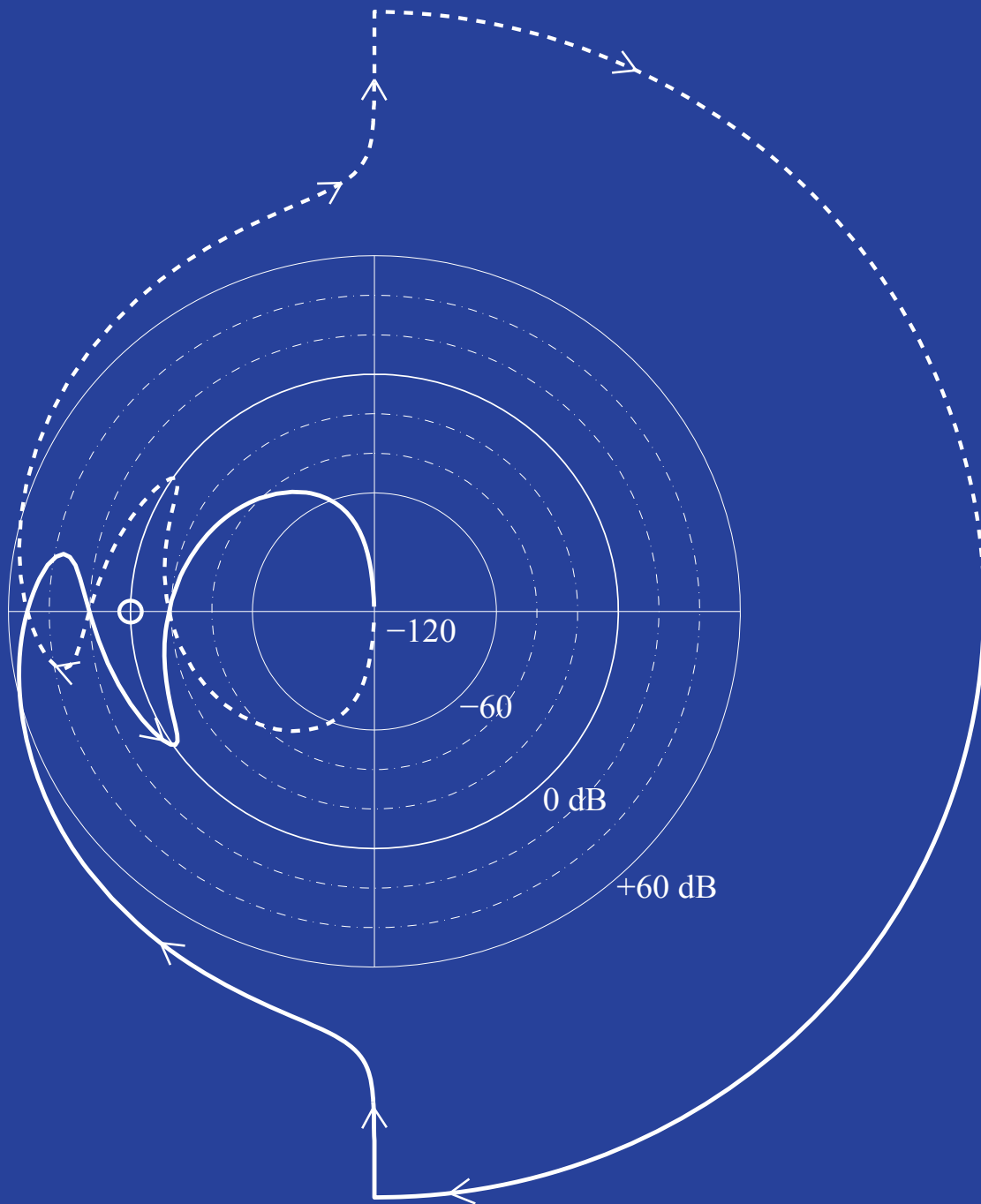


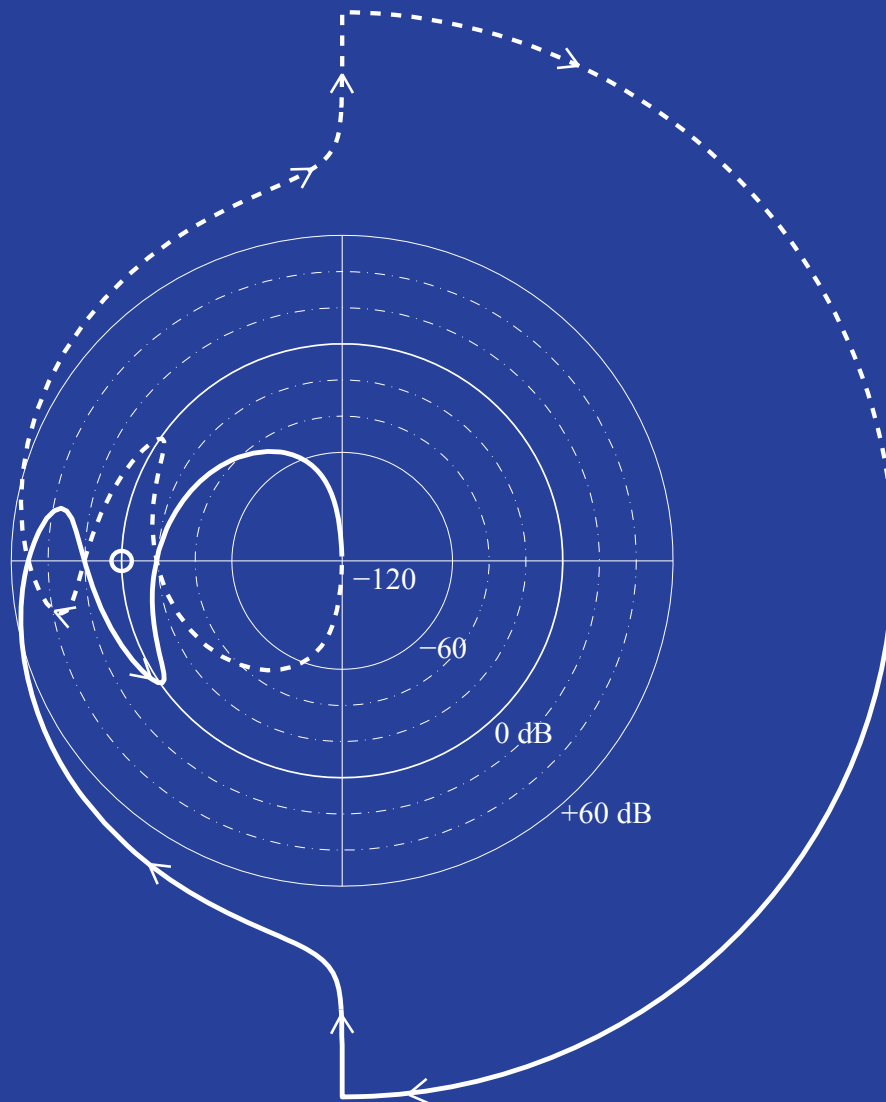
Jens G. Balchen

Trond Andresen

Bjarne A. Foss



Reguleringsteknikk



De fleste tekniske universiteter har i de siste ca. 60 år innført grunnkurs i regulerings-teknikk (engelsk: Control Systems, tysk: Regelungstechnik) for mange av sine studenter. Dette skyldes at det moderne høyt teknologiske samfunn er stadig mer avhengig av slike innsikter og løsninger som regulerings-teknikken tilbyr. Regulerings-teknikk (også kalt kybernetikk) er blitt et av basisfagene som mange ingeniører må beherske.

Boka inneholder mer stoff enn det er naturlig å kreve som pensum i et grunnkurs i regulerings-teknikk, fordi noen ønsker å gå litt dypere i stoffet på egen hånd, eller ha boka for hånden som en oppslagsbok.

Leseren vil ha fordel av noen matematikk-kunnskaper: differensiallikninger, lineær algebra og Laplacetransformasjon, men det er ikke noen betingelse, fordi denne matematikken i stor grad forklares fra bunnen av. Det forutsettes grunnkunnskaper i fysikk, som varmelære, mekanikk med dynamikk og litt elektroteknikk.

Reguleringsteknikk

**Jens G. Balchen
Trond Andresen
Bjarne A. Foss**

6. utgave

Institutt for teknisk kybernetikk

2016

Ny korrektur med noen mindre feil rettet 2021

©INSTITUTT FOR TEKNISK KYBERNETIKK, NTNU, TRONDHEIM 2016

ISBN 978-82-7842-202-1

*Institutt for teknisk kybernetikk
Norges teknisk-naturvitenskapelige universitet (NTNU)
Odd Bragstads plass 2D
7491 Trondheim*

<https://www.ntnu.no/itk>

Innhold

Forord

KAPITTEL 1

Introduksjon til reguleringsteknikk

1.1	Innledning	7
1.2	Reguleringsteknikkens historie	9
1.3	Vesentlige begreper og definisjoner	11
1.4	To metoder for regulering: Forover- og tilbakekopling	15
1.4.1	Statisk prosess	18
1.4.2	Dynamisk prosess	20
1.5	Aktuelle reguleringstekniske anvendelser	22
1.5.1	Styring og regulering av maskiner	22
1.5.2	Styring av fartøyer	23
1.5.3	Industriell prosessregulering	24
1.5.4	Regulering i energisystemer	26
1.5.5	Styring og regulering i fiskeriteknologi	27
1.5.6	Regulering i ikke-tekniske systemer	28
1.5.7	Oppsummering av reguleringstekniske anvendelser	28
1.6	Drivkrefter i industriell automatisering	30

KAPITTEL 2

Matematiske modeller av dynamiske systemer

2.1	Innledning og motivering	33
2.1.1	Stasjonær modell	39
2.2	Differensiallikninger – introduksjon	40
2.2.1	1. ordens differensiallikninger	40
2.3	Lineære systemer, linearitet	46
2.4	Impulsrespons og foldingsintegral	48
2.5	Sett av differensiallikninger – modeller av høyere orden	50
2.5.1	Fysiske balanser	50
2.5.2	Massebalanse (stoffbalanse)	51

2.5.3	Energibalanse	54
2.5.4	Kraftbalanse	54
2.5.5	Elektrisk balanse	58
2.5.6	Prosess med både elektrisk- og mekanisk balanse	61
2.5.7	Bruk av benevninger	64
2.6	Én differensiallikning av høyere orden	65
2.7	Tidsvariante differensiallikninger	68
2.8	Blokkdiagrammer	70

KAPITTEL 3

Tilstandsromanalyse

3.1	Innledning og motivering	73
3.2	Vektordifferensiallikninger	75
3.2.1	Tilstandsvektor og tilstandsrom	75
3.2.2	Et “rektangulært skjema” for matrise- og vektormultiplikasjon	76
3.3	Løsning av lineære vektordifferensiallikninger	77
3.3.1	Transisjonsmatrisen ved dekopling	77
3.3.2	Transisjonsmatrisen ved rekkeutvikling	80
3.3.3	Transisjonsmatrisen og Cayley-Hamiltons teorem	88
3.4	Monovariabel og multivariabel systemer	91
3.4.1	Impulsrespons for monovariabelt system av orden n	92
3.5	Valg av tilstandsvariable	92
3.6	Ulineære vektordifferensiallikninger. Linearisering	99

KAPITTEL 4

Laplacetransformasjon

4.1	Innledning og motivering	107
4.2	Definisjoner	108
4.2.1	Vanlig forekommende Laplacetransformer. Regler og teoremer	110
4.3	Laplacetransformasjon av differensiallikninger	113
4.3.1	1. ordens lineær differensiallikning	113
4.3.2	Høyere ordens lineær differensiallikning	114
4.4	Invers Laplacetransformasjon ved residuregning	115
4.4.1	Residuregning	115
4.4.2	Invers Laplacetransformasjon	117
4.5	Transferfunksjoner og blokkdiagrammer	119
4.5.1	Transferfunksjoner	119
4.5.2	Blokkdiagrammer	122
4.5.3	Blokkdiagram i tidsplan $-s$ som derivasjonsoperator	128

4.5.4	Bruk av benevninger for å sjekke transferfunksjoner	128
4.6	Lineære vektordifferensiallikninger og Laplacetransformasjonen	129
4.6.1	Transfermatrise	132
4.6.2	Fra transferfunksjon til tilstandsrombeskrivelse	135
4.7	Sammenhengen poler/nullpunkter og tidsrespons	138
4.7.1	Rasjonale transferfunksjoner	138
4.7.2	Egenskapene til 2. ordens system	140
4.7.3	Signaler og systemer	152
4.7.4	Modellreduksjon	153
4.7.5	Forflytning av poler/nullpunkter når regulator endres – rotkurver . . .	156
4.7.6	Intuitiv introduksjon til PI-regulator	167
4.7.7	Transferfunksjoner med tidsforsinkelse	168

KAPITTEL 5

Styrbarhet og observerbarhet

5.1	Innledning	175
5.2	Styrbarhet	176
5.3	Observerbarhet	179
5.4	Styrbarhet og observerbarhet i Laplacetransformerte systemer	182
5.5	Funksjonell styrbarhet (FS) av monovariabel lineære systemer	190

KAPITTEL 6

Frekvensanalyse

6.1	Innledning og motivering	193
6.2	Utleddning av frekvensrespons	195
6.2.1	Frekvensrespons av multivariable systemer	200
6.3	Grafisk representasjon av frekvensrespons for monovariabel systemer	201
6.4	Asymptotisk amplitude/fase/frekvens-diagram (Bodediagram)	209
6.4.1	Prosedyre for tegning av asymptotisk AFF-diagram	215
6.4.2	Ikke-minimum-fase-systemer	219
6.5	Skissering av tidsrespons på basis av frekvensrespons	224
6.5.1	Effektivverdien av et lineært systems respons på et stasjonært “tilfeldig” pådrag	232

KAPITTEL 7

Tilbakekoplede systemer

7.1	Innledning og motivering	237
7.2	Tilbakekopling i monovariabel systemer	240

7.2.1	Følgeforhold. Avviksforhold. Reguleringsgrad	241
7.2.2	Å finne transferfunksjon fra v til e	244
7.2.3	Frekvensanalyse av monovariabel tilbakeløst system	245
7.2.4	Nicholsdiagram	249
7.2.5	System med dynamisk tilbakeløst system kontra enhetstilbakeløst system	253
7.3	Følsomhet	260

KAPITTEL 8

Stabilitet

8.1	Innledning og motivering	263
8.2	Definisjoner på stabilitet	264
8.2.1	Inngangs-/utgangsstabilitet	265
8.2.2	Stabilitet i tilstandsrommet	267
8.3	Algebraiske kriterier for stabilitet	270
8.3.1	Rouths kriterium	270
8.3.2	Hurwitz's kriterium	274
8.3.3	Algebraiske kriterier for et system gitt ved sin transferfunksjon	276
8.3.4	Algebraiske kriterier ved transportforsinkelse i transferfunksjonen	279
8.4	Nyquists stabilitetskriterium	284
8.4.1	Utledning av Nyquists stabilitetskriterium	284
8.4.2	Nyquist-prosedyre – oppsummert	287
8.4.3	Nyquists kriterium for åpent stabilt system	288
8.4.4	Det logaritmiske Nyquist-diagram	292
8.4.5	Nyquists kriterium for åpent ustabilt system	294
8.4.6	Nyquists stabilitetskriterium ved ikke-rasjonale funksjoner	298
8.4.7	Bruk av amplitude/fase-diagram til stabilitetsundersøkelse	299
8.4.8	Bode-Nyquist-kriteriet	302
8.4.9	Fasemargin. Forsterkningsmargin	303

KAPITTEL 9

Syntesemetoder for lineære reguleringsystemer

9.1	Innledning og motivering	313
9.2	Systemspesifikasjoner	315
9.2.1	Spesifikasjon for prosessregulering	315
9.2.2	Spesifikasjon for følgesystem (servomekanismer)	318
9.3	Regulatorsyntese ved seriekompensasjon	327
9.3.1	Seriekompensasjon ved hjelp av standardregulatorer for prosessregulering	345

9.3.2	Eksperimentell innstilling av standardregulatorer for prosessregulering – Ziegler-Nichols metode	348
9.3.3	Eksperimentell innstilling av PI-regulator – SIMC-metoden	355
9.3.4	Standardregulatorer for et utvalg av prosesser. Parametervalg.	360
9.3.5	Seriekompensasjon av følgesystemer (servomekanismer)	365
9.4	Når trenger vi integrerende regulator i en reguleringsløyfe?	386
9.4.1	Sammenhengen mellom sluttverditeoremet og frekvensrespons	388
9.5	Kompensasjon med intern tilbakekopling	391
9.6	Foroverkopling i monovariabel systemer	406
9.6.1	Foroverkopling (fra forstyrrelsen) i prosessreguleringssystemer	407
9.6.2	Forholdsregulering	415
9.6.3	Foroverkopling (fra referansen) i følgesystemer	417
9.6.4	Prediktiv foroverkopling	426
9.7	Estimatorbasert regulering	428
9.7.1	Forenklet “tilstandsestimator” ved regulering av prosess med stor forsinkelse	428
9.7.2	Estimatorbasert regulering – Otto-Smith-regulator	432
9.8	Parametrisk optimering av systemer med gitt struktur	436
9.8.1	Kvadratisk objektfunksjon	436
9.8.2	Objektfunksjon basert på systemets frekvensrespons	443
9.8.3	Regulatorinnstilling ved numerisk optimalisering	448
9.9	Robust regulering	453
9.10	Sluttverdiregulering	459

KAPITTEL 10

Multivariable systemer

10.1	Innledning og motivering	467
10.2	Stabilitet og respons	471
10.2.1	Stabilitet og respons av et 2×2 multivariabelt system med diagonalregulator	472
10.3	Seriekompensasjon av en multivariabel reguleringsløyfe	481
10.4	Foroverkopling i multivariable systemer	487
10.5	Optimalregulering og modellbasert regulering	488
10.6	Singulærverdianalyse av lineære multivariable dynamiske systemer	491
10.7	Funksjonell styrbarhet (FS) av multivariable systemer	497

KAPITTEL 11

Diskret regulering av kontinuerlige systemer

11.1	Innledning og motivering	501
------	------------------------------------	-----

11.2	Analyse ved kontinuerlig tilnærming	505
11.2.1	Diskret PI-regulator	505
11.2.2	Frekvensanalyse, kontinuerlig tilnærming	506
11.2.3	Diskret regulator ved hjelp av tidsforskyvingsoperator	510
11.3	Filtre og filtrering	512
11.3.1	Analoge og diskrete filtre	512
11.3.2	Butterworth-filteret – et lavpassfilter	513
11.3.3	Høypassfilter fra Butterworth-filteret	514
11.3.4	Flere lav-ordens filtre eller et med høy orden?	515
11.3.5	Båndpass- og båndstopp-filtre	516
11.3.6	Filtre i reguleringstekniske anvendelser	517
11.4	Diskret analyse og syntese – z-transformasjonen	518
11.4.1	Diskret signal og system	518
11.4.2	z-transformasjon	520
11.4.3	Invers z-transformasjon	522
11.4.4	Transferfunksjon i z	523
11.4.5	Stabilitet for et diskret system	524
11.4.6	Frekvensrespons for diskrete systemer	527
11.4.7	Tasteteoremet (“Samplingsteoremet”)	532
11.5	Analyse av diskrete systemer ved w-transformasjon	534
11.6	Bruk av \underline{w} -transformasjon ved syntese av diskret regulator	545
11.7	Tidsrespons av diskret system	548
11.8	Den multivariable \underline{w} -transformasjonen	550

KAPITTEL 12

Begrensninger i lineære reguleringssystemer

12.1	Innledning og motivering	553
12.2	Metning. Amplitudebegrensning	555
12.2.1	Undersøkelse av metning i tidsplanet	556
12.2.2	Undersøkelse av metning i frekvensplanet	558
12.3	Effektbruk i reguleringssystemer	561
12.4	Bestemmelse av den realiserbare båndbredden ω_c av et reguleringssystem	564
12.5	Pådragsmetning i systemer med integrerende regulator	566

APPENDIKS A

Matrisealgebra

A.1	Lineære transformasjoner. Matriser	572
A.1.1	Eigenverdier og egenvektorer	572
A.1.2	Matriseoperasjoner	575

A.1.3 Similaritetstransformasjon. Diagonalisering	577
A.2 Derivasjon av matriser og vektorer	580

APPENDIKS B

Laplacetransformasjons-par

APPENDIKS C

Utleddning av styrbarhetskriteriet

APPENDIKS D

Manuell skissering av Nicholsdiagram ved hjelp av Bode-diagrammer

APPENDIKS E

Brønmesteren

Bibliografi	607
------------------------------	------------

Stikkordsliste	611
---------------------------------	------------

Forord

De fleste tekniske universiteter i verden har i de siste ca. 50 år innført et grunnkurs i regulerings-teknikk (engelsk: Control Systems, tysk: Regelungstechnik) for mange av sine studenter. Dette skyldes at det moderne høyteknologiske samfunn er stadig mer avhengig av slike innsikter og løsninger som regulerings-teknikken tilbyr. Regulerings-teknikk (også kalt kybernetikk) er blitt et av basisfagene som mange ingeniører må beherske.

Læreboka REGULERINGSTEKNIKK av Jens G. Balchen (f. 1926 – d. 2009) kom først ut i 1963. I 1999 kom en fullstendig omarbeidet utgave. I forhold til de gamle utgavene representerte denne mange forbedringer både med hensyn til faglig innhold og grafisk presentasjon. Disse bøkene har siden 1963 vært brukt i undervisningen ved Norges teknisk-naturvitenskapelige universitet (NTNU – tidligere NTH) og flere andre universiteter og høyskoler.

2016-utgaven er en oppdatering av en seinere utgave fra 2003. Innholdet er en del utvidet, noen pedagogiske forbedringer er gjennomført, og noen feil er rettet. Vi takker sivilingeniør Asle Hammerdal for å ha gjort en stor jobb med dette, og å ha bragt boka over til moderne L^AT_EX-format. Hammerdal har i tillegg designet omslaget på boka.

Leseren vil ha en fordel av noen matematikk-kunnskaper i differensiallikninger, lineær algebra og Laplacetransformasjon, men det er ikke noen betingelse fordi denne matematikken i stor grad forklares fra bunnen av. Det forutsettes basiskunnskaper i fysikk, herunder varmelære, mekanikk med dynamikk og helt grunnleggende elektroteknikk.

Boka inneholder mer stoff enn det er naturlig å kreve som pensum i et grunnkurs i regulerings-teknikk, fordi det er mange som gjerne vil gå litt dypere i stoffet på egen hånd, og dessuten ha boka for hånden som en oppslagsbok. Boka inneholder et stort antall eksempler. Det finnes på markedet mange øvings-samlinger som kan brukes som støtte utover eksemplene som finnes i boka. Det er også utarbeidet en tabell som beskriver de anvendelsesorienterte eksemplene i boka. Denne skal gjøre det enklere for leseren å se sammenhengen mellom de gjennomgående eksemplene i den.

Institutt for teknisk kybernetikk

Norges teknisk-naturvitenskapelige universitet

Trondheim, årsskiftet 2016/2017

Trond Andresen
amanuensis

Bjarne A. Foss
professor

Forord 2

Denne læreboka finnes fra 2021 av bare som (gratis) pdf.
Boka vil ikke bli trykket noe mer i papirutgave.

Asle Hammerdal takkes nok en gang, for å ha luket ut de siste kjente trykkfeil.

Trondheim, november 2021.

Trond Andresen
førsteamanuensis emeritus

TABELL 1 Oversikt over eksempler*

	Elektrisk krets	Økonomi	Blanding og blandetanker	Varmtvannskjel	Likestrømsmotor	Varmeveksler	Svingende masse	Biltrafikk
Kap. 2 Matematiske modeller	2.2 ¹ 2.7 ²	2.3 ¹	2.4 ³ 2.12 ⁴	2.1 ²	2.8 ²		2.5 ² 2.9 ⁵ 2.11 ⁴	
Kap. 3 Tilstandsrom-analyse					3.1 ⁶		3.1 ⁶ 3.2 ⁶ 3.3 ⁷ 3.4 ¹	
Kap. 4 Laplace-transformasjon	4.3 ⁸ 4.4				4.10 ³⁷ 4.11 ³⁷ 4.12 ⁹ 4.13 ³⁷ 4.14 ⁹ 4.15 ⁹		4.1 4.7 ¹⁰ 4.9 ^{8,11} 4.10 ¹¹	
Kap. 5 Styrbarhet og observerbarhet	5.4 ¹³							
Kap. 6 Frekvensanalyse					6.1 ¹⁴			
Kap. 7 Tilbakekoblede systemer	7.6 7.7				7.1, 7.2 ¹⁵ 7.3 7.4 ¹⁹ 7.5 ²⁰ 7.8 ¹⁶			
Kap. 8 Stabilitet					8.1 ³⁶ 8.9 ¹⁷ 8.14 ¹⁸ 8.16 ¹⁹ 8.18 ²¹			
Kap. 9 Konstruksjon av regulator					9.8 ²⁴ 9.13 ²⁵ 9.21 ²⁷	9.3 ²⁴ 9.4 ²⁴		9.11 ²⁴
Kap. 10 Multivariable systemer			10.2 ³¹					
Kap. 11 Diskret regulering av kontinuerlige systemer					11.10 ³³			
Kap. 12 Begrensninger i lineære reguleringsystemer					12.3 ³⁵			

TABELL 2 Oversikt over eksempler (forts.)*

Fjær- system i bil	Fartøy- styring	Invertert pendel	Kjemisk reaktor (flere typer)	Antenne- styring	Turbiner og generatorer i kraft- produksjon	Papir- produksjon	Løpekatt	
2.6 ²								Kap. 2 Matematiske modeller
			3.6 ³					Kap. 3 Tilstandsrom- analyse
4.6 ¹²								Kap. 4 Laplace- transformasjon
		5.5 ¹³						Kap. 5 Styrbarhet og observerbarhet
								Kap. 6 Frekvensanalyse
								Kap. 7 Tilbakekoblede systemer
			8.15 ¹⁸ 8.20 ²³		8.19 ²²			Kap. 8 Stabilitet
	9.9 ²⁴ 9.22 ²⁷		9.16 ²⁶ 9.20 ²⁸	9.10 ²⁴ 9.14 ²⁵	9.15 ²⁵ 9.17 ²⁶	9.19 ²⁷ 9.23 ²⁹	9.27 ³⁰	Kap. 9 Konstruksjon av regulator
	10.3 ³²		10.1 ³²					Kap. 10 Multivariable systemer
								Kap. 11 Diskret regulering av kontinuerlige systemer
	12.2 ³⁴							Kap. 12 Begrensninger i lineære reguleringssystemer

*Kode i tabellen: 2.1, 4.12 = eksempler. 2.1²⁴ der tallet 24 er tema gitt i oversikt på neste side.

Tema for eksempler i tabell 2

1. Tidsrespons beregnes
2. Utvikling av dynamisk modell
3. Utvikling av dynamisk modell og linearisering av denne
4. Blokkdiagram
5. Transformasjon av dynamisk modell
6. Tilstandsrommodell
7. Beregning av transisjonsmatrise
8. Laplacetransformasjon av tidsfunksjon og residuregning
9. Regulering, rotkurver og tidsrespons
10. Laplacetransformasjon av tilstandsrommodell
11. Laplacetransformasjon av modell og beregning av tidsrespons ved residuregning
12. Reduksjon av blokkdiagram
13. Styrbarhet og observerbarhet
14. Frekvensrespons ved ulike frekvenser
15. Følgeforhold og reguleringsgrad
16. Tilbakekopling
17. Rouths kriterium og sløyfetransferfunksjon
18. Nyquists stabilitetskriterium
19. Bruk av amplitude/fase-diagram
20. Bruk av Nicholsdiagram
21. Bruk av Bode-Nyquist kriteriet
22. Fase- og forsterkningsmargin
23. Åpent ustabil system undersøkt i Bode-diagram
24. Regulatorkonstruksjon
25. Regulering med intern tilbakekopling
26. Kaskaderegulering
27. Foroverkopling
28. Forholdsregulering
29. Bruk av estimator i regulering
30. Sluttverdi regulering
31. Enkelsløyferegulering av multivariabelt system
32. Multivariabel regulering
33. Bruk av w -transformasjon
34. Metning og dimensjonering av pådrag
35. Effektmessig dimensjonering
36. Styring av likestrømsmotor
37. Modellreduksjon

Introduksjon til reguleringsteknikk

1.1 Innledning 7

1.2 Reguleringsteknikkens historie 9

1.3 Vesentlige begreper og definisjoner 11

1.4 To metoder for regulering: Forover- og tilbakekopling 15

1.5 Aktuelle reguleringstekniske anvendelser 22

1.6 Drivkrefter i industriell automatisering 30

1.1 INNLEDNING

Den teknologiske utviklingen har forandret menneskenes livsbetingelser på en dramatisk måte i løpet av en periode på noen hundre år.

Fram til slutten av det 19. århundre var tekniske nyskapninger i stor grad knyttet til energi-produksjon, transport og den nye kjemiske- og elektrometallurgiske industrien. Behovet for ingeniører som kunne konstruere nytt utstyr og stå for drift av kompliserte fabrikker, økte kolossalt.

Fram til den andre verdenskrig (1939-45) var teknologien og industrielle produkter preget av mekaniske og elektriske kraftinnretninger. Radio og telefon var riktignok utbredt, og fjernsynet var oppfunnet uten å være i bruk. Informasjonsvitenskap og informasjonsteknologi var imidlertid ikke etablert skikkelig ennå. Det skjedde med stor kraft under den andre verdenskrig, som blant annet var preget av en enorm satsing på teknologisk utvikling.

En lang rekke nye teknologiske utviklingsretninger startet i denne perioden. Vi kan nevne:

- Atomenergi og atombomben.
- Rakett-teknologi som dannet grunnlaget for utvikling av romfart.
- Gassturbin-teknologi som ledet til jetmotoren og moderne flyteknikk.
- Elektroniske regnemaskiner som ledet til moderne datamaskiner.
- Elektronisk styring av våpen som ledet til moderne reguleringsteknikk og industriell automatisering.

Transistoren ble oppfunnet omkring 1948. Dette ledet til halvlederteknologien, som har gitt oss mikroelektronikk og en eksplosiv utvikling av hastigheten og ytelsen til datatekniske komponenter.

Basert på metoder utviklet av blant annet professor Norbert Wiener ved Massachusetts Institute of Technology (MIT), USA, ble det under den andre verdenskrig utviklet metoder for styring av luftvernsskyts som ga en dramatisk økning i sannsynligheten for treff. Metodene var basert på en matematisk beskrivelse av flyets bevegelse og usikkerheten man har ved observasjon av denne bevegelsen. Dette var en del av bakgrunnen da Wiener (1947) formulerte begrepet **kybernetikk**.

Ordet kybernetikk har greske røtter (kybernetes: rormann, den som styrer). Kybernetikk har siden vært et samlebegrep for flere disipliner med teknologisk dominans (teknisk kybernetikk), tyngde på biologiske aspekter (biokybernetikk) og hovedvekt på sosiologiske problemstillinger (sosiokybernetikk).

En hovedsak for Wiener var begrepet *informasjon* som til da hadde hatt liten plass i naturvitenskap og teknologi, men som han mente var like viktig som energi og materialers egenskaper.

Kybernetikk er et *informasjonsfag*, fordi fysiske størrelser angis som *signaler*. Det er framfor alt *informasjonen* i signalene som er viktig. I et annet fagfelt som for eksempel energiteknikk, vil fysiske størrelser være interessante ut fra deres innhold av *energi*.

Teknisk kybernetikk er en disiplin som omfatter regulering, kommunikasjon og databehandling i tekniske systemer. Reguleringsdelen av teknisk kybernetikk kalles ofte for **reguleringsteknikk**. Reguleringsteknikk har som formål å generere en automatisk påvirkning av et teknisk system eller en prosess slik at man oppnår et ønsket resultat.

Når et system for *automatisk regulering* skal realiseres, må mange teknologiske fagområder tas i bruk:

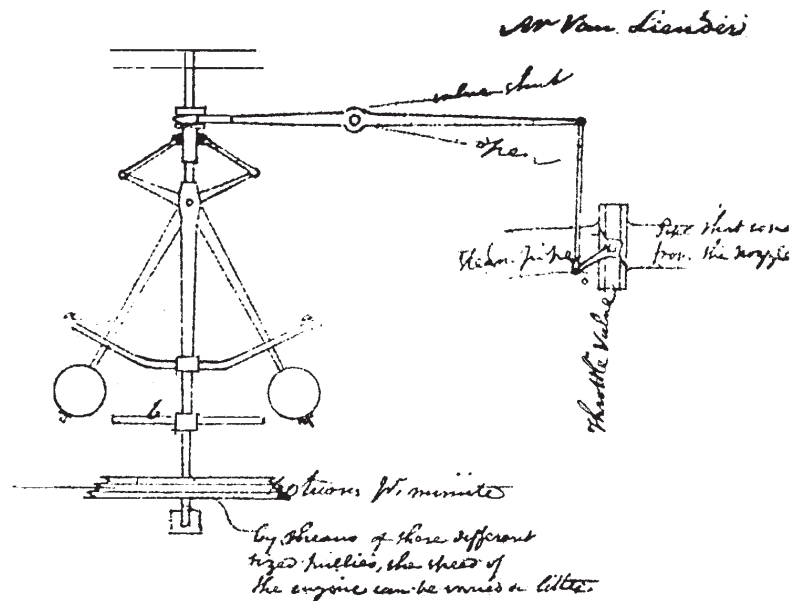
- *Mekanisering* som utnytter mekaniske og elektriske elementer (ventiler, motorer o.l.), for å påvirke farkosters og prosessers atferd.
- *Instrumentering* (eller måleteknikk) som kan generere og formidle informasjon om tilstanden til den regulerte prosessen.
- *Databehandling* (datateknikk og elektronikk) for implementasjon av beregninger og beslutningsregler som brukes ved automatisk regulering.
- *Reguleringsteknikk* som omfatter det teoretiske fundament og ingeniørmessige prinsipper for styring av prosesser av vidt forskjellig karakter.

1.2 REGULERINGSTEKNIKKENS HISTORIE

Mennesker har syslet med reguleringsteknikk i lange tider. Funn fra oldtiden viser for eksempel at egypterne hadde utpekulerte hydrauliske reguleringsmekanismer i sine vannklokker for at de skulle få stabil gang.

Men det skulle gå lang tid før de tekniske nyvinninger som dannede grunnlaget for *den industrielle revolusjon*, ble oppfunnet. James Watt, som hadde konstruert atskillige forbedringer på datidens dampmaskiner, fikk i 1769 patent på en hastighetsregulator. Grunnprinsippet bak hastighetsregulatoren spiller en betydelig rolle i mekanisk regulering av maskiner den dag i dag. Figur 1.1 viser den originale tegningen av James Watts hastighetsregulator.

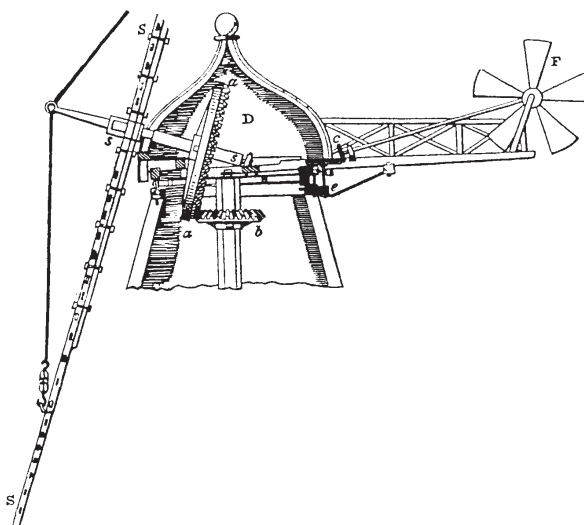
Figur 1.1
James Watts
hastighets-
regulator.



Den består av to kuler som under rotasjon får en bevegelse utover på grunn av sentrifugalkraften. Dette medfører at et rør utenpå den roterende sentrale aksen beveges opp og ned. Denne bevegelsen er et mål for rotasjonens hastighet og benyttes til å styre åpningen i ventilen som bestemmer tilførselen av damp til maskinen som igjen bestemmer rotasjonshastigheten, osv. Vi har dermed en lukket sløyfe som vi kaller *tilbakekopling*.

I 1745 (altså før James Watt) tok Edmund Lee ut patent på et system for regulering av en vindmølle slik at den alltid dreide seg mot vinden, som vist i figur 1.2. En liten vindfanger er utformet slik at den dreier til den blir stående på langs av vindretningen. Når den lille vindfangeren roterer, skjer det en mekanisk overføring over et stort og et lite tannhjul, slik at den store vindfangeren dreies til den blir stående på tvers av vindretningen.

Figur 1.2
System for dreining
av vindmølle mot
vinden.



Det 19. århundre byr på et vell av konstruksjoner for automatisk regulering innen dampteknikk, gruvedrift og metallurgi. J. C. Maxwell (1831-79), som er mest kjent for sine teorier om bølgeforsplantning, la i 1868 fram den første brukbare teorien basert på differensiallikninger for beskrivelse av et system som regulerte hastighet. Senere kom Hurwitz (1875), Routh (1884) og Liapunov (1892) med grunnleggende bidrag til studiet av stabilitet for dynamiske systemer. Disse og andre arbeider sammen med bidrag fra kjente matematikere som P. S. Laplace (1749-1827), J. B. Fourier (1768-1830) og A. L. Cauchy (1789-1857) ga et vesentlig fundament for den første fase av utviklingen av en moderne reguleringsteori. Senere har man tatt i bruk ideer introdusert av L. Euler (1707-83) og J. L. Lagrange (1736-1813).

H. Nyquist ga i 1932 et viktig bidrag til den første del av moderne reguleringsteknikk. Hans arbeid om stabilitet av tilbakekoblede systemer var banebrytende. Etter den tid, og særlig under og etter den andre verdenskrig, har både den teoretiske og praktiske forskningen vært svært mangfoldig. Russeren L. S. Pontryagin og amerikanerne R. E. Bellman og R. E. Kalman har levert viktige grunnleggende bidrag til den moderne reguleringsteknikkens andre fase (1950-80).

En tredje fase i reguleringsteknikkens teori kan påvises fra omkring 1970. Denne fasen preges av algebraiske metoder og pågår fremdeles (1997).

Det typiske reguleringstekniske utstyr omfatter gjerne enheter for måling, signalbehandling, beregning og prosesspåvirkning (pådrag). Fra de enkleste mekaniske og hydrauliske innretninger har utviklingen gått mot elektriske og pneumatiske konstruksjoner av til dels høy kompleksitet. Utviklingen av elektronikken har vært av størst betydning. Den har fra 1950-årene frambragt alle de elementene som i dag er uunnværlige ved bygging av avanserte reguleringssystemer. Datamaskinen spiller en helt spesiell rolle. Fra omkring 1960, da den første

datamaskinen ble tatt i bruk for industriell prosessregulering og fram til i dag, har vi sett en dramatisk utvikling av datateknologien som blant annet har medført en vesentlig reduksjon i pris, størrelse og vekt pr. enhet regnekapasitet. Dette har for eksempel medført at en lettere kan realisere teoretiske løsninger som krever stor databehandlingskapasitet, i praktiske systemer.

1.3 VESENTLIGE BEGREPER OG DEFINISJONER

Mye av terminologien som brukes innen fagområdet reguleringsteknikk er av utenlandsk, særlig engelsk, opprinnelse. Det har i lang tid vært en målsetning i norsk språkpolitikk å skape gode norske betegnelser og begreper i tide, før det danner seg en uakseptabel fremmedspråklig “ukultur”. I de tilfeller (og det er kanskje de fleste) der fremmedspråklige betegnelser og begreper har sitt utgangspunkt i et av de klassiske språkene (latin, gresk) er det naturlig og ønskelig at det norske begrepet bygger på den samme stammen. Men det hender at et fremmedspråklig begrep er basert på en språklig forvanskning, en misforståelse eller en sjargong og da er det bedre å skape et godt norsk ord. To begreper er sentrale i regulerings-teknikk: regulering og styring (og tilsvarende regulere, styre og regulator, styreenhet).¹

Forskjellen i bruk av begrepene regulering og styring er ofte vanskelig å se. Vi bruker gjerne begrepet *prosessregulering*, men *fartøystyring*. På tysk er man nøye med at *steuern* er en påvirkning av en prosess etter et program eller liknende uten at det foreligger *tilbakekopling*, mens *regeln* omfatter styring på basis av tilbakekopling. Vi kan derfor gjerne si at regulering og styring i stor grad brukes om hverandre på norsk om påvirkning av en prosess både med og uten tilbakekopling.

Vi vil forklare en del begreper som er sentrale i reguleringsteknikken.

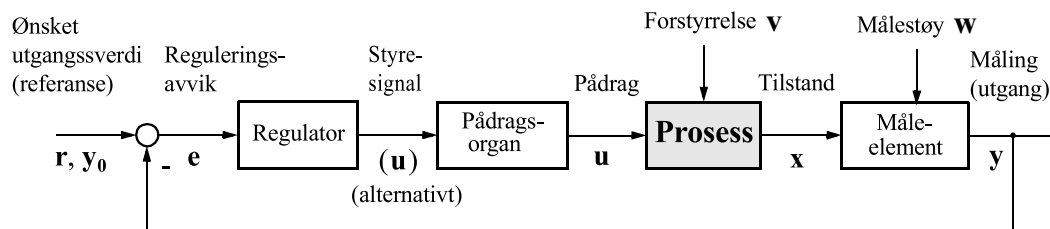
¹På tysk er begrepene *Regelung*, *regeln*, *Regler*, *Steuerung*, *steuern* og *Steuer* Gert. Oversatt til engelsk blir de *regulation*, *regulate*, *regulator* og *steering*, *steer*, *steering device*. Men det dominerende begrepet for regulering/styring er blitt *control*, og regulator blir kalt *controller*. Dette er egentlig en språklig misforståelse som går mange hundre år tilbake. Ordet kommer fra det gamle franske ordet *contre roller* som nokså nøyaktig svarer til det norske begrepet *kontrabok* som vi kjenner fra *dobbel bokføring*. Det norske begrepet *kontroll* betyr da også inspisere, verifisere, etterprøve. Det er derfor viktig å unngå å bruke det norske ordet *kontroll* i betydningen *regulere*, *styre*. Det engelske ordet *control* brukes forøvrig også i dets korrekte betydning i for eksempel *quality control* som betyr *inspeksjon av kvalitet*. Et eksempel på språklig forvirring har vi i det engelske ordet *birth control* som betyr *fødselsregulering*, det vil si påvirkning av antall fødsler, mens det norske begrepet *fødselskontroll* refererer til inspeksjon av mor og barn. Ordet *kontroll* har forøvrig en korrekt plass i reguleringsteknisk terminologi for eksempel i forbindelse med *kontrollrom* idet informasjon presenteres på instrumenter og dataskjermer for *kontroll*, dvs. inspeksjon, verifikasjon, overvåking o.l. Det har altså ikke noe å gjøre med selve regulerings- eller styringsfunksjonen, men er en samtidig aktivitet med denne.

Med et **system** skal vi i denne boka forstå en samling av enkeltenheter som påvirker hverandre gjensidig og som til sammen har en funksjon. Med andre ord er det *vårt formål* som bestemmer hva vi anser å utgjøre systemet, det vil si hvordan vi avgrenser systemet mot omverdenen (i virkeligheten påvirker jo alle “enheter” rundt oss hverandre, selv om påvirkningene kan være svake).

Dynamiske systemer er den undergruppen av systembegrepet som er viktigst for reguleringsteknikere. At et system er dynamisk betyr at dets indre tilstander endrer seg med tiden på grunn av vekselvirkningene mellom systemets enkeltenheter.

Figur 1.3 under gir en oversikt over noen sentrale begreper.

Figur 1.3
Begreper i et
regulerings-
system.



Det som utsettes for et *pådrag* og *forstyrrelser* karakteriseres ved *tilstander* og observeres gjennom *målinger*, vil betegnes som en **prosess**. Vi bruker “prosess” om det fysiske system som vi betrakter eller skal styre, mens helheten som består av prosess + regulerings-elementer da vil utgjøre et (**regulerings**)system. Eksempel: Et skip kan defineres som en prosess, mens skipet med autopilot er et system¹. En prosess kan være en del av en framstillingskjede for et produkt, f.eks. en kjemisk reaktor, metallurgisk reaktor eller en maskin som bearbeider egenskaper hos materialer eller enkeltkomponenter. En prosess kan i overført betydning være et fartøy eller en kraftmaskin. Et **reguleringsobjekt** vil ofte være et ekvivalent begrep til en prosess.

Tilstander (eller **tilstandsvariable**) er sentrale i beskrivelsen av prosesser. En tilstand endrer sin verdi som resultat av en forutgående påvirkning. Et eksempel på en tilstand er temperaturen i en reaktor. Dersom vi endrer oppvarmingen/avkjølingen eller graden av kjemiske reaksjoner i reaktoren, vil temperaturen endre seg. Et annet eksempel: hvis en motor virker på en robotarm, vil robotarmens tilstander være dens posisjon og hastighet, fordi disse størrelsene karakteriserer nå-situasjonen til robotarmen og kan endres. Siden et sett av variable kompakt kan uttrykkes i vektorform, vil begrepet **tilstandsvektor** spille en sentral rolle i resten av boka. Vi vil heretter bruke $x(t)$ for å betegne en tilstand og $\mathbf{x}(t)$ for å betegne en tilstandsvektor. Det siste bruker vi når et system har flere tilstandsvariable. En enkel

¹Nå er ikke dette helt konsekvent i litteraturen. Man bruker også “system” om “prosess”, men det går da fram av sammenhengen hva som menes. Dette forekommer også i denne boka.

fartøymodell f.eks., vil ha *to* tilstander, hastighet $x_1(t)$ og posisjon $x_2(t)$, dvs.

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Et dynamisk system vil oftest være utsatt for påvirkning fra omgivelsene, i vår terminologi kalt **forstyrrelse**. Slike påvirkninger kan være av forskjellig karakter, men har det til felles at de er *utenfor vår kontroll*. Eksempler på forstyrrelser er endring i sammensetningen av råvarer og endring i omgivelsestemperaturen til en reaktor, endring i omgivelsestemperatur rundt en bygning hvor temperaturen skal styres, endring i bølgepåvirkningen eller vindpåvirkningen på et fartøy. Vi bruker $v(t)$ for å betegne en forstyrrelse og $\mathbf{v}(t)$ for å betegne en forstyrrelsesvektor.

I motsetning til forstyrrelser har vi *tilsiktete* påvirkninger som skal utføre en styrende funksjon, blant annet å kompensere virkningen av de samme forstyrrelser. En slik påvirkning kalles gjerne et **pådrag**, se figur 1.3. Eksempler på pådrag er effekttilførsel eller kraftpåvirkning. Vi bruker symbolet $u(t)$. Dersom det finnes flere pådrag i systemet, kan pådragene samles i en **pådragsvektor** $\mathbf{u}(t)$. Noen ganger bruker vi betegnelsen **inngang(svariabel)** for pådrag.¹

Man tar ofte **målinger** av de fysiske eller matematiske tilstandene i et system (se figur 1.3). Men et system kan ha flere tilstander enn det som kan måles med det tekniske utstyret som er tilgjengelig. Målevariablene er som oftest funksjoner av systemets tilstander. Målingene kan være utsatt for **målestøy** (eller måleusikkerhet, målefeil) som forringer målingen. En “god” måling er gjerne et mål for en tilstand (eller noen få tilstander) i systemet og er lite påvirket av støy. Eksempler på målevariable er spenningen fra et termoelement (temperaturmåling), bevegelsen av en membran (trykkmåling), indusert spenning i en tacho-generator (hastighetsmåling) eller tidsavstanden mellom utsendt puls og ekko i radar og sonar (avstandsmåling). Et system som observeres gjennom flere målevariable har en **målevektor**. Vi vil heretter bruke $\mathbf{y}(t)$ for å betegne en målevektor og $y(t)$ for å betegne en måling. Noen ganger bruker vi betegnelsen **utgang(svariabel)** for måling.

I enkelte tilfeller er det behov for å skille mellom begrepene **egenskap** og tilstand. En egenskap er en funksjon av systemets tilstander. En egenskap vil ofte være det man forsøker å endre mot en ønsket verdi ved å endre systemets tilstander. Det er ofte lettere å realisere et reguleringsystem i de tilfeller der en måling gir et godt mål for en egenskap enn når egenskapen og målingen er forskjellig. Et eksempel på dette har en ved framstilling av papir der rivstyrke er en typisk egenskap, som er sterkt avhengig av cellulosefibrener lengde og

¹Noen ganger vil symbolet u bli brukt om *styresignalet* i stedet for pådraget, se figur 1.3. I slike tilfeller betrakter vi pådragsorganet som en del av den fysiske prosessen. Da vil vi ofte omtale styresignalet som “pådrag”.

fordeling (tilstand). Rivstyrken er ikke direkte målbar før papiret er ferdig framstilt, men størrelsen “freeness” som også er avhengig av fibrenes størrelsesfordeling, kan måles tidlig i prosessen. Vi bruker $z(t)$ for å betegne en egenskap og $\mathbf{z}(t)$ for å betegne en **egenskapsvektor**.

Referanse er betegnelsen på den ønskede verdien av en måling (eller en egenskap) og gis symbolet $y_0(\mathbf{y}_0)$ eller $r(\mathbf{r})$. Ofte brukes ordet **settpunkt** for det samme begrepet.

Forskjellen mellom $y_0(\mathbf{y}_0)$ og $y(\mathbf{y})$ kalles **avvik** og har symbolet $e(\mathbf{e})$.

Tabellen under gir en oversikt over de symbolene som er nevnt ovenfor. Merk at vi også bruker bestemte bokstaver for å angi dimensjoner på vektorer.

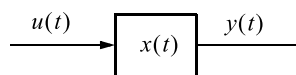
Øvrige begreper vil bli definert etterhvert som de dukker opp.

TABELL 1.1 Symboler for variable i reguleringsteknikk

Begrep	Symbol	Dimensjon
Tilstandsvariabel/tilstandsstørrelse	$x(t)$	1×1
Tilstandsvektor	$\mathbf{x}(t)$	$n \times 1$
Pådragsvariabel	$u(t)$	1×1
Pådragsvektor	$\mathbf{u}(t)$	$r \times 1$
Forstyrrelse	$v(t)$	1×1
Forstyrrelsesvektor	$\mathbf{v}(t)$	$s \times 1$
Måling	$y(t)$	1×1
Målevektor	$\mathbf{y}(t)$	$m \times 1$
Målestøy/målefeil	$w(t)$	1×1
Målestøyvektor	$\mathbf{w}(t)$	$m \times 1$
Egenskap, egenskapsvektor	$z(t), \mathbf{z}(t)$	$1 \times 1, p \times 1$
Referanse, settpunkt	$r(t), y_0(t)$	1×1
Referansevektor, settpunktsvektor	$\mathbf{r}(t), \mathbf{y}_0(t)$	$m \times 1$
Avvik, avviksvektor	$e(t), \mathbf{e}(t)$	$1 \times 1, m \times 1$

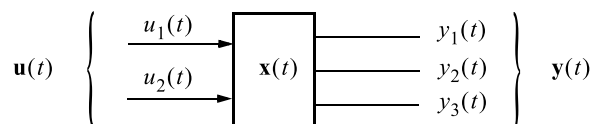
En prosess med *ett* inngangssignal og *ett* utgangssignal sier vi er **monovariabel**.

Figur 1.4
Monovariabel
prosess.



Når $\mathbf{y}(t)$ og/eller $\mathbf{u}(t)$ er en vektor, er prosessen **multivariabel**.

Figur 1.5
Multivariabel
prosess.



Vi skal i denne boka, med unntak av i kapittel 10, konsentrere oss om monovariabel systemer.

Merk at $\dim(\mathbf{y}) = m$ godt kan være forskjellig fra $\dim(\mathbf{u}) = r$.

Merk også at vi kan ha *flere* tilstander, det vil si $\mathbf{x}(t)$ er en vektor, selv om prosessen er monovariabel. Vi sier at et system har **orden** n , når antall tilstander er n ,

$$\dim(\mathbf{x}) = n \quad (1.1)$$

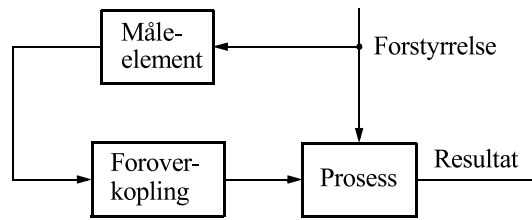
1.4 TO METODER FOR REGULERING: FOROVER- OG TILBAKEKOPLING

Vi skal presentere de to hovedprinsippene som anvendes ved styring av prosesser. Det første prinsippet er illustrert i figur 1.6 a). Et stoff, som skal bearbeides, tilføres gjennom en reguleringsventil. Prosessen utsettes for en målbar forstyrrelse slik at resultatet avviker fra det vi ønsker. I et slikt tilfelle kan vi lage en **foroverkopling** (også kalt *programstyring*) som består i at vi med utgangspunkt i den målte forstyrrelsen kan regne ut hvor mye reguleringsventilen må forstilles for at prosessens resultat skal bli det vi ønsker. Denne løsningen er svært utbredt i mange industrielle systemer og har en rekke fordeler, som vi senere skal behandle i detalj, i avsnitt 9.6. Vi kan også ha foroverkopling fra referansen

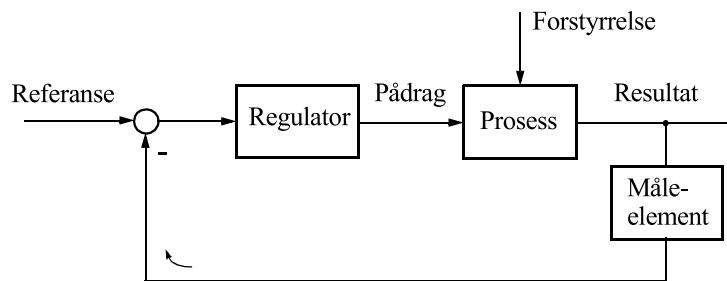
Det andre prinsippet, **tilbakekopling**, har vi allerede introdusert. Det er vist i figur 1.6 b). Vi måler prosessens tilstand ved hjelp av et *måleelement*, som gir ut et signal til en **regulator**. Der sammenliknes målingen med referansen (det ønskede resultatet). Deretter finner regulatoren ut hvor mye pådragsorganet må forstilles slik at prosessens tilstand gradvis endrer seg i riktig retning. Derved oppstår en **tilbakekopling** som medfører at prosessens resultat svinger seg inn mot den ønskede verdien.

