutazione del numero complesso $G(e^{j\theta})$ per un numero sufficientemente elevato di valori di θ . Anche questa determinazione è certamente laboriosa se compiuta manualmente, mentre non presenta difficoltà se si dispone di una macchina di calcolo, anche di modesta potenzialità. Tuttavia, per ricavare un andamento qualitativo del diagramma polare è possibile in casi semplici fare ricorso alla costruzione grafica descritta nel seguente.

Si consideri la funzione di trasferimento $G(z)$ data dalla (8.17), la cui risposta in frequenza è

$$G(e^{j\theta}) = \frac{\gamma \prod_{i=1}^q (e^{j\theta} - z_i) \prod_{i=1}^r (e^{j2\theta} - 2\varphi_i \cos(\zeta_i) e^{j\theta} + \varphi_i^2)}{(e^{j\theta} - 1)^s \prod_{i=1}^s (e^{j\theta} - p_i) \prod_{i=1}^h (e^{j2\theta} - 2\rho_i \cos(\theta_i) e^{j\theta} + \rho_i^2)}$$

con $\theta \in [0, \pi]$. È allora immediato verificare che risulta

$$G(e^{j\theta}) = M e^{j\gamma} \quad (8.56)$$

dove

$$M = \frac{\gamma \prod_{i=1}^q M_{qi} \prod_{i=1}^{2r} M_{ri}}{M_0^s \prod_{i=1}^s M_{si} \prod_{i=1}^{2h} M_{hi}}$$

e

$$\gamma = f_r + \sum_{i=1}^q f_{qi} + \sum_{i=1}^{2r} f_{ri} - g f_0 - \sum_{i=1}^s f_{si} - \sum_{i=1}^{2h} f_{hi}$$

In queste formule M_{qi} , M_{ri} , M_0 , M_{si} , M_{hi} sono i moduli e f_{qi} , f_{ri} , f_0 , f_{si} , f_{hi} sono le fasi dei vettori spiccati dalle corrispondenti singolarità di $G(z)$ e con il secondo estremo in $e^{j\theta}$ (Figura 8.13). Inoltre, f_r è l'argomento della costante di trasferimento γ . Analizzando al crescere di θ da zero a π il modulo e la fase dei singoli vettori, si può quindi determinare per punti l'andamento del numero complesso $G(e^{j\theta})$.

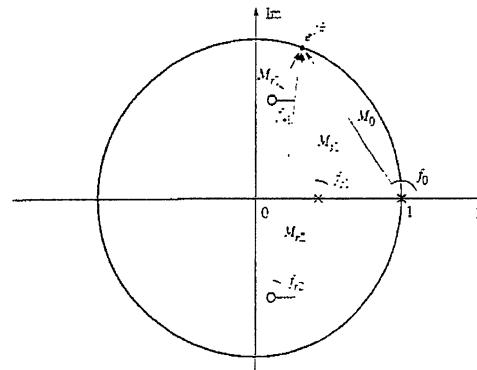


Figura 8.13
Valutazione di $G(e^{j\theta})$ per via grafica.

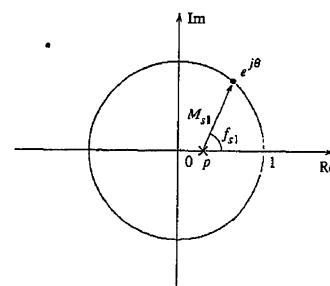
Esempio 8.4 Per il tracciamento del diagramma polare della risposta in frequenza associata alla funzione di trasferimento

$$G(z) = \frac{1-p}{z-p} \quad 0 < p < 1 \quad (8.57)$$

è necessario valutare il modulo M_{s1} e la fase f_{s1} del vettore mostrato nella Figura 8.14 al variare del punto corrente sulla semicirconferenza. In particolare risulta

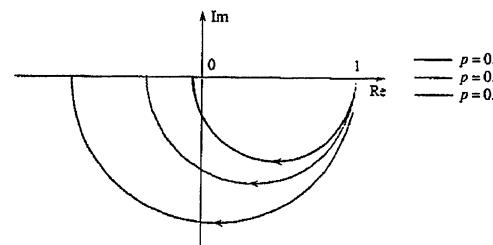
$$|G(e^{j\theta})| = \frac{1-p}{M_{s1}} \quad \arg G(e^{j\theta}) = -f_{s1}$$

Figura 8.14
Costruzione grafica per la valutazione del diagramma polare della risposta in frequenza associata alla (8.57).



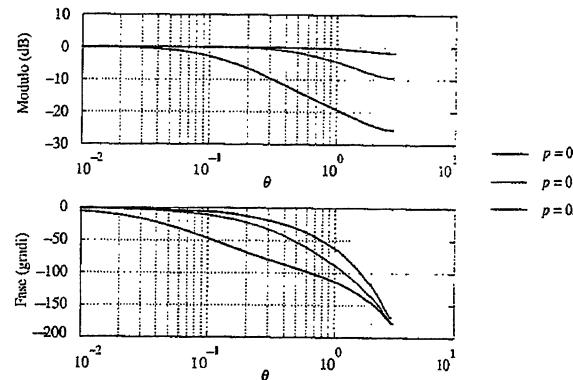
Dalla costruzione grafica di Figura 8.14, è immediato verificare che per $\theta = 0$ si ha $M_{s1} = 1 - p$ e $f_{s1} = 0^\circ$ e quindi $|G(e^{j0})| = 1$, $\arg G(e^{j0}) = 0^\circ$. Questo punto potrebbe essere dedotto valutando direttamente $G(1)$. Al crescere dell'angolo θ , sia il modulo M_{s1} sia l'argomento f_{s1} crescono monotonicamente, così che $|G(e^{j\theta})|$ e $\arg G(e^{j\theta})$ sono decrescenti. Per $\theta = \pi$ si ha $M_{s1} = 1 + p$ e $f_{s1} = 180^\circ$, cioè $|G(e^{j\pi})| = (1 - p)/(1 + p)$ e $\arg G(e^{j\pi}) = -180^\circ$, da cui $G(e^{j\pi}) = (p - 1)/(p + 1)$. Questo punto si potrebbe ricavare direttamente valutando $G(-1)$. Il diagramma polare complessivo associato alla (8.57) è riportato nella Figura 8.15 per diversi valori di p , mentre nella Figura 8.16 sono riportati i corrispondenti diagrammi di Bode.

Figura 8.15
Diagramma polare della risposta in frequenza associata alla (8.57) per $p = 0.1$, $p = 0.5$, $p = 0.9$.



L'analisi del diagramma di Bode del modulo di Figura 8.16 mostra come sia possibile considerare il sistema (8.57) come un filtro passa-basso con banda passante decrescente all'aumentare di p . Infatti, per il Teorema 8.5, applicando all'ingresso del sistema un segnale dotato di trasformata di Fourier $U(e^{j\theta})$, le armoniche in bassa frequenza vengono riprodotte in uscita praticamente inalterate, mentre l'ampiezza di quelle relative a frequenze via via crescenti è sempre più attenuata.

Figura 8.16
Diagrammi di Bode della risposta in frequenza associata alla (8.57) per $p = 0.1$, $p = 0.5$, $p = 0.9$.



Si noti infine che, in questo caso semplice, la valutazione analitica del numero complesso $G(e^{j\theta})$ al variare di θ non comporta difficoltà, in quanto risulta

$$|G(e^{j\theta})| = \frac{1-p}{\sqrt{1+p^2-2p \cos \theta}} \quad \arg G(e^{j\theta}) = -\arg(\cos \theta - p + j \sin \theta)$$

Esempio 8.5 (Seguito dell'Esempio 8.3) Per valutare la bontà dell'approssimante $G_c(z)$ (Equazione (8.36)) del sistema con funzione di trasferimento $G(z)$ data dalla (8.35), è possibile tracciare, con opportuni mezzi di calcolo, i diagrammi polari associati alle rispettive risposte in frequenza. Questi diagrammi sono riportati nella Figura 8.17, che mostra come anche nel dominio della frequenza l'approssimazione proposta sia soddisfacente.

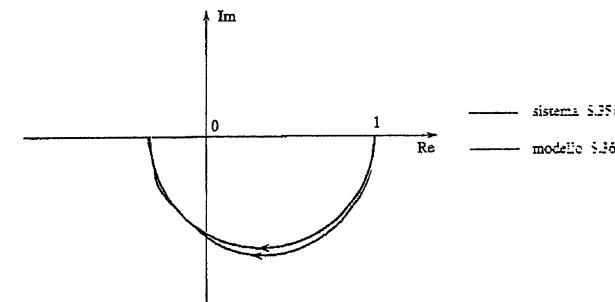


Figura 8.17
Diagrammi polari delle risposte in frequenza associate al sistema (8.35) e al modello (8.36).

8.9 Conclusioni

In questo capitolo sono state estese ai sistemi a tempo discreto lineari e stazionari le principali definizioni e proprietà relative all'analisi nel dominio della frequenza dei sistemi lineari e stazionari a tempo continuo. In particolare, in analogia con quanto fatto nei Capitoli 4-6 per i sistemi a tempo continuo, si è mostrato come sia possibile definire la funzione di trasferimento, la risposta in frequenza e le sue rappresentazioni mediante diagrammi di Bode e polari. Per alcuni sistemi del primo o del secondo ordine asintoticamente stabili, si è anche valutato l'andamento della risposta allo scalino in funzione della posizione delle singolarità. Successivamente, si è mostrato come ricavare sistemi approssimanti di tipo FIR a partire dalla risposta allo scalino di sistemi SISO asintoticamente stabili. Il problema della realizzazione di sistemi SISO è stato risolto con il ricorso a forme canoniche di struttura identica a quelle introdotte nel Capitolo 4. Si è anche mostrato che dai parametri della funzione di trasferimento è possibile dedurre con facilità una rappresentazione nel tempo ingresso-uscita che consente un'immediata valutazione del movimento forzato dell'uscita a fronte di un ingresso qualsiasi.

Esercizi

Esercizio 8.1 Si mostri che i coefficienti della risposta all'impulso di un sistema SISO possono essere ricavati direttamente dalla sua funzione di trasferimento $G(z)$ con il metodo della lunga divisione (Paragrafo C.3.3).

Suggerimento: per $u(k) = \text{imp}^*(k)$ sia $y(k) = g_y(k)$ l'uscita corrispondente e si verifichi che

$$Y(z) = G(z) = g_y(0) + g_y(1)z^{-1} + g_y(2)z^{-2} + \dots = \sum_{k=0}^{\infty} g_y(k)z^{-k}$$

Esercizio 8.2 Si verifichi che un sistema FIR con funzione di trasferimento (8.28) ammette le seguenti rappresentazioni nel tempo:

$$\begin{aligned} y(k) &= \sum_{i=0}^m \beta_i u(k+i-n) \\ y(k) &= \sum_{j=0}^n g_y(j) u(k-j) \end{aligned} \quad (8.58)$$

Esercizio 8.3 Definendo $\Delta u(k) = u(k) - u(k-1)$ e ricordando il legame che intercorre tra i coefficienti delle risposte all'impulso (g_y) e allo scalino (h_y), si verifichi che la (8.58) può essere scritta nella forma

$$y(k) = \sum_{j=0}^{n-1} h_y(j) \Delta u(k-j) + h_y(n) u(k-n)$$

Esercizio 8.4 Al variare di a e b , si determini la risposta all'impulso di un sistema con funzione di trasferimento

$$G(z) = \frac{z-b}{z-a}$$

Si verifichi il risultato ottenuto calcolando i primi tre coefficienti della risposta all'impulso con il metodo della lunga divisione (Appendice C.3.3).

Esercizio 8.5 Si determinino tre possibili rappresentazioni nel tempo del sistema descritto dalla funzione di trasferimento

$$G(z) = \frac{\beta_1 z + \beta_0}{z^2 + \alpha_1 z + \alpha_0}$$

Esercizio 8.6 Si verifichi che i due sistemi

$$y(k) = y(k-1) + u(k-1)$$

$$y(k) = y(k-n) + \sum_{i=1}^n u(k-i)$$

hanno la stessa funzione di trasferimento.

Esercizio 8.7 Si dimostrino le Equazioni (8.44)-(8.46).

Esercizio 8.8 Si determini la funzione di trasferimento di tipo FIR di un sistema la cui uscita, a fronte dell'ingresso

$$u(k) = \sin(k) + 2 \sin(2k) + 3 \sin(3k)$$

a transitorio esaurito assume l'andamento

$$y(k) = 10 \sin(3k)$$

Si calcoli la durata del transitorio.

Esercizio 8.9 Si traccino i diagrammi di Bode e il diagramma polare della risposta in frequenza associata alla funzione di trasferimento

$$G(z) = \frac{1-0.5z}{z-0.5}$$

Si dica perché questo sistema può essere considerato uno sfasatore puro.

Esercizio 8.10 Si determinino i diagrammi polari e di Bode delle risposte in frequenza associate alle funzioni di trasferimento

$$G_1(z) = \frac{0.5}{z-0.5}, \quad G_2(z) = \frac{0.5}{z(z-0.5)}, \quad G_3(z) = \frac{0.5}{z^2(z-0.5)}$$

e si valuti l'effetto della presenza di uno o più ritardi di tempo.

Sistemi di controllo a tempo continuo: stabilità

9

9.1 Introduzione

Con questo capitolo viene ripreso lo studio delle problematiche del controllo. le cui caratteristiche peculiari sono state introdotte già nel Capitolo 1 mediante esempi e in maniera puramente intuitiva.

Un sistema di controllo in anello chiuso può essere rappresentato mediante uno schema a blocchi retroazionato in cui compaiono diversi componenti (il processo, l'attuatore, il trasduttore, il regolatore) che in generale sono descrivibili come sistemi dinamici. L'analisi del funzionamento del sistema complessivo, ma anche il problema della sintesi del regolatore, possono essere quindi affrontati mediante l'uso delle tecniche di base della teoria dei sistemi, ampiamente trattate nei Capitoli 2-8. In questo e nei prossimi capitoli vengono inizialmente considerati sistemi di controllo in cui tutti i componenti sono modellizzati come sistemi dinamici lineari SISO, stazionari e a tempo continuo. Si assume in particolare che vi sia una sola variabile controllata e una sola variabile di controllo.

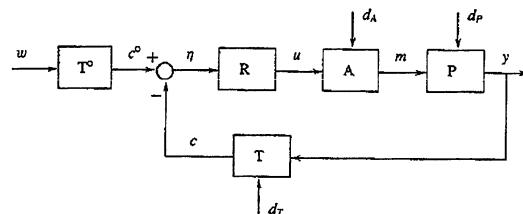
Diversi sono i requisiti che deve possedere un sistema di controllo: precisione statica e dinamica, velocità di risposta, capacità di neutralizzare l'effetto di eventuali disturbi ecc. Quello però di gran lunga più importante per il corretto funzionamento del sistema è il requisito di stabilità asintotica, senza il quale la variabile controllata potrebbe addirittura divergere in presenza di alcuni valori dello stato iniziale o in risposta a eventuali disturbi, anche di ampiezza limitata. Lo studio della stabilità di un sistema di controllo è quindi essenziale prima ancora di valutare le sue prestazioni, ed è reso complicato dal fatto che la stabilità di un sistema in anello chiuso non può essere garantita *a priori* connettendo tra loro sottosistemi asintoticamente stabili. A questo argomento è dedicata una parte consistente del presente capitolo, nel quale si discuteranno in particolare:

- le motivazioni che sono alla base dell'uso di modelli lineari nella descrizione di un sistema di controllo;
- la derivazione di uno schema generale di controllo in anello chiuso;
- i principali requisiti di un sistema di controllo;
- lo studio delle proprietà di stabilità asintotica in condizioni nominali, quando cioè il modello del sistema è perfettamente noto;
- la possibilità di garantire la stabilità del sistema di controllo anche in condizioni perturbate, cioè diverse da quelle nominali.

9.2 Controllo nell'intorno di un equilibrio

Nel Capitolo 1 è stato presentato lo schema generale di un sistema di controllo in retroazione (Figura 1.7). Questo schema, trascurando per semplicità l'azione di compensazione del disturbo, è analogo a quello di Figura 9.1 dove P rappresenta il processo, A l'attuatore, T e T^o i trasduttori e R il regolatore. L'uscita y è la variabile controllata e w il corrispondente andamento desiderato, o segnale di riferimento; c è la misura di y e c^o quella di w ; $\eta = c^o - c$ è l'errore valutato sulle misure, u la variabile di controllo, m la variabile manipolabile. Inoltre d_A , d_A e d_T descrivono altre variabili esogene di disturbo che agiscono rispettivamente sul processo, sull'attuatore e sul trasduttore. Nel seguito si assumerà che tutte queste grandezze siano scalari.

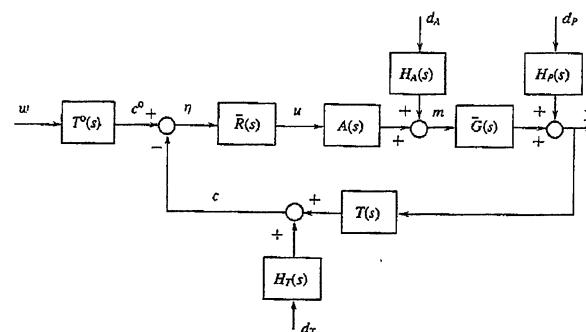
Figura 9.1
Schema generale di controllo in retroazione.



Poiché in generale gli elementi che compongono il sistema di controllo di Figura 9.1 sono sistemi dinamici non lineari e magari anche varianti nel tempo, l'analisi del sistema complessivo risulta decisamente difficile. Se invece tutti i blocchi dello schema potessero essere considerati come sistemi lineari e stazionari, a ognuno di essi si potrebbe sostituire la sua rappresentazione in termini di funzioni di trasferimento, pervenendo allo schema a blocchi di Figura 9.2. Sarebbe a questo punto possibile studiare le proprietà del sistema di controllo sfruttando la teoria dei sistemi lineari a tempo continuo sviluppata nei capitoli precedenti.

Nel seguito si supporrà di poter descrivere il sistema di controllo mediante lo schema di Figura 9.2. L'ipotesi di linearità sembra alquanto riduttiva, ma è in realtà assai ragionevole per un'importante categoria di problemi di controllo nei quali lo scopo del regolatore è quello di controllare il processo nell'intorno di una particolare condizione di equilibrio (si vedano gli Esempi 1.2 e 1.3). Se gli scostamenti delle variabili in gioco rispetto ai loro valori di equilibrio si mantengono "piccoli", è infatti lecito ricorrere a una descrizione linearizzata del processo ed

Figura 9.2
Schema di controllo in retroazione con elementi lineari e stazionari.



eventualmente degli altri componenti del sistema. In tal caso, le variabili che appaiono nello schema di Figura 9.2 andrebbero intese a rigor di termini come le deviazioni delle corrispondenti variabili rispetto ai valori di equilibrio. L'esempio che segue permetterà di chiarire meglio questa interpretazione.

Esempio 9.1 Si consideri il sistema idraulico di Figura 9.3, in cui il livello h dell'acqua nel serbatoio a pelo libero può essere regolato agendo sulla pompa volumetrica di immissione della portata d'ingresso q_I . La condotta di scarico genera una portata di uscita q_u verso un ambiente in cui è mantenuta una pressione costante p_u , inferiore alla pressione atmosferica. Si possono facilmente individuare i principali componenti di un sistema di controllo: l'attuatore rappresentato dalla pompa volumetrica, il processo sotto controllo costituito dal serbatoio, il trasduttore per la misura del livello. Inoltre, la variabile controllata è il livello h , la pressione p_u può essere interpretata come un disturbo agente sul sistema, q_i è la variabile manipolabile, mentre l'effettiva variabile di controllo u è il comando fornito alla pompa. Ora si ricaveranno i modelli matematici dei singoli componenti.

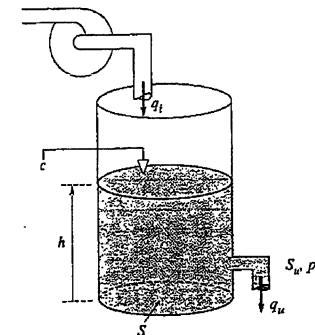


Figura 9.3
Serbatoio dell'Esempio 9.1.

Attuatore La pompa è azionata da un motore a corrente continua ed è rappresentata con buona accuratezza da un sistema dinamico lineare, invariante e asintoticamente stabile del secondo ordine. La funzione di trasferimento $A(s)$ associata a tale sistema abbia guadagno μ_A . Questa funzione rappresenta il legame tra il segnale di comando u , la corrente di alimentazione del motore, e la portata volumetrica q_I imposta. Va notato che, nel caso di ampie escursioni delle variabili, l'ipotesi di linearità andrebbe messa in discussione. Per esempio, la pompa ha comunque un comportamento non lineare dovuto al fatto che la portata imposta non può essere negativa e non può superare un valore massimo prefissato q_{IM} . Si parla in tal caso di un fenomeno di saturazione sulla variabile manipolabile. Esistono poi vari disturbi che influenzano il comportamento della pompa, tra cui improvvisi e temporanei cali della tensione di alimentazione, o riduzioni di portata dovute al surriscaldamento provocato da un prolungato funzionamento. Questi disturbi possono essere modellizzati dal segnale d_A di Figura 9.1, assumendo per esempio che la portata effettivamente entrante nel serbatoio sia la somma della portata imposta dalla pompa e di un termine dovuto ai disturbi, cioè $q_I + d_A$. È ragionevole supporre che d_A sia nullo in condizioni nominali.

Processo Il serbatoio è cilindrico con sezione di area S . Per il principio di conservazione della massa risulta

$$\dot{h}(t) = \frac{1}{S} (q_I(t) + d_A(t) - q_u(t)) \quad (9.1)$$

dove q_u è la portata di uscita, funzione del livello h , della sezione della condotta di scarico S_u e della pressione nel serbatoio di raccolta p_u ($p_u < p_a$). Dall'equazione di Bernoulli relativa al moto stazionario di un fluido in un condotto è possibile ricavare la relazione

$$q_u(t) = S_u \sqrt{2 \left[gh(t) + \frac{p_a - p_u(t)}{\rho} \right]} \quad (9.2)$$

dove g è l'accelerazione di gravità, ρ è la densità dell'acqua e p_a è la pressione atmosferica (che si suppone costante). Le Equazioni (9.1) e (9.2) definiscono il modello non lineare del serbatoio

$$\dot{h}(t) = \frac{1}{S} \left\{ q_I(t) + d_A(t) - S_u \sqrt{2 \left[gh(t) + \frac{p_a - p_u(t)}{\rho} \right]} \right\} \quad (9.3)$$

Trasduttore Per determinare il livello si usa un trasduttore che misura la capacità elettrica, funzione del livello stesso, tra elettrodi parzialmente sommersi nell'acqua del serbatoio. In prima approssimazione si può ritenere che il legame tra h e il segnale c fornito dal trasduttore sia non dinamico e rappresentato da una funzione monotona crescente e non lineare $c = f_T(h)$. Anche in questo caso è necessario considerare la presenza di incertezze che possono influenzare il comportamento del trasduttore in modo significativo. Per esempio, per il principio di funzionamento impiegato, la misura prodotta dipende certamente dalla salinità dell'acqua. L'effetto dei disturbi può essere modellizzato dal segnale d_T di Figura 9.1, supponendo per esempio che l'effettiva variabile trasmessa sia $c + d_T$, con $d_T = 0$ in condizioni nominali.

Per quanto riguarda il trasduttore del segnale di riferimento h^o , si ipotizza che anch'esso sia descritto dalla caratteristica statica non lineare f_T .

Condizioni di equilibrio Il sistema di controllo deve mantenere un livello costante $h^o = \bar{h}$ in presenza di una pressione allo scarico costante \bar{p}_u . Dall'Equazione (9.3), ricordando che in condizioni di funzionamento nominali $d_A = 0$, si deduce che la portata d'ingresso necessaria per ottenere il valore desiderato di livello è

$$\bar{q}_i = S_u \sqrt{2 \left(g\bar{h} + \frac{p_a - \bar{p}_u}{\rho} \right)}$$

Dal valore di \bar{q}_i è possibile ricavare direttamente il valore $\bar{u} = \bar{q}_i/\mu_A$ da impostare all'ingresso dell'attuatore. Infine, ipotizzando nullo il valore di d_T , la misura trasmessa all'equilibrio è data da $\bar{c} = f_T(\bar{h})$.

Linearizzazione Pur di definire i segnali alle variazioni

$$\begin{aligned}\delta u(t) &= u(t) - \bar{u} \quad , \quad \delta q_i(t) = q_i(t) - \bar{q}_i \quad , \quad \delta h(t) = h(t) - \bar{h} \\ \delta c(t) &= c(t) - \bar{c} \quad , \quad \delta p_u(t) = p_u(t) - \bar{p}_u \quad , \quad \delta d_A(t) = d_A(t) \\ \delta d_T(t) &= d_T(t)\end{aligned}$$

è ora possibile determinare il modello linearizzato del processo e del trasduttore, mentre, poiché si è ipotizzato che la pompa abbia un comportamento lineare, il legame tra le trasformate di Laplace di δu e δq_i è ancora dato da $A(s)$. Definendo

$$\tau(\bar{h}, \bar{p}_u) = \frac{S \sqrt{2 \left(g\bar{h} + \frac{p_a - \bar{p}_u}{\rho} \right)}}{S_u g}$$

la linearizzazione dell'Equazione (9.3) conduce a

$$\delta \dot{h}(t) = \frac{1}{S} \left[\delta q_i(t) + \delta d_A(t) - \frac{S}{\tau(\bar{h}, \bar{p}_u)} \delta h(t) + \frac{S}{\rho g \tau(\bar{h}, \bar{p}_u)} \delta p_u(t) \right] \quad (9.4)$$

Dall'Equazione (9.4) si ricavano le funzioni di trasferimento $\tilde{G}(s)$ tra $\delta q_i + \delta d_A$ e la variazione di livello δh e $H(s)$ tra δp_u e δh :

$$\tilde{G}(s) = \frac{\tau(\bar{h}, \bar{p}_u)/S}{1 + \tau(\bar{h}, \bar{p}_u)s} \quad , \quad H(s) = \frac{1/\rho g}{1 + \tau(\bar{h}, \bar{p}_u)s} \quad (9.5)$$

Sempre mediante una procedura di linearizzazione è infine possibile ricavare il legame

$$\delta c(t) = \mu_T(\bar{h}) \delta h(t) \quad , \quad \mu_T(\bar{h}) = \frac{df_T(h)}{dh} \Big|_{h=\bar{h}} \quad (9.6)$$

relativo al trasduttore.

In definitiva, nel progetto di un regolatore che debba mantenere il processo nell'intorno della condizione di equilibrio considerata si potrà fare riferimento allo schema retroazionato di Figura 9.4, in cui le funzioni di trasferimento del processo e dei trasduttori sono definite dalle Equazioni (9.5), (9.6). Nello schema si è anche evidenziata la possibilità di considerare delle (piccole) variazioni δh^o del segnale di riferimento attorno al valore nominale \bar{h} .

Le Equazioni (9.5), (9.6) mostrano chiaramente come, al variare del livello nominale, cambino le caratteristiche (polo e guadagni) dei blocchi che compaiono nello schema di Figura 9.4.

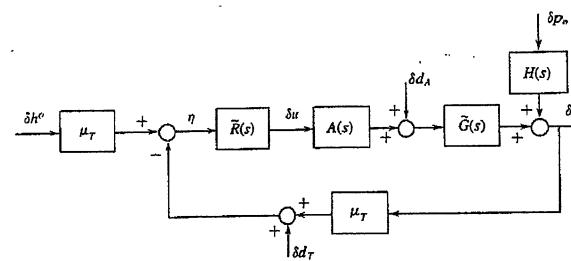


Figura 9.4
Schema di controllo in retroazione dell'Esempio 9.1.

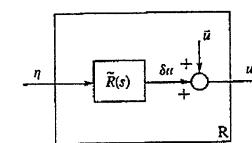


Figura 9.5
Realizzazione del regolatore nell'Esempio 9.1.

Si noti infine che, una volta determinata la funzione di trasferimento $\tilde{R}(s)$, il regolatore di Figura 9.1 può essere realizzato secondo lo schema di Figura 9.5, ricordando cioè che l'uscita di $\tilde{R}(s)$ rappresenta solo la variazione di corrente da applicare al motore rispetto al valore di equilibrio $\bar{u} = \bar{q}_i/\mu_A$.

9.3 Schema generale di controllo in retroazione

Come si è detto, lo schema di Figura 9.2 sarà utilizzato per studiare le proprietà di un generico sistema di controllo.

Per semplificare l'analisi successiva è possibile fare ricorso alle regole di elaborazione degli schemi a blocchi presentate nel Paragrafo 5.3 e modificare lo schema di Figura 9.2 come mostrato nella Figura 9.6. Si noti che la funzione di trasferimento d'anello $L(s) = T(s)\tilde{R}(s)A(s)\tilde{G}(s)$ dello schema retroazionato di Figura 9.6 è identica a quella dello schema di Figura 9.2. Pertanto, in base all'analisi svolta nel Paragrafo 5.4, le proprietà di stabilità dei due sistemi retroazionati corrispondenti sono equivalenti.

Si noti ora che nello schema di Figura 9.6 i segnali d e \tilde{d} possono essere considerati del tutto analoghi per quanto riguarda il loro legame con le variabili y e u di interesse per l'analisi. Pertanto, per semplicità, nel seguito si ipotizzerà

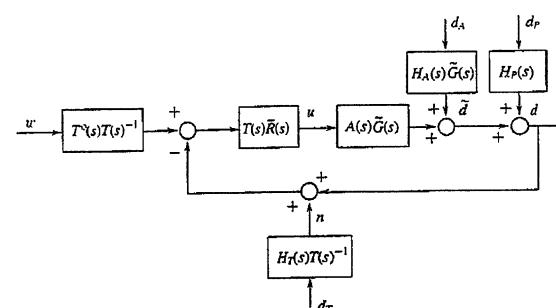
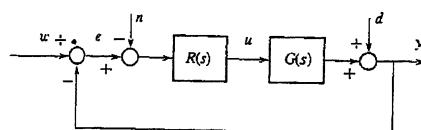


Figura 9.6
Schema di controllo della Figura 9.2 rielaborato.

Figura 9.7
Schema di controllo
in retroazione.



$\tilde{d} = 0$. Inoltre, nel caso frequente in cui $T^*(s) = T(s)$, il blocco con ingresso w ha funzione di trasferimento unitaria e si può quindi eliminare. Definendo infine le funzioni di trasferimento $G(s) = A(s)\tilde{G}(s)$ e $R(s) = T(s)\tilde{R}(s)$, con semplici sostituzioni ed elaborazioni lo schema di Figura 9.6 può essere rappresentato come nella Figura 9.7, dove è stata messa in evidenza la variabile $e = w - y$, che rappresenta l'errore del sistema di controllo.

Nel seguito del capitolo, per analizzare le proprietà del sistema retroazionato si farà sempre riferimento allo schema di Figura 9.7; naturalmente nel formulare ipotesi sull'andamento di d (o \tilde{d} nella Figura 9.6), così come di n , si ricorderà quando opportuno che essi sono l'uscita di sistemi dinamici con determinate caratteristiche, alimentati dagli effettivi segnali esogeni d_p (o d_A) e d_T . Inoltre, con un lieve abuso di linguaggio, $R(s)$ sarà detta la funzione di trasferimento del regolatore e $G(s)$ del sistema sotto controllo, ricordando soltanto quando necessario che esse comprendono anche la dinamica del trasduttore e dell'attuatore. Infine, se non specificato diversamente, si assumerà che $G(s)$ e $R(s)$ siano funzioni razionali con $G(s)$ strettamente propria, mentre $R(s)$ potrà rappresentare anche sistemi non strettamente propri ed eventualmente non dinamici. La funzione d'anello $L(s) = R(s)G(s)$ risulta in ogni caso strettamente propria.

9.4 Requisiti di un sistema di controllo

In questo paragrafo vengono presentati i principali requisiti che deve possedere un sistema di controllo in retroazione e si imposta il problema dell'analisi delle sue caratteristiche.

9.4.1 Stabilità

Stabilità in condizioni nominali Il requisito fondamentale del sistema retroazionato di Figura 9.7 è l'*asintotica stabilità*. Come discusso nel Paragrafo 3.4, se il sistema di controllo non godesse di tale proprietà, il suo comportamento per $t \rightarrow \infty$ dipenderebbe dalle condizioni iniziali. Inoltre, tutte le variabili presenti, compreso l'errore, potrebbero divergere anche a fronte di segnali esogeni di ampiezza limitata, in contrasto con l'obiettivo di far seguire alla variabile di uscita l'andamento desiderato w .

Stabilità in condizioni perturbate Come già altre volte sottolineato, la funzione di trasferimento $G(s)$ rappresenta spesso soltanto un modello approssimato del sistema sotto controllo. In particolare, le approssimazioni possono essere dovute:

- a ipotesi deliberatamente introdotte in fase di modellizzazione per rendere meno complessa la struttura del modello risultante (si veda al riguardo l'Esempio 2.7);
- a eventuali ulteriori semplificazioni, quali "quasi cancellazioni" di poli e zeri o l'eliminazione di singolarità al di fuori della banda di frequenze di interesse (per esempio poli a frequenza elevata);

- al fatto che la funzione di trasferimento del sistema sotto controllo sia stata ricavata da esperimenti eseguiti sul sistema, per esempio da una risposta allo scalino, e sia quindi inevitabilmente affetta da incertezza;
- a modifiche di $G(s)$ legate alla variazione del valore di qualche parametro;
- allo spostamento del sistema non lineare sotto controllo in un punto di funzionamento differente per il quale la funzione $G(s)$ ottenuta per linearizzazione non risulti più adeguata.

Esempio 9.2 (Seguito dell'Esempio 9.1) Poiché in genere la sezione del serbatoio è molto maggiore di quella dello scarico ($S \gg S_u$), la costante di tempo di $\tilde{G}(s)$ può essere di vari ordini di grandezza superiore a quelli di $A(s)$. Nel progetto del sistema di controllo, la pompa, il serbatoio e il trasduttore di livello possono essere allora descritti dalla funzione di trasferimento $\mu_A \mu_T \tilde{G}(s)$. Questo modello è ovviamente approssimato poiché non tiene conto della dinamica in alta frequenza dovuta ai poli della funzione di trasferimento $A(s)$ della pompa. Inoltre, come già sottolineato ed evidente dalle (9.5), (9.6), al variare del punto di funzionamento del sistema, cioè di \tilde{h} e $\tilde{\rho}_u$, le caratteristiche del modello cambiano. Infine la determinazione di alcuni parametri, per esempio l'area S_u , può essere difficoltosa e quindi il loro valore può essere notevolmente incerto.

Per tutte le ragioni elencate è importante garantire la proprietà di stabilità non soltanto in condizioni nominali, ma anche a fronte di errori di modello o variazioni dei parametri del sistema, cioè *in condizioni perturbate*. Se si riesce ad assicurare ciò, si dice che il sistema retroazionato gode della proprietà di *stabilità robusta*. Naturalmente, poiché è irrealistico richiedere la proprietà di stabilità per ogni possibile perturbazione, è necessario avere una stima dei margini di incertezza del modello impiegato e quindi assicurare la stabilità per tutte le perturbazioni ammissibili all'interno di tali margini.

9.4.2 Prestazioni

Una volta garantita la stabilità, è necessario che il sistema di controllo in condizioni nominali sia in grado di fornire determinate prestazioni, goda cioè delle proprietà di seguito elencate.

Prestazioni statiche in condizioni nominali Uno degli scopi tipici di un sistema di controllo è quello di garantire che per $t \rightarrow \infty$, cioè dopo che si è esaurito l'effetto del transitorio iniziale, l'errore e si mantenga limitato, o in particolare che sia nullo, a fronte di segnali esogeni w , d , n di caratteristiche specificate. Visto che si parla di comportamento a regime, queste prestazioni vengono di solito indicate come *prestazioni statiche*. I tipici requisiti riguardano:

- l'errore a regime dovuto a ingressi a scalino (con trasformata A/s): nel problema di controllo nell'intorno di una data condizione di equilibrio, il considerare segnali con andamento a scalino è giustificato dalla frequente esigenza di far passare il sistema retroazionato da un punto di funzionamento a un altro, prossimo al primo, e dal fatto che i disturbi sugli attuatori e/o sui trasduttori spesso possono essere modellizzati come dei segnali costanti nel tempo (con riferimento all'Esempio 9.1 si pensi a un errore sistematico di misura del livello causato da un'imperfetta taratura del trasduttore);
- l'errore a regime dovuto a ingressi che crescono linearmente nel tempo, modellizzabili come rampe (con trasformata A/s^2): a questo proposito vale la pena di ricordare che nel passare dallo schema di Figura 9.6 a quello di Figura 9.7 si

è definita con \tilde{d} l'uscita del sistema con funzione di trasferimento $H_A(s)\tilde{G}(s)$ e ingresso d_A ; pertanto se d_A è un segnale a scalino e $H_A(s)\tilde{G}(s)$ ha un polo nell'origine, oltre a eventuali altri poli a parte reale negativa, il disturbo \tilde{d} coincide asintoticamente con una rampa; vi sono poi casi in cui si richiede che la variabile controllata segua un andamento a rampa (si pensi, per esempio, a un'antenna radar che debba seguire il movimento di un satellite che voli a velocità costante);

- più in generale, il comportamento asintotico del sistema di controllo per segnali di ingresso di tipo canonico dotati di trasformata di Laplace A/s^i , con i intero positivo;
- il comportamento a transitorio esaurito del sistema di controllo in risposta a ingressi di tipo sinusoidale (Capitolo 6): se si dispone di una descrizione dei disturbi che agiscono sul sistema in termini delle loro proprietà spettrali, si può determinare l'effetto che ogni loro componente armonica produce asintoticamente sui segnali presenti lungo l'anello di retroazione e, in particolare, sul segnale errore; è per esempio di interesse verificare che l'ampiezza dell'errore asintotico prodotto da queste componenti si mantenga entro limiti predefiniti.

Prestazioni dinamiche in condizioni nominali Anche nel caso delle prestazioni dinamiche, cioè quelle relative al comportamento in transitorio del sistema di controllo, è conveniente fare riferimento a particolari segnali di ingresso. Nella grande maggioranza dei casi i requisiti di interesse sono relativi all'andamento della variabile di uscita y (o dell'errore e) a fronte di variazioni a scalino del segnale di riferimento w e dei disturbi d e n .

- **Risposta al segnale di riferimento.** Nel caso ideale ovviamente si vorrebbe una perfetta coincidenza di y con w a ogni istante di tempo. Dato che ciò non è possibile, i tipici requisiti fanno riferimento a limiti imposti al tempo di assestamento T_{ae} , al tempo di salita T_s , all'entità percentuale massima delle sovraetelongazioni $S\%$ (si veda il Paragrafo 4.4.2 per la loro definizione). Più in generale, si può richiedere che la risposta allo scalino del sistema controllato sia all'interno di una regione specificata, come mostrato nella Figura 9.8. A sua volta, questo requisito si tradurrà nel richiedere determinate caratteristiche per la risposta in frequenza associata alla funzione di trasferimento tra il segnale di riferimento e l'uscita, quali un andamento di tipo passa-basso, un'ampiezza specificata per la banda passante, la limitatezza di eventuali picchi di risonanza.
- **Risposta ai disturbi.** Idealmente si vorrebbe che l'effetto di d e n su y fosse nullo a ogni istante di tempo. Ciò in generale non è possibile, e ci si accontenta quindi di richiedere che il loro effetto produca variazioni di y contenute in ampiezza e che possano ritenersi esaurite in un tempo specificato (si veda ancora la Figura 9.8). Anche in questo caso il requisito di attenuazione dei disturbi verrà riformulato nel richiedere determinate caratteristiche per le risposte in frequenza associate alle funzioni di trasferimento tra i disturbi e l'uscita. Per esempio, se si conosce la banda di frequenze in cui è prevalentemente concentrata l'energia del disturbo, è bene che il modulo della risposta in frequenza tra il disturbo e l'uscita assuma valori "piccoli" in quella banda di frequenze.
- **Moderazione della variabile di controllo.** Con questo termine si intende la capacità del sistema retroazionato di produrre, a fronte dei segnali di riferimento e dei disturbi, una variabile di controllo contenuta in ampiezza. Questo requi-

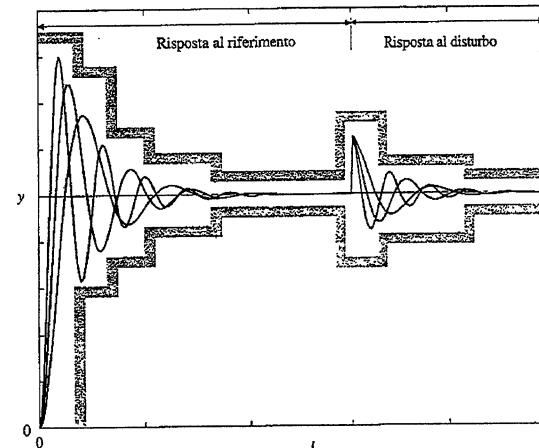


Figura 9.8
Famiglia di risposte desiderate per il sistema retroazionato a fronte di andamenti a scalino di u , d , e.

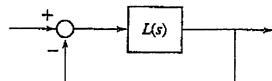
sito è di notevole importanza per varie ragioni. Prima di tutto, si ricordi che, come mostrato nel Paragrafo 9.2, spesso il sistema di Figura 9.7 rappresenta il sistema fisico soltanto nell'intorno della condizione di equilibrio attorno a cui è stata effettuata la linearizzazione; l'ingresso u è quindi la variazione della variabile di controllo rispetto al suo valore di equilibrio. Affinché il procedimento di linearizzazione sia significativo, è necessario che tutte le variabili che compaiano lungo l'anello di regolazione, e in particolare u , siano contenute in ampiezza. Inoltre, in ogni sistema reale, gli attuatori introducono delle non-linearità dovute alle saturazioni delle variabili manipolabili (si veda ancora al riguardo l'Esempio 9.1). È quindi opportuno che il sistema di controllo operi in modo che queste saturazioni non siano attive, cioè che la variabile u non assuma valori eccessivi. Oltre a ciò, è bene ricordare che alla variabile di controllo, corrisponde in genere un'elevata grandezza fisica, per esempio una portata di combustibile in problemi di controllo di temperatura, una forza o una coppia nel controllo di sistemi meccanici. Connesse quindi all'impiego di tali variabili vi possono essere considerazioni relative al "costo" dell'azione di regolazione. Si fa qui riferimento sia all'impossibilità di imporre alla variabile di controllo valori grandi a piacere, sia a una disponibilità limitata dell'azione di controllo dovuta all'uso di attuatori di dimensioni contenute, sia a esigenze di risparmio nella conduzione del sistema. Infine, va segnalato che in certi casi è opportuno evitare brusche variazioni della variabile di controllo al fine di ridurre le sollecitazioni cui è sottoposto l'attuatore, e quindi prolungarne la vita.

Prestazioni in condizioni perturbate Anche con riferimento alle prestazioni, è importante garantire il soddisfacimento di determinati requisiti non soltanto in condizioni nominali, ma anche a fronte di errori di modello o variazioni dei parametri del sistema, cioè in condizioni perturbate. Se si riesce ad assicurare ciò, si dice che il sistema retroazionato fornisce *prestazioni robuste*. Naturalmente, poiché anche in questo caso è irrealistico richiedere che le prestazioni siano garantite per ogni possibile perturbazione, è necessario disporre di una stima dei margini di incertezza del modello impiegato e quindi assicurare le prestazioni desiderate a fronte di tutte le perturbazioni ammissibili all'interno di tali margini.

9.5 Stabilità in condizioni nominali

Poiché la stabilità di un sistema lineare non dipende dagli ingressi (Paragrafo 3.4.1), nell'analisi che segue si farà riferimento al sistema di Figura 9.9 con $L(s) = R(s)G(s)$, anziché a quello di Figura 9.7.

Figura 9.9
Sistema retroazionato
considerato nell'analisi
di stabilità.



Si osservi che, per come sono state definite $R(s)$ e $G(s)$, la funzione di trasferimento d'anello è $L(s) = T(s)\tilde{R}(s)A(s)\tilde{G}(s)$. Una prima condizione necessaria per l'asintotica stabilità del sistema retroazionato di Figura 9.9 è che il sistema complessivo, formato dalla connessione in serie dei sottosistemi descritti da queste funzioni di trasferimento, non abbia autovalori a parte reale positiva o nulla associati a parti "nascoste" (non raggiungibili e/o non osservabili). Nell'ipotesi che i singoli sottosistemi non abbiano parti "nascoste", o che gli autovalori a esse associate siano a parte reale negativa, ciò implica che non vi debbano essere cancellazioni tra poli a parte reale maggiore o uguale a zero di $T(s)$, $\tilde{R}(s)$, $A(s)$ e $\tilde{G}(s)$ e corrispondenti zeri delle altre funzioni di trasferimento che concorrono a formare $L(s)$. Non viene invece violata la stabilità nel caso in cui avvengano cancellazioni tra poli e zeri a parte reale negativa. Nel seguito si assumerà sempre che la condizione necessaria precedente sia soddisfatta; inoltre si ipotizzerà che tutte le cancellazioni possibili siano effettuate così che il numeratore e il denominatore di $L(s)$ non abbiano radici in comune.

Come mostrato nel Paragrafo 5.4.3, l'asintotica stabilità del sistema retroazionato si analizza verificando che tutte le radici dell'equazione caratteristica di tale sistema, cioè

$$1 + L(s) = 0 \quad (9.7)$$

abbiano parte reale minore di zero. Questa verifica può essere effettuata mediante l'uso di strumenti di calcolo ormai ampiamente diffusi o, se $L(s)$ dipende da uno o più parametri, per mezzo del criterio di Routh (si veda l'Esempio 3.16). Tuttavia tali metodi, anche se efficaci per giudicare se un dato sistema di controllo è stabile o meno, risultano spesso di scarsa utilità pratica nella soluzione dei problemi di sintesi del regolatore. Per questa ragione è conveniente introdurre il fondamentale *criterio di Nyquist* che, oltre a rappresentare un valido strumento di analisi e di sintesi basato sulla risposta in frequenza della funzione di trasferimento d'anello, consentirà anche di valutare la proprietà di stabilità robusta del sistema stesso.

9.5.1 Diagramma di Nyquist

Si definisca preliminarmente come *diagramma di Nyquist*, associato alla generica funzione di trasferimento $M(s)$, l'immagine attraverso $M(s)$ della curva chiusa del piano s costituita dall'intero asse immaginario e da una semicirconferenza di raggio infinito giacente nel semipiano destro. Tale curva viene comunemente chiamata *percorso di Nyquist* ed è rappresentata nella Figura 9.10. In coerenza con quanto discusso nel Paragrafo 6.7 relativamente al tracciamento dei diagrammi polari, se $M(s)$ ha poli a parte reale nulla il percorso di Nyquist viene modificato evitando tali poli con semicirconferenze di raggio infinitesimo giacenti nel semipiano destro.

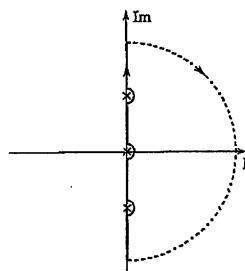


Figura 9.10
Percorso di Nyquist.

Per il tracciamento del diagramma di Nyquist di $M(s)$, si ricordi che $M(-j\omega) = M(j\omega)$ e, se $M(s)$ è strettamente propria, l'immagine attraverso $M(s)$ della semicirconferenza di raggio infinito corrisponde al punto $M(j\infty) = 0$. Il diagramma di Nyquist si può quindi determinare direttamente dal diagramma polare associato a $M(j\omega)$; esso infatti è costituito dal diagramma polare stesso e dalla sua immagine speculare rispetto all'asse reale. Il diagramma è quindi per costruzione una linea chiusa nel piano complesso, il cui orientamento è per convenzione fissato secondo valori crescenti di ω .

Esempio 9.3 (Seguito dell'Esempio 6.6) Nel sistema retroazionato di Figura 9.9 sia

$$L(s) = \frac{\mu}{1+Ts} \quad (9.8)$$

I diagrammi di Nyquist associati ai quattro casi ($\mu > 0, T > 0$), ($\mu < 0, T > 0$), ($\mu < 0, T < 0$) si ottengono direttamente dai diagrammi polari della Figura 6.19 e sono riportati nella Figura 9.11.

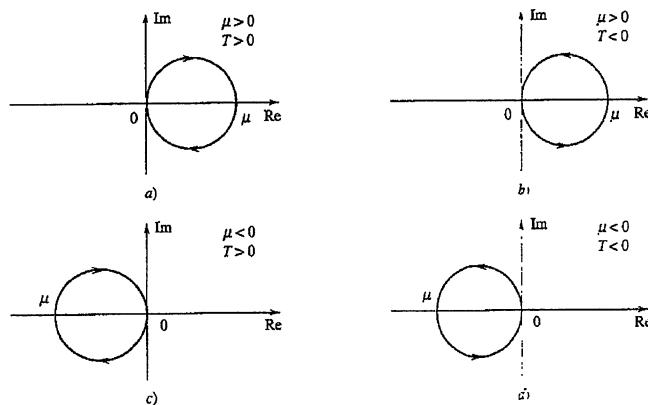


Figura 9.11
Diagramma di Nyquist
della (9.8) per diverse
combinazioni di segno
di μ e T .

9.5.2 Criterio di Nyquist

Sia P il numero di poli di $L(s)$ con parte reale maggiore di zero e N il numero di giri compiuti dal diagramma di Nyquist della funzione d'anello $L(s)$ attorno al punto -1 , conteggiati positivamente se compiuti in senso antiorario e negativamente se in senso orario. Se il diagramma passa per il punto -1 , il valore di N

non è ben definito. È possibile allora formulare il seguente teorema, che prende il nome di *criterio di Nyquist*.

Teorema 9.1 Condizione necessaria e sufficiente perché il sistema retroazionato di Figura 9.9 sia asintoticamente stabile è che N sia ben definito e risulti $N = P$. ■

Dimostrazione Si definisca preliminarmente la funzione

$$K(s) = 1 + L(s) \quad (9.9)$$

e si osservi che, se $L(s) = N_L(s)/D_L(s)$, risulta

$$K(s) = \frac{N_L(s) + D_L(s)}{D_L(s)}$$

Si noti che, per l'ipotesi introdotta in precedenza secondo cui $N_L(s)$ e $D_L(s)$ non hanno radici in comune, anche il numeratore e il denominatore di $K(s)$ non hanno radici in comune. Gli zeri di $K(s)$ corrispondono ai poli del sistema in anello chiuso, mentre i suoi poli coincidono con i poli di $L(s)$. Ciò implica che il percorso di Nyquist associato a $K(s)$ coincide con quello associato a $L(s)$. Inoltre, il numero di poli di $K(s)$ con parte reale positiva è pari a P .

Si consideri ora il diagramma di Nyquist di $K(s)$, relativamente al quale è opportuno distinguere due casi. Si analizzi dapprima il caso in cui, per uno o più valori $\hat{s} = \pm j\hat{\omega}$ al variare di s lungo il percorso di Nyquist, il diagramma passi per l'origine del piano complesso. La funzione $K(s)$ ha allora uno o più zeri sull'asse immaginario in posizione $\pm j\hat{\omega}$, cioè il sistema in anello chiuso ha poli a parte reale nulla e pertanto non è asintoticamente stabile. Si noti che, dalla (9.9), $K(\hat{s}) = 0$ corrisponde a $L(\hat{s}) = -1$, cioè il diagramma di Nyquist di $L(s)$ passa per il punto -1 e N non è ben definito.

Se invece $K(s)$ non si annulla per alcun valore $\hat{s} = \pm j\hat{\omega}$, il diagramma di Nyquist di $K(s)$ è una linea chiusa nel piano complesso non passante per l'origine. In particolare, quando la variabile corrente s compie un intero giro lungo il percorso di Nyquist, ogni zero di $K(s)$ interno al percorso stesso fornisce un contributo di fase pari a -2π , ogni polo interno al percorso dà un contributo pari a 2π , mentre poli e zeri esterni al percorso di Nyquist non producono alcun contributo di fase, come facilmente verificabile dall'analisi della Figura 9.12. Ciò

significa che uno zero interno al percorso fa sì che il diagramma di Nyquist di $K(s)$ compia un giro in senso orario attorno all'origine, mentre ogni polo causa un giro in senso antiorario.

Definendo con Z il numero di zeri di $K(s)$ contenuti all'interno della regione delimitata dal percorso di Nyquist, le osservazioni precedenti portano a concludere che il diagramma di Nyquist di $K(s)$ compie complessivamente un numero di giri in senso antiorario attorno all'origine pari a $P - Z$. La (9.9) implica che il numero N di rotazioni descritte da $K(s)$ intorno all'origine corrisponda al numero di rotazioni descritte da $L(s)$ intorno al punto -1 . In conclusione $N = P - Z$ e, poiché l'asintotica stabilità del sistema retroazionato è equivalente a $Z = 0$, il teorema è dimostrato. ■

Dalla dimostrazione del criterio di Nyquist è anche facile verificare che, quando N non è ben definito, il sistema retroazionato può essere o semplicemente stabile oppure instabile (ma nulla si può concludere in proposito). Se invece N è ben definito e diverso da P , il sistema è sicuramente instabile poiché ha $P - N$ poli a parte reale maggiore di zero.

Esempio 9.4 (Seguito dell'Esempio 9.3) Nel caso in cui $L(s)$ sia la (9.8), dall'equazione caratteristica (9.7) è facile verificare che il sistema in anello chiuso ha un solo autovalore $\hat{s} = -(1+\mu)/T$. Benché in questo caso sia banale determinare le condizioni che devono soddisfare μ e T affinché risulti $\hat{s} < 0$, a scopo illustrativo si analizza ora la proprietà di asintotica stabilità del sistema retroazionato applicando il criterio di Nyquist a partire dai diagrammi riportati nella Figura 9.11.

Caso 1: $\mu > 0$, $T > 0$. Poiché $T > 0$, si ha $P = 0$. Inoltre, dalla Figura 9.11a segue immediatamente $N = 0$. Le condizioni del criterio di Nyquist sono quindi verificate e il sistema in anello chiuso è asintoticamente stabile.

Caso 2: $\mu > 0$, $T < 0$. In questo caso $P = 1$ e, dalla Figura 9.11b, $N = 0$. Il sistema retroazionato è dunque instabile, così come $L(s)$.

Caso 3: $\mu < 0$, $T > 0$. Risulta $P = 0$ e, dall'esame della Figura 9.11c, è necessario distinguere tre sottocasi. Se $\mu > -1$, si ha $N = 0$ e il sistema retroazionato è asintoticamente stabile. Se $\mu = -1$, il valore di N non è ben definito e il sistema retroazionato non è asintoticamente stabile; in particolare, il polo in anello chiuso è $\hat{s} = 0$. Se $\mu < -1$, si ha $N = -1$ e il sistema retroazionato è instabile.

Caso 4: $\mu < 0$, $T < 0$. In questo caso $P = 1$ e ancora, dall'analisi della Figura 9.11d, è necessario distinguere tre sottocasi. Se $\mu > -1$, si ha $N = 0$ e il sistema retroazionato è instabile. Se $\mu = -1$, il valore di N non è ben definito e il sistema retroazionato non è asintoticamente stabile; in particolare, il polo in anello chiuso è $\hat{s} = 0$. Se $\mu < -1$, si ha $N = 1$ e il sistema retroazionato è asintoticamente stabile, malgrado $L(s)$ fosse instabile.

Esempio 9.5 (Seguito dell'Esempio 6.7) Il diagramma di Nyquist associato alla funzione di trasferimento $L(s) = \mu/s$, $\mu > 0$, si ricava direttamente dalla Figura 6.20 ed è riportato nella Figura 9.13. Poiché in questo caso risulta $P = N = 0$, applicando il criterio di Nyquist si conclude che il sistema costituito da un integratore retroazionato negativamente è asintoticamente stabile. Alla medesima conclusione si perviene osservando che l'unica soluzione dell'equazione caratteristica corrispondente è $s = -\mu$.

Figura 9.12
Contributo di fase di zeri e poli di $K(s)$ nel tracciamento del diagramma di Nyquist.

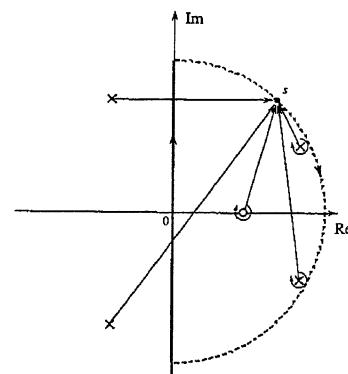
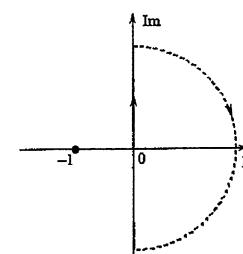


Figura 9.13
Diagramma di Nyquist associato a $L(s) = \mu/s$, $\mu > 0$.

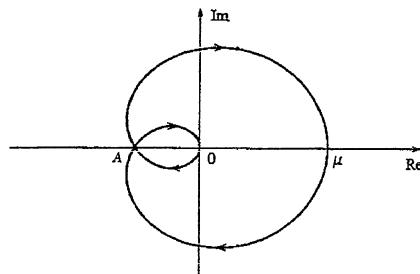


Esempio 9.6 Si consideri un sistema retroazionato negativamente con funzione d'anello

$$L(s) = \frac{\mu}{(1+s)^3}, \quad \mu > 0 \quad (9.10)$$

Si vuole determinare l'insieme dei valori del guadagno d'anello μ per cui il sistema è asintoticamente stabile. Poiché $P = 0$, dall'analisi del diagramma di Nyquist qualitativo di $L(s)$, riportato nella Figura 9.14, si può concludere che, se il punto A di intersezione del diagramma con il semiasse reale negativo ha ascissa $x_A > -1$, si ha $N = 0$ e il sistema retroazionato è asintoticamente stabile. Se invece risulta $x_A \leq -1$, il valore di N non è ben definito, per $x_A = -1$, o è pari a -2 ; in entrambi i casi il sistema retroazionato non è asintoticamente stabile.

Figura 9.14
Diagramma di Nyquist qualitativo associato alla (9.10).



Per la valutazione di x_A si può preliminarmente determinare la pulsazione ω_π per cui $\arg L(j\omega_\pi) = -180^\circ$. Dall'espressione di $L(s)$ questa relazione equivale a $-3 \arctan(\omega_\pi) = -180^\circ$, da cui $\omega_\pi = \sqrt{3}$. In corrispondenza si ha $|L(j\omega_\pi)| = \mu/8$, per cui $x_A = -\mu/8$. In conclusione, il sistema retroazionato negativamente è asintoticamente stabile per $\mu < 8$, mentre diventa instabile per valori maggiori del guadagno d'anello.

9.5.3 Estensioni del criterio di Nyquist

L'analisi condotta nell'Esempio 9.6 relativamente alla stabilità del sistema retroazionato al variare del guadagno della funzione d'anello (o *guadagno d'anello*) può essere generalizzata considerando il sistema di Figura 9.15, in cui $\mu \neq 0$ è un parametro e la cui equazione caratteristica è

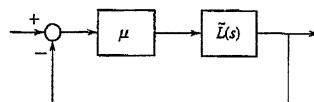
$$1 + \mu \tilde{L}(s) = 0 \quad (9.11)$$

Per tale sistema è possibile formulare il seguente risultato, la cui dimostrazione è analoga a quella del Teorema 9.1, pur di sostituire al punto -1 il punto $-1/\mu$ e alla funzione $K(s)$ definita dalla (9.9) la funzione

$$\tilde{K}(s) = \frac{1}{\mu} + \tilde{L}(s)$$

Corollario 9.1 Condizione necessaria e sufficiente perché il sistema retroazionato di Figura 9.15 sia asintoticamente stabile è che il numero di giri percorsi in senso antiorario dal diagramma di Nyquist di $\tilde{L}(s)$ attorno al punto $-1/\mu$ sia ben definito e risulti uguale al numero di poli di $\tilde{L}(s)$ a parte reale maggiore di zero. ■

Figura 9.15
Sistema retroazionato con funzione d'anello $\mu \tilde{L}(s)$.



Sistemi con retroazione positiva Il risultato precedente consente di estendere il criterio di Nyquist anche al caso dei sistemi caratterizzati da una retroazione positiva, come quello di Figura 9.16. L'equazione caratteristica di tale sistema è

$$1 - \tilde{L}(s) = 0 \quad (9.12)$$

che può essere derivata dalla (9.11) pur di porre $\mu = -1$, e il Corollario 9.1 può essere riformulato nel modo seguente.

Corollario 9.2 Condizione necessaria e sufficiente perché il sistema retroazionato di Figura 9.16 sia asintoticamente stabile è che il numero di giri percorsi in senso antiorario dal diagramma di Nyquist di $\tilde{L}(s)$ attorno al punto $+1$ sia ben definito e risulti uguale al numero di poli di $\tilde{L}(s)$ a parte reale maggiore di zero. ■

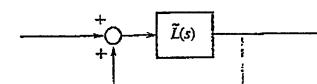


Figura 9.16
Sistema retroazionato positivamente.

Sistemi in anello aperto asintoticamente stabili Tornando al sistema di Figura 9.9, nel caso in cui la funzione di trasferimento $L(s)$ abbia soltanto poli a parte reale negativa, e risulti quindi $P = 0$, per il criterio di Nyquist il sistema retroazionato è asintoticamente stabile se e solo se $N = 0$. In vari casi di interesse questa condizione può essere garantita se sono soddisfatte alcune condizioni sufficienti, la cui verifica è immediata a partire dall'analisi dei diagrammi di Bode o polari associati alla funzione d'anello $L(s)$. Alcune di queste condizioni vengono presentate ora, mentre altre saranno introdotte nel Paragrafo 9.6.

In primo luogo si osservi che, se al variare di s lungo il percorso di Nyquist il modulo di $L(s)$ è comunque minore di 1, certamente il corrispondente diagramma di Nyquist non potrà né passare per il punto critico -1 né circondarlo, e pertanto risulterà $N = 0$, come richiesto. Analoghe considerazioni possono essere fatte circa l'asintotica stabilità del sistema con retroazione positiva di Figura 9.16. Quanto detto può essere formalmente enunciato nel seguente risultato.

Corollario 9.3 Dato un sistema retroazionato negativamente o positivamente con funzione di trasferimento d'anello $L(s)$ asintoticamente stabile, condizione sufficiente per l'asintotica stabilità del sistema in anello chiuso è che risulti $|L(j\omega)| < 1, \forall \omega$. ■

La verifica dell'ipotesi del Corollario 9.3 è immediata dall'analisi del diagramma di Bode del modulo di $L(j\omega)$; infatti è sufficiente verificare che il diagramma stesso giaccia interamente sotto l'asse a 0 dB. Come esempio di applicazione di questo risultato si veda il caso 3 ($\mu < 0, T > 0$) dell'Esempio 9.4. Inoltre, con riferimento allo schema di Figura 9.15, in cui si può supporre che μ rappresenti la funzione di trasferimento del regolatore e $\tilde{L}(s)$ la funzione di trasferimento del sistema sotto controllo, il Corollario 9.3 afferma che, se $\tilde{L}(s)$ è asintoticamente stabile, l'asintotica stabilità del sistema retroazionato è garantita pur di considerare un regolatore puramente proporzionale $R(s) = \mu$ con guadagno sufficientemente piccolo e tale per cui $|\mu \tilde{L}(j\omega)| < 1, \forall \omega$. Tuttavia, come si vedrà nel seguito, questa soluzione nella maggioranza dei casi non è soddisfacente per quanto riguarda i requisiti di prestazioni del sistema di controllo. Si osservi infine che nell'enunciato del corollario non è stata fatta alcuna ipotesi sul valore assunto dallo sfasamento di $L(j\omega)$.

Mentre il Corollario 9.3 esprime una condizione sufficiente relativa al modulo della funzione di trasferimento d'anello, un risultato simile si può formulare considerando solo la fase di $L(j\omega)$. Si osservi infatti che, per sistemi retroazionati negativamente, la condizione $N = 0$ è garantita purché risulti $|\arg L(j\omega)| < 180^\circ$, $\forall \omega$. In questo caso infatti il diagramma polare di $L(j\omega)$ non interseca mai il semiasse reale negativo. È allora possibile formulare il seguente risultato.

Corollario 9.4 Dato il sistema retroazionato negativamente di Figura 9.9 con funzione di trasferimento d'anello $L(s)$ asintoticamente stabile, condizione sufficiente per l'asintotica stabilità del sistema in anello chiuso è che sia verificata la condizione $|\arg L(j\omega)| < 180^\circ$, $\forall \omega$.

Anche in questo caso la verifica dell'ipotesi del Corollario 9.4 è immediata dall'analisi del diagramma di Bode della fase di $L(j\omega)$. Si osservi infine che nell'enunciato del corollario non è stata fatta alcuna ipotesi sul valore assunto dal modulo di $L(j\omega)$.

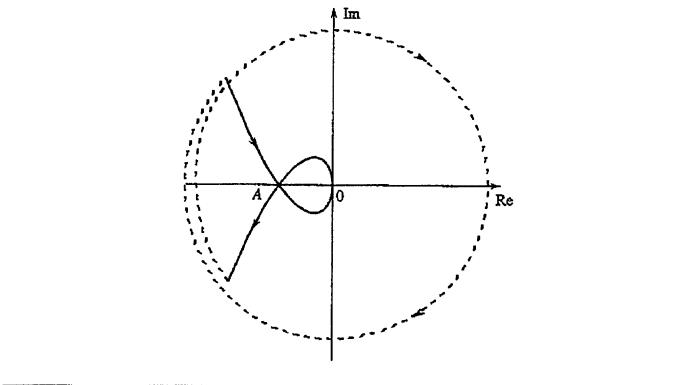
Sistemi a stabilità condizionata L'Esempio 9.6, e successivamente il Corollario 9.1, hanno mostrato come in molti casi, all'aumentare del guadagno μ di Figura 9.15, il sistema retroazionato perda la proprietà di asintotica stabilità. È tuttavia possibile anche il caso contrario, che cioè il sistema retroazionato, asintoticamente stabile per un determinato valore di μ , diventi instabile al diminuire dello stesso. I sistemi con queste caratteristiche sono comunemente chiamati *a stabilità condizionata*.

Esempio 9.7 Si consideri il sistema retroazionato di Figura 9.9 con funzione d'anello

$$L(s) = \frac{\mu(1+s)^2}{s^3}, \quad \mu > 0 \quad (9.13)$$

il cui diagramma di Nyquist associato è riportato nella Figura 9.17, dove è evidenziato il punto A di intersezione del diagramma con il semiasse reale negativo. Si noti che risulta $P = 0$, mentre il valore di N dipende dal valore dell'ascissa x_A di A. Se $x_A < -1$, contando anche le rotazioni all'infinito del diagramma di Nyquist risulta $N = 0$, cioè si ha asintotica stabilità del sistema retroazionato. Se $x_A = -1$, il valore di N non è definito, e se $x_A > -1$, risulta $N = -2$; in entrambi i casi il sistema retroazionato non è asintoticamente stabile. Come già visto nell'Esempio 9.6, il valore di x_A si determina dapprima valutando la pulsazione ω_π per cui $\arg L(j\omega_\pi) = -180^\circ$, e quindi calcolando $x_A = L(j\omega_\pi)$. Con riferimento alla (9.13) risulta $\omega_\pi = 1$, $x_A = -2\mu$. Per quanto detto, l'asintotica stabilità del sistema retroazionato si ha quindi per $\mu > 0.5$.

Figura 9.17
Diagramma di Nyquist della (9.13).



Sistemi con ritardo di tempo Il criterio di Nyquist è stato introdotto con riferimento al caso in cui la funzione di trasferimento d'anello $L(s)$ sia razionale; tuttavia è lecito applicarlo anche per l'analisi di stabilità dei sistemi retroazionati con funzione d'anello

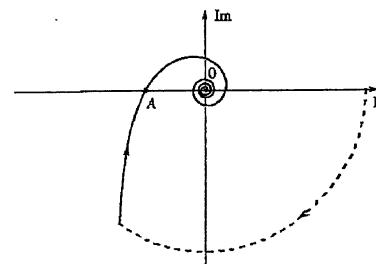
$$L(s) = e^{-\tau s} L'(s)$$

con $L'(s)$ razionale. In questo caso P è il numero di poli a parte reale positiva di $L'(s)$, mentre N va determinato dall'analisi del diagramma di Nyquist associato a $L(s)$.

Come già discusso nel Paragrafo 6.6.3, la presenza del ritardo di tempo introduce uno sfasamento negativo tanto maggiore quanto più è elevato il valore del ritardo τ . Questo sfasamento può facilmente produrre l'instabilità del sistema retroazionato.

Esempio 9.8 Nell'Esempio 9.5 si è mostrato che il sistema retroazionato negativamente con funzione d'anello $L(s) = \mu/s$, $\mu > 0$, è asintoticamente stabile. Si vuole ora valutare la stabilità del sistema retroazionato nel caso in cui $L(s) = e^{-\tau s} \mu/s$, $\mu > 0$, al variare di τ . Il corrispondente diagramma polare è riportato nella Figura 9.18 (per chiarezza si preferisce fare riferimento al diagramma polare anziché a quello di Nyquist), dove il punto A corrisponde alla prima intersezione, al crescere di ω , del diagramma con il semiasse reale negativo. In base al criterio di Nyquist, poiché $P = 0$, la stabilità asintotica del sistema retroazionato si ha soltanto se risulta $x_A > -1$, dove, come di consueto, x_A è l'ascissa di A. Per determinare il valore di x_A e delle altre intersezioni con il semiasse reale negativo è necessario preliminarmente valutare l'insieme dei valori ω_k , $k = 0, 1, 2, \dots$ per cui $\arg L(j\omega_k) = -90^\circ - \omega_k \tau 180^\circ/\pi = -180^\circ - k360^\circ$. Si ricava $\omega_k = (4k+1)\pi/2\tau$. Corrispondentemente si ha $|L(j\omega_k)| = 2\mu\tau/(4k+1)\pi$ e in particolare $x_A = -2\mu\tau/\pi$. Perciò, all'aumentare di τ il diagramma di Nyquist compie in senso orario un numero di giri sempre crescente attorno a -1 , rendendo così il sistema retroazionato instabile. In particolare, la stabilità asintotica si ha unicamente per $\tau < \pi/2\mu$.

Figura 9.18
Diagramma polare di $L(s) = e^{-\tau s} \mu/s$, $\mu > 0$.



9.6 Stabilità in condizioni perturate

Nello studio della stabilità di un sistema retroazionato si è finora immaginato di conoscere con esattezza la funzione di trasferimento d'anello $L(s)$. Applicando il criterio di Nyquist, o qualcuna delle sue estensioni, è possibile in tal caso giudicare la stabilità o meno di un sistema di controllo in condizioni nominali, cioè quando il sistema vero si comporta esattamente come descritto dal modello. È evidente però che, in tutti i casi pratici, è opportuno tener conto dell'inevitabile discrepanza tra il comportamento reale del sistema e quello ideale previsto dal modello. Proprio perché la stabilità è una proprietà fondamentale per il corretto

funzionamento di un sistema di controllo, è quindi interessante studiare la stabilità di un sistema retroazionato anche in condizioni perturbate, cioè quando la funzione d'anello è affetta da incertezza.

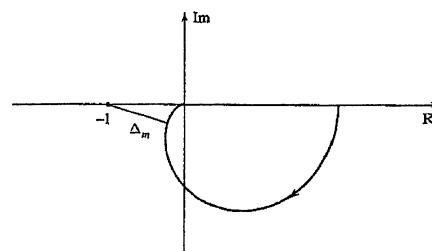
Supponendo di considerare un sistema retroazionato nominalmente stabile, si parlerà di *stabilità robusta* quando il sistema rimane stabile anche in presenza di "elevata" incertezza. Si introdurranno inoltre i cosiddetti *indicatori di robustezza*, cioè parametri numerici capaci di quantificare il grado di robustezza di un sistema, ovvero l'ampiezza delle massime perturbazioni sul modello che garantiscono che la proprietà di stabilità sia conservata. Dopo aver discusso un primo indicatore di stabilità robusta direttamente deducibile dal criterio di Nyquist (il cosiddetto *margine di stabilità vettoriale*), saranno poi introdotti il *margine di guadagno* e il *margine di fase*, che sono i due indicatori di robustezza più comunemente impiegati in pratica.

9.6.1 Margine di stabilità vettoriale

Il criterio di Nyquist, essendo espresso da una condizione di tipo geometrico che riguarda la posizione reciproca di un punto e di una curva nel piano complesso, suggerisce un modo naturale per caratterizzare la proprietà di stabilità robusta.

Si supponga per semplicità che la funzione d'anello $L(s)$ del sistema di Figura 9.9 non possieda poli con parte reale positiva (sia cioè $P = 0$) e si assuma che il sistema sia asintoticamente stabile in condizioni nominali. Ciò significa che $N = 0$, cioè che il diagramma di Nyquist non fa alcun giro intorno al punto -1 , come per esempio nella Figura 9.19, dove per comodità è tracciato solamente il diagramma polare. È ora evidente che, quanto più il diagramma polare si trova distante dal punto -1 , tanto maggiore è il margine di sicurezza rispetto a possibili perturbazioni della funzione d'anello. Se la distanza è elevata, risulta infatti improbabile che deformazioni del diagramma dovute a limitate variazioni di $L(s)$ portino a violare la condizione del criterio di Nyquist.

Figura 9.19
Definizione del margine di stabilità vettoriale.



In definitiva, nelle ipotesi fatte, si può valutare la stabilità robusta di un sistema retroazionato considerando la distanza Δ_m nel piano complesso tra il punto -1 e il diagramma polare associato alla funzione di trasferimento nominale $L(s)$. Tale distanza viene di solito chiamata *margine di stabilità vettoriale* e, per quanto detto, costituisce un utile indicatore di robustezza.

D'altra parte, il calcolo preciso di tale quantità richiede un tracciamento accurato del diagramma polare e non è direttamente ottenibile da una lettura dei diagrammi di Bode di modulo e fase di $L(j\omega)$. Appare quindi opportuno introdurre altri indicatori sintetici che diano informazioni anche solo parziali sul grado di robustezza, ma che, essendo calcolabili con maggiore facilità a partire dai diagrammi di Bode, siano più utili in pratica. Due di tali indicatori sono presentati nei paragrafi successivi.

9.6.2 Margine di guadagno

Con riferimento al sistema retroazionato di Figura 9.9, si supponga che la funzione di trasferimento d'anello $L(s)$ abbia guadagno positivo e non possieda poli a parte reale maggiore di zero (sia cioè $P = 0$); inoltre, il relativo diagramma polare attraversi una sola volta, e in un punto A al finito, il semiasse reale negativo, a parte il punto corrispondente a pulsazione infinita che cade nell'origine. Come illustrato nella Figura 9.20, si supponga infine che il punto A giaccia alla destra del punto -1 , cosicché il sistema retroazionato nominale risulti asintoticamente stabile in virtù del criterio di Nyquist.

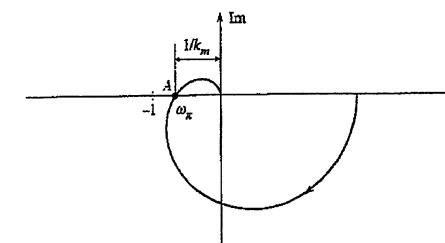


Figura 9.20
Definizione del margine di guadagno.

Una misura del *grado di stabilità* del sistema a fronte di possibili variazioni del modello nominale è certamente fornita dalla distanza tra i punti A e -1 . A questo riguardo, indicando con x_A l'ascissa del punto A , si definisce *margine di guadagno* del sistema retroazionato il valore $k_m = |x_A|^{-1}$. Si osservi che, denotando con ω_π la pulsazione corrispondente al punto A sul diagramma polare, il valore di k_m corrisponde all'inverso del modulo di $L(j\omega)$ in corrispondenza di tale pulsazione, ovvero

$$k_m = \frac{1}{|L(j\omega_\pi)|}, \quad \arg L(j\omega_\pi) = -180^\circ \quad (9.14)$$

La sua valutazione può quindi essere ottenuta direttamente dai diagrammi di Bode di modulo e fase di $L(j\omega)$, tenendo conto che $|k_m|_{dB} = -|L(j\omega_\pi)|_{dB}$, come esemplificato nella Figura 9.21.

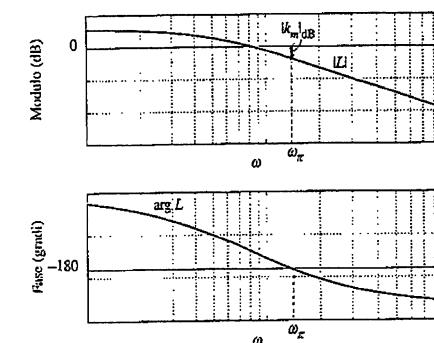
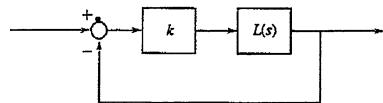


Figura 9.21
Calcolo del margine di guadagno dai diagrammi di Bode.

Per ricavare un'interpretazione del parametro k_m si osservi innanzitutto che, sotto le ipotesi formulate, e in particolare quella sulla posizione reciproca dei punti A e -1 , risulta $k_m > 1$. Facendo poi riferimento al sistema retroazionato di Figura 9.22, dove k rappresenta una costante positiva arbitraria, è facile rendersi conto

Figura 9.22
Sistema retroazionato con
guadagno d'anello
modificato.



che tale sistema risulta asintoticamente stabile per tutti i valori di k inferiori a k_m . Infatti, finché $k < k_m$, il diagramma di Nyquist della funzione d'anello $kL(s)$ non passa per il punto -1 né compie giri attorno a esso, e il criterio di Nyquist assicura l'asintotica stabilità. Il margine di guadagno k_m rappresenta dunque l'estremo superiore dei fattori moltiplicativi del guadagno d'anello che il sistema di Figura 9.9 può tollerare senza perdere la proprietà di stabilità asintotica. In altri termini, k_m fornisce una misura del grado di robustezza della stabilità a fronte di possibili incertezze sul guadagno d'anello.

Esempio 9.9 (Seguito dell'Esempio 9.6) Si consideri un sistema retroazionato negativamente con funzione d'anello

$$L(s) = \frac{4}{(1+s)^3}$$

Il margine di guadagno è pari a $k_m = 2$. Infatti il guadagno $\mu = 4$ della funzione d'anello è esattamente la metà del valore limite determinato nell'Esempio 9.6.

9.6.3 Margine di fase

Con riferimento al sistema retroazionato di Figura 9.9, si assuma ancora che P sia nullo e che il guadagno di $L(s)$ sia positivo; inoltre, il diagramma polare di $L(j\omega)$ attraversi una sola volta la circonferenza di raggio unitario e centro nell'origine, come mostrato nella Figura 9.23. Si osservi tra l'altro che, essendo $L(s)$ strettamente propria, quest'ultima ipotesi implica che l'attraversamento avvenga dall'esterno verso l'interno della circonferenza. Indicando con C il punto di intersezione, la corrispondente pulsazione ω_c viene chiamata *pulsazione critica* e rappresenta la pulsazione alla quale il modulo di $L(j\omega)$ è unitario, mentre l'angolo φ_c è detto *fase critica*. Si supponga che, come nella Figura 9.23, risulti $|\varphi_c| < 180^\circ$. Il sistema retroazionato risulta allora asintoticamente stabile, dato che il diagramma di Nyquist di $L(s)$ non può in nessun caso circondare il punto -1 . Si può a questo punto definire un nuovo indicatore di robustezza, associato alla distanza lungo la circonferenza tra il punto C e il punto -1 . Precisamente si chiama *margine di fase* del sistema retroazionato il valore $\varphi_m = 180^\circ - |\varphi_c|$, che può quindi essere calcolato attraverso le seguenti relazioni

$$\varphi_m = 180^\circ - |\varphi_c| \quad , \quad \varphi_c = \arg L(j\omega_c) \quad , \quad |L(j\omega_c)| = 1$$

Figura 9.23
Definizione del margine
di fase.

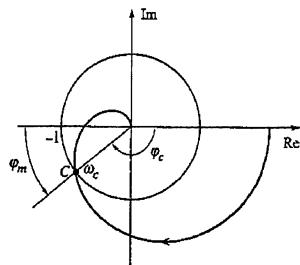
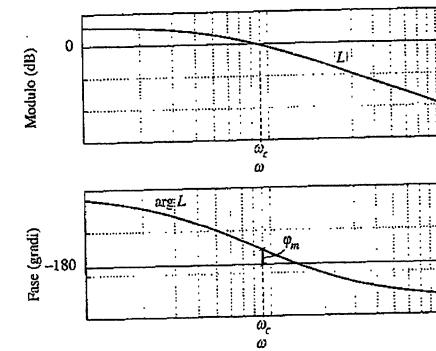


Figura 9.24
Calcolo del margine di fase
dai diagrammi di Bode.



La valutazione del margine di fase φ_m a partire dai diagrammi di Bode di modulo e fase di $L(j\omega)$ è quindi molto agevole ed è esemplificata nella Figura 9.24.

Per ricavare un'interpretazione del parametro φ_m si osservi innanzitutto che, sotto le ipotesi formulate, e in particolare quella sulla posizione del punto C , risotto $\varphi_m > 0^\circ$. Si consideri ora il sistema retroazionato di Figura 9.25, dove τ rappresenta un arbitrario ritardo di tempo. Come si sa, la presenza di un ritardo nella funzione d'anello non altera il valore del modulo, ma introduce alla generica pulsazione ω uno sfasamento negativo pari a $-\omega\tau$ radianti. Se quindi il sistema di Figura 9.9 possiede un margine di fase pari a $\varphi_m > 0^\circ$, il sistema di Figura 9.25 rimane asintoticamente stabile finché lo sfasamento aggiuntivo non porta il punto C a coincidere con il punto -1 . Evidentemente ciò avviene quando $\tau\omega_c = \varphi_m\pi/180$. Il fattore di conversione $\pi/180$ è necessario per tener conto del fatto che la pulsazione ω_c è misurata in radianti per unità di tempo, mentre φ_m è convenzionalmente espresso in gradi. In conclusione, il sistema di Figura 9.25 è asintoticamente stabile per tutti i valori di τ inferiori a $(\varphi_m/\omega_c)\pi/180$. Alla luce di queste considerazioni, il margine di fase φ_m di un sistema retroazionato può essere interpretato come una misura del grado di robustezza della stabilità a fronte di possibili incertezze sulla fase della funzione d'anello in corrispondenza della pulsazione critica, o, equivalentemente, a fronte della presenza di eventuali ritardi di tempo.

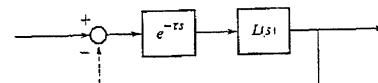


Figura 9.25
Sistema retroazionato
con ritardo.

Esempio 9.10 (Seguito dell'Esempio 9.9) Con riferimento al sistema retroazionato già considerato nell'Esempio 9.9, il valore della pulsazione critica ω_c si determina risolvendo l'equazione $|L(j\omega_c)| = 1$, ovvero $(1 + \omega^2)^{3/2} = 4$. Si ricava così $\omega_c \approx 1.23$ e corrispondentemente $\varphi_c = \arg L(j\omega_c) = -3 \arctan(\omega_c) \approx -153^\circ$. Il margine di fase è quindi $\varphi_m \approx 27^\circ$.

9.6.4 Significatività del margine di guadagno e di fase

A differenza di altri indicatori di stabilità robusta che tengono conto di perturbazioni sulla funzione di trasferimento d'anello appartenenti a classi relativamente ampie (come per esempio il margine di stabilità vettoriale Δ_m discusso nel Paragrafo 9.6.1), il margine di guadagno e il margine di fase sono indicatori più

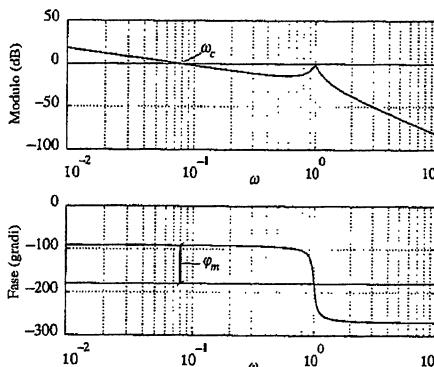
specifici, nel senso che si limitano a caratterizzare il grado di robustezza di un sistema retroazionato nei confronti di variazioni rispettivamente del modulo e della fase della funzione d'anello. In altri termini, entrambi sono basati sull'analisi di un singolo punto della risposta in frequenza. È però vero che, nella maggioranza dei casi di interesse pratico, la valutazione congiunta di questi due indicatori permette una ragionevole stima dell'entità delle perturbazioni di qualunque natura che il sistema può tollerare prima di diventare instabile. Si possono invece facilmente immaginare situazioni in cui una pericolosa vicinanza all'instabilità non viene adeguatamente evidenziata dal calcolo singolo di k_m o di φ_m . L'esempio che segue illustra una di queste circostanze.

Esempio 9.11 La funzione d'anello del sistema retroazionato di Figura 9.9 sia

$$L(s) = \frac{0.08}{s(1 + 0.1s + s^2)} \quad (9.15)$$

Come si può osservare dai diagrammi di Bode riportati nella Figura 9.26, il margine di fase è molto elevato ($\varphi_m \approx 90^\circ$) ma, a causa dell'ampio picco di risonanza dovuto ai poli complessi,

Figura 9.26
Diagrammi di Bode associati alla (9.15).



anche un lieve aumento del guadagno d'anello può rendere instabile il sistema. In questo esempio è solo il margine di guadagno, che in effetti vale $k_m = 1.25$, a mettere in luce questo ridotto grado di robustezza. Come si nota dalla Figura 9.27, il diagramma polare associato a $L(s)$, pur intersecando la circonferenza unitaria molto lontano dal punto -1 , si avvicina pericolosamente a tale punto per valori di ω prossimi a 1. Quindi, oltre all'incremento del guadagno, anche una piccola diminuzione dello smorzamento associato ai poli complessi potrebbe pregiudicare la stabilità del sistema retroazionato.

Figura 9.27
Diagramma polare associato alla (9.15).

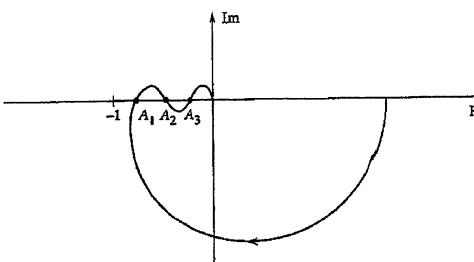
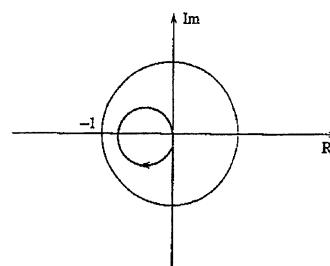


Figura 9.28
Diagramma polare con intersezioni multiple del semiasse reale negativo.

Le definizioni di margine di guadagno e margine di fase potrebbero essere estese, con opportune precisazioni, anche a situazioni come quelle mostrate nelle Figure 9.28 e 9.29, nelle quali il diagramma polare di $L(j\omega)$ attraversa più di una volta il semiasse reale negativo o la circonferenza di raggio unitario. Per conservare il significato di k_m e φ_m in termini di stabilità robusta, basterebbe precisare quale tra i punti di attraversamento deve essere preso in considerazione per il loro calcolo. È del tutto evidente, per esempio, che nel caso della Figura 9.28 il margine di guadagno k_m andrebbe valutato con riferimento al punto A_1 (ovvero $k_m = |x_{A_1}|^{-1}$ se x_{A_1} rappresenta l'ascissa di tale punto), mentre nel caso della Figura 9.29 il margine di fase andrebbe valutato con riferimento al punto C_3 (ovvero $\varphi_m = 180^\circ - |\varphi_3|$).

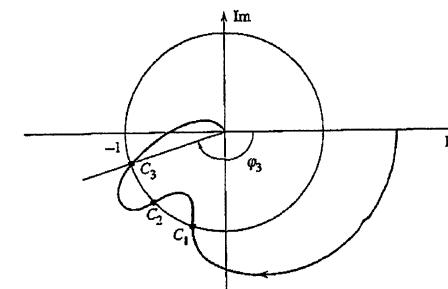


Figura 9.29
Diagramma polare con intersezioni multiple della circonferenza unitaria.

Si può infine osservare che, se il diagramma polare di $L(j\omega)$ non presenta punti di attraversamento con il semiasse reale negativo al di fuori dell'origine, come per esempio nella Figura 9.30, il sistema retroazionato rimane asintoticamente stabile anche in presenza di variazioni arbitrariamente elevate del guadagno d'anello (si ricordi a tale proposito il Corollario 9.4). Appare quindi ragionevole affermare che in casi come questo il sistema presenta un margine di guadagno infinito. Analogamente, se il diagramma polare di $L(j\omega)$ è interamente contenuto all'interno

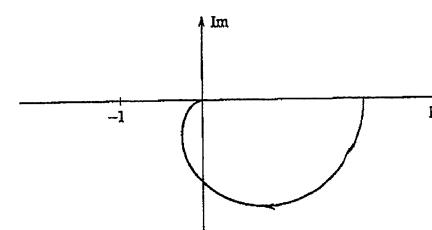
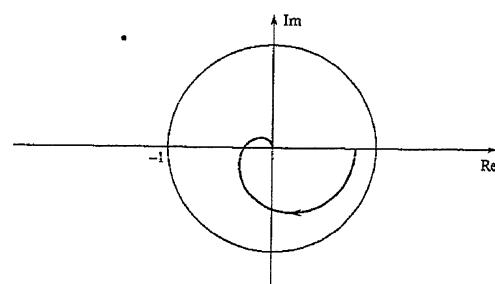


Figura 9.30
Diagramma polare privo di intersezioni con il semiasse reale negativo ($k_m = \infty$).

Figura 9.31
Diagramma polare privo di intersezioni con la circonferenza unitaria ($\varphi_m = \infty$).



della circonferenza unitaria, come nell'esempio di Figura 9.31, la presenza di un ritardo di tempo, per quanto grande esso sia, non è in grado di rendere instabile il sistema (si ricordi il Corollario 9.3). È lecito quindi in tal caso attribuire valore infinito al margine di fase.

9.6.5 Criterio di Bode

In generale, per verificare l'asintotica stabilità del sistema retroazionato di Figura 9.9 mediante il criterio di Nyquist (Teorema 9.1) è necessario tracciare il diagramma di Nyquist della funzione d'anello $L(s)$ e determinare il numero di giri che esso effettua intorno al punto -1. Quando però $L(s)$ non possiede poli a parte reale positiva ($P = 0$), la condizione espressa dal criterio si riduce a imporre che il punto critico -1 non venga circondato dalla curva. Tale condizione può essere verificata attraverso procedimenti che non richiedono il tracciamento del diagramma di Nyquist, ma si basano su un'opportuna analisi dei corrispondenti diagrammi di Bode. A questo scopo si possono utilizzare i concetti di margine di guadagno e margine di fase sopra definiti per giungere a formulare criteri di stabilità di applicazione molto più diretta, anche se validi solo per classi ristrette di sistemi. Il risultato espresso dal seguente teorema rappresenta uno di questi criteri, noto come *criterio di Bode*.

Teorema 9.2 Si consideri il sistema retroazionato di Figura 9.9 e si supponga che:

- $L(s)$ non abbia poli con parte reale maggiore di zero ($P = 0$);
- il diagramma di Bode del modulo di $L(j\omega)$ attraversi una sola volta l'asse a 0 dB.

Allora, indicando con μ il guadagno di $L(s)$ e con φ_m il margine di fase, condizione necessaria e sufficiente perché il sistema sia asintoticamente stabile è che risulti $\mu > 0$ e $\varphi_m > 0^\circ$. ■

La dimostrazione di questo risultato è immediata pur di riconoscere che, nelle ipotesi fatte, le due condizioni di positività su μ e φ_m equivalgono a imporre che il diagramma di Nyquist di $L(s)$ non compia giri intorno al punto -1. Si osservi in particolare che l'ipotesi che $L(s)$ sia strettamente propria implica che l'intersezione con l'asse a 0 dB della curva $|L(j\omega)|_{dB}$, se unica, avvenga dall'alto verso il basso. Inoltre la condizione (b) garantisce che sia univocamente definita la pulsazione critica ω_c , in corrispondenza della quale risulta $|L(j\omega_c)| = 1$, e dunque sia ben definito anche il margine di fase.

Esempio 9.12 Si desideri verificare l'asintotica stabilità del sistema di Figura 9.9 quando

$$L(s) = \frac{50e^{-0.1s}}{(1+s)(1+10s)} \quad (9.16)$$

I diagrammi di Bode di modulo e fase associati a $L(s)$ sono tracciati nella Figura 9.32. Si noti che il sistema soddisfa entrambe le condizioni (a) e (b) di applicabilità del criterio di Bode e inoltre è $\mu = 50 > 0$. Per affermare che il sistema retroazionato è asintoticamente stabile resta solo da verificare che il margine di fase φ_m sia positivo. Dalla Figura 9.32 si deduce che è $\omega_c \approx 2$ e $\varphi_c = -\arctan(\omega_c) - \arctan(10\omega_c) - 0.1\omega_c 180/\pi \approx -162^\circ$, e pertanto il margine di fase è $\varphi_m \approx 18^\circ > 0^\circ$.

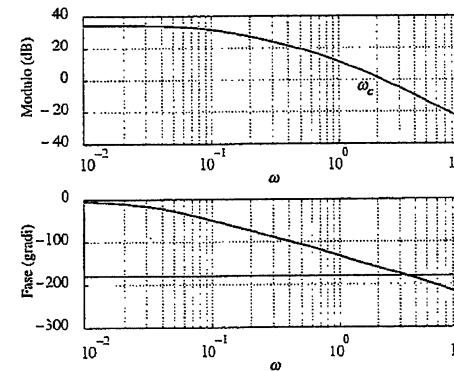


Figura 9.32
Diagrammi di Bode associati alla (9.16).

Sistemi a sfasamento minimo Se oltre a soddisfare le condizioni (a) e (b) del Teorema 9.2, la funzione d'anello $L(s)$ è a sfasamento minimo (e cioè, come visto al Paragrafo 6.6.4, non contiene ritardi di tempo e anche i suoi zeri, oltre ai poli, giacciono nel semipiano sinistro chiuso), esiste la possibilità di formulare una congettura attendibile sul segno del margine di fase φ_m da un esame diretto del diagramma asintotico del modulo di $L(s)$. Infatti se l'attraversamento di tale diagramma con l'asse a 0 dB avviene in un tratto con pendenza $-k$, in virtù dell'ipotesi di fase minima il diagramma asintotico della fase in corrispondenza della pulsazione critica assume il valore $-k90^\circ$ (si veda il Paragrafo 6.6.4). Alla luce di questa considerazione, pur ricordando che a volte si riscontra una scarsa aderenza tra il diagramma asintotico della fase e quello effettivo, soprattutto quando le singolarità della funzione hanno pulsazioni molto ravvicinate, appare allora altamente improbabile che il valore assoluto della fase critica φ_c risulti inferiore a 180° (e dunque sia $\varphi_m > 0^\circ$) tutte le volte che k è maggiore di 2. La situazione più favorevole ai fini dell'ottenimento di un elevato margine di fase è quella in cui la pendenza di attraversamento sia pari a -1 e il diagramma conservi tale pendenza su un intervallo di pulsazioni sufficientemente ampio (per esempio almeno una decade) centrato su ω_c . Questa osservazione, pur con la debita cautela suggerita dalla natura euristica del ragionamento, è comunque subordinata a un'effettiva verifica della sua validità, sarà impiegata più avanti quando si tratterà del problema di progetto di sistemi retroazionati sotto l'ipotesi di sfasamento minimo.

9.7 Complementi sulla stabilità in condizioni perturate

In questo paragrafo il tema della stabilità robusta viene approfondito con maggiore dettaglio, mostrando in particolare il ruolo svolto dai cosiddetti *modelli dell'incertezza*. Descrivendo in modo preciso l'insieme delle perturbazioni ammissibili sul sistema sotto controllo, si riescono in molti casi a enunciare semplici condizioni sotto le quali la stabilità è assicurata per ogni perturbazione ammessa. Risultati di questo genere vengono di solito chiamati *criteri di stabilità robusta*.

9.7.1 Criteri di stabilità robusta

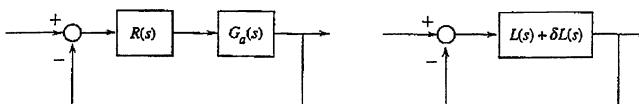
Per analizzare la proprietà di stabilità robusta del sistema retroazionato di Figura 9.9, si assuma che il sistema sotto controllo sia descritto dalla funzione di trasferimento

$$G_a(s) = G(s) + \delta G(s)$$

dove $G(s)$ è la funzione di trasferimento nominale e $\delta G(s)$ rappresenta una perturbazione incognita. L'incertezza descritta dal termine $\delta G(s)$ viene in questo caso chiamata *additiva*. Si ipotizzi inoltre che il numero di poli a parte reale maggiore di zero di $G_a(s)$ sia uguale a quello di $G(s)$ e che in nessun caso avvengano cancellazioni di singolarità nel semipiano destro chiuso tra $R(s)$ e $G_a(s)$. Infine, si supponga di conoscere una funzione maggiorante $\gamma(\omega)$ tale che

$$|\delta G(j\omega)| \leq \gamma(\omega) \quad \forall \omega \quad (9.17)$$

La funzione $\gamma(\omega)$, ricavata per esempio da ipotesi *a priori* sull'incertezza del modello oppure da rilevazioni sperimentali, dà una misura delle massime perturbazioni cui può essere soggetto il modulo della risposta in frequenza al variare della pulsazione.



Si definisca ora la perturbazione $\delta L(s) = R(s)\delta G(s)$ della funzione di trasferimento d'anello nominale $L(s) = R(s)G(s)$ e si ipotizzi che il sistema retroazionato nominale di Figura 9.9 sia asintoticamente stabile, cioè, con riferimento alla funzione $L(s)$, risulti $N = P$. Poiché si è supposto che $G(s)$ e $G_a(s)$ abbiano lo stesso numero di poli a parte reale positiva, il sistema retroazionato perturbato di Figura 9.33 è anch'esso asintoticamente stabile purché il numero di giri N_a compiuti dal diagramma di Nyquist della funzione d'anello $L_a(s) = L(s) + \delta L(s)$ sia uguale a N . Si osservi ora che $|1 + L(j\omega)|$ è la distanza tra il punto $L(j\omega)$ e il punto -1 nel piano complesso (si veda la Figura 9.34) e che, poiché $|\delta L(j\omega)| \leq |R(j\omega)|\gamma(\omega)$, $\forall \omega$, il punto $L_a(j\omega)$ giace certamente all'interno di una circonferenza centrale in $L(j\omega)$ e di raggio $|R(j\omega)|\gamma(\omega)$. Quindi, affinché N_a sia uguale a N per ogni possibile perturbazione δG che soddisfi la (9.17), è necessario e sufficiente che per ogni valore di ω tale circonferenza non contenga il punto -1 , cioè che sia soddisfatta la relazione

$$|R(j\omega)|\gamma(\omega) < |1 + L(j\omega)| \quad \forall \omega \quad (9.18)$$

Si veda ancora la Figura 9.34 per una rappresentazione grafica di questa condizione.

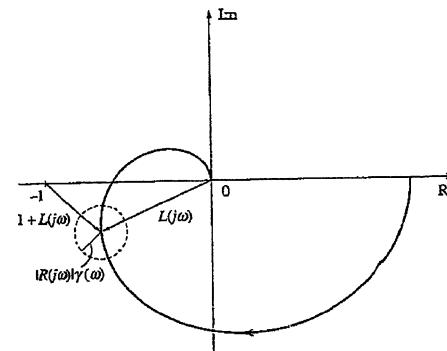


Figura 9.34
Condizione di robustezza (9.18).

La condizione (9.18), oltre a essere di utilità in fase di analisi una volta note $R(s)$ e una stima dell'incertezza massima $\gamma(\omega)$, può rappresentare uno strumento efficace nella definizione delle specifiche che deve possedere il sistema di controllo. In particolare, definendo la funzione $S(s) = (1 + L(s))^{-1}$ (che nel capitolo successivo sarà chiamata *funzione di sensitività*), la (9.18) mostra che, per garantire la stabilità asintotica del sistema perturbato per tutte le perturbazioni che soddisfano la (9.17), è necessario e sufficiente che risulti

$$|S(j\omega)| < \frac{1}{|R(j\omega)|\gamma(\omega)} \quad \forall \omega$$

e quindi che il modulo della risposta in frequenza della funzione di sensitività $S(s)$ sia "piccolo" dove l'incertezza $\gamma(\omega)$ è "grande". Più in generale, per avere stabilità robusta con ampi margini di incertezza è necessario progettare il regolatore $R(s)$ in modo da rendere piccola la quantità

$$S_M = \sup_{\omega} |S(j\omega)|$$

Si noti tra l'altro che S_M coincide con l'inverso del margine di stabilità vettoriale Δ_m definito nel Paragrafo 9.6.1. Infatti risulta

$$S_M = \sup_{\omega} |S(j\omega)| = \left(\inf_{\omega} |1 + L(j\omega)| \right)^{-1} = \Delta_m^{-1}$$

La (9.18) può essere ulteriormente elaborata con riferimento alla funzione $F(s) = L(s)(1 + L(s))^{-1}$, che nel seguito verrà chiamata *funzione di sensitività complementare*. Dividendo entrambi i termini della diseguaglianza per $|R(j\omega)G(j\omega)|$, la condizione (9.18) può essere riscritta nella forma

$$\frac{\gamma(\omega)}{|G(j\omega)|} < \frac{1}{|F(j\omega)|} \quad \forall \omega \quad (9.19)$$

La (9.19) mostra come l'incertezza relativa tollerabile $\gamma(\omega)/|G(j\omega)|$ dipenda dal modulo della risposta in frequenza $F(j\omega)$. In particolare, per garantire la stabilità asintotica anche a fronte di un elevato grado di incertezza relativa in una certa banda di pulsazioni è necessario fare sì che in questa banda il modulo di $F(j\omega)$ sia sufficientemente piccolo. Più in generale, per avere stabilità robusta con ampi

margini di incertezza il regolatore $R(s)$ deve essere progettato in modo da rendere piccolo il valore

$$F_M = \sup_{\omega} |F(j\omega)| \quad (9.20)$$

Esempio 9.13 Si consideri il sistema di Figura 9.9 con funzione di trasferimento nominale del processo data da

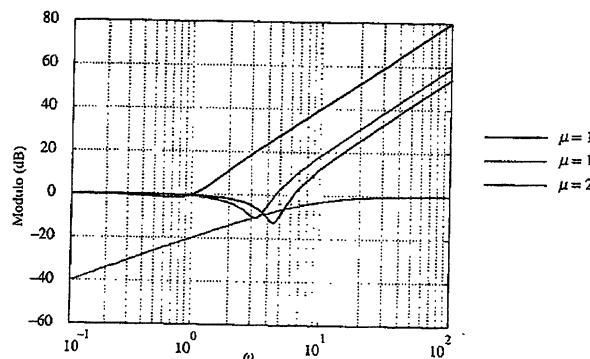
$$G(s) = \frac{1}{s(1+s)} \quad (9.21)$$

e regolatore $R(s) = \mu$. È facile verificare mediante il criterio di Nyquist (o il Corollario 9.4) che il sistema nominale in anello chiuso è asintoticamente stabile per ogni valore di $\mu > 0$. Inoltre risulta

$$F(s) = \frac{\mu}{s^2 + s + \mu} \quad (9.22)$$

Il diagramma di Bode di $1/|F(j\omega)|$ è riportato nella Figura 9.35 per diversi valori di μ . Poiché la condizione di stabilità robusta (9.19) corrisponde a richiedere che il diagramma di Bode di $\gamma(\omega)/|G(j\omega)|$ sia al di sotto di quello di $1/|F(j\omega)|$, la Figura 9.35 mostra che l'incertezza tollerabile in alta frequenza decresce all'aumentare di μ . Si osservi in particolare che, al crescere di μ , aumenta il valore del parametro F_M definito dalla (9.20).

Figura 9.35
Diagramma di Bode di $1/|F(j\omega)|$, con $F(s)$ data dalla (9.22) per diversi valori di μ , e diagramma di Bode del modulo della risposta in frequenza associata alla (9.24) (grigio chiaro).



Si supponga ora che il sistema perturbato differisca da quello nominale per la presenza di un polo ad alta frequenza, cioè risulti

$$G_a(s) = \frac{1}{s(1+s)(1+0.1s)} \quad (9.23)$$

Mediane il criterio di Routh si può verificare che il sistema di controllo retroazionato con funzione d'anello $R(s)G_a(s)$ è asintoticamente stabile purché $\mu < 11$. Dalle (9.21) e (9.23) segue anche che

$$\frac{\delta G(s)}{G(s)} = \frac{-0.1s}{1+0.1s} \quad (9.24)$$

Il diagramma di Bode del modulo della risposta in frequenza associata alla (9.24) è anch'esso riportato nella Figura 9.35. Si può verificare che per $\mu = 1$ la condizione di stabilità robusta (9.19) è soddisfatta, al contrario di quanto succede per $\mu = 11$ e ancor più per $\mu = 20$, come lecito attendersi in base alle considerazioni precedenti.

9.7.2 Legami tra indicatori di robustezza

Può essere interessante studiare più a fondo i legami tra i due parametri k_m (margine di guadagno) e φ_m (margine di fase) e gli indicatori di stabilità robusta introdotti

nel paragrafo precedente e definiti come

$$F_M = \sup_{\omega} |F(j\omega)| = \sup_{\omega} \frac{|L(j\omega)|}{|1 + L(j\omega)|}$$

$$S_M = \sup_{\omega} |S(j\omega)| = \sup_{\omega} \frac{1}{|1 + L(j\omega)|}$$

Si ricordi che elevati valori di F_M e S_M corrispondono a un basso grado di robustezza. Utilizzando le relazioni (9.14) si può scrivere $L(j\omega_\pi) = -1/k_m$ e pertanto, supponendo $k_m > 1$,

$$|F(j\omega_\pi)| = \frac{|L(j\omega_\pi)|}{|1 + L(j\omega_\pi)|} = \frac{1}{k_m - 1}$$

D'altra parte, per la stessa definizione di F_M , risulta $F_M \geq |F(j\omega_\pi)|$ e quindi si ottiene la diseguaglianza

$$k_m \geq 1 + \frac{1}{F_M} \quad (9.25)$$

Ragionando in modo analogo sulla funzione di sensitività $S(s)$ si ricava

$$S_M \geq |S(j\omega_\pi)| = \frac{1}{|1 + L(j\omega_\pi)|} = \frac{k_m}{k_m - 1}$$

e quindi

$$k_m \geq 1 + \frac{1}{S_M - 1} \quad (9.26)$$

Entrambe le relazioni (9.25), (9.26) mostrano che, in un sistema retroazionato avente un margine di guadagno di poco superiore a 1, i valori di F_M e S_M sono necessariamente elevati e quindi il sistema presenta un modesto grado di robustezza rispetto a incertezze di tipo additivo.

Quanto al margine di fase φ_m , si osservi dalla Figura 9.23 che, quando $\varphi_m > 0^\circ$, risulta

$$|1 + L(j\omega_c)| < \varphi_m \frac{\pi}{180} \quad (9.27)$$

Infatti, il primo membro della diseguaglianza rappresenta la distanza tra il punto C e il punto -1 , mentre il secondo membro è pari alla lunghezza dell'arco di circonferenza di raggio unitario che unisce gli stessi due punti. Utilizzando ancora le definizioni di F_M e S_M , e ricordando che $|L(j\omega_c)| = 1$, dalla (9.27) si ottiene

$$\varphi_m > \frac{1}{F_M} \frac{180}{\pi} \quad (9.28)$$

$$\varphi_m > \frac{1}{S_M} \frac{180}{\pi} \quad (9.29)$$

Anche in questo caso si può notare come valori del margine di fase prossimi a 0° implicino valori elevati sia di F_M sia di S_M .

Le relazioni (9.25), (9.26), (9.28), (9.29) possono essere impiegate anche per determinare valori di F_M e S_M in grado di garantire prefissati margini di guadagno e di fase. Per esempio, quando $F_M = 2$, la (9.25) e la (9.28) assicurano che $k_m \geq 1.5$ e $\varphi_m > 28^\circ$. Quando $S_M = 2$, la (9.26) e la (9.29) assicurano che $k_m \geq 2$ e $\varphi_m > 28^\circ$.

9.8 Conclusioni

Nel presente capitolo si è mostrato che un sistema di controllo in anello chiuso può, sotto opportune ipotesi, essere rappresentato come un sistema retroazionato in cui compaiono solo sottosistemi lineari e stazionari. Tra i requisiti di un sistema di controllo ci si è in particolare concentrati sull'importante proprietà di stabilità asintotica che, a causa della retroazione, non può essere accertata direttamente dall'esame dei poli dei vari sottosistemi. Sono stati quindi illustrati i principali metodi disponibili, che traggono tutta origine dal fondamentale criterio di Nyquist, per valutare le proprietà di stabilità di un sistema di controllo a partire dalla funzione di trasferimento d'anello, sia quando questa è perfettamente nota (cioè in condizioni nominali), sia quando è affetta da incertezza (stabilità robusta).

Esercizi

Esercizio 9.1 Attraverso il criterio di Nyquist si dimostri che un sistema retroazionato negativamente con funzione d'anello $L(s) = \mu/s^2$, $\mu > 0$ non è asintoticamente stabile. Si verifichi poi il risultato mediante il calcolo diretto delle radici dell'equazione caratteristica.

Esercizio 9.2 Mediante il criterio di Nyquist, oppure attraverso il Corollario 9.1, determinare per quali valori di μ il sistema di Figura 9.9 con funzione d'anello

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

risulta asintoticamente stabile.

Esercizio 9.3 Nel sistema di Figura 9.15 si supponga che $\tilde{L}(s)$ sia asintoticamente stabile e che μ sia un parametro positivo incerto, compreso tra μ_{\min} e μ_{\max} . Si dimostri che il sistema è asintoticamente stabile per qualunque valore ammissibile di μ se e solo se il diagramma di Nyquist di $\tilde{L}(s)$ non interseca e non circonda il segmento sull'asse reale avente gli estremi in $-1/\mu_{\min}$ e $-1/\mu_{\max}$.

Esercizio 9.4 Come verifica del Corollario 9.4, si analizzi la stabilità del sistema di Figura 9.9 con funzione d'anello

$$L(s) = \frac{\mu}{1+2\xi s/\omega_n + s^2/\omega_n^2} \quad \mu > 0, \omega_n > 0, 0 < \xi < 1$$

Si consideri poi anche il caso $\xi \geq 1$, che corrisponde a poli di $L(s)$ reali.

Esercizio 9.5 Si discuta, al variare dei parametri μ e τ , la stabilità di un sistema retroazionato negativamente con funzione d'anello

$$L(s) = \frac{\mu e^{-\tau s}}{(1+s)^3}$$

Esercizio 9.6 Calcolare il margine di stabilità vettoriale, il margine di guadagno e il margine di fase associati al sistema retroazionato di Figura 9.9, quando

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

Esercizio 9.7 Si consideri il sistema degli Esempi 9.3 e 9.4 e si discuta in quali casi è applicabile il criterio di Bode. Quando il criterio è applicabile, si confrontino i risultati con quelli ricavati nell'Esempio 9.4.

Esercizio 9.8 Enunciare una versione del criterio di Bode basata sulla valutazione del margine di guadagno, anziché del margine di fase. Si faccia attenzione alla corretta e completa formulazione sia delle condizioni di applicabilità sia delle ipotesi del criterio.

Esercizio 9.9 Con riferimento al sistema considerato negli Esempi 9.9 e 9.10, si calcolino analiticamente (o mediante qualche algoritmo numerico) gli indicatori di stabilità robusta F_M e S_M del Paragrafo 9.7 e si verifichi la correttezza delle diseguaglianze (9.25), (9.26), (9.28) e (9.29).

Esercizio 9.10 Si ripeta l'analisi di robustezza del Paragrafo 9.7.1 considerando il caso di *incertezza moltiplicativa*, cioè assumendo che il sistema sotto controllo sia descritto dalla funzione di trasferimento

$$G_a(s) = G(s)(1+\delta G(s))$$

con

$$|\delta G(j\omega)| \leq \gamma(\omega) \quad \forall \omega$$

Si dimostri in particolare che il sistema perturbato è asintoticamente stabile per tutte le perturbazioni ammissibili se e solo se risulta

$$|F(j\omega)| < \frac{1}{\gamma(\omega)} \quad \forall \omega$$

10.1 Introduzione

Dopo aver ampiamente discusso il problema della stabilità, in questo capitolo ci si occuperà delle prestazioni di un sistema di controllo, cioè di tutte quelle altre caratteristiche che permettono di valutarne l'efficacia da vari punti di vista. Per esempio, è importante poter garantire che, almeno a transitorio esaurito, l'errore tra il segnale di riferimento e la variabile controllata sia piccolo o nullo, o che il sistema dimostri la necessaria prontezza nel far seguire correttamente alla variabile controllata l'andamento prestabilido. Altri requisiti sono la capacità di attenuare o eliminare l'effetto dei disturbi e quella di mantenere contenuti i valori assunti dalla variabile di controllo. È poi estremamente opportuno che, così come per la stabilità, tutti questi requisiti di prestazione siano soddisfatti non soltanto quando il comportamento dei vari elementi del sistema retroazionato è perfettamente noto, cioè in condizioni nominali, ma anche in presenza di incertezze, cioè in condizioni perturbate.

L'analisi di tali proprietà sarà basata sullo studio delle cosiddette funzioni di sensitività, che sono le funzioni di trasferimento tra i segnali esogeni (segnaleti di riferimento e disturbi) e le variabili di interesse (variabile controllata, errore e variabile di controllo); da esse infatti dipendono le caratteristiche del sistema di controllo. Più precisamente, il capitolo è organizzato nei seguenti punti:

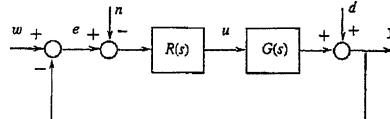
- la definizione delle tre funzioni di sensitività;
- lo studio delle singole funzioni, allo scopo di valutare le prestazioni del sistema di controllo in condizioni nominali;
- alcuni cenni al problema di garantire prestazioni robuste, cioè di assicurare che alcune proprietà continuino a valere anche in condizioni perturbate.

Obiettivo comune a tutta la trattazione è quello di dimostrare come l'esame della maggior parte delle proprietà di interesse di un sistema di controllo possa essere svolto essenzialmente mediante uno studio della funzione di trasferimento d'anello. Questo aspetto risulterà di grande utilità per la sintesi del controllore, di cui si parlerà nei capitoli successivi.

10.2 Funzioni di sensitività e limiti alle prestazioni

A parte il requisito di stabilità, le prestazioni di un sistema di controllo sono state definite nel Paragrafo 9.4.2 in relazione all'andamento delle variabili y , u ed e in risposta a determinati andamenti delle variabili w , d e n . Con riferimento allo schema di Figura 10.1, è quindi utile calcolare le corrispondenti funzioni di trasferimento. A tale scopo si definiscono preliminarmente le seguenti funzioni:

Figura 10.1
Schema di controllo in retroazione.



- *funzione di sensitività*

$$S(s) = \frac{1}{1 + R(s)G(s)} \quad (10.1)$$

- *funzione di sensitività complementare*

$$F(s) = \frac{R(s)G(s)}{1 + R(s)G(s)} \quad (10.2)$$

- *funzione di sensitività del controllo*

$$Q(s) = \frac{R(s)}{1 + R(s)G(s)} = F(s)G(s)^{-1} = R(s)S(s) \quad (10.3)$$

Impiegando le regole di elaborazione degli schemi a blocchi presentate nel Capitolo 5 e in base alle Equazioni (10.1)-(10.3), risulta

$$\begin{bmatrix} Y(s) \\ U(s) \\ E(s) \end{bmatrix} = \begin{bmatrix} F(s) & S(s) & -F(s) \\ Q(s) & -Q(s) & -Q(s) \\ S(s) & -S(s) & F(s) \end{bmatrix} \begin{bmatrix} W(s) \\ D(s) \\ N(s) \end{bmatrix} \quad (10.4)$$

dove, come di consueto, le lettere maiuscole rappresentano le trasformate di Laplace dei relativi segnali. Dall'analisi delle Equazioni (10.1)-(10.4), posto che la proprietà di asintotica stabilità sia verificata, segue immediatamente che, per soddisfare le specifiche sulle prestazioni nominali relative all'andamento della variabile di uscita y e dell'errore e , sarebbe auspicabile ottenere

$$F(s) = 1 \quad (10.5)$$

$$S(s) = 0 \quad (10.6)$$

per garantire l'effettiva coincidenza della variabile controllata y con il riferimento w e un effetto nullo del disturbo d su y . D'altra parte, in base alla (10.4), la (10.5) implicherebbe la totale mancanza di attenuazione dell'effetto del disturbo n sulla variabile di uscita y e sull'errore e . Inoltre, se fosse valida la (10.5), dalla (10.3) risulterebbe $Q(s) = G(s)^{-1}$. Poiché si è ipotizzato che $G(s)$ sia strettamente propria, il modulo di $Q(j\omega)$ risulterebbe crescente per valori di ω sufficientemente elevati. Alla luce della (10.4) ciò comporterebbe l'amplificazione di componenti in alta frequenza dei segnali esogeni sulla variabile di controllo u , in contrasto con il requisito di moderazione del controllo. Queste considerazioni confermano la necessità di ricorrere spesso a un compromesso tra diverse esigenze nella definizione delle specifiche che deve possedere un sistema di controllo.

È interessante inoltre osservare che non è comunque possibile assegnare indipendentemente le funzioni di trasferimento $F(s)$ e $S(s)$ mediante un'opportuna scelta di $R(s)$. In particolare, dalla loro definizione tramite le Equazioni (10.1), (10.2), risulta evidente che queste funzioni sono legate dalla relazione

$$F(s) + S(s) = 1$$

Inoltre, poiché $G(s)$ è strettamente propria e $R(s)$ è al più propria, la funzione d'anello $L(s) = R(s)G(s)$ è anch'essa strettamente propria, ed è facile verificare dalle (10.1), (10.2) che $F(s)$ è strettamente propria e $S(s)$ è propria (ma non strettamente), dato che, per $s \rightarrow \infty$, risulta $F(s) \rightarrow 0$, mentre $S(s) \rightarrow 1$. I requisiti ideali espressi dalle (10.5) e (10.6) sono dunque irrealizzabili.

A conclusione del paragrafo si osservi che $F(s)$ e $S(s)$ dipendono entrambe dalla funzione di trasferimento d'anello $L(s)$. Nel seguito del capitolo si mostrerà come molte specifiche di prestazioni, espresse originariamente in termini di funzioni di sensitività e sensitività complementare, possano essere analizzate più semplicemente a partire da $L(s)$.

10.3 Analisi della funzione di sensitività complementare

La funzione di sensitività complementare $F(s)$ è definita dalla (10.2), ma indicando al solito con $L(s) = R(s)G(s)$ la funzione d'anello, essa può essere più comodamente scritta come

$$F(s) = \frac{L(s)}{1 + L(s)} \quad (10.7)$$

Questa funzione gioca un ruolo importante nel caratterizzare il comportamento del sistema retroazionato di Figura 10.1 in quanto, come si deduce dalla (10.4), rappresenta:

- la funzione di trasferimento tra il riferimento w e l'uscita y ;
- la funzione di trasferimento, cambiata di segno, tra il disturbo n e l'uscita y ;
- la funzione di trasferimento tra il disturbo n e l'errore e .

In questo paragrafo si studieranno le caratteristiche di tale funzione di trasferimento a partire da quelle della funzione d'anello $L(s)$, analizzando i risultati in termini di prestazioni nominali del sistema retroazionato, alla luce delle interpretazioni (a), (b), (c) che si possono attribuire a $F(s)$.

Come già accennato, si può per prima cosa sottolineare che la situazione in cui fosse $F(s) = 1$ (peraltro irrealizzabile data la struttura della (10.7)) sarebbe ideale per quanto riguarda l'obiettivo di far coincidere l'uscita y con il riferimento w , ma non consentirebbe di attenuare in alcun modo l'effetto del disturbo n sull'uscita stessa e sull'errore e . Viceversa, ottenere $F(s) = 0$ significherebbe eliminare completamente l'effetto di questo disturbo, ma anche rendere del tutto inefficace il sistema di controllo. Fortunatamente, nella maggior parte dei casi pratici, l'intervallo di pulsazioni alle quali lo spettro del disturbo n presenta componenti significative è confinato verso le alte frequenze, mentre le specifiche sulle prestazioni di inseguimento del riferimento w da parte dell'uscita riguardano principalmente il comportamento a bassa frequenza. Nel progetto di un sistema di controllo è allora opportuno fare in modo che risulti

- $|F(j\omega)| \approx 1$ nella banda di pulsazioni in cui prevale l'esigenza di inseguimento del riferimento;
- $|F(j\omega)| \approx 0$ alle alte frequenze, dove di norma è concentrata l'energia del disturbo n .

Dopo queste considerazioni di carattere generale, si passa ora a studiare in dettaglio la funzione $F(s)$, prima con riferimento alle caratteristiche statiche e poi a quelle dinamiche.

10.3.1 Analisi statica

Le proprietà statiche della funzione $F(s)$ dipendono essenzialmente dal valore che tale funzione assume per s che tende a zero. Se per la funzione di trasferimento d'anello $L(s)$ si utilizza la generica rappresentazione (4.26), cioè

$$L(s) = \frac{\mu \prod_i (1 + \tau_i s) \prod_i (1 + 2\zeta_i s/\alpha_{ni} + s^2/\alpha_{ni}^2)}{s^g \prod_i (1 + T_i s) \prod_i (1 + 2\zeta_i s/\omega_{ni} + s^2/\omega_{ni}^2)} \quad (10.8)$$

e si ricorda la (10.7), è immediato verificare che risulta

$$\lim_{s \rightarrow 0} F(s) = \lim_{s \rightarrow 0} \frac{\mu}{s^g + \mu}$$

Possono ora presentarsi tre diverse situazioni a seconda del valore del tipo g della funzione d'anello, che viene anche spesso chiamato *tipo del sistema di controllo*. Precisamente, si ha

$$\lim_{s \rightarrow 0} F(s) = \begin{cases} \frac{\mu}{1 + \mu} & g = 0 \\ 1 & g > 0 \\ 0 & g < 0 \end{cases} \quad (10.9)$$

Ingresso a scalino Se si tiene presente che $F(s)$ rappresenta la funzione di trasferimento tra il riferimento w e l'uscita y , la (10.9) permette di valutare le prestazioni statiche del sistema di controllo quando l'ingresso w è uno scalino. Infatti, ammesso che il sistema sia asintoticamente stabile, il valore a transitorio esaurito dell'uscita a fronte dell'ingresso $w(t) = A \text{sca}(t)$ è dato, per il teorema del valore finale, da

$$y_\infty = \lim_{s \rightarrow 0} s F(s) \frac{A}{s} = A \lim_{s \rightarrow 0} F(s)$$

Si possono allora trarre le seguenti conclusioni.

- Nel caso $g > 0$, cioè quando la funzione d'anello possiede almeno un polo nell'origine, il sistema retroazionato fornisce le migliori prestazioni statiche, in quanto il guadagno μ_F della funzione $F(s)$ è unitario e $y_\infty = A = w_\infty$.
- Nel caso $g = 0$, il guadagno $\mu_F = \mu / (1 + \mu)$ è invece prossimo a 1 solo quando il valore assoluto di μ è sufficientemente elevato, e in particolare risulta $0 < \mu_F < 1$ quando μ è positivo. Se $g = 0$ e $\mu = -1$, la funzione $F(s)$ presenta almeno un polo nell'origine e non è asintoticamente stabile.
- Nel caso $g < 0$, $F(s)$ contiene almeno un'azione derivativa e quindi il sistema di controllo non offre buone prestazioni statiche, visto che a transitorio esaurito genera un'uscita nulla in corrispondenza di qualunque valore costante del riferimento w .

Tramite la (10.9) si potrebbe condurre lo stesso tipo di analisi nei riguardi del legame statico tra il disturbo n e l'errore e (o, solo con una differenza di segno, tra il disturbo n e l'uscita y), giungendo però a conclusioni opposte in merito ai giudizi sulla qualità del sistema di controllo. Per esempio, una funzione d'anello di tipo $g > 0$ fa sì che l'errore statico dovuto a un disturbo costante sul trasduttore (per esempio dovuto a una polarizzazione sistematica della misura) abbia la stessa entità del disturbo.

10.3.2 Poli e zeri

Per caratterizzare il comportamento dinamico della funzione $F(s)$, si può far riferimento al calcolo di poli e zeri. Se si rappresenta la funzione d'anello nella forma $L(s) = N_L(s)/D_L(s)$, la (10.7) porta a concludere che

$$F(s) = \frac{N_L(s)}{D_L(s) + N_L(s)}$$

per cui gli zeri di $F(s)$ coincidono con quelli di $L(s)$. In particolare $F(s)$ può avere zeri a parte reale positiva solo se ne ha $L(s)$. Quanto ai poli, essi sono le radici del polinomio $D_L(s) + N_L(s)$ (polinomio caratteristico del sistema retroazionato) e il loro valore non può essere ricavato direttamente dalla conoscenza di poli e zeri di $L(s)$. L'analisi di stabilità presentata nel Capitolo 9 può essere impiegata per verificare che tutti i poli di $F(s)$ giacciono nella regione di asintotica stabilità, ma nulla di più preciso si può dire a questo punto sulla loro esatta collocazione, a meno di effettuarne il calcolo esplicito. Rimandando al Capitolo 12 per un approfondimento di questa problematica mediante il cosiddetto *luogo delle radici*, l'analisi che segue permetterà tra le altre cose di ottenere indicazioni approssimate sulla posizione dei poli dominanti di $F(s)$.

10.3.3 Risposta in frequenza

In questo sottoparagrafo si intende indagare sulle relazioni che legano tra loro i diagrammi di Bode del modulo della risposta in frequenza d'anello $L(j\omega)$ e della funzione $F(j\omega)$. Si supporrà che, in accordo con le condizioni di applicabilità del criterio di Bode, $L(s)$ non abbia poli con parte reale positiva e che il diagramma di $|L(j\omega)|_{\text{dB}}$ presenti un solo attraversamento dell'asse a 0 dB, e abbia dunque l'aspetto mostrato nella Figura 10.2. La pulsazione corrispondente a tale intersezione è la pulsazione critica ω_c . Ricordando la (10.7) e osservando che $|1 + L(j\omega)| \approx |L(j\omega)|$ per $\omega \leq \omega_c$ e $|1 + L(j\omega)| \approx 1$ per $\omega > \omega_c$, si può allora

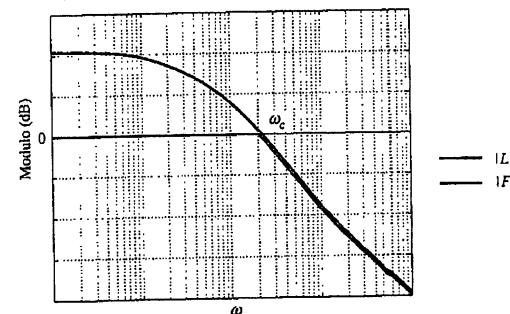


Figura 10.2
Valutazione approssimata di $|F(j\omega)|$.

adottare la seguente approssimazione:

$$|F(j\omega)| = \frac{|L(j\omega)|}{|1 + L(j\omega)|} \approx \begin{cases} 1 & , \omega \leq \omega_c \\ |L(j\omega)| & , \omega > \omega_c \end{cases} \quad (10.10)$$

che si traduce nel diagramma di Bode riportato nella Figura 10.2. Evidentemente, la qualità dell'approssimazione fornita dalla (10.10) può essere scadente in corrispondenza delle pulsazioni nelle quali il diagramma di Bode di $|L(j\omega)|$ è molto vicino all'asse a 0 dB. Sulla base della (10.10) si può concludere che:

- il modulo della risposta in frequenza $F(j\omega)$ appare simile a quello di un filtro passa-basso con guadagno unitario;
- il valore della pulsazione critica ω_c fornisce una stima approssimata dell'estremo superiore della banda passante di tale filtro.

Se si assume che il sistema retroazionato sia asintoticamente stabile e si ricorda che $F(s)$ coincide con la funzione di trasferimento tra w e y , il valore di ω_c rappresenta allora un utile indicatore delle prestazioni del sistema di controllo, in quanto tutte le componenti del segnale di riferimento w a pulsazione minore di ω_c vengono riprodotte fedelmente nel segnale di uscita y , almeno in termini di ampiezza. Naturalmente non va dimenticato che l'identica cosa avviene, a parte un cambio di segno, per tutte le componenti del disturbo n a pulsazione minore di ω_c . Pertanto, nel progetto di un sistema di controllo è in genere conveniente aumentare ω_c per incrementare la capacità del sistema di inseguire variazioni anche veloci del riferimento w , ma non fino al punto di invadere il campo di pulsazioni nelle quali il contributo del disturbo sul trasduttore n è significativo.

Un'altra conclusione che si può dedurre dall'esame della Figura 10.2 riguarda la collocazione dei poli dominanti del sistema retroazionato. Se infatti si presta fede al diagramma approssimato di $|F(j\omega)|$, si osserva come esso presenti un cambio di pendenza negativo proprio in corrispondenza della pulsazione ω_c . Ricordando le regole per il tracciamento dei diagrammi di Bode, è lecito quindi ritenere che la funzione $F(s)$ possieda uno o più poli in vicinanza di tale pulsazione. Poiché inoltre nel diagramma non vi è traccia alcuna di poli a pulsazioni inferiori (il grafico è orizzontale a bassa frequenza) i poli in corrispondenza di ω_c sono quelli dominanti.

Osservazione 10.1 L'ultima affermazione relativa alla collocazione dei poli dominanti richiede qualche precisazione. Infatti l'andamento quasi costante del diagramma del modulo per $\omega < \omega_c$ non esclude che $F(s)$ possa avere un polo a pulsazione minore di ω_c , la cui presenza è mascherata da uno zero a pulsazione molto vicina. Sotto l'ipotesi di asintotica stabilità del sistema retroazionato, il polo ha certamente parte reale negativa e possono allora verificarsi due casi. Se anche lo zero giace nel semipiano sinistro, la presenza della coppia polo/zero può essere in prima approssimazione trascurata, salvo considerare il fenomeno di lenta deriva nella risposta allo scalino associato a cancellazioni non perfette a bassa frequenza. Se invece lo zero si trova nel semipiano destro, la cancellazione non avviene e il polo a bassa frequenza risulterebbe essere quello dominante. Ricordando che gli zeri di $F(s)$ sono anche quelli di $L(s)$, è però facile convincersi che, nelle ipotesi fatte, la pulsazione associata allo zero non può essere sensibilmente inferiore a ω_c . Infatti il contributo negativo alla fase dovuto allo zero, insieme a quello di almeno due poli di $L(s)$ a pulsazione minore di ω_c necessari perché il diagramma di $|L(j\omega)|$ attraversi l'asse a 0 dB, porterebbe a valori negativi del margine di fase, violando l'ipotesi di asintotica stabilità. Dunque, in ogni caso appare corretto ritenere che i poli dominanti di $F(s)$ abbiano pulsazione prossima a ω_c . ■

Osservazione 10.2 Le argomentazioni finora svolte hanno riguardato l'andamento del modulo di $F(j\omega)$, prescindendo da quello della fase. In realtà, perché un filtro passa-basso $F(s)$ produca in uscita una fedele ricostruzione del segnale in ingresso, supposto che questo non presenti componenti significative al di fuori della banda passante, occorre che $\arg F(j\omega)$ sia prossimo a zero all'interno di tale banda, in modo da non provocare eccessive distorsioni di fase. Nel caso del sistema retroazionato in esame ci si deve aspettare, in base alla precedente discussione, che l'andamento della fase di $F(j\omega)$ sia simile a quello associato all'approssimazione a poli dominanti. Si noti quindi che, in vicinanza della pulsazione critica, $\arg F(j\omega)$ può assumere valori anche non irrilevanti. Ciò non è però sufficiente a intaccare la sostanza delle conclusioni prima riportate, permettendo comunque di affermare che, nelle ipotesi fatte, il sistema descritto da $F(s)$ approssima un filtro passa-basso ideale nell'intervallo $[0, \omega_c]$. ■

Esempio 10.1 Si consideri il sistema retroazionato di Figura 10.1 con

$$L(s) = R(s)G(s) = \frac{20}{(1+10s)(1+2s)(1+0.2s)} \quad (10.11)$$

Dai diagrammi di Bode di modulo e fase associati a $L(s)$, mostrati nella Figura 10.3, si deduce che

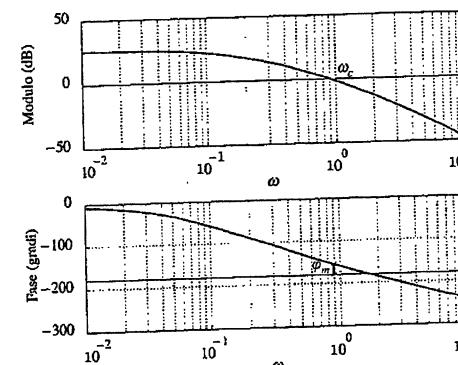


Figura 10.3
Diagrammi di Bode associati alla (10.11).

la pulsazione critica è $\omega_c \approx 0.9$ e il margine di fase vale $\varphi_m \approx 24^\circ$, cosicché, per il criterio di Bode, il sistema retroazionato è asintoticamente stabile. Inoltre, sulla base della (10.10), è lecito prevedere che la funzione di sensitività complementare $F(s)$ si comporti approssimativamente come un filtro passa-basso con guadagno unitario (in realtà dall'analisi statica risulta che $\mu_F = 20/21$) e banda passante $[0, 0.9]$. Il confronto, riportato nella Figura 10.4, tra il diagramma di Bode esatto di $|F(j\omega)|$ e quello derivato dall'approssimazione (10.10) dimostra che la congettura è abbastanza vicina alla realtà, anche se non consente di prevedere la presenza di un picco di risonanza in prossimità della pulsazione critica.

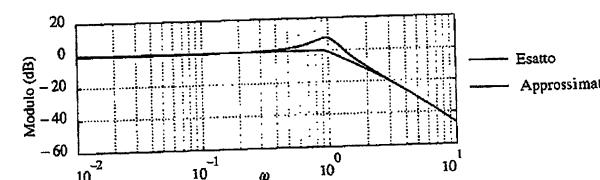


Figura 10.4
Diagramma di Bode di $|F(j\omega)|$ esatto e approssimato nell'Esempio 10.1.

10.3.4 Smorzamento e margine di fase

L'analisi fin qui condotta non permette di trarre altre conclusioni sui poli dominanti di $F(s)$, se non quella relativa alla loro collocazione sull'asse delle pulsazioni in vicinanza di ω_c . In particolare, non si è in grado di affermare alcunché sul fatto che si tratti di poli reali o complessi coniugati. La questione peraltro non è priva di interesse in quanto, come ben si sa, il comportamento dinamico di un sistema dipende in modo essenziale dalle caratteristiche dei suoi poli dominanti (si pensi, per esempio, alla diversa ampiezza e durata delle oscillazioni nella risposta allo scalino al variare dello smorzamento, anche a parità di pulsazione naturale).

Un modo per chiarire tale questione consiste nel valutare l'ampiezza dell'eventuale picco di risonanza presente nel diagramma di Bode di $|F(j\omega)|$ in corrispondenza dei poli dominanti, mettendola poi in relazione con lo smorzamento a essi associato. Supponendo perciò che ω_c rappresenti una buona approssimazione della pulsazione naturale ω_n dei poli dominanti (assunti per il momento complessi coniugati), si determini il valore di $|F(j\omega_c)|$. Tenendo conto che il numero complesso $L(j\omega_c)$ ha modulo unitario e argomento pari a φ_c , e quindi si può scrivere $L(j\omega_c) = e^{j\varphi_c}$, e ricordando che $\varphi_m = 180^\circ - |\varphi_c|$, risulta

$$\begin{aligned} |F(j\omega_c)| &= \frac{|L(j\omega_c)|}{|1 + L(j\omega_c)|} = \frac{1}{|1 + e^{j\varphi_c}|} = \frac{1}{\sqrt{(1 + \cos \varphi_c)^2 + \sin^2 \varphi_c}} = \\ &= \frac{1}{\sqrt{2(1 + \cos \varphi_c)}} = \frac{1}{\sqrt{2(1 - \cos \varphi_m)}} = \frac{1}{2 \sin(\varphi_m/2)} \quad (10.12) \end{aligned}$$

dove l'ultimo passaggio è basato sull'uso di una nota formula trigonometrica sotto l'ipotesi che sia $\varphi_m > 0^\circ$. D'altro canto, se utilizzando l'approssimazione a poli dominanti si assume che $F(s)$ sia una funzione di trasferimento asintoticamente stabile del secondo ordine a guadagno unitario priva di zeri e con poli complessi in $\omega_n = \omega_c$, deve risultare (si veda l'Equazione (6.29))

$$|F(j\omega_c)| = \frac{1}{2\xi} \quad (10.13)$$

dove ξ è lo smorzamento dei poli. Il confronto tra le espressioni (10.12) e (10.13) implica

$$\xi = \sin(\varphi_m/2) \quad (10.14)$$

Questa relazione è particolarmente interessante perché esprime un legame tra un'importante caratteristica del sistema retroazionato, quale è lo smorzamento associato ai poli dominanti, e il parametro φ_m che può essere ottenuto con facilità dall'esame dei diagrammi di Bode della funzione d'anello $L(s)$, e che tra l'altro serve anche in molti casi a giudicare le proprietà di stabilità del sistema di controllo. Naturalmente non va dimenticato che nel ricavare l'Equazione (10.14) sono state fatte varie ipotesi ($\omega_n = \omega_c$ e sistema del secondo ordine, privo di zeri, con guadagno unitario) che non sono sempre esattamente verificate. Il valore dello smorzamento ξ ottenuto dalla (10.14) deve essere dunque inteso come un'indicazione approssimata, tanto che spesso si preferisce usare al posto della (10.14) una diversa espressione ricavata sostituendo alla funzione seno il suo argomento espresso in radianti. Ricordando che φ_m è convenzionalmente misurato in gradi, la (10.14) diventa allora

$$\xi = \frac{\varphi_m \pi}{2 \cdot 180} \simeq \frac{\varphi_m}{100} \quad (10.15)$$

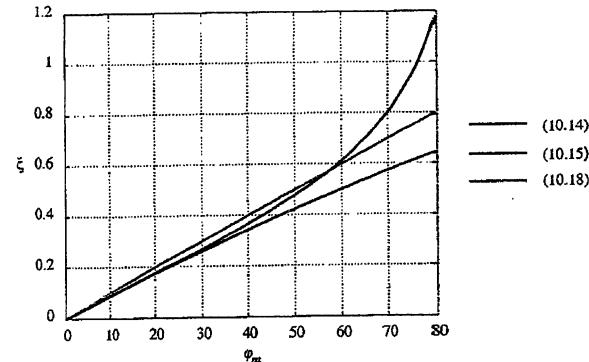


Figura 10.5
Smorzamento ξ in funzione del margine di fase φ_m , secondo le relazioni (10.14), (10.15) e (10.18).

A titolo di confronto, nella Figura 10.5 sono rappresentati i grafici della funzione (10.14) e della sua approssimazione (10.15).

Osservazione 10.3 La dipendenza che è stata messa in luce tra smorzamento e margine di fase ha una semplice giustificazione intuitiva. Innanzitutto, quando $\varphi_m = 0^\circ$, il diagramma polare di $L(s)$ passa per il punto -1 e perciò il sistema retroazionato possiede poli a smorzamento nullo (si veda al riguardo la parte della dimostrazione del criterio di Nyquist relativa al caso in cui N non sia ben definito). Inoltre, un piccolo valore dello smorzamento è associato alla presenza di poli del sistema in anello chiuso particolarmente vicini all'asse immaginario, e dunque suscettibili di diventare instabili in presenza di perturbazioni anche modeste sui parametri del sistema. Allora, nella misura in cui il margine di fase costituisce un indicatore di stabilità robusta del sistema retroazionato, è lecito aspettarsi che esso assuma valori decisamente bassi in una situazione di questo tipo. Il legame espresso dalla (10.15) appare dunque pienamente giustificato. ■

In realtà si può ritenere che la (10.15) fornisca una valutazione attendibile di ξ solo fino a valori di φ_m intorno a 75° . Si vedrà tra l'altro più avanti che, per elevati valori del margine di fase, la pulsazione critica ω_c rappresenta una stima per difetto della pulsazione a cui si trovano i poli dominanti, e dunque l'analisi precedente perde gran parte della sua validità. Quando $\varphi_m > 75^\circ$, è lecito in effetti congetturare che $F(s)$ possieda un polo dominante reale, e in tal caso ω_c costituisce una stima dell'inverso della costante di tempo associata. Il seguente esempio fornisce una giustificazione empirica di questo risultato.

Esempio 10.2 Nel sistema di Figura 10.1 con

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

la funzione di sensitività complementare vale

$$F(s) = \frac{\mu}{s^2 + s + \mu}$$

e lo smorzamento dei poli in anello chiuso è quindi

$$\xi = \frac{1}{2\sqrt{\mu}} \quad (10.16)$$

D'altra parte, è facile verificare analiticamente che la pulsazione critica ω_c e il margine di fase φ_m valgono rispettivamente

$$\omega_c = \sqrt{-\frac{1}{2} + \sqrt{\frac{1}{4} + \mu^2}}, \quad \varphi_m = \arctan\left(\frac{1}{\omega_c}\right) \quad (10.17)$$

Confrontando le Equazioni (10.16) e (10.17), si ricava che il legame tra ξ e φ_m è espresso in questo caso dalla relazione

$$\xi = \frac{1}{2} \frac{\sin(\varphi_m)}{\sqrt{\cos(\varphi_m)}} \quad (10.18)$$

la cui rappresentazione grafica è riportata nella Figura 10.5 insieme alle precedenti approssimazioni (10.14) e (10.15). Si noti in particolare che, in base alla (10.18), lo smorzamento raggiunge il valore $\xi = 1$, e quindi i poli di $F(s)$ diventano reali, per valori del margine di fase prossimi a 75° .

In definitiva, è lecito adottare la seguente regola approssimata:

- se $\varphi_m > 75^\circ$, allora $F(s)$ possiede un polo dominante reale con costante di tempo $\tau \approx 1/\omega_c$;
- se $0 < \varphi_m \leq 75^\circ$, allora $F(s)$ possiede una coppia di poli dominanti complessi con pulsazione naturale $\omega_n \approx \omega_c$ e smorzamento $\xi \approx \varphi_m/100$.

Esempio 10.3 (Seguito dell'Esempio 10.1) Per inferire le caratteristiche dei poli dominanti di $F(s)$, si identifichi la loro pulsazione naturale ω_n con la pulsazione critica $\omega_c \approx 0.9$ e si utilizzi la (10.15) per calcolare lo smorzamento come $\xi \approx 0.24$. Con opportuni strumenti di calcolo si mostrerebbe invece che i poli di $F(s)$ valgono $s_{1,2} = -0.196 \pm j0.985$ e $s_3 = -5.208$. La pulsazione naturale dei poli dominanti $s_{1,2}$ è quindi $\omega_n \approx 1$ e il loro smorzamento vale $\xi \approx 0.2$. Il livello di approssimazione appare dunque accettabile.

10.3.5 Valutazione esatta della banda passante

Una volta accertato attraverso l'analisi approssimata precedentemente svolta che, in molti casi significativi, la funzione $F(s)$ presenta un comportamento da filtro passa-basso, è possibile valutare in modo più preciso la banda passante utilizzando il diagramma polare di $L(j\omega)$. A tale scopo si ricordi che, per un filtro passa-basso $F(s)$ a guadagno unitario, l'intervallo di pulsazioni che costituisce la banda passante è definito dalla relazione (Paragrafo 6.8.1)

$$\frac{1}{\sqrt{2}} \leq |F(j\omega)| \leq \sqrt{2} \quad (10.19)$$

Inoltre, in base alla (10.7), si ha

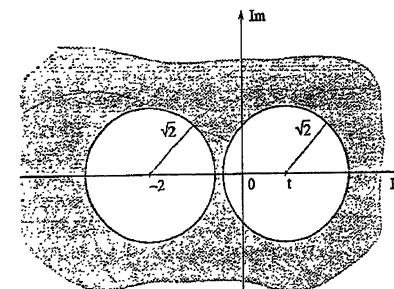
$$|F(j\omega)| = \frac{|L(j\omega)|}{|1 + L(j\omega)|} \quad (10.20)$$

D'altra parte, è facile verificare che l'insieme dei punti del piano complesso per i quali risulta

$$\frac{1}{\sqrt{2}} \leq \frac{|s|}{|1 + s|} \leq \sqrt{2} \quad (10.21)$$

è costituito dalla regione, indicata nella Figura 10.6, esterna alle due circonferenze ivi riportate. In effetti il luogo dei punti per i quali $|1 + s| = \sqrt{2}|s|$ è la circonferenza di centro 1 e raggio $\sqrt{2}$, mentre il luogo dei punti per i quali $|s| = \sqrt{2}|1 + s|$ è la circonferenza di centro -2 e raggio $\sqrt{2}$. Inoltre, le due rela-

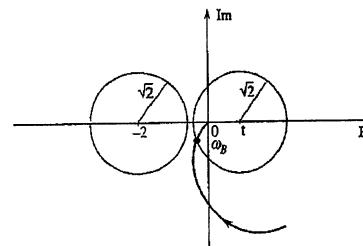
Figura 10.6
Luogo dei punti per cui vale la (10.21).



zioni precedenti scritte con il segno di \geq sono verificate da tutti i valori complessi interni rispettivamente al primo e al secondo dei due cerchi.

In base alla (10.20), è allora immediato concludere che i valori di ω per i quali la condizione (10.19) è verificata sono quelli per cui il valore $L(j\omega)$ cade nella regione colorata della Figura 10.6. Perciò, sovrapponendo alla stessa figura il diagramma polare di $L(j\omega)$ e supponendo di disporre della sua punteggiatura in ω , si ha la possibilità di determinare esattamente la banda passante del sistema retroazionato. A titolo di esempio, si supponga che il diagramma polare di $L(j\omega)$ sia quello mostrato nella Figura 10.7. Si dedurrebbe in tal caso che la banda passante di $F(s)$ è data dall'intervallo $[0, \omega_B]$.

Figura 10.7
Esempio di valutazione della banda passante.



Questo tipo di analisi consente di discutere più a fondo la validità delle approssimazioni introdotte in precedenza. A tale scopo si osservi la Figura 10.8, dove, oltre alle due circonferenze della Figura 10.6, è anche tracciata la circonferenza di raggio unitario e centro nell'origine. Nella figura è anche riportato il diagramma polare di un'ipotetica funzione $L(j\omega)$. Con riferimento ai punti P , Q e R indicati nella figura e con banali argomentazioni geometriche, si trova che l'angolo φ_P ha un'ampiezza pari a $2\arcsin(1/2\sqrt{2}) \approx 41^\circ$, che l'ascissa del punto Q è $x_Q = -2 + \sqrt{2} \approx -0.6$, e che il punto R di intersezione tra la circonferenza unitaria e quella di centro 1 e raggio $\sqrt{2}$ giace sull'asse immaginario. Per prima cosa, è quindi immediato concludere che condizione necessaria perché risulti $|F(j\omega)| \leq \sqrt{2}, \forall \omega$, è che sia

$$\varphi_m \geq \varphi_P \approx 41^\circ, \quad k_m \geq |x_Q|^{-1} \approx 1.7$$

In caso contrario, il diagramma di Bode di $|F(j\omega)|$ presenta un significativo picco di risonanza, coerentemente a quanto affermato in precedenza sullo smorzamento dei poli dominanti quando il margine di fase è inferiore a 40° . Dalla stessa

Figura 10.8
Esempio di valutazione della banda passante e legami con la pulsazione critica ω_c .

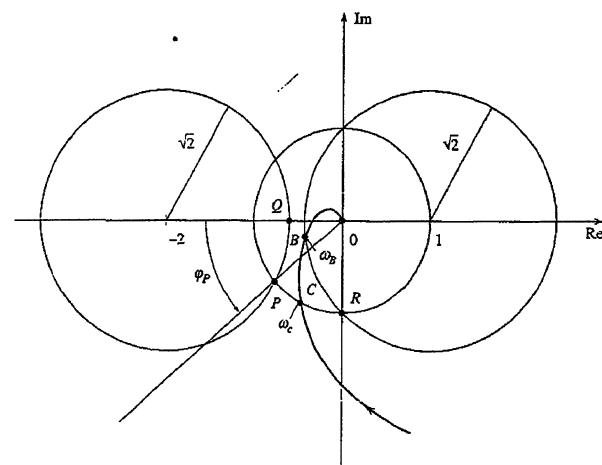


figura si osserva che, quando $41^\circ \leq \varphi_m < 90^\circ$, il valore ω_c della pulsazione critica rappresenta una stima per difetto dell'estremo superiore della banda passante, che è pari a ω_B nel caso illustrato in figura.

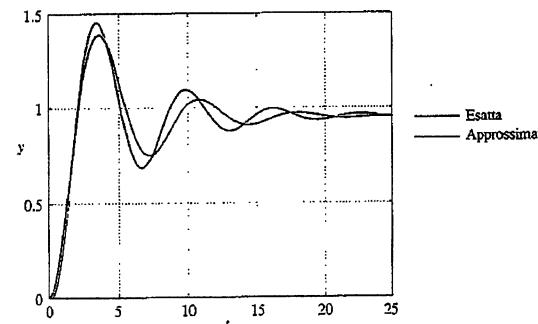
10.3.6 Risposta allo scalino

Dato che molto spesso i requisiti di un sistema di controllo sono formulati in relazione all'andamento desiderato della risposta a variazioni a scalino dei segnali esogeni, è conveniente reinterpretare sotto questo diverso punto di vista i risultati ottenuti ai punti precedenti, evidenziando come si possano ricavare valutazioni, anche approssimate, di alcuni parametri caratteristici della risposta allo scalino di $F(s)$ a partire dallo studio della funzione d'anello $L(s)$.

Assumendo che il sistema retroazionato sia asintoticamente stabile, già si è visto come calcolare il valore di regime della risposta allo scalino (si veda l'analisi statica, Paragrafo 10.3.1). Altre informazioni riguardanti il tempo di assettamento T_{as} , l'ampiezza percentuale massima della sovraelongazione $S\%$ e il periodo T_p delle eventuali oscillazioni (distanza temporale tra due picchi successivi) possono essere ricavate sostituendo a $F(s)$ la sua approssimazione a poli dominanti, ottenuta con i metodi precedentemente illustrati. È importante sottolineare come le caratteristiche di questi poli dominanti (costanti di tempo, se reali, pulsazione naturale e smorzamento, se complessi) possono essere dedotte con estrema facilità, anche se solo in prima approssimazione, dai valori della pulsazione critica ω_c e del margine di fase φ_m .

Esempio 10.4 (Seguito degli Esempi 10.1 e 10.3) Si voglia determinare l'andamento qualitativo della risposta allo scalino associata alla funzione $F(s)$. Sulla base dell'analisi statica e dell'approssimazione a poli dominanti, si prevede che tale risposta presenti un andamento oscillante smorzato e si assetti a regime sul valore del guadagno $\mu_F = 20/21 \approx 0.95$. Basandosi inoltre sulla Tabella 4.4 e utilizzando i valori approssimati $\omega_n = 0.9$ e $\xi \approx 0.24$ ricavati in precedenza, si deduce che il tempo di assettamento al 99% è $T_{99\%} \approx 21.3$, il periodo delle oscillazioni è $T_p \approx 7.2$ e la massima sovraelongazione percentuale è $S\% \approx 46$. Dalla Figura 10.9, in cui sono poste a confronto la risposta allo scalino esatta e quella approssimata, si constata che gli errori che si commettono con questo tipo di approssimazione sono decisamente modesti.

Figura 10.9
Risposta allo scalino associata a $F(s)$ esatta e approssimata nell'Esempio 10.4.



10.3.7 Effetto di un ritardo di tempo

Gran parte dei risultati ricavati nei punti precedenti con riferimento a una funzione d'anello $L(s)$ razionale sono generalizzabili anche al caso in cui $L(s)$ contenga un ritardo di tempo, sia cioè del tipo $L(s) = e^{-\tau s} L'(s)$, con $L'(s)$ razionale.

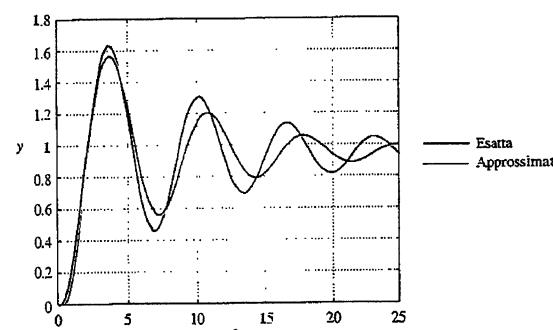
Riguardo all'analisi statica, la presenza del fattore $e^{-\tau s}$ non modifica il calcolo del limite (10.9) e quindi, assumendo che il teorema del valore finale continui a valere anche in presenza di trasformate di Laplace non razionali, i risultati ottenuti restano immutati. Quanto poi alle caratteristiche dinamiche, si osservi che la presenza del ritardo non modifica il valore della pulsazione critica ω_c , ma contribuisce a diminuire di una quantità pari a $\omega_c \tau 180/\pi$ il margine di fase φ_m . Pertanto, anche supponendo che la stabilità asintotica si conservi, la funzione $F(s)$ del sistema retroazionato, pur continuando a comportarsi sostanzialmente come un filtro passa-basso, presenta una dinamica con smorzamento minore.

Esempio 10.5 (Seguito degli Esempi 10.1, 10.3 e 10.4) Si supponga che la funzione d'anello dell'Esempio 10.1, per la presenza di un ritardo di tempo $\tau = 0.2$, si modifichi come segue:

$$L(s) = \frac{20e^{-0.2s}}{(1-10s)(1+2s)(1+0.2s)}$$

La pulsazione critica ω_c rimane naturalmente invariata e pari a $\omega_c \approx 0.9$, ma il margine di fase diminuisce della quantità $\omega_c \tau 180/\pi \approx 10^\circ$, riducendosi a $\varphi_m \approx 14^\circ$. Quindi il sistema retroazionato è ancora asintoticamente stabile, ma la risposta allo scalino di $F(s)$, oltre a essere ritardata di un intervallo pari a τ , contiene oscillazioni più ampie e meno smorzate.

Figura 10.10
Risposta allo scalino associata a $F(s)$ esatta e approssimata nell'Esempio 10.5.



In particolare, assumendo $\omega_n = \omega_c \approx 0.9$ e $\xi \approx \varphi_m/100 \approx 0.14$ si ricavano i seguenti valori dei parametri caratteristici: tempo di assettamento al 99% $T_{d1} \approx 36.5 + 0.2 = 36.7$; periodo delle oscillazioni $T_p \approx 7.05$; massima sovrelongazione percentuale $S\% \approx 64$. Per un utile raffronto con la risposta esatta si osservi la Figura 10.10.

10.4 Analisi della funzione di sensitività

La funzione di sensitività $S(s)$, definita nella (10.1), può anche essere scritta come

$$S(s) = \frac{1}{1 + L(s)} \quad (10.22)$$

e, come si è visto al Paragrafo 10.2, rappresenta contemporaneamente:

- a) la funzione di trasferimento tra il disturbo d e l'uscita y ;
- b) la funzione di trasferimento tra il riferimento w e l'errore e ;
- c) la funzione di trasferimento, cambiata di segno, tra il disturbo d e l'errore e .

Una situazione ideale per le prestazioni del sistema di controllo sarebbe quella in cui fosse $S(s) = 0$, poiché risulterebbe nullo l'effetto sull'errore degli ingressi w e d . D'altra parte, la forma stessa della (10.22) rivela che questo obiettivo è irrealizzabile e inoltre, come già osservato al Paragrafo 10.2, il fatto che $L(s)$ sia strettamente propria fa sì che comunque $S(s) \rightarrow 1$ per $s \rightarrow \infty$. Un requisito più realistico è invece quello di chiedere che la risposta in frequenza $S(j\omega)$ abbia modulo sufficientemente piccolo nella banda di pulsazioni dove si presume che lo spettro dei segnali w e d presenti componenti significative.

L'analisi che segue mostrerà come sia possibile inferire le principali caratteristiche di $S(s)$, e in particolare le sue proprietà filtranti, a partire dalla funzione d'anello $L(s)$.

10.4.1 Analisi statica

Se si assume che la funzione di trasferimento d'anello $L(s)$ abbia la generica forma dell'Equazione (10.8), dalla (10.22) si ricava

$$\lim_{s \rightarrow 0} S(s) = \lim_{s \rightarrow 0} \frac{s^g}{s^g + \mu} = \begin{cases} \frac{1}{1 + \mu}, & g = 0 \\ 0, & g > 0 \\ 1, & g < 0 \end{cases} \quad (10.23)$$

Pertanto, si riconoscono i seguenti casi:

- quando $g > 0$, la funzione $S(s)$ contiene un'azione derivativa;
- quando $g = 0$ e $\mu = -1$, la funzione $S(s)$ contiene un'azione integrale;
- in tutti gli altri casi, il limite calcolato nella (10.23) rappresenta il guadagno statico μ_S della funzione $S(s)$; risulta $\mu_S = 1/(1 + \mu)$, quando $g = 0$ e $\mu \neq -1$, e risulta invece $\mu_S = 1$, quando $g < 0$.

Ingresso a scalino È ora possibile valutare le prestazioni statiche del sistema di controllo calcolando l'effetto a transitorio esaurito sulla variabile errore di un ingresso w a scalino. Si noti che, a parte un cambio di segno, i risultati che si ottengono saranno validi anche per l'errore a transitorio esaurito dovuto a un andamento a scalino del disturbo d . Se si ipotizza che il sistema sia asintoticamente stabile, l'errore a transitorio esaurito quando $w(t) = A \text{ sca}(t)$, per il teorema del valore finale, è pari a

$$e_\infty = \lim_{s \rightarrow 0} s S(s) \frac{A}{s} = A \lim_{s \rightarrow 0} S(s) = \begin{cases} \frac{A}{1 + \mu}, & g = 0 \\ 0, & g > 0 \\ A, & g < 0 \end{cases}$$

Perciò

- se $g > 0$, cioè se la funzione d'anello possiede almeno un polo nell'origine, si ottiene la massima precisione statica perché risulta $e_\infty = 0$;
- se $g = 0$ e il guadagno d'anello μ è positivo, la precisione aumenta al crescere di μ e l'errore a transitorio esaurito può essere reso arbitrariamente piccolo scegliendo μ sufficientemente grande;
- se $g < 0$, cioè nel caso che $L(s)$ contenga un'azione derivativa, l'entità dell'errore è pari a quella dell'ingresso che lo ha generato.

Altri ingressi canonici È utile estendere la precedente analisi al caso in cui l'ingresso w sia un segnale canonico dotato di trasformata di Laplace A/s^i , con i intero positivo. Per $i = 1$ si ricade nel caso dello scalino appena considerato, $i = 2$ corrisponde a un ingresso a rampa, $i = 3$ a un ingresso a parabola e così via. Quando è $W(s) = A/s^i$, $i \geq 2$, dall'applicazione del teorema del valore finale si ricava

$$e_\infty = \lim_{s \rightarrow 0} s S(s) \frac{A}{s^i} = A \lim_{s \rightarrow 0} \frac{s^{g-i+1}}{s^g + \mu} = \begin{cases} \infty, & g < i - 1 \\ \frac{A}{\mu}, & g = i - 1 \\ 0, & g > i - 1 \end{cases}$$

Pertanto

- se $g < i - 1$, l'errore diverge;
- se $g = i - 1$, l'errore a transitorio esaurito è inversamente proporzionale al guadagno d'anello;
- se $g > i - 1$, l'errore è asintoticamente nullo.

Il numero di azioni integrali che devono essere quindi presenti nella funzione d'anello per garantire che e_∞ sia per lo meno limitato aumenta al crescere di i . I risultati dell'analisi sono sinteticamente riportati nella Tabella 10.1.

Per analizzare il comportamento asintotico dell'errore in risposta a un disturbo d di tipo canonico, si può utilizzare la stessa tabella a patto di cambiare segno agli elementi che vi compaiono.

Osservazione 10.4 È importante notare che alcuni dei risultati appena ricavati, in particolare quelli relativi all'annullamento dell'errore a transitorio esaurito in determinate circostanze, sono del tutto indipendenti dall'effettiva espressione della funzione di trasferimento $G(s)$ del sistema sotto controllo. Tali risultati godono quindi della proprietà di *robustezza*, essendo validi non solo in condizioni nominali, ma anche in presenza di incertezza. Si è visto per esempio che, sotto

Tabella 10.1
Errore a regime con ingresso w canonico, al variare del tipo g della funzione d'anello.

	$A_{\text{sc}}(t)$	$A_{\text{ram}}(t)$	$A_{\text{par}}(t)$
$g = 0$	$\frac{A}{1 + \mu}$	∞	∞
$g = 1$	0	$\frac{A}{\mu}$	∞
$g = 2$	0	0	$\frac{A}{\mu}$
$g = 3$	0	0	0

l'ipotesi di asintotica stabilità e con un ingresso a scalino, l'errore tende ad annullarsi asintoticamente purché il tipo g della funzione d'anello sia maggiore di zero (Tabella 10.1). Se dunque il regolatore $R(s)$ contiene almeno un'azione integrale, risulterà $e_\infty = 0$ anche in presenza di incertezze sulla funzione $G(s)$ del sistema da controllare, purché tali da non pregiudicare l'asintotica stabilità in anello chiuso. Si dice quindi che, in tal caso, il sistema di controllo assicura la *regolazione robusta a zero dell'errore*, o, nel caso del disturbo d , la *reiezione robusta* del disturbo a scalino. ■

10.4.2 Poli e zeri

Se si scrive la funzione d'anello nella forma fattorizzata $L(s) = N_L(s)/D_L(s)$, risulta dalla (10.22)

$$S(s) = \frac{D_L(s)}{D_L(s) + N_L(s)}$$

Come è ovvio, i poli di $S(s)$ coincidono con quelli della funzione di sensitività complementare $F(s)$. Delle loro proprietà si è discusso a lungo nel paragrafo precedente. Quanto agli zeri di $S(s)$, essi sono le radici di $D_L(s)$ e quindi coincidono con i poli di $L(s)$. In particolare, se la funzione d'anello è asintoticamente stabile, la funzione di sensitività non ha zeri nel semipiano destro chiuso.

10.4.3 Risposta in frequenza

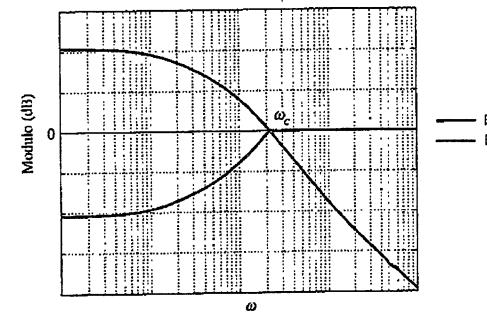
Verrà ora condotta un'analisi simile a quella che nel Paragrafo 10.3.3 era servita per evidenziare le caratteristiche della risposta in frequenza $F(j\omega)$ sulla base dei diagrammi di Bode associati alla funzione d'anello. Anche qui si supporrà che il diagramma di $|L(j\omega)|_{\text{dB}}$ attraversi una sola volta l'asse a 0 dB in corrispondenza di ω_c , come mostrato nella Figura 10.11. Usando la (10.22), in prima approssimazione si può scrivere

$$|S(j\omega)| = \frac{1}{|1 + L(j\omega)|} \approx \begin{cases} \frac{1}{|L(j\omega)|} \cdot \omega \leq \omega_c \\ 1 \quad , \omega > \omega_c \end{cases} \quad (10.24)$$

e quindi l'andamento del diagramma di Bode di $|S(j\omega)|$ può essere qualitativamente rappresentato come nella Figura 10.11, assumendo l'aspetto tipico di un filtro passa-alto. Si osservi a tale proposito che il diagramma di Bode di $1/|L(j\omega)|$ si ottiene dal ribaltamento di quello di $|L(j\omega)|$ rispetto all'asse a 0 dB.

Il risultato a cui si è pervenuti si presta a interessanti interpretazioni in termini di azione filtrante. Per esempio, guardando a $S(s)$ come alla funzione di trasferimento tra il disturbo d e l'uscita y , si deduce che, sotto l'ipotesi di asintotica

Figura 10.11
Valutazione approssimata di $|S(j\omega)|$.



stabilità, il sistema retroazionato produce un effetto di attenuazione sulla variabile controllata per tutte le componenti del disturbo a pulsazione minore di ω_c , effetto tanto più consistente quanto maggiore è il valore di $|L(j\omega)|$ in quella banda di pulsazioni. Invece, le componenti del disturbo d a pulsazione maggiore di ω_c non subiscono apprezzabili riduzioni in ampiezza.

In definitiva, nell'ottica di rendere $|S(j\omega)|$ sufficientemente piccolo per un intervallo di pulsazioni il più ampio possibile, un valore elevato della pulsazione critica ω_c è certamente vantaggioso.

Esempio 10.6 (Seguito degli Esempi 10.1 e 10.3) In base alla (10.24) e al diagramma di Bode del modulo di $L(j\omega)$ riportato nella Figura 10.3, si deduce che la funzione di sensitività $S(s)$ determina un'attenuazione dell'ingresso limitatamente all'intervallo di pulsazioni che si estende da $\omega = 0$ fino a circa $\omega_c \approx 0.9$. Per apprezzare la validità di tale conclusione, si osservi nella Figura 10.12 che il grafico approssimato di $|S(j\omega)|$ ottenuto dalla (10.24) e quello esatto sono quasi sovrapposti, tranne che in vicinanza della pulsazione critica ω_c . D'altra parte, anche la presenza del picco di $|S(j\omega)|$ che in vicinanza della pulsazione critica ω_c .

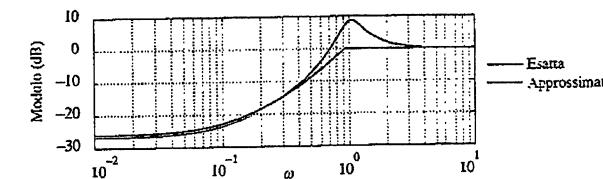


Figura 10.12
Diagramma di Bode di $|S(j\omega)|$ esatto e approssimato nell'Esempio 10.6.

poteva essere prevista sulla base del valore relativamente basso del margine di fase ($\phi_m \approx 24^\circ$) e quindi dello smorzamento associato ai poli di $S(s)$ che cadono in vicinanza di ω_c . A completamento dell'esempio, nella Figura 10.13 è riportato il grafico della risposta allo scalino della funzione $S(s)$.

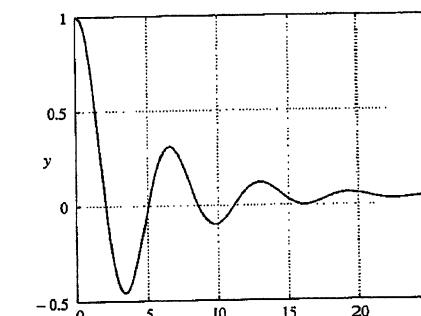
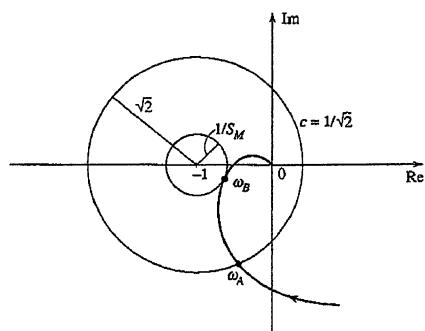


Figura 10.13
Risposta allo scalino associata a $S(s)$ nell'Esempio 10.6.

Il valore asintotico, in accordo con la Tabella 10.1, vale $1/(1+\mu) = 1/21$ e l'andamento oscillante è dovuto ai poli dominanti a basso smorzamento, già individuati nella discussione dell'Esempio 10.3.

Osservazione 10.5 Dall'esame del diagramma polare associato a $L(s)$ sarebbe possibile ricavare informazioni più precise sulle proprietà firanti della funzione $S(s)$. A tale scopo si noti che il luogo dei punti del piano complesso per i quali risulta $1/|1+s| = c$ è una circonferenza con centro in -1 e raggio pari a $1/c$. Allora, ricordando che $|S(j\omega)| = 1/|1+L(j\omega)|$, è immediato concludere che i valori di ω per i quali risulta $|S(j\omega)| < c$ sono quelli in corrispondenza dei quali il diagramma polare di $L(j\omega)$ è esterno al cerchio di centro -1 e raggio $1/c$. Nell'esempio della Figura 10.14, dove in particolare è tracciato il cerchio relativo a $c = 1/\sqrt{2} \approx -3$ dB, risulta $|S(j\omega)| < 1/\sqrt{2}$ per $\omega < \omega_A$.

Figura 10.14
Esempio di valutazione della banda passante di $S(s)$ e di S_M dal diagramma polare di $L(j\omega)$.



Si può impiegare la stessa procedura grafica per determinare l'eventuale valore di picco di $|S(j\omega)|$, già indicato con il simbolo S_M e collegato, come si sa, alla stabilità robusta del sistema. Esso coincide con l'inverso del raggio del cerchio di centro -1 che è tangente al diagramma polare di $L(j\omega)$ (si veda ancora la Figura 10.14, dove ω_B è il valore della pulsazione in corrispondenza del picco). ■

10.4.4 Altri limiti alle prestazioni

La situazione ideale per le prestazioni di un sistema di controllo sarebbe, come si è già detto, avere una funzione di sensitività $S(s)$ con un diagramma di Bode del modulo più basso possibile a tutte le pulsazioni, perché ciò garantirebbe un piccolo effetto sull'errore degli ingressi w e d , qualunque sia il loro contenuto spettrale. Si è peraltro già osservato che vi sono dei limiti invalicabili verso il raggiungimento di tale obiettivo, se non altro per il fatto che, per elevati valori della pulsazione ω , risulta in ogni caso $|S(j\omega)| \approx 1$. Ci sono poi certi casi in cui esistono vincoli ancora più stringenti sulla forma che può assumere il diagramma di Bode del modulo associato a $S(j\omega)$. Uno di questi casi è illustrato nel seguente teorema, valido nel caso in cui il grado relativo della funzione d'anello (cioè la differenza tra il grado del denominatore e quello del numeratore) sia maggiore o uguale a 2.

Teorema 10.1 Si supponga che il sistema retroazionato di Figura 10.1 sia asintoticamente stabile, si assuma che la funzione d'anello $L(s) = R(s)G(s)$ non

abbia poli con parte reale positiva ($P = 0$), e che possieda un grado relativo maggiore o uguale a 2. Allora risulta

$$\int_0^\infty |S(j\omega)|_{\text{dB}} d\omega = 0 \quad (10.25)$$

L'interpretazione di questo risultato è molto intuitiva. Secondo la formula (10.25), il valor medio di $|S(j\omega)|_{\text{dB}}$ sull'intero asse delle pulsazioni (se questo fosse tracciato con scala lineare anziché logaritmica) è nullo. Ciò significa che, se vi sono tratti del diagramma che si trovano al di sotto dell'asse a 0 dB, ve ne devono esserci altri al di sopra dello stesso asse. Ricordando che $S(s)$ necessariamente è reale, si conclude che, oltre a non riuscire a ottenere un'attenuazione del disturbo inferiore a 0 dB su tutto l'asse ω , una consistente attenuazione in una certa banda di pulsazioni è inevitabilmente compensata da una significativa amplificazione in un'altra banda.

Esempio 10.7 La Figura 10.15 illustra la proprietà enunciata nel Teorema 10.1, nel caso in cui sia

$$L(s) = \frac{k}{s^2 + 2s + 3}$$

per diversi valori di k . In accordo con la (10.25), le aree delle due regioni delimitate da ogni singola curva $|S(j\omega)|_{\text{dB}}$ e dall'asse a 0 dB (quella a bassa frequenza, dove c'è attenuazione, e quella ad alta frequenza, dove c'è amplificazione) sono necessariamente uguali. Aumentando il valore di k si ottiene una migliore attenuazione a bassa frequenza (oltre che un aumento della pulsazione critica ω_c), a scapito però di un'amplificazione sempre più elevata ad alta frequenza.

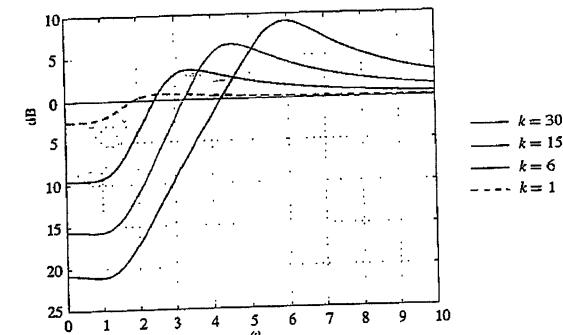


Figura 10.15
Diagrammi di $|S(j\omega)|$ per diversi valori di k nell'Esempio 10.7.

Altri risultati di questo genere si possono ricavare anche nel caso in cui $L(s)$ abbia poli nel semipiano destro, ma l'argomento è al di fuori degli scopi di questa trattazione.

10.5 Analisi della funzione di sensitività del controllo

Con riferimento al sistema retroazionato di Figura 10.1, è stata introdotta al Paragrafo 10.2 la funzione di sensitività del controllo $Q(s)$, che si può scrivere in

diverse forme equivalenti come

$$Q(s) = \frac{R(s)}{1 + R(s)G(s)} = R(s)S(s) = F(s)G(s)^{-1} \quad (10.26)$$

Dall'esame della (10.4), si vede come $Q(s)$ rappresenti, a parte un eventuale cambio di segno, la funzione di trasferimento tra gli ingressi w, d e n e la variabile di controllo u . Pertanto, l'analisi di tale funzione consente di valutare le caratteristiche del sistema retroazionato dal punto di vista della moderazione della variabile di controllo. Per i motivi già citati nel Paragrafo 9.4.2, sarebbe desiderabile che, a parità di altre prestazioni, il modulo di $Q(j\omega)$ risultasse il più piccolo possibile, per lo meno nell'insieme di pulsazioni in cui gli ingressi presentano componenti significative, in modo da ridurre le sollecitazioni a cui è sottoposta la variabile di controllo u . D'altra parte, si ricordi che in molti casi reali lo spettro del disturbo di misura n è prevalentemente concentrato verso le alte frequenze, mentre quelli dei segnali w e d sono significativamente diversi da zero solo a bassa frequenza. Per soddisfare il requisito di moderazione in ogni condizione di funzionamento, è quindi opportuno richiedere che $|Q(j\omega)|$ risulti sufficientemente basso per tutte le pulsazioni ω .

10.5.1 Analisi statica

Limitandosi a considerare per semplicità solo il caso in cui la funzione d'anello $L(s) = R(s)G(s)$ sia di tipo zero, e cioè non siano presenti azioni integrali o derivate né nel regolatore né nel processo da controllare, e indicando con μ_R , μ_G e μ i guadagni rispettivamente di $R(s)$, $G(s)$ e $L(s)$, si ricava dalle (10.9), (10.23) e (10.26)

$$\begin{aligned} \lim_{s \rightarrow 0} Q(s) &= \lim_{s \rightarrow 0} \mu_R S(s) = \frac{\mu_R}{1 + \mu} = \lim_{s \rightarrow 0} \frac{F(s)}{\mu_G} = \\ &= \frac{\mu}{\mu_G(1 + \mu)} \end{aligned}$$

In particolare, l'ultima espressione mette in luce che, quando $\mu \gg 1$, come di norma avviene, risulta $\mu_Q \approx \mu_G^{-1}$. In definitiva, se il sistema retroazionato è asintoticamente stabile e gli ingressi sono costanti, il valore a regime u_∞ della variabile di controllo è inversamente proporzionale al guadagno del processo da controllare. Ciò significa che, come è ovvio, è impossibile ottenere soddisfacenti prestazioni statiche nella regolazione di un processo con piccolo guadagno, se non accettando valori elevati della variabile di controllo.

10.5.2 Poli e zeri

Se $R(s) = N_R(s)/D_R(s)$ e $G(s) = N_G(s)/D_G(s)$ sono funzioni razionali, dalla (10.26) si ricava

$$Q(s) = \frac{N_R(s)D_G(s)}{D_R(s)D_G(s) + N_R(s)N_G(s)} = \frac{N_R(s)D_G(s)}{D_L(s) + N_L(s)} \quad (10.27)$$

dove, come di consueto, si sono indicati con $N_L(s)$ e $D_L(s)$ i polinomi al numeratore e al denominatore della funzione d'anello $L(s)$ e si è supposto che non avvengano cancellazioni tra poli e zeri di $R(s)$ e $G(s)$.

Sulle caratteristiche dei poli di $Q(s)$ è inutile dilungarsi poiché già abbondantemente discusse nei paragrafi precedenti. Quanto agli zeri, si osserva dalla (10.27) che essi sono gli zeri di $R(s)$ e i poli di $G(s)$. Di conseguenza, la risposta allo scalino della funzione $Q(s)$ può presentare sovraelongazioni di elevata ampiezza, potenzialmente incompatibili con le limitazioni naturali della variabile di controllo, tutte le volte che il regolatore $R(s)$ possiede zeri, o il processo $G(s)$ possiede poli, a pulsazioni inferiori a ω_c , che, come si è visto, approssima la pulsazione dei poli dominanti in anello chiuso (si veda la discussione del Paragrafo 4.4.4 a proposito dell'effetto degli zeri sulla risposta allo scalino).

10.5.3 Risposta in frequenza

A partire dalla (10.26), mediante le usuali approssimazioni si può scrivere

$$|Q(j\omega)| = \frac{|R(j\omega)|}{|1 + L(j\omega)|} \approx \begin{cases} \frac{1}{|G(j\omega)|}, & \omega \leq \omega_c \\ |R(j\omega)|, & \omega > \omega_c \end{cases} \quad (10.28)$$

A titolo di esempio, la Figura 10.16 illustra la procedura grafica che si può utilizzare per ricavare l'andamento di $|Q(j\omega)|$ basandosi sull'approssimazione (10.28). V'è notato che il diagramma di Bode di $1/|G(j\omega)|$ è il simmetrico di quello di $|G(j\omega)|$ rispetto all'asse a 0 dB, e che i diagrammi di $|R(j\omega)|$ e $1/|G(j\omega)|$ si intersecano in corrispondenza della pulsazione critica ω_c , dove risulta $|L(j\omega_c)| = |R(j\omega_c)| |G(j\omega_c)| = 1$.

Alla luce della (10.28), l'ampiezza di $|Q(j\omega)|$ a pulsazione minore di ω_c dipende solo dalla risposta in frequenza del sistema sotto controllo $G(j\omega)$, e dipende invece solo da quella del regolatore $R(j\omega)$ a pulsazione elevata. Si dovrà tenere conto di ciò in fase di progetto, evitando, per quanto possibile, di impiegare regolatori con $|R(j\omega)|$ elevato ad alta frequenza, se si vogliono limitare le sollecitazioni sulla variabile di controllo.

Per esaminare la stessa questione da un diverso punto di vista, si utilizzi l'ultima espressione della (10.26) per scrivere che $|Q(j\omega)| = |F(j\omega)| |G(j\omega)|^{-1}$. Si ricordi poi che il tipico aspetto della funzione di sensitività complementare $F(s)$ è quello di un filtro passa-basso con guadagno unitario e banda passante $[0, \omega_c]$. Ipotizzando allora che, come spesso accade, anche $G(s)$ abbia un comportamento da filtro passa-basso con banda passante $[0, \bar{\omega}]$, si arriva a concludere che, quando $\omega_c > \bar{\omega}$, il diagramma di Bode di $|Q(j\omega)|$, che si ottiene per differenza tra quelli di $|F(j\omega)|$ e di $|G(j\omega)|$, presenta un tratto con andamento crescente nell'intervallo $[\bar{\omega}, \omega_c]$ (Figura 10.17). Se quindi risulta $\omega_c \gg \bar{\omega}$, può accadere che $|Q(j\omega)|$

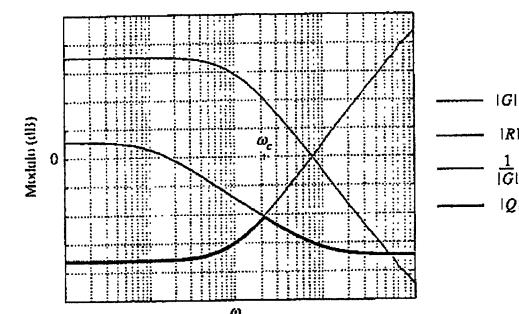
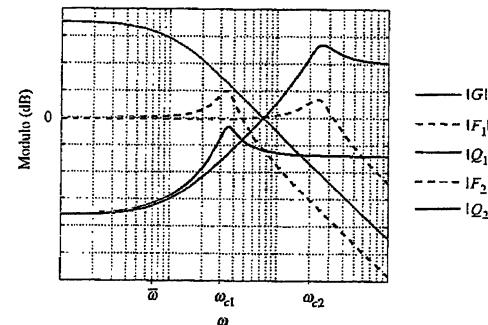


Figura 10.16
Valutazione approssimata di $|Q(j\omega)|$.

raggiunga valori eccessivi ad alta frequenza. Tutto ciò evidenzia il prezzo che si è costretti a pagare in termini di moderazione dell'azione di controllo quando si vuole spingere ω_c a valori molto più elevati rispetto alla banda passante del sistema sotto controllo, cioè quando si pretende di ottenere un sistema retroazionato "molto più veloce" del processo in anello aperto.

Figura 10.17
Effetto su $|Q(j\omega)|$ di diverse scelte di ω_c quando $G(s)$ è un filtro passa-basso con banda passante $[0, \bar{\omega}]$.



10.6 Prestazioni in condizioni perturate

Tutte le analisi svolte nei Paragrafi 10.3-10.5 per studiare le proprietà di un sistema di controllo sono basate sull'ipotesi di una conoscenza perfetta del modello del sistema, in particolare della funzione di trasferimento $G(s)$. Tuttavia, come si è già anticipato nel Paragrafo 9.4.2, è importante poter garantire che certe prestazioni del sistema di controllo non si modifichino, o almeno non subiscano un eccessivo degrado, in presenza di incertezze. A tale riguardo, sarebbe interessante capire come variano le caratteristiche salienti delle funzioni $F(s)$, $S(s)$ e $Q(s)$ quando la funzione di trasferimento effettiva $G_a(s)$ del sistema sotto controllo si discosta dal suo modello nominale $G(s)$. Questo tipo di analisi è tutt'altro che banale e una sua trattazione completa esula dagli scopi del presente testo. In questo paragrafo ci si limiterà a illustrare alcuni aspetti elementari, ma particolarmente significativi, del problema.

10.6.1 Regolazione robusta a zero dell'errore

Una prima proprietà che vale in maniera robusta è già stata discussa nell'Osservazione 10.4, e riguarda l'annullamento dell'errore a transitorio esaurito quando il sistema di controllo di Figura 10.1 è sottoposto all'azione di un riferimento w o di un disturbo d a scalino. Si è visto che, sotto l'ipotesi di asintotica stabilità, l'errore tende ad annullarsi asintoticamente se il tipo g della funzione d'anello è maggiore di zero (si veda la Tabella 10.1), indipendentemente da eventuali incertezze sulla funzione $G(s)$ del sistema da controllare, purché ovviamente tali da non pregiudicare l'asintotica stabilità in anello chiuso. Come già detto, si parla in tal caso di regolazione robusta a zero dell'errore, o, nel caso del disturbo d , di reiezione robusta del disturbo a scalino.

10.6.2 Reiezione robusta di disturbi sinusoidali

Si supponga che nello schema di Figura 10.1 sia $d(t) = A \sin(\bar{\omega}t)$. Assumendo che il sistema sia asintoticamente stabile, perché l'effetto di tale disturbo sull'er-

ore e sia nullo a transitorio esaurito occorre che risulti $|S(j\bar{\omega})| = 0$ (si ricordi infatti che la funzione di trasferimento tra d ed e è pari a $-S(s)$). Per far sì che ciò accada è necessario che $S(s)$ abbia una coppia di zeri in $s = \pm j\bar{\omega}$. D'altra parte, si è visto al Paragrafo 10.4.2 che gli zeri di $S(s)$ coincidono con i poli di $L(s)$. Allora, un modo per raggiungere l'obiettivo desiderato è quello di utilizzare un regolatore $R(s)$ che abbia una coppia di poli in $s = \pm j\bar{\omega}$, così da farli comparire anche tra i poli di $L(s)$. Anche in questo caso il risultato è indipendente dalla presenza di eventuali incertezze sulla funzione $G(s)$ del sistema sotto controllo, a patto che non alterino la proprietà di stabilità. Si ottiene in questo modo la *reiezione robusta* del disturbo sinusoidale.

Si noti che la soluzione adottata non assicura invece alcun grado di robustezza nei confronti di possibili incertezze sulla pulsazione $\bar{\omega}$ del disturbo. Inoltre, non va sottovalutata la difficoltà di riuscire a garantire l'asintotica stabilità del sistema retroazionato con un regolatore che introduce nella funzione d'anello due poli a smorzamento nullo, che determinano un contributo di fase di -180° dalla pulsazione $\bar{\omega}$ in avanti.

10.6.3 Attenuazione robusta di disturbi a banda limitata

In presenza di disturbi d a banda limitata, tali cioè da non possedere componenti spettrali a pulsazioni maggiori di un dato valore $\bar{\omega}$, un tipico requisito per il sistema di controllo di Figura 10.1 consiste nell'imporre $|S(j\omega)| < c, \forall \omega \in [0, \bar{\omega}]$, dove c è una costante minore di 1, in modo che tutte le componenti del disturbo vengano attenuate di un fattore pari a c . Nell'analisi in condizioni nominali, svolta al Paragrafo 10.4.3, si è visto come un vincolo di questo tipo si traduca nel richiedere che la parte del diagramma polare di $L(j\omega)$ corrispondente a $\omega \in [0, \bar{\omega}]$ giaccia all'esterno del cerchio di centro -1 e raggio $1/c$ (si veda la Figura 10.14). È facile estendere questo risultato al caso in cui la funzione di trasferimento del sistema da controllare sia affetta da incertezza di tipo additivo (si veda il Paragrafo 9.7.1). Si assuma cioè che l'effettiva funzione di trasferimento sia $G_a(s) = G(s) + \delta G(s)$, dove $G(s)$ rappresenta il modello nominale, e la perturbazione $\delta G(s)$ è incognita, ma soddisfa la condizione $|\delta G(j\omega)| \leq \gamma(\omega), \forall \omega$, con $\gamma(\omega)$ assegnata. Si indichi inoltre con $L_a(s) = R(s)G_a(s) = L(s) + R(s)\delta G(s)$ la funzione d'anello perturbata e con $S_a(s) = (1 + L_a(s))^{-1}$ la corrispondente funzione di sensibilità.

Se si desidera che il livello di attenuazione del disturbo prima descritto sia ottenuto in maniera robusta, ossia per tutte le perturbazioni ammissibili, occorre ovviamente che sia $|S_a(j\omega)| < c$, cioè $|1 + L_a(j\omega)| > 1/c$, per ogni pulsazione ω nell'intervallo $[0, \bar{\omega}]$ e per ogni $\delta G(s)$ ammissibile. Si osservi ora che

$$\begin{aligned} |1 + L_a(j\omega)| &= |1 + L(j\omega) + R(j\omega)\delta G(j\omega)| \geq \\ &\geq |1 + L(j\omega)| - |R(j\omega)| |\delta G(j\omega)| \geq \\ &\geq |1 + L(j\omega)| - \gamma(\omega) |R(j\omega)| \end{aligned} \quad (10.29)$$

dove si è sfruttata la proprietà che il modulo della somma di due numeri complessi è maggiore o uguale alla differenza dei moduli e si è utilizzata la funzione maggiorante $\gamma(\omega)$. Quindi, se si fa in modo che risulti

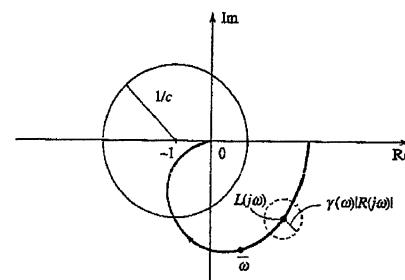
$$|1 + L(j\omega)| > \frac{1}{c} + \gamma(\omega) |R(j\omega)|, \quad \omega \leq \bar{\omega} \quad (10.30)$$

dalla (10.29) si ricava $|1 + L_a(j\omega)| > 1/c$ sullo stesso intervallo di pulsazioni, e la specifica di prestazioni robuste è rispettata. Si noti che la (10.30) è basata solo

sul valore nominale di $L(s)$ e sulle informazioni *a priori* sulla massima ampiezza dell'incertezza.

L'interpretazione geometrica della condizione (10.30) è illustrata nella Figura 10.18, dove è tracciato il diagramma polare nominale di $L(j\omega)$. Imporre la (10.30) equivale a richiedere che non vi sia intersezione tra il cerchio di centro -1 e raggio $1/c$ e tutti i cerchi di centro $L(j\omega)$ e raggio $\gamma(\omega)|R(j\omega)|$, con $\omega \leq \bar{\omega}$, che rappresentano la regione in cui può trovarsi il valore perturbato di $L_a(j\omega)$ al variare di ω , in base alle ipotesi fatte.

Figura 10.18
Condizione per l'attenuazione robusta di un disturbo a banda limitata.



10.6.4 Sensibilità rispetto a incertezze parametriche

Adottando un diverso punto di vista, si supponga di poter modellizzare il sistema sotto controllo mediante una funzione di trasferimento contenente un parametro incerto θ di cui è disponibile solamente una stima nominale $\bar{\theta}$. Per mettere in evidenza questa caratteristica anche nella notazione, si indichi con $G(s, \theta)$ la funzione nominale, con $G(s, \bar{\theta})$ quella perturbata, e con $\Delta\theta = \theta - \bar{\theta}$ lo scostamento del valore effettivo del parametro da quello nominale. Lo scopo della successiva analisi sarà quello di determinare l'effetto della variazione $\Delta\theta$ sulle caratteristiche del sistema retroazionato di Figura 10.19, almeno localmente, cioè per piccoli scostamenti del parametro rispetto al valore nominale. Per questo motivo si parla di *analisi di sensitività*.

Si cominci a considerare l'effetto di $\Delta\theta$ sulla funzione di trasferimento in anello aperto $G(s, \theta)$. Utilizzando lo sviluppo di Taylor arrestato al termine di primo ordine, si può affermare che, per $\Delta\theta$ sufficientemente piccolo, risulta

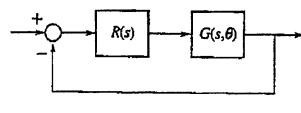
$$\Delta G(s, \theta, \bar{\theta}) = G(s, \theta) - G(s, \bar{\theta}) \simeq \frac{\partial G(s, \theta)}{\partial \theta} \Big|_{\theta=\bar{\theta}} \Delta\theta$$

L'espressione

$$E_G(s, \theta, \bar{\theta}) = \frac{\Delta G(s, \theta, \bar{\theta})}{G(s, \bar{\theta})} \quad (10.31)$$

rappresenta la variazione relativa della funzione di trasferimento in anello aperto, e la sua valutazione per $s = j\omega$, cioè $E_G(j\omega, \theta, \bar{\theta})$, descrive l'effetto dell'incertezza parametrica sulla risposta in frequenza associata. Elevati valori di

Figura 10.19
Sistema retroazionato affetto da incertezze parametriche.



$|E_G(j\omega, \theta, \bar{\theta})|$ in una certa banda di pulsazioni stanno dunque a indicare un elevato grado di sensitività della risposta in frequenza in quella banda rispetto alla variazione del parametro.

Passando ora a considerare il sistema in anello chiuso della Figura 10.19, si indichi con

$$F(s, \theta) = \frac{R(s)G(s, \theta)}{1 + R(s)G(s, \theta)} \quad (10.32)$$

la funzione di trasferimento tra w e y e con $F(s, \bar{\theta})$ la corrispondente funzione in condizioni nominali. Ponendo $\Delta F(s, \theta, \bar{\theta}) = F(s, \theta) - F(s, \bar{\theta})$ e osservando che risulta

$$\frac{\partial F(s, \theta)}{\partial \theta} = \frac{R(s)}{(1 + R(s)G(s, \theta))^2} \frac{\partial G(s, \theta)}{\partial \theta}$$

si ricava facilmente dalle (10.31), (10.32) che la variazione relativa della funzione $F(s, \theta)$ è data da

$$\begin{aligned} E_F(s, \theta, \bar{\theta}) &= \frac{\Delta F(s, \theta, \bar{\theta})}{F(s, \bar{\theta})} \simeq \frac{1}{F(s, \bar{\theta})} \frac{\partial F(s, \theta)}{\partial \theta} \Big|_{\theta=\bar{\theta}} \Delta\theta = \\ &= \frac{1}{1 + R(s)G(s, \bar{\theta})} \frac{\Delta G(s, \theta, \bar{\theta})}{G(s, \bar{\theta})} = S(s, \bar{\theta}) E_G(s, \theta, \bar{\theta}) \end{aligned}$$

dove $S(s, \bar{\theta})$ rappresenta la funzione di sensitività nominale. In particolare risulta

$$|E_F(j\omega, \theta, \bar{\theta})| = |S(j\omega, \bar{\theta})| |E_G(j\omega, \theta, \bar{\theta})|$$

Pertanto, il modulo di $S(j\omega, \bar{\theta})$ rappresenta il rapporto tra gli effetti dell'incertezza parametrica in anello chiuso e in anello aperto. Nell'intervallo di pulsazioni in cui è $|S(j\omega, \bar{\theta})| \ll 1$, la presenza della retroazione produce quindi una notevole attenuazione dell'effetto dell'incertezza. Poiché, come già discusso al Paragrafo 10.4.3, la funzione $S(j\omega, \bar{\theta})$ ha tipicamente modulo inferiore a 1 per valori di ω minori di ω_c , si conclude che la retroazione garantisce un buon livello di insensibilità rispetto alle variazioni parametriche all'interno della banda passante del sistema di controllo. Tra l'altro questo spiega perché la funzione $S(s, \bar{\theta})$ è chiamata "funzione di sensitività".

Esempio 10.8 Con riferimento al sistema di Figura 10.20, dove si suppone $\tilde{L}(0) = 1$, si voglia analizzare l'effetto di eventuali perturbazioni del guadagno d'anello rispetto al suo valore nominale $\bar{\mu}$. Indicando con $L(s, \bar{\mu}) = \bar{\mu}\tilde{L}(s)$ la funzione d'anello nominale, la sua variazione relativa in presenza di una perturbazione $\Delta\mu$ è ovviamente data da $E_L(s, \mu, \bar{\mu}) = \Delta\mu/\bar{\mu}$, mentre, in accordo con la precedente analisi, la variazione relativa della funzione $F(s, \mu)$ in anello chiuso è data da $E_F(s, \mu, \bar{\mu}) = S(s, \bar{\mu})\Delta\mu/\bar{\mu}$, se $S(s, \bar{\mu}) = (1 + L(s, \bar{\mu}))^{-1}$ è la funzione di sensitività nominale. In particolare, la variazione in corrispondenza di $s = 0$ rappresenta la perturbazione relativa del guadagno del sistema retroazionato $\Delta\mu_F/\bar{\mu}_F$. Dalla precedente espressione si deduce

$$E_F(0, \mu, \bar{\mu}) = \frac{\Delta\mu_F}{\bar{\mu}_F} = \frac{1}{1 + \bar{\mu}} \frac{\Delta\mu}{\bar{\mu}}$$

Si può allora concludere che l'effetto di una perturbazione del guadagno d'anello del sistema in anello chiuso risulta attenuato del fattore $1/(1 + \bar{\mu})$. Si confronti quest'ultimo risultato con quello discusso nel Paragrafo 10.4.1, con riferimento all'attenuazione a regime di disturbi costanti nei sistemi con tipo $g = 0$.

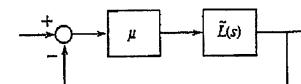


Figura 10.20
Sistema retroazionato con guadagno d'anello incerto, analizzato nell'Esempio 10.8.

Vale la pena di osservare che, nell'ambito della progettazione di circuiti elettronici, la proprietà messa in evidenza in questo esempio è sfruttata nei cosiddetti *amplificatori reazionati* per rendere il comportamento del circuito maggiormente insensibile rispetto a variazioni delle condizioni di funzionamento.

10.7 Conclusioni

Il filo conduttore dell'analisi svolta in questo capitolo sta nel fatto che la maggior parte delle proprietà di interesse di un sistema di controllo retroazionato possono essere valutate, così come si è fatto per la stabilità nel capitolo precedente, a partire dalle caratteristiche della funzione di trasferimento d'anello. Alcune conclusioni, per esempio quelle che riguardano la valutazione della banda passante del sistema retroazionato mediante ω_c , o dello smorzamento dei poli dominanti mediante φ_m , sono basate su procedimenti approssimati e forniscono solo informazioni indicative, anche se spesso molto aderenti alla realtà. La loro utilità potrà essere meglio apprezzata nel corso della discussione sulle tecniche di progetto del controllore illustrate nel prossimo capitolo.

Esercizi

Esercizio 10.1 Si consideri il sistema di Figura 10.1 con

$$R(s) = 10 \quad , \quad G(s) = \frac{1}{(1+s)^2}$$

e si confrontino i diagrammi di Bode esatti delle funzioni $F(s)$, $S(s)$ e $Q(s)$ con quelli ottenuti mediante le approssimazioni (10.10), (10.24) e (10.28).

Esercizio 10.2 Con riferimento al sistema dell'esercizio precedente, si confrontino i valori esatti di pulsazione naturale e smorzamento dei poli in anello chiuso con i valori deducibili in forma approssimata da ω_c e φ_m , secondo quanto discusso nei Paragrafi 10.3.3, 10.3.4.

Esercizio 10.3 Ancora riferendosi al sistema di controllo dell'Esercizio 10.1, e utilizzando le approssimazioni ricavate nel capitolo, si determinino le principali caratteristiche del movimento di $y(t)$, $u(t)$ ed $e(t)$, quando $w(t) = A \text{sca}(t)$, $d(t) = n(t) = 0$.

Esercizio 10.4 Si ripeta l'esercizio precedente assumendo ora che sia

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

Esercizio 10.5 Si dimostri che le funzioni di trasferimento $L(s)$, $F(s)$ e $S(s)$ sono "adimensionali", non risentono cioè delle unità di misura utilizzate per le diverse variabili che compaiono nello schema di un sistema di controllo. Si verifichi invece che ciò non è vero per la funzione di sensitività del controllo $Q(s)$. Si rifletta infine sulle conseguenze di tali caratteristiche sulla possibilità di attribuire un significato "concreto" all'asse a 0 dB nei rispettivi diagrammi di Bode del modulo.

Esercizio 10.6 Si disegni nel piano complesso il luogo dei punti per cui risulta

$$\frac{|s|}{|1+s|} = c$$

per un generico valore positivo c . Si verifichi che tale luogo è una circonferenza (che degenera in una retta solo quando $c = 1$). Prendendo spunto da quanto illustrato nel Paragrafo 10.3.5, si discuta poi come si possa usare l'insieme di questi luoghi (che prendono il nome di *cerchi di Hall*) per ricavare graficamente per punti il diagramma di $|F(j\omega)|$ a partire dal diagramma polare di $L(j\omega)$.

Esercizio 10.7 Si estenda l'analisi statica della funzione di sensitività del controllo $Q(s)$ al caso in cui il tipo della funzione d'anello è $g = 1$ (si distingua il caso in cui l'integratore è contenuto in $R(s)$ da quello in cui è contenuto in $G(s)$).

Esercizio 10.8 A partire dallo schema della Figura 9.2, in cui compaiono i disturbi d_P sul processo e d_A sull'attuatore, si calcolino le funzioni di trasferimento tra i due disturbi e l'uscita y , e, con argomentazioni simili a quelle del Paragrafo 10.4.3, si ricavino opportune approssimazioni per i corrispondenti diagrammi di Bode del modulo.

Esercizio 10.9 Con riferimento al sistema di controllo della Figura 10.21, si dimostri che, per garantire l'annullamento dell'errore a transitorio esaurito in risposta a un disturbo d_A a scalino, è necessario che $R(s)$ contenga almeno un integratore.

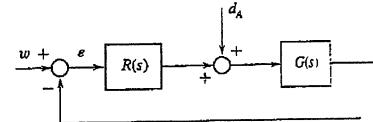


Figura 10.21
Sistema di controllo
dell'Esercizio 10.9.

Esercizio 10.10 Si ripeta l'analisi del Paragrafo 10.6.3 nel caso di incertezza moltiplicativa, cioè assumendo che la funzione di trasferimento effettiva del sistema sotto controllo sia $G_a(s) = G(s)(1 + \delta G(s))$, con $|\delta G(j\omega)| \leq \gamma(\omega)$, $\forall \omega$. In particolare, si dimostri che, in tal caso, la condizione per avere attenuazione robusta di un disturbo a banda limitata (cioè per fare in modo che risulti $|S_a(j\omega)| < c$ per ogni pulsazione ω nell'intervallo $[0, \hat{\omega}]$ e per ogni $\delta G(s)$ ammissibile) diventa

$$|1 + L(j\omega)| > \frac{1}{c} + \gamma(\omega) |L(j\omega)| \quad \hat{\omega} \leq \hat{\omega}$$

11.1 Introduzione

Seguendo l'impostazione presentata nei capitoli precedenti, il progetto di un sistema di controllo in retroazione si riconduce alla scelta della funzione di trasferimento del regolatore in modo che il sistema retroazionato possieda determinate caratteristiche. In particolare, devono essere rispettati i requisiti di stabilità e prestazione illustrati nel Paragrafo 9.4.

Si è visto che, in molti casi, le proprietà di un sistema retroazionato possono essere facilmente accertate a partire dalla risposta in frequenza associata alla funzione di trasferimento d'anello. Quindi, un primo approccio al problema della sintesi è quello di individuare una funzione d'anello che permetta di soddisfare tutti i requisiti, ricavando poi da questa la funzione di trasferimento incognita del regolatore. Questo modo di ragionare costringe però a formulare fin dall'inizio in modo preciso tutti gli obiettivi del controllo e può dar luogo a regolatori inutilmente complicati. Si preferisce perciò quasi sempre utilizzare una procedura per tentativi, in cui si prendono in considerazione in prima istanza regolatori di struttura molto semplice, complicandoli via via se necessario in base all'analisi delle prestazioni rispetto alle diverse specifiche di progetto.

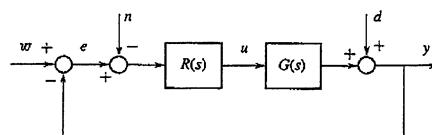
Questo capitolo è dedicato a illustrare, soprattutto tramite esempi, le principali tecniche di progetto basate sulla risposta in frequenza, limitandosi a considerare il caso, peraltro assai significativo, del controllo di sistemi asintoticamente stabili o comunque privi di poli nel semipiano destro. In particolare:

- si passeranno di nuovo in rassegna i principali requisiti di un sistema di controllo, con l'obiettivo di tradurli in un insieme di prescrizioni sulla funzione d'anello;
- si illustreranno i diversi approcci utilizzabili per la sintesi del regolatore;
- si discuteranno alcuni esempi di progetto, selezionati in modo da evidenziare le problematiche più significative e i principali criteri da seguire nella sintesi; in particolare, saranno mostrate le difficoltà che si incontrano nel progetto di regolatori per sistemi a fase non minima;
- si studieranno più in dettaglio alcune tipiche strutture di regolatori, come le cosiddette reti anticipatrici e ritardatrici.

11.2 Requisiti e specifiche

Nell'affrontare il problema del progetto di un controllore si farà riferimento allo schema di Figura 11.1, già utilizzato nei due capitoli precedenti per l'analisi dei sistemi retroazionati. In questo schema, w rappresenta il segnale di riferimento, n e d sono i disturbi, u è la variabile di controllo, y è la variabile controllata ed e rappresenta l'errore tra l'andamento desiderato e l'andamento effettivo dell'uscita y . I blocchi $R(s)$ e $G(s)$ descrivono rispettivamente il regolatore e il sistema sotto controllo. Supponendo di conoscere $G(s)$, il problema di progetto consiste nella scelta della funzione di trasferimento $R(s)$ in modo che il sistema di controllo fornisca determinate prestazioni. Si è visto nei capitoli precedenti che molte importanti proprietà di un sistema retroazionato sono legate alle caratteristiche della funzione d'anello. Pertanto, la maggior parte dei requisiti che deve possedere un sistema di controllo possono essere soddisfatti imponendo opportune specifiche sulla funzione $L(s) = R(s)G(s)$, data dal prodotto di un fattore noto $G(s)$ e di un fattore incognito $R(s)$, oggetto della sintesi.

Figura 11.1
Schema di un sistema di controllo.



Nel seguito del capitolo si assumerà che $L(s)$ soddisfi in ogni caso le condizioni di applicabilità del criterio di Bode (Teorema 9.2), in modo da poter sfruttare i risultati di analisi dei Capitoli 9 e 10. In particolare, sotto tale ipotesi $G(s)$ non può avere poli con parte reale maggiore di zero. Tali poli infatti non potrebbero venire cancellati da zeri corrispondenti del regolatore $R(s)$, pena la perdita dell'asintotica stabilità (si veda il Paragrafo 9.5), e dunque comparirebbero come poli della funzione d'anello. Il progetto di regolatori per sistemi instabili sarà rimandato ai successivi capitoli.

11.2.1 Requisiti principali

Saranno ora presi in esame i requisiti discussi nel Paragrafo 9.4 e, per ognuno di essi, si individueranno le corrispondenti condizioni sulla funzione d'anello $L(s)$.

Stabilità in condizioni nominali In base al criterio di Bode (Teorema 9.2), l'asintotica stabilità è garantita se il guadagno μ della funzione d'anello e il margine di fase φ_m sono entrambi positivi. Naturalmente, occorre anche evitare che nel prodotto $R(s)G(s)$ avvengano cancellazioni tra singolarità con parte reale maggiore o uguale a zero.

Stabilità in condizioni perturbate Un modo per ottenere un elevato grado di robustezza nei confronti di possibili incertezze sulla funzione $G(s)$ del sistema sotto controllo consiste nell'imporre valori elevati al margine di fase φ_m e al margine di guadagno k_m (si veda il Paragrafo 9.6). Bisogna inoltre prestare attenzione a non eccedere nella scelta della pulsazione critica ω_c , per evitare che eventuali ritardi di tempo, anche di piccola entità, di cui non si è tenuto conto possano condurre all'instabilità. Si ricordi infatti che il contributo negativo alla fase critica di un ritardo di tempo τ è dato da $-\omega_c \tau 180/\pi$.

Precisione statica Per ridurre l'ampiezza dell'errore a transitorio esaurito e_∞ in presenza di ingressi w e d canonici (scalino, rampa ecc.), occorre aumentare il tipo g e/o il guadagno μ della funzione d'anello (si vedano il Paragrafo 10.4.1 e la Tabella 10.1).

Precisione dinamica Per fare in modo che l'uscita y seguia fedelmente il riferimento w , anche quando questo varia rapidamente, occorre allargare la banda passante della funzione di sensitività complementare $F(s)$, e cioè ottenere una pulsazione critica ω_c sufficientemente elevata (si veda il Paragrafo 10.3.3). Inoltre, se si vogliono limitare le sovraelongazioni ed evitare eccessive oscillazioni nella risposta allo scalino, bisogna imporre uno smorzamento minimo ai poli dominanti di $F(s)$. Come si è visto nel Paragrafo 10.3.4, ciò equivale a richiedere che il margine di fase φ_m sia sufficientemente elevato. In definitiva, le tipiche specifiche "nel dominio del tempo" riguardanti il tempo di assettamento T_{ae} e l'entità massima delle sovraelongazioni $S\%$ possono essere direttamente tradotte in limiti inferiori sui valori di ω_c e φ_m .

Attenuazione dell'effetto del disturbo d Alla luce delle proprietà della funzione di sensitività $S(s)$ (si veda il Paragrafo 10.4.3), bisogna fare in modo che risulti $|L(j\omega)| \gg 1$ nell'intervallo di pulsazioni su cui insiste lo spettro del disturbo d . Tipicamente ciò implica un vincolo sul valore minimo della pulsazione critica ω_c .

Attenuazione dell'effetto del disturbo n Se lo spettro del disturbo n è confinato ad alta frequenza, come di solito accade per i disturbi associati alla misura, questo requisito implica un vincolo sul massimo valore di ω_c (si veda il Paragrafo 10.3).

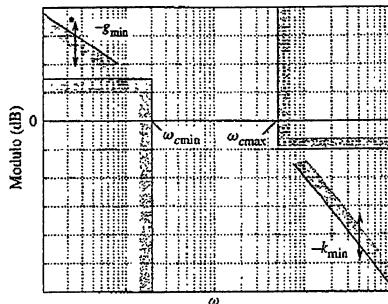
Moderazione della variabile di controllo Come si è visto nello studio della funzione di sensitività del controllo $Q(s)$ (Paragrafo 10.5), la proprietà di moderazione è legata al valore di $|R(j\omega)|$ a pulsazioni maggiori di ω_c . Quindi è opportuno mantenere $|R(j\omega)|$ limitato per $\omega > \omega_c$, o, equivalentemente, evitare che nello stesso insieme di pulsazioni risulti $|L(j\omega)| \gg |G(j\omega)|$. Si noti che questo requisito è l'unico tra quelli finora elencati che obbliga a considerare separatamente $R(s)$ e $G(s)$.

Realizzabilità del regolatore Per rispettare il vincolo che $R(s)$ sia la funzione di trasferimento di un sistema dinamico proprio, e quindi esista la possibilità di realizzare un dispositivo con tale comportamento, è necessario imporre che risulti $-k_L \leq -k_G$, dove $-k_L$ e $-k_G$ denotano la pendenza asintotica per $\omega \rightarrow \infty$ (negativa) dei diagrammi di Bode del modulo di $L(j\omega)$ e di $G(j\omega)$, rispettivamente.

11.2.2 Rappresentazione grafica dei vincoli

Sulla base delle considerazioni appena svolte, tutti i requisiti di un sistema di controllo possono essere tradotti in un insieme di vincoli sulla funzione d'anello, molti dei quali riconducibili a prescrizioni sull'andamento del diagramma di Bode di $|L(j\omega)|$. Una tipica situazione è quella rappresentata nella Figura 11.2, dove l'intersezione di alcuni vincoli determina una "regione di ammissibilità" entro la quale deve giacere il diagramma di Bode di $|L(j\omega)|$ per garantire un progetto soddisfacente. Nella Figura 11.2 si impone, per esempio, che la pulsazione critica

Figura 11.2
Rappresentazione grafica
dei vincoli di progetto.



sia compresa tra $\omega_{c\min}$ e $\omega_{c\max}$ e si tiene conto di opportuni vincoli su $|L(j\omega)|$: il limite inferiore a bassa frequenza soddisfa esigenze di precisione statica e di attenuazione del disturbo d , mentre il limite superiore ad alta frequenza garantisce una buona attenuazione del disturbo n e favorisce la robustezza della stabilità. Sono anche evidenziate nella figura altre restrizioni della regione ammissibile, connesse con eventuali specifiche sul minimo valore g_{\min} del tipo della funzione d'anello, che determinano la pendenza minima del tratto iniziale del diagramma, e con il vincolo di realizzabilità, che impone invece una pendenza minima $-k_{\min}$ del diagramma a pulsazione elevata. Si noti invece che questo tipo di rappresentazione, basata solo sul diagramma del modulo, non si presta a mettere in luce direttamente i vincoli imposti sul margine di fase φ_m , anche se, per funzioni d'anello a fase minima, ci si può aspettare che φ_m assuma valori sufficientemente elevati se si costringe il diagramma di $|L(j\omega)|$ ad attraversare l'asse a 0 dB con bassa pendenza. Ragionando sul diagramma asintotico la pendenza -1 è quella più favorevole (si ricordi la discussione che segue il criterio di Bode nel Paragrafo 9.6.5). Nel prossimo esempio si illustrerà la costruzione della regione di ammissibilità in un caso specifico.

Esempio 11.1 Si debba progettare un controllore per un sistema asintoticamente stabile descritto da una funzione di trasferimento $G(s)$ di tipo 0 e di grado relativo 2, tenendo conto dei seguenti requisiti:

- risulti $e_{\infty} = 0$ in presenza di un riferimento a scalino $w(t) = A \text{ sca}(t)$, con A arbitrario;
- la risposta di y a una variazione a scalino del riferimento w presenti un tempo d'assestamento $T_{a1} \leq 10$ e una sovraelongazione massima percentuale $S\% \leq 3$;
- un disturbo d con componenti spettrali nell'intervallo $[0, 0.1]$ risulti attenuato di almeno 20 dB;
- un disturbo n con componenti spettrali nell'intervallo $[10, 100]$ risulti attenuato di almeno 20 dB.

Il requisito (a) si traduce nella specifica $g \geq 1$ sul tipo della funzione d'anello, e quindi impone una pendenza iniziale negativa del diagramma di $|L(j\omega)|$.

Per garantire $S\% \leq 3$, occorre che lo smorzamento dei poli dominanti in anello chiuso sia superiore a $\xi = 0.75$ (si veda la Figura 4.11). In base all'Equazione (10.15), il margine di fase dovrà allora essere $\varphi_m \geq 75^\circ$. Con tali valori di φ_m , il sistema retroazionato è assimilabile a un sistema con un polo dominante reale in $s = -\omega_c$. Ricordando che in tal caso risulta $T_{a1} \approx 5/\omega_c$ (si veda la Tabella 4.2), l'ulteriore requisito sul tempo d'assestamento sarà soddisfatto pur di scegliere $\omega_c \geq 0.5$.

Per ottenere la voluta attenuazione del disturbo d , deve essere $|S(j\omega)|_{dB} \leq -20$ per $\omega \in [0, 0.1]$ e quindi, approssimativamente, $|L(j\omega)|_{dB} \geq 20$ nello stesso intervallo. Si noti tra l'altro che questo implica che sia $\omega_c \geq 0.1$, condizione già garantita dalla precedente specifica sulla pulsazione critica.

Infine, per soddisfare il requisito (d), deve essere $|F(j\omega)|_{dB} \leq -20$ per $\omega \in [10, 100]$ e quindi, approssimativamente, $|L(j\omega)|_{dB} \leq -20$ nello stesso intervallo. Anche questo vincolo non è in conflitto con la specifica su ω_c .

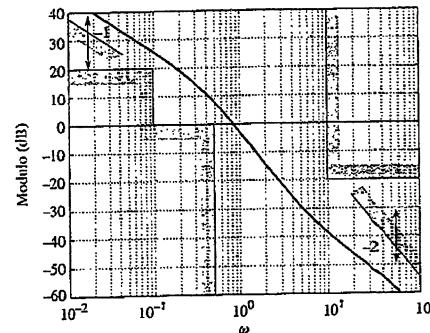


Figura 11.3
Rappresentazione grafica
dei vincoli di progetto
nell'Esempio 11.1, con un
esempio di diagramma di
 $|L(j\omega)|$ ammissibile.

Nella Figura 11.3 è rappresentata la regione nella quale deve trovarsi il diagramma di Bode di $|L(j\omega)|$ perché il sistema retroazionato fornisca le prestazioni desiderate. Si noti che, ad alta frequenza, la pendenza asintotica del diagramma deve essere almeno pari a -2 per assicurare la realizzabilità del regolatore, e che il vincolo sul margine di fase $\varphi_m \geq 60^\circ$ dovrà essere verificato a parte.

11.2.3 Altri requisiti

In molte applicazioni pratiche, il progettista di un sistema di controllo si trova a dover tener conto di altri tipi di vincoli connessi alla scelta del regolatore. Per esempio, normalmente si preferisce cercare una soluzione al problema di sintesi che non richieda un regolatore di ordine elevato. Infatti, dispositivi di regolazione di maggiore complessità comportano di solito costi superiori e sono reperibili più difficilmente sotto forma di componenti standard. D'altra parte, molti problemi di controllo possono essere risolti efficacemente con regolatori di struttura molto semplice, eventualmente anche privi di dinamica, come il regolatore proporzionale $R(s) = K_P$.

Inoltre, è spesso raccomandabile progettare regolatori che godano della proprietà di stabilità asintotica, anche se ciò che davvero importa è la stabilità del sistema retroazionato. Infatti, l'impiego di un regolatore non asintoticamente stabile può provocare indesiderati comportamenti divergenti nel caso di un malfunzionamento che porti il sistema di controllo a operare in anello aperto (per esempio in conseguenza di un guasto del trasduttore) e, come si vedrà nel Paragrafo 14.3.2, enfatizza gli effetti del cosiddetto fenomeno di wind-up in presenza di saturazione della variabile di controllo.

Limitazioni di questo genere sull'ordine o sulla struttura della funzione di trasferimento $R(s)$ non si prestano a essere facilmente incorporate nel quadro delle specifiche di progetto riguardanti la funzione d'anello. Come si vedrà, ciò suggerisce un diverso approccio al problema della sintesi del regolatore in presenza di questi vincoli o comunque tutte le volte che si voglia determinare una soluzione accettabile di minima complessità.

11.3 Approcci alla sintesi

Seguendo l'impostazione delineata nel precedente paragrafo, un primo approccio al problema della scelta del regolatore $R(s)$ potrebbe articolarsi nelle seguenti fasi:

- si traducono i requisiti del sistema di controllo in vincoli di progetto sulla funzione d'anello $L(s)$, del tipo di quelli esemplificati nella Figura 11.2, con l'aggiunta di ulteriori condizioni sul valore del margine di fase e/o del margine di guadagno e sull'assenza di cancellazioni tra singolarità con parte reale positiva o nulla;
- si individua una funzione $L^*(s)$ il cui diagramma di Bode del modulo giaccia all'interno della regione di ammissibilità e che soddisfi anche tutti i rimanenti vincoli;
- si determina la funzione di trasferimento del regolatore attraverso la relazione $R(s) = L^*(s)/G(s)$.

Il regolatore così ottenuto è una soluzione accettabile del problema di progetto, perché rispetta, per costruzione, tutti i requisiti posti inizialmente. D'altra parte, questo procedimento non permette di tener conto facilmente di vincoli aggiuntivi, quali quelli relativi alla struttura o all'ordine del regolatore. Anzi, come è facile capire, una scelta poco oculata della funzione $L^*(s)$ porta spesso a sintetizzare regolatori di complessità inutilmente elevata. Non va poi trascurato il fatto che nelle reali applicazioni non tutti gli obiettivi del controllo sono noti *a priori* al progettista in modo chiaro, preciso ed esauriente. La loro stessa formulazione è invece spesso frutto di scelte di compromesso, la cui significatività e le cui implicazioni possono essere valutate solo in base al confronto tra diverse soluzioni.

Un approccio alla sintesi che meglio si adatta a questo tipo di esigenze consiste invece nel procedere per successivi tentativi ragionati (*sintesi per tentativi*), cominciando a considerare come candidati alla soluzione regolatori di struttura molto semplice e complicandoli via via, se necessario, in modo da rispettare tutte le specifiche di progetto. In corrispondenza di ogni tentativo, gli strumenti dell'analisi dei sistemi retroazionati, e in particolare quelli basati sulla lettura dei diagrammi di Bode di $L(j\omega)$, possono essere impiegati per valutare se l'obiettivo è stato raggiunto o se sono invece necessari altri aggiustamenti.

Se si segue questa impostazione, è spesso comodo ragionare sulla funzione di trasferimento $R(s)$ del regolatore fattorizzata come $R(s) = R_1(s)R_2(s)$, dove

$$R_1(s) = \frac{\mu_R}{s^r} \quad (11.1)$$

viene chiamata *parte statica* del regolatore e

$$R_2(s) = \frac{\prod_i (1 + \tau_i s) \prod_i (1 + 2\zeta_i s/\omega_{ni} + s^2/\omega_{ni}^2)}{\prod_i (1 + T_i s) \prod_i (1 + 2\xi_i s/\omega_{ni} + s^2/\omega_{ni}^2)} \quad (11.2)$$

parte dinamica o anche *rete stabilizzatrice*. In effetti è facile intuire che solo il fattore $R_1(s)$ influenza sulle prestazioni statiche del sistema di controllo in risposta a segnali canonici, mentre $R_2(s)$ può essere utilizzata per modificare opportunamente le caratteristiche della funzione d'anello a pulsazione non nulla, in modo da rispettare i requisiti dinamici e garantire l'asintotica stabilità del sistema complessivo.

Il progetto allora si articola in due fasi:

- dapprima, assumendo per ipotesi che il sistema retroazionato al termine della sintesi sia asintoticamente stabile, si sceglie $R_1(s)$, cioè il tipo r e il guadagno μ_R del regolatore, in modo da soddisfare le specifiche statiche (*progetto statico*);
- in un secondo momento si considera la possibilità di introdurre un fattore $R_2(s) \neq 1$ per assicurare l'asintotica stabilità e il rispetto di tutti gli altri vincoli di progetto (*progetto dinamico*).

A volte, come si vedrà negli esempi seguenti, il valore del guadagno μ_R può rimanere indeterminato al termine del progetto statico, e in tal caso il problema della sua scelta fa parte del progetto dinamico.

La sintesi per tentativi consente una maggiore flessibilità al progettista, anche perché permette di concentrarsi in primo luogo sui requisiti fondamentali del sistema di controllo, raffinando eventualmente la soluzione in un secondo momento per soddisfare esigenze di importanza secondaria.

11.4 Esempi di progetto

Data l'ampia varietà di possibili situazioni e la difficoltà di elaborare una procedura di sintesi di validità generale, il problema del progetto per tentativi di un regolatore verrà illustrato in questo paragrafo attraverso una serie di esempi significativi. Peraltra, laddove sarà possibile, si evidenzieranno i principali criteri e alcune linee guida di vasta applicazione. Il primo esempio riguarda il controllo di un sistema a fase minima.

Esempio 11.2 Con riferimento allo schema di Figura 11.1, si debba progettare un regolatore per il sistema descritto da

$$G(s) = \frac{10}{(1+10s)(1+5s)(1+s)}$$

tenendo conto delle seguenti specifiche:

- $|e_{\text{sol}}| \leq 0.1$ in corrispondenza di $u(t) = A \text{ sca}(t)$ e $d(t) = B \text{ sca}(t)$, con $|A| \leq 1$ e $|B| \leq 5$;
- $\omega_c \geq 0.2$;
- $\varphi_m \geq 60^\circ$.

Si assume che la funzione di trasferimento del regolatore sia $R(s) = R_1(s)R_2(s)$, dove i fattori sono definiti dalle (11.1), (11.2), e si consideri dapprima la fase di progetto statico. Osservando che entrambi gli ingressi sono scalini, è immediato concludere che il vincolo (a) sull'errore a transitorio esaurito può essere soddisfatto, mediante un'opportuna scelta di $R_1(s)$, almeno in due modi. Infatti, con un regolatore di tipo $r = 0$ e guadagno μ_R positivo, considerando il caso pessimo in cui u e d hanno ampiezza massima e segno discorde, risulta (si veda la Tabella 10.1)

$$|e_{\infty}| = \frac{6}{1+10\mu_R}$$

e la condizione (a) è verificata, pur di prendere $\mu_R \geq 5.9$. Per garantire un certo grado di robustezza nei confronti di incertezze sul guadagno di $G(s)$, è conveniente scegliere un valore di μ_R un po' superiore al minimo, per esempio $\mu_R = 8$. D'altra parte, se si decide di utilizzare un regolatore di tipo $r = 1$, risulta comunque $e_{\infty} = 0$, indipendentemente dal valore del guadagno. Due possibili soluzioni al problema di progetto statico sono allora fornite da

$$R_1(s) = 8 \quad (11.3)$$

oppure

$$R_1(s) = \frac{\mu_R}{s} \quad (11.4)$$

con il parametro μ_R ancora libero.

Si completerà ora la sintesi con la fase di progetto dinamico, discutendo diverse soluzioni che originano dalla scelta preliminare di $R_1(s)$ secondo la (11.3) o la (11.4).

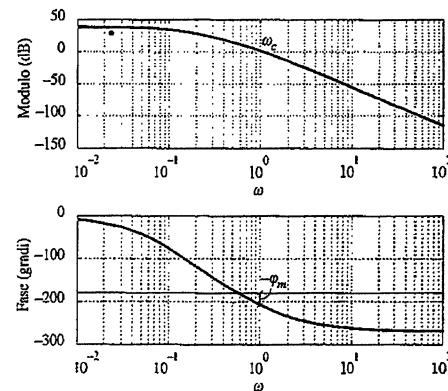
Progetto A

Se si sceglie $R_1(s)$ come nella (11.3), la funzione d'anello diventa $L(s) = L'(s)R_2(s)$ con

$$L'(s) = R_1(s)G(s) = \frac{80}{(1+10s)(1+5s)(1+s)}$$

Come primo tentativo si può porre $R_2(s) = 1$, che corrisponde a un regolatore puramente proporzionale, e verificare se le specifiche (b) e (c) sono soddisfatte. Dai diagrammi di Bode di $L'(j\omega)$ riportati nella Figura 11.4 si vede che $\omega_c \simeq 1$, ma $\varphi_m < 0^\circ$. Quindi, non solo è violata la condizione (c), ma il sistema di controllo è addirittura instabile. È perciò necessario utilizzare una rete

Figura 11.4
Diagrammi di Bode di $L'(j\omega)$ nell'Esempio 11.2 - Progetto A.



stabilizzatrice $R_2(s)$ che introduca un adeguato anticipo di fase in vicinanza della pulsazione critica desiderata.

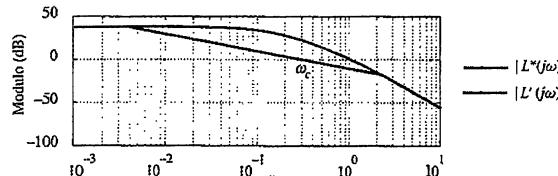
Uno dei modi per sintetizzare $R_2(s)$ consiste nel determinare il diagramma di Bode del modulo (asintotico) di una funzione d'anello di tentativo $L^*(s)$ che soddisfi i vincoli e ricavare quindi $R_2(s)$ come rapporto tra $L^*(s)$ e $L'(s)$. Si può costruire, per esempio, il diagramma di $|L^*(j\omega)|$ nel modo seguente. Si attribuisca alla pulsazione critica un valore di poco maggiore di quello minimo richiesto, per esempio $\omega_c = 0.3$, e si tracci una retta di pendenza -1 che intersechi l'asse a 0 dB proprio a quella pulsazione. Il diagramma di $|L^*(j\omega)|$ è poi ottenuto come mostrato nella Figura 11.5, raccordando questa retta con il diagramma originario di $|L'(j\omega)|$ a bassa e ad alta frequenza. Sotto l'ipotesi di fase minima, imporre l'attraversamento dell'asse con pendenza -1 dovrebbe favorire l'ottenimento di un soddisfacente margine di fase (si ricordi l'osservazione del Paragrafo 9.6.5). Attraverso questa costruzione grafica si giunge a determinare la funzione d'anello

$$L^*(s) = \frac{80}{(1+250s)(1+0.4s)^2}$$

che di fatto risolve il problema poiché assicura $\omega_c \simeq 0.3$ e $\varphi_m \simeq 76^\circ$. La rete stabilizzatrice $R_2(s)$ è quindi data da

$$R_2(s) = \frac{L^*(s)}{L'(s)} = \frac{(1+10s)(1+5s)(1+s)}{(1+250s)(1+0.4s)^2}$$

e il regolatore progettato ha funzione di trasferimento $R(s) = 8R_2(s)$.



Si osservi che la decisione di non modificare il diagramma di Bode della funzione d'anello a bassa frequenza è dovuta al fatto che il guadagno di $R_2(s)$ deve essere unitario, in modo da non interferire con il progetto statico. Non ci sono invece particolari tagioni per imporre che ad alta frequenza i diagrammi di $|L^*(j\omega)|$ e di $|L'(j\omega)|$ debbano coincidere. L'unica restrizione indotta dal vincolo di realizzabilità riguarderebbe le loro pendenze asintotiche. D'altra parte, diminuire la costante di tempo del polo doppio di $R_2(s)$ in alta frequenza implicherebbe una minore moderazione del controllo e un minor grado di robustezza della stabilità a fronte di incertezze su $G(s)$ ad alta frequenza, mentre aumentarla significherebbe ridurre il margine di fase. La scelta adottata può essere quindi interpretata come un ragionevole compromesso tra queste opposte esigenze.

Progetto B

La procedura di sintesi appena descritta ha portato a progettare una rete stabilizzatrice $R_2(s)$ che "cancella" completamente tramite i suoi zeri la dinamica del sistema sotto controllo, sostituendola

con altri tre poli opportunamente collocati. Si noti che tali cancellazioni sono lecite poiché riguardano singolarità con parte reale negativa. Se tra gli obiettivi del progetto fosse anche incluso il desiderio di ridurre la complessità del regolatore, si potrebbe considerare una seconda soluzione data da

$$R_2(s) = \frac{(1+10s)(1+5s)}{(1+250s)(1+0.4s)}$$

in cui si rinuncia a cancellare il polo ad alta frequenza di $G(s)$, quello cioè che meno contribuisce al ritardo di fase della funzione d'anello, ottenendo un regolatore $R(s) = 8R_2(s)$ del secondo ordine. Si lascia al lettore verificare che in corrispondenza di tale scelta risulta $\omega_c \simeq 0.3$ e $\varphi_m \simeq 67^\circ$, così che entrambe le specifiche (b) e (c) sono ancora rispettate.

Progetto C

Se non ci fossero controindicazioni ad ampliare la banda passante del sistema di controllo ben al di là del valore minimo richiesto, sarebbe proponibile il regolatore

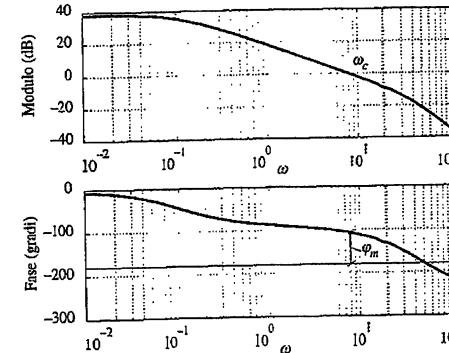
$$R(s) = \frac{8(1+5s)(1+s)}{(1+0.02s)^2}$$

i cui zeri cancellano due poli del sistema e i cui poli sono appropriatamente collocati ad alta frequenza. Si ottiene in tal caso la funzione d'anello

$$L(s) = \frac{80}{(1+10s)(1+0.02s)^2}$$

che, come si osserva nella Figura 11.6, presenta un margine di fase $\varphi_m \simeq 73^\circ$, con una pulsazione critica decisamente elevata ($\omega_c \simeq 7.8$). Non vanno però dimenticati i possibili inconvenienti connessi a questo genere di soluzione, specialmente in termini di moderazione dell'azione di controllo e di robustezza nei confronti di fenomeni dinamici in alta frequenza non modellizzati. Va inoltre sottolineato che, sebbene efficace nel velocizzare la risposta del sistema a una variazione del riferimento w , un sistema di controllo di questo tipo potrebbe non essere in grado di fare altrettanto nei riguardi di un disturbo. Si ricordi infatti che la variabile d che compare nello schema di Figura 11.1 non rappresenta di solito il vettore disturbo d_P agente sul sistema, ma va interpretata come l'uscita di un blocco con funzione di trasferimento $H_P(s)$ alimentato da d_P (si veda la discussione del Paragrafo 9.3). Perciò, denotando con $S(s) = (1+L(s))^{-1}$ la funzione di sensitività, la funzione di trasferimento tra d_P e y è $H_P(s)S(s)$. La risposta del sistema a una variazione a scalino del disturbo d_P dipende allora comunque dai poli di $H_P(s)$, che possono essere anche molto più "leni" di quelli di $S(s)$.

Figura 11.6
Diagrammi di Bode di $L(j\omega)$ nell'Esempio 11.2 - Progetto C.



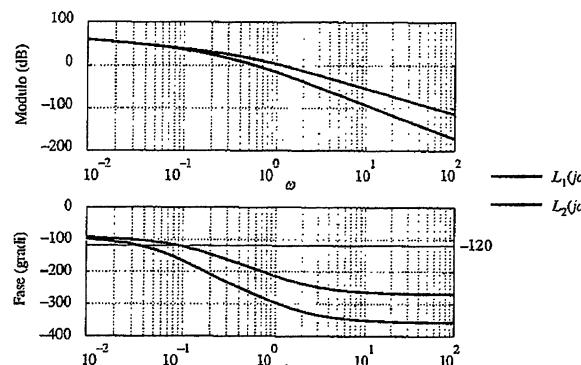
Progetto D

Si supponga ora di aver scelto la parte statica del regolatore come in (1.4). In tal caso, il guadagno μ_R è ancora impreciso e il suo valore può essere arbitrariamente fissato per soddisfare i requisiti dinamici (b) e (c). Si ponga come primo tentativo

$$R(s) = R_1(s) = \frac{\mu_R}{s}$$

ovvero $L(s) = \mu_R L_1(s)$, con $L_1(s) = \zeta(s)/s$, e si traccino i diagrammi di Bode di $L_1(j\omega)$ e di $L(j\omega)$ e di $L_1(j\omega)$ coincidono, mentre (Figura 11.7). Se $\mu_R > 0$, i diagrammi della fase di $L(j\omega)$ e di $L_1(j\omega)$ coincidono, mentre

Figura 11.7
Diagrammi di Bode
di $L_1(j\omega)$ e di $L_2(j\omega)$
nell'Esempio 11.2.
Progetto D.



il diagramma del modulo di $L(j\omega)$ risulta traslato verticalmente della quantità $|\mu_R|_{dB}$ rispetto a quello di $L_1(j\omega)$. Poiché dalla Figura 11.7 si osserva che $\arg L_1(j\omega)$ è una funzione monotona decrescente e risulta $\arg L_1(j\omega) = -120^\circ$ per $\omega \approx 0.03$, questa pulsazione rappresenta il massimo valore della pulsazione critica ottenibile con un'opportuna scelta di μ_R a fronte di un margine di fase $\varphi_m \geq 60^\circ$. Non è dunque possibile risolvere il problema di progetto senza introdurre nella funzione d'anello un opportuno anticipo di fase.

Prendendo spunto dai risultati del precedente Progetto A, si può condurre un tentativo dotando il regolatore di uno zero che cancelli il polo "lento" del sistema, cioè quello con costante di tempo 10. Si ipotizzi allora

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

e si considerino i diagrammi di Bode associati a $L_2(s) = R(s)G(s)/\mu_R$, anch'essi mostrati nella Figura 11.7. Poiché risulta $\arg L_2(j\omega) = -120^\circ$ per $\omega \approx 0.09$, l'azione anticipativa del regolatore permette di aumentare la pulsazione critica a parità di margine di fase, ma è ancora insufficiente per assicurare il rispetto delle specifiche.

Si consideri allora il regolatore

$$R(s) = \frac{\mu_R(1+10s)(1+5s)}{s(1+s)}$$

ottenuto dal precedente introducendo un altro zero che cancella il polo di $G(s)$ in $s = -0.2$ e un polo aggiuntivo in alta frequenza per rispettare il vincolo di realizzabilità. È immediato verificare che, scegliendo $\mu_R = 0.025$, si ottiene $\omega_c \approx 0.24$ e $\varphi_m \approx 63^\circ$, e il progetto è pertanto completato.

Confronto tra i regolatori progettati

I quattro regolatori progettati ai punti A, B, C e D rispettano tutte le specifiche originarie, ma generano comportamenti diversi del sistema retroazionato. Nella Figura 11.8 sono confrontati i rispettivi

Figura 11.8
Diagrammi di Bode dei
regolatori progettati
nell'Esempio 11.2.

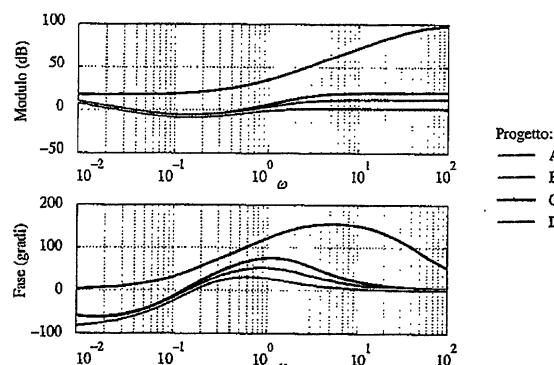
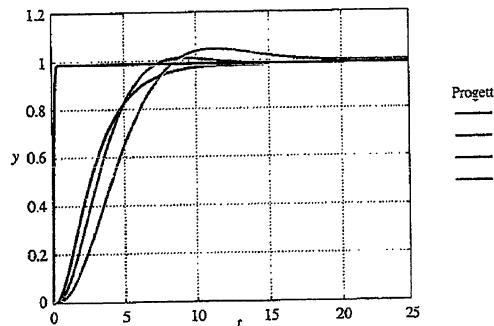


Figura 11.9
Risposta dell'uscita y a uno
scalino del riferimento w
per i sistemi di controllo
progettati nell'Esempio
11.2.



diagrammi di Bode. La soluzione C, che permette migliori prestazioni dinamiche, è sicuramente la più critica dal punto di vista della moderazione del controllo, visto che $|R(j\omega)|$ raggiunge valori molto elevati in alta frequenza.

Le risposte a uno scalino del riferimento w dei quattro sistemi di controllo sono riportate nella Figura 11.9, mentre la Figura 11.10 mette a confronto i corrispondenti andamenti della variabile di controllo u . Appare evidente come l'aumento della velocità di risposta nel caso del regolatore C sia ottenuto al prezzo di una maggiore sollecitazione della variabile di controllo (si noti che la curva C è scalata di un fattore 10^4). Gli altri tre regolatori offrono prestazioni molto simili tra loro, anche se il regolatore D, dotato di azione integrale, è l'unico a garantire errore nullo a transitorio esaurito, e oltre tutto in maniera robusta rispetto a incertezze sul guadagno di $G(s)$ (si veda il Paragrafo 10.4.1).

Benché i diversi regolatori assicurino valori analoghi per il margine di fase, si possono prevedere differenti caratteristiche di robustezza nei confronti di ritardi non modellizzati. In effetti, se si vuole conservare la stabilità, il sistema di controllo C può tollerare un ritardo inferiore a

$$\tau_C = \frac{\varphi_m}{\omega_c} \frac{\pi}{180} \approx 0.16$$

Progetto:
A
B
C
D

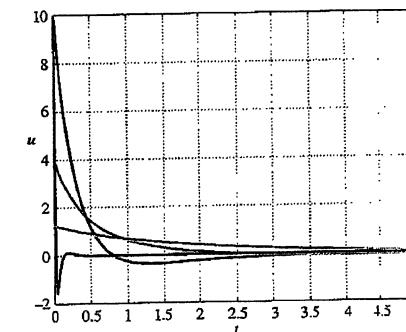


Figura 11.10
Risposta della variabile di
controllo u a uno scalino
del riferimento w per i
sistemi di controllo
progettati nell'Esempio
11.2. I valori della curva C
vanno moltiplicati per il
fattore 10^4 .

mentre in corrispondenza degli altri regolatori si ottengono i seguenti valori molto più elevati:

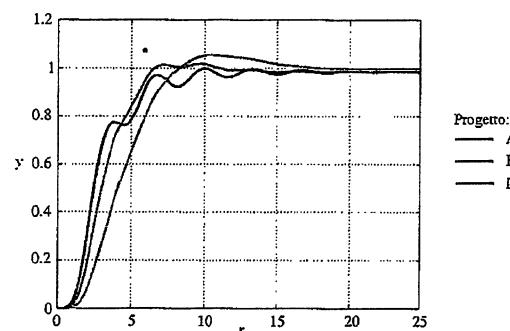
$$\tau_A \approx 4.42, \quad \tau_B \approx 3.90, \quad \tau_D \approx 4.58$$

Per quanto riguarda la robustezza nei confronti di incertezze sul guadagno di $G(s)$, i quattro valori del margine di guadagno sono:

$$k_{mA} \approx 15.67, \quad k_{mB} \approx 10.99, \quad k_{mC} \approx 12.55, \quad k_{mD} \approx 7.99$$

Si supponga poi che il sistema sotto controllo presenti in realtà una dinamica oscillatoria associata a una coppia di poli con pulsazione naturale $\omega_n = 2$ e smorzamento $\xi = 0.1$, di cui non si è tenuto conto in fase di progetto. È facile verificare che il regolatore C produce in tal caso un sistema instabile. Inoltre, osservando nella Figura 11.11 le risposte allo scalino dei rimanenti tre sistemi

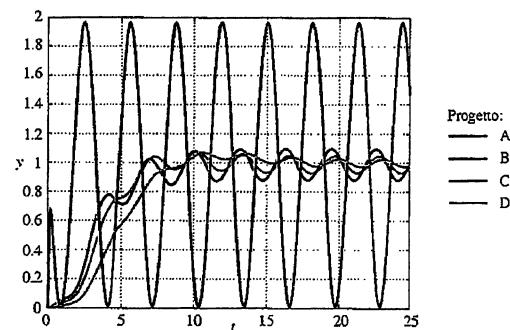
Figura 11.11
Risposta dell'uscita y a uno scalino del riferimento w in condizioni perturate per i sistemi di controllo progettati nell'Esempio 11.2.



di controllo, si nota che il regolatore che meno risente della presenza dei poli aggiuntivi è quello indicato con D, che tra tutti è quello che garantisce modulo della funzione d'anello più basso in $\omega = 2$.

Si consideri infine l'effetto di un disturbo sul trasduttore, modellizzato come $n(t) = \sin(2t)$. Alla luce della teoria svolta nel Capitolo 10, si deduce che anche in questo caso il regolatore capace di garantire la maggiore attenuazione di tale disturbo è il regolatore D, come confermato dalle simulazioni della risposta allo scalino riportate nella Figura 11.12.

Figura 11.12
Risposta dell'uscita y a uno scalino del riferimento w in presenza di un disturbo n per i sistemi di controllo progettati nell'Esempio 11.2.



L'esempio appena discusso suggerisce una possibile strategia di progetto valida nel caso di sistemi a fase minima. Si supponga infatti che $G(s)$ sia a fase minima, senza singolarità con parte reale nulla e con grado relativo v . Allora il regolatore

$$R(s) = \frac{1}{G(s)} \frac{\mu_R}{s(1+Ts)^{v-1}} \quad (11.5)$$

dove i poli in $s = -1/T$ con $0 < T \ll \mu_R^{-1}$ sono introdotti per rispettare il vincolo di realizzabilità, determina la funzione d'anello

$$L(s) = \frac{\mu_R}{s(1+Ts)^{v-1}}$$

che possiede una pulsazione critica $\omega_c \simeq \mu_R$ arbitrariamente elevata e un soddisfacente margine di fase $\varphi_m \simeq 90^\circ$. Dunque, in linea di principio, aumentando il guadagno μ_R del regolatore è possibile ottenere un sistema di controllo arbitrariamente veloce, asintoticamente stabile e privo di dinamica oscillante. I punti

debolli di questa tecnica di sintesi riguardano le prestazioni scadenti che ci si deve aspettare in termini di moderazione del controllo, di robustezza nei confronti di dinamica in alta frequenza non modellizzata, e di sensitività rispetto ai disturbi sulla linea di retroazione. Come si è visto nell'Esempio 11.2, è spesso più ragionevole allontanarsi dalla soluzione "ideale" (11.5) per venire incontro a queste esigenze.

Nel caso invece di sistemi a fase non minima, che contengono zeri con parte reale positiva o ritardi di tempo, esistono vincoli intrinseci sul massimo valore ottenibile per la banda passante con lo schema di Figura 11.1, come sarà illustrato nei due esempi che seguono.

Esempio 11.3 Il sistema da controllare nello schema di Figura 11.1 sia descritto da

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

e si voglia progettare un regolatore che garantisca un margine di fase $\varphi_m \geq 40^\circ$, cercando nel contempo di massimizzare la pulsazione critica ω_c .

Sia lo zero con parte reale positiva sia il polo nell'origine non sono cancellabili con corrispondenti singolarità del regolatore, pena la perdita della stabilità asintotica, e contribuiscono negativamente alla fase della funzione d'anello. La massima pulsazione ω_c ottenibile non potrà allora eccedere di molto il valore $\omega = 0.5$, pulsazione alla quale lo sfasamento dovuto a polo e zero è già di -135° .

Si ponga come primo tentativo $R(s) = \mu_R$. Analizzando i diagrammi di Bode di $G(j\omega)$ riportati nella Figura 11.13, risulta che $\arg G(j\omega) = -140^\circ$ in $\omega \simeq 0.085$. Visto l'andamento decrescente della fase, questa pulsazione rappresenta quindi il massimo valore di pulsazione critica ottenibile con il vincolo che sia $\varphi_m = 40^\circ$.

Le prestazioni possono essere migliorate da un regolatore che introduca un adeguato anticipo di fase, come per esempio

$$R(s) = \frac{\mu_R(1+10s)}{1+0.1s}$$

cui corrisponde la funzione d'anello $L(s) = \mu_R L_1(s)$, con

$$L_1(s) = \frac{0.1(1-2s)}{s(1+0.1s)^2}$$

Dalla Figura 11.13, si trova che $\arg L_1(j\omega) = -140^\circ$ in corrispondenza della pulsazione $\omega \simeq 0.49$, dove il modulo di $L_1(j\omega)$ vale circa -11 dB. Pertanto, la scelta $\mu_R = 11$ dB $\simeq 3.5$ permette di ottenere il margine di fase richiesto con $\omega_c \simeq 0.49$, che è un valore non lontano dal limite massimo prima individuato. Si noti peraltro che il margine di guadagno risulta abbastanza ridotto ($k_n \simeq 1.4$) e quindi la stabilità del sistema di controllo non è molto robusta nei confronti di eventuali

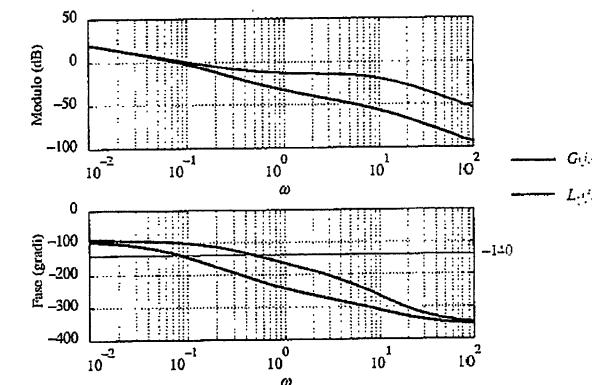


Figura 11.13
Diagrammi di Bode di $G(j\omega)$ e di $L_1(j\omega)$ nell'Esempio 11.3.

perturbazioni sul guadagno d'anello. La scarsa robustezza è anche messa in luce dal calcolo degli indicatori S_M e F_M , discusi nel Paragrafo 9.7, che assumono in questo caso valori relativamente elevati ($S_M \approx 12$ dB e $F_M \approx 8$ dB). Tutto ciò suggerirebbe di abbassare il valore del guadagno μ_R del regolatore rispetto a quello prima determinato. Per esempio, con $\mu_R = 2$, si ottengono $\omega_c \approx 0.22$, $\varphi_m \approx 64^\circ$, $k_m \approx 2.4$, $S_M \approx 5$ dB e $F_M \approx 0$ dB.

Esempio 11.4 Con riferimento allo schema di Figura 11.1, si debba progettare un regolatore per il sistema descritto da

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

rispettando le seguenti specifiche:

- a) $\epsilon_{\infty} = 0$ in presenza di una variazione a scalino del riferimento;
- b) $\omega_c \geq 0.1$;
- c) $\varphi_m \geq 30^\circ$.

Poiché il requisito (a) impone che il regolatore possieda un'azione integrale, la funzione d'anello risultante contiene in ogni caso due termini che introducono sfasamento negativo: il contributo del polo nell'origine è di -90° , mentre quello del ritardo di tempo è pari a $-4\omega_1 180/\pi$. Data la necessità di assicurare comunque l'asintotica stabilità del sistema retroazionato, esiste quindi anche in questo caso un vincolo sulla massima banda passante ottenibile. Infatti, alla luce del criterio di Bode, il massimo valore di pulsazione critica che garantisce asintotica stabilità sarà quello per cui la somma dei due contributi è -180° , ovvero $\omega_M = \pi/8 \approx 0.4$. Tenendo conto anche della specifica (b), la pulsazione critica ω_c deve allora cadere nell'intervallo $[0.1, 0.4]$. Se si utilizza il regolatore

$$R(s) = \frac{0.2}{s}$$

risulta $\omega_c \approx 0.2$, ma il margine di fase $\varphi_m \approx 22^\circ$ è insufficiente. Invece, con il regolatore

$$R(s) = \frac{0.2(1+s)}{s}$$

tutti i vincoli di progetto sono rispettati, in quanto risulta $\omega_c \approx 0.2$ e $\varphi_m \approx 33^\circ$.

Viene ora discusso un altro esempio di sintesi in cui il sistema sotto controllo presenta una dinamica oscillante poco smorzata.

Esempio 11.5 Si consideri il sistema di controllo della Figura 11.14, dove

$$G(s) = \frac{5}{(1+s)(1+0.01s+0.01s^2)}$$

Il sistema da controllare possiede quindi una coppia di poli complessi coniugati con $\omega_n = 10$ e $\xi = 0.05$. Nel progetto del regolatore si debbano rispettare le seguenti specifiche:

- a) $|\epsilon_{\infty}| \leq 0.05$ quando $w(t) = \text{sca}(t)$;
- b) $\varphi_m \geq 45^\circ$;
- c) $k_m \geq 4$.

Si desidera esplorare dapprima la possibilità di utilizzare un regolatore del primo ordine, esaminando in seguito altre soluzioni più complicate. In ogni caso si richiede, a progetto ultimato, di valutare le prestazioni del regolatore in termini di attenuazione del disturbo \tilde{d} sulla variabile di controllo e di robustezza rispetto a incertezze sulla pulsazione naturale ω_n dei poli complessi coniugati di $G(s)$.

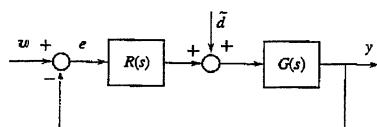
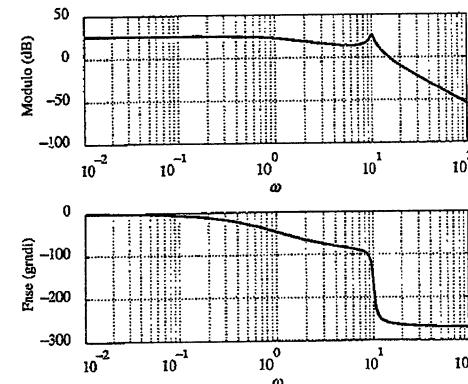


Figura 11.14
Sistema di controllo
dell'Esempio 11.5.

Figura 11.15
Diagrammi di Bode di
 $L'(j\omega)$ nell'Esempio 11.5 -
Progetto A.



Progetto A

È immediato verificare che, per rispettare il vincolo (a), occorre un regolatore dotato di azione integrale oppure un regolatore di tipo zero, ma con guadagno $\mu_R \geq 19/5$. Si fissi allora $\mu_R = 4$ e si traccino i diagrammi di Bode associati alla funzione $L'(s) = 4G(s)$ (Figura 11.15). La presenza della coppia di poli a basso smorzamento rende impossibile risolvere il problema di progetto con un regolatore del primo ordine, a meno di accontentarsi di valori della pulsazione critica molto inferiori a ω_n . In tal caso è opportuno collocare il polo del regolatore a bassa frequenza, per esempio in $s = -0.1$, facendo così in modo che il diagramma del modulo della risposta in frequenza d'anello attraversi l'asse a 0 dB a una pulsazione di poco superiore a $\omega = 1$. Con il regolatore

$$R(s) = \frac{4}{1+10s}$$

si ottengono $\omega_c \approx 1.26$, $\varphi_m \approx 42^\circ$ e $k_m \approx 12.66$. Poiché la specifica (b), anche se di poco, non è soddisfatta, si può introdurre uno zero che aumenti il margine di fase. Si tenga conto peraltro che, se collocato a pulsazione inferiore a ω_n , esso produce l'effetto secondario di avvicinare pericolosamente all'asse a 0 dB il picco di risonanza dovuto ai poli complessi, contribuendo a ridurre il margine di guadagno. Dopo alcuni tentativi si trova che il regolatore

$$R(s) = \frac{4(1+s/9)}{1+10s}$$

rispetta tutti i vincoli in quanto risulta $\varphi_m \approx 50^\circ$ e $k_m \approx 4.02$.

Per valutare la capacità del sistema di controllo progettato di attenuare l'effetto del disturbo \tilde{d} si osservi che, nello schema di Figura 11.14, la funzione di trasferimento tra \tilde{d} e y è data da $M(s) = G(s)S(s)$, dove $S(s) = (1+L(s))^{-1}$ è la consueta funzione di sensitività, e il relativo diagramma di Bode del modulo è riportato nella Figura 11.16 (linea in colore). Da esso si vede che le componenti

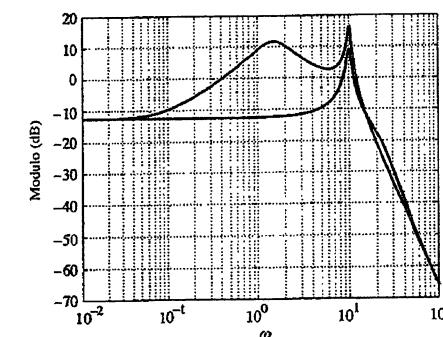
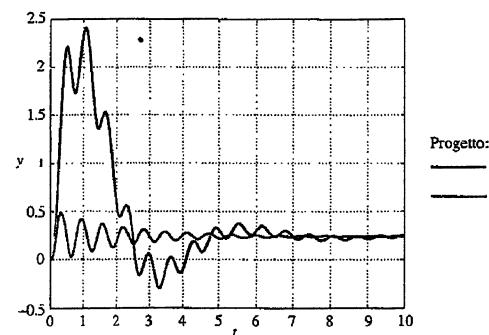


Figura 11.16
Diagrammi di Bode di
 $|M(j\omega)|$ nell'Esempio 11.5 -
Progetti A e B.

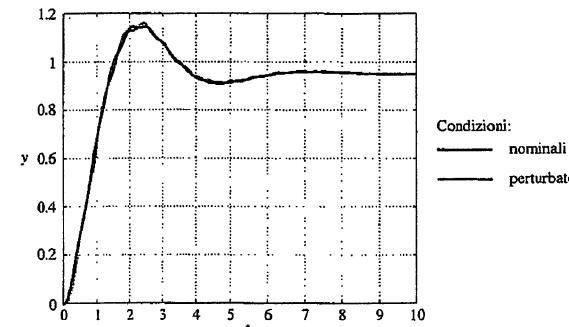
Progetto:
— A
— B

Figura 11.17
Risposta dell'uscita y a uno scalino del disturbo \tilde{d} nell'Esempio 11.5 - Progetti A e B.



del disturbo a pulsazioni comprese nell'intervallo $[0.4, 12]$ subiscono un'amplificazione, mentre quelle a bassa e ad alta frequenza vengono attenuate. Ciò si traduce nell'andamento della risposta allo scalino mostrato nella Figura 11.17 (linea in colore).

Figura 11.18
Risposta dell'uscita y a uno scalino del riferimento w in condizioni nominali e perturbate nell'Esempio 11.5 - Progetto A.



Quanto alla robustezza delle prestazioni in presenza di incertezze sulla pulsazione naturale dei poli, è lecito attendersi una relativa insensibilità del sistema di controllo a variazioni di ω_n poiché esse riguardano una dinamica a frequenza elevata rispetto a ω_c . Ciò è confermato dal confronto, riportato nella Figura 11.18, fra le risposte di y a uno scalino unitario del riferimento w nel caso nominale $\omega_n = 10$ e nel caso in cui sia in realtà $\omega_n = 12$.

Progetto B

Impiegando un regolatore di ordine più elevato è possibile migliorare le prestazioni rispetto alla soluzione precedentemente individuata. In particolare, può essere conveniente utilizzare un regolatore del tipo

$$R(s) = \frac{4(1 + 0.01s + 0.01s^2)}{(1 + 0.025s)^2}$$

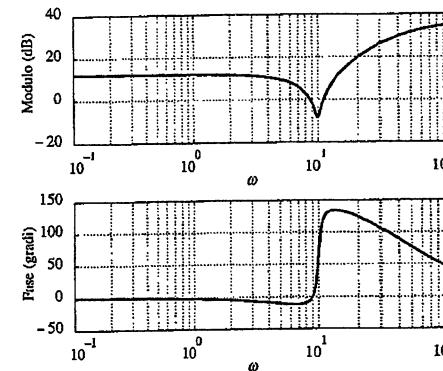
che cancella con i propri zeri i poli del sistema a basso smorzamento (Figura 11.19). La funzione d'anello corrispondente è

$$L(s) = \frac{20}{(1 + s)(1 + 0.025s)^2}$$

e risulta $\omega_c \approx 16.95$, $\varphi_m \approx 47^\circ$ e $k_m \approx 4.19$, cosicché le specifiche sono rispettate.

Benché per effetto della cancellazione poli/zeri la dinamica oscillante scompaia nella funzione di sensitività complementare, essa è invece presente nella funzione di trasferimento $M(s)$ tra il disturbo \tilde{d} e l'uscita y . È infatti facile verificare che tale funzione possiede tra i suoi poli, oltre alle radici di $1 + L(s)$, anche i poli complessi coniugati di $G(s)$. Quindi, la risposta a uno scalino di \tilde{d} presenta oscillazioni molto poco smorzate, come si vede nella Figura 11.17 (linea in nero). L'azione filtrante del sistema di controllo nei confronti del disturbo \tilde{d} può essere meglio apprezzata osservando il modulo della risposta in frequenza $M(j\omega)$, riportato come curva in nero nella Figura 11.19.

Figura 11.19
Diagrammi di Bode di $R(j\omega)$ nell'Esempio 11.5 - Progetto B.



Un aspetto poco soddisfacente del progetto B riguarda la sua scarsa robustezza nei confronti di eventuali incertezze sulla pulsazione naturale ω_n e sullo smorzamento ξ dei poli del sistema. Bastano infatti piccole deviazioni dai loro valori nominali per far sì che la cancellazione poli/zeri non sia più accurata. A titolo illustrativo, nella Figura 11.20 sono confrontati gli andamenti dell'uscita y e risposta a uno scalino unitario del riferimento w nel caso nominale e nel caso in cui l'effettiva pulsazione naturale sia $\omega_n = 12$, anziché $\omega_n = 10$. Come si vede, tale perturbazione determina scostamenti anche significativi rispetto alla risposta nominale e la presenza di oscillazioni poco smorzate.