Figur 1.6

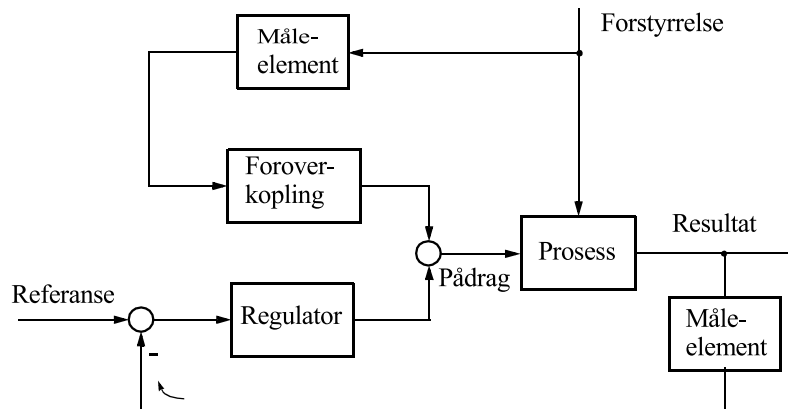
- a) Foroverkopling
- b) Tilbakekopling
- c) Kombinasjon av foroverkopling og tilbakekopling



a)



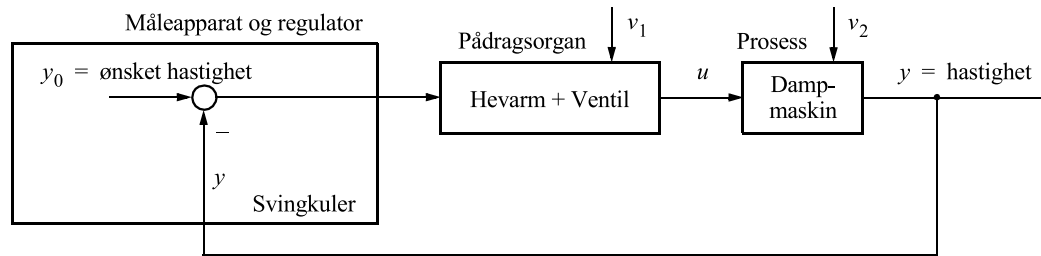
b)



c)

James Watt anvendte prinsippet med tilbakekopling da han konstruerte sin hastighetsregulator. De roterende kulene ble brukt som kombinert måleapparat og regulator. Figur 1.7 gir en skjematisk framstilling av Watts hastighetsregulator.

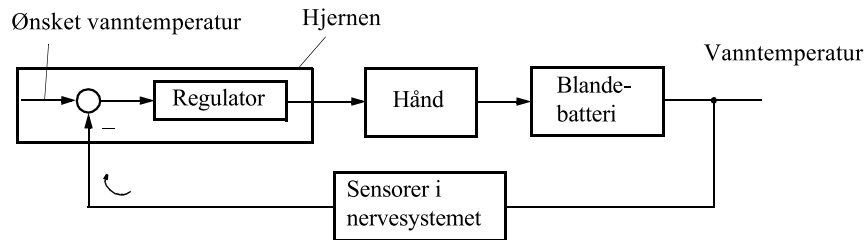
Figur 1.7
Hastighetsregulering ved tilbakekopling.



I dag er de fleste tilbakekoblede reguleringssystemer basert på måleinstrumenter som gir ut elektroniske signaler, og regulatorer implementert ved hjelp av analog elektronikk eller datamaskiner.

Vi mennesker benytter tilbakekopling i vårt daglige liv. Når vi dusjer, kjenner vi etter hvor varmt vannet er for deretter å skru på kranen som regulerer varmtvannstilførselen. Figur 1.8 viser et blokkdiagram som illustrerer dette.

Figur 1.8
Justering av vanntemperaturen i dusjen.



Vi bruker også tilbakekopling ved matlaging: Ved å smake på maten kan vi finne ut om vi skal tilsette mer av enkelte ingredienser. Vi benytter smaksorganene som sensorer, hjernen som komparator og regulator, og hånden som pådragsorgan for å tilsette salt, krydder og liknende.

En mangel ved foroverkoplingen er at vi ikke har noe mål for hva prosessens resultat egentlig er blitt. På den annen side har vi ved *tilbakekopling* ikke utnyttet et eventuelt kjennskap til forstyrrelsens størrelse som kan brukes til endring av ventilstillingen. Derfor er det egentlig svært logisk å kombinere disse to metodene slik som vist i figur 1.6 c). Vi kan gjerne si at foroverkoplingen gir den hurtige, men unøyaktige korreksjonen, mens tilbakekoplingen justerer reguleringsventilen slik at resultatet blir helt riktig.

1.4.1 Statisk prosess

En **statisk prosess** er en matematisk abstraksjon som ikke finnes i den fysiske virkelighet, men som gir oss innsikt i noen sterke og svake sider ved forover- og tilbakekopling:

Hvis sammenhengen mellom inngang u og utgang y for en prosess kan beskrives ved en vilkårlig funksjon $f(\cdot)$, er y ved tidspunktet t gitt av

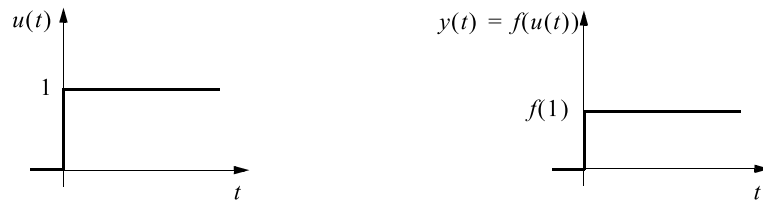
$$y(t) = f(u(t)) \quad (1.2)$$

Vi sier at denne prosessen er statisk (eller **momentan, ikke-dynamisk**, eller engelsk: **memoryless, static**).

Den statiske prosessen er, uansett formen på $f(\cdot)$, øyeblikkelig i sin respons på endringer i pådraget u . Den statiske prosessen har altså ingen "treghet", slik man finner i virkelige, *dynamiske* prosesser (som behandles nedenfor).

Anta at $u(t) = 0$ for $t < 0$ og $u(t) = 1$ for $t \geq 0$. Dette kaller vi et **enhetsprang**. Den statiske prosessen vil da få en **sprangrespons** som vist i figur 1.9.

Figur 1.9
Pådrag og sprang-
respons for en
statisk prosess.



Hvis den inverse funksjonen $f^{-1}(\cdot)$ eksisterer, kan den statiske prosessen styres perfekt ved

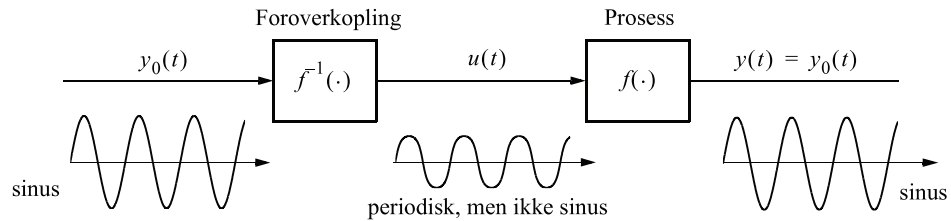
$$u(t) = f^{-1}(y_0(t)) \quad (1.3)$$

siden

$$y(t) = f(u(t)) = f(f^{-1}(y_0(t))) = y_0(t) \quad (1.4)$$

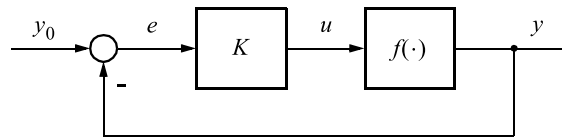
Anta at inngangssignalet (referansen) er en sinus som vist i figur 1.10. Pådraget og utgangen blir da som vist. Vi har perfekt styring siden $y(t) = y_0(t)$.

Figur 1.10
Statisk prosess
med foroverkopling
fra referansen.



Denne foroverkoplinga krever en eksakt modell av prosessen. Nå er det imidlertid svært sjelden vi har dette, så la oss prøve tilbakekopling. Den største fordelen med den er som vi skal se, at den ikke krever nøyaktig kjennskap til $f(u)$. Strukturen blir som vist i figur 1.11.

Figur 1.11
Statisk prosess
med tilbake-
kopling.



Vi har innført en forsterker K som sørger for å endre y , hvis *avviket* $e \neq 0$. Figuren viser at

$$e = y_0 - y \quad (1.5)$$

$$y = f(Ke) \quad (1.6)$$

Vi benytter den inverse funksjonen $f^{-1}(\cdot)$. Likning (1.6) løst med hensyn på e og innsatt i (1.5) gir da

$$y_0 - y = \frac{f^{-1}(y)}{K}$$

som gir det relative avviket for det regulerte systemet:

$$\frac{e}{y_0} = \frac{y_0 - y}{y_0} = \frac{1}{K} \frac{f^{-1}(y)}{y_0} \quad (1.7)$$

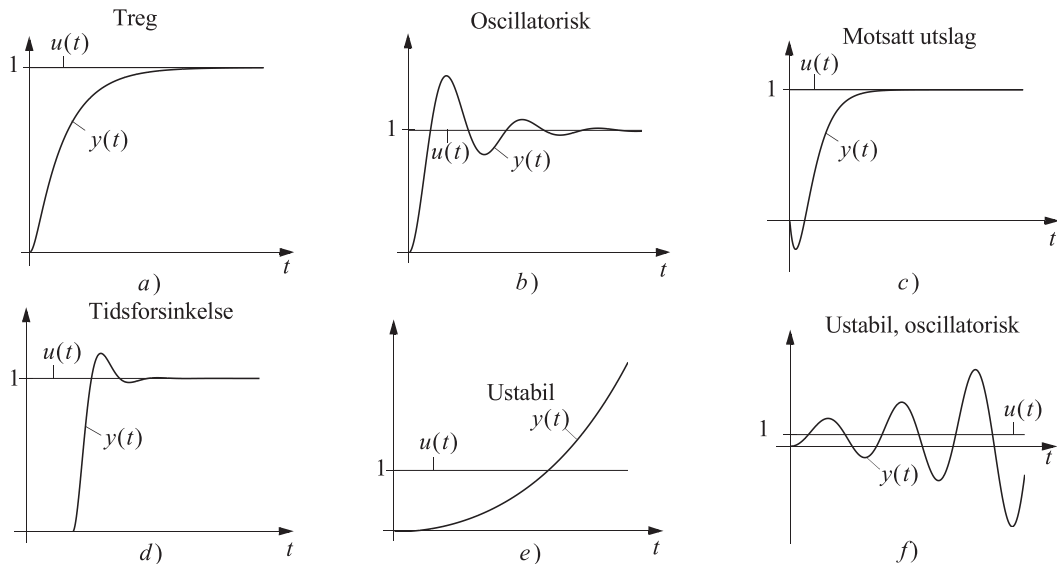
Vi ser at vi kan gjøre dette så lite vi vil ved å velge K stor nok. Dette resultatet er uavhengig av formen på funksjonen $f^{-1}(\cdot)$, d.v.s. også av $f(\cdot)$. Vi kan oppsummere at en statisk prosess tilsynelatende kan reguleres tilnærmet perfekt med tilbakekopling, ved å bruke $K \approx \infty$. Men virkelige prosesser er *dynamiske*, ikke statiske. Stor forsterkning K fører på grunn av dette til problemer, som vi skal se i neste avsnitt.

1.4.2 Dynamisk prosess

Virkelige fysiske prosesser følger til dels kompliserte bevegelseslover som gir et utall forskjellige typer responser, som har det til felles at de ikke er momentane slik den statiske prosessens respons var.

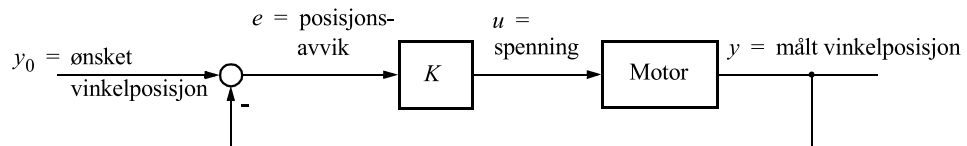
Figur 1.12 viser noen mulige responser $y(t)$ på enhetsprang i pådraget $u(t)$ for forskjellige dynamiske (dvs. ikke-momentane) prosesser.

Figur 1.12
Sprangresponser for forskjellige dynamiske prosesser.



Vellykket regulering av dynamiske prosesser krever kjennskap til prosessens bevegelseslover, dens *dynamikk*. Hvis vi for eksempel betrakter en elektrisk motor og definerer den som vår prosess, kommer vi sørgelig til kort hvis vi lager en tilbakekopling og skrur forsterkningen mot uendelig, slik (1.7) kunne friste oss til. Betrakt figur 1.13.

Figur 1.13
Motor i tilbakekoplet sløyfe.

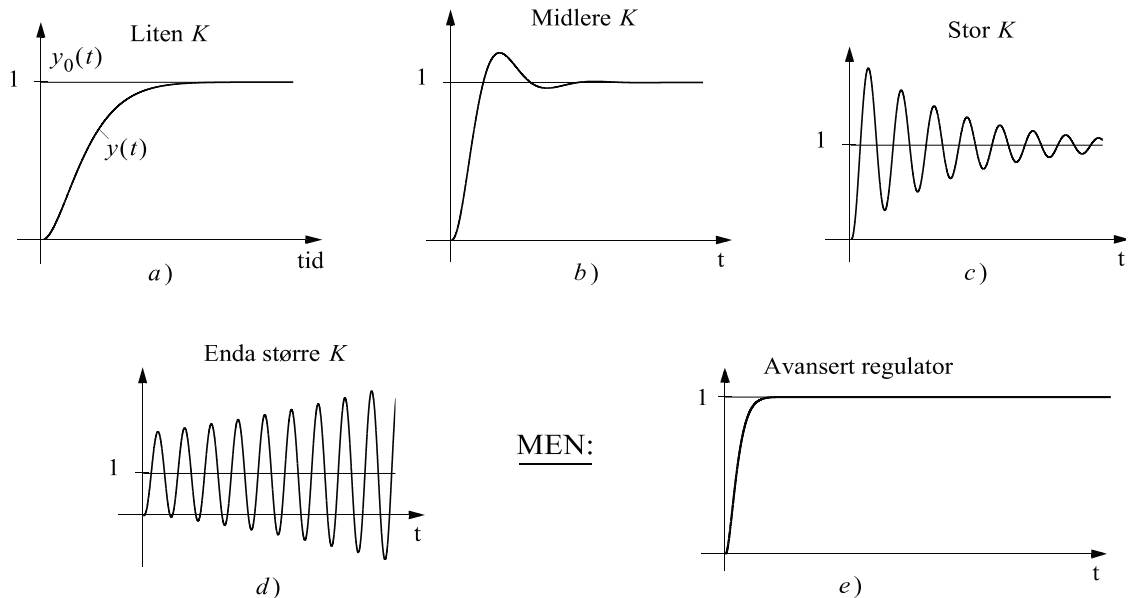


Oppgaven er å få motoren til å stille seg fort og nøyaktig inn til ønsket vinkel (dette er en aktuell problemstilling i robotteknikk). En forsterker gir spenning til motorens klemmer proporsjonalt med avviket. Sprangresponsen fra referansen $y_0(t)$ til målingen $y(t)$ for økende K blir som vist i figur 1.14.

Vi ser at responsen for liten K er langsom. I håp om raskere respons øker vi K , men da begynner utgangen y å svinge opp og ned (vi sier at y blir **oscillatorisk**). Responsen for den største K er helt ubrukelig, fordi vinkelposisjonen y får stadig kraftigere utsving. Vi sier at reguleringsystemet blir **ustabilt**.

Figur 1.14

Sprangrespons for systemet i figur 1.13 for forskjellige K .



Responsen i figur 1.14 e) indikerer hva vi kan oppnå med en annen og mer komplisert regulator enn en ren forsterker. Denne mer avanserte regulatoren er valgt og justert inn basert på en matematisk modell av motoren. Vi vil komme tilbake til dette eksemplet gjentatte ganger i denne boka.

Vi kan nå oppsummere: Skal vi regulere en dynamisk prosess, bør vi modellere den matematisk ("beskrive dens dynamikk"). På grunnlag av dette finnes så en god regulatorløsning.

Teorien som benyttes ved studier av dynamiske systemer har anvendelse innen vidt forskjellige praktiske felter. Det matematiske verktøy som brukes for å beskrive bevegelsesmønsteret til en enkel mekanisk innretning, kan være det samme som benyttes for å beskrive bevegelsesmønsteret til et romfartøy. Det samme verktøyet kan benyttes for modellering av en kjemisk reaktor og til beskrivelse av samspillet mellom flere enheter som inngår i et kompleks fabrikkasjonsystem. Dynamiske modeller av samme art kan også beskrive økonomiske, ressursmessige, befolkningsmessige og andre forhold i samfunnet.

1.5 AKTUELLE REGULERINGSTEKNISKE ANVENDELSER

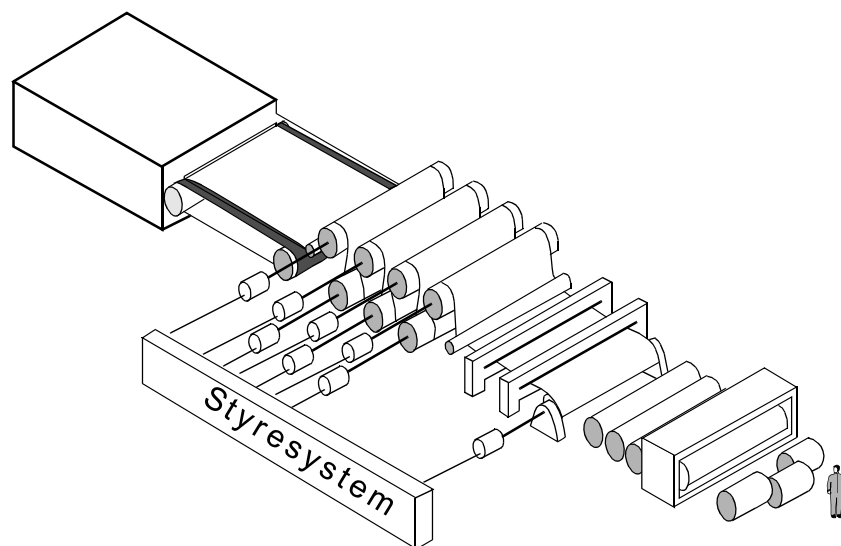
Reguleringsteoretiske metoder og reguleringsteknisk tanke sett brukes mye til styring av fartøyer og maskiner, regulering av industrielle prosesser og energisystemer og regulering i landbruk og fiske. Vi skal vise noen anvendelser innen forskjellige felter.

1.5.1 Styring og regulering av maskiner

James Watt var som nevnt tidlig ute med en automatisk regulator for å holde en dampmaskins hastighet konstant under varierende belastning. Mange tilsvarende problemer har siden oppstått og funnet sin løsning. Man kan vanskelig tenke seg en kraftstasjon med turbiner og generatorer der man ikke benytter turbinregulatorer for å holde frekvensen på nettet konstant, og å sørge for riktig produksjon av effekt til levering på nettet.

Roterende maskiner som driver produksjonsmaskineri av forskjellig art, har lenge vært gjenstand for automatisk regulering av hastighet. Dette er for eksempel tilfellet for de mange motorene som driver en moderne papirmaskin. Der er det om å gjøre at hastigheten til hver motor står i et bestemt forhold til de andre motorenes hastighet, for å oppnå høy kvalitet på papiret, og for å hindre at papirbanen rives over! Et slikt system er vist i figur 1.15 der mange og store stålsylindre drives av hver sin elektriske motor. Sylindrene tilføres vanndamp med høy temperatur innvendig, og de brukes til å tørke et kontinuerlig papirbånd som går fra den ene sylindren til den andre med stor fart. Hastigheten på hver roterende sylinder må reguleres meget nøyaktig slik at en liten, men konstant strekk-kraft opprettholdes i papiret. Dette er en meget krevende reguleringsteknisk oppgave.

Figur 1.15
Regulering av
hastigheten på en
papirmaskin.



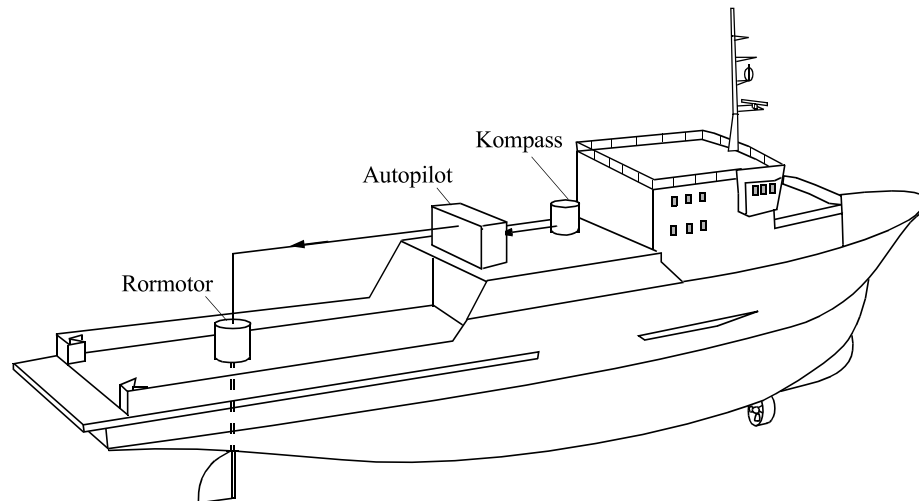
Det samme problemet forekommer også i valseverk og trykkeriprosesser og refereres gjerne til som **likeløpsregulering**.

En type motorstyring som har fått stor betydning i verkstedindustrien fra 1950-årene, er styring av verktøymaskiner. Man oppnår at verktøyet beveger seg svært nøyaktig i forhold til arbeidsstykket (eller omvendt), etter et på forhånd gitt mønster. De første styrte verktøymaskinene kunne kopiere en mekanisk mal, mens moderne løsninger benytter seg av digitale programmer som styrer motorer med stor presisjon, og utfører ønskede bevegelsesmønstre.

1.5.2 Styring av fartøyer

Automatisk styring av skip var et av de tidligste bruksområdene for reguleringstekniske grunnprinsipper. Seilskuter kan under forholdsvis stabile vind - og strømningsforhold holde kursen konstant ved hjelp av et lite styreseil som påvirker roret over en mekanisk forbindelse. En mer moderne *automatisk rorvakt* eller **autopilot** benytter seg av et kompass der avviket fra ønsket kurs påvirker en *styremaskin* (rormotor). I enkle systemer vil kompasset være et magnetkompass. Dersom man trenger stor presisjon, brukes gjerne et gyrokompass. Figur 1.16 antyder et skip utstyrt med kompass, autopilot (kursregulator), rormotor og ror. For å øke effektiviteten ytterligere ved styring av kostbare skip, er det utviklet datamaskinbaserte løsninger som styrer skipet, holder greie på trafikkforholdene og gir beskjed om mulig kollisjonsfare og andre viktige forhold.

Figur 1.16
Moderne
fartøystyring.



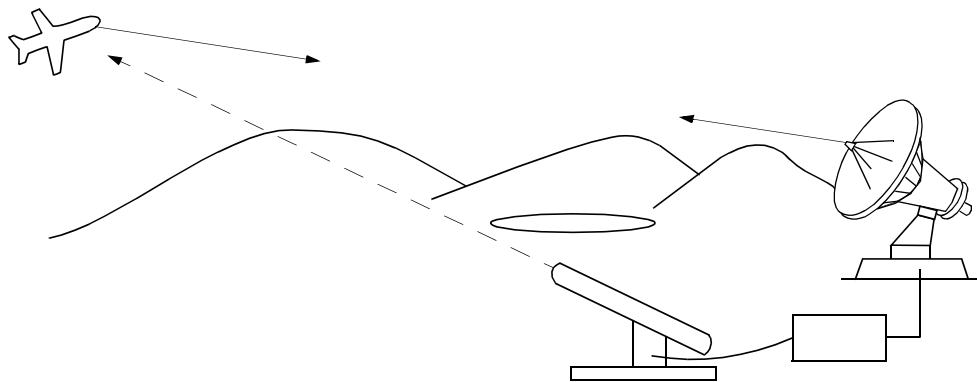
Styring og manøvrering av en moderne supertanker er ingen enkel sak, og innsikt i skipets *dynamiske egenskaper* (= bevegelseslover) er viktig for å kunne lykkes.

Maskinrommet på et moderne skip inneholder dieselmotorer, kompressorer, pumper, kjeler og tanker, og det er derfor et stort antall oppgaver som krever reguleringstekniske løsninger. Delvis ubemannede maskinrom er en realitet blant annet gjennom utstrakt bruk av reguleringstekniske metoder.

Mange andre typer fartøyer som vi møter i hverdagen ville ikke fungert uten bruk av reguleringstekniske virkemidler. Dette gjelder framfor alt moderne fly. Foruten automatisk regulering av hastighet, høyde og kurs har mange fly systemer for automatisk innflygning og landing.

Militære anvendelser av reguleringstekniske metoder er utbredt. Utvikling av moderne våpensystemer har også gitt avgjørende bidrag til moderne reguleringsteknisk teori og utstyr. Systemer for automatisk følgeing av fly ved hjelp av radar, og styring av luftskyts med omfattende datateknisk utstyr for baneberegninger (som vist i figur 1.17) er utviklet under og etter den andre verdenskrig. Etter dette følger utviklingen av styrte prosjektiler som enten kan fjernstyres mot et bevegelig mål, eller som selv inneholder komponenter for deteksjon av målet og muligheter for automatisk styring.

Figur 1.17
Styring av
luftskyts.



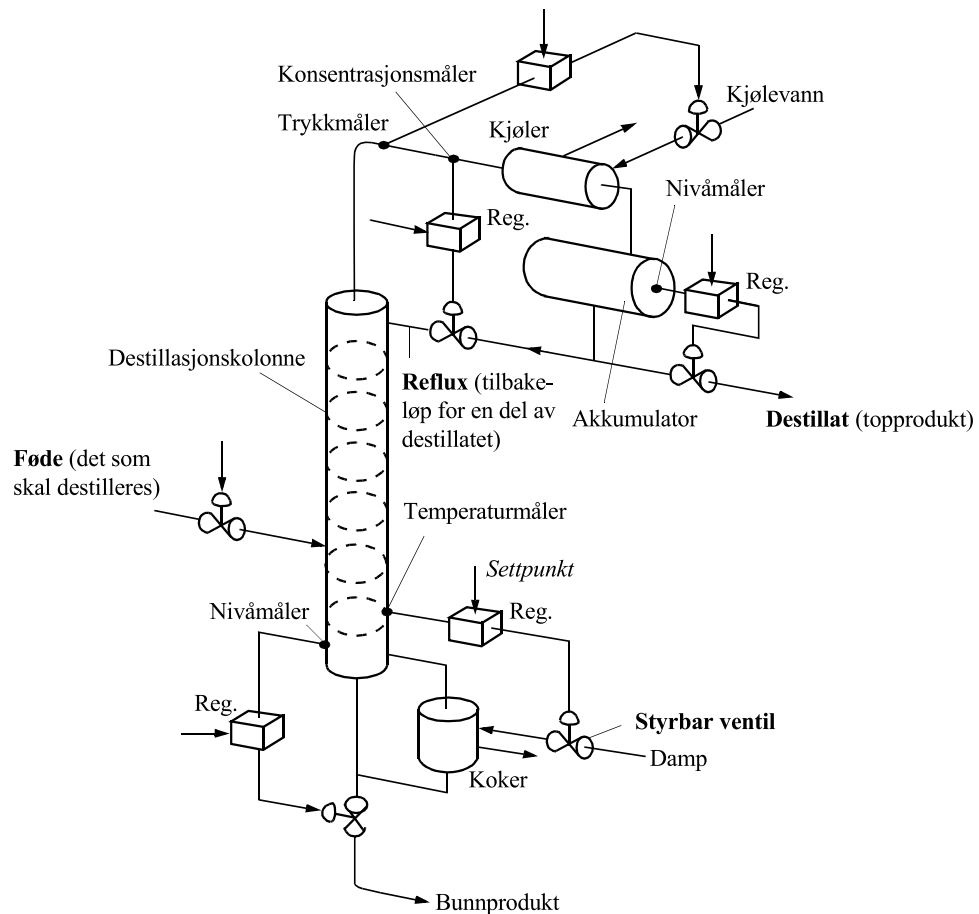
Romfart er den nyeste anvendelsen av fartøystyring. Reguleringstekniske teorier og metoder har vært avgjørende for konstruksjon av systemer for styring av raketter under oppskyting og for banestyring av romfartøyer. I likhet med militært teknisk forskning har romforskning gitt mange grunnleggende bidrag til reguleringsteoretiske metoder og komponenter.

1.5.3 Industriell prosessregulering

Papirindustri, oljeindustri, kjemisk industri, sementindustri og elektrometallurgisk industri kaller vi gjerne **prosessindustri**. Prosessindustrien er ofte karakterisert ved at råvarer og energi tilføres prosessapparatene i en mer eller mindre kontinuerlig strøm. Dette resulterer i mellomprodukter som igjen tilføres andre prosessapparater. I et stort og moderne fabrikk-kompleks for framstilling av kjemiske produkter vil en lang rekke fabrikkavsnitt være koplet tett sammen fordi avsnittene er avhengige av produkter fra de andre. Reguleringstekniske

løsninger benyttes mye i slike anlegg. En destillasjonskolonne (destillasjonstårn) der stoffer med forskjellig kokepunkt skilles fra hverandre, er et eksempel på en vanlig industriell prosess. Et slikt system er vist i figur 1.18 der det er antydnet 5 forskjellige reguleringssystemer som sørger for at driften av destillasjonskolonnen stabiliseres og holdes på optimale driftsbetingelser.

Figur 1.18
Styring av en destillasjonskolonne.



I mange tilfeller holder vi arbeidsbetingelsene i en prosess konstante ved automatisk regulering av stofftilførselen og energitilførselen basert på målinger av *prosessstilstander* (en tilstand kan for eksempel være en temperatur, en konsentrasjon eller liknende). I andre tilfeller kan det være aktuelt å benytte automatisk regulering til å stabilisere en prosess som ellers ikke kunne fungert, fordi den er ustabil. Et vanlig og moderne anlegg innen petrokjemisk industri vil kunne inneholde flere hundre reguleringstekniske undersystemer.

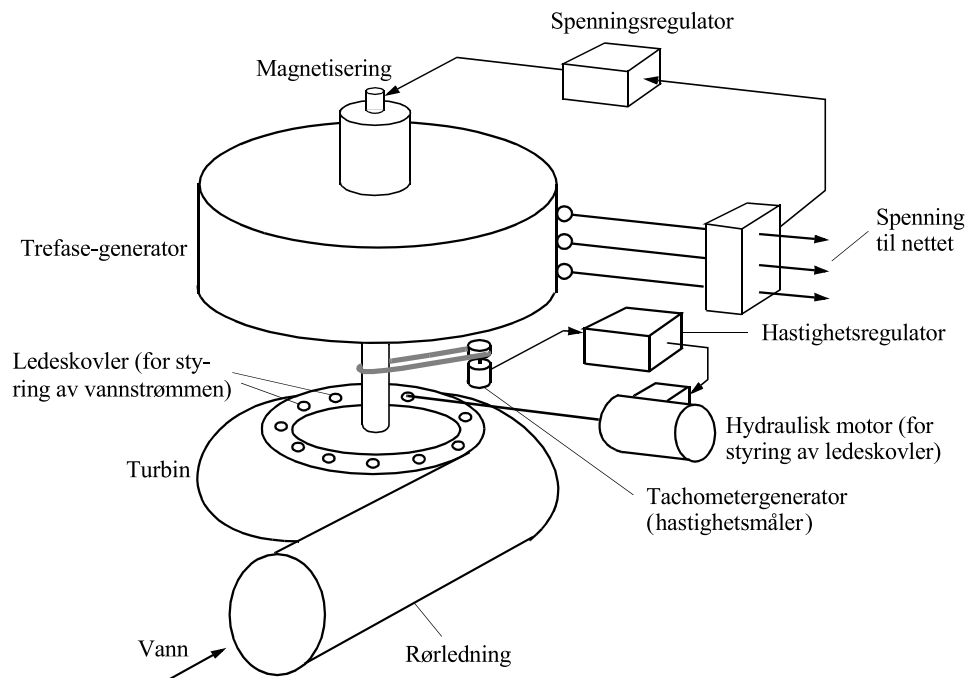
Økt produktivitet og kvalitet er viktige drivkrefter ved bruk av automatisk regulering i prosessindustrien, men hensynet til sikkerhet og arbeidsmiljø kan også være avgjørende.

Styring av et romfartøy kan i første omgang fortone seg som en helt annen type oppgave enn automatisk regulering av en framstillingsprosess f.eks. for plast. Dimensjonering av reguleringssystemer for disse to formålene bygger imidlertid på det samme teoretiske grunnlaget, noe som betyr at teorien som presenteres er anvendelig på et bredt sett av anvendelser.

1.5.4 Regulering i energisystemer

Produksjon og distribusjon av energi, i første rekke elektrisk energi, er en av de viktigste funksjoner i et moderne samfunn. Automatisk regulering spiller en avgjørende rolle i alle faser av denne prosessen. Figur 1.19 viser en **vannkraftturbin** som driver en elektrisk trefasegenerator. Turbinen er utstyrt med hastighetsregulering (omtrent som en dampmaskin) der hastigheten måles av en liten elektrisk generator, som gjennom en regulator påvirker en stor ventil i turbinens tilførselsledning. Trefasegeneratoren, som leverer sin elektriske effekt på et nett, er utstyrt med spenningsregulering ved hjelp av en elektrisk regulator som måler nettspenningen og regulerer på magnetiseringsmaskinen som igjen påvirker generatorspenningen. Det er en vanskelig oppgave å regulere kombinasjonen av turbin og generator når mange generatorer er koplet sammen i et stort nett.

Figur 1.19
Vannkraftturbin.



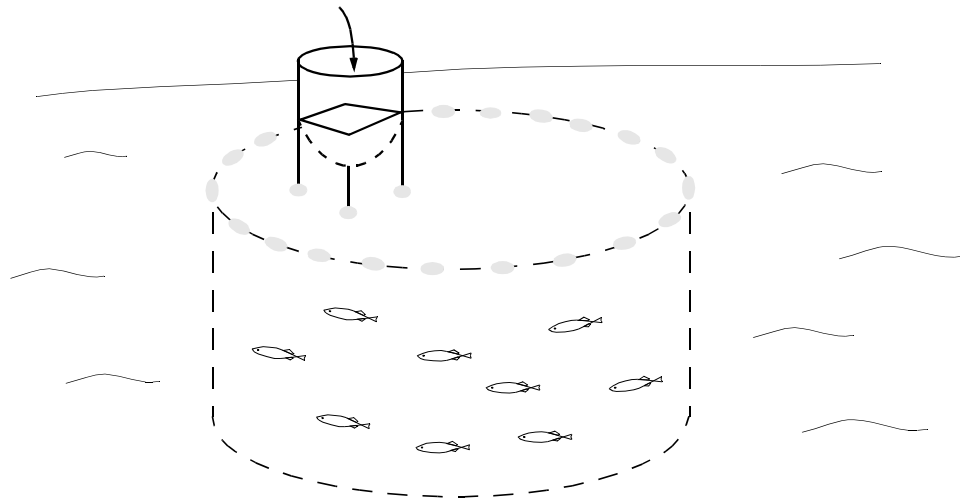
Den reguleringstekniske løsningen vil være helt avgjørende for kvaliteten på produktet som kommer fram til forbrukerne, men er også et virkemiddel til å holde et større elektrisk kraftnett sammen. Etter hvert som behovet for elektrisk energi øker, og størrelsen av produksjonsenhetene og forbrukerenhetene øker, vil det bli nødvendig med en større integrasjon av produksjonsenheter i et større geografisk område. Dette medfører et problem med hensyn til overføringsevne og overføringsstabilitet, og i denne sammenheng er automatisk regulering av stor betydning. Moderne reguleringsteknisk teori er et viktig hjelpemiddel ved løsning av slike problemer.

1.5.5 Styring og regulering i fiskeriteknologi

Reguleringstekniske metoder har en lang rekke anvendelser innen fiskeriteknologi og havbruk. Et eksempel på dette er styringen av en tråler som sleper en pelagisk trål, der bevegelsen av trålen er et resultat av trålerens hastighet, trålwirenes lengde, oterfjølenes stilling og flere andre faktorer. Skal trålen fange effektivt må disse faktorene styres, og det er et reguleringsteknisk problem.

Innen havbruk trenger vi automatisk regulering av førtilførselen til en fiskemerde avhengig av fiskens størrelse, temperaturen i vannet, tiden på døgnet og fiskens kondisjon. På basis av slike størrelser kan vi automatisk regne ut hvor mye fôr fisken skal ha, og dosere dette ut i merden ved hjelp av en fôringsautomat, som antydnet i figur 1.20. Tilsvarende løsninger utvikles innen landbruket for automatisering av fôring og andre funksjoner i husdyrhold.

Figur 1.20
Fôring av fisk.



1.5.6 Regulering i ikke-tekniske systemer

Utviklingen av teoretiske og praktiske metoder for styring av komplekse tekniske prosesser har vært en betydelig inspirasjonskilde for de som er interessert i å beskrive naturlige reguleringsmekanismer i for eksempel *biologiske og sosiale systemer*.

Det menneskelige legeme har sannsynligvis hundrevis av reguleringssystemer i seg. Noen av disse arbeider helt ned på celle-nivå, andre virker internt i et organ og atter andre styrer samspillet mellom forskjellige organer i kroppen. Fysiologer benytter i stigende grad metoder beslektet med de som brukes i reguleringsteknikken, for beskrivelse av komplekse **fysiologiske systemer**. Det bygges matematiske modeller og utføres simulering av fysiologiske mekanismer på samme måte som i reguleringsteknikken. Vi lærer dermed systemet å kjenne og kan for eksempel komme med forslag til kunstige inngrep dersom naturlige organer skulle svikte. **Estimering**¹ av fysiologiske systemers parametre ved hjelp av simuleringssystemer som et virkemiddel i diagnose, har allerede vært i bruk i noen tid. Implanterte tekniske reguleringsmekanismer til erstatning for defekte fysiologiske reguleringssystemer vil nok få økende utbredelse.

Et annet felt innen biologien er studium av for eksempel fiskens atferd i havet med sikte på utvikling av tekniske systemer for å styre atferden. Formålet med slike systemer kan være langtidslagring og gjenfangst av levende fisk for å oppnå jevn tilførsel av råstoff til fiskeforedlingsindustrien eller ved oppdrett av fisk i større fjordområder. Regulering basert på tilbakekopling fra målt atferd er da en mulighet.

De typer av dynamiske fenomener som forekommer i samspillet mellom forskjellige underenheter i et **sosialt system**, har mye til felles med fenomener i fysiske og tekniske systemer. Sosialforskere har derfor i noen grad tatt i bruk metoder fra reguleringsteknikken for å forklare og foreslå kunstige inngrep som kan forbedre situasjonen i samfunnet.

Økonomien er en side ved et sosialt system. De gjensidige påvirkninger i et komplekst **økonomisk system** har vært gjenstand for teoretiske studier i mange år. De teoretiske metodene som brukes der har mye til felles med de som forekommer i reguleringsteknikken. Resultater fra reguleringsteknisk teori har i noen grad påvirket økonomisk teori og omvendt.

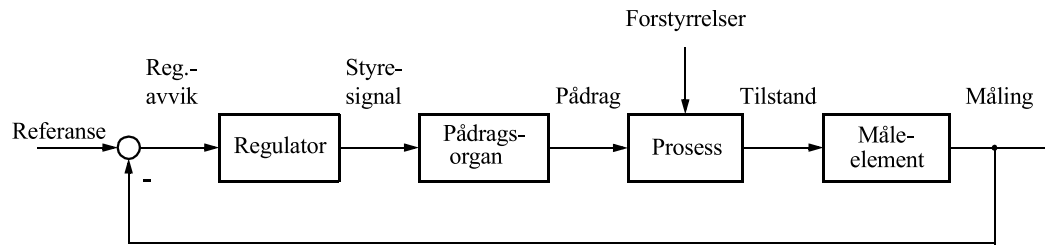
1.5.7 Oppsummering av reguleringstekniske anvendelser

Tabell 1.2 gir en oppsummering og utdypning av avsnitt 1.5. Den viser noen utvalgte eksempler på reguleringstekniske anvendelser.

Figur 1.21 viser et tilbakekoplet reguleringssystem. Begreper som benyttes i tabell 1.2 er angitt i figuren.

¹Å estimere betyr å *anslå*, det vil si vi antar en matematisk modell og bruker måledata for å finne rimelige verdier for parametre i modellen.

Figur 1.21
Tilbakekoplet
reguleringssystem.



TABELL 1.2 Utvalgte reguleringssystemer

Prosess	Referanse	Pådragsorgan/ Pådrag	Prosess- tilstand	Forstyrrelse
Destillasjonskolonne	Konstant konsentrasjon i destillat (utgående stoff)	Effekt til koker og returstrøm (reflux)	Konsentrasjoner og temperaturer	Varierende konsentrasjon og strøm i føden (tilført stoff) og varierende effekt
Luft i oppholdsrom	Konstant lufttemperatur	Effekt til ovner	Lufttemperatur	Utetemperatur og utluftning
Luftvernrakett	Målets varierende posisjon	Rakettmotor, bæreflater og ror	Posisjon, hastighet og akselerasjon	Vind, lufttrykk og avledningsbuss
Nasjonaløkonomi	Stabil utviklingsbane	Skatter, avgifter, overføringer og offentlig konsum	Bruttonasjonalprodukt og handelsbalanse	Konjunkturer, råvarepriser og endrede markedsforhold
Oljerigg med dynamisk posisjoneringssystem	Konstant posisjon, evt. varierende ved flytting	Vribare propeller og vinsj for ankerkjetting	Posisjon, hastighet og akselerasjon	Vind, strøm og bølger
Papirmaskin	Konstant papirfuktighet og papirtykkelse	Innsprøytningstrykk for papirmasse, papirbanehastighet og effekt til tørkesylindere	Fuktighet og papirtykkelse	Varierende fibregerenskaper og fuktighet i råstoff og varierende effekt

TABELL 1.2 Utvalgte reguleringssystemer (forts.)

Prosess	Referanse	Pådragsorgan/ Pådrag	Prosess- tilstand	Forstyrrelse
Robot	Bevegelses- mønster	Elektriske, hydrau- liske eller pneuma- tiske motorer	Posisjon, hastig- het og kraft	Tyngden av ro- botarm og motor og ytre krefter fra omgivelsene
Samfunn	Regjeringens målsetting	Media, propaganda, inntektsfordeling og sosial velferd	Produktivitet og sosial stabilitet	Internasjonale konjunkturer, krig og naturka- tastrofer
Undervisning av studenter	Ønsket utbytte av undervisnin- gen	Foreleser, stud. ass., øvingsoppgaver, lærebok og student- innsats	Ytelse og stem- ning	Timeplan-kol- lisjoner, for- kunnskapene bedre/verre enn antatt, UKA
Vannkraftstasjon	Konstant ampli- tude og frekvens på den genererte spenningen	Dyseåpning, mag- netiseringskrets og vanntilførsel	Spenning, turtall og effekt	Varierende vann- trykk og nettbe- lastning
Væsketank	Konstant væske- nivå	Ventil i tilførselsled- ningen	Væsknivå	Forbruk (uttap- ping) av væske

1.6 DRIVKREFTER I INDUSTRIELL AUTOMATISERING

Etter to hundre år med industriell virksomhet, kan vi se tilbake og skrive historien om hva som virkelig har skjedd. På basis av denne utviklingshistorien og en formulering av hva vi egentlig ønsker skal skje i samfunnet, burde det nå være lettere å trekke opp retningslinjer for hvordan mennesket kan påvirke den teknologiske og økonomiske utviklingen.

Enhver tidligere tidsepoke må betraktes ikke bare ut fra vår tids forutsetninger, men også ut fra datidens perspektiv. Slik vil vi bedre forstå hvorfor forskjellige hendelser og utviklingstendenser har inntruffet.

Vi kan forsøke å betrakte industrien fra tre forskjellige synsvinkler:

- Den utenforstående, ikke direkte berørte, økonomen, politiker
- Industrilederen, eieren, ingeniøren
- Industriarbeideren og hans nærmeste

Sett fra disse tre synsvinklene vil tilstanden i industrien og ønskene om framtidig utviklingsretning kunne fortone seg helt forskjellig.

Ingeniører i dag må se sin rolle i et videre perspektiv enn de gjorde i tidligere tider. Vi bør se på *industrien som et virkemiddel til å realisere menneskelige og samfunnsmessige målsetninger* snarere enn å se *mennesker som et virkemiddel til å realisere industrielle målsetninger*. Ut fra en slik holdning vil ingeniører kunne utforme tekniske systemer de er ansvarlige for på en helt annen måte.

Tre hoveddrivkrefter er dominerende i utviklingen av den industrielle automatisering:

- Ønske om økt **produktivitet** ut fra bedrifts- og samfunnsøkonomiske betraktninger. En rekke former for produktivitet, som er forholdet mellom utbytte og innsats, må defineres med referanse til forskjellige innsatsstørrelser. Vi vil møte *råvareproduktivitet*, *kapitalproduktivitet* og *arbeidskraftproduktivitet*. Av disse vil de to første kunne betegnes som *ressursproduktiviteter* og angår først og fremst samfunnet som helhet, mens *arbeidskraftproduktivitet* i høy grad også berører det enkelte mennesket.
- Ønske om bedre *kvalitet* på produkter som virkemiddel i den generelle samfunnsutviklingen, og som virkemiddel for den enkelte bedrift i en konkurransesituasjon.
- Ingeniørens, konstruktørens og utstyrsleverandørens *nysgjerrighet* og utfordring som leder til nye tekniske løsninger. En betydelig industri er vokst opp som nettopp lever av å levere automatiseringsteknikk til brukerindustrien.

I den tidligste fasen av industriell aktivitet var det ikke uttrykt noen eksplisitt målsetning om tilfredsstillende av menneskelige behov. Det menneskelige element kom inn som *bibetingelser* eller *begrensninger* framtvunget som følge av *konflikter*.

I vår tid ønsker vi å føye en fjerde og femte drivkraft til listen ovenfor som vi kan kalle:

- Ønske om å minske miljøbelastningen fra en gitt produksjonsprosess, både på miljøet i produksjonen (“det indre miljø”) og på “det ytre miljø”.
- Økt menneskelighet

I dag ønsker de fleste en eksplisitt formulering av kravet til *økt menneskelighet* allerede under planlegging og utforming av nye systemer for industriell automatisering. Dette betinger imidlertid at begrepet *menneskelighet* er nøye spesifisert. Det er for eksempel ikke sikkert at det å bli fritatt fra fysisk arbeid forbedrer folks situasjon. Det bør gjøres mer forskning på sammenhengen mellom menneskelighet og slike forhold som *fysisk og mental utfordring*, *medbestemmelsesrett*, *belønningsformer*, *prosessforståelse* osv. Men selv uten særlig mye ny innsikt kan vesentlige forbedringer gjøres ved en viss *holdningsendring* hos konstruktører av automatiseringssystemer.

Matematiske modeller av dynamiske systemer

- 2.1 Innledning og motivering 33
- 2.2 Differensiallikninger – introduksjon 40
- 2.3 Lineære systemer, linearitet 46
- 2.4 Impulsrespons og foldingsintegral 48
- 2.5 Sett av differensiallikninger – modeller av høyere orden 50
- 2.6 Én differensiallikning av høyere orden 65
- 2.7 Tidsvariante differensiallikninger 68
- 2.8 Blokkdiagrammer 70

2.1 INNLEDNING OG MOTIVERING

For å kunne forstå og regulere en dynamisk prosess utvikler vi *matematiske modeller* av den. Dette er likninger som beskriver prosessens oppførsel. En typisk matematisk modell er gitt som et sett av differensiallikninger. I dag er det vanlig å *simulere* prosesser ved hjelp av datamaskin for å få innsikt i dens egenskaper eller prøve ut reguleringsløsninger. Å simulere prosessens oppførsel (med eller uten regulering) betyr å finne den numeriske løsningen av dens matematiske modell. Det finnes i dag et stort antall data-assisterte metoder for dette.

Hensikten med å benytte matematiske modeller er for det første at det er mye enklere, billigere og sikrere å utføre tester på en matematisk modell enn på et virkelig system. Dette gjelder i særlig grad for store og kostbare systemer som et skip eller et oljeraffineri, eller følsomme systemer som et menneskelig hjerte. For det andre er det en rekke prosesser hvor det er *påkrevet* med en matematisk modell for overhodet å kunne regulere dem: Bare hvis man konstruerer regulatoren med basis i en tilstrekkelig god matematisk modell av prosessen, er man i stand til å lage et reguleringsystem som kan styre prosessen tilfredsstillende. Et eksempel her er å holde flytende oljeplattformer i posisjon over et punkt på havbunnen.

I dette kapitlet vil vi derfor gå gjennom en rekke eksempler på modeller av prosesser, og introdusere matematiske verktøy og begrep som vi trenger for modellering.

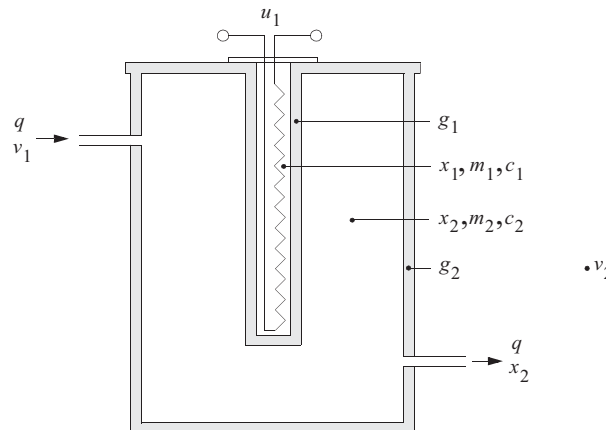
EKSEMPEL 2.1: Modellering av en varmtvannsbereder

Varmtvannsberedere er viktige både i den private husholdning og i industriell virksomhet. Alle hus har en varmtvannsbereder. Videre benytter så å si all industriell og offentlig virksomhet elektrokjeler eller oljefyrte kjeler til å framskaffe varmtvann i sin virksomhet. En slik bereder skal levere vann med en ønsket temperatur tross variasjoner i forbruk, innkommende vanntemperatur og omgivelsestemperatur.

Vi skal ikke her gå inn på hvordan velge regulatorløsning, men bare lage en matematisk modell av berederen, som er vår *prosess* i dette tilfellet. Vi skal bruke energibevaringsloven samt enkle sammenhenger for effektflyt og massestrøm.

Figur 2.1 viser varmtvannsberederen. En kald vannstrøm (q) med temperatur v_1 kommer inn øverst til venstre. Midt i beholderen sitter en varmekolbe med et varmeelement som tilføres en elektrisk effekt u_1 . Vi antar at varmekolben har en masse m_1 og en spesifikk varmekapasitet¹ c_1 . Temperaturen i varmekolben antas for enkelhets skyld å være den samme overalt, nemlig x_1 . Varmeovergangstallet mellom varmekolben og vannet er g_1 .

Figur 2.1
Varmtvanns-
bereder



Temperaturen x_2 i vannet antas å være den samme overalt i beholderen (dette kan oppnås med omrøring hvis det er behov for det, men det er uansett en rimelig forenkling², og den gjør den matematiske behandlingen enkel). Massen av vannet i beholderen betegnes med m_2 og den spesifikke varmekapasitet er c_2 . Det totale varmeovergangstallet mellom beholderen og den omgivende luften betegnes med g_2 og den omgivende luftens temperatur er v_2 . For dette systemet kan det settes opp to energibalanser: Én for varmekolben og én for beholderen med ytre isolasjon. Energibevaringsloven sier at for varmekolben gjelder følgende: Summen av den energien som tilføres, akkumuleres og bortføres må være lik null – den energi som

¹For de som ikke er kjent med disse begrepene henvises det til avsnitt 2.5.3

²“Everything should be made as simple as possible, but not simpler” – Albert Einstein

akkumuleres er lik differansen mellom den som tilføres og den som bortføres.