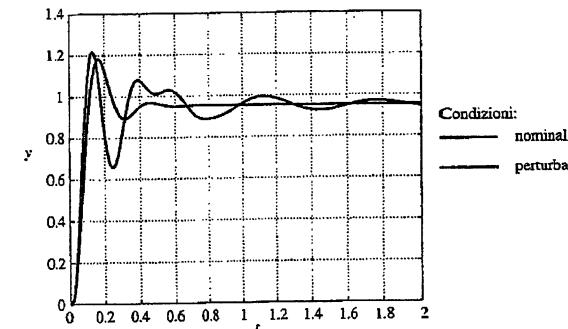


Figura 11.20
Risposta dell'uscita y a uno scalino del riferimento w in condizioni nominali e perturbate nell'Esempio 11.5 - Progetto B.

Si conclude questo paragrafo con un esempio di progetto più articolato.

Esempio 11.6 Si consideri il sistema di controllo descritto dallo schema di Figura 11.21, dove

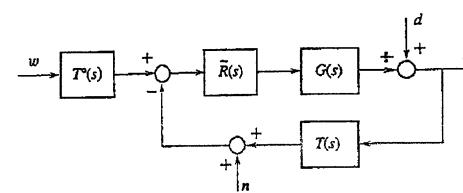


Figura 11.21
Sistema di controllo dell'Esempio 11.6.

la funzione di trasferimento del processo è

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

La costante di tempo τ non è conosciuta, ma si sa che è positiva e minore di 0.04. Il blocco

$$T(s) = \frac{0.1}{1+0.2s}$$

rappresenta il trasduttore; la sua uscita è affetta dal disturbo n il cui spettro è diverso da zero solo nell'intervallo di pulsazioni [50, 100]. Si assuma inoltre che il riferimento sia $w(r) = ram(t)$, mentre il disturbo d sia modellizzato come una sinusoide di ampiezza unitaria e di pulsazione compresa nell'intervallo [0.01, 0.5].

Lo scopo del progetto sia quello di scegliere le funzioni di trasferimento $T^o(s)$ e $\tilde{R}(s)$ in modo da rispettare i seguenti vincoli:

- a) l'errore a transitorio esaurito dovuto a w e d abbia modulo inferiore a 0.1;
- b) l'attenuazione del disturbo n sia di almeno 20 dB;
- c) il margine di fase sia $\varphi_m \geq 30^\circ$.

Secondo quanto discusso nel Paragrafo 9.3, è opportuno porre innanzitutto $T^o(s) = T(s)$, così da ricondurre lo schema a quello di Figura 11.1, con $R(s) = \tilde{R}(s)T(s)$.

Le ipotesi sui disturbi d e n suggeriscono poi di fissare il valore della pulsazione critica ω_c nell'intervallo [0.5, 50], in modo da poter garantire una buona attenuazione di entrambi. Tra l'altro, la scelta di limitare ω_c consente anche di rendere meno critico il progetto rispetto alla presenza della costante di tempo incerta τ , che contribuisce a modificare la funzione d'anello in alta frequenza.

Si scomponga ora la funzione di trasferimento $\tilde{R}(s)$ da sintetizzare nei fattori che ne rappresentano la parte statica e quella dinamica:

$$\tilde{R}(s) = \tilde{R}_1(s)\tilde{R}_2(s) = \frac{\mu_R}{s^r}\tilde{R}_2(s) \quad , \quad \tilde{R}_2(0) = 1$$

La funzione d'anello è allora data da

$$L(s) = \tilde{R}(s)T(s)G(s) = \frac{\mu_R}{s^r}\tilde{R}_2(s) \frac{-0.2(1+s)}{(1+10s)(1+0.2s)(1+\tau s)} \quad (11.6)$$

Progetto statico

Poiché il sistema da controllare ha guadagno negativo, è necessario che sia $\mu_R < 0$, in modo da garantire la positività del guadagno d'anello (condizione necessaria per l'asintotica stabilità secondo il criterio di Bode). Inoltre, per avere errore a regime finito con un riferimento a rampa è necessario che sia almeno $r = 1$. In tal caso, indicando con $e_{w\infty}$ ed $e_{d\infty}$ i contributi all'errore e_∞ dovuti rispettivamente a w e a d , risulta

$$|e_{w\infty}| = \frac{1}{0.2|\mu_R|} \quad , \quad |e_{d\infty}| \leq \max_{\omega \in [0.01, 0.5]} \frac{1}{|1+L(j\omega)|}$$

e il requisito (a) è soddisfatto se è verificata la seguente relazione:

$$|e_\infty| \leq |e_{w\infty}| + |e_{d\infty}| \leq 0.1$$

Decidendo allora per semplicità di equipartire l'errore massimo tra i due contributi, si ottengono le diseguaglianze

$$\begin{aligned} \frac{1}{0.2|\mu_R|} &\leq 0.05 \\ \frac{1}{|1+L(j\omega)|} &\leq 0.05 \quad , \quad \omega \in [0.01, 0.5] \end{aligned}$$

La prima impone il vincolo $|\mu_R| \geq 100$ sul guadagno del regolatore, mentre la seconda, osservando che i valori di pulsazione interessati sono comunque inferiori a ω_c e dunque $|1+L(j\omega)| \simeq |L(j\omega)|$, può essere approssimata dalla relazione

$$|L(j\omega)| \geq 20 \quad , \quad \omega \in [0.01, 0.5] \quad (11.7)$$

Il progetto statico si conclude quindi ponendo $\tilde{R}_1(s) = -100/s$, e ricordando che nel progetto dinamico si dovrà tenere conto della condizione (11.7), che si traduce nel richiedere che il diagramma di Bode del modulo della funzione d'anello sia superiore a 26 dB nell'intervallo $\omega \in [0.01, 0.5]$.

Progetto dinamico

Determinata $\tilde{R}_1(s)$ come appena visto, la funzione d'anello (11.6) diventa

$$L(s) = \frac{20(1+s)}{s(1+10s)(1+0.2s)(1+\tau s)} \tilde{R}_2(s) = L'(s) \frac{\tilde{R}_2(s)}{(1+\tau s)}$$

Nel progetto di $\tilde{R}_2(s)$ si dovranno considerare le specifiche su ω_c e φ_m , la condizione (11.7) e il requisito (b). Poiché la funzione di trasferimento tra il disturbo n e l'errore e è la funzione di sensitività complementare $F(s) = L(s)/(1+L(s))$ e lo spettro del disturbo è concentrato nell'intervallo $\omega \in [50, 100]$, al di là della pulsazione critica, quest'ultimo requisito impone che sia

$$|L(j\omega)| \leq 0.1 \quad , \quad \omega \in [50, 100] \quad (11.8)$$

ovvero che il diagramma del modulo della funzione d'anello sia inferiore a -20 dB in quell'intervallo di pulsazioni.

Il diagramma di Bode di $|L'(j\omega)|$, tracciato nella Figura 11.22, mostra che la scelta $\tilde{R}_2(s) = 1$ non permette comunque di rispettare il vincolo (11.7), indipendentemente dal valore di τ . Quindi, se non si vuole modificare il valore di μ_R , è indispensabile introdurre nel regolatore uno zero a bassa frequenza. Per esempio, ponendo

$$\tilde{R}_2(s) = \frac{1+10s}{1+2s}$$

e trascurando per il momento il polo con costante di tempo τ , il diagramma associato alla funzione d'anello $L''(s) = L'(s)\tilde{R}_2(s)$ (si veda ancora la Figura 11.22, linea in colore chiaro) soddisfa le condizioni (11.7) e (11.8) e rispetta il vincolo sulla pulsazione critica ($\omega_c \simeq 6.3$). È però immediato verificare che il margine di fase risultante è modesto ($\varphi_m \simeq 34^\circ$), e tale da non consentire di rispettare il requisito (c) quando la costante di tempo del polo aggiuntivo è maggiore di $\tan(4^\circ)/\omega_c \simeq 0.011$ (in questo calcolo si presume, come in effetti accade, che la presenza del polo non alteri di molto il valore di ω_c). Pertanto, è conveniente cancellare con uno zero del regolatore il polo in $s = -5$ e aggiungere un polo in $s = -20$ per non violare la condizione (11.8). Il progetto risultante è allora descritto da

$$\tilde{R}_2(s) = \frac{(1+10s)(1+0.2s)}{(1+2s)(1+0.05s)}$$

cuì è associata la funzione d'anello $L'''(s) = L'(s)\tilde{R}_2(s)$, rappresentata con linea nera nella Figura 11.22. Si trova $\omega_c \simeq 9.1$ e $\varphi_m \simeq 62^\circ$ e, nel caso più sfavorevole in cui sia $\tau = 0.04$, la pulsazione critica cambia di poco ($\omega_c \simeq 8.7$) e il margine di fase rimane comunque accettabile ($\varphi_m \simeq 44^\circ$).

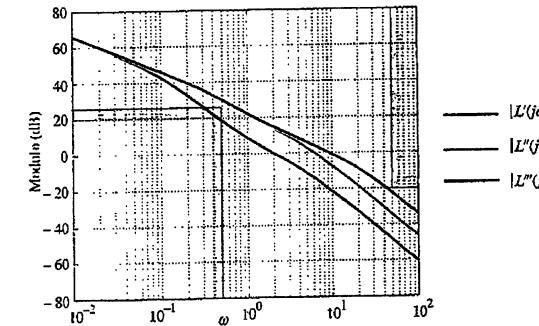


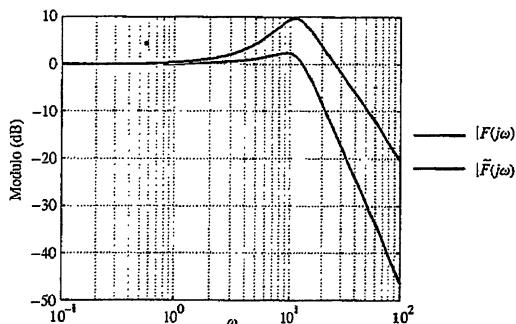
Figura 11.22
Diagrammi di Bode
di $|L'(j\omega)|$, $|L''(j\omega)|$ e
 $|L'''(j\omega)|$ nell'Esempio 11.6.

Discussione sul progetto di $T^o(s)$
Il problema di sintesi è stato risolto scegliendo

$$T^o(s) = \frac{0.1}{1+0.2s} \quad , \quad \tilde{R}(s) = \frac{-100(1+10s)(1+0.2s)}{s(1+2s)(1+0.05s)} \quad (11.9)$$

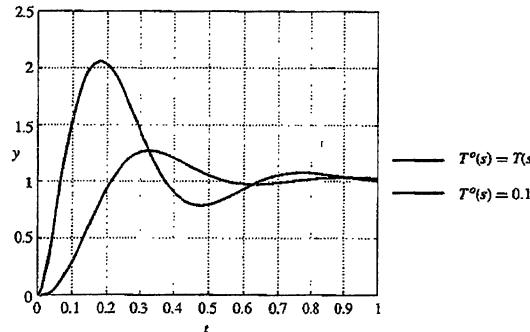
Si vuole ora discutere l'effetto di una diversa scelta di $T^o(s)$. Si osservi innanzitutto dalla Figura 11.21 che tale effetto riguarda solo la funzione di trasferimento tra il riferimento w e l'uscita y , che vale $\tilde{F}(s) = T^o(s)T(s)^{-1}F(s)$. Se si sceglie $T^o(s) = T(s)$ come nella (11.9), risulta ovviamente $\tilde{F}(s) = F(s)$. In ogni caso, il guadagno di $T^o(s)$ deve essere uguale a quello di $T(s)$ in modo che risulti $\tilde{F}(0) = F(0) = 1$, e quindi $\tilde{F}(s)$ abbia guadagno unitario. Sarebbe quindi possibile porre

Figura 11.23
Diagrammi di Bode
di $|F(j\omega)|$ e $|\tilde{F}(j\omega)|$
nell'Esempio 11.6.



per semplicità $T^0(s) = 0.1$, trascurando la dinamica del trasduttore. L'effetto però sarebbe quello di introdurre in $\tilde{F}(s)$ uno zero aggiuntivo in $s = -5$, che allontanerebbe il comportamento di $\tilde{F}(s)$ da quello di un filtro passa-basso. Infatti il diagramma di Bode del modulo di $\tilde{F}(j\omega)$ assumerebbe l'andamento mostrato nella Figura 11.23, dove è anche riportato per confronto quello di $F(j\omega)$ (in entrambi i casi si assume $\tau = 0.04$). Le conseguenze di ciò sulla risposta del sistema a uno scalino del riferimento sono evidenti nel grafico della Figura 11.24.

Figura 11.24
Risposta dell'uscita y a uno
scalino del riferimento w
con $T^0(s) = T(s)$
 $T^0(s) = 0.1$ nell'Esempio
11.6.



Si noti che il fenomeno appena discusso si presenta tutte le volte che il trasduttore possiede poli a pulsazione inferiore a ω_c che non vengono replicati nella funzione di trasferimento $T^0(s)$, in quanto i poli di $T(s)$ sono zeri di $\tilde{F}(s)$.

Vale anche la pena di osservare che la scelta $T^0(s) = 0.1$ non consentirebbe in questo caso di rispettare il vincolo di progetto (a). Infatti, indicando con

$$E(s) = W(s) - Y(s) = (1 - \tilde{F}(s)) W(s)$$

la trasformata dell'errore, in corrispondenza di $W(s) = 1/s^2$ si ricava

$$e_{w\infty} = \lim_{s \rightarrow 0} s E(s) = -\frac{3}{20}$$

e il vincolo (a) è sicuramente violato.

11.5 Principali reti stabilizzatrici

Nel risolvere i problemi di progetto illustrati negli esempi del precedente paragrafo si è fatto uso di regolatori $R(s)$ dalla struttura molto semplice, riconducibili a poche e ben individuabili tipologie. Le principali classi di reti stabilizzatrici saranno ora descritte più in dettaglio, fornendo per ciascuna di esse le ragioni che

ne motivano l'impiego e i principali criteri di progetto. Mentre nel precedente paragrafo si era parlato di rete stabilizzatrice con riferimento alla funzione di trasferimento (11.2), che ha guadagno unitario, nella trattazione che segue si userà tale termine in un'accezione più estesa, ammettendo che una rete stabilizzatrice possa anche avere un guadagno diverso da 1.

11.5.1 Rete anticipatrice

Una rete *anticipatrice* (o *lead network*) è descritta dalla funzione di trasferimento

$$R(s) = \mu_R \frac{1 + Ts}{1 + \alpha Ts} \quad (11.10)$$

con $\mu_R > 0$, $T > 0$, $0 < \alpha < 1$. Come si vede dai diagrammi di Bode associati alla (11.10), mostrati nella Figura 11.25, questo tipo di regolatore permette di ottenere un anticipo di fase che raggiunge il suo massimo in corrispondenza della pulsazione $\bar{\omega} = 1/(T\sqrt{\alpha})$, che rappresenta il punto medio (in scala logaritmica) tra le pulsazioni $1/T$ e $1/\alpha T$ associate rispettivamente allo zero e al polo. È immediato verificare che

$$\arg R(j\bar{\omega}) = \arctan \frac{1}{\sqrt{\alpha}} - \arctan \sqrt{\alpha} \quad (11.11)$$

Se il parametro T del regolatore viene scelto in modo che la pulsazione $\bar{\omega}$ cada in vicinanza della pulsazione critica desiderata, è lecito attendersi un effetto stabilizzante sul sistema di controllo, in quanto l'anticipo di fase introdotto nella funzione d'anello in corrispondenza di ω_c tende a favorire l'ottenimento di valori più elevati del margine di fase. Naturalmente, l'azione anticipativa può essere resa più efficace riducendo il valore di α , fino ad arrivare a un contributo alla fase prossimo a 90° , ma questo determina sul diagramma del modulo una maggiore amplificazione ad alta frequenza. Come più volte discusso, quest'ultima caratteristica produce effetti negativi sulle prestazioni del controllore in termini di moderazione del controllo, di robustezza e di attenuazione dei disturbi di misura in alta frequenza. La scelta spesso adottata di collocare lo zero e il polo alla distanza relativa di una decade (ponendo cioè $\alpha = 0.1$) rappresenta un buon compromesso tra le diverse esigenze e comporta un massimo anticipo di fase di circa 55° in corrispondenza di $\bar{\omega}$ (si veda la (11.11)). Ovviamente, nell'ampliare la banda passante mediante l'uso di una rete anticipatrice, va posta una certa cautela per evitare che ω_c si sposti a valori di pulsazione troppo elevati, in corrispondenza dei quali il sistema

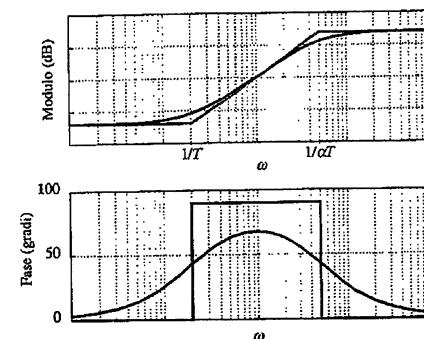
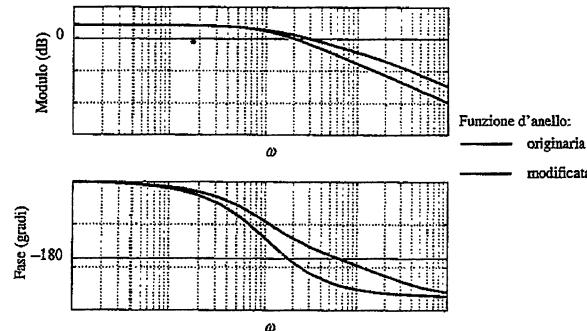


Figura 11.25
Diagrammi di Bode di una
rete anticipatrice.

Figura 11.26
Effetto dell'introduzione di una rete anticipatrice: diagrammi di Bode della funzione d'anello originaria e di quella modificata.



da controllare può presentare uno sfasamento così negativo da rendere inefficace l'antico di fase e pregiudicare addirittura la stabilità.

L'effetto dell'introduzione di una rete anticipatrice sulle caratteristiche della funzione d'anello è esemplificato dai diagrammi di Bode della Figura 11.26, dove le curve in colore rappresentano i diagrammi di una tipica funzione d'anello e quelle in nero mostrano le modifiche derivanti dall'impiego di una rete anticipatrice a guadagno unitario opportunamente progettata.

Nel caso limite $\alpha = 0$, la rete anticipatrice assume l'aspetto di un regolatore PD (ad azione proporzionale e derivativa), che sarà oggetto di un'analisi specifica nel Capitolo 14.

11.5.2 Rete ritardatrice

Una *rete ritardatrice* (o lag network) è descritta dalla funzione di trasferimento

$$R(s) = \mu_R \frac{1 + Ts}{1 + \alpha Ts}$$

con $\mu_R > 0$, $T > 0$, $\alpha > 1$. In questo caso i diagrammi di Bode hanno l'aspetto mostrato nella Figura 11.27 e mettono in luce che un regolatore di questo tipo può contribuire ad aumentare il modulo della funzione d'anello a bassa frequenza, al prezzo di introdurre uno sfasamento negativo che raggiunge il suo minimo in $\bar{\omega} = 1/(T\sqrt{\alpha})$.

L'impiego di una rete ritardatrice è indicato tutte le volte che si intende migliorare la precisione statica o garantire una maggiore attenuazione delle com-

Figura 11.27
Diagrammi di Bode di una rete ritardatrice.

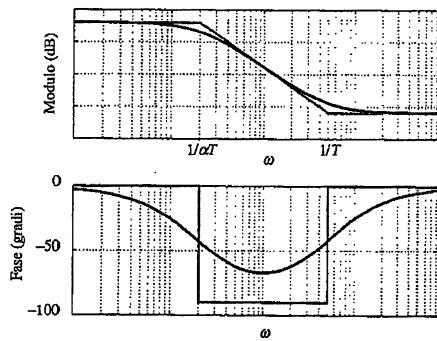
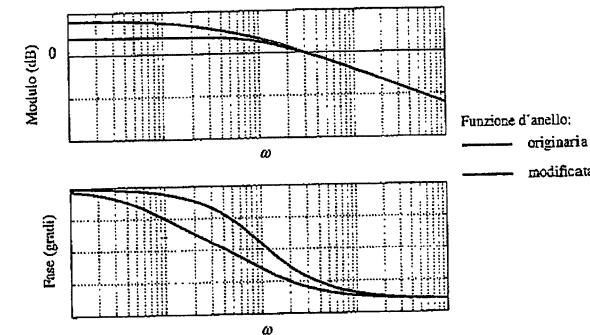
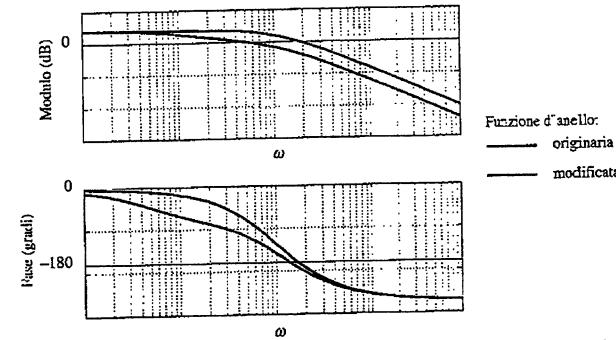


Figura 11.28
Effetto dell'introduzione di una rete ritardatrice: diagrammi di Bode della funzione d'anello originaria e di quella modificata.



ponenti a bassa frequenza del disturbo d . Ovviamente, per evitare che il ritardo di fase introdotto si ripercuota troppo sul valore del margine di fase φ_m , occorre scegliere $T > 1/\omega_c$. Per esempio, se $T = 10/\omega_c$, lo sfasamento negativo prodotto dalla rete ritardatrice in corrispondenza della pulsazione critica è in valore assoluto comunque inferiore a 6° . Un esempio di impiego di una rete ritardatrice con $\mu_R = \alpha$, così da non alterare il comportamento in alte frequenze, è illustrato con la Figura 11.28. Da un differente punto di vista (si osservi la Figura 11.29), si può notare che l'introduzione nella funzione d'anello di una rete ritardatrice con guadagno μ_R unitario può servire ad abbassare il diagramma di $|L(j\omega)|$ in modo che l'attraversamento dell'asse a 0 dB avvenga a una pulsazione per la quale il valore della fase della funzione d'anello sia meno sfavorevole, permettendo così di migliorare il margine di fase.

Figura 11.29
Effetto dell'introduzione di una rete ritardatrice con guadagno unitario: diagrammi di Bode della funzione d'anello originaria e di quella modificata.



Al crescere del parametro α verso l'infinito, il polo della rete ritardatrice si sposta sempre più a bassa frequenza fino a trasformarsi in un polo nell'origine. In tal caso, il regolatore diventa un regolatore PI (ad azione proporzionale e integrale), che sarà oggetto di studio nel Capitolo 14.

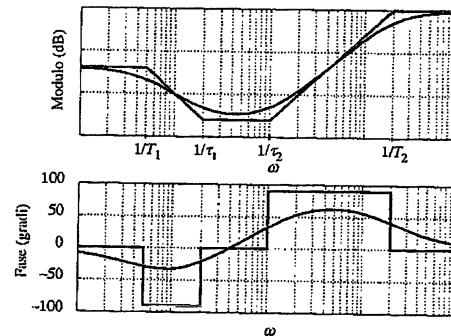
11.5.3 Rete a sella

La combinazione di una rete anticipatrice e di una ritardatrice dà luogo alla cosiddetta *rete a sella* (o lead-lag network), descritta dalla funzione di trasferimento

$$R(s) = \mu_R \frac{(1 + \tau_1 s)(1 + \tau_2 s)}{(1 + T_1 s)(1 + T_2 s)}$$

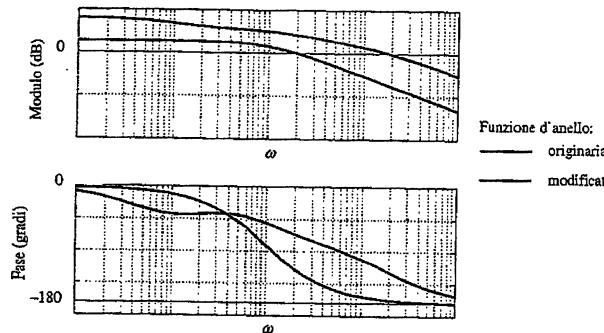
con $\mu_R > 0$, $T_1 > \tau_1 \geq \tau_2 > T_2 > 0$, e rappresentata dai diagrammi di Bode della Figura 11.30. Spesso si pone $T_1 T_2 = \tau_1 \tau_2$, cosicché il valore di $|R(j\omega)|$ per ω che tende all'infinito coincide con il guadagno μ_R .

Figura 11.30
Diagrammi di Bode di una rete a sella.



L'uso di una rete a sella consente di sfruttare contemporaneamente i vantaggi della parte ritardatrice e di quella anticipatrice, modificando i diagrammi di Bode della funzione d'anello come esemplificato nella Figura 11.31. Per ottenere benefici consistenti dall'anticipo di fase introdotto da questo tipo di regolatore, è opportuno scegliere le costanti di tempo di poli e zeri in modo che risulti $\tau_2 > 1/\omega_c > T_2$, dove ω_c rappresenta il valore desiderato della pulsazione critica.

Figura 11.31
Effetto dell'introduzione di una rete a sella: diagrammi di Bode della funzione d'anello originaria e di quella modificata.



Si vedrà nel Capitolo 14 che un regolatore PID (nella sua forma "reale") possiede la tipica struttura di una rete a sella.

11.6 Conclusioni

Le tecniche di analisi di sistemi retroazionati basate sui diagrammi di Bode della funzione d'anello costituiscono, come si è visto, il supporto per efficaci procedure di progetto del regolatore, almeno in quei casi in cui è possibile riformulare gli obiettivi del controllo in termini di specifiche sul valore di indicatori sintetici, quali la pulsazione critica, il margine di fase e di guadagno e l'errore a transitorio esaurito. Non sono invece direttamente affrontabili con i metodi illustrati in questo capitolo i problemi di progetto di controlleri per sistemi instabili, la cui trattazione è rimandata al Capitolo 15, dove si illustreranno anche architetture del sistema di

controllo più complesse di quelle finora analizzate. Nei Capitoli 12 e 13 saranno invece discusse tecniche di sintesi differenti, non basate sull'uso della risposta in frequenza.

Esercizi

Esercizio 11.1 Si debba progettare un sistema di controllo con le seguenti caratteristiche:

- risultati $|\epsilon_{oo}| < 0.1$ in presenza di un riferimento a scalino $w(t) = \text{sca}(t)$;
- la banda passante della funzione di trasferimento tra il riferimento w e l'uscita y si estenda almeno fino a $\omega = 0.02$:
- un disturbo sinusoidale $d(t) = \sin(0.001t)$ sull'uscita risulti attenuato di almeno 40 dB;
- un disturbo di misura $n(t) = \sin(0.8t)$ risulti attenuato di almeno 40 dB.

Si rappresenti la regione di ammissibilità in cui deve giacere il diagramma di Bode del modulo della funzione d'anello corrispondente.

Esercizio 11.2 Con riferimento al sistema di controllo di Figura 11.1, sia

$$G(s) = \frac{20(1+0.5s)}{s(1+s)^2}$$

e si debba progettare un regolatore che assicuri $|\epsilon_{oo}| < 0.2$ in corrispondenza di $w(t) = 0.1 \text{ ram}(t)$, $d(t) = \pm \text{ram}(t)$, $n(t) = \pm 0.01 \text{ sca}(t)$. Determinare il tipo e il guadagno del regolatore.

Esercizio 11.3 Facendo ancora riferimento al sistema dell'esercizio precedente, si progetti un regolatore di tipo zero che permetta di ottenere un valore arbitrariamente elevato della pulsazione critica.

Esercizio 11.4 Si consideri il sistema dell'Esempio 11.2 e si valutino le prestazioni ottenibili con un regolatore del tipo

$$R(s) = \mu_R \frac{(1+10s)(1+s)}{(1+250s)(1+0.4s)}$$

Esercizio 11.5 Sempre con riferimento all'Esempio 11.2, si valutino le prestazioni ottenibili con un regolatore del tipo

$$R(s) = \mu_R \frac{(1+10s)(1+5s)(1+s)}{s(1+0.4s)^2}$$

e si discutano eventuali limitazioni.

Esercizio 11.6 Si progetti un regolatore per il sistema con funzione di trasferimento

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

dove μ e τ sono parametri incerti che possono assumere un valore arbitrario, rispettivamente negli intervalli $10 < \mu < 20$ e $1 < \tau < 2$. Il progetto deve garantire in ogni caso l'annullamento dell'errore a transitorio esaurito in corrispondenza di riferimenti a scalino, una pulsazione critica $\omega_c > 0.1$ e un margine di fase $\varphi_m > 50^\circ$.

Esercizio 11.7 Si ripeta il progetto dell'Esempio 11.6 supponendo che nel progetto statico si decida di porre

$$|\epsilon_{wool}| \leq 0.08, \quad |\epsilon_{doo}| \leq 0.02$$

anziché equipartire l'errore tra i due contributi.

Esercizio 11.8 Si progetti una rete anticipatrice che assicuri un anticipo di fase di 70° in corrispondenza di $\omega = 1$.

Esercizio 11.9 Si supponga di dover progettare una rete ritardatrice per controllare un sistema con funzione di trasferimento

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

Determinare i parametri della rete in modo che il guadagno d'anello sia maggiore di 100, la pulsazione critica sia maggiore di 5, e il margine di fase sia maggiore di 60° .

Esercizio 11.10 Si traccino i diagrammi asintotici di Bode associati a una rete a sella in cui $\tau_1 = \tau_2$, cioè con zeri reali coincidenti.

12.1 Introduzione

Oltre ai metodi basati sulla risposta in frequenza, ampiamente discussi nei precedenti due capitoli, sono disponibili altre tecniche per l'analisi e la sintesi dei sistemi di controllo in retroazione. Tra di esse è particolarmente utile, nel caso dei sistemi SISO, quella incentrata sul cosiddetto luogo delle radici, che attraverso una sintetica rappresentazione grafica permette di ricavare informazioni sulla esatta collocazione dei poli del sistema in anello chiuso a partire dalle caratteristiche della funzione d'anello. A questo riguardo, si ricorderà che i metodi precedentemente descritti sono invece in grado di individuare solamente il valore dei poli dominanti e, per giunta, in maniera approssimata. Inoltre, la tecnica del luogo delle radici è applicabile anche nel caso di sistemi di controllo per processi instabili e, più in generale, in situazioni nelle quali non sono soddisfatte le condizioni di applicabilità del criterio di Bode. D'altra parte, poiché permette di trattare unicamente problemi nei quali la funzione di trasferimento d'anello è razionale, il metodo non è utilizzabile in presenza di ritardi di tempo.

Il capitolo presenta:

- la definizione di luogo delle radici, insieme alla sua caratterizzazione geometrica;
- alcune semplici regole per il tracciamento qualitativo del luogo;
- la discussione di una serie di esempi significativi di analisi e di sintesi.

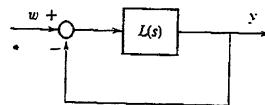
12.2 Definizione e proprietà

In questo capitolo si farà riferimento al sistema retroazionato di Figura 12.1, la cui funzione di trasferimento d'anello è espressa come

$$L(s) = \rho \frac{\prod_{i=1}^m (s + z_i)}{\prod_{i=1}^n (s + p_i)} = \rho \frac{N^*(s)}{D(s)} \quad (12.1)$$

Nella (12.1) ρ rappresenta la costante di trasferimento e gli scalari z_i e p_i (eventualmente complessi o nulli) corrispondono rispettivamente a zeri e poli di $L(s)$ cambiati di segno. Lo schema di Figura 12.1 è sufficientemente generale, anche se

Figura 12.1
Sistema retroazionato
considerato nell'analisi.



non vi compaiono i consueti ingressi relativi ai disturbi, perché l'attenzione sarà concentrata esclusivamente sui poli del sistema in anello chiuso, che coincidono comunque con le radici dell'equazione caratteristica

$$1 + L(s) = 0 \quad (12.2)$$

In particolare, il problema di cui ci si intende occupare è quello di determinare come si modifichino i poli in anello chiuso al variare della costante di trasferimento ρ , il cui valore è proporzionale al guadagno d'anello μ . L'interesse per questo problema è motivato dal fatto che la conoscenza della posizione dei poli nel piano complesso permette di ricavare utili informazioni sul comportamento dinamico del sistema retroazionato in termini di stabilità e qualità dei transitori. A quest'ultimo proposito, si ricordi che gli zeri delle funzioni di trasferimento in anello chiuso non dipendono invece da ρ e sono ricavabili in modo assai semplice (si vedano i Paragrafi 10.4.2, 10.5.2 e 10.6.2).

12.2.1 Caratterizzazione del luogo

In relazione al problema cui si è accennato, si introduca la seguente definizione.

Definizione 12.1 Per il sistema di Figura 12.1, con funzione di trasferimento d'anello (12.1), si chiama *luogo delle radici* il luogo descritto nel piano complesso dalle radici dell'equazione caratteristica (12.2) al variare del parametro reale ρ da $-\infty$ a $+\infty$, con $\rho \neq 0$. Precisamente, la parte del luogo corrispondente a $\rho > 0$ prende il nome di *luogo diretto* (LD), mentre si chiama *luogo inverso* (LI) quella corrispondente a $\rho < 0$.

Si noti che il valore $\rho = 0$, che viene escluso nella precedente definizione, corrisponde di fatto alla situazione in cui la retroazione è assente e pertanto i poli del sistema di Figura 12.1 coincidono con quelli di $L(s)$.

Allo scopo di determinare la forma del luogo delle radici, si osservi che, in base alla (12.1), l'Equazione (12.2) è equivalente a

$$\frac{N^*(s)}{D(s)} = -\frac{1}{\rho} \quad (12.3)$$

Perciò, i valori di s che soddisfano la (12.3) per qualche valore reale di ρ corrispondono a punti appartenenti al luogo. D'altra parte, la relazione complessa (12.3) equivale alle due seguenti equazioni, espresse in termini di modulo e fase:

$$\frac{|N^*(s)|}{|D(s)|} = \frac{1}{|\rho|} \quad (12.4)$$

$$\arg N^*(s) - \arg D(s) = 180^\circ - \arg \rho = \\ = \begin{cases} (2k+1)180^\circ, & \rho > 0, k \text{ intero (LD)} \\ 2k180^\circ, & \rho < 0, k \text{ intero (LI)} \end{cases} \quad (12.5)$$

Come si vedrà, la (12.5) è sufficiente a caratterizzare completamente l'aspetto geometrico del luogo, mentre la (12.4) serve a determinare la punteggiatura rispetto a ρ . In effetti si osservi che, dalla (12.1),

$$\arg N^*(s) = \arg \prod_{i=1}^m (s + z_i) = \sum_{i=1}^m \arg(s + z_i) = \sum_{i=1}^m \theta_i$$

$$\arg D(s) = \arg \prod_{i=1}^n (s + p_i) = \sum_{i=1}^n \arg(s + p_i) = \sum_{i=1}^n \varphi_i$$

dove gli angoli θ_i e φ_i hanno una diretta interpretazione geometrica. Infatti, come si vede nella Figura 12.2, dove è mostrata una generica disposizione di poli e zeri, θ_i (oppure φ_i) rappresenta l'angolo formato con il semiasse reale positivo dal vettore che congiunge lo zero $-z_i$ (oppure il polo $-p_i$) al generico punto s del piano.

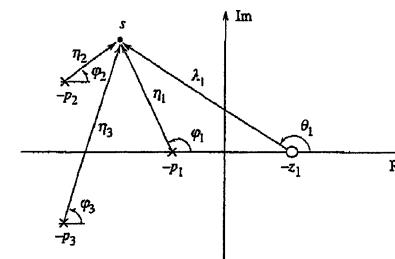


Figura 12.2
Valutazione degli angoli θ_i e φ_i e dei moduli λ_i e η_i .

In accordo con la (12.5), un punto s del piano appartiene al luogo diretto se e solo se la quantità $\sum_i \theta_i - \sum_i \varphi_i$ risulta uguale a un multiplo dispari di 180° , mentre appartiene al luogo inverso se e solo se essa è uguale a un multiplo pari di 180° . In linea di principio, esplorando punto per punto il piano complesso e selezionando i punti per i quali una di tali relazioni è verificata, si ottiene dunque il luogo cercato.

Riguardo invece alla punteggiatura del luogo, si noti che

$$|N^*(s)| = \prod_{i=1}^m |s + z_i| \quad , \quad |D(s)| = \prod_{i=1}^n |s + p_i|$$

e quindi, basandosi ancora sulla Figura 12.2, la quantità $|N^*(s)|$ (oppure $|D(s)|$) si può interpretare come il prodotto delle distanze del generico punto s dagli zeri (oppure dai poli). Perciò, una volta individuato un punto s appartenente al luogo e definite con $\lambda_i = |\bar{s} + z_i|$ e $\eta_i = |\bar{s} + p_i|$ le sue distanze da zeri e poli di $L(s)$, il corrispondente valore del modulo di ρ si ricava dalla (12.4) come

$$|\rho| = \frac{\prod_{i=1}^n |\bar{s} + p_i|}{\prod_{i=1}^m |\bar{s} + z_i|} = \frac{\prod_{i=1}^n \eta_i}{\prod_{i=1}^m \lambda_i} \quad (12.6)$$

Ovviamente, il segno di ρ dipende dal fatto che \bar{s} appartenga al luogo diretto (ρ positivo) o inverso (ρ negativo).

La procedura appena descritta verrà ora utilizzata in un semplice esempio, rimandando al seguito la formulazione di regole pratiche di tracciamento di più facile applicazione.

Esempio 12.1 Nel sistema di Figura 12.1 sia

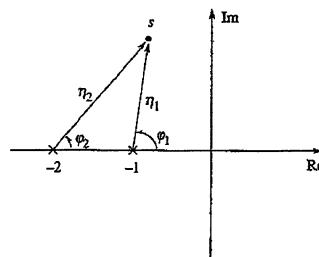
$$L(s) = \frac{\rho}{(s+1)(s+2)}$$

Con riferimento alla Figura 12.3 e applicando la (12.5), un generico punto s del piano complesso appartiene al luogo diretto se e solo se

$$-\varphi_1 - \varphi_2 = (2k+1)180^\circ$$

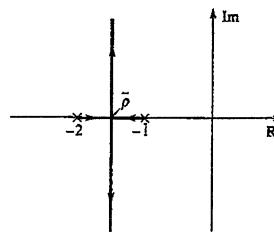
per qualche k intero. È immediato verificare che tale condizione è soddisfatta da tutti i punti del segmento $(-2, -1)$ sull'asse reale, per i quali $\varphi_1 = 180^\circ$, $\varphi_2 = 0^\circ$, nonché dai punti appartenenti all'asse dello stesso segmento, per i quali φ_1 e φ_2 sono angoli supplementari.

Figura 12.3
Valutazione di φ_i e η_i
nell'Esempio 12.1.



Il luogo diretto è allora quello rappresentato nella Figura 12.4. Utilizzando la (12.6) per ricavare la punteggiatura in ρ , è possibile evidenziare con delle frecce sul luogo il verso corrispondente a valori crescenti di ρ . Per piccoli valori di ρ le radici sono reali e prossime ai poli della funzione d'anello, mentre diventano complesse coniugate quando $\rho > \bar{\rho}$, con $\bar{\rho} = \eta_1 \eta_2 = 0.25$, come si vede dalla Figura 12.4 e dalla (12.6).

Figura 12.4
Luogo delle radici diretto
nell'Esempio 12.1.



Passando ora a considerare il luogo inverso, si osservi dalla Figura 12.3 che gli unici punti del piano per cui risulta

$$-\varphi_1 - \varphi_2 = 2k180^\circ, \quad k \text{ intero}$$

sono quelli della semiretta $(-\infty, -2)$ sull'asse reale, per i quali $\varphi_1 = \varphi_2 = 180^\circ$, e quelli della semiretta $(-1, +\infty)$, per i quali $\varphi_1 = \varphi_2 = 0^\circ$. L'aspetto risultante del luogo inverso è mostrato nella Figura 12.5, dove le frecce indicano valori crescenti di $|\rho|$. Mediante la (12.6) si può poi ricavare la punteggiatura. Per esempio, il luogo passa per l'origine quando $\rho = -\eta_1 \eta_2 = -2$.

Il lettore può verificare facilmente la correttezza dei risultati ottenuti, discutendo al variare di ρ la posizione dei poli in anello chiuso, che in questo caso si possono calcolare esplicitamente come radici di un'equazione di secondo grado e valgono

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

Dal punto di vista delle proprietà del sistema retroazionato, i grafici riportati nelle Figure 12.4 e 12.5 mostrano che il sistema di Figura 12.1 diventa instabile per valori di ρ inferiori a -2 , in quanto uno dei suoi poli si sposta nel semipiano destro, mentre rimane asintoticamente stabile per ogni valore positivo di ρ , anche se lo smorzamento dei poli si riduce al crescere di ρ .

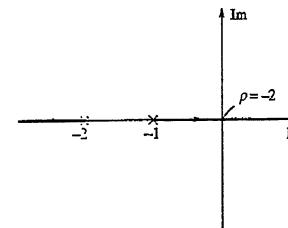


Figura 12.5
Luogo delle radici inverso
nell'Esempio 12.1.

12.2.2 Regole di tracciamento

Il tracciamento del luogo delle radici è agevole se si dispone di ordinari strumenti di calcolo e di grafica, uniti a un programma per la soluzione numerica dell'equazione complessa (12.3) o, equivalentemente, della coppia di equazioni (12.4), (12.5). Tuttavia, è di particolare interesse individuare alcune semplici regole che permettono il tracciamento qualitativo del luogo. Innanzitutto perché in questo modo si possono ricavare a prima vista le informazioni principali sul comportamento del sistema retroazionato, anche senza utilizzare particolari mezzi di calcolo. In secondo luogo, perché la conoscenza di queste regole permette più facilmente di congetturare l'effetto sulle prestazioni di un sistema retroazionato derivante da eventuali modifiche della funzione d'anello, quali l'aggiunta di un polo o di uno zero o la variazione del guadagno. In questo senso, il tracciamento approssimato del luogo può essere di grande aiuto per la fase di progetto.

Nel seguente, coerentemente con la (12.1), n e m rappresentano i gradi dei polinomi al denominatore e al numeratore di $L(s)$, e $v = n - m$ è il grado relativo. Come già ipotizzato nei capitoli precedenti, si assumerà che $L(s)$ sia strettamente propria e quindi risulti $v > 0$. Inoltre, per un dato valore di ρ , si indicherà come *baricentro* la somma dei poli in anello chiuso divisa per n .

Regola 1 Il luogo delle radici è costituito da $2n$ rami: n di questi fanno parte del luogo diretto e gli altri n del luogo inverso.

Tale affermazione deriva dalla constatazione che la (12.2) è equivalente all'equazione polinomiale di ordine n

$$D(s) + \rho N^*(s) = 0 \quad (12.7)$$

le cui radici in campo complesso sono sempre in numero pari a n e dipendono con continuità dal parametro ρ .

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

Poiché infatti l'Equazione (12.7) ha coefficienti reali, le sue radici sono reali oppure complesse coniugate.

Regola 3 I rami "partono" dai poli di $L(s)$.

Infatti, per $|\rho| \rightarrow 0$ le radici di (12.7) convergono verso i poli della funzione d'anello.

Regola 4 Sia nel luogo diretto sia nel luogo inverso, per $|\rho| \rightarrow \infty$, m rami "arrivano" negli zeri di $\tilde{L}(s)$ e i restanti v rami tendono all'infinito.

La giustificazione di questo risultato si basa sull'osservazione che, quando $|\rho| \rightarrow \infty$, l'Equazione (12.7) degenera nell'equazione di ordine m

$$N^*(s) = 0$$

che ha solo m radici, coincidenti con gli zeri di $L(s)$. Per $|\rho| \rightarrow \infty$, queste sono allora le uniche radici al finito dell'Equazione (12.7).

Regola 5 I rami che tendono all'infinito lo fanno lungo asintoti che si intersecano sull'asse reale nel punto di ascissa

$$x_a = \frac{1}{v} \left(\sum_{i=1}^m z_i - \sum_{i=1}^n p_i \right)$$

e formano con l'asse reale angoli pari a

$$\psi_{ak} = \begin{cases} \frac{(2k+1)180^\circ}{v}, & k = 0, 1, \dots, v-1, \rho > 0 \text{ (LD)} \\ \frac{2k180^\circ}{v}, & k = 0, 1, \dots, v-1, \rho < 0 \text{ (LI)} \end{cases}$$

La dimostrazione di questo risultato, qui non riportata per motivi di semplicità, passa attraverso l'approssimazione $\tilde{L}(s) = \rho/(s - x_a)^v$ della funzione d'anello, valida per valori di $|\rho|$ e di $|s|$ sufficientemente grandi.

Regola 6 Tutti i punti dell'asse reale, tranne quelli corrispondenti a singolarità di $L(s)$, appartengono al luogo delle radici. Precisamente, fanno parte del luogo diretto tutti i punti a sinistra di un numero dispari di singolarità di $L(s)$; fanno parte invece del luogo inverso tutti i punti a sinistra di un numero pari di singolarità di $L(s)$.

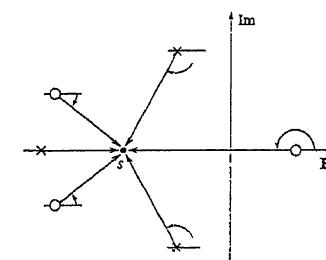
La dimostrazione di questa proprietà è immediata se si ricorda la condizione (12.5) e si osserva la Figura 12.6. Preso un punto generico sull'asse reale, ogni singolarità reale alla sua sinistra e ogni coppia di singolarità complesse coniugate produce un contributo ad $\arg N^*(s) - \arg D(s)$ che è nullo o pari a $\pm 360^\circ$. Viceversa, ogni singolarità reale alla sua destra dà un contributo di $\pm 180^\circ$. Affinché la somma dei vari termini sia pari a un multiplo dispari (oppure pari) di 180° , il punto deve essere quindi a sinistra di un numero dispari (oppure pari) di singolarità.

Regola 7 Quando $v \geq 2$, il baricentro del luogo non dipende da ρ e coincide con il punto dell'asse reale di ascissa

$$x_b = -\frac{1}{n} \sum_{i=1}^n p_i$$

Per giustificare questo risultato si osservi che, in un polinomio di grado n , il rapporto tra il coefficiente del termine di grado $n-1$ e quello di grado n , cambiato di

Figura 12.6
Derivazione geometrica
della Regola 6.



segno, è sempre uguale alla somma delle radici (e quindi alla somma delle parti reali quando il polinomio è a coefficienti reali). D'altra parte, quando $v \geq 2$, i primi due coefficienti del polinomio al primo membro della (12.7) non dipendono da ρ , e quindi rimane costante al variare di ρ la somma delle radici, che coincide per definizione con il baricentro moltiplicato per n . Poiché per $\rho = 0$ il baricentro coincide con il punto di ascissa x_b , la proprietà è dimostrata.

Regola 8 Si consideri il polo $-p_j$ di $L(s)$ e si supponga che abbia molteplicità h_j . Gli h_j rami del luogo diretto e gli h_j rami del luogo inverso che partono da questo polo hanno in quel punto tangenti che formano con l'asse reale angoli uguali a

$$\alpha_{jk} = \begin{cases} \frac{1}{h_j} \left[(2k+1)180^\circ + \sum_{i=1}^m \theta_i - \sum_{i=j}^{i=h_j} \varphi_i \right], & k = 0, 1, \dots, h_j-1, \rho > 0 \text{ (LD)} \\ \frac{1}{h_j} \left[2k180^\circ + \sum_{i=1}^m \theta_i - \sum_{i=j}^{i=h_j} \varphi_i \right], & k = 0, 1, \dots, h_j-1, \rho < 0 \text{ (LI)} \end{cases}$$

dove gli angoli φ_i e θ_i sono quelli calcolati nel modo consueto, considerando i vettori che congiungono il punto $s = -p_j$ agli altri poli e agli zeri (Figura 12.7). In particolare, se il polo $-p_j$ è semplice, la tangente del ramo uscente presenta un angolo

$$\alpha_j = \begin{cases} 180^\circ + \sum_{i=1}^m \theta_i - \sum_{i=j}^{i=h_j} \varphi_i, & \rho > 0 \text{ (LD)} \\ \sum_{i=1}^m \theta_i - \sum_{i=j}^{i=h_j} \varphi_i, & \rho < 0 \text{ (LI)} \end{cases}$$

Per verificare la correttezza di questa regola basta applicare la condizione (12.5) nell'intorno del punto $-p_j$.

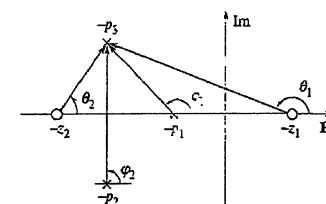


Figura 12.7
Valutazione degli angoli θ_i e φ_i per l'applicazione
della Regola 8.

Regola 9 Si consideri lo zero $-z_j$ di $L(s)$ e si supponga che abbia molteplicità h_j . Gli h_j rami del luogo diretto e gli h_j rami del luogo inverso che arrivano in questo zero hanno in quel punto tangentini che formano con l'asse reale angoli uguali a

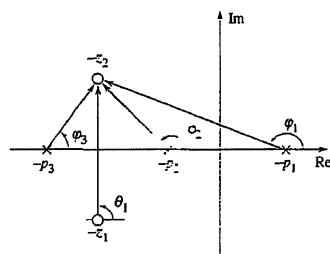
$$\beta_{jk} = \begin{cases} \frac{1}{h_j} \left[(2k+1)180^\circ - \sum_{i \neq j} \theta_i + \sum_{i=1}^n \varphi_i \right], & k = 0, 1, \dots, h_j - 1, \rho > 0 \\ \frac{1}{h_j} \left[2k180^\circ - \sum_{i \neq j} \theta_i + \sum_{i=1}^n \varphi_i \right], & k = 0, 1, \dots, h_j - 1, \rho < 0 \end{cases} \quad (\text{LD})$$

dove gli angoli θ_i e φ_i sono quelli calcolati nel modo consueto, considerando i vettori che congiungono il punto $s = -z_j$ agli altri zeri e ai poli (Figura 12.8). In particolare, se lo zero $-z_j$ è semplice, la tangente del ramo entrante presenta un angolo

$$\beta_j = \begin{cases} 180^\circ - \sum_{i \neq j} \theta_i + \sum_{i=1}^n \varphi_i, & \rho > 0 \text{ (LD)} \\ -\sum_{i \neq j} \theta_i + \sum_{i=1}^n \varphi_i, & \rho < 0 \text{ (LI)} \end{cases}$$

Questa proprietà si dimostra in maniera analoga alla precedente.

Figura 12.8
Valutazione degli angoli θ_i e φ_i per l'applicazione della Regola 9.



Regola 10 Eventuali punti di incrocio di rami sull'asse reale si possono determinare trovando i minimi e massimi relativi della funzione

$$\gamma(x) = -\frac{D(x)}{N^*(x)}$$

con x reale. Precisamente, se \bar{x} è un punto di minimo e il punto $s = \bar{x}$ appartiene al luogo diretto, esistono due rami complessi che confluiscono sull'asse reale in \bar{x} ; se \bar{x} è invece un punto di massimo e il punto $s = \bar{x}$ appartiene al luogo diretto, esistono due rami reali che si incontrano in \bar{x} e poi si separano diventando complessi (Figura 12.9). La situazione è rovesciata se si considera il luogo inverso.

La dimostrazione è semplice se si osserva che la funzione $\gamma(x)$ coincide con $\rho(x)$ quando x si muove sull'asse reale (si veda la (12.3)).

Regola 11 In ogni punto del luogo il valore di $|\rho|$ è dato da

$$|\rho| = \frac{\prod_{i=1}^n \eta_i}{\prod_{i=1}^m \lambda_i}$$

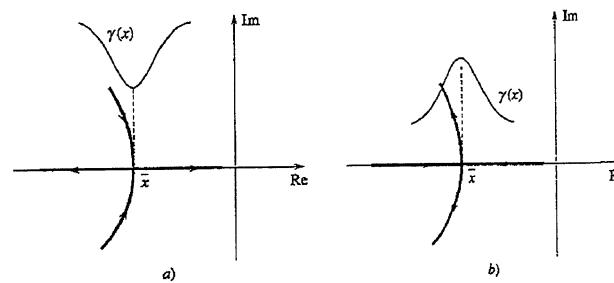


Figura 12.9
Illustrazione della Regola 10 per il luogo diretto:
a) punto di confluenza;
b) punto di diramazione.

dove λ_i e η_i rappresentano le distanze del punto dagli zeri e dai poli di $L(s)$. Questa formula coincide con la (12.6).

12.3 Uso del luogo delle radici nell'analisi

In questo paragrafo si illustrerà, mediante un certo numero di esempi significativi, l'uso dei luoghi delle radici per valutare alcune caratteristiche importanti di un sistema retroazionato.

Esempio 12.2 Con riferimento allo schema di Figura 12.1, sia

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

e si desideri valutare la posizione dei poli in anello chiuso al variare di ρ .

Utilizzando le regole del Paragrafo 12.2.2, si deduce che il luogo diretto è costituito da tre rami, che partono dai poli di $L(s)$ e tendono tutti all'infinito lungo tre asintoti che hanno inclinazioni di 60° , 180° e 300° e si incontrano sull'asse reale nel punto di ascissa $x_a = -8/3$ (Figura 12.10). Inoltre, in base alla Regola 6, fanno parte del luogo diretto i punti del segmento $(-3, 0)$ e quelli della semiretta $(-\infty, -5)$. Quindi, i due rami che partono dai poli in $s = -3$ e in $s = 0$ devono incontrarsi in un punto intermedio, di cui è possibile calcolare l'ascissa mediante la Regola 10. Infatti è facile verificare che la funzione $\gamma(x) = -x(x+3)(x+5)$ presenta un massimo in $\bar{x} = (-8 + \sqrt{19})/3 \approx -1.21$. Questa è dunque l'ascissa del punto Q_1 nella Figura 12.10 in cui i due rami si distaccano dall'asse. Si noti che, a partire da questo punto, i rami devono spostarsi verso

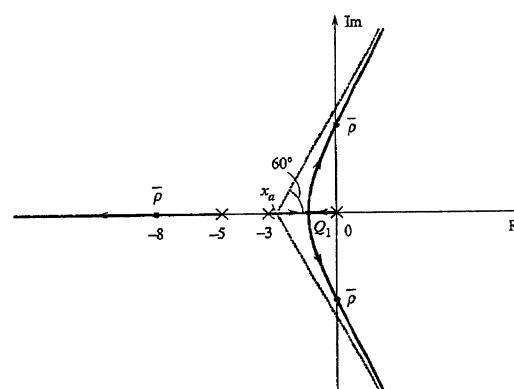


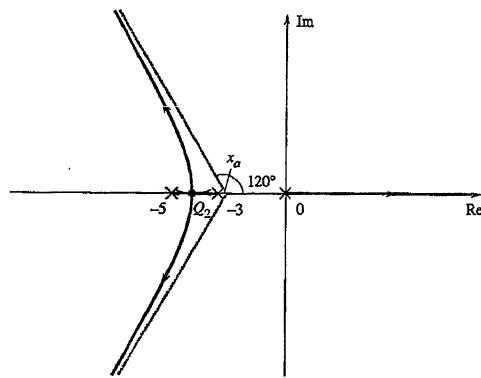
Figura 12.10
Luogo diretto nell'Esempio 12.2.

destra perché il baricentro del luogo rimanga fisso nel punto di ascissa $x_b = -8/3$, come deve essere per la Regola 7, visto che $\nu = 3$.

Dall'andamento del luogo si deduce che esiste un valore positivo di ρ oltre il quale il sistema in anello chiuso è instabile. Naturalmente, tale estremo corrisponde al valore $\bar{\rho}$ per il quale i due rami complessi attraversano l'asse immaginario. La valutazione precisa di $\bar{\rho}$ si può in questo caso ricavare osservando che, grazie alla proprietà di invarianza del baricentro, il terzo polo del sistema in anello chiuso per $\rho = \bar{\rho}$ è collocato nel punto di ascissa $3x_b = -8$. Applicando allora la regola della punteggiatura in quest'ultimo punto, si ottiene $\bar{\rho} = 120$.

Nella Figura 12.11 è riportato il grafico del luogo inverso, il cui tracciamento qualitativo è ricavabile anch'esso a partire dalle regole del Paragrafo 12.2.2. In particolare, si noti che il punto Q_2 di diramazione ha ascissa $(-8 - \sqrt{19})/3 \approx -4.12$, valore in corrispondenza del quale $\gamma(x)$ presenta un punto di minimo, e che uno dei rami si sviluppa interamente nel semipiano destro, indicando così che il sistema non può essere asintoticamente stabile per valori negativi di ρ .

Figura 12.11
Luogo inverso
nell'Esempio 12.2.

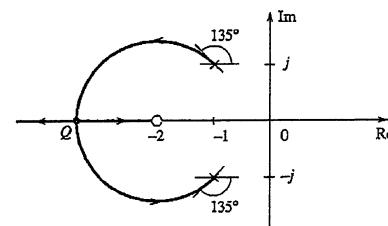


Esempio 12.3 Si vogliono studiare le caratteristiche del sistema di Figura 12.1 quando la funzione d'anello è

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

che possiede uno zero in -2 e due poli complessi in $-1 \pm j$. Considerando dapprima il luogo diretto, si noti che, in base alla Regola 6, esso deve includere la semiretta $(-\infty, -2)$ sull'asse reale. Si può quindi concludere che i due rami uscenti dai poli devono incontrarsi sull'asse reale alla sinistra di -2 , per poi tendere l'uno allo zero e l'altro verso l'infinito. Le direzioni di uscita dei rami dai poli sono valutabili mediante la Regola 8: per il polo a parte immaginaria positiva risulta $\alpha_1 = 180^\circ + 45^\circ - 90^\circ = 135^\circ$ e α_2 si ricava per simmetria. L'aspetto del luogo diretto è perciò quello mostrato nella Figura 12.12. Volendo, è anche possibile determinare con la Regola 10 l'ascissa del punto Q di intersezione dei rami.

Dalla Figura 12.12 si possono trarre utili indicazioni sulle proprietà del sistema retroazionato per valori positivi di ρ , che in questo caso coincide con il guadagno d'anello μ . Innanzitutto, il sistema è asintoticamente stabile $\forall \mu > 0$, come si sarebbe anche potuto inferire dall'applicazione del Corollario 9.4, visto che risulta $|\arg L(j\omega)| < 180^\circ, \forall \omega$. Inoltre, appare evidente che al crescere di μ lo smorzamento dei poli in anello chiuso tende ad aumentare fino a che si ottengono poli reali per μ sufficientemente elevato. Anche questo risultato è coerente con la teoria precedentemente svolta, dato che per μ crescente il margine di fase, come è facile verificare, tende ad avvicinarsi a 90° . Poiché infine insieme a μ cresce anche la pulsazione critica ω_c , ci si dovrebbe aspettare, in base alle argomentazioni del Paragrafo 10.3.3, che il polo dominante del sistema retroazionato si sposti sempre più verso sinistra. In effetti, esaminando la Figura 12.12 si scopre che, per valori elevati del guadagno d'anello, uno dei poli diventa sempre più "veloce", mentre l'altro tende allo zero di $L(s)$. Pertanto, almeno nei riguardi della funzione di sensitività complementare $F(s) = L(s)/(1+L(s))$, i cui zeri, come è noto, coincidono con quelli di $L(s)$, il polo "lento" risulta quasi cancellato dallo zero e ha scarsa influenza sul comportamento dinamico del sistema, se non per la generazione dei fenomeni di lenta deriva nella risposta allo scalino di cui si è discusso nel Paragrafo 4.4. Si ponga



attenzione al fatto che questa conclusione non è invece applicabile ad altre funzioni di trasferimento in anello chiuso, come per esempio la funzione di sensitività $S(s) = 1/(1+L(s))$.

Per completare la discussione dell'esempio, è riportato nella Figura 12.13 l'andamento del luogo inverso, corrispondente a valori negativi del guadagno d'anello. La caratteristica saliente di tale grafico è che uno dei rami penetra nel semipiano destro per valori di $|\rho|$ sufficientemente elevati, indicando la perdita della proprietà di stabilità. Ciò avviene per $\rho < \bar{\rho}$, con

$$\bar{\rho} = -\frac{\eta_1 \eta_2}{\lambda_1} = -\frac{\sqrt{2}\sqrt{2}}{2} = -1$$

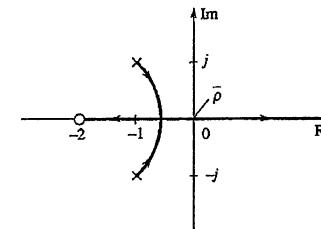


Figura 12.13
Luogo inverso
nell'Esempio 12.3.

Esempio 12.4 Il luogo delle radici diretto associato alla funzione d'anello

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

è riportato nella Figura 12.14. Dei quattro rami del luogo che escono dal punto -2 , due convergono verso la coppia di zeri in -1 e due tendono all'infinito lungo l'asse reale. In accordo con la Regola 8, gli angoli di partenza dei rami uscenti dal polo di molteplicità $h = 4$ sono uguali ad $\alpha_k = 45^\circ + k90^\circ$, $k = 0, 1, 2, 3$, mentre, in accordo con la Regola 9, gli angoli di arrivo dei due rami nello zero doppio valgono $\beta_k = 90^\circ + k180^\circ$, $k = 0, 1$. Il sistema in anello chiuso risulta asintoticamente stabile per tutti i valori positivi di ρ .

Il grafico del luogo inverso è invece tracciato nella Figura 12.15. In questo caso i rami escono dai poli in $s = -2$ con tangenti parallele agli assi. È facile verificare mediante l'applicazione della Regola 10 che i due rami complessi tornano a intersecarsi sull'asse reale in corrispondenza

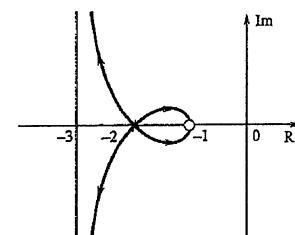
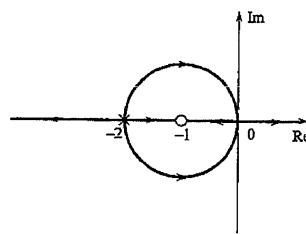


Figura 12.14
Luogo diretto
nell'Esempio 12.4.

dell'origine, dove risulta $|\rho| = 16$. Quindi, il sistema retroazionato è asintoticamente stabile per $\rho > -16$ e diventa instabile quando $\rho < -16$, perché uno dei rami del luogo si sposta nel semipiano destro.

Figura 12.15
Luogo inverso
nell'Esempio 12.4.



Benché il metodo del luogo delle radici sia stato originariamente ideato per studiare il comportamento di un sistema retroazionato al variare della costante di trasferimento d'anello ρ , esso può essere adattato con poca difficoltà anche al caso in cui interessi discutere la posizione dei poli in anello chiuso in funzione di un generico parametro reale che compare nella funzione d'anello $L(s)$. L'esempio seguente mira a illustrare questo tipo di estensione.

Esempio 12.5 Nel sistema retroazionato di Figura 12.1 sia

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

e si voglia determinare il luogo dei poli in anello chiuso al variare dello smorzamento ξ da zero a infinito. L'equazione caratteristica del sistema retroazionato è

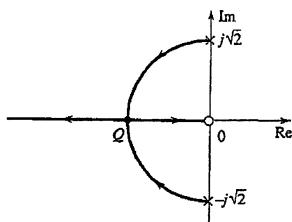
$$s^2 + 2\xi s + 2 = 0$$

che, definendo $\tilde{N}^*(s) = s$, $\tilde{D}(s) = s^2 + 2\xi s + 2$, può essere anche riscritta come

$$\frac{\tilde{N}^*(s)}{\tilde{D}(s)} = -\frac{1}{\tilde{\rho}} \quad (12.8)$$

Poiché quest'ultima equazione è formalmente identica alla (12.3), il luogo cercato può essere costruito applicando alla (12.8) la consueta tecnica di analisi del luogo delle radici. In particolare, interessa tracciare il luogo diretto, che, come mostrato nella Figura 12.16, è costituito da due rami che partono dalle radici di $\tilde{D}(s)$, si incrociano sull'asse reale nel punto Q e poi divergono, l'uno tendendo all'unica radice di $\tilde{N}^*(s)$ e l'altro verso $-\infty$. È facile verificare tramite l'applicazione delle Regole 10 e 11 che il punto Q ha ascissa $-\sqrt{2}$ e che il valore di $\tilde{\rho}$ associato a tale punto è $\tilde{\rho}_1 = 2\sqrt{2}$. Corrispondentemente, $\xi_1 = \sqrt{2}$ e si conclude che i poli del sistema retroazionato sono complessi coniugati per $0 < \xi < \xi_1$ e diventano reali per $\xi \geq \xi_1$. In ogni caso l'asintotica stabilità è garantita per ogni valore di ξ strettamente positivo.

Figura 12.16
Luogo diretto
nell'Esempio 12.5.



12.4 Uso del luogo delle radici nella sintesi

Come altri strumenti per l'analisi dei sistemi retroazionati, il metodo del luogo delle radici può essere utilmente impiegato per il progetto del regolatore, seguendo l'approccio della sintesi per tentativi. In particolare, il metodo ben si presta a trattare problemi nei quali gli obiettivi del controllo sono direttamente esprimibili in termini di posizione nel piano complesso dei poli in anello chiuso, come per esempio succede quando l'unico requisito riguarda la stabilità asintotica del sistema di controllo o quando si vogliono prescrivere le caratteristiche dei poli dominanti. Per esempio:

- vincoli del tipo $\xi \geq \tilde{\xi}$ riguardanti lo smorzamento dei poli dominanti si traducono nell'imporre che i poli in anello chiuso siano confinati nel settore del piano complesso indicato nella Figura 12.17a;
- la condizione $\omega_n \geq \tilde{\omega}_n$ sulla pulsazione naturale dei poli dominanti corrisponde invece a richiedere che essi giacciono nella regione di Figura 12.17b, esterna alla circonferenza di raggio $\tilde{\omega}_n$ e centro nell'origine;
- un'eventuale specifica sul massimo tempo di assolvimento si può esprimere come un vincolo sulla parte reale $-\sigma$ dei poli in anello chiuso, che definisce il semipiano $-\sigma \leq -\tilde{\sigma}$ di Figura 12.17c.

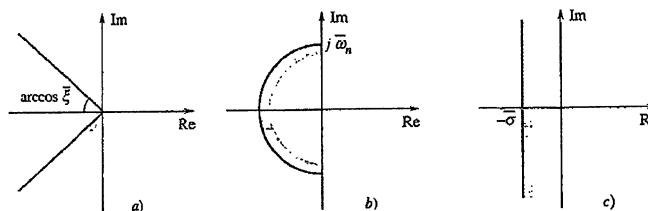


Figura 12.17
Esempi di vincoli di progetto sulla posizione dei poli in anello chiuso.

Naturalmente, la formulazione di problemi in cui si richiede il soddisfacimento contemporaneo di condizioni di questo genere porta a costruire regioni ammissibili nel piano complesso di forma più articolata, ottenute dall'intersezione dei vari vincoli.

Seguono ora alcuni esempi che illustrano l'impiego del luogo delle radici in problemi di progetto.

Esempio 12.6 Con riferimento al sistema di controllo di Figura 12.18, in cui

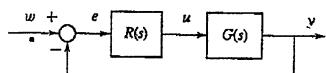
$$G(s) = \frac{s+3}{s(s+1)(s+10)}$$

si debba progettare un regolatore puramente proporzionale $R(s) = \mu_R$ in modo che il tempo di assolvimento della risposta a uno scalino del riferimento sia minore o uguale a $T_{d1} = 2.5$.

Poiché una specifica sul tempo di assolvimento può essere facilmente interpretata come una prescrizione sulla posizione dei poli dominanti in anello chiuso, il metodo del luogo delle radici si rivela adatto alla soluzione di questo problema di sintesi. Poiché $L(s) = \mu_R G(s)$, il guadagno μ_R coincide con la costante di trasferimento d'anello ρ .

Ricordando che in un sistema del secondo ordine con poli complessi in $-\sigma \pm j\omega$ risulta $T_{d1} \approx 5/\sigma$ (si veda la Tabella 4.4), il requisito sul tempo di assolvimento si traduce nell'imporre che i poli dominanti siano collocati a sinistra della retta verticale con ascissa $-\sigma \approx -5/T_{d1} = -2$. Si tracci allora il grafico del luogo diretto associato a $L(s)$ (Figura 12.19) e si osservi che esiste in effetti la possibilità di determinare μ_R in modo che il sistema retroazionato abbia due poli sulla retta prestabilita. Infatti, i due rami complessi del luogo tendono all'infinito lungo un asintoto verticale di ascissa $x_a = (3 - 10 - 1)/2 = -4$ e devono pertanto necessariamente intersecare la retta di ascissa

Figura 12.18
Sistema di controllo
considerato nella sintesi.

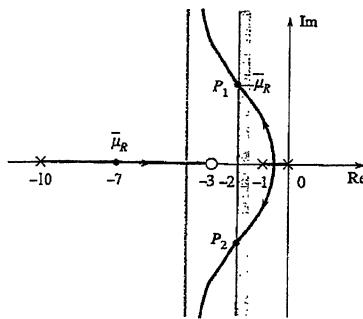


$-\bar{\sigma}$ nei punti P_1 e P_2 . Per calcolare il valore di μ_R in corrispondenza di questi punti è conveniente ragionare come nell'Esempio 12.2, osservando che il baricentro del luogo è fisso, perché $v = 2$, e si trova nel punto di ascissa $x_b = -11/3$. Perciò, quando la parte reale dei due poli complessi vale -2 , il terzo polo è collocato in $s = -7$. Grazie alla formula per la punteggiatura del luogo, il valore di μ_R in quest'ultimo punto è dato da

$$\bar{\mu}_R = \frac{7 \cdot 6 \cdot 3}{4} = 31.5$$

e corrisponde al valore da attribuire al guadagno del regolatore per soddisfare la specifica.

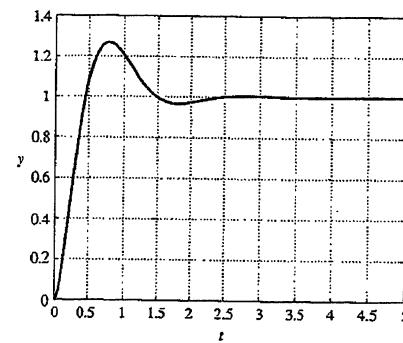
Figura 12.19
Luogo diretto
nell'Esempio 12.6.



Si lascia al lettore verificare che l'ordinata del punto P_1 è uguale a $\sqrt{9.5} \approx 3.08$ (lo si può fare per via geometrica, imponendo il rispetto della condizione (12.5), oppure per via algebrica a partire dall'equazione caratteristica) e che quindi i poli dominanti del sistema di controllo progettato hanno una pulsazione naturale $\omega_n \approx 3.67$ e uno smorzamento pari a $\xi \approx \arccos(3.08/2) \approx 0.54$.

La risposta dell'uscita a uno scalino del riferimento in corrispondenza del regolatore progettato (Figura 12.20) dimostra che il vincolo di progetto sul tempo di assestamento viene in effetti rispettato.

Figura 12.20
Risposta di y a uno scalino
del riferimento w per
il sistema di controllo
progettato
nell'Esempio 12.6.



Esempio 12.7 Si consideri il sistema di controllo di Figura 12.18, dove si utilizza un regolatore puramente proporzionale $R(s) = \mu_R$ per stabilizzare il sistema descritto dalla funzione di trasferimento

$$G(s) = \frac{260(s+1)}{(s-1)(s^2 + 10s + 26)}$$

Si vogliono determinare i valori di μ_R per cui il sistema risulta asintoticamente stabile.

I poli della funzione d'anello $L(s) = \mu_R G(s)$ valgono 1 e $-5 \pm j$. Lo zero vale -1 e la costante di trasferimento è $\rho = 260\mu_R$. Osservando che uno dei rami del luogo inverso coincide con la semiretta $(1, +\infty)$, si può immediatamente concludere che non è possibile rendere stabile il sistema con valori negativi di μ_R . Si consideri allora solo il luogo diretto, rappresentato nella Figura 12.21. Mentre uno dei rami parte dal polo reale e converge verso lo zero, gli altri due rami che originano dai poli complessi tendono all'infinito lungo una retta verticale di ascissa $x_a = -4$ (si veda la Regola 5). Il sistema in anello chiuso è allora asintoticamente stabile quando $\rho > \bar{\rho}$, dove $\bar{\rho} = 26$ è il valore corrispondente all'origine del piano complesso. In conclusione, il guadagno μ_R del regolatore deve essere maggiore di $\bar{\mu}_R = \bar{\rho}/260 = 0.1$.

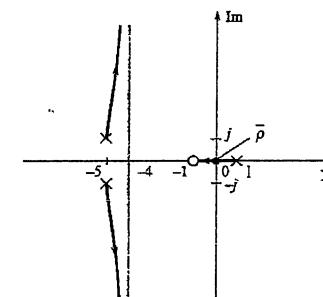


Figura 12.21
Luogo diretto
nell'Esempio 12.7.

Esempio 12.8 Si intenda progettare un regolatore $R(s)$ per il sistema di Figura 12.18, dove

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

cercando di soddisfare i seguenti requisiti:

- risulti $|e_\infty| < 0.1$ quando $w(t) = \text{ram}(t)$;
- il sistema retroazionato presenti due poli dominanti complessi con pulsazione naturale $\omega_n = 2$ e smorzamento $\xi \geq 0.5$.

Si noti innanzitutto che la specifica (a) impone, a meno di impiegare un regolatore dotato di azione integrale, che il guadagno d'anello sia superiore a 10 (si veda la Tabella 10.1) e che di conseguenza il guadagno μ_R del regolatore sia maggiore di 1 .

La sintesi del regolatore verrà ora condotta per tentativi successivi, utilizzando a ogni passo il metodo del luogo delle radici per verificare il rispetto delle specifiche. Si consideri dapprima il regolatore $R_0(s) = \mu_R$ ad azione puramente proporzionale, e l'associata funzione d'anello

$$L_0(s) = \frac{\rho}{s(s+1)} \quad , \quad \rho = 10\mu_R$$

Dopo aver osservato che per l'asintotica stabilità deve essere comunque $\mu_R > 0$, si esamini il luogo diretto relativo a $L_0(s)$, riportato nella Figura 12.22. Per ottenere $\xi \geq 0.5$, occorre che i poli in anello chiuso cadano all'interno del settore indicato. Quindi, il massimo valore di ω_n ottenibile con questo genere di regolatore è quello corrispondente al punto P nella figura, cui è associato il valore $\bar{\rho}$ della costante di trasferimento. Poiché si trova $\omega_n = 1$ e $\bar{\rho} = 1$, da cui $\bar{\mu}_R = 0.1$, si deduce che nessuna delle condizioni (a) e (b) è soddisfatta.

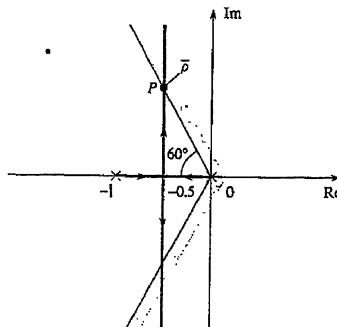
Si provi allora a considerare un regolatore costituito da una rete anticipatrice (si veda il paragrafo 11.5.1), che consenta di ampliare la banda passante del sistema retroazionato (cioè di aumentare ω_n) a parità di margine di fase (e quindi di smorzamento ξ):

$$R_1(s) = \mu_R \frac{1+Ts}{1+\alpha Ts} \quad , \quad \alpha < 1$$

La nuova funzione d'anello è

$$L_1(s) = \rho \frac{s + \frac{1}{T}}{s(s+1)(s + \frac{1}{\alpha T})} \quad , \quad \rho = \frac{10\mu_R}{\alpha}$$

Figura 12.22
Luogo diretto associato a $L_0(s)$ nell'Esempio 12.8.

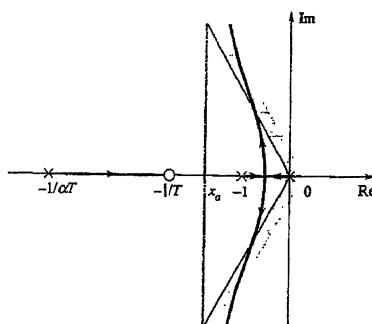


e il corrispondente luogo delle radici dirette è mostrato nella Figura 12.23, nell'ipotesi che si scelga $T < 1$. Da questo grafico si nota che l'effetto benefico della rete anticipatrice riguarda la possibilità di modificare la posizione dell'asintoto verticale, spostandolo verso sinistra i rami associati alle radici complesse. In questo modo una più ampia parte del luogo rimane confinata all'interno della regione ammissibile, corrispondente al vincolo sullo smorzamento. In effetti, l'ascissa dell'asintoto vale

$$x_a = \frac{1}{2} \left(\frac{1}{T} - 1 - \frac{1}{\alpha T} \right) = -\frac{1}{2} + \frac{1}{2T} \left(1 - \frac{1}{\alpha} \right) < -\frac{1}{2}$$

e può essere resa arbitrariamente negativa pur di diminuire α . Scegliendo per esempio $T = 0.5$ e $2/5 < \alpha < 2/3$, si ottiene il luogo mostrato nella Figura 12.24, con x_a compreso tra -2 e -1 .

Figura 12.23
Luogo diretto associato a $L_1(s)$ con $T < 1$ e $0 < \alpha < 1$ nell'Esempio 12.8.



Restano ora da determinare il valore preciso di α e il guadagno μ_R del regolatore. A tale scopo si imponga che il sistema in anello chiuso abbia due poli in $s = -1 \pm j\sqrt{3}$, corrispondenti ai punti P_1 e P_2 nella Figura 12.24, in modo da rispettare le specifiche su ω_n e ξ . Perché il punto P_1 appartenga al luogo deve essere

$$\theta_1 - \varphi_1 - \varphi_2 - \varphi_3 = (2k+1)180^\circ$$

per qualche k intero. Ma poiché con semplici argomentazioni geometriche risulta $\theta_1 = 60^\circ$, $\varphi_1 = 120^\circ$ e $\varphi_2 = 90^\circ$, tale condizione richiede che sia $\varphi_3 = 30^\circ$ e di conseguenza $-2/\alpha = -4$. Il valore cercato è allora $\alpha = 0.5$. Inoltre, in corrispondenza di P_1 è

$$\tilde{\rho} = \frac{\eta_1 \eta_2 \eta_3}{\lambda_1} = \frac{2 \cdot \sqrt{3} \cdot 2\sqrt{3}}{2} = 6$$

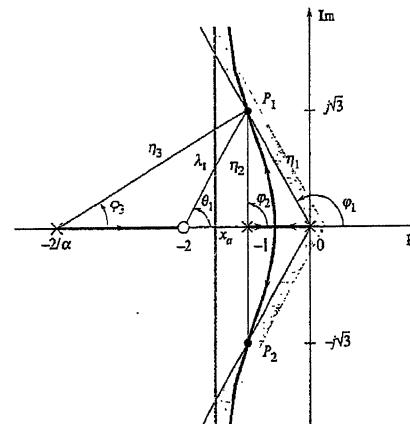
da cui $\mu_R = \alpha \tilde{\rho}/10 = 0.3$.

In definitiva, il regolatore progettato è descritto da

$$R_1(s) = 0.3 \frac{1+0.5s}{1+0.25s} \quad (12.9)$$

e soddisfa per costruzione la condizione (b). La specifica (a) sulla precisione statica non è invece rispettata perché è $\mu_R < 1$ e un ulteriore aumento del guadagno porterebbe i poli al di fuori della regione ammissibile.

Figura 12.24
Luogo diretto associato a $L_1(s)$ con $T = 0.5$ e $2/5 < \alpha < 2/3$ nell'Esempio 12.8.



Si potrebbe ovviare al problema decidendo di ricalibrare i parametri della rete anticipatrice, per esempio diminuendo α , a scapito però della moderazione dell'azione di controllo (si veda il Paragrafo 11.5.1). Un'altra soluzione praticabile è quella di introdurre nel regolatore anche una rete ritardatrice, che permetta di migliorare la precisione statica senza alterare troppo le caratteristiche dinamiche del sistema di controllo fino a questo punto ottenute. Sia allora

$$L_1(s) = \tilde{\rho} \frac{s+2}{s(s+1)(s+4)}$$

la funzione d'anello connessa all'impiego del regolatore (12.9), e si supponga di aggiungere la rete ritardatrice

$$R'(s) = \rho' \frac{s + \frac{1}{T'}}{s + \frac{1}{\alpha' T'}} \quad , \quad \alpha' > 1 \quad (12.10)$$

Scegliendo T' sufficientemente grande, il luogo delle radici associato alla funzione d'anello modificata $L_2(s) = R'(s)L_1(s)$ non presenta sensibili differenze rispetto a quello considerato in precedenza, se non in prossimità dell'origine dove appare una nuova coppia polo/zero. Ponendo per esempio $T' = 10$, il luogo si modifica come nella Figura 12.25, e la parte del luogo "lontana" dall'origine non risente in pratica dell'effetto dell'aggiunta della rete ritardatrice, né nella forma geometrica, né riguardo alla punteggiatura in ρ . Infatti, a distanza sufficientemente grande dall'origine, la coppia polo/zero di $R'(s)$ produce contributi nelle formule (12.4) e (12.5) che approssimativamente si elidono.

Perciò, con $\rho' = 1$ nella (12.10), il sistema in anello chiuso possiede ancora due poli complessi coniugati, all'incirca collocati in $-1 \pm j\sqrt{3}$ quando la costante di trasferimento di $L_2(s)$ è $\rho = \tilde{\rho} = 6$. Grazie però alla rete ritardatrice, il guadagno d'anello è adesso $\mu = \alpha' \tilde{\rho}/2 = 3\alpha'$. Scegliendo

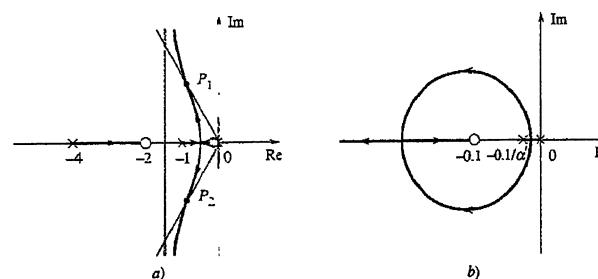


Figura 12.25
a) Luogo diretto associato a $L_2(s)$ con $T' = 10$ e $\alpha' > 1$ nell'Esempio 12.8;
b) andamento del luogo in vicinanza dell'origine.

allora per esempio $\alpha' = 4$, si è in grado di garantire il rispetto della condizione (a). Il regolatore risultante è descritto da

$$R_2(s) = R_1(s)R'(s) = \frac{(1+10s)(1+0.5s)}{(1+40s)(1+0.25s)}$$

e ha la struttura di una rete a sella (si veda il Paragrafo 11.5.3).

Si noti che, in realtà, l'aggiunta della rete ritardatrice fa sì che il sistema retroazionato abbia un quarto polo collocato lungo il ramo del luogo che converge nello zero in -0.1 (Figura 12.25b). Pur essendo questo polo associato a una dinamica "tenta", il suo effetto sul legame tra w e y è trascurabile a causa della quasi perfetta cancellazione con lo stesso zero. È dunque fatto concludere che, almeno nei riguardi della funzione di sensitività complementare, i poli dominanti sono effettivamente quelli corrispondenti ai punti P_1 e P_2 .

Esempio 12.9 (Seguito dell'Esempio 11.3) Si voglia utilizzare il metodo del luogo delle radici per rivisitare il problema di sintesi già discusso nell'Esempio 11.3, relativo al progetto di un controllore per il sistema con funzione di trasferimento

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

Si noti innanzitutto che il requisito originario sul margine di fase ($\varphi_m \geq 40^\circ$) può essere interpretato come un vincolo sullo smorzamento minimo dei poli in anello chiuso. In prima approssimazione è equivalente a impostare $\xi \geq 0.4$ (si veda la formula (10.15)). Inoltre, poiché la pulsazione critica ω_c è un'approssimazione della pulsazione naturale dei poli dominanti, l'esigenza di massimizzare ω_c si traduce nel richiedere che i poli dominanti siano alla massima distanza possibile dall'origine. Con queste premesse, procedendo in parallelo con quanto svolto nell'Esempio 11.3, si analizzino dapprima le prestazioni che può offrire un regolatore proporzionale $R(s) = \mu_R$, con $\mu_R > 0$. La costante di trasferimento della funzione d'anello è $\rho = -0.2\mu_R$, e dunque interessa il grafico del luogo inverso, riportato nella Figura 12.26, limitatamente alla regione intorno all'origine. In realtà, il luogo possiede un terzo ramo che percorre l'asse reale dal polo in $s = -10$ verso $-\infty$, che è però irrilevante ai fini della determinazione dei poli dominanti. È immediato osservare che, finché la stabilità è garantita, il valore della pulsazione naturale dei poli dominanti in anello chiuso è comunque modesto (minore di 0.1), anche a causa dello zero che attira verso destra i rami a essi relativi.

Come già discusso nell'Esempio 11.3, le prestazioni possono essere migliorate impiegando come regolatore la rete anticipatrice

$$R(s) = \frac{\mu_R(1+10s)}{1+0.1s} \quad (12.11)$$

che modifica il luogo delle radici come mostrato nella Figura 12.27, e fa sì che risulti $\rho = -20\mu_R$. Si vede dalla figura che, per tutti i valori di ρ compresi nell'intervallo $-98 < \rho < -55$, il sistema è asintoticamente stabile e la distanza dei poli dominanti dall'origine rimane pressoché invariata. Poiché il vincolo sullo smorzamento impone che i poli in anello chiuso siano comunque all'interno del settore indicato in figura, bisogna rispettare anche la condizione $\rho \geq -76$. In effetti, il punto P corrispondente a $\rho = \bar{\rho} = -76$ è quello più lontano dall'origine tra tutti i punti dell'arco PQ e la sua distanza è $\omega_n \simeq 1.4$. La migliore taratura del guadagno del regolatore è allora $\mu_R = |\bar{\rho}|/20 =$

Figura 12.26
Luogo delle radici con regolatore proporzionale nell'Esempio 12.9.

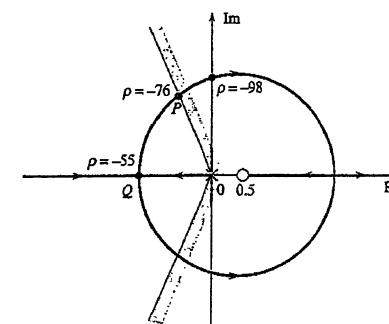
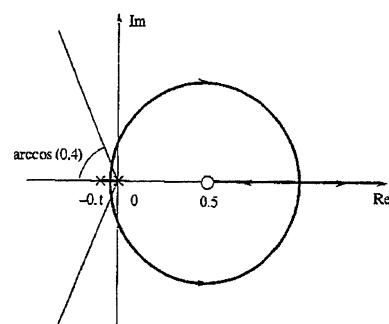


Figura 12.27
Luogo delle radici (in vicinanza dell'origine) con rete anticipatrice (12.11) nell'Esempio 12.9.

3.8, da confrontare con quella ricavata nel corso dell'Esempio 11.3. In realtà, si verifica attraverso i diagrammi di Bode che il regolatore (12.11) con $\mu_R = 3.8$ produce la pulsazione critica $\omega_c \simeq 0.6$ e il margine di fase $\varphi_m \simeq 33^\circ$, leggermente inferiore a quanto richiesto. Sarebbe pertanto opportuno ridurre un poco il guadagno μ_R per rispettare la specifica iniziale sul margine di fase. Si osservi infine che, in questo caso, la pulsazione critica fornisce un'approssimazione piuttosto imprecisa dell'effettivo valore della pulsazione naturale dei poli dominanti $\omega_n \simeq 1.4$. Ciò si spiega con il fatto che il diagramma di Bode del modulo della funzione d'anello risultante rimane prossimo all'asse a 0 dB per un lungo tratto al di là del punto di attraversamento e dunque le conclusioni basate sui ragionamenti approssimati del Paragrafo 10.3.3 sono inevitabilmente soggette a errori.

12.5 Conclusioni

Il luogo delle radici consente di valutare la posizione dei poli di un sistema retroazionato al variare della costante di trasferimento della funzione d'anello, e quindi fornisce importanti indicazioni sulla stabilità e sul comportamento dinamico del sistema. Grazie al fatto che si riescono a determinare semplici regole per il tracciamento qualitativo del luogo, pur senza l'ausilio di particolari strumenti di calcolo, il metodo è di grande utilità pratica e può essere impiegato anche in fase di progetto del regolatore, in alternativa o in combinazione con le tecniche basate sulla risposta in frequenza. A questo riguardo, vale la pena di sottolineare che il metodo del luogo delle radici è applicabile anche in situazioni nelle quali non si può fare ricorso al criterio di Bode e alle tecniche su di esso fondate, come per esempio nella sintesi di sistemi di controllo per sistemi instabili. Viceversa, risulta alquanto difficile affrontare in questo ambito questioni di grande importanza nello studio dei sistemi retroazionati, quali quelle riguardanti la stabilità robusta (se non indirettamente attraverso la valutazione dello smorzamento dei poli dominanti), la moderazione del controllo e l'attenuazione dell'effetto dei disturbi.

Esercizi

Esercizio 12.1 Si verifichi che, rimuovendo l'ipotesi che $L(s)$ sia strettamente propria, alcuni rami del luogo delle radici tendono all'infinito per ρ finito.

Esercizio 12.2 Si dimostri la Regola 5 del Paragrafo 12.2.2, relativa alle proprietà degli asintoti del luogo delle radici.

Esercizio 12.3 Si dimostrino le Regole 8 e 9 del Paragrafo 12.2.2, relative alle tangenti al luogo delle radici nei punti di partenza e di arrivo.

Esercizio 12.4 Si verifichino i risultati di stabilità dell'Esempio 12.2 attraverso l'uso del criterio di Routh o mediante il calcolo dai diagrammi di Bode del margine di guadagno associato alla funzione $L(s)/\rho$.

Esercizio 12.5 Si dimostri che, per una generica funzione d'anello a sfasamento minimo con grado relativo $v = 1$, esiste sempre un valore di ρ sufficientemente grande capace di garantire l'asintotica stabilità del sistema retroazionato.

Esercizio 12.6 Si tracci il luogo delle radici (diretto e inverso) associato alla funzione d'anello

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

e si verifichi, al crescere di $|\rho|$, l'effetto "destabilizzante" dello zero nel semipiano destro.

Esercizio 12.7 Considerando il sistema retroazionato con funzione d'anello

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

si dimostri mediante il luogo delle radici che non esistono valori di ρ per cui il sistema è asintoticamente stabile.

Esercizio 12.8 Utilizzando il luogo delle radici, si progetti un opportuno regolatore che stabilizzi il sistema descritto da

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

Esercizio 12.9 Mediante le tecniche di analisi basate sulla risposta in frequenza, si verifichi il progetto dell'Esempio 12.6, valutando il tempo di assottigliamento del sistema di controllo in base ai valori di pulsazione critica e margine di fase.

Esercizio 12.10 Si confrontino i risultati dell'Esempio 12.8 con quelli ricavabili utilizzando tecniche di progetto basate sulla risposta in frequenza.

Assegnamento degli autovalori

13

13.1 Introduzione

I metodi di progetto del controllore finora visti fanno tutti riferimento a una descrizione nel dominio delle trasformate del sistema da controllare, e il risultato della sintesi è la funzione di trasferimento del regolatore. Queste tecniche però non sono le uniche disponibili. Ne esistono altre che, invece, sfruttano la rappresentazione di stato del sistema sotto controllo, cioè si basano su modelli nel dominio del tempo. Per questo motivo, sono dette anche tecniche di sintesi nello spazio di stato.

In questo capitolo ne sarà presentata una, la tecnica ad assegnamento degli autovalori, che mira a costruire un regolatore capace di far sì che gli autovalori del sistema retroazionato abbiano valori prestabiliti, scelti ad arbitrio. In questo modo si riesce a imporre al sistema di controllo un comportamento dinamico assegnato, oltre eventualmente a stabilizzare mediante retroazione un sistema instabile.

Gli argomenti trattati nel capitolo sono i seguenti:

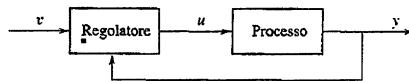
- una discussione preliminare sui vantaggi della retroazione dallo stato rispetto a quella dall'uscita;
- la sintesi del regolatore che assegna gli autovalori quando tutte le variabili di stato del sistema possono essere misurate;
- il progetto del cosiddetto osservatore (o ricostruttore) dello stato, cioè di un sistema dinamico in grado di fornire una stima attendibile dello stato quando questo non è misurabile direttamente;
- la sintesi del regolatore che assegna gli autovalori nel caso in cui lo stato non sia misurabile;
- un'interpretazione nel dominio delle trasformate dei risultati ottenuti.

Per ragioni di semplicità, si supporrà che il sistema sotto controllo sia un sistema SISO, anche se molti dei metodi presentati in questo capitolo si applicano, con poche modifiche, anche a sistemi multivariabili.

13.2 Retroazione statica dall'uscita o dallo stato

Finora si sono considerati sistemi di controllo con retroazione dall'uscita, immaginando cioè che la variabile di uscita fosse l'unica misura sul processo da controllare disponibile al regolatore. La struttura del sistema di controllo è dunque

Figura 13.1
Sistema di controllo con retroazione dall'uscita.

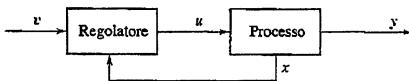


quella di Figura 13.1, dove la variabile v è un ulteriore ingresso del regolatore, tipicamente il segnale di riferimento.

In particolare, mediante il luogo delle radici si è visto come si possano valutare i poli del sistema in anello chiuso quando il regolatore è puramente statico, cioè quando la variabile di controllo è legata istantaneamente a quella di uscita. Al variare del guadagno d'anello, i poli in anello chiuso si spostano lungo il luogo delle radici e possono quindi essere in qualche misura modificati per ottenere le prestazioni desiderate. Essi però non possono essere assegnati arbitrariamente mediante una retroazione puramente statica dall'uscita. Come si sa, per raggiungere determinati obiettivi è di solito fondamentale utilizzare regolatori dinamici.

Se invece il regolatore dispone di un'informazione completa sulle variabili di stato del processo, come mostrato nello schema di Figura 13.2, è intuitivo che ci si troverebbe in una situazione più favorevole. In effetti, nel seguito del capitolo si mostrerà che con una retroazione statica dallo stato è possibile, sotto determinate ipotesi, fare in modo che gli autovalori del sistema in anello chiuso (e quindi i suoi poli) assumano valori arbitrari.

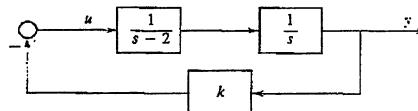
Figura 13.2
Sistema di controllo con retroazione dallo stato.



Un confronto tra le prestazioni ottenibili con una retroazione statica dall'uscita oppure dallo stato è discusso nel seguente semplice esempio.

Esempio 13.1 Con riferimento al sistema di Figura 13.3, si supponga di poter manipolare il guadagno k del regolatore nell'intento di stabilizzare il sistema.

Figura 13.3
Sistema di controllo dell'Esempio 13.1 con retroazione statica dall'uscita.



Tracciando il luogo delle radici associato alla funzione d'anello

$$L(s) = \frac{k}{s(s-2)}$$

ci si rende immediatamente conto (Figura 13.4) che non esiste alcun valore di k che rende il sistema asintoticamente stabile, poiché sia nel luogo diretto ($k > 0$) sia nel luogo inverso ($k < 0$) c'è sempre

Figura 13.4
Luogo delle radici associato al sistema di Figura 13.3.

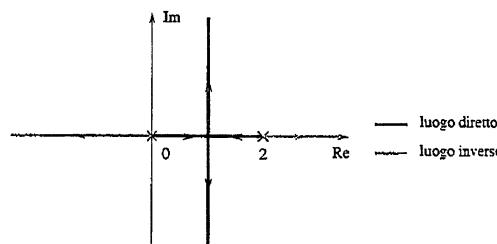
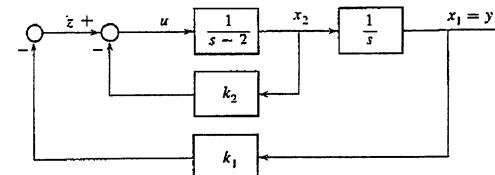


Figura 13.5
Sistema di controllo dell'Esempio 13.1 con retroazione statica dallo stato.



almeno uno dei due poli in anello chiuso che giace nel semipiano destro. In definitiva, il sistema non può essere stabilizzato mediante una retroazione statica dall'uscita.

Si consideri ora lo schema di Figura 13.5, in cui si è ipotizzato che il regolatore abbia accesso anche alla misura della variabile intermedia x_2 , oltre che alla misura dell'uscita $y = x_1$, ed effettui una retroazione statica $u(t) = -k_1 x_1(t) - k_2 x_2(t)$ a partire da queste due variabili. Le variabili x_1 e x_2 sono così indicate perché, come è facile verificare, possono essere scelte come variabili di stato del sistema. Analizzando l'anello più interno, la funzione di trasferimento da z a x_2 vale $F_2(s) = 1/(s+p)$, con $p = k_2 - 2$. Al variare di k_2 , si può modificare il valore di p a piacimento.

Ragionando ora sull'anello più esterno, la funzione d'anello è

$$L(s) = \frac{k_1}{s} F_2(s) = \frac{k_1}{s(s+p)}$$

e l'associato luogo delle radici diretto, supponendo $k_2 > 2$ e quindi $p > 0$, è mostrato nella Figura 13.6. Al variare di $k_1 > 0$, i poli in anello chiuso si spostano lungo tale luogo.

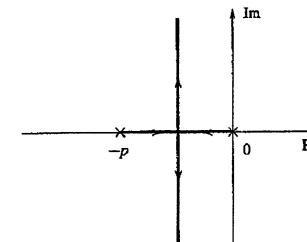


Figura 13.6
Luogo delle radici associato al sistema di Figura 13.5, con $k_1 > 0$ e $p = k_2 - 2$.

È facile a questo punto convincersi che, agendo simultaneamente sui guadagni k_1 e k_2 , si può non solo stabilizzare il sistema, ma anche allocare i poli in anello chiuso in posizioni arbitrarie del piano complesso (ovviamente conservando la simmetria rispetto all'asse reale). D'altra parte, si osservi che il polinomio caratteristico del sistema di Figura 13.5 è

$$\varphi(s) = s^2 + (k_2 - 2)s + k_1$$

e, con opportuni valori dei parametri k_1 e k_2 , si può fare in modo che esso abbia radici qualsiasi.

In conclusione, un'opportuna retroazione statica dallo stato permette in questo esempio di assegnare arbitrariamente i poli in anello chiuso.

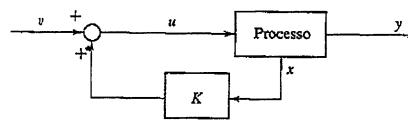
Il risultato dell'esempio ha validità molto più generale, come sarà mostrato nei paragrafi successivi.

13.3 Assegnamento degli autovalori con stato misurabile

In questo paragrafo si discuterà la possibilità di assegnare arbitrariamente gli autovalori del sistema in anello chiuso effettuando una retroazione statica dallo stato. Precisamente, si consideri il sistema di ordine n descritto dall'equazione di stato

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (13.1)$$

Figura 13.7
Schema per l'assegnamento degli autovalori con stato misurabile.



dove per semplicità si suppone che l'ingresso u sia scalare. La legge di controllo che si vuole utilizzare è

$$u(t) = Kx(t) + v(t) \quad (13.2)$$

in cui, per le ipotesi fatte,

$$K = [k_0 \ k_1 \ \dots \ k_{n-1}] \quad (13.3)$$

è un vettore riga (d'ora in poi detto anche *matrice guadagno*), mentre $v(t)$ rappresenta il nuovo ingresso del sistema retroazionato (Figura 13.7). La ragione dell'uso anomalo degli indici nel denotare gli elementi di K sarà chiara tra breve. Il sistema che si ottiene sostituendo la legge di controllo (13.2) nell'Equazione (13.1) è dunque descritto dall'equazione di stato

$$\dot{x}(t) = (A + BK)x(t) + Bv(t) = Fx(t) + Bv(t)$$

avendo definito $F = A + BK$.

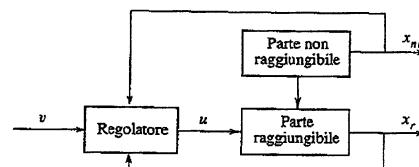
Il problema che si vuole ora considerare è se sia possibile, mediante un'opportuna scelta della matrice K , fare in modo che la matrice della dinamica F del sistema retroazionato abbia autovalori arbitrari. Ovviamenete, per trattare questo problema, si potrebbe porre $v(t) = 0$, ma la presenza dell'ingresso $v(t)$ è utile per altri scopi, come si discuterà più avanti. Si mostra ora che il problema ammette soluzione se e solo se il sistema (13.1) è completamente raggiungibile (Paragrafo 3.6.2). Più precisamente, vale il seguente risultato.

Teorema 13.1 Siano date le matrici reali A e B e un insieme arbitrario $\Lambda^0 = \{\lambda_1^0, \lambda_2^0, \dots, \lambda_n^0\}$ di numeri complessi (reali o complessi coniugati a coppie). Condizione necessaria e sufficiente perché esista una matrice reale K tale che gli autovalori di $F = A + BK$ coincidano con gli elementi di Λ^0 è che la coppia (A, B) sia completamente raggiungibile. ■

Che la condizione espressa dal Teorema 13.1 sia necessaria è abbastanza semplice da dimostrare. Se infatti il sistema (13.1) non fosse completamente raggiungibile, conterebbe una parte che non può essere in alcun modo influenzata dall'ingresso u , come indicato nella Figura 13.8. È allora del tutto evidente che gli autovalori di tale parte non potrebbero essere modificati da una qualunque retroazione che agisca su u .

Per mostrare invece che la raggiungibilità di (A, B) è anche una condizione sufficiente, è conveniente riferirsi inizialmente a un sistema in forma canonica di raggiungibilità (Paragrafo 4.5.1).

Figura 13.8
Retroazione dello stato per un sistema non completamente raggiungibile.



13.3.1 Sistema in forma canonica

Si supponga che il sistema (13.1) sia completamente raggiungibile e sia in forma canonica di raggiungibilità, ovvero si assuma che le matrici A e B abbiano la seguente struttura:

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-1} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (13.4)$$

Come si era osservato nel Paragrafo 4.5.1, è immediato verificare che in tal caso il polinomio caratteristico della matrice A è

$$\varphi(s) = s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_1s + \alpha_0$$

Inoltre, la matrice $F = A + BK$ conserva la struttura di A . Infatti, in base alle Equazioni (13.3) e (13.4), risulta

$$F = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 + k_0 & -\alpha_1 + k_1 & -\alpha_2 + k_2 & \dots & -\alpha_{n-1} + k_{n-1} \end{bmatrix}$$

Il polinomio caratteristico in anello chiuso è quindi

$$\varphi_F(s) = s^n + (\alpha_{n-1} - k_{n-1})s^{n-1} + \dots + (\alpha_1 - k_1)s + (\alpha_0 - k_0) \quad (13.5)$$

A questo punto, scegliendo opportunamente gli n parametri liberi k_i , $i = 0, \dots, n-1$, si può rendere il polinomio (13.5) identico a un qualunque polinomio preassegnato di grado n (con il primo coefficiente uguale a 1). In altre parole, si possono assegnare arbitrariamente gli autovalori del sistema in anello chiuso. Più precisamente, si indichino con λ_i^0 , $i = 1, \dots, n$ gli autovalori che si vogliono assegnare (eventualmente complessi, ma in tal caso coniugati a coppie) e si costruisca il polinomio caratteristico desiderato

$$\begin{aligned} \varphi^0(s) &= (s - \lambda_1^0)(s - \lambda_2^0) \cdots (s - \lambda_n^0) = \\ &= s^n + \beta_{n-1}s^{n-1} + \dots + \beta_1s + \beta_0 \end{aligned} \quad (13.6)$$

Per fare in modo che F abbia gli autovalori desiderati, basta porre

$$k_i = \alpha_i - \beta_i, \quad i = 0, \dots, n-1$$

e la matrice guadagno che risolve il problema è

$$K = [\alpha_0 - \beta_0 \ \alpha_1 - \beta_1 \ \dots \ \alpha_{n-1} - \beta_{n-1}] \quad (13.7)$$

13.3.2 Sistema non in forma canonica

Si passi ora a considerare la situazione più generale in cui il sistema (13.1) sia si completamente raggiungibile, ma non sia nella forma canonica di raggiungibilità (13.4). In tal caso, per poter applicare gli argomenti precedenti, è necessaria una trasformazione preliminare di variabili che ponga il sistema in forma canonica. In linea di principio, occorre cioè determinare una matrice T tale che, operando la trasformazione $\hat{A} = TAT^{-1}$, $\hat{B} = TB$, la coppia (\hat{A}, \hat{B}) sia nella forma canonica di raggiungibilità, cioè risulti

$$\hat{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-1} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Si noti che tale trasformazione esiste sempre sotto l'ipotesi di completa raggiungibilità. Procedendo in questo modo si arriva a dimostrare che la matrice guadagno che assegna arbitrariamente gli autovalori in anello chiuso è data da

$$K = \hat{K} \hat{M}_r M_r^{-1} \quad (13.8)$$

dove

$$M_r = [B \ AB \ \dots \ A^{n-1}B]$$

è la matrice di raggiungibilità del sistema originario,

$$\hat{M}_r = [\hat{B} \ \hat{A}\hat{B} \ \dots \ \hat{A}^{n-1}\hat{B}]$$

è la matrice di raggiungibilità del sistema in forma canonica, e

$$\hat{K} = [\alpha_0 - \beta_0 \quad \alpha_1 - \beta_1 \quad \dots \quad \alpha_{n-1} - \beta_{n-1}]$$

I parametri β_i presenti in quest'ultima formula sono i coefficienti del polinomio caratteristico desiderato (13.6). L'inversione di matrice che appare nella formula (13.8) è sempre possibile visto che, grazie all'ipotesi di completa raggiungibilità, la matrice M_r è non singolare. Si noti infine che i calcoli non richiedono la costruzione esplicita della matrice di trasformazione T .

Esempio 13.2 Si consideri il sistema descritto dalle matrici

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

che, come è facile verificare, è completamente raggiungibile e ha autovalori in -1 e 2 . Si voglia determinare la matrice guadagno K in modo che gli autovalori di $F = A + BK$ valgano -3 e -4 . Risulta

$$M_r = \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}, \quad M_r^{-1} = \begin{bmatrix} -\frac{2}{3} & \frac{5}{6} \\ \frac{1}{3} & -\frac{1}{6} \end{bmatrix}$$

Inoltre, il polinomio caratteristico in anello aperto è $\varphi(s) = s^2 - s - 2$ e quindi si ha $\alpha_0 = -2$, $\alpha_1 = -1$. Il sistema in forma canonica è allora rappresentato dalla coppia di matrici

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

e la sua matrice di raggiungibilità è

$$\hat{M}_r = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Poiché il polinomio caratteristico desiderato in anello chiuso è $\varphi^*(s) = s^2 + 7s + 12$, risulta $\beta_0 = 12$, $\beta_1 = 7$. Quindi

$$\hat{K} = [-14 \ -8]$$

Infine, dalla (13.8) si ricava

$$K = \hat{K} \hat{M}_r M_r^{-1} = [-2 \ -3]$$

Si lascia al lettore la verifica che, in effetti, la matrice $F = A + BK$ ha autovalori in -3 e -4 .

La formula (13.8) non è l'unica possibile per risolvere il problema di assegnamento degli autovalori. Esistono altri metodi equivalenti che, per esempio, non richiedono l'inversione di matrici.

La teoria finora discussa può essere opportunamente estesa al caso in cui le variabili di ingresso siano più di una, cioè B sia una matrice con un numero di colonne maggiore di 1, e la matrice guadagno K abbia un corrispondente numero di righe. Se l'obiettivo resta quello di assegnare arbitrariamente gli autovalori in anello chiuso, è abbastanza facile intuire che in tal caso vi sono maggiori gradi di libertà nel progetto della legge di controllo. Tali gradi di libertà possono essere sfruttati, per esempio, per sviluppare metodi di numericamente meglio condizionati oppure per ottenere, tra le tante soluzioni, quella più robusta rispetto a eventuali incertezze nella descrizione del sistema. In ogni caso, si dimostra che l'ipotesi fondamentale rimane quella della completa raggiungibilità della coppia (A, B) . Su questi aspetti del problema si rimanda il lettore a testi più avanzati.

13.4 Osservatore dello stato

Si è finora ipotizzato che il regolatore disponga di un'informazione completa su tutte le variabili di stato. Poiché, come è facile intuire, tale ipotesi è poco realistica nella maggior parte delle applicazioni pratiche, diventa importante chiedersi se, dato un sistema dinamico, sia possibile determinare un'accettabile approssimazione dello stato a partire dalla sola conoscenza delle variabili di ingresso e di uscita. Tale problema viene chiamato di *osservazione* (o *ricostruzione*) dello stato e, come si vedrà, la sua soluzione presenta forti analogie con i risultati del paragrafo precedente.

Il sistema di cui si vuole osservare lo stato sia descritto dalle equazioni

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (13.9)$$

$$y(t) = Cx(t) + Du(t) \quad (13.10)$$

dove si ipotizza per semplicità che ingresso e uscita siano scalari. Quindi B è un vettore colonna, C è un vettore riga, e D è uno scalare. Supponiamo inoltre che il sistema evolva a partire dallo stato iniziale $x(0) = x_0$, non noto.

Come indicato nello schema di Figura 13.9, l'osservatore è un sistema (eventualmente dinamico) che, avendo come ingressi $u(t)$ e $y(t)$, fornisce come uscita

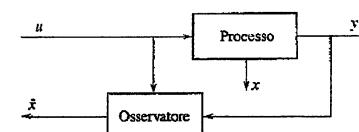


Figura 13.9
Schema per l'osservazione dello stato.

una stima $\hat{x}(t)$ dello stato $x(t)$ del sistema. La qualità della stima può essere misurata in termini dell'entità dell'errore di osservazione $e(t) = \hat{x}(t) - x(t)$.

13.4.1 Osservatore banale

Una soluzione banale al problema di osservazione è costituita dall'uso di una "replica" del sistema originario, alimentata dal medesimo ingresso, cioè

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t), \quad \hat{x}(0) = \hat{x}_0 \quad (13.11)$$

Il valore \hat{x}_0 va interpretato come la stima *a priori* dello stato iniziale. Si noti, tra l'altro, che tale osservatore non fa uso dell'informazione sull'uscita $y(t)$. Se lo stato iniziale x_0 fosse perfettamente noto (cioè se si potesse porre $\hat{x}_0 = x_0$) e non ci fosse alcuna incertezza sulle matrici che descrivono il sistema, l'osservatore (13.11) sarebbe in grado di riprodurre con assoluta precisione l'evoluzione dello stato del sistema originario. Se invece $\hat{x}_0 \neq x_0$, facendo la differenza tra le Equazioni (13.11) e (13.9) si verifica che l'evoluzione temporale dell'errore di stima $e(t) = \hat{x}(t) - x(t)$ è descritta dall'equazione differenziale

$$\dot{e}(t) = Ae(t), \quad e(0) = \hat{x}_0 - x_0 \neq 0$$

Si possono allora presentare due casi:

- la matrice A è asintoticamente stabile: l'errore di stima tende a zero, al crescere di t , per qualunque valore dell'errore iniziale $e(0)$, ma la rapidità di convergenza dipende dagli autovalori di A ;
- la matrice A non è asintoticamente stabile: l'errore di stima non converge a zero, o addirittura diverge nel caso di matrice instabile.

Entrambe le conclusioni precedenti non sono particolarmente soddisfacenti, dato che nel progetto di un osservatore sarebbe gradita la possibilità di governare la dinamica dell'errore attraverso opportune scelte progettuali, che portino, per esempio, ad aumentare la velocità di convergenza dello stato stimato $\hat{x}(t)$ allo stato "vero" $x(t)$. Con l'osservatore banale (13.11), la convergenza a zero dell'errore e la sua dinamica dipendono invece solo dalle proprietà del sistema di partenza. Soprattutto per questo motivo è opportuno fare riferimento a una diversa struttura dell'osservatore.

13.4.2 Osservatore asintotico con dinamica arbitraria

Si è già notato che la formula di ricostruzione (13.11) non sfrutta in alcun modo l'informazione sull'uscita $y(t)$ del sistema, che invece si suppone misurabile. Per ovviare agli inconvenienti prima sottolineati si può ricorrere alla seguente formulazione dell'osservatore:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + H(\hat{y}(t) - y(t)), \quad \hat{x}(0) = \hat{x}_0 \quad (13.12)$$

$$\hat{y}(t) = C\hat{x}(t) + Du(t) \quad (13.13)$$

In sostanza, nell'equazione di stato si è aggiunto un termine correttivo proporzionale alla discrepanza tra l'uscita stimata $\hat{y}(t)$, calcolata mediante la (13.13), e l'uscita vera $y(t)$. Ciò appare ragionevole, in quanto valori assoluti elevati di tale discrepanza suggeriscono un'azione correttiva più consistente, tale da far riavvicinare la stima $\hat{x}(t)$ al valore corretto $x(t)$.

La matrice H (che nelle ipotesi fatte è in realtà un vettore colonna di dimensione n) viene chiamata *matrice guadagno dell'osservatore* e rappresenta un parametro libero, a disposizione del progettista. Si vedrà ora come la si possa scegliere per migliorare le prestazioni dell'osservatore. A tale scopo si osservi che, facendo la differenza tra le Equazioni (13.12) e (13.9) e usando anche le trasformazioni di uscita (13.13) e (13.10), si ottiene con semplici passaggi

$$\dot{e}(t) = (A + HC)e(t), \quad e(0) = \hat{x}_0 - x_0$$

Come si vede, la dinamica dell'errore di stima dipende questa volta, oltre che da A e C , anche dal parametro di progetto H . Per imporre una determinata dinamica, si è allora indotti a scegliere H così da assegnare in modo appropriato gli autovalori della matrice $N = A + HC$. Se, per esempio, si riesce a garantire che tutti gli autovalori di N abbiano parte reale negativa, allora l'errore converge asintoticamente a zero indipendentemente dall'errore iniziale. Per questo motivo il sistema (13.12), (13.13) viene chiamato *osservatore asintotico dello stato*.

Il problema che qui si presenta (cioè se sia possibile scegliere H in modo da assegnare arbitrariamente gli autovalori di $N = A + HC$) ha molte affinità con il problema di assegnamento degli autovalori mediante retroazione dallo stato, considerato al Paragrafo 13.3. In effetti, la soluzione può essere ricondotta a quella già nota osservando che:

- gli autovalori di una matrice coincidono con quelli della sua trasposta;
- per le note proprietà delle matrici trasposte, risulta $(A + HC)' = A' + C'H'$;
- una coppia di matrici (A, C) è completamente osservabile se e solo se la coppia (A', C') è completamente raggiungibile.

Grazie a queste osservazioni è facile dimostrare il seguente risultato, "duale" di quello del Teorema 13.1.

Teorema 13.2 Siano date le matrici reali A e C e un insieme arbitrario $\Lambda^o = \{\lambda_1^o, \lambda_2^o, \dots, \lambda_n^o\}$ di numeri complessi (reali o complessi coniugati a coppie). Condizione necessaria e sufficiente perché esista una matrice reale H tale che gli autovalori di $N = A + HC$ coincidano con gli elementi di Λ^o è che la coppia (A, C) sia completamente osservabile. ■

Inoltre, la matrice H' può essere calcolata con la stessa procedura con cui era stato calcolato il guadagno K al Paragrafo 13.3, pur di sostituire A con A' e B con C' . Così facendo, si trova che la formula per la determinazione di H è data da

$$H = (\tilde{H}' \tilde{M}_o M_o^{-1})' \quad (13.14)$$

dove

$$M_o = [C' \ A'C' \ \dots \ (A')^{n-1} C']$$

è la matrice di osservabilità del sistema originario,

$$\tilde{M}_o = [\tilde{C}' \ \tilde{A}'\tilde{C}' \ \dots \ (\tilde{A}')^{n-1} \tilde{C}']$$

è la matrice di osservabilità del sistema in forma canonica di osservabilità (Paragrafo 4.5.2), e

$$\tilde{H}' = [\alpha_0 - \beta_0 \quad \alpha_1 - \beta_1 \quad \dots \quad \alpha_{n-1} - \beta_{n-1}]$$

I parametri α_i e β_i presenti in quest'ultima formula hanno lo stesso significato visto in precedenza, cioè rappresentano rispettivamente i coefficienti del polinomio caratteristico di A (o di A') e quelli del polinomio caratteristico desiderato per la matrice N .

Esempio 13.3 Si consideri il sistema (13.9), (13.10) con

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

Tale sistema è completamente osservabile dato che la matrice di osservabilità

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

è non singolare. Inoltre, gli autovalori di A sono -1 e 2 . Pertanto, a causa dell'autovalore instabile, l'osservatore banale dello stato produrrebbe in generale un errore divergente.

Si determini allora un osservatore asintotico del tipo (13.12), (13.13), tale che la matrice $N = A + HC$ abbia entrambi gli autovalori in -10 .

Ponendo il sistema in forma canonica di osservabilità, risulta

$$\tilde{A} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \tilde{M}_o = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Inoltre, il polinomio caratteristico desiderato è $\varphi^*(s) = s^2 + 20s + 100$ e quindi $\beta_0 = 100$, $\beta_1 = 20$. Infine si ha, come nell'Esempio 13.2, $\alpha_0 = -2$, $\alpha_1 = -1$. Dalla formula (13.14) si ricava dunque, con semplici passaggi,

$$H = \begin{bmatrix} 3 \\ -24 \end{bmatrix}$$

Si lascia al lettore verificare che la matrice $N = A + HC$ ha effettivamente due autovalori coincidenti in -10 , come desiderato.

La teoria dell'osservatore, qui derivata nel caso di un'uscita scalare, può essere opportunamente estesa al caso in cui le variabili di uscita siano più di una, cioè C sia una matrice con un numero di righe maggiore di 1, e la matrice guadagno H abbia un corrispondente numero di colonne. Anche in questo caso, i gradi di libertà aggiuntivi possono essere sfruttati per ottenere una soluzione più robusta rispetto a eventuali incertezze sul modello del sistema. In ogni caso, si dimostra che l'ipotesi fondamentale rimane quella della completa osservabilità della coppia (A, C) . Il lettore interessato può approfondire tali aspetti su testi più avanzati dedicati all'argomento.

C'è infine da notare che l'osservatore (13.12), (13.13) è un sistema dinamico di ordine n , cioè uguale a quello del sistema originario. In realtà, sotto l'ipotesi di osservabilità, si può dimostrare che, per un sistema di ordine n con p uscite, è possibile progettare un osservatore asintotico con dinamica arbitraria di ordine $n - p$, anziché n , cioè un *osservatore di ordine ridotto*. Tale risultato ha una spiegazione intuitiva, visto che ogni componente del vettore $y(t) - Du(t)$ è in realtà una combinazione lineare delle variabili di stato che si vogliono stimare. Non è necessario quindi ricostruire l'intero vettore di stato, ma solamente $n - p$ sue componenti. Comunque, anche questo argomento esula dagli scopi della presente trattazione.

13.5 Assegnamento degli autovalori con stato non misurabile

Riassumendo quanto visto nei due paragrafi precedenti, dato il sistema (13.9), (13.10) e supponendo che esso sia completamente raggiungibile e osservabile, è possibile

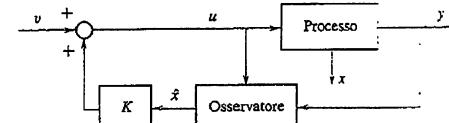


Figura 13.10
Schema per l'assegnamento degli autovalori con stato non misurabile.

- progettare una legge di controllo $u(t) = Kx(t) + v(t)$ che assegna arbitrariamente gli autovalori del sistema retroazionato;
- progettare un osservatore asintotico del tipo (13.12), (13.13), che fornisce una stima $\hat{x}(t)$ dello stato, assegnando arbitrariamente la dinamica dell'errore di stima.

A questo punto è legittimo chiedersi quali proprietà abbia il sistema che si ottiene combinando i due progetti, come mostrato nella Figura 13.10, effettuando cioè la retroazione a partire dallo stato stimato $\hat{x}(t)$ anziché da quello vero. L'analisi che segue mostrerà che il sistema risultante gode in effetti di alcune interessanti proprietà.

Si consideri dunque la legge di controllo

$$u(t) = K\hat{x}(t) + v(t) \quad (13.15)$$

Se si sostituisce la (13.15) nella (13.9) e nella (13.12), in base anche alle relazioni (13.10) e (13.13), si ottiene

$$\dot{x}(t) = Ax(t) + BK\hat{x}(t) + Bv(t) \quad (13.16)$$

$$\dot{\hat{x}}(t) = (A + BK + HC)\hat{x}(t) - HCx(t) - Bv(t) \quad (13.17)$$

Sottraendo la prima di queste relazioni dalla seconda e ricordando la definizione di $e(t)$, si ricava

$$\begin{aligned} \dot{e}(t) &= \dot{\hat{x}}(t) - \dot{x}(t) = (A + HC)\hat{x}(t) - (A - HC)x(t) = \\ &= (A + HC)e(t) \end{aligned} \quad (13.18)$$

Si noti ora che la (13.16), con la sostituzione $\hat{x}(t) = x(t) + e(t)$, può essere riscritta come

$$\dot{x}(t) = (A + BK)x(t) + BKe(t) - Bu(t) \quad (13.19)$$

In definitiva, in base alle (13.18) e (13.19), e definendo il vettore di stato allargato (di dimensione $2n$)

$$z(t) = \begin{bmatrix} x(t) \\ e(t) \end{bmatrix}$$

il sistema di Figura 13.10 può essere descritto dalla seguente rappresentazione di stato:

$$\dot{z}(t) = \begin{bmatrix} A + BK & BK \\ 0 & A + HC \end{bmatrix} z(t) + \begin{bmatrix} B \\ 0 \end{bmatrix} v(t) = \bar{A}z(t) + \bar{B}v(t)$$

Si osservi che la matrice \bar{A} è triangolare a blocchi. Pertanto, l'insieme dei suoi autovalori è costituito dalla riunione dell'insieme degli autovalori di $F = A + BK$ e dell'insieme degli autovalori di $N = A + HC$. Ciò autorizza a concludere che:

- il sistema complessivo è asintoticamente stabile se e solo se sono asintoticamente stabili le matrici F e N ;

- se il sistema originario è completamente raggiungibile e osservabile, gli autovalori del sistema complessivo sono assegnabili in modo arbitrario;
- il progetto del guadagno K della legge di controllo e il progetto del guadagno H dell'osservatore possono essere condotti indipendentemente l'uno dall'altro. Precisamente, il guadagno K può essere progettato come se lo stato $x(t)$ fosse misurabile, e il guadagno H può essere progettato come se il sistema di cui si vuole stimare lo stato fosse in anello aperto. Questo importante risultato va sotto il nome di *principio di separazione*.

In definitiva, combinando la legge di controllo e l'osservatore dello stato, progettati in modo indipendente, è possibile risolvere il problema dell'assegnamento degli autovalori anche quando lo stato non è direttamente misurabile. L'unica differenza è che, in questo tipo di progetto, bisogna scegliere, oltre agli n autovalori del sistema in anello chiuso, anche gli n autovalori associati alla dinamica dell'osservatore. Di solito è conveniente scegliere questi ultimi in modo che le costanti di tempo a essi associate siano minori di quelle associate agli autovalori del sistema in anello chiuso, in modo che l'errore di stima converga a zero più rapidamente rispetto alla dinamica del sistema retroazionato.

Il fatto che la tecnica di progetto qui illustrata permetta di assegnare in modo arbitrario tutti gli autovalori del sistema in anello chiuso potrebbe far pensare che, sotto le ipotesi di completa raggiungibilità e osservabilità, non vi sia alcuna limitazione nel rendere il sistema di controllo arbitrariamente veloce. In realtà è facile rendersi conto che, quanto più si tenta di aumentare la velocità del sistema, rendendo sempre più negativa la parte reale degli autovalori, tanto più viene sollecitata la variabile di controllo, la cui azione diviene sempre meno moderata. Una tecnica di progetto del controllore nello spazio di stato per certi versi analoga a quella qui presentata, ma in grado di tenere esplicitamente conto della moderazione del controllo, oltre che delle prestazioni dinamiche del sistema, è quella che va sotto il nome di *controllo ottimo*.

13.6 Interpretazione in termini di funzioni di trasferimento

Il sistema di controllo della Figura 13.10 è stato progettato usando una rappresentazione nel dominio del tempo del sistema da controllare. Esso sembra quindi avere poche analogie con i sistemi di controllo studiati nei capitoli precedenti, descritti nel dominio delle trasformate. In realtà, quando u e y sono scalari, è facile interpretare il sistema di controllo della Figura 13.10 mediante funzioni di trasferimento. Infatti, si può isolare il regolatore come mostrato nella Figura 13.11a e vederlo come un sistema lineare con due ingressi scalari, v e y , e un'uscita scalare u , ridisegnandone lo schema come nella Figura 13.11b. Il segno meno con cui viene fatto entrare nel nodo sommatore il ramo che dipende dall'ingresso y è stato posto per conformità con il classico schema a retroazione negativa. Inoltre si assumerà, come già nei Capitoli 9 e 10, che il sistema sotto controllo sia strettamente proprio, cioè $D = 0$.

Per calcolare le funzioni di trasferimento $P(s)$ e $R(s)$ della Figura 13.11b, si possono utilizzare le relazioni del paragrafo precedente. Precisamente, dalla (13.17), ricordando le (13.10), (13.15), si ricava

$$\begin{aligned}\dot{\hat{x}}(t) &= (A + BK + HC)\hat{x}(t) - Hy(t) + Bu(t) \\ u(t) &= K\hat{x}(t) + v(t)\end{aligned}$$

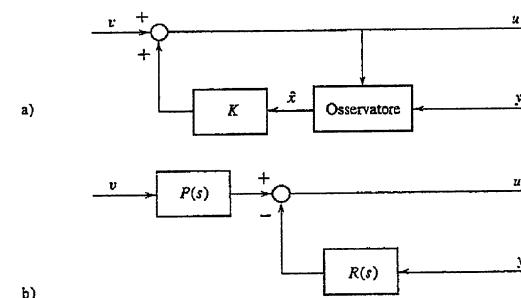


Figura 13.11
Schema del regolatore per l'assegnamento degli autovalori: a) in termini di stato; b) in termini di funzioni di trasferimento.

Pertanto, posto $\Psi = A + BK + HC$, si ottiene

$$\begin{aligned}P(s) &= \frac{U(s)}{V(s)} = K(sI - \Psi)^{-1}B + 1 \\ R(s) &= -\frac{U(s)}{Y(s)} = K(sI - \Psi)^{-1}H\end{aligned}$$

Si noti che non è affatto detto che tali funzioni di trasferimento siano asintoticamente stabili. Infatti, mentre le matrici $F = A + BK$ e $N = A + HC$ hanno tutti gli autovalori nel semipiano sinistro per costruzione, nulla si può dire in generale sugli autovalori della matrice Ψ .

Una volta che il regolatore è stato così descritto, l'analisi delle prestazioni del sistema di controllo può essere condotta facendo riferimento allo schema di Figura 13.12, nel quale $G(s) = C(sI - A)^{-1}B$ è la funzione di trasferimento del sistema sotto controllo, e utilizzando gli strumenti classici dell'analisi di sistemi retroazionati, illustrati nei capitoli precedenti.

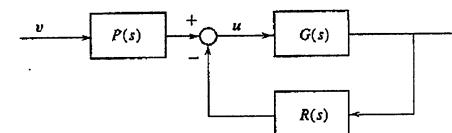


Figura 13.12
Schema per l'assegnamento degli autovalori in termini di funzioni di trasferimento.

Esempio 13.4 (Seguito degli Esempi 13.2 e 13.3) Combinando i progetti dei due precedenti esempi e supponendo $D = 0$, si ottiene

$$\Psi = A + BK + HC = \begin{bmatrix} 0 & 3 \\ -28 & -28 \end{bmatrix}$$

e le funzioni di trasferimento $P(s)$ e $R(s)$ del regolatore sono

$$\begin{aligned}P(s) &= K(sI - \Psi)^{-1}B + 1 = \frac{s^2 + 20s + 102}{s^2 + 28s + 84} \\ R(s) &= K(sI - \Psi)^{-1}H = \frac{66s + 228}{s^2 + 28s + 84}\end{aligned}$$

Si noti che i poli di tali funzioni hanno parte reale negativa. Il regolatore risulta quindi asintoticamente stabile, anche se ciò non era garantito *a priori*.

La funzione di trasferimento del sistema da controllare è invece

$$G(s) = C(sI - A)^{-1}B = \frac{3s + 6}{s^2 - s - 2}$$

Pertanto, la funzione d'anello del sistema retroazionato vale

$$L(s) = R(s)G(s) = \frac{(66s + 228)(3s + 6)}{(s^2 + 28s + 84)(s^2 - s - 2)}$$

In questo caso, visto che la funzione d'anello presenta un polo instabile, per valutare le prestazioni del sistema di controllo non si possono utilizzare gli strumenti basati sul criterio di Bode. Si può però verificare che i poli del sistema retroazionato sono posizionati come desiderato. Infatti, il polinomio caratteristico in anello chiuso vale

$$\begin{aligned} \hat{\phi}(s) &= (s^2 + 28s + 84)(s^2 - s - 2) + (66s + 228)(3s + 6) = \\ &= s^4 + 27s^3 + 252s^2 + 940s + 1200 = \\ &= (s + 3)(s + 4)(s + 10)^2 \end{aligned}$$

ovvero i quattro poli del sistema coincidono con i valori $-3, -4, -10$ (doppio) che erano stati scelti in fase di progetto.

Va sottolineato il fatto che il progetto di un regolatore in grado di assegnare i poli in anello chiuso del sistema di Figura 13.12 può essere condotto anche direttamente nel dominio delle trasformate, senza fare riferimento alla procedura nello spazio di stato che è stata presentata nel paragrafo precedente, ma ricorrendo ai cosiddetti *metodi polinomiali*, di cui si vedrà un esempio di applicazione (a tempo discreto) nel Paragrafo 18.5.4.

Regolatore con azione integrale Interpretando la variabile v come il segnale di riferimento, il sistema di controllo di Figura 13.12 non garantisce alcun tipo di precisione statica. Quando questo requisito risulta importante, è conveniente fare ricorso a una variante della tecnica ad assegnamento degli autovalori vista finora, con l'introduzione di un'azione integrale. Il relativo schema è mostrato nella Figura 13.13. Grazie alla presenza dell'integratore, se il sistema va a regime con un riferimento w a scalino, il valore a transitorio esaurito dell'errore di inseguimento $e_w = w - y$ è nullo. Il progetto del guadagno \bar{k} dell'integratore può essere condotto congiuntamente a quello della matrice guadagno K con l'obiettivo di assegnare arbitrariamente gli $n+1$ autovalori del sistema in anello chiuso, assumendo che lo stato x del processo sia misurabile. Per fare ciò si deve considerare il sistema "allargato", costituito dal processo e dall'integratore, descritto, se si ipotizza ancora $D = 0$, dalle equazioni di stato

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (13.20)$$

$$\dot{e}_w(t) = w(t) - Cx(t) \quad (13.21)$$

Si dimostra che l'assegnamento arbitrario degli autovalori per il sistema (13.20), (13.21) mediante la legge di controllo

$$u(t) = [K \bar{k}] \begin{bmatrix} x(t) \\ e_w(t) \end{bmatrix}$$

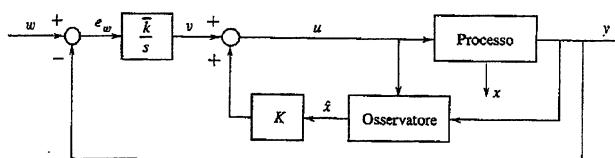


Figura 13.13
Schema per
l'assegnamento degli
autovalori con
l'introduzione di un'azione
integrale.

è possibile a patto che il sistema (13.9), (13.10) sia completamente raggiungibile e non possieda zeri nell'origine. Invece, il progetto dell'osservatore resta invariato rispetto al metodo precedente, perché non è necessario stimare la variabile di stato e_w del sistema allargato, che è perfettamente nota. In definitiva, con lo schema di Figura 13.13 si ottiene un sistema di controllo che, oltre a possedere una dinamica arbitraria, assicura un guadagno statico unitario tra il riferimento w e l'uscita y .

13.7 Conclusioni

La tecnica ad assegnamento degli autovalori, unita a quella per il progetto di un osservatore dello stato, permette di sintetizzare un sistema di controllo a retroazione dall'uscita con preassegnati valori dei poli in anello chiuso. Come tutti i metodi di sintesi nello spazio di stato, questa tecnica ha il vantaggio di essere facilmente adattabile al caso di sistemi multivariabili. Nel caso però di sistemi a un ingresso e un'uscita, essa non consente con semplicità, come invece i metodi di sintesi nel dominio della frequenza, di tener contemporaneamente conto dei vari aspetti del progetto, quali quelli legati ai requisiti di moderazione, di robustezza e di insensibilità ai disturbi. Inoltre, il regolatore risultante può anche avere caratteristiche non prevedibili e non gradite, come per esempio zeri, o addirittura poli, nel semipiano destro. Per tutti questi motivi, la principale applicazione della tecnica ad assegnamento degli autovalori nel caso SISO è quella del progetto di un regolatore stabilizzante per un sistema instabile, problema nel quale, come è noto, le tecniche in frequenza soffrono di alcune limitazioni.

In tutto il capitolo si è supposto di conoscere una rappresentazione di stato del sistema da controllare. Va comunque osservato che, anche se il modello del sistema fosse dato originariamente sotto forma di una funzione di trasferimento, è relativamente semplice ricavare una rappresentazione di stato, per esempio in forma canonica di raggiungibilità, che permette di applicare la procedura di sintesi del Paragrafo 13.5.

Si noti infine che le tecniche di progetto dell'osservatore discusse nel Paragrafo 13.4 possono essere utili anche in problemi diversi da quelli di controllo considerati in questo capitolo. Per esempio, la possibilità di disporre di una stima affidabile di grandezze non misurabili è di notevole importanza nella supervisione di sistemi potenzialmente pericolosi o per sviluppare procedure di diagnostica dei guasti negli impianti industriali.

Esercizi

Esercizio 13.1 Con riferimento all'Esempio 13.1, si trovino i valori di k_1 e k_2 in modo che il sistema con retroazione dallo stato abbia autovalori in -4 e -10 . Si verifichi inoltre che non è possibile ottenere lo stesso risultato con una retroazione statica dall'uscita.

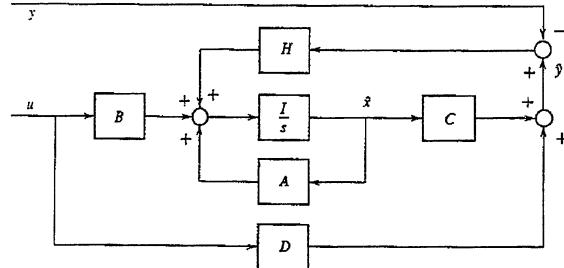
Esercizio 13.2 Usando la scomposizione canonica del sistema descritto dalla coppia (A, B) in parte raggiungibile e parte non raggiungibile, si mostri che gli autovalori della parte non raggiungibile sono autovalori della matrice $F = A + BK$, qualunque sia K .

In modo duale, a partire dalla scomposizione canonica del sistema descritto dalla coppia (A, C) in parte osservabile e parte non osservabile, si mostri che gli autovalori della parte non osservabile sono autovalori della matrice $N = A + HC$, qualunque sia H .

Esercizio 13.3 Si dimostri la correttezza della formula (13.8) per la matrice guadagno che assegna arbitrariamente gli autovalori nel caso che il sistema sia in forma canonica.

(Suggerimento: si verifichi dapprima che la trasformazione T che mette il sistema in forma canonica di raggiungibilità è $T = M_r M_r^{-1}$; si mostri poi che il legame tra il guadagno \hat{K} che assegna gli autovalori per il sistema in forma canonica e quello K per il sistema originario è $K = \hat{K}T$.)

Figura 13.14
Schema espanso
dell'osservatore asintotico
(13.12), (13.13).



Esercizio 13.6 Con riferimento allo schema di Figura 13.12, si verifichi che, quando il sistema sotto controllo non è strettamente proprio, le funzioni di trasferimento $P(s)$ e $R(s)$ del regolatore progettato mediante l'assegnamento degli autovalori valgono

$$P(s) = K(sI - \bar{\Psi})^{-1}(B + HD) + 1$$

$$R(s) = K(sI - \bar{\Psi})^{-1}H$$

con $\bar{\Psi} = A + BK + HC + HDK$.

Esercizio 13.7 Si ripeta il progetto descritto negli Esempi 13.2-13.4 allocando gli autovalori della matrice F in -1 e -4 . Si noti che, in questo caso, la funzione di trasferimento $R(s)$ del regolatore contiene uno zero in -1 che, ai fini del calcolo della funzione d'anello $L(s)$, cancella esattamente uno dei poli del sistema. Si verifichi che ciò avviene tutte le volte che tra gli autovalori desiderati di F appare uno degli autovalori di A , cioè quando si impone che uno degli autovalori del sistema non si sposti.

Esercizio 13.8 Mediante le tecniche viste nel capitolo, si progetti un regolatore per il sistema descritto dalla funzione di trasferimento

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

in modo da garantire una pulsazione critica $\omega_c > 2$. Si valuti quindi il margine di fase φ_m .

Esercizio 13.9 Mediante le tecniche viste nel capitolo, si progetti un regolatore stabilizzante per il sistema descritto dalla funzione di trasferimento

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

Esercizio 13.10 Si dimostri che, se il sistema (13.9), (13.10) è completamente raggiungibile e non possiede zeri nell'origine, allora il sistema allargato (13.20), (13.21), con $w(t) = 0$, risulta completamente raggiungibile.

(Suggerimento: si verifichi dapprima che, considerando solo l'ingresso u , la matrice di raggiungibilità del sistema (13.20), (13.21) può essere fattorizzata come

$$\tilde{M}_r = \begin{bmatrix} -A & -B \\ C & 0 \end{bmatrix} \begin{bmatrix} 0 & -M_r \\ -1 & 0 \end{bmatrix} = \Lambda \Theta$$

Pertanto, osservando che $\det(\Theta) = -\det(M_r)$, si ha $\det(\tilde{M}_r) = -\det(\Lambda) \det(M_r)$. Inoltre, è abbastanza facile verificare che $\det(\Lambda) = -CA^{-1}B = G(0)$. A tale scopo si può immaginare che le matrici A , B , e C siano in forma canonica di raggiungibilità. In conclusione, le ipotesi $\det(M_r) \neq 0$ e $G(0) \neq 0$ implicano che $\det(\tilde{M}_r) \neq 0$, cioè che il sistema allargato è completamente raggiungibile.)

Regolatori PID

14

14.1 Introduzione

I regolatori lineari più usati in ambito industriale sono certamente i PID, o regolatori ad azione Proporzionale, Integrale, Derivativa. Le ragioni del loro successo sono varie: prima di tutto il loro impiego consente di controllare in modo soddisfacente un'ampia gamma di processi; in secondo luogo, negli anni sono state sviluppate e largamente utilizzate semplici regole per la loro taratura automatica, applicabili con buoni risultati anche nel caso in cui non sia disponibile un modello matematico preciso del sistema sotto controllo. Inoltre, per la loro semplicità, i regolatori PID possono essere realizzati con le tecnologie più varie: meccanica, pneumatica, idraulica, elettronica analogica e digitale. Questo implica una grande disponibilità commerciale, che permette la realizzazione in tempi brevi e con costi contenuti di schemi di controllo complessi.

Nel capitolo saranno trattati i seguenti argomenti:

- il modello dei regolatori PID e i parametri che li caratterizzano;
- le problematiche relative al loro impiego in ambito industriale e le soluzioni più comunemente adottate;
- le regole di taratura automatica e la loro interpretazione in base ai metodi di analisi e sintesi presentati nei Capitoli 9-12.

14.2 Modello dei regolatori PID

Tradizionalmente la struttura dei PID viene introdotta in base a considerazioni empiriche secondo le quali è opportuno che la variabile di controllo u sia generata come la somma di tre contributi:

- il primo, di significato intuitivo, è proporzionale all'errore e tra il segnale di riferimento w e la variabile di uscita y del sistema sotto controllo;
- il secondo è proporzionale all'integrale di e (cioè di fatto al suo valor medio) ed è richiesto per impostare che l'errore si annulli asintoticamente a fronte di segnali di riferimento o disturbi additivi costanti;
- il terzo è proporzionale alla derivata di e e ha lo scopo di tentare di "anticipare" l'andamento dell'errore negli istanti futuri: se per esempio la derivata dell'errore è positiva, così come il guadagno del sistema, è opportuno aumentare u per provocare un aumento di y e quindi una diminuzione di e .

La legge di controllo, cioè il legame tra e e u , è quindi

$$u(t) = K_P e(t) + K_I \int_{t_0}^t e(\tau) d\tau + K_D \frac{de(t)}{dt} \quad (14.1)$$

dove K_P , K_I e K_D sono costanti positive o nulle (nell'ipotesi che il guadagno del processo sia positivo). Il coefficiente K_P è detto *coefficiente dell'azione proporzionale*, mentre K_I e K_D sono rispettivamente il *coefficiente dell'azione integrale* e il *coefficiente dell'azione derivativa*. Nella letteratura tecnica si usa anche fare riferimento alla banda proporzionale $PB = 100/K_P$.

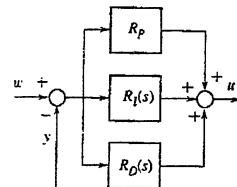
È facile verificare che i PID, almeno nella loro *forma ideale* (14.1), sono sistemi dinamici SISO, lineari, stazionari a tempo continuo e impropri. Applicando la trasformata di Laplace alla (14.1) con $t_0 = 0$, si deduce immediatamente che essi sono descritti dalla funzione di trasferimento

$$R_{PID}(s) = K_P + \frac{K_I}{s} + K_D s = \frac{K_D s^2 + K_P s + K_I}{s} \quad (14.2)$$

in cui si individuano il termine proporzionale ($R_P = K_P$), quello integrale ($R_I(s) = K_I/s$) e quello derivativo ($R_D(s) = K_D s$). In base a quanto esposto nei Capitoli 10 e 11, si può concludere che la presenza dell'azione integrale garantisce errore nullo a fronte di segnali di riferimento e disturbi additivi costanti, mentre l'azione derivativa introduce uno zero, che a sua volta genera un anticipo di fase e quindi una maggiore prontezza del sistema di controllo.

Lo schema a blocchi di un PID ideale è riportato nella Figura 14.1.

Figura 14.1
Struttura del PID ideale.



Una diversa rappresentazione dei PID, forse anche più utilizzata della (14.2), è data da

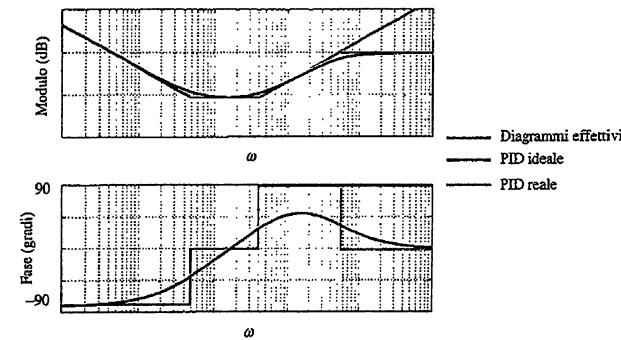
$$R_{PID}(s) = K_P \left(1 + \frac{1}{T_I s} + T_D s \right) = K_P \frac{T_I T_D s^2 + T_I s + 1}{T_I s} \quad (14.3)$$

in cui $T_I = K_P/K_I$ è il *tempo integrale* (o di *reset*) e $T_D = K_D/K_P$ è il *tempo derivativo*.

Il PID ideale con funzione di trasferimento (14.2), o (14.3), è un sistema con due zeri a parte reale negativa e un solo polo nell'origine; i diagrammi di Bode asintotici della sua risposta in frequenza sono riportati nella Figura 14.2. Il PID è quindi un sistema improprio per la presenza del termine derivativo $R_D(s)$. Per questo motivo nella pratica l'azione derivativa è ottenuta per mezzo della funzione di trasferimento

$$R_D^a(s) = \frac{K_P T_D s}{1 + \frac{T_D}{N} s} = \frac{K_D s}{1 + \frac{K_D}{K_P N} s} \quad (14.4)$$

dove la costante positiva N è scelta in modo che il polo $s = -N/T_D$, aggiunto per la realizzabilità, sia all'esterno della banda di frequenze di interesse nel controllo.



Valori tipici di N sono $5 \div 20$. Il PID in *forma reale* ha quindi la funzione di trasferimento

$$R_{PID}^a(s) = K_P \left(1 + \frac{1}{T_I s} + \frac{T_D}{1 + \frac{T_D}{N} s} s \right) = K_P + \frac{K_I}{s} + \frac{K_D s}{1 + \frac{K_D}{K_P N} s} \quad (14.5)$$

Si osservi che il polo in $s = -N/T_D$ di $R_D^a(s)$ modifica anche la posizione degli zeri della funzione di trasferimento complessiva del PID. Tuttavia, per N sufficientemente grande, si può verificare che gli zeri della (14.5) approssimativamente coincidono con quelli della (14.2). I diagrammi di Bode asintotici della risposta in frequenza di un PID reale sono anch'essi riportati nella Figura 14.2 unitamente ai diagrammi effettivi. Nel seguito, se non altrimenti specificato, si farà sempre riferimento alla forma ideale (14.2), sottintendendo comunque la presenza di un polo aggiuntivo in alta frequenza.

Naturalmente non tutte le azioni devono essere contemporaneamente presenti: in particolare è possibile impiegare soltanto una di esse o combinazioni di due. Trascurando i regolatori caratterizzati unicamente dalla presenza dell'azione derivativa, o delle azioni integrale e derivativa soltanto, dal generico PID si possono ottenere come casi particolari i regolatori di seguito elencati.

Regolatori P Ponendo $K_I = K_D = 0$ nella (14.2), o $T_I \rightarrow \infty$ e $T_D = 0$ nella (14.3), il PID può essere utilizzato come un semplice regolatore proporzionale con funzione di trasferimento $R_P = K_P$. L'impiego di un regolatore proporzionale (si vedano i Capitoli 9-12) è normalmente limitato al controllo di processi asintoticamente stabili, o semplicemente stabili, e nel caso in cui le prestazioni statiche richieste al sistema di controllo non rendano necessario l'inserimento di un'azione integrale.

Regolatori I Un PID può esercitare unicamente un'azione di tipo integrale pur di porre $K_P = K_D = 0$ nella (14.2), ottenendo così la funzione di trasferimento $R_I(s)$ precedentemente definita. Regolatori puramente integrali possono essere interpretati come reti ritardatrici in cui il polo è posto a pulsazione zero, mentre lo zero è all'infinito. In questo modo si accentuano le caratteristiche delle reti ritardatrici che, come discusso nel Paragrafo 11.5.2, portano a un restrimento della banda passante per far fronte al contributo negativo di fase dovuto all'impiego del regolatore. Questi regolatori sono utilizzati quando è necessaria un'azione

Figura 14.2
Diagrammi di Bode asintotici di un PID ideale e reale e diagrammi effettivi.

integrale per il soddisfacimento dei requisiti sull'errore a transitorio esaurito, ma non si richiedono prestazioni elevate in termini di velocità di risposta del sistema.

Regolatori PI Si ricavano dalla (14.2) ponendo $K_D = 0$, o dalla (14.3) con $T_D = 0$. Anch'essi possono essere visti come reti ritardatrici con il polo nell'origine del piano complesso e lo zero in $s = -1/T_I$. La loro funzione di trasferimento è quindi

$$R_{PI}(s) = \frac{K_P s + K_I}{s} = K_P \frac{1 + T_I s}{T_I s} \quad (14.6)$$

I regolatori PI vengono impiegati quando è indispensabile l'azione integrale per le prestazioni statiche, ma è necessaria anche la presenza di uno zero per avere una banda passante più ampia rispetto a quella ottenibile con un semplice regolatore integrale. I PI sono molto diffusi nel controllo dei processi industriali, spesso descrivibili come la serie di una funzione di trasferimento del primo ordine e di un eventuale ritardo, per i quali la presenza dell'ulteriore zero nella funzione di trasferimento di un PID completo, anziché migliorare le prestazioni, porterebbe soltanto a una maggiore difficoltà di taratura per un eccessivo allargamento della banda passante del sistema retroazionato.

Regolatori PD Ponendo $K_I = 0$ nella (14.2) si ottiene una rete anticipatrice in cui lo zero è posto in $s = -K_P/K_D = -1/T_D$ e il polo, nel caso ideale, può essere immaginato all'infinito. Corrispondentemente la funzione di trasferimento è

$$R_{PD}(s) = K_P + K_D s = K_P(1 + T_D s)$$

Nei PD reali il polo aggiunto all'azione derivativa (si ricordino le (14.4) e (14.5)) viene comunque posto al di fuori della banda di pulsazioni di interesse per il controllo. Come di consueto, l'impiego di una rete anticipatrice è tipico di quei casi in cui non vi siano problemi di stabilità o di prestazioni statiche, ma sia invece necessario ottenere la banda passante più ampia possibile.

Regolatori PID I PID hanno un polo nell'origine del piano complesso, e, nella loro forma ideale, due zeri in posizione

$$s = \frac{-T_I \pm \sqrt{T_I(T_I - 4T_D)}}{2T_I T_D} \quad (14.7)$$

Se si considera la forma reale, gli zeri si modificano (di poco) e si aggiunge un polo a pulsazione più elevata. Pertanto il PID può essere interpretato come una particolare rete a sella e il suo impiego consente il controllo di vaste classi di sistemi (si veda il Paragrafo 11.5.3). Dalla (14.7) si osservi inoltre che gli zeri sono reali per $T_I \geq 4T_D$ e che quando $T_I = 4T_D$ essi sono coincidenti in $s = -1/2T_D$; questa scelta è spesso effettuata per semplificare la taratura automatica, come discusso nel Paragrafo 14.4.

14.3 Realizzazione dei regolatori PID

Nella realizzazione pratica dei regolatori, e in particolare dei PID, vengono adottati vari accorgimenti volti a migliorare le prestazioni del sistema di controllo. In questo paragrafo si presentano i principali problemi connessi con l'uso dei PID e le tecniche utilizzate per la loro soluzione.

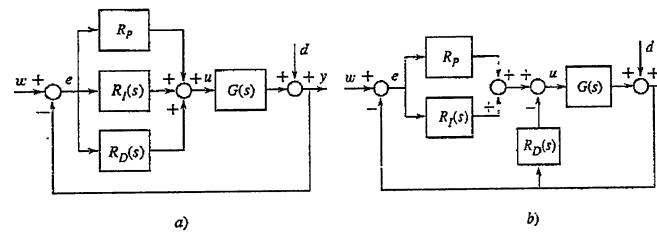


Figura 14.3
a) PID con derivazione dell'errore; b) PID con derivazione dell'uscita.

14.3.1 Limitazione dell'azione derivativa

Nel classico schema di controllo in retroazione di Figura 14.3a l'azione derivativa è effettuata sull'errore e ; in questo caso, in presenza di uno scalino del segnale di riferimento w , l'uscita del derivatore, e di conseguenza la variabile di controllo u , hanno un andamento di tipo impulsivo. Questa brusca variazione è in contrasto con il requisito di moderazione del controllo e può provocare la saturazione dell'attuatore e l'allontanamento del sistema dalla condizione di linearità con riferimento alla quale normalmente si progetta il regolatore. Per queste ragioni frequentemente l'azione derivativa è esercitata sulla sola variabile di uscita y , come mostrato nella Figura 14.3b. Poiché y è l'uscita di un sistema che usualmente ha le caratteristiche di un filtro passa-basso, le sue variazioni istantanee (e quindi la sua derivata) sono in genere contenute e la presenza dell'azione derivativa non provoca il suddetto andamento impulsivo di u .

Esempio 14.1 Per il sistema sotto controllo con funzione di trasferimento

$$G(s) = \frac{1}{(s+1)^3} \quad (14.8)$$

si consideri un PID ideale in cui $K_P = 2$, $K_I = K_D = 1$ ($T_I = 1$, $T_D = 0.5$). La funzione di trasferimento del regolatore è quindi

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

e la funzione d'anello è

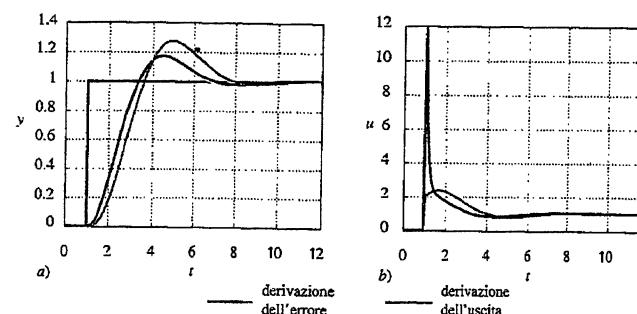
$$L(s) = R_{PID}(s)G(s) = \frac{1}{s(s+1)}$$

a cui corrisponde la pulsazione critica $\omega_c \approx 0.78$, il margine di fase $\varphi_m \approx 52^\circ$ e un margine di guadagno k_m infinito. La funzione di trasferimento del PID reale, con $N = 5$, cambia per l'aggiunta di un polo in $s = -N/T_D = -10$, ma la pulsazione critica risultante è $\omega_c \approx 0.8$, il margine di fase è $\varphi_m \approx 51^\circ$ e il margine di guadagno è $k_m \approx 22$. Le prestazioni del sistema complessivo cambiano quindi soltanto marginalmente tra un caso e l'altro.

Il PID reale così progettato è stato realizzato secondo i due schemi di Figura 14.3; le uscite y del sistema corrispondenti a uno scalino del segnale di riferimento $w = 1$ sono riportate nella Figura 14.4a e i corrispondenti andamenti di u sono rappresentati nella Figura 14.4b. Da queste figure si può notare che l'aver eliminato il picco di tipo impulsivo della variabile di controllo (Figura 14.4b) in corrispondenza della variazione a scalino di w produce soltanto un lieve deterioramento della risposta del sistema (Figura 14.4a).

Per quanto riguarda la scelta del parametro N , cioè del polo aggiunto per la realizzabilità del PID, si osservi che, al crescere di N , il PID reale si comporta in modo sempre più simile a quello ideale. D'altra parte, all'aumentare di N aumenta anche $|R_{PID}(j\omega)|$ per $\omega \rightarrow \infty$. Si ricordi anche che, come discusso nel Paragrafo 10.5.3, il modulo della risposta in frequenza associata alla funzione di sensitività del controllo $Q(s)$ coincide approssimativamente con $|R_{PID}(j\omega)|$ quando

Figura 14.4
Esempio 14.1:
a) Andamento di y e u con lo schema a derivazione dell'errore e con quello a derivazione dell'uscita;
b) andamento delle rispettive variabili di controllo.



$|R_{PID}(j\omega)G(j\omega)| \ll 1$. Per evitare che eventuali componenti in alta frequenza del disturbo d vengano eccessivamente amplificate sulla variabile di controllo, è opportuno allora scegliere un valore di N basso, sempre compatibilmente con l'esigenza di porre il polo in $s = -N/T_D$ al di fuori della banda di interesse per il controllo.

Esempio 14.2 (Seguito dell'Esempio 14.1) Nel sistema di controllo di Figura 14.3b, con $G(s)$ data dalla (14.8) e regolatore PID definito da $K_P = 2$, $K_I = K_D = 1$, al tempo $t = 1$ è stato imposto un andamento a scalino unitario al segnale di riferimento w , mentre il disturbo d è stato generato come un segnale con spettro costante. Le variabili y e u , a fronte del medesimo andamento temporale di d e per $N = 30$ e $N = 5$, sono riportate nelle Figure 14.5 e 14.6. La Figura 14.5 illustra l'andamento di y , praticamente indistinguibile nei due casi, mentre il confronto della Figura 14.6a ($N = 30$) con la Figura 14.6b ($N = 5$) mostra che per $N = 30$ la variabile di controllo è eccessiva-

Figura 14.5
Andamento di w e y nell'Esempio 14.2 con lo schema a derivazione dell'uscita per $N = 5$ e $N = 30$.

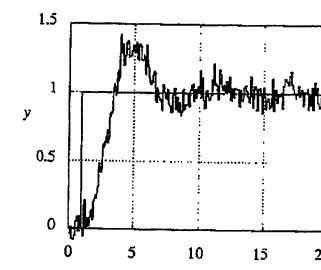
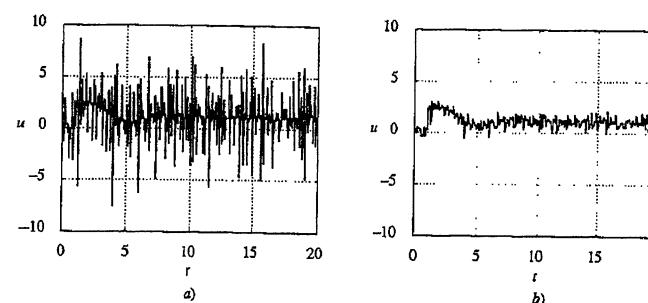


Figura 14.6
Andamento di u nell'Esempio 14.2 con lo schema a derivazione dell'uscita: a) per $N = 30$; b) per $N = 5$.



vamente e inutilmente sollecitata, in contrasto con il requisito di moderazione; questo risultato è in accordo con il fatto che in alta frequenza $|Q(j\omega)|$ è crescente con N .

I risultati dell'esempio precedente mettono in luce una difficoltà connessa con l'impiego dei regolatori PID, soprattutto se i loro parametri non vengono tarati a partire da un modello accurato dell'impianto: l'inserimento di un'azione derivativa può produrre una variabile di controllo eccessivamente sollecitata senza che questo comporti un miglioramento significativo delle prestazioni. Per questo motivo nella pratica industriale spesso si preferisce utilizzare regolatori PI nonostante ciò implichia una limitazione delle prestazioni potenzialmente ottenibili.

14.3.2 Desaturazione dell'azione integrale

La presenza combinata dell'azione integrale e di una saturazione dovuta all'attuatore provoca un effetto di tipo non lineare che può deteriorare significativamente le prestazioni del sistema di controllo. Per analizzare più in dettaglio questo fenomeno, si ipotizzi inizialmente e per semplicità che il regolatore sia puramente integrale e che l'attuatore, con ingresso u e uscita m , sia descritto dalla relazione

$$m(t) = \begin{cases} -u_M, & u(t) < -u_M \\ u(t), & |u(t)| \leq u_M \\ u_M, & u(t) > u_M \end{cases} \quad (14.9)$$

come mostrato nella Figura 14.7, dove $G(s)$ è la funzione di trasferimento del processo.

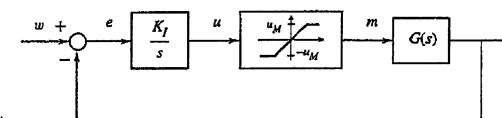


Figura 14.7
Schema di controllo con attuatore saturante.

Il fenomeno del wind-up Quando l'errore e si mantiene dello stesso segno per un certo periodo, lo stato dell'integratore, che coincide con la sua uscita u , cresce in modulo sempre più. Ciò avviene anche se l'effettiva variabile di ingresso m del sistema sotto controllo viene limitata al valore u_M o $-u_M$ dalla saturazione dovuta all'attuatore. Quando questo accade, se l'errore cambia segno è necessario attendere che lo stato u dell'integratore torni ad assumere valori in modulo inferiori a u_M prima che l'attuatore riprenda a operare in zona lineare, cioè si abbia $m(t) = u(t)$. In altri termini, si deve attendere la scarica dell'azione integrale. Sarebbe invece molto più opportuno che la variabile di controllo effettiva lasciasse il valore di saturazione non appena l'errore cambia segno. Il fenomeno sopra descritto prende il nome di *carica integrale* o, più comunemente dall'inglese, di *integral wind-up*.

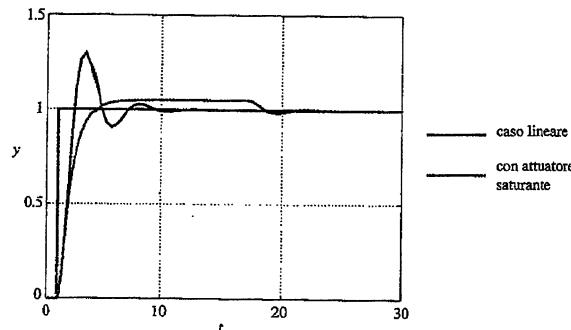
Esempio 14.3 Si consideri lo schema di Figura 14.7 in cui

$$G(s) = \frac{1}{s+1} \quad (14.10)$$

e $K_I = 2$. In assenza di limitazioni dovute all'attuatore il sistema retroazionato è lineare con $\omega_c \approx 1.25$ e $\varphi_m \approx 39^\circ$. Si ipotizzi invece che l'attuatore sia caratterizzato da $u_M = 1.05$. Il

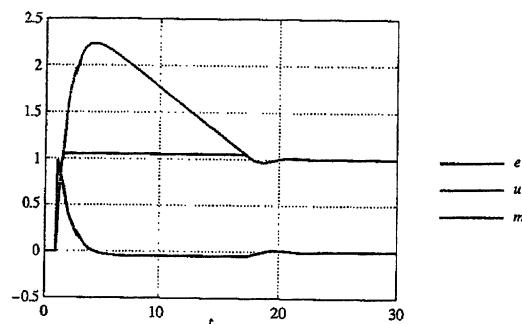
sistema di controllo sollecitato da un andamento a scalino unitario del segnale di riferimento a $t = 1$ produce, nel caso lineare e nel caso di saturazione dell'attuatore, i transitori della variabile di uscita riportati nella Figura 14.8. È evidente che, in presenza della saturazione, la variabile controllata si mantiene costante al valore u_M per un lungo periodo, assestandosi sul valore unitario desiderato solo dopo l'istante $t = 20$.

Figura 14.8
Andamento di w e y
nell'Esempio 14.3 nel caso
lineare e in presenza
dell'attuatore saturante.



Per comprendere meglio il fenomeno del wind-up, si analizzino gli andamenti, riportati nella Figura 14.9, delle variabili e , u e m in presenza della saturazione. Nell'istante di variazione del riferimento l'errore e è pari a 1; successivamente decresce finché a $t \approx 4.4$ diventa negativo. A partire da questo istante sarebbe opportuno che la variabile di controllo m , già saturata, cominciasse a decrescere e il sistema tornasse a operare in zona di linearità. Tuttavia, poiché la variabile u è arrivata ad assumere valori elevati, è necessario attendere che essa decresca fino a che, a $t \approx 17.4$, torna a essere minore di u_M .

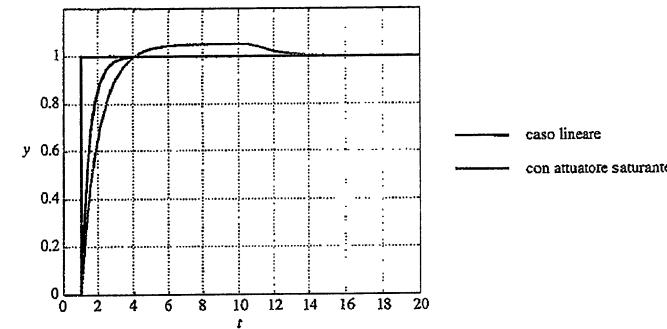
Figura 14.9
Andamento di e , u e m
nell'Esempio 14.3 in
presenza dell'attuatore
saturante.



Naturalmente il fenomeno del wind-up è presente anche nel caso di regolatori più complessi, in particolare PI o PID, purché caratterizzati dalla presenza di un'azione integrale.

Esempio 14.4 Si consideri ancora il sistema (14.10) e un regolatore PI con $K_P = K_I = 2$, a cui, in assenza di saturazioni, corrisponde un sistema retroazionato con $\omega_c = 2$ e $\varphi_m = 90^\circ$. Sollecitando il sistema di controllo con una variazione a scalino unitario del segnale di riferimento a $t = 1$, si ottengono, nel caso lineare e in presenza della saturazione con $u_M = 1.05$, i transitori riportati nella Figura 14.10. Anche in questo caso è evidente il deterioramento delle prestazioni dovuto alla presenza della saturazione sulla variabile di controllo.

Figura 14.10
Andamento di w e y
nell'Esempio 14.4 nel caso
lineare e in presenza
dell'attuatore saturante.



Uno schema di desaturazione per regolatori PID Il fenomeno del wind-up è dovuto al fatto che la dinamica del regolatore non è influenzata dall'eventuale presenza di limitazioni sulla sua variabile di uscita. Se queste limitazioni sono attive si usa dire che lo stato del regolatore *non è congruente* con l'effettiva variabile m . Per ovviare a ciò, tutti i vari schemi proposti in letteratura per attenuare il wind-up hanno in comune la caratteristica di alimentare il regolatore anche con il segnale a valle della saturazione, in modo che il suo stato possa evolvere coerentemente con l'andamento della variabile che effettivamente agisce sul processo. Viene ora presentato un possibile schema di desaturazione per i regolatori PID, rimandando al paragrafo successivo la sua generalizzazione a regolatori di struttura generica.

Si ipotizzi che l'attuatore sia descritto dalla (14.9), si consideri lo schema di Figura 14.11 e si noti preliminarmente che, poiché usualmente l'azione derivativa è esercitata solo sull'uscita y , lo schema di desaturazione riguarda unicamente le azioni proporzionale e integrale. Si osservi inoltre che all'interno del regolatore viene replicata la caratteristica non lineare dell'attuatore.

Per comprendere il principio di funzionamento dello schema di Figura 14.11, si osservi preliminarmente che, se si opera in zona di linearità, la funzione di trasferimento complessiva tra l'errore e e la variabile c coincide con quella di un

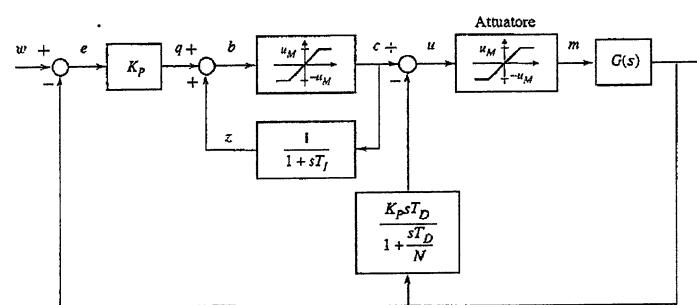


Figura 14.11
Schema di realizzazione
dei regolatori PID con
desaturazione.

PI. Infatti, in tal caso, risulta

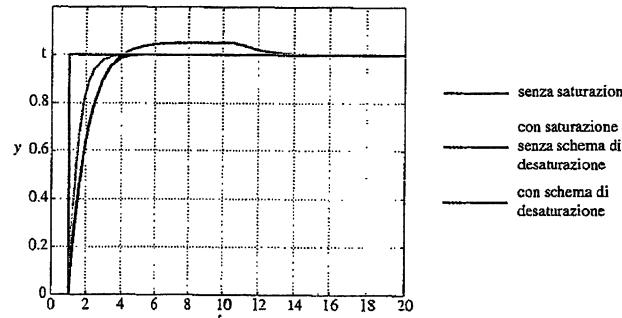
$$C(s) = \frac{1}{1 - \frac{1}{1 + sT_I}} Q(s) = \frac{1 + sT_I}{sT_I} Q(s) = K_P \frac{1 + sT_I}{sT_I} E(s)$$

dove, come di consueto, $C(s)$, $Q(s)$ e $E(s)$ sono le trasformate di Laplace di c , q , e .

Si supponga ora $K_P > 0$, che l'errore si mantenga positivo per un certo periodo di tempo e che la variabile b sia saturata al valore u_M ; allora anche c è uguale a u_M e la variabile z tende al valore costante u_M con la dinamica di un sistema del primo ordine con costante di tempo T_I . Se poi e cambia segno, anche q assume segno negativo e la variabile $b = q + z$ diventa inferiore al limite di saturazione u_M , cioè il sistema torna a funzionare con comportamento lineare.

Esempio 14.5 (Seguito dell'Esempio 14.4) Il regolatore PI con $K_P = K_I = 2$ è stato realizzato secondo lo schema di Figura 14.11. Il transitorio della variabile di uscita fornita da questo schema è confrontato nella Figura 14.12 con i risultati già presentati nella Figura 14.10. Questi andamenti mostrano chiaramente il miglioramento che si può ottenere con un efficace schema di desaturazione.

Figura 14.12
Andamento di u e y nell'Esempio 14.5 in assenza di saturazione, con saturazione e senza schema di desaturazione.



Uno schema generale di desaturazione Come già accennato, il problema del wind-up è sempre presente quando il regolatore contiene un'azione integrale e la variabile di controllo è limitata. È quindi opportuno generalizzare lo schema di Figura 14.11 al caso in cui il regolatore sia descritto dalla generica funzione di trasferimento

$$R(s) = \frac{N_R(s)}{D_R(s)} \quad (14.11)$$

in cui, per la presenza dell'azione integrale, $D_R(0) = 0$. Si consideri ora lo schema di desaturazione rappresentato nella Figura 14.13a, dove il polinomio $\Gamma(s)$ deve essere scelto in modo che $N_R(0)/\Gamma(0) > 0$ e la funzione di trasferimento

$$\Psi(s) = \frac{\Gamma(s) - D_R(s)}{\Gamma(s)} \quad (14.12)$$

sia asintoticamente stabile e strettamente propria. Si noti che tale funzione ha guadagno unitario ($\Psi(0) = 1$).

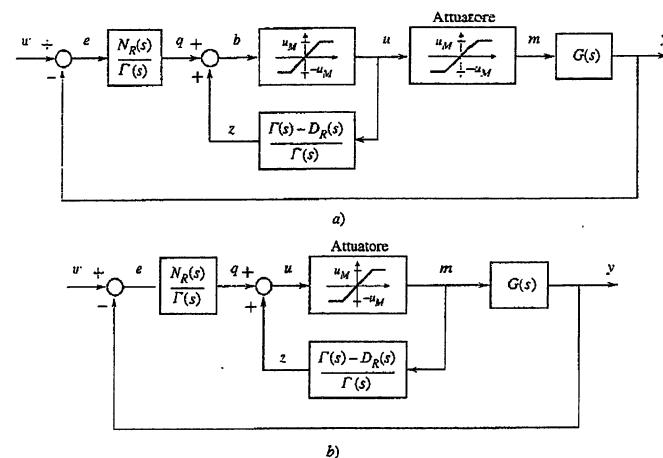


Figura 14.13
Schema di desaturazione per un regolatore generico con azione integrale.

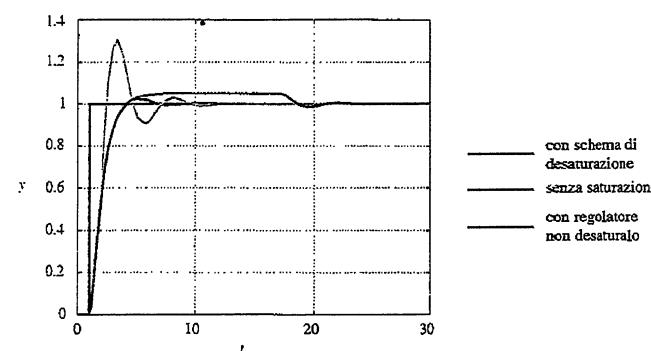
Quando si opera in zona di linearità, dallo schema di Figura 14.13a si deduce che la funzione di trasferimento complessiva tra l'errore e e la variabile u coincide con la (14.11). Se invece l'errore si mantiene di segno costante, per esempio positivo, per un certo periodo di tempo e corrispondentemente l'errore varia lentamente rispetto alla dinamica associata alle radici di $\Gamma(s)$, la variabile q tende ad assumere valori positivi. Si immagini quindi che b sia saturata, per esempio al valore u_M ; allora anche u è uguale a u_M e, poiché $\Psi(0) = 1$, anche la variabile z tende al valore costante u_M con una dinamica funzione di $\Gamma(s)$. Se poi e cambia segno, anche q assume segno negativo e la variabile $b = q + z$ diventa inferiore al limite di saturazione u_M , cioè il sistema torna a funzionare con comportamento lineare. Da questa analisi segue che il rientro in zona lineare di b dovuto al cambio di segno di e è tanto più veloce quanto più rapido è il transitorio dovuto alle radici di $\Gamma(s)$; questa considerazione può essere un valido criterio per la scelta di $\Gamma(s)$.

Si noti infine che, quando l'uscita m dell'attuatore è misurabile, è possibile realizzare lo schema di Figura 14.13b, in cui non è necessario replicare all'interno del regolatore la caratteristica non lineare dell'attuatore.

Esempio 14.6 (Seguito dell'Esempio 14.3) Si consideri ancora il problema di regolare il sistema (14.10) con un regolatore puramente integrale con $K_I = 2$. Lo schema di Figura 14.13b è stato realizzato con $\Gamma(s) = s + 10$, a cui, per la (14.12), corrisponde la funzione di trasferimento $\Psi(s) = 10/(s + 10)$. Il transitorio della variabile di uscita fornita da questo schema è confrontato nella Figura 14.14 con i risultati già presentati nella Figura 14.8. È evidente che con lo schema di desaturazione si conseguono delle prestazioni che, in termini di sovraetensione della risposta allo scalino, sono da considerarsi perfino migliori di quelle fornite dal sistema di controllo in assenza di saturazioni.

Un'analisi rigorosa delle caratteristiche di stabilità e delle prestazioni degli schemi delle Figure 14.11 e 14.13 non è praticabile con le nozioni sin qui introdotte per la presenza degli elementi non lineari. Inoltre, mentre per la stabilità è possibile condurre un'analisi con strumenti più sofisticati (del tipo di quelli presentati nel Paragrafo 16.3), la valutazione teorica delle prestazioni dei diversi schemi di desaturazione è un problema di ardua soluzione.

Figura 14.14
Andamento di w e y
nell'Esempio 14.6 ottenuto
con lo schema di
desaturazione, senza
saturazione e con
regolatore non desaturato.



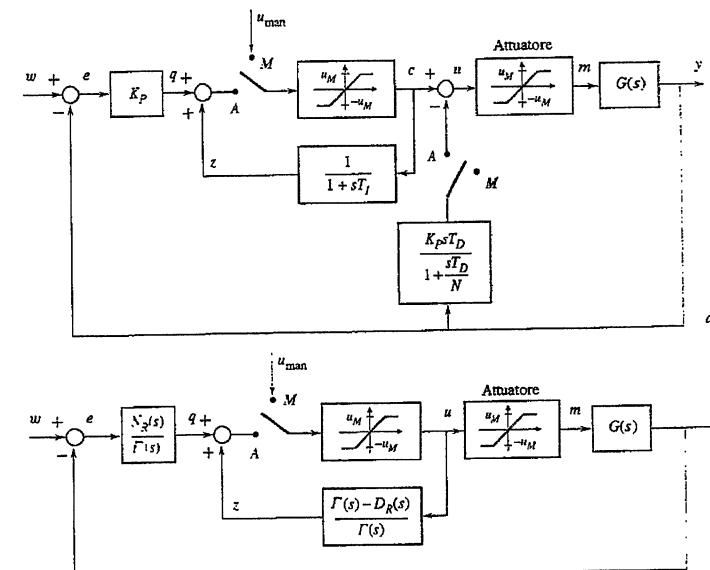
14.3.3 Inserimento "morbido" della regolazione automatica

Come discusso nel Paragrafo 9.2, il progetto di un regolatore descritto dalla funzione di trasferimento $R(s)$ è nella maggior parte dei casi effettuato assumendo che il sistema di controllo operi nell'intorno di un punto di funzionamento nominale. In molte applicazioni industriali, durante la fase di avviamento dell'impianto, questa ipotesi spesso non è verificata (si pensi al problema di regolazione di livello dell'Esempio 9.1 a partire da una condizione iniziale in cui il serbatoio è vuoto). In queste circostanze, l'impiego di un regolatore progettato per condizioni molto diverse può portare a prestazioni del tutto insoddisfacenti. È invece più opportuno controllare inizialmente il sistema con altre tecniche, per esempio con controllo manuale, e quindi commutare sulla regolazione automatica soltanto quando si è raggiunto un intorno del punto di funzionamento nominale. All'atto della commutazione è importante fare in modo che il regolatore sia in grado di fornire instantaneamente un valore della variabile di controllo identico, o almeno molto simile, a quello impiegato fino a quel momento, cioè è necessario che la regolazione automatica sia inserita in modo "morbido", senza che si verifichino brusche variazioni della variabile di controllo. Poiché il regolatore è un sistema dinamico, la sua uscita dipende dallo stato, che quindi deve assumere un valore opportuno.

Questo problema è analogo a quello relativo alla desaturazione dell'azione integrale; anche in quel caso infatti, quando l'attuatore satura, l'anello di regolazione in pratica si apre e il sistema opera come in controllo manuale con $u = \pm u_M$.

Per questo motivo si può fare ricorso a uno schema analogo a quello di Figura 14.11 per un regolatore PID o a quello di Figura 14.13 per un generico regolatore. I corrispondenti schemi sono riportati nella Figura 14.15, dove gli interruttori sono posti su A quando si utilizza la regolazione automatica e su M quando si impiega il controllo manuale e u_{man} è il valore del controllo manuale (con $|u_{man}| \leq |u_M|$). La logica sottostante a questi schemi è quella precedentemente descritta, così come il significato dei simboli: se la commutazione da M ad A avviene con l'errore e che si mantiene circa nullo, si ha $q = 0$, $z = u_{man}$ e, nello schema di Figura 14.15a, l'uscita dell'azione derivativa è anch'essa nulla; pertanto le variabili u e m conservano il valore desiderato u_{man} anche subito dopo la commutazione.

Figura 14.15
Schema di desaturazione
e inserimento "morbido"
della regolazione
automatica:
a) per un PID;
b) per un generico
regolatore.



Esempio 14.7 (Seguito degli Esempi 14.4 e 14.5) Si consideri un classico schema di controllo a retroazione unitaria, con funzione di trasferimento del sistema sotto controllo (14.10), regolatore PI definito da $K_P = K_I = 2$ e si impieghi inizialmente il controllo manuale $u_{man} = 1$. Poiché $G(s)$ ha guadagno 1 e una costante di tempo unitaria, dall'istante $t = 5$ l'uscita assume in pratica il valore di regime $y = 1$. A $t = 10$ si commuta sul controllo automatico con $w = 1$. Lo stato dell'integratore evidentemente non ha in quell'istante il valore richiesto e, benché l'errore e sia nullo, u varia così come l'uscita (Figura 14.16). In questo caso il passaggio al controllo automatico ha indotto una perturbazione indesiderata. Se invece il regolatore è realizzato secondo lo schema di Figura 14.15a, per $t \geq 10$ l'uscita mantiene il valore 1 precedentemente raggiunto.

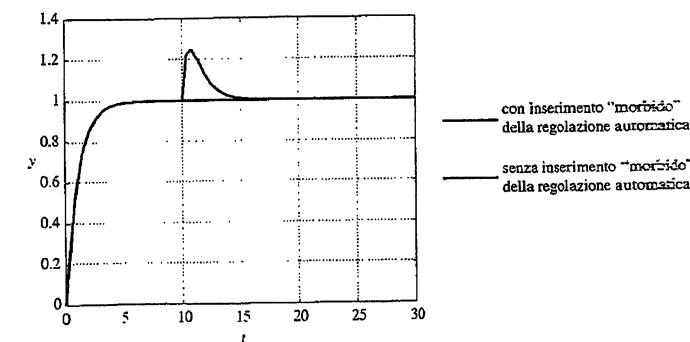


Figura 14.16
Andamento di y
nell'Esempio 14.7 senza
schema di inserimento
"morbido" della
regolazione automatica
e con lo schema proposto.

14.4 Metodi di taratura automatica

Quando la funzione di trasferimento del sistema sotto controllo è nota, i parametri del PID possono essere tarati per mezzo delle tecniche di sintesi presentate nei Capitoli 11 e 12 (si veda l'Esempio 11.4 in cui il regolatore sintetizzato è un PI). Spesso tuttavia la determinazione di un modello a partire dalle leggi che governano i fenomeni in esame può richiedere un eccessivo impegno rispetto all'esigenza di progettare un regolatore in grado di fornire prestazioni accettabili. In questi casi è frequente l'uso di metodi automatici di taratura che consentono di pervenire direttamente alla sintesi del regolatore a partire da specifiche prove effettuate sul processo. Queste prove sono volte, in modo più o meno esplicito, a stimare un modello matematico approssimato del sistema sotto controllo.

Le prime regole di taratura furono introdotte nel 1942 da Ziegler e Nichols. Da allora nella letteratura specializzata sono state proposte molte altre tecniche, alcune delle quali già implementate in controllori industriali commercializzati su larga scala. Esistono oggi molti prodotti che consentono la taratura automatica dei regolatori PID con un modestissimo sforzo richiesto all'operatore per la loro messa in opera.

Gli esperimenti necessari per ricavare il modello approssimato devono essere effettuati, in base alla tecnica impiegata, direttamente sul processo (non regolato) o sul sistema retroazionato con un particolare regolatore. Nel primo caso si parla di metodi in anello aperto, nel secondo di metodi in anello chiuso.

Nel seguito ci si limiterà a presentare i principali metodi di taratura automatica, sforzandosi quando possibile di inquadrarli nell'ambito della teoria del controllo fin qui sviluppata. Inoltre, per semplicità, si assumerà sempre che il sistema sotto controllo sia asintoticamente stabile e con guadagno positivo.

14.4.1 Metodi in anello chiuso

Sarà ora descritta la più nota tecnica di sintesi, il *metodo di Ziegler e Nichols in anello chiuso*, e si mostrerà come sia possibile modificare questo approccio per assegnare il margine di guadagno o di fase del sistema retroazionato.

Metodo di Ziegler e Nichols in anello chiuso Questo metodo richiede di attivare preliminarmente solo l'azione proporzionale, innalzando il coefficiente K_P finché il sistema retroazionato viene portato al limite di stabilità, cioè fino a quando, a fronte di variazioni a scalino imposte al segnale di riferimento w , l'uscita y del sistema è in oscillazione permanente di periodo \bar{T} . Il corrispondente valore \bar{K}_P è chiamato *guadagno critico*. L'andamento qualitativo di w e y in un esperimento per la determinazione di \bar{T} e \bar{K}_P è riportato nella Figura 14.17; si osservi che l'utilizzo di un regolatore puramente proporzionale fa sì che l'errore a transitorio esauria sia in generale diverso da zero e decrescente all'aumentare di K_P .

Una volta determinati \bar{T} e \bar{K}_P , i parametri del regolatore P, PI o PID, vengono tarati secondo la Tabella 14.1. Si noti in particolare che si pone $T_I = 4T_D$, cioè gli zeri del regolatore PID sono coincidenti in posizione $s = -4/T_D$.

Interpretazioni Il guadagno critico \bar{K}_P non è altro che il margine di guadagno k'_m del sistema sotto controllo descritto dalla funzione di trasferimento $G(s)$; pertanto questa tecnica può essere utilizzata unicamente quando il margine di guadagno del processo è finito. Inoltre risulta $\bar{T} = 2\pi/\omega'_\pi$ dove ω'_π è la pulsazione corrispondente al punto A di intersezione del diagramma polare di $G(j\omega)$ con

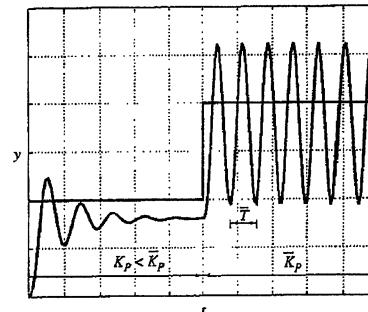


Figura 14.17
Andamento di w e y durante la fase di taratura con il metodo di Ziegler e Nichols in anello chiuso.

	K_P	T_I	T_D
P	$0.5\bar{K}_P$		
PI	$0.45\bar{K}_P$	$0.8\bar{T}$	
PID	$0.6\bar{K}_P$	$0.5\bar{T}$	$0.125\bar{T}$

Tabella 14.1
Regole di taratura di Ziegler e Nichols in anello chiuso.

il semiasse reale negativo (si veda la Figura 14.18). Il metodo proposto consiste quindi nell'individuare le caratteristiche del punto $G(j\omega'_\pi) = -1/k'_m$ della risposta in frequenza del sistema sotto controllo e nell'utilizzare solo questo dato per effettuare la taratura.

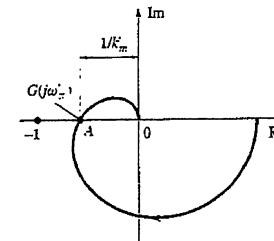


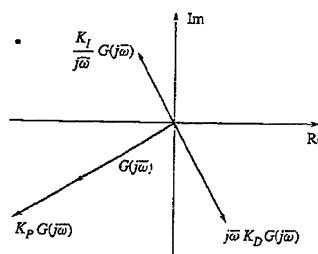
Figura 14.18
Caratteristiche del punto A.

Esempio 14.8 Si è mostrato nell'Esempio 9.6 che per il sistema con funzione di trasferimento

$$G(s) = \frac{1}{(1+s)^2} \quad (14.13)$$

si ha $\omega'_\pi = \sqrt{3}$ e $k'_m = 8$, da cui $\bar{K}_P = 8$ e $\bar{T} \approx 3.63$. Applicando la Tabella 14.1 si sintetizza un PI con $K_P = 3.6$ e $T_I = 2.9$; il sistema di controllo corrispondente ha margine di fase $\varphi_m \approx 14^\circ$, pulsazione critica $\omega_c \approx 1.2$ e margine di guadagno $k_m \approx 1.5$. Le regole precedenti portano quindi alla taratura di un regolatore con prestazioni poco soddisfacenti. Se invece si sintetizza un PID con $N = 10$, si ottiene $K_P = 4.8$, $T_I = 1.815$ e $T_D = 0.454$, a cui corrisponde un sistema retroazionato con un buon margine di guadagno ($k_m \approx 9$), con un margine di fase $\varphi_m \approx 29^\circ$ ancora relativamente basso e con $\omega_c \approx 1.4$. Si osservi inoltre che gli zeri del PID sono coincidenti in posizione $s = -1.1$ e prossimi ai poli di $G(s)$; il regolatore fornito dalle regole automatiche di taratura cancella quasi perfettamente i poli del processo e il progetto risultante è simile a quelli discussi nel Capitolo 11.

Figura 14.19
Spostamento di un punto
del piano complesso per
effetto delle azioni del PID.



Per valutare le prestazioni che ci si possono attendere da un regolatore progettato secondo la Tabella 14.1, si osservi preliminarmente che un regolatore puramente proporzionale $R_P = K_P$ non introduce alcuno sfasamento e la funzione d'anello risultante $K_P G(s)$ è tale che

$$K_P G(j\omega'_\pi) = 0.5 \bar{K}_P G(j\omega'_\pi) = -0.5$$

Il diagramma polare di $K_P G(j\omega)$ attraversa allora l'asse reale in -0.5 e il sistema retroazionato con il regolatore proporzionale ha il margine di guadagno $k_m = 2$.

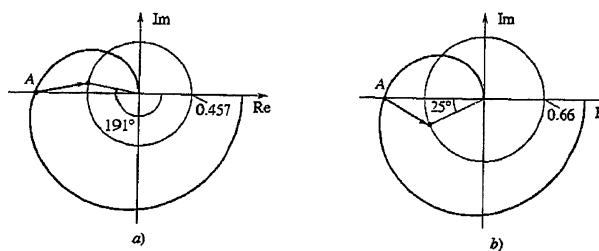
Passando invece a considerare il caso generale, si noti che, dato un arbitrario valore di pulsazione $\bar{\omega}$, è possibile spostare il punto $R_{PID}(j\bar{\omega})G(j\bar{\omega})$ nel piano complesso mediante un'opportuna scelta dei parametri del PID. Infatti, come mostrato nella Figura 14.19, l'azione proporzionale muove $G(j\bar{\omega})$ radialmente rispetto all'origine del piano complesso, mentre le azioni integrale e derivativa consentono movimenti ortogonali introducendo degli sfasamenti di $\pm 90^\circ$.

Questa considerazione consente di determinare come viene spostato il punto A della Figura 14.18 con l'uso di un PI o un PID tarato secondo la Tabella 14.1. In particolare per un PI risulta

$$\begin{aligned} R_{PI}(j\omega'_\pi)G(j\omega'_\pi) &= 0.45 \bar{K}_P \left(1 + \frac{1}{j1.6\pi} \right) \frac{-1}{\bar{K}_P} \simeq \\ &\simeq -0.45 + j0.08 = 0.457 e^{-j(180^\circ + 11^\circ)} \end{aligned}$$

L'impiego di una rete ritardatrice (quale è il PI) introduce quindi alla pulsazione ω'_π uno sfasamento negativo pari a -11° . Inoltre, benché risulti $|R_{PI}(j\omega)G(j\omega)| = 1$ per un valore di ω inferiore a ω'_π , il margine di fase complessivo è in molti casi insufficiente. Lo spostamento del punto A prodotto dall'impiego di un PI è rappresentato nella Figura 14.20a.

Figura 14.20
Spostamento del punto A
nel piano complesso causato da: a) un PI;
b) un PID.



Se invece si utilizza un PID, alla pulsazione ω'_π si ha

$$\begin{aligned} R_{PID}(j\omega'_\pi)G(j\omega'_\pi) &= 0.6 \bar{K}_P \left(1 + \frac{1}{j\pi} + \frac{1}{4} \right) \frac{-1}{\bar{K}_P} \simeq \\ &\simeq -0.6(1 + j0.47) = 0.66 e^{-j(180^\circ - 25^\circ)} \quad (14.14) \end{aligned}$$

Il PID introduce quindi un anticipo a ω'_π di 25° (si veda la Figura 14.20b). Non si può, in questo caso e nel precedente, trarre alcuna conclusione di carattere generale sul margine di fase ottenuto; tuttavia, anche in base al risultato riportato nell'Equazione (14.14), è lecito affermare che se il diagramma polare di $G(j\omega)$ è sufficientemente regolare nell'intorno di ω'_π , la Tabella 14.1 porta a sintetizzare un PID in grado di fornire $30^\circ \div 40^\circ$ di margine di fase.

Prima di esaminare più in dettaglio le possibilità offerte in fase di taratura dalla conoscenza del punto A , vale la pena di sottolineare ancora che il metodo di Ziegler e Nichols richiede di portare il sistema, retroazionato con $R_P = K_P$, al limite di stabilità. In molti casi non è opportuno seguire questo procedimento, per esempio nel caso di sistemi particolarmente delicati o potenzialmente pericolosi. Una diversa tecnica per la determinazione delle caratteristiche del punto A sarà mostrata nel Paragrafo 16.4.6.

Si osservi infine che alcuni sistemi, per esempio del primo o del secondo ordine senza zeri, hanno margine di guadagno infinito; in questi casi la procedura di Ziegler e Nichols non può essere applicata.

Assegnamento del margine di guadagno Poiché l'uso di un PID consente di spostare $R_{PID}(j\omega)G(j\omega)$, per un valore fissato di ω , in un punto arbitrario del piano complesso, è anche possibile assegnare il margine di guadagno k_m o di fase φ_m del sistema retroazionato con funzione d'anello $R_{PID}(j\omega)G(j\omega)$ muovendo il punto A individuato con la procedura di taratura delle regole di Ziegler e Nichols in anello chiuso, come mostrato nella Figura 14.21.

Per assegnare il margine di guadagno k_m , imponendo che la pulsazione ω_π in cui il diagramma polare di $R_{PID}(j\omega)G(j\omega)$ interseca il semiasse reale negativo coincida con ω'_π , è sufficiente utilizzare il regolatore puramente proporzionale

$$R_P = K_P = \frac{\bar{K}_P}{k_m} \quad (14.15)$$

che sposta il punto A in A_1 .

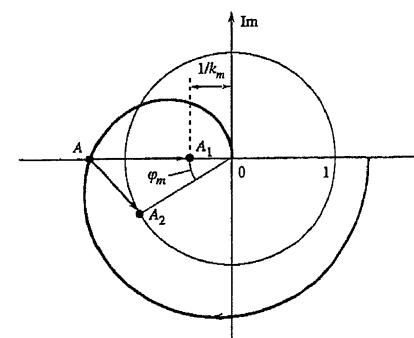


Figura 14.21
Spostamento del punto A in A_1 per l'assegnamento
del margine di guadagno
o in A_2 per l'assegnamento
del margine di fase.

Volendo impiegare anche l'azione integrale per considerazioni relative alle prestazioni statiche, e sempre ponendo $\omega_n = \omega'_n$, è necessario utilizzare anche l'azione derivativa e fare in modo che lo sfasamento prodotto da questi due termini in corrispondenza di ω'_n sia nullo. A tale scopo deve essere verificata la relazione

$$\frac{1}{j\omega'_n T_I} + j\omega'_n T_D = 0$$

ovvero

$$1 - \omega'^2_n T_I T_D = 0 \quad (14.16)$$

Questa relazione non specifica univocamente i valori di T_I e T_D , ma è prassi comune, come visto in precedenza, scegliere

$$T_I = 4T_D \quad (14.17)$$

in modo che gli zeri del PID siano reali e coincidenti. Le relazioni (14.15)-(14.17) definiscono ora univocamente i parametri del PID e risulta

$$T_I = \frac{2}{\omega'_n}, \quad T_D = \frac{1}{2\omega'_n}$$

Esempio 14.9 (Seguito dell'Esempio 14.8) Applicando le regole di taratura (14.15)-(14.17) al sistema con funzione di trasferimento (14.13) per avere il margine di guadagno $k_m = 9$, si ottiene $K_P = 0.89$, $T_I = 1.16$, $T_D = 0.29$, a cui corrisponde il margine di fase $\varphi_m \approx 38^\circ$ e la pulsazione critica $\omega_c \approx 0.56$.

Assegnamento del margine di fase Nel caso in cui si voglia imporre un dato margine di fase φ_m con il vincolo che la pulsazione critica di $R_{PID}(j\omega)G(j\omega)$ coincida con ω'_n , deve risultare

$$\arg(R_{PID}(j\omega'_n)G(j\omega'_n)) = \left(\frac{\varphi_m}{180} - 1\right)\pi \quad (14.18)$$

$$|R_{PID}(j\omega'_n)G(j\omega'_n)| = 1 \quad (14.19)$$

in modo che il punto A identificato con la procedura di Ziegler e Nichols in anello chiuso venga spostato nel punto A_2 della Figura 14.21. Per soddisfare le (14.18), (14.19) è necessario utilizzare l'azione derivativa del PID per poter disporre dell'antico di fase necessario in corrispondenza di ω'_n . Qualunque sia $K_P > 0$ si ha $\arg(K_P G(j\omega'_n)) = -\pi$ e allora dalle (14.18), (14.19) si ottiene

$$\begin{aligned} \arg\left(1 + \frac{1}{j\omega'_n T_I} + j\omega'_n T_D\right) &= \varphi_m \frac{\pi}{180} \\ \left|K_P\left(1 + \frac{1}{j\omega'_n T_I} + j\omega'_n T_D\right)\frac{-1}{K_P}\right| &= 1 \end{aligned}$$

cioè

$$\omega'_n T_D - \frac{1}{\omega'_n T_I} = \tan(\varphi_m) \quad (14.20)$$

$$K_P = \bar{K}_P \cos(\varphi_m) \quad (14.21)$$

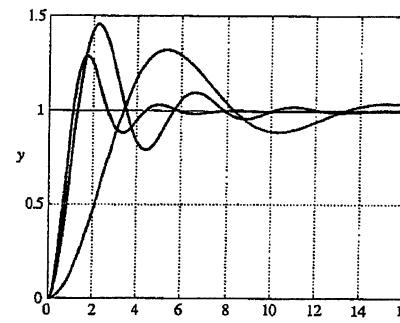


Figura 14.22
Risposta allo scalino del sistema retroazionato nell'Esempio 14.10 con il PID sintetizzato mediante la Tabella 14.1, assegnando il margine di guadagno e il margine di fase.

Le (14.20), (14.21), insieme ancora alla (14.17), definiscono univocamente i parametri del PID cercato.

Come già osservato, questo procedimento per l'assegnamento del margine di fase fa sì che la pulsazione critica di $R_{PID}(j\omega)G(j\omega)$ coincida con ω'_n , valore che in alcuni casi può risultare inadeguato.

Esempio 14.10 (Seguito degli Esempi 14.8, 14.9) Applicando le regole di taratura (14.17), (14.20), (14.21) al sistema (14.13) per imporre $\varphi_m = 45^\circ$, si ottiene $K_P = 5.66$, $T_I = 2.79$, $T_D = 0.697$. Impiegando un PID reale con questi valori dei parametri e con $N = 10$, risulta $\varphi_m \approx 39^\circ$, $\omega_c \approx 1.8$ e $k_m \approx 5.69$; la riduzione del margine di fase è ovviamente dovuta al polo in alta frequenza.

Sempre con riferimento al sistema (14.13), nella Figura 14.22 sono confrontate le risposte allo scalino dei sistemi retroazionati nei tre casi in cui il regolatore PID, con $N = 10$, è sintetizzato con le regole della Tabella 14.1 (I caso), assegnando il margine di guadagno $k_m = 9$ (II caso) e il margine di fase $\varphi_m = 45^\circ$ (III caso). Nonostante i valori di φ_m ottenuti nel II e III caso siano molto simili, il PID sintetizzato con il metodo basato sull'assegnamento del margine di fase consente di ottenere una velocità di risposta decisamente superiore a quella fornita dal PID tarato con la tecnica di assegnamento del margine di guadagno. Questo risultato è giustificato dal diverso valore di ω_c ottenuto nei due progetti.

In conclusione, dalle (14.15)-(14.17) e dalle (14.17), (14.20), (14.21) si possono ottenere le regole di taratura, riportate nella Tabella 14.2, di un regolatore PID che assegna il margine di guadagno k_m o il margine di fase φ_m a partire dai parametri caratteristici del processo \bar{K}_P e \bar{T} ricavati con la prova sperimentale richiesta dal metodo di Ziegler e Nichols in anello chiuso. Queste regole, a differenza di quelle della Tabella 14.1, consentono di specificare un parametro di progetto (k_m o φ_m).

	K_P	T_I	T_D
Assegnamento di k_m	$\frac{\bar{K}_P}{k_m}$	$\frac{\bar{T}}{\pi}$	$\frac{\bar{T}}{4\pi}$
Assegnamento di φ_m	$\bar{K}_P \cos(\varphi_m)$	$\frac{\bar{T}}{\pi} \frac{1 + \sin(\varphi_m)}{\cos(\varphi_m)}$	$\frac{\bar{T}}{4\pi} \frac{1 + \sin(\varphi_m)}{\cos(\varphi_m)}$

Tabella 14.2
Taratura del PID con assegnamento del margine di guadagno o di fase.

14.4.2 Metodi in anello aperto

Questi metodi di taratura dei parametri dei regolatori PID, sviluppati per sistemi con risposta allo scalino essenzialmente non oscillante, si basano sull'impiego di

un modello approssimato del sistema da controllare descritto dalla funzione di trasferimento

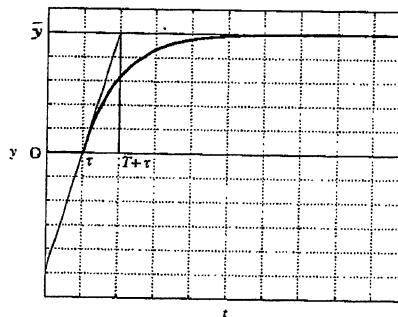
$$G_a(s) = \frac{\mu}{1+Ts} e^{-\tau s} \quad (14.22)$$

in cui il parametro τ è comunemente chiamato *ritardo equivalente*, mentre T è definita *costante di tempo equivalente*.

Prima di presentare i principali metodi di sintesi del PID per sistemi descritti dalla funzione di trasferimento $G_a(s)$, sono ora illustrate le tecniche più comunemente impiegate per determinare i parametri del modello approssimato (14.22) a partire dalla risposta allo scalino sperimentale del processo.

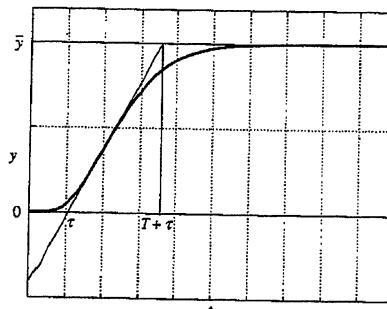
Metodo della tangente I due metodi più comunemente utilizzati per ricavare i parametri μ , τ e T del modello approssimato (14.22) a partire dalla risposta allo scalino sperimentale del processo sono il metodo della tangente e il metodo delle aree. Per introdurre il primo metodo, si consideri preliminarmente la risposta allo scalino, riportata nella Figura 14.23, del sistema con funzione di trasferimento (14.22). Da questa è possibile ricavare il valore del guadagno μ come rapporto tra il valore di regime \bar{y} dell'uscita e l'ampiezza \bar{u} dello scalino d'ingresso. Inoltre, il punto di massima pendenza della curva corrisponde all'istante $t = \tau$ e la tangente alla curva in questo punto interseca l'asintoto corrispondente al valore di regime dell'uscita all'istante $t = T + \tau$.

Figura 14.23
Parametri caratteristici della risposta allo scalino del sistema (14.22) impiegati nel metodo della tangente.



Supponendo ora che la risposta allo scalino sperimentale sia del tipo mostrato nella Figura 14.24, dal grafico (o per via numerica) si possono valutare:

Figura 14.24
Determinazione dei parametri del modello (14.22) a partire dalla risposta allo scalino sperimentale.



- il guadagno μ , dato da \bar{y}/\bar{u} ;
- la tangente nel punto di massima pendenza;
- i parametri τ e T .

Esempio 14.11 (Seguito degli Esempi 14.8, 14.9, 14.10) Il metodo della tangente applicato alla risposta allo scalino del sistema (14.13) porta alla determinazione del seguente modello approssimato

$$G_a(s) = \frac{e^{-0.8s}}{1+3.7s} \quad (14.23)$$

Nella Figura 14.25 sono riportate le risposte allo scalino del sistema (14.13) e del modello (14.23), mentre i rispettivi diagrammi di Bode sono confrontati nella Figura 14.26.

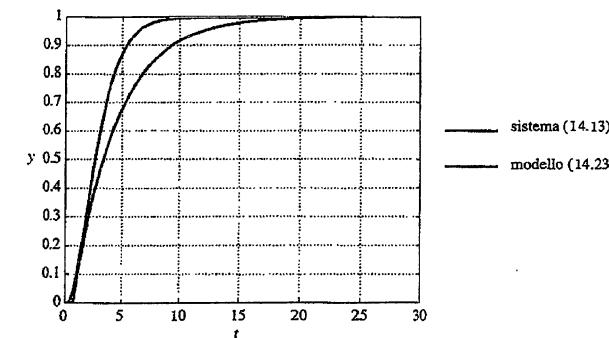


Figura 14.25
Risposta allo scalino del sistema (14.13) e del modello (14.23).

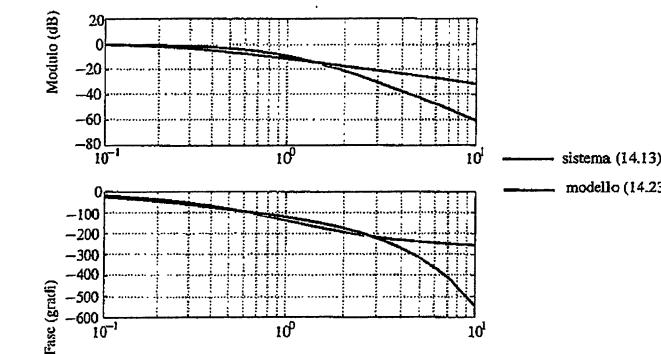
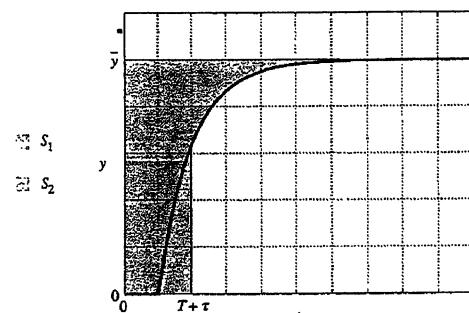


Figura 14.26
Diagrammi di Bode del sistema (14.13) e del modello (14.23).

Si osservi che l'applicazione del metodo a partire da una risposta allo scalino ricavata sperimentalmente può essere resa problematica dalla presenza di errori di misura. In questi casi, infatti, l'individuazione del punto di massima pendenza della curva e la valutazione della tangente sono difficili, soprattutto se effettuate numericamente, rendendo ancora minore la significatività del modello ricavato.

Metodo delle aree Si consideri ancora la risposta a uno scalino di ampiezza \bar{u} del sistema descritto dalla funzione di trasferimento (14.22). Come nel metodo

Figura 14.27
Parametri caratteristici della risposta allo scalino del sistema (14.22) impiegati nel metodo delle aree.



della tangente, il guadagno del sistema si può determinare direttamente come $\mu = \bar{y}/\bar{u}$. Inoltre, l'area S_1 compresa tra la risposta e l'asintoto della curva (si veda la Figura 14.27) è data da

$$S_1 = \mu \bar{u} \tau + \int_0^\infty \mu \bar{u} e^{-t/T} dt = \mu \bar{u} (\tau + T) = \bar{y} (\tau + T)$$

Dalla sua valutazione si ricava quindi il valore di $T + \tau$. Considerando ora l'area S_2 compresa tra la curva, l'asse dei tempi e una retta verticale passante per il punto di ascissa $T + \tau$, si ha

$$S_2 = \int_0^{T+\tau} \mu \bar{u} (1 - e^{-t/T}) dt = \frac{\mu \bar{u} T}{e} = \frac{\bar{y} T}{e}$$

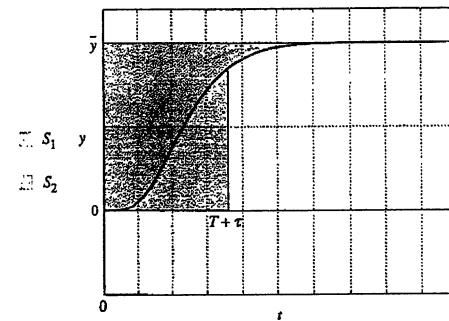
da cui

$$T = \frac{e S_2}{\bar{y}}$$

Data la risposta allo scalino del processo di cui si vuole determinare il modello approssimato (14.22) (si veda per esempio la risposta riportata nella Figura 14.28), il metodo delle aree consiste quindi nel ricavare:

- il guadagno $\mu = \bar{y}/\bar{u}$;
- l'area S_1 ;
- il valore $T + \tau = S_1/\bar{y}$;
- l'area S_2 ;
- i parametri $T = e S_2/\bar{y}$ e $\tau = (S_1 - \bar{y} T)/\bar{y}$.

Figura 14.28
Determinazione dei parametri del modello (14.22) a partire dalla risposta allo scalino sperimentale.



Si osservi che, soprattutto in presenza di segnali rumorosi, il calcolo numerico di un'area è meno critico della valutazione della tangente a una curva. Sotto questo aspetto il metodo delle aree si può ritenere più affidabile del metodo della tangente.

Esempio 14.12 (Seguito degli Esempi 14.8, 14.9, 14.10, 14.11) Applicando il metodo delle aree alla risposta allo scalino del sistema (14.13) si ottiene il seguente modello approssimato

$$G_d(s) = \frac{e^{-1.17s}}{1 + 1.83s} \quad (14.24)$$

La Figura 14.29 illustra le risposte allo scalino del sistema (14.13) e del modello (14.24), mentre i rispettivi diagrammi di Bode sono rappresentati nella Figura 14.30.

Figura 14.29
Risposta allo scalino del sistema (14.13) e del modello (14.24).

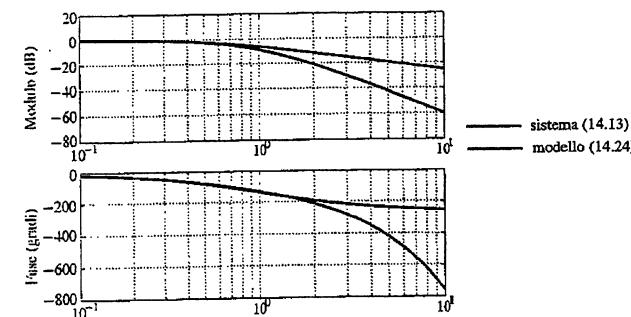
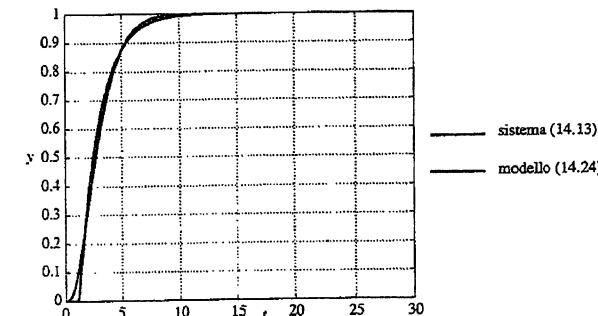


Figura 14.30
Diagrammi di Bode del sistema (14.13) e del modello (14.24).

Metodo di Ziegler e Nichols in anello aperto A partire dal modello descritto dalla (14.22), il *metodo di Ziegler e Nichols in anello aperto* consiste nel progettare un regolatore P, PI o PID secondo la Tabella 14.3.

Si osservi che, anche in questo caso, nella taratura del PID i due zeri sono coincidenti in posizione $s = -1/\tau$.

Tabella 14.3
Regole di taratura di Ziegler e Nichols in anello aperto.

	K_P	T_I	T_D
P	$\frac{T}{\mu\tau}$		
PI	$\frac{0.9T}{\mu\tau}$	3τ	
PID	$\frac{1.2T}{\mu\tau}$	2τ	0.5τ

Esempio 14.13 (Seguito degli Esempi 14.11, 14.12) Si consideri ancora il sistema (14.13) e i modelli approssimati (14.23) e (14.24) ricavati con i metodi della tangente e delle aree. Applicando a questi modelli le regole di taratura della Tabella 14.3 si ottengono i risultati riportati nella Tabella 14.4.

Tabella 14.4
Prestazioni del metodo di Ziegler e Nichols in anello aperto.

	K_P	T_I	T_D	k_m	φ_m	ω_c
PI - metodo della tangente	4.16	2.4	—	1.4	5.2°	1.3
PID - metodo della tangente	5.55	1.6	0.4	17	21°	1.5
PI - metodo delle aree	1.4	3.51	—	12	63°	0.58
PID - metodo delle aree	1.88	2.34	0.58	26.7	62°	0.74

I risultati di questo esempio sono di carattere generale: il metodo della tangente porta spesso alla determinazione di un modello approssimato di validità limitata, che a sua volta produce un sistema di controllo con bassi margini di fase e guadagno.

Altre tecniche di taratura La grande diffusione dei regolatori PID in ambito industriale, la semplicità del modello (14.22) facilmente ricavabile con prove d'impianto, e le prestazioni non sempre soddisfacenti fornite dalle regole riportate nella Tabella 14.3, hanno fatto sì che negli anni siano stati proposti diversi criteri per la taratura dei parametri K_P , T_I e T_D . Senza voler qui spiegare nel dettaglio i motivi che hanno portato alla loro definizione, e rimandando per questo ai testi citati in bibliografia, si riportano nel seguito i metodi di Cohen e Coon e IMC, due tra i criteri più noti e utilizzati nelle applicazioni.

Il *metodo di Cohen e Coon* consiste nell'utilizzare le relazioni riportate nella Tabella 14.5, ricavate imponendo un rapporto di smorzamento pari a 0.25 tra due picchi successivi della risposta a un disturbo a scalino sulla variabile di controllo.

Tabella 14.5
Regole di taratura di Cohen e Coon.

	K_P	T_I	T_D
P	$\frac{3T + \tau}{3\mu\tau}$		
PI	$\frac{10.8T + \tau}{12\mu\tau}$	$\frac{30T + 3\tau}{\tau(9T + 20\tau)}$	
PID	$\frac{16T + 3\tau}{12\mu\tau}$	$\frac{32T + 6\tau}{\tau(12\tau)}$	$\frac{4T\tau}{11T + 2\tau}$

La tecnica di sintesi chiamata *Internal Model Control*, o IMC, suggerisce di utilizzare le regole di taratura sintetizzate nella Tabella 14.6. Si noti che nelle regole IMC compare anche il termine T_f , che rappresenta un ulteriore parametro positivo di progetto. Il valore di T_f può essere scelto liberamente in fase di sintesi osservando che a un suo aumento corrisponde una riduzione della banda passante del sistema in anello chiuso e un aumento dei margini di fase e di guadagno.

Tabella 14.6
Regole di taratura IMC.

	K_P	T_I	T_D
PI	$\frac{T}{\mu(\tau + T_f)}$	T	
PID	$\frac{T + 0.5\tau}{\mu(\tau + T_f)}$	$T + 0.5\tau$	$\frac{0.5\tau T}{0.5\tau + T}$

In generale si può affermare che il metodo di Cohen e Coon porta a una taratura simile a quella fornita dalle regole di Ziegler e Nichols di Tabella 14.3, mentre la taratura IMC produce dei sistemi di controllo più moderati.

Esempio 14.14 (Seguito degli Esempi 14.8-14.12) L'impiego congiunto del modello (14.24) e dei metodi di taratura di Cohen e Coon e IMC consente di tarare dei PI che, applicati al sistema (14.13), producono i risultati riportati nella Tabella 14.7. Da questi risultati si può notare l'utilità nel metodo IMC di poter disporre del parametro libero di taratura T_f : all'aumentare del suo valore il sistema in anello chiuso diventa sempre più lento, ma con migliori caratteristiche di robustezza in termini di margine di fase e di guadagno.

Tabella 14.7
Prestazioni dei metodi Cohen e Coon e IMC nell'Esempio 14.14.

	φ_m	k_m	ω_c
Cohen e Coon	31°	7.9	0.74
IMC ($T_f = 0.4$)	45°	10.8	0.6
IMC ($T_f = 0.8$)	53°	12.8	0.5
IMC ($T_f = 1.2$)	59°	14.4	0.42

Metodi di ottimizzazione Alcuni metodi di taratura consistono nel determinare i parametri del regolatore in modo da minimizzare opportune funzioni obiettivo caratterizzanti le risposte del sistema in anello chiuso a fronte di andamenti a scalino del segnale di riferimento o dei disturbi. Tra i vari funzionali normalmente considerati, i più comuni sono i seguenti

$$IAE = \int_0^\infty |e(t)| dt \quad ITAE = \int_0^\infty t |e(t)| dt$$

$$ISE = \int_0^\infty e^2(t) dt \quad ISTE = \int_0^\infty t^2 e^2(t) dt$$

Il funzionale *IAE* (*Integral Absolute Error*) penalizza il modulo dell'errore, mentre nel funzionale *ITAE* (*Integral Time Absolute Error*) viene poco penalizzato il modulo dell'errore nei primi istanti del transitorio. Analogamente, il criterio *ISE* (*Integral Square Error*) penalizza l'integrale del quadrato dell'errore, mentre il funzionale *ISTE* (*Integral Square Time Error*) può essere utilizzato quando siano accettabili errori anche elevati nei primi istanti della risposta.

Naturalmente è difficile determinare analiticamente i valori ottimi dei parametri per una generica funzione di trasferimento $G(s)$ del processo, ma è necessario impiegare caso per caso metodi di ottimizzazione di tipo evolutivo. Tuttavia, facendo riferimento al modello (14.22), ottenibile come di consueto con il metodo delle aree o della tangente, e a risposte a scalino del segnale di riferimento, sono state proposte regole empiriche di taratura di regolatori PI o PID ottenute interpolando i risultati di specifiche prove di ottimizzazione. Per esempio, per il funzionale *ITAE* si possono impiegare le formule riportate nella Tabella 14.8.

Tabella 14.8
Regole di taratura ITAE.

	K_P	T_I	T_D
PI_{ITAE}	$\frac{0.586}{\mu} \left(\frac{T}{\tau}\right)^{0.916}$	T^2	$1.03T - 0.165\tau$
PID_{ITAE}	$\frac{0.965}{\mu} \left(\frac{T}{\tau}\right)^{0.855}$	T^2	$0.308T \left(\frac{\tau}{T}\right)^{0.929}$

14.5 Conclusioni

L'ampio spazio dedicato ai regolatori PID è giustificato dal loro vasto impiego in ambito industriale. Inoltre, la possibilità di tarare i loro parametri con semplici procedimenti, che non richiedono la conoscenza di un modello matematico preciso del processo, li rende particolarmente adatti alla soluzione di problemi in cui è opportuno abbreviare i tempi del progetto del sistema di controllo e della sua messa in servizio. Le regole presentate nel Paragrafo 14.4 sono spesso impiegate in una prima fase, e la taratura dei parametri del PID è completata successivamente con aggiustamenti sull'impianto reale.

Nella realizzazione pratica dei PID in particolare, e in generale dei regolatori, hanno anche grandissima importanza le considerazioni riportate nel Paragrafo 14.3. Per un'implementazione accettabile di qualsiasi regolatore è indispensabile dare una soluzione soddisfacente ai problemi di desaturazione della variabile di controllo e della commutazione dalla configurazione di controllo manuale a quella di controllo automatico. Inoltre, è importante utilizzare correttamente l'azione derivativa presente nel regolatore per migliorare la velocità di risposta del sistema retroazionato senza produrre eccessive sollecitazioni della variabile di controllo.

Esercizi

Esercizio 14.1 Si consideri un sistema con funzione di trasferimento

$$G(s) = \frac{1+8s}{(1+s)(1+100s)}$$

e i due regolatori PI seguenti

$$R_{PI_1}(s) = K_{P_1} \frac{1+s}{s} \quad , \quad R_{PI_2}(s) = \frac{K_{P_2}}{100} \frac{1+100s}{s}$$

Si discutano, al variare di K_{P_1} e K_{P_2} :

- le prestazioni ottenibili per il sistema retroazionato in termini di margine di fase e di pulsazione critica;
- la configurazione di poli e zeri del sistema retroazionato (si suggerisce di impiegare il metodo del luogo delle radici);
- l'andamento qualitativo della risposta a una variazione a scalino del segnale di riferimento.

Esercizio 14.2 Si progetti un regolatore PI per un sistema descritto da

$$G(s) = \frac{\mu}{s(1+Ts)} \quad , \quad T > 0$$

in modo da ottenere un dato margine di fase $\bar{\phi}_m$.

Esercizio 14.3 Si progetti un regolatore PID stabilizzante per un sistema descritto da

$$G(s) = \frac{\mu}{s(1+Ts)} \quad , \quad T < 0$$

(si suggerisce di impiegare il metodo del luogo delle radici).

Esercizio 14.4 Con riferimento agli schemi (a) e (b) della Figura 14.3, si valutino le funzioni di trasferimento tra gli ingressi w e d e le uscite y e u nei due casi, ricordando che le azioni proporzionale e integrale sono rappresentate dalla funzione di trasferimento (14.6) mentre l'azione derivativa ha la funzione di trasferimento (14.4). A partire dalle funzioni di trasferimento trovate si verifichi che:

- le proprietà di stabilità dei due schemi sono identiche;
- la funzione di trasferimento tra d e y è la stessa nei due casi e coincide con la funzione di sensitività $S(s)$;
- la funzione di trasferimento tra d e u è la stessa nei due casi e coincide, a meno del segno, con la funzione di sensitività del controllo $Q(s)$;
- la funzione di trasferimento tra w e y ha in entrambi i casi guadagno unitario e la funzione di sensitività ha uno zero nell'origine;
- cambiano gli zeri delle funzioni di trasferimento tra w e le variabili y e u .

Si discutano le implicazioni delle osservazioni precedenti in termini di risposta al disturbo d , risposta al segnale di riferimento w , errore a transitorio esaurito a fronte di segnali di riferimento o disturbi costanti.

Esercizio 14.5 Si consideri il particolare regolatore PID descritto da

$$U(s) = K_P E_P(s) + \frac{K_I}{s} E(s) - K_D s E_D(s)$$

dove $E(s) = W(s) - Y(s)$, mentre

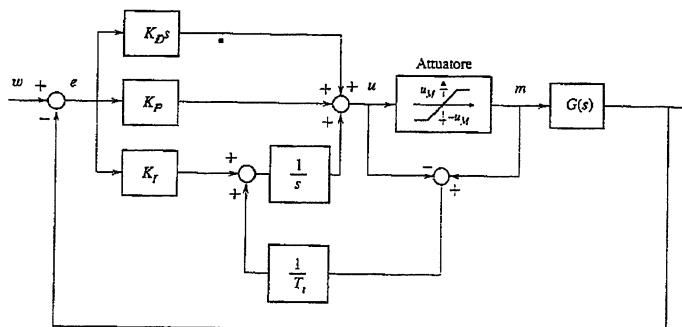
$$E_P(s) = \alpha W(s) - Y(s) \quad , \quad E_D(s) = \beta W(s) - Y(s) \quad (14.25)$$

in cui $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$. Si verifichi che:

- per ogni coppia (α, β) ammissibile le proprietà di stabilità del sistema retroazionato e le sue prestazioni statiche non cambiano rispetto al caso in cui $\alpha = \beta = 1$;
- per $\alpha = 1$ e $\beta = 0$ si ha lo schema di Figura 14.3b;
- per $\alpha = \beta = 0$ sia l'azione derivativa che quella proporzionale sono esercitate unicamente sulla variabile di uscita;
- per $\beta \neq 1$ vengono modificate le caratteristiche della risposta al segnale di riferimento (si discuta l'effetto di diverse scelte di β).

Esercizio 14.6 Si consideri lo schema di realizzazione dei regolatori PID con desaturazione riportato nella Figura 14.31, in cui T_t è la cosiddetta costante di tempo d'inseguimento, e se ne analizzino le caratteristiche di funzionamento.

Figura 14.31
Schema di desaturazione
dei PID.



Esercizio 14.7 Le prestazioni fornite da un regolatore PI tarato secondo la Tabella 14.3 e impiegato per il controllo del modello (14.22) con $\tau/T \in [0.1, 1]$, possono essere determinate in modo approssimato mediante le relazioni seguenti

$$\varphi_m = 78^\circ \frac{\tau}{T} + 15^\circ \quad (14.26)$$

$$k_m = 0.9 \frac{\tau}{T} + 1.5 \quad (14.27)$$

Si verifichi per diversi valori di T e τ il grado di approssimazione delle (14.26) e (14.27).

Esercizio 14.8 Dato un sistema descritto dalla funzione di trasferimento (14.22) con μ, T positivi, si sintetizzi un regolatore PI che garantisca un margine di fase maggiore o uguale a un dato valore $\bar{\varphi}_m$ e una pulsazione critica non superiore a un dato valore $\bar{\omega}_c$.

Esercizio 14.9 Si valuti l'opportunità di utilizzare anche l'azione derivativa del PID per un sistema descritto dalla funzione di trasferimento (14.22).

Esercizio 14.10 Dato il sistema dell'Esempio 14.8 e il modello approssimato ricavato nell'Esempio 14.11 con il metodo della tangente, si progetti un regolatore PI con le regole di taratura delle Tabelle 14.5, 14.6 e se ne valutino le prestazioni.

Schemi di controllo avanzati

15

15.1 Introduzione

Nei precedenti capitoli sono state presentate le principali tecniche di analisi e di sintesi del classico sistema di controllo in retroazione in cui il regolatore ha come ingresso l'errore, eventualmente corrotto da un disturbo di misura, e produce come uscita la variabile di controllo, ingresso del processo. Questo schema è certamente molto utilizzato e consente di risolvere un gran numero di problemi. Tuttavia esso non è l'unico possibile e spesso viene inserito in strutture più complesse, realizzate per migliorare le prestazioni dell'intero sistema di controllo e per sfruttare al meglio tutta l'informazione disponibile sul processo, come per esempio la misura dei disturbi.

In questo capitolo si presenteranno alcuni schemi di controllo frequentemente impiegati per la soluzione di problemi specifici. In particolare, saranno trattati i seguenti argomenti:

- il progetto di regolatori in anello aperto, o compensatori del segnale di riferimento e dei disturbi misurabili, spesso impiegati congiuntamente agli schemi in retroazione per migliorare le prestazioni complessive;
- lo schema a predittore di Smith per sistemi affetti da ritardo di tempo, utilizzato per rimuovere la limitazione imposta al massimo valore della pulsazione critica dalla presenza del ritardo;
- gli schemi di controllo in cascata, largamente impiegati per la stabilizzazione di sistemi instabili o quando il sistema sotto controllo è composto dalla serie di due sottosistemi, uno con dinamica "veloce", quale quella che tipicamente caratterizza gli attuatori, e uno con dinamica "lenta" e presenza di elementi a fase non minima, usualmente associata al processo;
- il progetto di sistemi di controllo per sistemi MIMO basato sull'impiego di regolatori di disaccoppiamento, che consentono di estendere, per quanto possibile, le tecniche sviluppate per sistemi SISO, oppure basato su schemi di controllo decentralizzato.

15.2 Regolatori in anello aperto

Nel Paragrafo 1.5.1 si è mostrato che l'uso di elementi in anello aperto può migliorare le prestazioni di un sistema di controllo in anello chiuso, per esempio per compensare l'effetto di disturbi misurabili. Inoltre, a volte è opportuno filtrare

il segnale di riferimento w mediante *regolatori in anello aperto* o *compensatori*, prima che esso venga inviato allo schema di controllo in retroazione. In questo modo si possono modificare le caratteristiche della funzione di trasferimento tra w e l'uscita y per conseguire determinate prestazioni, statiche o dinamiche.

Per queste ragioni spesso sono realizzati schemi di controllo in cui sono presenti sia un'azione di regolazione in anello aperto sia una in anello chiuso, combinando così gli effetti positivi di entrambe.

Nel seguito del paragrafo si analizzeranno dapprima schemi in cui compaiono regolatori in anello aperto alimentati dal segnale di riferimento; successivamente si presenteranno le modalità di progetto di compensatori in grado di eliminare o ridurre l'effetto dei disturbi misurabili.

15.2.1 Prefiltraggio del segnale di riferimento

Uno schema di controllo in cui il segnale di riferimento w è posto all'ingresso di un opportuno compensatore in anello aperto è quello riportato nella Figura 15.1, dove $C(s)$ è la funzione di trasferimento del compensatore e si suppone che il sistema retroazionato sia asintoticamente stabile. Il progetto di $C(s)$ può essere effettuato secondo modalità e con fini diversi, con l'unico e ovvio vincolo che esso sia un sistema asintoticamente stabile e proprio.

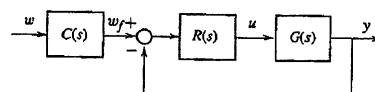


Figura 15.1
Schema di controllo con prefiltro del segnale di riferimento.

A questo proposito si osservi preliminarmente che, come già discusso nel Paragrafo 1.3.4 (si veda la Figura 1.7) e nel Paragrafo 9.2, w deve normalmente essere filtrato da un trasduttore (l'elemento T^0 nella Figura 9.1) per produrre un segnale da confrontare con la misura dell'uscita e generare così l'errore; in questo caso $C(s)$ rappresenta quindi il trasduttore.

Inoltre, $C(s)$ può essere scelto anche con altre finalità. A questo riguardo, si noti che

$$Y(s) = \frac{R(s)G(s)}{1 + R(s)G(s)}C(s)W(s) = F(s)C(s)W(s) \quad (15.1)$$

$$U(s) = \frac{R(s)}{1 + R(s)G(s)}C(s)W(s) = Q(s)C(s)W(s) \quad (15.2)$$

Il compensatore $C(s)$ consente quindi di modificare:

- la funzione di trasferimento tra w e y , per esempio variando il guadagno e restringendo o allargando la banda passante, cioè modificando la velocità di risposta del sistema di controllo;
- la funzione di trasferimento tra w e u , in modo da evitare brusche sollecitazioni della variabile di controllo.

Compensazione statica Quando il segnale di riferimento è uno scalino, affinché a transitorio esaurito l'uscita y coincida con w , la funzione di trasferimento $Y(s)/W(s)$ deve avere guadagno unitario. Dalla (15.1) segue che, se $F(0) \neq 1$, cioè se $R(s)G(s)$ non contiene almeno un'azione integrale, è possibile assicurare la condizione sul guadagno per mezzo del compensatore statico

$$C(s) = \mu_C = F(0)^{-1} \quad (15.3)$$

Si osservi tuttavia che la presenza di una o più azioni integrali nella funzione d'anello $R(s)G(s)$ garantisce comunque che la condizione $F(0) = 1$ sia verificata anche a fronte di variazioni dei parametri del processo o di incertezze sul modello (purché il sistema di controllo resti asintoticamente stabile); in questo caso si ha cioè la cosiddetta regolazione robusta a zero dell'errore (si veda al riguardo il Paragrafo 10.6.1). Se invece il guadagno di $Y(s)/W(s)$ è reso unitario mediante l'impiego del compensatore $C(s)$, eventuali variazioni dei parametri di $G(s)$, e quindi di $F(s)$, dovrebbero comportare una modifica anche di $C(s)$ per soddisfare la (15.3). Poiché di norma il compensatore, una volta determinato, non viene più modificato, la regolazione a zero dell'errore ottenuta tramite l'impiego di $C(s)$ non è robusta.

Filtraggio passa-basso Il compensatore $C(s)$ viene molto spesso scelto come un filtro di tipo passa-basso con lo scopo di far sì che, a fronte di variazioni a scalino del segnale di riferimento w , il sistema retroazionato sia sollecitato dal segnale w_f di Figura 15.1, caratterizzato da una dinamica meno veloce. Tipicamente w_f ha un andamento di tipo esponenziale ottenuto mediante l'uso di una funzione di trasferimento $C(s)$ del primo ordine. In questo modo gli organi di attuazione vengono sottoposti a sollecitazioni inferiori, è possibile soddisfare più facilmente il requisito di moderazione della variabile di controllo e si riducono eventuali problemi dovuti alla presenza di saturazioni (si veda a questo proposito il Paragrafo 14.3.2), mantenendo per il sistema di controllo un comportamento di tipo lineare.

Naturalmente, poiché come mostrato dalla (15.1) la funzione di trasferimento complessiva tra w e y è costituita dalla serie di $C(s)$ e della funzione $F(s)$ del sistema retroazionato, quando l'estremo superiore della banda passante di $C(s)$ è inferiore alla pulsazione critica associata a $R(s)G(s)$, la risposta dell'uscita y al segnale di riferimento w risulta rallentata rispetto alla situazione in cui il compensatore è assente.

Filtraggio passa-alto In alcune circostanze il compensatore $C(s)$ di Figura 15.1 viene progettato per ragioni opposte a quelle discusse nel punto precedente, ovvero con lo scopo di rendere più rapida la risposta del sistema di controllo a fronte di variazioni del segnale di riferimento. Si consideri per esempio il caso in cui nel progetto del regolatore in retroazione $R(s)$ sia necessario tener conto di un'elevata incertezza sul valore assunto da $G(j\omega)$ a pulsazioni $\omega > \bar{\omega}$. Per avere un sistema di controllo con elevata robustezza, è allora opportuno determinare $R(s)$ in modo che la risposta in frequenza associata a $R(s)G(s)$ garantisca una sufficiente attenuazione per $\omega > \bar{\omega}$, ovvero è necessario fare sì che la pulsazione critica sia minore di $\bar{\omega}$, limitando così la velocità di risposta del sistema retroazionato. In questo caso è possibile progettare $C(s)$ come una rete anticipatrice, con lo zero prossimo alla pulsazione critica e il polo a pulsazione $\bar{\omega}$, così che complessivamente la funzione di trasferimento $F(s)C(s)$ tra segnale di riferimento e uscita abbia modulo circa unitario fino a $\bar{\omega}$. Si noti che in questo modo le proprietà di stabilità non vengono alterate.

Un secondo caso significativo in cui è opportuno progettare $C(s)$ come un filtro passa-alto è quello in cui eventuali restrizioni sulla banda passante massima ottenibile per il sistema di controllo in retroazione siano dovute a vincoli di tipo strutturale sul regolatore $R(s)$ da impiegare, per esempio perché è richiesto che esso abbia le caratteristiche di un PI o di un PID. Anche in questi casi un compensatore $C(s)$ opportunamente progettato può consentire di conseguire prestazioni migliori, come mostrato nell'esempio successivo.

Esempio 15.1 Si voglia progettare un regolatore in retroazione di tipo PI per un sistema descritto dal modello nominale

$$G(s) = \frac{1}{(1+s)(1+0.05s)^2} \quad (15.4)$$

in cui la collocazione dei poli (ad alta frequenza) in $s = -20$ è parzialmente incerta. A causa della struttura imposta al regolatore e per la presenza dell'incertezza, la pulsazione critica deve necessariamente essere limitata.

Il regolatore

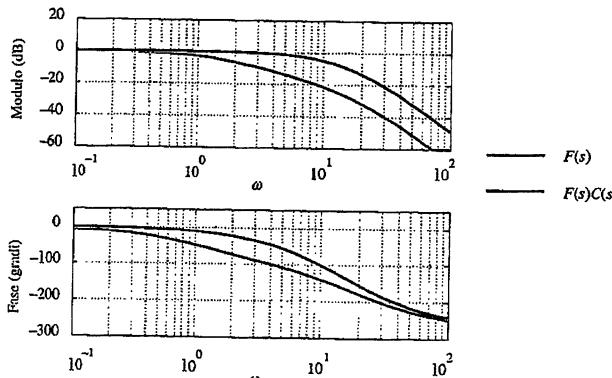
$$R(s) = \frac{1+s}{s} \quad (15.5)$$

soddisfa la specifica di struttura e garantisce un elevato margine di fase ($\varphi_m \approx 84^\circ$). Tuttavia, la corrispondente pulsazione critica è $\omega_c \approx 0.9$ e il sistema retroazionato si comporta in prima approssimazione come un filtro passa-basso del primo ordine con banda passante [0, 0.9]. Il compensatore $C(s)$ può quindi essere progettato come una rete anticipatrice con lo zero in $s = -1.1$ e il polo a frequenza più alta. Per esempio la scelta

$$C(s) = \frac{1+0.9s}{1+0.1s} \quad (15.6)$$

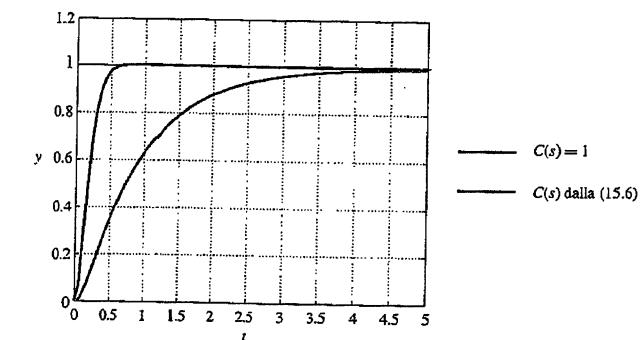
fa sì che la funzione di trasferimento $F(s)C(s)$ abbia la banda passante [0, 7], come mostrato dai diagrammi di Bode riportati nella Figura 15.2; nella stessa figura sono rappresentati anche i dia-

Figura 15.2
Diagrammi di Bode associati a $F(s)$ e a $F(s)C(s)$ nell'Esempio 15.1.



grammi di $F(s)$ definita tramite le (15.4), (15.5). Nella Figura 15.3 sono invece riportate le risposte allo scalino dei sistemi di controllo con $C(s) = 1$ e $C(s)$ dato dalla (15.6); come è ovvio, la presenza del compensatore velocizza sensibilmente la risposta del sistema retroazionato.

Figura 15.3
Risposta allo scalino del sistema retroazionato con $C(s) = 1$ e $C(s)$ dato dalla (15.6) nell'Esempio 15.1.



15.2.2 Compensazione del segnale di riferimento

Un diverso schema di compensazione del segnale di riferimento, che mette in luce anche con maggiore evidenza il ruolo svolto dai regolatori in anello aperto, è quello riportato nella Figura 15.4, dove $\hat{C}(s)$ può essere progettato con l'obiettivo di rendere unitaria la risposta in frequenza della funzione di trasferimento tra w e y nella banda più ampia possibile.

Dalla Figura 15.4 con semplici elaborazioni è facile verificare che risulta

$$Y(s) = \frac{R(s)G(s) + \hat{C}(s)G(s)}{1 + R(s)G(s)} W(s)$$

Pertanto, se si pone

$$\hat{C}(s) = G(s)^{-1} \quad (15.7)$$

si ha $Y(s) = W(s)$ come idealmente richiesto.

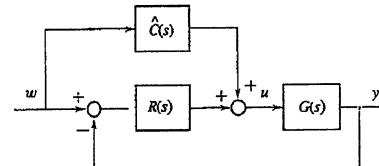


Figura 15.4
Schema di controllo con compensazione del segnale di riferimento.

Per garantire la realizzabilità e la stabilità asintotica del compensatore, la scelta ideale di $\hat{C}(s)$ data dalla (15.7) non può essere effettuata. Infatti, poiché $G(s)$ è strettamente propria, $\hat{C}(s)$ è comunque impropria. Inoltre, quando $G(s)$ ha uno o più zeri nel semipiano destro, il compensatore (in anello aperto) (15.7) è instabile. Infine, se $G(s)$ contiene un termine di ritardo, la formula (15.7) porta comunque a sintetizzare un compensatore dotato del termine non causale e^{ts} . Tuttavia, anche in questi casi, la (15.7) fornisce un'utile indicazione per il progetto di $\hat{C}(s)$. Infatti è evidente che è opportuno realizzare un compensatore in modo che la sua risposta in frequenza sia "la più simile possibile" a quella associata a $G(s)^{-1}$ nella banda di frequenze in cui il segnale di riferimento ha un contenuto armonico significativo.

Esempio 15.2 (Seguito dell'Esempio 15.1). Si consideri ancora il processo descritto dalla funzione di trasferimento (15.4) e il regolatore PI dato dalla (15.5). Volendo impiegare lo schema di Figura 15.4, dalle (15.4), (15.7) segue immediatamente che il compensatore ideale, con tre zeri e senza poli, è improprio. Tuttavia, supponendo che il segnale di riferimento non abbia un contenuto armonico apprezzabile per $\omega > 10$, un compensatore efficace e realizzabile è

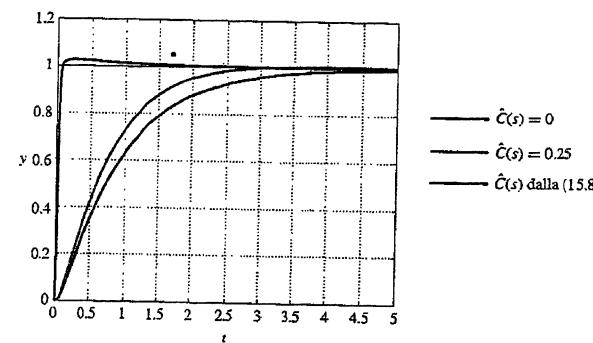
$$\hat{C}(s) = \frac{(1+s)(1+0.05s)^2}{(1+0.01s)^3} \quad (15.8)$$

La risposta allo scalino unitario del sistema di Figura 15.4, nei due casi $\hat{C}(s) = 0$ e $\hat{C}(s)$ dato dalla (15.8), è riportata nella Figura 15.5. Si osservi che la presenza del compensatore, pur velocizzando in modo considerevole la risposta, produce una piccola sovraelongazione dovuta a una non perfetta cancellazione di un polo con uno zero più vicino all'origine del piano complesso (si veda al riguardo il Paragrafo 4.4.4). Infatti con $\hat{C}(s)$ dato dalla (15.8), si può verificare che gli zeri della funzione di trasferimento tra w e y sono in $s = -1.08, s = -16.8, s = -22.84$, mentre i poli sono in $s = -1.12, s = -14.8$ e $s = -24$, oltre ai tre poli di $\hat{C}(s)$ in $s = -100$.

È opportuno anche segnalare che molto spesso è sufficiente utilizzare un compensatore $\hat{C}(s)$ di struttura molto semplice per ottenere un netto miglioramento delle prestazioni. Nell'esempio considerato è facile verificare che, se si impiega il compensatore statico $\hat{C}(s) = \mu_C$, risulta

$$Y(s) = \frac{1 + (1 + \mu_C)s}{(1 + s(1 + 0.05s)^2)(1 + s)} W(s) \quad (15.9)$$

Figura 15.5
Risposta allo scalino con $\hat{C}(r) = 0$, $\hat{C}(s)$ dato dalla (15.8) e $\hat{C}(s) = \mu_C = 0.25$ nell'Esempio 15.2.



Nella (15.9), se si pone $\mu_C = 0$, si ritrova la funzione di trasferimento tra riferimento e uscita in assenza del compensatore, mentre mediante un'opportuna scelta di μ_C si può fare in modo che il termine

$$\frac{1 + (1 + \mu_C)s}{(1 + s)}$$

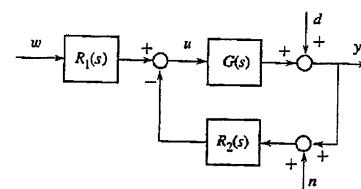
abbia le caratteristiche di una rete anticipatrice, da scegliersi con considerazioni analoghe a quelle impiegate nel paragrafo precedente a proposito del compensatore passa-alto. In particolare, la risposta allo scalino unitario del sistema con $\mu_C = 0.25$, anch'essa riportata nella Figura 15.5, evidenzia le migliori prestazioni ottenibili con una semplice compensazione di tipo statico. Un aumento del valore di μ_C d'altra parte genererebbe nella funzione di trasferimento del sistema (15.9) uno zero a frequenza nettamente più bassa rispetto ai poli e conseguentemente una risposta allo scalino con una sensibile sovraelongazione.

A conclusione del paragrafo, è opportuno osservare che il compensatore ideale (15.7) è indipendente dal regolatore in retroazione $R(s)$ utilizzato ed è quello che naturalmente verrebbe progettato se lo schema di controllo fosse soltanto in anello aperto. Quando invece non si utilizza il compensatore ideale, la scelta di $\hat{C}(s)$ dipende anche dal regolatore $R(s)$ utilizzato, come mostrato nell'esempio precedente relativamente alla determinazione di μ_C .

15.2.3 Schemi di controllo a due gradi di libertà

Gli schemi di controllo riportati nelle Figure 15.1 e 15.4 vengono comunemente detti *a due gradi di libertà*, in quanto sono caratterizzati dal fatto che la funzione di trasferimento tra il segnale di riferimento w e la variabile di controllo u non coincide con quella (cambiata di segno) tra l'uscita y e u , come invece avviene nello schema classico nel quale l'azione del controllore è basata sull'errore. Uno schema generale di controllo a due gradi di libertà è riportato nella Figura 15.6. I regolatori $R_1(s)$ e $R_2(s)$ possono essere progettati separatamente per tener conto di diverse esigenze, quali la robustezza, la risposta al riferimento, l'attenuazione

Figura 15.6
Schema di controllo a due gradi di libertà.



dei disturbi. Si osservi al riguardo che lo schema di Figura 14.3b del Paragrafo 14.3.1, impiegato nella realizzazione dei regolatori PID per la limitazione dell'azione derivativa, può essere interpretato come un altro esempio di schema a due gradi di libertà.

Il progetto degli schemi di controllo a due gradi di libertà, benché di notevole interesse per la soluzione di svariati problemi, non sarà ulteriormente sviluppato nel seguito. Il lettore interessato è rimandato alla letteratura specializzata.

15.2.4 Compensazione dei disturbi misurabili

Si consideri il sistema sotto controllo descritto da

$$Y(s) = G(s)U(s) + H(s)D(s)$$

Quando il disturbo d è misurabile, è possibile utilizzare un regolatore in anello aperto, con funzione di trasferimento $M(s)$ e ingresso d , per agire direttamente sulla variabile di controllo u e annullare o ridurre l'effetto di d sulla variabile controllata y . Per semplicità, nel seguito si ipotizzerà che $M(s)$ includa anche l'eventuale dinamica del trasduttore del disturbo. Lo schema di controllo corrispondente è riportato nella Figura 15.7, dove è presente anche il regolatore in retroazione $R(s)$, come sempre impiegato per far fronte a possibili incertezze o eventualmente per garantire la stabilità dell'intero sistema. Poiché il compensatore $M(s)$ opera in anello aperto, il requisito che esso necessariamente deve possedere è quello della stabilità asintotica, oltre a quello ovvio della realizzabilità.

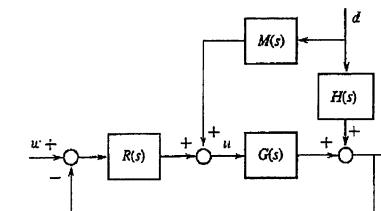


Figura 15.7
Schema di controllo con compensazione del disturbo.

Per determinare la forma che, almeno idealmente, deve avere $M(s)$ si noti che nello schema di Figura 15.7 risulta

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

Pertanto, per annullare l'effetto del disturbo sull'uscita è necessario porre

$$M(s) = -H(s)G(s)^{-1} \quad (15.10)$$

La (15.10) non può essere soddisfatta quando $H(s)G(s)^{-1}$ non è propria o quando $G(s)$ ha un ritardo di tempo o zeri a parte reale positiva. Tuttavia spesso è possibile realizzare il compensatore in modo che la risposta in frequenza associata a $M(s)$ sia simile, o uguale, a quella associata a $-H(s)G(s)^{-1}$ per le pulsazioni in cui d ha un contenuto armonico significativo. Se per esempio d è un segnale sinusoidale di pulsazione $\bar{\omega}$ sarà sufficiente impiegare un compensatore $M(s)$ tale che

$$M(j\bar{\omega}) = -H(j\bar{\omega})G(j\bar{\omega})^{-1} \quad (15.11)$$

per annullare asintoticamente l'effetto di d su y . In particolare, per $\bar{\omega} = 0$ si può fare ricorso al compensatore statico

$$M(s) = \mu_M = -\frac{H(0)}{G(0)} \quad (15.12)$$

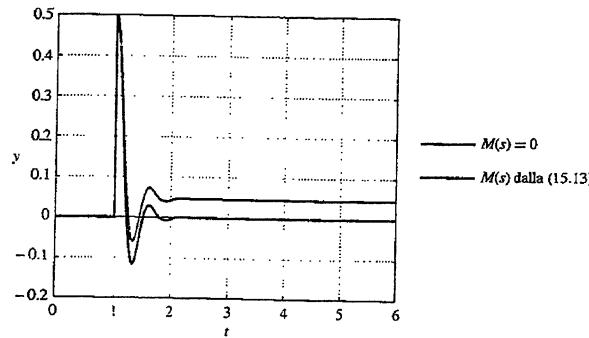
Tuttavia, nel caso di disturbi a scalino l'impiego del compensatore (15.12) non garantisce la reiezione robusta (si veda il Paragrafo 10.6.1), che invece è garantita da una scelta opportuna del regolatore in retroazione.

Esempio 15.3 Per la funzione di trasferimento $G(s)$ descritta dalla (15.4) e $H(s) = 1$, si consideri il regolatore in retroazione $R(s) = 10$. Il sistema retroazionato risultante ha $\varphi_m \simeq 51^\circ$ e $\omega_c \simeq 8.4$. Poiché la funzione d'anello non possiede alcuna azione integrale, l'errore a transitorio esaurito dovuto a un disturbo costante è asintoticamente non nullo, come mostrato nella Figura 15.8, in cui è riportato l'andamento dell'uscita a fronte di un disturbo a scalino di ampiezza 0.5 all'istante $t = 1$ e di un segnale di riferimento nullo. Se invece si utilizza il compensatore statico

$$M(s) = \mu_M = -1 \quad (15.13)$$

progettato secondo la (15.12), l'errore dovuto a questo disturbo asintoticamente si annulla: si veda ancora la Figura 15.8, dove è anche rappresentato il transitorio dell'uscita in presenza del compensatore (15.13).

Figura 15.8
Risposta a un disturbo a scalino con $M(s) = 0$ e $M(s)$ dato dalla (15.13) nell'Esempio 15.3.



A conclusione del paragrafo è opportuno osservare che il compensatore (15.10) è indipendente dal regolatore in retroazione $R(s)$ utilizzato ed è quello che naturalmente verrebbe progettato se lo schema di controllo fosse soltanto in anello aperto.

15.3 Predittore di Smith

Nei Capitoli 9-11 si è mostrato che la presenza di un ritardo di tempo limita le prestazioni ottenibili con il tradizionale schema in retroazione di Figura 15.9 in cui, per semplicità, non sono rappresentati i disturbi che agiscono sul sistema di controllo. Infatti il contributo negativo di fase introdotto dal ritardo impedisce di aumentare il valore della pulsazione critica, pena una drastica riduzione del margine di fase o addirittura la perdita della stabilità asintotica del sistema retroazionato (si veda al riguardo l'Esempio 11.4). Si mostra ora che, se il processo è asintoticamente stabile, queste restrizioni possono essere attenuate mediante l'adozione di un opportuno schema di controllo.

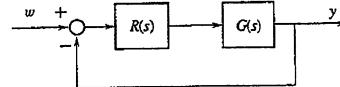


Figura 15.9
Schema "classico" di controllo in retroazione.

15.3.1 Schema a predittore di Smith

Si assume che il processo sia SISO, asintoticamente stabile e descritto da

$$G(s) = G'(s)e^{-\tau s} = \frac{N(s)}{D(s)}e^{-\tau s}$$

dove $N(s)$ e $D(s)$ sono i polinomi a numeratore e denominatore della funzione di trasferimento razionale $G'(s)$. Si consideri quindi lo schema di controllo di Figura 15.10, in cui il regolatore complessivo è costituito dai blocchi con funzione di trasferimento

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

e

$$P(s) = (1 - e^{-\tau s}) G'(s)$$

Mediante una semplice elaborazione è facile verificare che la funzione di trasferimento d'anello $L'(s)$ dello schema di Figura 15.10 è

$$L'(s) = (G(s) + P(s))R'(s) = G'(s)R'(s)$$

Il progetto di $R'(s)$ può quindi essere effettuato con riferimento a $G'(s)$ senza tener conto delle restrizioni dovute alla presenza del ritardo di tempo di $G(s)$.

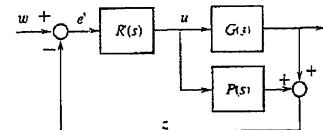


Figura 15.10
Schema a predittore di Smith.

Lo schema di Figura 15.10 è detto schema a *predittore di Smith* e fu originariamente introdotto nel 1958 da O.J.M. Smith. Esso deve il suo nome al fatto che la variabile z coincide con l'effettiva variabile di uscita y del processo a meno del ritardo τ e ne rappresenta quindi una predizione. Risulta infatti

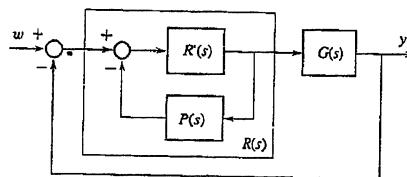
$$Z(s) = (G(s) + P(s))U(s) = \frac{N(s)}{D(s)}U(s) = e^{\tau s}G(s)U(s) = e^{\tau s}Y(s)$$

In altri termini, è come se nello schema venisse retroazionata non l'effettiva variabile sotto controllo $y(t)$, ma la sua predizione $z(t) = y(t + \tau)$.

Riguardo allo schema a predittore di Smith, è opportuno fare alcune considerazioni che ne chiariscano il funzionamento e le modalità d'uso. Si osservi in primo luogo che una sua rappresentazione alternativa è quella riportata nella Figura 15.11; esso è pertanto del tutto equivalente al classico schema di Figura 15.9 pur di porre

$$R(s) = \frac{R'(s)}{1 + R'(s)P(s)} = \frac{N_R(s)D(s)}{D(s)D_R(s) + N(s)N_R(s) - N(s)N_R(s)e^{-\tau s}}$$

Figura 15.11
Rappresentazione alternativa dello schema a predittore di Smith.



Questo regolatore ha tra gli zeri le radici di $D(s)$ ed effettua quindi una cancellazione dei poli del processo: si giustifica così l'ipotesi di asintotica stabilità di $G(s)$ inizialmente introdotta. Il regolatore $R'(s)$ è tuttavia un sistema con funzione di trasferimento non razionale per la presenza del termine $e^{-\tau s}$ a denominatore.

Si osservi inoltre che risulta $P(0) = 0$; pertanto, nell'ipotesi di stabilità asintotica del sistema retroazionato e a fronte di segnali di riferimento a scalino, nello schema di Figura 15.10 la variabile z coincide asintoticamente con l'uscita y e la variabile e' coincide con l'errore $e = w - y$. Perciò, per avere errore nullo a transitorio esaurito a fronte dei suddetti segnali esogeni, il regolatore $R'(s)$ deve includere almeno un'azione integrale e , più in generale, tutte le considerazioni fatte nel Capitolo 10 circa il legame tra l'errore a transitorio esaurito e la struttura del regolatore $R(s)$ nello schema di Figura 15.9 si estendono immutate al regolatore $R'(s)$ nello schema di Figura 15.10.

Posto

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

si può inoltre facilmente verificare che la funzione di trasferimento tra w e y è

$$F(s) = \frac{R'(s)G(s)}{1 + R'(s)P(s) + R'(s)G(s)} = F'(s)e^{-\tau s} \quad (15.14)$$

Come mostra la (15.14), benché $R'(s)$ vada progettato unicamente con riferimento a $G'(s)$, anche il ritardo di tempo $e^{-\tau s}$ compare come fattore di $F(s)$ e influenza la velocità del sistema di controllo, che pertanto non può essere arbitrariamente elevata. Esaminando l'uscita y conseguente a uno scalino di w , si osserverà in ogni caso un ritardo causato da $e^{-\tau s}$.

Infine, nel progetto di $R'(s)$ è comunque opportuno tener conto anche dei motivi che suggeriscono di non allargare eccessivamente la banda passante del sistema retroazionato, come per esempio il requisito di moderazione della variabile di controllo o la presenza di incertezze di modello in alta frequenza. Per comprendere meglio, seppure in modo intuitivo, questo aspetto, si noti per esempio che eventuali errori di valutazione del ritardo del sistema fanno sì che la variabile z non coincida più con la predizione di y e pertanto il segnale e' non possa essere considerato equivalente all'errore $w - y$ a meno del ritardo del processo.

15.3.2 Approssimanti di Padé

Nella realizzazione dello schema a predittore di Smith, e più in generale nello studio dei sistemi dinamici caratterizzati dalla presenza di ritardi di tempo, può essere conveniente sostituire al ritardo una sua approssimante costruita come una funzione razionale della variabile complessa s . A questo scopo sono normalmente impiegate le *approssimanti di Padé*, derivate come descritto nel seguito. Si definisce in primo luogo la struttura della funzione approssimante, razionale in s , che si desidera considerare e il valore \hat{s} della variabile complessa s per cui si vuole

ottenere l'approssimante. Il valore dei q parametri liberi della funzione approssimante viene quindi determinato facendo coincidere i primi q termini del suo sviluppo in serie attorno a \hat{s} con i corrispondenti valori dello sviluppo in serie di $e^{-\tau s}$ attorno a \hat{s} . Poiché normalmente si è interessati ad approssimazioni in bassa frequenza, di solito si sceglie $\hat{s} = 0$.

Esempio 15.4 Si voglia determinare l'approssimante di Padé del primo ordine e di struttura

$$G_I(s) = \mu \frac{1 + as}{1 + bs} \quad (15.15)$$

attorno a $\hat{s} = 0$. Poiché i parametri liberi della (15.15) sono μ , a , b , è possibile far coincidere i primi tre coefficienti dello sviluppo di McLaurin di

$$e^{-\tau s} \approx 1 - \tau s + \frac{\tau^2 s^2}{2} + \dots$$

con i corrispondenti termini dello sviluppo della (15.15)

$$\mu \frac{1 + as}{1 + bs} = \mu + \mu(a - b)s + \mu b(b - a)s^2 + \dots$$

Risulta pertanto

$$\mu = 1, \quad b - a = \tau, \quad b(b - a) = \tau^2/2$$

da cui

$$G_I(s) = \frac{1 - 0.5\tau s}{1 + 0.5\tau s} \quad (15.16)$$

Si osservi che la (15.16) corrisponde a uno sfasatore puro (si veda il Paragrafo 6.6.4), asintomaticamente stabile, con risposta in frequenza di modulo unitario e fase decrescente all'aumentare della frequenza.

Con considerazioni analoghe a quelle del caso precedente è possibile verificare che l'approssimante di Padé del secondo ordine con due zeri, per $\hat{s} = 0$, è data da

$$G_{II}(s) = \frac{1 - 0.5\tau s + \tau^2 s^2/12}{1 + 0.5\tau s + \tau^2 s^2/12} \quad (15.17)$$

Anche in questo caso si ottiene uno sfasatore puro asintomaticamente stabile, a modulo unitario e fase decrescente con la frequenza. Per valutare la significatività in frequenza delle (15.16) e (15.17), nella Figura 15.12 sono riportati i diagrammi di Bode della fase delle risposte in frequenza associate al ritardo di tempo $e^{-\tau s}$, a $G_I(s)$ e a $G_{II}(s)$ in funzione della pulsazione normalizzata $\omega\tau$. Dalla figura si conclude che $G_I(s)$ può essere considerata una buona approssimazione in frequenza del ritardo fino a $\omega \approx 1/\tau$, mentre la validità di $G_{II}(s)$ si estende sino a $\omega \approx 2/\tau$. Naturalmente l'uso di approssimanti razionali di struttura più complessa, e quindi con un maggior numero di parametri liberi, consentirebbe di ottenere approssimanti sempre migliori.

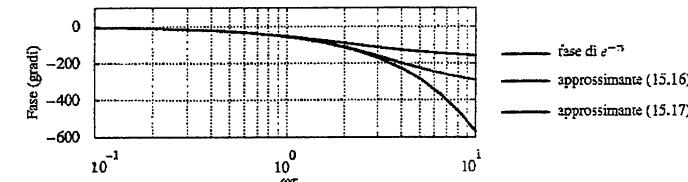


Figura 15.12
Diagramma di Bode della fase di $e^{-\tau s}$ e delle approssimanti (15.16) e (15.17).