Akkumuleringen av energi i varmekolben foregår ved en temperaturstigning og vi finner at den akkumulerte energien pr. tidsenhet vil være

$$\frac{dE_1}{dt} = m_1 c_1 \frac{dx_1}{dt}$$

Den tilførte effekten er lik u_1 , pådraget i dette tilfellet. Den effekten som føres bort fra kolben, antas å være proporsjonal med temperaturdifferansen mellom kolben og det omgivende vannet ($x_1 - x_2$) og med varmeovergangstallet g_1 , altså $g_1(x_1 - x_2)$.

Vi finner derfor for varmekolben

$$\frac{dx_1}{dt} = \frac{1}{m_1 c_1} (u_1 - g_1(x_1 - x_2)) \quad (2.1)$$

Kolbetemperaturen x_1 er vår ene *tilstand*. For vannet i beholderen bruker vi energibevaringsloven på tilsvarende måte. Vi har at akkumulert energi pr. tidsenhet i vannet er

$$\frac{dE_2}{dt} = m_2 c_2 \frac{dx_2}{dt}$$

Tilført effekt til beholderen består av:

- Den effekten som kommer fra varmekolben, $g_1(x_1 - x_2)$.
- Den effekten som følger med det tilførte vannet til beholderen. Dersom væskestrømmen q er volumstrøm og vannets tetthet er ρ , vil den tilførte energien pr. tidsenhet med vannet bli $q\rho c_2 v_1$ ($v_1 =$ temperaturen av tilført vann, som i vår terminologi blir en *forstyrrelse*, fordi vi ikke har innflytelse på denne påvirkningen).

Den effekten som bortføres fra beholderen, består av:

- Effekten som bortføres med utstrømmende vann, $q\rho c_2 x_2$
- Effekttapet gjennom isolasjonen til omgivelsene på grunn av varmeledning, -konveksjon og -stråling. Effekttapet til omgivelsene antas å være proporsjonalt med temperaturdifferansen mellom vannet og den omgivende luften, $x_2 - v_2$, og med varmeovergangstallet g_2 . Omgivelsestemperaturen v_2 er også å betrakte som en *forstyrrelse*.

Varmebalansen for beholderen vil da gi oss

$$\frac{dx_2}{dt} = \frac{1}{m_2 c_2} (g_1(x_1 - x_2) + q\rho c_2 v_1 - q\rho c_2 x_2 - g_2(x_2 - v_2)) \quad (2.2)$$

Vanntemperaturen x_2 er vår andre *tilstand*. På basis av (2.1) og (2.2) kan vi nå stille opp

følgende matematiske beskrivelse av prosessen

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + b_{11}u_1 \quad (2.3)$$

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + e_{21}v_1 + e_{22}v_2 \quad (2.4)$$

der¹

$$\begin{aligned} a_{11} &= -\frac{g_1}{m_1c_1}, & a_{12} &= \frac{g_1}{m_1c_1} = -a_{11} \\ a_{21} &= \frac{g_1}{m_2c_2}, & a_{22} &= \frac{1}{m_2c_2}(-g_1 - q\rho c_2 - g_2) \\ b_{11} &= \frac{1}{m_1c_1}, & e_{21} &= \frac{q\rho}{m_2} & e_{22} &= \frac{g_2}{m_2c_2} \end{aligned}$$

(2.3) og (2.4) viser at vi for dette systemet har fått to koplede 1. ordens differensiallikninger av en type som vi seinere skal møte i mange sammenhenger.

La oss nå gjøre følgende betraktninger:

- Vi har foretatt *noen antakelser* for å komme fram til de to differensiallikningene ovenfor. Temperaturen i berederen og varmekolbe antas homogent fordelt. Modellens parametre antas å være konstante, noe som innebærer at volumstrømmen q som inngår i a_{22} og e_{21} , antas konstant.
- Differensiallikningene er *lineære* idet temperaturen i varmekolben og utløpstemperaturen, og effekten og omgivelsestemperaturen, inngår lineært i (2.3) og (2.4).
- Modellen trenger *initialverdier*. Siden vi har to 1.ordens differensiallikninger, trengs to initialverdier; for x_1 og x_2 . Dersom vi også har informasjon om temperaturen på tilførselsvannet v_1 og omgivelsestemperaturen v_2 , kan differensiallikningene løses.
- Der er i prinsippet mulig å finne en analytisk løsning av modellen i dette enkle eksemplet. Det er imidlertid vanlig å benytte en *numerisk integrasjonsmetode* for å løse differensiallikninger. Som vi vil se gjentatte ganger i denne boka, benytter vi programvare (MATLAB, Simulink) som løser differensiallikninger, uten at vi går inn på hvordan løsningsmetodene virker. Dette temaet behandles f.eks. i Cheney and Kincaid (1994).

¹Koeffisient-symbolene a , b og e brukes i tilknytning til henholdsvis tilstand x , pådrag u , og forstyrrelse v . Dette er valgt som standard i denne boka, og samsvarer også med mye av øvrig litteratur.

TABELL 2.1 Parameterverdier

Parameterverdier	
g_1	$3 \left[\frac{\text{kW}}{^\circ\text{C}} \right]$
m_1	10 [kg]
c_1	$5 \left[\frac{\text{kJ}}{^\circ\text{C}\cdot\text{kg}} \right]$
g_2	$0.1 \left[\frac{\text{kW}}{^\circ\text{C}} \right]$
m_2	300 [kg]
c_2	$4.184 \left[\frac{\text{kJ}}{^\circ\text{C}\cdot\text{kg}} \right]$
ρ	$1000 \left[\frac{\text{kg}}{\text{m}^3} \right]$
q	$5 \left[\frac{\text{liter}}{\text{s}} \right] = 0.005 \left[\frac{\text{m}^3}{\text{s}} \right]$

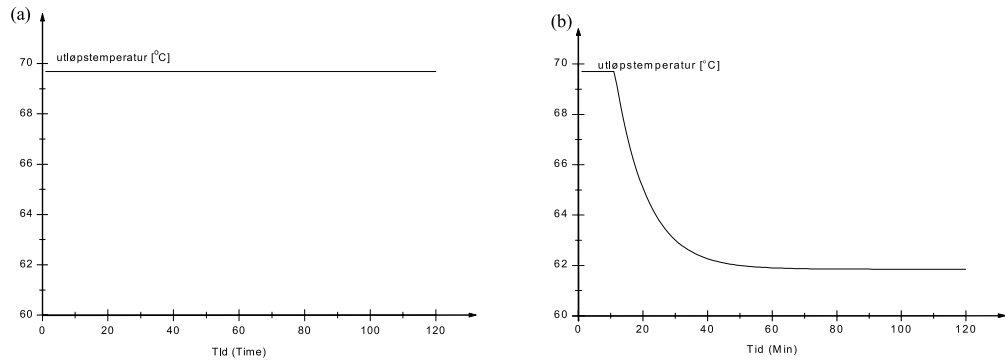
Vi skal nå vise resultatet av noen løsninger av modellen. Dette kalles gjerne å *simulere* systemet. Figur 2.2 viser løsningen for to tilfeller hvor modellparametrene er hentet fra tabell 2.1.

- Vi antar at omgivelsestemperaturen er konstant lik $20 [^\circ\text{C}]$ og tilført effekt er konstant lik 5.0 [kW] . Videre er temperaturen på tilført vann lik $5 [^\circ\text{C}]$ og forbruksraten er lik $5 \left[\frac{\text{liter}}{\text{s}} \right] = 0.005 \left[\frac{\text{m}^3}{\text{s}} \right]$. Temperaturen i kolben og vanntemperaturen i berederen har vært konstante en tid før simuleringene starter. Løsningen viser at utløpstemperaturen forblir konstant i dette tilfellet. Dette benevnes gjerne en *stasjonær løsning* til modellen.
- I dette tilfellet reduseres effekten momentant ned til 3 kW etter 10 minutter. De dynamiske fenomenene er tydelige på utløpstemperaturen.

I reguleringsteknikken vil et typisk problem være å finne en regulator som sørger for at utløpstemperaturen holder seg på et spesifisert nivå, som gjerne kan variere over døgnet. Det kan oppnås ved å benytte en måling av utløpstemperaturen til å regulere effekten på varmekolben.

I det følgende skal vi se på hvordan eksemplet ovenfor gjerne presenteres i reguleringsteknikken – det vi kaller å gi en *reguleringsteknisk beskrivelse* av problemet.

Figur 2.2
En stasjonær løøsning (a) og en dynamisk løøsning (b).

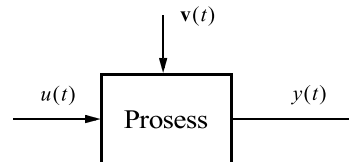


Figur 2.3 viser et generelt skjema med pådrag u , prosessutgang y , og forstyrrelser v . Dette har følgende sammenheng med eksemplet ovenfor:

- Pådraget er den eller de variable som kan påvirkes av en regulator. Dette er effekten på varmekolben.
- Prosessutgangen, eller det som vi ønsker å regulere, er utløpstemperaturen. Dette betyr at $y = x_2$. Legg merke til at temperaturen x_1 i kolben ikke måles, den er en “intern” tilstand i prosessen.
- Forstyrrelsene i dette eksemplet er innløpstemperaturen og omgivelsestemperaturen,

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Figur 2.3
Prosess med pådrag u , prosessutgang y og forstyrrelse v .



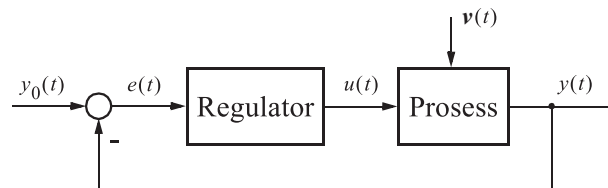
I tillegg til begrepene ovenfor, introduseres *referansen eller settpunktet* $y_0(t)$ som defineres som den ønskede verdien på prosessutgangen. I dette eksemplet vil settpunktet være ønsket utløpstemperatur.

Tilbakekoplinga vist i figur 2.4, virker som følger:

- Prosessen påvirkes av pådraget og forstyrrelsen som forklart ovenfor.
- Prosessutgangen, det vil si utløpstemperaturen, måles og *avviket* e fra settpunktet beregnes; $e = y_0 - y$.
- *Regulatoren* inneholder en algoritme som beregner pådraget, det vil si effekten, som funksjon av avviket. Den meste brukte regulatortype for dette formål er en termostat. Den kan i sin enkleste utgave beskrives som

$$u(t) = \begin{cases} 1 & e(t) > 0 \\ 0 & e(t) \leq 0 \end{cases}$$

hvor $u(t) = 0$ betyr avslått ovn, og $u(t) = 1$ angir effekt på varmekolben.



Figur 2.4
Tilbakekoplet
reguleringsløyfe
for en prosess.

■

2.1.1 Stasjonær modell

I eksemplet over har vi studert både **stasjonære** såvel som dynamiske aspekter ved en modell. Et stasjonært tidsforløp har vi når alle variable i et system holder seg på konstante verdier¹. Sammenlign figur 2.2a med 2.2b: I a-figuren vises et stasjonært forløp, mens b-figuren beskriver et dynamisk forløp. I noen tilfeller er vi bare interesserte i en stasjonær modell, det vil si en modell som gir sammenhørende verdier for prosessens variable når alle er konstante. Slike modeller kan beskrives av et sett av algebraiske likninger. Den stasjonære utgave av temperaturmodellen i (2.3) og (2.4) blir

$$a_{11}x_1 + a_{12}x_2 + b_{11}u_1 = 0$$

$$a_{21}x_1 + a_{22}x_2 + e_{21}v_1 + e_{22}v_2 = 0$$

I denne boka vil vi ta for oss dynamiske modeller. Stasjonære modeller er spesialtilfeller av disse – de følger av differensiallikningene til den dynamiske modell ved å sette alle deriverte lik null.

¹I kapittel 6 vil vi utvide dette begrepet til også å omfatte sinusformede tidsforløp.

2.2 DIFFERENSIALLIKNINGER – INTRODUKSJON

Differensiallikninger utgjør et av de mest sentrale verktøy i beskrivelse av dynamiske systemers atferd. Systemer av en type som kan karakteriseres ved at de har **konsentrerte parametre (-elementer)**, vil være beskrevet av **ordinære differensiallikninger** med *en* fri variabel (nemlig tiden). Differensiallikningene kan ha konstante eller varierende koeffisienter og de kan være lineære eller ulineære i de avhengige variablene.

Systemer med **fordelte parametre**¹, som for eksempel varmetransport gjennom kontinuerlige media, kombinert masse - og varmetransport i rørreaktorer, bølgeforplantning på elektriske linjer og i elastiske media og liknende, beskrives med **partielle differensiallikninger** der de frie variable i tillegg til tiden også kan være romlige koordinater (1, 2 eller 3 etter forholdene).

I reguleringsteknisk sammenheng vil bruk av partielle differensiallikninger sjelden forekomme, fordi en sluttet løsning av partielle differensiallikninger bare eksisterer i særdeles enkle tilfeller. Når de fordelte fenomenene ikke kan ignoreres, tar man i stedet utgangspunkt i en beskrivelse basert på partielle differensiallikninger, og approksimerer modellen ved hjelp av ordinære differensiallikninger. *I denne boka skal vi bare behandle prosesser og systemer som kan beskrives med ordinære differensiallikninger*, idet vi allerede ved oppstilling av systemlikningene skal gjøre de forenklingene som er nødvendige for å betrakte systemet som konsentrert i stedet for fordelt.

2.2.1 1. ordens differensiallikninger

Vi skal nedenfor og i avsnitt 2.5 se på utledningen av noen flere matematiske modeller av enkle prosesser, i form av differensiallikninger. Den enkleste differensiallikningen vi skal betrakte er en *1. ordens ordinær differensiallikning*. Fram til og med avsnitt 2.4 skal vi bare betrakte slike. Vi skriver

$$\frac{dx}{dt} = \dot{x} = f(x, u, t) \quad (2.5)$$

Vi har her innført et nytt symbol (en prikk over x) for den tidsderiverte av den avhengige variabelen x , og vil i stor utstrekning bruke dette i det følgende. Høyresiden i likningen er en funksjon av den avhengige variabelen x , av en ytre drivfunksjon u som antas å være en funksjon av tiden (for eksempel et pådrag). $f(\cdot)$ er også en eksplisitt funksjon av tiden selv. Funksjonen $f(\cdot)$ kan anta mange forskjellige former. Slik den står i (2.5) kalles differensiallikningen **ulineær**, **drevet** (på grunn av den ytre drivfunksjonen u) og **tidsvariant**.

¹Vanntemperaturen i eksempel 2.1 ble *antatt* å være jevnt fordelt i tanken. Da blir dette et system med konsentrerte elementer. Men hvis vi skulle vært pinlig nøyaktige (som vi ikke trengte i dette eksemplet), måtte vi modellert vanntemperaturen som avhengig både av tid og romkoordinater – ikke bare av tid. Da blir dette et system med fordelte parametre.

Dersom funksjonen (2.5) ikke endres med tiden, sier vi at den er **tidsinvariant** og skriver $\dot{x} = f(x, u)$.

Et eksempel på en ulineær, drevet og tidsvariant differensiallikning er $\dot{x} = tx + \tan(u)$, mens en ulineær, drevet og tidsinvariant differensiallikning er $\dot{x} = 2x + \tan(u)$. En annen er $\dot{x} = \sqrt{x} + u$. Her er det leddet med \sqrt{x} som gjør den ulineær.

Dersom (2.5) er lineær med hensyn til x og u ,

$$\dot{x} = a(t)x + b(t)u \quad (2.6)$$

sier vi at differensiallikningen er *lineær, drevet og tidsvariant*. (x, u betyr her som vanlig $x(t), u(t)$, slik at avhengigheten av t er underforstått.)

Dersom koeffisientene a og b er konstante, vil likningen anta formen

$$\dot{x} = ax + bu \quad (2.7)$$

og likningen refereres da til som *lineær, drevet og tidsinvariant*. Dersom den generelle likningen i (2.5) antar formen

$$\dot{x} = f(x) \quad (2.8)$$

(det vil si vi har ingen ytre påvirkning), sier vi at systemet er **autonomt**. Et spesialtilfelle har vi i det *lineære, 1. ordens, autonome* systemet

$$\dot{x} = ax \quad (2.9)$$

som representerer den enkleste differensiallikningen vi kan tenke oss.

For at vi skal kunne løse differensiallikningen i (2.5) eller en av typene som er avledet av denne, må vi i tillegg til selve differensiallikningen spesifisere begynnelsesverdien av den avhengige variabelen $x_0 = x(t_0)$.

Vi skal i det følgende for det meste konsentrere oss om tidsinvariante lineære differensiallikninger¹ og vil derfor ta utgangspunkt i (2.7).

Denne likningen løses ved først å finne løsningen til den *homogene* (autonome) likningen (2.9). Denne vet vi er

$$x_h(t) = ce^{a(t-t_0)} \quad (2.10)$$

der c er en konstant som er avhengig av begynnelsesbetingelsen x_0 ². Løsningen av den inhomogene (drevne) likningen (2.7) kan finnes ved å anta at c er en funksjon av tiden, altså

$$x(t) = c(t)e^{a(t-t_0)} \quad (2.11)$$

¹Systemer som beskrives av slike likninger kalles på engelsk *linear time invariant systems*, vanligvis omtalt som “**LTI** systems”.

²Vi kan kontrollere ved å derivere (2.10) og se at resultatet stemmer med (2.9).

Vi kan nå bestemme funksjonen $c(t)$ ved derivasjon av (2.11) og sammenlikning med (2.7). Derivasjon av (2.11) gir

$$\dot{x}(t) = \dot{c}e^{a(t-t_0)} + ac(t)e^{a(t-t_0)} \quad (2.12)$$

mens innsetting av (2.11) i (2.7) gir

$$\dot{x}(t) = ac(t)e^{a(t-t_0)} + bu \quad (2.13)$$

Sammenlikning av de to siste likningene gir

$$\dot{c}(t) = e^{-a(t-t_0)}bu \quad (2.14)$$

$c(t)$ finnes ved integrasjon av (2.14)

$$c(t) = c_0 + \int_{t_0}^t e^{-a(\tau-t_0)}bu(\tau)d\tau \quad (2.15)$$

Innsetting av (2.15) i (2.11) gir da

$$x(t) = c_0e^{a(t-t_0)} + \int_{t_0}^t e^{a(t-\tau)}bu(\tau)d\tau \quad (2.16)$$

Setter vi $t = t_0$ i (2.16), finner vi at $c_0 = x_0$. Den endelige løsningen blir da

$$x(t) = e^{a(t-t_0)}x_0 + \int_{t_0}^t e^{a(t-\tau)}bu(\tau)d\tau \quad (2.17)$$

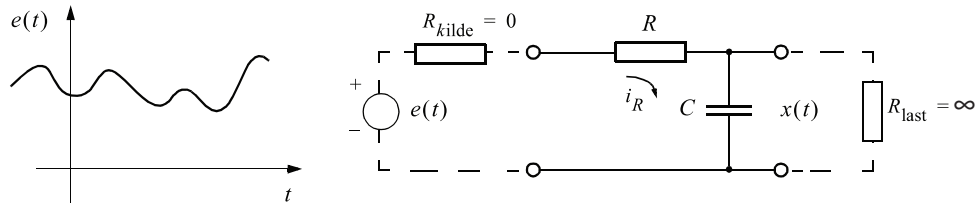
Første ledd i (2.17) ser vi er løsningen av den homogene likningen. I det andre leddet inngår den ytre drivfunksjonen $u(t)$, og vi ser at dette leddet har formen av et **foldingsintegral**. Løsningen i (2.17) er meget viktig og skal få stor betydning for det videre arbeidet. Den skal i avsnitt 3.3 generaliseres til å omfatte løsningen av et *sett* av lineære, koblede, 1. ordens differensiallikninger.

Et eksempel vil illustrere løsningen i (2.17).

EKSEMPEL 2.2: Responsen til en elektrisk krets

Figur 2.5 viser en enkel elektrisk krets som påtrykkes en vilkårlig tidsvarierende spenning e . Vi forutsetter innledningsvis at belastningsimpedansen, R_L , på utgangsklemmene er uendelig stor, slik at det ikke går strøm via utgangsklemmene. Vi forutsetter også at den indre motstanden i spenningskilden e er lik null slik at den ikke forårsaker spenningsfall. Vi skal undersøke hvordan spenningen x over kondensatoren varierer.

Figur 2.5
En RC-krets



Siden spenningsfallet i en lukket sløyfe er lik null, har vi at

$$e = e_R + e_C = Ri_R + e_C \quad (2.18)$$

Når impedansen over utgangsklemmene er uendelig, vil strømmen gjennom motstanden være lik strømmen gjennom kondensatoren

$$i_R = i_C = C \frac{de_C}{dt} \quad (2.19)$$

Likning (2.19) innsatt i (2.18) gir

$$e = RC \frac{de_C}{dt} + e_C \quad (2.20)$$

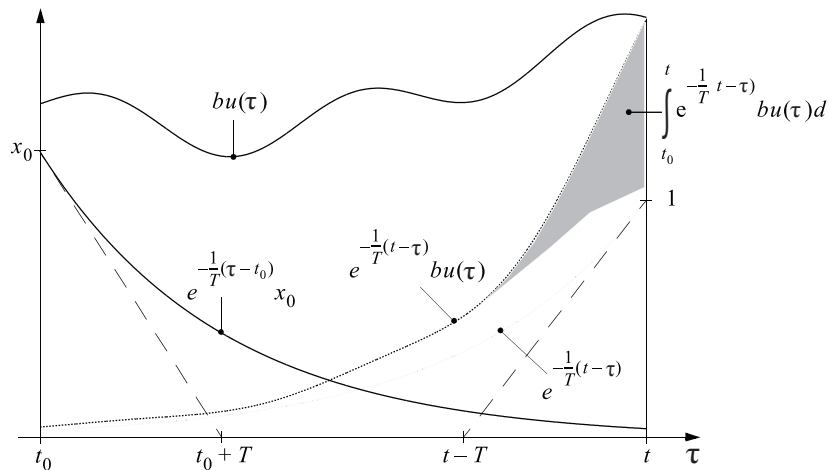
Vi setter $e_C = x$ og $e = u$ inn i (2.20) og finner

$$\dot{x} = \frac{1}{RC}(-x + u) = -\frac{1}{T}x + \frac{1}{T}u = ax + bu \quad (2.21)$$

Vi har en lineær, drevet og tidsinvariant differensiallikning.

Vi har her innført størrelsen $T = RC = -1/a$ som kalles en **tidskonstant**. Løsningen i (2.17) gjelder direkte for dette systemet. Figur 2.6 viser de faktorene som inngår i (2.17).

Figur 2.6
Faktorene som
inngår i (2.17)



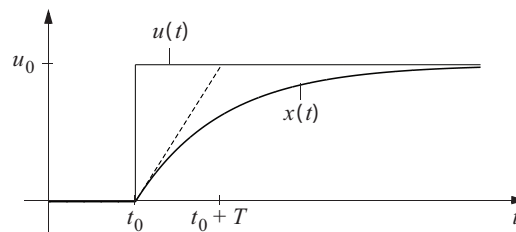
Den eksponentielt utdøende virkning av begynnelsesbetingelsen x_0 er vist til venstre der vi ser betydningen av tidskonstanten¹ T . Foldingsintegralet framgår i høyre del av figuren som arealet under den eksponentielt veide pådragsfunksjonen i tidsintervallet fra t_0 til t . Summen av disse to bidragene utgjør altså den endelige verdien x . Vi merker oss at begynnelsesbetingelsen x_0 får mindre innflytelse når t blir større.

For den elektriske kretsen vil pådraget bli et **sprang** når spenningen slås på. At $x_0 = 0$ og u er en tidsfunksjon som undergår et sprang fra null til u_0 ved $t = t_0$ er derfor et viktig spesialtilfelle. Vi får ved bruk av (2.17)

$$x = \int_{t_0}^t e^{-\frac{t-\tau}{T}} d\tau = u_0 \left[1 - e^{-\frac{t-t_0}{T}} \right] \text{ for } t > t_0, 0 \text{ ellers} \quad (2.22)$$

Figur 2.7 viser denne løsningen. Spenningen x bygger seg gradvis opp (karakterisert ved tidskonstanten T) fra null mot u_0 .

Figur 2.7
Respons ved
sprang i $u(t)$



■

¹I figur 2.6, merk at tangenten til den eksponentielt utdøende delen av responsen i punktet (t_0, x_0) skjærer tidsaksen i $t_0 + T$. Vis at det er slik! Denne sammenhengen er nyttig ved rask skissering av eksponentielle tidsforløp.

EKSEMPEL 2.3: Kapitalakkumulasjon

Vi skal lage en matematisk modell av formuesveksten til *en rentenist*.

En rentenist er karakterisert ved at han/hun utelukkende har finansielle inntekter, det vil si renter på utlån og avkastning på investeringer. For enkelhets skyld antar vi at rentenisten bare har *en* type plassering av sine penger, med *en* konstant rentesats, *en* nedbetalingstid, etc.

Vi definerer

x = formue [kr]

r = rentesats [1/år] (Vi bruker ikke %, men desimaltall; dvs. 0.1 i stedet for 10%)

d = tilbakebetalingsfaktor [1/år] (Andel av restgjeld betalt pr. år)

s = rentenistens sparetilbøyelighet [·]

c = rentenistens konsumtilbøyelighet [·]

s og c er dimensjonsløse størrelser. De antas konstante. s angir andelen av løpende inntekt som rentenisten vil reinvestere (for å få formuen til å vokse), mens c angir den resterende andel, som rentenisten vil bruke til konsum (=forbruk), altså penger som ikke anvendes slik at de bidrar til å øke formuen ytterligere. Vi har $s + c = 1$.

Vi får følgende sammenheng:

$$\begin{aligned} \text{[formuesøkning/tidsenhet]} &= \text{[samlet pengestrøm inn]} \\ \dot{x} & \quad (r+d)x \\ & - \text{[pengestrøm til konsum]} \\ & \quad c(r+d)x \\ & - \text{[reduksjon p.g.a. tilbakebetaling av lån]} \\ & \quad dx \end{aligned}$$

$$\dot{x} = ((r+d) - c(r+d) - d)x \quad (2.23)$$

Som kan skrives slik

$$\dot{x} = (-d + s(r+d))x \quad (2.24)$$

Dette er en autonom 1. ordens differensiallikning, jfr. (2.9). Løsningen blir

$$x = e^{\lambda t} x(0) \quad (2.25)$$

der

$$\lambda = -d + s(r+d) \quad (2.26)$$

Hvis $\lambda > 0$, vil formuen vokse. Dette svarer til betingelsen (som kan formuleres på forskjellige måter):

$$s > \frac{d}{r+d}, \quad \text{eller} \quad (2.27)$$

$$c < \frac{r}{r+d}, \quad \text{eller} \quad (2.28)$$

$$rs > (1-s)d \quad (2.29)$$

Vi ser at høy rentesats gir formuesvekst, som forventet.

Det er mer interessant at en høy sparetilbøyelighet har en kraftigere effekt enn rentesatsen, jfr. (2.29). Vi ser også at tilbakebetalingstiden, som er $1/d$, betyr mye: Hvis rentenistens plasseringer har lang tilbakebetalingstid (liten d), vil formuen vokse kraftigere enn om d er stor. Hvis plasseringen er i aksjer (som har $d=0^1$), vil formuen vokse *uansett* avkastningsnivå (rentesats) og sparetilbøyelighet (her ser vi bort fra inflasjon og eventuelle tap).



2.3 LINEÆRE SYSTEMER, LINEARITET

I denne boka vil vi for det aller meste ta for oss *lineære dynamiske systemer*. Et system er **lineært** når det tilfredsstiller følgende betingelse:

Anta to vilkårlige inngangssignaler u_1 og u_2 , som hver for seg fører til responsene x_1 og x_2 . Vi danner et nytt inngangssignal

$$u = \varphi u_1 + \rho u_2$$

der φ og ρ er vilkårlige konstante koeffisienter. Hvis den resulterende responsen da blir

$$x = \varphi x_1 + \rho x_2$$

er systemet lineært.

Vi sier at lineære systemer tilfredsstillers **superposisjonsprinsippet**, fordi vi kan legge sammen (superponere) delresponsene for å få responsen på et sammensatt inngangssignal. Dette er en svært nyttig egenskap som vi vil benytte mye.

¹En aksjepost i en bedrift kan betraktes som et lån til bedriften, et lån som ikke betales tilbake. Aksjeutbyttet svarer da til lånerenta.

Hvis vi nå betrakter (2.6) med u_1 og u_2 som inngangssignaler og x_1 , henholdsvis x_2 , som utgangssignaler, har vi

$$\begin{aligned} \dot{x}_1 &= ax_1 + bu_1 \\ \dot{x}_2 &= ax_2 + bu_2 \end{aligned} \tag{2.30}$$

(a og b kan være konstante eller tidsavhengige).

Vi danner $x = \varphi x_1 + \rho x_2$ og setter inn i (2.30)

$$\begin{aligned} \dot{x} &= \varphi \dot{x}_1 + \rho \dot{x}_2 \\ &= \varphi ax_1 + \varphi bu_1 + \rho ax_2 + \rho bu_2 \\ &= a(\varphi x_1 + \rho x_2) + b(\varphi u_1 + \rho u_2) \\ &= ax + bu \end{aligned} \tag{2.31}$$

Lineære differensiallikninger gir altså lineære systemer!¹

Hvis vi derimot betrakter likningene

$$\dot{x} = -x^2 + u \quad \text{eller} \quad \dot{x} = x + \tan(u)$$

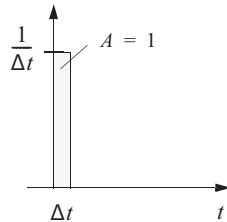
så er linearitetsbetingelsene ikke oppfylt: Ulineære differensiallikninger gir ulineære systemer.

¹Det finnes også lineære systemer som *ikke* beskrives ved lineære differensiallikninger. Et enkelt eksempel er systemet $x(t) = u(t - \tau)$ der τ er et konstant tidsintervall, det vil si hvor utgangssignalet ganske enkelt er en tidsforsinket utgave av inngangssignalet. Også dette systemet tilfredsstiller linearitetsbetingelsene. *Impulsresponsen* til systemet, se avsnitt 2.4, er $h(t) = \delta(t - \tau)$.

2.4 IMPULSRESPONS OG FOLDINGSINTEGRAL

Et spesielt inngangssignal spiller en viktig rolle i reguleringsteknisk teori: **impulsfunksjonen** (den kalles også **deltafunksjon** eller **Dirac-puls**). For dette signalet bruker vi symbolet $\delta(t)$. Funksjonen $\delta(t)$ kan defineres som grenseverdien av signalet vist i figur 2.8 når $\Delta t \rightarrow 0$; med andre ord en uendelig høy og smal puls med areal $A = 1$.

Figur 2.8
Dirac-puls



$\delta(t)$ er åpenbart en matematisk abstraksjon, siden ikke noe fysisk signal kan ha uendelig amplitude og null varighet. Men enkelte fysiske signaler kan tilnærmet betraktes som impulsfunksjoner, for eksempel elastiske støt i mekanikken og spenningspulser i elektroteknikken.

Vi betrakter nå det generelle systemet i figur 2.9. Systemet er i ro og $x(t) = 0$ for $t \leq 0$.

Figur 2.9
System som er i ro eksiteres av en impuls



Hvis $u(t) = \delta(t)$, kaller vi $x(t)$ **impulsresponsen** for systemet. Denne symboliseres generelt med betegnelsen $h(t)$.

Vi betrakter nå (2.17), og forutsetter at vårt system har $x_0 = 0$ ved $t = 0$. Systemet utsettes for impulsen $u(t) = \delta(t)$. Da får vi impulsresponsen

$$h(t) = \int_0^t e^{a(t-\tau)} b \delta(\tau) d\tau \quad (2.32)$$

På grunn av $\delta(\tau)$ er integranden null for alle $\tau \neq 0$. For $\tau = 0$ får vi integralet av en impuls (areal 1) multiplisert med faktoren $e^{at}b$ som kan settes utenfor integralet. Derfor blir impulsresponsen i dette tilfellet

$$h(t) = e^{at}b \quad (2.33)$$

og vi kan skrive (2.17) slik

$$x(t) = e^{a(t-t_0)}x_0 + \int_{t_0}^t h(t-\tau)u(\tau)d\tau \quad (2.34)$$

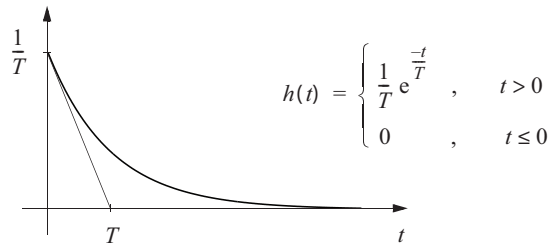
Når $x_0 = 0$ for $t = 0$, får vi

$$x(t) = \int_0^t h(t - \tau)u(\tau)d\tau = h(t) * u(t) \quad (2.35)$$

Med andre ord: utgangssignalet er lik inngangssignalet foldet med impulsresponsen. Systemet er dermed entydig gitt gjennom sin impulsrespons (Folding kalles også **konvolusjon**. Operatoren $*$ indikerer folding).

Vi vil senere få se i avsnitt 3.4.1 at (2.35) gjelder for vilkårlige lineære systemer av høyere orden, ikke bare for slike som kan beskrives ved hjelp av en 1. ordens differensiallikning. Impulsresponsen for den elektriske kretsen i eksempel 2.2 blir i følge (2.33) som vist i figur 2.10.

Figur 2.10
Impulsrespons for
den elektriske
kretsen



Vi kan i dette tilfellet tenke på $\delta(t)$ som et kraftig kortvarig spenningsstøt på kretsens inngang. Spenningen over kondensatoren vil da få et forløp som i figur 2.10.

Ved hjelp av $h(t)$ kan vi nå gi en intuitiv forklaring på fenomenet folding: Anta at systemet utsettes for et *vilkårlig* inngangssignal $u(t)$, og at impulsresponsen $h(t)$ er kjent. Vi spalter opp tidsaksen fra 0 til t i N like store, korte intervaller Δt , slik at $t = N\Delta t$. Vi antar at vi kan betrakte $u(t)$ som konstant innen hvert intervall.

For lineære systemer gjelder *superposisjonsprinsippet*. Dermed kan vi finne responsen $y(t)$ som *summen* av *delresponsene* $y_i(t)$, hvor hver delrespons skyldes en smal rektangulær puls $u(t_i)$ som antydnet i figur 2.11.

Dersom Δt er liten, vil hver av disse smale pulsene tilnærmet være en deltafunksjon med areal $= u(t_i)\Delta t$. Delresponsen $y_i(t)$ blir da $h(t - t_i)u(t_i)\Delta t$. Totalresponsen blir

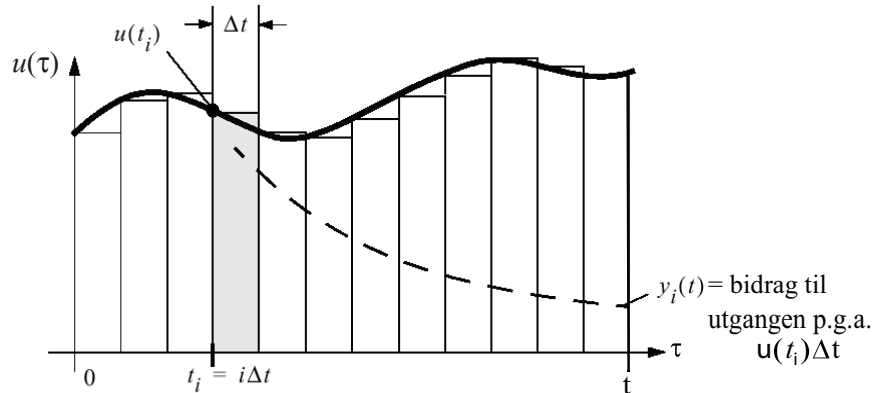
$$y(t) = \sum_{i=0}^{N-1} y_i(t) = \sum_{i=0}^{N-1} h(t - t_i)u(t_i)\Delta t \quad (2.36)$$

som når $\Delta t \rightarrow 0$ og $N \rightarrow \infty$ blir *foldingsintegralet*

$$\int_0^t h(t-\tau)u(\tau)d\tau = h(t) * u(t) \quad (2.37)$$

Figur 2.11

Stykkevis konstant tilnærming til $u(t)$.
(Merk at løpende tid er τ , mens t er et gitt punkt på tidsaksen, dvs. “nå”)



2.5 SETT AV DIFFERENSIALLIKNINGER – MODELLER AV HØYERE ORDEN

Med unntak av det innledende eksemplet med varmtvannsberederen, har vi til nå betraktet **modeller av 1. orden**, dvs. slike som kan beskrives av én differensiallikning av 1. orden. Varmtvannsberederen ble beskrevet av en *modell av 2. orden* (vi sier også et “2. ordens system”), dvs. et *sett* av to 1. ordens differensiallikninger i to variable. Vi kan selvsagt ha systemer av enda høyere orden. I dette avsnittet skal vi studere flere modeller, av orden 2 eller 3.

Før vi gir eksempler på slike, skal vi oppsummere de fysiske lover vi bruker for å kunne stille opp slike sett av differensiallikninger.

2.5.1 Fysiske balanser

En matematisk beskrivelse av fysiske prosesser (reaktorer, maskiner, fartøy og liknende) vil i store trekk omfatte oppstilling av flere typer **fysiske balanser** som foreligger på forskjellige steder i prosessen og som til sammen bestemmer den dynamiske oppførselen. Det kan formuleres en rekke forskjellige typer av fysiske balanser, men for de prosessmodellene som er mest aktuelle for reguleringstekniske problemstillinger, er det fire typer fysiske balanser som er dominerende.

Disse er:

- Masse (stoff)-balanse
- Energibalanse
- Kraftbalanse (som har en parallell i momentbalanse ved roterende bevegelse)
- Elektrisk balanse

Masse- og energibalanse bygger på et konserveringsprinsipp. Konserveringslovene uttrykkes som følger

$$\begin{aligned} \left[\begin{array}{c} \text{Netto akkumulert} \\ \text{i et fast volum} \end{array} \right] &= \left[\begin{array}{c} \text{Netto transportert} \\ \text{inn i volumet} \end{array} \right] - \left[\begin{array}{c} \text{Netto transportert} \\ \text{ut av volumet} \end{array} \right] \\ &+ \left[\begin{array}{c} \text{Netto generert} \\ \text{i volumet} \end{array} \right] - \left[\begin{array}{c} \text{Netto forbrukt} \\ \text{i volumet} \end{array} \right] \end{aligned} \quad (2.38)$$

Kraftbalanse og elektrisk balanse bygger på likevektslover

$$\sum F_i = 0 \quad (2.39)$$

F_i kan være krefter, strømmer inn mot et knutepunkt osv.

Vi skal ved hjelp av eksempler vise hvordan disse fire typene balanser kan være til hjelp ved oppsetting av matematiske relasjoner for en del enkle prosesser. I mange prosesser vil ikke *en* type balanse alene være tilstrekkelig for å beskrive prosessens oppførsel. Flere typer balanser og andre fenomener må i mange tilfeller trekkes inn.

2.5.2 Massebalanse (stoffbalanse)

Massebalansen uttrykker i sin enkleste form at stoff ikke forsvinner, men kan forflyttes, akkumuleres og i visse tilfeller (for eksempel i kjemiske reaktorer) omdannes til andre stoffer via kjente reaksjoner¹. Et enkelt eksempel vil illustrere fenomenet stoffbalanse.

¹Generelt - og mer komplisert enn vi går inn på i denne boka - har vi balanselikningen

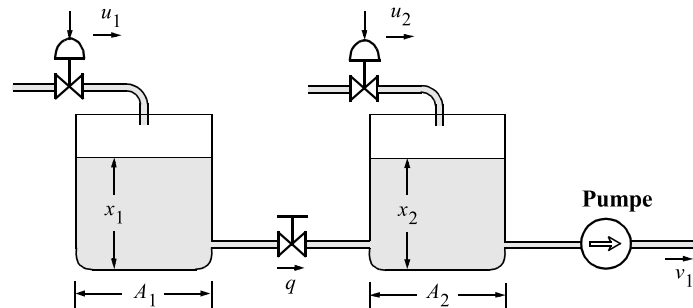
$$\frac{d}{dt} \iiint_V \rho dV = - \iint_A \rho \mathbf{v}^T \mathbf{n} dA + m'$$

der dV er et volumelement, ρ er tettheten til substansen vi betrakter, dA er et overflateelement til V , \mathbf{n} har lengde 1 og er den utadrettede normalen til overflaten, \mathbf{v} er hastighetsvektoren for substansen i et vilkårlig punkt og m' er genereringsraten minus forbruksraten til substansen.

EKSEMPEL 2.4: Volumbalanse for en væskestrøm

Figur 2.12 viser et hydraulisk system der to kar som kommuniserer gjennom en ventil (væskestrøm q), tilføres uavhengige væskestrømmer (u_1 og u_2). Fra karet til høyre bortføres en væskestrøm (v_1) ved hjelp av en pumpe.

Figur 2.12
To kar koplet
i serie



For dette systemet kan vi sette opp en stoffbalanse for hvert av de to karene der vi benytter oss av det enkle forholdet at summen av væskemengdene som tilføres, bortføres og akkumuleres må være lik null. Dette kan også uttrykkes på følgende måte:

- Differansen mellom det som tilføres og det som bortføres er lik det som akkumuleres.

Vi antar først at de to karene er sylindriske med grunnflate henholdsvis A_1 og A_2 . Når nivåene i de to karene betegnes med henholdsvis x_1 og x_2 , finner vi at den stoffmengden (volumet) som akkumuleres i venstre kar per tidsenhet, må være

$$A_1 \frac{dx_1}{dt}$$

Stoffbalansen for venstre kar gir da

$$\frac{dx_1}{dt} = \frac{1}{A_1} (u_1 - q) \quad (2.40)$$

For høyre kar finner vi tilsvarende

$$\frac{dx_2}{dt} = \frac{1}{A_2} (u_2 + q - v_1) \quad (2.41)$$

Vi legger merke til at væskestrømmen (volumstrømmen) som bortføres fra det venstre karet, tilføres det høyre.

Væskestrømmen som utveksles mellom de to karene (q) er et resultat av den hydrostatiske trykkdifferansen som oppstår på grunn av nivåforskjellen mellom de to karene. Væskestrømmen er derfor en funksjon av differansen mellom de to nivåene.

Fra hydrodynamikken har vi at sammenhengen mellom væskestrøm og trykkfall med god tilnærming kan skrives

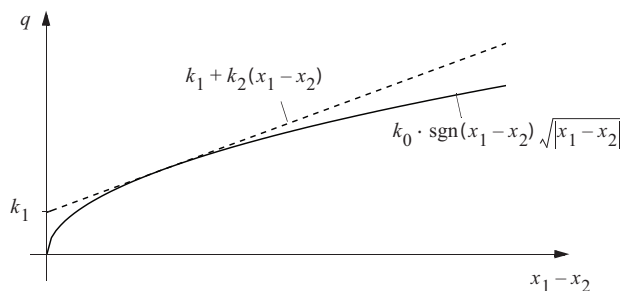
$$q = k_0 \cdot \operatorname{sgn}(x_1 - x_2) \sqrt{|x_1 - x_2|} \quad (2.42)$$

Dette¹ er en ulineær relasjon der parameteren k_0 er bestemt av ventilåpningen, rørdimensjoner og liknende. I nærheten av en bestemt verdi av differansen $(x_1 - x_2)$ kan (2.42) tilnærmes med følgende lineære sammenheng

$$q \approx k_1 + k_2(x_1 - x_2) \quad (2.43)$$

Dette er vist i figur 2.13.

Figur 2.13
Ventilkarakteristikk



Figur 2.13 er generert i MATLAB. Koden finnes i figur 2.14.

Benytter vi (2.43) i (2.40) og (2.41) finner vi

$$\frac{dx_1}{dt} = \frac{1}{A_1} (-k_2 x_1 + k_2 x_2 + u_1 - k_1) \quad (2.44)$$

$$\frac{dx_2}{dt} = \frac{1}{A_2} (k_2 x_1 - k_2 x_2 + u_2 - v_1 + k_1) \quad (2.45)$$

Vi er kommet fram til to 1. ordens differensiallikninger som er koblede fordi de avhengige variablene forekommer i begge likningene. Hvordan disse likninger kan løses skal behandles i senere avsnitt.

(2.40) og (2.41) var uttrykk for stoffbalansen i de to karene mens (2.42) i virkeligheten er et uttrykk for kraftbalansen (trykkbalansen over forbindelsesrøret). Vi har derfor her et eksempel på forekomsten av to typer fysiske balanser i samme system.

(2.44) og (2.45) er et eksempel på et sett av lineære differensiallikninger som er *en tilnærming til et system som er ulineært*. Dette kaller vi **linearisering**, som er behandlet nærmere i avsnitt 3.6.

¹ $\operatorname{sgn} x = +1$ når $x > 0$ og -1 når $x < 0$.

Figur 2.14
MATLAB-kode

```

1 k0 = 2;
2 k1 = 0.5;
3 k2 = 2;
4 N = 100;
5 dx = 0.01;
6 x1_x2 = [0: dx: N*dx]; ← Lager vektoren x1-x2
7
8 q1 = k0*sqrt(x1_x2);
9 q2 = k1+k2*x1_x2;
10
11 plot(x1_x2, q1, x1_x2, q2, '--', 'linewidth', 1.6)

```

Plotter q₁ og q₂ som funksjon av vektoren x₁-x₂



2.5.3 Energibalanse

Energibalansen uttrykker at energi ikke forsvinner, men kan forflyttes, akkumuleres eller omdannes til en annen form.

For de energibalansene vi skal sette opp i denne boka, er det kun nødvendig å kjenne til begrepene *varmekapasitet* og *varmeovergangstall*. **Varmekapasitet** har benevning [Joule/grad K] = [J/(°K)] og symboliseres ofte med bokstaven *C*. Varmekapasiteten forteller hvor mye energi (Joule) som må tilføres en gitt mengde stoff eller gjenstand for at dens temperatur skal øke med en grad Kelvin (Celcius)¹. Man kan også oppgi *spesifikk varmekapasitet*, ofte symbolisert med liten *c*, og med benevning [J/(°K · kg)]. Vi har da $C = mc$ der $m[\text{kg}]$ er mengden av stoff.

Varmeovergangstallet symboliseres ofte med *g* og har benevning [W/(°K)] = [J/(s · °K)]. Varmeovergangstallet forteller hvor stor effekt [W] som strømmer over et grensesjikt ved en temperaturforskjell på en grad Kelvin.

2.5.4 Kraftbalanse

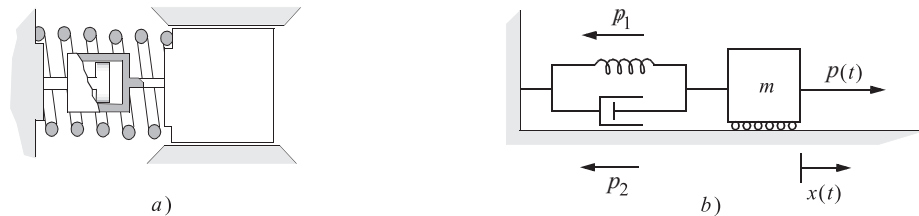
Kraftbalanse uttrykker det velkjente resultat fra klassisk mekanikk at i et mekanisk system bestående av masselementer, fjærende elementer og dempende elementer, vil summen av kraft og motkraft, i følge Newtons 3. lov, alltid være lik null. Dette resultatet gjelder både for translasjonsbevegelse og rotasjonsbevegelse. Ved rotasjonsbevegelse erstattes begrepet *kraft* med *moment*. To eksempler vil illustrere kraftbalanse.

¹Det er likegyldig om vi bruker Kelvin eller Celsius i denne sammenheng fordi de to skalaene har samme avstand mellom hver grad – de er bare forskjøvet i forhold til hverandre, [K] = [C]+273.

EKSEMPEL 2.5: Kraftbalansen til en svingende masse

Vi skal finne posisjonen x til en masse m . En kraft p påvirker massen som er fastspent til en vegg via en fjær og en demper. Massen glir friksjonsfritt på underlaget.

Figur 2.15
a) Svingende masse
b) Prinsippsskisse



Vi setter opp kraftbalansen til massen ved bruk av Newtons 2. lov

$$ma = \sum_{i=1}^n p_i \quad (2.46)$$

Fjærkraften, p_1 , er

$$p_1 = kx$$

der k er fjærkonstanten og x er forlengelsen i fjæren, som også er lik forskyvningen til massen. Dempekraften, p_2 , er

$$p_2 = f \frac{dx}{dt}$$

der f er dempekonstanten. Vi benytter likning (2.46)

$$\begin{aligned} m \frac{d^2x}{dt^2} &= p - p_2 - p_1 \\ &= p - f \frac{dx}{dt} - kx \end{aligned}$$

Dette er en 2. ordens differensiallikning

$$\ddot{x} + \frac{f}{m} \dot{x} + \frac{k}{m} x = \frac{1}{m} p \quad (2.47)$$

idet vi har innført notasjonen

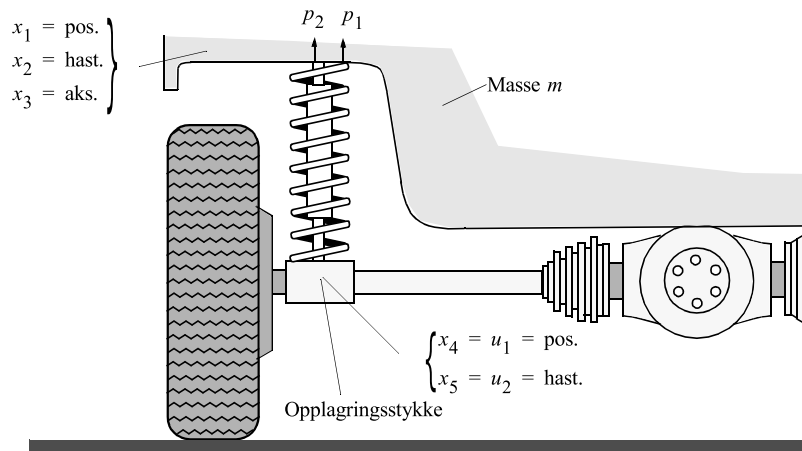
$$\frac{dx}{dt} = \dot{x}, \quad \frac{d^2x}{dt^2} = \ddot{x}, \quad \dots \quad \frac{d^n x}{dt^n} = x^{(n)}$$

■

EKSEMPEL 2.6: Kraftbalanse for fjærsystemet i en bil

Dette eksemplet er bare en litt mer virkelighetsnær utgave av forrige eksempel.

Figur 2.16
Fjærsystemet i en
bil



Figur 2.16 viser fjærsystemet i en bil. For enkelhets skyld ignorerer vi de tre andre hjulene. En fjerdedel av bilens masse er forbundet med et opplagringsstykke gjennom en fjær og en støtdemper. Opplagringsstykket tenkes å bevege seg opp og ned tilsvarende veidekket. Problemet er å bestemme hvordan massen vil bevege seg. Når alt er i ro, vil vekten av massen utøve en kraft på fjæren som vil sammentrykkes, inntil det oppstår en likevekt der fjærkraften er lik vekten av massen. Posisjonen til massen (x_1) defineres lik 0 når fjæren er spenningsfri. Likeledes regnes posisjonen av opplagringspunktet (underlaget) (x_4) relativt til det stasjonære underlaget nevnt ovenfor. For enkelhets skyld antar vi at hjulet er laget av svært hard gummi (dvs. vi ignorerer fjæring og dempning i dekket).

Vi kan sette opp tre kraftbalanser for systemet:

- En kraftbalanse for kreftene som virker på massen
- En kraftbalanse for kreftene som virker på fjæren
- En kraftbalanse for kreftene som virker på demperen

På massen virker fire krefter:

- En gravitasjonskraft proporsjonal med massen (m) og tyngdens akselerasjon (g)
- En akselerasjonskraft proporsjonal med massen (m) og akselerasjonen (x_3)
- En kraft (p_1) fra fjæren
- En kraft (p_2) fra demperen

Summen av disse fire kreftene vil være lik null. Vi får derfor

$$-mg - mx_3 + p_1 + p_2 = 0 \quad (2.48)$$

I (2.48) har vi valgt positiv retning oppover¹ på krefter og akselerasjoner.

På fjæren virker det to krefter:

- En kraft (p_1) fra kontaktflaten med massen
- En kraft som skyldes sammentrykningen av fjæren, og som derfor er proporsjonal med sammentrykningen ($x_1 - x_4$) og fjærkonstanten (k).

Disse to kreftene vil være i balanse, det vil si deres sum er lik null. Vi får derfor

$$-p_1 + k(x_4 - x_1) = 0 \quad (2.49)$$

På demperen virker to krefter:

- En kraft (p_2) fra kontaktflaten med massen
- En kraft som skyldes hastigheten hvormed stempelet beveger seg i dempesylinderen og som derfor er proporsjonal med hastigheten ($x_5 - x_2$) og dempekonstanten (f).

Disse to kreftene vil være i balanse og deres sum settes derfor lik null. Vi får derfor

$$-p_2 + f(x_5 - x_2) = 0 \quad (2.50)$$

Disse tre likningene sammen med den kunnskapen vi har om sammenhengen mellom posisjon, hastighet og akselerasjon,

$$\frac{dx_1}{dt} = x_2 \quad (2.51)$$

$$\frac{dx_2}{dt} = x_3 \quad (2.52)$$

$$\frac{dx_4}{dt} = x_5 \quad (2.53)$$

gjør det mulig for oss å sette opp en komplett matematisk beskrivelse av systemet i figur 2.16.