15.3.3 Impiego delle approssimanti di Padé nel predittore di Smith

È possibile ora valutare tramite un esempio le prestazioni ottenibili con lo schema a predittore di Smith nel caso in cui il ritardo sia riprodotto in modo corretto e

quando invece nella costruzione del predittore il ritardo venga realizzato mediante un'approssimante di Padé.

Esempio 15.5 (Seguito dell'Esempio 11.4) Dato il sistema descritto da

$$G(s) = \frac{1}{(1+s)^2} e^{-4s} \quad (15.18)$$

si è precedentemente mostrato che, mediante lo schema di Figura 15.9, per ottenere il margine di fase $\varphi_m \geq 30^\circ$ con un regolatore PI è necessario limitare la banda passante del sistema retroazionato. In particolare si è progettato il regolatore

$$R(s) = \frac{0.2(1+s)}{s} \quad (15.19)$$

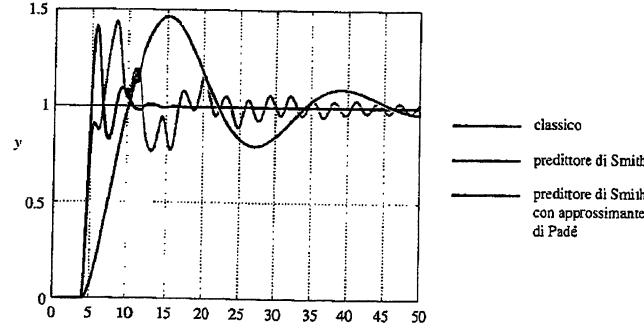
a cui corrisponde la pulsazione critica $\omega_c \approx 0.2$ e il margine di fase $\varphi_m \approx 33^\circ$.

Se invece si considera lo schema di Figura 15.10 in cui

$$P(s) = (1 - e^{-4s}) \frac{1}{(s+1)^2}, \quad R'(s) = \frac{3.46(s+1)}{s} \quad (15.20)$$

si ottiene la pulsazione critica $\omega_c \approx 1.73$ e il margine di fase $\varphi_m \approx 30^\circ$. La risposta allo scalino del sistema retroazionato di Figura 15.9 con $G(s)$ e $R(s)$ date dalle (15.18), (15.19) è confrontata nella Figura 15.13 con quella dello schema di Figura 15.10 dove $P(s)$ e $R'(s)$ sono date dalla (15.20).

Figura 15.13
Risposta allo scalino del sistema di controllo dell'Esempio 15.5: schema "classico", schema a predittore di Smith e schema a predittore di Smith con approssimante di Padé.



Per analizzare l'effetto di una realizzazione imprecisa del predittore, che può essere interpretata come un errore di modellizzazione, si sostituisca al ritardo nel blocco $P(s)$ di Figura 15.10 l'approssimante di Padé del primo ordine ottenendo

$$P(s) = \left(1 - \frac{1-2s}{1+2s}\right) \frac{1}{(s+1)^2} = \frac{4s}{(1+2s)(s+1)^2} \quad (15.21)$$

La risposta allo scalino del sistema retroazionato di Figura 15.10, dove $P(s)$ è data dalla (15.21) e $R'(s)$ dalla (15.20), è anch'essa riportata nella Figura 15.13; è evidente che l'errore di modello dovuto all'uso dell'approssimante di Padé è tale da deteriorare notevolmente le prestazioni rispetto al caso ideale. A questo riguardo si ricorda che l'approssimante di Padé utilizzata si può considerare una buona approssimazione in frequenza del ritardo fino alla pulsazione $\omega = 0.25$, nettamente inferiore alla pulsazione critica $\omega_c = 1.73$ considerata nell'esempio.

Si osservi infine che, in ogni caso, la risposta allo scalino presenta un ritardo pari a quello di $G(s)$, come già evidenziato nell'analisi della funzione di trasferimento (15.14).

15.4 Regolatori in cascata

Può accadere che il sistema sotto controllo sia composto da due sottosistemi in serie descritti dalle funzioni di trasferimento $G_1(s)$ e $G_2(s)$, come mostrato nella

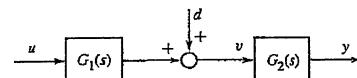


Figura 15.14
Processo composto da sottosistemi in serie.

Figura 15.14. Si assuma poi che sull'uscita di $G_1(s)$ agisca un disturbo d e che la variabile intermedia v sia misurabile. Si ipotizzi anche che la risposta in frequenza $G_1(j\omega)$ abbia uno sfasamento trascurabile nella banda in cui invece lo sfasamento di $G_2(j\omega)$ è significativo, o che comunque le prestazioni che si possono ottenere nel progetto di un sistema di controllo solo per $G_1(s)$ siano migliori di quelle conseguibili nella sintesi di un regolatore per l'intero processo $G_1(s)G_2(s)$. Questa ipotesi è per esempio verificata quando $G_1(s)$ è a sfasamento minimo e $G_2(s)$ ha zeri a parte reale positiva o contiene un ritardo di tempo. Una seconda situazione significativa è quella in cui, benché sia $G_1(s)$ sia $G_2(s)$ siano a sfasamento minimo, il requisito di moderazione della variabile di controllo pone limitazioni sulla banda passante ottenibile nel progetto, mentre la dinamica della sola $G_1(s)$ (più "veloce" di quella di $G_2(s)$) consentirebbe il conseguimento di prestazioni migliori. Queste ipotesi sono spesso verificate quando $G_1(s)$ rappresenta l'attuatore e $G_2(s)$ il processo sotto controllo; in questo caso è frequente che la dinamica dell'attuatore sia rapida e non introduca uno sfasamento negativo nella banda di frequenze in cui l'effetto di poli, zeri o del ritardo del processo è non trascurabile.

Nelle ipotesi precedenti è conveniente utilizzare lo schema di *controllo in cascata* di Figura 15.15, dove il regolatore $R_1(s)$ è progettato con riferimento al solo anello interno (e quindi a $G_1(s)$) con lo scopo di attenuare quanto più possibile l'effetto di d su v (e quindi sulla variabile controllata y), mentre la sintesi di $R_2(s)$ viene effettuata con riferimento a $G_2(s)$ assumendo che sia $v \approx v^*$.

Per comprendere meglio quanto esposto si osservi che risulta

$$V(s) = \frac{1}{1 + R_1(s)G_1(s)} D(s) + \frac{R_1(s)G_1(s)}{1 + R_1(s)G_1(s)} V^*(s) \quad (15.22)$$

dove, come di consueto, $V(s)$, $V^*(s)$ e $D(s)$ sono le trasformate di Laplace di v , v^* e d . Se lo spettro di d ha componenti armoniche significative fino a una data pulsazione $\bar{\omega}$ e se $R_1(s)$ può essere progettato in modo che $|R_1(j\omega)G_1(j\omega)| \gg 1$ per ogni $\omega \leq \bar{\omega}$, si può ritenere che l'effetto di d su v sia attenuato, o in pratica annullato, e che quindi nella (15.22) risulti

$$\left| \frac{1}{1 + R_1(j\omega)G_1(j\omega)} \right| \simeq 0 \quad (15.23)$$

D'altra parte, poiché $G_2(s)$ pone dei vincoli stringenti sulla pulsazione critica ω_c ottenibile per l'intero sistema di controllo, se risulta $\omega_c \ll \bar{\omega}$, nel progetto di $R_2(s)$ si può assumere

$$\left| \frac{R_1(j\omega)G_1(j\omega)}{1 + R_1(j\omega)G_1(j\omega)} \right| \simeq 1 \quad (15.24)$$

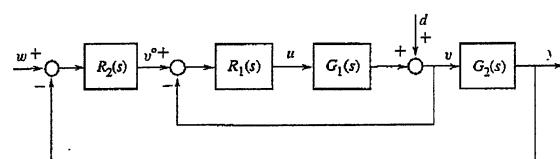


Figura 15.15
Schema di controllo in cascata.

nell'intervallo di pulsazioni di interesse $[0, \omega_c]$, e quindi $R_2(s)$ può essere sintetizzato unicamente con riferimento a $G_2(s)$. In altri termini, $R_1(s)$ e $R_2(s)$ sono progettati secondo un criterio di *disaccoppiamento in frequenza*, la cui validità è bene comunque che sia confermata da un'analisi a posteriori una volta completata la fase di sintesi.

Esempio 15.6 Si consideri il sistema di Figura 15.14 con

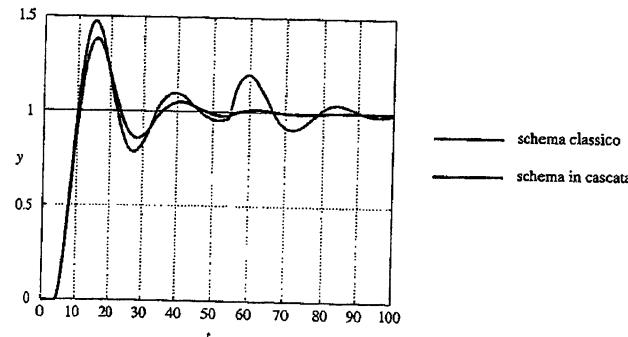
$$G_1(s) = \frac{1}{1 + 0.05s}, \quad G_2(s) = \frac{e^{-4s}}{(1 + s)^2}$$

Per il controllo di $G(s) = G_1(s)G_2(s)$ si utilizzi dapprima lo schema "classico" di Figura 15.9, dove $R(s)$ ha la struttura di un PI ed è descritto da

$$R(s) = \frac{0.2(s+1)}{s} \quad (15.25)$$

Il sistema retroazionato che si ottiene ha pulsazione critica $\omega_c \approx 0.2$ e margine di fase $\varphi_m \approx 33^\circ$; si confronti questo risultato con quelli degli Esempi 11.4 e 15.5 per verificare che la presenza di $G_1(s)$ è di fatto irrilevante ai fini della determinazione di $R(s)$. Nella Figura 15.16 è riportato l'andamento dell'uscita y a fronte di uno scalino unitario del segnale di riferimento a $t = 0$ e di uno scalino di ampiezza 0.2 del disturbo d che interviene all'istante $t = 50$.

Figura 15.16
Esempio 15.6: risposta allo scalino dello schema classico e dello schema in cascata.



Si realizzi ora lo schema di controllo in cascata di Figura 15.15, in cui $R_2(s)$ coincide con la funzione di trasferimento (15.25) e

$$R_1(s) = 10 \quad (15.26)$$

Il sistema di controllo risultante ha pulsazione critica $\omega_c \approx 0.18$ e margine di fase $\varphi_m \approx 38^\circ$, valori simili a quelli ottenuti con il progetto precedente. È quindi lecito attendersi che la risposta allo scalino del segnale di riferimento sia analoga a quella fornita dallo schema di Figura 15.9. Si osservi tuttavia che nello schema di controllo in cascata, per la particolare scelta di $R_1(s)$ data dalla (15.26), risulta

$$|R_1(j\omega)G_1(j\omega)| \approx 10$$

per valori di ω anche nettamente superiori alla pulsazione critica $\omega_c \approx 0.18$. Si può quindi ritenere che per $\omega \in [0, 20]$ risulti

$$\left| \frac{1}{1 + R_1(j\omega)G_1(j\omega)} \right| \approx 0.1, \quad \left| \frac{R_1(j\omega)G_1(j\omega)}{1 + R_1(j\omega)G_1(j\omega)} \right| \approx 1$$

e che pertanto l'effetto del disturbo d sull'uscita del sistema venga sensibilmente attenuato. A questo riguardo si ricorda che la trasformata di Fourier del disturbo è $D(j\omega) = 0.2[(1/\omega) + \pi \operatorname{imp}(\omega)]$ (si veda l'Esempio B.24), il cui contenuto armonico è prevalentemente concentrato nell'intorno di $\omega = 0$.

Queste considerazioni sono confermate dall'andamento dell'uscita, anch'esso riportato nella Figura 15.16, a fronte di andamenti a scalino del segnale di riferimento e del disturbo identici a quelli considerati per lo schema classico di Figura 15.9. Dal confronto dei transitori è evidente che nello schema di controllo in cascata l'effetto di d su y è irrilevante.

15.5 Controllo di sistemi instabili

Quando il sistema sotto controllo è instabile per la presenza di poli a parte reale positiva, non è possibile utilizzare i metodi di sintesi in frequenza basati sul criterio di Bode e presentati nel Capitolo 11. In questi casi il progetto del regolatore può essere effettuato per mezzo del criterio di Nyquist, tramite il metodo del luogo delle radici (Capitolo 12) o mediante il metodo dell'assegnamento degli autovalori (Capitolo 13). Tuttavia, la definizione delle specifiche risulta spesso poco agevole in quanto requisiti sulla velocità di risposta, sulla limitatezza delle sovraelongazioni e sulla robustezza non sono più direttamente legati a indicatori sintetici come il margine di fase o di guadagno e la pulsazione critica. Per ricordarsi alle tecniche di sintesi in frequenza sin qui usate, è allora conveniente ricorrere allo schema di controllo di Figura 15.17, in cui il regolatore $R_1(s)$ viene progettato con lo scopo principale di stabilizzare l'anello interno con funzione d'anello $R_1(s)G(s)$, e il regolatore $R_2(s)$ è successivamente determinato con riferimento alla funzione di sensitività complementare dell'anello interno

$$F_1(s) = \frac{R_1(s)G(s)}{1 + R_1(s)G(s)} \quad (15.27)$$

per soddisfare i consueti requisiti sull'errore a transitorio esaurito, sulla pulsazione critica, sul margine di fase.

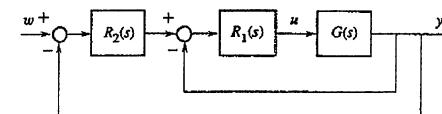


Figura 15.17
Schemi di controllo di sistemi instabili.

Esempio 15.7 Dato il processo descritto da

$$G(s) = \frac{100}{(1 - 0.1s)(1 + 0.2s)} = \frac{-5000}{(s - 10)(s + 5)}$$

si voglia progettare un sistema di controllo in anello chiuso, con la struttura di Figura 15.17, tale che l'errore a transitorio esaurito sia nullo a fronte di segnali di riferimento a scalino e che risulti $\omega_c \geq 10$ e $\varphi_m \geq 45^\circ$.

Il progetto di $R_1(s)$ può essere effettuato con il metodo del luogo delle radici. Con questa tecnica è immediato verificare che non è possibile stabilizzare l'anello interno con un regolatore puramente proporzionale $R_1(s) = k$ (si veda il luogo di Figura 15.18a), ma è necessario "sostare" i rami del luogo verso il semipiano reale negativo. Per questo è conveniente utilizzare il regolatore

$$R_1(s) = k \frac{s + 5}{s + p} \quad (15.28)$$

dove i parametri liberi $p > 0$ e k determinano rispettivamente la forma e la punteggiatura del luogo, riportato nella Figura 15.18b.

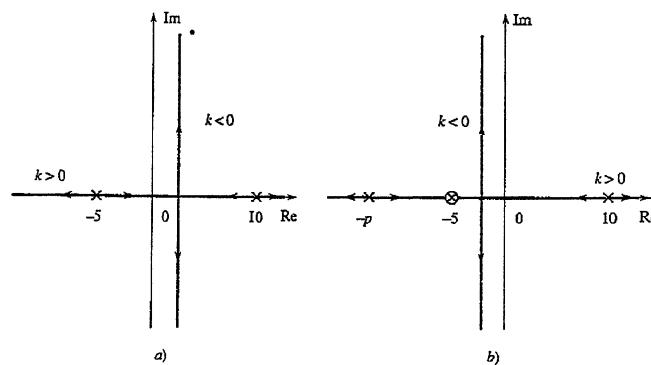
Se per esempio si desidera avere due poli coincidenti in $s = -10$, è necessario porre $p = 30$ e $k = -0.08$. In corrispondenza, la (15.27) diventa

$$F_1(s) = \frac{400}{(s + 10)^2} = \frac{4}{(1 + 0.1s)^2} \quad (15.29)$$

Con riferimento al sistema con funzione d'anello $R_2(s)F_1(s)$, dove $F_1(s)$ è data dalla (15.29), è ora possibile soddisfare le specifiche sull'errore a transitorio esaurito, sulla pulsazione critica e sul margine di fase scegliendo per esempio il regolatore PID (ideale)

$$R_2(s) = \frac{\mu(1 + 0.1s)^2}{s}$$

Figura 15.18
Luogo delle radici del sistema retroazionato dell'Esempio 15.7 con funzione d'anello
a) $R_1(s)G(s)$ con $R_1(s) = k$;
b) $R_1(s)$ data dalla (15.28).



con $\mu \geq 2.5$. Risulta infatti $\omega_c = 4\mu$ e $\varphi_m = 90^\circ$; inoltre la presenza nel regolatore dell'azione integrale annulla asintoticamente l'errore in risposta ai segnali di riferimento specificati.

Si osservi infine che, benché lo scopo precipuo del progetto di $R_1(s)$ sia quello di stabilizzare l'anello interno, una cattiva scelta di questo regolatore si riflette negativamente nella successiva fase di sintesi di $R_2(s)$. Nell'esempio considerato, prendendo ancora $R_1(s)$ dato dalla (15.28) con k negativo, ma di valore assoluto "grande", si ottiene una funzione $F_1(s)$ con una coppia di poli complessi coniugati con smorzamento decrescente al crescere di $|k|$ (si veda ancora il luogo delle radici di Figura 15.18b). Come mostrato nell'Esempio 11.5, il progetto di un regolatore per un sistema con poli complessi coniugati a basso smorzamento è particolarmente difficoltoso se effettuato mediante tecniche basate sul criterio di Bode.

15.6 Regolatori di disaccoppiamento

Le tecniche sin qui sviluppate fanno riferimento unicamente al controllo di sistemi SISO. In generale, benché gli stessi principi possano essere estesi ai sistemi MIMO, l'analisi e la sintesi dei sistemi retroazionati con più ingressi e più uscite richiede metodi più complessi di quelli presentati in questo testo. Tuttavia, spesso è possibile ricondurre il problema della sintesi dei sistemi di controllo MIMO a un insieme di problemi SISO. In questo paragrafo si analizzerà più in dettaglio questo approccio facendo riferimento in particolare a sistemi con due variabili di controllo e due variabili controllate, o sistemi 2×2 , e generalizzando quanto verrà esposto al caso di sistemi $m \times m$, cioè con m variabili di controllo e m variabili controllate, soltanto al termine della presentazione. Non si tratterà mai invece il caso, più complicato, in cui il processo abbia un numero di variabili di controllo diverso da quello delle variabili controllate. Inoltre, sempre per semplicità, si ignorerà la presenza di eventuali disturbi.

Per chiarezza, nel seguito la funzione di trasferimento $G(s)$ di un sistema $m \times m$ sarà chiamata *matrice di trasferimento*, mentre i suoi elementi $G_{ij}(s)$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, m$, saranno ancora definiti funzioni di trasferimento. È anche importante sottolineare il fatto che, date due matrici di trasferimento $G(s)$ e $H(s)$, entrambe di dimensione $m \times m$, in generale risulta $G(s)H(s) \neq H(s)G(s)$; ciò implica che nell'elaborazione degli schemi a blocchi dei sistemi MIMO è necessario considerare correttamente l'ordine dei prodotti. Infine, un sistema sarà detto *triangolare* se caratterizzato da una matrice di trasferimento $G(s)$ triangolare (si veda l'Appendice A).

15.6.1 Disaccoppiamento di sistemi triangolari

Si consideri il sistema 2×2 descritto da

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & 0 \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix}$$

in cui, per la struttura triangolare della matrice di trasferimento $G(s)$, la prima variabile di uscita y_1 dipende unicamente dalla prima variabile di ingresso u_1 , mentre la seconda variabile di uscita y_2 dipende da entrambi gli ingressi u_1 e u_2 . Per questo sistema è naturale considerare lo schema di controllo di Figura 15.19, in cui l'anello superiore ha il compito di regolare y_1 agendo su u_1 . Il progetto del regolatore SISO $R'_1(s)$ può essere effettuato con le tecniche già introdotte a partire dalla funzione di trasferimento del processo $G_{11}(s)$, in quanto la variabile u_2 non ha alcuna influenza su y_1 . Al contrario, u_1 agisce su y_2 tramite $G_{21}(s)$; tuttavia, nel progetto di un regolatore per l'anello inferiore di Figura 15.19, il segnale u_1 può essere considerato come un disturbo misurabile (la variabile di controllo u_1 è ovviamente nota in quanto generata dal regolatore $R'_1(s)$). Si può quindi realizzare il compensatore del "disturbo" u_1 secondo quanto esposto nel Paragrafo 15.2.4 e determinare la funzione di trasferimento

$$M(s) = -G_{21}(s)G_{22}(s)^{-1}$$

che compare nello schema, o una sua approssimante in frequenza se essa non è realizzabile. Infine il regolatore SISO $R'_2(s)$ viene progettato con riferimento alla funzione di trasferimento $G_{22}(s)$.

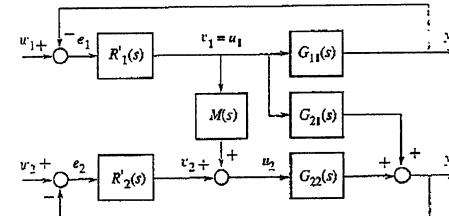


Figura 15.19
Schema di controllo con disaccoppiamento di un sistema triangolare 2×2 .

Utilizzando lo schema di Figura 15.19 si è in sostanza effettuato il *disaccoppiamento* dei due problemi di controllo.

In generale, anche per sistemi $m \times m$, con $m > 2$, in cui la matrice di trasferimento $G(s)$ è triangolare, il problema di progetto può essere ricondotto alla sequenza di m problemi SISO.

15.6.2 Disaccoppiamento di sistemi generici

Lo schema di controllo con disaccoppiamento di Figura 15.19 può essere rappresentato come in Figura 15.20 e il corrispondente regolatore può essere descritto, con notazione matriciale, dalle relazioni

$$U(s) = \Delta(s)V(s) \quad , \quad V(s) = R'(s)E(s) \quad (15.30)$$

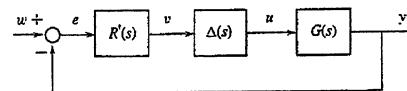


Figura 15.20
Schema di controllo con disaccoppiatore.

dove

$$\Delta(s) = \begin{bmatrix} 1 & 0 \\ -\frac{G_{21}(s)}{G_{22}(s)} & 1 \end{bmatrix}, R'(s) = \begin{bmatrix} R'_1(s) & 0 \\ 0 & R'_2(s) \end{bmatrix}, E(s) = \begin{bmatrix} W_1(s) - Y_1(s) \\ W_2(s) - Y_2(s) \end{bmatrix} \quad (15.31)$$

Il regolatore complessivo, con ingresso e e uscita u , ha quindi matrice di trasferimento $R(s) = \Delta(s)R'(s)$. Il blocco $\Delta(s)$ rappresenta il cosiddetto *disaccoppiatore* e nella (15.31) è progettato in modo che la matrice di trasferimento $G(s)\Delta(s)$ assuma la struttura diagonale

$$\tilde{G}_d(s) = G(s)\Delta(s) = \begin{bmatrix} G_{11}(s) & 0 \\ 0 & G_{22}(s) \end{bmatrix}$$

Una volta determinata $\Delta(s)$, gli elementi della matrice diagonale $R'(s)$ sono sintetizzati a partire dai corrispondenti elementi di $\tilde{G}_d(s)$.

Ci si può chiedere ora se è possibile impiegare, anche nel caso in cui $G(s)$ non sia triangolare, il procedimento precedente, che consiste nel progettare dapprima il disaccoppiatore $\Delta(s)$ che soddisfi l'equazione

$$G(s)\Delta(s) = G_d(s) \quad (15.32)$$

dove $G_d(s)$ è una matrice diagonale opportunamente definita, e quindi nel risolvere due ($o m$) problemi di controllo indipendenti SISO per la determinazione di $R'(s)$. Per questo problema, nel seguito saranno fornite due soluzioni alternative.

15.6.3 Disaccoppiamento "in avanti"

Si osservi innanzitutto che formalmente si può riscrivere la (15.32) come

$$\Delta(s) = G(s)^{-1}G_d(s) \quad (15.33)$$

che, per una generica $G_d(s)$, rappresenta la forma ideale del disaccoppiatore.

Per comprendere le condizioni sotto cui è possibile impiegare la (15.33), si supponga temporaneamente che il processo sia SISO. In questo caso, data una generica $G_d(s)$, il regolatore complessivo

$$R(s) = \Delta(s)R'(s) = G(s)^{-1}G_d(s)R'(s)$$

è causale solamente se $G(s)$ non contiene un termine di ritardo di tempo. Inoltre, $R(s)$ effettua la cancellazione dei poli e degli zeri del processo, che pertanto deve essere a fase minima e senza singolarità con parte reale nulla. In caso contrario, $G_d(s)$ dovrà contenere le singolarità di $G(s)$ a parte reale maggiore o uguale a zero e l'eventuale termine di ritardo. Condizioni analoghe possono essere espresse anche nel caso di sistemi MIMO; tuttavia, una trattazione completa non può essere qui presentata poiché in questo testo non sono state introdotte le nozioni di *zero* e di *struttura dei ritardi* per sistemi MIMO. Per procedere comunque secondo la (15.33) si ipotizzerà nel prosieguo che valgano le seguenti condizioni sufficienti:

- a) $G(s)$ è razionale;
- b) $G(s)$ è asintoticamente stabile;
- c) $\det G(s) \neq 0$, per ogni s con $\operatorname{Re}(s) \geq 0$.

Se queste condizioni non sono tutte verificate, la procedura descritta nel seguito può risultare inapplicabile, ma ciò tuttavia non significa che non si possa determinare comunque $\Delta(s)$ con altri procedimenti che richiedono una scelta appropriata di $G_d(s)$.

Nell'ipotesi in cui le condizioni sufficienti siano soddisfatte, posto

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

per una generica $G_d(s)$ la condizione di disaccoppiamento (15.32) richiede che siano soddisfatte le due equazioni scalari

$$\Delta_{11}(s)\Delta_{12}(s) + \Delta_{12}(s)\Delta_{22}(s) = 0 \quad (15.34)$$

$$\Delta_{21}(s)\Delta_{11}(s) + \Delta_{22}(s)\Delta_{21}(s) = 0 \quad (15.35)$$

nelle quattro incognite $\Delta_{11}(s)$, $\Delta_{12}(s)$, $\Delta_{21}(s)$ e $\Delta_{22}(s)$. Tra le infinite soluzioni del sistema (15.34), (15.35), una comunemente utilizzata è

$$\Delta_{11}(s) = \Delta_{22}(s) = 1 \quad (15.36)$$

$$\Delta_{12}(s) = -G_{12}(s)/G_{11}(s) \quad (15.37)$$

$$\Delta_{21}(s) = -G_{21}(s)/G_{22}(s) \quad (15.38)$$

a cui corrisponde lo schema di controllo complessivo di Figura 15.21.

A partire dalle (15.36)-(15.38) è facile verificare che risulta

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) - \frac{G_{12}(s)G_{21}(s)}{G_{22}(s)} & 0 \\ 0 & G_{22}(s) - \frac{G_{21}(s)G_{12}(s)}{G_{11}(s)} \end{bmatrix} \begin{bmatrix} V_1(s) \\ V_2(s) \end{bmatrix}$$

Si è quindi assicurato il disaccoppiamento desiderato, ma le funzioni di trasferimento diagonali ottenute, che devono essere considerate nella sintesi di $R'(s)$, possono risultare di ordine elevato.

A conclusione di questo punto si osservi che in alcuni casi le relazioni (15.36)-(15.38) non possono essere soddisfatte per problemi di realizzabilità delle funzioni

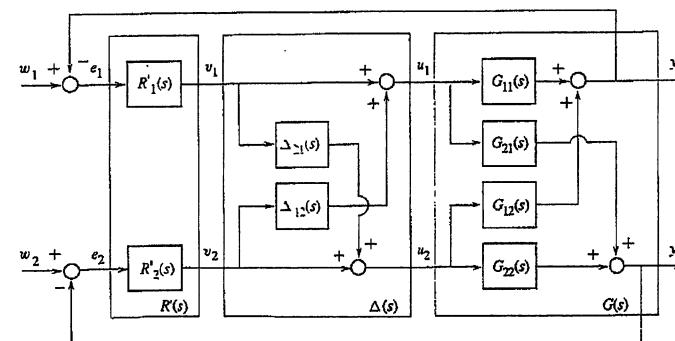


Figura 15.21
Schema di controllo con disaccoppiamento "in avanti" di un sistema 2×2 .

di trasferimento $\Delta_{ij}(s)$, $i, j = 1, 2$. Se questo accade, è possibile effettuare scelte diverse da quelle specificate tramite le (15.36)-(15.38), ma che tuttavia verificino le condizioni (15.34), (15.35). In alternativa, prendendo ancora le funzioni $\Delta_{ij}(s)$, $i, j = 1, 2$, in accordo con le (15.36)-(15.38), è possibile effettuare un disaccoppiamento limitatamente all'intervallo di pulsazioni di interesse nel progetto del sistema di controllo, secondo le modalità già discusse a proposito della compensazione in anello aperto, per esempio aggiungendo un adeguato numero di poli in alta frequenza all'esterno di tale intervallo. In particolare, a volte può bastare usare un disaccoppiatore $\Delta(s)$ tale che

$$\Delta(j\bar{\omega}) = G(j\bar{\omega})^{-1}$$

per disaccoppiare il sistema in corrispondenza di una specifica pulsazione $\bar{\omega}$, tipicamente scelta uguale a zero.

15.6.4 Disaccoppiamento "all'indietro"

Mediante un diverso schema di disaccoppiamento è possibile ottenere una matrice di trasferimento $G(s)\Delta(s)$ più semplice e generalizzare il procedimento di determinazione del disaccoppiatore ai sistemi $m \times m$. A questo scopo si assume preliminarmente che le condizioni (a)-(c) introdotte in precedenza siano verificate e si imponga che gli elementi della matrice diagonale $G_d(s)$ siano

$$G_{di}(s) = G_{ii}(s) \quad , \quad i = 1, 2, \dots, m \quad (15.39)$$

Si ipotizzi poi che il disaccoppiatore sia descritto da

$$U(s) = \Gamma(s)U(s) + V(s) \quad (15.40)$$

in cui $\Gamma(s)$ è da scegliersi opportunamente. Dalla (15.40) e ricordando le (15.30) è immediato verificare che risulta

$$\Delta(s) = (I - \Gamma(s))^{-1} \quad (15.41)$$

Dalla (15.41), con semplici elaborazioni, si mostra che la (15.32) è equivalente a

$$\Gamma(s) = G_d(s)^{-1}(G_d(s) - G(s)) \quad (15.42)$$

Ricordando la (15.39), l'Equazione (15.42) definisce gli elementi $\Gamma_{ij}(s)$, $i = 1, \dots, m$, $j = 1, \dots, m$, di $\Gamma(s)$ come segue:

$$\Gamma_{ij}(s) = \begin{cases} 0 & , \quad i = j \\ -G_{ij}(s)/G_{ii}(s) & , \quad i \neq j \end{cases} \quad (15.43)$$

Naturalmente anche in questo caso può essere necessario risolvere, secondo quanto discusso in precedenza, eventuali problemi dovuti alla realizzabilità delle funzioni di trasferimento $\Gamma_{ij}(s)$.

Si osservi che il disaccoppiatore, nota $\Gamma(s)$, deve essere realizzato secondo la (15.40) che non richiede il calcolo di $\Delta(s)$. In particolare, per un sistema 2×2 si ha

$$\Gamma(s) = \begin{bmatrix} 0 & \Gamma_{12}(s) \\ \Gamma_{21}(s) & 0 \end{bmatrix} = \begin{bmatrix} 0 & -G_{12}(s)/G_{11}(s) \\ -G_{21}(s)/G_{22}(s) & 0 \end{bmatrix}$$

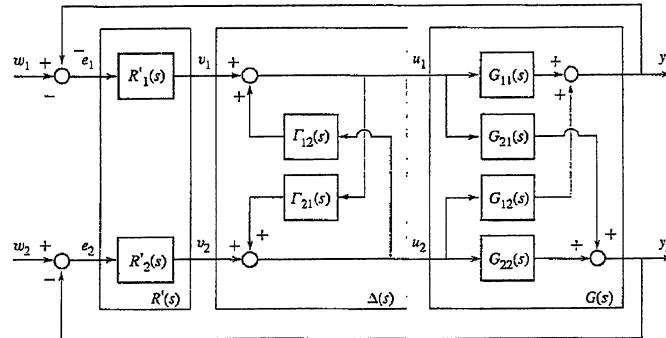


Figura 15.22
Schema di controllo con disaccoppiamento "all'indietro" di un sistema 2×2 .

e lo schema di controllo corrispondente è riportato nella Figura 15.22. Oltre alla maggiore semplicità delle funzioni $Y_1(s)/V_1(s)$ e $Y_2(s)/V_2(s)$ ottenute, lo schema di Figura 15.22 è da preferirsi a quello di Figura 15.21 in quanto i blocchi $\Gamma_{12}(s)$ e $\Gamma_{21}(s)$ sono alimentati dalle effettive variabili di controllo u_1 e u_2 . Questo consente di risolvere più facilmente le difficoltà connesse con la messa in servizio del disaccoppiatore, quali il problema della saturazione degli attuatori o quello dell'inserimento morbido della regolazione automatica (si veda il Paragrafo 14.3 e lo schema di Figura 14.15).

15.7 Controllo decentralizzato

Lo schema generale di disaccoppiamento di Figura 15.20 prevede l'impiego del regolatore $R(s) = \Delta(s)R'(s)$, i cui elementi sono in generale tutti diversi da zero, in quanto $\Delta(s)$ ha una struttura non diagonale. Esso è comunemente detto di tipo *centralizzato*, poiché *a priori* tutti gli elementi e_i , $i = 1, 2, \dots, m$, del vettore e di ingresso al regolatore concorrono a determinare ognuna delle sue uscite, cioè le componenti u_i , $i = 1, 2, \dots, m$, del vettore u .

Per il controllo dei sistemi MIMO è anche possibile utilizzare uno schema di controllo *decentralizzato*, ottenuto dalla Figura 15.20 semplicemente ponendo $\Delta(s) = I$, in cui, con ovvio significato dei simboli, risulta

$$U_i(s) = R'_i(s)E_i(s) \quad , \quad i = 1, 2, \dots, m$$

L' i -esimo regolatore agisce quindi su una singola variabile di controllo osservando solamente l'errore associato alla i -esima variabile controllata.

Lo schema di controllo decentralizzato dei sistemi 2×2 è riportato nella Figura 15.23.

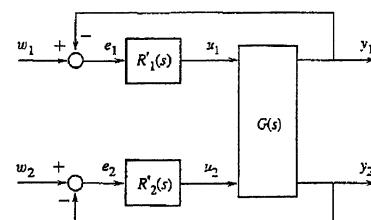


Figura 15.23
Schema di controllo decentralizzato.

Si potrebbe pensare di progettare i regolatori SISO $R'_i(s)$ con riferimento alle singole funzioni di trasferimento $G_{ii}(s)$. Tuttavia, anche se ogni regolatore $R'_i(s)$, $i = 1, \dots, m$, garantisce la stabilità e le prestazioni di un sistema di controllo in retroazione SISO in cui il processo è descritto da $G_{ii}(s)$, non è detto che il sistema multivariabile complessivo che si ottiene sia asintoticamente stabile e caratterizzato da buone prestazioni. Infatti le interazioni tra ingressi e uscite del processo possono modificare notevolmente le caratteristiche del sistema di controllo risultante.

Per tenere conto delle interazioni, un possibile procedimento consiste nel progettare sequenzialmente le singole $R'_i(s)$, $i = 1, \dots, m$, considerando anche la presenza dei regolatori SISO precedentemente determinati. Nel caso di un sistema 2×2 , il metodo consiste nel sintetizzare dapprima la funzione di trasferimento $R'_1(s)$ con riferimento a $G_{11}(s)$, quindi nel determinare $R'_2(s)$ a partire dalla funzione di trasferimento $G'_{22}(s)$, che si ha tra u_2 e y_2 quando $R'_1(s)$ è già inserito nell'anello di regolazione. È facile verificare che tale funzione è data da

$$G'_{22}(s) = G_{22}(s) - \frac{G_{12}(s)G_{21}(s)R'_1(s)}{1 + R'_1(s)G_{11}(s)} \quad (15.44)$$

Nel caso generale di sistemi $m \times m$, la procedura di sintesi del j -esimo regolatore SISO $R'_j(s)$ viene condotta a partire dalla funzione di trasferimento tra u_j e y_j che si ottiene quando i regolatori $R'_i(s)$, $i = 1, 2, \dots, j-1$, precedentemente progettati, sono già stati inseriti nello schema di controllo. Anche in questo caso tuttavia non vi è alcuna garanzia *a priori* che il procedimento conduca alla determinazione di un regolatore decentralizzato in grado di fornire le prestazioni richieste.

Nonostante le difficoltà connesse con il loro impiego, gli schemi di controllo decentralizzato sono molto utilizzati nelle applicazioni industriali in quanto implicano una maggiore semplicità del regolatore, spesso costituito da semplici PI o PID, e una minore complessità del sistema di trasmissione dati.

15.7.1 Matrice dei guadagni relativi

Per valutare la possibilità di realizzare un efficace schema di controllo decentralizzato è utile avere una misura del grado di interazione tra gli ingressi e le uscite del processo e un criterio per la determinazione degli accoppiamenti (u_i, y_j) in base ai quali progettare poi il regolatore $R'(s)$. È infatti evidente l'opportunità di regolare la generica variabile di uscita y_j con la variabile di controllo u_i che ha maggiore influenza su di essa.

Al di là di considerazioni basate sul buon senso o sull'esperienza ingegneristica, in letteratura sono disponibili tecniche piuttosto sofisticate per valutare il grado di interazione e determinare gli accoppiamenti migliori tra ingressi e uscite. Per semplicità qui si presenta soltanto il criterio empirico basato sulla *matrice dei guadagni relativi*, o *Relative Gain Array (RGA)*, che, benché non sia fondato su solidi risultati teorici, fornisce spesso utili indicazioni ed è pertanto molto usato. Si noti che la determinazione dei migliori accoppiamenti (u_i, y_j) può essere di aiuto anche nel progetto dei regolatori di disaccoppiamento $\Delta(s)$ precedentemente introdotti; infatti l'azione del disaccoppiatore è tanto più "facile" quanto meno forti sono i legami tra le variabili di controllo e quelle controllate da annullare mediante l'uso di $\Delta(s)$.

La matrice dei guadagni relativi è stata originariamente introdotta come una misura delle interazioni alla pulsazione $\omega = 0$ nei processi asintoticamente stabili e con matrice dei guadagni statici non singolare, cioè per cui sia verificata la

condizione

$$\det G(0) \neq 0$$

Supponendo che il sistema sia in una condizione di equilibrio corrispondente a un dato valore costante delle variabili di controllo, se si dà una variazione a scalino di ampiezza δu_i alla generica componente d'ingresso u_i , ogni uscita y_j a regime ha una variazione δy_{jAA} rispetto al suo valore di equilibrio e risulta

$$\frac{\delta y_{jAA}}{\delta u_i} = G_{ji}(0) = g_{ji}$$

In altri termini, g_{ji} è il guadagno della funzione di trasferimento in anello aperto tra u_i e y_j . Si supponga ora di dare ancora la medesima variazione δu_i a u_i , e che, agendo opportunamente sulle altre variabili di controllo, a regime tutte le uscite tranne la componente y_j , che subisce una variazione δy_{jAC} , assumano ancora il valore precedente alla variazione stessa. Si definisca quindi il "guadagno in anello chiuso"

$$h_{ji} = \frac{\delta y_{jAC}}{\delta u_i}$$

È ora possibile introdurre il *guadagno relativo* della coppia (u_i, y_j) come

$$\lambda_{ji} = \frac{g_{ji}}{h_{ji}} \quad (15.45)$$

e la matrice dei guadagni relativi

$$\Lambda = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_{m1} & \dots & \lambda_{mm} \end{bmatrix}$$

Si dimostra che il calcolo di Λ può essere effettuato semplicemente mediante la formula

$$\Lambda = G(0) \odot (G(0)^{-1})^T \quad (15.46)$$

dove il simbolo \odot rappresenta il *prodotto elemento per elemento*, o *prodotto di Schur*, e l'apice indica la trasposizione di matrice.

La correttezza della (15.46) può essere facilmente verificata nel caso di un sistema 2×2 . Infatti per definizione risulta

$$\begin{bmatrix} \delta y_{1AC} \\ 0 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} \delta u_1 \\ \delta u_{2AC} \end{bmatrix} \quad (15.47)$$

dove δu_{2AC} è la variazione da impostare a u_2 perché δy_2 sia nulla quando si applica δu_1 . Supponendo per semplicità $g_{22} \neq 0$, dalla (15.47) è immediato ricavare

$$h_{11} = \frac{\delta y_{1AC}}{\delta u_1} = \frac{\det G(0)}{g_{22}}$$

e quindi dalla (15.45)

$$\lambda_{11} = \frac{g_{11}g_{22}}{\det G(0)}$$

che corrisponde all'elemento (1, 1) della (15.46). Analoghi ragionamenti si possono applicare per ricavare gli altri elementi di Λ .

La matrice (15.46) possiede varie interessanti proprietà, alcune delle quali sono di seguito elencate:

- la somma degli elementi di una generica riga è pari a 1;
- la somma degli elementi di una generica colonna è pari a 1;
- i suoi elementi sono indipendenti dalle unità di misura usate in fase di modellizzazione di $G(s)$;
- essa coincide con la matrice identità se $G(s)$ è triangolare (superiore o inferiore) e quindi, in particolare, se $G(s)$ è diagonale.

Nei sistemi 2×2 , dalla definizione e da queste proprietà segue che

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

In questo caso è sufficiente quindi un unico parametro λ per descrivere completamente la matrice. Si noti inoltre che:

- se risulta $\lambda = 1$, dalla (15.45) segue che il fatto che l'uscita y_2 sia regolata o no è irrilevante ai fini del legame tra u_1 e y_1 : non vi è quindi alcuna interazione a regime tra u_1 e y_2 e tra u_2 e y_1 , gli accoppiamenti corretti sono quindi (u_1, y_1) e (u_2, y_2) ;
- se $\lambda = 0$, si devono ovviamente considerare gli accoppiamenti (u_1, y_2) e (u_2, y_1) ;
- quando $0 < \lambda < 1$, il guadagno in anello chiuso tra u_1 e y_1 è maggiore di quello in anello aperto; esiste quindi un'interazione tra le variabili che è tanto più critica quanto più λ è prossimo a 0;
- se $\lambda > 1$ c'è interazione, tanto maggiore quanto più è grande λ ;
- se $\lambda < 0$ il problema di controllo è particolarmente critico, perché la chiusura di un anello provoca il cambio del segno del guadagno nell'altro anello; per questo motivo è assolutamente sconsigliabile l'accoppiamento (u_1, y_1) .

Queste considerazioni possono essere estese al caso di un generico sistema $m \times m$ e portano a formulare il seguente criterio: è opportuno accoppiare le variabili di ingresso e di uscita in modo che i guadagni relativi corrispondenti siano positivi e il più possibile prossimi a 1.

Si osservi infine che la matrice dei guadagni relativi definita tramite la (15.46) rappresenta una misura delle interazioni a pulsazione $\omega = 0$, di particolare interesse per le prestazioni statiche del sistema. D'altra parte è possibile estendere i concetti precedentemente introdotti e analizzare le interazioni a una generica pulsazione $\bar{\omega}$.

Esempio 15.8 Si consideri il processo 2×2 descritto da

$$G(s) = \begin{bmatrix} \frac{2}{1+s} & \frac{1}{1+5s} \\ \frac{-1}{1+0.5s} & \frac{2}{1+2s} \end{bmatrix} \quad (15.48)$$

a cui, come è facile verificare, corrisponde

$$\Lambda = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

In base al criterio precedente e volendo impiegare una struttura di controllo decentralizzata, è opportuno considerare gli accoppiamenti (u_1, y_1) e (u_2, y_2) . In particolare si possono utilizzare i regolatori PI descritti da

$$U_1(s) = 0.5 \frac{1+s}{s} E_1(s), \quad U_2(s) = 0.5 \frac{1+2s}{s} E_2(s) \quad (15.49)$$

dove i simboli hanno il consueto significato. I due regolatori sono stati progettati in modo che, se venissero inseriti in due anelli di controllo singoli ognuno dei quali con funzione di trasferimento

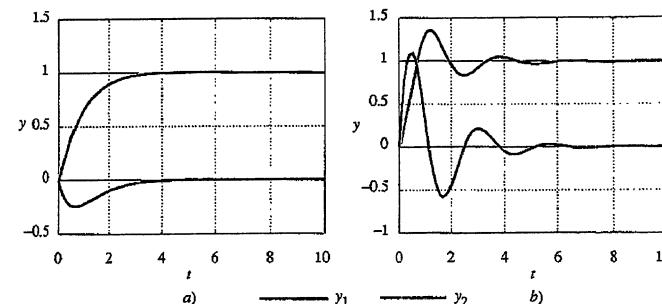


Figura 15.24
Risposta del sistema di controllo dell'Esempio 15.8 a una variazione a scalino di u_1 : a) con i regolatori (15.49); b) con i regolatori (15.50).

del processo data dagli elementi (1, 1) e (2, 2) della (15.48), produrrebbero $\omega_c = 1$ e $\varphi_m = 90^\circ$. La risposta del sistema di controllo risultante a uno scalino unitario del segnale di riferimento w_1 , con $w_2 = 0$, è riportata nella Figura 15.24a. I transitori mostrano che l'andamento di y_1 è molto simile a quello che si otterrebbe se gli elementi $G_{12}(s)$ e $G_{21}(s)$ di $G(s)$ fossero nulli e che la variazione indotta in y_2 è abbastanza contenuta.

Contrariamente a quanto suggerito dal criterio basato sull'analisi di Λ , si considerino ora gli accoppiamenti (u_1, y_2) e (u_2, y_1) e si utilizzino i regolatori descritti da

$$U_1(s) = -\frac{1+0.5s}{s} E_2(s), \quad U_2(s) = \frac{1+5s}{s} E_1(s) \quad (15.50)$$

ancora progettati per ottenere $\omega_c = 1$ e $\varphi_m = 90^\circ$, nell'ipotesi fittizia di assenza dei legami (u_1, y_1) e (u_2, y_2) . Sottponendo il corrispondente sistema di controllo alla medesima variazione di w_1 , con $w_2 = 0$, si ottengono i transitori riportati nella Figura 15.24b. Il confronto di questi risultati con quelli di Figura 15.24a evidenzia che l'effetto delle interazioni in questo caso è molto più sensibile.

Esempio 15.9 Si consideri il sistema descritto da

$$G(s) = \begin{bmatrix} \frac{1}{1+s} & \frac{2}{1+5s} \\ \frac{1}{1+0.5s} & \frac{1}{1+2s} \end{bmatrix} \quad (15.51)$$

a cui corrisponde

$$\Lambda = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$$

Poiché in questo caso l'elemento (1, 1) di Λ è negativo, il progetto di un sistema di controllo decentralizzato basato sugli accoppiamenti (u_1, y_1) e (u_2, y_2) è da evitarsi. A conferma di ciò, il sistema di controllo che si ottiene impiegando il regolatore decentralizzato descritto da

$$U_1(s) = \frac{1+s}{s} E_1(s), \quad U_2(s) = \frac{1+2s}{s} E_2(s) \quad (15.52)$$

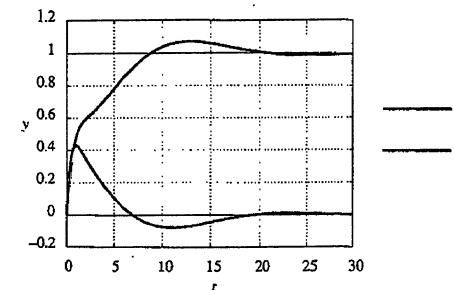


Figura 15.25
Risposta del sistema (15.51), (15.53) a uno scalino di u_1 nell'Esempio 15.9.

è instabile. Nella (15.52) i regolatori SISO sono stati progettati, come nell'esempio precedente, per ottenere $\omega_c = 1$ e $\varphi_m = 90^\circ$ nell'ipotesi che le funzioni di trasferimento tra le coppie (u_1, y_2) e (u_2, y_1) siano nulle.

Si consideri invece il regolatore

$$U_1(s) = \frac{1 + 0.5s}{s} E_2(s) \quad U_2(s) = 0.5 \frac{1 + 5s}{s} E_1(s) \quad (15.53)$$

progettato secondo gli stessi criteri precedenti. La risposta del sistema di controllo risultante a uno scalino unitario di w_1 con $w_2 = 0$ è riportata nella Figura 15.25. È chiaro che l'effetto dell'interazione è significativo, come anticipato dagli elementi di Λ ($\lambda_{12} = \lambda_{21} = 2$). In questo caso sarebbe quindi conveniente utilizzare un regolatore di disaccoppiamento per eliminare l'effetto delle interazioni.

Anche quando l'analisi di Λ consiglia di utilizzare una particolare struttura, mediante un criterio di progetto sequenziale può essere possibile ricavare un regolatore decentralizzato e stabilizzante, come illustrato nel seguente esempio.

Esempio 15.10 (Seguito dell'Esempio 15.9). Per il sistema descritto dalla (15.51) si voglia determinare un regolatore decentralizzato progettato secondo gli accoppiamenti (u_1, y_1) e (u_2, y_2) , che, come già discusso, sono particolarmente critici. Si consideri dapprima il regolatore

$$U_1(s) = \frac{1 + s}{s} E_1(s) \quad (15.54)$$

progettato a partire da $G_{11}(s)$ per ottenere $\omega_c = 1$ e $\varphi_m = 90^\circ$. È facile a questo punto verificare che, una volta inserito nello schema di controllo il regolatore (15.54), la funzione di trasferimento $G'_{22}(s)$ tra u_2 e y_2 , data dalla (15.44), in questo caso è

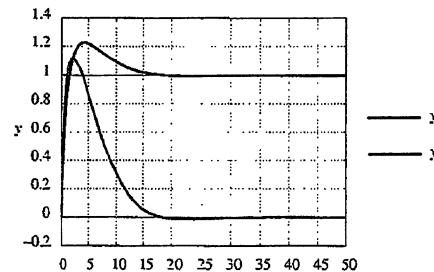
$$G'_{22}(s) = \frac{-(1 - 2.5s)(1 + s)}{(1 + 0.5s)(1 + 5s)(1 + 2s)} \quad (15.55)$$

Si osservi che, come già noto dall'analisi di Λ , i guadagni di $G_{22}(s)$ e $G'_{22}(s)$ sono di segno opposto. In base alla (15.55) si può quindi sintetizzare il regolatore

$$U_2(s) = -0.12 \frac{1 + 5s}{s} E_2(s) \quad (15.56)$$

che, con riferimento alla (15.55), consente di ottenere $\omega_c \approx 0.12$ e $\varphi_m \approx 62^\circ$. La risposta del sistema di controllo risultante a uno scalino unitario di w_1 con $w_2 = 0$ è riportata nella Figura 15.26. Il sistema retroazionato è quindi asintoticamente stabile (contrariamente a quanto ottenuto nell'esempio precedente), anche se le sue prestazioni non sono molto soddisfacenti a causa del notevole effetto dell'interazione.

Figura 15.26
Risposta del sistema (15.51), (15.54), (15.56) a una variazione a scalino di w_1 nell'Esempio 15.10.



15.8 Conclusioni

Nella illustrazione e nello studio degli schemi di controllo avanzati presentati in questo capitolo sono state largamente impiegate le tecniche in frequenza basate sul criterio di Bode, già sviluppate nei Capitoli 9-11. È chiaro quindi che questi metodi sono lo strumento fondamentale per l'analisi di strutture, anche complesse, per il controllo dei sistemi SISO.

Anche nell'analisi dei sistemi di controllo MIMO è spesso possibile ricorrere con successo a queste tecniche. Il controllo dei sistemi MIMO può essere affrontato con strumenti più evoluti ed efficaci di quelli qui discussi, estendendo i metodi di analisi in frequenza o facendo ricorso ad algoritmi basati su modelli nel dominio del tempo (per esempio, è facile impiegare anche nel caso MIMO il metodo di sintesi ad assegnamento degli autovalori presentato nel Capitolo 13). Tuttavia, gli approcci introdotti nell'ultima parte del capitolo sono utilizzati molto di frequente nel controllo di processi industriali anche complessi.

Esercizi

Esercizio 15.1 Si consideri il sistema retroazionato di Figura 15.1 in cui

$$R(s) = 2 \quad G(s) = \frac{1}{s} \quad C(s) = C_1(s) = 1$$

e $u(t) = sca(t)$.

- si valuti l'andamento di $u(t)$ e si determini il valore massimo assunto dalla variabile di controllo;
- si ripeta il punto precedente quando $C(s) = C_2(s) = \frac{1}{1+s}$ e si confronti il valore massimo di u nei due casi;
- si ripeta il punto precedente quando $C(s) = C_3(s) = \frac{1}{1+0.1s}$ e si confrontino i risultati ottenuti;
- si discuta l'eventuale riduzione della banda passante dovuta a $C_2(s)$ o $C_3(s)$ rispetto al caso $C(s) = 1$.

Esercizio 15.2 Si verifichi che lo schema per il prefiltraggio del segnale di riferimento mostrato nella Figura 15.27 fa sì che, a fronte di uno scalino w di ingresso, la sua uscita w_f abbia essenzialmente un andamento a rampa troncata con pendenza che dipende dall'elemento di saturazione. Si confronti questa soluzione con quella proposta nel Paragrafo 15.2.1 in termini di sollecitazione del sistema di controllo.

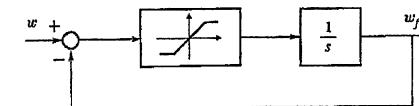


Figura 15.27
Prefiltraggio del segnale di riferimento.

Esercizio 15.3 Con riferimento allo schema di Figura 15.7, supponendo il sistema retroazionato asintoticamente stabile, $H(s) = 1$ e d'un segnale a rampa, si mostri come con un regolatore $R(s)$ di tipo PI e un compensatore statico $M(s) = M$ opportuno è possibile annullare l'errore a transitorio esaurito dovuto al disturbo. Si valuti la robustezza della regolazione a zero dell'errore così ottenuta.

Esercizio 15.4 Si consideri lo schema di Figura 15.7 in cui

$$G(s) = \frac{1}{(1 + s)^2} \quad H(s) = \frac{10}{1 + 5s}$$

e si progetti il compensatore statico M per annullare a transitorio esaurito l'effetto di un disturbo sinusoidale con pulsazione $\omega = 0.2$.

Esercizio 15.5 Dato il sistema asintoticamente stabile descritto dalla funzione di trasferimento

$$G(s) = \frac{\mu}{1 + Ts} e^{-\tau s}$$

si consideri uno schema a preditore di Smith in cui il regolatore $R'(s)$ è un PI descritto da

$$R'(s) = \frac{K(1+Ts)}{s}$$

e il ritardo è approssimato dall'espressione

$$e^{-\tau s} = 1 - \tau s$$

Si mostri che il regolatore equivalente ha ancora la struttura di un PI con funzione di trasferimento

$$R(s) = \frac{K}{1+\mu K\tau} \frac{(1+Ts)}{s}$$

e si confronti questo risultato con la taratura IMC di Tabella 14.6 con $T_f = 1/\mu K$. Si ripeta l'esercizio trovando la funzione di trasferimento $R(s)$ quando il ritardo è approssimato dalla funzione di trasferimento (15.16).

Esercizio 15.6 Con riferimento allo schema di controllo in cascata di Figura 15.15, si valuti l'opportunità di utilizzare per l'anello interno un regolatore con azione integrale (si ricordino le (15.23), (15.24)).

Esercizio 15.7 Si consideri il sistema con ingresso u e uscita y descritto da

$$\begin{aligned}\dot{x}(t) &= \alpha x(t) - u(t) \\ y(t) &= x(t-10)\end{aligned}$$

in cui x è misurabile. Nei due casi $\alpha = -1$ e $\alpha = 1$, utilizzando alcuni degli schemi presentati nel capitolo, si progetti un sistema di controllo che garantisca errore nullo a transitorio esaurito per segnali di riferimento costanti, pulsazione critica $\omega_c \geq 10$ e margine di fase $\varphi_m \geq 60^\circ$.

Esercizio 15.8 Si estenda al caso di un sistema triangolare con tre ingressi e tre uscite lo schema di disaccoppiamento presentato nel Paragrafo 15.6.1.

Esercizio 15.9 Dato il sistema con matrice di trasferimento

$$G(s) = \begin{bmatrix} \frac{1}{1+s} & 0 \\ \frac{1}{1-s} & \frac{1}{1+s} \end{bmatrix}$$

si valuti la possibilità di ottenere un sistema retroazionato asintoticamente stabile mediante uno schema di controllo decentralizzato.

Esercizio 15.10 Si verifichi tramite la (15.41) che, se $G_{12}(s) = 0$ (sistema triangolare), il disaccoppiatore risultante $\Delta(s)$ coincide con quello precedentemente ottenuto e riportato nella (15.31). Più in generale, si dimostri che se il sistema è triangolare (inferiore o superiore) il disaccoppiatore $\Delta(s)$ definito dalle (15.41) e (15.43) è triangolare.

Sistemi di controllo non lineari

16

16.1 Introduzione

Nel Paragrafo 9.2 è stato illustrato il problema del controllo di un sistema nell'intorno di una sua condizione di equilibrio per motivare lo studio approfondito dei sistemi retroazionati lineari (Capitoli 9-15). In realtà ci sono situazioni in cui, anche per affrontare lo stesso problema, risulta inevitabile tenere in esplicita considerazione la presenza di qualche non linearità all'interno dell'anello di controllo. Essa può essere relativa al processo o alla strumentazione: tipico è il caso di una saturazione all'uscita dell'attuatore. In altre circostanze una non linearità potrebbe anche essere posta intenzionalmente nel controllore, come accade per i relè.

La teoria generale dei sistemi di controllo non lineari, per la sua vastità e complessità, non può certo essere trattata qui in modo estensivo. In questa sede, invece, è possibile presentare le principali problematiche che sorgono quando nel sistema retroazionato compare un singolo elemento non lineare essenzialmente privo di dinamica, insieme ai più classici metodi per il loro studio.

Innanzitutto, verrà trattato un problema di stabilità, i cui termini precisi dovranno essere specificati in modo adeguato, visto che, in un contesto non lineare, la stabilità asintotica è per definizione una proprietà locale dei singoli movimenti di un sistema e non una proprietà globale del sistema nel suo complesso. La proprietà cui ci si riferirà è la cosiddetta stabilità assoluta, che coincide con la stabilità globale di un particolare stato di equilibrio e contiene un concetto di robustezza insito nella sua stessa definizione. Poi, si discuterà una questione tipica dei sistemi non lineari, che consiste nell'accertare l'eventuale esistenza di oscillazioni permanenti, cioè di movimenti periodici, anche in assenza di ingressi. Tra l'altro, lo studio di questi movimenti è di interesse perché essi sono proprio quelli che si ottengono quando si utilizzano i controllori a relè.

In termini specifici, i temi trattati sono i seguenti:

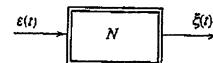
- l'introduzione e la motivazione dello schema del sistema canonico cui si farà riferimento nel resto del capitolo;
- la stabilità assoluta, e in particolare una condizione sufficiente per la sua esistenza, detta criterio del cerchio;
- il metodo della funzione descrittiva per la determinazione della presenza e delle caratteristiche delle oscillazioni permanenti;
- l'applicazione di questo metodo alla taratura automatica di un controllore PID.

16.2 Considerazioni preliminari

16.2.1 Schemi a blocchi di sistemi non lineari

In termini generali, i sistemi non lineari costituiti da sottosistemi interconnessi si possono descrivere mediante schemi a blocchi del tipo di quelli considerati nel Capitolo 5, pur di ampliare l'insieme dei componenti. Qui basta aggiungere a quelli introdotti al Paragrafo 5.2 l'elemento N riportato in Figura 16.1, che rappresenta un sistema non lineare SISO. In particolare, si suppone che il legame

Figura 16.1
Sistema non lineare.

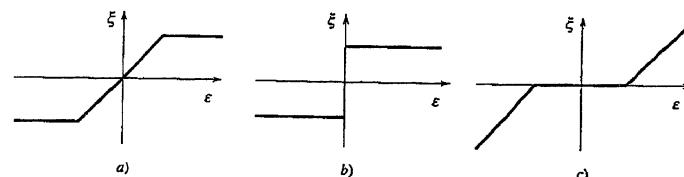


tra l'ingresso ε e l'uscita ξ sia non dinamico e descritto dalla relazione

$$\xi(t) = \varphi(\varepsilon(t)) \quad (16.1)$$

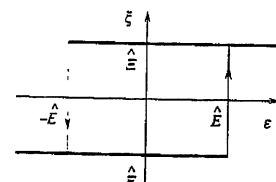
che mostra come il valore di ξ a un istante generico dipenda solo dal valore di ε allo stesso istante secondo quanto specificato dalla funzione reale φ . La Figura 16.2 riporta il grafico di tre funzioni non lineari di interesse: la *saturazione*, il

Figura 16.2
Tipici elementi non lineari:
a) saturazione; b) relè
senza isteresi; c) zona
morta.



relè senza isteresi (sostanzialmente la funzione segno, ma la definizione del valore di $\varphi(0)$ è irrilevante) e la *zona morta*. In realtà si accetta pure che N sia un elemento un poco più complesso di quelli di Equazione (16.1), quale per esempio il *relè con isteresi*, o semplicemente *relè*, descritto dalla Figura 16.3. L'uscita di tale elemento può assumere solo due valori, denotati con $-\hat{E}$ e \hat{E} , e commuta dall'uno all'altro quando l'ingresso supera le soglie $-\hat{E}$ e \hat{E} , come indicato dalle linee verticali e dalle frecce poste su di esse. Ciò specifica il comportamento del sistema durante il suo normale funzionamento. Se all'avvio l'ingresso è contenuto nell'intervallo $[-\hat{E}, \hat{E}]$, detto *finestra di isteresi*, occorre inoltre assegnare l'uscita iniziale, che non risulta determinata in modo univoco dalla Figura 16.3. In modo analogo, possono essere presi in considerazione altri elementi nei quali l'uscita può assumere solo un numero finito di valori, commutando dall'uno all'altro quando l'ingresso supera determinate soglie. Tutti i sistemi qui considerati sono detti *elementi non lineari da caratteristica*. Quest'ultima può essere a un sol valore, come nel caso dell'Equazione (16.1), o a più valori, come per il relè con isteresi di Figura 16.3; non necessariamente è a simmetria dispari come

Figura 16.3
Relè con isteresi.



in tutte le esemplificazioni precedenti (si veda a questo proposito la Figura 1.16 dell'Esempio 1.18).

16.2.2 Sistema canonico

Le tecniche presentate in questo capitolo fanno riferimento allo schema di Figura 16.4, dove il sistema lineare SISO strettamente proprio con funzione di trasferimento $\Gamma(s)$ sarà descritto anche dalla sua realizzazione minima

$$\dot{\eta}(t) = A\eta(t) + B\xi(t) \quad (16.2)$$

$$\chi(t) = C\eta(t) \quad (16.3)$$

Il sistema di Figura 16.4 viene chiamato *sistema canonico* perché, pur nella sua semplicità, può rappresentare parecchie situazioni interessanti quali le seguenti.

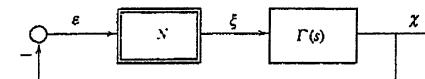


Figura 16.4
Sistema canonico.

Saturazione dell'attuatore Con riferimento alla discussione del Paragrafo 9.2, si voglia studiare il comportamento di un sistema di controllo attorno a un suo equilibrio e si assuma che tutti i componenti siano descrivibili adeguatamente con modelli linearizzati, salvo l'attuatore, per il quale sia opportuno tenere conto in modo esplicito dell'esistenza di una saturazione a valle di una funzione di trasferimento che rappresenta la sua dinamica. Supponendo che le variazioni rispetto ai valori di equilibrio di tutti i segnali esterni siano nulle e utilizzando gli stessi simboli della Figura 9.2, si perviene allora allo schema a blocchi di Figura 16.5, dove N è come in Figura 16.2a. È evidente che questo schema coincide con quello del sistema canonico pur di porre

$$\Gamma(s) = \tilde{G}(s)\tilde{T}(s)\tilde{R}(s)A(s)$$

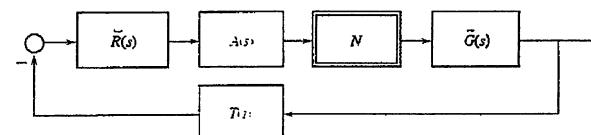


Figura 16.5
Sistema di controllo con
saturazione della variabile
manipolabile.

Controlleri a relè Si supponga di voler analizzare il comportamento di un processo controllato mediante un relè, con isteresi o no. Se il processo, l'attuatore e il trasduttore sono linearizzabili nell'intorno dello stato di equilibrio di interesse, essi, nel complesso, si possono descrivere con un'unica funzione di trasferimento $\Gamma(s)$. Allora, quando gli ingressi esterni sono nulli, l'intero sistema di controllo assume la configurazione di Figura 16.4, dove N rappresenta il relè.

16.3 Stabilità assoluta

16.3.1 Generalità

Si consideri il sistema canonico di Figura 16.4, nel quale l'elemento non lineare N è costituito da una funzione a un sol valore

$$\xi(t) = \varphi(\varepsilon(t)) \quad (16.4)$$

che soddisfa la condizione

$$\varphi(0) = 0 \quad (16.5)$$

Ricordando le Equazioni (16.2) e (16.3), che rappresentano una realizzazione minima di $\Gamma(s)$, è facile verificare che l'equazione di stato del sistema complessivo è

$$\dot{\eta}(t) = A\eta(t) + B\varphi(-C\eta(t)) \quad (16.6)$$

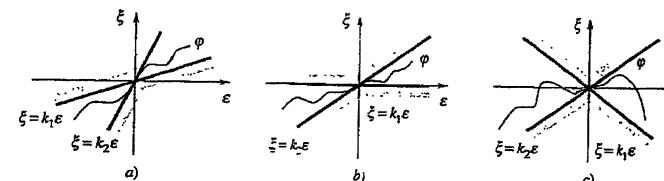
e palesemente $\eta = 0$ è un suo stato di equilibrio per la (16.5).

È allora spontaneo chiedersi quali siano le proprietà di stabilità di tale equilibrio. Nel concetto di stabilità che si vuole introdurre si fa riferimento alla sua stabilità globale, equivalente alla stabilità asintotica con corrispondente regione di attrazione pari all'intero insieme di stato (Paragrafi 2.6.1 e 2.6.2). Si assume, cioè, che lo stato iniziale del sistema canonico possa essere qualunque (quindi anche molto "lontano" dall'equilibrio $\eta = 0$) e ci si domanda se il movimento dello stato che ne consegue tende in ogni caso allo stato nullo oppure no. Inoltre, si ipotizza che, a differenza della parte lineare $\Gamma(s)$, la funzione (16.4) sia parzialmente incerta, ovvero non sia del tutto specificata. Di essa si suppone di sapere soltanto che appartiene alla particolare classe funzionale $\Phi[k_1, k_2]$, contenente tutte le funzioni φ che soddisfano l'Equazione (16.5) e la condizione

$$k_1 \leq \frac{\varphi(\varepsilon)}{\varepsilon} \leq k_2, \quad \varepsilon \neq 0$$

dove k_1 e $k_2 > k_1$ sono due scalari assegnati. In altri termini, nel piano (ε, ξ) il grafico della funzione φ è compreso tra le rette, passanti per l'origine, di coefficienti angolari k_1 e k_2 , come illustrato nella Figura 16.6 per alcuni casi particolarmente significativi. Si può allora dare la seguente definizione.

Figura 16.6
Classe funzionale $\Phi[k_1, k_2]$:
a) $k_1 > 0$; b) $k_1 = 0$;
c) $k_1 < 0$ e $k_2 > 0$.



Definizione 16.1 Il sistema canonico di Equazione (16.6) si dice *assolutamente stabile nel settore $[k_1, k_2]$* se lo stato di equilibrio $\eta = 0$ è globalmente stabile qualunque sia la funzione φ nella classe $\Phi[k_1, k_2]$. ■

Si osservi come, in questa definizione, sia perfettamente specificata solo la parte lineare del sistema considerato, mentre per la funzione φ si ipotizzi solo la sua appartenenza a una classe opportuna. Pertanto la proprietà di stabilità assoluta in un settore risulta robusta rispetto a variazioni di φ all'interno della classe stessa. D'altra parte è chiaro che la proprietà di stabilità assoluta è legata in modo indissolubile a un settore di riferimento; se questo non viene specificato essa rimane priva di significato.

Per lo studio della stabilità assoluta non sono disponibili condizioni necessarie e sufficienti. Vengono allora presentate qui di seguito una semplice condizione solo necessaria e poi una efficace condizione solo sufficiente, strettamente legate tra loro.

16.3.2 Una condizione necessaria

Condizioni necessarie di stabilità assoluta in un dato settore $[k_1, k_2]$ si possono ottenere imponendo la stabilità globale dello stato nullo per una classe di funzioni φ più ristretta di $\Phi[k_1, k_2]$. Conviene in particolare riferirsi alle funzioni lineari

$$\xi(t) = k\varepsilon(t), \quad k_1 \leq k \leq k_2$$

perché, così facendo, il sistema non lineare originario si riduce al sistema lineare di Figura 16.7, per il quale la stabilità globale dello stato di equilibrio corrisponde alla stabilità asintotica dell'intero sistema (si vedano i Paragrafi 3.4.1 e 3.4.2). Quest'ultima può essere accertata tramite la condizione necessaria e sufficiente espressa dal Corollario 9.1. Per enunciare questa condizione in una forma comoda per il seguito della trattazione, è opportuno prima introdurre la nozione di segmento $\rho[k_1, k_2]$ come l'insieme dei numeri reali $-1/k$ con $k_1 \leq k \leq k_2$, e $k \neq 0$ se del caso. Si osservi che se $k_1 = 0$ oppure $k_2 = 0$ allora ciò che è stato chiamato "segmento" $\rho[k_1, k_2]$ è in realtà una semiretta, mentre se $k_1 < 0$ e $k_2 > 0$ esso è costituito da due semirette.

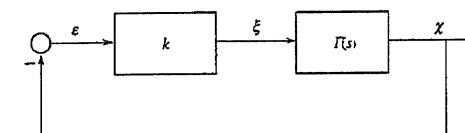


Figura 16.7
Sistema lineare usato per enunciare una condizione necessaria di stabilità assoluta.

Teorema 16.1 Condizione necessaria perché il sistema canonico di Equazione (16.6) sia assolutamente stabile nel settore $[k_1, k_2]$ è che il numero di giri percorsi in senso antiorario dal diagramma di Nyquist di $\Gamma(s)$ attorno al segmento $\rho[k_1, k_2]$ sia ben definito e risulti uguale al numero di poli di $\Gamma(s)$ a parte reale maggiore di zero. ■

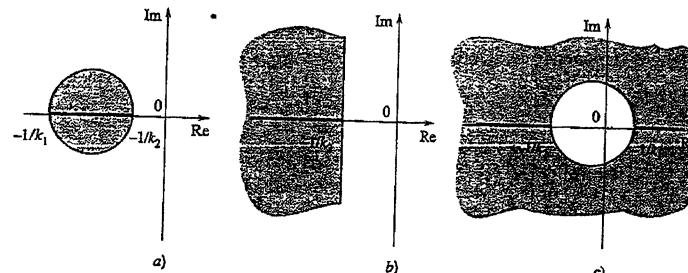
Si osservi che, quando $k_1 \leq 0 \leq k_2$, tra i sistemi in anello chiuso descritti dallo schema di Figura 16.4 vi è anche quello costituito dal solo sistema $\Gamma(s)$ in anello aperto, che corrisponde al caso in cui $k = 0$. Quindi, in questa particolare situazione, la stabilità asintotica di $\Gamma(s)$ è condizione necessaria per la stabilità assoluta.

16.3.3 Una condizione sufficiente

La condizione del Teorema 16.1 è anche sufficiente per la stabilità asintotica dello stato $\eta = 0$ qualunque sia la funzione φ nella classe $\Phi[k_1, k_2]$, purché il sistema N sia linearizzabile in $\varepsilon = 0$. Non è invece sufficiente se N non è linearizzabile e, in ogni caso, per la stabilità globale.

La determinazione di condizioni sufficienti di stabilità assoluta è molto difficile, in quanto non è possibile ricordare l'analisi del sistema non lineare di Figura 16.4 a quella di alcun sistema lineare. La non linearità, in altri termini, deve essere trattata in quanto tale. Tra i tanti risultati esistenti in materia viene qui riportato il cosiddetto *criterio del cerchio*, che costituisce, in un certo senso, un'estensione della parte sufficiente del criterio di Nyquist (si vedano il Teorema 9.1 e il Corollario 9.1). Di esso non si darà alcuna dimostrazione, vista la sua notevole complessità. Per enunciarlo è necessario introdurre prima la nozione di

Figura 16.8
Cerchio $\sigma [k_1, k_2]$: a) $k_1 > 0$;
b) $k_1 = 0$; c) $k_1 < 0$ e $k_2 > 0$.



cerchio $\sigma [k_1, k_2]$ come la regione del piano complesso che contiene il segmento $\rho [k_1, k_2]$ e delimitata dalla circonferenza che ha centro sull'asse reale e passa per i punti $-1/k_1$ e $-1/k_2$. Si osservi che $\sigma [k_1, k_2]$ è un cerchio nel senso geometrico del termine quando $\rho [k_1, k_2]$ è propriamente un segmento. Invece, se $\rho [k_1, k_2]$ è una semiretta, $\sigma [k_1, k_2]$ è un intero semipiano e, se $\rho [k_1, k_2]$ è costituito da due semirette, allora $\sigma [k_1, k_2]$ è l'intero piano complesso con l'esclusione dei punti interni alla circonferenza che delimita la regione. La Figura 16.8 illustra alcuni casi significativi.

Teorema 16.2 Condizione sufficiente perché il sistema canonico di Equazione (16.6) sia assolutamente stabile nel settore $[k_1, k_2]$ è che il numero di giri percorsi in senso antiorario dal diagramma di Nyquist di $\Gamma(s)$ attorno al cerchio $\sigma [k_1, k_2]$ sia ben definito e risulti uguale al numero di poli di $\Gamma(s)$ a parte reale maggiore di zero. ■

Considerando nell'insieme i Teoremi 16.1 e 16.2, i casi in cui non si può decidere circa l'esistenza o meno della stabilità assoluta sono solo quelli in cui ci sia un'intersezione tra il diagramma polare di $\Gamma(j\omega)$ e il cerchio $\sigma [k_1, k_2]$, senza che vi siano invece punti in comune tra il diagramma e il segmento $\rho [k_1, k_2]$, attorno al quale anzi il diagramma di Nyquist di $\Gamma(s)$ compia un numero di giri corretto.

16.3.4 Esempi di applicazione

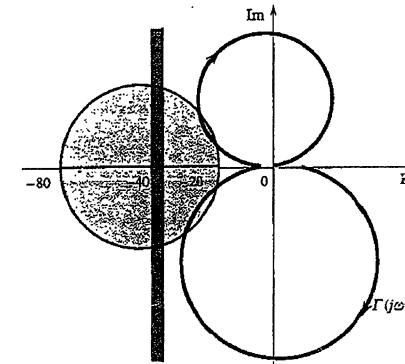
Le nozioni introdotte e, in special modo, l'uso delle condizioni espresse dai Teoremi 16.1 e 16.2 sono ora illustrati mediante alcuni esempi.

Esempio 16.1 La funzione di trasferimento

$$\Gamma(s) = \frac{100}{(s^2 + 0.2s + 1)(s^2 + 0.1s + 8)}$$

ha tutti i poli con parte reale negativa. Pertanto, per il Teorema 16.2, il sistema canonico corrispondente è assolutamente stabile nel settore $[0, 0.025]$, come mostra la Figura 16.9 che riporta il diagramma polare di $\Gamma(j\omega)$ insieme al cerchio $\sigma [0, 0.025]$ (degenero nel semipiano dei punti con parte reale minore o uguale a -40). La stessa figura indica che, sulla base dei Teoremi 16.1 e 16.2, nulla si può dire invece circa l'assoluta stabilità nel settore $[0.0125, 0.05]$ in quanto il diagramma di Nyquist non compie giri attorno al segmento $\rho [0.0125, 0.05]$, che coincide con l'intervallo $[-80, -20]$, mentre interseca il cerchio $\sigma [0.0125, 0.05]$.

Figura 16.9
Condizioni di stabilità assoluta per il sistema canonico dell'Esempio 16.1.



Esempio 16.2 La Figura 16.10 mostra che il sistema canonico con

$$\Gamma(s) = \frac{100(s+1)}{(s-1)(s+10)}$$

è assolutamente stabile nel settore $[0.125, 0.5]$. Infatti il numero di giri in senso antiorario compiuti dal diagramma di Nyquist di $\Gamma(s)$ attorno al cerchio $\sigma [0.125, 0.5]$ è unitario, così come il numero dei poli di $\Gamma(s)$ con parte reale positiva.

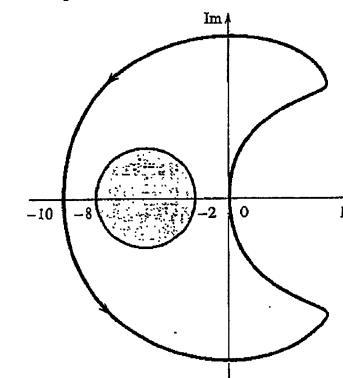


Figura 16.10
Condizione sufficiente di stabilità assoluta per il sistema canonico dell'Esempio 16.2.

L'uso del concetto di stabilità assoluta è illustrato dall'esempio seguente, nel quale si mostra come possa accadere di pervenire allo studio di un sistema privo di ingressi, come il sistema canonico.

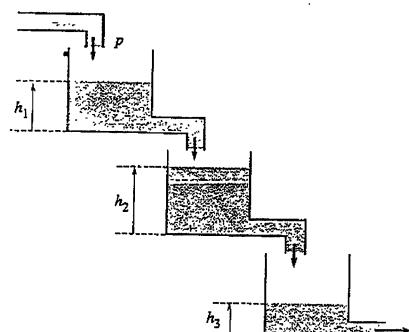
Esempio 16.3 Processo Si consideri il processo di Figura 16.11, costituito da tre serbatoi cilindrici identici posti in cascata e contenenti un liquido. Detti h_1 , h_2 e h_3 i livelli, p la portata di alimentazione del primo serbatoio, q la portata uscente dal terzo, e indicate con α e β due costanti positive, il modello complessivo è

$$\dot{h}_1(t) = -\alpha\sqrt{h_1(t)} + \beta p(t) \quad (16.7)$$

$$\dot{h}_2(t) = \alpha(\sqrt{h_1(t)} - \sqrt{h_2(t)}) \quad (16.8)$$

$$\dot{h}_3(t) = \alpha(\sqrt{h_2(t)} - \sqrt{h_3(t)}) \quad (16.9)$$

Figura 16.11
Processo dell'Esempio 16.3.



$$q(t) = \alpha\sqrt{h_3(t)} \quad (16.10)$$

Si assume poi che la portata di alimentazione debba soddisfare la condizione

$$0 \leq p(t) \leq P \quad (16.11)$$

cioè debba essere non negativa e non superiore al valore noto P .

Equilibrio Supposta costante e pari a \bar{p} la portata afferente, all'equilibrio si ha

$$\bar{h}_1 = \bar{h}_2 = \bar{h}_3 = \left(\frac{\beta}{\alpha}\right)^2 \bar{p}^2 \quad (16.12)$$

$$\bar{q} = \beta \bar{p} \quad (16.13)$$

Linearizzando il sistema attorno a questi valori, si ottiene

$$\dot{x}_1(t) = -\gamma x_1(t) + \beta u(t) \quad (16.14)$$

$$\dot{x}_2(t) = \gamma(x_1(t) - x_2(t)) \quad (16.15)$$

$$\dot{x}_3(t) = \gamma(x_2(t) - x_3(t)) \quad (16.16)$$

$$y(t) = \gamma x_3(t) \quad (16.17)$$

dove

$$u(t) = p(t) - \bar{p}$$

$$x_i(t) = h_i(t) - \bar{h}_i \quad , \quad i = 1, 2, 3$$

$$y(t) = q(t) - \bar{q}$$

$$\gamma = \frac{\alpha^2}{2\beta\bar{p}} > 0$$

Pertanto il modello (16.14)-(16.17) si può usare per studiare il sistema (16.7)-(16.10) attorno all'equilibrio (16.12), (16.13).

Per semplicità, si assume in particolare che in corrispondenza dei valori costanti di stato e uscita di interesse risulti

$$\bar{p} = P/2 \quad (16.18)$$

Controllore Si supponga ora che, per tenere il sistema attorno al valore costante \bar{w} di y , si utilizzi un regolatore puramente proporzionale con guadagno δ . Ricordando le Equazioni (16.11) (saturation dell'attuatore) e (16.18) si ha allora

$$u(t) = \psi(e(t)) \quad (16.19)$$

$$e(t) = \bar{w} - y(t) \quad (16.20)$$

dove

$$\psi(e) = \begin{cases} -\bar{p}, & e < -\bar{p}/\delta \\ \delta e, & -\bar{p}/\delta \leq e \leq \bar{p}/\delta \\ \bar{p}, & e > \bar{p}/\delta \end{cases} \quad (16.21)$$

Dalle Equazioni (16.14)-(16.21), all'equilibrio si ha

$$\bar{e} = \bar{w} - \beta\psi(\bar{e})$$

Quest'equazione ha un'unica soluzione \bar{e} qualunque sia \bar{w} e, da essa, si possono ricavare i valori costanti \bar{u} di u , \bar{x}_i di x_i , $i = 1, 2, 3$, e \bar{y} di y . Si osservi che $\bar{e} \neq 0$ per ogni $\bar{w} \neq 0$ perché nel sistema non sono presenti azioni integrali. Si assuma per semplicità che \bar{w} sia tale che risulti

$$|\bar{e}| \leq \bar{p}/\delta \quad (16.22)$$

Stabilità Le proprietà di stabilità dello stato di equilibrio appena determinato possono essere accerte verificando quali siano le proprietà dello stato nullo del sistema alle variazioni. Esso è descritto dal sistema canonico di Figura 16.4, dove

$$\Gamma(s) = \frac{\beta\gamma^3}{(s + \gamma)^3} \quad (16.23)$$

è la funzione di trasferimento del sistema (16.2), (16.3) con

$$\xi(t) = u(t) - \bar{u}$$

$$\eta_i(t) = x_i(t) - \bar{x}_i \quad , \quad i = 1, 2, 3$$

$$x(t) = y(t) - \bar{y}$$

$$A = \begin{bmatrix} -\gamma & 0 & 0 \\ \gamma & -\gamma & 0 \\ 0 & \gamma & -\gamma \end{bmatrix}, \quad B = \begin{bmatrix} \beta \\ 0 \\ 0 \end{bmatrix}, \quad C = [0 \ 0 \ \gamma]$$

Inoltre, definendo

$$\varepsilon(t) = e(t) - \bar{e} = \bar{y} - y(t) = -x(t)$$

il blocco N è descritto dalla relazione

$$\xi = \varphi(\varepsilon) = \begin{cases} -\bar{p} - \delta\bar{e}, & \varepsilon < -\bar{p}/\delta - \bar{e} \\ \delta\varepsilon, & -\bar{p}/\delta - \bar{e} \leq \varepsilon \leq \bar{p}/\delta - \bar{e} \\ \bar{p} - \delta\bar{e}, & \varepsilon > \bar{p}/\delta - \bar{e} \end{cases} \quad (16.24)$$

e rappresentato nella Figura 16.12 per un generico valore di \bar{w} tale che l'Equazione (16.22) sia soddisfatta. La stabilità asintotica dello stato nullo del sistema canonico si può verificare sostituendo la caratteristica non lineare (16.24) con la sua linearizzazione

$$\xi(t) = \delta\varepsilon(t)$$

e ragionando, per esempio con il criterio di Nyquist, sul sistema lineare di funzione di trasferimento d'anello

$$L(s) = \frac{\beta\gamma^3\delta}{(s + \gamma)^3}$$

Per accettare poi la stabilità globale dello stato di equilibrio, si può innanzitutto notare che la funzione φ appartiene alla classe $\Phi[0, \delta]$. Pertanto, se si riesce a dimostrare che il sistema canonico con $\Gamma(s)$ dato dall'Equazione (16.23) è assolutamente stabile in tale settore, allora lo stato nullo è globalmente stabile per definizione, qualunque sia la funzione non lineare nella classe $\Phi[0, \delta]$ e quindi anche per quella di Equazione (16.24). Quindi, in questo caso, non solo si dimostra quanto desiderato, ma in più si conclude che la stabilità globale è una proprietà robusta rispetto a incertezze

Figura 16.12
Elemento non lineare
di Equazione (16.24).

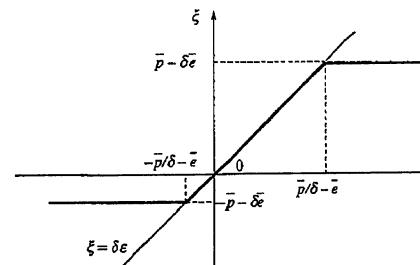
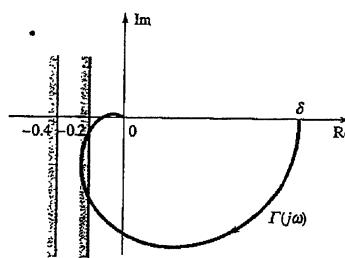


Figura 16.13
Condizioni di stabilità assoluta per il sistema canonico dell'Esempio 16.3.



sulla funzione φ : infatti, qualunque sua variazione che la lasci nella classe $\Phi[0, \delta]$ non fa venire a mancare la stabilità globale dello stato di equilibrio. Nulla si può dire se, invece, non si riesce a dimostrare la stabilità assoluta. Si veda, a questo proposito, la Figura 16.13 nella quale sono considerati i casi $\delta = 2.5$ e $\delta = 5$ per $\beta = \gamma = 1$ e, per semplicità, è riportato il diagramma di $\Gamma(j\omega)$ solo per $\omega \geq 0$. Nel primo caso la stabilità globale è dimostrata per il Teorema 16.2; nel secondo rimane indimostrabile con gli strumenti presentati.

16.4 Oscillazioni permanenti

16.4.1 Generalità

Si consideri ancora il sistema canonico di Figura 16.4, in cui l'elemento non lineare N è uno qualunque dei sistemi introdotti al Paragrafo 16.2.1, purché a simmetria dispari. L'obiettivo di questo paragrafo è quello di discutere il problema dell'esistenza di oscillazioni permanenti compatibili con il sistema canonico, determinandone, in caso affermativo, le caratteristiche principali. In altri termini, si vogliono stabilire le proprietà degli eventuali movimenti periodici dello stato del sistema canonico.

Il problema trattato non sarebbe significativo in un contesto puramente lineare. Infatti, in questo caso, la presenza di oscillazioni permanenti in un sistema privo di ingressi equivale alla stabilità non asintotica, con presenza quindi di autovalori dotati di parte reale nulla. Questa proprietà non si può dire robusta rispetto a variazioni parametriche, che potrebbero "spostare" fuori dall'asse immaginario gli autovalori con parte reale nominalmente nulla. Si vedrà invece come un certo grado di robustezza sia ottenibile grazie alla non linearità presente nel sistema canonico.

Quanto ai metodi di analisi, la ricerca dell'esistenza stessa delle soluzioni periodiche di un'equazione differenziale non lineare, com'è l'equazione di stato del sistema canonico, è un problema così complesso da giustificare per la sua soluzione anche approcci di tipo euristico, quale il *metodo della funzione descrittiva* cui è dedicato il resto di questo paragrafo. Per le sue caratteristiche, esso può anche essere utilizzato per affrontare problemi di sintesi.

16.4.2 Funzione descrittiva

Si supponga innanzitutto che l'elemento non lineare N sia sollecitato dall'ingresso sinusoidale

$$\varepsilon(t) = E \cos(\omega t) \quad (16.25)$$

definito dalla coppia di numeri reali positivi (E, ω) e si consideri l'uscita periodica corrispondente di periodo $T = 2\pi/\omega$. Se il sistema è descritto da una relazione

come la (16.1), l'uscita periodica esiste, è unica ed è data da $\xi(t) = \varphi(E \cos(\omega t))$. Quando invece la caratteristica dell'elemento non lineare N è a più valori, le uscite corrispondenti all'ingresso (16.25) possono essere molteplici; si assuma allora che tra queste ve ne sia una e una sola, periodica, almeno per determinati valori di E . Per eliminare situazioni prive di interesse si considererà solo il caso in cui le uscite periodiche cui si sta facendo riferimento non siano costanti.

Esempio 16.4 Il sistema N sia costituito dal relè senza isteresi di Figura 16.2b con

$$\xi(t) = \begin{cases} -\hat{\varepsilon}, & \varepsilon(t) < 0 \\ \hat{\varepsilon}, & \varepsilon(t) > 0 \end{cases}$$

Di conseguenza l'uscita corrispondente all'ingresso (16.25) è l'onda quadra di periodo T definita da

$$\xi(t) = \xi_a(t) = \begin{cases} -\hat{\varepsilon}, & -T/2 < t < -T/4 \\ \hat{\varepsilon}, & -T/4 < t < T/4 \\ -\hat{\varepsilon}, & T/4 < t < T/2 \end{cases} \quad (16.26)$$

Esempio 16.5 Se il sistema N è il relè con isteresi di Figura 16.3, in corrispondenza dell'ingresso (16.25) non vi è alcuna uscita periodica non costante quando $E \leq \hat{\varepsilon}$. Se invece $E > \hat{\varepsilon}$, come illustrato dalla Figura 16.14, l'unica uscita periodica è ancora un'onda quadra T -periodica definita da

$$\xi(t) = \xi_b(t) = \begin{cases} -\hat{\varepsilon}, & -T/2 < t < -i \\ \hat{\varepsilon}, & -i < t < T/2 - i \\ -\hat{\varepsilon}, & T/2 - i < t < T/2 \end{cases} \quad (16.27)$$

$$i = \frac{1}{\omega} \arccos(\hat{\varepsilon}/E) \quad (16.28)$$

Si noti che, se l'ingresso viene applicato in un istante in cui il suo modulo è minore di $\hat{\varepsilon}$, l'uscita all'inizio rimane indipendente dall'ingresso stesso e conserva il suo valore iniziale finché ε raggiunge una delle due soglie di commutazione. Pertanto, essa assume immediatamente l'andamento periodico solo quando il suo valore iniziale è coerente con quello fornito dalle (16.27), (16.28). In caso contrario prende tale forma solo dopo un intervallo di tempo finito.

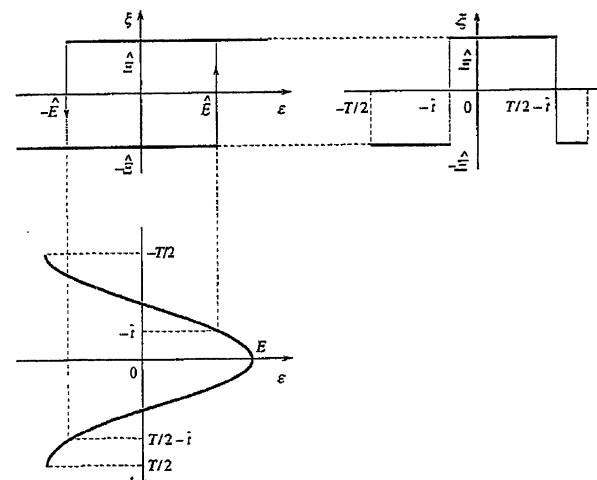


Figura 16.14
Uscita periodica del relè con isteresi sollecitato da una cosinusoida.

Sotto l'ipotesi che l'uscita del sistema N sia sviluppabile in serie di Fourier, ricordando l'Equazione (B.31) si può scrivere

$$\xi(t) = \sum_{n=1}^{+\infty} |\Xi_n(E)| \cos(n\omega t + \arg \Xi_n(E))$$

perché il valor medio della funzione $\xi(t)$ è nullo, essendo N un elemento a simmetria dispari e avendo $\varepsilon(t)$ un valor medio nullo. Inoltre, in questa relazione si è messa in evidenza la dipendenza dei coefficienti di Fourier dall'ampiezza E dell'ingresso ε e si è tenuto conto del fatto che essi, invece, non dipendono dalla sua pulsazione ω , in quanto l'uscita periodica di un qualunque elemento non lineare da caratteristica ha una "forma" che non dipende dal periodo dell'ingresso che l'ha generata. È interessante notare che, nel caso in cui N sia costituito da una funzione a un solo valore come la (16.1), i coefficienti Ξ_n , $n = 1, 2, \dots$, sono reali perché $\varphi(E \cos(\omega t))$ è una funzione pari di t .

A questo punto è possibile definire la funzione

$$D(E) = \frac{\Xi_1(E)}{E} = \frac{|\Xi_1(E)|}{E} e^{j \arg \Xi_1(E)} \quad (16.29)$$

detta *funzione descrittiva* di N . Essa "descrive" N quanto basta per risolvere il problema di calcolare la prima armonica della sua uscita periodica generata da un ingresso del tipo (16.25). In questo senso estende il concetto di risposta in frequenza di un sistema lineare stazionario. Si osservi però che, se N fosse effettivamente lineare, $D(E)$ non dipenderebbe da E . La funzione descrittiva D è in generale complessa con

$$|D(E)| = \frac{|\Xi_1(E)|}{E} \quad (16.30)$$

$$\arg D(E) = \arg \Xi_1(E) \quad (16.31)$$

ma, per quanto detto sopra, essa è reale se N è descritto da una funzione a un solo valore.

Esempio 16.6 (Seguito dell'Esempio 16.4) Lo sviluppo in serie di Fourier della funzione periodica (16.26) si ricava sulla falsariga di quanto riportato nell'Esempio B.14 (il fatto che essa non sia definita in un numero finito di istanti in un periodo è irrilevante). Risulta in particolare $\Xi_1 = \Xi_{1a} = 4\hat{\Xi}/\pi$ e allora, per la (16.29),

$$D(E) = \frac{4\hat{\Xi}}{\pi E}$$

che indica come il rapporto tra l'ampiezza della prima armonica del segnale periodico $\xi(t)$ e l'ampiezza E di $\varepsilon(t)$ (Equazione (16.25)) sia inversamente proporzionale a quest'ultima.

Esempio 16.7 (Seguito dell'Esempio 16.5) Il coefficiente di Fourier relativo alla prima armonica dell'uscita periodica (16.27), (16.28) del relè con isteresi si può calcolare notando che questa funzione non è altro che la (16.26) ritardata di $T/4 - \bar{t}$; allora, visti l'Equazione (B.24) e il risultato dell'esempio precedente, si ha

$$\begin{aligned} \Xi_1 &= \Xi_{1b} = \frac{2}{T} \int_T \xi_b(t) e^{-j\omega t} dt = \frac{2}{T} \int_T \xi_a \left(t - \frac{T}{4} + \bar{t} \right) e^{-j\omega t} dt = \\ &= e^{-j\omega(\frac{T}{4} - \bar{t})} \frac{2}{T} \int_T \xi_a(\theta) e^{-j\omega\theta} d\theta = e^{-j(\frac{\pi}{2} - \omega\bar{t})} \Xi_{1a} = \frac{4\hat{\Xi}}{\pi} e^{-j(\frac{\pi}{2} - \omega\bar{t})} = \\ &= \frac{4\hat{\Xi}}{\pi} \left[\cos\left(\frac{\pi}{2} - \omega\bar{t}\right) - j \sin\left(\frac{\pi}{2} - \omega\bar{t}\right) \right] = \frac{4\hat{\Xi}}{\pi} [\sin(\omega\bar{t}) - j \cos(\omega\bar{t})] = \end{aligned}$$

$$\begin{aligned} &= \frac{4\hat{\Xi}}{\pi} [\sin(\arccos(\hat{E}/E)) - j \cos(\arccos(\hat{E}/E))] = \\ &= \frac{4\hat{\Xi}}{\pi} \left[\sqrt{1 - (\hat{E}/E)^2} - j \hat{E}/E \right] \end{aligned}$$

Pertanto

$$D(E) = \frac{4\hat{\Xi}}{\pi E} \left[\sqrt{1 - (\hat{E}/E)^2} - j \hat{E}/E \right], \quad E > \hat{E}$$

Gli esempi di calcolo appena riportati avevano lo scopo di illustrare il concetto stesso di funzione descrittiva. Tuttavia, per le applicazioni riguardanti le non linearità più comuni, non è necessario calcolare volta per volta le funzioni descrittive, vista la disponibilità di tabelle, quali la Tabella 16.1, che fa riferimento agli elementi di Figura 16.15. Si vedrà tra breve che spesso è utile saper ricavare il valore di E che corrisponde a un dato valore di D . In questo caso si preferisce riferirsi a rappresentazioni grafiche delle funzioni descrittive.

N	$D(E)$
Relè senza isteresi (Figura 16.15a)	$\frac{4\hat{\Xi}}{\pi E}$
Saturazione (Figura 16.15b)	$\begin{cases} \frac{\hat{\Xi}}{\hat{E}} & , E \leq \hat{E} \\ \frac{\hat{\Xi}}{\hat{E}} \frac{2}{\pi} \left[\arcsin\left(\frac{\hat{E}}{E}\right) + \left(\frac{\hat{E}}{E}\right) \sqrt{1 - \left(\frac{\hat{E}}{E}\right)^2} \right] & , E > \hat{E} \end{cases}$
Zona morta (Figura 16.15c)	$k \left[1 - \frac{2}{\pi} \left[\arcsin\left(\frac{\hat{E}}{E}\right) + \left(\frac{\hat{E}}{E}\right) \sqrt{1 - \left(\frac{\hat{E}}{E}\right)^2} \right] \right] , E > \hat{E}$
Relè senza isteresi con zona morta (Figura 16.15d)	$\frac{4\hat{\Xi}}{\pi E} \sqrt{1 - \left(\frac{\hat{E}}{E}\right)^2} , E > \hat{E}$
Relè con isteresi (Figura 16.15e)	$\frac{4\hat{\Xi}}{\pi E} \sqrt{1 - \left(\frac{\hat{E}}{E}\right)^2} - j \frac{\hat{E}}{E} , E > \hat{E}$

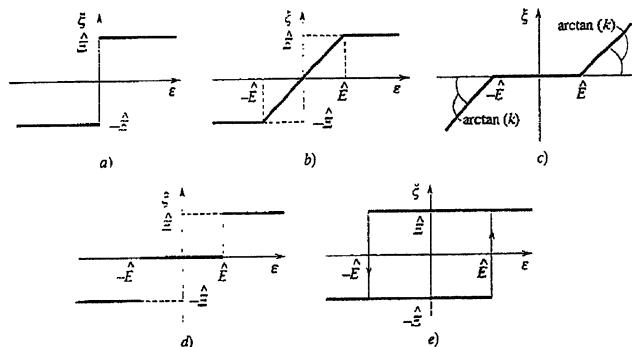


Tabella 16.1
Elementi non lineari e corrispondenti funzioni descrittive.

Figura 16.15
Elementi non lineari da caratteristica considerati nella Tabella 16.1: a) relè senza isteresi; b) saturazione; c) zona morta; d) relè senza isteresi con zona morta; e) relè con isteresi.

16.4.3 Metodo della funzione descrittiva

Si assuma che sia compatibile con il sistema canonico di Figura 16.4 un movimento periodico di pulsazione $\omega > 0$, o di periodo $T = 2\pi/\omega$; in particolare, la variabile ξ abbia valor medio nullo e possieda lo sviluppo in serie di Fourier

$$\xi(t) = \sum_{n=1}^{+\infty} |\Xi_n| \cos(n\omega t + \arg \Xi_n)$$

Ipotizzando che la funzione di trasferimento $\Gamma(s)$ non abbia alcun polo con parte reale nulla, per il Teorema 6.6 lo sviluppo in serie del corrispondente movimento periodico dell'uscita x è

$$x(t) = \sum_{n=1}^{+\infty} |\Gamma(jn\omega) \Xi_n| \cos(n\omega t + \arg \Gamma(jn\omega) + \arg \Xi_n)$$

Se in questa equazione risulta

$$|\Gamma(jn\omega) \Xi_n| \ll |\Gamma(j\omega) \Xi_1| , \quad n = 2, 3, \dots \quad (16.32)$$

allora si può ritenere che $x(t)$ sia approssimativamente costituita solo dalla prima armonica. L'Equazione (16.32) si dice *ipotesi dell'azione filtrante* ed è alla base del metodo della funzione descrittiva per l'accertamento dell'esistenza e dei parametri delle oscillazioni permanenti in un sistema canonico. Essa non è verificabile in modo diretto sui dati del problema ($\Gamma(s)$ e N), ma dipende anche dalla sua soluzione ($\Xi_n, n = 1, 2, \dots$). Solo quando, assumendo senz'altro soddisfatta la condizione (16.32), si sarà giunti a determinare le caratteristiche che le oscillazioni dovrebbero possedere, si potrà discutere la plausibilità di ciò che si è ipotizzato. E proprio di "plausibilità" si deve parlare, perché la condizione di "molto minore" contenuta nell'Equazione (16.32) fa comunque sì che l'ipotesi dell'azione filtrante sia definita in modo sfumato. Tutto ciò rende euristico il metodo che si sta trattando, nel senso che esso non è rigorosamente fondato su un piano formale.

Vale la pena di ribadire che la condizione (16.32) non significa ipotizzare che il movimento periodico di ξ sia di tipo puramente sinusoidale, cosa che avrebbe molto poco senso (se, per esempio, N è un relè, $\xi(t)$ è un'onda quadra non certo approssimabile con una sinusode). Invece, almeno se la risposta in frequenza $\Gamma(j\omega)$ della parte lineare è di tipo passa-basso, è realistico supporre che $x(t)$ sia approssimativamente sinusoidale. A ogni modo, il soddisfacimento dell'Equazione (16.32) dipende dalle caratteristiche di $\xi(t)$ e di $\Gamma(s)$ in maniera congiunta.

Portando al limite l'ipotesi dell'azione filtrante, supponendo cioè

$$\Gamma(jn\omega) \Xi_n = 0 , \quad n = 2, 3, \dots$$

si ha

$$x(t) = |\Gamma(j\omega) \Xi_1| \cos(\omega t + \arg \Gamma(j\omega) + \arg \Xi_1)$$

e di conseguenza

$$\varepsilon(t) = -x(t) = |\Gamma(j\omega) \Xi_1| \cos(\omega t + \arg \Gamma(j\omega) + \arg \Xi_1 + \pi) \quad (16.33)$$

Siccome l'origine dell'asse dei tempi si può fissare in maniera del tutto arbitraria, senza perdere la generalità si può pensare di sceglierla in modo che $\arg \Xi_1$ sia tale che valga la relazione

$$\arg \Gamma(j\omega) + \arg \Xi_1 + \pi = 0 \quad (16.34)$$

e quindi in definitiva ε abbia la forma

$$\varepsilon(t) = E \cos(\omega t) , \quad E > 0 \quad (16.35)$$

Pertanto, l'ipotesi dell'azione filtrante implica che il segnale periodico all'ingresso del blocco non lineare sia una pura sinusode, cosicché N si può descrivere mediante la sua funzione descrittiva D , almeno al fine di calcolare la prima armonica di $\xi(t)$.

Equazione di congruenza Perché un'oscillazione come la (16.35) sia compatibile con il sistema canonico, essa deve rispettare tutte le condizioni poste dagli elementi che compaiono in questo e pertanto, procedendo a ritroso lungo l'anello, per le (16.30), (16.31), (16.33) si trova quanto segue:

$$\begin{aligned} \varepsilon(t) &= E \cos(\omega t) = |\Gamma(j\omega) \Xi_1| \cos(\omega t + \arg \Gamma(j\omega) + \arg \Xi_1 + \pi) = \\ &= |\Gamma(j\omega) D(E)| E \cos(\omega t + \arg \Gamma(j\omega) + \arg D(E) + \pi) \end{aligned}$$

Insieme con l'Equazione (16.34) che si può riscrivere nella forma

$$\arg \Gamma(j\omega) + \arg D(E) + \pi = 0$$

deve pertanto risultare

$$|\Gamma(j\omega) D(E)| = 1$$

e perciò

$$1 + \Gamma(j\omega) D(E) = 0$$

Questa discussione può allora essere conclusa con la seguente affermazione, enunciata sotto forma di proposizione invece che di teorema perché la dimostrazione precedente è basata sull'ipotesi euristica dell'azione filtrante.

Proposizione 16.1 Se la funzione di trasferimento $\Gamma(s)$ è priva di poli con parte reale nulla, sotto l'ipotesi dell'azione filtrante (16.32) il sistema canonico può ammettere un'oscillazione della forma

$$\varepsilon(t) = \bar{E} \cos(\bar{\omega}t) \quad (16.36)$$

se la coppia $(\bar{E}, \bar{\omega})$ è una soluzione dell'equazione

$$1 + \Gamma(j\omega) D(E) = 0 \quad (16.37)$$

La relazione (16.37) costituisce un'equazione di congruenza; essa viene detta *equazione pseudocaratteristica* a causa della sua somiglianza con l'equazione caratteristica di un sistema retroazionato lineare: come già notato, la funzione descrittiva estende all'elemento non lineare il concetto di risposta in frequenza, e quindi $\Gamma(j\omega) D(E)$ sostituisce la risposta in frequenza d'anello.

Si noti che la (16.37) è una sorta di condizione necessaria per l'esistenza di oscillazioni permanenti, per giunta basata su un'ipotesi euristica. Pertanto, l'effettiva esistenza di oscillazioni, la loro forma e il valore dei corrispondenti parametri dovrebbero essere confermati da altri tipi di analisi. Per esempio, si potrebbe verificare *a posteriori* la validità dell'ipotesi dell'azione filtrante calcolando lo spettro del movimento periodico di $\xi(t)$, e quindi quello di $x(t)$, conseguenti all'oscillazione individuata per ε (però si osservi che, nel fare tale calcolo, N

non può essere rappresentato dalla semplice funzione descrittiva). Allora, vista la difficoltà d'uso dei metodi analitici esatti, la prassi più comune consiste nel rivolgersi alla simulazione, cosa semplice almeno per le oscillazioni che, godendo della proprietà di stabilità di cui si discuterà nel paragrafo seguente, sono le uniche di interesse nelle applicazioni. Si può comunque affermare che le conclusioni cui si perviene mediante il metodo della funzione descrittiva sono molto spesso confermate nella loro sostanza: tipicamente, ciò accade quando l'ampiezza della più significativa armonica di $\varepsilon(t)$ successiva alla fondamentale non è maggiore del 5 ÷ 10% dell'ampiezza di quest'ultima.

Per quanto riguarda la risoluzione dell'equazione pseudocaratteristica, si osservi innanzitutto che essa equivale a due equazioni reali in due incognite. Tuttavia, vista la sua natura non lineare, ciò non significa che esistano semplici condizioni per l'esistenza e l'unicità della soluzione. Nemmeno esistono formule risolutive in forma chiusa, e perciò di solito essa viene risolta per via numerica. Dovendo procedere in questa maniera, può essere utile sfruttare una semplice interpretazione grafica dell'equazione. Innanzitutto, visto che si è interessati al caso in cui $D(E) \neq 0$, si può porre

$$\Lambda(E) = -\frac{1}{D(E)}$$

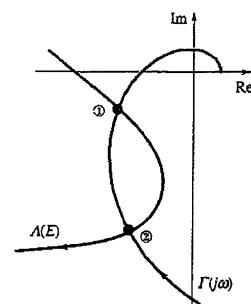
per cui l'Equazione (16.37) prende la forma

$$\Gamma(j\omega) = \Lambda(E)$$

Quindi, è immediato rendersi conto del fatto che ogni soluzione $(\bar{E}, \bar{\omega})$ di quest'ultima equazione corrisponde a un'intersezione tra il diagramma polare di $\Gamma(j\omega)$ e la rappresentazione nel piano complesso di $\Lambda(E)$, detta *luogo dei punti critici*. Il valore \bar{E} dell'ampiezza E può essere letto sulla curva che rappresenta Λ , che è punteggiata in E ; quello $\bar{\omega}$ della pulsazione ω sul diagramma di Γ , che è punteggiato in ω . Per esempio, la Figura 16.16 riporta un caso in cui l'equazione pseudocaratteristica ha due soluzioni (E_1, ω_1) e (E_2, ω_2) , corrispondenti ai punti indicati con 1 e 2.

È ragionevole attendersi che, se i diagrammi di Λ e Γ si intersecano, essi continuino a farlo anche in presenza di piccole variazioni dovute a perturbazioni dei relativi sistemi. Pertanto, l'esistenza di un'oscillazione permanente gode di un certo grado di robustezza. Si confronti questa conclusione con quanto affermato in precedenza a proposito delle oscillazioni permanenti compatibili con sistemi lineari privi di ingressi.

Figura 16.16
Soluzione grafica
dell'equazione
pseudocaratteristica.



16.4.4 Stabilità delle oscillazioni

Una volta che si sia accertata la compatibilità con il sistema canonico di un'oscillazione permanente, è interessante accettare se essa gode di qualche proprietà di stabilità. Le difficoltà che si incontrano se si cerca di affrontare il problema in maniera rigorosa sono ancora superiori a quelle della ricerca pura e semplice delle soluzioni periodiche dell'equazione di stato. Pertanto, riferendosi a un concetto di stabilità specifico, definito in modo informale, è del tutto ragionevole affidarsi ancora a una tecnica euristica di verifica, che si può formulare rimanendo nell'ambito del metodo della funzione descrittiva e sarà qui enunciata senza darne alcuna giustificazione.

Si supponga innanzitutto che il sistema canonico ammetta l'oscillazione permanente caratterizzata dall'Equazione (16.36), dove la coppia $(\bar{E}, \bar{\omega})$ è una soluzione dell'equazione pseudocaratteristica (16.37). Essa si dice *asintoticamente stabile* se si ottengono movimenti di ε che differiscono "poco" da quello nominale e tendono a coincidere con esso in modo asintotico, eventualmente a meno di uno sfasamento temporale, pur di assumere che lo stato iniziale del sistema abbia variazioni "sufficientemente piccole" rispetto a quello che produce il movimento (16.36). Si osservi che tale definizione coincide solo in apparenza con la Definizione 2.6 di asintotica stabilità applicata a un movimento periodico. Qui infatti si accetta che tra il movimento nominale e ogni movimento perturbato asintotico esista uno sfasamento, cosicché le due oscillazioni nell'insieme di stato abbiano la stessa forma, ma siano invece differenti i tempi in cui le variabili di stato assumono i singoli valori.

Si definiscono ora due vettori spiccati dal punto del piano complesso in cui i diagrammi di Λ e Γ si intersecano, indicando così la presenza della soluzione $(\bar{E}, \bar{\omega})$ dell'Equazione (16.37): il vettore \vec{t} è tangente al diagramma di Λ e orientato nel senso dei valori crescenti di E ; il vettore \vec{n} è ortogonale al diagramma di Γ e orientato verso destra per chi lo percorre nel senso dei valori crescenti di ω (si veda la Figura 16.17 per un esempio). Vale allora il seguente risultato.

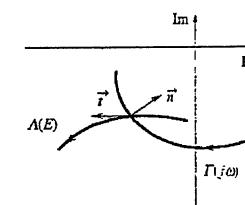


Figura 16.17
Vettori \vec{t} ed \vec{n} usati per
la verifica di stabilità
dell'oscillazione.

Proposizione 16.2 L'oscillazione di Equazione (16.36) è asintoticamente stabile se e solo se il prodotto scalare tra i vettori \vec{t} ed \vec{n} è negativo. cioè

$$\vec{t} \times \vec{n} < 0$$

Pertanto, la stabilità asintotica di un'oscillazione dipende dal fatto che i due vettori \vec{t} ed \vec{n} formino un angolo maggiore di 90° oppure no. Per esempio, nel caso di Figura 16.17 l'oscillazione è asintoticamente stabile; invece, nel caso di Figura 16.16 solo l'oscillazione che corrisponde al punto 2 è asintoticamente stabile, mentre quella corrispondente al punto 1 non lo è. Si deve comunque ricordare ancora che il metodo di verifica applicato è di tipo empirico e quindi soggetto a errori, almeno in casi critici.

16.4.5 Esempi di applicazione

In questo paragrafo verranno presentati alcuni esempi di uso del metodo della funzione descrittiva sia in problemi di analisi sia in problemi di sintesi.

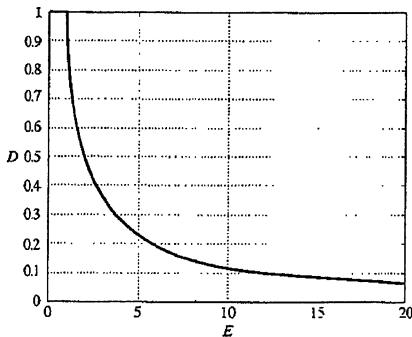
Esempio 16.8 Si consideri il sistema canonico, dove N è dato dalla saturazione di Tabella 16.1 con $\hat{E} = \hat{\Sigma} = 1$ e

$$\Gamma(s) = \frac{\mu}{(1+s)^3}, \quad \mu > 0$$

È facile verificare che "in piccolo" l'origine dell'insieme di stato è asintoticamente stabile se e solo se $\mu < 8$. A questo risultato conduce infatti l'applicazione di una qualunque condizione necessaria e sufficiente di stabilità asintotica al sistema linearizzato ottenuto sostituendo a N la semplice relazione di identità tra ε e ξ (si veda a questo proposito l'Esempio 9.6). Con la stessa tecnica utilizzata al termine dell'Esempio 16.3 si può poi accettare che lo stato nullo è addirittura globalmente stabile se $\mu < 4$. Infatti, si osservi dapprima che N appartiene alla classe $\Phi[0, 1]$; inoltre il minimo della parte reale di $\Gamma(j\omega)$ è pari a -0.25μ e quindi, in base al Teorema 16.2, il sistema canonico è assolutamente stabile nel settore $[0, 1]$.

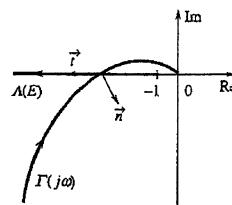
Se $\mu = 8$, il diagramma polare di $\Gamma(j\omega)$ passa per il punto -1 : l'origine dell'insieme di stato è stabile non asintoticamente e il sistema linearizzato ha due autovalori sull'asse immaginario in $s = \pm j\sqrt{3}$. Quindi il sistema linearizzato può sostenere oscillazioni sinusoidali di pulsazione $\bar{\omega} = \sqrt{3}$ e ampiezza qualunque. Allo stesso risultato si giunge applicando il metodo della funzione descrittiva, che tuttavia, nell'ambito non lineare, indica in 1 il massimo valore assumibile dall'ampiezza (si veda la Tabella 16.1).

Figura 16.18
Funzione descrittiva della saturazione unitaria.



Infine, se $\mu > 8$, il punto di equilibrio diventa instabile, come si può dedurre dalla presenza nel sistema linearizzato di un autovalore a parte reale positiva; anche in questo caso il metodo della funzione descrittiva indica l'esistenza di un'oscillazione permanente. Per dimostrarlo si inizi con l'osservare in Figura 16.18 il diagramma della funzione descrittiva in oggetto. Da esso si deduce che il luogo dei punti critici giace sull'asse reale tra il punto -1 e $-\infty$ (Figura 16.19). Poiché il diagramma polare di $\Gamma(j\omega)$ attraversa il semiasse reale negativo a sinistra del punto -1 , l'equazione pseudocaratteristica ammette una soluzione $(\bar{E}, \bar{\omega})$ con $\bar{\omega} = \sqrt{3}$: infatti la pulsazione per cui la parte immaginaria di $\Gamma(j\omega)$ è nulla non dipende da μ . Per quanto riguarda l'ampiezza, si può innanzitutto calcolare che $\Gamma(j\sqrt{3}) = -\mu/8$. Allora deve essere $D(\bar{E}) = -1/\Lambda(\bar{E}) = -1/\Gamma(j\sqrt{3}) = 8/\mu$. Quindi per ogni valore di μ si determina $D(\bar{E})$ e dal diagramma di Figura 16.18 si giunge a \bar{E} . Per esempio per $\mu = 16$ si trova $\bar{E} \approx 2$.

Figura 16.19
Equazione pseudocaratteristica per l'Esempio 16.8 con $\mu > 8$.



Visto che il luogo dei punti critici è percorso da destra verso sinistra all'aumentare di E , mentre il diagramma polare di $\Gamma(j\omega)$ attraversa il semiasse reale negativo dal basso verso l'alto, l'oscillazione è asintoticamente stabile in base alla Proposizione 16.2 (si veda ancora la Figura 16.19). Nella Figura 16.20 si mostra l'andamento esatto della variabile ε per $\mu = 16$ e per una condizione iniziale assegnata. Si può notare come i risultati trovati descrivano con efficacia quanto accade in realtà per ciò che riguarda il periodo dell'oscillazione (pari a circa 3.8 invece che a $2\pi/\sqrt{3} \approx 3.6$), mentre l'ampiezza risulta un poco sottostimata.

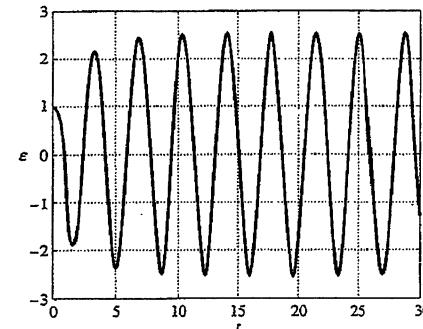


Figura 16.20
Oscillazione del sistema dell'Esempio 16.8 con $\mu = 16$.

L'esempio che segue mostra come il metodo della funzione descrittiva possa essere utilizzato anche per il progetto, in particolare quando si voglia adottare un relè come controllore di un dato processo. Con riferimento a quest'ultimo caso, si noti che nelle applicazioni ciò che conta è soprattutto l'ampiezza dell'oscillazione, mentre la pulsazione è bene solo che non sia troppo elevata per evitare un'usura eccessiva degli organi di attuazione; la forma dell'oscillazione si può poi quasi sempre ritenere inessenziale.

Esempio 16.9 Si consideri ancora il sistema canonico, nel quale questa volta N è il relè con istersi di Tabella 16.1 con $\hat{E} = \hat{\Sigma} = 1$ e

$$\Gamma(s) = R(s)G(s), \quad G(s) = \frac{5}{(1+s)^2} \quad (16.38)$$

dove $G(s)$ rappresenta la funzione di trasferimento del sistema sotto controllo, mentre $R(s)$ è da determinare in modo che il sistema sostenga un'oscillazione permanente asintoticamente stabile caratterizzata da

$$\varepsilon(t) = 2 \cos(0.6t)$$

L'equazione pseudocaratteristica (16.37) implica che la soluzione del problema soddisfi la condizione

$$R(j0.6) \frac{5}{(1+j0.6)^2} = \Lambda(2) = -\frac{1}{D(2)}$$

da cui, ricordando quanto riportato nella Tabella 16.1,

$$R(j0.6) \approx 0.43e^{-j1.54}$$

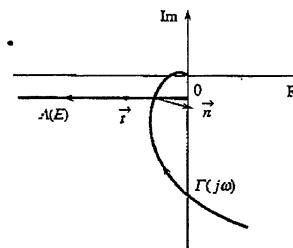
Questa relazione può, per esempio, essere verificata dal regolatore del primo ordine

$$R(s) = \frac{\mu}{1+s\tau}, \quad \mu > 0, \tau > 0$$

pur di porre

$$|R(j0.6)| = \frac{\mu}{\sqrt{1+(0.6\tau)^2}} = 0.43, \quad \arg R(j0.6) = -\arctan(0.6\tau) = -1.54$$

Figura 16.21
Equazione pseudocaratteristica per l'Esempio 16.9.



La seconda di queste relazioni impone $\tau \approx 54.1$ e quindi dalla prima si ricava $\mu \approx 14$; pertanto il metodo della funzione descrittiva suggerisce di adottare il regolatore

$$R(s) = \frac{14}{1 + 54.1s} \quad (16.39)$$

Le rappresentazioni di $\Lambda(E)$ e $\Gamma(j\omega)$ sono riportate nella Figura 16.21 insieme ai vettori r ed \hat{r} , che mostrano come l'oscillazione sia asintoticamente stabile. Si osservi in particolare che il luogo dei punti critici è una semiretta caratterizzata da parte immaginaria negativa indipendente da E e parte reale che, all'aumentare di E , varia tra 0 e $-\infty$; infatti, per un generico relè con isteresi risulta

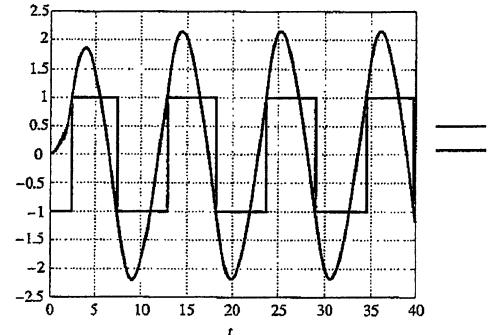
$$\Lambda(E) = -\frac{\pi}{4\hat{E}} \left(\sqrt{E^2 - \hat{E}^2} + j\hat{E} \right), \quad E > \hat{E} \quad (16.40)$$

Conviene ora accettare mediante una simulazione che i risultati trovati siano corretti. Di fatto dalla Figura 16.22, che riporta anche il grafico di $\xi(t)$, si deduce che ε ha un andamento quasi sinusoidale con ampiezza 2.14 e pulsazione 0.56.

Il metodo della funzione descrittiva permette quindi di pervenire a una soluzione del problema decisamente accettabile. Ciò avviene perché, in questo caso, l'ipotesi dell'azione filtrante è "ben verificata". Infatti, a partire dall'andamento di ξ nella figura, ricordando lo sviluppo dell'onda quadra dell'Esempio B.14 e le Equazioni (16.38) e (16.39), si può calcolare che l'ampiezza della terza armonica di $\varepsilon(t)$ è circa il 3.7% dell'ampiezza della prima, mentre la seconda è nulla.

Il progetto potrebbe essere raffinato facendo variare i parametri del regolatore attorno ai valori già individuati e valutando le prestazioni via via ottenute mediante l'aiuto della simulazione.

Figura 16.22
Oscillazione di ε e ξ nel sistema dell'Esempio 16.9.

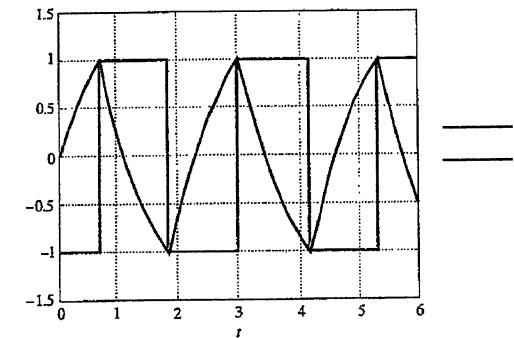


È istruttivo trattare anche un caso in cui il metodo della funzione descrittiva fornisce indicazioni errate.

Esempio 16.10 Si consideri il sistema canonico, nel quale N sia di nuovo il relè con isteresi di Tabella 16.1 e

$$\Gamma(s) = \frac{\mu}{1 + s\tau} \quad \mu > \frac{\hat{E}}{\hat{\Xi}}, \quad \tau > 0$$

Figura 16.23
Oscillazione di ε e ξ nel sistema dell'Esempio 16.10.



Visto che $\text{Re}(\Gamma(j\omega)) > 0$ per ogni ω finito, mentre $\text{Re}(\Lambda(E)) \leq 0$ per ogni $E \geq \hat{E}$ per l'Equazione (16.40), l'equazione pseudocaratteristica non ammette soluzioni. Invece il sistema è capace di sostenere un'oscillazione asintoticamente stabile, come si può dimostrare risolvendo le equazioni che ne governano il funzionamento a partire da una condizione iniziale qualunque. La Figura 16.23 riporta una simulazione relativa al caso $\hat{E} = \tilde{E} = 1$, $\mu = 2$ e $\tau = 1$.

Si può osservare che l'errore nel quale induce il metodo della funzione descrittiva è causato dal fatto che l'ipotesi dell'azione filtrante in questo caso non si può ritenere soddisfatta neppure in prima approssimazione, come dimostra nella Figura 16.23 l'andamento di ε , ben diverso da una sinusoide. A maggior conferma, dal calcolo del movimento si può intanto verificare che la pulsazione dell'oscillazione è

$$\bar{\omega} = \frac{\pi}{\tau \ln \Delta}, \quad \Delta = \frac{\mu \tilde{E} + \hat{E}}{\mu \tilde{E} - \hat{E}}$$

Quindi, ricordando ancora i risultati dell'Esempio B.14, semplici calcoli permettono di maggiornare il rapporto $\gamma_{3,1}$ tra le ampiezze della terza e della prima armonica di $\varepsilon(t)$, che è quello più interessante perché la seconda armonica è nulla. Si trova così

$$\begin{aligned} \gamma_{3,1} &= \frac{|\Gamma(j3\bar{\omega})| \Xi_3}{|\Gamma(j\bar{\omega})| \Xi_1} = \frac{1}{3} \sqrt{\frac{(\ln \Delta)^2 + \pi^2}{(\ln \Delta)^2 + 9\pi^2}} = \\ &= \frac{1}{3} \sqrt{1 - \frac{8\pi^2}{(\ln \Delta)^2 + 9\pi^2}} > \frac{1}{3} \sqrt{1 - \frac{8\pi^2}{9\pi^2}} = \frac{1}{3} \end{aligned}$$

Il valore ricavato corrisponde al limite di $\gamma_{3,1}$ per $\mu \rightarrow +\infty$, mentre si può anche verificare che in ogni caso si ha

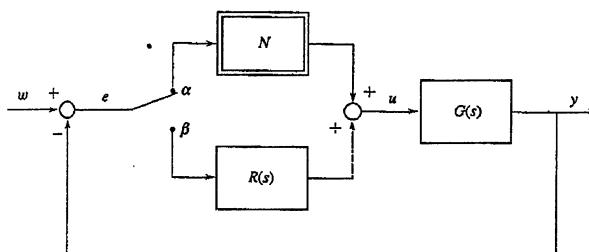
$$\frac{1}{9} < \gamma_{3,1} < \frac{1}{3}$$

Si confronti quanto sopra con l'indicazione del Paragrafo 16.4.3 secondo cui le conclusioni raggiunte con il metodo della funzione descrittiva si devono ritenere affidabili solo se il rapporto tra le ampiezze della più significativa armonica trascurata (sulla base dell'ipotesi dell'azione filtrante) e della fondamentale è al massimo $0.05 \div 0.1$.

16.4.6 Taratura automatica di un PID

Il metodo della funzione descrittiva può essere utilizzato per tarare un controllore PID mediante la tecnica di Ziegler e Nichols in anello chiuso descritta al Paragrafo 14.4.1. Per rendersi conto di come ciò sia possibile, si consideri lo schema a blocchi di Figura 16.24; in esso $G(s)$ è la funzione di trasferimento del sistema sotto controllo, che si assume non abbia poli con parte reale nulla, e $R(s)$ è la funzione di trasferimento del PID; inoltre N è un relè senza isteresi con funzione descrittiva $D(E) = 4\hat{E}/\pi E$ (Tabella 16.1). Nello schema compare anche

Figura 16.24
Schema per la taratura automatica di un PID.



un commutatore che permette di chiudere l'anello su N (posizione α) oppure su $R(s)$ (posizione β). La posizione α consente di determinare i parametri necessari per la taratura, cioè il guadagno critico \bar{K}_P e il periodo \tilde{T} dell'oscillazione che si otterrebbe utilizzando il controllore proporzionale $K_P = \bar{K}_P$; la posizione β corrisponde al normale funzionamento in automatico.

Con il commutatore in posizione α e $w = 0$, assumendo che il diagramma polare della risposta in frequenza $G(j\omega)$ attraversi il semiasse reale negativo tagliandolo dal basso verso l'alto, il metodo della funzione descrittiva indica la presenza di un'oscillazione permanente asintoticamente stabile, perché il luogo dei punti critici occupa tutto il semiasse reale negativo (Figura 16.25). Se la coppia $(\tilde{E}, \tilde{\omega})$ corrisponde al punto di intersezione, si ha

$$e(t) \simeq \tilde{E} \cos(\tilde{\omega}t)$$

Sotto l'ipotesi dell'azione filtrante, tale coppia deve risolvere l'equazione pseudocaratteristica e quindi

$$G(j\tilde{\omega}) = \Lambda(\tilde{E}) = -\frac{\pi \tilde{E}}{4\hat{E}}$$

Si supponga ora di mandare il sistema in oscillazione e rilevare in modo sperimentale i valori di \tilde{E} e $\tilde{\omega}$, magari tramite una misura del periodo \tilde{T} . È immediato rendersi conto del fatto che il periodo dell'oscillazione generata dal relè è proprio quello dell'oscillazione ("lineare") richiesta dal metodo di taratura, che si può avere solo per $\arg G(j\omega) = -180^\circ$; allora risulta

$$\tilde{T} = \tilde{\tau} = \frac{2\pi}{\tilde{\omega}}$$

In altri termini il punto A di Figura 14.18 coincide con il punto A di Figura 16.25. Inoltre, ricordando che il guadagno critico coincide con il margine di guadagno di $G(s)$, si ricava

$$\bar{K}_P = \frac{1}{|G(j\tilde{\omega})|} = \frac{4\hat{E}}{\pi \tilde{E}}$$

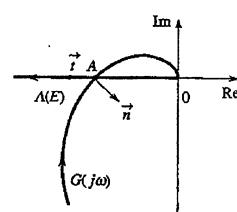


Figura 16.25
Equazione pseudocaratteristica per la taratura di un PID.

Una volta noti \bar{K}_P e \tilde{T} , la Tabella 14.1 fornisce i valori dei parametri del controllore e, assegnati questi ultimi a $R(s)$, si può porre il commutatore nella posizione β .

La tecnica di rilevamento dei parametri qui descritta risulta preferibile a quella originaria del Paragrafo 14.4.1 perché, pur continuando a richiedere che si instauri un'oscillazione, non impone di portare al limite di stabilità il processo retroazionato in maniera lineare, com'è invece necessario secondo l'esperimento proposto da Ziegler e Nichols, con il rischio che si giunga addirittura a una situazione di instabilità. Oltre a tutto, un'accorta scelta del parametro \hat{E} permette di limitare a piacimento l'ampiezza dell'oscillazione, mentre essa è praticamente imprevedibile in ambito lineare.

Il metodo proposto prevede che il diagramma polare di $G(j\omega)$ attraversi il semiasse reale negativo, così come del resto l'approccio di Ziegler e Nichols. Se ciò non accade, la tecnica illustrata può essere ancora impiegata pur di utilizzare un relè con isteresi. Per tale estensione si rimanda alla letteratura specializzata.

In chiusura, conviene notare che esistono in commercio molte apparecchiature, dette *PID ad autosintonia*, che, proprio secondo lo schema di principio di Figura 16.24, sono in grado di procedere in maniera autonoma alla taratura dei parametri di un PID.

16.5 Conclusioni

In questo capitolo sono stati trattati i problemi principali inerenti i sistemi di controllo che contengono un singolo elemento non lineare da caratteristica. Essi consistono nel verificare la stabilità globale dell'equilibrio in condizioni di incertezza circa la non linearità e nell'accertare l'esistenza di oscillazioni permanenti. Si è fatto riferimento alle tecniche di studio più tradizionali, che, oltre all'analisi, consentono la sintesi per tentativi, in analogia con quanto accade per le tecniche classiche del controllo dei sistemi lineari.

Molteplici sono le estensioni disponibili delle questioni e dei metodi trattati. Solo a titolo di esempio, si citano la possibilità di affrontare il problema della stabilità globale di un movimento non costante con tecniche interpretabili come derivazioni del criterio del cerchio, e quella di studiare le oscillazioni permanenti in sistemi con ingressi non nulli mediante una modifica del metodo della funzione descrittiva che tiene conto di un valore medio delle oscillazioni diverso da zero.

Esercizi

Esercizio 16.1 Si rifletta sul fatto che, in generale, la stabilità assoluta nei settori $[k_1, k_2]$ e $[k_2, k_3]$ non implica quella nel settore $[k_1, k_3]$.

Esercizio 16.2 Si consideri il sistema canonico di Equazione (16.6) con

$$\Gamma(s) = \frac{10}{(1+s)^3} R(s)$$

Si determini $R(s)$, di tipo zero, guadagno unitario e grado relativo nullo, in modo che esso sia assolutamente stabile nel settore $[0, 1]$.

Esercizio 16.3 Si mostri che la verifica della stabilità assoluta in un generico settore $[k_1, k_2]$ può sempre essere ricondotta a quella nel particolare settore $[0, k]$, con $k = k_2 - k_1$, pur di sostituire all'elemento lineare $\Gamma(s)$ la funzione di trasferimento

$$\tilde{\Gamma}(s) = \frac{\Gamma(s)}{1 + \Gamma(s)}$$

(Suggerimento: si noti che $\xi(t) = \varphi(\varepsilon(t)) = \varphi(\varepsilon(t)) - k_1 \varepsilon(t) + k_1 \varepsilon(t) = \varphi(\varepsilon(t)) - k_1 \varepsilon(t) - k_1 X(t)$.)

Esercizio 16.4 Oltre che con il Teorema 16.2, l'esistenza della proprietà di stabilità assoluta in un settore del tipo $[0, k]$ può essere accertata pure mediante il criterio di Popov, più potente, anche se meno semplice da utilizzare. Esso afferma che il sistema canonico di Equazione (16.6) è assolutamente stabile nel settore $[0, k]$ se il sistema (16.2)-(16.3) è asintoticamente stabile e inoltre esiste un numero reale q tale che

$$\frac{1}{k} + \operatorname{Re}((1+j\omega q)\Gamma(j\omega)) > 0, \quad \omega \geq 0 \quad (16.41)$$

Si dia un'interpretazione grafica della (16.41) in termini della posizione relativa assunta nel piano complesso tra il diagramma della funzione

$$\Gamma^*(j\omega) = \operatorname{Re}(\Gamma(j\omega)) + j\omega \operatorname{Im}(\Gamma(j\omega)), \quad \omega \geq 0$$

e una retta opportuna.

Successivamente, si mostri che il sistema canonico di Equazione (16.6) è assolutamente stabile nel settore $[0, k]$, qualunque sia $k > 0$ e qualunque sia $\Gamma(s)$ della forma

$$\Gamma(s) = \frac{\mu}{1+as+bs^2}, \quad \mu > 0, \quad a > 0, \quad b > 0$$

Esercizio 16.5 Si considerino la zona morta e il relè senza isteresi con zona morta di Figura 16.15. Si ricavino le funzioni descrittive riportate nella Tabella 16.1. Quindi si traccino i grafici di D in funzione di E e si individuino le regioni del piano complesso nelle quali giacciono i corrispondenti luoghi dei punti critici.

Esercizio 16.6 Si verifichi che il luogo dei punti critici per il relè con isteresi di Tabella 16.1 e Figura 16.15 è una semiretta caratterizzata da parte immaginaria negativa indipendente da E e parte reale che, all'aumentare di E , varia tra 0 e $-\infty$.

Esercizio 16.7 Si consideri il sistema canonico, dove N è dato dal relè senza isteresi con zona morta di Figura 16.15 con $\hat{E} = \hat{S} = 1$ e

$$\Gamma(s) = \frac{\mu}{(1+s)(1+0.4s)(1+2s)}, \quad \mu > 0$$

Si individui l'insieme dei valori di μ per i quali esistono oscillazioni permanenti compatibili con il sistema. Quindi, per alcuni valori di μ in tale insieme, si determinino le caratteristiche principali delle oscillazioni, cioè l'ampiezza \tilde{E} e la pulsazione $\tilde{\omega}$ di $s(t)$, precisando se esse sono asintoticamente stabili.

Esercizio 16.8 Si consideri il sistema canonico, dove N è dato dalla zona morta di Figura 16.15 con $\hat{E} = k = 1$. Si determini una funzione di trasferimento $\Gamma(s)$ in modo che sia compatibile con il sistema un'oscillazione permanente asintoticamente stabile.

(*Suggerimento:* dall'Esercizio 16.5, si ricordino il grafico di D in funzione di E e la regione del piano complesso nella quale giace il corrispondente luogo dei punti critici.)

Esercizio 16.9 Si consideri il sistema canonico, dove N è dato dalla saturazione di Figura 16.15 con $\hat{E} = \hat{S} = 1$ e

$$\Gamma(s) = \frac{\mu(1-s)}{(1+s)^2} R(s), \quad \mu > 0$$

Dapprima, per $R(s) = 1$, si trovi un valore di μ per il quale sia compatibile con il sistema un'oscillazione permanente, e se ne precisino le caratteristiche. Quindi, in corrispondenza di tale valore di μ , si determini $R(s)$, di tipo zero, guadago unitario e grado relativo nullo, in modo che con il sistema non siano più compatibili oscillazioni permanenti. Si discutano, infine, le proprietà di stabilità dello stato nullo del sistema.

(*Suggerimento:* si ricordino il grafico di D in funzione di E di Figura 16.18 e la regione del piano complesso nella quale giace il corrispondente luogo dei punti critici di Figura 16.19.)

Esercizio 16.10 Utilizzando il metodo della funzione descrittiva, si tari un controllore PID mediante la tecnica di Ziegler e Nichols in anello chiuso descritta al Paragrafo 14.4.1 per il processo descritto dalla funzione di trasferimento

$$G(s) = \frac{30}{(1+0.1s)(1+0.25s)(1+0.4s)}$$

Sistemi di controllo digitale: analisi e sintesi a tempo continuo

17

17.1 Introduzione

La disponibilità di sistemi di elaborazione con prestazioni crescenti e a costi sempre più contenuti fa sì che la maggior parte dei sistemi di controllo sia ormai realizzata con tecnologia digitale. L'impiego di calcolatori programmabili consente una grande flessibilità e la facoltà di realizzare algoritmi di controllo anche complicati, nei quali, per esempio, elaborate funzioni matematiche possono coesistere con sequenze di istruzioni logiche. Inoltre, negli apparati digitali è semplice integrare le funzionalità proprie di un sistema di controllo con altre, quali il monitoraggio delle variabili di interesse, l'impiego di metodi automatici per la conduzione ottimale del processo, la supervisione e la diagnostica. Poiché gli apparati digitali sono in grado di elaborare unicamente segnali di tipo discreto, mentre il processo è usualmente un sistema a tempo continuo, i sistemi di controllo digitale appartengono alla categoria dei sistemi ibridi.

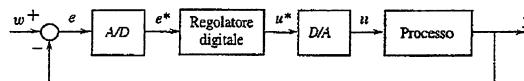
Nel capitolo saranno considerati i sistemi di controllo digitale di sistemi SISO e si introdurranno i principali elementi che li costituiscono. In dettaglio, saranno trattati i temi seguenti:

- l'analisi delle strutture di controllo digitale più comunemente impiegate;
- la descrizione delle caratteristiche dei dispositivi necessari alla realizzazione dei sistemi di controllo digitale: il convertitore analogico/digitale (A/D), o campionatore, e il convertitore digitale/analogico (D/A), o mantenitore;
- lo studio delle problematiche connesse con il campionamento di un segnale a tempo continuo e del mantenimento di un segnale a tempo discreto;
- l'analisi dei sistemi di controllo nel dominio della frequenza effettuata mediante lo studio di un sistema retroazionato a tempo continuo opportunamente definito;
- la scelta del periodo di campionamento con cui devono operare i convertitori per la realizzazione di un sistema di controllo ibrido;
- lo studio delle tecniche di discretizzazione per la determinazione di un regolatore realizzato con tecnologia digitale, o regolatore digitale, che emuli il comportamento di un regolatore a tempo continuo di riferimento, chiamato nel seguito anche regolatore analogico, progettato secondo i criteri discussi nei Capitoli 11, 12 e 14;
- gli schemi realizzativi di regolatori digitali per tener conto dei problemi connessi con il fenomeno del wind-up.

17.2 Schemi di controllo digitale

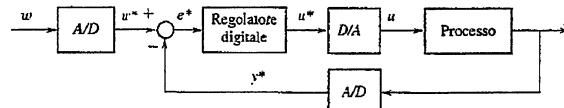
Lo schema di un possibile sistema di controllo digitale è riportato nella Figura 17.1, dove, come di consueto, $u(t)$ è la variabile di controllo, $y(t)$ è la variabile di uscita, $w(t)$ è il riferimento ed $e(t) = w(t) - y(t)$ rappresenta l'errore. Poiché il *regolatore digitale* può ricevere in ingresso soltanto sequenze di numeri, è necessario prevedere la presenza del *convertitore analogico/digitale A/D* che trasformi il segnale a tempo continuo $e(t)$ nel segnale a tempo discreto $e^*(k)$; tale blocco è detto anche *campionatore*. Analogamente il *convertitore digitale/analogico D/A*, detto anche *mantenitore*, trasforma il segnale a tempo discreto $u^*(k)$, generato dal regolatore digitale, in quello a tempo continuo $u(t)$.

Figura 17.1
Schema di controllo digitale a campionamento dell'errore.



I sistemi di controllo digitale possono essere realizzati secondo schemi diversi da quello della Figura 17.1, nel quale il campionamento è effettuato su e . Per esempio, come mostrato nella Figura 17.2, è possibile eseguire il campionamento separatamente su y e w , per ottenere i segnali a tempo discreto y^* e w^* e calcolare successivamente l'errore campionato $e^*(k) = u^*(k) - y^*(k)$. Nello schema di Figura 17.2, molto spesso il segnale di riferimento è originato direttamente come un segnale a tempo discreto w^* e il campionamento si effettua soltanto sull'uscita y del processo. In tutti gli schemi è comunque necessaria la presenza di uno o più campionatori e del mantenitore. Per semplicità, nel seguito si assumerà che il sistema sotto controllo sia SISO e si presenteranno le caratteristiche dei convertitori e del regolatore digitale facendo riferimento allo schema della Figura 17.1. Le considerazioni che saranno svolte potranno essere replicate anche per lo schema della Figura 17.2 con ovvia sostituzione di simboli.

Figura 17.2
Schema di controllo digitale a campionamento dell'uscita e del riferimento.



17.2.1 Campionatore

Il campionatore *ideale* e *periodico*, utilizzato nel seguito e denotato negli schemi a blocchi con una delle rappresentazioni della Figura 17.3, è un convertitore che, dato in ingresso il segnale a tempo continuo $e(t)$, produce in uscita il segnale a tempo discreto

$$e^*(k) = e(kT_s + \tau_s) \quad (17.1)$$

dove T_s è il *periodo* (o *passo*, o *intervallo*) di campionamento, e rappresenta la durata dell'intervalllo di tempo continuo che intercorre tra gli istanti in cui vengono prelevati due campioni successivi. A partire da T_s si possono definire anche la

Figura 17.3
Rappresentazioni grafiche del campionatore.

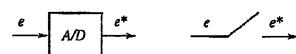
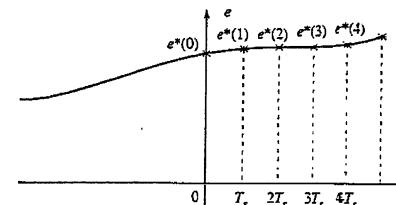


Figura 17.4
Campionamento periodico.



frequenza f_s e la pulsazione ω_s di campionamento come

$$f_s = \frac{1}{T_s} \quad , \quad \omega_s = \frac{2\pi}{T_s}$$

Nella (17.1) il parametro τ_s ($0 \leq \tau_s < T_s$) tiene conto del fatto che tra gli istanti di campionamento può essere assente l'istante $t = 0$; tuttavia, poiché la trattazione sviluppata nel seguito farà riferimento unicamente al controllo di sistemi stazionari per i quali l'origine dell'asse dei tempi può essere fissata arbitrariamente, d'ora in poi per semplicità si porrà $\tau_s = 0$ senza ledere la generalità. Il significato dell'Equazione (17.1), con $\tau_s = 0$, è rappresentato nella Figura 17.4.

Problema fondamentale del campionamento Il problema più importante connesso con l'impiego di segnali campionati è dovuto alla perdita di informazione che l'operazione di campionamento può comportare. Infatti è chiaro che *a priori* la stessa sequenza $e^*(k)$ può essere generata da diverse funzioni a tempo continuo $e(t)$ campionate con periodo T_s , come mostrato nel seguente esempio.

Esempio 17.1 Si consideri la famiglia di segnali

$$e_h(t) = \sin((\alpha + 2h)\pi t)$$

in cui α è un numero reale e h è un intero. Campionando i segnali e_h con periodo $T_s = 1$, si ha

$$e^*(k) = e_h(kT_s) = \sin((\alpha + 2h)\pi k) = \sin(\alpha\pi k) = e_0(kT_s)$$

cioè, indipendentemente dal valore di h , il segnale risultante è comunque quello che si ottiene dal campionamento di e_0 . Pertanto, dall'analisi del segnale campionato $e^*(k)$ non è possibile risalire univocamente al segnale a tempo continuo $e_h(t)$ che lo ha generato. A titolo di esempio, nella Figura 17.5 sono riportati i segnali $e_0(t) = \sin(0.6\pi t)$, $e_{-1}(t) = \sin(-1.4\pi t)$ e il segnale campionato e^* da essi ottenuto.

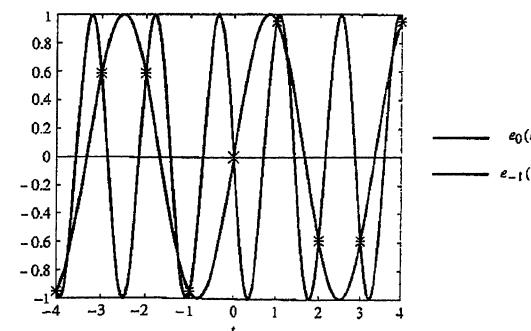


Figura 17.5
Campionamento di $e_0(t) = \sin(0.6\pi t)$, $e_{-1}(t) = \sin(-1.4\pi t)$ con $T_s = 1$.

Ritornando all'analisi dello schema della Figura 17.1, affinché il regolatore digitale possa operare correttamente, è indispensabile che l'errore campionato e^* rappresenti in modo completo e non ambiguo l'effettivo andamento di e . È necessario, cioè, che il campionamento non produca una perdita di informazione. Le problematiche connesse con il campionamento saranno discusse più a fondo nel Paragrafo 17.3.

Quantizzazione Il campionatore qui considerato è ideale in quanto si suppone che negli istanti di campionamento il valore di e^* coincida perfettamente con quello di e . In realtà, nei sistemi digitali i valori reali e^* sono rappresentati con *parole* di lunghezza finita, cioè con un numero finito di *cifre binarie*, o *bit*. Il campionamento implica quindi inevitabilmente un'approssimazione, detta *quantizzazione*, dovuta alla codifica del segnale. A causa della quantizzazione, che può essere effettuata per *troncamento* o per *arrotondamento*, non sono distinguibili dopo il campionamento valori di e che differiscono tra loro di una quantità inferiore al livello di quantizzazione impiegato, cioè al cosiddetto *zero macchina*.

La quantizzazione è un fenomeno intrinsecamente non lineare e difficilmente analizzabile in modo rigoroso. Qui ci si limita a dire che il suo effetto sul comportamento di un sistema di controllo è tanto maggiore quanto minore è la lunghezza della parola con cui si codifica e^* e dipende fortemente dal fatto che si usi una rappresentazione *in virgola fissa* o *in virgola mobile*. Inoltre il fenomeno è tanto più critico quanto più piccolo è T_s .

17.2.2 Mantenitore

Il mantenitore, di solito rappresentato con uno dei simboli riportati nella Figura 17.6, ha il compito di trasformare il segnale a tempo discreto $u^*(k)$ nel segnale a



tempo continuo $u(t)$. Questa operazione può essere effettuata in diverse maniere, tra le quali si considererà il caso più semplice in cui

$$u(t) = u^*(k) \quad , \quad kT_m + \tau_m \leq t < (k+1)T_m + \tau_m \quad (17.2)$$

dove T_m è il *periodo* (o *passo*, o *intervallo*) di *mantenimento*, o *aggiornamento*. A partire da T_m si possono definire anche la *frequenza* f_m e la *pulsazione* ω_m di *mantenimento* come

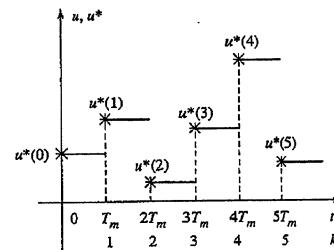
$$f_m = \frac{1}{T_m} \quad , \quad \omega_m = \frac{2\pi}{T_m}$$

Nella (17.2) il parametro τ_m ($0 \leq \tau_m < T_m$) tiene conto del fatto che tra gli istanti di aggiornamento di $u(t)$ può essere assente l'istante $t = 0$.

Il mantenitore (17.2) viene chiamato *mantenitore di ordine zero*, o *ZOH* (Zero Order Hold), in quanto mantiene per $u(t)$ un andamento polinomiale di ordine zero, cioè costante, compatibile con l'ultimo valore di $u^*(k)$. Il funzionamento dello ZOH per $\tau_m = 0$ è rappresentato nella Figura 17.7.

Per l'analisi delle caratteristiche di uno schema di controllo digitale è necessario esaminare l'effetto prodotto sul sistema dalla presenza del mantenitore. A questo proposito, il legame prodotto dallo ZOH tra le trasformate di Fourier di u^* e u sarà esaminato nel Paragrafo 17.4.

Figura 17.7
Mantenimento di ordine zero.



17.2.3 Regolatore digitale

Il *regolatore digitale* è un elaboratore che calcola *in tempo reale* la sequenza $u^*(k)$ in funzione della sequenza $e^*(k)$. In questo senso, esso può essere descritto come un sistema dinamico a tempo discreto con ingresso e^* e uscita u^* .

Poiché la relazione tra e^* e u^* , cioè la *legge di controllo*, può essere rappresentata da un generico algoritmo realizzato in software, è chiaro che vi è una grande flessibilità nell'uso di leggi di controllo anche molto complesse. Tuttavia, per rimanere nell'ambito della trattazione sin qui sviluppata, nel seguito ci si limiterà a considerare relazioni di tipo lineare e stazionario. Pertanto, denotando con $E^*(z)$ e $U^*(z)$ le trasformate Zeta di e^* e u^* , si assumerà che

$$U^*(z) = R^*(z)E^*(z) \quad (17.3)$$

dove

$$R^*(z) = \frac{\beta_v z^v + \beta_{v-1} z^{v-1} + \dots + \beta_1 z + \beta_0}{z^v + \alpha_{v-1} z^{v-1} + \dots + \alpha_1 z + \alpha_0} \quad (17.4)$$

Si osservi che, come mostrato nel Paragrafo 8.4, alla luce della 17.4, la legge di controllo (17.3), espressa nel dominio delle trasformate, può essere riscritta nel dominio del tempo discreto come

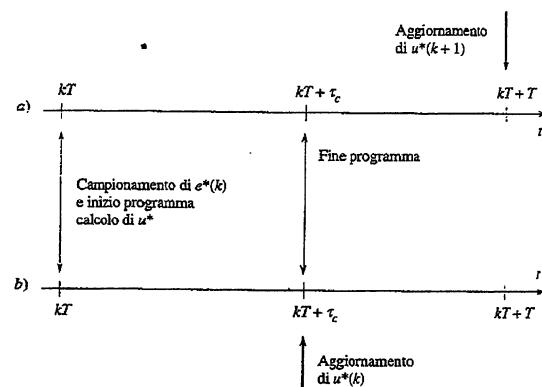
$$\begin{aligned} u^*(k) = & -\alpha_{v-1}u^*(k-1) - \dots - \alpha_1u^*(k-v+1) - \alpha_0u^*(k-v) + \\ & + \beta_v e^*(k) + \beta_{v-1}e^*(k-1) + \dots + \beta_1e^*(k-v+1) + \beta_0e^*(k-v) \end{aligned} \quad (17.5)$$

La (17.5) mostra alcuni fatti di interesse. In primo luogo, è evidente la semplicità con cui è possibile realizzare la legge di controllo anche nel caso di un regolatore $R^*(z)$ di ordine elevato; infatti sono sufficienti $2v$ registri in cui memorizzare i valori passati di e^* e u^* e alcune operazioni di somma e sottrazione. In secondo luogo, se $R^*(z)$ è propria, ma non strettamente ($\beta_v \neq 0$), $u^*(k)$ dipende da $e^*(k)$. È chiaro tuttavia che, dato il campione $e^*(k) = e(kT_s)$, è comunque necessario un tempo di calcolo finito τ_c per valutare tramite la (17.5) il valore di $u^*(k)$, che successivamente verrà trasformato nel segnale a tempo continuo u dal mantenitore (17.2). Pertanto, in questo caso, la dipendenza istantanea di $u^*(k)$ da $e^*(k)$ rappresenta un'astrazione, i cui potenziali effetti saranno esaminati nei paragrafi successivi.

17.2.4 Temporizzazione

Per il buon funzionamento del sistema di controllo digitale, le diverse operazioni di conversione e calcolo eseguite dagli elementi precedentemente introdotti devono essere opportunamente temporizzate. Nel seguito si ipotizzerà che il periodo

Figura 17.8
Modalità di campionamento di e^* e aggiornamento di u^* :
a) $R^*(z)$ strettamente propria; b) $R^*(z)$ propria, non strettamente.



di campionamento e di mantenimento coincidano, cioè che risulti

$$T_s = T_m = T$$

e T sarà chiamato ancora periodo di campionamento. Si assumerà inoltre, come già detto, $\tau_s = 0$ e che il tempo τ_c necessario per calcolare $u^*(k)$ sia inferiore a un periodo di campionamento, cioè

$$\tau_c < T \quad (17.6)$$

Per quanto riguarda il tempo di ritardo τ_m relativo al mantenimento, sono possibili due casi:

- se la funzione di trasferimento $R^*(z)$ del regolatore è strettamente propria, in vista della (17.6), all'interno del passo di campionamento è possibile, a partire da $e^*(k) = e(kT)$, completare tramite la (17.5) il calcolo di $u^*(k+1)$ e quindi porre $\tau_m = 0$ nella (17.2), come simbolicamente rappresentato nella Figura 17.8a;
- se invece $R^*(z)$ è solo propria (ma non strettamente), come ipotizzato nella Figura 17.8b, τ_m coincide con τ_c e nel sistema di controllo si introduce di fatto un ritardo di calcolo che può modificarne in modo significativo le prestazioni; per tener conto in fase di analisi di questo fenomeno, è possibile rappresentare questo ritardo mediante un blocco interposto tra il mantenitore e il processo, come mostrato nella Figura 17.9. Nel seguito si assumerà che questo ritardo sia inglobato nel processo e pertanto non sarà più considerato esplicitamente. Naturalmente, quando τ_c è trascurabile rispetto alla dinamica propria del sistema retroazionato, la presenza di questo ritardo può essere ignorata.

Per completezza si osservi infine che τ_c può includere anche tutti i ritardi intrinseci nelle operazioni di conversione, il cui effetto è comunque irrilevante nella maggioranza dei casi.

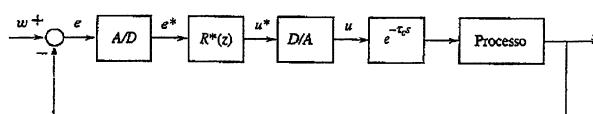


Figura 17.9
Schema equivalente con ritardo computazionale τ_c .

17.3 Campionamento

Un aspetto fondamentale nel campionamento riguarda la determinazione di un periodo T tale da garantire che non vadano perse informazioni utili relative al segnale $f(t)$ da campionare. In questo paragrafo si presenteranno dapprima alcuni risultati teorici necessari per la determinazione di T nel caso ideale di campionamento di un segnale $f(t)$ a banda limitata (si veda l'Appendice B.4.2) e quindi si forniranno indicazioni sul corretto modo di operare quando $f(t)$ non gode di tale proprietà.

17.3.1 Trasformata di un segnale campionato

Per comprendere i legami che intercorrono tra un generico segnale a tempo continuo $f(t)$ e il corrispondente segnale campionato $f^*(k) = f(kT)$, ottenuto tramite un campionatore ideale con periodo di campionamento T , è opportuno analizzare le relazioni tra le rispettive trasformate di Fourier $F(j\omega) = \mathcal{F}[f(t)]$ e $F^*(e^{j\omega}) = \mathcal{F}^*[f^*(k)]$ (si vedano le Appendici B e C). A questo riguardo vale il seguente risultato.

Teorema 17.1 Assumendo che $f(t)$ sia un segnale continuo negli istanti di campionamento, risulta

$$F^*(e^{j\omega T}) = \frac{1}{T} F_s(j\omega) \quad (17.7)$$

dove

$$F_s(j\omega) = \sum_{h=-\infty}^{+\infty} F(j(\omega + h\omega_s)) \quad , \quad \omega_s = \frac{2\pi}{T} \quad (17.8)$$

Dimostrazione Dalla definizione di trasformata di Fourier per un segnale a tempo discreto (Equazione (C.17)) si ha

$$F^*(e^{j\omega T}) = \sum_{k=-\infty}^{+\infty} f(kT) e^{-jk\omega T} \quad (17.9)$$

Si ricordi ora che, per la proprietà (B.3) della funzione impulso, data la generica funzione $g(t)$ continua in $t = \tau$, si può scrivere

$$g(\tau) = \int_{-\infty}^{+\infty} g(t) \text{imp}(t - \tau) dt$$

e quindi in particolare

$$f(kT) e^{-jk\omega T} = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \text{imp}(t - kT) dt \quad (17.10)$$

Sostituendo la (17.10) nella (17.9) si ottiene

$$\begin{aligned} F^*(e^{j\omega T}) &= \sum_{k=-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \text{imp}(t - kT) dt = \\ &= \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \sum_{k=-\infty}^{+\infty} \text{imp}(t - kT) dt \end{aligned} \quad (17.11)$$

La funzione T -periodica.

$$c(t) = \sum_{k=-\infty}^{+\infty} \text{imp}(t - kT) \quad (17.12)$$

rappresenta un treno di impulsi e, come mostrato nell'Esempio B.16, può essere espressa in serie di Fourier nella forma

$$c(t) = \frac{1}{T} \sum_{h=-\infty}^{+\infty} e^{j h \omega_s t} \quad (17.13)$$

Grazie alle (17.12), (17.13), e ricordando la definizione di trasformata di Fourier del segnale a tempo continuo $f(t)$ (Equazione (B.33)), la (17.11) è equivalente a

$$\begin{aligned} F^*(e^{j\omega T}) &= \frac{1}{T} \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \sum_{h=-\infty}^{+\infty} e^{j h \omega_s t} dt = \frac{1}{T} \sum_{h=-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(t) e^{-j(\omega-h\omega_s)t} dt = \\ &= \frac{1}{T} \sum_{h=-\infty}^{+\infty} F(j(\omega - h\omega_s)) = \frac{1}{T} \sum_{h=-\infty}^{+\infty} F(j(\omega + h\omega_s)) = \frac{1}{T} F_s(j\omega) \end{aligned}$$

che dimostra l'assunto. ■

Si osservi che $F_s(j\omega)$ è simmetrica e periodica di periodo ω_s , in quanto sono verificate le relazioni

$$\begin{aligned} F_s(-j\omega) &= \bar{F}_s(j\omega) \\ F_s(j(\omega + h\omega_s)) &= F_s(j\omega) \quad , \quad \forall \omega, h \text{ intero} \end{aligned}$$

Pertanto, $F_s(j\omega)$ è nota quando sia noto il suo andamento in $[0, \omega_N]$, dove

$$\omega_N = \frac{\omega_s}{2}$$

è chiamata *pulsazione di Nyquist*.

17.3.2 Aliasing

Il legame che intercorre tra $F(j\omega)$ e $F_s(j\omega)$ può essere maggiormente apprezzato mediante un'espressiva rappresentazione grafica. Si supponga (fittiziamente, poiché in realtà $F(j\omega)$ e $F_s(j\omega)$ sono complesse) che $F(j\omega)$ sia una funzione reale, quale per esempio quella di Figura 17.10a. Dalla (17.8) segue che, per ricavare $F_s(j\omega)$, è necessario traslare $F(j\omega)$ infinite volte della quantità $h\omega_s$, con h intero, come mostrato nella Figura 17.10b. Infine, il valore di $F_s(j\omega)$ è determinato in corrispondenza di ogni ω , sommando i contributi delle singole funzioni ottenute per traslazione, come illustrato nella Figura 17.10c. Come la Figura 17.10 evidenzia chiaramente, nell'intervallo $[-\omega_N, \omega_N]$ la funzione $F_s(j\omega)$ può essere anche molto diversa dalla $F(j\omega)$ che l'ha generata. Questo accade quando $F(j\omega)$ non è nulla per $\omega \geq \omega_N$ (si veda la Figura 17.10a); infatti, in questo caso, le armoniche non nulle di $F(j\omega)$, con $\omega \geq \omega_N$, contribuiscono a formare lo spettro $F_s(j\omega)$, per $\omega \leq \omega_N$. Questo fenomeno di sovrapposizione di spettri prende il nome di *aliasing*, o *equivocazione in frequenza*: quando esso è presente non è possibile dedurre dallo spettro del segnale campionato quello del segnale a tempo continuo che lo ha prodotto. Infatti, per l'aliasing, armoniche dello spettro ad alta

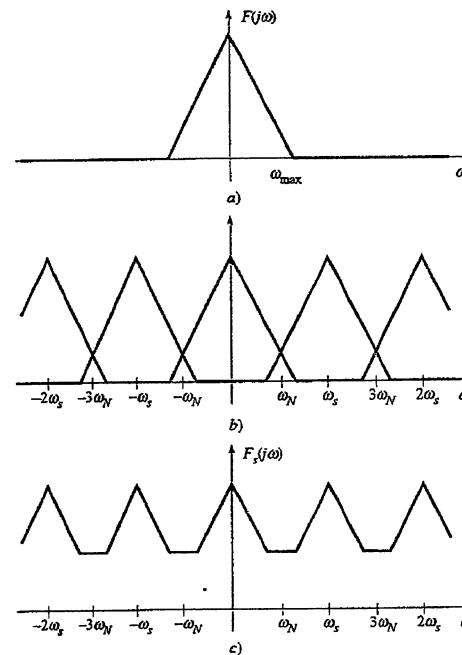


Figura 17.10
Esempio di costruzione di $F_s(j\omega)$ a partire da $F(j\omega)$.

frequenza si confondono irrimediabilmente con quelle a bassa frequenza. Queste considerazioni, sviluppate per semplicità e chiarezza nell'ipotesi che $F(j\omega)$ sia reale, sono valide anche nel caso in cui $F(j\omega)$ sia una funzione complessa.

Esempio 17.2 (Seguito dell'Esempio 17.1) Nella Figura 17.11a è riportato lo spettro d'ampiezza del segnale $e_0(t) = \sin(0.6\pi t)$. Esso è formato da due impulsi di ampiezza π in corrispondenza di $\omega = \pm 0.6\pi$ (si veda l'Esempio B.21). Lo spettro d'ampiezza del corrispondente segnale campionato $e_0^*(k)$, ottenuto con $T = 1$ e riportato nella Figura 17.11b, è formato da impulsi di ampiezza π/T in corrispondenza di $\omega = \pm 0.6\pi, \pm 1.4\pi, \pm 2.2\pi, \pm 3.4\pi, \dots$

È allora facile verificare che esso coincide con lo spettro del segnale ottenuto dal campionamento di $e_{-1}(t) = \sin(-1.4\pi t)$ con il medesimo valore di T .

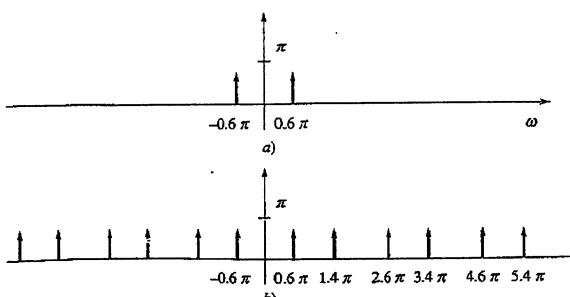


Figura 17.11
a) Spettro d'ampiezza di $e_0(t) = \sin(0.6\pi t)$; b) spettro d'ampiezza di $e_0^*(k)$ con $T = 1$ e $\theta = \omega T$.

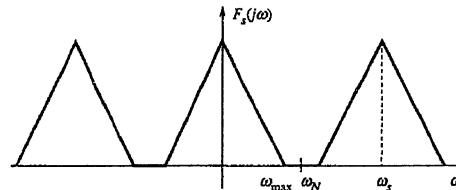
17.3.3 Teorema del campionamento

Per determinare sotto quali condizioni è possibile evitare il fenomeno dell'aliasing, si supponga che il segnale da campionare $f(t)$ sia a banda limitata (si veda il Paragrafo B.5.2) con pulsazione massima ω_{\max} , cioè che lo spettro $F(j\omega)$ sia nullo per $\omega > \omega_{\max}$. Se allora si utilizza un periodo di campionamento T tale che $\omega_{\max} < \omega_N$, in $[0, \omega_N]$ lo spettro $F_s(j\omega)$ coincide con $F(j\omega)$ e risulta

$$F_s(j\omega) = \begin{cases} F(j\omega) = TF^*(e^{j\omega T}) & , 0 \leq \omega \leq \omega_N \\ 0 & , \omega > \omega_N \end{cases}$$

Pertanto, in questo caso la conoscenza di $F^*(e^{j\omega T})$ in $[0, \omega_N]$ consente di determinare univocamente $F(j\omega)$ e quindi, in virtù della biunivocità della trasformata di Fourier, da $f^*(k) = f(kT)$ è possibile ricostruire $f(t)$. Tornando all'esempio della Figura 17.10a, il segnale è a banda limitata e scegliendo un passo di campionamento opportuno, si ottiene la $F_s(j\omega)$ riportata nella Figura 17.12.

Figura 17.12
 $F_s(j\omega)$ con $\omega_N > \omega_{\max}$.



Queste considerazioni rappresentano una dimostrazione del seguente fondamentale risultato, detto *teorema del campionamento*, o di *Shannon*.

Teorema 17.2 Sia $f(t)$ un segnale a banda limitata con pulsazione massima ω_{\max} . Allora se

$$\omega_{\max} < \omega_N = \frac{\pi}{T} \quad (17.14)$$

il segnale $f(t)$, assunto continuo negli istanti di campionamento, può essere ricostruito esattamente a partire dai campioni $f^*(k) = f(kT)$. ■

Esempio 17.3 (Seguito degli Esempi 17.1 e 17.2) Dato il segnale $e(t) = \sin(0.6\pi t)$, con $\omega_{\max} = 0.6\pi$, affinché sia verificata la condizione (17.14) del teorema del campionamento deve risultare

$$T = \frac{\pi}{\omega_N} < \frac{1}{0.6}$$

Questa diseguaglianza è soddisfatta dal periodo $T = 1$ usato negli Esempi 17.1, 17.2. Quindi, il fenomeno dell'aliasing non si manifesta.

Infine, è opportuno osservare che, come mostrato nell'esempio precedente, data una funzione $f(t)$ a banda limitata, la scelta di un valore di T in accordo con il teorema del campionamento non significa affatto che non possano esistere altre funzioni a tempo continuo $g(t)$ tali che $g(kT) = f(kT)$, che anzi esistono in numero infinito, ma implica invece che queste funzioni non siano a banda limitata con pulsazione massima minore di ω_N . A riprova di ciò si riesamineranno gli Esempi 17.1, 17.2.

17.3.4 Filtri anti-aliasing

Dato che, quando si campiona un segnale $f(t)$, non si può in genere garantire *a priori* che esso sia a banda limitata, per ridurre gli effetti dell'aliasing è conveniente prefiltrare $f(t)$ con un filtro anti-aliasing, cioè con un filtro passabasso con opportuna banda passante $[0, \omega_f]$, come illustrato nello schema della Figura 17.13.

Filtri anti-aliasing comunemente utilizzati sono quelli di Butterworth (si veda il Paragrafo 6.8.1). Il loro impiego fa sì che lo spettro del segnale da campionare non abbia componenti significative a pulsazioni superiori a ω_f . Pertanto, deve risultare $\omega_f < \omega_N$ e, tenendo conto che il filtro non è comunque mai ideale, è bene che ω_f sia decisamente inferiore a ω_N .

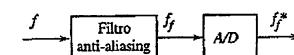


Figura 17.13
Campionamento con filtro anti-aliasing.

17.4 Mantenitore di ordine zero

Si vuole determinare ora la relazione che intercorre, in termini di trasformate, tra il segnale $f^*(k)$ di ingresso allo ZOH e il corrispondente segnale di uscita $f(t)$ (Figura 17.14). Anziché condurre l'analisi con riferimento alle corrispondenti

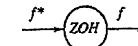


Figura 17.14
ZOH.

trasformate di Fourier, è conveniente analizzare inizialmente la relazione fra la trasformata Zeta $F^*(z)$ di $f^*(k)$ e la trasformata di Laplace $F(s)$ di $f(t)$. A questo scopo, si assuma dapprima che il segnale d'ingresso sia un impulso all'istante $k = 0$, cioè $f^*(k) = \text{imp}(k)$ e $F^*(z) = 1$. Come mostrato nella Figura 17.15, l'uscita dello ZOH è l'impulso rettangolare

$$f(t) = h_0(t) = \text{sca}(t) - \text{sca}(t - T)$$

la cui trasformata di Laplace è

$$H_0(s) = \frac{1 - e^{-sT}}{s} \quad (17.15)$$

In generale, dato il segnale di ingresso $f^*(k)$ con trasformata

$$F^*(z) = \sum_{k=0}^{\infty} f^*(k) z^{-k}$$

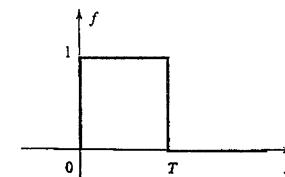


Figura 17.15
Risposta all'impulso dello ZOH.

il corrispondente segnale di uscita (Figura 17.7) è

$$f(t) = \sum_{k=0}^{\infty} f^*(k) h_0(t - kT)$$

con trasformata di Laplace

$$F(s) = \sum_{k=0}^{\infty} f^*(k) e^{-skT} H_0(s) = H_0(s) F^*(e^{sT}) \quad (17.16)$$

La (17.16) mostra che $H_0(s)$ rappresenta il rapporto tra la trasformata di Laplace del segnale di uscita e la trasformata Zeta del segnale di ingresso valutata in $z = e^{sT}$. Pertanto $H_0(s)$ può essere interpretata, in senso lato, come la "funzione di trasferimento" del mantenitore di ordine zero. Inoltre, posto $s = j\omega$ e denotando con $F(j\omega)$ e $F^*(e^{j\omega T})$ le trasformate di Fourier dei segnali $f(t)$ e $f^*(k)$, dalla (17.16) si può calcolare la "risposta in frequenza" dello ZOH, intesa come rapporto tra $F(j\omega)$ e $F^*(e^{j\omega T})$, e data da

$$H_0(j\omega) = \frac{1 - e^{-j\omega T}}{j\omega} = e^{-j\omega T/2} \frac{e^{j\omega T/2} - e^{-j\omega T/2}}{2j} \frac{2j}{j\omega} = T e^{-j\omega T/2} \frac{\sin(\omega T/2)}{\omega T/2} \quad (17.17)$$

Dalla (17.17) risulta

$$|H_0(j\omega)| = T \left| \frac{\sin(\omega T/2)}{\omega T/2} \right|, \quad \arg H_0(j\omega) = -\frac{\omega T}{2} + \arg \left(\sin \left(\frac{\omega T}{2} \right) \right) \quad (17.18)$$

I diagrammi, in scala lineare e di Bode, di $|H_0(j\omega)|/T$ e $\arg H_0(j\omega)$ sono riportati nelle Figure 17.16 e 17.17, in cui i punti di discontinuità della fase sono dati dalle variazioni da 0° a 180° e viceversa del contributo del termine reale $\sin(\omega T/2)$.

Dalle espressioni analitiche (17.17), (17.18), così come dalle Figure 17.16 e 17.17 è chiaro che nell'intervallo di pulsazioni $[0, \omega_N]$ il modulo di $H_0(j\omega)$ si mantiene all'incirca costante con

$$\frac{|H_0(j\omega_N)|}{T} = \frac{2}{\pi} \simeq -3 \text{ dB}$$

Figura 17.16
Risposta in frequenza dello ZOH: $H_0(j\omega)/T$ in scala lineare.

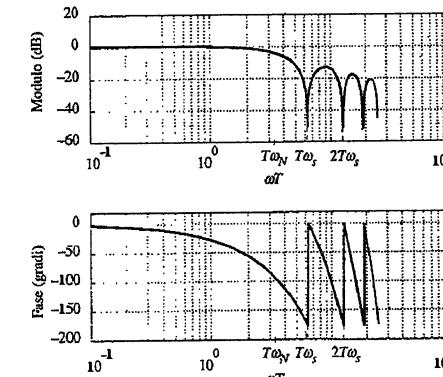
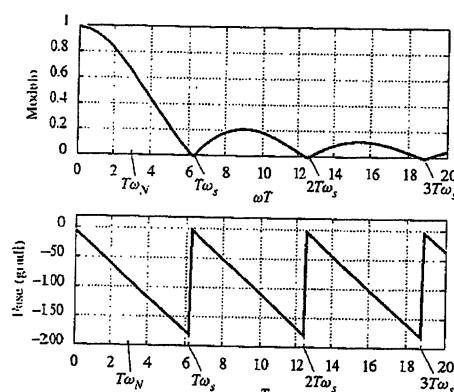


Figura 17.17
Risposta in frequenza dello ZOH: diagrammi di Bode di $H_0(j\omega)/T$.

per poi diminuire rapidamente, mentre la sua fase è identica a quella che sarebbe fornita da un ritardo di tempo pari a $T/2$. Pertanto, almeno in prima approssimazione, lo ZOH può considerarsi come un filtro reale passa-basso con banda passante $[0, \omega_N]$ e, sempre in questa banda, si può ritenere valida l'approssimazione

$$H_0(j\omega) \simeq T e^{-j\omega T/2}, \quad \omega \in [0, \omega_N] \quad (17.19)$$

17.5 Analisi a tempo continuo dei sistemi di controllo ibridi

Lo studio dei sistemi di controllo ibridi viene condotto nel seguito interpretando la serie costituita dal campionatore, dal regolatore digitale e dal mantenitore come un unico sistema il cui ingresso e la cui uscita sono segnali a tempo continuo. In quest'ottica, è necessario determinare le relazioni che intercorrono tra le trasformate di Fourier di questi segnali, cioè ricavare la relativa "risposta in frequenza". A questo scopo, si riesamini lo schema della Figura 17.1, riportato per comodità nella Figura 17.18 in cui il campionatore è ideale e il mantenitore è lo ZOH, e si ipotizzi che tutti i segnali in gioco siano dotati della trasformata di Fourier.

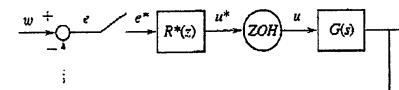


Figura 17.18
Sistema di controllo ibrido.

Ricordando il Teorema 17.1, la relazione tra le trasformate di Fourier di $e(t)$ ed $e^*(k)$, con ovvio significato dei simboli, è

$$E^*(e^{j\omega T}) = \frac{1}{T} E_s(j\omega) \quad (17.20)$$

dove

$$E_s(j\omega) = \sum_{h=-\infty}^{+\infty} E(j(\omega + h\omega_s)), \quad \omega_s = \frac{2\pi}{T} \quad (17.21)$$

D'altra parte, dato il regolatore digitale $R^*(z)$, dalla (17.3) risulta

$$U^*(e^{j\omega T}) = R^*(e^{j\omega T})E^*(e^{j\omega T}) \quad (17.22)$$

mentre, per la (17.16), si ha

$$U(j\omega) = H_0(j\omega)U^*(e^{j\omega T}) \quad (17.23)$$

Combinando ora le (17.20)-(17.23), si ottiene

$$U(j\omega) = \frac{1}{T}H_0(j\omega)R^*(e^{j\omega T})E_s(j\omega) \quad (17.24)$$

Per ricavare un legame diretto tra $U(j\omega)$ e $E(j\omega)$ si noti che:

- se e è un segnale a banda limitata con pulsazione massima $\omega_{\max} < \omega_N$, per il teorema del campionamento nell'intervallo $[0, \omega_N]$ lo spettro $E_s(j\omega)$ coincide con $E(j\omega)$;
- come già osservato, $H_0(s)$ può essere interpretata come la funzione di trasferimento di un filtro passa-basso con guadagno pari a T e banda passante $[0, \omega_N]$.

Pertanto, in base a queste considerazioni, si può ritenere

$$H_0(j\omega)E_s(j\omega) \simeq H_0(j\omega)E(j\omega) \quad \omega \in [0, \infty)$$

e, quindi, dalla (17.24) risulta

$$U(j\omega) \simeq \frac{1}{T}H_0(j\omega)R^*(e^{j\omega T})E(j\omega) \quad \omega \in [0, \infty) \quad (17.25)$$

Questa formula mostra come il sistema a tempo continuo costituito da campionatore, regolatore e mantenitore possa essere approssimativamente rappresentato dalla funzione di trasferimento

$$R(s) = \frac{1}{T}H_0(s)R^*(e^{sT}) \quad (17.26)$$

Per semplificare ulteriormente l'analisi, si osservi che in $[0, \omega_N]$ è possibile impiegare per il mantenitore l'approssimazione fornita dalla (17.19). Quindi, dalle (17.19), (17.25) segue

$$U(j\omega) \simeq e^{-j\omega T/2}R^*(e^{j\omega T})E(j\omega) \quad \omega \in [0, \omega_N] \quad (17.27)$$

per cui il comportamento di campionatore, regolatore e mantenitore in $[0, \omega_N]$ può essere approssimativamente descritto da

$$R(s) = e^{-sT/2}R^*(e^{sT}) \quad (17.28)$$

La (17.25) e la (17.27) mostrano che, dato un generico regolatore digitale $R^*(z)$ e nell'ipotesi che il segnale da campionare e soddisfi il teorema del campionamento, almeno in prima approssimazione l'analisi del corrispondente sistema di controllo digitale nella banda $[0, \omega_N]$ può essere ricondotta a quella di un sistema a tempo continuo in cui il regolatore $R(s)$ è dato dalla (17.26), o dalla (17.28).

La trasformazione

$$z = e^{sT} \quad (17.29)$$

che serve a determinare l'equivalente a tempo continuo di un regolatore digitale, riveste un ruolo di fondamentale importanza nell'analisi e nel progetto dei sistemi di controllo digitale. Essa è chiamata *trasformazione di campionamento*, ed esprime un legame tra i punti del piano S , di definizione della variabile complessa s , e i punti del piano Z , di definizione della variabile complessa z .

Esempio 17.4 Nel sistema di controllo a tempo continuo rappresentato nella Figura 17.19a si supponga che il processo sia descritto da

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

e il regolatore, puramente proporzionale, sia

$$R^*(s) = \mu = 1.5$$

In corrispondenza, risulta $\omega_c \simeq 1$ e $\varphi_m \simeq 44^\circ$.

È possibile realizzare lo stesso sistema di controllo in modo digitale, per esempio con periodo di campionamento $T = \pi/5$, come illustrato nella Figura 17.19b.

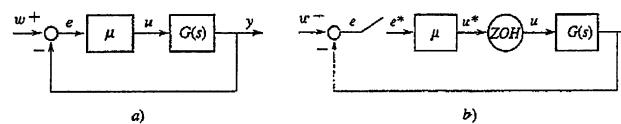
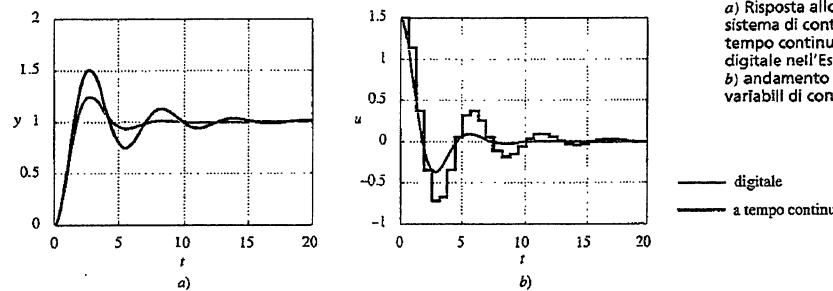


Figura 17.19
a) Sistema di controllo a tempo continuo; b) sistema di controllo digitale nell'Esempio 17.4.

Nella Figura 17.20a sono messe a confronto le risposte dei due sistemi della Figura 17.19 a uno scalino unitario del riferimento: è evidente che la realizzazione digitale del regolatore ha prodotto un notevole deterioramento delle prestazioni in termini di sovraetensione e di tempo di assolvimento. Nella Figura 17.20b è anche riportato l'andamento delle corrispondenti variabili di controllo.



Per comprendere questo risultato, è possibile impiegare l'analisi a tempo continuo del sistema di controllo digitale e confrontare i diagrammi di Bode associati alla funzione di trasferimento d'anello

$$L(s) = R^*(s)G(s) = \frac{1.5}{s(s+1)}$$

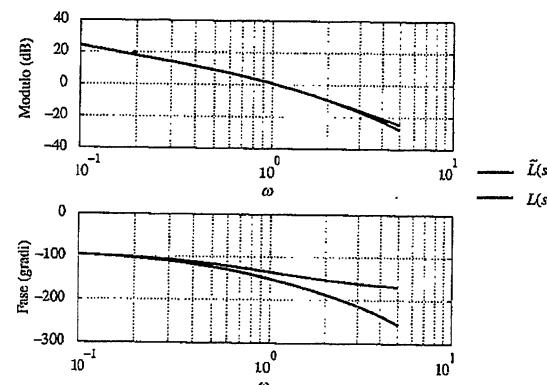
del sistema di controllo della Figura 17.19a con quelli associati alla funzione di trasferimento d'anello (approssimata)

$$\tilde{L}(s) = \frac{1}{T}H_0(s)R^*(e^{sT})G(s) \approx \frac{1 - e^{-0.2Ts}}{0.2s} \frac{1.5}{s(s+1)}$$

ottenuta a partire dalla (17.26) e relativa al sistema di controllo digitale di Figura 17.19b. I diagrammi, riportati nella Figura 17.21, sono rappresentati solamente sino a $\omega_N = \pi/T = 5$. Come è lecito attendersi dalle considerazioni precedenti, i diagrammi di Bode del modulo delle due funzioni

Figura 17.20
a) Risposta allo scalino del sistema di controllo a tempo continuo e di quello digitale nell'Esempio 17.4; b) andamento delle variabili di controllo.

Figura 17.21
Diagrammi di Bode associati a $L(s)$ e $\tilde{L}(s)$ nell'Esempio 17.4.



sono in pratica identici, mentre quelli della fase differiscono in modo sostanziale per la presenza del campionatore e dello ZOH. In particolare, in corrispondenza della pulsazione critica ω_c risulta

$$\arg \tilde{L}(j\omega_c) - \arg L(j\omega_c) \approx -18^\circ$$

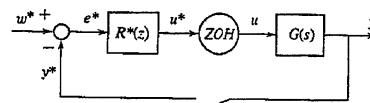
Questo valore si sarebbe potuto calcolare direttamente mediante la (17.28).

In sostanza, è come se la realizzazione digitale del regolatore avesse prodotto una riduzione del margine di fase da 44° a 26° , tale da giustificare pienamente le notevoli sovraetensioni presenti nella corrispondente risposta allo scalino di Figura 17.20.

17.6 Scelta del periodo di campionamento

L'analisi condotta nel Paragrafo 17.5 e i risultati dell'Esempio 17.4 mostrano che la scelta del periodo di campionamento è di fondamentale importanza per la realizzazione di sistemi di controllo digitale. Il problema della determinazione di T è analizzato nel seguito con riferimento allo schema a campionamento dell'uscita di Figura 17.22, in cui si può immaginare che il filtro anti-aliasing che precede il campionatore sia inglobato nella funzione di trasferimento $G(s)$. Si assume inoltre che entrambi i convertitori operino con un comune periodo di campionamento T e che l'eventuale ritardo computazionale τ_c (si veda il Paragrafo 17.2.4) sia anch'esso incorporato come fattore di $G(s)$.

Figura 17.22
Schema di controllo digitale a campionamento dell'uscita.



La scelta di T , o equivalentemente della pulsazione di campionamento $\omega_s = 2\pi/T$, va effettuata tenendo conto di varie esigenze, tra le quali saranno ora discusse le più significative.

- **Costo dei dispositivi** All'aumentare della pulsazione di campionamento ω_s , il costo dei convertitori A/D e D/A cresce in modo più che lineare e aumenta nel contempo la potenza di calcolo richiesta per l'unità di elaborazione su cui realizzare la legge di controllo. Ciò suggerisce di non utilizzare valori di T inutilmente bassi.

- **Problemi di tipo numerico** Si è già accennato nel Paragrafo 17.2.1 al fatto che un valore basso del periodo di campionamento enfatizza l'effetto degli inevitabili errori di quantizzazione connessi alla codifica digitale del segnale campionato.

- **Campionamento e informazione** La pulsazione di campionamento ω_s va in ogni caso commisurata alla dinamica desiderata del sistema retroazionato. Immaginando infatti che, nei confronti del segnale di riferimento, il sistema di controllo si comporti come un filtro passa-basso, e indicando con $\bar{\omega}$ l'estremo superiore della banda passante che si vuole realizzare, è evidente che il contenuto informativo utile ai fini dell'azione di controllo del segnale y che viene campionato è sostanzialmente collocato in una banda di pulsazioni che ha 0 come estremo inferiore e si estende poco al di là della pulsazione $\bar{\omega}$. La teoria del campionamento (Paragrafo 17.3) garantisce allora che tale informazione venga preservata pur di utilizzare una pulsazione di campionamento ω_s superiore a $2\bar{\omega}$. Nei casi in cui si faccia riferimento a un regolatore a tempo continuo e la pulsazione critica ω_c rappresenti un'accettabile approssimazione di $\bar{\omega}$ (Paragrafi 10.3.3 e 10.3.5), la prescrizione precedente si traduce nel vincolo $\omega_s > 2\omega_c$.

- **Effetto del filtro anti-aliasing** Si è visto nel Paragrafo 17.3.4 che è buona norma che l'estremo superiore ω_f della banda passante del filtro anti-aliasing sia inferiore alla pulsazione di Nyquist $\omega_N = \omega_s/2$. D'altra parte, la natura non ideale del filtro fa sì che anche a pulsazioni prossime a ω_f esso introduca uno sfasamento negativo non trascurabile (per esempio, un filtro di Butterworth di ordine n presenta una fase di $-n45^\circ$ in corrispondenza di ω_f). Perché tale sfasamento non produca inaccettabili diminuzioni del margine di fase, pregiudicando così le prestazioni del sistema di controllo, è opportuno fare in modo che ω_f sia molto superiore alla pulsazione critica ω_c . In definitiva deve risultare $\omega_s \gg 2\omega_c$.

In base alle precedenti considerazioni, si possono elaborare regole euristiche per la determinazione di un valore adeguato di ω_s , da applicare in tutte quelle situazioni in cui sia disponibile una stima *a priori* della pulsazione critica ω_c associata al sistema di controllo che si intende realizzare. Tipicamente si suggeriscono valori di ω_s che rispettino un vincolo del tipo

$$\alpha\omega_c \leq \omega_s \leq 10\alpha\omega_c \quad (17.30)$$

con α di solito compreso tra 5 e 10. Tale indicazione di massima non va ovviamente utilizzata in modo acritico, ma adattata caso per caso anche alla luce di altre considerazioni specifiche.

Può essere interessante valutare cosa comporta l'applicazione della regola (17.30) in termini di numero di campioni per tempo di assestamento del sistema retroazionato. A tale riguardo si osservi che, assumendo un valore tipico dello smorzamento dei poli dominanti in anello chiuso pari a $\xi \approx 0.7$, il tempo di assestamento al 99% della risposta allo scalino del sistema retroazionato è $T_{a1} \approx 5/\xi\omega_c \approx 6.6/\omega_c$. Dalla (17.30), ricordando che $T = 2\pi/\omega_s$, si ottiene allora approssimativamente

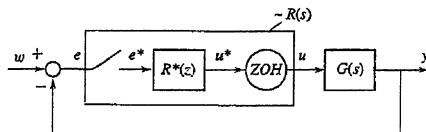
$$\frac{T_{a1}}{10\alpha} \leq T \leq \frac{T_{a1}}{\alpha} \quad (17.31)$$

Quando, per esempio, si pone $\alpha = 5$, la condizione (17.31) suggerisce che, a partire dal valore desiderato dell'indicatore T_{a1} , la scelta del periodo di campionamento T va fatta in modo da avere dai 5 ai 50 campioni nell'intervallo di tempo di assestamento.

17.7 Discretizzazione di un regolatore a tempo continuo

Una volta effettuata la scelta del periodo di campionamento, il problema della sintesi di un regolatore digitale si riduce alla determinazione della sua funzione di trasferimento $R^*(z)$.

Figura 17.23
Schema di controllo digitale a campionamento dell'errore.



Per introdurre un possibile approccio a tale problema, è conveniente riferirsi allo schema a campionamento dell'errore di Figura 17.23. Mediante l'analisi a tempo continuo del Paragrafo 17.5, si è già visto come il sottosistema costituito dalla serie di campionatore, regolatore digitale e mantenitore possa essere descritto approssimativamente dalla funzione di trasferimento

$$R(s) = e^{-sT/2} R^*(e^{sT}) \quad (17.32)$$

almeno nell'intervallo di pulsazioni $[0, \omega_N]$. Si può allora impostare il problema di progetto di $R^*(z)$ cercando di fare in modo che in tale intervallo la risposta in frequenza associata alla (17.32) approssimi quella di un regolatore analogico $R^*(s)$ di riferimento.

Il termine di ritardo $e^{-sT/2}$, dovuto alla presenza dello ZOH, contribuisce a tale risposta in frequenza solamente in termini di sfasamento, e in particolare il suo contributo alla fase critica, espresso in gradi, è

$$\varphi_0 = -\omega_c \frac{T}{2} \frac{180}{\pi} = -\frac{\omega_c}{\omega_s} 180^\circ$$

Se quindi si utilizza la regola euristica (17.30) per la scelta del periodo di campionamento T , risulta

$$-\frac{180^\circ}{\alpha} \leq \varphi_0 \leq -\frac{18^\circ}{\alpha}$$

Per esempio, con $\alpha = 5$, il valore assoluto di φ_0 è compreso tra 3.6° (per $\omega_s = 50\omega_c$) e 36° (per $\omega_s = 5\omega_c$) e può dunque risultare tutt'altro che trascurabile.

Un'appropriata strategia di sintesi consiste allora nel progettare il regolatore analogico di riferimento $R^*(s)$ in modo da assicurare, rispetto alle specifiche originarie, un'adeguata eccedenza nel margine di fase in grado di compensare l'effetto del ritardo introdotto dal mantenitore. In tal caso, in base alla Equazione (17.32), la determinazione della funzione di trasferimento $R^*(z)$ del regolatore digitale può poi essere condotta con l'unico obiettivo che risulti

$$R^*(e^{j\omega T}) \simeq R^*(j\omega) \quad , \quad 0 \leq \omega \leq \omega_N \quad (17.33)$$

A questo punto, per soddisfare la (17.33) si potrebbe pensare di calcolare $R^*(z)$ da $R^*(s)$ utilizzando la trasformazione inversa di campionamento (Equazione (17.29))

$$s = \frac{1}{T} \ln z$$

ponendo cioè

$$R^*(z) = R^*(\frac{1}{T} \ln z) \quad (17.34)$$

Però, in questo modo, si otterebbe una funzione $R^*(z)$ non razionale nella variabile z , e quindi non associabile a un sistema dinamico a tempo discreto a dimensione finita. Nel seguito di questo paragrafo si presenteranno invece tecniche approssimate per risolvere il problema descritto dalla (17.33) che danno luogo a soluzioni $R^*(z)$ razionali.

17.7.1 Metodo della trasformazione bilineare

Il problema di approssimazione in frequenza associato alla (17.33) ammette una facile interpretazione nel dominio del tempo. Infatti, si tratta di determinare un sistema dinamico a tempo discreto il cui legame ingresso-uscita, descritto da un sistema di equazioni alle differenze, approssimi quello imposto da un dato sistema dinamico a tempo continuo, formulato in termini di un sistema di equazioni differenziali. In questo senso, la soluzione del problema può quindi essere ottenuta mediante l'uso di opportune tecniche di integrazione numerica. Qui si esamineranno solo alcune di queste tecniche, che danno origine a procedure di discretizzazione particolarmente semplici.

Si supponga di disporre di una realizzazione in forma di stato della funzione di trasferimento $R^*(s)$ da approssimare; sia cioè

$$R^*(s) = C(sI - A)^{-1} B + D \quad (17.35)$$

per opportuni valori di A , B , C e D . Dato che l'ingresso del regolatore analogico è l'errore e e la sua uscita è la variabile di controllo u , si può allora descrivere il sistema come

$$\dot{x}(t) = Ax(t) + Be(t) \quad (17.36)$$

$$u(t) = Cx(t) + De(t) \quad (17.37)$$

Integrando ambo i membri dell'equazione di stato (17.36) sull'intervallo $[kT, (k+1)T]$, e denotando con $x^*(k) = x(kT)$ il vettore di stato al k -esimo istante di campionamento, si ottiene

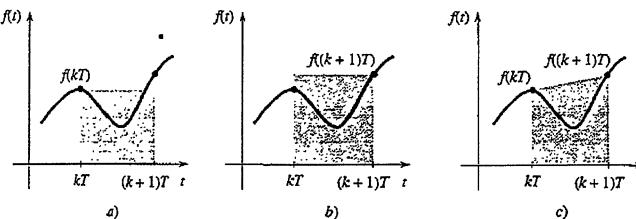
$$x^*(k+1) - x^*(k) = A \int_{kT}^{(k+1)T} x(t) dt + B \int_{kT}^{(k+1)T} e(t) dt \quad (17.38)$$

Affinché anche al secondo membro compaia un'espressione contenente solo i valori di $x(t)$ ed $e(t)$ negli istanti di campionamento, è necessario approssimare gli integrali che vi compaiono. A tale scopo si può utilizzare la seguente formula di integrazione numerica, qui scritta per una generica funzione integranda $f(t)$:

$$\int_{kT}^{(k+1)T} f(t) dt \simeq [(1-\alpha) f(kT) + \alpha f((k+1)T)] T \quad , \quad 0 \leq \alpha \leq 1 \quad (17.39)$$

Secondo tale formula, l'integrale viene approssimato mediante la combinazione lineare dei valori della funzione negli estremi dell'intervallo di integrazione. Per un'interpretazione geometrica della (17.39) nel caso in cui $f(t)$ sia scalare e si ponga $\alpha = 0$, $\alpha = 1$ oppure $\alpha = 0.5$ si veda la Figura 17.24. Con $\alpha = 0$ o $\alpha = 1$ l'area sottesa dalla curva $f(t)$ nell'intervallo $[kT, (k+1)T]$ viene rispettivamente

Figura 17.24
Interpretazione dei metodi di integrazione numerica:
a) Eulero in avanti ($\alpha = 0$);
b) Eulero all'indietro ($\alpha = 1$);
c) Tustin ($\alpha = 0.5$).



approssimata dall'area del rettangolo rappresentata nella Figura 17.24a o nella Figura 17.24b. Si parla in tal caso di *metodo di Eulero in avanti* ($\alpha = 0$) e di *metodo di Eulero all'indietro* ($\alpha = 1$). Quando invece si pone $\alpha = 0.5$, la (17.39) corrisponde al cosiddetto *metodo del trapezio* o *metodo di Tustin*, in cui si calcola l'area del trapezio mostrato nella Figura 17.24c.

Utilizzando l'approssimazione (17.39) nell'Equazione (17.38), e denotando con $e^*(k) = e(kT)$ l'errore negli istanti di campionamento, si ottiene

$$\begin{aligned} x^*(k+1) - x^*(k) &= A[(1-\alpha)x^*(k) + \alpha x^*(k+1)]T + \\ &+ B[(1-\alpha)e^*(k) + \alpha e^*(k+1)]T \end{aligned} \quad (17.40)$$

Inoltre, dall'Equazione (17.37) valutata in corrispondenza degli istanti di campionamento, si ha

$$u^*(k) = Cx^*(k) + De^*(k) \quad (17.41)$$

dove ovviamente si è posto $u^*(k) = u(kT)$. Le due relazioni (17.40) e (17.41) costituiscono una rappresentazione approssimata del sistema originario e possono essere interpretate come le equazioni di un sistema dinamico lineare e stazionario a tempo discreto. Applicando la trasformata Zeta all'Equazione (17.40), si ricava con semplici passaggi

$$\left[\frac{1}{T} \frac{z-1}{\alpha z + 1 - \alpha} I - A \right] X^*(z) = BE^*(z)$$

da cui, usando anche l'Equazione (17.41) e ricordando l'Equazione (17.35), si conclude che la funzione di trasferimento $R^*(z) = U^*(z)/E^*(z)$ è data da

$$R^*(z) = C \left[\frac{1}{T} \frac{z-1}{\alpha z + 1 - \alpha} I - A \right]^{-1} B + D = R^o \left(\frac{1}{T} \frac{z-1}{\alpha z + 1 - \alpha} \right)$$

In conclusione, si può determinare un'approssimante digitale di un dato regolatore analogico di riferimento $R^o(s)$ utilizzando la semplice regola di sostituzione di variabile

$$s = \frac{1 - z - 1}{T \alpha z + 1 - \alpha} \quad (17.42)$$

dove α è un arbitrario valore compreso tra 0 e 1. Tale metodo viene chiamato *metodo della trasformazione bilineare* a causa della struttura della trasformazione (17.42). Esso si specializza nei metodi di Eulero in avanti e all'indietro (d'ora in poi indicati con le sigle EA ed EI) e nel metodo di Tustin (TU) quando α è uguale

a 0, 1, e 0.5, dando luogo rispettivamente alle trasformazioni

$$s = \frac{z-1}{T} \quad (\text{EA})$$

$$s = \frac{z-1}{Tz} \quad (\text{EI})$$

$$s = \frac{2z-1}{Tz+1} \quad (\text{TU})$$

In particolare, il metodo EA si presta a una interpretazione intuitiva osservando che moltiplicare per s nel dominio della trasformata di Laplace corrisponde all'operazione di derivazione, mentre il prodotto per $(z-1)/T$ nel dominio della trasformata Zeta rappresenta l'operazione di calcolo del rapporto incrementale. È ovvio che per $T \rightarrow 0$ i due operatori tendono a coincidere.

Vale la pena di sottolineare che la struttura bilineare della (17.42) garantisce che, se $R^o(s)$ è razionale, la funzione di trasferimento del regolatore digitale risultante è anch'essa razionale e quindi associabile a una legge di controllo lineare e stazionaria.

Un'altra interessante analisi riguarda la collocazione nel piano complesso dei poli (e degli zeri) del regolatore digitale. A questo riguardo, si osservi che la trasformazione inversa della (17.42) è data da

$$z = \frac{1 + (1-\alpha)Ts}{1 - \alpha Ts} = \psi(s) \quad (17.43)$$

e che, se $R^o(s)$ ha un polo (o uno zero) in s_1 , allora $R^*(z)$ ha un polo (o uno zero) in $z_1 = \psi(s_1)$. In particolare, un polo nell'origine viene trasformato in un polo in $z = 1$ per qualunque valore di α e di T . È anche possibile visualizzare la corrispondenza tra regioni del piano S e del piano Z indotta dalla trasformazione (17.43). Per esempio, può essere interessante determinare l'immagine attraverso la (17.43) dei punti appartenenti al semipiano sinistro di S , per verificare che poli stabili del regolatore analogico si trasformino in poli stabili del regolatore digitale.

Nella Figura 17.25 è mostrata l'immagine attraverso $\psi(s)$ del semipiano sinistro del piano S nei tre casi particolari $\alpha = 0$, $\alpha = 1$ e $\alpha = 0.5$:

- per $\alpha = 0$ la regione risultante è il semipiano alla sinistra della retta verticale di ascissa $z = 1$;
- per $\alpha = 1$ è costituita dai punti del cerchio di raggio 0.5 e centro in $z = 0.5$;
- per $\alpha = 0.5$ essa coincide con il cerchio di raggio unitario e centro in $z = 0$.

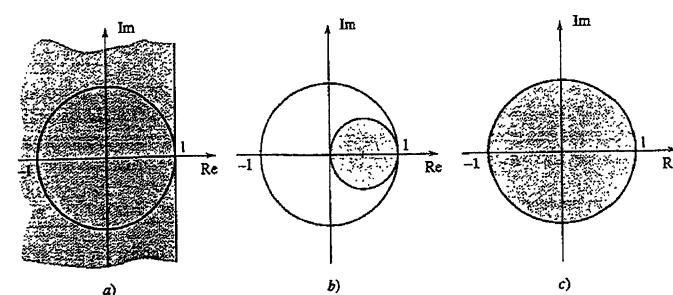


Figura 17.25
Immagine nel piano Z del semipiano sinistro del piano S :
a) Eulero in avanti;
b) Eulero all'indietro;
c) Tustin.

Pertanto, solo nel caso del metodo di Tustin ($\alpha = 0.5$) esiste un'esatta corrispondenza tra le regioni di stabilità a tempo continuo e a tempo discreto. Si osservi, inoltre, che per $\alpha = 0.5$ la (17.42) coincide, a parte il fattore $2/T$, con la trasformazione (7.47) utilizzata nel Paragrafo 7.5.4 per studiare la stabilità di un sistema a tempo discreto con i metodi propri dei sistemi a tempo continuo. Se invece si usa il metodo di Eulero in avanti, o comunque quando si sceglie $\alpha < 0.5$, può accadere che un regolatore $R^0(s)$ asintoticamente stabile venga approssimato con un regolatore $R^*(z)$ instabile. Questo è uno dei motivi per cui nell'uso del metodo della trasformazione bilineare si preferisce di solito adottare un valore di α maggiore o uguale a 0.5.

La formula (17.43) permette anche di riconoscere che i tre metodi di discretizzazione appena illustrati forniscono soluzioni approssimate della relazione (17.34), che era basata sull'idea di utilizzare la trasformazione di campionamento $z = e^{sT}$. Infatti, tramite lo sviluppo in serie dell'esponenziale intorno a $s = 0$, si ricavano le due espressioni alternative

$$z = e^{sT} \approx 1 + sT$$

$$z = e^{sT} = \frac{1}{e^{-sT}} \approx \frac{1}{1 - sT}$$

che, come si vede dalla (17.43), corrispondono rispettivamente ai metodi di Eulero in avanti e all'indietro.

Per quanto riguarda il metodo di Tustin, si osservi che lo sviluppo di Taylor della funzione (17.43) intorno a $s = 0$ è

$$\psi(s) = 1 + sT + s^2\alpha T^2 + \dots$$

i cui primi tre termini, per $\alpha = 0.5$, coincidono con quelli dello sviluppo in serie di e^{sT} .

Esempio 17.5 La versione digitale del regolatore PI

$$R^0(s) = K_P \left(1 + \frac{1}{sT_I} \right) \quad (17.44)$$

ottenuta con il metodo della trasformazione bilineare è descritta dalla funzione di trasferimento

$$R^*(z) = K_P \left(1 + \frac{T(\alpha z + 1 - \alpha)}{T_I(z - 1)} \right)$$

In particolare, i metodi EA, EI e TU portano a sintetizzare i seguenti regolatori digitali:

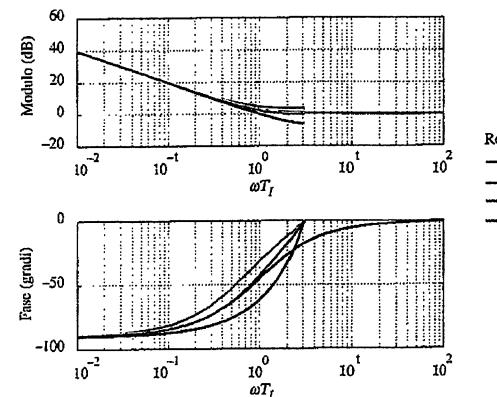
$$R^*(z) = K_P \frac{z - 1 + T/T_I}{z - 1} \quad (\text{EA})$$

$$R^*(z) = K_P \frac{(1 + T/T_I)z - 1}{z - 1} \quad (\text{EI})$$

$$R^*(z) = K_P \frac{(1 + T/2T_I)z - (1 - T/2T_I)}{z - 1} \quad (\text{TU})$$

Essi hanno in comune il polo in $z = 1$ e il guadagno (generalizzato) $\mu^* = K_P T / T_I$, mentre differiscono solo per il valore dello zero.

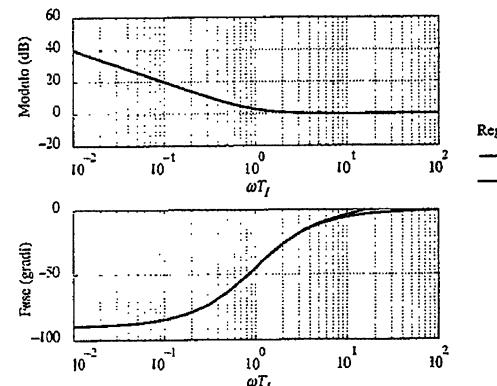
Nella Figura 17.26 sono riportati i diagrammi di Bode di $R^*(e^{j\omega T})$ relativi alle tre soluzioni (EA), (EI) e (TU) con $T = T_I$, confrontati con quelli del regolatore analogico di riferimento. Ovviamente il confronto è significativo solo nell'intervallo di pulsazioni $[0, \pi/T]$. Diminuendo il



Regolatori:
— analogico
— digitale EA
— digitale EI
— digitale TU

Figura 17.26
Diagrammi di Bode di un regolatore PI analogico e digitale (metodi EA, EI e TU), con $T = T_I$, $K_P > 0$. La scala delle pulsazioni è normalizzata e sul diagramma del modulo è riportato $|R^*(e^{j\omega T})| T_I / K_P$.

valore di T , si osserverebbe una migliore aderenza tra i diagrammi, come esemplificato nella Figura 17.27, relativa al regolatore (TU) con $T = 0.2T_I$. I diagrammi associati agli altri due regolatori (EA) e (EI) non sono riportati poiché in questo caso non presenterebbero differenze significative rispetto a quelli della Figura 17.27.



Regolatori:
— PI analogico
— digitale TU

Figura 17.27
Diagrammi di Bode di un regolatore PI analogico e digitale (metodo TU), con $T = 0.2T_I$, $K_P > 0$. La scala delle pulsazioni è normalizzata e sul diagramma del modulo è riportato $|R^*(e^{j\omega T})| T_I / K_P$.

17.7.2 Esempio di sintesi

Per illustrare le tecniche di sintesi basate sulla discretizzazione, verrà ora ripreso uno degli esempi del Capitolo 11, confrontando le prestazioni del regolatore analogico là sintetizzato con quelle ottenibili con una sua versione digitale. L'esempio mira anche a mostrare che l'analisi approssimata introdotta nel Paragrafo 17.5 può condurre, a volte, a conclusioni erronee sulle prestazioni del sistema, e addirittura sulla stabilità.

Esempio 17.6 (Seguito dell'Esempio 11.3) Per controllare il sistema descritto dalla funzione

$$G(s) = \frac{0.1(1 - 2s)}{s(1 + 10s)(1 + 0.1s)} \quad (17.45)$$

è stato progettato nell'Esempio 11.3 il regolatore a tempo continuo

$$R(s) = \frac{2(1+10s)}{(1+0.1s)} \quad (17.46)$$

che assicura la pulsazione critica $\omega_c \approx 0.22$, e il margine di fase $\varphi_m \approx 64^\circ$. Si noti che tale valore di φ_m è già ampiamente superiore al valore richiesto nelle specifiche ($\varphi_m = 40^\circ$) e dunque non è necessario rivedere il progetto analogico per tener conto del ritardo introdotto dal mantenitore. Quanto alla scelta della pulsazione di campionamento, in base alla regola (17.30) con $\alpha = 5$, un valore adeguato per $\omega_s = 2\pi/T$ deve trovarsi nell'intervallo [1, 10]; si può quindi scegliere $T = 1$.

Utilizzando i metodi di discretizzazione EA, EI, TU sulla funzione (17.46), si ottengono rispettivamente i regolatori

$$R^*(z) = \frac{200z - 180}{z + 9} \quad (\text{EA}) \quad (17.47)$$

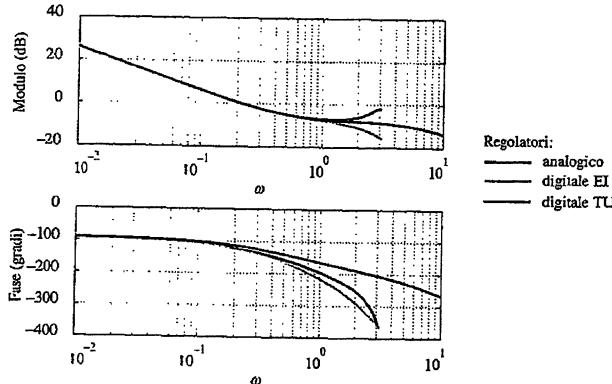
$$R^*(z) = \frac{20z - 18.18}{z - 0.09} \quad (\text{EI}) \quad (17.48)$$

$$R^*(z) = \frac{35z - 31.67}{z + 0.67} \quad (\text{TU}) \quad (17.49)$$

È possibile verificare per simulazione che il regolatore (17.47) non è in realtà utilizzabile poiché genera un sistema di controllo instabile.

Per quanto riguarda le due rimanenti soluzioni, si consideri la Figura 17.28, in cui il diagramma di Bode della funzione d'anello originaria $L(j\omega) = R(j\omega)G(j\omega)$ è confrontato, per entrambi i regolatori, con quello della funzione $\tilde{L}(j\omega) = H_0(j\omega)R^*(e^{j\omega T})G(j\omega)/T$, che, alla luce della (17.25) e della discussione del Paragrafo 17.5, fornisce una descrizione approssimata della risposta in frequenza d'anello associata al sistema di controllo digitale, valida fino alla pulsazione di Nyquist $\omega_N = \pi/T \approx 3.14$. Si può notare che, a fronte di un valore della pulsazione critica ω_c praticamente invariato, il margine di fase ottenuto con i due regolatori digitali è leggermente minore. In effetti, si ricava $\varphi_m \approx 58^\circ$ nel caso TU e $\varphi_m \approx 53^\circ$ nel caso EI, ma il requisito sul margine di fase è comunque soddisfatto.

Figura 17.28
Diagrammi di Bode della funzione d'anello nell'Esempio 17.6.



Nella Figura 17.29 sono invece messi a confronto gli andamenti delle risposte a uno scalino del riferimento in corrispondenza del regolatore analogico (17.46) e dei regolatori TU e EI. In particolare, si può notare che la soluzione TU è addirittura migliore di quella originaria in quanto, a dispetto della diminuzione del margine di fase, permette di ottenere un assottigliamento più rapido al valore di regime con uno sforzo ridotto della variabile di controllo. Peraltra, questo risultato inaspettato è strettamente legato alla specificità di questo esempio, e in particolare al fatto che il sistema sotto controllo non è a sfasamento minimo e il regolatore è una rete anticipatrice. In questo caso, le prestazioni del regolatore EI appaiono leggermente inferiori, anche se è quello che sollecita in misura minore la variabile di controllo.

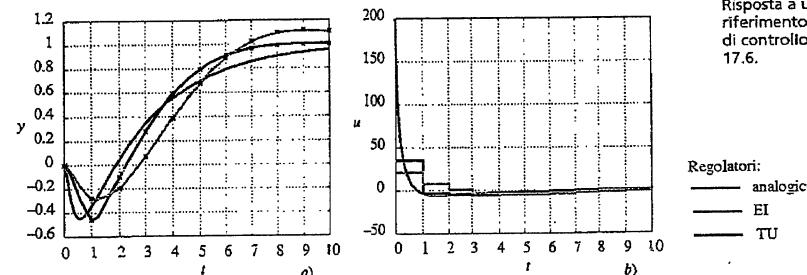


Figura 17.29
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 17.6.

Regolatori:
 — analogico
 - - EI
 - · TU

17.8 Problemi realizzativi

Tra le varie questioni connesse alla realizzazione pratica di un regolatore digitale, ci si limita ad accennare a un paio di problemi che nascono in presenza di regolatori contenenti un'azione integrale.

17.8.1 Quantizzazione

Fenomeni di quantizzazione dovuti alla codifica digitale del segnale errore possono dar luogo a comportamenti inaspettati del sistema di controllo. Si consideri, per esempio, la discretizzazione di un regolatore ad azione puramente integrale, $R(s) = K_I/s$, effettuata mediante il metodo EI:

$$R^*(z) = \frac{K_I T z}{z - 1}$$

La corrispondente legge di controllo nel dominio del tempo è

$$u^*(k) = u^*(k-1) + K_I T e^*(k) \quad (17.50)$$

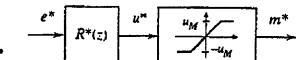
Si supponga ora che l'errore e^* si assesti su un valore costante non nullo $\bar{e}^* < \varepsilon/K_I T$, dove ε rappresenta il livello di quantizzazione impiegato nella codifica digitale. La legge di controllo (17.50) fornisce, allora, un valore costante di u^* , poiché il termine a secondo membro contiene l'errore \bar{e}^* inferiore allo zero-macchina ε . Pertanto, l'errore rimane polarizzato al valore \bar{e}^* , senza annullarsi asintoticamente, come invece ci si aspetterebbe per la presenza dell'azione integrale.

È interessante notare che il valore di soglia $\varepsilon/K_I T$ è inversamente proporzionale al periodo di campionamento T , e quindi l'inconveniente di cui si è parlato risulta tanto più critico quanto minore è T . A tale circostanza si è già accennato nel Paragrafo 17.2.

17.8.2 Desaturazione dell'azione integrale

Anche nel caso dei regolatori digitali, la presenza combinata di un'azione integrale e di una saturazione sulla variabile di controllo, dovuta per esempio a vincoli

Figura 17.30
Regolatore con saturazione sulla variabile di controllo.

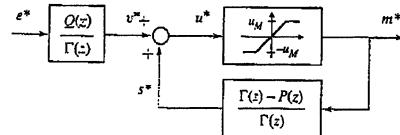


sull'attuatore, può provocare il fenomeno di *carica integrale*, o *wind-up*, discusso nel Paragrafo 14.3.2 con riferimento a sistemi di controllo a tempo continuo. Precisamente, si consideri il regolatore con saturazione indicato nella Figura 17.30, dove $R^*(z) = Q(z)/P(z)$ e

$$m^*(k) = \text{sat}(u^*(k)) = \begin{cases} -u_M, & u^*(k) < -u_M \\ u^*(k), & |u^*(k)| \leq u_M \\ u_M, & u^*(k) > u_M \end{cases}$$

Si ipotizzi poi che $R^*(z)$ contenga un integratore discreto, sia cioè $P(1) = 0$. Si indichi inoltre con n il grado del polinomio $P(z)$ e si supponga unitario il coefficiente del termine di grado n . Come si è visto nel Paragrafo 14.3.2, il fenomeno del *wind-up* consiste in una crescita anomala dello stato dell'integratore, che durante la fase di saturazione perde coerenza con l'effettiva variabile di controllo m^* . Ciò provoca una scarsa prontezza a ritornare nella zona di funzionamento lineare non appena si verifica un cambio di segno dell'errore, con conseguente degrado delle prestazioni.

Figura 17.31
Schema di desaturazione.



Per ovviare a questo inconveniente, si può ricorrere allo schema di desaturazione rappresentato nella Figura 17.31, in cui $\Gamma(z)$ è un polinomio di grado n con tutte le radici in modulo minori di uno e con il coefficiente di grado n pari a uno. Una scelta tipica consiste nel porre $\Gamma(z) = z^n$. La logica di funzionamento di questo schema è del tutto analoga a quella dello schema di Figura 14.13. Infatti, finché $|u^*(k)|$ rimane al di sotto della soglia di saturazione u_M , il legame tra e^* e m^* è descritto dalla funzione di trasferimento

$$\frac{M^*(z)}{E^*(z)} = \frac{Q(z)}{\Gamma(z)} \frac{1}{1 - \frac{\Gamma(z) - P(z)}{\Gamma(z)}} = \frac{Q(z)}{P(z)} = R^*(z)$$

mentre, quando la saturazione è attiva, s^* raggiunge asintoticamente (dopo n passi di campionamento se $\Gamma(z) = z^n$) il valore $\pm u_M$, e fa sì che il controllore sia in grado di tornare in zona di funzionamento lineare non appena v^* cambia segno. Si osservi inoltre che il sistema con funzione di trasferimento $(\Gamma(z) - P(z))/\Gamma(z)$, grazie alle ipotesi fatte, risulta strettamente proprio.

Esempio 17.7 Dato il regolatore

$$R^*(z) = \frac{Q(z)}{P(z)} = \frac{\mu z}{z - 1}$$

e posto $\Gamma(z) = z$, una sua realizzazione coerente con lo schema di desaturazione di Figura 17.31 è la seguente:

$$\begin{aligned} v^*(k) &= \mu e^*(k) \\ v^*(k) &= v^*(k) + m^*(k-1) \\ m^*(k) &= \text{sat}(u^*(k)) \end{aligned}$$

Esempio 17.8 (Seguito degli Esempi 14.4 e 14.5) Si consideri il sistema descritto dalla funzione di trasferimento

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

e il regolatore PI con $K_P = K_I = 2$, discretizzato con il metodo EI e periodo di campionamento $T = 0.1$. Si ipotizzi che sulla variabile di uscita del regolatore agisca una saturazione con $u_M = 1.05$. Sollecitando il sistema di controllo con una variazione a scalino unitario del segnale di riferimento a $t = 1$, si ottengono, in assenza di uno schema di desaturazione e quando invece si adotta lo schema della Figura 17.31 con $\Gamma(z) = z$ e $\Gamma(z) = z - 0.9$, i transitori della variabile di uscita riportati nella Figura 17.32.

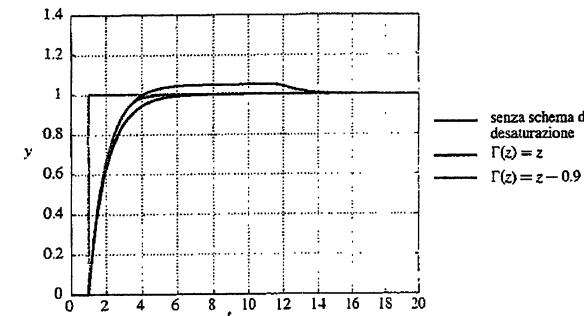


Figura 17.32
Andamento di w e y nell'Esempio 17.8 senza schema di desaturazione, con $\Gamma(z) = z$ e $\Gamma(z) = z - 0.9$.

Infine, si noti che lo schema rappresentato nella Figura 17.31, con opportune modifiche, si presta anche a risolvere il problema della commutazione "morbida" dalla regolazione manuale a quella automatica, come nell'analogo schema di Figura 14.15 per il caso a tempo continuo.

17.9 Conclusioni

In questo capitolo sono stati introdotti i sistemi di controllo ibridi, in cui il regolatore è realizzato con tecnologia digitale. In particolare, si è mostrato che la realizzazione di un sistema di controllo digitale richiede l'uso di convertitori A/D e D/A, le cui caratteristiche influenzano fortemente le prestazioni del sistema. Si è anche osservato che il campionamento di un segnale, a causa del fenomeno dell'aliasing, può provocare perdita di informazione. Inoltre, la presenza di un mantenitore di ordine zero introduce nel sistema di controllo uno sfasamento negativo, che deve essere considerato in fase di sintesi del regolatore. Queste considerazioni hanno condotto alla determinazione di opportuni criteri di scelta del periodo di campionamento.

L'analisi approssimata di un sistema di controllo ibrido può essere condotta in frequenza riconducendosi a quella di un equivalente sistema a tempo continuo, e il regolatore digitale può essere ricavato discretizzando con uno dei metodi presentati (EA, EI, TU) un regolatore a tempo continuo progettato con le tecniche dei Capitoli 11-14. È stata infine mostrata una possibile soluzione ai problemi di tipo realizzativo connessi con il fenomeno del wind-up e con l'implementazione del regolatore.

Esercizi

Esercizio 17.1 Si determini l'andamento nel tempo di un segnale $u(t)$ prodotto da un *mantenitore di ordine uno*, o *First Order Hold (FOH)*, che genera un segnale con andamento polinomiale di ordine uno compatibile con gli ultimi due valori dell'ingresso u^* .

Esercizio 17.2 Si analizzino i problemi, e le possibili soluzioni, relativi alla realizzazione di uno schema di controllo digitale in cui il tempo τ_c necessario per il calcolo di $u^*(k)$ sia maggiore del periodo di campionamento T che si intende utilizzare, per esempio $\tau_c = T + \tau'_c$, $0 < \tau'_c < T$.

Esercizio 17.3 Nell'ipotesi che venga rispettato il teorema del campionamento, noto $f^*(k) = f(kT)$, il segnale $f(t)$ può essere determinato mediante la seguente formula di Shannon, o del decampionatore ideale

$$f(t) = \sum_{k=-\infty}^{+\infty} f^*(k) \operatorname{sinc}(\omega_N t - k\pi), \quad \operatorname{sinc}(\varphi) = \begin{cases} \frac{\sin(\varphi)}{\varphi}, & \varphi \neq 0 \\ 1, & \varphi = 0 \end{cases}$$

1. Si verifichi che il decampionatore ideale è di fatto un filtro ideale passa-basso con banda passante $[0, \omega_N]$ e si confronti la sua risposta in frequenza con quella dello ZOH.

(Suggerimento: si mostri che la formula di antirtrasformazione di $f(t)$, opportunamente elaborata sfruttando l'Equazione (17.7), può essere scritta nella forma precedente).

2. Si analizzino le caratteristiche di causalità del decampionatore ideale.

3. Si valuti la possibilità di utilizzare il decampionatore ideale in problemi di controllo digitale.

Esercizio 17.4 Con riferimento all'Esempio 17.4, si determini un periodo di campionamento T , con cui realizzare il regolatore in modo digitale, tale per cui la riduzione del margine di fase rispetto al progetto nel continuo sia di 4° .

Esercizio 17.5 Si interpretino i risultati dell'esercizio precedente e dell'Esempio 17.4 in funzione del criterio euristico di scelta del periodo di campionamento dato dalla regola (17.30).

Esercizio 17.6 Si verifichi che l'immagine nel piano Z del semipiano sinistro del piano S , mediante la trasformazione inversa dei metodi EA, EI e TU, è quella illustrata nella Figura 17.25.

Esercizio 17.7 Si verifichi che i metodi EI e TU portano comunque alla determinazione di un regolatore digitale non strettamente proprio e si discutano le implicazioni di questo risultato in termini di problemi di implementazione della legge di controllo.

Esercizio 17.8 Per il sistema descritto da

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

si progetti un regolatore PI che garantisca il tempo di assettamento $T_{a1} = 5$, con smorzamento $\xi = 0.7$ dei poli in anello chiuso. Si determini il periodo di campionamento da utilizzare nella discretizzazione del PI in modo da avere 10 campioni per tempo di assettamento. Si valutino quindi l'attenuazione a ω_s garantita dal progetto a tempo continuo e la riduzione del margine di fase dovuta alla presenza dello ZOH. Si determini infine la funzione di trasferimento $R^*(z)$ del regolatore digitale utilizzando il metodo di Tustin e la sua realizzazione per tener conto del fenomeno del wind-up.

Esercizio 17.9 Con riferimento all'Esempio 15.6, si valuti la possibilità di realizzare in digitale i due regolatori $R_1(s)$ e $R_2(s)$. Si determini il corrispondente sistema di controllo ibrido scegliendo il periodo T con cui campionare i segnali v e y della Figura 15.15.

Esercizio 17.10 Dato un regolatore analogico con funzione di trasferimento $R^0(s)$, si mostri che è possibile determinare la funzione di trasferimento $R^*(z)$ di un regolatore digitale tale che $R^*(e^{j\omega T}) = R^0(j\bar{\omega})$, per una data pulsazione $\bar{\omega}$, pur di utilizzare la trasformazione

$$s = \frac{\bar{\omega}}{\tan(0.5\bar{\omega}T)} \frac{z-1}{z+1} \quad (17.51)$$

La (17.51) definisce il *metodo di Tustin con prewarping* ed è particolarmente utile quando $R^0(s)$ contiene singolarità con basso smorzamento. Per verifica, si discretizzi il regolatore

$$R^0(s) = \frac{1}{s^2 + 0.1s + 1}$$

con il metodo TU e con il metodo di Tustin con prewarping assumendo $\bar{\omega} = 1$; si confrontino poi i relativi diagrammi di Bode del modulo con quello associato a $R^0(s)$.

Sistemi di controllo digitale: analisi e sintesi a tempo discreto

18

18.1 Introduzione

In questo capitolo si presenterà un diverso approccio alla sintesi del regolatore digitale per sistemi SISO, che consiste dapprima nel ricavare il sistema a tempo discreto che descrive, negli istanti di campionamento, la serie del mantenitore, del processo e del campionatore. Quindi, a partire dalla corrispondente funzione di trasferimento $G^*(z)$, si determinerà direttamente la funzione di trasferimento $R^*(z)$ del regolatore digitale con opportuni algoritmi di sintesi. Per introdurre questi metodi è necessario preliminarmente estendere ai sistemi retroazionati a tempo discreto alcuni risultati di analisi e sintesi discussi nei Capitoli 9-13 per sistemi a tempo continuo. L'analisi condotta consentirà inoltre di ricavare altre tecniche di discretizzazione di un regolatore digitale, da utilizzare in alternativa ai metodi EA, EI e TU discussi nel capitolo precedente.

Nel capitolo saranno trattati i seguenti argomenti:

- la determinazione della funzione di trasferimento $G^*(z)$ e l'analisi delle sue caratteristiche (poli, zeri e guadagno);
- la presentazione di altre tecniche di discretizzazione di regolatori a tempo continuo;
- l'analisi dei sistemi retroazionati a tempo discreto in termini di stabilità, prestazioni statiche e dinamiche;
- alcuni metodi di sintesi di regolatori a tempo discreto per sistemi a tempo discreto descritti mediante la funzione di trasferimento: la sintesi con il luogo delle radici, il metodo ad assegnamento del modello, il metodo ad assegnamento dei poli.

18.2 Sistema a segnali campionati

Nello schema a campionamento dell'uscita della Figura 18.1, la serie del mantenitore, del processo e del campionatore può essere interpretata come un unico

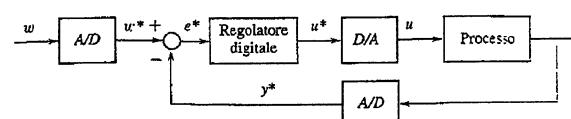
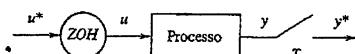


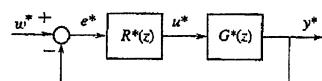
Figura 18.1
Schema di controllo digitale a campionamento dell'uscita e del riferimento.

Figura 18.2
Sistema a segnali campionati.



sistema, riportato nella Figura 18.2 e comunemente chiamato *sistema a segnali campionati*, che ha come ingresso e uscita le variabili a tempo discreto u^* e y^* . Esso può quindi essere rappresentato come un sistema a tempo discreto in grado di descrivere compiutamente i legami che intercorrono tra u^* e y^* . Nel caso in cui il processo sia un sistema dinamico lineare e stazionario a tempo continuo, si mostrerà in questo paragrafo che il sistema a segnali campionati è un sistema dinamico lineare e stazionario a tempo discreto descritto da una funzione di trasferimento $G^*(z)$. Pertanto, lo studio del sistema di controllo di Figura 18.1 può essere ricondotto a quello del sistema retroazionato a tempo discreto di Figura 18.3, in cui compaiono solamente sottosistemi a tempo discreto, cioè il regolatore digitale e il sistema a segnali campionati. L'analisi e la sintesi di tale sistema di controllo potranno poi essere affrontate estendendo al caso in esame molti dei risultati e dei metodi precedentemente introdotti per i sistemi retroazionati a tempo continuo, quali il criterio di Nyquist, il luogo delle radici e la tecnica ad assegnamento degli autovalori.

Figura 18.3
Sistema a tempo discreto equivalente.



18.2.1 Calcolo della funzione di trasferimento

Nel sistema a segnali campionati di Figura 18.2 si assume che il processo sia descritto dal sistema lineare e stazionario

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (18.1)$$

$$y(t) = Cx(t) + Du(t) \quad (18.2)$$

Il movimento dello stato (si veda l'Equazione (3.3)) a partire dall'istante iniziale $t = t_0$ è

$$x(\tau) = e^{A(\tau-t_0)}x(t_0) + \int_{t_0}^{\tau} e^{A(\tau-\xi)}Bu(\xi)d\xi \quad (18.3)$$

Si ponga ora $t_0 = kT$, $\tau = kT + T$ e si osservi che, ipotizzando $\tau_m = 0$, per la presenza dello ZOH

$$u(\xi) = u(kT) = u^*(k) \quad , \quad \xi \in [kT, kT + T] \quad (18.4)$$

dove $u^*(k)$ è il segnale in ingresso al mantenitore. Per la (18.4), l'Equazione (18.3) implica che

$$x(kT + T) = e^{AT}x(kT) + \int_{kT}^{kT+T} e^{A(kT+\tau-\xi)}Bu(\xi)d\xi \quad (18.5)$$

Posto $\eta = kT + T - \xi$, si ha

$$\int_{kT}^{kT+T} e^{A(kT+\tau-\xi)}Bu(\xi)d\xi = \int_0^T e^{A\eta}Bd\eta \quad (18.6)$$

Definendo

$$x^*(k) = x(kT) \quad (18.7)$$

e ricordando che $y^*(k) = y(kT)$, dalle (18.4)-(18.7), la relazione tra u^* e y^* è descritta da

$$x^*(k+1) = A^*x^*(k) + B^*u^*(k) \quad (18.8)$$

$$y^*(k) = Cx^*(k) + Du^*(k) \quad (18.9)$$

con

$$A^* = e^{AT} \quad , \quad B^* = \int_0^T e^{A\eta}Bd\eta \quad (18.10)$$

Il sistema a tempo discreto (18.8), (18.9) è lineare e stazionario, e rappresenta compiutamente il comportamento del processo e dei convertitori negli istanti di campionamento, mentre non fornisce alcuna informazione sull'andamento delle variabili a tempo continuo x e y tra due istanti di campionamento successivi. Il sistema a segnali campionati (18.8), (18.9) può essere anche descritto dalla funzione di trasferimento

$$G^*(z) = C(zI - A^*)^{-1}B^* + D \quad (18.11)$$

18.2.2 Presenza di ritardi di tempo

Nella determinazione del sistema a segnali campionati (18.8), (18.9) si è assunto implicitamente che il processo (18.1), (18.2) non contenesse ritardi di tempo. Questa ipotesi è restrittiva per due ragioni. In primo luogo, in molti casi i sistemi sotto controllo hanno una dinamica caratterizzata dalla presenza di ritardi. Inoltre, come visto nel Paragrafo 17.2.4, è possibile modellizzare con un ritardo $\tau_m = \tau_c \neq 0$ lo sfasamento temporale che intercorre tra l'acquisizione del segnale con il campionatore e la determinazione della variabile di controllo tramite il mantenitore.

Per maggiore generalità, si consideri, quindi, il sistema

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

$$y(t) = Cx(t) + Du(t - \tau)$$

dove il ritardo τ è tale che

$$\tau = hT + \varepsilon$$

con h intero non negativo e $\varepsilon \in [0, T]$. Si possono ora presentare due casi, a seconda che τ sia un multiplo intero di T oppure no, che vanno analizzati separatamente.

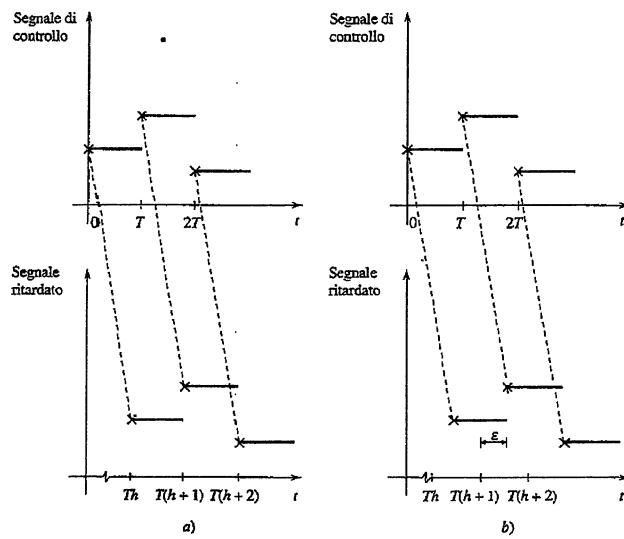
Caso $\varepsilon = 0$ In questo caso, come rappresentato simbolicamente nella Figura 18.4a, $u(t - \tau)$ è ancora costante nell'intervallo di integrazione $[kT, kT + T]$ e risulta $u(t - \tau) = u^*(k-h)$ per $t \in [kT, kT + T]$. Quindi, è possibile ripercorrere i passaggi precedenti per ottenere il sistema a segnali campionati

$$x^*(k+1) = A^*x^*(k) + B^*u^*(k-h) \quad (18.12)$$

$$y^*(k) = Cx^*(k) + Du^*(k-h) \quad (18.13)$$

dove A^* e B^* sono ancora definite dalle (18.10). Il sistema (18.12), (18.13) con ingresso u^* e uscita y^* può poi essere descritto tramite la sua funzione di trasferimento che, come mostrato nel Paragrafo 8.2.4, coincide con la (18.11) divisa per z^h .

Figura 18.4
Trasiazione del segnale di ingresso dovuta al ritardo t : a) $\epsilon = 0$; b) $\epsilon \neq 0$.



Caso $\epsilon \neq 0$ In questo caso, come mostrato nella Figura 18.4b, l'ingresso u non si mantiene più costante nell'intervalle di integrazione, che, pertanto, deve essere suddiviso in due parti. Si ricava quindi

$$\begin{aligned} & \int_{kT}^{kT+T} e^{A(kT+T-\xi)} Bu(\xi - \tau) d\xi = \\ &= \int_{kT}^{kT+\epsilon} e^{A(kT+T-\xi)} Bd\xi u(kT - (h+1)T) + \int_{kT+\epsilon}^{kT+T} e^{A(kT+T-\xi)} Bd\xi u(kT - hT) \end{aligned}$$

Ponendo

$$\begin{aligned} B_0^* &= \int_{kT}^{kT+\epsilon} e^{A(kT+T-\xi)} Bd\xi = \int_{T-\epsilon}^T e^{A\eta} Bd\eta, \\ B_1^* &= \int_{kT+\epsilon}^{kT+T} e^{A(kT+T-\xi)} Bd\xi = \int_0^{T-\epsilon} e^{A\eta} Bd\eta \end{aligned}$$

con elaborazioni analoghe alle precedenti si determina il sistema a segnali campionati

$$\begin{aligned} x^*(k+1) &= A^*x^*(k) + B_1^*u^*(k-h) + B_0^*u^*(k-h-1) \\ y^*(k) &= Cx^*(k) + Du^*(k-h-1) \end{aligned}$$

Per ricondurre questo sistema a una forma corrispondente a quella del sistema (18.12), (18.13), è necessario allargare il vettore di stato e assumere come ingresso

$u^*(k-h)$ per ottenere

$$\begin{aligned} \begin{bmatrix} x^*(k+1) \\ u^*(k-h) \end{bmatrix} &= \begin{bmatrix} A^* & B_0^* \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^*(k) \\ u^*(k-h-1) \end{bmatrix} + \begin{bmatrix} B_1^* \\ 1 \end{bmatrix} u^*(k-h) \\ y^*(k) &= [C \ D] \begin{bmatrix} x^*(k) \\ u^*(k-h-1) \end{bmatrix} \end{aligned}$$

La funzione di trasferimento tra l'ingresso u^* e l'uscita y^* ha quindi $h+1$ poli nell'origine.

Nel seguito della trattazione, per semplicità, si assumerà comunque $\tau = 0$, lasciando al lettore il compito di estendere i risultati al caso $\tau > 0$.

18.2.3 Autovalori del sistema a segnali campionati

In questo paragrafo e nei successivi saranno esaminate le connessioni tra le proprietà del sistema a tempo continuo (18.1), (18.2) e quelle della sua versione a segnali campionati (18.8), (18.9). Si consideri dapprima il legame tra i rispettivi autovalori. La relazione (18.10) tra A e A^* implica che all'autovalore s_i , $i = 1, \dots, n$, di A corrisponda l'autovalore

$$z_i = e^{s_i T}$$

di A^* ; gli autovalori seguono quindi la cosiddetta trasformazione di campionamento, già introdotta al Paragrafo 17.5,

$$z = e^{s T} \quad (18.14)$$

tra i punti del piano S , di definizione della variabile complessa s , e i punti del piano Z , di definizione della variabile complessa z . La corrispondenza tra s_i e z_i è univoca, ma non biunivoca. Infatti, se si considerano due autovalori s_p e s_q di A tali che

$$s_q = s_p + j \frac{2\pi}{T} \quad (18.15)$$

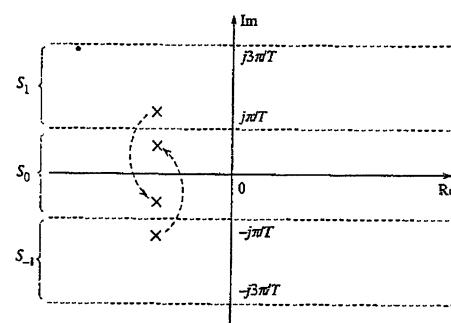
con h intero, applicando la (18.14) si ottiene

$$z_q = e^{s_q T} = e^{s_p T} e^{j2\pi h} = e^{s_p T} = z_p$$

cioè due autovalori diversi del sistema a tempo continuo generano due autovalori coincidenti del sistema a segnali campionati associato. Per interpretare meglio questo fenomeno, che può essere visto come una diversa manifestazione dell'aliasing, si divida il piano S in strisce S_h , con h intero, parallele all'asse reale, di ampiezza $2\pi/T$ e con S_0 simmetrica rispetto all'asse reale, come mostrato nella Figura 18.5.

La situazione critica rappresentata dalla (18.15) ha luogo quando due autovalori del sistema a tempo continuo differiscono solo nella parte immaginaria per quantità multiple di $2\pi/T$. In pratica, si può pensare che nel campionamento tutte le strisce S_h , e gli autovalori di A in esse presenti, vengano dapprima traslate su S_0 e che quindi vengano trasformati con la (18.14) gli autovalori ormai contenuti tutti in S_0 (si veda ancora la Figura 18.5). In questo modo, autovalori che vanno a sovrapporsi in S_0 diventano indistinguibili dopo il campionamento, e comunque autovalori con modulo della parte immaginaria maggiore di π/T vengono equivocati come autovalori a più bassa frequenza. A questo riguardo, si osservi che al diminuire di T aumenta l'ampiezza di S_0 e, per T sufficientemente piccolo rispetto alla dinamica del sistema, S_0 contiene tutti gli autovalori di A , e in questo caso la trasformazione (18.14) si può ritenere biunivoca.

Figura 18.5
Strisce del piano S di trasformazione degli autovalori.



Le considerazioni precedenti possono far intuire la correttezza del seguente risultato, qui non dimostrato.

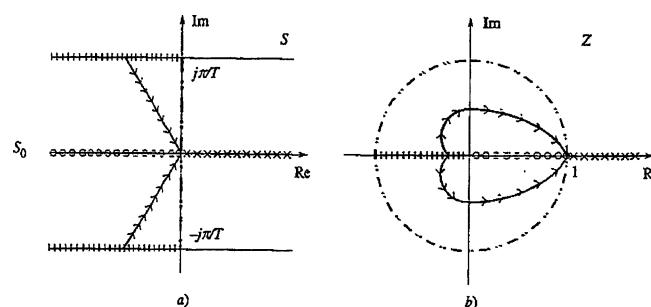
Teorema 18.1 Il sistema a segnali campionati (18.8), (18.9) è raggiungibile e osservabile se e solo se:

- a) il sistema (18.1), (18.2) è raggiungibile e osservabile;
- b) non esistono autovalori di A , s_p e s_q , tali che

$$s_p = s_q + j \frac{2\pi}{T} h, \quad h \neq 0 \text{ intero}$$

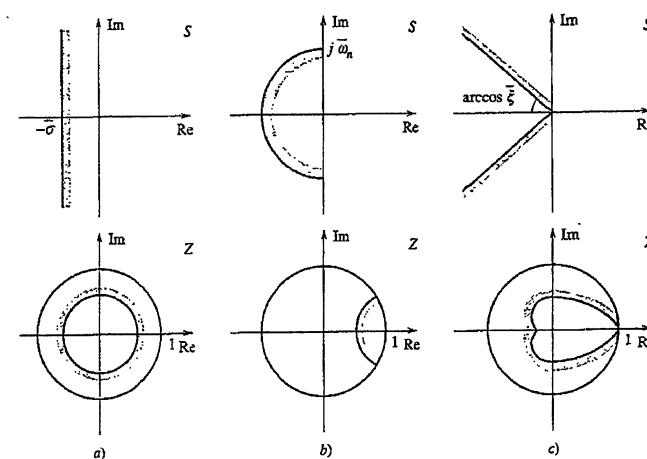
La trasformazione di campionamento (18.14) applicata agli autovalori consente anche di stabilire una corrispondenza tra particolari luoghi di punti del piano S e regioni del piano Z . Dapprima si osservi che punti del piano S a parte reale negativa vengono trasformati in punti del piano Z con modulo minore di 1. Pertanto, autovalori del sistema a tempo continuo (18.1), (18.2) posti all'interno della regione di asintotica stabilità del piano S diventano autovalori del sistema a segnali campionati (18.8), (18.9) appartenenti alla regione di asintotica stabilità del piano Z . Inoltre, limitandosi per le ragioni viste a considerare la striscia S_0 , nella Figura 18.6 sono riportate le corrispondenze più significative.

Figura 18.6
Corrispondenza piano S -piano Z .



Sempre mediante la (18.14) è possibile determinare come alcune regioni del piano S , di particolare interesse nel caratterizzare la dinamica del sistema (Figura 12.17), si trasformino nel piano Z . Nella Figura 18.7 si riportano le immagini nel piano Z dei luoghi dei punti in S con parte reale minore di un determinato valore $-\bar{\sigma}$, con pulsazione naturale maggiore di $\bar{\omega}_n$ e parte reale negativa, con smorzamento maggiore di ξ .

Figura 18.7
Luoghi in S e corrispondenze in Z con:
a) $\operatorname{Re}(s) < -\bar{\sigma}$; b) $\omega_n > \bar{\omega}_n$;
c) $\xi > \bar{\xi}$.



Periodo di campionamento e quantizzazione La scelta di un valore di T troppo piccolo rispetto alla dinamica del sistema può creare difficoltà di tipo numerico per effetto della quantizzazione. Infatti, alla luce della (18.14), per $T \rightarrow 0$ tutti gli autovalori di A^* tendono al punto 1, rendendo indistinguibili, in una rappresentazione con aritmetica finita, autovalori anche molto diversi.

Esempio 18.1 Si consideri un sistema a tempo continuo del secondo ordine con autovalori $s_1 = -1$ e $s_2 = -7.83$. Il sistema a segnali campionati a esso associato, ricavato utilizzando il periodo di campionamento $T = 0.001$, ha gli autovalori

$$z_1 = e^{-0.001} = 0.999 \quad z_2 = e^{-0.00783} = 0.9922$$

Per codificare z_1 e z_2 si impieghi un'aritmetica in virgola fissa, una parola di 8 bit. Si ricorra a un'approssimazione per troncamento e si ipotizzi che, poiché la parte intera di z_1 e z_2 è nulla, gli 8 bit siano tutti utilizzati per rappresentare la loro parte frazionaria. In questa ipotesi, il risultato della conversione è

$$z_{b1} = z_{b2} = .1111110$$

I valori z_{b1} e z_{b2} sono quindi indistinguibili, benché gli autovalori s_1 e s_2 che li hanno generati siano molto diversi.

Se invece si impiega il periodo di campionamento $T = 0.1$, si ha

$$z_1 = e^{-0.1} = 0.9048 \quad z_2 = e^{-0.783} = 0.457$$

la cui codifica binaria è

$$z_{b1} = .1100111 \quad z_{b2} = .0111010$$

Questi due valori, ritrasformati in codifica decimale, equivalgono a

$$\tilde{z}_1 = 0.9023 \quad \tilde{z}_2 = 0.4531$$

Il loro confronto con z_1 e z_2 mostra che comunque l'impiego di una parola di lunghezza finita comporta degli errori dovuti al troncamento, ma che una scelta appropriata di T consente, nonostante la quantizzazione, di distinguere autovalori sostanzialmente diversi. A riprova di ciò si noti che applicando la trasformazione di campionamento inversa a \tilde{z}_1 e \tilde{z}_2 si ottengono i valori

$$\tilde{s}_1 = \frac{1}{T} \ln \tilde{z}_1 = -1.028 \quad \tilde{s}_2 = \frac{1}{T} \ln \tilde{z}_2 = -7.916$$

poco dissimili dagli autovalori s_1 e s_2 di partenza.

18.2.4 Determinazione diretta della funzione di trasferimento

Quando il sistema sotto controllo è rappresentato tramite la sua funzione di trasferimento $G(s)$, è possibile determinare direttamente $G^*(z)$ senza passare dalla rappresentazione di stato (18.1), (18.2). Per questo basta osservare che la risposta a uno scalino unitario (a tempo continuo) di $G^*(z)$ deve coincidere con i valori ottenuti dal campionamento della risposta a uno scalino unitario (a tempo continuo) di $G(s)$. Il calcolo di $G^*(z)$ è compiuto quindi secondo i seguenti passi:

1. si determina la trasformata della risposta allo scalino del sistema a tempo continuo: $Y(s) = G(s)/s$;
2. si antitransforma $Y(s)$ e si valuta l'uscita $y(kT)$ negli istanti di campionamento;
3. si calcola la trasformata Zeta $Y^*(z)$ di $y^*(k) = y(kT)$;
4. si ricava, infine, $G^*(z)$ come rapporto tra $Y^*(z)$ e la trasformata dell'ingresso $U^*(z)$, cioè

$$G^*(z) = Y^*(z) \frac{z - 1}{z}$$

Esempio 18.2 Si determini la funzione di trasferimento del sistema a segnali campionati corrispondente al processo descritto da

$$G(s) = \frac{a}{s + a} \quad (18.16)$$

La risposta allo scalino unitario di $G(s)$ ha come trasformata

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

e, quindi, vale

$$y(t) = 1 - e^{-at}, \quad t \geq 0$$

Campionando $y(t)$ e trasformando, risulta

$$Y^*(z) = \sum_{k=0}^{\infty} z^{-k} - \sum_{k=0}^{\infty} (e^{-aT} z^{-1})^k = \frac{z}{z - 1} - \frac{z}{z - e^{-aT}}$$

In conclusione

$$G^*(z) = \left(\frac{z}{z - 1} - \frac{z}{z - e^{-aT}} \right) \frac{z - 1}{z} = \frac{1 - e^{-aT}}{z - e^{-aT}} \quad (18.17)$$

Dall'analisi delle (18.16), (18.17) si può notare che, come atteso, il polo di $G(s)$ si trasforma nel polo di $G^*(z)$ mediante la (18.14). Inoltre si ha $G^*(1) = G(0)$, cioè si conserva il guadagno.

Esempio 18.3 Applicando all'integratore

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

la procedura precedente, si ottiene $Y(s) = 1/s^2$, quindi $y(t) = t$, $t \geq 0$, e

$$Y^*(z) = \sum_{k=0}^{\infty} T k z^{-k} = \frac{Tz}{(z - 1)^2}$$

Pertanto

$$G^*(z) = \frac{Tz}{(z - 1)^2} \frac{z - 1}{z} = \frac{T}{z - 1} \quad (18.18)$$

Si noti che, mentre a tempo continuo la trasformata di Laplace $1/s$ dello scalino unitario coincide con la funzione di trasferimento dell'integratore con guadagno unitario, a tempo discreto la trasformata Zeta dello scalino unitario è diversa dalla funzione di trasferimento del sistema a segnali campionati (18.18).

Esempio 18.4 Con semplici calcoli si può verificare che, dato il sistema

$$G(s) = \frac{a}{s(s + a)} \quad (18.19)$$

la funzione di trasferimento del sistema a segnali campionati è

$$G^*(z) = \frac{(aT - 1 + e^{-aT})z + (1 - e^{-aT}(1 + aT))}{a(z^2 - (1 + e^{-aT})z + e^{-aT})} \quad (18.20)$$

Si noti che $G^*(z)$ ha uno zero, mentre $G(s)$ non ne ha. Il legame tra gli zeri del sistema a tempo continuo e quello del corrispondente sistema a segnali campionati è, quindi, non banale e merita un approfondimento.

18.2.5 Zeri del sistema a segnali campionati

Per analizzare le caratteristiche degli zeri dei sistemi a segnali campionati, è opportuno preliminarmente generalizzare il risultato dell'Esempio 18.3.

Integratori multipli Mediante la procedura di calcolo di $G^*(z)$ a partire da $G(s)$, è possibile verificare che il sistema a segnali campionati corrispondente a

$$G(s) = \frac{1}{s^n}$$

ha la funzione di trasferimento

$$G^*(z) = \frac{T^n}{n!} \frac{A_n(z)}{(z - 1)^n} \quad (18.21)$$

dove $A_n(z)$ è un polinomio di ordine $n - 1$. Nella Tabella 18.1 sono riportati, per $n \leq 4$, i polinomi $A_n(z)$ unitamente alle loro radici, che costituiscono quindi gli zeri di $G^*(z)$. Si osservi che questi zeri sono tutti reali, negativi e alcuni di essi sono esterni alla regione di stabilità del piano Z . Inoltre si può dimostrare che per ogni n risulta $A_n(1) = n!$.

Tabella 18.1
Polinomi $A_n(z)$.

n	A_n	Radici
1	1	—
2	$z - 1$	-1
3	$z^2 + 4z - 1$	-0.268, -3.73
4	$z^3 + 11z^2 + 11z - 1$	-0.1, -1, -9.89

Zeri del campionamento La determinazione del numero e della collocazione degli zeri di $G^*(z)$ è una questione tutt'altro che banale, per la quale ci si limiterà a discutere alcuni importanti risultati basandosi su considerazioni prevalentemente qualitative. A questo scopo, nel seguito si considererà il problema di ricavare, dato il sistema

$$G(s) = \rho \frac{\prod_{i=1}^m (s + a_i)}{\prod_{j=1}^n (s + b_j)} = \frac{\rho}{s^{n-m}} \frac{\prod_{i=1}^m (1 + a_i/s)}{\prod_{j=1}^n (1 + b_j/s)}, \quad m < n \quad (18.22)$$

un'espressione approssimata, per $T \rightarrow 0$, della funzione di trasferimento $G^*(z)$ del corrispondente sistema a segnali campionati. Dall'analisi di questa funzione approssimante si potranno poi dedurre informazioni utili sul numero e sulle caratteristiche degli zeri di $G^*(z)$.

In primo luogo, si noti che, per $|s| \rightarrow \infty$, $G(s)$ può essere approssimata da

$$\tilde{G}(s) = \frac{\rho}{s^{n-m}} \quad (18.23)$$

a cui, per la (18.21), corrisponde il sistema a segnali campionati

$$\tilde{G}^*(z) = \frac{\rho T^{n-m}}{(n-m)!} \frac{A_{n-m}(z)}{(z-1)^{n-m}} \quad (18.24)$$

con $n-m$ poli nel punto $z=1$ e $n-m-1$ zeri coincidenti con le radici di $A_{n-m}(z)$.

Nelle ipotesi del Teorema 18.1, la funzione di trasferimento $G^*(z)$ ha n poli e quindi, per ricavare una sua approssimante dello stesso ordine, è opportuno riscrivere la (18.24) nella forma

$$\tilde{G}^*(z) = \frac{\rho T^{n-m}}{(n-m)!} \frac{A_{n-m}(z)(z-1)^m}{(z-1)^n} \quad (18.25)$$

Si osservi ora che l'ipotesi $|s| \rightarrow \infty$ corrisponde nel dominio della frequenza a considerare $\omega \rightarrow \infty$ e pertanto, per elevati valori di ω , la funzione di trasferimento $G(s)$ può essere rappresentata dalla (18.23). Se ora nella determinazione del sistema a segnali campionati associato alla (18.22) si utilizza un periodo di campionamento $T \rightarrow 0$, che corrisponde a $\omega_N \rightarrow \infty$, in base alle considerazioni precedenti si può ritenere che la (18.25) sia proprio l'approssimante di $G^*(z)$ cercata. Si giustificano così i seguenti risultati.

Proposizione 18.1 Gli m zeri di $G(s)$ generano altrettanti zeri di $G^*(z)$ che, per $T \rightarrow 0$, tendono al punto $z=1$; inoltre, sempre per $T \rightarrow 0$, gli m zeri di $G^*(z)$ seguono anch'essi la trasformazione di campionamento (18.14).

Proposizione 18.2 Quando $n-m > 1$ nascono $n-m-1$ zeri, detti *zeri del campionamento*, che per $T \rightarrow 0$ tendono alle radici del polinomio $A_{n-m}(z)$ (Tabella 18.1).

Questi risultati, benché asintotici, forniscono utili indicazioni anche per T sufficientemente piccolo. In particolare, per valori di ω_N elevati rispetto alla dinamica di $G(s)$, cioè alla collocazione delle sue singolarità, tutti i poli e gli zeri di $G(s)$ si trasformano in singolarità di $G^*(z)$ prossime al punto $z=1$, creando problemi notevoli di carattere numerico e "quasi cancellazioni" che possono rendere critico il progetto di un regolatore digitale effettuato a partire da $G^*(z)$. Inoltre, eventuali zeri del campionamento si trovano sempre sull'asse reale negativo e, ancora per T sufficientemente piccolo, sono molto vicini alle radici del polinomio $A_{n-m}(z)$ corrispondente; quindi $G^*(z)$ può avere zeri esterni alla regione di stabilità del piano Z anche se $G(s)$ è a fase minima.

La nascita degli zeri del campionamento si può anche spiegare come segue: se $G(s)$ è razionale, la sua risposta allo scalino $y(t)$ è genericamente diversa da zero per $t > 0$ indipendentemente dal grado relativo $n-m$. Affinché la risposta allo scalino $y^*(k)$ di $G^*(z)$ coincida con $y(t)$ negli istanti di campionamento, è necessario che il tempo di latenza sia unitario, e quindi che il grado relativo di $G^*(z)$ sia 1 (si vedano al riguardo il Paragrafo 8.3.1 e il metodo di anttrasformazione della lunga divisione riportato nel Paragrafo C.3.3). Pertanto, $G^*(z)$ deve avere $n-1$ zeri, cioè devono nascere gli $n-m-1$ zeri del campionamento.

18.2.6 Guadagno e tipo del sistema a segnali campionati

Un ultimo risultato riguarda il legame tra i guadagni di $G(s)$ e $G^*(z)$.

Se $G(s)$ è di tipo $g=0$, risulta

$$G^*(1) = G(0)$$

e cioè il guadagno si conserva, come verificato nell'Esempio 18.2. La correttezza di questo risultato può essere dimostrata formalmente in base alle (18.10), (18.11) o, più intuitivamente, osservando che se il sistema a tempo continuo (18.1), (18.2) è all'equilibrio con ingresso \bar{u} e uscita \bar{y} , il sistema a segnali campionati (18.8), (18.9) è anch'esso all'equilibrio con ingresso $\bar{u}^* = \bar{u}$ e uscita $\bar{y}^* = \bar{y}$.

Se invece $G(s)$ è di tipo $g > 0$ e le ipotesi del Teorema 18.1 sono soddisfatte, $G^*(z)$ è anch'essa di tipo g . Inoltre, posto

$$\mu = \lim_{s \rightarrow 0} s^g G(s) \quad , \quad \mu^* = \lim_{z \rightarrow 1} (z-1)^g G^*(z)$$

si ha

$$\mu^* = T^g \mu \quad (18.26)$$

come negli Esempi 18.3 e 18.4.

18.2.7 Esempio riassuntivo

Il prossimo esempio evidenzia come le caratteristiche di $G^*(z)$ cambino in funzione del periodo di campionamento T adottato.

Esempio 18.5 Per il sistema descritto da

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

si determini $G^*(z)$ utilizzando i periodi di campionamento $T_a = 1$ e $T_b = 0.1$. Mediante uno dei procedimenti visti, è possibile ricavare le funzioni di trasferimento

$$G_a^*(z) = \frac{0.2642z + 0.1353}{(z - 0.3679)^2}$$

per $T_a = 1$ e

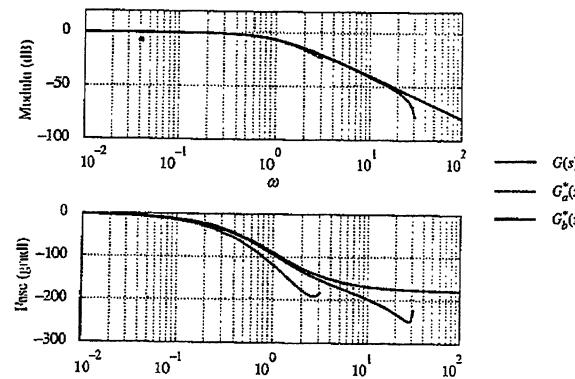
$$G_b^*(z) = \frac{0.0047z + 0.0044}{(z - 0.9048)^2}$$

per $T_b = 0.1$, che si prestano ad alcune interessanti considerazioni.

- Poiché il grado relativo di $G(s)$ è 2, nasce uno zero del campionamento, $z_a = -0.5121$ per $G_a^*(z)$ e $z_b = -0.9362$ per $G_b^*(z)$, che in entrambi i casi è reale negativo. Inoltre, dato che $T_b < T_a$, z_b è più vicino a z_a alla radice -1 del polinomio $A_2(z)$; un'ulteriore riduzione di T farebbe avvicinare ancora di più lo zero del sistema a segnali campionati al punto $z=-1$.
- Il guadagno di $G(s)$, $G_a^*(z)$ e $G_b^*(z)$ è lo stesso, cioè $G(0) = G_a^*(1) = G_b^*(1) = 1$.
- I due poli di $G_a^*(z)$ e $G_b^*(z)$ si ottengono da quelli di $G(s)$ mediante la trasformazione di campionamento (18.14) e quelli di $G_b^*(z)$ sono più vicini al punto $z=1$ rispetto a quelli di $G_a^*(z)$. Ciò è in accordo con il fatto che i poli del sistema a segnali campionati tendono al punto 1 al decrescere del periodo di campionamento.