¹For å ikke rote seg bort når man setter opp kraftbalansen, er det viktig å bestemme seg for hvilken retning som defineres som positiv og så holde seg konsekvent til den.

Innsetting leder til at variabelen x_3 kan elimineres og vi finner følgende likninger

$$\frac{dx_1}{dt} = x_2 \quad (2.54)$$

$$\frac{dx_2}{dt} = \frac{1}{m}(-kx_1 - fx_2) + \frac{1}{m} \left(ku_1 + f \frac{du_1}{dt} \right) - g \quad (2.55)$$

Vi ser at dette er to 1. ordens differensiallikninger og de er koplet sammen idet begge de avhengige variablene x_1 og x_2 forekommer i begge likningene. Den ytre påvirkningen på systemet ser vi utgjøres av posisjonen $u_1 = x_4$ av underlaget og dens tidsderiverte (hastigheten).

$$u_2 = \frac{du_1}{dt} = x_5 \quad (2.56)$$

I tillegg kommer den konstante tyngdens akselerasjon g .

Vi skal siden komme tilbake til et detaljert studium av differensiallikninger av den typen som er gitt i (2.54) og (2.55).

De tre kraftbalansene som er benyttet i det foregående, kan selvsagt uttrykkes på mange forskjellige måter. Det vesentlige er midlertid at man bryter ned problemet til elementære fenomener som siden knyttes sammen med relasjoner av den typen som er gitt i (2.48)-(2.53).



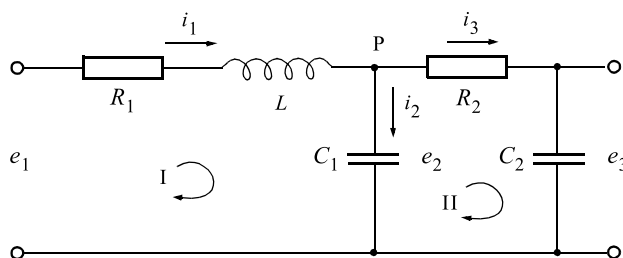
2.5.5 Elektrisk balanse

Anvendelse av balansebetraktninger av samme type som i de foregående avsnittene er velkjente i analyse av elektriske kretser. Kirchhoffs lover uttrykker som kjent at summen av strømmene inn til et knutepunkt er lik null og at summen av enkeltspenningene over alle komponentene i en maske (sløyfe) er lik null.

EKSEMPEL 2.7: Elektrisk balanse for en krets

Vi tar for oss kretsen som er vist i figur 2.17.

Figur 2.17
Elektrisk krets



Kretsen påtrykkes en varierende spenning $e_1(t)$. Det forutsettes at spenningskilden $e_1(t)$ har null indre motstand. Vi ønsker å finne strømmene og spenningene i kretsen, særlig spenningen e_3 som blir kretsens utgangsvariabel. Det forutsettes at denne måles med åpne klemmer, dvs. uendelig belastningsimpedans.

De generelle sammenhengene mellom spenning v og strøm i for en kondensator, henholdsvis induktans, er

$$i = C \frac{dv}{dt}, \quad \text{og} \quad v = L \frac{di}{dt}$$

I en krets der en induktans inngår, lønner det seg derfor å velge strømmen gjennom induktansen som avhengig variabel – i vårt eksempel, i_1 . Når en kapasitet forekommer i en krets, lønner det seg å bruke spenningen over kapasiteten som avhengig variabel – i dette tilfellet e_2 og e_3 . Spenningsbalansen rundt masken I blir da

$$L \frac{di_1}{dt} = e_1 - e_2 - R_1 i_1 \quad (2.57)$$

Spenningsbalanse rundt masken II gir

$$R_2 i_3 = R_2 C_2 \frac{de_3}{dt} = e_2 - e_3 \quad (2.58)$$

Strømbalanse i forgreningspunktet P gir

$$\begin{aligned} i_1 &= i_2 + i_3 = C_1 \frac{de_2}{dt} + C_2 \frac{de_3}{dt} \\ &= C_1 \frac{de_2}{dt} + \frac{e_2 - e_3}{R_2} \end{aligned} \quad (2.59)$$

Vi innfører en notasjon tilsvarende den som er brukt i de foregående eksemplene,

$$i_1 = x_1$$

$$e_2 = x_2 \quad \text{og} \quad e_1 = u_1$$

$$e_3 = x_3$$

Da vil de foregående likningene bli:

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + b_{11}u_1 \quad (2.60)$$

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \quad (2.61)$$

$$\frac{dx_3}{dt} = a_{32}x_2 + a_{33}x_3 \quad (2.62)$$

der

$$\begin{aligned} a_{11} &= -\frac{R_1}{L}, & a_{12} &= -\frac{1}{L}, & b_{11} &= \frac{1}{L} \\ a_{21} &= \frac{1}{C_1}, & a_{22} &= -\frac{1}{R_2C_1}, & a_{23} &= \frac{1}{R_2C_1} \\ a_{32} &= \frac{1}{R_2C_2}, & a_{33} &= -\frac{1}{R_2C_2} \end{aligned}$$

Likning (2.60)-(2.62) viser at vi igjen har fått et sett med 1. ordens koplede tidsinvariante (dvs. konstante koeffisienter) lineære differensiallikninger¹.

■

¹ Dette er ikke den mest effektive måten å sette opp matematiske modeller på for slike enkle kretser, men det er en fordel å være i stand til å gjøre dette hvis man for eksempel skulle møte på en krets med ulineære elementer (dioder etc.). Da er det nødvendig å bruke en slik teknikk. Men hvis alle elementer er lineære, er det som regel enklest å bruke **impedans**begrepet, behandlet i eksempel 4.4.

2.5.6 Prosess med både elektrisk- og mekanisk balanse

Noen prosesser inneholder elementer fra *flere* av de fysiske delområder som vi har sett på til nå. En kategori er elektriske maskiner, som både har elektriske og mekaniske komponenter. Vi skal se nærmere på dette.

Dette eksemplet er særskilt sentralt – og vil gå igjen mange steder seinere i boka.

EKSEMPEL 2.8: Likestrømsmotor

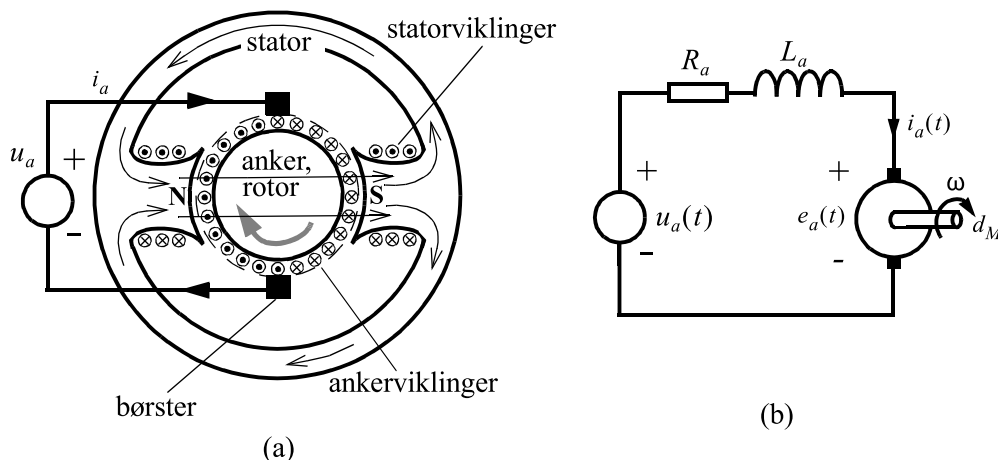
Vi skal presentere og utvikle en matematisk modell av **likestrømsmotoren**. Den brukes i mange sammenhenger, der det er behov for hastighetsregulering (elektriske kjøretøyer, trikker, tog, papirmaskiner, valseverk) eller posisjonsregulering (roboter, skrivere). Motorene kan ha effekter fra MW-området og ned til brøkdeler av en Watt. De brukes også som generatorer, da settes det et dreiemoment på akselen og man tar ut elektrisk effekt fra motorens *anker*krets (eller *rotor*krets – se nedenfor). Vi snakker ofte om likestrømsmaskin for å indikere at den, som de fleste andre typer¹ elektriske maskiner, kan brukes både som motor og generator. En spesiell generatoranvendelse av en liten likestrømsmaskin er som **tachometer** (turtallsmåler); da kopler man maskinen til en aksel som driver den og lar den virke som måleinstrument. Den genererte spenningen blir proporsjonal med turtallet.

Vi skal forklare likestrømsmotorens virkemåte ved å referere til de to skissene i figur 2.18. Delfigur (a) forklarer hvorfor motoren roterer (men du trenger ikke skjønne detaljene i delfigur (a) for å kunne modellere likestrømsmotoren – du kan nøye deg med delfigur (b)). Motoren består av to hovedkomponenter, den ytre og stillestående *stator*en, og den indre og roterende *rotor*en – også kalt *ankeret*. Begge er laget av magnetisk ledende blikk. Statorviklingene er påsatt en spenning som kalles *feltspenningen*. Den gir opphav til *feltstrømmen*, som lager et magnetfelt i statorspolene. (Disse statorspolene kan også være permanentmagneter; da trenger man ikke feltviklinger for å magnetisere motoren.) Magnetfeltet fra statorspolene går gjennom rotoren som indikert ved de tynne pilene i delfigur (a). Vinkelrett på dette magnetfeltet og parallelt med rotorakselen, går viklingene i rotoren. Disse mates med likestrøm, *ankerstrømmen* i_a . Dette skjer via kullstifter som kalles *børster*, som forholdsvis friksjonsfritt kan overføre strøm til rotorviklingene, som glir forbi børstene. Magnetfeltet fra statoren

¹Likestrømsmotoren er bare én i en flora av elektriske motortyper. Den mest utbredte er **asynkronmotoren**, som fullstendig dominerer i industrien. Denne motoren er billig og slitesterk (mest fordi den ikke har børster), og den går direkte på trefase vekselspenning fra nettet, uten omforming. Ulempen er at den i motsetning til likestrømsmotoren er vanskeligere å hastighets- eller posisjonsregulere. Den anvendes derfor der hvor man bare trenger konstant hastighet, som i ventilasjonsanlegg, pumper, transportbånd og de aller fleste kraner. Men man kan ved hjelp av moderne kraftelektronikk bruke en asynkronmotor til avanserte servo-anvendelser. Etterhvert som denne elektronikken utvikler seg i ytelse/pris, vinner asynkronmotoren fram på bekostning av likestrømsmotoren. Vi vil likevel nøye oss med å behandle likestrømsmotoren i denne boka, fordi den fortsatt dominerer i regulerings-tekniske anvendelser, er enkel å forstå, og fordi innsikt i én type motor gjør det lett å forstå de andre typene, da det er store likhetstrekk mellom de ulike motortypene.

utøver ei kraft på hver vikling som er proporsjonal med magnetfeltets styrke, med viklingens lengde på tvers av feltet, og med ankerstrømmen i_a . Med de strømrøtninger i stator som er antydnet (“kryss” betyr at strømmen går inn i papirets plan, “prikk” at den kommer ut av planet), vil de samlede krefter på viklingene gi et dreiemoment d_M [Nm] i den retning som er antydnet med den krumme pila i rotoren. Den detaljerte viklingsgeometri framgår ikke av figuren, men poenget er at viklingene legges slik at strømmen i rotor alltid går en vei foran den ene polen, og motsatt vei foran den andre.

Figur 2.18
Likestrømsmotor



For våre anvendelser vil vi anta at statorfeltet er konstant, enten fordi man har konstant feltstrøm, eller fordi statorfeltet settes opp av permanentmagneter. Da er motorens dreiemoment gitt av i_a . Vi har en *ankerstyrt motor*, noe som er svært vanlig i reguleringsteknikken. (Men man kan også ha feltstyrte motorer, eller motorer hvor begge viklinger reguleres samtidig). For en ankerstyrt motor vil man ha følgende enkle formel for dreiemomentet:

$$d_M = K_T i_a \quad (2.63)$$

K_T kalles *motormomentkonstanten*.

Nå til delfigur (b): Dette er en blanding av et elektrisk krets-skjema og en mekanisk skisse. Alt i denne skissen har med rotoren å gjøre, siden vi bare betrakter alternativet med ankerstyring og konstant statorfelt. Virkningen av statorens konstante magnetfelt er inkorporert i motormomentkonstanten, og vi kan derfor se bort fra statoren i det følgende. Ankerspenningen u_a gir opphav til ankerstrømmen i_a . Rotorens elektriske egenskaper modelleres som en seriekopling av *ankerinduktans* L_a og *ankermotstand* R_a . Når motoren roterer, oppstår en *motindusert spenning* e_a over ankerviklingene, proporsjonal med vinkelhastigheten ω . Sammenhengen er

$$e_a = K_v \omega \quad (2.64)$$

K_v kalles *hastighetskonstanten* (eller *spenningskonstanten*). Det kan vises (se bakerst i dette eksemplet) at vi under rimelige forutsetninger har $K_v = K_T$. Den motinduserte spenninga

e_a virker begrensende på ankerstrømmen, slik det framgår av fortegnene i delfigur (b). For spenningsbalansen i ankerkretsen har vi

$$u_a = R_a i_a + L_a \frac{di_a}{dt} + e_a \quad (2.65)$$

Likestrømsmotoren er et *elektromekanisk* system. Den mekaniske delen er også vist i delfigur (b). For den gjelder Newtons 2. lov for rotasjon (som svarer til $F = ma$ for rettlinjet bevegelse). Generelt kan denne uttrykkes som

$$d = J \dot{\omega} \quad (2.66)$$

Her er d netto dreiemoment [Nm], J er treghetsmoment [kgm^2] for den ekvivalente roterende masse som motoren driver $\dot{\omega}$ og er vinkelakselerasjon [rad/s^2]. Med “ekvivalent” menes at alle de masser som motoren beveger, inklusive rotor og aksel, er omregnet til ett ekvivalent treghetsmoment. **Momentbalansen** (som er analog til kraftbalanse for rettlinjet bevegelse) blir da

$$d_M + v = J \dot{\omega} \quad (2.67)$$

Her er v belastningsmomentet på motoren, som skyldes det arbeid motoren gjør. Belastningsmomentet v vil være negativt når motoren yter arbeid. (Når den går som generator, vil v være positivt. Videre: Ved motordrift er $u_a > e_a$, ved generatordrift er $e_a > u_a$.)

Vi har til nå forutsatt at motoren er friksjonsløs. Dette vil for de fleste formål være akseptabelt. Hvis vi likevel skal ta hensyn til friksjon, kan det meste av friksjonsvirkninger modelleres ved å innføre en viskøs dempekonstant B [Nm s /rad] som virker via ω og gir et bremsende tilleggsmoment. Da blir momentbalansen endret til

$$d_M + v - B\omega = J \dot{\omega} \quad (2.68)$$

Vi innfører igjen en notasjon tilsvarende den som er brukt i de foregående eksemplene. Tilstandene blir rotasjonshastighet og ankerstrøm:

$$\omega = x_1$$

$$i_a = x_2$$

Pådraget er ankerspenningen $u_a = u$. Forstyrrelsen er lastmomentet, som allerede er gitt standardbetegnelse v .

Da vil de foregående likningene gi:

$$\frac{dx_1}{dt} = -\frac{B}{J}x_1 + \frac{K_T}{J}x_2 + \frac{1}{J}v \quad (2.69)$$

$$\frac{dx_2}{dt} = -\frac{K_v}{L_a}x_1 - \frac{R_a}{L_a}x_2 + \frac{1}{L_a}u \quad (2.70)$$

Vi har nok en gang fått et sett med 1. ordens koplede, tidsinvariante lineære differensiallikninger.

For spesielt interesserte:

Vi skal vi vise at $K_v = K_T$. Vi forutsetter at den elektriske effekt som motoren avleverer på akselen, svarer til den elektriske effekt som motoren tilføres, fratrukket tap i rotoren på grunn av dens ohmske motstand R . Andre tap kan vi se bort fra. Avlevert mekanisk effekt P på akselen er

$$P = d_M \omega = (K_T i_a) \omega \quad (2.71)$$

Denne skal svare til netto elektrisk effekt tilført motoren, slik at vi også har

$$P = e_a i_a = (K_v \omega) i_a \quad (2.72)$$

Dermed får vi $K_v = K_T$.



2.5.7 Bruk av benevninger

Benevninger er nyttige for å sjekke om de formler vi kommer fram til er konsistente (dvs. korrekte). La oss illustrere dette ved å utlede ligning (2.2). Vi setter opp energibalansen for et lite tidsintervall Δt :

<p>Akkumulert energi i vannet</p> $\underbrace{m_2}_{[\text{kg}]} \underbrace{c_2}_{\left[\frac{\text{kWs}}{\text{kg}\cdot^\circ\text{C}}\right]} \underbrace{\Delta x_2}_{[^\circ\text{C}]}$ <p>= [kWs] = [kJ] = [energi]</p>	<p>Tilført fra varmeelement</p> $\underbrace{g_1}_{\left[\frac{\text{kW}}{^\circ\text{C}}\right]} \underbrace{(x_1 - x_2)}_{[^\circ\text{C}]}$	<p>Bortført av vannstrøm</p> $\underbrace{q}_{\left[\frac{\text{m}^3}{\text{s}}\right]} \underbrace{\rho}_{\left[\frac{\text{kg}}{\text{m}^3}\right]} \underbrace{c_2}_{\left[\frac{\text{kWs}}{\text{kg}\cdot^\circ\text{C}}\right]} \underbrace{(x_2 - v_1)}_{[^\circ\text{C}]}$	<p>Varmetap til omgivelser fra yttervegg</p> $\underbrace{g_2}_{\left[\frac{\text{kW}}{^\circ\text{C}}\right]} \underbrace{(x_2 - v_2)}_{[^\circ\text{C}]}$
$= \left(\underbrace{g_1 (x_1 - x_2)}_{\left[\frac{\text{kW}}{^\circ\text{C}}\right] [^\circ\text{C}]} - \underbrace{q \rho c_2 (x_2 - v_1)}_{\left[\frac{\text{m}^3}{\text{s}}\right] \left[\frac{\text{kg}}{\text{m}^3}\right] \left[\frac{\text{kWs}}{\text{kg}\cdot^\circ\text{C}}\right] [^\circ\text{C}]} - \underbrace{g_2 (x_2 - v_2)}_{\left[\frac{\text{kW}}{^\circ\text{C}}\right] [^\circ\text{C}]} \right) \cdot \underbrace{\Delta t}_{[\text{s}]}$ <p>= [kWs] = [kJ] = [energi]</p>			

Hvis vi deler på Δt og $m_2 c_2$ på begge sider av likhetstegnet, får vi (2.2). Benevninga på begge sider blir

$$\frac{[\text{kWs}]}{[\text{kg}] \cdot \left[\frac{\text{kWs}}{\text{kg} \cdot ^\circ\text{C}} \right] \cdot [\text{s}]} = \left[\frac{^\circ\text{C}}{\text{s}} \right] = \left[\frac{\text{temperatur}}{\text{tidsenhet}} \right]$$

som stemmer med (2.2):

$$\frac{dx_2}{dt} = \dots$$

2.6 ÉN DIFFERENSIALLIKNING AV HØYERE ORDEN

Vi har til nå, med unntak av eksempel (2.5) utelukkende betraktet én differensiallikning, eller et sett av 1. ordens differensiallikninger. Men en differensiallikning kan også inneholde høyere ordens deriverte av den avhengige variabelen med hensyn på tiden t . Da kalles den en høyere ordens differensiallikning. En slik likning av n 'te orden, med én avhengig variabel x og en ytre påvirkning u , kan på generell form skrives

$$g(x, \dot{x}, \ddot{x}, \dots, x^{(n)}, u, t) = 0 \quad (2.73)$$

der $g(\cdot)$ også kan være en eksplisitt funksjon av t som vist (det vil si at $g(\cdot)$ kan ha koeffisienter som varierer med tiden).

I (2.73) har vi brukt følgende symboler

$$\dot{x} \equiv \frac{dx}{dt}, \quad \dots \quad x^{(n)} \equiv \frac{d^n x}{dt^n}$$

Likning (2.73) kan omformes til

$$x^{(n)} = f(x, \dot{x}, \ddot{x}, \dots, x^{(n-1)}, u, t) \quad (2.74)$$

Vi skal sjelden løse slike høyere ordens differensiallikninger, fordi det ofte er mer hensiktsmessig å gjøre dem om til et sett av 1. ordens differensiallikninger¹. *Og vi kan alltid gjøre dette.*

¹Men ved bruk av **Laplacetransformasjon** kan vi løse høyere ordens tidsinvariante lineære differensiallikninger direkte. Dette vil bli gjort i kapittel 4.

Metoden går ut på å innføre nye variable, én for hver av de høyere ordens deriverte av x . Vi innfører i (2.74)

$$\begin{aligned}x_1 &= x \\x_2 &= \dot{x} = \dot{x}_1 \\x_3 &= \ddot{x} = \dot{x}_2 \\&\cdot \\&\cdot \\x_n &= x^{(n-1)} = \dot{x}_{n-1}\end{aligned}$$

Dette gir følgende sett av 1. ordens differensiallikninger

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\cdot \\ &\cdot \\ \dot{x}_n &= f(x_1, x_2, \dots, x_n, u, t)\end{aligned}\tag{2.75}$$

Vi har gjort om én n 'te ordens differensiallikning til n førsteordens koblede differensiallikninger. Variablene x_1, x_2, \dots, x_n i (2.75) kalles **tilstandsvariable**. For et hvert system karakterisert med én n 'te ordens differensiallikning kan man altså lage et ekvivalent system med n tilstandsvariable. Begge systemer er av n 'te orden.

EKSEMPEL 2.9: Omgjøring av en 2. ordens differensiallikning

Vi betrakter 2. ordens differensiallikningen som beskriver bevegelsen av en masse som er festet til en fjær og en demper – eksempel 2.5

$$\ddot{x} + \frac{f}{m}\dot{x} + \frac{k}{m}x = \frac{1}{m}p \quad (2.76)$$

Vi skal gjøre om likning (2.76) til et sett av 1. ordens likninger. Vi setter $x = x_1$ og $\dot{x} = x_2$. Den påtrykte kraften, p , blir vårt pådrag, u . Likning (2.76) kan da skrives slik

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k}{m}x_1 - \frac{f}{m}x_2 + \frac{1}{m}u \end{aligned} \quad (2.77)$$



Tilstandsvariable som danner den spesielle strukturen i (2.75) og som (2.77) er et eksempel på, kalles gjerne **fasevariable**. De består av en variabel og de første $n - 1$ deriverte av denne. Et eksempel på en modell hvor de tilstandsvariable *ikke* er “fasevariable” er den elektriske kretsen gitt ved likningene (2.60)-(2.62). Tilstandsvariablene i de fleste eksemplene i avsnitt 2.5 passet ikke inn i den fasevariable strukturen (2.75).

Likning (2.78) gjengir en generell struktur av n koplede 1. ordens differensiallikninger

$$\left. \begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \dots, x_n, u, t) \\ \dot{x}_2 &= f_2(x_1, x_2, \dots, x_n, u, t) \\ &\cdot \\ &\cdot \\ \dot{x}_n &= f_n(x_1, x_2, \dots, x_n, u, t) \end{aligned} \right\}, \quad \text{eller kompakt uttrykt: } \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u, t) \quad (2.78)$$

Likning (2.78) representerer også ulineære og tidsvariante systemer.

Et spesialtilfelle som dominerer i denne boka er når funksjonene av tilstandsvariablene og pådraget på høyresiden i likning (2.78) er lineære og tidsinvariante. Alle eksempler til nå unntatt 2.4 (massebalanse for væskestrøm, som er ulineært), er i denne kategorien.

Et slikt sett av første-ordens koplede, lineære differensiallikninger med konstante koeffisien-

ter og én pådragsvariabel kan skrives

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_{11}u_1 \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_{21}u_1 \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + b_{n1}u_1 \end{aligned} \quad (2.79)$$

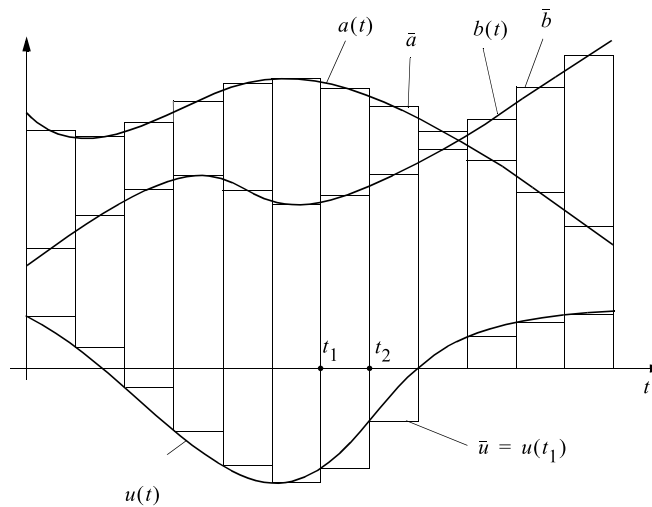
Likningssettet i (2.79) kan løses ved hjelp av de metodene som utvikles i kapittel 3.

2.7 TIDSVARIANTE DIFFERENSIALLIKNINGER

Likning (2.17) gir løsningen av en lineær differensiallikning med *konstante* koeffisienter. Den er for enkelhets skyld gjengitt i (2.80)

$$x(t) = e^{a(t-t_0)}x_0 + \int_{t_0}^t e^{a(t-\tau)}bu(\tau)d\tau \quad (2.80)$$

Figur 2.19
Diskretisering av
pådrag og tidsvari-
erende konstanter



Når koeffisientene er *tidsvarierende*, som i likning (2.6), finnes det en løsning som likner på (2.80). For å løse lineære differensiallikninger med tidsvarierende koeffisienter, tar vi i bruk numeriske teknikker. Vi deler opp tidsaksen i intervaller. Lengden på intervallene velges slik at koeffisientene $a(t)$ og $b(t)$ og pådraget $u(t)$ kan betraktes som tilnærmet konstante i intervallene. Dette er illustrert i figur 2.19.

Vi kan finne ut hvordan systemet utvikler seg fra en begynnelsestilstand ved tidspunktet t_1 til tilstanden ved tidspunktet t_2 , der $t_2 - t_1 = \Delta t$ (**diskretiseringsintervallet**), ved å benytte (2.80)

$$x(t_2) = e^{\bar{a}\Delta t}x(t_1) + \int_{t_1}^{t_2} e^{\bar{a}(t_2-\tau)}\bar{b}\bar{u}d\tau \quad (2.81)$$

\bar{a} , \bar{b} og \bar{u} er middelverdiene av de respektive tidsfunksjonene i intervallet fra t_1 til t_2 .

Løsningen av (2.81) er

$$\begin{aligned} x(t_2) &= e^{\bar{a}\Delta t}x(t_1) + \frac{\bar{b}}{\bar{a}}(e^{\bar{a}\Delta t} - 1)\bar{u} \\ &= \varphi(t, \Delta t)x(t_1) + \delta(t, \Delta t)u(t_1) \end{aligned} \quad (2.82)$$

Vi har for enkelhets skyld antatt at den midlere verdien av pådraget $u(t)$ over tidsintervallet Δt er lik verdien pådraget hadde ved tidspunktet t_1 , altså $\bar{u} = u(t_1)$.

Dersom koeffisientene a og b er konstante, slik tilfellet var i avsnitt 2.2.1, vil likning (2.82) forenkles til

$$x(t_2) = \varphi(\Delta t)x(t_1) + \delta(\Delta t)u(t_1) \quad (2.83)$$

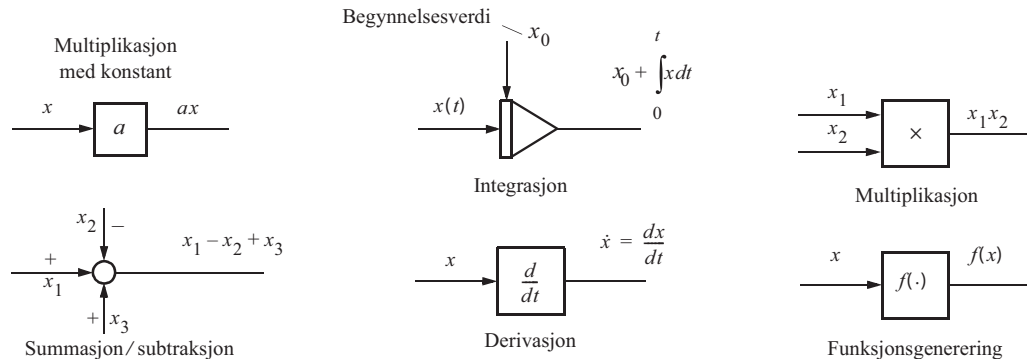
der $\varphi(\Delta t)$ og $\delta(\Delta t)$ er konstante. Det knytter seg spesiell interesse til funksjonen $\varphi(t, \Delta t)$ i (2.82). Den kalles **transisjonsfunksjonen**, fordi den beskriver hvordan systemets tilstand gjennomgår en *transisjon*¹ fra $x(t_1)$ ved tidspunktet t_1 til $x(t_2)$ ved tidspunktet t_2 , når systemet ikke er påvirket utenfra. Transisjonsfunksjonen er en funksjon av tiden og tidsintervallet når systemet er tidsvariant. Når systemet er tidsinvariant, er den bare avhengig av tidsintervallets lengde.

¹Egentlig halv-engelsk: “transition” = overgang.

2.8 BLOKKDIAGRAMMER

Et blokkdiagram gir en skjematisk representasjon av en matematisk modell.

Figur 2.20
Vanlige symboler



I et blokkdiagram vil tilstander og andre tidsvariable være representert ved linjer (omtrent som et elektrisk signal på en ledning) mens matematiske operasjoner som funksjonsberegning, integrasjon, derivasjon og liknende er karakterisert ved firkantede blokker eller spesielle symboler som vist i figur 2.20. Addisjon og subtraksjon av en eller flere tidsvariable er representert ved et symbol der fortegnet angir om det er addisjon eller subtraksjon. De mest vanlige symbolene er vist i figur 2.20.

Fordelen med blokkdiagrammer - som inneholder akkurat samme informasjon som et sett av differensiallikninger - er at det er lettere å se av et slikt diagram hvordan vekselvirkningene i systemet skjer; dvs. tilbakekoplinger og koplinger mellom deler av systemet.

Vi vil illustrere hvordan man lager blokkdiagrammer med noen eksempler. Fra nå av vil vi stort sett sløyfe plusstegnene i summasjonspunktene. Det er underforstått at vi mener pluss på innganger der det ikke står noe annet.

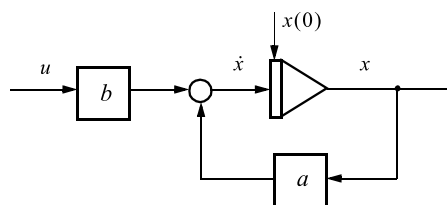
EKSEMPEL 2.10: Blokkdiagram for et 1. ordens system

Vi tar utgangspunkt i følgende 1. ordens differensiallikning

$$\dot{x} = ax + bu \quad (2.84)$$

Likning (2.84) kan illustreres ved å sette sammen to blokker for forsterkning, et summasjonspunkt og en integrator. Resultatet kan sees i blokkdiagrammet i figur 2.21.

Figur 2.21
Blokkdiagram for
en 1. ordens
differensiallikning



En vesentlig egenskap ved blokkdiagrammet i figur 2.21 er at løsningen av differensiallikningen i (2.84) foregår ved integrasjon, dvs. vi bestemmer en variabel ved å integrere den deriverte. I prinsippet kunne vi gjort det motsatte, ved å bestemme den deriverte ved derivasjon av variabelen. Det siste er imidlertid både numerisk og fysisk uheldig fordi små hurtige forstyrrelser ville forårsaket store beregningsfeil.

■

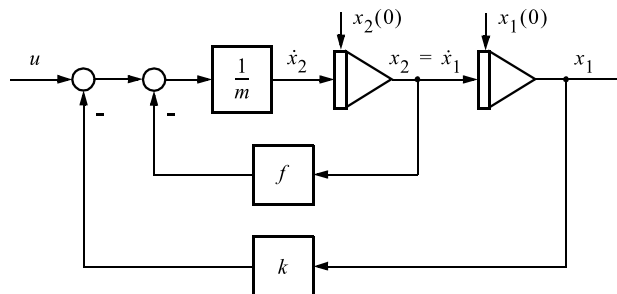
EKSEMPEL 2.11: Blokkdiagram for et 2. ordens system

Vi tar utgangspunkt i det problemet som er beskrevet i eksempel 2.5 og 2.9, og skal illustrere innholdet i likning (2.77),

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k}{m}x_1 - \frac{f}{m}x_2 + \frac{1}{m}u \end{aligned} \quad (2.85)$$

Vi benytter to integratorer, to summeringspunkter og flere blokker for multiplikasjon med konstanter for å illustrere innholdet av (2.85). Resultatet er vist i figur 2.22.

Figur 2.22
Elementært blokk-
diagram for
svingende-masse-
systemet.



Blokkdiagrammet i figur 2.22 kalles et **elementært blokkdiagram** fordi den matematiske operasjonen, som inngår i hver enkelt blokk, kan kalles en elementær operasjon. I kapittel 4 skal vi introdusere blokkdiagrammer der noen blokker representerer mer sammensatte operasjoner enn de elementære, ved å bruke Laplacetransformasjon.

■

Vi kan enkelte ganger sette opp et bilde av de fysiske balansene med et elementært blokkdiagram uten å gå veien om en matematisk formulering av differensiallikningene. Forutsetningen for dette er at vi kan gjøre fornuftige valg av matematiske tilstandsvariable. Ofte kan vi med fordel velge tilstander som utganger av et tilsvarende antall integratorer. Inngangen til disse integratorene vil være de tidsderiverte av tilstandsvariablene. Som eksempel på dette velger vi systemet som er gitt ved figur 2.12 i eksempel 2.4.

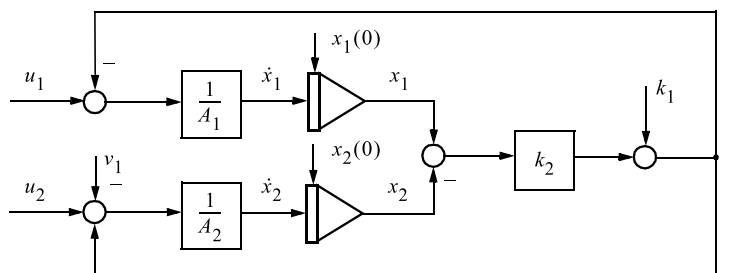
EKSEMPEL 2.12: Elementært blokkdiagram av et system med to tanker

Nivåene x_1 og x_2 varierer som funksjon av tiden. Vi velger disse som tilstandsvariable og som utganger av hver sin integrator. Det er logisk at stoffvolumet (væskehøyden ganget med tankarealet) i en tank er gitt ved utgangen av en integrator, fordi innholdet i tanken er et resultat av volumstrømmen inn til og ut fra tanken fra $t = 0$ og fram til t (og av initialtilstanden).

Vi lar volumstrømmene inn på og ut av de to tankene gå inn på summasjonspunkter slik at utgangen til summasjonspunktene blir netto volumstrøm. Utgangen fra summasjonspunktene korrigeres for arealet i de respektive tankene og går deretter inn på hver sin integrator. Dette er vist i figur 2.23. Vi gjenkjenner alle detaljene som inngår i likningene (2.44) og (2.45).

Dersom vi ønsker å representere rørforbindelsen mellom de to karene med en riktigere matematisk modell enn den som baserer seg på linearisering, kan vi erstatte blokken til høyre i figur 2.23, som inneholder konstanten k , med en funksjonsblokk som inneholder en generell funksjon, $f(\cdot)$. Det er uansett differansen mellom nivåene x_1 og x_2 som er avgjørende for strømmen av vann.

Figur 2.23
Blokkdiagram for
systemet med to
kar



■

Tilstandsromanalyse

3.1 Innledning og motivering 73

3.2 Vektordifferensiallikninger 75

3.3 Løsning av lineære vektordifferensiallikninger 77

3.4 Monovariabel og multivariabel systemer 91

3.5 Valg av tilstandsvariable 92

3.6 Ulineære vektordifferensiallikninger. Linearisering 99

3.1 INNLEDNING OG MOTIVERING

I forrige kapittel viste vi hvordan matematiske modeller for dynamiske systemer kunne utvikles. Disse modellene ble beskrevet ved bruk av differensiallikninger. I dette kapitlet skal vi presentere *tilstandsrommodellen* som er en generell, kompakt og særdeles anvendbar måte å formulere matematiske modeller for dynamiske systemer. Bruken av tilstandsrommodellen har følgende fordeler:

- Det gir en kompakt standard beskrivelse av differensiallikningene som danner modellen av systemet.
- Det kan effektivisere løsningen av differensiallikningene ved at formen blir ekvivalent med det som foreligger for en enkel 1. ordens differensiallikning.
- Mange reguleringsalgoritmer er basert på tilstandsromanalyse.

Vi skal illustrere tilstandsrom-strukturen et innledende eksempel.

EKSEMPEL 3.1: Modell av likestrømsmotor og modell av svingende masse

Modellen av likestrømsmotoren som ble utviklet i eksempel 2.8, ble beskrevet av to 1. ordens lineære differensiallikninger (2.69) og (2.70),

$$\begin{aligned}\dot{x}_1 &= -\frac{B}{J}x_1 + \frac{K_T}{J}x_2 + 0 \cdot u + \frac{1}{J}v \\ \dot{x}_2 &= -\frac{K_v}{L_a}x_1 - \frac{R_a}{L_a}x_2 + \frac{1}{L_a}u + 0 \cdot v\end{aligned}\tag{3.1}$$

mens modellen for den svingende massen i eksempel 2.5 ble beskrevet av én 2. ordens likning

$$\ddot{x} = \frac{f}{m}\dot{x} + \frac{k}{m}x = \frac{1}{m}p\tag{3.2}$$

Men i eksempel 2.9 viste vi hvordan denne siste modellen kan omgjøres til to 1. ordens likninger

$$\begin{aligned}\dot{x}_1 &= 0 \cdot x_1 + x_2 + 0 \cdot u + 0 \cdot v \\ \dot{x}_2 &= -\frac{k}{m}x_1 - \frac{f}{m}x_2 + \frac{1}{m}u + 0 \cdot v\end{aligned}\tag{3.3}$$

Her har vi innført standardsymbolet for pådrag, u . Vi ser at disse to forskjellige prosessene kan gis representasjoner med samme struktur, (3.1) og (3.3). Når en prosessmodell beskrives av et slikt sett med 1.ordens differensiallikninger, kalles den en **tilstandsrommodell**. Den unødig omstendelige skrivemåten med ledd hvor man multipliserer med null, vil få mening når vi i neste avsnitt innfører matriser og vektorer i slike modeller.



3.2 VEKTORDIFFERENSIALLIKNINGER

3.2.1 Tilstandsvektor og tilstandsrom

Som nevnt ovenfor, kan et dynamisk system beskrives ved et sett av koblede 1. ordens differensiallikninger. Generelt har vi

$$\begin{aligned}
 \dot{x}_1 &= f_1(x_1, x_2 \dots x_n, u_1, u_2 \dots u_r, t) \\
 \dot{x}_2 &= f_2(x_1, x_2 \dots x_n, u_1, u_2 \dots u_r, t) \\
 &\cdot \\
 &\cdot \\
 &\cdot \\
 \dot{x}_n &= f_n(x_1, x_2 \dots x_n, u_1, u_2 \dots u_r, t)
 \end{aligned} \tag{3.4}$$

Med vektornotasjon kan vi skrive dette som

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{3.5}$$

Vektoren \mathbf{x} er en tilstandsvektor og dens elementer er **tilstandsvariable**. Vektorrommet som utspennes av tilstandsvektoren kalles **tilstandsrommet** og (3.5) er en **tilstandsrommodell**.

Vektoren \mathbf{u} har dimensjon r . Den inneholder variable som påvirker systemet utenfra¹. Hvis differensiallikningene i (3.4) er lineære med konstante koeffisienter, vil (3.5) kunne skrives på formen

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{3.6}$$

der \mathbf{A} og \mathbf{B} er konstante matriser.

¹For enkelhets skyld ser vi her bort fra andre påvirkninger enn pådraget \mathbf{u} . Man kan også trekke inn prosessforstyrrelse \mathbf{v} og målestøy \mathbf{w} . Men dette skiller seg ikke prinsipielt fra hvordan virkningen av \mathbf{u} representeres, så vi lar det ligge her. Det behandles i avsnitt 3.4.

EKSEMPEL 3.2: System med svingende masse på tilstandsromform

I likning (3.3) i eksempel 3.1 ble systemet med svingende masse beskrevet med to koblede førsteordens differensiallikninger. Vi skal nå uttrykke (3.3) på formen (3.6). Tilstandsvektoren blir

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Siden vi bare har ett pådrag, vil pådragsvektoren degenerere til en skalar. Når pådraget er skalart, blir matrisen \mathbf{B} en vektor, \mathbf{b} . Vi har da for vårt system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \quad (3.7)$$

Innsatt verdiene i likning (3.3), får vi

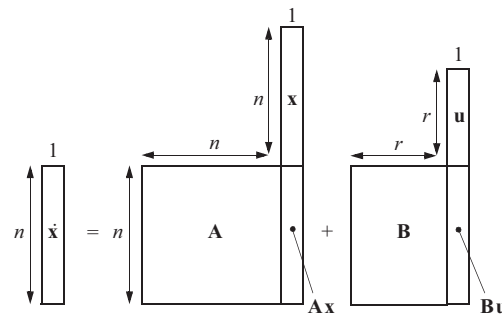
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{f}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \quad (3.8)$$

■

3.2.2 Et “rektangulært skjema” for matrise- og vektormultiplikasjon

For å holde styr på dimensjoner, og for å se klarere hva matrise- og vektormultiplikasjon innebærer, kan likning (3.6) framstilles ved hjelp av det grafiske skjemaet i figur 3.1.

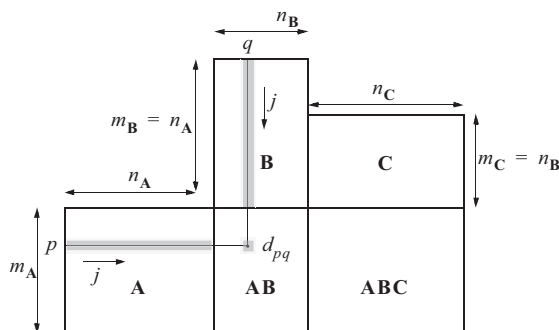
Figur 3.1
“Rektangulært skjema”



Resulterende elementer framkommer ved å multiplisere matrise-rader med vektorer.

Dette grafiske skjemaet er generelt nyttig ved multiplikasjon av matriser. Anta nå at tre vilkårlige matriser \mathbf{A} , \mathbf{B} og \mathbf{C} skal multipliseres med hverandre, se figur 3.2. (Merk at vi må ha overensstemmelser med hensyn på dimensjonen til matrisene: $m_{\mathbf{B}} = n_{\mathbf{A}}$ og $m_{\mathbf{C}} = n_{\mathbf{B}}$.)

Figur 3.2
Matrise-
multiplikasjon



Vi kaller mellomresultatet $\mathbf{AB} = \mathbf{D}$. Elementet d_{pq} i \mathbf{D} , framkommer i skjæringen mellom forlengelsen av en radvektor og en kolonnevektor fra henholdsvis \mathbf{A} og \mathbf{B} . Vi har

$$d_{pq} = \sum_{j=1}^{n_A (=m_B)} a_{pj} b_{jq} \quad (3.9)$$

Forøvrig: merk hvordan vektorene \mathbf{x} og \mathbf{u} i figur 3.1 kan betraktes som $n \times 1$ - og $r \times 1$ -“matriser”.

Mer om matriser i appendiks A.

3.3 LØSNING AV LINEÆRE VEKTORDIFFERENSIALLIKNINGER

3.3.1 Transisjonsmatrisen ved dekopling

Vi ønsker å løse vektordifferensiallikningen

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \quad \text{med begynnelsesverdi } \mathbf{x}(t_0) \quad (3.10)$$

for å finne ut hvordan tilstandene endrer seg som funksjon av tiden. Vi går ut fra at \mathbf{A} har *distinkte (forskjellige) egenverdier* og derfor lineært uavhengige egenvektorer (se appendiks A om matrisealgebra). Vi kan innføre en ny tilstandsvektor \mathbf{q} ved å foreta en lineær transformasjon av den opprinnelige tilstandsvektoren

$$\mathbf{x} = \mathbf{Mq} \quad , \quad \mathbf{q} = \mathbf{M}^{-1}\mathbf{x} \quad (3.11)$$

\mathbf{M} er en **egenvektormatrise** som er forklart nærmere i A.1.3. Innført i (3.10) gir dette

$$\mathbf{M}\dot{\mathbf{q}} = \mathbf{AMq} + \mathbf{Bu}$$

Vi formultipliserer med \mathbf{M}^{-1} og bruker at $\mathbf{M}^{-1}\mathbf{M} = \mathbf{I}$

$$\dot{\mathbf{q}} = \mathbf{M}^{-1}\mathbf{AMq} + \mathbf{M}^{-1}\mathbf{Bu} \quad (3.12)$$

Vi bruker likning (A.20) i appendiks A

$$\mathbf{M}^{-1}\mathbf{AM} = \mathbf{\Lambda} \quad (3.13)$$

der $\Lambda =$ **egenverdimatrisen** til A

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Likning (3.12) kan nå skrives slik

$$\dot{\mathbf{q}} = \Lambda \mathbf{q} + \mathbf{M}^{-1} \mathbf{B} \mathbf{u} \quad (3.14)$$

Vi innfører $\mathbf{M}^{-1} \mathbf{B} \mathbf{u} = \mathbf{p}$ i (3.14). Likning (3.14) kan skrives på komponentform, fordi matrisen Λ er diagonal og \mathbf{p} er en vektor

$$\begin{aligned} \dot{q}_1 &= \lambda_1 q_1 + p_1 \\ \dot{q}_2 &= \lambda_2 q_2 + p_2 \\ &\cdot \\ &\cdot \\ \dot{q}_n &= \lambda_n q_n + p_n \end{aligned} \quad (3.15)$$

Dette gir n uavhengige differensiallikninger. Hver av likningene i (3.15) er av samme form som likning (2.7) og vil ha samme løsning som angitt i (2.17). Vi får derfor for likning nr. i

$$q_i(t) = e^{\lambda_i(t-t_0)} q_i(t_0) + \int_{t_0}^t e^{\lambda_i(t-\tau)} p_i(\tau) d\tau \quad (3.16)$$

Vi innfører notasjonen

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & & \\ \vdots & & \ddots & \\ 0 & & & e^{\lambda_n t} \end{bmatrix} \quad (3.17)$$

Fordi likning (3.16) er dekoplet¹ i sin form, kan vi skrive løsningen av (3.12) på følgende vis

$$\mathbf{q}(t) = e^{\Lambda(t-t_0)}\mathbf{q}(t_0) + \int_{t_0}^t e^{\Lambda(t-\tau)}\mathbf{p}(\tau)d\tau \quad (3.18)$$

Benytter vi oss av transformasjonen i (3.11), finner vi

$$\mathbf{x}(t) = \mathbf{M}\mathbf{q}(t) = \mathbf{M}e^{\Lambda(t-t_0)}\mathbf{M}^{-1}\mathbf{x}(t_0) + \int_{t_0}^t \mathbf{M}e^{\Lambda(t-\tau)}\mathbf{M}^{-1}\mathbf{B}u(\tau)d\tau \quad (3.19)$$

Vi innfører på samme måte som i likning (3.13)

$$\mathbf{M}e^{\Lambda t}\mathbf{M}^{-1} = e^{\mathbf{A}t} \quad (3.20)$$

$$\mathbf{M}^{-1}e^{\Lambda t}\mathbf{M} = e^{\mathbf{A}t} \quad (3.21)$$

Likning (3.19) gir

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{B}u(\tau)d\tau \quad (3.22)$$

Matrisen $e^{\mathbf{A}t}$ kalles **transisjonsmatrisen** og spiller en sentral rolle i tilstandsromanalysen. Den betegnes også med symbolet $\Phi(t)$.

$\Phi(t)$ ($= e^{\mathbf{A}t}$) kan finnes av uttrykket i (3.21). Dens elementer er tidsvarierende. Det er viktig å være klar over at

$$e^{\mathbf{A}t} \neq \{e^{a_{ij}t}\}$$

Løsningen i (3.22) er en generalisert versjon av løsningen for det skalare tilfellet i (2.17)².

¹En tilstand \mathbf{q} i et dekoplet (“diagonalt”) system kalles med en spesialbetegnelse for en **modus**, flertall: modi (engelsk entall: *mode*)

²Merk at (3.22) er en vektoriell versjon av (2.17), slik som vi kunne forvente. Men (3.22) er funnet ved å dekoppe differensiallikningen, mens (2.17) er funnet ved hjelp av “metoden med variasjon av parametre”. Vi kunne utmerket godt ha brukt denne metoden også i det vektorielle tilfellet i stedet for å dekoppe. Vi må da bruke regneregler for derivasjon av matriser og vektorer og derivasjon av produkter av disse. Se appendiks A.

3.3.2 Transisjonsmatrisen ved rekkeutvikling

En **autonom** vektordifferensiallikning er kun en funksjon av tilstandene i systemet. Det lineære autonome systemet

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \quad , \quad \mathbf{x}(t_0) = \mathbf{x}(0) = \mathbf{x}_0 \quad (3.23)$$

har en generell løsning som kan skrives

$$\mathbf{x}(t) = \Phi(t)\mathbf{x}_0 \quad (3.24)$$

Dette kan betraktes som en generalisering av den *skalare* differensiallikning

$$\dot{x} = ax, \quad x(0) = x_0 \quad \text{med løsning} \quad x(t) = e^{at}x_0 \quad (3.25)$$

Vi vet at e^{at} kan uttrykkes som

$$e^{at} = 1 + at + \frac{a^2t^2}{2!} + \frac{a^3t^3}{3!} + \dots \quad (3.26)$$

På denne bakgrunn *antar* vi nå at løsningen $\Phi(t)$ kan uttrykkes som

$$\Phi(t) = e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \dots + \frac{1}{i!}\mathbf{A}^i t^i + \dots \quad (3.27)$$

Før vi prøver om (3.27) oppfyller differensiallikninga med begynnelsesbetingelse, likning (3.23), bemerker vi at (3.27) er et legalt uttrykk i forhold til de regler som gjelder for matriseregning: Alle ledd er $n \times n$ -matriser, og kan summeres.

For det første oppfyller (3.27) begynnelsesbetingelsen i (3.23), siden (3.24) med $t = 0$ blir

$$\mathbf{x}(0) = \Phi(0)\mathbf{x}_0 = \left(\mathbf{I} + \mathbf{A}0 + \frac{1}{2!}\mathbf{A}^2 0^2 + \dots \right) \mathbf{x}_0 = \mathbf{x}_0 \quad (3.28)$$

For det andre oppfyller (3.27) differensiallikninga i (3.23), siden

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \dot{\Phi}(t)\mathbf{x}_0 = \frac{d}{dt} \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \dots + \frac{1}{i!}\mathbf{A}^i t^i + \dots \right) \mathbf{x}_0 = \\ & \left(\mathbf{0} + \mathbf{A} + \frac{2}{2!}\mathbf{A}^2t^1 + \frac{3}{3!}\mathbf{A}^3t^2 + \dots + \frac{i}{i!}\mathbf{A}^i t^{i-1} + \dots \right) \mathbf{x}_0 = \\ & \mathbf{A} \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \dots + \frac{1}{(i-1)!}\mathbf{A}^{i-1}t^{i-1} + \dots \right) \mathbf{x}_0 = \mathbf{A}\Phi(t)\mathbf{x}_0 = \mathbf{A}\mathbf{x}(t) \end{aligned} \quad (3.29)$$

Dermed har vi vist at

$$\Phi(t) = e^{At} = \mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2t^2 + \dots + \frac{1}{i!}\mathbf{A}^i t^i + \dots \quad (3.30)$$

er løøsning på (3.23).

Likning (3.30) gir en ny måte å formulere transisjonsmatrisen = matriseeksponentialfunksjonen e^{At} på, og

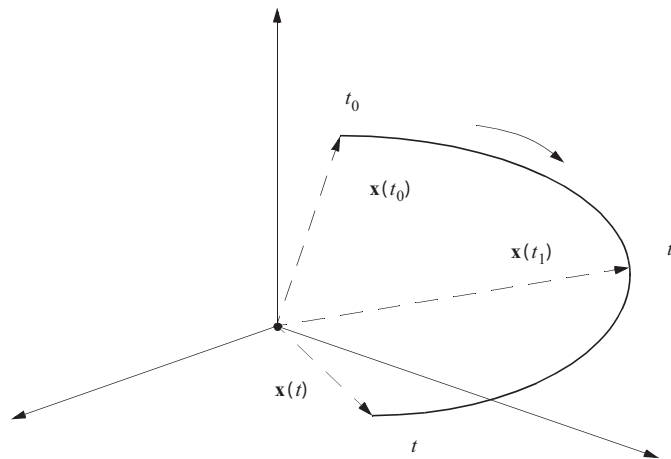
Egenverdiene til \mathbf{A} trenger ikke å være distinkte i dette tilfellet.

Vi innfører $\Phi(t)$ i likning (3.22)

$$\mathbf{x}(t) = \Phi(t-t_0)\mathbf{x}(t_0) + \int_{t_0}^t \Phi(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau \quad (3.31)$$

Hvis $\mathbf{u}(\tau) = 0$ i (3.31), kan vi illustrere løsningen for det autonome systemet som i figur 3.3.