Si osservi inoltre che, esaminando $G_a^*(z)$ e $G_b^*(z)$ senza ricordare che sono frutto del campionamento con periodi diversi, si potrebbe essere indotti a ritenere che $G_a^*(z)$ abbia una dinamica più veloce di $G_b^*(z)$, in quanto i suoi poli sono più vicini all'origine del piano Z . In effetti, indicando con n_a e con n_b il numero di campioni necessario perché la risposta allo scalino di $G_a^*(z)$ e di $G_b^*(z)$ arrivi, per esempio, al 90% del valore di regime, si ha $n_a < n_b$. Tuttavia le due risposte coincidono entrambe con il campionamento (con periodi diversi) della medesima risposta e risulta evidentemente $n_a T_a = n_b T_b$.

Figura 18.8
Diagrammi di Bode
associati a $G(s)$,
 $G_a^*(z)$ e $G_b^*(z)$.



Infine, può essere utile confrontare i diagrammi di Bode, riportati nella Figura 18.8, di $G(j\omega)$, $G_a^*(e^{j\omega T_a})$ e $G_b^*(e^{j\omega T_b})$. Per quanto già detto, i diagrammi di $G_a^*(e^{j\omega T_a})$ e $G_b^*(e^{j\omega T_b})$ sono rappresentati soltanto sino alle rispettive pulsazioni di Nyquist $\omega_{Na} = \pi$ e $\omega_{Nb} = 10\pi$. Da tali diagrammi si nota che la risposta in frequenza del sistema a segnali campionati approssima discretamente quella del sistema originario solo a bassa frequenza, tanto meglio quanto più piccolo è il periodo di campionamento.

18.3 Altri metodi di discretizzazione del regolatore

Prima di procedere nello studio dei sistemi a segnali campionati e delle tecniche di analisi e sintesi a essi associati, è opportuno riprendere il problema già trattato nel Paragrafo 17.7, relativo alla discretizzazione di un regolatore a tempo continuo. I risultati appena presentati motivano infatti l'introduzione di due nuove tecniche di discretizzazione. Si ricorda preliminarmente che, dato un regolatore descritto dalla funzione di trasferimento $R^o(s)$, l'obiettivo dei metodi approssimati di discretizzazione consiste nel determinare un regolatore digitale descritto da $R^*(z)$ e tale che

$$R^*(e^{j\omega T}) \simeq R^o(j\omega) \quad , \quad 0 \leq \omega \leq \omega_N \quad (18.27)$$

18.3.1 Tenuta e campionamento

Per determinare un regolatore digitale che soddisfi la (18.27) si può sostituire al controllore analogico di riferimento $R^o(s)$ la sua versione a segnali campionati. La funzione di trasferimento $R^*(z)$ viene quindi calcolata sulla base dello schema di Figura 18.9 con le tecniche presentate nel Paragrafo 18.2.

Malgrado l'apparente ragionevolezza di tale scelta, questo metodo di discretizzazione può portare a soluzioni non del tutto soddisfacenti. Infatti, lo schema equivalente del sistema di controllo risultante, che è riportato nella Figura 18.10,

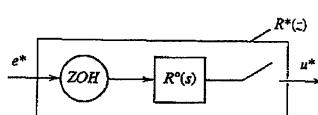


Figura 18.9
Regolatore analogico
a segnali campionati.

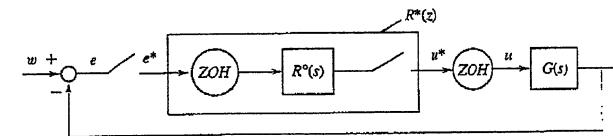


Figura 18.10
Schema equivalente
del sistema di controllo
digitale ottenuto con
il metodo HE.

contiene una doppia coppia di campionatore e mantenitore in aggiunta rispetto alla configurazione del sistema a tempo continuo di riferimento. Come si sa, ognuna di queste coppie di convertitori è assimilabile in prima approssimazione a un ritardo di tempo pari a $T/2$ (si veda il Paragrafo 17.5) e dunque l'anello di Figura 18.10 presenta un ritardo aggiuntivo T rispetto allo schema puramente analogico. Perché l'effetto di tale ritardo non infici le prestazioni del sistema di controllo, è allora necessario che il progetto preliminare di $R^o(s)$ garantisca un'eccedenza di margine di fase doppia rispetto a quella normalmente richiesta quando si utilizza l'approccio della discretizzazione.

Il metodo appena visto prende il nome di *tenuta e campionamento* o di *hold equivalence* (HE).

18.3.2 Trasformazione diretta di poli e zeri

Vista la natura in ogni caso approssimata del metodo HE appena descritto, è legittimo individuarne opportune varianti che consentono una sintesi ancora più semplice del regolatore digitale. Una di queste versioni semplificate va sotto il nome di *trasformazione diretta di poli e zeri*, o *pole-zero mapping* (nel seguito PZ). Essa si basa sull'osservazione che le uniche caratteristiche della funzione $R^*(z)$ associate allo schema di Figura 18.9 che non possono essere dedotte in modo banale da $R^o(s)$ sono quelle relative agli zeri. Infatti, per quanto visto nel Paragrafo 18.2.3, i poli di $R^*(z)$ si ottengono direttamente da quelli di $R^o(s)$ attraverso la trasformazione di campionamento (18.14). Inoltre, vale la relazione (18.26) tra il guadagno di $R^*(z)$ e quello di $R^o(s)$.

Il metodo PZ consiste nel determinare la posizione degli zeri ricorrendo ai risultati asintotici del Paragrafo 18.2.5, che a rigore sono validi solo per $T \rightarrow 0$. In definitiva, data una generica $R^o(s)$ di ordine n , con m zeri, tipo g e guadagno μ , si calcola $R^*(z)$ in questo modo:

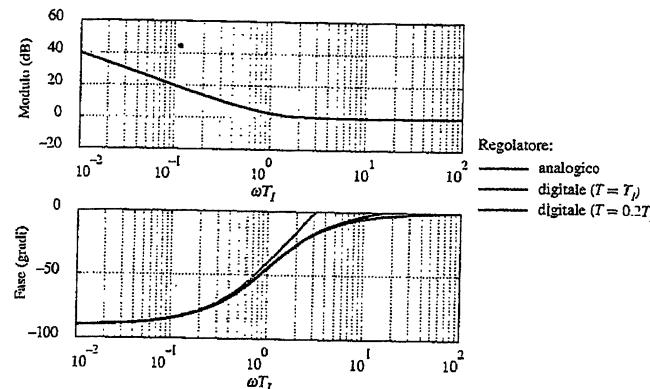
- il guadagno è $\mu^* = T^g \mu$;
- gli n poli sono ottenuti da quelli di $R^o(s)$ mediante la trasformazione $z = e^{sT}$;
- m zeri sono ottenuti da quelli di $R^o(s)$ mediante la trasformazione $z = e^{sT}$;
- se $m < n - 1$, vengono aggiunti $n - m - 1$ zeri (del campionamento) coincidenti con le radici del polinomio $A_{n-m}(z)$ della Tabella 18.1.

A volte viene proposta una versione ulteriormente semplificata del metodo, che prevede di collocare tutti gli $n - m - 1$ zeri del campionamento nel punto $z = -1$, cioè in corrispondenza dell'unica radice del polinomio $A_2(z)$, anche quando $n - m > 2$. Se poi si desidera ottenere un regolatore non strettamente proprio, gli zeri da aggiungere sono in numero $n - m$.

Esempio 18.6 (Seguito dell'Esempio 17.5) Utilizzando la procedura del Paragrafo 18.2.4, è facile verificare che la versione a segnali campionati del PI analogico (17.44) è

$$R^*(z) = K_P \frac{z - 1 + T/T_I}{z - 1}$$

Figura 18.11
Diagrammi di Bode di un regolatore PI analogico e digitale (metodo PZ), con $T = T_I$ e $T = 0.2T_I$. La scala delle pulsazioni è normalizzata e sul diagramma del modulo è riportato $|R^*(e^{j\omega T})| T_I / K_P$.



Questa è dunque la funzione di trasferimento del regolatore digitale progettato con il metodo HE, coincidente in questo caso con quella ricavata con il metodo di Eulero in avanti.

Il metodo PZ porterebbe invece a sintetizzare il regolatore

$$R^*(z) = \rho \frac{z - e^{-T/T_I}}{z - 1} \quad (18.28)$$

dove la costante ρ va tarata in modo da rispettare la condizione sui guadagni $\mu^* = K_P T / T_I$, e quindi vale

$$\rho = K_P \frac{T/T_I}{1 - e^{-T/T_I}}$$

I diagrammi di Bode relativi alla soluzione (18.28) sono confrontati nella Figura 18.11 con quelli del regolatore analogico per due diversi valori di T .

Esempio 18.7 (Seguito degli Esempi 11.3 e 17.6) La discretizzazione del regolatore a tempo continuo

$$R(s) = \frac{2(1+10s)}{(1+0.1s)}$$

con periodo di campionamento $T = 1$, e utilizzando i metodi di discretizzazione HE e PZ, porta alla determinazione dei regolatori

$$R^*(z) = \frac{200z - 198}{z - e^{-10}} \quad (\text{HE}) \quad (18.29)$$

$$R^*(z) = \frac{21.02z - 19.02}{z - e^{-10}} \quad (\text{PZ}) \quad (18.30)$$

Il regolatore (18.29) produce un sistema di controllo instabile, mentre il regolatore (18.30) è analogo a quello ottenuto con il metodo di Eulero all'indietro (EI), le cui prestazioni sono già state analizzate nell'Esempio 17.6.

18.4 Analisi di sistemi retroazionati a tempo discreto

Quando si vuole affrontare il problema della sintesi del regolatore digitale basandosi sulla rappresentazione a segnali campionati del sistema di controllo, è necessario fare ricorso alle tecniche di analisi dei sistemi retroazionati a tempo discreto. Questo paragrafo sarà dedicato a una sintetica presentazione di alcuni di questi strumenti, le cui analogie con quanto discusso nei Capitoli 9-12 a proposito dei sistemi a tempo continuo risulteranno evidenti.

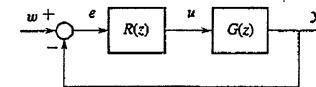


Figura 18.12
Sistema retroazionato a tempo discreto.

Per semplicità, si farà riferimento allo schema di Figura 18.12, dove $R(z)$ rappresenta il regolatore e $G(z)$ il sistema sotto controllo, che si suppone strettamente proprio. La funzione $L(z) = R(z)G(z)$ sarà chiamata *funzione di trasferimento d'anello*. In tutto il paragrafo si abbandonerà temporaneamente la convenzione finora adottata di indicare con l'asterisco i segnali e le funzioni di trasferimento, anche perché i risultati qui riportati sono applicabili in generale a un sistema puramente a tempo discreto, cioè non necessariamente derivato dalla rappresentazione a segnali campionati di un sistema di controllo digitale.

18.4.1 Stabilità

Una prima importante questione riguardante lo schema di Figura 18.12 è quella della stabilità. Da quanto visto nei Capitoli 7 e 8, è immediato osservare che, in perfetta analogia con il caso di sistemi a tempo continuo, il sistema retroazionato di Figura 18.12 è asintoticamente stabile se e solo se

- nella determinazione delle singole funzioni di trasferimento $R(z)$ e $G(z)$ non sono avvenute cancellazioni di singolarità con modulo maggiore o uguale a 1;
- nel prodotto $R(z)G(z)$ non avvengono cancellazioni tra singolarità con modulo maggiore o uguale a 1;
- tutte le radici dell'equazione caratteristica

$$1 + L(z) = 0 \quad (18.31)$$

hanno modulo minore di 1.

Per verificare l'ultima condizione, sono ovviamente disponibili tutti gli strumenti descritti nel Paragrafo 7.5, come per esempio il criterio di Jury, ma anche a tempo discreto si possono introdurre criteri legati più direttamente alle proprietà della funzione di trasferimento d'anello $L(z)$ e più facilmente utilizzabili in fase di sintesi e per l'analisi di robustezza. Il più importante di questi sarà illustrato qui di seguito.

18.4.2 Criterio di Nyquist

Si consideri il sistema retroazionato di Figura 18.13 con funzione d'anello $L(z)$ e si definisca come *diagramma di Nyquist* l'immagine attraverso $L(z)$ della circonferenza di raggio unitario e centro nell'origine del piano complesso. In altri termini, esso è costituito dai punti $L(e^{j\theta})$ con $\theta \in [-\pi, \pi]$. Se $L(z)$ ha poli su questa circonferenza, la definizione va modificata considerando delle indentature infinitesime rivolte verso l'esterno della circonferenza che consentono di aggirare tali singolarità (si veda la Figura 18.14, analoga alla Figura 8.12 del Paragrafo 8.8).

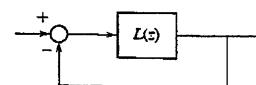
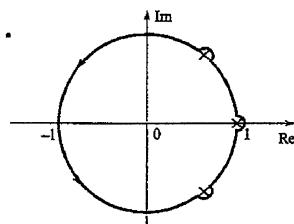


Figura 18.13
Sistema retroazionato considerato nell'analisi di stabilità.

Figura 18.14
Luogo dei punti modificato per il tracciamento del diagramma di Nyquist.



Ricordando la nozione di diagramma polare associato a una funzione di trasferimento a tempo discreto e la proprietà $L(z) = \bar{L}(z)$, è immediato concludere che il diagramma di Nyquist si ottiene dal diagramma polare di $L(z)$ completandolo con la sua immagine speculare rispetto all'asse reale. Esso risulta quindi costituito da una linea chiusa nel piano complesso, il cui orientamento è per convenzione fissato per valori crescenti di θ .

Con argomenti del tutto analoghi a quelli usati nel Paragrafo 9.5.2 per derivare il criterio di Nyquist a tempo continuo, si può pervenire a formulare il corrispondente risultato a tempo discreto. A tale scopo, si indichi con P il numero di poli di $L(z)$ con modulo maggiore di 1 e con N il numero di giri del diagramma di Nyquist di $L(z)$ attorno al punto -1 , conteggiati positivamente se compiuti in senso antiorario e negativamente se in senso orario. Se il diagramma passa per il punto -1 si dice che N non è ben definito. Sussiste allora il seguente fondamentale risultato, detto *criterio di Nyquist a tempo discreto*.

Teorema 18.2 Condizione necessaria e sufficiente perché il sistema retroazionato di Figura 18.13 sia asintoticamente stabile è che N sia ben definito e risulti $N = P$. Inoltre, quando N è ben definito e diverso da P , il sistema ha $P - N$ poli con modulo maggiore di 1. ■

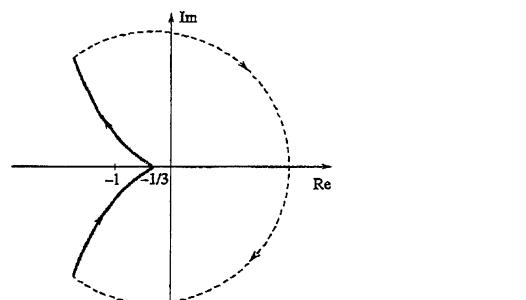
Esempio 18.8 Nel sistema di Figura 18.13 sia

$$L(z) = \frac{\rho z}{(z-1)(z-0.5)} \quad (18.32)$$

Poiché $L(z)$ non ha poli con modulo maggiore di 1, la condizione per l'asintotica stabilità del sistema retroazionato è che il diagramma di Nyquist associato a $L(z)$ non compia giri attorno al punto -1 .

Nella Figura 18.15 è riportato tale diagramma per $\rho = 1$, evidenziando anche il tratto di curva "all'infinito". Il diagramma interseca l'asse reale nel punto di ascissa $-1/3$, corrispondente

Figura 18.15
Diagramma di Nyquist associato a $L(z)$, nell'Esempio 18.8.



a $\theta = \pi$, e quindi non circonda il punto -1 . Poiché il coefficiente ρ , quando positivo e diverso da 1, produce solo una modifica del fattore di scala del diagramma, senza cambiarne la forma, si conclude che il sistema è asintoticamente stabile per $0 < \rho < 3$. Con analoghe considerazioni si deduce che non c'è stabilità per $\rho < 0$.

Per i sistemi a tempo discreto sono replicabili molte delle considerazioni svolte nel Capitolo 9 a proposito di estensioni e specializzazioni del criterio di Nyquist. In particolare, possono essere dimostrati risultati simili a quelli dei Corollari 9.1-9.4, pur di apportare le dovute ovvie varianti agli enunciati. Inoltre, il criterio di Nyquist suggerisce anche qui un modo per valutare la robustezza della stabilità di un sistema retroazionato a fronte di incertezze sulla funzione d'anello. Infatti, alla luce del Teorema 18.2, la distanza dal punto -1 del diagramma polare associato alla funzione d'anello nominale rappresenta un utile indicatore del grado di robustezza, e si possono adattare al caso in esame anche i concetti di margine di guadagno e margine di fase, ampiamente discussi nel Paragrafo 9.6 per sistemi a tempo continuo.

Ciò che rende poco frequente l'uso di tali nozioni nella pratica è la difficoltà, a cui si è già fatto cenno nel Paragrafo 8.8, di ottenere senza adeguati strumenti di calcolo una rappresentazione, anche solo approssimata, dei diagrammi di Bode a tempo discreto direttamente a partire dai poli e dagli zeri della funzione di trasferimento d'anello.

18.4.3 Luogo delle radici

Nel Capitolo 12 si è studiata la tecnica grafica del luogo delle radici per valutare la posizione nel piano complesso dei poli di un sistema retroazionato a tempo continuo al variare della costante di trasferimento d'anello. Poiché il problema del calcolo dei poli in anello chiuso per un sistema retroazionato a tempo discreto è formalmente identico, non vi è difficoltà a estendere quel tipo di analisi al contesto di cui ci si sta occupando.

Si consideri allora il sistema di Figura 18.13, dove

$$L(z) = \rho \frac{\prod_{i=1}^m (z + z_i)}{\prod_{i=1}^n (z + p_i)} = \rho \frac{N^*(z)}{D(z)}$$

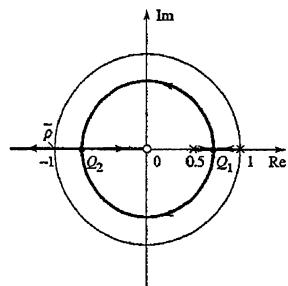
e si voglia determinare la posizione dei poli del sistema, cioè delle radici dell'equazione caratteristica (18.31) al variare del parametro reale ρ . Si chiama *luogo delle radici* l'insieme dei punti del piano complesso che sono radici dell'Equazione (18.31) per qualche valore di $\rho \neq 0$.

Vista l'identità formale tra le Equazioni (18.31) e (12.2), per il tracciamento del luogo si può ricorrere alle medesime regole enunciate nel Capitolo 12. L'unica sostanziale differenza riguarda ovviamente l'interpretazione che si può dare ai diagrammi che in questo modo si ottengono. Innanzitutto, le conclusioni sulla stabilità asintotica del sistema sono da collegare al fatto che i punti che costituiscono il luogo siano o meno interni al cerchio di raggio unitario. Inoltre, le caratteristiche dinamiche del sistema, quali il tempo di assestamento, lo smorzamento e il periodo delle eventuali oscillazioni in risposta allo scalino, vanno valutate alla luce dei risultati del Paragrafo 8.3. Si ricordi, per esempio, che poli con modulo minore di 1, ma prossimi alla circonferenza unitaria, danno luogo a una dinamica particolarmente lenta, mentre quando tutti i poli sono nell'origine le

risposte all'impulso e allo scalino si esauriscono in un numero finito di passi (si vedano i sistemi FIR, Paragrafi 7.4.6 e 8.3.5). Per inferire le proprietà del sistema dalla posizione dei poli a tempo discreto può anche essere conveniente riferirsi alla mappa della trasformazione di campionamento (Figura 18.6), immaginando che i poli provengano dal campionamento di un fitizio sistema a tempo continuo e valutando le caratteristiche dinamiche di quest'ultimo in termini di pulsazione naturale ω_n e di smorzamento ξ .

Esempio 18.9 (Seguito dell'Esempio 18.8) Per analizzare la stabilità del sistema retroazionato con funzione d'anello (18.32), si tracci il luogo delle radici. Nella Figura 18.16 è riportato il luogo diretto, costituito da due rami che, partendo dai poli, si incrociano nel punto Q_1 , diventano complessi, tornano a intersecarsi in Q_2 per poi dirigersi l'uno a $-\infty$ e l'altro verso lo zero nell'origine. Mediante la Regola 10 del Paragrafo 12.2.2 si trova che le ascisse di Q_1 e Q_2 sono rispettivamente $\pm\sqrt{0.5} \approx \pm 0.7$. Inoltre, il tracciamento preciso del luogo rivela che nel tratto tra Q_1 e Q_2 i rami descrivono due semicirconferenze e quindi non escono dal cerchio unitario.

Figura 18.16
Luogo delle radici dirette
nell'Esempio 18.9.



Sulla base della Figura 18.16, si deduce che il sistema retroazionato è asintoticamente stabile per $0 < \rho < \bar{\rho}$, dove $\bar{\rho}$ corrisponde al punto -1 in cui uno dei rami attraversa la circonferenza unitaria. Grazie alla regola sulla punteggiatura si ricava $\bar{\rho} = 3$ e si trova conferma del risultato dell'Esempio 18.8. Si osserva poi immediatamente che il sistema è instabile per valori negativi di ρ , dato che uno dei rami del luogo inverso è costituito dalla semiretta $(1, \infty)$.

Quanto alle caratteristiche dinamiche del sistema al variare di $\rho > 0$, si può osservare che il modulo del polo dominante non è mai inferiore a 0.7 e pertanto il tempo di assennamento al 99% finché il sistema è asintoticamente stabile, è sempre superiore a $-5/\ln(0.7) \approx 14$ (Tabella 8.1). Inoltre, per piccoli valori di ρ il sistema non presenta dinamica oscillante poiché i poli sono entrambi reali e positivi. Poi, al crescere di ρ oltre il valore $\rho_1 = 1.5 - \sqrt{2} \approx 0.86$, corrispondente al punto Q_1 , lo smorzamento diminuisce poiché i rami del luogo intersecano regioni a smorzamento via via minore.

Attraverso il luogo delle radici è facile ricavare un interessante risultato che riguarda la stabilità del sistema retroazionato di Figura 18.13 quando il valore assoluto della costante di trasferimento d'anello ρ (o equivalentemente del guadagno d'anello) tende all'infinito. Poiché infatti si è ipotizzato che $L(z)$ sia strettamente propria, e quindi $n > m$, sia il luogo diretto sia quello inverso possiedono almeno un ramo che tende all'infinito. Ciò significa che per valori di $|\rho|$ sufficientemente grandi esiste almeno un polo in anello chiuso che giace all'esterno del cerchio unitario, provocando così l'instabilità del sistema retroazionato. In altri termini, non esistono sistemi retroazionati a tempo discreto con "margine di guadagno infinito".

18.4.4 Funzioni di sensitività

A partire dallo schema di Figura 18.12 si possono ricavare le seguenti funzioni di trasferimento in anello chiuso

$$F(z) = \frac{Y(z)}{W(z)} = \frac{L(z)}{1 + L(z)} \quad (18.33)$$

$$S(z) = \frac{E(z)}{W(z)} = \frac{1}{1 + L(z)} \quad (18.34)$$

$$Q(z) = \frac{U(z)}{W(z)} = \frac{R(z)}{1 + L(z)} \quad (18.35)$$

che prendono il nome rispettivamente di *sensibilità complementare*, *sensibilità*, e *sensibilità del controllo*.

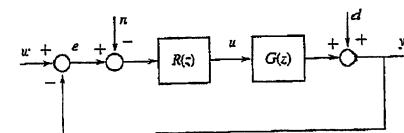


Figura 18.17
Sistema retroazionato
a tempo discreto con
disturbi.

Queste stesse funzioni di trasferimento, eventualmente cambiate di segno, entrano in gioco nel descrivere l'effetto su y , u ed e di eventuali disturbi sulla linea di andata o sulla linea di retroazione. A tale riguardo si consideri lo schema più generale di Figura 18.17, nel quale i disturbi a tempo discreto d e n hanno lo stesso ruolo delle variabili con lo stesso nome nello schema a tempo continuo di Figura 9.7. È immediato osservare che risulta

$$\begin{bmatrix} Y(z) \\ U(z) \\ E(z) \end{bmatrix} = \begin{bmatrix} F(z) & S(z) & -F(z) \\ Q(z) & -Q(z) & -Q(z) \\ S(z) & -S(z) & F(z) \end{bmatrix} \begin{bmatrix} W(z) \\ D(z) \\ N(z) \end{bmatrix}$$

in perfetta analogia con l'Equazione (10.4).

L'analisi statica e dinamica del sistema retroazionato è pertanto riconducibile allo studio delle caratteristiche delle tre funzioni $F(z)$, $S(z)$ e $Q(z)$. In particolare, in assenza del disturbo n , sarebbe auspicabile ottenere

$$F(z) = 1, \quad S(z) = 0 \quad (18.36)$$

in modo da garantire la perfetta identità tra la variabile controllata y e il riferimento w e il completo annullamento dell'effetto del disturbo d sull'uscita. Peraltra, l'eventuale presenza di n indurrebbe a prestare una maggiore cautela nel perseguire l'obiettivo (18.36). Allo stesso modo, gran parte delle considerazioni svolte nel Paragrafo 10.2 per sistemi a tempo continuo sono applicabili al caso in esame. Per esempio, vale comunque la relazione

$$F(z) + S(z) = 1$$

che esprime un vincolo intrinseco sulla possibilità di assegnare indipendentemente le funzioni $F(z)$ e $S(z)$ mediante un'opportuna scelta del regolatore $R(z)$.

Una sostanziale differenza rispetto al caso a tempo continuo deriva invece dal fatto che, anche sotto l'ipotesi che $L(z)$ sia strettamente propria, può risultare $|L(e^{j\theta})| \gg 1, \forall \theta$, così da consentire che le risposte in frequenza $F(e^{j\theta})$ e $S(e^{j\theta})$ siano arbitrariamente prossime alle soluzioni ideali corrispondenti alla (18.36).

Nei sottoparagrafi successivi ci si limiterà a presentare sinteticamente i principali risultati deducibili dallo studio delle funzioni di sensitività.

18.4.5 Analisi statica

Uno dei principali requisiti di un sistema di controllo è la sua precisione a regime. Prendendo in esame il sistema di Figura 18.17, di cui si assumerà senz'altro la stabilità asintotica, è per esempio interessante calcolare il valore e_{∞} dell'errore a transitorio esaurito quando il riferimento è $w(k) = A \text{ sca}^*(k)$. Utilizzando il teorema del valore finale (si veda l'Equazione (C.6)), ricordando che la funzione di trasferimento tra w ed e è $S(z)$, mentre la trasformata dello scalino unitario è $z/(z - 1)$, si ottiene

$$e_{\infty} = \lim_{k \rightarrow \infty} e(k) = \lim_{z \rightarrow 1} (z - 1) S(z) \frac{Az}{z - 1} = A \lim_{z \rightarrow 1} S(z)$$

Tale limite dipende dal tipo g della funzione d'anello. Risulta infatti

$$e_{\infty} = \begin{cases} A & , g < 0 \\ \frac{A}{1 + \mu} & , g = 0 \\ 0 & , g > 0 \end{cases} \quad (18.37)$$

dove μ è il guadagno di $L(z)$, o *guadagno d'anello*. L'Equazione (18.37) mette in evidenza che l'errore a transitorio esaurito può essere reso piccolo, quando $g = 0$, aumentando il guadagno d'anello e si annulla per $g > 0$, ovvero quando la funzione d'anello possiede poli in $z = 1$. In quest'ultimo caso il guadagno $\mu_F = F(1)$ della funzione di sensitività complementare è uguale a 1. Non dovrebbe sfuggire la perfetta analogia di tale conclusione con quella presentata nel Capitolo 10 per i sistemi a tempo continuo.

Si potrebbero estendere allo stesso modo i risultati sull'analisi statica in presenza di altri segnali canonici (quali rampa o parabola), nonché lo studio della risposta a regime del sistema ai disturbi d e n dello schema di Figura 18.17.

18.4.6 Analisi dinamica

Le caratteristiche dinamiche del sistema di Figura 18.17 dipendono essenzialmente dalla posizione dei poli in anello chiuso. Già si è visto nel Paragrafo 18.4.2 come sia possibile verificare che essi siano tutti interni alla regione di asintotica stabilità mediante il criterio di Nyquist, cioè con uno studio effettuato sulla funzione d'anello $L(z)$. Inoltre, il tracciamento del luogo delle radici permette di avere indicazioni più dettagliate sulla posizione dei poli. Quanto agli zeri delle funzioni di sensitività $F(z)$, $S(z)$ e $Q(z)$, valgono considerazioni simili a quelle sviluppate nel Capitolo 10 per le corrispondenti funzioni a tempo continuo.

Esiste però anche la possibilità di valutare le prestazioni del sistema retroazionato attraverso lo studio delle risposte in frequenza associate a tali funzioni di sensitività. Per esempio, dall'Equazione (18.33) si ricava

$$|F(e^{j\theta})| = \frac{|L(e^{j\theta})|}{|1 + L(e^{j\theta})|} \simeq \begin{cases} 1 & , |L(e^{j\theta})| \geq 1 \\ |L(e^{j\theta})| & , |L(e^{j\theta})| < 1 \end{cases} \quad (18.38)$$

Si deduce che, per avvicinarsi a un filtro ideale "passa-tutto" tra w e y , conviene fare in modo che sia $|L(e^{j\theta})| \gg 1$, $\forall \theta$. Poiché tale esigenza è di solito in contrasto con quella di stabilità (è facile infatti che tale condizione implichi che il diagramma di Nyquist di $L(z)$ circondi il punto -1) ci si accontenta spesso che risulti $|L(e^{j\theta})| > 1$ in un opportuno intervallo $[0, \theta_c]$, con $\theta_c < \pi$. Si è così

in grado di assicurare, prescindendo da eventuali distorsioni di fase, una fedele riproduzione in uscita delle componenti a pulsazione (discreta) minore di θ_c del segnale di riferimento w .

Analogamente, dall'Equazione (18.34) si può ricavare

$$|S(e^{j\theta})| = \frac{1}{|1 + L(e^{j\theta})|} \simeq \begin{cases} \frac{1}{|L(e^{j\theta})|} & , |L(e^{j\theta})| \geq 1 \\ 1 & , |L(e^{j\theta})| < 1 \end{cases} \quad (18.39)$$

Ricordando che la funzione di sensitività $S(z)$ rappresenta la relazione tra il disturbo d e l'uscita y , si conclude che il sistema retroazionato provoca una riduzione dell'effetto del disturbo nella banda di pulsazioni (discrete) in cui è $|L(e^{j\theta})| > 1$, e l'attenuazione è tanto più consistente quanto maggiore è $|L(e^{j\theta})|$.

Anche la funzione di sensitività del controllo $Q(z)$ dell'Equazione (18.35) si presta a essere analizzata da questo punto di vista, pervenendo all'approssimazione

$$|Q(e^{j\theta})| = \frac{|R(e^{j\theta})|}{|1 + L(e^{j\theta})|} \simeq \begin{cases} \frac{1}{|G(e^{j\theta})|} & , |L(e^{j\theta})| \geq 1 \\ |R(e^{j\theta})| & , |L(e^{j\theta})| < 1 \end{cases} \quad (18.40)$$

che può essere utilizzata per discutere la proprietà di moderazione della variabile di controllo.

18.5 Sintesi diretta a tempo discreto

Come già accennato, le tecniche di progetto a tempo discreto di un regolatore digitale sono quelle che fanno riferimento alla rappresentazione di Figura 18.18, dove $G^*(z)$ è la funzione di trasferimento del sistema a segnali campionati costituito dalla connessione in serie del mantenitore, del processo sotto controllo e del campionatore. Nello schema sono stati messi in evidenza anche i segnali di ingresso d^* e n^* , rappresentativi di eventuali disturbi sulla linea di andata (disturbi additivi sull'uscita) o su quella di retroazione (disturbi di misura). Il problema di sintesi consiste nel determinare $R^*(z)$ in modo che il sistema retroazionato a tempo discreto possieda determinate proprietà.

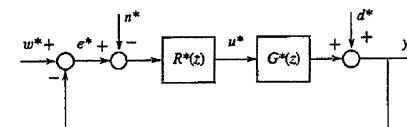
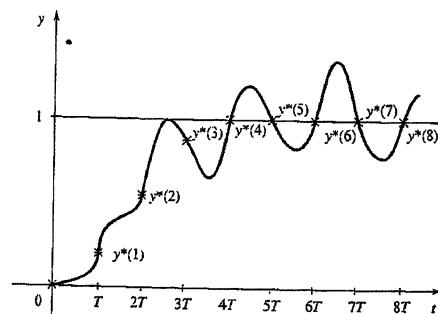


Figura 18.18
Schema equivalente a tempo discreto di un sistema di controllo digitale.

Va in ogni caso sottolineato che lo schema di Figura 18.18 è una rappresentazione parziale del sistema di controllo digitale originario, poiché descrive l'evoluzione del sistema solo negli istanti di campionamento. Come esemplificato nella Figura 18.19, andamenti soddisfacenti dell'uscita a tempo discreto possono pertanto coesistere con comportamenti sgraditi o addirittura inaccettabili dell'effettiva variabile a tempo continuo che si desidera controllare. Per prevenire fenomeni di questo genere, che prendono il nome di *oscillazioni nascoste*, oltre a operare una scelta adeguata del periodo di campionamento secondo i criteri discussi nel Paragrafo 17.6, è conveniente adottare opportune precauzioni di cui si parlerà fra poco.

Figura 18.19
Esempio di oscillazioni nascoste.



In questo paragrafo, dopo avere brevemente illustrato il modo in cui si possono formulare i principali requisiti del sistema di controllo, saranno presentate alcune semplici tecniche di progetto a tempo discreto. La trattazione esauriente di questo argomento esula dagli scopi del presente testo e può essere trovata dal lettore sui riferimenti bibliografici dedicati al controllo digitale.

18.5.1 Requisiti e specifiche

Limitandosi per semplicità al caso di prestazioni nominali (in assenza cioè di incertezze sul modello $G^*(z)$ del sistema), si riassumono qui i requisiti cui deve soddisfare di norma il sistema di Figura 18.18.

Stabilità Per assicurare l'asintotica stabilità del sistema di controllo è necessario evitare cancellazioni tra singolarità di $R^*(z)$ e $G^*(z)$ non interne al cerchio di raggio unitario. Inoltre, i poli in anello chiuso devono avere tutti modulo minore di 1. Per effettuare quest'ultima verifica sono disponibili gli strumenti di analisi presentati nel Paragrafo 18.4, quali il criterio di Nyquist e il luogo delle radici.

Precisione statica Tale requisito si riferisce alla capacità del sistema di controllo di mantenere limitato (o addirittura nullo) l'errore a transitorio esaurito in presenza di ingressi canonici. Per esempio, una tipica specifica riguarda il massimo valore tollerabile di $|e_\infty^*|$ quando il riferimento w^* è uno scalino. Dall'analisi svolta nel Paragrafo 18.4.5, si sa che il rispetto di vincoli di questo genere dipende dal tipo della funzione d'anello (cioè dal numero di poli in $z = 1$ di $L^*(z) = R^*(z)G^*(z)$) e dal guadagno d'anello. Per esempio, la presenza di un polo in $z = 1$ garantisce l'annullamento a regime di e_∞^* quando $w^*(k) = A \text{ sca}^*(k)$, con A arbitrario.

Precisione dinamica A volte le specifiche dinamiche sono formulate come vincoli sulle risposte in frequenza associate alle funzioni di sensitività, che, come discusso nel Paragrafo 18.4.6, si traducono in prescrizioni sull'andamento di $L^*(e^{j\theta})$. Più spesso gli obiettivi del controllo vengono espressi in termini di limitazioni sulla posizione dei poli in anello chiuso, miranti a impostare determinati comportamenti in transitorio della variabile di uscita. A tale scopo, è utile fare riferimento alla corrispondenza tra piano S e piano Z indotta dalla trasformazione di campionamento $z = e^{jT}$ (Paragrafo 18.2.3). Un vincolo sul tempo di assestamento del tipo $T_{a1} \leq \bar{T}_{a1}$, che a tempo continuo si tradurrebbe approssimativamente nella richiesta che la parte reale dei poli sia inferiore a $-\bar{\sigma} = -5/\bar{T}_{a1}$ (si vedano le Tabelle 4.1-4.3), si trasforma nella condizione che i poli a tempo

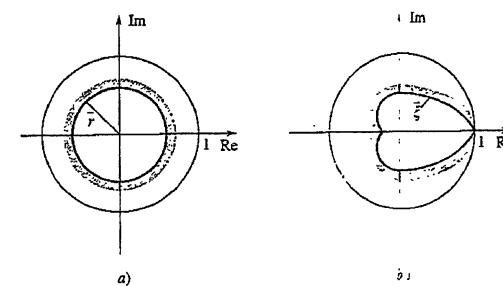


Figura 18.20
Regioni di ammissibilità per i poli in anello chiuso:
a) vincolo sul tempo di assestamento; b) vincolo sullo smorzamento.

discreto abbiano modulo minore di $\bar{r} = e^{-5T/\bar{T}_{a1}}$ (Figura 18.20a). Analogamente, per rispettare una specifica sul minimo smorzamento accettabile ξ , è necessario che i poli a tempo discreto giacciono all'interno della regione indicata nella Figura 18.20b.

Attenuazione dell'effetto del disturbo d^* Alla luce della (18.39), si può garantire un'efficace reiezione del disturbo se si fa in modo che risulti $|L^*(e^{j\theta})| \gg 1$ in corrispondenza delle pulsazioni alle quali lo spettro del segnale d^* presenta componenti significative.

Attenuazione dell'effetto del disturbo n^* Si supponga di conoscere l'intervallo di pulsazioni in cui è concentrato lo spettro di n^* . Poiché la funzione di trasferimento tra n^* e y^* è data da $F^*(z)$, occorre che in quell'intervallo $|L^*(e^{j\theta})|$ non sia troppo elevato (si ricordi la (18.38)).

Moderazione della variabile di controllo Per evitare inutili sollecitazioni di u^* , e quindi dell'effettiva variabile di controllo u all'uscita del mantenitore, è bene evitare, in base alla (18.40), che $|R^*(e^{j\theta})|$ risulti troppo elevato in corrispondenza dei valori di θ per cui $|L^*(e^{j\theta})| < 1$. Si è inoltre visto nel Paragrafo 10.5 che la pretesa di ottenere un sistema retroazionato la cui dinamica è "molto più veloce" di quella propria del sistema sotto controllo si scontra quasi sempre con il requisito di mantenere limitati i valori assunti dalla variabile di controllo. Ciò induce, quindi, ad adottare anche nel caso di un sistema di controllo digitale una certa cautela nella scelta dei valori dei poli in anello chiuso, almeno di quelli dominanti.

Assenza di oscillazioni nascoste Fenomeni indesiderati come quello illustrato nella Figura 18.19 sono di solito causati dalla presenza di poli poco smorzati del regolatore $R^*(z)$, poli cioè che giacciono all'esterno di regioni come quella di Figura 18.20b. Tali poli possono dar luogo a una dinamica oscillante della variabile di controllo, che a sua volta produce oscillazioni persistenti di y . Un buon criterio è dunque quello di evitare, per quanto possibile, una circostanza di questo tipo. Per analoghi motivi, è bene che poli a basso smorzamento del processo non vengano cancellati da zeri del regolatore.

Realizzabilità del regolatore È ovviamente necessario che la procedura di sintesi porti a determinare una funzione di trasferimento $R^*(z)$ del regolatore digitale che rappresenti un sistema dinamico proprio, tale cioè che il grado del numeratore non superi il grado del denominatore. Se si opta per l'uso di un regolatore proprio, ma non strettamente, si ricordi che è necessario tener conto esplicitamente del ritardo computazionale τ_c nel calcolo della funzione $G^*(z)$ del sistema sotto controllo (si vedano i Paragrafi 17.2.4 e 18.2.2).

18.5.2 Sintesi mediante il luogo delle radici

Le tecniche di sintesi del regolatore digitale basate sul luogo delle radici sono particolarmente indicate quando si è in grado di esprimere le specifiche di progetto sotto forma di vincoli sulla posizione dei poli in anello chiuso. La scelta di $R^*(z)$ viene di solito effettuata attraverso successivi tentativi, in cui si arricchisce via via la struttura del regolatore, fino a soddisfare tutti i requisiti. Si mostrerà ora un esempio di questa procedura.

Esempio 18.10 Si voglia progettare un regolatore digitale per il sistema con funzione di trasferimento

$$G(s) = \frac{0.2}{(1+s)(1+0.2s)}$$

adottando il periodo di campionamento $T = 0.1$ e cercando di rispettare i seguenti requisiti relativi alla risposta a uno scalino del riferimento:

- a) errore a transitorio esaurito nullo;
- b) tempo di assestamento al 99% minore di $\bar{T}_{a1} = 2.5$;
- c) massima sovralungazione inferiore al 25%.

Si calcoli innanzitutto con uno dei metodi del Paragrafo 18.2 la funzione di trasferimento $G^*(z)$ del sistema a segnali campionati:

$$G^*(z) = \frac{0.0041(z + 0.819)}{(z - 0.905)(z - 0.607)} \quad (18.41)$$

La specifica (a) richiede che il regolatore $R^*(z)$ possieda almeno un polo in $z = 1$. Il vincolo (b) è equivalente a impostare un tempo (discreto) di assestamento inferiore a $\bar{k} = 25$ passi di campionamento. A sua volta, ciò si traduce nell'imporre che i poli del sistema retroazionato abbiano modulo minore di $\bar{r} = e^{-5/\bar{k}} \simeq 0.819$. Il requisito (c), infine, restinge ulteriormente la regione ammissibile per i poli in anello chiuso, come mostrato nella Figura 18.20b, poiché, in base alla Figura 4.11, equivale a richiedere uno smorzamento superiore a $\bar{\xi} = 0.4$.

Si analizzi dapprima la possibilità di risolvere il problema di sintesi con il regolatore

$$R^*(z) = \frac{\rho_R}{z - 1} \quad (18.42)$$

con un'opportuna scelta di ρ_R . Dall'esame del luogo delle radici diretto di $L^*(z) = R^*(z)G^*(z)$ (Figura 18.21) si nota che per valori positivi elevati di ρ_R il sistema diventa instabile e, comunque, per $\rho_R > 0$ due dei poli in anello chiuso sono sempre esterni alla regione ammissibile. D'altra parte, un ramo del luogo inverso è costituito dalla semiretta $(1, \infty)$, mostrando così che anche la soluzione $\rho_R < 0$ non è accettabile.

Come secondo tentativo si ponga

$$R^*(z) = \frac{\rho_R(z - 0.905)}{z - 1} \quad (18.43)$$

Figura 18.21
Luogo delle radici con il regolatore (18.42) nell'Esempio 18.10.

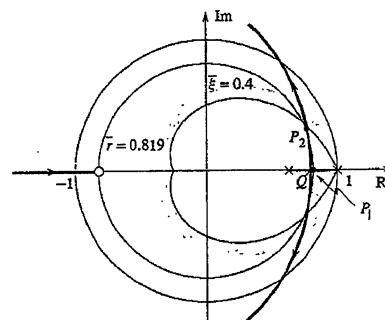
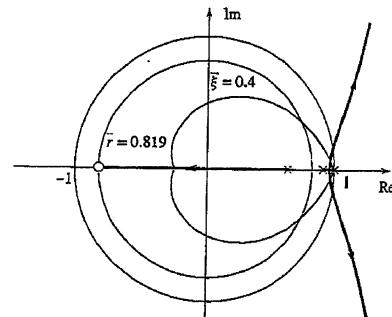


Figura 18.22
Luogo delle radici con il regolatore (18.43) nell'Esempio 18.10.

in modo da cancellare il polo "più lento" di $G^*(z)$. La funzione d'anello vale

$$L^*(z) = \frac{\rho(z + 0.819)}{(z - 0.607)(z - 1)}$$

con $\rho = 0.0041\rho_R$. Il corrispondente luogo delle radici, riportato nella Figura 18.22, mostra che entrambi i poli in anello chiuso sono ammissibili quando il valore di ρ è prossimo a quello corrispondente al punto Q di diramazione, la cui ascissa è $Q \simeq 0.79$ (si veda la Regola 10 del Paragrafo 12.2.2). Più precisamente, i valori accettabili di ρ sono quelli compresi tra $\rho_1 \simeq 0.023$ e $\rho_2 \simeq 0.08$ corrispondenti ai punti P_1 (quasi coincidente con Q) e P_2 della Figura 18.22. Pertanto, il progetto è concluso scegliendo il regolatore (18.43) con ρ_R nell'intervallo $[5.6, 19.5]$. La Figura 18.23, che mostra la risposta del sistema di controllo con $\rho_R = 14$ a uno scalino del riferimento, evidenzia che le specifiche di progetto sono effettivamente soddisfatte.

Si noti infine che la funzione di trasferimento (18.43) del regolatore non è strettamente propria. Alla luce di quanto discusso nei Paragrafi 17.2.4 e 18.5.1, bisogna quindi assicurarsi che il ritardo computazionale τ_c sia trascurabile rispetto al periodo di campionamento $T = 0.1$.

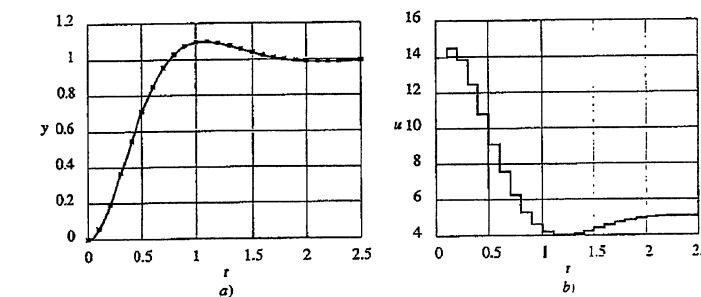


Figura 18.23
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 18.10: a) uscita y ; b) variabile di controllo u .

18.5.3 Assegnamento del modello

Si consideri il sistema retroazionato di Figura 18.18 e si immagini di essere riusciti a formalizzare tutte le specifiche di progetto attraverso l'individuazione di una funzione di sensitività complementare $F^*(z)$ ottimale. L'idea che sta alla base del cosiddetto metodo ad *assegnamento del modello*, noto anche come *metodo di Ragazzini*, è quella di ricavare la funzione di trasferimento incognita $R^*(z)$ del regolatore a partire dalla conoscenza di $G^*(z)$, ottenuta dal modello del processo sotto controllo, e di $F^*(z)$, fissata *a priori* dal progettista.

Ricordando l'Equazione (18.33), si ha

$$F^*(z) = \frac{R^*(z)G^*(z)}{1 + R^*(z)G^*(z)}$$

e, risolvendo quest'ultima espressione rispetto a $R^*(z)$, si ottiene

$$R^*(z) = \frac{F^*(z)}{G^*(z)[1 - F^*(z)]} \quad (18.44)$$

che può essere interpretata come la formula risolutiva del metodo. Perché questa formula sia realmente applicabile, è necessario che la funzione $R^*(z)$ risultante goda di opportune proprietà connesse alla realizzabilità del regolatore e alla stabilità asintotica del sistema complessivo. Si deve quindi corredare la (18.44) di adeguate precisazioni sui vincoli cui deve soddisfare la scelta iniziale di $F^*(z)$. Tali precisazioni costituiscono anzi, in certo senso, il vero nocciolo della tecnica di progetto adottata.

Per sviluppare lo studio, si rappresentino $G^*(z)$, $R^*(z)$ e $F^*(z)$ in forma fattorizzata come

$$G^*(z) = \frac{N(z)}{D(z)} \quad , \quad R^*(z) = \frac{Q(z)}{P(z)} \quad , \quad F^*(z) = \frac{B(z)}{A(z)} \quad (18.45)$$

dove si assume che i polinomi al numeratore e al denominatore delle singole funzioni siano primi tra loro. Si indichino inoltre con n_N , n_D , n_Q , n_P , n_B , e n_A i gradi dei rispettivi polinomi che compaiono in (18.45), supponendo anche che sia $n_B < n_A$, e cioè che $F^*(z)$ sia strettamente propria.

Grazie alla (18.45), la formula risolutiva (18.44) diventa

$$R^*(z) = \frac{Q(z)}{P(z)} = \frac{D(z)B(z)}{N(z)[A(z) - B(z)]} \quad (18.46)$$

e i parametri di progetto sono i polinomi $A(z)$ e $B(z)$. Va sottolineato che i polinomi $Q(z)$ e $P(z)$ sono quelli che si ottengono dopo aver semplificato eventuali fattori comuni tra il numeratore $D(z)B(z)$ e il denominatore $N(z)[A(z) - B(z)]$.

Realizzabilità Per rispettare il vincolo che $R^*(z)$ corrisponda a un sistema dinamico proprio, deve essere $n_Q \leq n_P$. D'altra parte, dalla (18.46) e dall'ipotesi $n_B < n_A$, risulta

$$n_P - n_Q = (n_N + n_A) - (n_D + n_B)$$

La condizione che assicura la realizzabilità può essere allora scritta come

$$n_A - n_B \geq n_D - n_N \quad (18.47)$$

In altri termini, il grado relativo di $F^*(z)$ non può essere inferiore a quello di $G^*(z)$. Tale vincolo ammette una semplice interpretazione intuitiva se si ricorda che in un sistema a tempo discreto il grado relativo rappresenta il tempo di latenza tra l'ingresso e l'uscita (Paragrafo 8.3.1). La condizione (18.47) esprime dunque la naturale impossibilità di ridurre mediante la retroazione il tempo di latenza presente nel sistema.

Se si desidera progettare un regolatore strettamente proprio, basta scegliere i gradi n_A e n_B in modo che la diseguaglianza (18.47) sia rispettata con il segno di maggiore.

Stabilità Il sistema di controllo è asintoticamente stabile se e solo se le radici di $A(z)$ hanno modulo minore di 1 e nello schema di Figura 18.18 non avvengono cancellazioni tra singolarità di $R^*(z)$ e di $G^*(z)$ collocate al di fuori della regione di stabilità. Mentre la prima condizione può essere facilmente soddisfatta con un'oculata scelta di $A(z)$, alla luce dell'Equazione (18.46) per evitare cancellazioni illecite devono essere verificate le seguenti condizioni.

- Ogni zero z_i di $G^*(z)$ con $|z_i| \geq 1$ deve comparire tra le radici del polinomio $B(z)$, ovvero

$$B(z_i) = 0 \quad (18.48)$$

Infatti, uno zero di $G^*(z)$ (cioè una radice di $N(z)$) non appare come polo di $R^*(z)$ (cioè radice di $P(z)$) solo quando esso è anche una radice di $B(z)$, poiché solo in tal caso esso scompare nella semplificazione del secondo membro della (18.46).

- Ogni polo p_i di $G^*(z)$ con $|p_i| \geq 1$ deve essere anche radice del polinomio $A(z) - B(z)$, e cioè

$$A(p_i) = B(p_i) \quad (18.49)$$

Infatti, un polo di $G^*(z)$ (cioè una radice di $D(z)$) non appare come zero di $R^*(z)$ (cioè radice di $Q(z)$) solo quando esso è anche una radice del polinomio $A(z) - B(z)$.

In definitiva, le relazioni (18.48) e (18.49) introducono limitazioni aggiuntive sulla scelta dei polinomi di progetto $A(z)$ e $B(z)$.

Precisione statica Se si vuole ottenere un errore nullo a transitorio esaurito quando il riferimento è uno scalino, bisogna scegliere $F^*(z)$ in modo che $F^*(1) = 1$, ovvero

$$A(1) = B(1) \quad (18.50)$$

Condizioni analoghe, anche se più complicate, nascono pure nel caso di specifiche statiche riguardanti altri ingressi canonici.

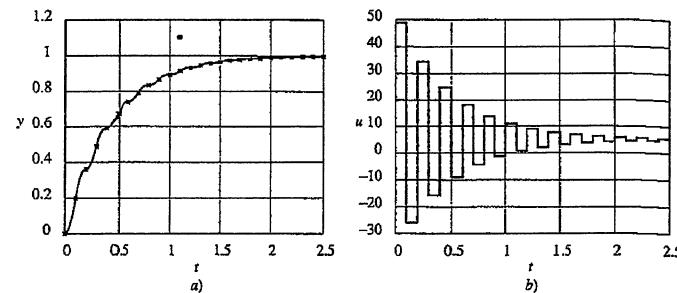
Precisione dinamica Al di là del rispetto delle condizioni (18.47)-(18.50), il progettista ha completa libertà nella scelta della funzione $F^*(z)$. Normalmente, si opera in modo da collocare poli e zeri in anello chiuso in posizioni del piano complesso favorevoli per quanto concerne le caratteristiche dinamiche del sistema retroazionato. Se, per esempio, si devono rispettare requisiti sul tempo di assestamento e sull'ampiezza delle sovraelongazioni, la regione del piano Z in cui posizionare le radici del polinomio $A(z)$ viene costruita come intersezione di luoghi geometrici come quelli della Figura 18.20. Considerazioni sulla modulazione del controllo suggeriscono poi di evitare di collocare i poli dominanti di $F^*(z)$ a una distanza dall'origine molto minore rispetto a quella dei poli in anello aperto (cioè delle radici di $D(z)$).

Purtroppo, però, le conseguenze di scelte di questo genere, in combinazione con quelle relative al polinomio $B(z)$, su altre proprietà del sistema di controllo, come per esempio la robustezza e l'inibizione di eventuali oscillazioni nascoste, sono spesso difficilmente prevedibili.

Esempio 18.11 (Seguito dell'Esempio 18.10) Si voglia risolvere il problema di progetto utilizzando il metodo ad assegnamento del modello. Dall'Equazione (18.41) risulta

$$\begin{aligned} N(z) &= 0.0041(z + 0.819) \\ D(z) &= (z - 0.905)(z - 0.607) \end{aligned}$$

Figura 18.24
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 18.11 con il regolatore (18.51): a) uscita y ; b) variabile di controllo u .



con $n_D - n_N = 1$. Inoltre, $G^*(z)$ non possiede alcun polo o zero con modulo maggiore o uguale a 1. Pertanto, tutte le singolarità sono potenzialmente cancellabili senza rischi per la stabilità. Tenendo conto della condizione (18.47), del vincolo (18.50), necessario per rispettare il requisito statico, e della specifica sul tempo di assestamento, una possibile scelta per $F^*(z)$ è

$$F^*(z) = \frac{B(z)}{A(z)} = \frac{0.2}{z - 0.8}$$

Il corrispondente regolatore, ricavato con la formula (18.46), è

$$R^*(z) = \frac{48.8(z - 0.905)(z - 0.607)}{(z + 0.819)(z - 1)} \quad (18.51)$$

che, quando il riferimento è un scalino, produce la risposta di Figura 18.24. Il comportamento poco soddisfacente del transitorio, con ampie oscillazioni della variabile di controllo u a dispetto di un andamento più che accettabile dell'uscita y negli istanti di campionamento, è dovuto all'esistenza nella funzione (18.51) del regolatore di un polo a basso smorzamento in $z = -0.819$ (si veda il commento sulle oscillazioni nascoste nel Paragrafo 18.5.1).

Per evitare tale fenomeno, è sufficiente impedire che lo zero $z_1 = -0.819$ di $G^*(z)$ risultì cancellato da un polo del regolatore, imponendo che sia $B(z_1) = 0$, come è stato fatto nella (18.48) nei confronti degli eventuali zeri non interni al cerchio unitario. Ciò richiede di aumentare di uno il grado dei polinomi $A(z)$ e $B(z)$. Si ponga allora, per esempio,

$$F^*(z) = \frac{B(z)}{A(z)} = \frac{\rho(z + 0.819)}{(z - 0.8)^2}$$

avendo deciso di collocare anche il secondo polo in anello chiuso in $z = 0.8$. Per soddisfare la condizione (18.50), deve essere $\rho = 0.022$. A seguito di questa nuova scelta di $F^*(z)$, dalla (18.46) si ricava il regolatore

$$R^*(z) = \frac{5.37(z - 0.905)(z - 0.607)}{(z - 0.622)(z - 1)} \quad (18.52)$$

le cui prestazioni, decisamente migliori rispetto al progetto precedente, possono essere valutate osservando la Figura 18.25.

Figura 18.25
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 18.11 con il regolatore (18.52): a) uscita y ; b) variabile di controllo u .

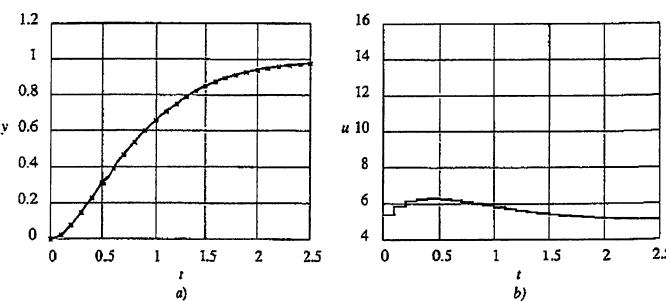
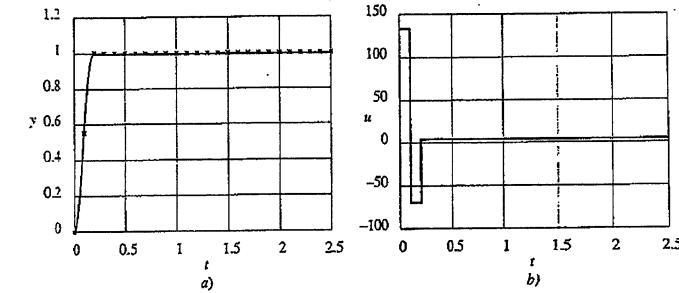


Figura 18.26
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 18.11 con il regolatore (18.53): a) uscita y ; b) variabile di controllo u .



Per evidenziare l'opportunità della scelta adottata di collocare il polo dominante di $F^*(z)$ in $z = 0.8$ e non più vicino all'origine, si ripeta il progetto con

$$F^*(z) = \frac{1}{1.819} \frac{z + 0.819}{z^2}$$

ottenendo

$$R^*(z) = \frac{134(z - 0.905)(z - 0.607)}{(z + 0.45)(z - 1)} \quad (18.53)$$

Apparentemente tale scelta è vanaggiosa poiché corrisponde a un sistema di controllo di tipo FIR (si vedano i Paragrafi 7.4.6 e 8.3.5) che presenta risposte all'impulso o allo scalino che si esauriscono in tempo finito. I transitori della risposta allo scalino riportati nella Figura 18.26 mostrano però chiaramente i difetti di tale soluzione in termini di moderazione della variabile di controllo.

18.5.4 Assegnamento dei poli

Come si è visto, il metodo ad assegnamento del modello, a causa della presenza del fattore $G^*(z)$ al denominatore della formula (18.46), induce a progettare un regolatore che cancella poli e zeri del sistema sotto controllo, a meno di impedirlo esplicitamente con vincoli del tipo (18.48) per gli zeri e (18.49) per i poli. Le cancellazioni dei poli di $G^*(z)$ con zeri di $R^*(z)$ sono a volte indesiderate per una serie di motivi. Innanzitutto, ogni cancellazione contribuisce a creare una parte "nascosta" (si vedano i Paragrafi 4.2.4, 4.2.5 e 5.5) e il sistema di controllo è incapace di neutralizzare qualunque disturbo che agisca su tale parte. Inoltre, l'incertezza che inevitabilmente affligge il modello del processo da controllare fa sì che in realtà queste cancellazioni non siano mai esatte, e la non perfetta semplificazione polo-zero possa dar luogo a spiacevoli fenomeni (si veda a tempo continuo l'effetto di una "quasi cancellazione" a bassa frequenza nella Figura 4.7).

Ciò motiva l'interesse per diverse tecniche di sintesi, come per esempio quella basata sul luogo delle radici, in cui l'obiettivo è quello di "spostare", piuttosto che di "cancellare", i poli del sistema. Tra queste tecniche è di particolare importanza il cosiddetto metodo ad *assegnamento degli autovalori, o dei poli*, già presentato nel Capitolo 13 e basato sull'idea di determinare il regolatore in modo che gli autovalori (i poli) del sistema in anello chiuso assumano valori prefissati. A questo problema è stata data una soluzione nel Capitolo 13 con tecniche di progetto nel dominio del tempo, che potrebbero facilmente essere estese al caso di sistemi a tempo discreto. Nel seguito sarà invece proposto un diverso metodo che consente, rimanendo nel dominio della variabile complessa Zeta, di determinare la funzione di trasferimento del regolatore in modo sistematico risolvendo un sistema di equazioni lineari, e garantisce *a priori* l'assenza di cancellazioni, a meno che queste vengano esplicitamente imposte.

Con riferimento al sistema di Figura 18.18, si ponga

$$G^*(z) = \frac{N(z)}{D(z)} = \frac{n_{v-1}z^{v-1} + n_{v-2}z^{v-2} + \dots + n_1z + n_0}{z^v + d_v z^{v-1} + \dots + d_1 z + d_0} \quad (18.54)$$

$$R^*(z) = \frac{Q(z)}{P(z)} = \frac{q_m z^m + q_{m-1} z^{m-1} + \dots + q_1 z + q_0}{p_m z^m + p_{m-1} z^{m-1} + \dots + p_1 z + p_0} \quad (18.55)$$

Si ipotizzi, cioè, che $G^*(z)$ rappresenti un sistema dinamico strettamente proprio di ordine v e che il regolatore sia invece un sistema proprio di ordine m da determinarsi. Nelle (18.54), (18.55) si assuma inoltre che $N(z)$ e $D(z)$, così come $O(z)$ e $P(z)$, non abbiano fattori in comune.

Il polinomio caratteristico del sistema retroazionato è dato da $D(z)P(z) + N(z)Q(z)$ ed è di grado $\nu + m$. Si voglia ora impostare che esso coincida con un polinomio prefissato

$$C(z) = z^{v+m} + c_{v+m-1}z^{v+m-1} + \dots + c_1z + c_0 \quad (18.56)$$

ovviamente scelto in modo da garantire favorevoli posizioni dei poli in anello chiuso. Per la scelta di $C(z)$ ci si può basare su considerazioni simili a quelle viste a proposito dei metodi di sintesi precedenti. Ci si riconduce allora a considerare l'equazione polinomiale

$$D(z)P(z) + N(z)Q(z) = C(z) \quad (18.57)$$

che, noti $D(z)$, $N(z)$ e $C(z)$, interessa risolvere nelle incognite $P(z)$ e $Q(z)$ o, equivalentemente, nei coefficienti incogniti p_i e q_i .

L'Equazione (18.57) prende il nome di *equazione diofantina*. Si può dimostrare che essa ammette un'unica soluzione per ogni generico $C(z)$ se e solo se i polinomi $D(z)$ e $N(z)$, come ipotizzato, sono primi tra loro.

Utilizzando le (18.54)-(18.56), l'Equazione (18.57) si può riscrivere come

$$\begin{aligned} & (z^v + d_{v-1}z^{v-1} + \dots + d_1z + d_0)(p_m z^m + p_{m-1}z^{m-1} + \dots + p_1z + p_0) + \\ & + (n_{v-1}z^{v-1} + n_{v-2}z^{v-2} + \dots + n_0)(q_m z^m + q_{m-1}z^{m-1} + \dots + q_0) = \\ & = z^{v+m} + c_{v+m-1}z^{v+m-1} + \dots + c_1z + c_0 \end{aligned} \quad (18.58)$$

Poiché dalla (18.55) il numero di parametri incogniti è $2m+2$, e poiché i $v+m+1$ coefficienti del polinomio a primo membro della (18.58) sono funzioni lineari di tali incognite, è immediato concludere che, per la risolubilità del problema, in generale è necessario che sia $v+m+1 \leq 2m+2$, in modo che la (18.57) implichi un numero di relazioni non superiore al numero di incognite. Si deduce quindi che l'ordine del regolatore deve essere almeno $m = v - 1$. Supponendo di effettuare proprio questa scelta, è facile verificare, uguagliando i coefficienti di pari grado, che l'Equazione (18.58) è equivalente al sistema lineare

$$\left[\begin{array}{ccccccccc} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ d_{v-1} & 1 & 0 & \cdots & 0 & n_{v-1} & 0 & \cdots & 0 \\ d_{v-2} & d_{v-1} & 1 & \cdots & 0 & n_{v-2} & n_{v-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ d_0 & d_1 & d_2 & \cdots & d_{v-1} & n_0 & n_1 & \cdots & n_{v-1} \\ 0 & d_0 & d_1 & \cdots & d_{v-2} & 0 & n_0 & \cdots & n_{v-2} \\ 0 & 0 & d_0 & \cdots & d_{v-3} & 0 & 0 & \cdots & n_{v-3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & d_0 & 0 & 0 & \cdots & n_0 \end{array} \right] \left[\begin{array}{c} p_m \\ p_{m-1} \\ p_{m-2} \\ \vdots \\ p_0 \\ q_m \\ q_{m-1} \\ \vdots \\ q_0 \end{array} \right] = \left[\begin{array}{c} 1 \\ c_{v-2} \\ c_{v-3} \\ \vdots \\ c_{v-1} \\ c_{v-2} \\ c_{v-3} \\ \vdots \\ c_0 \end{array} \right] \quad (18.59)$$

nelle incognite p_i e q_i . Si dimostra che la matrice dei coefficienti, grazie alla sua particolare struttura, è invertibile se e solo se i polinomi $D(z)$ e $N(z)$ non hanno fattori comuni. Dalla (18.59) si ricavano quindi i parametri del regolatore.

Per assicurare un'accettabile robustezza numerica alla soluzione del sistema (18.59) è necessario che la matrice dei coefficienti non solo sia invertibile, ma sia anche ben condizionata, cioè sufficientemente "lontana" dalla singolarità. A questo riguardo si dimostra che essa è tanto peggio condizionata quanto più i polinomi $D(z)$ e $N(z)$ hanno radici prossime tra loro. D'altra parte, questi polinomi derivano dalla rappresentazione a segnali campionati del sistema sotto controllo e, come si è visto nel Paragrafo 18.2, possono avere radici che si addensano tutte intorno al punto 1 quando il periodo di campionamento è molto piccolo. Per garantire una buona precisione numerica alla soluzione del sistema (18.59) bisogna allora evitare valori troppo bassi del passo di campionamento T .

Precisione statica Per come è stato finora illustrato, il metodo ad assegnamento dei poli non fornisce alcuna garanzia sulle prestazioni statiche del sistema di controllo. È peraltro molto facile mettere a punto opportune varianti atte a tener conto di eventuali requisiti di questo genere. Per esempio, se è richiesta la regolazione a zero dell'errore in presenza di ingressi a scalino e se il processo $G^*(z)$ non contiene già di per sé un polo in $z = 1$, è possibile effettuare il progetto del regolatore secondo lo schema di Figura 18.27. In tale schema il regolatore è costituito dalla serie di un blocco $\tilde{R}(z)$ da progettare e di un integratore a tempo discreto, descritto dalla funzione di trasferimento $z/(z - 1)$. Si può quindi procedere alla sintesi di $\tilde{R}(z)$ con il metodo ad assegnamento dei poli discusso in precedenza, facendo però riferimento al "processo allargato" con funzione di trasferimento

$$\tilde{G}(z) = \frac{zG^*(z)}{z-1} = \frac{zN(z)}{(z-1)D(z)} \quad (18.60)$$

Poiché l'ordine di $\tilde{G}(z)$ supera di 1 quello del sistema originario v , la funzione di trasferimento $\tilde{R}(z)$ da sintetizzare dovrà questa volta avere ordine almeno pari a v , e il regolatore complessivo risultante $R^*(z) = z\tilde{R}(z)/(z - 1)$ sarà in generale di ordine $v + 1$.

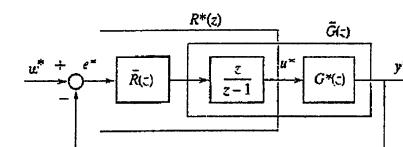


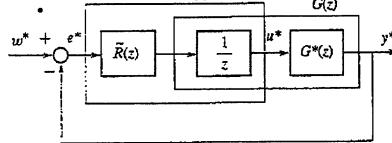
Figura 18.27
Schema utilizzato per
introdurre un'azione
integrale.

Regolatore strettamente proprio Il regolatore (18.55) risulta in generale non strettamente proprio. Con un procedimento analogo a quello appena visto relativo all'introduzione dell'azione integrale, si può imporre che il grado relativo di $R^*(z)$ sia pari a 1. A tale scopo basta riferirsi allo schema di Figura 18.28, dove questa volta il "processo allargato" rispetto al quale si conduce il progetto ha funzione di trasferimento

$$\tilde{G}(z) = \frac{G^*(z)}{z} = \frac{N(z)}{zD(z)}$$

Una volta determinata $\tilde{R}(z)$ per questo sistema allargato, si ricava il regolatore $R^*(z) = \tilde{R}(z)/z$, che ha un grado relativo almeno pari a 1 per costruzione.

Figura 18.28
Schema utilizzato per ottenere un regolatore strettamente proprio.



Esempio 18.12 (Seguito degli Esempi 18.10 e 18.11) Usando l'Equazione (18.41) e volendo imporre la presenza di un'azione integrale, si definisca, in accordo con la (18.60), la funzione di trasferimento del processo allargato

$$\tilde{G}(z) = \frac{zG^*(z)}{z-1} = \frac{0.0041z(z+0.819)}{(z-1)(z-0.905)(z-0.607)} = \\ = \frac{0.0041z^2 + 0.0034z}{z^3 - 2.51z^2 + 2.06z - 0.549} = \frac{\tilde{N}(z)}{\tilde{D}(z)}$$

Poiché il grado di $\tilde{D}(z)$ è $v = 3$, si adotti il regolatore di ordine $m = 2$

$$\tilde{R}(z) = \frac{Q(z)}{P(z)} = \frac{q_2 z^2 + q_1 z + q_0}{p_2 z^2 + p_1 z + p_0}$$

di cui vanno determinati i parametri. Si scelga inoltre il polinomio caratteristico in anello chiuso, di grado $v+m = 5$,

$$C(z) = (z-0.6)^3(z-0.8)^2$$

in modo da rispettare le specifiche dinamiche. Risolvendo il sistema (18.59) si ricava

$$P(z) = z^2 - 0.9186z + 0.2516 \quad Q(z) = 7.4668z^2 - 11.1257z + 4.0022 \quad (18.61)$$

e pertanto il regolatore risultante è

$$R^*(z) = \frac{z}{z-1} \tilde{R}(z) = \frac{7.4668z^3 - 11.1257z^2 + 4.0022z}{(z-1)(z^2 - 0.9186z + 0.2516)} \quad (18.62)$$

Le prestazioni di tale regolatore in presenza di uno scalino del riferimento sono mostrate nella Figura 18.29. L'andamento non del tutto soddisfacente dell'uscita y , con una leggera sovraelongazione e un lento assottigliamento sul valore di regime, è dovuto alla presenza di uno zero "a bassa frequenza" in $z \approx 0.883$ del regolatore, e quindi anche della funzione di sensitività complementare $F^*(z)$ (si veda il Paragrafo 8.3.4).

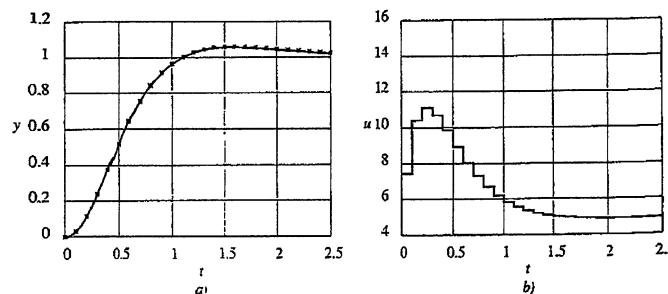
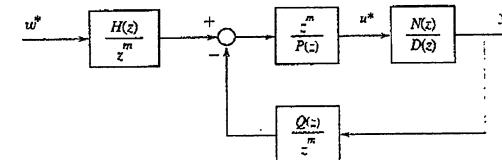


Figura 18.29
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 18.12 con il regolatore (18.62); a) uscita y ; b) variabile di controllo u .

Figura 18.30
Schema utilizzato per l'assegnamento degli zeri.



Assegnamento degli zeri Come si è visto nell'Esempio 18.12, il metodo ad assegnamento dei poli non permette di prevedere la posizione degli zeri della funzione di trasferimento tra w^* e y^* , e ciò può dar luogo a fenomeni indesiderati, come la sovraelongazione della Figura 18.29. Esistono però varianti del metodo, basate sull'uso di strutture a due gradi di libertà (si veda il Paragrafo 15.2.3), che consentono di fissare oltre al valore dei poli anche quello degli zeri, o di alcuni zeri, delle funzioni di trasferimento in anello chiuso. Per esempio, si può ricorrere allo schema di Figura 18.30, dove i polinomi $P(z)$ e $Q(z)$ sono ricavati come in precedenza risolvendo il sistema (18.59), mentre

$$H(z) = h_m z^m + h_{m-1} z^{m-1} + \dots + h_1 z + h_0$$

è un polinomio da determinare. Si osservi dapprima che, se si pone $H(z) = Q(z)$, la funzione di trasferimento tra w^* e y^* è la stessa che si calcola sul sistema di Figura 18.18 quando $G^*(z)$ e $R^*(z)$ sono date dalle (18.54), (18.55), e quindi ha tra i suoi zeri le radici di $Q(z)$. Invece, in corrispondenza di un arbitrario $H(z)$, dalla Figura 18.30 risulta

$$\frac{Y^*(z)}{W^*(z)} = \frac{N(z)H(z)}{D(z)P(z) + N(z)Q(z)} = \frac{N(z)H(z)}{C(z)}$$

e dunque, dopo aver assegnato i poli con $C(z)$, è possibile fissare anche alcuni zeri di tale funzione di trasferimento con un'opportuna scelta di $H(z)$. Naturalmente, perché i requisiti statici del sistema di controllo siano soddisfatti, occorre che il guadagno risulti unitario e cioè sia $H(1) = C(1)/N(1)$. L'utilità di questa soluzione sarà illustrata nell'esempio seguente.

Esempio 18.13 (Seguito degli Esempi 18.11 e 18.12) Con riferimento allo schema di Figura 18.31, dove $P(z)$ e $Q(z)$ sono definiti come in (18.61), si ponga

$$H(z) = 8.582z - 0.8z^2$$

in modo che gli zeri di $H(z)$ cancellino i due poli in anello chiuso in $z = 0.8$ e venga rispettata la condizione che garantisce guadagno unitario tra w^* e y^* . La corrispondente risposta a uno scalino del riferimento è riportata nella Figura 18.32 e appare decisamente migliore rispetto a quella della Figura 18.29.

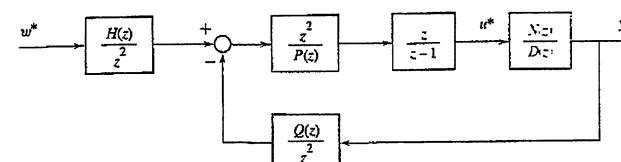
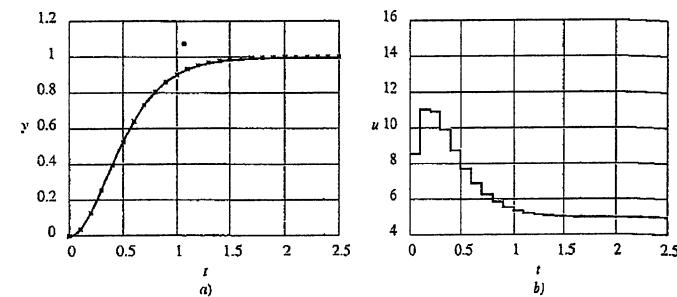


Figura 18.31
Schema del sistema di controllo dell'Esempio 18.13.

Figura 18.32
Risposta a uno scalino del riferimento per il sistema di controllo dell'Esempio 18.13: a) uscita y ; b) variabile di controllo u .



A conclusione del paragrafo, è opportuno fare alcune osservazioni. In primo luogo, si noti che il metodo ad assegnamento dei poli qui introdotto, e basato essenzialmente sulla soluzione dell'equazione diofantina (18.57), è stato sviluppato per sistemi a tempo discreto. Tuttavia, pur di sostituire la variabile complessa z con la variabile complessa s ed eventualmente apportare ovvie modifiche per inserire un'azione integrale, esso può essere immediatamente applicato a sistemi a tempo continuo per i quali una soluzione è già stata fornita nel Capitolo 13. In secondo luogo, il regolatore sintetizzato con il procedimento qui descritto è di ordine $n - 1$, mentre l'approccio presentato nel Capitolo 13, è basato su una rappresentazione di stato del sistema, porta alla sintesi di un regolatore di ordine n . A questo proposito, rimandando ai testi citati in bibliografia per una trattazione più approfondita, si può qui semplicemente notare che $n - 1$ è l'ordine minimo necessario per risolvere il problema di determinare un regolatore per un sistema con retroazione unitaria che assegna i poli del sistema retroazionato. Peraltro, allo stesso risultato si potrebbe giungere anche con una semplice estensione delle tecniche presentate nel Capitolo 13.

18.6 Conclusioni

In questo capitolo e nel precedente si è mostrato che l'analisi e la sintesi di un sistema di controllo digitale possono essere condotte secondo due approcci differenti. Il primo prevede di descrivere, almeno in modo approssimato, il sottosistema costituito da campionatore, regolatore digitale e mantenitore mediante la funzione di trasferimento $R(s)$ di un equivalente regolatore a tempo continuo. In questo caso, il progetto del regolatore è effettuato con le tecniche relative ai sistemi a tempo continuo e la funzione di trasferimento del regolatore digitale $R^*(z)$ viene poi determinata da $R(s)$ con metodi approssimati di discretizzazione. Il secondo approccio consiste invece nel determinare la funzione di trasferimento $G^*(z)$ del sistema a segnali campionati formato dalla serie del mantenitore, del processo e del campionatore. Il sistema di controllo risultante è descritto da un sistema retroazionato a tempo discreto in cui compaiono $G^*(z)$ e la funzione di trasferimento $R^*(z)$ del regolatore digitale, che viene determinato con metodi di sintesi basati sull'assegnamento di un modello del sistema retroazionato o, più semplicemente, dei suoi poli.

Un aspetto differenzia sostanzialmente l'analisi (e la sintesi) "nel continuo" da quella "nel discreto". La prima, condotta nel dominio della frequenza, for-

nisce informazioni sul comportamento del sistema di controllo digitale su tutto l'asse del tempo continuo; la seconda consente invece di conoscere unicamente l'andamento delle variabili di ingresso, stato e uscita negli istanti di campionamento, ma non è in grado di descrivere il loro comportamento tra due istanti di campionamento successivi.

Infine, va segnalato che l'approccio a tempo discreto offre l'ambito più favorevole allo sviluppo di tecniche di progetto più avanzate, a cui non si è fatto cenno in questa trattazione, tra le quali si possono citare, a titolo di esempio, quelle di controllo ottimo, controllo predittivo, controllo stocastico, controllo adattativo, controllo "a logica fuzzy".

Esercizi

Esercizio 18.1 Si determini la funzione di trasferimento $G^*(z)$ del sistema a segnali campionati associato al processo descritto da

$$G(s) = \frac{b}{s+a} e^{-es}$$

in cui $0 \leq e \leq T$, dove T è il periodo di campionamento. Si valuti quindi la posizione dello zero di $G^*(z)$ al variare di e con particolare riferimento ai casi $e = 0$ e $e = T$.

Esercizio 18.2 Dato il sistema a tempo continuo descritto da

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

1. si verifichi che, con il periodo di campionamento $T = \pi$, la funzione di trasferimento del corrispondente sistema a segnali campionati è

$$G^*(z) = \frac{1 + e^{-\pi}}{z + e^{-\pi}}$$

2. si analizzi il risultato precedente e si giustifichi il fatto che $G^*(z)$ è di ordine inferiore rispetto a $G(s)$;

3. si determini un diverso possibile sistema a tempo continuo, e la relativa funzione di trasferimento $G(s)$, a cui corrisponde, sempre per $T = \pi$, la medesima $G^*(z)$.

Esercizio 18.3 Si ricavi la (18.20) a partire dalla (18.19).

Esercizio 18.4 Si discretizzi il regolatore

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

utilizzando la tecnica PZ e la sua variante consistente nel porre tutti gli zeri del campionamento in $z = 1$. Si confrontino poi le risposte in frequenza associate a $R^*(s)$ e ai due regolatori digitali ottenuti.

Esercizio 18.5 Si consideri un sistema con retroazione unitaria negativa in cui il processo e il regolatore sono descritti da

$$G(s) = \frac{1}{s} \quad R(s) = \mu > 0$$

Si mostri che se il regolatore è realizzato a tempo continuo, il sistema retroazionato è sempre stabile. Utilizzando poi i risultati di questo capitolo, si verifichi che se il regolatore è implementato secondo lo schema di controllo digitale di Figura 18.1, per qualsiasi valore del periodo di campionamento utilizzato esiste un valore massimo di μ oltre il quale il sistema retroazionato è instabile. Si interpreti questo risultato con l'analisi nel continuo presentata nel Paragrafo 17.5.

Esercizio 18.6 Per il sistema a segnali campionati descritto dalla funzione di trasferimento

$$G^*(z) = \frac{1}{1.1 - z}$$

si progetti con il metodo ad assegnamento del modello un regolatore che garantisca errore a transitorio esaurito nullo a fronte di segnali di riferimento a scalino, e tale per cui tutti i poli del sistema retroazionato siano in $z = 0$. Si interpreti il risultato ottenuto in termini di luogo delle radici.

Esercizio 18.7 Si ripeta l'esercizio precedente utilizzando il metodo dell'assegnamento dei poli e si confrontino le soluzioni ottenute. Si ripeta ancora l'esercizio imponendo che il regolatore sia strettamente proprio.

Esercizio 18.8 Dato il sistema a tempo continuo descritto da

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

si utilizzi il metodo dell'assegnamento dei poli, presentato in questo capitolo ed esteso al caso dei sistemi a tempo continuo, per determinare il regolatore $R(s)$ tale per cui tutti i poli del sistema retroazionato siano in $s = -1$.

Esercizio 18.9 Si dimostri che se, nel metodo dell'assegnamento dei poli discusso nel Paragrafo 18.5.4, si sceglie $C(z)$ in modo che contenga dei fattori di $D(z)$, essi compariranno necessariamente anche in $Q(z)$. Ciò equivale quindi a imporre una cancellazione tra zeri del regolatore e poli del sistema. Viceversa, se $C(z)$ contiene fattori di $N(z)$, essi risulteranno anche fattori di $P(z)$ e si realizzerà una cancellazione tra poli del regolatore e zeri del sistema.

Esercizio 18.10 Per il sistema a segnali campionati descritto dalla funzione di trasferimento

$$G^*(z) = \frac{N(z)}{D(z)} = \frac{n_1 z + n_0}{z^2 + d_1 z + d_0}$$

si estenda la tecnica ad assegnamento dei poli utilizzando un regolatore del secondo ordine (cioè pari all'ordine del sistema). Si verifichi che in questo caso vi è un grado di libertà che può essere sfruttato per ottenere un regolatore strettamente proprio.

A.1 Introduzione

Questa appendice raccoglie definizioni e risultati di calcolo matriciale utili nello studio di molte questioni riguardanti i sistemi dinamici. La trattazione non ha alcuna pretesa di generalità o completezza e tende a presentare le nozioni utili in questo testo nella maniera più semplice e funzionale possibile.

Sinteticamente, i temi trattati sono:

- le definizioni e le operazioni fondamentali sulle matrici;
- gli autovalori, gli autovettori, la forma diagonale e la forma di Jordan;
- le matrici potenza ed esponenziale;
- i limiti, le derivate e gli integrali.

A.2 Definizioni e operazioni fondamentali

A.2.1 Generalità

Una *matrice* è una tabella di numeri o funzioni organizzati in righe e colonne. Se n ed m sono il numero di righe e quello delle colonne della tabella, cioè le *dimensioni* della matrice, quest'ultima, sinteticamente, si dice $n \times m$. Una matrice 1×1 si dice *scalare*.

Di solito le matrici si indicano con lettere maiuscole e i loro elementi con le minuscole corrispondenti, opportunamente indicate. Denotando allora con a_{ij} l'elemento sulla riga i e sulla colonna j di una matrice A , si scrive

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

Se $n = 1$ la matrice è un *vettore riga*, se $m = 1$ la matrice è un *vettore colonna*: col semplice termine *vettore* di solito si fa riferimento a un vettore colonna.

I vettori riga e colonna si indicano a volte con lettere minuscole e, se a_j è l'elemento sulla colonna j del vettore riga a , si scrive

$$a = [a_1 \ a_2 \ \cdots \ a_m]$$

Una notazione analoga si usa per i vettori colonna.

Una matrice $A, n \times m$, e un vettore colonna $a, n \times 1$, costituiti da elementi reali o da funzioni che assumano valori reali si indicano sinteticamente con i simboli $A \in R^{n \times m}$ e $a \in R^n$; se invece A e a sono complessi si scrive $A \in C^{n \times m}$ e $a \in C^n$.

Se $n = m$ la matrice è *quadrata*, altrimenti è *rettangolare*.

Gli elementi a_{ii} , cioè di indici uguali, di una matrice quadrata A costituiscono la sua *diagonale*; quelli $a_{i,i+1}$, cioè con indice della colonna pari all'indice della riga più uno, formano invece la *sopradiagonale*. La stessa matrice quadrata si dice *triangolare inferiore* se sono nulli tutti i suoi elementi che si trovano al di sopra (triangolare inferiore) o al di sotto (triangolare superiore) della diagonale, cioè se sono nulli gli elementi $a_{ij}, i < j$, o rispettivamente gli elementi $a_{ij}, i > j$. Una particolare matrice triangolare è quella *diagonale*, in cui tutti gli elementi fuori dalla diagonale sono nulli: detto a_i il generico elemento sulla diagonale, si scrive

$$A = \text{diag} \{a_1, a_2, \dots, a_n\} \quad (\text{A.1})$$

Una matrice può essere vista come costituita da *blocchi*, o *sottomatrici*, cioè da altre matrici semplicemente giustapposte una all'altra in modo opportuno. I blocchi sono indicati seguendo la stessa convenzione adottata per i singoli elementi di una matrice qualunque.

Per esempio, con le matrici $B, n_1 \times m_1$, $C, n_1 \times m_2$, $D, n_2 \times m_1$, $E, n_2 \times m_2$, si può costruire la matrice a blocchi

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$

dotata di $n_1 + n_2$ righe e $m_1 + m_2$ colonne; B si può denotare con A_{11} , C con A_{12} e così via.

Per una *matrice a blocchi* quadrata con blocchi di pari indice anch'essi quadrati, si definiscono i concetti di matrice *diagonale a blocchi* e matrice *triangolare a blocchi*, nonché di *diagonale dei blocchi*, estendendo in modo ovvio le definizioni e la notazione (A.1) sopra riportate, cioè riferendole ai blocchi invece che ai singoli elementi.

La matrice *nulla* è costituita esclusivamente da zeri; può avere dimensioni qualunque e si indica col simbolo 0_{nm} o, quando non vi sia possibilità di equivoco, col simbolo 0. La matrice *identità* I_n è diagonale con tutti i suoi elementi non nulli pari a 1; spesso è indicata semplicemente con I .

Due matrici A e B sono *uguali* se hanno le stesse dimensioni e sono uguali tutti i loro elementi di stessi indici; si scrive $A = B$.

A.2.2 Operazioni su una matrice

La *trasposta* $B = A'$ di una matrice $A, n \times m$, è la matrice $m \times n$, in cui $b_{ij} = a_{ji}$.

A una matrice quadrata $A, n \times n$, è associato il *determinante* $\det(A)$, definito dall'algoritmo ricorsivo di calcolo seguente. Per ogni coppia (i, j) si denoti innanzitutto con \hat{A}_{ij} il *complemento algebrico* dell'elemento a_{ij} , cioè il determinante della matrice ottenuta da A eliminandone la riga i e la colonna j moltiplicato per $(-1)^{i+j}$. Allora, se $n = 1$ si pone

$$\det(A) = a_{11}$$

e se $n > 1$ si definisce

$$\det(A) = \sum_{i=1}^n a_{ij} \hat{A}_{ij} = \sum_{j=1}^n a_{ij} \hat{A}_{ij}$$

dove nella prima formula di calcolo j è arbitrario, mentre nella seconda è arbitrario i . Così facendo, il calcolo del determinante di una matrice viene ricondotto al calcolo di determinanti associati a matrici di dimensioni inferiori.

In particolare, per n uguale a 2 risulta

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

e per n uguale a 3

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{13}a_{22} - a_{32}a_{23}a_{11} - a_{12}a_{21}a_{33}$$

Il determinante di una matrice triangolare a blocchi con blocchi quadrati sulla diagonale è il prodotto dei determinanti dei blocchi posti sulla diagonale.

Una matrice quadrata A con determinante non nullo si dice *non singolare*, o anche *invertibile*. In questo caso, infatti, si definisce la matrice $B, n \times n$, *inversa* di A , ponendo

$$b_{ij} = \frac{1}{\det(A)} \hat{A}_{ji}$$

e si scrive $B = A^{-1}$. Risulta poi

$$(A^{-1})^{-1} = A$$

Per esempio, per $n = 2$ si ha

$$A^{-1} = \begin{bmatrix} \frac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}} & \frac{-a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \\ \frac{-a_{21}}{a_{11}a_{22} - a_{12}a_{21}} & \frac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}} \end{bmatrix}$$

L'inversa della matrice diagonale (A.1), assunta non singolare, è

$$A^{-1} = \text{diag} \left\{ \frac{1}{a_1}, \frac{1}{a_2}, \dots, \frac{1}{a_n} \right\}$$

A proposito della condizione di invertibilità $\det(A) \neq 0$, si osservi che, se A è una matrice numerica, il suo determinante è un numero che non deve essere nullo. Se invece A è costituita da funzioni, il suo determinante è anch'esso una funzione che occorre non sia identicamente nulla; questa funzione però può annullarsi per particolari valori delle variabili indipendenti.

Si trova poi che

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

Il *prodotto* di una matrice A per uno scalare α si effettua semplicemente moltiplicando per lo scalare tutti gli elementi della matrice. Per questa operazione valgono le proprietà usuali dell'algebra: in particolare $\alpha A = A\alpha$. Se A è quadrata, $n \times n$, si può verificare che

$$\det(\alpha A) = \alpha^n \det(A)$$

Da una matrice $A, n \times m$, si possono costruire altre matrici selezionandone solo gli elementi posti su alcune righe e colonne, come in fondo si è già fatto nel definire il concetto di complemento algebrico. Tra queste matrici si considerino quelle quadrate, di dimensioni $\bar{n} \times \bar{n}$. Il massimo valore di \bar{n} per cui almeno una di

queste ultime matrici è non singolare si dice *rango* di A e si denota con $\rho(A)$. Naturalmente risulta

$$0 \leq \rho(A) \leq \min\{n, m\}$$

Se $\rho(A) = \min\{n, m\}$ la matrice A è di *rango massimo*; se, inoltre, $n = m$ la matrice A è di rango massimo quando $\det(A) \neq 0$.

Un qualunque insieme di k righe o colonne A_i di una matrice A , $n \times m$, si dicono *linearmente indipendenti* quando risultă

$$\sum_{i=1}^k \gamma_i A_i \neq 0$$

comunque si scelgano, non tutti nulli, gli scalari γ_i , quando cioè nessuna di esse si possa ottenere come arbitraria *combinazione lineare* delle altre. Si dimostra che il massimo numero di righe e colonne linearmente indipendenti estraibili da una matrice è pari al suo rango; le $\rho(A)$ righe e le $\rho(A)$ colonne con i cui elementi si può costruire una matrice non singolare sono linearmente indipendenti.

La *traccia* di una matrice quadrata A , $n \times n$, è la somma degli elementi posti sulla diagonale

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}$$

Naturalmente

$$\text{tr}(A') = \text{tr}(A)$$

Per ogni scalare α risulta inoltre

$$\text{tr}(\alpha A) = \alpha \text{tr}(A)$$

Per un vettore $w \in R^n$ si definisce la *norma* secondo la relazione

$$\|w\| = (w_1^2 + w_2^2 + \dots + w_n^2)^{1/2}$$

Se $n = 1$ la norma coincide con il modulo dello scalare w .

A.2.3 Operazioni tra matrici

La *somma* (algebrica) di due matrici di identiche dimensioni A e B è una matrice C delle stesse dimensioni il cui generico elemento è pari alla somma dei corrispondenti elementi dei due addendi; si scrive $C = A + B$. Per la somma tra matrici valgono le usuali proprietà dell'algebra. Qualunque sia la matrice A risulta $A + 0 = A$.

Il *prodotto* tra due matrici A , $n \times m$, e B , $p \times q$, si definisce quando $m = p$, cioè il numero di colonne del primo fattore coincide col numero di righe del secondo; il risultato dell'operazione è una matrice $C = AB$, $n \times q$, il cui generico elemento è

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}$$

Si constata che il prodotto tra matrici, mentre gode della proprietà associativa e di quella distributiva, non gode della proprietà commutativa; anzi, se si può definire il prodotto AB , non è neppure detto che si possa definire il prodotto BA . Qualunque sia A però risulta

$$AI_m = I_n A = A$$

e, se $n = m$,

$$AA^{-1} = A^{-1}A = I$$

Queste relazioni giustificano, tra l'altro, il modo in cui sono state definite la matrice identità e la matrice inversa.

È interessante notare che, se A e B sono due matrici quadrate, non singolari e delle stesse dimensioni, allora anche il loro prodotto è non singolare e risulta

$$(AB)^{-1} = B^{-1}A^{-1}$$

La *potenza* k -esima, $k > 0$ intero, di una matrice quadrata A si definisce in modo ovvio come

$$A^k = \underbrace{AA \dots A}_{k \text{ volte}}$$

mentre si ha

$$A^{-k} = (A^k)^{-1} = (A^{-1})^k, \quad A^0 = I$$

In particolare, quando A è diagonale ed espressa dalla (A.1), qualunque sia k intero

$$A^k = \text{diag}\{a_1^k, a_2^k, \dots, a_n^k\}$$

Per finire, vale la pena citare, tra le tante esistenti, le proprietà che seguono. Date due matrici A , $n \times m$, e B , $m \times n$, risulta

$$\text{tr}(AB) = \text{tr}(BA), \quad \det(AB) = \det(BA)$$

Se poi A e B sono quadrate

$$\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B), \quad \det(AB) = \det(A)\det(B)$$

A.3 Autovalori e autovettori

A.3.1 Generalità

Detta λ una variabile complessa, a una matrice quadrata A , $n \times n$, si possono associare il *polinomio caratteristico*, di grado n ,

$$\varphi(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \alpha_2 \lambda^{n-2} + \dots + \alpha_{n-1} \lambda + \alpha_n \quad (\text{A.2})$$

e l'*equazione caratteristica*

$$\varphi(\lambda) = 0 \quad (\text{A.3})$$

Le n soluzioni λ_i dell'equazione caratteristica si dicono *autovalori* di A ; se quest'ultima è costituita da numeri reali, anche i coefficienti α_i sono reali e quindi gli autovalori sono reali oppure complessi coniugati a coppie.

Si verifica facilmente che

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

Gli autovalori di una matrice diagonale a blocchi, o anche solamente triangolare a blocchi, sono gli autovalori dei blocchi posti sulla diagonale. In particolare, gli autovalori di una matrice diagonale o triangolare sono gli elementi posti sulla diagonale della matrice.

Si osservi che, per ogni matrice $T, n \times n$, non singolare, la matrice TAT^{-1} ha gli stessi autovalori di A . Infatti

$$\begin{aligned}\det(\lambda I - TAT^{-1}) &= \det(T\lambda T^{-1} - TAT^{-1}) = \det(T(\lambda I - A)T^{-1}) = \\ &= \det(T)\det(\lambda I - A)\det(T^{-1}) = \det(\lambda I - A)\end{aligned}$$

Le matrici A e TAT^{-1} si dicono *simili*.

Si può dimostrare che la matrice A soddisfa la sua equazione caratteristica. Cioè, se si valuta in A il polinomio caratteristico (A.2), ponendo

$$\varphi(A) = A^n + \alpha_1 A^{n-1} + \alpha_2 A^{n-2} + \dots + \alpha_{n-1} A + \alpha_n I$$

si ottiene

$$\varphi(A) = 0$$

Questo risultato va sotto il nome di *teorema di Cayley-Hamilton*.

A ogni autovalore λ_i si può associare un vettore colonna $v_i, n \times 1$, non nullo che soddisfa l'equazione

$$(\lambda_i I - A) v_i = 0 \quad (\text{A.4})$$

Questo vettore si dice *autovettore* e non è definito in modo unico.

A.3.2 Forma diagonale e forma di Jordan

Si supponga che gli autovalori di A siano tutti *distinti* tra loro.

Si verifica subito che le n equazioni del tipo (A.4) si possono scrivere in maniera compatta come

$$A [v_1 \ v_2 \ \dots \ v_n] = [v_1 \ v_2 \ \dots \ v_n] \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$$

Si dimostra che la *matrice degli autovettori*, $n \times n$, è non singolare e perciò si può porre

$$T_D^{-1} = [v_1 \ v_2 \ \dots \ v_n]$$

cosicché, definendo la *forma diagonale* di A

$$A_D = \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} \quad (\text{A.5})$$

si ottiene infine

$$A_D = T_D A T_D^{-1} \quad (\text{A.6})$$

e quindi la matrice diagonale A_D dell'Equazione (A.5) è simile alla matrice A originaria. Si dice in tal caso che A è *diagonalizzabile*. Naturalmente se A è reale, ma qualche suo autovalore non lo è, anche il corrispondente autovettore e la matrice A_D risultano non reali.

Se si elimina l'ipotesi che A abbia autovalori distinti, se cioè ne ha alcuni *molti*, essa può ancora risultare diagonalizzabile secondo un procedimento del tutto simile a quello che ha portato alla (A.6). Ciò accade quando esistono n autovettori che soddisfino le n equazioni della forma (A.4) e generino una matrice degli autovettori non singolare, cosa quest'ultima non garantibile in generale. Se, invece, tali autovettori non esistono, con una trasformazione del tipo di quella definita dalla (A.6) la matrice A si può porre in una forma A_J , detta *forma di Jordan*, che ha elementi non nulli solo sulla diagonale e sulla sopradiagonale. Si dimostra cioè l'esistenza di una matrice $T_J, n \times n$, non singolare, tale che

$$A_J = T_J A T_J^{-1} \quad (\text{A.7})$$

e inoltre, denotando con μ il numero di autovalori di A distinti tra loro e con n_i la *moltiplicità* dell'autovalore λ_i , cioè il numero di radici della (A.3) pari a λ_i , risulti

$$A_J = \operatorname{diag}\{A_{J1}, A_{J2}, \dots, A_{J\mu}\} \quad (\text{A.8})$$

$$A_{Ji} = \operatorname{diag}\{A_{Ji1}, A_{Ji2}, \dots, A_{Ji\eta_{ih}}\}, \quad i = 1, 2, \dots, \mu \quad (\text{A.9})$$

$$A_{Jih} = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_i & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{bmatrix} \in R^{\eta_{ih} \times \eta_{ih}} \quad (\text{A.10})$$

$$i = 1, 2, \dots, \mu, \quad h = 1, 2, \dots, \nu_i$$

dove $\sum_{h=1}^{\nu_i} \eta_{ih} = n_i$. La forma di Jordan A_J , pertanto, è diagonale a blocchi; la matrice A_{Ji} è il *blocco di Jordan* corrispondente all'autovalore λ_i , mentre le matrici A_{Jih} sono i *miniblocchi di Jordan* corrispondenti a λ_i .

A.4 Potenza

Per ogni matrice $A, n \times n$, e ogni scalare intero $k \geq 0$ si è visto che la matrice potenza $\Phi(k)$ è definita dalla relazione

$$\Phi(k) = A^k \quad (\text{A.11})$$

che estende la classica definizione valida per gli scalari.

Se A è diagonalizzabile, per la (A.6) esiste una matrice T_D tale che $A = T_D^{-1} A_D T_D$, dove A_D ha la forma (A.5). Sostituendo nella (A.11) si trova

$$\Phi(k) = (T_D^{-1} A_D T_D)^k = T_D^{-1} A_D^k T_D = T_D^{-1} \operatorname{diag}\{\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k\} T_D$$

Pertanto T_D mette anche $\Phi(k)$ nella sua forma diagonale

$$\Phi_D(k) = \operatorname{diag}\{\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k\}$$

nel senso che

$$T_D \Phi(k) T_D^{-1} = \Phi_D(k)$$

e il generico elemento di posizione (α, β) di $\Phi(k)$ assume la forma

$$\varphi_{\alpha\beta}(k) = \sum_{i=1}^n r_{\alpha\beta i} \lambda_i^k \quad (\text{A.12})$$

dove i coefficienti $r_{\alpha\beta i}$ sono indipendenti da k .

Nel caso più generale in cui A non sia diagonalizzabile, utilizzando le Equazioni (A.7)-(A.11), si ricava

$$\Phi(k) = T_J^{-1} \operatorname{diag}\{A_{J1}^k, A_{J2}^k, \dots, A_{J\mu}^k\} T_J$$

dove

$$A_{ji}^k = \text{diag} \{ A_{ji1}^k, A_{ji2}^k, \dots, A_{ji\nu_i}^k \}, \quad i = 1, 2, \dots, \mu$$

$$A_{Jih}^k = \begin{bmatrix} \xi_0 & \xi_1 & \xi_2 & \cdots & \xi_{\eta_{ih}-2} & \xi_{\eta_{ih}-1} \\ 0 & \xi_0 & \xi_1 & \cdots & \xi_{\eta_{ih}-3} & \xi_{\eta_{ih}-2} \\ 0 & 0 & \xi_0 & \cdots & \xi_{\eta_{ih}-4} & \xi_{\eta_{ih}-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \xi_0 & \xi_1 \\ 0 & 0 & 0 & \cdots & 0 & \xi_0 \end{bmatrix}, \quad \xi_l = \begin{cases} \binom{k}{l} \lambda_i^{k-l}, & l \leq k \\ 0, & l > k \end{cases}$$

$$i = 1, 2, \dots, \mu, \quad h = 1, 2, \dots, \nu_i, \quad l = 0, 1, \dots, \eta_{ih}-1$$

La potenza $\Phi(k)$ conserva perciò la struttura di blocchi e miniblocchi di Jordan di A e il suo generico elemento (α, β) ha la forma

$$\varphi_{\alpha\beta}(k) = \sum_{i=1}^{\mu} \sum_{l=0}^{\min\{k, \hat{\eta}_i-1\}} r_{\alpha\beta il}(k) \lambda_i^{k-l} \quad (\text{A.13})$$

dove i termini $r_{\alpha\beta il}(k)$ sono polinomi in k di grado non maggiore di l , mentre

$$\hat{\eta}_i = \max_{h=1,2,\dots,\nu_i} \{\eta_{ih}\}$$

rappresenta la massima tra le dimensioni dei miniblocchi di Jordan associati a λ_i .

Se A è reale, evidentemente anche $\Phi(k)$ è reale in vista della sua definizione (A.11); questo però non è evidente dalle Equazioni (A.12) e (A.13) nelle quali i termini $r_{\alpha\beta i}$ e $r_{\alpha\beta il}(k)$ e gli autovalori λ_i possono non essere reali. Tuttavia, se $\lambda_i = \rho_i e^{j\theta_i}$ non è reale, tra gli autovalori di A esiste anche il suo complesso coniugato, cioè esiste $\lambda_h = \bar{\lambda}_i$, e per di più si dimostra che $r_{\alpha\beta i} = \bar{r}_{\alpha\beta i}$, $r_{\alpha\beta il}(k) = \bar{r}_{\alpha\beta il}(k)$. Allora, per la formula di Eulero,

$$\begin{aligned} & r_{\alpha\beta i} \lambda_i^k + r_{\alpha\beta h} \lambda_h^k = \\ & = \{ [\text{Re}(r_{\alpha\beta i}) + j \text{Im}(r_{\alpha\beta i})] \rho_i^k e^{j\theta_i k} + [\text{Re}(r_{\alpha\beta i}) - j \text{Im}(r_{\alpha\beta i})] \rho_i^k e^{-j\theta_i k} \} = \\ & = \rho_i^k \{ [\text{Re}(r_{\alpha\beta i}) + j \text{Im}(r_{\alpha\beta i})] [\cos(\theta_i k) + j \sin(\theta_i k)] + \\ & \quad + [\text{Re}(r_{\alpha\beta i}) - j \text{Im}(r_{\alpha\beta i})] [\cos(\theta_i k) - j \sin(\theta_i k)] \} = \\ & = 2\rho_i^k [\text{Re}(r_{\alpha\beta i}) \cos(\theta_i k) - \text{Im}(r_{\alpha\beta i}) \sin(\theta_i k)] \end{aligned}$$

e analogamente

$$\begin{aligned} & r_{\alpha\beta il}(k) \lambda_i^{k-l} + r_{\alpha\beta hl}(k) \lambda_h^{k-l} = \\ & = 2\rho_i^{k-l} [\text{Re}(r_{\alpha\beta il}(k)) \cos(\theta_i(k-l)) - \text{Im}(r_{\alpha\beta il}(k)) \sin(\theta_i(k-l))] \end{aligned}$$

Pertanto gli elementi $\varphi_{\alpha\beta}(k)$ di Equazioni (A.12) e (A.13) assumono proprio valori reali.

Queste ultime due equazioni possono essere ancora utilmente elaborate, cosa che qui si farà innanzitutto nel caso (A.12) in cui la matrice A sia diagonalizzabile. Si ponga

$$r_{\alpha\beta i} = |r_{\alpha\beta i}| e^{j \arg r_{\alpha\beta i}}$$

Allora

$$\begin{aligned} & 2\rho_i^k [\text{Re}(r_{\alpha\beta i}) \cos(\theta_i k) - \text{Im}(r_{\alpha\beta i}) \sin(\theta_i k)] = \\ & = 2|r_{\alpha\beta i}| \rho_i^k \left\{ \left[\frac{\text{Re}(r_{\alpha\beta i})}{|r_{\alpha\beta i}|} \right] \cos(\theta_i k) - \left[\frac{\text{Im}(r_{\alpha\beta i})}{|r_{\alpha\beta i}|} \right] \sin(\theta_i k) \right\} = \\ & = 2|r_{\alpha\beta i}| \rho_i^k \{ \cos(\arg r_{\alpha\beta i}) \cos(\theta_i k) - \sin(\arg r_{\alpha\beta i}) \sin(\theta_i k) \} = \\ & = 2|r_{\alpha\beta i}| \rho_i^k \cos(\theta_i k + \arg r_{\alpha\beta i}) \end{aligned}$$

Perciò, mentre un autovalore reale produce nella (A.12) una potenza, una coppia di autovalori complessi coniugati genera una potenza moltiplicata per una cosinusoida.

Quando A non è diagonalizzabile, elaborazioni analoghe alle precedenti possono essere condotte sulla (A.13) facendo riferimento ai coefficienti $r_{\alpha\beta il\eta}$ delle singole potenze k^η , $0 \leq \eta \leq \hat{\eta}_i - 1$, nei polinomi $r_{\alpha\beta il}(k)$. Si trovano così termini del tipo

$$2|r_{\alpha\beta il\eta}| k^\eta \rho_i^{k-l} \cos(\theta_i(k-l) + \arg r_{\alpha\beta il\eta})$$

i quali mostrano che, mentre autovalori reali multipli producono nella (A.13) potenze degli autovalori moltiplicate per polinomi in k , coppie multiple di autovalori complessi coniugati generano potenze dei moduli degli autovalori moltiplicate per cosinusoidi e per opportuni polinomi in k .

A.5 Esponenziale

Per ogni matrice A , $n \times n$, e ogni scalare reale $t \geq 0$ la matrice *esponenziale* $\Psi(t)$ è definita dalla relazione

$$\Psi(t) = e^{At} = \sum_{i=0}^{\infty} \frac{1}{i!} (At)^i = I_n + At + \frac{(At)^2}{2!} + \dots \quad (\text{A.14})$$

che estende la classica definizione valida per gli scalari.

Se A è diagonalizzabile, per la (A.6) esiste una matrice T_D tale che $A = T_D^{-1} A_D T_D$, dove A_D ha la forma (A.5). Sostituendo nella (A.14) si trova

$$\begin{aligned} \Psi(t) &= \sum_{i=0}^{\infty} \frac{1}{i!} (T_D^{-1} A_D t T_D)^i = T_D^{-1} \left(\sum_{i=0}^{\infty} \frac{1}{i!} (A_D t)^i \right) T_D = \\ &= T_D^{-1} \text{diag} \{ e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t} \} T_D \end{aligned}$$

Pertanto T_D mette anche $\Psi(t)$ nella sua forma diagonale

$$\Psi_D(t) = \text{diag} \{ e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t} \}$$

nel senso che

$$T_D \Psi(t) T_D^{-1} = \Psi_D(t)$$

e il generico elemento di posizione (α, β) di $\Psi(t)$ assume la forma

$$\psi_{\alpha\beta}(t) = \sum_{i=1}^n p_{\alpha\beta i} e^{\lambda_i t} \quad (\text{A.15})$$

dove i coefficienti $p_{\alpha\beta i}$ sono indipendenti da t .

Nel caso più generale in cui A non sia diagonalizzabile, utilizzando le Equazioni (A.7)-(A.10), (A.14), si ricava

$$\Psi(t) = T_J^{-1} \operatorname{diag} \{e^{A_{J1}t}, e^{A_{J2}t}, \dots, e^{A_{J\mu}t}\} T_J$$

dove

$$e^{A_{Jit}} = \operatorname{diag} \{e^{A_{J1it}}, e^{A_{J2it}}, \dots, e^{A_{J\mu it}}\} \quad i = 1, 2, \dots, \mu$$

$$e^{A_{Jit}} = \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{\eta_{ih}-2}}{(\eta_{ih}-2)!} & \frac{t^{\eta_{ih}-1}}{(\eta_{ih}-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{\eta_{ih}-3}}{(\eta_{ih}-3)!} & \frac{t^{\eta_{ih}-2}}{(\eta_{ih}-2)!} \\ 0 & 0 & 1 & \cdots & \frac{t^{\eta_{ih}-4}}{(\eta_{ih}-4)!} & \frac{t^{\eta_{ih}-3}}{(\eta_{ih}-3)!} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & t \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} e^{\lambda_i t}$$

$$i = 1, 2, \dots, \mu \quad h = 1, 2, \dots, v_i$$

L'esponenziale $\Psi(t)$ conserva perciò la struttura di blocchi e miniblocchi di Jordan di A e il suo generico elemento (α, β) ha la forma

$$\psi_{\alpha\beta}(t) = \sum_{i=1}^{\mu} p_{\alpha\beta i}(t) e^{\lambda_i t} \quad (\text{A.16})$$

dove i termini $p_{\alpha\beta i}(t)$ sono polinomi in t di grado non maggiore di $\hat{\eta}_i - 1$ e inoltre

$$\hat{\eta}_i = \max_{h=1,2,\dots,v_i} \{\eta_{ih}\}$$

rappresenta la massima tra le dimensioni dei miniblocchi di Jordan associati a λ_i .

Se A è reale, evidentemente anche $\Psi(t)$ è reale in vista della sua definizione (A.14); questo però non è evidente dalle Equazioni (A.15) e (A.16) nelle quali i coefficienti $p_{\alpha\beta i}$ e gli autovalori λ_i possono non essere reali. Tuttavia, se $\lambda_i = \sigma_i + j\omega_i$ non è reale, tra gli autovalori di A esiste anche il suo complesso coniugato, cioè esiste $\lambda_h = \bar{\lambda}_i$, e per di più si dimostra che $p_{\alpha\beta h}(t) = \bar{p}_{\alpha\beta i}(t)$. Allora, per la formula di Eulero,

$$\begin{aligned} p_{\alpha\beta i}(t) e^{\lambda_i t} + p_{\alpha\beta h}(t) e^{\lambda_h t} &= \\ &= e^{\sigma_i t} [\operatorname{Re}(p_{\alpha\beta i}(t)) + j \operatorname{Im}(p_{\alpha\beta i}(t))] [\cos(\omega_i t) + j \sin(\omega_i t)] + \\ &\quad + [\operatorname{Re}(p_{\alpha\beta i}(t)) - j \operatorname{Im}(p_{\alpha\beta i}(t))] [\cos(\omega_i t) - j \sin(\omega_i t)] = \\ &= 2e^{\sigma_i t} [\operatorname{Re}(p_{\alpha\beta i}(t)) \cos(\omega_i t) - \operatorname{Im}(p_{\alpha\beta i}(t)) \sin(\omega_i t)] \end{aligned}$$

Pertanto gli elementi $\psi_{\alpha\beta}(t)$ di Equazioni (A.15) e (A.16) assumono proprio valori reali.

L'equazione precedente può essere ancora utilmente elaborata, cosa che qui si farà innanzitutto nel caso (A.15) in cui la matrice A sia diagonalizzabile. Si ponga

$$p_{\alpha\beta i} = |p_{\alpha\beta i}| e^{j \arg p_{\alpha\beta i}}$$

Allora

$$\begin{aligned} 2e^{\sigma_i t} [\operatorname{Re}(p_{\alpha\beta i}) \cos(\omega_i t) - \operatorname{Im}(p_{\alpha\beta i}) \sin(\omega_i t)] &= \\ &= 2 |p_{\alpha\beta i}| e^{\sigma_i t} \left\{ \left[\frac{\operatorname{Re}(p_{\alpha\beta i})}{|p_{\alpha\beta i}|} \right] \cos(\omega_i t) - \left[\frac{\operatorname{Im}(p_{\alpha\beta i})}{|p_{\alpha\beta i}|} \right] \sin(\omega_i t) \right\} = \\ &= 2 |p_{\alpha\beta i}| e^{\sigma_i t} \{ \cos(\arg p_{\alpha\beta i}) \cos(\omega_i t) - \sin(\arg p_{\alpha\beta i}) \sin(\omega_i t) \} = \\ &= 2 |p_{\alpha\beta i}| e^{\sigma_i t} \cos(\omega_i t + \arg p_{\alpha\beta i}) \end{aligned}$$

Perciò, mentre un autovalore reale produce nella (A.15) un termine esponenziale, una coppia di autovalori complessi coniugati genera un esponenziale moltiplicato per una cosinusoida.

Quando A non è diagonalizzabile, elaborazioni analoghe alle precedenti possono essere condotte sulla (A.16) facendo riferimento ai coefficienti $p_{\alpha\beta i\eta}$ delle singole potenze t^η , $0 \leq \eta \leq \hat{\eta}_i - 1$, nei polinomi $p_{\alpha\beta i}(t)$. Si trovano così termini del tipo

$$2 |p_{\alpha\beta i\eta}| t^\eta e^{\sigma_i t} \cos(\omega_i t + \arg p_{\alpha\beta i\eta})$$

i quali mostrano che, mentre autovalori reali multipli producono nella (A.16) termini esponenziali moltiplicati per polinomi in t , coppie multiple di autovalori complessi coniugati generano funzioni esponenziali moltiplicate per cosinusoidi e per opportuni polinomi in t .

A.6 Limiti, derivate e integrali

Alle matrici costituite da funzioni di una singola variabile scalare, reale o complessa, si possono estendere senza problemi le classiche nozioni di *limite*, *derivata*, *integrale definito* e *indefinito*, semplicemente applicando ai singoli elementi della matrice le definizioni e le regole di calcolo valide per le funzioni scalari. Si estende pure, in modo ovvio, il concetto di *derivata parziale* di una matrice funzione di più variabili scalari.

Vale la pena di spiegare invece che cosa si intende per *derivata di un vettore rispetto a un vettore*: dato un vettore colonna f , $n \times 1$, funzione di un vettore colonna x , $m \times 1$, la derivata di f rispetto a x è la matrice $n \times m$ il cui elemento (i, j) è costituito dalla derivata parziale della funzione scalare $f_i(x)$ rispetto alla variabile scalare x_j , cioè

$$\frac{df(x)}{dx} = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_m} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \cdots & \frac{\partial f_2(x)}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \cdots & \frac{\partial f_n(x)}{\partial x_m} \end{bmatrix}$$

La *derivata parziale di un vettore rispetto a un vettore* si effettua estendendo quanto sopra in maniera del tutto naturale al caso in cui un vettore di funzioni dipenda da più vettori di variabili. Per esempio, dato un vettore colonna f , $n \times 1$,

funzione di un vettore colonna $x, m \times 1$, e di un vettore y , la derivata parziale di f rispetto a x è la matrice $n \times m$ seguente:

$$\frac{\partial f(x, y)}{\partial x} = \begin{bmatrix} \frac{\partial f_1(x, y)}{\partial x_1} & \frac{\partial f_1(x, y)}{\partial x_2} & \dots & \frac{\partial f_1(x, y)}{\partial x_m} \\ \frac{\partial f_2(x, y)}{\partial x_1} & \frac{\partial f_2(x, y)}{\partial x_2} & \dots & \frac{\partial f_2(x, y)}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n(x, y)}{\partial x_1} & \frac{\partial f_n(x, y)}{\partial x_2} & \dots & \frac{\partial f_n(x, y)}{\partial x_m} \end{bmatrix}$$

Segnali a tempo continuo

B

B.1 Introduzione

L'obiettivo di questa appendice consiste nel presentare gli strumenti più comuni per l'analisi nel dominio della variabile complessa ω , come pure si dice, nel dominio della frequenza dei segnali a tempo continuo, cioè delle funzioni della variabile reale t che rappresenta il tempo. Questi strumenti sono costituiti dalla trasformata di Laplace, dallo sviluppo in serie di Fourier e dalla trasformata di Fourier, e saranno presentati non solo per le funzioni definite in senso classico, ma anche per la funzione generalizzata detta impulso di Dirac. Rinunciando a volte al rigore formale e alla completezza, ci si limiterà a discutere le questioni fondamentali relative al tema indicato in modo semplice e intuitivo, trattando specificatamente:

- i cosiddetti segnali canonici e in particolare l'impulso di Dirac;
- la definizione e le proprietà della trasformata di Laplace, con particolare riferimento al metodo di antitrasformazione che utilizza lo sviluppo di Heaviside;
- la definizione, il significato e le proprietà dello sviluppo in serie di Fourier;
- la definizione, il significato e le proprietà della trasformata di Fourier, nonché le sue relazioni con la trasformata di Laplace.

B.2 Impulso e altri segnali canonici

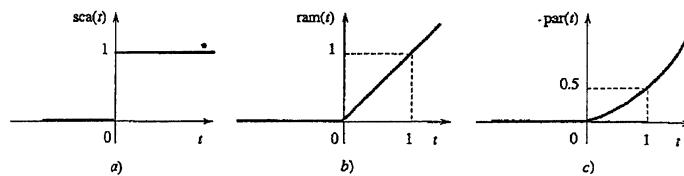
Nello studio dei sistemi dinamici a tempo continuo, in particolare di quelli lineari e stazionari, si fa spesso riferimento alle seguenti funzioni reali appartenenti alla classe dei *segnali canonici*:

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

$$\text{ram}(t) = \begin{cases} 0, & t < 0 \\ t, & t \geq 0 \end{cases}$$

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

Figura B.1
Principali segnali canonici a tempo continuo: a) scalino; b) rampa; c) parabola.



Essi sono detti rispettivamente *scalino*, *rampa* e *parabola* e sono rappresentati nella Figura B.1. È evidente che

$$\int_{-\infty}^t \text{sca}(\tau) d\tau = \text{ram}(t), \quad \int_{-\infty}^t \text{ram}(\tau) d\tau = \text{par}(t)$$

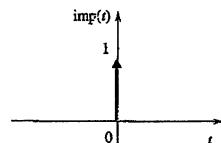
e, d'altra parte, la rampa è la derivata della parabola, mentre lo scalino è la derivata della rampa per $t \neq 0$. Altri segnali canonici si possono ottenere per successive integrazioni della parabola. Palesemente, invece, non esiste alcuna funzione, definita in senso classico come corrispondenza tra due insiemi, che abbia come integrale lo scalino. Una tale funzione, detta *impulso di Dirac*, o semplicemente *impulso*, e denotata col simbolo $\text{imp}(t)$, si può definire solo in senso generalizzato facendo riferimento alla teoria delle distribuzioni. Senza approfondire la questione, qui ci si limiterà a definire l'impulso mediante le relazioni

$$\text{imp}(t) = 0, \quad t \neq 0 \quad (\text{B.1})$$

$$\int_{-\infty}^{+\infty} \text{imp}(t) dt = 1 \quad (\text{B.2})$$

nella seconda delle quali l'integrale può essere limitato a un qualunque intervallo contenente il punto $t = 0$, in vista della prima. Si noti che, se l'impulso fosse una funzione definita in senso classico, l'Equazione (B.2) sarebbe incompatibile con la (B.1).

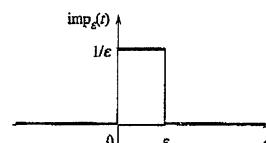
Figura B.2
Impulso di Dirac.



Questa funzione si rappresenta convenzionalmente col simbolo di Figura B.2 e, in via puramente intuitiva, si può ritenere sia un segnale di durata estremamente breve e ampiezza estremamente elevata ("infinita"). Infatti, considerando un impulso rettangolare (Figura B.3) definito dalla relazione

$$\text{imp}_\varepsilon(t) = \begin{cases} 1/\varepsilon, & 0 \leq t < \varepsilon \\ 0, & \text{altrove} \end{cases}$$

Figura B.3
Impulso rettangolare.



si può assumere che risulti

$$\lim_{\varepsilon \rightarrow 0} \text{imp}_\varepsilon(t) = \text{imp}(t)$$

Si osservi che, data una qualunque funzione $\varphi(t)$ continua in $t = \tau$, risulta per la (B.1)

$$\varphi(t) \text{imp}(t - \tau) = \varphi(\tau) \text{imp}(t - \tau)$$

e quindi per la (B.2)

$$\int_{-\infty}^{+\infty} \varphi(t) \text{imp}(t - \tau) dt = \varphi(\tau) \int_{-\infty}^{+\infty} \text{imp}(t - \tau) dt = \varphi(\tau) \quad (\text{B.3})$$

Questa relazione, in cui l'intervallo di integrazione potrebbe essere ridotto a piacemento col solo vincolo di contenere il punto $t = \tau$, costituisce una definizione alternativa dell'impulso, basata non tanto sui valori che esso assume, ma piuttosto sull'effetto prodotto su altre funzioni.

Dalle Equazioni (B.1), (B.2) risulta immediatamente che, almeno per $t \neq 0$,

$$\int_{-\infty}^t \text{imp}(\tau) d\tau = \text{sca}(t)$$

Inoltre, per una qualunque funzione φ continua e derivabile in $t = 0$, si ha

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{d(\text{sca}(t))}{dt} \varphi(t) dt &= [\text{sca}(t) \varphi(t)]|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \text{sca}(t) \frac{d(\varphi(t))}{dt} dt = \\ &= \varphi(\infty) - 0 - \int_0^{+\infty} \frac{d(\varphi(t))}{dt} dt = \varphi(\infty) - \varphi(t)|_0^{+\infty} = \varphi(0) \end{aligned}$$

Si conclude, allora, che la derivata dello scalino soddisfa la proprietà (B.3) e quindi è un impulso, cioè

$$\frac{d(\text{sca}(t))}{dt} = \text{imp}(t)$$

In termini più ampi si può affermare che, nell'ambito delle funzioni generalizzate, la derivata di una funzione discontinua di prima specie è definibile e, in corrispondenza di ogni punto di discontinuità, essa contiene un impulso di ampiezza pari alla differenza tra i limiti destro e sinistro della funzione.

B.3 Trasformata di Laplace

B.3.1 Generalità

Sia data una funzione complessa f della variabile reale t ; sia poi $s = \sigma + j\omega \in C$ una variabile complessa con parte reale σ e coefficiente dell'unità immaginaria j pari a ω . Se la funzione

$$F(s) = \int_0^{+\infty} f(t) e^{-st} dt \quad (\text{B.4})$$

esiste almeno per qualche valore di s , essa si dice *trasformata di Laplace* di $f(t)$. Di solito si denota la trasformata con la stessa lettera, però maiuscola, che indica la funzione del tempo trasformanda e si usa scrivere $F(s) = \mathcal{L}[f(t)]$ e

$f(t) = \mathcal{L}^{-1}[F(s)]$. La definizione (B.4) richiede che $f(t)$ sia definita almeno per $t \geq 0$; essa può essere usata pure per funzioni definite per $t < 0$, ma i corrispondenti valori di f non concorrono comunque a determinare la trasformata; tipicamente, ma non necessariamente, si assume $f(t) = 0$ per $t < 0$. Inoltre, il primo estremo dell'integrale deve intendersi pari a 0^- , nel senso che se la funzione f contiene impulsi applicati al tempo $t = 0$ questi devono essere tenuti in conto nell'integrazione.

Quanto sopra e i risultati che seguono si estendono in modo naturale a funzioni f vettoriali.

L'operazione di *trasformazione di Laplace* ha quindi successo quando esistono valori di s tali che l'integrale nell'Equazione (B.4) converga. Sia $\tilde{\sigma} > -\infty$ l'estremo inferiore delle parti reali di tali valori di s ; si dimostra che la trasformata esiste in un semipiano definito dalla condizione $\operatorname{Re}(s) > \tilde{\sigma}$; per questo motivo $\tilde{\sigma}$ si dice *ascissa di convergenza*. La trasformata di Laplace risulta essere una *funzione analitica* e, grazie alle particolari proprietà delle funzioni analitiche, la sua definizione può essere estesa anche in punti s tali che $\operatorname{Re}(s) \leq \tilde{\sigma}$, indipendentemente dal fatto che l'integrale nella (B.4) non converga. La problematica della convergenza, comunque, non sarà qui approfondita e, anzi, si assumerà tacitamente che esistano tutte le trasformate di cui si avrà occasione di discutere.

Di particolare importanza sono le *trasformate razionali*, cioè quelle in cui

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

con $N(s)$ e $D(s)$ polinomi primi tra loro. Le radici dell'equazione $N(s) = 0$ si dicono *zeri* e quelle dell'equazione $D(s) = 0$ si dicono *poli*: nell'insieme, poli e zeri si chiamano *singolarità*. Quando f è reale i coefficienti dei polinomi N e D sono reali e l'ascissa di convergenza è pari alla massima tra le parti reali dei poli; è chiaro peraltro che $F(s)$ è ben definita per tutti i valori di s diversi dai poli.

La funzione trasformanda può essere ricavata dalla sua trasformata mediante la *formula di antitrasformazione*

$$f(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} F(s) e^{st} ds \quad (B.5)$$

dove σ è un qualunque numero reale maggiore di $\tilde{\sigma}$; per precisione occorre dire che $f(t)$ è fornita per $t \geq 0$, perché solo tale parte di f contribuisce a determinare F , e solo nei punti di continuità; la (B.5) dà invece $f(t) = 0$ per $t < 0$. Quindi, limitandosi a considerare funzioni nulle per tempi negativi, la corrispondenza tra funzioni del tempo trasformabili secondo Laplace e trasformate è biumivoca, fermo restando che due funzioni devono considerarsi uguali anche se differiscono per i valori assunti in un numero, anche infinito, di punti isolati.

L'Equazione (B.5) è molto importante da un punto di vista concettuale perché insieme alla (B.4) afferma l'esistenza di una corrispondenza biumivoca tra una funzione del tempo trasformabile $f(t)$, $t \geq 0$, e la sua trasformata funzione di variabile complessa $F(s)$: in altri termini, queste due funzioni hanno lo stesso contenuto informativo, cioè non sono altro che due modi diversi di rappresentare la stessa entità. La formula (B.5) è invece poco usata per effettuare l'antitrasformazione in pratica; per risolvere questo problema nel seguito del paragrafo sarà presentato un metodo molto semplice utilizzabile per le funzioni razionali.

Esempio B.1 Trasformata dell'impulso Sia $f(t) = \operatorname{imp}(t)$; allora, qualunque sia s , per l'Equazione (B.3) si ottiene

$$\mathcal{L}[\operatorname{imp}(t)] = \int_0^{+\infty} \operatorname{imp}(t) e^{-st} dt = e^{-s0} = 1$$

Esempio B.2 Trasformata dello scalino Sia $f(t) = \operatorname{sca}(t)$. In questo caso si ha

$$\begin{aligned} \mathcal{L}[\operatorname{sca}(t)] &= \int_0^{+\infty} \operatorname{sca}(t) e^{-st} dt = \int_0^{+\infty} e^{-st} dt = \\ &= \left. \frac{e^{-st}}{-s} \right|_0^{+\infty} = \left[\frac{e^{-\sigma t}}{-(\sigma + j\omega)} (\cos(\omega t) - j \sin(\omega t)) \right]_0^{+\infty} \end{aligned}$$

e, per $\sigma > 0$,

$$\mathcal{L}[\operatorname{sca}(t)] = \frac{0 - 1}{-(\sigma + j\omega)} = \frac{1}{s}$$

Pertanto risulta $\tilde{\sigma} = 0$, ma la funzione $1/s$ è definita per ogni $s \neq 0$.