Figur 3.3
Endring av tilstandsvektoren som funksjon av tiden



Systemets tilstandsvektor endrer seg fra tidspunktet t_0 til sin nye verdi ved tidspunktet t . Vektorens endepunkt gjennomløper et spor (trajektor). Dersom systemet hadde startet ved tidspunktet t_1 med en vektor for starttilstander tilsvarende $\mathbf{x}(t_1)$ på det første sporet, ville vi ende opp i samme punkt i tidspunktet t . Dette kan uttrykkes slik

$$\begin{aligned}\mathbf{x}(t) &= \Phi(t - t_0)\mathbf{x}(t_0) = \Phi(t - t_1)\mathbf{x}(t_1) \\ &= \Phi(t - t_1)\Phi(t_1 - t_0)\mathbf{x}(t_0)\end{aligned}\quad (3.32)$$

Likning (3.32) viser at

$$\Phi(t - t_0) = \Phi(t - t_1)\Phi(t_1 - t_0) \quad (3.33)$$

Løsningen til likning (3.31) kan bare finnes når pådragsvektoren $\mathbf{u}(t)$ er kjent i hele det aktuelle tidsrommet. Dersom vi deler opp tidsrommet i mindre intervaller av konstant lengde, slik det ble gjort i avsnitt 2.7, kan en tilnærmet løsning av (3.31) bestemmes. Vi antar at pådragsvektoren er stykkevis konstant i hvert av de små tidsintervallene, og kaller begynnelsen og slutten av et intervall henholdsvis t_1 og t_2 . Vi finner når $t_2 - t_1 = \Delta t$

$$\begin{aligned}\mathbf{x}(t_2) &= \Phi(\Delta t)\mathbf{x}(t_1) + \int_{t_1}^{t_2} \Phi(t_2 - \tau)\mathbf{B}\mathbf{u}(\tau)d\tau \\ &= \Phi(\Delta t)\mathbf{x}(t_1) + \Delta(\Delta t)\mathbf{u}(t_1)\end{aligned}\quad (3.34)$$

Den konstante verdien av pådraget i tidsintervallet t_1 til t_2 er betegnet med $\mathbf{u}(t_1)$. Bestemmelsen av integralet $\Delta(\Delta t)$ i (3.34) kan gjøres ved å diagonalisere som i (3.19). Vi benytter (3.13) og at $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$ og finner

$$\begin{aligned}\Delta(\Delta t)\mathbf{u}(t_1) &= \mathbf{M} \int_{t_1}^{t_2} e^{\Lambda(t_2 - \tau)} d\tau \mathbf{M}^{-1} \mathbf{B} \mathbf{u}(t_1) \\ &= \mathbf{M}(-\Lambda^{-1})(\mathbf{I} - e^{\Lambda(t_2 - t_1)})\mathbf{M}^{-1}\mathbf{B}\mathbf{u}(t_1) \\ &= \mathbf{M}\Lambda^{-1}(e^{\Lambda\Delta t} - \mathbf{I})\mathbf{M}^{-1}\mathbf{B}\mathbf{u}(t_1) \\ &= \mathbf{M}\Lambda^{-1}\mathbf{M}^{-1}\mathbf{M}(e^{\Lambda\Delta t} - \mathbf{I})\mathbf{M}^{-1}\mathbf{B}\mathbf{u}(t_1) \\ &= \mathbf{A}^{-1}(e^{\Lambda\Delta t} - \mathbf{I})\mathbf{B}\mathbf{u}(t_1)\end{aligned}\quad (3.35)$$

Ved innsetting av (3.30) får vi

$$\begin{aligned}\Delta(\Delta t)\mathbf{u}(t_1) &= \mathbf{A}^{-1} \left(\mathbf{I} + \mathbf{A}\Delta t + \frac{1}{2!}\mathbf{A}^2\Delta t^2 + \dots - \mathbf{I} \right) \mathbf{B}\mathbf{u}(t_1) \\ &= \mathbf{A}^{-1} \left(\mathbf{A}\Delta t + \frac{1}{2!}\mathbf{A}^2\Delta t^2 + \dots \right) \mathbf{B}\mathbf{u}(t_1) \\ &= \Delta t \left(\mathbf{I} + \frac{1}{2!}\mathbf{A}\Delta t + \frac{1}{3!}\mathbf{A}^2\Delta t^2 + \dots \right) \mathbf{B}\mathbf{u}(t_1)\end{aligned}\quad (3.36)$$

Sluttresultatet i (3.36) viser at \mathbf{A} gjerne kan være singulær ($\det \mathbf{A} = 0$).

Når pådraget er konstant i intervallet Δt , har vi funnet at transisjonen av et system fra tidspunktet t_1 til t_2 , er gitt av

$$\mathbf{x}(t_2) = \Phi(\Delta t)\mathbf{x}(t_1) + \Delta(\Delta t)\mathbf{u}(t_1) \quad (3.37)$$

der

$$\Phi(\Delta t) = e^{\mathbf{A}\Delta t} \quad \text{og} \quad \Delta(\Delta t) = \mathbf{A}^{-1}(\Phi(\Delta t) - \mathbf{I})\mathbf{B} \quad (3.38)$$

Når tidsintervallet Δt er svært kort, er det tilstrekkelig å bare bruke de første leddene i rekken ved innsetting av rekkeutviklingen i (3.30).

Hvis vi for eksempel nøyer oss med ledd t.o.m. 3. potens, blir

$$\Phi(\Delta t) \approx \mathbf{I} + \mathbf{A}\Delta t + \frac{1}{2!}\mathbf{A}^2\Delta t^2 + \frac{1}{3!}\mathbf{A}^3\Delta t^3 \quad (3.39)$$

I dette tilfellet kan $\Delta(\Delta t)$ skrives slik

$$\Delta(\Delta t) \approx \left(\mathbf{I} + \frac{1}{2!}\mathbf{A}\Delta t + \frac{1}{3!}\mathbf{A}^2\Delta t^2 \right) \Delta t \mathbf{B}$$

(3.37) er en rekursiv formel som kan brukes til å finne tilstandsvektorens forløp i et vilkårlig langt tidsintervall $(0, t)$. En "liten" Δt velges, og $\Phi(\Delta t)$ og $\Delta(\Delta t)$ beregnes numerisk. Så deles intervallet $(0, t)$ opp i like lange biter Δt , og dette gir tidspunktene $t_0 = 0$, $t_1 = \Delta t$, $t_2 = 2\Delta t$, ..., $t_N = t$, der $N = t/\Delta t$. Så brukes (3.37) gjentatte ganger, slik:

$$\begin{aligned}\mathbf{x}(t_1) &= \Phi(\Delta t)\mathbf{x}(t_0) + \Delta(\Delta t)\mathbf{u}(t_0) \\ \mathbf{x}(t_2) &= \Phi(\Delta t)\mathbf{x}(t_1) + \Delta(\Delta t)\mathbf{u}(t_1) \\ &\dots\text{OSV}\dots\end{aligned}\quad (3.40)$$

Numerisk beregning for å finne $\mathbf{x}(t)$ via denne rekursive metoden, ikke algebraisk løsning, er den absolutt mest vanlige framgangsmåte, feks. i programsystemer som MATLAB.

EKSEMPEL 3.3: Å finne transisjonsmatrisen til et gitt system

Vi skal betrakte svingende-masse-systemet i eksempel 2.5, 2.9 og 3.2.

Vi benytter initialverdiene $x_1(0) = x_2(0) = 0$. Dersom $k = 84$ [N/m], $f = 40$ [Ns/m] og $m = 4$ [kg] vil likning (3.8) bli

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -21 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0.25 \end{bmatrix} u \quad (3.41)$$

Vi skal finne det eksakte uttrykket for transisjonsmatrisen og benytter oss av diagonalisering. Avsnitt A.1.1 og A.1.3 forklarer nærmere hvordan vi finner en matrisens egenverdier og egenvektormatrise.

Vi finner egenverdiene ved å løse $|\mathbf{A} - \lambda \mathbf{I}| = 0$.

$$\begin{vmatrix} -\lambda & 1 \\ -21 & -10 - \lambda \end{vmatrix} = 0 \quad (3.42)$$

Dette gir

$$\lambda^2 + 10\lambda + 21 = 0 \quad (3.43)$$

Røttene til (3.43) er $\lambda_1 = -7$ og $\lambda_2 = -3$.

Vi finner egenvektorer ved hjelp av likningen $\mathbf{A}\mathbf{m} = \lambda\mathbf{m}$. For λ_1 gir dette

$$\begin{aligned} m_{21} &= -7m_{11} \\ -21m_{11} - 10m_{21} &= -7m_{21} \end{aligned} \quad (3.44)$$

Vi velger $m_{11} = 1$ og får $m_{21} = -7$. Egenverdien λ_2 gir en egenvektor

$$m_{12} = -0.33 \quad \text{og} \quad m_{22} = 1$$

Egenvektormatrisen $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2]$ og dens inverse er da gitt ved

$$\mathbf{M} = \begin{bmatrix} 1 & -\frac{1}{3} \\ -7 & 1 \end{bmatrix} \quad \text{og} \quad \mathbf{M}^{-1} = \frac{1}{4} \begin{bmatrix} -3 & -1 \\ -21 & -3 \end{bmatrix} \quad (3.45)$$

Vi finner transisjonsmatrisen ved å benytte (3.20)

$$\begin{aligned}\Phi(t) &= e^{At} = \mathbf{M}e^{\Lambda t}\mathbf{M}^{-1} = \frac{1}{4} \begin{bmatrix} 1 & -\frac{1}{3} \\ -7 & 1 \end{bmatrix} \begin{bmatrix} e^{-7t} & 0 \\ 0 & e^{-3t} \end{bmatrix} \begin{bmatrix} -3 & -1 \\ -21 & -3 \end{bmatrix} \\ &= \frac{1}{4} \begin{bmatrix} -3e^{-7t} + 7e^{-3t} & -e^{-7t} + e^{-3t} \\ 21e^{-7t} - 21e^{-3t} & 7e^{-7t} - 3e^{-3t} \end{bmatrix}\end{aligned}\quad (3.46)$$

Vi vil nå benytte oss av løsningen i (3.37) og dele opp det totale tidsintervallet i små stykker av lengde Δt . Vi vil så undersøke tilnærmelsen i (3.39) for forskjellige størrelser av Δt . Skritt lengden Δt må velges liten i forhold til den inverse tallverdien til den egenverdien som har størst tallverdi; i dette eksemplet er denne $\lambda_1 = -7$. Vi velger da

$$\Delta t = 0.2 \frac{1}{|\lambda_1|} \approx 0.03$$

Dette gir ved bruk av (3.46)

$$\Phi(0.03) = \begin{bmatrix} 0.991441 & 0.025837 \\ -0.542571 & 0.733074 \end{bmatrix}$$

Bruker vi approksimasjonen i (3.39), får vi

$$\Phi(0.03) = \begin{bmatrix} 0.991495 & 0.025856 \\ -0.542966 & 0.732940 \end{bmatrix}$$

Feilen som oppstår ved å benytte approksimasjonen, er ikke så stor. Vi velger nå et større tidsskritt, for eksempel $\Delta t = 0.3$, og finner den nøyaktige transisjonsmatrisen av (3.46)

$$\Phi(0.3) = \begin{bmatrix} 0.619655 & 0.071028 \\ -1.49159 & -0.090628 \end{bmatrix}$$

Approksimasjonen i (3.39) gir

$$\Phi(0.3) = \begin{bmatrix} 1.000000 & 0.205500 \\ -4.315500 & -1.055000 \end{bmatrix}$$

Feilen som oppstår ved å benytte approksimasjonen, er stor i dette tilfellet på grunn av at tidsskrittet er for langt.

Transisjonsmatrisen kan finnes ved rekkeutvikling i MATLAB:

Figur 3.4
MATLAB-kode for
å finne transisjons-
matrisen

```

1 dt = 0.03;
2 A = [0, 1; -21, -10];
3 PHI = expm(A*dt);
4
5 %
6 % Se også funksjonen c2d i MATLAB

```



Når transisjonsmatrisen for et gitt tidsskritt er bestemt med stor nøyaktighet og begynnelses-tilstanden er kjent, kan tilstandsvektorens utvikling i følge (3.37) bestemmes skrittvis:

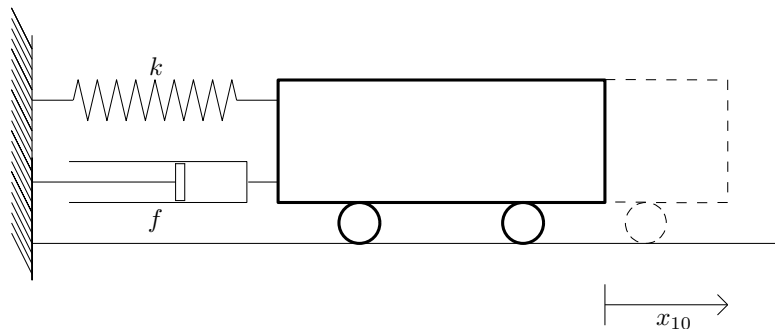
$$\begin{aligned}
 \mathbf{x}(\Delta t) &= \Phi(\Delta t)\mathbf{x}(0) + \Delta(\Delta t)\mathbf{u}(0) \\
 \mathbf{x}(2\Delta t) &= \Phi(\Delta t)\mathbf{x}(\Delta t) + \Delta(\Delta t)\mathbf{u}(\Delta t) \\
 \cdot &= \cdot \\
 \cdot &= \cdot \\
 \mathbf{x}(i\Delta t) &= \Phi^i(\Delta t)\mathbf{x}(0) + \Phi^{i-1}(\Delta t)\Delta(\Delta t)\mathbf{u}(0) + \dots + \Delta(\Delta t)\mathbf{u}(i-1)\Delta t
 \end{aligned} \tag{3.47}$$

Vi ser at feil i $\mathbf{x}(i\Delta t)$ vil bygge seg opp fordi transisjonsmatrisen potenseres mange ganger. Det bør derfor tas med mange ledd for mest mulig nøyaktig $\Phi(\Delta t)$, når den beregnes ved rekkeutvikling.

EKSEMPEL 3.4: Svingende masse – autonomt forløp

Fjærkonstanten og dempekonstanten er som oppgitt i eksempel 3.3, og det antas ingen friksjon mot underlaget. Massen trekkes så en lengde x_{10} til høyre og holdes i ro der. Dette gjør at x_2 er null (ingen hastighet). Ved tiden $t = 0$ slippes den. Det er da ingen pådrag i systemet, dvs at $\mathbf{u}(t) = \mathbf{0}$ og systemet sies derfor å være autonomt.

Figur 3.5
Svingende masse
fra eksempel 3.3



Fra likning (3.22), får vi da følgende uttrykk for $\mathbf{x}(t)$:

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

Med $\mathbf{x}(t) = [x_1(t) \ x_2(t)]^T$, initialbetingelser $x_0 = [x_{10} \ 0]^T$ og matrisen $\Phi(t)$ som oppgitt i likning (3.46), får vi

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -3e^{-7t} + 7e^{-3t} & -e^{-7t} + e^{-3t} \\ 21e^{-7t} - 21e^{-3t} & 7e^{-7t} - 3e^{-3t} \end{bmatrix} \begin{bmatrix} x_{10} \\ 0 \end{bmatrix}$$

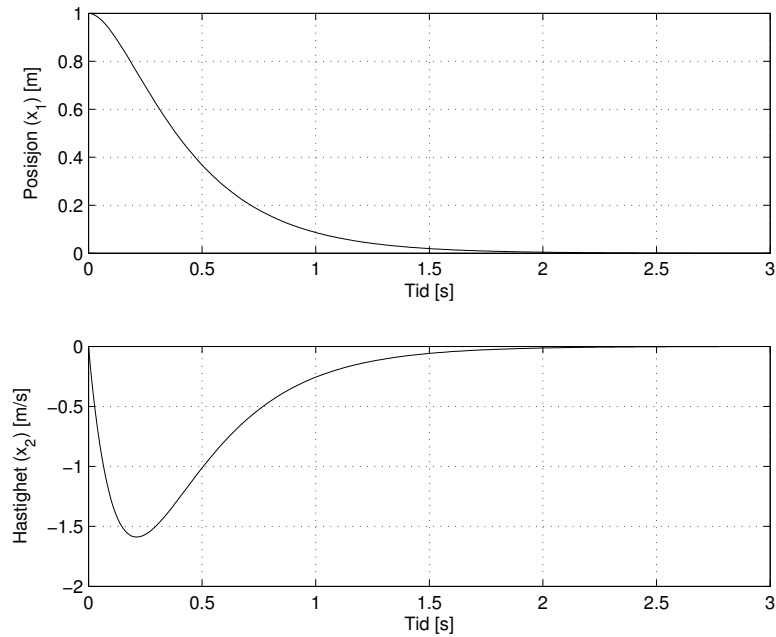
Ganger vi dette ut, får vi følgende uttrykk for $x_1(t)$ og $x_2(t)$:

$$x_1(t) = \frac{1}{4}(-3e^{-7t} + 7e^{-3t})x_{10}$$

$$x_2(t) = \frac{1}{4}(21e^{-7t} - 21e^{-3t})x_{10}$$

I figur 3.6 vises tidsforløpet for x_1 og x_2 for startposisjonen $x_{10} = 1$ m. Merk at hastigheten $x_2(t)$ både starter og ender i null.

Figur 3.6
Forløp av x_1 og x_2



■

3.3.3 Transisjonsmatrisen og Cayley-Hamiltons teorem

En annen metodikk for å finne transisjonsmatrisen algebraisk benytter seg av *Cayley-Hamiltons teorem*. Vi rekapitulerer at det karakteristiske polynomet for en matrise \mathbf{A} er gitt ved:

$$\rho(\lambda) = |\lambda\mathbf{I} - \mathbf{A}| = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 \quad (3.48)$$

og ifølge Cayley-Hamiltons teorem oppfyller \mathbf{A} likningen

$$\rho(\mathbf{A}) = \mathbf{A}^n + \alpha_{n-1}\mathbf{A}^{n-1} + \dots + \alpha_1\mathbf{A} + \alpha_0\mathbf{I} = \mathbf{0} \quad (3.49)$$

La oss nå vise at dette er tilfelle.

$$\begin{aligned} \rho(\mathbf{A}) &= ((\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1})^n + \alpha_{n-1}(\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1})^{n-1} + \dots + \alpha_1\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1} + \alpha_0\mathbf{I}) \\ &= \mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1}\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1}\dots\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1} + \dots + \alpha_1\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1} + \alpha_0\mathbf{I} \end{aligned}$$

Siden $\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$ har vi at

$$\begin{aligned} \rho(\mathbf{A}) &= \mathbf{M}\mathbf{\Lambda}^n\mathbf{M}^{-1} + \alpha_{n-1}\mathbf{M}\mathbf{\Lambda}^{n-1}\mathbf{M}^{-1} + \dots + \alpha_1\mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1} + \alpha_0\mathbf{I} \\ &= \mathbf{M}(\mathbf{\Lambda}^n + \alpha_{n-1}\mathbf{\Lambda}^{n-1} + \dots + \alpha_1\mathbf{\Lambda} + \alpha_0\mathbf{I})\mathbf{M}^{-1} \\ &= \mathbf{M} \begin{bmatrix} \lambda_1^n + \dots + \alpha_0 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2^n + \dots + \alpha_0 & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \lambda_n^n + \dots + \alpha_0 \end{bmatrix} \mathbf{M}^{-1} \\ &= \mathbf{M}[\mathbf{0}]\mathbf{M}^{-1} = [\mathbf{0}] \end{aligned} \quad (3.50)$$

□

Etter likning (3.49) kan vi nå uttrykke \mathbf{A}^n ved

$$\mathbf{A}^n = - \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \quad (3.51)$$

det vil si at \mathbf{A}^n kan uttrykkes ved en endelig potensrekke i \mathbf{A} av grad mindre enn n . Vi multipliserer med \mathbf{A} på begge sider i (3.51) og får

$$\mathbf{A}^{n+1} = - \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^{i+1} = -(\alpha_0\mathbf{A} + \dots + \alpha_{n-2}\mathbf{A}^{n-1} + \alpha_{n-1}\mathbf{A}^n)$$

Dersom vi setter inn \mathbf{A}^n fra (3.51), får vi

$$\mathbf{A}^{n+1} = - \sum_{i=0}^{n-1} \beta_i \mathbf{A}^i \quad (3.52)$$

(3.52) forteller at også \mathbf{A}^{n+1} kan uttrykkes som en endelig potensrekke i \mathbf{A} av grad mindre enn n . Ved så å gå løs på (3.52) finner vi det samme om \mathbf{A}^{n+2} , og generelt følger at enhver potens i \mathbf{A} kan uttrykkes ved en endelig potensrekke i \mathbf{A} av grad mindre enn n .

Av dette følger at matriseeksponentialfunksjonen kan uttrykkes som en endelig potensrekke i \mathbf{A} av grad mindre enn n . Men koeffisientene vil bli tidsfunksjoner. Vi kan derfor skrive

$$\Phi(t) = e^{\mathbf{A}t} = \alpha_0(t)\mathbf{I} + \alpha_1(t)\mathbf{A} + \dots + \alpha_{n-1}(t)\mathbf{A}^{n-1} \quad (3.53)$$

Dersom egenverdiene til \mathbf{A} er distinkte, er det nå enkelt å bestemme koeffisientene $\alpha_0(t)$, ..., $\alpha_n(t)$. Ved bruk av matrisen \mathbf{M} får vi da

$$\mathbf{M}^{-1} e^{\mathbf{A}t} \mathbf{M} = e^{\Lambda t} = \alpha_0(t)\mathbf{I} + \alpha_1(t)\Lambda + \dots + \alpha_{n-1}(t)\Lambda^{n-1} \quad (3.54)$$

som på komponentform simpelthen uttrykker

$$\begin{aligned} e^{\lambda_1 t} &= \alpha_0(t) + \alpha_1(t)\lambda_1 + \alpha_2(t)\lambda_1^2 + \dots + \alpha_{n-1}(t)\lambda_1^{n-1} \\ e^{\lambda_2 t} &= \alpha_0(t) + \alpha_1(t)\lambda_2 + \alpha_2(t)\lambda_2^2 + \dots + \alpha_{n-1}(t)\lambda_2^{n-1} \\ &\cdot \\ &\cdot \\ &\cdot \\ e^{\lambda_n t} &= \alpha_0(t) + \alpha_1(t)\lambda_n + \alpha_2(t)\lambda_n^2 + \dots + \alpha_{n-1}(t)\lambda_n^{n-1} \end{aligned} \quad (3.55)$$

Siden egenverdiene $\lambda_1, \dots, \lambda_n$ er antatt å være distinkte, har likningssystemet (3.55) en og bare en løsning. Vi skal illustrere framgangsmåten med et eksempel.

EKSEMPEL 3.5: Beregning av $\Phi(t)$ ved bruk av Cayley-Hamiltons teorem (endelig potensrekke)

Vi tar for oss systemet

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & 1 \\ 0 & -2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$|\lambda \mathbf{I} - \mathbf{A}| = 0$ gir oss egenverdiene $\lambda_1 = -1$ og $\lambda_2 = -2$. Likning (3.55) gir da

$$\begin{aligned} e^{-t} &= \alpha_0(t) - \alpha_1(t) \\ e^{-2t} &= \alpha_0(t) - 2\alpha_1(t) \end{aligned} \tag{3.56}$$

Vi subtraherer den andre linjen fra den første og finner

$$\alpha_1(t) = e^{-t} - e^{-2t}$$

Vi setter denne inn i (3.56) og får

$$\alpha_0(t) = 2e^{-t} - e^{-2t}$$

Likning (3.53) gir da

$$\begin{aligned} \Phi(t) = e^{\mathbf{A}t} &= (2e^{-t} - e^{-2t}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (e^{-t} - e^{-2t}) \begin{bmatrix} -1 & 1 \\ 0 & -2 \end{bmatrix} \\ &= \begin{bmatrix} e^{-t} & e^{-t} - e^{-2t} \\ 0 & e^{-2t} \end{bmatrix} \end{aligned} \tag{3.57}$$

■

Hittil har vi antatt distinkte egenverdier. Men hva om vi ikke har distinkte egenverdier og ikke et fullt sett av lineært uavhengige egenvektorer? (Distinkte egenverdier impliserer alltid et fullt sett av lineært uavhengige egenvektorer). Her kunne vi benyttet den såkalte Jordan-kanoniske form som kan betraktes som en generalisert diagonalisering, se Strang (1988). Men siden vi kan gjøre alle egenverdier distinkte ved en ubetydelig endring av parametre i \mathbf{A} , vil vi i denne boka for enkelhets skyld anta at matrisen \mathbf{A} alltid kan diagonaliseres når $\Phi(t)$ skal finnes algebraisk. Men vi vil senere se at vi kan håndtere sammenfallende egenverdier algebraisk ved hjelp av Laplacetransformasjonen.

Når $\Phi(\Delta t)$ derimot skal finnes *numerisk* for et fast tidsintervall, bruker vi rekkeutvikling, jfr. (3.40). Da trenger vi ikke forutsette distinkte egenverdier.

3.4 MONOVARIABLE OG MULTIVARIABLE SYSTEMER

Et dynamisk system er **multivariabelt** når det utsettes for flere uavhengige pådrag (\mathbf{u}) og/eller observeres gjennom flere målinger (\mathbf{y}). Hvor mange tilstandsvariable som finnes, er ikke avgjørende for om systemet er **monovariabelt** eller multivariabelt. Vi kan altså ha et monovariabelt system av høy orden.

Et monovariabelt system er et spesialtilfelle av et multivariabelt system. Et monovariabelt system har ett pådrag og én måling (\mathbf{u} og \mathbf{y} er skalare). Mange multivariabel systemer kan betraktes som et antall uavhengige monovariabel systemer dersom det ikke finnes noen vesentlig kopling mellom de forskjellige systemdelene.

Et monovariabelt system kan uttrykkes som

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}u + \mathbf{E}\mathbf{v} \\ y &= \mathbf{C}\mathbf{x} + w\end{aligned}\tag{3.58}$$

Størrelsen w (som i (3.58) er en skalar) representerer usikkerhet (eller **målestøy**) knyttet til målingene \mathbf{y} . Vektoren \mathbf{v} inneholder et sett **prosessforstyrrelser** (**prosessstøy**).

Fordi \mathbf{u} og \mathbf{y} er skalare størrelser i monovariabel systemer, innfører vi ofte i stedet for \mathbf{B} og \mathbf{C}

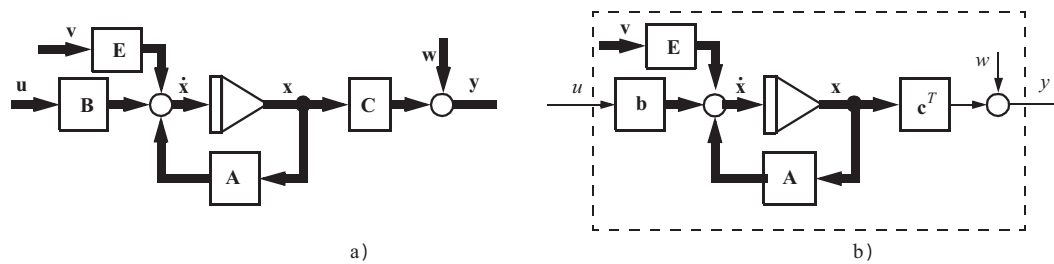
$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \text{og} \quad \mathbf{c}^T = [c_1 \quad c_2 \quad \cdots \quad c_n]$$

Vi kan da skrive

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u + \mathbf{E}\mathbf{v} \\ y &= \mathbf{c}^T \mathbf{x} + w\end{aligned}\tag{3.59}$$

Et blokkdiagram kan også brukes til å representere vektorer og matriser i likhet med det som er gjort i avsnitt 2.8 for skalare størrelser. Figur 3.7 a viser et blokkdiagram der de tjukke linjene representerer vektorer, de firkantede blokkene representerer matriser og integrator-symbolet er et sett av n integratorer.

Figur 3.7
 Blokkdiagram for
 a) Multivariabelt system
 b) Monovariabelt system



En av de store fordelene med blokkdiagram-representasjon er at forskjellige undersystemer med hvert sitt blokkdiagram kan kombineres og koples sammen, for å beskrive og gi oversikt over kompliserte strukturer. Dette vil bli illustrert senere.

Figur 3.7 b) viser hvordan blokkdiagrammet er for det monovariabelt systemet. Dersom $\mathbf{v} = \mathbf{0}$ og $\mathbf{w} = \mathbf{0}$, har vi et n -te ordens system som påvirkes av en variabel u til venstre og gir en respons y til høyre, med dynamikk gitt av de n tilstandsvariable innenfor den stiplede konturen.

Vi skal nå se på dette monovariabelt systemets impulsrespons, et begrep som vi introduserte i avsnitt 2.4:

3.4.1 Impulsrespons for monovariabelt system av orden n

I avsnitt 2.4 utledet vi impulsresponsen for et 1. ordens system, likning (2.33). Vi skal nå utlede impulsresponsen for et monovariabelt system av orden n .

Likning (3.22) for et monovariabelt system (skalar u) blir

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{\mathbf{A}(t-\tau)}\mathbf{b}u(\tau)d\tau \quad (3.60)$$

Impulsresponsen er definert som systemets utgang når pådraget er $u(t) = \delta(t)$, og starttilstanden $\mathbf{x}(t_0) = \mathbf{0}$. Vi får

$$h(t) = y(t) = \mathbf{c}^T \mathbf{x}(t) = \mathbf{c}^T \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{b}\delta(\tau)d\tau = \mathbf{c}^T e^{\mathbf{A}t}\mathbf{b} \quad (3.61)$$

Det siste trinnet til høyre følger av at integranden bare er $\neq 0$ når $\tau = 0$. Arealet under integranden blir da $\mathbf{c}^T e^{\mathbf{A}t}\mathbf{b}$ ganger arealet under en impuls (som har areal = 1).

Vi kan konstatere at $h(t) = \mathbf{c}^T e^{\mathbf{A}t}\mathbf{b}$ er en generalisering av formelen for impulsresponsen til et 1. ordens system, jfr. likning (2.33).

3.5 VALG AV TILSTANDSVARIABLE

Et valg av tilstandsvariable for beskrivelse av et dynamisk system er ikke entydig. For lineære systemer med konstante koeffisienter gjelder det at hvilken som helst lineær og ikke-singulær

transformasjon av en tilstandsvektor gir en ny mulig tilstandsvektor. Som oftest ønsker vi å velge matematiske tilstandsvariable som er lik iøynefallende fysiske tilstandsvariable ved oppsetting av matematiske modeller, for eventuelt siden å foreta transformasjoner.

I varmtvannsberederen i eksempel 2.1 kunne vi for eksempel valgt energiinnholdet i hhv. vannet og varmekolben som tilstander, i stedet for temperaturen tilsvarende steder. Men det er viktig å merke seg at dimensjonen på \mathbf{x} , det vil si *antallet* differensiallikninger, er entydig. Dette antallet kan ikke endres for en modell. For berederen betyr det at modellen med de antakelser som ligger inne, alltid vil være gitt av to 1. ordens differensiallikninger.

Hvis vi i det følgende ser bort fra målestøy og prosessstøy, er den matematiske modellen for et lineært multivariabelt dynamisk system med konstante koeffisienter

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx}\end{aligned}\tag{3.62}$$

\mathbf{x} er tilstandsvektor, \mathbf{u} er pådragsvektor, \mathbf{y} er målevektor¹. De forskjellige matrisene gis følgende betegnelser

A = systemmatrise

B = pådragsmatrise

C = målematrise

Ved forskjellige valg av tilstandsvariable påvirkes alle matrisene i (3.62). Formen på systemmatrisen **A** er spesielt interessant. Ved en bestemt systemkonfigurasjon med gitte pådrag og målinger vil den totale sammenhengen mellom pådrag og målinger være uforandret selv om man endrer valget av tilstandsvariable.

¹Vi har her valgt en notasjon som svarer til den som er mest utbredt internasjonalt. Mange andre notasjoner finnes i bruk.

Vi skal se på fem viktige systemstrukturer som gir karakteristiske former på systemmatrisen. For enkelhets skyld studerer vi disse bare for det *monovariabel* tilfellet. Etter valg av tilstandsvariable skal vi referere til disse som:

Form 1: Generelle variable

Form 2: **Kanonisk form**¹ med diagonaliserte variable

Form 3: Kanonisk form med fasevariable

Form 4: Kanonisk form, en annen variant

Form 5: Kaskadevariable

Form 1: “Generelle variable” indikerer et valg av variable som gjør at systemmatrisen ikke får noen karakteristisk struktur. De fleste eller alle elementene i matrisen kan ha en verdi, og plasseringen av nuller (elementer med verdi null) er ikke strengt systematisk. Denne formen vil vi ofte møte når fysiske variable velges som tilstandsvariable. I det monovariabel tilfellet vil systemmodellen (3.62) bli

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u \\ y &= \mathbf{c}^T \mathbf{x}\end{aligned}\tag{3.63}$$

Form 2: En **diagonalisering** av systemet i (3.63) ved hjelp av transformasjonen i (3.11), $\mathbf{x} = \mathbf{M}\mathbf{q}$, gir

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{\Lambda}\mathbf{q} + \mathbf{h}u = \mathbf{\Lambda}\mathbf{q} + \mathbf{p} \\ y &= \mathbf{f}^T \mathbf{q}\end{aligned}\tag{3.64}$$

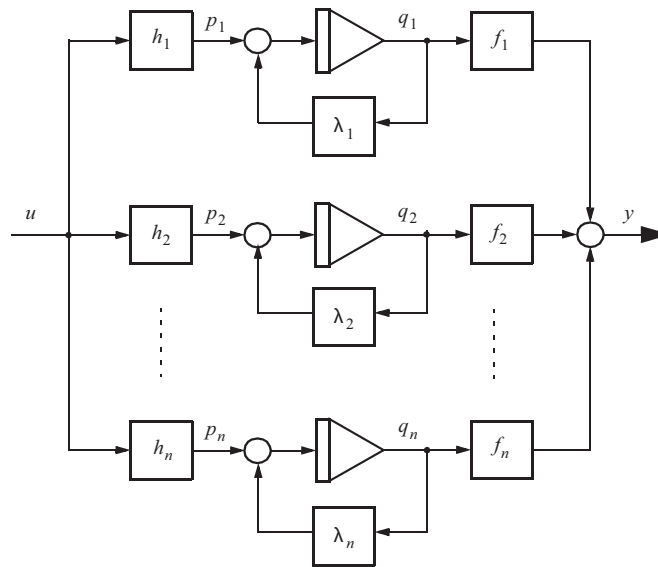
der $\mathbf{\Lambda} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}$, $\mathbf{h} = \mathbf{m}^{-1}\mathbf{b}$ og $\mathbf{f}^T = \mathbf{c}^T\mathbf{M}$

Figur 3.8 viser et elementært blokkdiagram for (3.64).

Det fremgår tydelig hvordan de kanoniske tilstandene opptrer i parallell og må summeres for å gi systemutgangen (målingen). Med denne representasjonsformen, som baserer seg på en matematisk transformasjon, vil ikke tilstandene ha noen fysisk tolkning.

¹**Kanonisk form** kan forenklet sies å være et felles navn på alle tilstandsromstrukturer hvor matrisene inneholder det minimale antall koeffisienter som skal til for å representere et gitt system.

Figur 3.8
Diagonalt system



Form 3: I forbindelse med omtalen av likningssystemet i (2.77) har vi allerede introdusert begrepet **fasevariable**. Figur 3.9 viser et system som er karakterisert ved følgende differensiallikninger:

$$\begin{aligned}
 \dot{\eta}_1 &= \eta_2 \\
 \dot{\eta}_2 &= \eta_3 \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 \dot{\eta}_n &= -\alpha_0 \eta_1 - \alpha_1 \eta_2 - \dots - \alpha_{n-1} \eta_n + u \\
 y &= \tilde{c}_1 \eta_1 + \tilde{c}_2 \eta_2 + \dots + \tilde{c}_n \eta_n
 \end{aligned} \tag{3.65}$$

Skrevet på vektorform vil dette bli

$$\begin{aligned}
 \dot{\boldsymbol{\eta}} &= \tilde{\mathbf{A}}\boldsymbol{\eta} + \tilde{\mathbf{b}}u \\
 y &= \tilde{\mathbf{c}}^T \boldsymbol{\eta}
 \end{aligned} \tag{3.66}$$

der

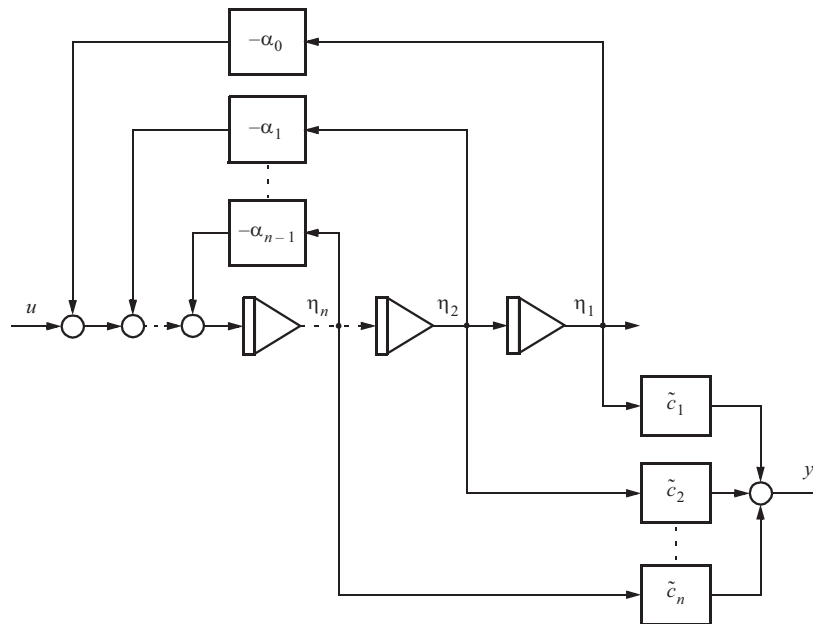
$$\tilde{\mathbf{A}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-2} & -\alpha_{n-1} \end{bmatrix}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} \quad (3.67)$$

$$\tilde{\mathbf{c}}^T = [\tilde{c}_1 \quad \tilde{c}_2 \quad \tilde{c}_3 \quad \cdots \quad \tilde{c}_n]$$

Den formen systemmatrisen $\tilde{\mathbf{A}}$ her antar er spesiell fordi den bare har n karakteriserende elementer. Hvis vi forsøker å bestemme egenverdiene til matrisen $\tilde{\mathbf{A}}$, vil vi finne at matrisens karakteristiske likning vil anta formen

$$\lambda^n + \alpha_{n-1}\lambda^{n-1} + \alpha_{n-2}\lambda^{n-2} + \dots + \alpha_1\lambda + \alpha_0 = 0 \quad (3.68)$$

Figur 3.9
Fasevariable



Elementene i $\tilde{\mathbf{A}}$ er altså koeffisienter i matrisens karakteristiske likning. Vi skal siden se at denne spesielle formen leder til et særlig enkelt svar, når vi foretar Laplacetransformasjon av systemet.

Det vil i mange tilfeller være naturlig å velge en slik representasjonsform når det forekommer fysiske variable som er tidsderiverte av hverandre, som for eksempel posisjon, hastighet og akselerasjon. Systemet med svingende masse i eksempel 3.2 er et eksempel på fasevariabel form. Det kan forøvrig vises at et generelt system som i (3.62) under visse betingelser (styrbarhet, se avsnitt 5.2) har en lineær transformasjon $\mathbf{x} = \mathbf{S}h$, som kan bringe systemet fra den opprinnelige formen over på formen (3.66).

Form 4: Figur 3.10 viser et system med en tredje **kanonisk form** som likner på den foregående formen.

Differensiallikningen for dette systemet vil være av formen

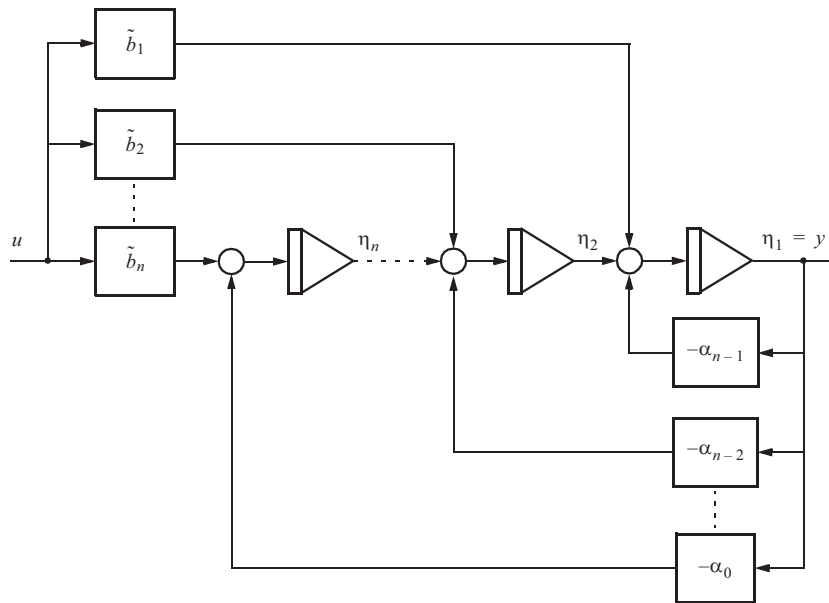
$$\dot{\boldsymbol{\eta}} = \tilde{\mathbf{A}}\boldsymbol{\eta} + \tilde{\mathbf{b}}u \quad , \quad y = \tilde{\mathbf{c}}^T \boldsymbol{\eta} \quad (3.69)$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} -\alpha_{n-1} & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ -\alpha_{n-2} & 0 & 1 & & & & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -\alpha_1 & 0 & 0 & \cdot & \cdot & 0 & 1 \\ -\alpha_0 & 0 & 0 & \cdot & \cdot & 0 & 0 \end{bmatrix} , \quad \tilde{\mathbf{b}} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \cdot \\ \cdot \\ \cdot \\ \tilde{b}_{n-1} \\ \tilde{b}_n \end{bmatrix}$$

$$\tilde{\mathbf{c}}^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

De samme kommentarene gjelder for denne matrisen $\tilde{\mathbf{A}}$ som den foregående, men tilstandsvariablene kalles i dette tilfellet ikke for fasevariable.

Figur 3.10
Kanonisk form 3



Form 5: Figur 3.11 viser et system som består av flere kaskadekoblede (seriekoblede) 1. ordens systemer. Denne formuleringen forekommer ofte i beskrivelsen av fysiske prosesser.

Differensiallikningene for systemet vil bli

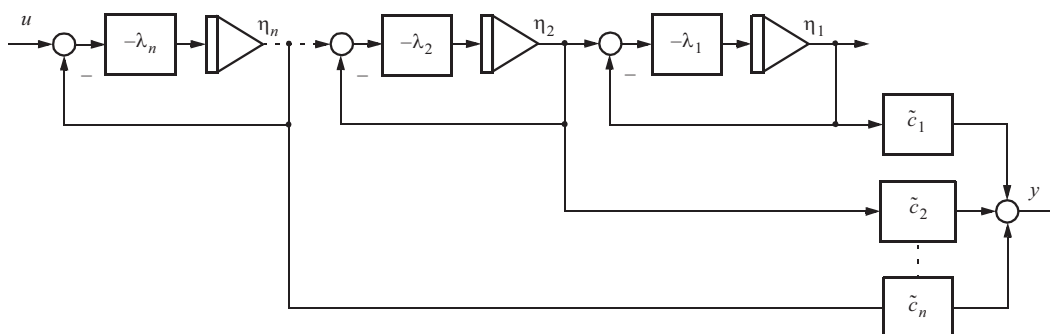
$$\dot{\boldsymbol{\eta}} = \tilde{\mathbf{A}}\boldsymbol{\eta} + \tilde{\mathbf{b}}u \quad , \quad y = \tilde{\mathbf{c}}^T \boldsymbol{\eta} \quad (3.70)$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} \lambda_1 & -\lambda_1 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & -\lambda_2 & 0 & \cdot & \cdot & 0 \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ -\lambda_n \end{bmatrix}$$

$$\tilde{\mathbf{c}}^T = [\tilde{c}_1 \quad \tilde{c}_2 \quad \cdots \quad \tilde{c}_n]$$

Karakteristisk for matrisen $\tilde{\mathbf{A}}$ er at den har *båndstruktur* idet det bare finnes betydningsfulle parametre i et smalt bånd (i dette tilfellet med bredde på to elementer) omkring diagonalen.

Figur 3.11
Kaskadevariable



Ved formulering av matematiske modeller er det mest naturlig å benytte fysisk relevante variable \mathbf{x} som tilstandsvariable. Dersom den matematiske modellen får en uhensiktsmessig form for videre behandling, kan en lineær transformasjon til en annen tilstandsvektor foretas ved hjelp av en fritt valgt kvadratisk matrise \mathbf{Z} som bare trenger å oppfylle kravet om å være inverterbar. Transformasjonen er

$$\mathbf{x} = \mathbf{Z}\boldsymbol{\gamma} \quad \text{eller} \quad \boldsymbol{\gamma} = \mathbf{Z}^{-1}\mathbf{x} \quad (3.71)$$

for derved å oppnå en systemmatrise som har en mer hensiktsmessig form. Denne vil bli

$$\tilde{\mathbf{A}} = \mathbf{Z}^{-1}\mathbf{A}\mathbf{Z} \quad (3.72)$$

Mer om dette i avsnitt 4.6.2

3.6 ULINEÆRE VEKTORDIFFERENSIALLIKNINGER. LINEARISERING

En prosess kan være lineær innenfor et visst område og ulineær utenfor området, dersom den matematiske karakteristikken for systemet endres utenfor området. En fjær er et eksempel på dette. Dersom man trekker så hardt i fjæren at man strekker den utenfor det elastiske området, risikerer man at fjæren mister elastisiteten og blir ulineær (dvs. fjærkonstanten endrer seg).

Fysiske prosesser er aldri helt lineære, men vi kan ofte **linearisere** prosessen for små svingninger (perturbasjoner) omkring et **arbeidspunkt**. Et arbeidspunkt er en bestemt tilstand \mathbf{x}^p og ett bestemt pådrag \mathbf{u}^p som vi velger å betrakte. Når vi “lineariserer rundt arbeidspunktet” betyr dette at vi lager en lineær differensiallikning, som er en god tilnærming til den ulineære “i nærheten av” \mathbf{x}^p og \mathbf{u}^p .

Vi ønsker ofte å linearisere ulineære prosesser fordi

- Lineære prosesser er vesentlig enklere å analysere enn ulineære prosesser.
- Laplacetransformasjon, som ofte brukes til å representere en matematisk modell av en prosess, krever at prosessmodellen er lineær.
- Mange teknikker for valg og dimensjonering av regulatorer forutsetter lineære prosesser.

Gitt en ulineær vektordifferensiallikning (svarer til (3.5), men vi forutsetter nå tidsinvariant prosess).

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad (3.73)$$

Vi betrakter først en av likningene i (3.73)

$$\frac{dx_i}{dt} = f_i(x_1, \dots, x_n, u_1, \dots, u_r) \quad (3.74)$$

Vi ønsker å linearisere (3.74) omkring arbeidspunktet p , det vil si $(x_1^p, \dots, x_n^p, u_1^p, \dots, u_r^p)$. Dersom vi har en liten forskyvning $\Delta \mathbf{x}$ og $\Delta \mathbf{u}$ vekk fra arbeidspunktet, kan (3.74) skrives

$$\frac{d}{dt}(x_i^p + \Delta x_i) = f_i(x_1^p + \Delta x_1, \dots, x_n^p + \Delta x_n, u_1^p + \Delta u_1, \dots, u_r^p + \Delta u_r) \quad (3.75)$$

Vi Taylor-rekkeutvikler (3.75) og neglisjerer alle ledd med orden større enn 1:

$$\begin{aligned} \dot{x}_i^p + \Delta \dot{x}_i \approx & f_i(x_1^p, \dots, x_n^p, u_1^p, \dots, u_r^p) + \\ & \left. \frac{\partial f_i}{\partial x_1} \right|_p \Delta x_1 + \dots + \left. \frac{\partial f_i}{\partial x_n} \right|_p \Delta x_n + \left. \frac{\partial f_i}{\partial u_1} \right|_p \Delta u_1 + \dots + \left. \frac{\partial f_i}{\partial u_r} \right|_p \Delta u_r \end{aligned} \quad (3.76)$$

(Merk: notasjonen " $\left. \right|_p$ " er kompakt skrivemåte for " $\left. \right|_{\mathbf{x}=\mathbf{x}^p, \mathbf{u}=\mathbf{u}^p}$ ").

Hvis vi benytter at

$$\dot{x}_i^p = f_i(x_1^p, \dots, x_n^p, u_1^p, \dots, u_r^p)$$

kan vi stryke \dot{x}_i^p og $f_i(x_1^p, \dots, x_n^p, u_1^p, \dots, u_r^p)$ på hver side av likhetstegnet i (3.76). En modell for små utslag rundt arbeidspunktet blir da

$$\Delta \dot{x}_i = \left. \frac{\partial f_i}{\partial x_1} \right|_p \Delta x_1 + \dots + \left. \frac{\partial f_i}{\partial x_n} \right|_p \Delta x_n + \left. \frac{\partial f_i}{\partial u_1} \right|_p \Delta u_1 + \dots + \left. \frac{\partial f_i}{\partial u_r} \right|_p \Delta u_r \quad (3.77)$$

Dette kan skrives på følgende måte

$$\Delta \dot{x}_i = a_{i1} \Delta x_1 + \dots + a_{in} \Delta x_n + b_{i1} \Delta u_1 + \dots + b_{ir} \Delta u_r \quad (3.78)$$

Dersom vi benytter likning (3.78) på alle differensiallikningene i (3.73), kan vi samle alle de lineariserte differensiallikningene i en lineær vektordifferensiallikning

$$\Delta \dot{\mathbf{x}} = \mathbf{A} \Delta \mathbf{x} + \mathbf{B} \Delta \mathbf{u} \quad (3.79)$$

der \mathbf{A} og \mathbf{B} er gitt ved

$$\mathbf{A} = \begin{bmatrix} \left. \frac{\partial f_1}{\partial x_1} \right|_p & \left. \frac{\partial f_1}{\partial x_2} \right|_p & \dots & \left. \frac{\partial f_1}{\partial x_n} \right|_p \\ \left. \frac{\partial f_2}{\partial x_1} \right|_p & \left. \frac{\partial f_2}{\partial x_2} \right|_p & & \left. \frac{\partial f_2}{\partial x_n} \right|_p \\ \vdots & & & \vdots \\ \left. \frac{\partial f_n}{\partial x_1} \right|_p & \left. \frac{\partial f_n}{\partial x_2} \right|_p & \dots & \left. \frac{\partial f_n}{\partial x_n} \right|_p \end{bmatrix} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_p = \left\{ \left. \frac{\partial f_i}{\partial x_j} \right|_p \right\} \quad (3.80)$$

$$\mathbf{B} = \begin{bmatrix} \left. \frac{\partial f_1}{\partial u_1} \right|_p & \left. \frac{\partial f_1}{\partial u_2} \right|_p & \dots & \left. \frac{\partial f_1}{\partial u_n} \right|_p \\ \left. \frac{\partial f_2}{\partial u_1} \right|_p & \left. \frac{\partial f_2}{\partial u_2} \right|_p & & \left. \frac{\partial f_2}{\partial u_n} \right|_p \\ \vdots & & & \vdots \\ \left. \frac{\partial f_n}{\partial u_1} \right|_p & \left. \frac{\partial f_n}{\partial u_2} \right|_p & \dots & \left. \frac{\partial f_n}{\partial u_n} \right|_p \end{bmatrix} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right|_p = \left\{ \left. \frac{\partial f_i}{\partial u_j} \right|_p \right\} \quad (3.81)$$

Likning (3.79) er altså den lineariserte versjonen av likning (3.73) omkring et arbeidspunkt. I hvilke tilfeller kan vi benytte den lineariserte modellen i stedet for den ulineære modellen?

- Generelt kan vi benytte den lineariserte versjonen av prosessmodellen over et større område for \mathbf{x} og \mathbf{u} dersom prosessen ikke er sterkt ulineær. Ja, en lineær modell av typen (3.6) kan betraktes som en “linearisering” rundt “arbeidspunktet” $\mathbf{x}^p = \mathbf{0}$, $\mathbf{u}^p = \mathbf{0}$, med ubegrenset variasjonsområde for tilstander og pådrag.
- I reguleringssammenheng ønsker vi ofte at en ulineær prosess skal befinne seg i ett bestemt punkt, det vil si at \mathbf{x}^p og \mathbf{u}^p er konstante hele tiden. Dette er et spesialtilfelle, fordi vi til nå ikke har forutsatt konstant \mathbf{x}^p og \mathbf{u}^p . Men dette spesialtilfellet er meget vanlig, fordi vi i prosessregulering nettopp ønsker å holde forskjellige variable på konstante verdier (se eksempel 3.6 på neste side).

Begrepet “**arbeidspunkt**” reserveres ofte i litteraturen for spesialtilfellet $\mathbf{x}^p = \text{konstant}$ og $\mathbf{u}^p = \text{konstant}$. Da har vi $\dot{\mathbf{x}}^p = \mathbf{0}$ og (3.73) blir dermed

$$\mathbf{f}(\mathbf{x}^p, \mathbf{u}^p) = \mathbf{0} \quad (3.82)$$

det vil si et ulineært statisk likningssystem.

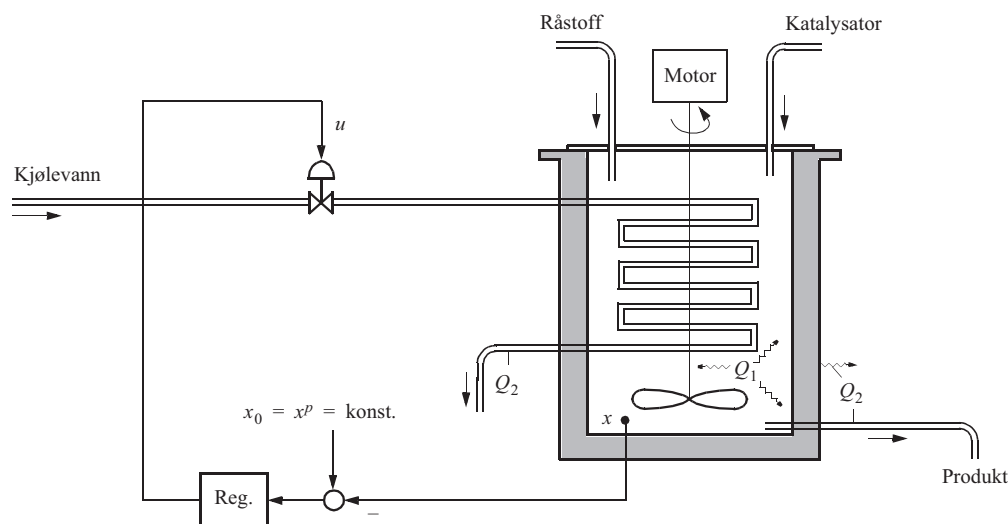
Dette kan brukes til å finne den konstante \mathbf{u}^p som svarer til den konstante tilstandsvektoren \mathbf{x}^p . Et punkt $\mathbf{x}^p, \mathbf{u}^p$ som oppfyller (3.82), kalles ofte også et **singulært punkt**.

I eksempel 8.15 skal vi se at slike punkter kan være *stabile* eller *ustabile*. I det første tilfellet vil prosessen komme tilbake til det singulære punktet (arbeidspunktet) selv om den utsettes for små forstyrrelser. Vi sier at vi har en stabil **likevektstilstand**. I det andre tilfellet vil den drive vekk fra punktet etter en liten forstyrrelse.

EKSEMPEL 3.6: Linearisering av reaktormodell

Figur 3.12 viser en prosess hvor råstoff og katalysator reagerer i en reaktor og danner et produkt som tappes ut. Temperaturen x i reaktoren antas å være den samme overalt (oppnås ved omrøring, se “propellen” i figuren). Reaksjonen er **eksoterm**, det vil si den utvikler effekten $Q_1(x)$. Denne effekten motvirkes av økt kjøling som skyldes varmeavgang fra reaktorens ytre og bortledet effekt via det utstrømmende produktet. I tillegg har vi som vist i figur 3.12 installert en regulert kjølevannstilførsel, u . Tilsammen utgjør disse kjøleeffekten Q_2 .

Figur 3.12
Kjemisk reaktor



Kjølevannsventilen har posisjonen u [0 - 100%]. Da blir den totale kjøleeffekten en funksjon $Q_2(x, u)$.

Varmebalansen (effektbalansen) for reaktoren blir

$$C\dot{x} = Q_1(x) + Q_2(x, u) \quad (3.83)$$

C er her varmekapasiteten for væsken i reaktoren.

Likning (3.83) svarer til (3.73) med $\mathbf{x} = x$ og $\mathbf{u} = u$, dvs. vi har en skalar differensiallikning

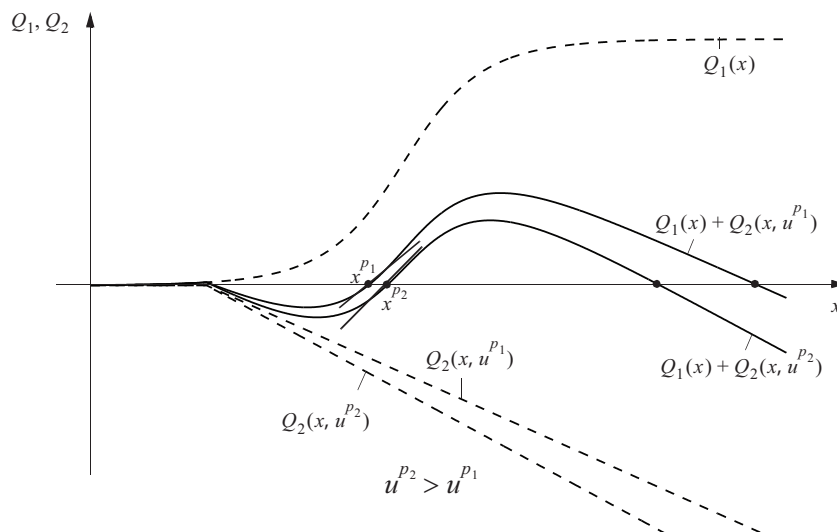
$$\dot{x} = f(x, u) = \frac{1}{C}[Q_1(x) + Q_2(x, u)] \quad (3.84)$$

Likning (3.83) viser at det til enhver statisk verdi av reaktortemperaturen x^p , må svare en verdi u^p , som oppfyller $\dot{x} = 0$ og som gir

$$Q_2(x^p, u^p) = -Q_1(x^p) \quad (3.85)$$

x^p, u^p er et arbeidspunkt vi kan velge for prosessen. Målet er å holde x lik $x^p = x_0$, og det gjør vi ved å holde pådraget u rundt u^p . Funksjonene $Q_1(x)$ og $Q_2(x, u)$ er vist i figur 3.13.

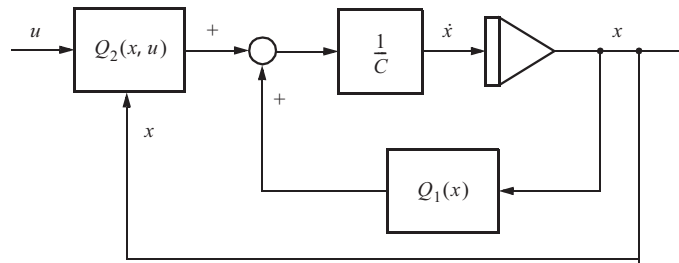
Figur 3.13
Linearisering
omkring arbeids-
punktene p_1 og p_2



$Q(x, u)$ er vist for to arbeidspunkter p_1 og p_2 , med korresponderende verdier x^{p_1} og x^{p_2} .

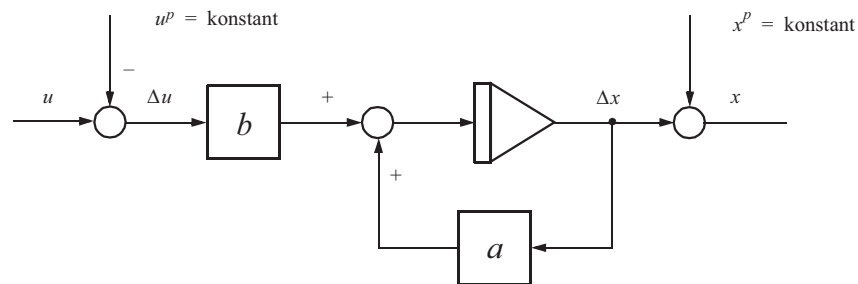
Den ulineære modellen (3.83) er vist i blokkdiagrammet i figur 3.14.

Figur 3.14
Blokkdiagram for
den ulineære
modellen



Modellen i (3.83) og i figur 3.12 er ulineær både med hensyn til reaktortemperatur x og pådraget u . En detaljert analyse av et slikt system skal ikke utføres i denne boka. Men når systemets variable endrer seg lite omkring arbeidspunktet, vil vi kunne betrakte det som lineært, og en lineær analyse er som nevnt vesentlig enklere enn en "ulineær analyse". For at vi skal kunne bruke reguleringssteori som bygger på lineære forutsetninger på den aktuelle prosessen, omformer vi modellen i figur 3.14 til den lineære modellen i figur 3.15.

Figur 3.15
Blokkdiagram for
den lineariserte
modellen



Den konstante u^p er beregnet på forhånd ut fra (3.85), eller funnet ved hjelp av en manuell innstilling av prosessen til konstant ønsket temperatur x^p .

Vi søker nå de ukjente koeffisientene a og b i

$$\Delta \dot{x} = a\Delta x + b\Delta u \quad (3.86)$$

(3.80) og (3.84) gir

$$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_p = \left. \frac{\partial f}{\partial x} \right|_p = \frac{1}{C} \left. \frac{\partial Q_1}{\partial x} \right|_p + \frac{1}{C} \left. \frac{\partial Q_2}{\partial x} \right|_p \quad (3.87)$$

$$\mathbf{B} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right|_p = \left. \frac{\partial f}{\partial u} \right|_p = \frac{1}{C} \left. \frac{\partial Q_2}{\partial u} \right|_p \quad (3.88)$$

Da blir

$$\Delta \dot{x} = \underbrace{\frac{1}{C} \left(\left. \frac{\partial Q_1}{\partial x} \right|_{x=x^p} + \left. \frac{\partial Q_2}{\partial x} \right|_{x=x^p, u=u^p} \right)}_a \Delta x + \underbrace{\frac{1}{C} \left. \frac{\partial Q_2}{\partial u} \right|_{x=x^p, u=u^p}}_b \Delta u = a\Delta x + b\Delta u \quad (3.89)$$

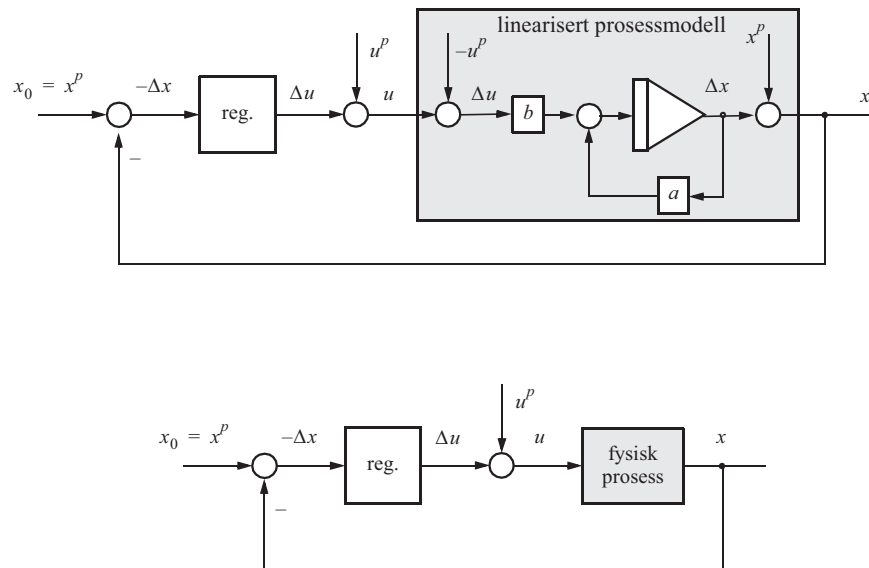
Vi benytter a og b i stedet for \mathbf{A} og \mathbf{B} siden systemet er skalart.

$C \cdot a$ i (3.89) svarer til vinkelkoeffisienten til tangenten i figur 3.13. To tangenter er vist, for arbeidspunkt x^{p1}, u^{p1} og x^{p2}, u^{p2} . Når a er positiv slik som i dette tilfellet, er prosessen *ustabil* rundt arbeidspunktet. Dette kommer vi tilbake til i en stabilitetsanalyse av samme prosess i eksempel 8.15. Merk at $b < 0$, fordi $Q_2(x, u)$ er monotont fallende med u , uavhengig av x : Økt u svarer til økt vanntilførsel og dermed økt kjøling.

Den endelige reguleringsstrukturen med tilbakekopling blir som vist i figur 3.16. Den øverste delfiguren viser den lineariserte prosess (3.89), som ligger til grunn for valg av regulatoren, den nederste viser den fysiske prosess med regulator.

Merk fortegnet, at det blir $-\Delta x$ foran regulatoren i blokkdiagrammet. Dette gjelder generelt, fordi det følger av definisjonen $x = x^p + \Delta x$.

Figur 3.16
Linearisert modell som brukes til å velge regulator (øverst), og blokkdiagram for den endelige (fysiske) reguleringsstrukturen, med den valgte regulator (nederst). Merk at pådraget inneholder en permanent konstant komponent u^p .



■

Laplacetransformasjon

4.1 Innledning og motivering 107

4.2 Definisjoner 108

4.3 Laplacetransformasjon av differensiallikninger 113

4.4 Invers Laplacetransformasjon ved residuregning 115

4.3.2 Høyere ordens lineær differensiallikning 114

4.5 Transferfunksjoner og blokkdiagrammer 119

4.6 Lineære vektordifferensiallikninger og Laplacetransformasjonen 129

4.6.1 Transfermatrise 132

4.7 Sammenhengen poler / nullpunkter og tidsrespons 138

4.1 INNLEDNING OG MOTIVERING

Laplacetransformasjon gjør det mulig å løse en lineær høyere ordens differensiallikning eller et sett av lineære differensiallikninger på en enkel måte: Ved å Laplacetransformere differensiallikningen (-ene) blir de omgjort til en eller et sett av *algebraiske* likninger. Disse løses med hensyn på variablene vi er interessert i, og så kan vi inverstransformere uttrykkene for å finne tidsforløpene til de samme variablene.

EKSEMPEL 4.1: Modell av svingende masse

Modellen av den svingende massen som ble introdusert i eksempel 2.5 og som er benyttet i flere etterfølgende eksempler, er på tilstandsromform gitt som

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k}{m}x_1 - \frac{f}{m}x_2 + \frac{1}{m}u \end{aligned} \tag{4.1}$$

hvor x_1 og x_2 er posisjonen og hastigheten til massen, og u er den påtrykte kraften.

Erfaring viser at løsningen av differensiallikninger er lang mer krevende enn løsningen av algebraiske likninger. Ved å benytte Laplacetransformasjonen viser det seg at en modell for sammenhengen mellom posisjon og påtrykt kraft kan skrives som¹

$$x_1(s) = \frac{\frac{1}{m}}{s^2 + \frac{f}{m}s + \frac{k}{m}} u(s) \quad (4.2)$$

hvor x_1 og u er Laplacetransformerte funksjoner. Dette er markert ved at de er funksjoner av Laplace-variabelen s i stedet for den tidsvariable t .

Vi kan observere at den Laplacetransformerte modellen er algebraisk. Denne typen modeller viser seg å være meget hensiktsmessige for å analysere lineære dynamiske systemer, og for å konstruere regulatorer.

■

4.2 DEFINISJONER

Laplacetransformasjonen er definert som

$$F(s) = \int_{-\infty}^{\infty} f(t)e^{-st} dt \quad (4.3)$$

Vi skriver også dette som

$$\mathcal{L}(f(t)) = F(s) \quad (4.4)$$

eller

$$f(t) \xleftrightarrow{\mathcal{L}} F(s) \quad (4.5)$$

$f(\cdot)$ og $F(\cdot)$ kalles et *Laplacetransform-par*. Den siste formuleringen indikerer også at det finnes en *invers Laplacetransformasjon*

$$f(t) = \mathcal{L}^{-1}(F(s)) \quad (4.6)$$

som vi kommer tilbake til.

¹Dette gjelder når systemet har initialverdier $x(0) = 0$, og $\dot{x}(0) = 0$. Initialverdier $\neq 0$ kan også håndteres. Mer om det på de etterfølgende sider.

(Uttrykket til venstre i (4.3) kan leses “Laplacetransformen til $f(t)$ ”, eller “Laplacetransformasjonen av $f(t)$ ”. Begrepene “transform” og “transformasjon” har noe forskjellig betydning siden “transform” er betegnelsen på resultatet av en transformasjon.)

Vi vil fra nå av ikke benytte stor bokstav for Laplacetransformer. For å “spare” på symbolene vil denne boka i de fleste tilfeller bruke samme symbol for å betegne både en tidsfunksjon og dens Laplacetransform, slik at $f(\cdot)$ kan symbolisere både $f(t)$ og $f(s)$. Det vil framgå av sammenhengen hva som menes.

Alle tidsforløp som vi betrakter i denne boka, vil ha $f(t) = 0$ for $t < 0$. Dermed vil vi alltid ha

$$f(s) = \int_0^{\infty} f(t)e^{-st} dt \quad (4.7)$$

(4.7) og (4.3) kalles ofte den ensidige henholdsvis den tosidige Laplacetransformasjon. Vi skal som sagt bare benytte (4.7).

Betingelsen for at en ensidig Laplacetransform, $f(s)$, skal eksistere, er at

$$\int_0^{\infty} f(t)e^{-\sigma_a t} dt < \infty \quad (4.8)$$

der σ_a er en reell størrelse. Den minste verdien av σ_a som tilfredsstiller ulikheten i (4.8), angir det som kalles *konvergensaksen til $f(t)$* . De aller fleste funksjoner som kommer på tale i reguleringsteknikken, er Laplacetransformerbare.

Mens t i $f(t)$ er en reell variabel, så er s i $f(s)$ en kompleks variabel; $s = \sigma + j\omega$. Vi snakker i den forbindelse om “ s -planet”, og tilsvarende om “tidsplanet” (selv om dette siste er språklig litt misvisende, siden t alltid er reell). Konvergensaksen ligger i s -planet parallelt med den imaginære akse, med $\sigma = \sigma_a$.

Den **inverse Laplacetransformasjon** gjør det mulig å transformere en funksjon fra s -planet tilbake til tidsplanet:

$$\mathcal{L}^{-1}(f(s)) = f(t) = \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} f(s)e^{ts} ds \quad (4.9)$$

Uttrykket til venstre for likhetstegnet i (4.9) leses: “Den inverse Laplacetransformen til $f(s)$ ”. Den dannes ved en integrasjon i det komplekse planet (s -planet) langs en rett linje parallell med den imaginære akse i en avstand σ .

σ kan være en vilkårlig konstant, bare den oppfyller $\sigma > \sigma_a$. Dette er alltid mulig for de tilfellene vi betrakter.

Integralet i (4.9) løses enkelt ved **residuregning**, som følger av **Cauchys integralsats** fra kompleks funksjonsteori. Dette behandles i avsnitt 4.4.1.

4.2.1 Vanlig forekommende Laplacetransformer. Regler og teoremer

I dette avsnittet presenterer vi noen regler for Laplacetransformasjon av tidsfunksjoner:

1. **Linearitet:**

$$\mathcal{L}(af(t)) = af(s) \quad (4.10)$$

a er her en konstant

$$\mathcal{L}(f_1(t) \pm f_2(t)) = f_1(s) \pm f_2(s) \quad (4.11)$$

2. **Derivasjon:**

$$\mathcal{L}(\dot{f}(t)) = sf(s) - f(t=0) \quad (4.12)$$

og

$$\mathcal{L}(\ddot{f}(t)) = s^2 f(s) - sf(t=0) - \dot{f}(t=0) \quad (4.13)$$

osv.

3. **Integrasjon:**

$$\mathcal{L}\left(\int_0^t f(t)dt\right) = \frac{1}{s}f(s) \quad (4.14)$$

og

$$\mathcal{L}\left(\int_{-\infty}^t f(t)dt\right) = \frac{1}{s}f(s) + \frac{1}{s} \int_{-\infty}^0 f(t)dt \quad (4.15)$$

4. **Tidsmålestokk:**

$$\mathcal{L}\left(f\left(\frac{t}{a}\right)\right) = af(as) \quad (4.16)$$

5. Reell translasjon (tidsforsinkelse):

$$\mathcal{L}(f(t - \tau)\mu(t - \tau)) = e^{-\tau s} f(s), \quad \tau = \text{konstant, og}$$

$$\mu(t - \tau) = \begin{cases} 0 & \text{for } t < \tau \\ 1 & \text{for } t > \tau \end{cases} \quad (4.17)$$

6. Kompleks translasjon:

$$\mathcal{L}(e^{-at} f(t)) = f(s + a) \quad (4.18)$$

der a er en konstant

7. Lineær tidsveiging:

$$\mathcal{L}(tf(t)) = -\frac{df(s)}{ds} \quad (4.19)$$

8. Reell multiplikasjon (kompleks folding):

$$\mathcal{L}(f_1(t)f_2(t)) = \frac{1}{2\pi j} \int_{\gamma_2 - j\infty}^{\gamma_2 + j\infty} f_1(s - w)f_2(w)dw \quad (4.20)$$

9. Kompleks multiplikasjon (reell folding):

$$\mathcal{L}\left(\int_0^t f_1(t - \tau)f_2(\tau)d\tau\right) = f_1(s)f_2(s) \quad (4.21)$$

10. Sluttverdi:

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sf(s) \quad (4.22)$$

når $sf(s)$ er analytisk på og til høyre for den imaginære akse. Dette svarer til at grensen til $f(t)$ eksisterer når $t \rightarrow \infty$, noe som ikke alltid er tilfelle.

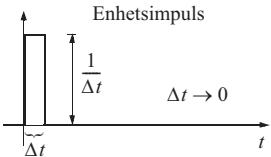
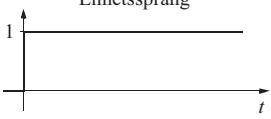
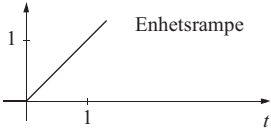
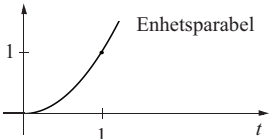
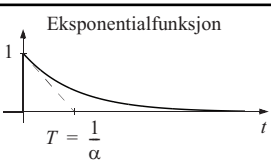
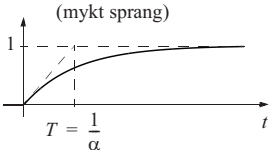
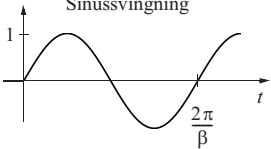
11. Begynnelsesverdi:

$$\lim_{t \rightarrow 0} f(t) = \lim_{s \rightarrow \infty} sf(s) \quad (4.23)$$

I appendiks B finnes en omfattende liste over Laplacetransform-par.

Laplacetransformene av noen enkle og ofte forekommende tidsforløp er framstilt i tabell 4.1.

TABELL 4.1 $f(s)$ for noen vanlig forekommende tidsforløp

Tidsforløp	$f(t)$	$\mathcal{L}(f(t)) = f(s)$
 <p>Enhetsimpuls</p>	$\delta(t) = \lim_{\Delta t \rightarrow 0} g(t, \Delta t)$ $g(\cdot) = \begin{cases} \frac{1}{\Delta t} & , \quad 0 < t < \Delta t \\ 0 & , \quad t < 0, t > \Delta t \end{cases}$	1
 <p>Enhetsprang</p>	$\mu_1(t) = \begin{cases} 1 & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$	$\frac{1}{s}$
 <p>Enhetsrampe</p>	$\mu_2(t) = \begin{cases} t & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$	$\frac{1}{s^2}$
 <p>Enhetsparabel</p>	$\mu_3(t) = \begin{cases} t^2 & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$	$\frac{2}{s^3}$
 <p>Ekspontialfunksjon</p>	$\begin{cases} e^{-\alpha t} & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$	$\frac{1}{s + \alpha} = \frac{T}{1 + Ts}$
 <p>Ekspontielt sprang (mykt sprang)</p>	$\begin{cases} 1 - e^{-\alpha t} & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$	$\frac{\alpha}{s(s + \alpha)} = \frac{1}{s(1 + Ts)}$
 <p>Sinussvingning</p>	$\begin{cases} \sin \beta t & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases}$	$\frac{\beta}{s^2 + \beta^2} = \frac{\beta}{(s + j\beta)(s - j\beta)}$

4.3 LAPLACETRANSFORMASJON AV DIFFERENSIALLIKNINGER

4.3.1 1. ordens lineær differensiallikning

Ved hjelp av transformasjonsreglene som er gitt for linearitet og derivasjon i (4.10)-(4.13), kan vi nå Laplacetransformere en differensiallikning av formen

$$\dot{x}(t) = ax(t) + bu(t) \quad (4.24)$$

Vi får da

$$sx(s) - x(t=0) = ax(s) + bu(s) \quad (4.25)$$

som løst med hensyn på $x(s)$ gir

$$x(s) = \frac{1}{s-a}x_0 + \frac{b}{s-a}u(s) \quad (4.26)$$

der vi for enkelhets skyld skriver x_0 for $x(t=0)$.

EKSEMPEL 4.2: Sprangrespons for (4.24)

Anta nå at $u(t)$ er et enhetsprang $\mu_1(t)$.

Tabell 4.1 gir $\mathcal{L}(\mu_1(t)) = \frac{1}{s}$, og (4.26) gir

$$x(s) = \frac{1}{s-a}x_0 + \frac{b}{s(s-a)} \quad (4.27)$$

Dette kan inverstransformeres ledd for ledd, for eksempel ved residuegning som beskrevet i avsnitt 4.4.1. Men her bruker vi linje 7 og 8 i tabellen i appendiks B. Dette gir løsningen

$$x(t) = e^{at}x(0) + \frac{b}{a}(e^{at} - 1) \quad (4.28)$$

Hvis vi setter $a = -1/T$ og $b = 1/T$, svarer (4.24) til differensiallikningen for den enkle RC-kretsen i eksempel 2.2. Likning (4.28) med $x(0) = 0$ svarer da til (2.22).



Vender vi igjen tilbake til situasjonen med en generell $u(t)$, så kan vi sammenholde ledd for ledd (4.26) og (2.17) med $t_0 = 0$:

$$\begin{aligned}x(s) &= \frac{1}{s-a}x_0 + \frac{b}{s-a}u(s) \\x(t) &= e^{at}x_0 + \int_0^t e^{a(t-\tau)}bu(\tau)d\tau\end{aligned}\tag{4.29}$$

Av dette bekreftes at

$$e^{at} \xleftrightarrow{\mathcal{L}} \frac{1}{s-a}$$

som vi fant i tabellen i appendiks B. Videre har vi at integralet til høyre i (4.29) er et foldingsintegral. Dette gir, fordi impulsresponsen $h(t) = e^{at}b$ (jfr. (2.33)),

$$h(t) * u(t) \xleftrightarrow{\mathcal{L}} h(s)u(s)$$

Det vil si “folding i tidsplanet svarer til multiplikasjon i s -planet”; vi kjenner igjen regel (4.21). $\mathcal{L}[h(t)] = h(s)$, kaller vi *transferfunksjonen*. Mer om den i avsnitt 4.5.

Vi har i dette tilfellet brukt tabell for å finne den inverse Laplacetransform. I avsnitt 4.4 skal vi lære å bruke residuregning til dette.

4.3.2 Høyere ordens lineær differensiallikning

En høyere ordens differensiallikning med én avhengig variabel og en ytre påvirkning u (inhomogen likning) kan i følge (2.73) skrives

$$g(x, \dot{x}, \ddot{x}, \dots, x^{(n)}, u, t) = 0\tag{4.30}$$

I tilfelle likningen er lineær og har konstante koeffisienter vil vi kunne skrive

$$\alpha_0 x(t) + \alpha_1 \dot{x}(t) + \alpha_2 \ddot{x}(t) + \dots + \alpha_n x^{(n)}(t) + \beta u(t) = 0\tag{4.31}$$

Laplacetransformasjon av (4.31) gir, i det vi benytter reglene (4.10) – (4.13):

$$\begin{aligned}\alpha_0 x(s) + \alpha_1 [sx(s) - x_0] + \alpha_2 [s^2x(s) - sx_0 - \dot{x}_0] + \dots \\ \dots + \alpha_n [s^n x(s) - s^{n-1}x_0 - \dots - x_0^{(n-1)}] + \beta u(s) = 0\end{aligned}\tag{4.32}$$

Løses (4.32) med hensyn på $x(s)$, finner vi

$$x(s) = \frac{1}{\alpha_0 + \alpha_1 s + \dots + \alpha_n s^n} (Q(s) - \beta u(s)) \quad (4.33)$$

der $Q(s)$ er et polynom i s av grad $n - 1$.

Likning (4.26) er et spesialtilfelle av (4.33) med $n = 1$, $\alpha_1 = 1$, $\alpha_0 = -a$ og $\beta = -b$.

Hadde alle begynnelsesbetingelsene som inngår i $Q(s)$ vært lik null, ville (4.33) blitt særlig enkel, for da er $Q(s) = 0$. Videre kan vi i (4.31) dividere med α_n på begge sider av likhetstegnet, slik at koeffisienten foran leddet med s^n blir 1. Da kan vi skrive

$$x(s) = h(s)u(s), \text{ der } h(s) = \frac{-\frac{\beta}{\alpha_n}}{(s-s_1)(s-s_2)\dots(s-s_n)} \quad (4.34)$$

s_1, \dots, s_n er røttene i nevnerpolynomet. Uttrykk av typen (4.34) vil vi benytte seinere.

4.4 INVERS LAPLACETRANSFORMASJON VED RESIDUREGNING

4.4.1 Residuregning

Det skal vises hvordan man kan finne den inverse Laplacetransform for funksjoner som er kontinuerlige i tidsplanet (tilsvarende metode kan brukes for tidsdiskrete funksjoner; kapittel 11).

Den **inverse Laplacetransformasjon** $x(t)$ av en funksjon $x(s)$ er i følge (4.9)

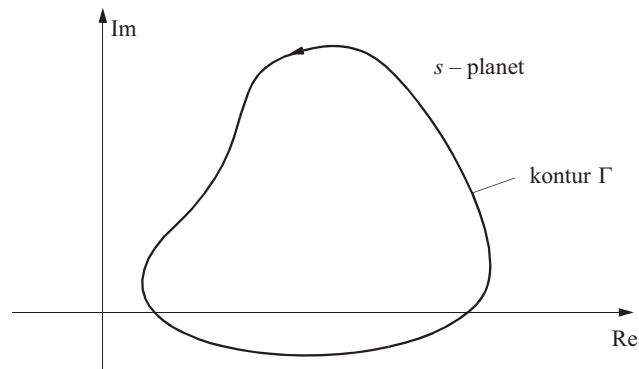
$$x(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} x(s)e^{ts} ds \quad (4.35)$$

Integralet (4.35) er vanskelig å løse direkte. Vi skal nå forklare hvordan man kan benytte **residuregning** for å løse (4.35). Denne metoden baserer seg på resultater fra kompleks funksjonsteori. Der dreier det seg om å finne integralet av en funksjon $f(s)$ i en kompleks variabel s , når integrasjonsbanen Γ er en lukket kontur *mot* urviseren i det komplekse plan. Et eksempel er vist i figur 4.1.

Det kan da vises at hvis $f(s)$ er **analytisk** innenfor konturen (“analytisk” $\Leftrightarrow f(s)$ og alle deriverte av $f(s)$, eksisterer overalt innenfor konturen), så er konturintegralet

$$\oint_{\Gamma} f(s) ds = 0 \quad (4.36)$$

Figur 4.1
En integrasjons-
bane Γ



Hvis $f(s)$ derimot har en **pol** av orden m innenfor konturen Γ , slik at $f(s)$ kan skrives

$$f(s) = \frac{g(s)}{(s-a)^m}, \text{ der } g(s) \text{ ikke har pol(er) innenfor konturen,} \quad (4.37)$$

sier den såkalte **residusatsen** at

$$\frac{1}{2\pi j} \oint_{\Gamma} f(s) ds = \operatorname{res}_{s=a}[f(s)] = \frac{1}{(m-1)!} \left[\frac{d^{m-1}}{ds^{m-1}} [(s-a)^m f(s)] \right]_{s=a} \quad (4.38)$$

der $\operatorname{res}_{s=a}[f(s)]$ uttales “residuet til $f(s)$ i a ”¹.

For $m = 1$ blir (4.38) atskillig enklere

$$\frac{1}{2\pi j} \oint_{\Gamma} f(s) ds = \operatorname{res}_{s=a}[f(s)] = [(s-a)f(s)]_{s=a} \quad (4.39)$$

Hvis det er flere distinkte poler innenfor konturen, har vi

$$\frac{1}{2\pi j} \oint_{\Gamma} f(s) ds = \sum_i \operatorname{res}_{s=a_i}(f(s)) \quad (4.40)$$

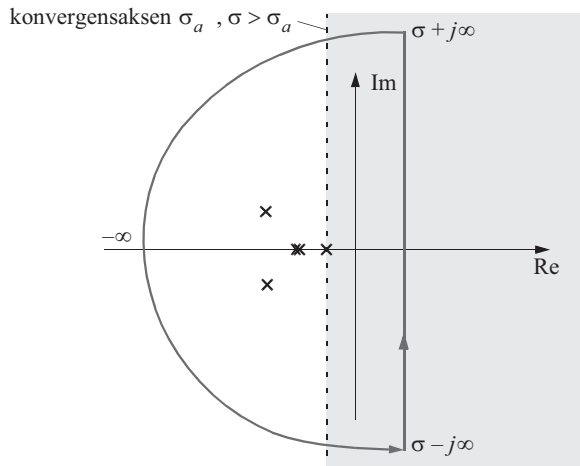
I (4.40) anvender vi (4.39) eller (4.38) avhengig av om polen a_i er enkel eller av høyere orden (dvs. **multippel pol**; flere identiske poler).

¹I det generelle tilfellet skal høyresida i (4.38) være $\lim_{s \rightarrow a} \frac{1}{(m-1)!} \left(\frac{d^{m-1}}{ds^{m-1}} [(s-a)^m f(s)] \right)$, men for slike funksjoner som er aktuelle for oss, trenger vi ikke foreta noen slik grenseovergang.

4.4.2 Invers Laplacetransformasjon

Hvordan kan så dette brukes for å finne en invers Laplacetransform? Venstresiden i (4.40) likner på (4.35), med $f(s) = x(s)e^{ts}$. Problemet er integrasjonsbanen til (4.35), som ikke er lukket. Vi må derfor bruke et knep for å lukke integrasjonsbanen. Betrakt figur (4.2):

Figur 4.2
Lukket
integrasjonsbane



Gitt integranden $x(s)e^{ts}$ i (4.35). Et sett poler i denne integranden er antydnet i figur 4.2. Som det framgår av figur 4.2 har vi modifisert integrasjonsbanen. Vi har føyd til en uendelig stor halvsirkel inn i venstre halvplan for å få en lukket kontur.

Dette kan vi tillate oss uten å endre verdien av integralet (4.35), så lenge $|x(s)e^{ts}| \rightarrow 0$ når $|s| \rightarrow \infty$.

Hvis integranden er så liten langs halvsirkelen at dette oppveier den uendelige lengden på denne banen, gir integrasjonen langs halvsirkelen intet bidrag til integralet. Hvis $x(s)$ er et rasjonalt uttrykk, så er dette oppfylt når graden av nevneren er minst *en* høyere enn graden av telleren. (Dette er ikke helt trivielt og må bevises. Det gjøres ikke her.) Vi kaller da $x(s)$ **strengt propert**. $x(s)$ vil alltid være strengt propert for de kontinuerlige signalene og systemene vi skal arbeide med. Vi kan altså konkludere med at den lodrette integrasjonsbanen i (4.35) uten videre kan lukkes med en uendelig stor halvsirkel, uten at dette endrer verdien av integralet. *Dermed kan vi bruke residusatsen ved invers Laplacetransformasjon.*

Nå er vi i stand til å angi den endelige formelen for invers Laplacetransformasjon ved hjelp av residuegning:

$$x(t) = \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} x(s)e^{ts} ds = \sum_i \operatorname{res}_{s=a_i}(x(s)e^{ts}) \quad (4.41)$$

der hvert residu i summen til høyre svarer til (4.38) eller (4.39) avhengig av om a_i er multiplert eller enkel.

EKSEMPEL 4.3: Residuregning på en RC-krets

Gitt RC-kretsen i eksempel 2.2. Anta nå at den påsatte spenningen, som er vårt pådrag $u(t)$, er

$$u(t) = 1 - e^{-\frac{t}{T}} \quad (4.42)$$

det vil si at pådraget settes på som et “mykt sprang” med samme form som i figur 2.7. Hva blir nå utgangsspenningen? For enkelhets skyld setter vi spenningen over kondensatoren $x(t) = 0$ for $t = 0$. Likning (4.29) med $a = -1/T$ og $b = 1/T$, og (4.42) gir da:

$$x(s) = \frac{\frac{1}{T}}{s + \frac{1}{T}} \cdot \mathcal{L}\left(1 - e^{-\frac{t}{T}}\right) = \frac{1}{1 + Ts} \cdot \frac{1}{s(1 + Ts)} = \frac{\frac{1}{T^2}}{s\left(s + \frac{1}{T}\right)^2} \quad (4.43)$$

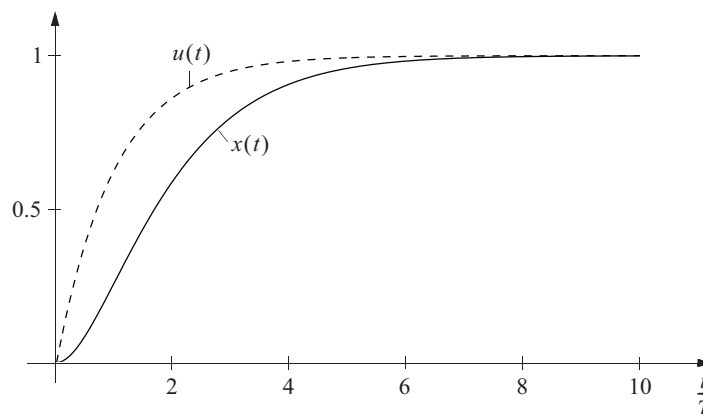
Polen $s = 0$ har multiplisitet $m = 1$ og polen $s = -1/T$ har multiplisitet $m = 2$. (4.39) og (4.38) med $f(s) = x(s)e^{ts}$ gir da

$$\begin{aligned} x(t) &= \mathcal{L}^{-1}(x(s)) = \operatorname{res}_{s=0} \left(\frac{e^{ts}}{T^2 s \left(s + \frac{1}{T}\right)^2} \right) + \operatorname{res}_{s=-\frac{1}{T}} \left(\frac{e^{ts}}{T^2 s \left(s + \frac{1}{T}\right)^2} \right) \\ &= \left(\frac{e^{ts}}{T^2 \left(s + \frac{1}{T}\right)^2} \right) \Bigg|_{s=0} + \left(\frac{d}{ds} \left(\frac{e^{ts}}{T^2 s} \right) \right) \Bigg|_{s=-\frac{1}{T}} \\ &= 1 - \frac{e^{ts}}{T^2 s^2} \Bigg|_{s=-\frac{1}{T}} + \frac{t e^{ts}}{T^2 s} \Bigg|_{s=-\frac{1}{T}} = 1 - \left(1 + \frac{t}{T}\right) e^{-\frac{t}{T}} \end{aligned} \quad (4.44)$$

Figur 4.3 viser responsen.¹

¹Merk at vi i løsningen (4.44) nå får med ledd av typen ... + konst. $\cdot te^{at}$ + Dette skyldes at vi har to sammenfallende poler i Laplacetransformen. Se ellers appendiks B, linje 42 for tilfellet med enda flere sammenfallende poler.

Figur 4.3
 Respons i 1.
 ordens system på
 mykt sprang



4.5 TRANSFERFUNKSJONER OG BLOKKDIAGRAMMER

4.5.1 Transferfunksjoner

Når alle begynnelsesbetingelser er lik null, er Laplacetransformen til et lineært systems respons entydig gitt av inngangssignalet Laplacetransform og funksjonen $h(s)$, som vi kaller systemets **transferfunksjon**. Responsen gis da som

$$y(s) = h(s)u(s) \quad (4.45)$$

Ordet “transfer” er egentlig engelsk, men er også blitt norsk terminologi. “Transfer” betyr ordrett “overføre”, vi kan si at $h(s)$ forteller hvordan $u(s)$ blir “overført” til $y(s)$. Vi bruker også (sjeldnere) **overføringsfunksjon**.

Ofte skriver man en transferfunksjon som en brøk mellom utgang og inngang, med s i “felles parentes”,

$$h(s) = \frac{y(s)}{u(s)} = \frac{y}{u}(s) \quad (4.46)$$

Hvis inngangssignalet er en impuls (deltafunksjon), blir utgangssignalet impulsresponsen, vi får $y(t) = h(t)$.

I s -planet har vi $u(s) = \mathcal{L}[\delta(t)] = 1$. Vi får $y(s) = h(s) \cdot 1 = h(s)$. Derfor:

Transferfunksjonen $h(s)$ inneholder akkurat samme informasjon om et systems egenskaper som dens inverse Laplacetransform, impulsresponsen $h(t)$.

Men transferfunksjonen er mye enklere å bruke, siden folding i tidsplanet blir til multiplikasjon i s -planet, jfr. (4.21):

$$h(t) * u(t) \xleftrightarrow{\mathcal{L}} h(s)u(s)$$

Transferfunksjoner er ytterst sentrale i reguleringsteknikken, og vil fra nå av bli brukt svært mye.

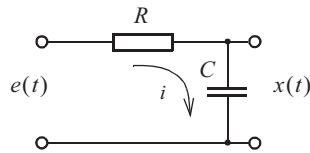
Vi skal seinere ta for oss den vektorielle generalisering av transferfunksjon, når systemet er multivariabelt. Da har vi $\mathbf{y}(s) = \mathbf{H}(s)\mathbf{u}(s)$, hvor $\mathbf{H}(s)$ er systemets **transfermatrise** (avsnitt 4.6.1).

EKSEMPEL 4.4: Impedans

I elektrisitetstæren bruker man begrepet **impedans**, gjerne symbolisert med stor bokstav Z . Dette er en generalisering av ohmsk motstand R . Ohms lov for sammenhengen mellom strøm og spenning uttrykt i s -planet, som vi skal benytte for å definere impedans, er $e(s) = Ri(s)$. Impedansen er en transferfunksjon, $Z(s)$, slik at

$$e(s) = Z(s)i(s), \text{ dvs. } Z(s) = \frac{e(s)}{i(s)} \quad (4.47)$$

Figur 4.4
RC-krets



Vi skal finne impedansen til RC-kretsen fra eksempel 4.3, se figur 4.4. Vi har tidligere funnet transferfunksjonen fra inngangsspenningen til utgangsspenningen. Fra (4.43), midtre del, har vi at den er

$$\frac{x}{e}(s) = \frac{1}{1 + Ts} \quad (4.48)$$

Vi søker nå impedansen, som er en annen transferfunksjon,

$$Z(s) = \frac{e}{i}(s) = \frac{e}{x}(s) \cdot \frac{x}{i}(s) = \frac{1 + Ts}{1} \cdot \frac{1}{sC} = \frac{1}{sC} + R \quad (4.49)$$

Vi har her benyttet at sammenhengen mellom spenning og strøm for en kondensator er

$$i = C \frac{dx}{dt}, \text{ som Laplacetransformert blir } i = sCx, \Rightarrow Z_C(s) = \frac{1}{sC} \quad (4.50)$$

Vi har også benyttet at $T = RC$.

Summen i høyre ledd i (4.49) uttrykker en seriekopling av to impedanser, én for motstanden og én for kondensatoren. Vi kan regne ut total impedans basert på serie- og parallellkoplinger av del-impedanser, etter samme regler som for ohmske motstander.

Vi vil til slutt nevne at impedansen for en *induktans* blir

$$Z_L(s) = sL, \text{ fordi } x = L \frac{di}{dt} \quad (4.51)$$

Se også eksempel 2.7, hvor du nå kan finne transferfunksjonen fra f.eks. e_1 til e_3 , ved å benytte impedanser.

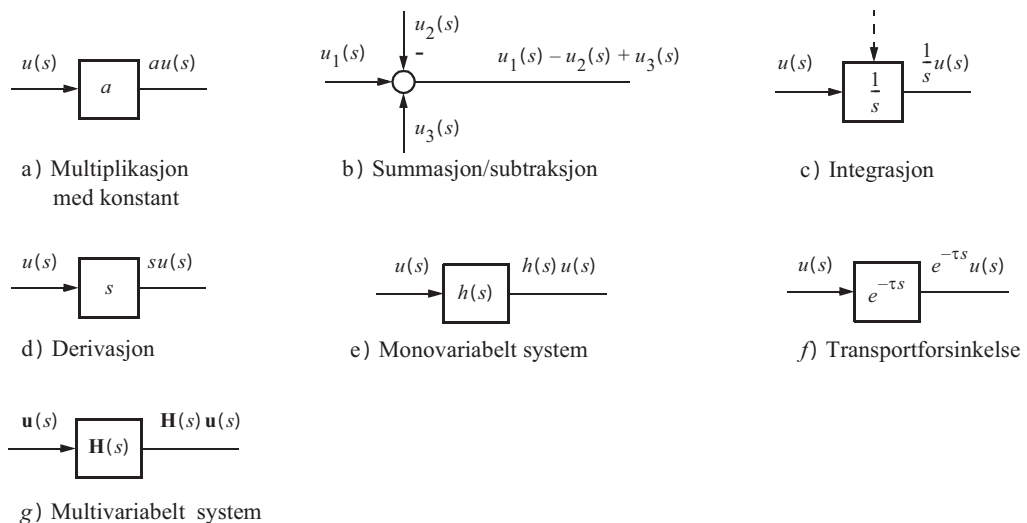


4.5.2 Blokkdiagrammer

Transferfunksjoner brukes ofte sammen med **blokkdiagrammer**. Avsnitt 2.8 omhandlet blokkdiagrammer av systemer i *tidsplanet*. Vi skal i nå ta for oss blokkdiagrammer i *s-planet*. De symboler som brukes i slike diagrammer er vist i figur 4.5. I såkalte **elementære blokkdiagrammer** brukes bare de fire første symbolene, (a) - (d). Siden Laplacetransformasjonen bare gjelder for lineære systemer, er det ikke aktuelt med symboler for ulineære operasjoner.

Integrasjon og derivasjon symboliseres med blokkene $1/s$ og s som vist i figur 4.5 c) og d). En stiplet pil inn på integratorblokken i figur 4.5 c) indikerer begynnelsesverdien til utgangen på integratoren (men svært ofte vil vi anta at den er lik 0). Figur 4.5 e) viser et blokkdiagram for et vilkårlig monovariabelt system. Dette blokkdiagrammet er ekvivalent med systemets transferfunksjon. Figur 4.5 g) viser det vektorielle tilfellet, med transfermatrise $\mathbf{H}(s)$.

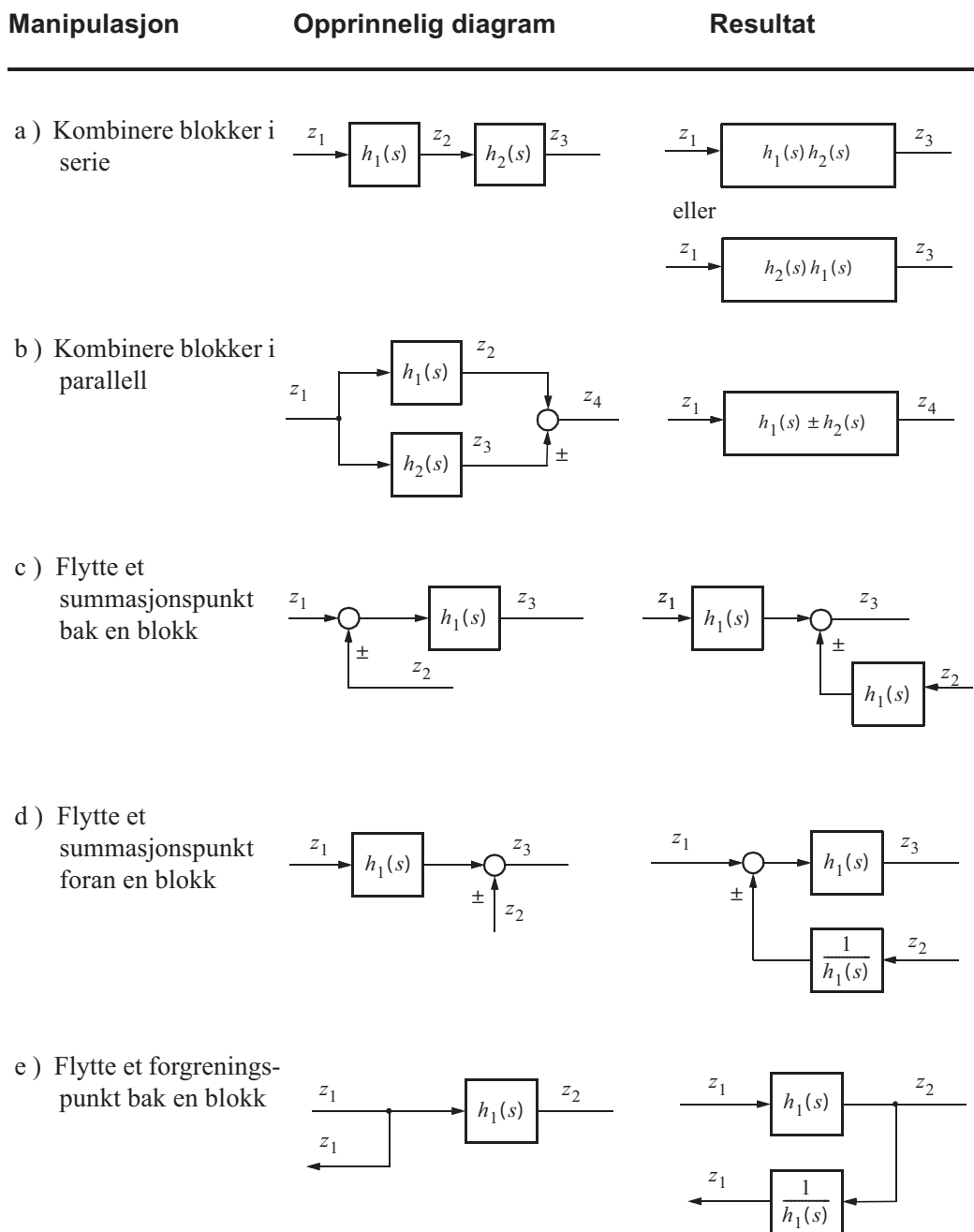
Figur 4.5
Blokkdiagrammer
med Laplace-
transformer

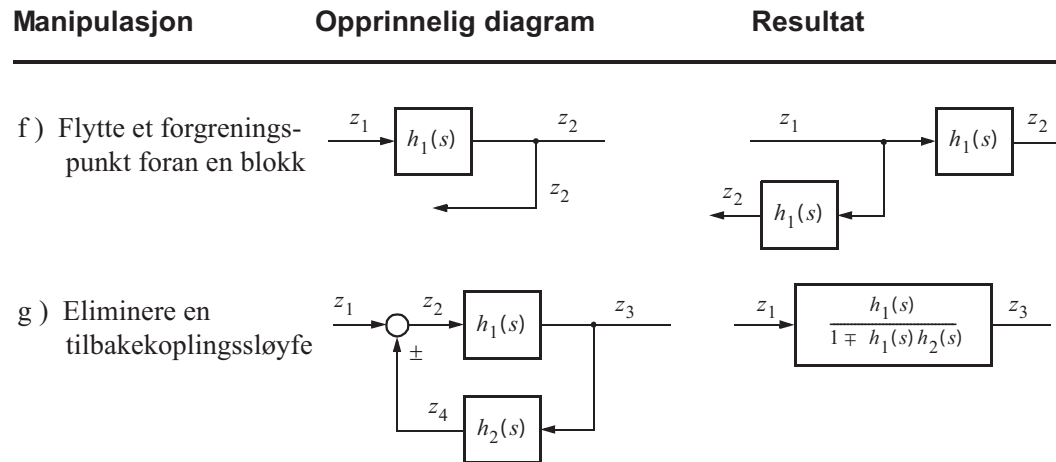


Med bakgrunn i teorien for Laplacetransformasjon og de symbolene som er innført i figur 4.5, kan vi utvikle et sett av regler for manipulasjon og **reduksjon av blokkdiagrammer** som inneholder transferfunksjoner i blokkene. Målet med en reduksjonsprosedyre er å forenkle systemet til én eller noen få blokker.

Figur 4.6 a) og b) viser at seriekopling av transferfunksjoner medfører multiplikasjon av transferfunksjonene, og parallellkopling av transferfunksjoner medfører summasjon av transferfunksjonene. Det er dette som utgjør Laplacetransformasjonens store fordel; den gjør differensiallikninger om til algebraiske likninger.

Figur 4.6
Manipulering av
blokkdiagrammer





Vær oppmerksom på at reglene b)-g) i figur 4.6 også gjelder dersom transferfunksjonen $h(s)$ erstattes med en transfermatrise $\mathbf{H}(s)$. Regel a) gjelder også for $\mathbf{H}(s)$ dersom det nederste ekvivalentskjemaet benyttes.

Reglene i figur 4.6 skal illustreres med noen eksempler.