B.3.2 Proprietà principali

La trasformazione di Laplace gode di parecchie proprietà. Le più importanti per le applicazioni sono riportate nel seguito, senza le spesso semplici dimostrazioni, e sono usate per ricavare alcune utili trasformate evitando di ricorrere alla definizione (B.4).

Linearità Si abbiano due funzioni f e g . Allora $\forall \alpha \in C, \forall \beta \in C$ risulta

$$\mathcal{L}[\alpha f(t) + \beta g(t)] = \alpha F(s) + \beta G(s)$$

cioè la trasformazione di Laplace è un'operazione lineare.

Traslazione nel dominio del tempo Per un qualunque $\tau > 0$ si consideri la funzione $\hat{f}(t) = f(t - \tau)$ ottenuta traslando in avanti la funzione $f(t)$, supposta nulla per tempi negativi, di un tempo pari a τ . Si trova

$$\mathcal{L}[\hat{f}(t)] = \mathcal{L}[f(t - \tau)] = e^{-\tau s} F(s)$$

Esempio B.3 Si consideri la funzione $f(t) = \alpha \operatorname{sca}(t - \tau_1) + \beta \operatorname{sca}(t - \tau_2)$ riportata nella Figura B.4 in un caso in cui $\tau_2 > \tau_1 > 0$, $\alpha > 0$ e $\beta < 0$. Per le proprietà di linearità e traslazione nel dominio del tempo si ha

$$\begin{aligned} \mathcal{L}[f(t)] &= \mathcal{L}[\alpha \operatorname{sca}(t - \tau_1)] + \mathcal{L}[\beta \operatorname{sca}(t - \tau_2)] = \\ &= \alpha \mathcal{L}[\operatorname{sca}(t - \tau_1)] + \beta \mathcal{L}[\operatorname{sca}(t - \tau_2)] = \alpha \frac{e^{-\tau_1 s}}{s} + \beta \frac{e^{-\tau_2 s}}{s} \end{aligned}$$

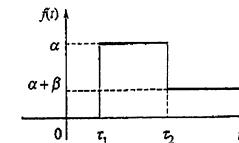


Figura B.4
Funzione dell'Esempio B.3.

Traslazione nel dominio della variabile complessa Per un qualunque $\alpha \in C$ si consideri la funzione $\hat{f}(t) = e^{\alpha t} f(t)$. La sua trasformata è

$$\mathcal{L}[\hat{f}(t)] = \mathcal{L}[e^{\alpha t} f(t)] = F(s - \alpha) \quad (\text{B.6})$$

Esempio B.4 Trasformata dell'esponenziale Ponendo $f(t) = \text{sca}(t)$ nella (B.6) e ricordando la trasformata del scalino con la sua ascissa di convergenza, si trova che la trasformata di Laplace della funzione $e^{\alpha t} \text{sca}(t)$, qualunque sia $\alpha \in C$, è

$$\mathcal{L}[e^{\alpha t} \text{sca}(t)] = \frac{1}{s - \alpha} \quad (\text{B.7})$$

cui corrisponde $\tilde{\sigma} = \operatorname{Re}(\alpha)$, pur esistendo essa per $s \neq \alpha$.

Esempio B.5 Trasformata della cosinusoida Si può calcolare facilmente la trasformata della funzione $\cos(\omega t) \text{sca}(t)$ utilizzando la formula della trasformata dell'esponenziale con $\alpha = j\omega$ e $\alpha = -j\omega$ e sfruttando la proprietà di linearità. Infatti, per $s \neq \pm j\omega$,

$$\begin{aligned} \mathcal{L}[\cos(\omega t) \text{sca}(t)] &= \mathcal{L}\left[\frac{e^{j\omega t} + e^{-j\omega t}}{2} \text{sca}(t)\right] = \\ &= \frac{1}{2} \left(\frac{1}{s - j\omega} + \frac{1}{s + j\omega} \right) = \frac{s}{s^2 + \omega^2} \end{aligned}$$

cui corrisponde $\tilde{\sigma} = 0$.

Derivazione nel dominio del tempo Si supponga che la funzione $f(t)$ sia derivabile, nel senso delle funzioni generalizzate, per tutti gli $t \geq 0$, o almeno dotata di derivate sinistra (per $t > 0$) e destra. Risulta allora

$$\mathcal{L}[\dot{f}(t)] = sF(s) - f(0) \quad (\text{B.8})$$

Se la funzione f è discontinua di prima specie in $t = 0$, allora \dot{f} contiene un impulso applicato al tempo 0 e $f(0)$ si deve interpretare come $f(0^-)$.

Si può dimostrare che, quando $F(s)$ è razionale con grado del denominatore maggiore del grado del numeratore, la derivata \dot{f} esiste ed è una funzione trasformabile definita in senso classico sommata eventualmente a un impulso di Dirac.

È evidente come la formula (B.8) sia iterabile così da consentire il calcolo delle trasformate delle derivate successive di f , ammesso che esistano; per esempio si ha

$$\mathcal{L}[\ddot{f}(t)] = s\mathcal{L}[\dot{f}(t)] - \dot{f}(0) = s^2 F(s) - sf(0) - \dot{f}(0)$$

e in generale

$$\mathcal{L}\left[\frac{d^n f(t)}{dt^n}\right] = s^n F(s) - \sum_{i=1}^n s^{n-i} \left. \frac{d^{i-1} f(t)}{dt^{i-1}} \right|_{t=0}$$

Quindi, se la funzione e le sue derivate sono nulle in $t = 0$, moltiplicare per s nel dominio della variabile complessa equivale a derivare nel dominio del tempo, cioè s si può interpretare come *operatore di derivazione*.

Esempio B.6 Trasformata della sinusoida Oltre che in maniera analoga a quella seguita per il calcolo della trasformata della funzione coseno, la trasformata della funzione $\sin(\omega t) \text{sca}(t)$ si può determinare notando innanzitutto che

$$\begin{aligned} \frac{d(\cos(\omega t) \text{sca}(t))}{dt} &= -\omega \sin(\omega t) \text{sca}(t) + \cos(\omega t) \text{imp}(t) = \\ &= -\omega \sin(\omega t) \text{sca}(t) + \text{imp}(t) \end{aligned}$$

Allora, per $s \neq \pm j\omega$,

$$\begin{aligned} \mathcal{L}[\sin(\omega t) \text{sca}(t)] &= \mathcal{L}\left[\frac{1}{\omega} \left(\text{imp}(t) - \frac{d(\cos(\omega t) \text{sca}(t))}{dt} \right)\right] = \\ &= \frac{1}{\omega} \left[1 - \left(\frac{s}{s^2 + \omega^2} \right) \right] = \frac{\omega}{s^2 + \omega^2} \end{aligned}$$

cui corrisponde $\tilde{\sigma} = 0$.

Derivazione nel dominio della variabile complessa Si supponga che la funzione $F(s)$ sia derivabile per tutti gli s , con l'eventuale esclusione di un numero finito di valori. Risulta

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

formula estendibile al calcolo della trasformata di $t^n f(t)$.

Esempio B.7 Trasformata della rampa Si noti che $\text{ram}(t) = t \text{sca}(t)$ e allora

$$\mathcal{L}[\text{ram}(t)] = -\frac{d\left(\frac{1}{s}\right)}{ds} = \frac{1}{s^2}$$

Integrazione nel dominio del tempo Si supponga che la funzione $f(t)$ sia integrabile tra 0 e $+\infty$. Allora

$$\mathcal{L}\left[\int_0^t f(\tau) d\tau\right] = \frac{1}{s} F(s)$$

cosicché dividere per s nel dominio della variabile complessa equivale a integrare nel dominio del tempo e $1/s$ si può interpretare come *operatore di integrazione*.

Esempio B.8 Trasformata della parabola Visto che

$$\text{par}(t) = \int_0^t \text{ram}(\tau) d\tau$$

si ottiene

$$\mathcal{L}[\text{par}(t)] = \frac{1}{s} \mathcal{L}[\text{ram}(t)] = \frac{1}{s^3}$$

Convoluzione nel dominio del tempo Date due funzioni f_1 e f_2 , il loro *prodotto di convoluzione* è

$$\begin{aligned} f_1(t) * f_2(t) &= \int_{-\infty}^{+\infty} f_1(\tau) f_2(t - \tau) d\tau = \int_{-\infty}^{+\infty} f_1(t - \eta) f_2(\eta) d\eta = \\ &= f_2(t) * f_1(t) \end{aligned}$$

Se esse sono nulle per tempi negativi, allora

$$f_1(t) * f_2(t) = \int_0^t f_1(\tau) f_2(t - \tau) d\tau = \int_0^t f_1(t - \eta) f_2(\eta) d\eta = f_2(t) * f_1(t)$$

e si trova

$$\mathcal{L}[f_1(t) * f_2(t)] = F_1(s) F_2(s)$$

Esempio B.9 Per le funzioni $f_1(t) = e^{\alpha t} \text{sca}(t)$ e $f_2(t) = \text{sca}(t)$ si ha

$$\mathcal{L}[f_1(t) * f_2(t)] = \frac{1}{s(s - \alpha)}$$

Teorema del valore iniziale Se una funzione reale f ha trasformata razionale F con grado del denominatore maggiore del grado del numeratore, allora

$$f(0) = \lim_{s \rightarrow \infty} s F(s) \quad (\text{B.9})$$

Se la funzione f è discontinua di prima specie in $t = 0$, $f(0)$ si deve interpretare come $f(0^+)$.

L'equazione (B.9) vale sotto larghe ipotesi anche se F non è razionale, purché $f(0)$, o almeno $f(0^+)$, esista.

Teorema del valore finale Se una funzione reale f ha trasformata razionale F con grado del denominatore maggiore del grado del numeratore e poli nulli o con parte reale negativa, allora

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} s F(s) \quad (\text{B.10})$$

L'equazione (B.10) vale sotto larghe ipotesi anche se F non è razionale, purché esista il $\lim_{t \rightarrow \infty} f(t)$.

Esempio B.10 Data la funzione $e^{\alpha t} \text{sca}(t)$, $\alpha \in \mathbb{R}$, in base al teorema del valore iniziale e alla (B.7) si trova

$$[e^{\alpha t} \text{sca}(t)]|_{t=0^+} = \lim_{s \rightarrow \infty} s \frac{1}{s - \alpha} = 1$$

e, in base al teorema del valore finale, applicabile per $\alpha \leq 0$,

$$\lim_{t \rightarrow \infty} e^{\alpha t} \text{sca}(t) = \lim_{s \rightarrow 0} s \frac{1}{s - \alpha} = \begin{cases} 0, & \alpha < 0 \\ 1, & \alpha = 0 \end{cases}$$

La Tabella B.1 raccoglie le trasformate di Laplace ricavate fino a questo punto insieme ad alcune altre la cui determinazione è lasciata al lettore.

Tabella B.1
Segnali e corrispondenti trasformate di Laplace.

$f(t)$	$F(s)$
imp(t)	1
sca(t)	$\frac{1}{s}$
ram(t)	$\frac{1}{s^2}$
par(t)	$\frac{1}{s^3}$
$e^{\alpha t} \text{sca}(t)$	$\frac{1}{s - \alpha}$
$t e^{\alpha t} \text{sca}(t)$	$\frac{1}{(s - \alpha)^2}$
$\sin(\omega t) \text{sca}(t)$	$\frac{\omega}{s^2 + \omega^2}$
$\cos(\omega t) \text{sca}(t)$	$\frac{s}{s^2 + \omega^2}$
$t \sin(\omega t) \text{sca}(t)$	$\frac{2\omega s}{(s^2 + \omega^2)^2}$
$t \cos(\omega t) \text{sca}(t)$	$\frac{s^2 - \omega^2}{(s^2 + \omega^2)^2}$
$e^{\sigma t} \sin(\omega t) \text{sca}(t)$	$\frac{\omega}{(s - \sigma)^2 + \omega^2}$
$e^{\sigma t} \cos(\omega t) \text{sca}(t)$	$\frac{s - \sigma}{(s - \sigma)^2 + \omega^2}$
$t e^{\sigma t} \sin(\omega t) \text{sca}(t)$	$\frac{2\omega(s - \sigma)}{[(s - \sigma)^2 + \omega^2]^2}$
$t e^{\sigma t} \cos(\omega t) \text{sca}(t)$	$\frac{(s - \sigma)^2 - \omega^2}{[(s - \sigma)^2 + \omega^2]^2}$

B.3.3 Sviluppo di Heaviside e antitrasformazione di trasformate razionali

Si vuole ora mostrare che è possibile antitrasformare una trasformata razionale con una tecnica semplice che non fa ricorso alla formula (B.5). Mediante il classico algoritmo per la divisione dei polinomi, una qualsiasi funzione razionale si può scrivere come somma di un'altra funzione razionale con grado del denominatore maggiore del grado del numeratore e un polinomio. Per linearità l'antitrasformata della funzione è la somma delle antitrasformate dei due addendi. Si può però dimostrare che un qualunque polinomio non può essere una trasformata di Laplace

di una funzione definita in senso classico e contenente eventualmente degli impulsi, a meno che il polinomio non si riduca a una costante. In quest'ultimo caso l'antitrasformata è un impulso applicato al tempo 0 di ampiezza pari alla costante stessa. Pertanto, all'interno della problematica in oggetto, l'unico caso interessante è quello in cui il denominatore della trasformata ha grado maggiore di quello del numeratore.

Nel seguito si affronterà perciò il problema dell'antitrasformazione della funzione razionale

$$F(s) = \frac{N(s)}{D(s)} \quad (\text{B.11})$$

in cui il grado di $D(s)$ è maggiore del grado di $N(s)$. Si assumerà anche che i coefficienti dei polinomi N e D siano reali, cosicché per ogni zero e polo non reale esisterà il complesso coniugato e la funzione antitrasformata $f(t)$ sarà reale.

Poli distinti In termini generali, l'idea alla base del metodo che si vuole presentare consiste nello sviluppare $F(s)$ nella somma di funzioni la cui antitrasformata è nota e poi, per la proprietà di linearità, ottenere l'antitrasformata come somma delle antitrasformate dei singoli addendi.

Occorre quindi mostrare come si possa ottenere lo sviluppo indicato, detto *sviluppo di Heaviside*, per la qual cosa si assumerà inizialmente che i poli di $F(s)$ siano tutti distinti tra loro. Se nella (B.11) si pone

$$D(s) = \prod_{i=1}^n (s + p_i) \quad , \quad p_h \neq p_j \text{ per } h \neq j \quad (\text{B.12})$$

con i $p_i \in C$, si possono trovare i coefficienti $P_i \in C$ in modo che sia soddisfatta l'identità

$$\frac{N(s)}{\prod_{i=1}^n (s + p_i)} \equiv \sum_{i=1}^n \frac{P_i}{s + p_i} \quad (\text{B.13})$$

Infatti, moltiplicando per $s + p_i$ si ottiene

$$(s + p_i) F(s) = \frac{N(s)}{\prod_{\substack{j=1 \\ j \neq i}}^n (s + p_j)} \equiv (s + p_i) \sum_{j=1}^n \frac{P_j}{s + p_j} + P_i$$

e valutando per $s = -p_i$ risulta

$$P_i = \frac{N(-p_i)}{\prod_{\substack{j=1 \\ j \neq i}}^n (p_j - p_i)} = \left. \frac{N(-p_i)}{dD(s)} \right|_{s=-p_i} = [(s + p_i) F(s)]|_{s=-p_i} \quad (\text{B.14})$$

Ricavati così i coefficienti P_i , detti *residui* di $F(s)$ nei poli $-p_i$, ricordando la Tabella B.1 e assumendo nulla l'antitrasformata per $t < 0$, si ottiene in definitiva

$$f(t) = \mathcal{L}^{-1}[F(s)] = \mathcal{L}^{-1} \left[\sum_{i=1}^n \frac{P_i}{s + p_i} \right] = \left(\sum_{i=1}^n P_i e^{-p_i t} \right) \text{sca}(t) \quad (\text{B.15})$$

Un metodo alternativo all'Equazione (B.14) per il calcolo dei residui consiste nell'eliminare i denominatori dalla (B.13) e imporre l'identità dei coefficienti delle varie potenze di s , pervenendo così a un sistema di n equazioni nelle n incognite P_i che, si dimostra, ammette una e una sola soluzione. A un tale sistema si giunge

pure valutando i due polinomi identici per n valori di s arbitrari e non necessariamente uguali ai poli, come è stato fatto, nella sostanza, quando si è ricavata l'Equazione (B.14).

Esempio B.11 Se

$$F(s) = \frac{s - 10}{(s + 2)(s + 5)}$$

si può scrivere

$$\frac{s - 10}{(s + 2)(s + 5)} = \frac{P_1}{s + 2} + \frac{P_2}{s + 5} \quad (\text{B.16})$$

dove per la (B.14)

$$P_1 = \frac{-2 - 10}{-2 + 5} = -4 \quad , \quad P_2 = \frac{-5 - 10}{-5 + 2} = 5$$

Alternativamente, eliminando i denominatori della (B.16), si ottiene

$$s - 10 \equiv P_1(s + 5) + P_2(s + 2) \quad (\text{B.17})$$

che, uguagliando i coefficienti delle potenze di s , produce il sistema di equazioni

$$\begin{aligned} 1 &= P_1 + P_2 \\ -10 &= 5P_1 + 2P_2 \end{aligned}$$

che ha come soluzione $P_1 = -4$, $P_2 = 5$. Valutando invece la (B.17) per $s = 0$ e $s = 1$, due valori scelti arbitrariamente, si ha

$$\begin{aligned} -10 &= 5P_1 + 2P_2 \\ -9 &= 6P_1 + 3P_2 \end{aligned}$$

I risultati naturalmente non cambiano e, in definitiva,

$$f(t) = \mathcal{L}^{-1} \left[\frac{s - 10}{(s + 2)(s + 5)} \right] = (-4e^{-2t} + 5e^{-5t}) \text{sca}(t)$$

Se non tutti i poli di $F(s)$ sono reali, l'Equazione (B.15) esprime la funzione reale f facendo ricorso a numeri non reali; questo però può essere evitato notando che per ogni polo non reale vi è anche il polo complesso coniugato e operando nella maniera seguente. Si supponga che sia $p_j = \bar{p}_i$; dalla (B.14) si trova allora che $P_j = \bar{P}_i$ e pertanto

$$\begin{aligned} \frac{P_i}{s + p_i} + \frac{P_j}{s + p_j} &= \frac{P_i}{s + p_i} + \frac{\bar{P}_i}{s + \bar{p}_i} = \\ &= 2 \frac{\operatorname{Re}(P_i)s + \operatorname{Re}(P_i)\operatorname{Re}(p_i) + \operatorname{Im}(P_i)\operatorname{Im}(p_i)}{s^2 + 2\operatorname{Re}(p_i)s + |p_i|^2} \end{aligned}$$

nonché, per la formula di Eulero,

$$\begin{aligned} P_i e^{-p_i t} + P_j e^{-p_j t} &= P_i e^{-p_i t} + \bar{P}_i e^{-\bar{p}_i t} = \\ &= 2 e^{\operatorname{Re}(-p_i)t} [\operatorname{Re}(P_i) \cos(\operatorname{Im}(-p_i)t) - \operatorname{Im}(P_i) \sin(\operatorname{Im}(-p_i)t)] = \\ &= 2 |P_i| e^{\operatorname{Re}(-p_i)t} \cos(\operatorname{Im}(-p_i)t + \arg P_i) \end{aligned} \quad (\text{B.18})$$

Sostituendo nella (B.15) i termini sopra ricavati per ogni coppia di poli complessi coniugati, la formula stessa rimane espressa in ambito puramente reale. Alternativamente, nello sviluppare $D(s)$ si può introdurre senz'altro un termine di secondo grado con numeratore di primo grado per ogni coppia di poli complessi coniugati

in luogo dei due di primo grado corrispondenti ai singoli poli; poi, eliminando i denominatori nella (B.13) si trova un'identità che, valutata per n differenti valori di s , fornisce n equazioni indipendenti in n incognite; infine si usa la Tabella B.1 per l'anttrasformazione dei singoli termini.

Esempio B.12 La funzione

$$F(s) = \frac{100}{(s+1)(s^2+4s+13)}$$

ha i suoi poli in $s = -1$ e $s = -2 \pm j3$. Secondo l'Equazione (B.14) il residuo del primo polo è

$$P_1 = \frac{100}{(2-j3-1)(2+j3-1)} = 10$$

mentre quello di $s = -2 + j3$ è

$$P_2 = \frac{100}{(1-2+j3)(2+j3-2+j3)} = -\frac{5}{3}(3-j)$$

e allora utilizzando la (B.18)

$$\begin{aligned} f(t) &= \mathcal{L}^{-1}\left[\frac{100}{(s+1)(s^2+4s+13)}\right] = \\ &= 10\left[e^{-t} + \frac{\sqrt{10}}{3}e^{-2t}\cos(3t + \arg(j-3))\right] \text{sca}(t) \end{aligned} \quad (\text{B.19})$$

Alternativamente si può porre

$$\frac{100}{(s+1)(s^2+4s+13)} = \frac{P_1}{s+1} + \frac{\Phi s + \Psi}{s^2+4s+13}$$

da cui, ricordando il valore di P_1 già noto,

$$100 = 10(s^2+4s+13) + (\Phi s + \Psi)(s+1)$$

Valutando, per esempio, per $s = 0$ e $s = 1$ si ottiene

$$\begin{aligned} 100 &= 130 + \Psi \\ 100 &= 180 + (\Phi + \Psi)2 \end{aligned}$$

e infine

$$\Phi = -10, \quad \Psi = -30$$

Si ha poi

$$\begin{aligned} f(t) &= \mathcal{L}^{-1}\left[\frac{100}{(s+1)(s^2+4s+13)}\right] = \\ &= \mathcal{L}^{-1}\left[\frac{10}{s+1} - \frac{10(s+3)}{s^2+4s+13}\right] = 10\mathcal{L}^{-1}\left[\frac{1}{s+1} - \frac{s+3}{s^2+4s+13}\right] = \\ &= 10\mathcal{L}^{-1}\left[\frac{1}{s+1} - \frac{s+2}{(s+2)^2+3^2} - \frac{1}{3}\frac{3}{(s+2)^2+3^2}\right] = \\ &= 10\left[e^{-t} - e^{-2t}\left(\cos(3t) + \frac{1}{3}\sin(3t)\right)\right] \text{sca}(t) \end{aligned}$$

da cui con semplici calcoli si ritrova la (B.19).

Pol multipli Quando $F(s)$ ha poli *multipli* il suo denominatore non ha più l'espressione (B.12). Assumendo che i poli distinti siano μ e che il polo $-p_i$ abbia molteplicità n_i , esso si può invece scrivere come

$$D(s) = \prod_{i=1}^{\mu} (s + p_i)^{n_i}, \quad p_h \neq p_j \text{ per } h \neq j$$

In corrispondenza la trasformata si sviluppa nella forma

$$\frac{N(s)}{\prod_{i=1}^{\mu} (s + p_i)^{n_i}} \equiv \sum_{i=1}^{\mu} \sum_{h=1}^{n_i} \frac{P_{i,h}}{(s + p_i)^h} \quad (\text{B.20})$$

dove $P_{i,1}$ è ancora detto *residuo* del polo $-p_i$. Questa formula generalizza la (B.13). Moltiplicando per $(s + p_i)^{n_i}$ si ottiene

$$\begin{aligned} (s + p_i)^{n_i} F(s) &= \frac{N(s)}{\prod_{j=1, j \neq i}^{\mu} (s + p_j)^{n_j}} \equiv \\ &\equiv (s + p_i)^{n_i} \sum_{j=1, j \neq i}^{\mu} \sum_{h=1}^{n_j} \frac{P_{j,h}}{(s + p_j)^h} + \sum_{h=1}^{n_i} (s + p_i)^{n_i-h} P_{i,h} \end{aligned} \quad (\text{B.21})$$

e valutando per $s = -p_i$

$$P_{i,n_i} = \frac{N(-p_i)}{\prod_{j=1, j \neq i}^{\mu} (p_j - p_i)^{n_j}} = [(s + p_i)^{n_i} F(s)]|_{s=-p_i}$$

Poi, derivando la (B.21) rispetto a s e valutando ancora per $s = -p_i$, si ha

$$P_{i,n_i-1} = \left. \frac{d[(s + p_i)^{n_i} F(s)]}{ds} \right|_{s=-p_i}$$

e più in generale si trova

$$P_{i,h} = \frac{1}{(n_i - h)!} \left. \frac{d^{n_i-h}[(s + p_i)^{n_i} F(s)]}{ds^{n_i-h}} \right|_{s=-p_i} \quad (\text{B.22})$$

Ricavati così i coefficienti $P_{i,h}$ e ricordando la Tabella B.1 nonché la proprietà di derivazione nel dominio della variabile complessa, si ottiene in definitiva

$$\begin{aligned} f(t) &= \mathcal{L}^{-1}[F(s)] = \mathcal{L}^{-1}\left[\sum_{i=1}^{\mu} \sum_{h=1}^{n_i} \frac{P_{i,h}}{(s + p_i)^h}\right] = \\ &= \left(\sum_{i=1}^{\mu} \sum_{h=1}^{n_i} P_{i,h} \frac{t^{h-1} e^{-p_i t}}{(h-1)!}\right) \text{sca}(t) \end{aligned} \quad (\text{B.23})$$

Anche per le funzioni con poli multipli è possibile determinare i coefficienti dello sviluppo di Heaviside identificando i polinomi che si ottengono dopo aver eliminato i denominatori dell'Equazione (B.20), nonché trattare in ambito reale il caso di poli complessi coniugati.

Esempio B.13 Se

$$F(s) = \frac{s+18}{s(s+3)^2}$$

si può scrivere

$$\frac{s+18}{s(s+3)^2} = \frac{P_{1,1}}{s} + \frac{P_{2,1}}{s+3} + \frac{P_{2,2}}{(s+3)^2}$$

dove per la (B.22)

$$\begin{aligned} P_{1,1} &= [sF(s)]|_{s=0} = \left. \frac{s+18}{(s+3)^2} \right|_{s=0} = 2 \\ P_{2,1} &= \left. \frac{d[(s+3)^2 F(s)]}{ds} \right|_{s=-3} = -2 \\ P_{2,2} &= \left. [(s+3)^2 F(s)] \right|_{s=-3} = -5 \end{aligned}$$

Allora in definitiva, per l'Equazione (B.23),

$$\begin{aligned} f(t) &= \mathcal{L}^{-1} \left[\frac{s+18}{s(s+3)^2} \right] = \mathcal{L}^{-1} \left[\frac{2}{s} - \frac{2}{s+3} - \frac{5}{(s+3)^2} \right] = \\ &= (2 - 2e^{-3t} - 5te^{-3t}) \text{ sca}(t) \end{aligned}$$

B.4 Serie di Fourier

B.4.1 Forma esponenziale

Si consideri una funzione complessa f della variabile reale t definita per $-\infty < t < +\infty$ e periodica di periodo $T > 0$ (o T -periodica), cioè tale per cui

$$f(t+T) = f(t) \quad , \quad \forall t$$

Se vale questa relazione è anche vero che $f(t+mT) = f(t)$, $\forall t$, per ogni intero m ; si assumerà invece che per nessun reale positivo $T' < T$ risulti $f(t+T') = f(t)$, $\forall t$, a meno che f non sia costante, nel qual caso T è arbitrario. La pulsazione del segnale f è allora definita come

$$\omega_0 = \frac{2\pi}{T}$$

Sotto larghe ipotesi è possibile associare a f i coefficienti complessi

$$F_n = \frac{1}{T} \int_T f(t) e^{-jn\omega_0 t} dt \quad , \quad n = \dots, -2, -1, 0, 1, 2, \dots \quad (\text{B.24})$$

dove il simbolo \int_T significa che l'integrale è esteso a un qualunque intervallo di tempo della lunghezza di un periodo. Questi scalari sono detti *coefficienti di Fourier*, la successione $\{F_n\}$ si dice *spettro di f* e con riferimento alle successioni $\{|F_n|\}$ e $\{\arg(F_n)\}$ si parla di *spettro di ampiezza* e di *spettro di fase*. Essi si rappresentano come mostrato nella Figura B.5 e si dicono *discreti*, nel senso che sono funzioni della variabile intera n , anche se a volte sono visti invece come funzioni della variabile reale $\omega = n\omega_0$, che può comparire sull'asse delle ascisse in luogo di n .

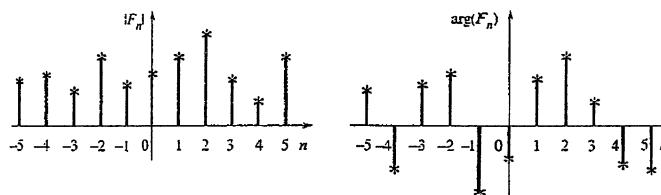


Figura B.5
Spettro di ampiezza
e spettro di fase.

Quanto sopra e i risultati che seguono si estendono in modo naturale a funzioni f vettoriali.

Almeno nei punti in cui f è continua e derivabile, si ha

$$f(t) = \sum_{n=-\infty}^{+\infty} F_n e^{jn\omega_0 t} \quad (\text{B.25})$$

cioè lo spettro di una funzione è in corrispondenza biunivoca con la funzione stessa, fermo restando che due funzioni devono considerarsi uguali pure se differiscono per i valori assunti in un numero, anche infinito, di punti isolati.

Questa serie si dice *serie di Fourier in forma esponenziale*, o semplicemente *serie di Fourier*, associata alla funzione f e ne costituisce una rappresentazione alternativa.

Particolarmente rilevante è il caso in cui f sia reale, caso al quale d'ora in avanti si farà riferimento salvo contrario avviso: in questa situazione infatti

$$F_{-n} = \bar{F}_n \quad , \quad n = 1, 2, \dots$$

per cui risulta

$$f(t) = F_0 + \sum_{n=1}^{+\infty} [F_n e^{jn\omega_0 t} + \bar{F}_n e^{-jn\omega_0 t}] \quad (\text{B.26})$$

È chiaro quindi che i valori di F_n con $n = 0, 1, 2, \dots$ definiscono completamente lo spettro. Di questa proprietà si fa uso per semplificare le rappresentazioni grafiche.

B.4.2 Forma trigonometrica

La forma esponenziale (B.24) e (B.25) esprime l'operazione di *sviluppo in serie di Fourier* nel modo più sintetico e comodo per molte applicazioni. Sotto il punto di vista dell'espressività è però da preferirsi una forma differente. Dalla (B.26) si ha

$$\begin{aligned} f(t) &= F_0 + \sum_{n=1}^{+\infty} F_n [\cos(n\omega_0 t) + j \sin(n\omega_0 t)] + \\ &\quad + \sum_{n=1}^{+\infty} \bar{F}_n [\cos(n\omega_0 t) - j \sin(n\omega_0 t)] = \\ &= F_0 + 2 \sum_{n=1}^{+\infty} [\operatorname{Re}(F_n) \cos(n\omega_0 t) - \operatorname{Im}(F_n) \sin(n\omega_0 t)] \end{aligned}$$

Poi dalla (B.24) si ricava

$$F_0 = \frac{1}{T} \int_T f(t) dt \quad (\text{B.27})$$

e ponendo

$$F_{cn} = 2 \operatorname{Re}(F_n) = \frac{2}{T} \int_T f(t) \cos(n\omega_0 t) dt, \quad n = 1, 2, \dots \quad (\text{B.28})$$

$$F_{sn} = -2 \operatorname{Im}(F_n) = \frac{2}{T} \int_T f(t) \sin(n\omega_0 t) dt, \quad n = 1, 2, \dots \quad (\text{B.29})$$

si ottiene in definitiva

$$f(t) = F_0 + \sum_{n=1}^{+\infty} [F_{cn} \cos(n\omega_0 t) + F_{sn} \sin(n\omega_0 t)] \quad (\text{B.30})$$

Le Equazioni (B.27)-(B.30) esprimono la *serie di Fourier in forma trigonometrica*. La (B.30) si può anche scrivere in modo più compatto come

$$f(t) = F_0 + 2 \sum_{n=1}^{+\infty} |F_n| \cos(n\omega_0 t + \arg F_n) \quad (\text{B.31})$$

La forma trigonometrica (B.27)-(B.30), o (B.24), (B.31), mostra quindi come i segnali *sviluppabili in serie di Fourier* siano scomponibili nella somma di una costante F_0 , chiamata *componente a pulsazione nulla*, e di un'infinità numerabile di componenti cosinusoidali, dette *armoniche*, di pulsazioni multiple di ω_0 : si parla di *armonica fondamentale* per riferirsi proprio a quella di pulsazione ω_0 e di *armonica n-esima* per denotare la generica di pulsazione $n\omega_0$. Allora $|F_n|$ rappresenta il "peso" in f dell'armonica n-esima e $\arg F_n$ la sua fase. La *pulsazione minima* di un segnale è quella corrispondente al minimo valore n_1 di n per cui $F_n \neq 0$; la *pulsazione massima* corrisponde invece all'estremo superiore n_2 dei valori di n per cui $F_n \neq 0$; la *banda* è l'intervallo di pulsazioni compreso tra la minima $n_1\omega_0$ e la massima $n_2\omega_0$; se $n_2 < \infty$ il segnale è *a banda limitata* con *larghezza di banda* $(n_2 - n_1)\omega_0$, altrimenti è *a banda illimitata*. Si usa dire che sviluppare in serie di Fourier un segnale equivale a effettuarne una *analisi armonica*, cioè a rappresentarlo nel dominio della frequenza.

Esempio B.14 Sviluppo in serie dell'onda quadra Si consideri la funzione f , di periodo T , definita da

$$f(t) = \begin{cases} 0, & -T/2 \leq t < -T/4 \\ 1, & -T/4 \leq t < T/4 \\ 0, & T/4 \leq t < T/2 \end{cases}$$

rappresentata nella Figura B.6 e detta onda quadra. Dalla (B.24) si ha

$$F_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-jn\omega_0 t} dt = \frac{1}{T} \int_{-T/4}^{T/4} e^{-jn\omega_0 t} dt$$

per cui

$$F_0 = \frac{1}{2}$$

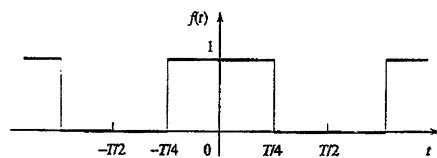


Figura B.6
Onda quadra dell'Esempio B.14.

mentre per $n = \dots, -2, -1, 1, 2, \dots$ ricordando che $\omega_0 = 2\pi/T$,

$$F_n = \frac{1}{T} \frac{e^{-jn\omega_0 t}}{-jn\omega_0} \Big|_{-T/4}^{T/4} = \frac{e^{-jn\pi/2} - e^{jn\pi/2}}{-jn2\pi} = \frac{1}{2} \frac{\sin(n\pi/2)}{n\pi/2}$$

Da qui per $n = 1, 2, \dots$ segue

$$|F_n| = \left| \frac{1}{2} \frac{\sin(n\pi/2)}{n\pi/2} \right| = \begin{cases} 1/n\pi, & n \text{ dispari} \\ 0, & n \text{ pari} \end{cases}$$

$$\arg F_n = \arg \left(\frac{1}{2} \frac{\sin(n\pi/2)}{n\pi/2} \right) = \arg(\sin(n\pi/2)) = \begin{cases} 0, & n = 1, 5, 9, \dots \\ \pi, & n = 3, 7, 11, \dots \end{cases}$$

$$F_{cn} = \frac{\sin(n\pi/2)}{n\pi/2}, \quad F_{sn} = 0$$

Il segnale è pertanto a banda illimitata anche se il peso delle armoniche diminuisce al crescere della pulsazione; il suo spettro è mostrato nella Figura B.7.

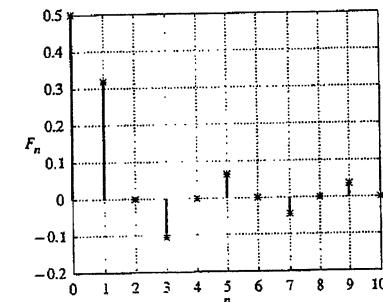


Figura B.7
Spettro dell'onda quadra dell'Esempio B.14.

Esempio B.15 La funzione T-periodica

$$f(t) = \frac{2}{T}t, \quad -T/2 \leq t < T/2$$

è rappresentata nella Figura B.8. Per essa è facile ricavare

$$F_0 = 0, \quad F_n = j \frac{1}{n\pi} \cos(n\pi) \quad n = \dots, -2, -1, 1, 2, \dots$$

da cui

$$F_{cn} = 0, \quad F_{sn} = -\frac{2}{n\pi} \cos(n\pi) \quad n = 1, 2, \dots$$

cosicché anche questo segnale è a banda illimitata.

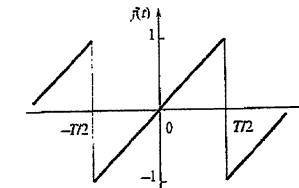


Figura B.8
Segnale dell'Esempio B.15.

Esempio B.16 Sviluppo in serie del treno di impulsi Si prenda ora in esame la funzione T -periodica

$$f(t) = \sum_{k=-\infty}^{+\infty} \text{imp}(t - kT)$$

detto treno di impulsi e rappresentato nella Figura B.9. L'Equazione (B.24) insieme alla (B.3) implica che, qualunque sia n ,

$$\begin{aligned} F_n &= \frac{1}{T} \int_{-T/2}^{T/2} \sum_{k=-\infty}^{+\infty} \text{imp}(t - kT) e^{-jn\omega_0 t} dt = \\ &= \frac{1}{T} \sum_{k=-\infty}^{+\infty} \int_{-T/2}^{T/2} \text{imp}(t - kT) e^{-jn\omega_0 t} dt = \\ &= \frac{1}{T} \int_{-T/2}^{T/2} \text{imp}(t) e^{-jn\omega_0 t} dt = \frac{1}{T} \end{aligned}$$

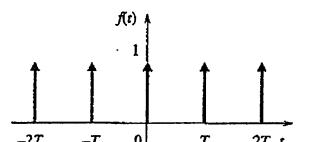
da cui, per $n = 1, 2, \dots$,

$$F_{cn} = 2/T, \quad F_{sn} = 0$$

Quindi si ha

$$f(t) = \sum_{n=-\infty}^{+\infty} \frac{1}{T} e^{jn\omega_0 t} = \frac{1}{T} \left(1 + \sum_{n=1}^{+\infty} 2 \cos(n\omega_0 t) \right)$$

Ancora una volta il segnale è a banda illimitata, in particolare il suo spettro è costante al variare della pulsazione.



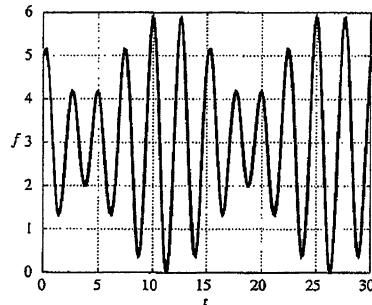
Esempio B.17 Sia ora

$$f(t) = 3 + \sin\left(\frac{2}{3}\pi t\right) + 2 \cos\left(\frac{4}{5}\pi t\right)$$

la funzione periodica, rappresentata nella Figura B.10, con

$$T = 15, \quad \omega_0 = \frac{2}{15}\pi$$

Figura B.9
Treno di impulsi
dell'Esempio B.16.



Le due componenti del segnale ne costituiscono la quinta e la sesta armonica. Pertanto, dalla (B.30) risulta $F_0 = 3$ e poi

$$F_{s5} = 1, \quad F_{s6} = 2$$

da cui, per le (B.28), (B.29),

$$F_5 = -j0.5, \quad F_6 = 1$$

mentre tutti gli altri coefficienti F_n sono nulli. Il segnale ha quindi una banda limitata con pulsazione minima 0 e pulsazione massima $4\pi/5$.

B.4.3 Proprietà principali

Tra le proprietà dell'operazione di sviluppo in serie di Fourier, le seguenti possono risultare particolarmente utili.

Linearità Si abbiano due funzioni complesse f e g con spettri $\{F_n\}$ e $\{G_n\}$. Allora $\forall \alpha \in C, \forall \beta \in C$ lo spettro della funzione $\alpha f + \beta g$ è $\{\alpha F_n + \beta G_n\}$, cioè lo sviluppo in serie di Fourier è un'operazione lineare.

Funzioni pari e dispari Una funzione f pari, cioè tale che $f(-t) = f(t), \forall t$, è sviluppabile in serie di soli coseni: in altri termini il suo spettro è reale, cioè risulta

$$F_{sn} = 0, \quad n = 1, 2, \dots$$

Invece una funzione f dispari, cioè tale che $f(-t) = -f(t), \forall t$, è sviluppabile in serie di soli seni: in altri termini il suo spettro è immaginario, cioè risulta

$$F_{0n} = 0, \quad F_{cn} = 0, \quad n = 1, 2, \dots$$

A illustrazione di quanto sopra si rivedano gli Esempi B.14-B.16.

Uguaglianza di Parseval Data una funzione complessa f con spettro $\{F_n\}$ risulta

$$\frac{1}{T} \int_T |f(t)|^2 dt = \sum_{n=-\infty}^{+\infty} |F_n|^2 \quad (\text{B.32})$$

purché l'integrale al primo membro esista finito.

Questa relazione, detta *uguaglianza di Parseval*, ha un'interessante interpretazione se f è reale. Il primo membro della (B.32) rappresenta la potenza media associata al segnale f e tale quantità, quando è finita, può essere calcolata facendo riferimento alle caratteristiche frequenziali del segnale piuttosto che a quelle temporali. Inoltre, la (B.32) consente di definire il concetto di *banda essenziale* del segnale, che è costituita dal più piccolo intervallo di pulsazioni tra la minima e quella cui corrisponde una data percentuale, per esempio il 95%, della potenza media totale.

B.5 Trasformata di Fourier

B.5.1 Forma esponenziale

Sia data una funzione complessa f della variabile reale tempo t definita per $-\infty < t < +\infty$. Se la funzione complessa della variabile reale ω

$$F(j\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad (\text{B.33})$$

esiste, essa si dice *trasformata di Fourier in forma esponenziale*, o semplicemente *trasformata di Fourier*, di $f(t)$. Per convenzione si denota la trasformata con la stessa lettera, però maiuscola, che indica la funzione del tempo trasformanda e si usa scrivere $F(j\omega) = \mathcal{F}[f(t)]$ e $f(t) = \mathcal{F}^{-1}[F(j\omega)]$. La funzione F si dice anche *spettro di f* e con riferimento alle funzioni $|F(j\omega)|$ e $\arg F(j\omega)$ si parla di *spettro di ampiezza* e di *spettro di fase*. Questi spettri si dicono *continui*, nel senso che sono funzioni di ω che è una variabile reale.

Quanto sopra e i risultati che seguono si estendono in modo naturale a funzioni f vettoriali.

La funzione trasformanda può essere ricavata dalla sua trasformata, sotto larghe ipotesi e almeno nei punti in cui f è continua e derivabile, mediante la *formula di antitrasformazione*

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(j\omega) e^{j\omega t} d\omega \quad (\text{B.34})$$

Lo spettro di una funzione ne costituisce allora una descrizione alternativa e dunque la *trasformazione di Fourier* stabilisce una corrispondenza biunivoca tra $f(t)$ e $\mathcal{F}[f(t)]$, fermo restando che due funzioni devono considerarsi uguali pure se differiscono per i valori assunti in un numero, anche infinito, di punti isolati.

Particolarmente rilevante è il caso in cui f è reale, al quale d'ora in avanti si farà riferimento salvo contrario avviso: in questa situazione infatti è facile dimostrare che

$$F(-j\omega) = \bar{F}(j\omega)$$

per cui risulta

$$f(t) = \frac{1}{2\pi} \int_0^{+\infty} [F(j\omega) e^{j\omega t} + \bar{F}(j\omega) e^{-j\omega t}] d\omega \quad (\text{B.35})$$

È chiaro che in questo caso bastano i valori di $F(j\omega)$ con $\omega \geq 0$ a definire compiutamente lo spettro.

B.5.2 Forma trigonometrica

La forma esponenziale (B.33) e (B.34) esprime l'operazione di trasformazione di Fourier nel modo più sintetico e comodo per molte applicazioni. Sotto il punto di vista dell'espressività è però da preferirsi una forma differente. Dalla (B.35) si ha

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_0^{+\infty} [F(j\omega) (\cos(\omega t) + j \sin(\omega t)) + \\ &\quad + \bar{F}(j\omega) (\cos(\omega t) - j \sin(\omega t))] d\omega = \\ &= \frac{1}{\pi} \int_0^{+\infty} [\operatorname{Re}(F(j\omega)) \cos(\omega t) - \operatorname{Im}(F(j\omega)) \sin(\omega t)] d\omega \end{aligned}$$

Quindi, ponendo

$$F_c(j\omega) = 2 \operatorname{Re}(F(j\omega)) = 2 \int_{-\infty}^{+\infty} f(t) \cos(\omega t) dt \quad (\text{B.36})$$

$$F_s(j\omega) = -2 \operatorname{Im}(F(j\omega)) = 2 \int_{-\infty}^{+\infty} f(t) \sin(\omega t) dt \quad (\text{B.37})$$

si ottiene in definitiva

$$f(t) = \frac{1}{2\pi} \int_0^{+\infty} [F_c(j\omega) \cos(\omega t) + F_s(j\omega) \sin(\omega t)] d\omega \quad (\text{B.38})$$

Le Equazioni (B.36)-(B.38) esprimono la *formula di antitrasformazione in forma trigonometrica*, che si può anche scrivere in modo più compatto come

$$f(t) = \frac{1}{\pi} \int_0^{+\infty} |F(j\omega)| \cos(\omega t + \arg F(j\omega)) d\omega \quad (\text{B.39})$$

La forma trigonometrica di Equazioni (B.36)-(B.38), o (B.33), (B.39), mostra quindi come i segnali *trasformabili con Fourier* siano scomponibili in una infinità non numerabile di componenti cosinusoidali, dette *armoniche*. Si osservi che anche i segnali di durata limitata vengono visti come somme di cosinusoidi che invece hanno durata illimitata. Interpretando l'integrale nella (B.39) come il limite di una sommatoria, si può dire che l'armonica di pulsazione ω ha un "peso" in $f(t)$ pari a $|F(j\omega)| d\omega/\pi$ e una fase pari ad $\arg F(j\omega)$. La *pulsazione minima* di un segnale è quella corrispondente all'estremo inferiore ω_1 di ω per cui $F(j\omega) \neq 0$; la *pulsazione massima* corrisponde invece all'estremo superiore ω_2 dei valori di ω per cui $F(j\omega) \neq 0$; la *banda* è l'intervallo di pulsazioni compreso tra la minima ω_1 e la massima ω_2 ; se $\omega_2 < \infty$ il segnale è a *banda limitata*, con *larghezza di banda* $\omega_2 - \omega_1$, altrimenti è a *banda illimitata*. Si usa dire che trasformare con Fourier un segnale equivale a effettuarne un'*analisi armonica*, cioè a rappresentarlo nel dominio della frequenza. Un segnale trasformabile può contenere armoniche di qualunque pulsazione reale non negativa a differenza di un segnale periodico sviluppabile in serie di Fourier, che può contenere solo armoniche di pulsazioni multiple di quella dell'armonica fondamentale.

Esempio B.18 Trasformata dell'impulso Sia $f(t) = \operatorname{imp}(t)$; allora dalla (B.3) si trova

$$\mathcal{F}[\operatorname{imp}(t)] = \int_{-\infty}^{+\infty} \operatorname{imp}(t) e^{-j\omega t} dt = e^{-j\omega 0} = 1 \quad (\text{B.40})$$

Esempio B.19 Trasformata delle funzioni esponenziali Si consideri dapprima la funzione $f(t) = e^{\sigma t} \operatorname{sca}(t)$ con σ reale negativo. Si ha

$$\begin{aligned} \mathcal{F}[e^{\sigma t} \operatorname{sca}(t)] &= \int_{-\infty}^{+\infty} e^{\sigma t} \operatorname{sca}(t) e^{-j\omega t} dt = \\ &= \int_0^{+\infty} e^{(\sigma-j\omega)t} dt = \frac{e^{(\sigma-j\omega)t}}{\sigma-j\omega} \Big|_0^{+\infty} = \frac{1}{j\omega-\sigma} \end{aligned} \quad (\text{B.41})$$

Se invece si prende $f(t) = e^{-\sigma t} \operatorname{sca}(-t)$, con σ reale negativo, si trova subito

$$\mathcal{F}[e^{-\sigma t} \operatorname{sca}(-t)] = \int_{-\infty}^0 e^{-(\sigma+j\omega)t} dt = \frac{-1}{j\omega+\sigma} \quad (\text{B.42})$$

Questi segnali sono a banda illimitata, anche se il loro spettro di ampiezza decresce al crescere della pulsazione. Le trasformate delle funzioni sopra considerate invece non esistono se σ è reale positivo.

Sia ora $f(t) = e^{j\omega_0 t}$ con ω_0 reale. La sua trasformata di Fourier è

$$\mathcal{F}[e^{j\omega_0 t}] = 2\pi \operatorname{imp}(\omega - \omega_0) \quad (\text{B.43})$$

e la banda del segnale si riduce al punto $\omega = \omega_0$. Infatti

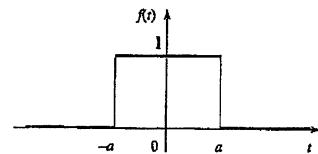
$$\mathcal{F}^{-1}[2\pi \operatorname{imp}(\omega - \omega_0)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} 2\pi \operatorname{imp}(\omega - \omega_0) e^{j\omega t} d\omega = e^{j\omega_0 t}$$

Esempio B.20 Trasformata dell'impulso rettangolare Si consideri l'impulso rettangolare f definito da

$$f(t) = \begin{cases} 1, & -a \leq t < a \\ 0, & \text{altrimenti} \end{cases}$$

dove $a > 0$, che è rappresentato in Figura B.11. Dalla (B.33) si ha

Figura B.11
Impulso rettangolare
dell'Esempio B.20.

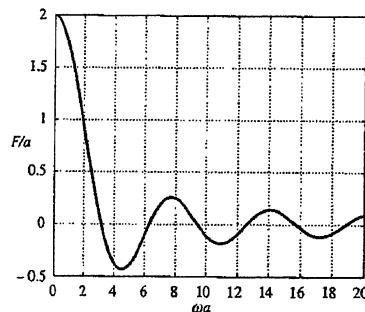


$$F(j0) = 2a$$

$$\begin{aligned} F(j\omega) &= \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt = \int_{-a}^a e^{-j\omega t} dt = \\ &= \frac{e^{-j\omega t}}{-j\omega} \Big|_{-a}^a = 2a \frac{\sin(\omega a)}{\omega a}, \quad \omega \neq 0 \end{aligned}$$

Il segnale è pertanto a banda illimitata; il suo spettro è mostrato nella Figura B.12.

Figura B.12
Spettro dell'impulso
rettangolare dell'Esempio
B.20



B.5.3 Proprietà principali

Tra le proprietà della trasformata di Fourier, le seguenti possono risultare particolarmente utili.

Linearità Si abbiano due funzioni complesse f e g . Allora $\forall \alpha \in C, \forall \beta \in C$ risulta

$$\mathcal{F}[\alpha f(t) + \beta g(t)] = \alpha F(j\omega) + \beta G(j\omega)$$

cioè la trasformazione di Fourier è un'operazione lineare.

Esempio B.21 Trasformata della sinusoide, della cosinusoide e delle funzioni sviluppabili in serie di Fourier Sia $f(t) = \cos(\omega_0 t)$; allora per la (B.43)

$$\begin{aligned} \mathcal{F}[\cos(\omega_0 t)] &= \mathcal{F}[0.5(e^{j\omega_0 t} + e^{-j\omega_0 t})] = \\ &= 0.5(\mathcal{F}[e^{j\omega_0 t}] + \mathcal{F}[e^{-j\omega_0 t}]) = \\ &= 0.5(2\pi \operatorname{imp}(\omega - \omega_0) + 2\pi \operatorname{imp}(\omega + \omega_0)) = \\ &= \pi[\operatorname{imp}(\omega - \omega_0) + \operatorname{imp}(\omega + \omega_0)] \end{aligned}$$

Analogamente si trova

$$\mathcal{F}[\sin(\omega_0 t)] = j\pi(\operatorname{imp}(\omega + \omega_0) - \operatorname{imp}(\omega - \omega_0))$$

Ricordando poi la (B.30), è immediato ricavare le trasformate dei segnali periodici sviluppabili in serie di Fourier. Essenzialmente, lo spettro discreto degli sviluppi in serie diventa uno spettro continuo costituito da impulsi di Dirac.

Esempio B.22 Trasformata della funzione segno La funzione segno è definita dalla relazione

$$\operatorname{sgn}(t) = \begin{cases} \frac{t}{|t|}, & t \neq 0 \\ 0, & t = 0 \end{cases}$$

e gode della proprietà

$$\operatorname{sgn}(t) = \lim_{\sigma \rightarrow 0^-} (e^{\sigma t} \operatorname{sca}(t) - e^{-\sigma t} \operatorname{sca}(-t)), \quad t \neq 0$$

Si assume ora che l'operazione di passaggio al limite si possa invertire con quella di trasformazione di Fourier. Tale ipotesi, non corretta in generale, consente in questo caso di pervenire al risultato esatto. Allora, ricordando le (B.41), (B.42) e notando che la trasformata non dipende dal valore della funzione in un singolo punto, si ha

$$\begin{aligned} \mathcal{F}[\operatorname{sgn}(t)] &= \lim_{\sigma \rightarrow 0^-} (\mathcal{F}[e^{\sigma t} \operatorname{sca}(t)] - \mathcal{F}[e^{-\sigma t} \operatorname{sca}(-t)]) = \\ &= \lim_{\sigma \rightarrow 0^-} \left(\frac{1}{j\omega - \sigma} + \frac{1}{j\omega + \sigma} \right) = \frac{2}{j\omega} \end{aligned} \quad (\text{B.44})$$

Dualità Data una funzione complessa f , se

$$F(j\omega) = \mathcal{F}[f(t)]$$

allora

$$f(-\omega) = \frac{1}{2\pi} \mathcal{F}[F(t)]$$

Esempio B.23 Trasformata della costante Sia $f(t) = 1$. Ricordando la trasformata dell'impulso (B.40), per la proprietà di dualità si ha

$$\mathcal{F}[1] = 2\pi \operatorname{imp}(-\omega) = 2\pi \operatorname{imp}(\omega) \quad (\text{B.45})$$

Esempio B.24 Trasformata dello scalino Visto che

$$\operatorname{sgn}(t) = 0.5(1 + \operatorname{sgn}(t)), \quad t \neq 0$$

e che la trasformata non dipende dal valore della funzione in un singolo punto, per linearità le Equazioni (B.44), (B.45) producono

$$\mathcal{F}[\operatorname{sca}(t)] = \pi \operatorname{imp}(\omega) + \frac{1}{j\omega}$$

Funzioni pari e dispari Lo spettro di una funzione f pari è reale, cioè risulta

$$F_p(j\omega) = 0 \quad , \quad \omega \geq 0$$

Invece lo spettro di una funzione f dispari è immaginario, cioè risulta

$$F_d(j\omega) = 0 \quad , \quad \omega \geq 0$$

A illustrazione di quanto sopra si rivedano gli Esempi B.18, B.20-B.23.

Uguaglianza di Parseval Data una funzione complessa f con spettro F risulta

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |F(j\omega)|^2 d\omega \quad (B.46)$$

purché l'integrale al primo membro esista finito.

Questa relazione, detta *uguaglianza di Parseval*, ha un'interessante interpretazione se f è reale. Il primo membro della (B.46) rappresenta l'energia associata al segnale f e tale quantità, quando è finita, può essere calcolata facendo riferimento alle caratteristiche frequenziali del segnale piuttosto che a quelle temporali.

La relazione (B.46) consente di definire il concetto di *banda essenziale* del segnale che è costituita dall'intervallo di pulsazioni tra la minima e quella cui corrisponde una data percentuale, per esempio il 95%, dell'energia totale.

La Tabella B.2 raccoglie alcune delle trasformate di Fourier sopra ricavate insieme ad altre rilevanti.

B.5.4 Relazioni con la trasformata di Laplace

Confrontando le definizioni (B.4) e (B.33) si notano subito le strette relazioni esistenti tra la trasformata di Laplace e quella di Fourier, per la quale, in aggiunta a quelle appena descritte, valgono parecchie proprietà molto simili a quelle presentate per la trasformata di Laplace. Conviene però osservare anche le differenze tra i due strumenti matematici.

Si considerino le funzioni f nulle per $t < 0$. Allora, se la trasformata di Laplace esiste e ha ascissa di convergenza $\sigma < 0$, vale la relazione

$$\mathcal{F}[f(t)] = \mathcal{L}[f(t)]|_{s=j\omega}$$

In altri termini, esiste anche la trasformata di Fourier che si ottiene subito ponendo $s = j\omega$ in quella di Laplace. Ciò giustifica l'uso di denotare con $F(s)$ e $F(j\omega)$ le due trasformate, laddove quella di Fourier si sarebbe a buon diritto dovuta indicare come funzione di ω piuttosto che di $j\omega$.

Dalle Tabelle B.1 e B.2 si confrontino, per esempio, le trasformate dell'impulso di Dirac e della funzione $e^{\sigma t} \text{sca}(t)$, σ reale, $\sigma < 0$, ricordando anche le corrispondenti ascisse di convergenza ricavate negli Esempi B.1 e B.4.

Quando invece $\sigma \geq 0$, l'esistenza della trasformata di Laplace non implica l'esistenza di quella di Fourier.

Ancora dall'Esempio B.4 si ricordi che la trasformata di Laplace di $e^{\sigma t} \text{sca}(t)$, $\sigma > 0$, ha ascissa di convergenza $\bar{\sigma} = \sigma$, mentre nell'Esempio B.19 si è affermato che la trasformata di Fourier non esiste.

Infatti è chiaro che

$$\mathcal{L}[f(t)] = \mathcal{F}[f(t)e^{-\sigma t}]$$

Tabella B.2
Segnali e corrispondenti trasformate di Fourier.

$f(t)$	$F(j\omega)$
$\text{imp}(t)$	1
1	$2\pi \text{imp}(\omega)$
$e^{j\omega_0 t}$	$2\pi \text{imp}(\omega - \omega_0)$
$\sin(\omega_0 t)$	$j\pi [\text{imp}(\omega + \omega_0) - \text{imp}(\omega - \omega_0)]$
$\cos(\omega_0 t)$	$\pi [\text{imp}(\omega - \omega_0) + \text{imp}(\omega + \omega_0)]$
$\text{sgn}(t)$	$\frac{2}{j\omega}$
$\text{sca}(t)$	$\pi \text{imp}(\omega) + \frac{1}{j\omega}$
$e^{\sigma t} \text{sca}(t)$, $\sigma < 0$	$\frac{1}{j\omega - \sigma}$
$e^{-\sigma t} \text{sca}(-t)$, $\sigma < 0$	$-\frac{1}{j\omega + \sigma}$
$\sin(\omega_0 t) \text{sca}(t)$	$\frac{\pi}{2j} [\text{imp}(\omega - \omega_0) - \text{imp}(\omega + \omega_0)] + \frac{\omega_0}{\omega_0^2 - \omega^2}$
$\cos(\omega_0 t) \text{sca}(t)$	$\frac{\pi}{2} [(\text{imp}(\omega - \omega_0) + \text{imp}(\omega + \omega_0))] + \frac{j\omega}{\omega_0^2 - \omega^2}$
$e^{\sigma t} \sin(\omega_0 t) \text{sca}(t)$, $\sigma < 0$	$\frac{\omega_0}{(j\omega - \sigma)^2 + \omega_0^2}$
$e^{\sigma t} \cos(\omega_0 t) \text{sca}(t)$, $\sigma < 0$	$\frac{j\omega - \sigma}{(j\omega - \sigma)^2 + \omega_0^2}$

cioè la trasformata di Laplace coincide con la trasformata di Fourier della funzione moltiplicata per un'esponenziale che si può scegliere come decrescente, limitando se è necessario la regione del piano complesso nel quale la trasformata di Laplace è definita. Per questo motivo, funzioni che per $t \rightarrow \infty$ tendono all'infinito, ma "non troppo rapidamente", cioè non più velocemente di un'esponenziale, possono essere trasformabili con Laplace ma non con Fourier.

Tuttavia, anche quando risulta $\bar{\sigma} \geq 0$ la trasformata di Fourier può esistere ugualmente; essa però non si ottiene semplicemente ponendo $s = j\omega$ in quella di Laplace. Il fatto che la prima trasformata sia definita sull'asse immaginario implica che essa sia formalmente più complicata della seconda, che invece non lo è, tipicamente per la presenza di funzioni impulsive.

Si confrontino nelle Tabelle B.1 e B.2 le trasformate delle funzioni seno e coseno (nulle per $t < 0$), tenendo conto che l'ascissa di convergenza indicata negli Esempi B.5 e B.6 è $\bar{\sigma} = 0$. Si osservi pure che se nelle trasformate di Fourier si pone $\omega = s/j$ e le si valuta per $\text{Re}(s) > \bar{\sigma}$ si ottengono proprio le trasformate di Laplace.

Se, per quanto sopra, l'operazione di trasformazione di Laplace appare più potente di quella di Fourier, non bisogna però sottovalutare i due fatti seguenti, che in molti casi fanno preferire la seconda trasformata alla prima.

Innanzitutto con la trasformata di Fourier è possibile trattare segnali definiti e non identicamente nulli anche per $t < 0$, cosa per la verità fattibile pure con un'estensione, detta *trasformata bilatera*, della trasformata di Laplace (*unilatera*) qui descritta.

Soprattutto, poi, la trasformata di Fourier, nella sua forma trigonometrica, consente di interpretare le funzioni di una vasta classe come costituite dalla somma di un'infinità non numerabile di armoniche, estendendo così ai segnali non periodici l'importante risultato conseguito con lo sviluppo dei segnali periodici in serie di Fourier, cioè nella somma di un'infinità numerabile di armoniche. Tali interpretazioni risultano utilissime da un punto di vista ingegneristico. La trasformata di Laplace permetterebbe solo di sviluppare i segnali nella somma di un'infinità non numerabile di armoniche moltiplicate per esponenziali, un fatto questo che risulta molto meno trasparente.

Pertanto le due trasformate devono ritenersi assolutamente complementari.

Segnali a tempo discreto



C.1 Introduzione

L'obiettivo di questa appendice consiste nel presentare gli strumenti più comuni per l'analisi nel dominio della variabile complessa ω , come pure si dice, nel dominio della frequenza dei segnali a tempo discreto, cioè delle funzioni della variabile intera k che rappresenta il tempo. Questi strumenti sono costituiti dalla trasformata Zeta, dallo sviluppo di Fourier e dalla trasformata di Fourier discreti. Rinunciando a volte al rigore formale e alla completezza, ci si limiterà a discutere le questioni fondamentali relative al tema indicato in modo semplice e intuitivo, trattando specificatamente:

- i cosiddetti segnali canonici discreti e in particolare l'impulso;
- la definizione e le proprietà della trasformata Zeta, con particolare riferimento ai metodi di antitrasformazione che utilizzano lo sviluppo di Heaviside e la lunga divisione;
- la definizione, il significato e le proprietà dello sviluppo di Fourier discreto;
- la definizione, il significato e le proprietà della trasformata di Fourier discreta, nonché le sue relazioni con la trasformata Zeta.

C.2 Impulso e altri segnali canonici discreti

Nello studio dei sistemi dinamici a tempo discreto, in particolare di quelli lineari e stazionari, si fa spesso riferimento alle seguenti funzioni appartenenti alla classe dei *segnali canonici discreti*:

$$\text{sca}^*(k) = \begin{cases} 0, & k < 0 \\ 1, & k \geq 0 \end{cases}$$

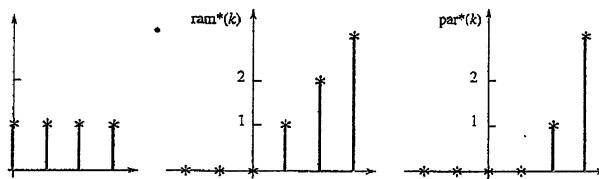
$$\text{ram}^*(k) = \begin{cases} 0, & k < 0 \\ k, & k \geq 0 \end{cases}$$

$$\text{par}^*(k) = \begin{cases} 0, & k < 0 \\ k(k-1)/2, & k \geq 0 \end{cases}$$

Essi sono detti rispettivamente *scalino*, *rampa* e *parabola* e sono rappresentati nella Figura C.1. È evidente che

$$\sum_{i=-\infty}^{k-1} \text{sca}^*(i) = \text{ram}^*(k) \quad , \quad \sum_{i=-\infty}^{k-1} \text{ram}^*(i) = \text{par}^*(k)$$

Figura C.1
Principali segnali canonici a tempo discreto:
a) scalino; b) rampa;
c) parabola.



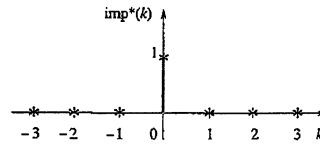
Altri segnali canonici si ottengono per successive iterazioni di queste formule. Si definisce poi un *impulso discreto*, o semplicemente *impulso*, o anche *simbolo di Kronecker*, come

$$\text{imp}^*(k) = \begin{cases} 0, & k \neq 0 \\ 1, & k = 0 \end{cases} \quad (\text{C.1})$$

Esso è rappresentato nella Figura C.2 e, a differenza dell'impulso continuo di Dirac, è una funzione definita in senso classico. Si può inoltre scrivere

$$\sum_{i=-\infty}^k \text{imp}^*(i) = \text{sca}^*(k)$$

Figura C.2
Impulso.



C.3 Trasformata Zeta

C.3.1 Generalità

Sia data una funzione complessa f della variabile intera k ; sia poi $z = \rho e^{j\theta} \in C$ una variabile complessa con modulo $\rho \geq 0$ e fase θ . Se la funzione

$$F(z) = \sum_{k=0}^{+\infty} f(k) z^{-k} \quad (\text{C.2})$$

esiste almeno per qualche valore di z , essa si dice *trasformata Zeta* di $f(k)$. Di solito si denota la trasformata con la stessa lettera, però maiuscola, che indica la funzione del tempo trasformanda e si usa scrivere $F(z) = \mathcal{Z}[f(k)]$ e $f(k) = \mathcal{Z}^{-1}[F(z)]$. La definizione (C.2) richiede che $f(k)$ sia definita almeno per $k \geq 0$; essa può essere usata anche per funzioni definite per $k < 0$, ma i corrispondenti valori di f non concorrono comunque a determinare la trasformata: tipicamente, ma non necessariamente, si assume $f(k) = 0$ per $k < 0$.

Quanto sopra e i risultati che seguono si estendono in modo naturale a funzioni f vettoriali.

L'operazione di *trasformazione Zeta* ha quindi successo quando esistono valori di z tali che la serie nell'Equazione (C.2) converga. Sia $\bar{\rho} > 0$ l'estremo inferiore dei moduli di tali valori di z ; si dimostra che la trasformata esiste all'esterno di un cerchio definito dalla condizione $\rho > \bar{\rho}$; per questo motivo $\bar{\rho}$ si dice

raggio di convergenza. La trasformata Zeta risulta essere una *funzione analitica* e, grazie alle particolari proprietà delle funzioni analitiche, la sua definizione può essere estesa anche in punti z tali che $\rho \leq \bar{\rho}$, indipendentemente dal fatto che la serie nella (C.2) non converga. La problematica della convergenza comunque non verrà qui approfondita e, anzi, si assumerà tacitamente che esistano tutte le trasformate di cui si avrà occasione di discutere.

Di particolare importanza sono le *trasformate razionali*, cioè quelle in cui

$$F(z) = \frac{N(z)}{D(z)}$$

con $N(z)$ e $D(z)$ polinomi primi tra loro. Le radici dell'equazione $N(z) = 0$ si dicono *zeri* e quelle dell'equazione $D(z) = 0$ si dicono *poli*: nell'insieme poli e zeri si chiamano *singolarità*. Quando f è reale i coefficienti dei polinomi N e D sono reali e il raggio di convergenza è pari al massimo tra i moduli dei poli; è chiaro peraltro che $F(z)$ è ben definita per tutti i valori di z diversi dai poli.

La funzione trasformanda può essere ricavata dalla sua trasformata mediante la *formula di antitrasformazione*

$$f(k) = \frac{1}{2\pi j} \oint F(z) z^{k-1} dz \quad (\text{C.3})$$

dove l'integrazione deve essere effettuata lungo una circonferenza centrata nell'origine, di raggio maggiore di $\bar{\rho}$ e percorsa in senso antiorario; per precisione occorre dire che $f(k)$ è fornita per $k \geq 0$, perché solo tale parte di f contribuisce a determinare F ; la (C.3) dà invece $f(k) = 0$ per $k < 0$. Quindi, limitandosi a considerare funzioni nulle per tempi negativi, la corrispondenza tra funzioni del tempo trasformabili Zeta e trasformate è biunivoca.

L'Equazione (C.3) è molto importante da un punto di vista concettuale perché insieme alla (C.2) afferma l'esistenza di una corrispondenza biunivoca tra una funzione del tempo trasformabile $f(k)$, $k \geq 0$, e la sua trasformata funzione di variabile complessa $F(z)$: in altri termini, queste due funzioni hanno lo stesso contenuto informativo, cioè non sono altro che due modi diversi di rappresentare la stessa entità. La formula (C.3) è invece poco usata per effettuare l'antitrasformazione in pratica; per risolvere questo problema nel seguito del paragrafo saranno presentati metodi molto semplici utilizzabili per le funzioni razionali.

Esempio C.1 Trasformata dell'impulso Sia $f(k) = \text{imp}^*(k)$; allora, qualunque sia z , per l'Equazione (C.1) si ottiene

$$\mathcal{Z}[\text{imp}^*(k)] = \sum_{k=0}^{+\infty} \text{imp}^*(k) z^{-k} = 1$$

Esempio C.2 Trasformata dello scalino Sia $f(k) = \text{sca}^*(k)$. In questo caso si ha

$$\mathcal{Z}[\text{sca}^*(k)] = \sum_{k=0}^{+\infty} z^{-k}$$

e, per le proprietà della serie geometrica, per $|z| > 1$,

$$\mathcal{Z}[\text{sca}^*(k)] = \frac{z}{z-1}$$

Pertanto risulta $\bar{\rho} = 1$, ma la funzione $z/(z-1)$ è definita per ogni $z \neq 1$.

C.3.2 Proprietà principali

La trasformazione Zeta gode di parecchie proprietà. Le più importanti per le applicazioni sono riportate nel seguito, senza le spese semplici dimostrazioni, e sono usate per ricavare alcune utili trasformate evitando di ricorrere alla definizione (C.2).

Linearità Si abbiano due funzioni f e g . Allora $\forall \alpha \in C, \forall \beta \in C$ risulta

$$\mathcal{Z}[\alpha f(k) + \beta g(k)] = \alpha F(z) + \beta G(z)$$

cioè la trasformazione Zeta è un'operazione lineare.

Ritardo Si consideri la funzione $\hat{f}(k) = f(k-1)$ ottenuta traslando in avanti di un istante di tempo la funzione $f(k)$, supposta nulla per tempi negativi. Si trova

$$\mathcal{Z}[\hat{f}(k)] = \mathcal{Z}[f(k-1)] = \frac{1}{z}F(z) \quad (\text{C.4})$$

cosicché dividere per z nel dominio della variabile complessa equivale a ritardare di un passo nel dominio del tempo e $1/z$ si può interpretare come *operatore di ritardo unitario*. La generalizzazione della formula al caso di ritardi maggiori è ovvia.

Anticipo Si consideri la funzione $\hat{f}(k) = f(k+1)$ ottenuta traslando indietro di un istante di tempo la funzione $f(k)$. Si trova

$$\mathcal{Z}[\hat{f}(k)] = \mathcal{Z}[f(k+1)] = zF(z) - zf(0)$$

Quindi, se la funzione è nulla in $k = 0$, moltiplicare per z nel dominio della variabile complessa equivale ad anticipare di un passo nel dominio del tempo, cioè z si può interpretare come *operatore di anticipo unitario*. La generalizzazione della formula al caso in cui $\hat{f}(k) = f(k+m)$, $m > 1$, è

$$\mathcal{Z}[\hat{f}(k)] = \mathcal{Z}[f(k+m)] = z^m F(z) - z^m \sum_{k=0}^{m-1} f(k) z^{-k}$$

Prodotto per un'esponenziale Per un qualunque $\alpha \in C$ si consideri la funzione $\alpha^k f(k)$. La sua trasformata è

$$\mathcal{Z}[\alpha^k f(k)] = F\left(\frac{z}{\alpha}\right) \quad (\text{C.5})$$

Esempio C.3 Trasformata dell'esponenziale Ricordando la trasformata dello scalino con il suo raggio di convergenza e ponendo $f(k) = \text{sca}^*(k)$ nell'Equazione (C.5), si trova che la trasformata Zeta della funzione $\alpha^k \text{sca}^*(k)$, qualunque sia $\alpha \in C$, è

$$\mathcal{Z}[\alpha^k \text{sca}^*(k)] = \frac{z}{z-\alpha}$$

cui corrisponde $\tilde{\rho} = |\alpha|$, pur esistendo essa per $z \neq \alpha$.

Derivazione nel dominio della variabile complessa Si supponga che la funzione $F(z)$ sia derivabile per tutti gli z , con l'eventuale esclusione di un numero finito di valori. Risulta

$$\mathcal{Z}[kf(k)] = -z \frac{dF(z)}{dz}$$

formula ovviamente estendibile al calcolo della trasformata di $k^n f(k)$.

Esempio C.4 Trasformata della rampa Si noti che $\text{ram}^*(k) = k \text{sca}^*(k)$ e allora

$$\mathcal{Z}[\text{ram}^*(k)] = -z \frac{d\left(\frac{z}{z-1}\right)}{dz} = \frac{z}{(z-1)^2}$$

Esempio C.5 Trasformata della parabola Si osservi innanzitutto che

$$\text{par}^*(k) = 0.5[k \text{ram}^*(k) - \text{ram}^*(k)]$$

e che

$$\mathcal{Z}[k \text{ram}^*(k)] = -z \frac{d\left(\frac{z}{(z-1)^2}\right)}{dz} = \frac{z(z+1)}{(z-1)^3}$$

Allora si trova

$$\mathcal{Z}[\text{par}^*(k)] = \frac{z}{(z-1)^3}$$

Convoluzione nel dominio del tempo Date due funzioni f_1 e f_2 il loro *prodotto di convoluzione* è

$$f_1(k) * f_2(k) = \sum_{i=-\infty}^{+\infty} f_1(i) f_2(k-i) = \sum_{l=-\infty}^{+\infty} f_1(k-l) f_2(l) = f_2(k) * f_1(k)$$

Se esse sono nulle per tempi negativi, allora

$$f_1(k) * f_2(k) = \sum_{i=0}^{+\infty} f_1(i) f_2(k-i) = \sum_{l=0}^{+\infty} f_1(k-l) f_2(l) = f_2(k) * f_1(k)$$

e si trova

$$\mathcal{Z}[f_1(k) * f_2(k)] = F_1(z) F_2(z)$$

Esempio C.6 Per le funzioni $f_1(k) = \alpha^k \text{sca}^*(k)$ e $f_2(k) = \text{sca}^*(k)$ si ha

$$\mathcal{Z}[f_1(k) * f_2(k)] = \frac{z^2}{(z-\alpha)(z-1)}$$

Teorema del valore iniziale Il valore iniziale di una funzione f è

$$f(0) = \lim_{z \rightarrow \infty} F(z)$$

Teorema del valore finale Se una funzione reale f ha trasformata razionale F con poli in $z = 1$ o con modulo minore di 1, allora

$$\lim_{k \rightarrow \infty} f(k) = \lim_{z \rightarrow 1} (z - 1) F(z) \quad (C.6)$$

L'Equazione (C.6) vale sotto larghe ipotesi anche se F non è razionale purché $\lim_{k \rightarrow \infty} f(k)$ esista.

Esempio C.7 Data la funzione $a^k \text{sca}^*(k)$, $a \in \mathbb{R}$, in base al teorema del valore iniziale si trova

$$\left[a^k \text{sca}^*(k) \right]_{k=0} = \lim_{z \rightarrow \infty} \frac{z}{z - a} = 1$$

e in base al teorema del valore finale, applicabile per $-1 < a \leq 1$,

$$\lim_{k \rightarrow \infty} a^k \text{sca}^*(k) = \lim_{z \rightarrow 1} (z - 1) \frac{z}{z - a} = \begin{cases} 0, & -1 < a < 1 \\ 1, & a = 1 \end{cases}$$

La Tabella C.1 raccoglie le trasformate Zeta ricavate fino a questo punto insieme ad alcune altre la cui determinazione è lasciata al lettore.

Tabella C.1
Segnali e corrispondenti trasformate Zeta.

$f(k)$	$F(z)$
$\text{imp}^*(k)$	1
$\text{sca}^*(k)$	$\frac{z}{z - 1}$
$\text{ram}^*(k)$	$\frac{z}{(z - 1)^2}$
$\text{par}^*(k)$	$\frac{z}{(z - 1)^3}$
$a^k \text{sca}^*(k)$	$\frac{z}{z - a}$
$a^k \text{ram}^*(k)$	$\frac{az}{(z - a)^2}$
$\sin(\theta k) \text{sca}^*(k)$	$\frac{z \sin(\theta)}{z^2 - 2 \cos(\theta) z + 1}$
$\cos(\theta k) \text{sca}^*(k)$	$\frac{z [z - \cos(\theta)]}{z^2 - 2 \cos(\theta) z + 1}$

C.3.3 Sviluppo di Heaviside, lunga divisione e antitrasformazione di trasformate razionali

Si vuole ora mostrare che è possibile antitrasformare una trasformata razionale mediante una tecnica semplice che non fa ricorso alla formula (C.3). Per mezzo del classico algoritmo per la divisione dei polinomi, una qualsiasi funzione razionale si può scrivere come somma di un'altra funzione razionale con grado

del denominatore maggiore o uguale al grado del numeratore e un polinomio di grado maggiore di zero. Per linearità l'antitrasformata della funzione è la somma delle antitrasformate dei due addendi. Si può però dimostrare che un qualunque polinomio di grado maggiore di zero non può essere una trasformata Zeta. Pertanto, all'interno della problematica in oggetto, l'unico caso interessante è quello in cui il denominatore della trasformata ha grado maggiore o uguale a quello del numeratore.

Nel seguito si affronterà perciò il problema dell'antitrasformazione della funzione razionale

$$F(z) = \frac{N(z)}{D(z)} \quad (C.7)$$

dove il grado di $D(z)$ non è minore del grado di $N(z)$. Si assumerà anche che i coefficienti dei polinomi N e D siano reali, cosicché per ogni zero e polo non reale esisterà il complesso coniugato e la funzione antitrasformata $f(k)$ sarà reale.

In termini generali, l'idea alla base del metodo che si vuole presentare consiste nello sviluppare $F(z)$ nella somma di funzioni la cui antitrasformata è nota e, per la proprietà di linearità, ottenere l'antitrasformata come somma delle antitrasformate dei singoli addendi. La tecnica da seguire è pertanto del tutto analoga a quella illustrata a proposito dell'antitrasformazione delle trasformate di Laplace al Paragrafo B.3.3. Tuttavia, la Tabella C.1 mostra come le trasformate di interesse abbiano tutte un fattore z al numeratore. Allora, per fare in modo che i singoli termini dello sviluppo assumano proprio tale forma, conviene applicare il classico sviluppo di Heaviside alla funzione $F(z)/z$ invece che alla funzione $F(z)$. In alternativa, si può procedere in maniera tradizionale e poi antitrasformare ricordando l'Equazione (C.4). La procedura suggerita sarà illustrata con esclusivo riferimento al caso in cui i poli di $F(z)$ siano tutti distinti tra loro e non nulli. Se si pone

$$zD(z) = \prod_{i=0}^n (z + p_i), \quad p_h \neq p_j \text{ per } h \neq j$$

dove $p_i \in \mathbb{C}$ e $p_0 = 0$, dall'Equazione (C.7) si possono trovare i coefficienti $P_i \in \mathbb{C}$ in modo che sia soddisfatta l'identità

$$\frac{N(z)}{\prod_{i=0}^n (z + p_i)} = \sum_{i=0}^n \frac{P_i}{z + p_i}$$

Infatti, moltiplicando per $z + p_i$ si ottiene

$$(z + p_i) \frac{F(z)}{z} = \frac{N(z)}{\prod_{j=0}^n (z + p_j)} \equiv (z + p_i) \sum_{j=0}^n \frac{P_j}{z + p_j} + P_i$$

e valutando per $z = -p_i$ risulta

$$P_i = \frac{N(-p_i)}{\prod_{j=0}^n (p_j - p_i)} = \frac{N(-p_i)}{\left. \frac{d(zD(z))}{dz} \right|_{z=-p_i}} = \left[(z + p_i) \frac{F(z)}{z} \right]_{z=-p_i} \quad (C.8)$$

Ricavati così i coefficienti P_i , detti *residui* di $F(z)/z$ nei poli $-p_i$, ricordando la Tabella C.1 e assumendo nulla l'antitrasformata per $k < 0$, si ottiene innanzitutto

$$\frac{F(z)}{z} = \sum_{i=0}^n \frac{P_i}{z + p_i}$$

e quindi in definitiva •

$$\begin{aligned} f(k) &= \mathcal{Z}^{-1}[F(z)] = \mathcal{Z}^{-1}\left[P_0 + \sum_{i=1}^n \frac{P_i z}{z + p_i}\right] = \\ &= P_0 \text{imp}^*(k) + \left[\sum_{i=1}^n P_i (-p_i)^k\right] \text{sca}^*(k) \end{aligned} \quad (\text{C.9})$$

Esempio C.8 Se

$$F(z) = \frac{z - 10}{(z + 2)(z + 5)} \quad (\text{C.10})$$

si può scrivere

$$\frac{z - 10}{z(z + 2)(z + 5)} \equiv \frac{P_0}{z} + \frac{P_1}{z + 2} + \frac{P_2}{z + 5}$$

dove per la (C.8)

$$P_0 = \frac{-10}{2 \cdot 5} = -1 \quad , \quad P_1 = \frac{-2 - 10}{-2(-2 + 5)} = 2 \quad , \quad P_2 = \frac{-5 - 10}{-5(-5 + 2)} = -1$$

Si ottiene così per la (C.9)

$$f(k) = \mathcal{Z}^{-1}\left[\frac{z - 10}{(z + 2)(z + 5)}\right] = -\text{imp}^*(k) + [2(-2)^k - (-5)^k] \text{sca}^*(k) \quad (\text{C.11})$$

Se invece si ponesse

$$\frac{z - 10}{(z + 2)(z + 5)} = \frac{-4}{z + 2} + \frac{5}{z + 5}$$

dalla Tabella C.1 e ricordando l'Equazione (C.4) si otterebbe

$$f(k) = \mathcal{Z}^{-1}\left[\frac{z - 10}{(z + 2)(z + 5)}\right] = [-4(-2)^{k-1} + 5(-5)^{k-1}] \text{sca}^*(k - 1) \quad (\text{C.12})$$

Palesemente le Equazioni (C.11) e (C.12) descrivono la medesima funzione.

La trattazione delle particolarità connesse con l'esistenza di poli di F non reali e/o multipli, oltre che nell'origine, si può condurre in analogia con quanto fatto al Paragrafo B.3.3 a proposito dell'anttrasformazione delle trasformate di Laplace.

Se invece si rinuncia a ottenere l'anttrasformata di $F(z)$ in forma chiusa e ci si accontenta di calcolare i valori assunti dalla funzione $f(k)$ nei singoli istanti di tempo, si può applicare il metodo della *tunga divisione*. Esso consiste nell'effettuare la divisione, con il classico algoritmo euclideo, tra il numeratore N e il denominatore D di F , scrivendo quindi tale funzione nella forma (C.2), cosicché il coefficiente di z^{-k} assume il significato di valore della funzione f al tempo k . La procedura è illustrata di seguito con un esempio.

Esempio C.9 (Seguito dell'Esempio C.8) Dividendo il numeratore per il denominatore della funzione (C.10), si ottiene

$$F(z) = z^{-1} - 17z^{-2} + 109z^{-3} - 593z^{-4} + \dots$$

e pertanto

$$f(0) = 0, \quad f(1) = 1, \quad f(2) = -17, \quad f(3) = 109, \quad f(4) = -593$$

C.4 Sviluppo di Fourier discreto

C.4.1 Forma esponenziale

Si consideri una funzione complessa f della variabile intera k definita per $-\infty < k < +\infty$ e periodica di periodo $N > 0$ (o N -periodica), cioè tale per cui

$$f(k + N) = f(k) \quad , \quad \forall k$$

Se vale tale relazione è anche vero che $f(k + mN) = f(k)$, $\forall k$, per ogni intero m ; si assumerà invece che per nessun intero positivo $N' < N$ risulti $f(k + N') = f(k)$, $\forall k$, a meno che f non sia costante, nel qual caso N è arbitrario. La pulsazione del segnale f è allora definita come

$$\theta_0 = \frac{2\pi}{N} \quad (\text{C.13})$$

È possibile associare a f i coefficienti complessi

$$F_n = \frac{1}{N} \sum_N f(k) e^{-jn\theta_0 k}, \quad n = \dots, -2, -1, 0, 1, 2, \dots \quad (\text{C.14})$$

dove il simbolo \sum_N significa che la somma è estesa a un qualunque intervallo di tempo della lunghezza di un periodo. Questi scalari sono detti *coefficienti di Fourier*, la successione $\{F_n\}$ si dice *spettro di f* e con riferimento alle successioni $\{|F_n|\}$ e $\{\arg F_n\}$ si parla di *spettro di ampiezza* e di *spettro di fase*. Essi si dicono *discreti*, nel senso che sono funzioni della variabile intera n , anche se a volte sono visti invece come funzioni della variabile reale $\theta = n\theta_0$. È facile verificare che lo spettro è una funzione periodica di periodo N , per cui la sua conoscenza su un qualunque intervallo di valori di n di lunghezza N equivale alla conoscenza dello spettro intero. Questa proprietà deriva dal fatto che, per la (C.13),

$$e^{-jn\theta_0 k} = e^{-j(n+N)\theta_0 k}$$

Quanto sopra e i risultati che seguono si estendono in modo naturale a funzioni f vettoriali.

Si ha poi

$$f(k) = \sum_N F_n e^{jn\theta_0 k} \quad (\text{C.15})$$

dove la somma è estesa a un qualunque intervallo di valori di n di lunghezza N , cioè lo spettro di una funzione è in corrispondenza biunivoca con la funzione stessa.

La somma (C.15) si dice *sviluppo di Fourier discreto in forma esponenziale*, o semplicemente *sviluppo di Fourier discreto*, associato alla funzione f e ne costituisce una rappresentazione alternativa. Si usa tralasciare l'aggettivo "discreto" quando ciò non possa dar luogo a equivoci. A volte, con terminologia leggermente impropria, si parla anche di *DFT* (dall'inglese *Discrete Fourier Transform*).

Tra le proprietà dell'operazione di sviluppo di Fourier discreto, notevole è quella di linearità. Essa afferma che, date due funzioni complesse f e g con spettri $\{F_n\}$ e $\{G_n\}$, lo spettro della funzione $\alpha f + \beta g$ è $\{\alpha F_n + \beta G_n\}$, $\forall \alpha \in C$, $\forall \beta \in C$.

Particolarmente rilevante è il caso in cui f sia reale: in questa situazione infatti

$$F_{-n} = F_n, \quad n = 1, 2, \dots$$

ed è allora chiaro che per definire compiutamente lo spettro (N -periodico) bastano i valori di F_n per $n = 0, 1, 2, \dots, \hat{N} - 1$, dove \hat{N} è il più piccolo intero maggiore di $N/2$. Della periodicità e di quest'ultima proprietà si fa ovviamente uso per semplificare le rappresentazioni grafiche, del tutto simili a quelle usate per gli spettri di segnali a tempo continuo (Figura B.5).

C.4.2 Forma trigonometrica

Per le funzioni reali si può ricavare anche la *forma trigonometrica dello sviluppo di Fourier discreto* e la funzione f si può scrivere come

$$f(k) = \begin{cases} F_0 + 2 \sum_{n=1}^{\hat{N}-2} |F_n| \cos(n\theta_0 k + \arg F_n) + (-1)^k F_{N/2}, & N \text{ pari} \\ F_0 + 2 \sum_{n=1}^{\hat{N}-1} |F_n| \cos(n\theta_0 k + \arg F_n), & N \text{ dispari} \end{cases}$$

Questa formula mostra come i segnali dotati di *sviluppo di Fourier* siano scomponibili nella somma di una costante F_0 , chiamata *componente a pulsazione nulla*, e di $\hat{N} - 1$ componenti cosinusoidali, dette *armoniche*, di pulsazioni multiple di θ_0 : si parla di *armonica fondamentale* per riferirsi proprio a quella di pulsazione θ_0 e di *armonica n -esima* per denotare la generica di pulsazione $n\theta_0$. Allora $|F_n|$ rappresenta il “peso” in f dell’armonica n -esima e $\arg F_n$ la sua fase. La *pulsazione minima* di un segnale è quella corrispondente al minimo valore n_1 di n , $0 \leq n \leq \hat{N} - 1$, per cui $F_n \neq 0$; la *pulsazione massima* corrisponde invece al massimo valore n_2 dei valori di n , $0 \leq n \leq \hat{N} - 1$, per cui $F_n \neq 0$; la *banda* è l’intervallo di pulsazioni compreso tra la minima $n_1\theta_0$ e la massima $n_2\theta_0$. Si usa dire che sviluppare secondo Fourier un segnale equivale a effettuarne un’*analisi armonica*, cioè a rappresentarlo nel dominio della frequenza.

Esempio C.10 Sviluppo della sinusoida Si consideri la funzione

$$f(k) = \sin(\theta k)$$

Essa è N -periodica se e solo se $\theta/2\pi$ è razionale, cioè esiste un intero m tale che $\theta N = m2\pi$. In questo caso, ponendo $\theta_0 = 2\pi/N$, si ha

$$f(k) = \sin(m\theta_0 k) = \frac{e^{jm\theta_0 k} - e^{-jm\theta_0 k}}{2j}$$

Si ottiene quindi, per la (C.15),

$$F_{m+hN} = \tilde{F}_{-(m+hN)} = \frac{1}{2j} = -j0.5 \quad h = \dots, -2, -1, 0, 1, 2, \dots$$

mentre tutte le altre componenti dello spettro sono nulle.

Esempio C.11 Sviluppo dell’onda quadra Si consideri la funzione periodica f di periodo $N = 4a$, a intero, definita da

$$f(k) = \begin{cases} 0 & -2a \leq k < -a \\ 1 & -a \leq k \leq a \\ 0 & a < k < 2a \end{cases}$$

e detta onda quadra discreta. Dalla (C.14) si ha

$$F_n = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} f(k) e^{-jn\theta_0 k} = \frac{1}{4a} \sum_{k=-a}^a e^{-jn\theta_0 k}$$

per cui

$$F_n = \frac{2a+1}{4a} = \frac{0.5N+1}{N}$$

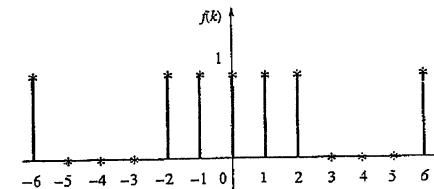


Figura C.3
Onda quadra dell’Esempio C.11.

quando n è un multiplo di N ; altrimenti

$$\begin{aligned} F_n &= \frac{1}{4a} \frac{e^{-jn(a+1)\theta_0} - e^{jna\theta_0}}{e^{-jn\theta_0} - 1} = \\ &= \frac{1}{N} \frac{e^{-jn(0.5\theta_0)} (e^{-jn(a+0.5)\theta_0} - e^{jn(a+0.5)\theta_0})}{e^{-jn0.5\theta_0} (e^{-jn0.5\theta_0} - e^{jn0.5\theta_0})} = \\ &= \frac{1}{N} \frac{\sin(n(0.25N + 0.5)\theta_0)}{\sin(n0.5\theta_0)} \end{aligned}$$

in quanto

$$\sum_{i=\alpha}^{\beta} y^i = \frac{\gamma^{\beta+1} - \gamma^\alpha}{\gamma - 1}, \quad \gamma \neq 1 \quad (\text{C.16})$$

L’onda quadra e il suo spettro sono riportati nelle Figure C.3 e C.4 per $N = 8$.

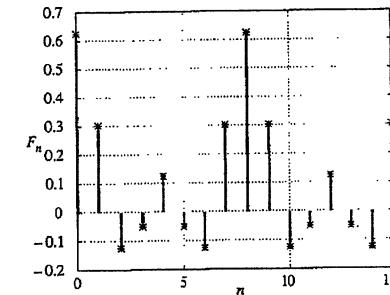


Figura C.4
Spettro dell’onda quadra dell’Esempio C.11.

C.5 Trasformata di Fourier discreta

C.5.1 Forma esponenziale

Sia data una funzione complessa f della variabile intera k definita per $-\infty < k < +\infty$. Se la funzione complessa della variabile reale θ

$$F(e^{j\theta}) = \sum_{k=-\infty}^{+\infty} f(k) e^{-jk\theta} \quad (\text{C.17})$$

esiste almeno per qualche valore di θ , essa si dice *trasformata di Fourier discreta in forma esponenziale*, o semplicemente *trasformata di Fourier discreta*, di $f(k)$. Si usa tralasciare l’aggettivo “discreta” quando ciò non possa dar luogo a equivoci. Per convenzione si denota la trasformata con la stessa lettera, però maiuscola, che

indica la funzione del tempo trasformanda e si usa scrivere $F(e^{j\theta}) = \mathcal{F}^*[f(k)]$ e $f(k) = \mathcal{F}^{*-1}[F(e^{j\theta})]$. La funzione F si dice anche *spettro* di f e con riferimento alle funzioni $|F(e^{j\theta})|$ e $\arg F(e^{j\theta})$ si parla di *spettro di ampiezza* e di *spettro di fase*. Questi spettri si dicono *continui*, nel senso che sono funzioni di θ che è una variabile reale. È facile verificare che lo spettro è una funzione periodica di periodo 2π , per cui la sua conoscenza su un qualunque intervallo di valori di θ di lunghezza 2π equivale alla conoscenza dello spettro intero. Questa proprietà deriva dal fatto che

$$e^{-j\theta k} = e^{-j(\theta+2\pi)k}$$

Quanto sopra e i risultati che seguono si estendono in modo naturale a funzioni f vettoriali.

La funzione trasformanda può essere ricavata dalla sua trasformata mediante la *formula di antitrasformazione*

$$f(k) = \frac{1}{2\pi} \int_{2\pi} F(e^{j\theta}) e^{jk\theta} d\theta \quad (\text{C.18})$$

dove il simbolo $\int_{2\pi}$ significa che l'integrale è esteso a un qualunque intervallo di valori di θ di lunghezza 2π , cioè la *trasformazione di Fourier* stabilisce una corrispondenza biunivoca tra $f(k)$ e $\mathcal{F}^*[f(k)]$. Lo spettro costituisce perciò una rappresentazione alternativa di $f(k)$.

Tra le proprietà dell'operazione di trasformazione di Fourier discreta, notevole è quella di linearità. Essa afferma che, date due funzioni complesse f e g con trasformate $F(e^{j\theta})$ e $G(e^{j\theta})$, risulta $\mathcal{F}^*[\alpha f + \beta g] = \alpha F(e^{j\theta}) + \beta G(e^{j\theta})$, $\forall \alpha \in C, \forall \beta \in C$.

Particolarmente rilevante è il caso in cui f sia reale: in questa situazione infatti

$$F(e^{-j\theta}) = \bar{F}(e^{j\theta})$$

ed è allora chiaro che per definire compiutamente lo spettro (2π -periodico) bastano i valori di $F(e^{j\theta})$ per $0 \leq \theta \leq \pi$. Della periodicità e di quest'ultima proprietà, si fa ovviamente uso per semplificare le rappresentazioni grafiche.

C.5.2 Forma trigonometrica

Per le funzioni reali, in alternativa alla (C.18) si può ricavare anche la *forma trigonometrica della trasformazione di Fourier discreta* e la funzione f si può scrivere come

$$f(k) = \frac{1}{\pi} \int_0^\pi |F(e^{j\theta})| \cos(\theta k + \arg F(e^{j\theta})) d\theta \quad (\text{C.19})$$

Questa formula mostra come i segnali discreti *trasformabili con Fourier* siano scomponibili in una infinità non numerabile di componenti cosinusoidali, dette *armoeniche*. Si osservi che anche i segnali di durata limitata vengono visti come somme di cosinusoidi che invece hanno durata illimitata. Interpretando l'integrale nella (C.19) come il limite di una sommatoria, si può dire che l'armonica di pulsazione θ ha un "peso" in $f(k)$ pari a $|F(e^{j\theta})| d\theta / \pi$ e una fase pari ad $\arg F(e^{j\theta})$. La *pulsazione minima* di un segnale è quella corrispondente all'estremo inferiore θ_1 di θ nell'intervallo $[0, \pi]$ per cui $F(e^{j\theta}) \neq 0$; la *pulsazione massima* corrisponde invece all'estremo superiore θ_2 dei valori di θ nell'intervallo $[0, \pi]$ per cui $F(e^{j\theta}) \neq 0$; la *banda* è l'intervallo di pulsazioni compreso tra la minima θ_1 e la massima θ_2 . Si usa dire che trasformare con Fourier un segnale equivale a

effettuarne un'*analisi armonica*, cioè a rappresentarlo nel dominio della frequenza. Un segnale trasformabile può contenere armoniche di qualunque pulsazione reale compresa tra 0 e π a differenza di un segnale periodico sviluppabile secondo Fourier, che può contenere solo armoniche di pulsazioni multiple di quella dell'armonica fondamentale.

Esempio C.12 Trasformata dell'impulso Sia $f(k) = \text{imp}^*(k)$; allora per le Equazioni (C.1) e (C.17) si trova

$$\mathcal{F}^*[\text{imp}^*(k)] = \sum_{k=-\infty}^{+\infty} f(k) e^{-j\theta k} = e^{-j\theta 0} = 1$$

Esempio C.13 Trasformata dell'impulso rettangolare Si consideri la funzione f definita da

$$f(k) = \begin{cases} 1, & -a \leq k \leq a \\ 0, & \text{altrove} \end{cases}$$

dove a è un intero positivo, che è detta *impulso rettangolare*. Dalla (C.17) si ha

$$F(e^{j\theta}) = \sum_{k=-\infty}^{+\infty} f(k) e^{-j\theta k} = \sum_{k=-a}^a e^{-j\theta k}$$

e quindi

$$F(e^{j0}) = 2a + 1$$

mentre per $0 < \theta \leq \pi$, ricordando la (C.16),

$$\begin{aligned} F(e^{j\theta}) &= \frac{e^{-j(a+1)\theta} - e^{-ja\theta}}{e^{-j\theta} - 1} = \\ &= \frac{e^{-j0.5\theta} (e^{-j(a+0.5)\theta} - e^{j(a+0.5)\theta})}{e^{-j0.5\theta} (e^{-j0.5\theta} - e^{j0.5\theta})} = \frac{\sin((a+0.5)\theta)}{\sin(0.5\theta)} \end{aligned}$$

Le Figure C.5 e C.6 riportano l'impulso rettangolare e il suo spettro per $a = 2$.

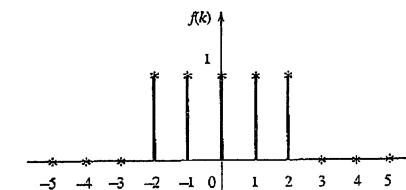


Figura C.5
Impulso rettangolare
dell'Esempio C.13.

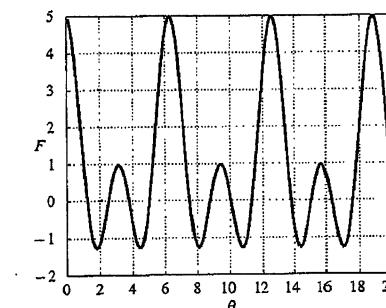


Figura C.6
Spettro dell'impulso
rettangolare dell'Esempio
C.13.

C.5.3 Relazioni con la trasformata Zeta

Confrontando le definizioni (C.2) e (C.17) si notano subito le strette relazioni esistenti tra la trasformata Zeta e quella di Fourier discreta. Conviene però osservare anche le differenze tra i due strumenti matematici.

Si considerino le funzioni f nulle per $k < 0$. Allora, se la trasformata Zeta esiste e ha raggio di convergenza $\bar{\rho} < 1$, vale la relazione

$$\mathcal{F}^*[f(k)] = \mathcal{Z}[f(k)]|_{z=e^{j\theta}} \quad (\text{C.20})$$

In altri termini, esiste anche la trasformata di Fourier discreta che si ottiene subito ponendo $z = e^{j\theta}$ in quella Zeta. Ciò giustifica l'uso di denotare con $F(z)$ e $F(e^{j\theta})$ le due trasformate, laddove quella di Fourier si sarebbe a buon diritto dovuta indicare come funzione di θ piuttosto che di $e^{j\theta}$. Quando invece $\bar{\rho} \geq 1$, l'esistenza della trasformata Zeta non implica l'esistenza di quella di Fourier discreta. Quest'ultima può ancora esistere, senza però che valga l'Equazione (C.20).

Se quindi l'operazione di trasformazione Zeta appare più potente di quella di Fourier discreta, non bisogna però sottovalutare i due fatti seguenti, che in molti casi fanno preferire la seconda trasformata alla prima. Innanzitutto con la trasformata di Fourier discreta è possibile trattare segnali definiti e non identicamente nulli anche per $k < 0$, cosa per la verità fattibile anche con un'estensione, detta *trasformata bilatera*, della trasformata Zeta (*unilatera*) qui descritta.

Soprattutto, poi, la trasformata di Fourier discreta, nella sua forma trigonometrica, consente di interpretare le funzioni di una vasta classe come costituite dalla somma di un'infinità non numerabile di armoniche, estendendo così ai segnali non periodici l'importante risultato conseguito con lo sviluppo discreto dei segnali periodici secondo Fourier, cioè nella somma di un numero finito di armoniche. La trasformata Zeta permetterebbe solo di sviluppare i segnali in una somma di un'infinità non numerabile di armoniche moltiplicate per esponenziali, un fatto questo che risulta molto meno trasparente.

Pertanto le due trasformate devono ritenersi assolutamente complementari.

Riferimenti bibliografici

Data l'intensa attività di ricerca sui temi dell'automatica sviluppatasi negli ultimi decenni e il vasto spettro di argomenti trattati in questo testo, la bibliografia potrebbe essere molto estesa. Ci si limiterà invece a riportare, secondo uno schema che sostanzialmente ricalca la struttura del volume, i riferimenti ai libri disponibili in italiano e ai principali testi in lingua inglese, confidando che ciò sia sufficiente a indicare un valido punto di partenza per l'approfondimento delle singole tematiche. Saranno inoltre elencati alcuni riferimenti bibliografici che, benché relativi ad argomenti non trattati in questo volume, forniscono utili indicazioni su vari importanti settori di sviluppo della teoria del controllo, quali il controllo ottimo, robusto, predittivo, adattativo, a struttura variabile, fuzzy. Si segnalano infine alcuni testi per lo studio delle tecniche di identificazione e di trattamento e analisi dei segnali.

Modellistica La determinazione del modello di un processo viene normalmente effettuata a partire dalle leggi della fisica, dell'ingegneria, o delle scienze economiche e sociali che governano il fenomeno in esame. Approcci di carattere più generale per lo sviluppo di modelli matematici utili ai fini del controllo sono riportati in [1.1]-[1.4].

- [1.1] V. Kecman, *State Space Models of Lumped and Distributed Systems*. Springer-Verlag, 1988.
- [1.2] J.L. Shearer, B.T. Kulakowski e J.F. Gardner, *Dynamic Modeling and Control of Engineering Systems*, Prentice Hall, 1997.
- [1.3] P. Thomas, *Simulation of Industrial Processes for Control Engineers*, Butterworth-Heinemann, 1999.
- [1.4] P.E. Wellstead, *Introduction to Physical Modelling*, Academic Press, 1980.

Teoria dei sistemi Lo studio delle caratteristiche e proprietà dei sistemi dinamici è trattato in testi di teoria dei sistemi quali [2.1], [2.2], [2.4], [2.7], [2.8], [2.9]. L'analisi dei sistemi lineari è anche esaurientemente condotta in [2.3], mentre per un approfondimento sui sistemi non lineari si rimanda a [2.11]. Altri testi di teoria dei sistemi utili per la consultazione sono [2.5], [2.6], [2.10].

- [2.1] S. Chiaverini, F. Caccavale, L. Villani e L. Sciaffico, *Fondamenti di Sistemi Dinamici*, McGraw-Hill, 2003.
- [2.2] E. Fornasini e G. Marchesini, *Appunti di Teoria dei Sistemi*, Edizioni Libreria Progetto, 2003.
- [2.3] T. Kailath, *Linear Systems*, Prentice Hall, 1980.
- [2.4] G. Marro, *Teoria dei Sistemi e del Controllo*, Zanichelli, 1989.
- [2.5] K. Ogata, *System Dynamics*, Prentice Hall, 2004.
- [2.6] K. Ogata, *Discrete-Time Control Systems*, Prentice Hall, 1995.
- [2.7] S. Rinaldi, *Teoria dei Sistemi*, Clup, 1973.
- [2.8] S. Rinaldi, C. Piccardi, *I Sistemi lineari: Teoria, Modelli, Applicazioni*, Città Studi, 1997.
- [2.9] A. Ruberti e A. Isidori, *Teoria dei Sistemi*, Boringhieri, 1979.
- [2.10] W.J. Rugh, *Linear System Theory*, Prentice Hall, 1996.
- [2.11] M. Vidyasagar, *Nonlinear Systems Analysis*, Prentice Hall, 1993.

Tecniche classiche di controllo I metodi di controllo di sistemi SISO trattati nei Capitoli 9-13 sono analizzati diffusamente nei testi in italiano [3.2], [3.3], [3.5], [3.6], [3.9]. Libri in lingua inglese in cui si segue un'impostazione tradizionale sono [3.4], [3.7], [3.10], mentre una trattazione della teoria classica del controllo secondo un approccio più recente è riportata in [3.1]. Infine, una sorta di enciclopedia del controllo di rapida ed efficace consultazione è [3.8].

- [3.1] J.C. Doyle, B.A. Francis e A.R. Tannenbaum, *Feedback Control Theory*, Maxwell MacMillan Int. Editions, 1992.
- [3.2] A. Ferrante, A. Lepischy e U. Viaro, *Introduzione ai Controlli Automatici*, UTET, 2000.
- [3.3] G. Fiorio, *Controlli Automatici con Elementi di Teoria dei Sistemi*, CLUT, 1992.
- [3.4] G.F. Franklin, J.D. Powell e A. Emami-Naeini, *Feedback Control of Dynamic Systems*, Prentice Hall, 2002.
- [3.5] G.O. Guardabassi, *Note di Controlli Automatici*, Pitagora Editrice, 1997.
- [3.6] A. Isidori, *Sistemi di Controllo*, Siderea, 1992.
- [3.7] B. C. Kuo, F. Golnaragh, *Automatic Control Systems*, Wiley, 2002.
- [3.8] W. S. Levine (Ed.), *The Control Handbook*, IEEE Press, 1996.
- [3.9] G. Marro, *Controlli Automatici*, Zanichelli, 1997.
- [3.10] K. Ogata, *Modern Control Engineering*, Prentice Hall, 2002.

Regolatori PID Il testo in cui sono diffusamente analizzate le caratteristiche dei regolatori PID, le problematiche relative alla loro realizzazione e le più recenti tecniche per la taratura automatica dei loro parametri è [4.1]. Per ulteriori approfondimenti si rimanda a [4.2].

- [4.1] K.J. Åström e T. Hagglund, *PID Controllers: Theory, Design and Tuning*, ISA, 1995.
- [4.2] A. Datta, M. Ho e S.P. Bhattacharyya, *Structure and Synthesis of PID Controllers*, Springer, 2000.

Controllo dei processi Gli schemi presentati nel Capitolo 15 sono frequentemente impiegati nel controllo dei processi industriali, in particolare nell'industria

chimica e petrolchimica. Libri specificamente dedicati al controllo dei processi sono [5.1]-[5.5].

- [5.1] M.A. Henson e D.E. Seborg (a cura di), *Nonlinear Process Control*, Prentice Hall, 1997.
- [5.2] W.L. Luyben, *Process Modeling, Simulation and Control for Chemical Engineers*, McGraw-Hill, 1973.
- [5.3] B.A. Ogunnaike e W.H. Ray, *Process Dynamics Modeling and Control*, Oxford University Press, 1994.
- [5.4] D.E. Seborg, T.F. Edgar e D.A. Mellichamp, *Process Dynamics and Control*, Wiley, 2003.
- [5.5] F.G. Shinskey, *Process Control Systems*, McGraw-Hill, 1996.

Controllo non lineare Per un approfondimento delle tematiche riguardanti lo studio dei sistemi di controllo non lineari si rimanda ai testi in lingua inglese [6.1]-[6.5].

- [6.1] A. Isidori, *Nonlinear Control Systems*, Springer-Verlag, 1995.
- [6.2] H.K. Khalil, *Nonlinear Systems*, Prentice Hall, 2002.
- [6.3] R. Marino e P. Tomei, *Nonlinear Control Design*, Prentice Hall, 1995.
- [6.4] R. Sepulchre, M. Jankovic e P. Kokotovic, *Constructive Nonlinear Control*, Springer-Verlag, 1997.
- [6.5] E. Slotine e W. Li, *Applied Nonlinear Control*, Prentice Hall, 1991.

Controllo digitale Tra i libri in italiano relativi al progetto e alla realizzazione di sistemi di controllo digitale si segnalano [7.2], [7.4], [7.6]. Per approfondimenti si rimanda anche a [7.1], [7.3], [7.5], [7.7].

- [7.1] K.J. Åström e B. Wittenmark, *Computer-Controlled Systems: Theory and Design*, Prentice Hall, 1997.
- [7.2] C. Bonivento, C. Melchiorri e R. Zanasi, *Sistemi di Controllo Digitale*, Esculapio, 1999.
- [7.3] G.F. Franklin, J.D. Powell e M.L. Workman, *Digital Control of Dynamic Systems*, Prentice Hall, 1998.
- [7.4] G.O. Guardabassi, *Elementi di Controllo Digitale*, Clup, 1992.
- [7.5] V. Kucera, *Analysis and Design of Discrete Linear Control Systems*, Prentice Hall, 1991.
- [7.6] G. Marro, *Complementi di Controlli Automatici, Controlli Digitali, e Approccio nello Spazio degli Stati*, Zanichelli, 1998.
- [7.7] C.L. Phillips e T.N. Nagle, *Digital Control System Analysis and Design*, Prentice Hall, 1995.

Identificazione Le principali metodologie per l'identificazione di modelli a partire da dati sperimentali sono presentate in [8.1], [8.2]. Ulteriori libri per l'approfondimento è la consultazione sono [8.3], [8.4].

- [8.1] S. Bittanti, *Identificazione dei Modelli e Sistemi Adattativi*, Pitagora, 2002.

- [8.2] S. Bittanti, *Teoria della Predizione e del Filtraggio*, Pitagora, 2000.
- [8.3] L. Ljung, *System Identification. Theory for the User*, Prentice Hall, 1999.
- [8.4] T. Söderström e P. Stoica, *System Identification*, Prentice Hall, 1994.

Controllo ottimo Un approccio ormai classico al problema della sintesi del regolatore, non trattato in questo libro, basato sulla minimizzazione di un funzionale di costo, è presentato in [9.6], [9.7]. Sullo stesso argomento si raccomandano i testi [9.1], [9.2], [9.3], [9.5] e, per una trattazione di carattere introduttivo, il libro [9.4].

- [9.1] B.D.O. Anderson e J.B. Moore, *Optimal Control - Linear Quadratic Methods*, Prentice Hall, 1990.
- [9.2] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, 2000.
- [9.3] A.E. Bryson e H.C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*, Hemisphere Pub., 1981.
- [9.4] B. Friedland, *Control System Design: an Introduction to State-Space Methods*, McGraw-Hill, 1987.
- [9.5] H. Kwakernaak e R. Sivan, *Linear Optimal Control Systems*, Wiley, 1972.
- [9.6] A. Locatelli, *Controllo Ottimo*, Pitagora, 1996.
- [9.7] A. Locatelli, *Optimal Control: an Introduction*, Birkhäuser, 2001.

Controllo robusto Alcune recenti tecniche per la sintesi di regolatori che garantiscono determinate caratteristiche di robustezza a fronte di errori di modello e di attenuazione dei disturbi sono presentate in [10.1]. Ulteriori libri in lingua inglese sono [10.2]-[10.7].

- [10.1] P. Colaneri e A. Locatelli, *Controllo in RH_2/RH_∞* , Pitagora, 1993.
- [10.2] P. Colaneri, J.C. Gerome e A. Locatelli, *Control Theory and Design: an RH_2 and RH_∞ Viewpoint*, Academic Press, 1997.
- [10.3] M. Green e D.J.N. Limebeer, *Linear Robust Control*, Prentice Hall, 1995.
- [10.4] J. Maciejowski, *Multivariable Feedback Design*, Addison-Wesley, 1989.
- [10.5] M. Morari e E. Zafriou, *Robust Process Control*, Prentice Hall, 1989.
- [10.6] S. Skogestad e I. Postlethwaite, *Multivariable Feedback Control*, Wiley, 1996.
- [10.7] K. Zhou, J.C. Doyle e K. Glover, *Robust and Optimal Control*, Prentice Hall, 1996.

Tecniche avanzate di controllo Tra le più recenti tecniche di controllo, il controllo adattativo ha suscitato molto interesse in ambito sia accademico sia industriale. Per un'introduzione ai principali approcci adottati nel settore si segnalano i libri [11.1], [11.5], [11.7]. Il controllo predittivo è discusso in [11.2], [11.6]. Un metodo di controllo non lineare di grande efficacia è il cosiddetto controllo a struttura variabile, per il quale due testi di riferimento sono [11.4], [11.8]. Le metodologie e i campi applicativi del controllo fuzzy sono presentati in [11.3], [11.9].

- [11.1] K.J. Åström e B. Wittenmark, *Adaptive Control*, Prentice Hall, 1995.

- [11.2] E.F. Camacho e C. Bordons, *Model Predictive Control*, Springer-Verlag, 2004.
- [11.3] D. Driankov, H. Hellendoorn e M. Reinfrank, *An Introduction to Fuzzy Control*, Springer-Verlag, 1996.
- [11.4] C. Edwards e S.K. Spurgeon, *Sliding Mode Control: Theory & Applications*, Taylor & Francis, 1998.
- [11.5] G.C. Goodwin e K.S. Sin, *Adaptive Filtering, Prediction and Control*, Prentice Hall, 1984.
- [11.6] J.M. Maciejowski, *Predictive Control*, Prentice Hall, 2002.
- [11.7] E. Mosca, *Optimal, Predictive and Adaptive Control*, Prentice Hall, 1995.
- [11.8] V.I. Utkin, *Sliding Modes in Control and Optimization*, Springer-Verlag, 1992.
- [11.9] J. Yen, R. Langari e L. Zadeh, *Industrial Applications of Fuzzy Logic and Intelligent Systems*, IEEE Press, 1995.

Sistemi a eventi discreti La modellistica e il controllo di sistemi a eventi discreti sono diffusamente trattati nei testi [12.1]-[12.3].

- [12.1] C.G. Cassandras e S. Lafortune, *Introduction to Discrete Event Systems*, Kluwer, 1999.
- [12.2] A. Di Febbraro e A. Giua, *Sistemi ad Eventi Discreti*, McGraw-Hill, 2002.
- [12.3] L. Ferrarini, *Automazione Industriale: Controllo Logico con Reti di Petri*, Pitagora, 2001.

Componenti I principali componenti dei sistemi di controllo e i più diffusi apparati di strumentazione sono presentati in [13.1]-[13.4].

- [13.1] R.N. Bateson, *Introduction to Control System Technology*, Prentice Hall, 2002.
- [13.2] G. Magnani, *Tecnologie dei Sistemi di Controllo*, McGraw-Hill, 2000.
- [13.3] G. Marro, *Componenti dei Sistemi di Controllo*, Zanichelli, 1984.
- [13.4] M. Petternella e R. Vitelli, *Strumentazione Industriale*, UTET, 1981.

Analisi dei segnali Libri che, per la loro completezza e autorevolezza, costituiscono un importante riferimento sul trattamento dei segnali a tempo continuo e discreto sono [14.1]-[14.4].

- [14.1] H. Kwakernaak e R. Sivan, *Modern Signals and Systems*, Prentice Hall, 1991.
- [14.2] B.P. Lathi, *Linear Systems and Signals*, Oxford University Press, 2004.
- [14.3] A.V. Oppenheim, R.W. Shafer e J.R. Buck, *Discrete-Time Signal Processing*, Prentice Hall, 1999.
- [14.4] A.V. Oppenheim, A.S. Willsky e S. Hamid, *Signals and Systems*, Prentice Hall, 1997.

Indice analitico

- A**
- aliasing, 438
 - analisi, 323
 - armonica, 522, 527, 542, 544, 545
 - di sensitività, 284
 - approssimanti di Padé, 388
 - assegnamento
 - degli autovalori, 335
 - degli zeri, 491
 - dei poli, 487
 - del modello, 483
 - attenuazione di un disturbo, 291. 481
 - attuatore, 9
 - automatica, 12
 - automazione, 21
 - autovalore, 54, 63, 178, 184, 463. 499
 - autovettore, 499
 - azione
 - derivativa, 355
 - filtrante, 158
 - integrale, 357, 455
- B**
- banda, 522, 527, 542, 544
 - passante, 159, 162, 266, 270
 - proporzionale, 352
- C**
- campionamento, 437
 - campionatore, 431, 432
 - cancellazione, 91, 93, 205, 206
 - carica integrale, 357, 456
 - cerchi di Hall, 286
 - coefficiente
 - dell'azione derivativa, 352
 - dell'azione integrale, 352
 - dell'azione proporzionale, 352
 - combinazione lineare, 498
- D**
- compensatore, 7, 380
 - compensazione, 7, 383, 385
 - statica, 380
 - comportamento
 - dinamico, 14
 - statico, 14
 - controllabilità, 79, 195
 - controlli automatici, 5
 - controllo
 - a due gradi di libertà, 384, 491
 - adattativo, 21
 - automatico, 4
 - centralizzato, 399
 - decentralizzato, 399
 - digitale, 9
 - in anello aperto, 6, 379
 - in anello chiuso, 6
 - in cascata, 391
 - in retroazione, 6
 - manuale, 4
 - nell'intorno di un equilibrio, 230
 - controllore, 4
 - dinamico, 19
 - in anello aperto, 6, 380
 - in anello chiuso, 7, 19
 - non dinamico, 19
 - statico, 19
 - convertitore
 - analogico/digitale, 431
 - digitale/analogico, 431
 - costante
 - di tempo, 94, 96
 - di tempo equivalente, 370
 - di trasferimento, 94, 207
 - criterio
 - del cerchio, 411
 - di Bode, 252
 - di Jury, 187
 - di Kharitonov, 68
 - di Nyquist, 239, 473
- E**
- elemento non lineare da caratteristica, 408
 - equazione
 - caratteristica, 64, 124, 185, 499
 - di stato, 29, 168
 - diofantina, 488
 - pseudocaratteristica, 421
 - equilibrio, 38, 60, 172, 175, 182
 - asintoticamente stabile, 43, 173
 - instabile, 42, 173
 - nominale, 41, 70, 173, 190
 - stabile, 41, 173
 - errore, 5
 - esponenziale di matrice, 503

F
 fase critica, 248
 feedback, 7
 feedforward, 6
 filtro, 158
 - a spillo, 166
 - anti-aliasing, 441, 447
 - di Butterworth, 160
 - passa-alto, 161
 - passa-basso, 158
 FOH, 458
 forma
 - canonica
 -- di controllo, 110
 - di osservabilità, 111, 217
 - di raggiungibilità, 110, 217
 -- di ricostruzione, 111
 - minima, 81, 89, 197, 205
 formula
 - di Lagrange, 50
 - di Shannon, 458
 frequenza
 - di campionamento, 433
 - di mantenimento, 434
 funzione
 - descrittiva, 416, 418
 - di sensitività, 255, 262, 274, 477
 -- complementare, 255, 262, 263, 477
 - del controllo, 262, 279, 477
 - di trasferimento, 87, 88, 203
 -- d'anello, 119, 473

G
 grado relativo, 89, 205, 209
 guadagno, 94, 95, 207, 469
 - critico, 364
 - d'anello, 242, 478
 - generalizzato, 95, 208
 - statico, 60, 95, 183, 207

I
 IAE, 375
 identificazione, 141
 impulso, 507, 533
 incertezza
 - additiva, 254
 - moltiplicativa, 259
 - parametrica, 284
 integratore, 96, 208
 Internal Model Control, 375
 ipotesi dell'azione filtrante, 420
 ISE, 375
 ISTE, 375
 ITAE, 375

L
 legge di controllo, 12, 352, 435
 linearizzazione, 70, 190
 lunga divisione, 538
 luogo
 - dei punti critici, 422
 - delle radici, 315, 475, 482
 - diretto, 316
 - inverso, 316
 M
 mantenitore, 431, 434
 - di ordine uno, 458
 - di ordine zero, 434, 441
 margine
 - di fase, 248, 249, 268, 368
 - di guadagno, 247, 249, 367
 - di stabilità vettoriale, 246
 matrice, 495
 - dei guadagni relativi, 400
 - della dinamica, 36, 172
 - di osservabilità, 79, 196
 - di raggiungibilità, 77, 194
 - di trasferimento, 394
 - guadagno
 - del regolatore, 338
 - dell'osservatore, 343
 - in forma di Jordan, 55, 180, 500
 - in forma diagonale, 54, 178, 500
 matrici simili, 500
 metodo
 - della funzione descrittiva, 416, 420
 - della tangente, 370
 - delle aree, 371
 - di Cohen e Coon, 374
 - di Euler all'indietro, 450
 - di Euler in avanti, 450
 - di tenuta e campionamento, 470
 - di trasformazione diretta di poli e zeri, 471
 - di Tustin, 450
 - di Ziegler e Nichols, 364, 373
 modello
 - approssimante, 215
 - matematico, 11, 18
 moderazione del controllo, 6, 236, 280, 291, 481
 modo, 54, 178
 movimento, 49, 175
 - asintoticamente stabile, 46, 175
 - dell'uscita, 30, 169
 - dello stato, 30, 168
 - forzato, 50, 58, 175, 181
 - instabile, 45, 174
 - libero, 50, 62, 175, 184
 - nominale, 45, 174
 - perturbato, 41, 45, 173, 174
 - stabile, 45, 174

N
 norma, 498
 O
 ordine, 29, 92, 168
 oscillazioni
 - nasconde, 479, 481
 - permanenti, 416
 osservabilità, 73, 79, 93, 193, 195, 206
 osservatore
 - asintotico, 342
 - dello stato, 341
 - di ordine ridotto, 344
 ottimizzazione, 375
 P
 parabola, 507, 533
 parametro, 2
 parte
 - non osservabile, 80, 197
 - non raggiungibile, 78, 195
 - osservabile, 80, 93, 197, 206
 - raggiungibile, 78, 93, 195, 206
 percorso di Nyquist, 238
 periodo
 - dell'oscillazione, 99
 - di campionamento, 432, 446, 465
 - di mantenimento, 434
 polinomio caratteristico, 64, 185, 499
 polo, 89, 94, 205, 207, 510, 535
 - dominante, 107, 163, 214, 266, 268
 potenza di matrice, 499, 501
 precisione
 - dinamica, 18, 291, 480, 485
 - statica, 15, 291, 478, 480, 485, 489
 preditore di Smith, 386
 prestazioni, 235
 - dinamiche, 18, 236, 478
 - robuste, 237, 282
 - statiche, 15, 235, 264, 274, 280, 478
 prewarping, 458
 principio
 - di separazione, 346
 - di sovrapposizione degli effetti, 51, 176
 problemi di controllo, 2
 processo, 2
 progetto
 - dinamico, 294
 - statico, 294
 proprietà bloccante, 139
 pulsazione
 - critica, 248, 265, 277
 - di campionamento, 433
 - di mantenimento, 434
 - di Nyquist, 438
 - naturale, 94, 96

Q
 quantizzazione, 434, 447, 455, 465
 R
 raggiungibilità, 73, 76, 93, 193, 206
 ramo, 319
 rampa, 507, 533
 rango, 498
 rappresentazione, 111
 - di stato, 30, 169
 - esterna, 88
 - ingresso-stato-uscita, 30, 169
 - ingresso-uscita, 84, 88, 199, 204, 217
 - interna, 30, 169
 realizzabilità, 291, 481, 484
 realizzazione, 110, 216
 regione
 - di asintotica stabilità, 63, 184
 - nell'insieme dei parametri, 67, 188
 - di attrazione, 44, 174
 regolatore, 4
 - digitale, 432, 435
 - I, 353
 - in cascata, 390
 - P, 353
 - PD, 354
 - PI, 354
 - PID, 351, 354, 427
 regolazione
 - a zero dell'errore, 489
 - robusta, 276, 282, 381
 reiezione robusta, 276, 282
 relè, 20, 409
 - con isteresi, 408, 419
 - senza isteresi, 408, 419
 rete
 - a sella, 311
 - anticipatrice, 309
 - ritardatrice, 310
 - stabilizzatrice, 294, 308
 retroazione
 - negativa, 119
 - positiva, 119
 RGA, 400
 ricostruità, 81, 197
 risonanza
 - picco di, 145
 - pulsazione di, 145
 risposta
 - all'impulso, 58, 181
 - allo scalino, 98, 99, 209, 272
 - esponenziale, 139, 222
 - in frequenza, 131, 134, 218, 220
 - inversa, 102
 - sinusoidale, 131
 ritardo di tempo, 37, 93, 107, 151, 206, 245, 273, 461
 ritardo equivalente, 370
 S
 saturazione, 408, 419
 scalino, 507, 533
 schema a blocchi, 115, 217, 408
 scomposizione canonica (o di Kalman), 73, 81, 193, 197
 segnale
 - a banda limitata, 522, 527
 - a tempo continuo, 507
 - a tempo discreto, 533
 - canonico, 507, 533
 - di riferimento, 2
 serie di Fourier, 520
 sfasatore puro, 152
 simulazione, 13
 singolarità, 89, 205, 510, 535
 sintesi, 289, 327
 - diretta a tempo discreto, 479
 - nello spazio di stato, 335
 - per tentativi, 294
 sistema
 - a dimensione finita, 38, 172
 - a dimensione infinita, 38, 172
 - a parametri concentrati, 38, 172
 - a parametri distribuiti, 38, 172
 - a segnali campionati, 460
 - a sfasamento minimo, 152, 300
 - a sfasamento non minimo, 301
 - a tempo continuo, 29, 49
 - a tempo discreto, 168, 472
 - canonico, 409
 - del primo ordine, 99, 210
 - del secondo ordine, 100, 211
 - di controllo, 1, 4, 14, 407
 - dinamico, 29, 168
 - equivalente, 52, 177
 - FIR, 181, 213, 215
 - ibrido, 431
 - improprio, 89
 - instabile, 393
 - invariante nel tempo, 35, 171
 - lineare, 35, 49, 167, 172
 - linearizzato, 71, 191
 - monovariabile (o SISO), 34, 171
 - multivariabile (o MIMO), 34, 171
 - non dinamico, 35, 171
 - non lineare, 35, 172, 407
 - proprio, 35, 171
 - retroazionato, 472
 - sotto controllo, 2
 - statico, 35, 171
 - stazionario, 35, 49, 167, 171
 - strettamente proprio, 35, 171
 - triangolare, 394
 - variante nel tempo, 35, 171

T
 taratura
 - automatica, 364, 427
 tempo
 - derivativo, 352
 - di assestamento, 99, 211
 - di latenza, 209, 468, 484
 - di massima sovraelongazione, 99
 - di reset, 352
 - di ritardo, 99
 - di salita, 99
 - integrale, 352
 temporizzazione, 435
 teorema di Shannon (o del campionamento), 440
 teoria
 - dei sistemi, 11
 - del controllo, 11
 tipo, 94, 207, 469
 - del sistema di controllo, 264
 trasduttore, 9
 trasformata
 - di Fourier, 525, 543
 - di Laplace, 509
 - Zeta, 534

556 Indice analitico

- trasformazione
 - bilineare, 186, 449
 - d'uscita, 29, 168
 - di campionamento, 445, 463
- U**
 - uscita di equilibrio, 38, 172
- V**
 - valore
- W**
 - di regime, 99
 - finale, 98, 209
 - iniziale, 98, 209
 - massimo, 99
- variabile
 - controllata, 2
 - di controllo, 2
 - di ingresso, 27, 168
 - di stato, 29, 168
 - di uscita, 27, 168
 - manipolabile, 10
- Z**
 - zero, 89, 94, 205, 207, 467, 510, 535
 - del campionamento, 467
- ZOH, 434, 441
- zona morta, 408, 419

Finito di stampare
nel mese di giugno 2004

Questo volume, sprovvisto del talloncino a fronte, è da considerarsi copia saggio-campione gratuito fuori commercio.
Fuori campo applicazione IVA ed esente da bolla di accompagnamento (art. 22 L. 67/1987, art. 2, lett. I D.P.R. 633/1972 e art. 4 n. 6 D.P.R. 627/1978).

ibern • R. Scattolini • N. Schiavoni
FONDAMENTI
VROLLI AUTOMATI
onda edizione
"Companies
1999