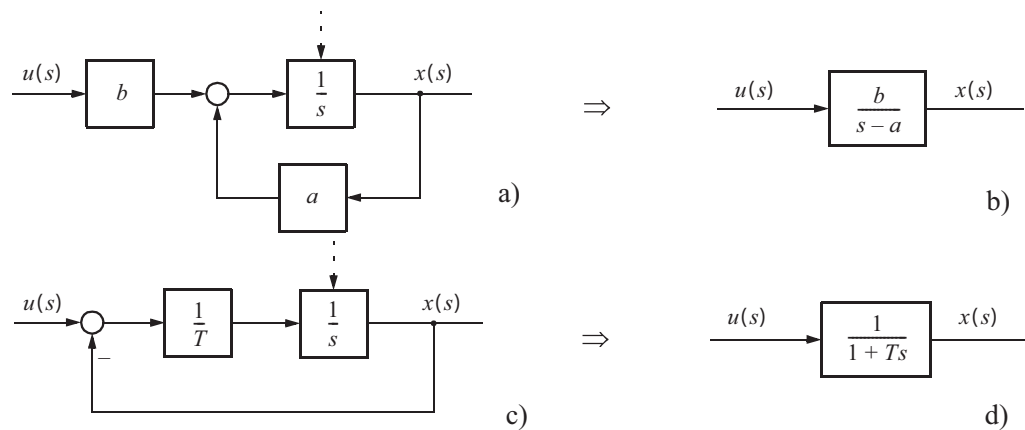
EKSEMPEL 4.5: Blokkdiagram for 1. ordens lineær differensiallikning

Et system beskrevet av en 1. ordens lineær differensiallikning

$$\dot{x} = ax + bu$$

vil ha et elementært blokkdiagram som vist i figur 4.7 a).

Figur 4.7
Elementært blokk-
diagram for
1. ordens
differensiallikning



Bruk av regel a) og g) i figur 4.6 gir diagrammet i figur 4.7 b).

Det enkle RC-nettverket i figur 2.5 som er beslektet med systemet i a) og b), har differensiallikningen

$$T\dot{x} = -x + u$$

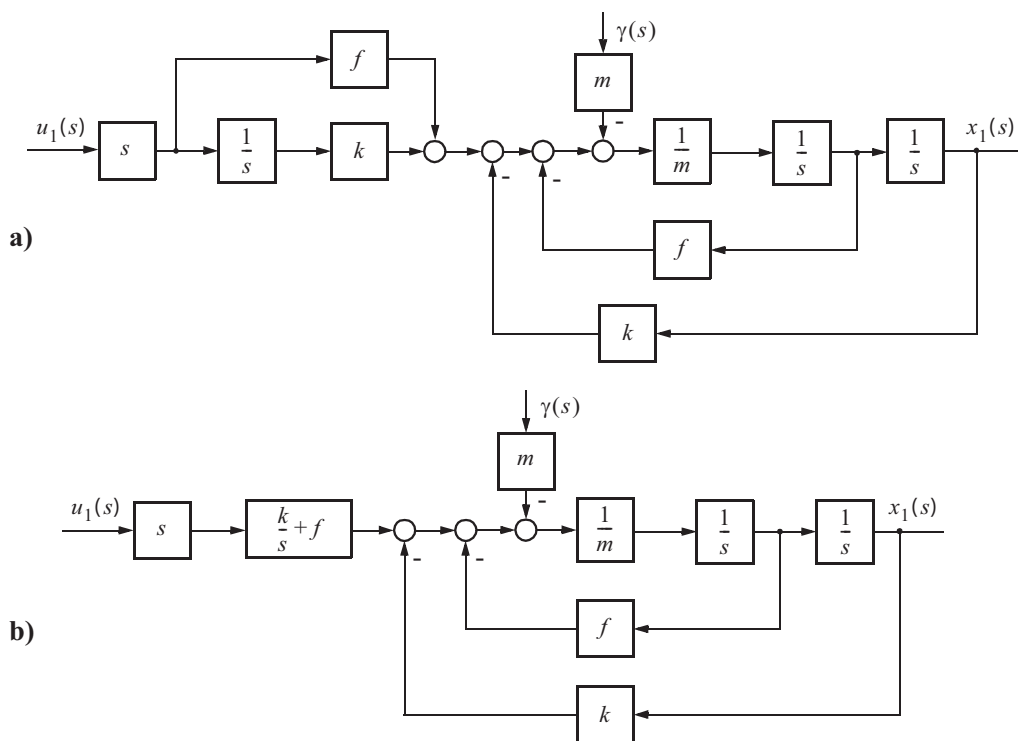
Det elementære blokkdiagrammet for dette systemet er vist i figur 4.7 c). Reduksjon ved hjelp av regel g) i figur 4.6 gir transferfunksjonen i figur 4.7 d).¹



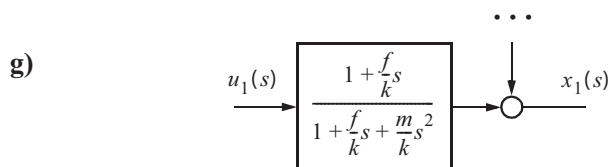
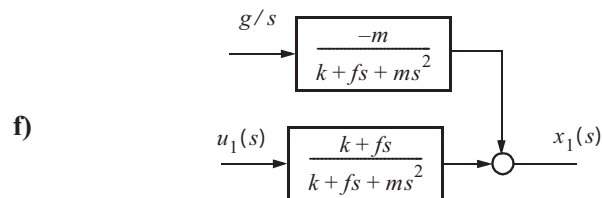
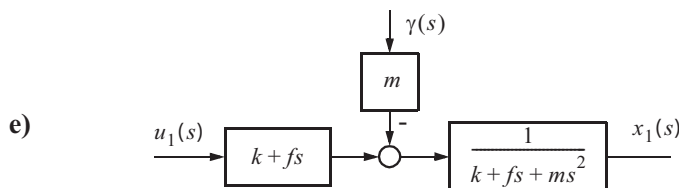
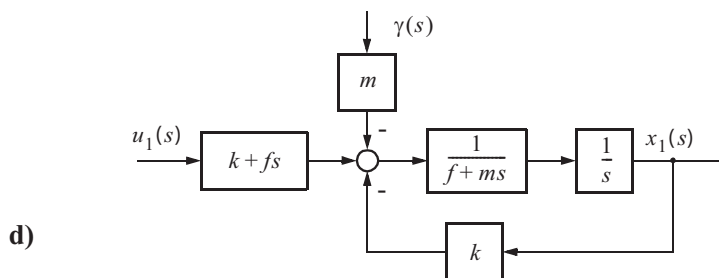
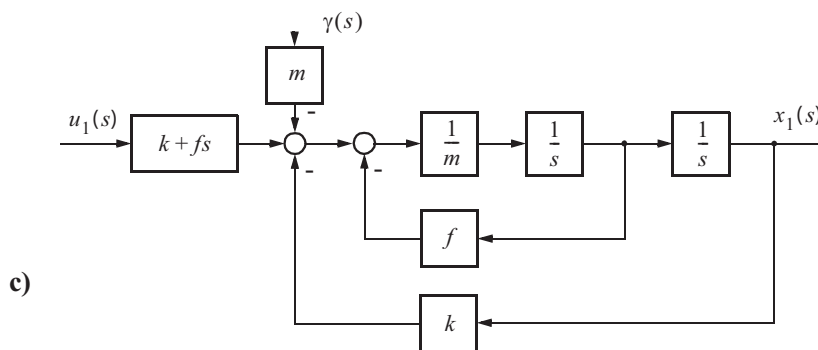
EKSEMPEL 4.6: Reduksjon av blokkdiagrammet til et fjærsystem for bil

Systemet med masse og støddemper fra eksempel 2.6 kan nå representeres som vist i figur 4.8 a). Ved hjelp av reglene i figur 4.6 kan så blokkdiagrammet reduseres suksessivt. Vi skal illustrere hvordan:

Figur 4.8
Reduksjon av
blokkdiagram



¹De stiplede linjene som angir begynnelsesverdien av integratorutgangen i figur 4.5 c), 4.7 a) og 4.7 c) vil som oftest bli sløffet for enkelhets skyld når man arbeider med blokkdiagrammer. De er da underforstått.



Tyngdens akselerasjon er innledningsvis antatt å være en tidsvarierende funksjon $\gamma(t)$ med Laplacetransform $\gamma(s)$, et knep som vi forklarer nedenfor.

De suksessive reduksjonene foregår slik: I overgangen mellom **a)** og **b)** er regel b) og d) fra figur 4.6 benyttet. I overgangen mellom **b)** og **c)** er regel a) fra figur 4.6 benyttet. I tillegg er tyngdens akselerasjon flyttet to summasjonspunkter lenger mot venstre. Dette endrer ikke påvirkningen fra forstyrrelsen på de to sløyfene som kommer etter. Regel g) er benyttet både i overgangen mellom **c)** og **d)** og i overgangen mellom **d)** og **e)**. Mellom **e)** og **f)** er regel a) og c) benyttet.

Videre er innvirkningen av tyngdens akselerasjon, g , nå spesifisert. Siden $g = \text{konstant} > 0$ for alle t (også $t < 0$), vil den ikke uten videre kunne Laplacetransformeres. Vi tenkte oss innledningsvis at tyngdens akselerasjon var en eller annen tidsvarierende akselerasjon som startet ved $t = 0$, med Laplacetransform $\gamma(s)$, som vist t.o.m. figur 4.8 **e)**. Siden den virkelige $g = \text{konstant}$, kan vi finne innvirkningen ved å anta at den settes på som et sprang ved $t = 0$ som vist i figur 4.8 **f)**, og finne den stasjonære responsen ved hjelp av sluttverditeoremet (se (4.22)). Mer “fysisk” kan man forestille seg at dette skjer ved at massen holdes oppe ved starttilstanden, så høyt at fjæra er akkurat ubelastet. Så slippes taket, og massen vil etter en stund innta posisjonen

$$\lim_{t \rightarrow \infty} \mathcal{L}^{-1} \left\{ \frac{-m}{k + fs + ms^2} \gamma(s) \right\} = \lim_{s \rightarrow 0} \left\{ s \left(\frac{-m}{k + fs + ms^2} \frac{g}{s} \right) \right\} = -\frac{gm}{k} \quad (4.52)$$

Men siden tyngdekrafta i virkeligheten jo ble “slått på” ved $t = -\infty$, vil dens virkning for *alle* t være en konstant forskyving av posisjonen $x_1(t)$ med $-gm/k$. Vi ignorerer derfor først den øvre greina i figur 4.8 **f)**, og regner ut x_1 som om den bare var bestemt av u_1 , for så etterpå å korrigere x_1 med $-gm/k$. Vi bruker altså superposisjonsprinsippet.

Sluttverditeoremet sier oss at resultatet av en konstant påvirkning på et system – *hvis systemets respons går mot en konstant verdi p.g.a. påvirkningen, noe som må undersøkes separat* – finnes ved å multiplisere denne konstanten med transferfunksjonen mellom påvirkningen og responsen, innsatt $s = 0$. Vi har i vårt eksempel

$$\frac{x_1(s)}{\gamma(s)} = h(s) = -\frac{m}{k + fs + ms^2}, \text{ som gir } h(0) = -\frac{m}{k} \quad (4.53)$$

Sluttverditeoremet er nyttig ved beskrivelse av dynamiske systemer fordi systemer ofte påvirkes av *konstante* forstyrrelser, referanseverdier eller liknende.

Til slutt: I figur 4.8 **g)** er transferfunksjonen mellom $u_1(s)$ og $x_1(s)$ også vist omformet slik at de første leddene i telleren og nevneren er lik 1. Dette er en fordel når vi seinere skal finne tidskonstantene i systemet.



4.5.3 Blokkdiagram i tidsplan $-s$ som derivasjonsoperator

De regler for manipulasjon og reduksjon av blokkdiagrammer som er beskrevet i avsnitt 4.5, følger av visse viktige egenskaper ved Laplacetransformasjonen, som superposisjonsprinsippet og at utgangsvariabelen er inngangsvariabelens Laplacetransform multiplisert med transferfunksjonen til den aktuelle blokk.

Hvis vi nå betrakter et blokkdiagram i tidsplanet, slik som i figur 3.7, vil multiplikasjon med $1/s$ erstattes av et integral. Men også i tidsplanet kan vi manipulere og redusere blokkdiagrammet etter de samme regler, hvis integratorsymbolene erstattes med blokker $1/s$, og vi nå tolker symbolet “ s ” som en **derivasjonsoperator**. Vi definerer nå notasjonen

$$sx = \frac{d}{dt}x, \quad s^2x = \frac{d^2}{dt^2}x, \quad \text{osv.} \quad (4.54)$$

Med denne tolkninga vil et blokkdiagram slik de er skissert i figur 4.6, ha variable $z_i(t)$, dvs. de er tidsfunksjoner. Hvis man reduserer et blokkdiagram og finner en transferfunksjon fra u til x :

$$h(s) = \frac{\beta_0 + \beta_1s + \dots + \beta_{n-1}s^{n-1}}{\alpha_0 + \alpha_1s + \alpha_2s^2 + \dots + \alpha_ns^n} = \frac{x(t)}{u(t)} \quad (4.55)$$

så vil dette i *tidsplanet* (jfr. høyre likhetstegn) bli å tolke som en alternativ skrivemåte for den høyere ordens differensiallikning

$$\alpha_0x(t) + \alpha_1\dot{x}(t) + \alpha_2\ddot{x}(t) + \dots + \alpha_n \overset{(n)}{x}(t) = \beta_0u(t) + \beta_1\dot{u}(t) + \dots + \beta_{n-1} \overset{(n-1)}{u}(t) \quad (4.56)$$

I mye av litteraturen bruker man symbolet p for derivasjonsoperator, for å unngå forveksling med Laplacetransformasjonens s . I denne boka brukes s i begge betydninger, og det går fram av sammenhengen hvilken tolkning som gjelder.

4.5.4 Bruk av benevninger for å sjekke transferfunksjoner

På samme måte som benevninger er nyttige ved sjekk av differensiallikninger, kan de også brukes ved sjekk av transferfunksjoner. Først må vi da avklare hva benevninga på “ s ” er. Laplacetransformasjonen av en tidsfunksjon er

$$h(s) = \int_0^{\infty} e^{-ts} h(t) dt$$

” $-ts$ ” i e^{-ts} må være dimensjonsløs. Da blir benevninga til s [1/tidsenhet], for eksempel [1/sek] hvis t måles i sekunder.

La oss prøve dette på svingende-masse-systemet:

Med $y[m]$ som utgang, og $u[N]$ som pådrag, har vi

$$\frac{y}{u}(s) = h(s) = \frac{1}{ms^2 + fs + k}$$

Erstatter symbolene med deres benevninger:

$$\frac{1}{[\text{kg}] \cdot \left[\frac{1}{\text{sek}^2}\right] + \left[\frac{\text{N}}{\text{m}}\right] \cdot \left[\frac{1}{\text{sek}}\right] + \left[\frac{\text{N}}{\text{m}}\right]} = \frac{1}{\left[\frac{\text{kg}\cdot\text{m}}{\text{sek}^2\cdot\text{m}}\right] + \left[\frac{\text{N}}{\text{m}}\right] + \left[\frac{\text{N}}{\text{m}}\right]} = \frac{1}{\left[\frac{\text{N}}{\text{m}}\right] + \left[\frac{\text{N}}{\text{m}}\right] + \left[\frac{\text{N}}{\text{m}}\right]}$$

$$\text{siden } [\text{N}] = [\text{kg m/s}^2] \implies \frac{y}{u} = \left[\frac{\text{m}}{\text{N}}\right] \text{ som er riktig!}$$

Merk at alle ledd i nevneren, og i enhver sum for den saks skyld, må ha samme benevning, noe som er nyttig for kontrollformål.

4.6 LINEÆRE VEKTORDIFFERENSIALLIKNINGER OG LAPLACETRANSFORMASJONEN

Vi skal nå ta for oss vektordifferensiallikningen

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t)\end{aligned}\tag{4.57}$$

Hvert element i en vektor som er en funksjon av tiden, kan Laplacetransformeres. Resultatet blir en vektor av Laplacetransformerte.

$$\mathcal{L}(\mathbf{x}(t)) = \mathbf{x}(s)$$

Laplacetransformasjonen av (4.57) gir

$$\begin{aligned}s\mathbf{x}(s) - \mathbf{x}(t=0) &= \mathbf{A}\mathbf{x}(s) + \mathbf{B}\mathbf{u}(s) \\ \mathbf{y}(s) &= \mathbf{C}\mathbf{x}(s)\end{aligned}\tag{4.58}$$

Løsning av (4.58) med hensyn på $\mathbf{x}(s)$ gir

$$\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(t=0) + (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(s)\tag{4.59}$$

og

$$\mathbf{y}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(t=0) + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(s)\tag{4.60}$$

Vi henter fram resultatet i (3.22) og antar at $t_0 = 0$. For systemet i (4.57) får vi da

$$\begin{aligned}\mathbf{x}(t) &= e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau \\ &= \Phi(t)\mathbf{x}(0) + \int_0^t \Phi(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau\end{aligned}\quad (4.61)$$

Sammenlikning av (4.59) og (4.61) gir

$$\mathcal{L}(\Phi(t)) = \mathcal{L}(e^{\mathbf{A}t}) = \Phi(s) = (s\mathbf{I} - \mathbf{A})^{-1} \quad (4.62)$$

$$\mathcal{L}\left(\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau\right) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(s) \quad (4.63)$$

$(s\mathbf{I} - \mathbf{A})^{-1}$ kalles **resolventmatrisen** og den er lik Laplacetransformasjonen av transisjonsmatrisen $\Phi(t)$. Dette er den vektorielle utgaven av betraktningene vi gjorde rundt likning (4.29).

Dersom matrisen \mathbf{A} har lineært uavhengige egenvektorer, kan den diagonaliseres ved hjelp av en egenvektormatrise¹ $= \mathbf{M}$. Vi kan skrive

$$\mathbf{A} = \mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1} \quad (4.64)$$

der $\mathbf{\Lambda}$ = egenverdimatrisen. Dette gir

$$\begin{aligned}(s\mathbf{I} - \mathbf{A})^{-1} &= (s\mathbf{M}\mathbf{M}^{-1} - \mathbf{M}\mathbf{\Lambda}\mathbf{M}^{-1})^{-1} = (\mathbf{M}(s\mathbf{I} - \mathbf{\Lambda})\mathbf{M}^{-1})^{-1} \\ &= \mathbf{M}(s\mathbf{I} - \mathbf{\Lambda})^{-1}\mathbf{M}^{-1} \\ &= \mathbf{M} \begin{bmatrix} \frac{1}{s - \lambda_1} & 0 & \cdot & \cdot & 0 \\ 0 & \frac{1}{s - \lambda_2} & & & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \frac{1}{s - \lambda_n} \end{bmatrix} \mathbf{M}^{-1}\end{aligned}\quad (4.65)$$

¹ \mathbf{M} kalles også **modalmatrisen**.

Vi innfører

$$\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n] \quad (4.66)$$

$\mathbf{m}_i =$ egenvektor

og

$$\mathbf{M}^{-1} = \mathbf{N} = \begin{bmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \\ \cdot \\ \cdot \\ \mathbf{n}_n^T \end{bmatrix} \quad (4.67)$$

der \mathbf{n}_i kalles **resiprok basisvektor**.

Likning (4.65) antar formen

$$(s\mathbf{I} - \mathbf{A})^{-1} = \sum_{i=1}^n \mathbf{m}_i \mathbf{n}_i^T \frac{1}{(s - \lambda_i)} \quad (4.68)$$

der høyresiden uttrykker summen av n matriser. Elementærleddene, $1/(s - \lambda_i)$, forekommer derfor i alle elementene i $(s\mathbf{I} - \mathbf{A})^{-1}$. Innfører vi (4.65) i (4.59) og foretar inverstransformasjon, får vi en løsning som blir av samme form som den som tidligere er presentert i (3.19). Vi legger merke til at egenverdiene $\lambda_1, \lambda_2, \dots, \lambda_n$ til matrisen \mathbf{A} spiller samme rolle i disse uttrykkene som singularitetene s_1, s_2, \dots, s_n spiller i uttrykket (4.34).

EKSEMPEL 4.7: Laplacetransformasjon av tilstandsvektoren i svingende-masse-systemet

Vi skal betrakte systemet i eksempel 3.3 som ble introdusert i eksempel 2.5, og som består av en masse som er festet til en fjær og en demper. Systemet er karakterisert ved følgende likning

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -21 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0.25 \end{bmatrix} u$$

Vi bruker likning (4.59) for å finne $\mathbf{x}(s)$. $(s\mathbf{I} - \mathbf{A})^{-1}$ finnes ved hjelp av egenverdiene og egenvektormatrisen til \mathbf{A} .

Egenvektorene, egenvektormatrisen og den inverse til egenvektormatrisen ble funnet i eksempel 3.3. Egenverdiene var $\lambda_1 = -7$ og $\lambda_2 = -3$. Egenvektormatrisen (modalmatrisen) og dens inverse var

$$\mathbf{M} = \begin{bmatrix} 1 & -\frac{1}{3} \\ -7 & 1 \end{bmatrix} \text{ og } \mathbf{M}^{-1} = \frac{1}{4} \begin{bmatrix} -3 & -1 \\ -21 & -3 \end{bmatrix}$$

Likning (4.65) benyttes for å finne $(s\mathbf{I} - \mathbf{A})^{-1}$

$$\frac{1}{4} \begin{bmatrix} 1 & -\frac{1}{3} \\ -7 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{s+7} & 0 \\ 0 & \frac{1}{s+3} \end{bmatrix} \begin{bmatrix} -3 & -1 \\ -21 & -3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} \frac{-3}{s+7} + \frac{7}{s+3} & \frac{-1}{s+7} + \frac{1}{s+3} \\ \frac{21}{s+7} - \frac{21}{s+3} & \frac{7}{s+7} - \frac{3}{s+3} \end{bmatrix}$$

Vi benytter formel (4.59) og kan se bort i fra leddet som inneholder initialtilstandene, siden disse er 0.

$$\begin{aligned} \mathbf{x}(s) &= \frac{1}{4} \begin{bmatrix} \frac{-3}{s+7} + \frac{7}{s+3} & \frac{-1}{s+7} + \frac{1}{s+3} \\ \frac{21}{s+7} - \frac{21}{s+3} & \frac{7}{s+7} - \frac{3}{s+3} \end{bmatrix} \begin{bmatrix} 0 \\ 0.25 \end{bmatrix} u(s) \\ &= \begin{bmatrix} \frac{-1}{16(s+7)} + \frac{1}{16(s+3)} \\ \frac{7}{16(s+7)} + \frac{3}{16(s+3)} \end{bmatrix} u(s) = \begin{bmatrix} \frac{1}{84(\frac{s}{3} + 1)(\frac{s}{7} + 1)} \\ \frac{s}{84(\frac{s}{3} + 1)(\frac{s}{7} + 1)} \end{bmatrix} u(s) \end{aligned} \quad (4.69)$$

Likning (4.69) viser at $x_2(s) = s x_1(s)$. Dette betyr i følge (4.12) at $x_2(t) = \dot{x}_1(t)$.

■

4.6.1 Transfermatrise

Uttrykkene i (4.59) og (4.61) inneholder to ledd:

- Det første skyldes systemets begynnelsestilstand og vil i tidsplanet være en eksponentiell innsvingning mot null, så sant systemets egenverdier (λ_i) er slik at $\text{Re } \lambda_i < 0$.
- Den andre delen skyldes den ytre påvirkningen (pådraget).

Det samme gjelder representasjonen systemets tilstander får gjennom målevektoren, enten denne har få eller mange elementer.

Antar vi nå at systemets initialtilstand $\mathbf{x}(0) = \mathbf{0}$, vil (4.60) bli

$$\mathbf{y}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} \mathbf{u}(s) = \mathbf{H}(s) \mathbf{u}(s) \quad (4.70)$$

der $\mathbf{H}(s)$ er **transfermatrisen** og der elementene $h_{ij}(s)$ blir transferfunksjoner. En transfermatrise kan selvfølgelig etableres mellom pådragsvektoren og en hvilken som helst avledning av systemtilstanden, for eksempel tilstandsvektoren selv.

For å være mer presis med hensyn til hvilken transfermatrise det dreier seg om, kunne vi i (4.70) ha benyttet skriveformen $\mathbf{H}_{yu}(s)$.

Vi setter inn resultatet fra (4.65) i (4.70)

$$\mathbf{H}(s) = \mathbf{C}\mathbf{M} \begin{bmatrix} \frac{1}{s-\lambda_1} & 0 & \cdot & \cdot & 0 \\ 0 & \frac{1}{s-\lambda_2} & & & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \frac{1}{s-\lambda_n} \end{bmatrix} \mathbf{M}^{-1}\mathbf{B} \quad (4.71)$$

Dette viser at i det generelle tilfellet vil alle transferfunksjonene $h_{ij}(s)$ inneholde alle faktorene av typen

$$\frac{1}{s-\lambda_k}, \quad k = 1, 2, \dots, n \quad (4.72)$$

slik at vi får

$$h_{ij}(s) = \frac{g_{ij}(s)}{(s-\lambda_1)(s-\lambda_2)\dots(s-\lambda_n)} \quad (4.73)$$

der $g_{ij}(s)$ er et polynom i s av orden høyst lik $n-1$. Transfermatrisen $\mathbf{H}(s)$ og dens elementer kan selvsagt også bestemmes direkte av (4.70) uten å gå veien om diagonalisering, men ved bruk av regelen for invertering av en matrise. En får da ved å bruke resultatet i (A.10)

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \mathbf{C} \frac{\text{adj}(s\mathbf{I} - \mathbf{A})}{|s\mathbf{I} - \mathbf{A}|} \mathbf{B} = \frac{1}{|s\mathbf{I} - \mathbf{A}|} \mathbf{C}(\text{adj}(s\mathbf{I} - \mathbf{A}))\mathbf{B} \quad (4.74)$$

Brukes (4.74) vil transferfunksjonene $h_{ij}(s)$ ikke ha nevner oppdelt i enkeltfaktorer, men en fellesnevner som er et polynom i s av orden n ; det karakteristiske polynom for systemet.

EKSEMPEL 4.8: Å finne et systems transfermatrise

Vi betrakter systemet

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & 2 \\ 0 & -4 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \mathbf{u}$$

$$\mathbf{y} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \mathbf{x}$$

Eigenverdiene er $\lambda_1 = -1$ og $\lambda_2 = -4$. Ett sett egenvektorer kan bestemmes som gir egenverdimatrisen

$$\mathbf{M} = \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}$$

Den inverse av egenverdimatrisen er

$$\mathbf{M}^{-1} = \frac{1}{-3} \begin{bmatrix} -3 & -2 \\ 0 & 1 \end{bmatrix}$$

Vi benytter (4.71):

$$\mathbf{H}(s) = \frac{1}{-3} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} \frac{1}{s+1} & 0 \\ 0 & \frac{1}{s+4} \end{bmatrix} \begin{bmatrix} -3 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \Rightarrow$$

$$\mathbf{H}(s) = \begin{bmatrix} \frac{4(s+3)}{(s+1)(s+4)} & \frac{2(s+2)}{(s+1)(s+4)} \\ \frac{7s+31}{(s+1)(s+4)} & \frac{s+7}{(s+1)(s+4)} \end{bmatrix}$$

(4.75)

I dette tilfellet ville det vært enklere å bruke (4.74) direkte. Dette gir

$$\mathbf{H}(s) = \frac{1}{(s+1)(s+4)} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} s+4 & 2 \\ 0 & s+1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

(4.76)

Utrekning av dette uttrykket gir samme svar som (4.75). Vi legger merke til at alle transferfunksjonene i transfermatrisen $\mathbf{H}(s)$ har de samme polene.

■

En transferfunksjon for et system beskrevet med lineære differensiallikninger med konstante koeffisienter vil, som vi ser av (4.73), bli en rasjonal funksjon av den komplekse variabelen s der telleren og nevneren er polynomer i s . Det som framfor alt karakteriserer en slik *rasjonal funksjon* er røttene i nevnerens og tellerens polynomer. Disse røttene kalles henholdsvis **poler** og **nullpunkter**. Polene til en transferfunksjon vil være **singulariteter** i denne komplekse funksjonen, fordi funksjonen går mot uendelig når s nærmer seg en slik verdi (nevneren blir lik null).

Polene for transfermatrisen blir identiske med egenverdiene til systemmatrisen

Vi skal nå diskutere forskjellige valg av tilstandsvariable og går for enkelhets skyld ut fra at systemet har ett pådrag, én måling og ingen forstyrrelser.

4.6.2 Fra transferfunksjon til tilstandsrombeskrivelse

Hvis systemet er monovariabelt (\mathbf{u} og \mathbf{y} skalare), vil transfermatrisen $\mathbf{H}(s)$ i (4.74) reduseres til å bli en (skalar) transferfunksjon; et rasjonalt uttrykk i s ,

$$h(s) = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} \quad (4.77)$$

Transferfunksjonen vil bli den samme for et stort antall forskjellige tilstandsrommodeller av typen

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}u \\ y &= \mathbf{c}^T \mathbf{x} \end{aligned} \quad (4.78)$$

Grunnen til dette er at vi kan foreta en vilkårlig lineær, ikke-singulær transformasjon av tilstandsvektoren \mathbf{x} uten å endre overføringen mellom u og y . Vi skal demonstrere dette for den spesielle transformasjonen som diagonaliserer et system, jfr (3.64). Ved hjelp av transformasjonen $\mathbf{x} = \mathbf{M}\mathbf{q}$, dvs. $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$, fikk vi

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{\Lambda}\mathbf{q} + \mathbf{b}u \\ y &= \mathbf{f}^T \mathbf{q} \end{aligned} \quad , \text{ der } \mathbf{\Lambda} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}, \quad \mathbf{h} = \mathbf{M}^{-1}\mathbf{b} \text{ og } \mathbf{f}^T = \mathbf{c}^T \mathbf{M} \quad (4.79)$$

Transferfunksjon for dette systemet blir $\hat{h}(s) = h(s)$, fordi

$$\begin{aligned} \hat{h}(s) &= \mathbf{f}^T (s\mathbf{I} - \mathbf{\Lambda})^{-1} \mathbf{h} = \mathbf{c}^T \mathbf{M} (s\mathbf{M}^{-1}\mathbf{I}\mathbf{M} - \mathbf{M}^{-1}\mathbf{A}\mathbf{M})^{-1} \mathbf{M}^{-1} \mathbf{b} \\ &= \mathbf{c}^T \mathbf{M} [\mathbf{M}^{-1} (s\mathbf{I} - \mathbf{A}) \mathbf{M}]^{-1} \mathbf{M}^{-1} \mathbf{b} = \mathbf{c}^T \mathbf{M} \mathbf{M}^{-1} (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{M} \mathbf{M}^{-1} \mathbf{b} \\ &= \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} = h(s) \end{aligned} \quad (4.80)$$

Vi har nå vist dette for den spesielle transformasjonen $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$. Men akkurat det samme gjelder for en *vilkårlig* transformasjon $\boldsymbol{\eta} = \mathbf{Z}^{-1}\mathbf{x}$, som vi var inne på i (3.71) og (3.72). Det

eneste som kreves er at \mathbf{Z} er kvadratisk og inverterbar. Det finnes uendelig mange matriser som oppfyller dette kravet.

Dette betyr at om vi kjenner et systems transferfunksjon, kan vi ikke entydig avlede en tilstandsrombeskrivelse. Det finnes uendelig mange tilstandsrombeskrivelser som gir samme transferfunksjon. Men da kan vi alltid velge oss en eller annen tilstandsrombeskrivelse som svarer til en gitt transferfunksjon.

I avsnitt 3.5 har vi introdusert forskjellige typer kanoniske former av et monovariabelt system. Det knytter seg i denne forbindelse særlig interesse til de kanoniske formene som er gitt i (3.66) og (3.69). Parametrene som karakteriserer disse formene, står nemlig i et særlig enkelt forhold til parametrene i de tilsvarende transferfunksjonene. Antar vi at det monovariabelt systemet er beskrevet av transferfunksjonen

$$h(s) = \frac{\rho_{n-1}s^{n-1} + \dots + \rho_1s + \rho_0}{s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_1s + \alpha_0} \quad (4.81)$$

finner vi at begge systemtypene vist i figur 3.9 og figur 3.10 vil ha en slik transferfunksjon.

For den første representasjonsformen (“fasevariabel form”; på engelsk “controllable canonical form”) vil vi få, referert til uttrykket (4.78),

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 1 & & & 0 & 0 \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \cdot & \cdot & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdot & \cdot & -\alpha_{n-2} & -\alpha_{n-1} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} \quad (4.82)$$

$$\mathbf{c}^T = [\rho_0 \quad \rho_1 \quad \rho_2 \quad \cdot \quad \cdot \quad \cdot \quad \rho_{n-1}]$$

Dette kan enklest vises ved å tegne elementært blokkdiagram i s -planet basert på denne tilstandsrommodellen, og så redusere blokkdiagrammet. Da vil man få (4.81).

Bruker vi den andre representasjonsformen (figur 3.10), får vi

$$\mathbf{A} = \begin{bmatrix} -\alpha_{n-1} & 1 & 0 & \cdot & \cdot & 0 & 0 \\ -\alpha_{n-2} & 0 & 1 & & & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -\alpha_1 & 0 & 0 & \cdot & \cdot & 0 & 1 \\ -\alpha_0 & 0 & 0 & \cdot & \cdot & 0 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \rho_{n-1} \\ \rho_{n-2} \\ \cdot \\ \cdot \\ \cdot \\ \rho_1 \\ \rho_0 \end{bmatrix}$$

$$\mathbf{c}^T = [1 \quad 0 \quad 0 \quad \cdot \quad \cdot \quad \cdot \quad 0]$$

Vi legger merke til at i begge tilfellene ovenfor skulle det $2n$ parametre til for å beskrive systemets transferfunksjon, og de samme parametre går igjen i hver av de to tilstandsrombeskrivelsene. Hvis vi finner røttene til nevneren i (4.81), kan vi skrive

$$h(s) = \frac{\rho_{n-1}s^{n-1} + \dots + \rho_1s + \rho_0}{(s - \lambda_1)(s - \lambda_2)\dots(s - \lambda_n)} \quad (4.83)$$

λ_i tilsvarer systemets egenverdier. Dersom egenverdiene er distinkte, kan vi på en enkel måte foreta en delbrøksoppspaltning, slik at vi får

$$h(s) = \frac{f_1}{(s - \lambda_1)} + \frac{f_2}{(s - \lambda_2)} + \dots + \frac{f_n}{(s - \lambda_n)} \quad (4.84)$$

der

$$f_i = (s - \lambda_i)h(s)|_{s=\lambda_i} \quad (4.85)$$

Dette resulterer i en diagonalform som vist i figur 3.8, der tilstandsrombeskrivelsen med notasjon som i (3.64) blir

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

$$\mathbf{f}^T = [f_1 \quad f_2 \quad \cdot \quad \cdot \quad f_n]$$

4.7 SAMMENHENGEN POLER/NULLPUNKTER OG TIDSRESPONS

4.7.1 Rasjonale transferfunksjoner

Transferfunksjonen for et monovariabelt system beskrevet med ordinære lineære differensiallikninger med konstante koeffisienter vil være en **rasjonal funksjon** i s , det vil si telleren og nevneren vil være polynomer i s . Ordenstallet til polynomet i telleren vil for et fysisk realiserbart system høyst være lik og vanligvis være mindre enn ordenstallet til nevnerens polynom. Vi benytter samme notasjon som i (4.81):

$$h(s) = \frac{\rho_p s^p + \dots + \rho_1 s + \rho_0}{s^n + \alpha_{n-1} s^{n-1} + \dots + \alpha_1 s + \alpha_0} = \frac{\rho_p (s - v_1)(s - v_2) \dots (s - v_p)}{(s - \lambda_1)(s - \lambda_2) \dots (s - \lambda_n)} \quad (4.86)$$

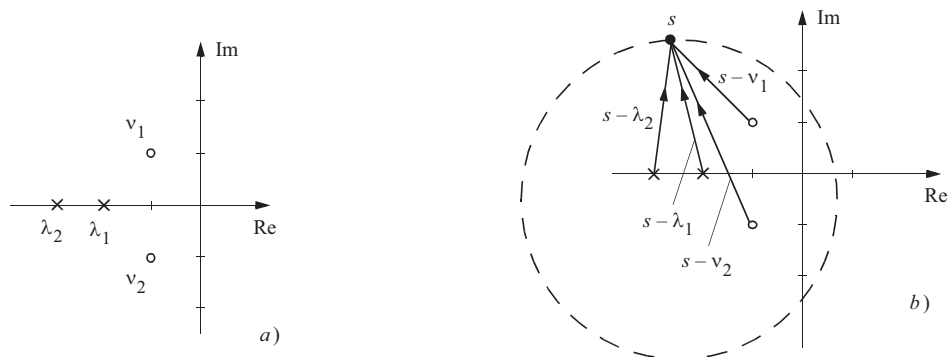
Vi kaller røttene i henholdsvis teller og nevner for **nullpunkter** og **poler**. I (4.86) har vi antatt at det er p nullpunkter og n poler. Videre ser vi at polene er gitt betegnelsen λ_i fordi de er identiske med egenverdiene¹.

Når vi skisserer i det komplekse plan velger vi å angi nullpunkter ved hjelp av små sirkler og poler ved hjelp av små kryss. Et eksempel vil illustrere dette: Transferfunksjonen

$$h(s) = \frac{2s^2 + 4s + 4}{s^2 + 5s + 6} = 2 \frac{(s + 1 + j)(s + 1 - j)}{(s + 2)(s + 3)} \quad (4.87)$$

har kompleks konjugerte nullpunkter med negativ realdel. Polene er reelle og negative. Plasingen av nullpunkter og poler er vist i figur 4.9 a).

Figur 4.9
a) Nullpunkter og poler til $h(s)$
b) Kontur som karakteriserer $h(s)$



¹Dette gjelder så lenge systemet er **styrbart** og **observerbart**. Disse begrepene behandles i kapittel 5. De fleste systemer er styrbare og observerbare.

Figur 4.9 *a*) vil entydig karakterisere transferfunksjonen hvis vi ser bort fra den konstante faktoren i (4.87).

Vi kan finne transferfunksjonen og dens egenskaper ved å betrakte plasseringen av dens nullpunkter og poler i det komplekse plan.

Omvendt kan det vises, blant annet ved hjelp av Cauchys integralsats, at dersom transferfunksjonens verdi er kjent i alle punkter s (transferfunksjonen $h(s)$ er en kompleks funksjon i s) langs en vilkårlig lukket kontur som omslutter samtlige nullpunkter og poler – se eksempel på dette i figur 4.9 *b*) – vil konfigurasjonen av nullpunkter og poler være entydig kjent.

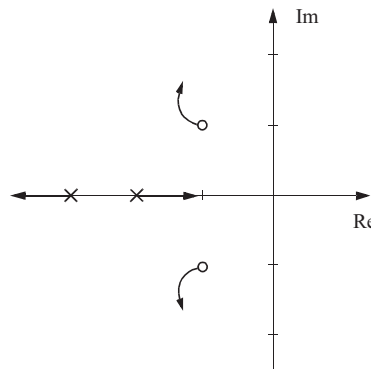
En annen kontur, som følger den imaginære akse fra $-j\infty$ til $+j\infty$ og omslutter venstre del av det komplekse plan (venstre halvplan), tilfredsstiller også dette kravet. Dette skal senere utnyttes til beskrivelse av dynamiske systemers egenskaper ved hjelp av frekvensanalyse.

I resten av dette kapitlet skal vi også drøfte et annet viktig poeng:

Plasseringen av nullpunkter og poler endrer seg når parametrene til et system endres.

Slik bevegelse av nullpunkter og poler er antydnet i figur 4.10. I dette eksemplet medfører en endring av parametre at polene går fra hverandre, mens nullpunktene beveger seg etter kurver på en slik måte at de alltid er **kompleks konjugerte**.¹

Figur 4.10
Et eksempel på polenes og nullpunktens bevegelse i det komplekse plan



Vi har diverse metoder til å tolke konsekvensen av bevegelsen til nullpunkter og poler, for eksempel ved hjelp av de motsvarende sprangresponser. Betrakninger av pol- og nullpunktsplassering er et nyttig hjelpemiddel for beskrivelse av et dynamisk systems egenskaper.

¹Polynomene i teller og nevner har alltid reelle koeffisienter. For polynomer med reelle koeffisienter kan eventuelle komplekse røtter bare forefinnes som konjugerte par.

4.7.2 Egenskapene til 2. ordens system

Svært mange prosesser kan modelleres som om de er av 2. orden. En gjennomgang av dynamikken til 2. ordens systemer vil også gi innsikt i de viktigste egenskaper som karakteriserer systemer av orden *større* enn 2. Vi skal derfor gå grundig inn på 2. ordens systemer, og gjør det i form av et representativt eksempel.

EKSEMPEL 4.9: System med svingende masse

Svingende-masse-systemet i eksempel 2.5 var karakterisert ved følgende 2. ordens differensiallikning

$$\ddot{x}(t) + \frac{f}{m}\dot{x}(t) + \frac{k}{m}x(t) = \frac{1}{m}p(t) \quad (4.88)$$

Vi definerer bevegelsen $x(t)$ som en måling $y(t)$, definerer kraften $p(t)$ som et pådrag $u(t)$, og setter $y(0) = 0$ og $\dot{y}(0) = 0$. Laplacetransformasjon av (4.88) gir da

$$s^2y(s) + \frac{f}{m}sy(s) + \frac{k}{m}y(s) = \frac{1}{m}u(s) \quad (4.89)$$

Transferfunksjonen blir

$$\frac{y(s)}{u(s)} = h(s) = \frac{1}{ms^2 + fs + k} \quad (4.90)$$

$h(s)$ kan også uttrykkes som

$$h(s) = \frac{\frac{1}{m}}{s^2 + \frac{f}{m}s + \frac{k}{m}} = \frac{\frac{1}{m}}{(s - \lambda_1)(s - \lambda_2)} \quad (4.91)$$

der λ_1 og λ_2 er systemets poler. Av (4.91) finner vi

$$\left. \begin{array}{l} \lambda_1 \\ \lambda_2 \end{array} \right\} = -\frac{f}{2m} \left(1 \pm \sqrt{1 - 4\frac{km}{f^2}} \right) \quad (4.92)$$

Siden m , f og k er positive, vil polene (egenverdiene) ha negativ realverdi og følgelig ligge i venstre halvplan.

$$\begin{aligned}
4\frac{km}{f^2} < 1 & \text{ gir reelle og forskjellige poler} \\
4\frac{km}{f^2} = 1 & \text{ gir reelle og sammenfallende poler} \\
4\frac{km}{f^2} > 1 & \text{ gir kompleks konjugerte poler}
\end{aligned} \tag{4.93}$$

Antar vi nå at kraften som påvirker massen, $u(t)$, er et enhetssprang ved $t = 0$, vet vi at $u(s) = 1/s$. Laplacetransformen av massens posisjon $y(t)$ blir da

$$y(s) = \frac{1}{s(s - \lambda_1)(s - \lambda_2)} \tag{4.94}$$

Ved bruk av residusatsen finner vi

$$y(t) = \sum_{i=1}^3 [\text{res}(y(s)e^{ts})]_{s=s_i} \tag{4.95}$$

der s_i er singulariteter av $y(s)$, $s_1 = 0$, $s_2 = \lambda_1$ og $s_3 = \lambda_2$.

Avhengig av om polene er reelle og forskjellige, reelle og sammenfallende, eller kompleks konjugerte, får vi forskjellige løsninger av (4.95):

Dersom polene λ_1 og λ_2 ikke er sammenfallende, blir sprangresponsen

$$\begin{aligned}
y(t) &= \frac{1}{m} \left[\frac{1}{\lambda_1 \lambda_2} + \frac{e^{\lambda_1 t}}{\lambda_1(\lambda_1 - \lambda_2)} + \frac{e^{\lambda_2 t}}{\lambda_2(\lambda_2 - \lambda_1)} \right] \\
&= \frac{1}{k} \left[1 + \frac{\lambda_2}{\lambda_1 - \lambda_2} e^{\lambda_1 t} - \frac{\lambda_1}{\lambda_1 - \lambda_2} e^{\lambda_2 t} \right]
\end{aligned} \tag{4.96}$$

der verdiene av λ_1 og λ_2 fra (4.92) skal brukes.

Videre drøfting vil skje i tre underavsnitt, jfr. (4.93)

Tilfelle 1: Reelle og forskjellige poler (overdempet system)

Hvis vi har reelle og forskjellige poler kan vi innføre tidskonstantene

$$T_1 = -\frac{1}{\lambda_1} \text{ og } T_2 = -\frac{1}{\lambda_2}$$

Da vil transferfunksjonen anta formen

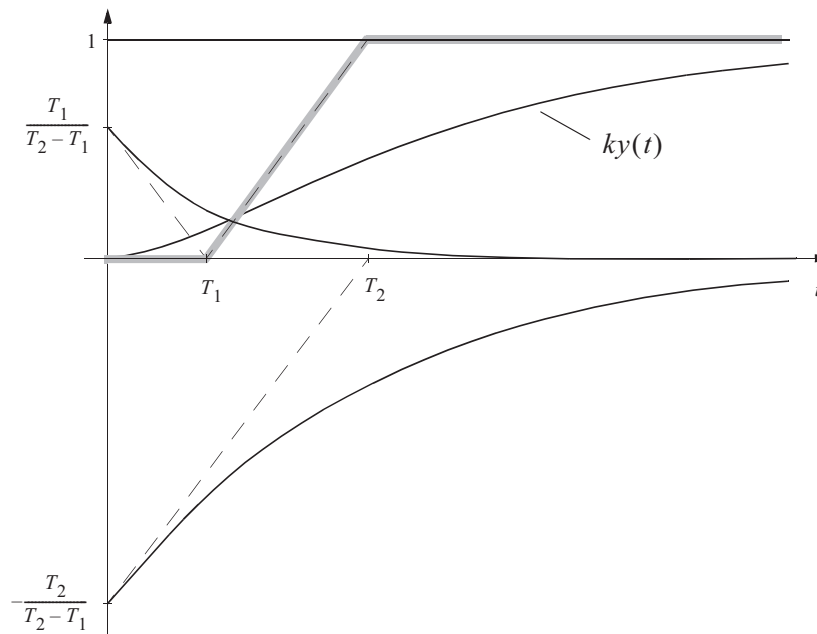
$$h(s) = \frac{\frac{1}{k}}{(1 + T_1s)(1 + T_2s)} \quad (4.97)$$

Dette er ofte en mer hensiktsmessig form enn den som er gitt i (4.90). Omskrevet vil (4.96) da bli

$$y(t) = \frac{1}{k} \left[1 + \frac{T_1}{T_2 - T_1} e^{-\frac{t}{T_1}} - \frac{T_2}{T_2 - T_1} e^{-\frac{t}{T_2}} \right] \quad (4.98)$$

der $0 < T_1 < T_2$. Innholdet av (4.98) er vist i figur 4.11.

Figur 4.11
Tidsrespons ved
reelle og forskjell-
lige poler



De tre enkeltbidragene er vist hver for seg og med (grove) rettlinjede approksimasjoner. Den totale responsen og dens stykkevis rettlinjede (grove) approksimasjon er også vist. Vi kan konstruere den rettlinjede approksimasjonen ved først å angi dens stasjonære verdi ($1/k$), avmerke de to tidskonstantene (T_1 og T_2) og deretter forbinde punktene med rette linjer som vist.

Tilfelle 2: Reelle og sammenfallende poler (kritisk dempet system)

Dersom polene i transferfunksjonen (4.91) og (4.97) er reelle og sammenfallende, dvs.

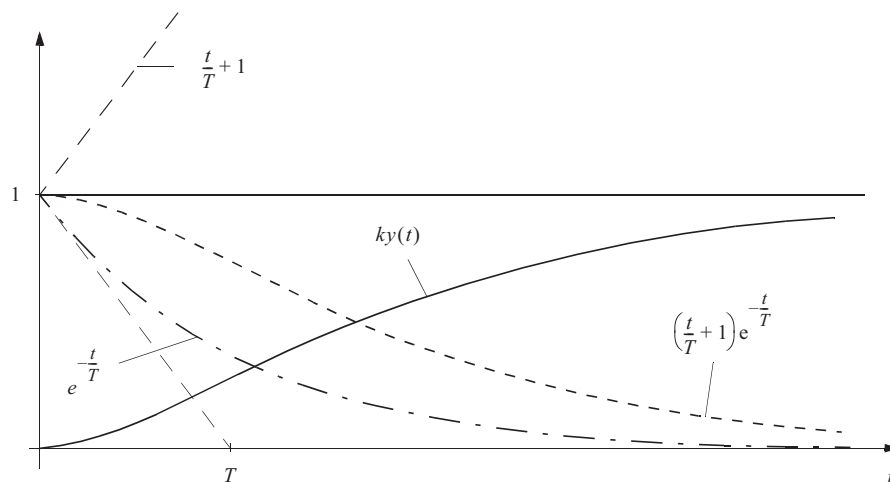
$$\lambda_1 = \lambda_2 = \lambda = -\frac{1}{T}$$

som oppstår når $4km/f^2 = 1$, vil en ved anvendelse av (4.38) få

$$\begin{aligned} y(t) &= \frac{1}{k} \left[1 - (-\lambda t + 1) e^{\lambda t} \right] \\ &= \frac{1}{k} \left[1 - \left(\frac{t}{T} + 1 \right) e^{-\frac{t}{T}} \right] \end{aligned} \quad (4.99)$$

Denne responsen er skissert i figur 4.12.¹

Figur 4.12
Tidsrespons ved
reelle og sammen-
fallende poler

Tilfelle 3: Kompleks konjugerte poler (underdempet system)

Når polene er kompleks konjugerte, innfører vi symbolene α og β i realdelen og imaginærdelen av en pol. Vi skriver polene som

$$\lambda_1 = -\alpha - j\beta \quad \text{og} \quad \lambda_2 = -\alpha + j\beta \quad (4.100)$$

¹En slik respons har vi allerede regnet ut for et annet tilfelle, eksempel 4.3

Vi kan nå skrive $h(s)$ som

$$\begin{aligned}
 h(s) &= \frac{\frac{1}{m}}{s^2 + \frac{f}{m}s + \frac{k}{m}} = \frac{\frac{1}{m}}{(s - \lambda_1)(s - \lambda_2)} = \frac{\frac{1}{m}}{(s + \alpha + j\beta)(s + \alpha - j\beta)} \\
 &= \frac{\frac{1}{m}}{s^2 + 2\alpha s + (\alpha^2 + \beta^2)} = \frac{\frac{1}{m}}{s^2 + 2\zeta\omega_0 s + \omega_0^2} = \frac{\frac{1}{k}}{1 + 2\zeta\frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}
 \end{aligned} \tag{4.101}$$

Vi har her også innført to nye størrelser, $\zeta = \frac{1}{2} \frac{f}{\sqrt{km}}$ og $\omega_0 = \sqrt{\frac{k}{m}}$. Sammenholder vi de to ledd til venstre i 2. linje i (4.101), finner vi at

$$\alpha^2 + \beta^2 = \omega_0^2 \tag{4.102}$$

$$\text{og } \zeta = \frac{\alpha}{\omega_0} = \sin \varphi \text{ (for mer om vinkelen } \varphi, \text{ se nedenfor)} \tag{4.103}$$

Følgende betegnelser benyttes for de definerte størrelser:

$$\alpha = \zeta \omega_0 = \frac{f}{2m} = \text{absolutt dempningsfaktor} \tag{4.104}$$

$$\beta = \sqrt{\frac{k}{m} - \left(\frac{f}{2m}\right)^2} = \text{svingefrekvens} \tag{4.105}$$

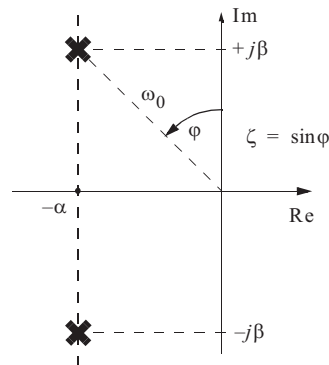
$$\omega_0 = \sqrt{\frac{k}{m}} = \text{udempet resonansfrekvens} \tag{4.106}$$

$$\zeta = \frac{1}{2} \frac{f}{\sqrt{km}} = \sin \varphi = \text{relativ dempningsfaktor} \tag{4.107}$$

$$\varphi = \arcsin(\zeta) = \text{faseforskyvning} \tag{4.108}$$

Figur 4.13 viser de to komplekskonjugerte polene, og angir størrelsene som er definert i (4.104)-(4.108). φ er vinkelen mellom vektoren ut til en av polene og den imaginære akse. ω_0 angir lengden av vektoren ut til en av polene.

Figur 4.13
Grafisk framstilling
i det komplekse
plan av sentrale
størrelser



Vi kan nå benytte svaret i (4.96) og sette inn $\lambda_1 = -\alpha - j\beta$, $\lambda_2 = -\alpha + j\beta$. Vi får

$$y(t) = \frac{1}{k} \left[1 - \frac{1}{2j\beta} \left\{ (-\alpha + j\beta)e^{(-\alpha - j\beta)t} - (-\alpha - j\beta)e^{(-\alpha + j\beta)t} \right\} \right] \quad (4.109)$$

$e^{-\alpha t}$ kan settes utafør klammeparentesen:

$$y(t) = \frac{1}{k} \left[1 - \frac{e^{-\alpha t}}{2j\beta} \left\{ (-\alpha + j\beta)e^{-j\beta t} - (-\alpha - j\beta)e^{j\beta t} \right\} \right] \quad (4.110)$$

Vi samler ledd med α og β hver for seg:

$$\begin{aligned} y(t) &= \frac{1}{k} \left[1 - \frac{e^{-\alpha t}}{2j\beta} \left\{ \alpha(e^{j\beta t} - e^{-j\beta t}) + j\beta(e^{j\beta t} + e^{-j\beta t}) \right\} \right] = \\ &= \frac{1}{k} \left[1 - \frac{e^{-\alpha t}}{\beta} \left\{ \alpha \frac{(e^{j\beta t} - e^{-j\beta t})}{2j} + j\beta \frac{(e^{j\beta t} + e^{-j\beta t})}{2j} \right\} \right] \end{aligned} \quad (4.111)$$

Eulers formler gir da:

$$y(t) = \frac{1}{k} \left[1 - \frac{e^{-\alpha t}}{\beta} (\beta \cos \beta t + \alpha \sin \beta t) \right] \quad (4.112)$$

Dette siste kan igjen omskrives til:

$$y(t) = \frac{1}{k} \left[1 - \frac{\omega_0}{\beta} e^{-\alpha t} \left(\frac{\beta}{\omega_0} \cos \beta t + \frac{\alpha}{\omega_0} \sin \beta t \right) \right] \quad (4.113)$$

Fra figur 4.13 har vi at $\cos \varphi = \beta/\omega_0$ og $\sin \varphi = \alpha/\omega_0$

Dette gir

$$y(t) = \frac{1}{k} \left[1 - \frac{\omega_0}{\beta} e^{-\alpha t} (\cos \varphi \cos \beta t + \sin \varphi \sin \beta t) \right] \quad (4.114)$$

Vi benytter nå den velkjente formelen $\cos(u \pm v) = \cos u \cos v \mp \sin u \sin v$ og finner svaret,

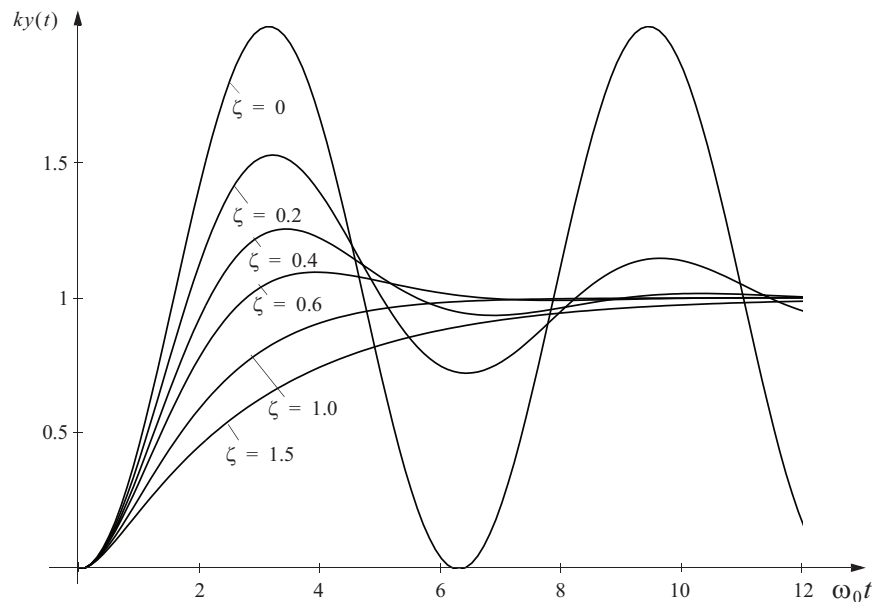
$$y(t) = \frac{1}{k} \left[1 - \frac{\omega_0}{\beta} e^{-\alpha t} \cos(\beta t - \varphi) \right], \text{ alternativt uttrykt som} \quad (4.115)$$

$$y(t) = \frac{1}{k} \left[1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta \omega_0 t} \cos(\sqrt{1-\zeta^2} \omega_0 t - \varphi) \right] \quad (4.116)$$

der α , β , ω_0 og φ er som definert tidligere.

Sprangresponser (4.115)/(4.116) er vist i figur 4.14 for forskjellige verdier av ζ , i dimensjonsløs tidsskala $\omega_0 t$. I tillegg er sprangresponser vist for sammenfallende poler, og for ett tilfelle av forskjellige reelle poler.

Figur 4.14
Sprangresponser
for forskjellige ζ



Den nederste kurven ($\zeta > 1$) tilsvare at systemets poler er reelle og forskjellige. Når $\zeta > 1$ kan man ikke bruke uttrykket $\zeta = \sin \varphi$. Men man kan alltid, så lenge $|\lambda_1|$ og $|\lambda_2| > 0$, uttrykke ζ som funksjon av de to polene, slik:

$$\zeta = \frac{-(\lambda_1 + \lambda_2)}{2\sqrt{\lambda_1 \lambda_2}} \quad (4.117)$$

Figur 4.15 viser hvordan man kan plote en av tidsresponsene i figur 4.14 ved hjelp av MATLAB.

Figur 4.15
MATLAB-kode
for plotting

```

1 w0 = 1.414; zeta = 0.2; k = 1;
2 t = [0:0.1:12]/w0;
3 phi = asin(zeta*ones(1,length(t)));
4 h = sqrt(1-zeta^2);
5 x = (1/k)*(1-(1/h)*exp(-zeta*w0*t).*cos(h*w0*t-phi));
6 plot(t,x,'linewidth',1)

```

$$\frac{1}{k} \left[1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_0 t} \cos(\sqrt{1-\zeta^2}\omega_0 t - \phi) \right]$$

Vanligvis vil vi ikke bruke MATLAB på denne måten, altså å plote en tidsrespons som vi først har beregnet algebraisk. I stedet vil vi la MATLAB finne den numeriske løsningen på problemet: Gitt en transferfunksjon $h(s)$ (eventuelt en tilstandsrommodell \mathbf{A} , \mathbf{B} , \mathbf{C}) og et pådrag som er et enhetsprang. Hva blir sprangresponsen? Figur 4.16 viser MATLAB-koden når systemet er gitt ved $h(s)$, og figur 4.17 viser koden når systemet er gitt ved en av flere mulige tilstandsrommodeller, for eksempel

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\zeta\omega_0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \quad \mathbf{c}^T = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (4.118)$$

(Dette er en fasevariabel kanonisk form (4.82), med koeffisienter fra (4.101)).

Figur 4.16
MATLAB-kode for
simulering og plot-
ting ved bruk av
transferfunksjon

```

1 k = 1; w0 = 1.414; zeta = 0.2;
2 t = [0:0.1:12]/w0;
3 t0 = 1/k;
4 n0 = [1/(w0^2) 2*zeta/w0 1];
5 step(t0,n0,t)

```

Figur 4.17

MATLAB-kode for simulering og plotting ved bruk av tilstandsromform

```

1 w0 = 1.414; zeta = 0.2; m = 0.5;
2 t = [0:0.1:12]/w0;
3 A = [0, 1; -w0^2, -2*zeta*w0];
4 b = [0; 1/m];
5 c = [1, 0];
6 d = 0,
7 step(A,b,c,d,1,t)

```

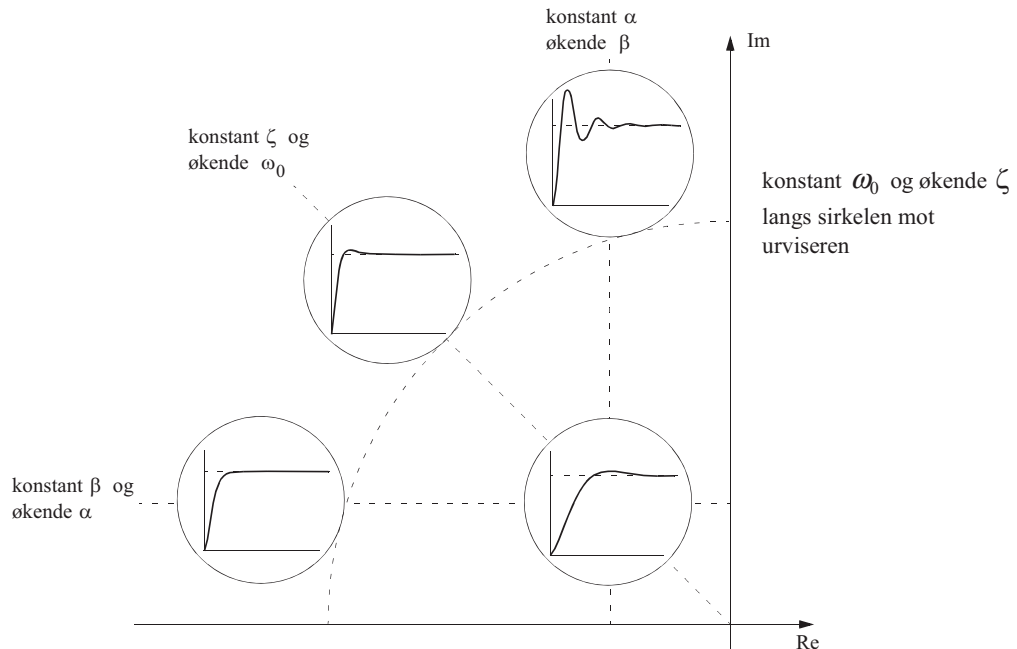
Beregner og plotter sprangresponsen ved bruk av tilstandsromform.

Figur 4.18 illustrerer hvordan polenes bevegelse henger sammen med α , β , ζ og ω_0 .

- Dersom dette 2. ordens systemet med to komplekse konjugerte poler forandres, slik at polene beveger seg parallelt med den reelle akse, altså konstant β , vil responsens svingefrekvens forbli konstant, mens den absolute dempningen av svingningsforløpet vil forandres.
- Hvis polene imidlertid beveger seg parallelt med den imaginære akse, vil den absolute dempningen være konstant, mens svingefrekvensen β av responsen vil forandre seg.
- Beveger polene seg langs en stråle fra origo, slik at ζ er konstant, vil både den absolute dempningen og svingefrekvensen forandre seg, men på en slik måte at formen på innsvingningen blir den samme (i forandret tidsskala).

Figur 4.18

Sammenhengen mellom polplassering, sentrale størrelser og tidsresponser



α , β , ζ og ω_0 påvirker *innsvingningsforløpet* som følger:

- Den **absolutte dempningsfaktoren** α sørger for at svingningene reduseres. Dette kan forklares ut fra leddet $e^{-\alpha t}$ i (4.115). Et system med stor α vil være sterkt dempet og reagere raskt.
- **Svingefrekvensen** β angir i følge (4.115) vinkelfrekvensen [rad/s] på svingningene. Jo større β er, desto raskere oscillerer systemet.
- Oversvinget i sprangresponsen er avhengig av den **relative dempningsfaktoren** ζ . Dersom ζ er liten, blir oversvinget stort.

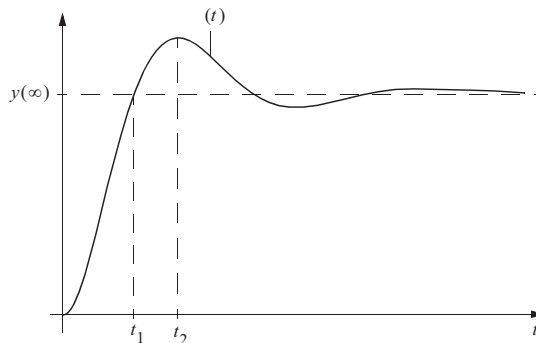
Sprangresponsen blir mer oscillatorisk (altså mindre dempet) når ζ er liten. Dersom $\zeta = 0$, blir den en stående svingning¹.

ζ angir den *absolutte* dempningen α relativt til ω_0 , altså $\zeta = \alpha/\omega_0$.

- Dersom $\zeta = 0$, vil responsen bli en stående svingning med frekvens lik den **udempede resonansfrekvensen** ω_0 .

Ved skissering av tidsresponsen kan det være hensiktsmessig å kjenne de to tidspunktene som fremgår av figur 4.19, nemlig t_1 som angir første krysning av den stasjonære verdien og t_2 som angir første toppunkt.

Figur 4.19
Sentrale punkter ved skissering av tidsresponser



Det kan vises at

$$\omega_0 t_1 = \frac{\frac{\pi}{2} + \varphi}{\sin\left(\frac{\pi}{2} + \varphi\right)} \quad (4.119)$$

¹Dersom $\zeta < 0$, (noe som ikke er mulig for et masse-fjær system fordi det forutsetter en negativ "fjærkonstant") har vi et *ustabilt* system. Da ligger polene i *høyre* halvplan. Et eksempel på et ustabilt system behandles første gang i eksempel 4.14.

og

$$\omega_0 t_2 = \frac{\pi}{\sin\left(\frac{\pi}{2} + \varphi\right)} \quad (4.120)$$

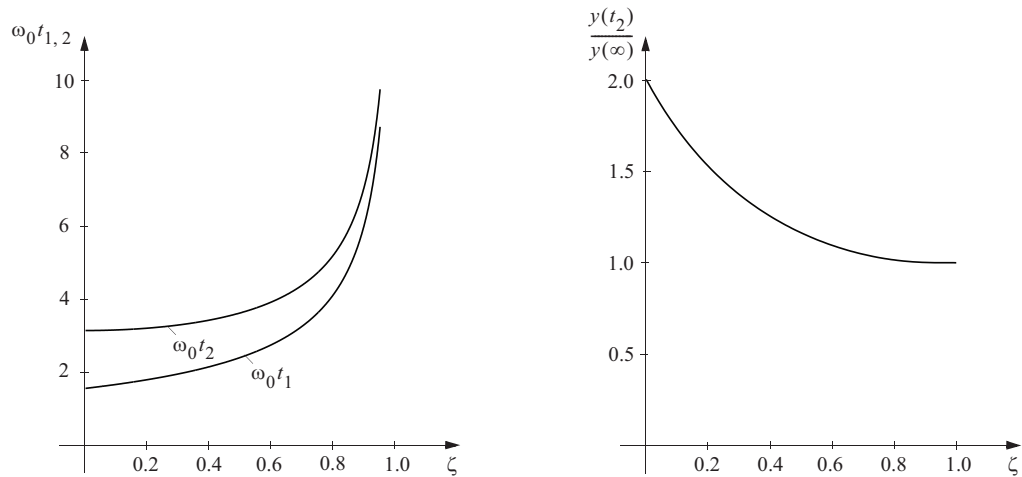
Innfører vi $\sin \varphi = \zeta$, kan innholdet av (4.119) og (4.120) angis grafisk som vist i figur 4.20.

Det kan også være nyttig å ha et uttrykk for hvor høy den første toppen er relativt til den stasjonære verdien av responsen. Dette forholdet kan vises å bli

$$\frac{y(t_2)}{y(\infty)} = 1 + e^{-\frac{\zeta}{\sqrt{1-\zeta^2}}\pi} \quad (4.121)$$

Det er også skissert i figur 4.20.

Figur 4.20
Grafisk framstilling
av sentrale trekk
ved tidsresponsen



■
Grunnen til at det er brukt såpass mye plass til å behandle dette eksemplet, er at det spiller en betydelig rolle i praksis, og at responsen til mer kompliserte lineære systemer kan betraktes som summen av responsene til et antall 1. ordens systemer og et antall 2. ordens systemer. Og alle 2. ordens systemer har en dynamikk som i hovedtrekkene er lik den som dette eksemplet oppviser.

Tabell 4.2 viser transferfunksjonen, nullpunkter, poler og sprangrespons for en del ofte forekommende systemer av 1. og 2. orden – siste kategori under den tykke skillelinjen.

TABELL 4.2 Noen vanlig forekommende transferfunksjoner

Transferfunksjon $h(s)$	Nullpunkter og poler	Sprangrespons
K		
$K \frac{1}{Ts}$		
$K \frac{1+Ts}{Ts}$		
$K \frac{1}{1+Ts}$		
$K \frac{Ts}{1+Ts}$		
$K \frac{1+T_2s}{1+T_1s}$ $T_2 > T_1$		
$K \frac{1+T_2s}{1+T_1s}$ $T_2 < T_1$		
$\frac{K}{1 + \left(\frac{s}{\omega_0}\right)^2}$		
$\frac{K}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$		
$\frac{Ks}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$		
$\frac{K(1+Ts)}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$		

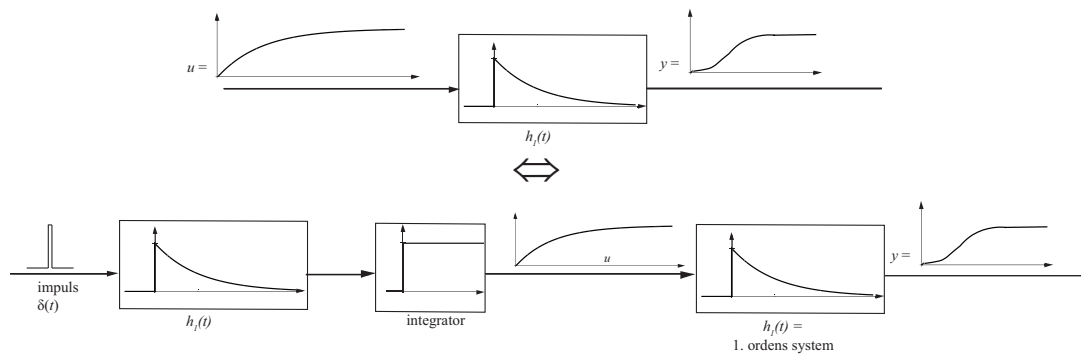
4.7.3 Signaler og systemer

Dette avsnittet har til hensikt å vise at det matematisk sett ikke er noe skille mellom et signal og et lineært system. Et lineært system kan defineres ved hjelp av dets impulsrespons, som jo er et tidsforløp eller et signal. Mens et signal kan betraktes som utgangen av et system eksitert av en impuls $\delta(t)$. La oss illustrere dette ved å sammenlikne de to eksemplene 4.3 og 4.9 - tilfelle 2. De har samme respons bortsett fra den trivielle multiplikative faktoren $1/k$ i eksempel 4.9, som ikke er tatt med her,

$$y(t) = \left(1 - \left(1 + \frac{t}{T}\right) e^{-\frac{t}{T}}\right). \quad (4.122)$$

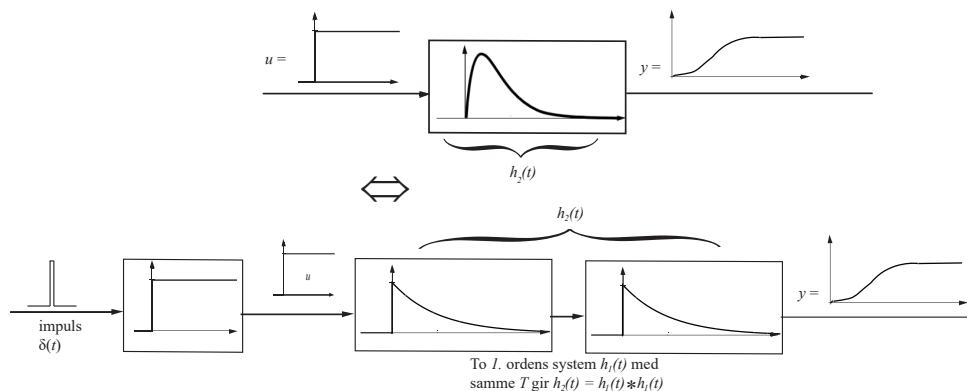
Hvis vi nå karakteriserer et system med sin impulsrespons, så kan eksempel 4.3 fremstilles slik:

Figur 4.21
Eksempel 4.3.



Eksempel 4.9 - tilfelle 2 blir slik:

Figur 4.22
Eksempel 4.9 -
tilfelle 2.



Vi ser at begge eksempler kan oppfattes som generert av $\delta(t)$. Legg merke til at impulsresponsen til en integrator er et enhetssprang. Hvis vi kaller impulsresponsen til en integrator

$h_i(t)$, så gir eksempel 4.3

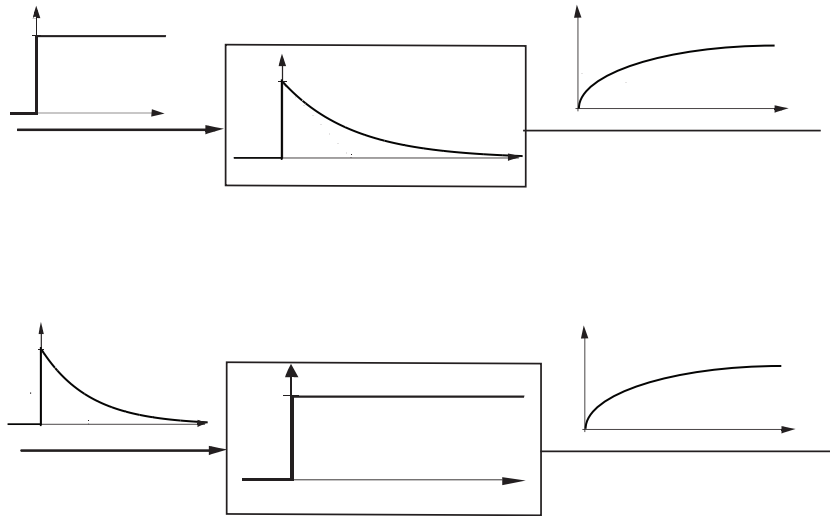
$$y(t) = h_1(t) \star h_i(t) \star h_1(t) \star \delta(t), \quad (4.123)$$

og for eksempel 4.9 - tilfelle 2:

$$y(t) = h_1(t) \star h_1(t) \star h_i(t) \star \delta(t). \quad (4.124)$$

Konvolusjon (folding) er kommutativ, dvs rekkefølgen er likegyldig. Derfor er (4.123) og (4.124) like. Denne egenskapen kan benyttes ytterligere til å markere at signaler og systemer kan beskrives om hverandre. Betrakt følgende to eksempler:

Figur 4.23
Bytte av rekkefølge
til signal og system



Oppgave: Vis hvordan et sinussignal

$$y(t) = \begin{cases} \sin(t), & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (4.125)$$

kan oppfattes som utgangen av et system eksitert av en impuls $\delta(t)$. Hva blir systemets transferfunksjon?

4.7.4 Modellreduksjon

Til nå har vi studert polenes plassering og sammenhørende tidsrespons som funksjon av parametrene i en prosess – vi har ikke sett på prosesser i *et reguleringsystem*. Vi skal i neste avsnitt studere eksempler på hvordan poler/nullpunkter i et reguleringsystem forflytter seg når regulatoren endres. Disse eksemplene vil alle dreie seg om likestrømsmotoren. Men først skal vi introdusere begrepet **modellreduksjon**. Hvis man har en modell av en prosess, og

erstatter denne med en enklere modell som likevel (for våre formål) ivaretar de vesentligste dynamiske egenskapene til prosessen, så kalles dette modellreduksjon. Vi forklarer dette med to eksempler basert på motormodellen vi utviklet i eksempel 2.8.

EKSEMPEL 4.10: Likestrømsmotor - reduksjon av modellens orden

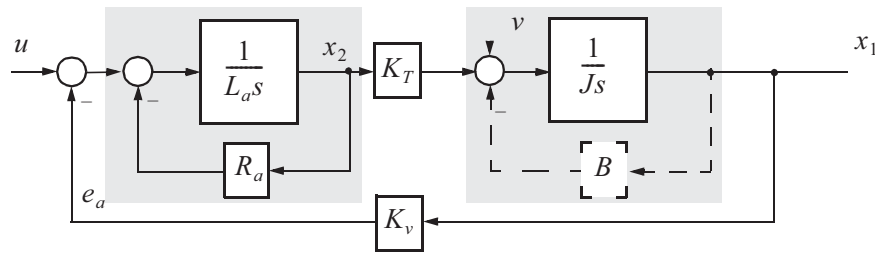
Likningene (2.69) - (2.70) for likestrømsmotoren var:

$$\dot{x}_1 = \frac{1}{J}(-Bx_1 + K_T x_2 + v) \quad (4.126)$$

$$\dot{x}_2 = \frac{1}{L_a}(-K_v x_1 - R_a x_2 + u) \quad (4.127)$$

Blokkdiagrammet for likestrømsmotoren blir ut fra dette

Figur 4.24
Blokkdiagram for
likestrømsmotor



Vi ser i det følgende bort fra forstyrrelsen v . De elektriske og mekaniske “undersystemene” er antydnet med skraverte rektangler. Vi antar ubetydelig friksjon, $B = 0$; dette er *ett* trinn i en modellforenkling. Men vi ønsker nå å gå mer drastisk til verks – å redusere modellens orden, det vil si finne en 1. ordens tilnærming som representerer motoren på en akseptabel måte.

Vi finner først transferfunksjonen fra u til x_1 ved reduksjon av blokkdiagrammet:

$$\frac{x_1}{u}(s) = h_u(s) = \frac{K_T}{(R_a + L_a s)Js} = \frac{1}{K_v \left(1 + \frac{JR_a}{K_v K_T} s + \frac{JL_a}{K_v K_T} s^2 \right)} \quad (4.128)$$

Hvis L_a/R_a er liten, kan vi se bort fra leddet foran s^2 , og vi får 1. ordens-modellen

$$h_u(s) = \frac{1}{K_v \left(1 + \frac{JR_a}{K_v K_T} s \right)} = \frac{1}{K_v (1 + T_m s)} \quad (4.129)$$

Der T_m kalles motorens **mekaniske tidskonstant**. En slik modellreduksjon er ofte akseptabel når motoren inngår som en komponent i et større system som har langsommere dynamikk enn motorens ankerkrets.

Vi kunne også ha gjennomført modellreduksjonen ved å bruke tilstandsrommodellen (4.126) - (4.127). Det elektriske undersystemets dynamikk er gitt av (4.127), som kan formuleres som

$$T_a \dot{x}_2 = -x_2 + \frac{1}{R_a}(-K_v x_1 + u), \text{ hvor vi nå har innført } T_a = \frac{L_a}{R_a} \quad (4.130)$$

T_a er motorens **elektriske tidskonstant**. Når T_a er så liten¹ at den kan forutsettes lik null, går (4.130) over i en algebraisk likning, $x_2 = (-K_v x_1 + u)/R_a$. Ved å sette dette inn i (4.126), kan vi eliminere x_2 . Vi står da igjen med en 1. ordens differensiallikning i x_1 . Når vi også tar med forenklingen $B = 0$, blir denne

$$\dot{x}_1 = -\frac{K_v K_T}{J R_a} x_1 + \frac{K_T}{J R_a} u = \frac{1}{T_m} \left(-x_1 + \frac{1}{K_v} u \right) \quad (4.131)$$

som svarer til transferfunksjonen (4.129).

Virkningen av en liten T_a kan også tolkes i blokkdiagrammet i figur 4.24: Det elektriske undersystemets transferfunksjon

$$\frac{1}{R_a + L_a s} \quad (4.132)$$

erstattes² med konstanten $1/R_a$. Ved så å redusere blokkdiagrammet med dette som utgangspunkt, får man (4.129).



¹I motoranvendelser kan T_a være i størrelsesorden 10 - 1000 ganger mindre enn T_m

²Dette er i seg sjøl et enkelt eksempel på modellreduksjon – en hurtig 1. ordens modell, (4.132), forenkles til en "0. ordens modell", dvs. en konstant. En konstant er det mest trivielle tilfelle av en momentan (statisk) prosess, se avsnitt 1.4.1.

EKSEMPEL 4.11: Likestrømsmotor - samme orden, men forenklet struktur

Noen ganger kan vi ønske å forenkle en modells struktur, uten å redusere dens orden. I *frekvensanalysen* (kapittel 6 og utover) kan det være hensiktsmessig å eliminere “ytre” tilbakekoplinger, slik som vi har via K_v i vårt eksempel; se blokkdiagrammet i figur 4.24. Dermed framstår prosessen som en enklere “seriekopling” av lavere ordens del-prosesser.

Det eksakte uttrykket for transferfunksjonen var (4.128)

$$h_u(s) = \frac{1}{K_v \left(1 + \frac{JR_a}{K_v K_T} s + \frac{JL_a}{K_v K_T} s^2 \right)} \quad (4.133)$$

Vi søker nå å faktorisere nevneren i to 1.ordens-ledd, ved å benytte at $T_a \ll T_m$, dvs.

$$\frac{L_a}{R_a} \ll \frac{JR_a}{K_v K_T}$$

Da kan vi tillate oss tilnærmingen

$$h_u(s) \approx \frac{1}{K_v \left(1 + \left[\frac{JR_a}{K_v K_T} + \frac{L_a}{R_a} \right] s + \frac{JL_a}{K_v K_T} s^2 \right)} = \frac{1}{K_v (1 + T_a s)(1 + T_m s)} \quad (4.134)$$

som gir den ønskede “seriekoplings”-struktur. Dette er en mindre grov forenkling av motor-modellen enn i forrige eksempel.

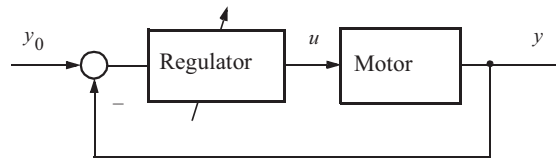
**4.7.5 Forflytning av poler/nullpunkter når regulator endres – rotkurver**

Vi skal nå drøfte hvordan poler/nullpunkter i et reguleringssystem forflytter seg når regulatoren endres. Vi benytter nok en gang likestrømsmotoren. I stedet for motorens *vinkelhastighet* betrakter vi nå dens *vinkelposisjon* som prosessutgang. (Dette er aktuelt f.eks. i robotanvendelser, hvor vinkelposisjoner til motorer i robotens ledd omsettes til posisjoner for ledde-
ne.) Vi kaller vinkelposisjonen y . Transferfunksjonen fra ankerspenning til posisjon blir da (4.128) i serie med en integrator, fordi vinkelposisjon er integralet av vinkelhastighet:

$$\frac{y}{u}(s) = h_u(s) = \frac{1}{K_v} \frac{1}{s \left(1 + \frac{JR_a}{K_v K_T} s + \frac{JL_a}{K_v K_T} s^2 \right)} \quad (4.135)$$

Systemet vi skal betrakte gjennom tre eksempler, er variasjoner over den strukturen som er vist i figur 4.25. Den skrå pila bak blokken “regulator” symboliserer at vi kan justere regulatorens parametre (i motsetning til i prosessen, hvor jo parametrene ligger fast).

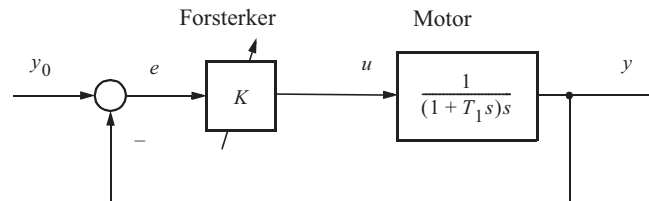
Figur 4.25
Blokkdiagram for motor i reguleringsystem



EKSEMPEL 4.12: System med enkel prosessmodell og P-regulering

Gitt reguleringsystemet i figur 4.26:

Figur 4.26
Blokkdiagram for enkel motormodell med P-regulator



Her brukes den enkleste type regulator – en **proporsjonalregulator** (“P-regulator”). Den består av en spenningsforsterker som leverer en ankerspenning $K \cdot e$ til motoren proporsjonalt med målespenningen e , som uttrykker avviket mellom ønsket og virkelig vinkelposisjon. Vi har benyttet den enkleste modellen (4.129) fra eksempel 4.10. Tidskonstanten T_1 er identisk med den mekaniske tidskonstanten T_m . Vi betrakter for enkelhets skyld den konstante faktoren $1/K_v$ som inkorporert i regulatorens forsterkning K .

Vi finner transferfunksjonen mellom referansen y_0 og målingen y :

$$\begin{aligned} \frac{y}{y_0}(s) &= \frac{\frac{K}{s(1+T_1s)}}{1 + \frac{K}{s(1+T_1s)}} = \frac{1}{\frac{T_1}{K}s^2 + \frac{1}{K}s + 1} = \frac{1}{\frac{T_1}{K}(s - \lambda_1)(s - \lambda_2)} \\ &= \frac{1}{\frac{T_1}{K}(s + \alpha + j\beta)(s + \alpha - j\beta)} = \frac{1}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \end{aligned} \quad (4.136)$$

Vi har fått en 2. ordens transferfunksjon av samme type som i masse-fjær-systemet. Dette systemets dynamiske egenskaper endrer seg når forsterkningen K i regulatoren endres. Vi finner fra nevneren i (4.136) at polene i det **lukkede**¹ system blir

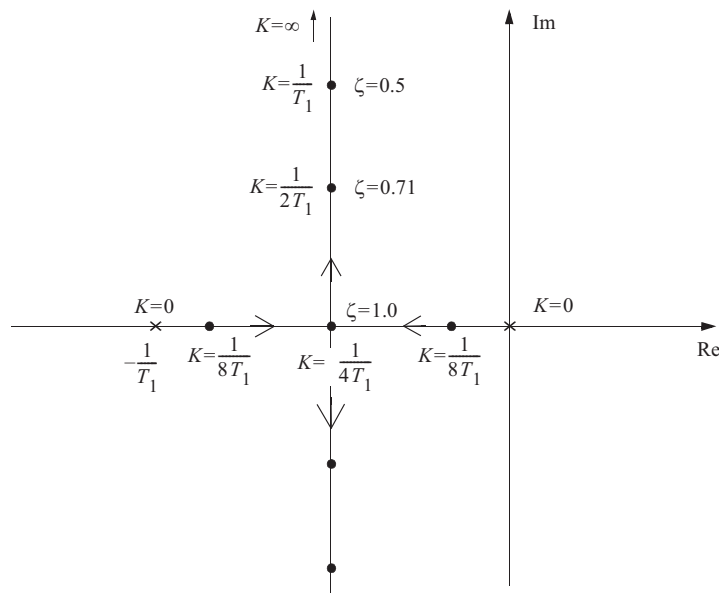
¹Et system sies å være **åpent** når man ikke har tilbakekopling; $K = 0$. Når man har tilbakekopling, dvs. $K > 0$, sier vi at systemet er **lukket** – dette henspiller på at tilbakekopplingsløyfen da er sluttet.

$$\left. \begin{matrix} \lambda_1 \\ \lambda_2 \end{matrix} \right\} = -\frac{1}{2T_1} (1 \pm \sqrt{1 - 4T_1 K}) \quad \text{for } K \leq \frac{1}{4T_1} \quad (4.137)$$

$$\text{og } \alpha = \frac{1}{2T_1}, \quad \beta = \frac{1}{2T_1} \sqrt{4T_1 K - 1} \quad \text{for } K \geq \frac{1}{4T_1} \quad (4.138)$$

Figur 4.27 viser hvordan de to polene beveger seg i det komplekse planet når K endres. For $K = 0$ starter vi med det åpne systemets poler ($\lambda_1 = 0$, $\lambda_2 = -1/T_1$). Når K øker, beveger polene seg imot hverandre langs den negative reelle akse inntil de møtes for $K = 1/(4T_1)$. Økes K ytterligere, beveger polene seg fra hverandre langs en linje parallell med den imaginære akse. Dette betyr at når polene er kompleks konjugerte, vil de som det også framgår av (4.138) ha konstant absolutt dempning $\alpha = 1/(2T_1)$.

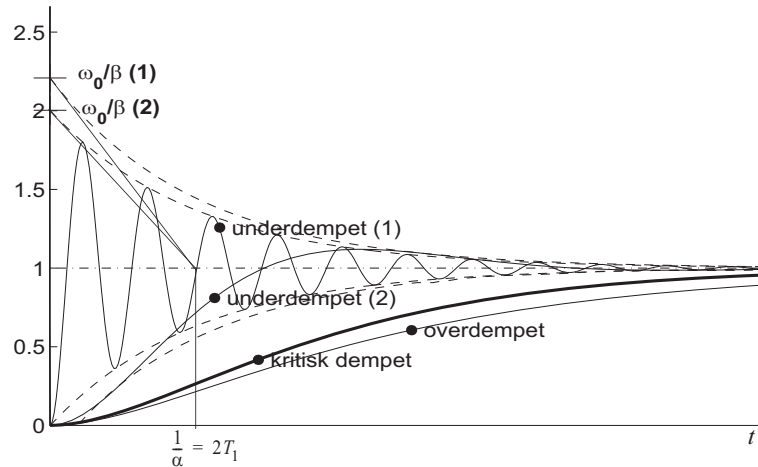
Figur 4.27
Polenes bevegelse
i det komplekse
plan



Polers (og eventuelle nullpunkters) baner i det komplekse plan som funksjon av en endret regulatorparameter (forsterkning) kalles **rotkurver** (engelsk: **Root locus**).

Nå til virkningen på systemets tidsforløp ved økende K : med økende forsterkning slik at polene blir komplekskonjugerte ($\zeta < 1$), vil sprangresponansen som vist i figur 4.28 få samme omhyllingskurve, men stadig stigende svingefrekvens β

Figur 4.28
Sprangresponser



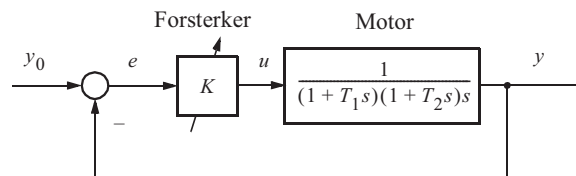
Vi skal i kapittel 9 komme tilbake til spørsmålet om gunstig innstilling av regulatorer, men her bare påpeke at sterke oscillasjoner, som gjerne oppstår med stor verdi av K , er uakseptable i reguleringsystemer. Vi ønsker heller ikke en svært liten verdi av K , som gir langsom innsvingning. En verdi av K som gir $\zeta \approx 0.5$, vil ofte være et hensiktsmessig kompromiss.

■

EKSEMPEL 4.13: System med mer nøyaktig motormodell og P-regulering

Vi benytter nå den mer nøyaktige modellen (4.134) fra eksempel 4.11. Tidskonstanten T_2 er identisk med den elektriske tidskonstanten T_a . Vi betrakter fortsatt den konstante faktoren $1/K_v$ som inkorporert i forsterkningen K . Blokkdiagrammet er vist i figur 4.29

Figur 4.29
Blokkdiagram for mer nøyaktig motormodell med P-regulator



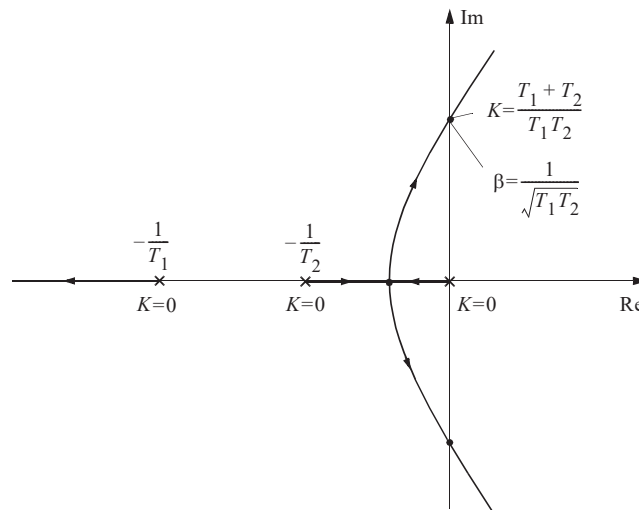
Transferfunksjonen fra referansen y_0 til målingen y blir:

$$\begin{aligned} \frac{y}{y_0}(s) &= \frac{1}{\frac{T_1 T_2}{K} s^3 + \frac{T_1 + T_2}{K} s^2 + \frac{1}{K} s + 1} = \frac{1}{\frac{T_1 T_2}{K} (s - \lambda_1)(s - \lambda_2)(s - \lambda_3)} \\ &= \frac{1}{\frac{T_1 T_2}{K} (s - \lambda_1)(s + \alpha + j\beta)(s + \alpha - j\beta)} \end{aligned} \quad (4.139)$$

Vi ser at vi nå har fått en 3. ordens transferfunksjon. Polene i denne vil enten alle være reelle, eller en reell og to kompleks konjugerte som antydnet i (4.139).

Dersom forsterkningen K endres, vil polene flytte seg i det komplekse planet. Det er imidlertid arbeidskrevende å finne et analytisk uttrykk for polenes verdi avhengig av K , fordi vi må bestemme røttene i et 3. ordens polynom. men dette gjøres enkelt med numerisk beregning i programpakker som MATLAB. Figur 4.30 viser bevegelsen av polene.

Figur 4.30
Polenes bevegelse
i det komplekse
plan



MATLAB-koden som gir figur 4.30 er gjengitt i figur 4.31.

Figur 4.31
MATLAB-kode for
simulering og plot-
ting ved bruk av
tilstandsromform

```

1 T1 = 0.5;
2 T2 = 1;
3 t = 1;
4 n = [T1*T2 (T1+T2) 1 0]; ← nevner = s(1+T1*s)(1+T2*s)
5 rlocus(t,n)

```

Polene vil også i dette tilfellet starte sin bevegelse fra polene for det åpne systemet ($K = 0$). Da er $\lambda_1 = 0$, $\lambda_2 = -1/T_1$ og $\lambda_3 = -1/T_2$. Når K økes, vil den av polene som opprinnelig hadde størst negativ verdi, bevege seg mot venstre, mens de to andre beveger seg mot hverandre langs den reelle akse inntil de møtes. Økes K ytterligere fra denne verdien, vil den første polen fortsette mot venstre, mens de to andre vil bli komplekse og bevege seg fra hverandre. De to komplekse polene vil til å begynne med bevege seg parallelt med den imaginære akse.

Deretter vil de bøye av mot høyre og bevege seg mot høyre halvplan i stedet for å fortsette parallelt med den imaginære akse som i det forrige tilfellet. Ved verdien

$$K = \frac{T_1 + T_2}{T_1 T_2}$$

vil polene krysse gjennom den imaginære akse i et punkt karakterisert ved

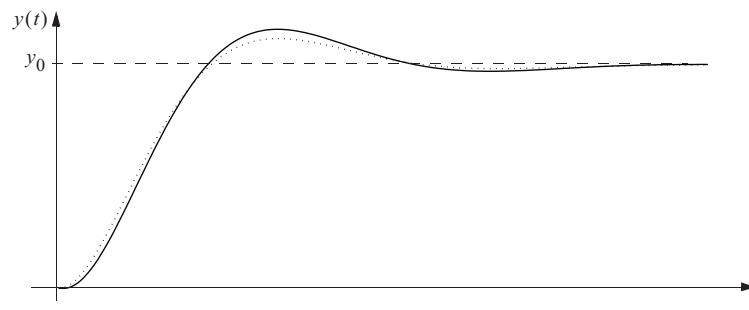
$$\beta = \frac{1}{\sqrt{T_1 T_2}}$$

Vi skal ikke her vise hvordan disse tallverdiene framkommer, men henviser til avsnitt 8.3 der en utledning foreligger.

Når de komplekse polene befinner seg i venstre halvplan, vil den absolutte dempningen være positiv, hvilket medfører at svingningene vil dempes ut. Når polene ligger i høyre halvplan, vil den absolutte dempningen være negativ. Dette betyr at svingningene vil øke i amplitude. Da er systemet *ustabilt*.

Figur 4.32 viser sprangresponsen for systemet heltrukket, mens sprangresponsen fra forrige eksempel (med kun to poler) er vist prikket. Innflytelsen av den tredje polen vil være en liten forskyvning av sprangresponsen mot høyre (forsinkelse).

Figur 4.32
Sprangrespons



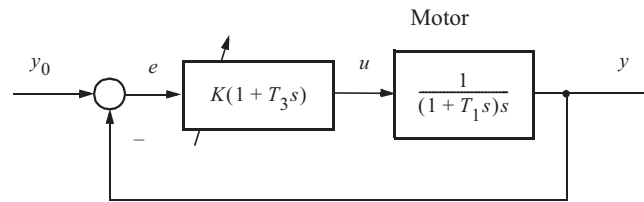
■

EKSEMPEL 4.14: System med enkel prosessmodell og PD-regulering

Vi går tilbake til den forenklede prosessbeskrivelsen som ble brukt i eksempel 4.12. Vi modifierer forsterkeren slik at den får transferfunksjonen $K(1 + T_3s)$. En slik type regulator vil bli grundig behandlet i kapittel 9. Den kalles en **proporsjonal+derivatregulator (PD-regulator)**. Med slik regulator får vi et blokkdiagram som vist i figur 4.33.

Figur 4.33

Blokkdiagram for enkel motormodell med PD-regulator



Transferfunksjonen fra y_0 til y blir

$$\begin{aligned} \frac{y}{y_0}(s) &= \frac{K(1 + T_3s)}{T_1s^2 + (1 + KT_3)s + K} = \frac{KT_3}{T_1} \frac{s + \frac{1}{T_3}}{(s - \lambda_1)(s - \lambda_2)} \\ &= \frac{KT_3}{T_1} \frac{s + \frac{1}{T_3}}{(s + \alpha + j\beta)(s + \alpha - j\beta)} \end{aligned} \quad (4.140)$$

Vi får igjen et 2. ordens system, men i tillegg til de to polene har det nå også et *nullpunkt*. Nullpunktet ligger fast i posisjonen $-1/T_3$, mens polenes plassering avhenger av forsterkingen K . Vi finner

$$\left. \begin{array}{l} \lambda_1 \\ \lambda_2 \end{array} \right\} = -\frac{1 + T_3K}{2T_1} \left(1 \pm \sqrt{1 - \frac{4T_1K}{(1 + T_3K)^2}} \right) \quad \text{for} \quad \frac{4T_1K}{(1 + T_3K)^2} \leq 1 \quad (4.141)$$

og

$$\alpha = \frac{1 + T_3K}{2T_1}, \quad \beta = \frac{1 + T_3K}{2T_1} \sqrt{\frac{4T_1K}{(1 + T_3K)^2} - 1} \quad \text{for} \quad \frac{4T_1K}{(1 + T_3K)^2} \geq 1$$

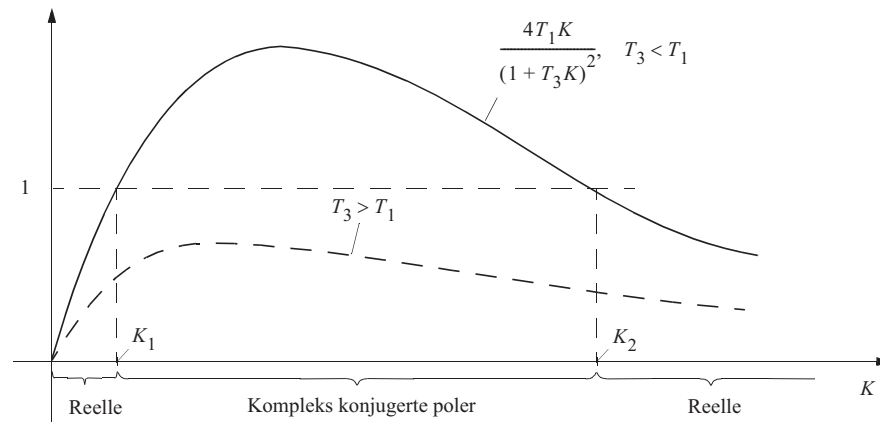
Vi ser at sammenhengen mellom K og plasseringen av polene nå er noe mer sammensatt.

Det kan være nyttig å se på en grafisk framstilling som vist i figur 4.34, av sammenhengen mellom K og leddet fra likning (4.141),

$$\frac{4T_1K}{(1 + T_3K)^2}$$

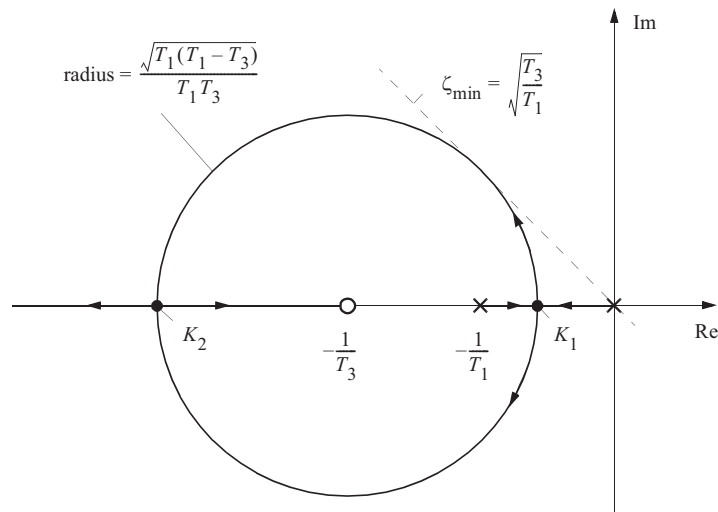
Den heltrukne kurven som gjelder for $T_3 < T_1$, er av spesiell interesse. Dersom $0 < K < K_1$ eller $K > K_2$, er polene i transferfunksjonen reelle og forskjellige, og når $K_1 < K < K_2$, er polene kompleks konjugerte. For $K = K_1$ og $K = K_2$ er polene reelle og like.

Figur 4.34
Sammenhengen
mellom K og
polene



Figur 4.35 viser hvordan polene flytter seg når K varierer fra null mot uendelig. De begynner som i de to foregående eksemplene i verdier som er polene for det "åpne" systemet ($\lambda_1 = 0$ og $\lambda_2 = -1/T_1$). Økes K , går polene først mot hverandre langs den reelle aksen inntil de møtes for verdien $K = K_1$. Figur 4.35 viser at dersom K økes ytterligere, vil polene bli komplekse og bevege seg etter en sirkel med sentrum i punktet $-1/T_3$, inntil de igjen møtes lenger til venstre på den reelle aksen når $K = K_2$. Økes K enda mer vil den ene polen fortsette ut til venstre mens den andre går i retning av nullpunktet og til slutt absorberes i dette. Følgende er karakteristisk for dette systemet:

Figur 4.35
Polenes bevegelse
når K beveger seg
fra 0 mot uendelig



Når K øker fra null, vil den relative dempningen ζ av de komplekse polene først avta mot en minimumsverdi gitt av vinkelen for tangenten til sirkelen fra origo, for så igjen å øke. For meget store verdier av K vil den ene polen tilnærmet absorberes av nullpunktet, slik at det resulterende systemet tilnærmet blir et 1. ordens system med en pol langt ute til venstre (liten tidskonstant).

Vi har fått to virkninger av å innføre nullpunktet i transferfunksjonen. For det første har vi fått en forskyvning av polene mot venstre i forhold til det vi hadde i det første av disse eksemplene som vist i figur 4.27. Den absolutte og relative dempningen er dermed blitt større hvilket medfører et system med mindre oscillasjoner for en gitt verdi av K , dvs. bedre regulering.

Vi har også fått en annen type transferfunksjon som gjerne kan skrives på formen

$$\begin{aligned} \frac{y}{y_0}(s) &= \frac{1 + T_3 s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \\ &= \frac{1}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} + \frac{T_3 s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \end{aligned} \quad (4.142)$$

Transferfunksjonen kan betraktes som summen av to deler som gir hvert sitt bidrag til responsen i y

$$y(t) = y_a(t) + y_b(t) \quad (4.143)$$

der faktorene er ordnet i samme rekkefølge som i (4.142). Dette medfører at

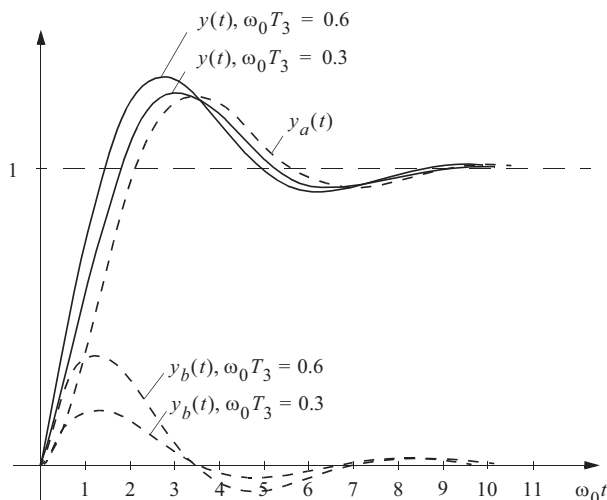
$$y_b(t) = T_3 \frac{dy_a(t)}{dt}$$

Vi kan derfor bestemme totalresponsen $y(t)$ ved først å bestemme $y_a(t)$, slik det ble gjort i eksempel 4.9, og deretter legge til et ledd som er proporsjonalt med den tidsderivate av $y_a(t)$. Dette er illustrert i figur 4.36 for det tilfellet at $y_0(t)$ er et sprang og den relative dempningsfaktor er lik $\zeta = 0.4$. Det er antydning to tilfeller, $\omega_0 T_3 = 0.6$ og $\omega_0 T_3 = 0.3$

Tilsvarende betraktninger kan gjøres gjeldende dersom transferfunksjonen har et polynom av høyere orden enn 1 i telleren (flere enn et nullpunkt). Vi må da generere høyere ordens deriverte av grunnresponsen $y_a(t)$.

Figur 4.36

Å finne tidsrespon-
sen $y(t)$ ved hjelp
av $y_a(t)$



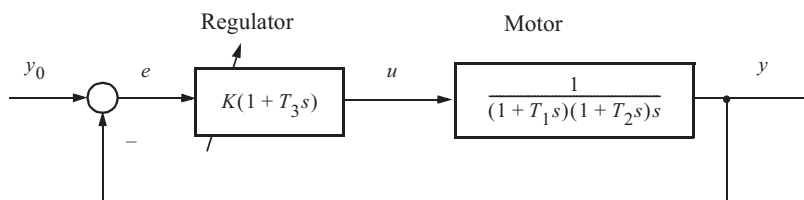
■

EKSEMPEL 4.15: System med mer nøyaktig motormodell og PD-regulering

Vi benytter igjen den mer nøyaktige modellen (4.134) fra eksempel 4.11. Blokkdiagrammet er vist i figur 4.37

Figur 4.37

Blokkdiagram for
mer nøyaktig
motormodell med
PD-regulator



Transferfunksjonen fra y_0 til y blir

$$\begin{aligned} \frac{y}{y_0}(s) &= \frac{K(1 + T_3s)}{T_1T_2s^3 + (T_1 + T_2)s^2 + (1 + KT_3)s + K} = \frac{KT_3}{T_1T_2} \left(s + \frac{1}{T_3} \right) \\ &= \frac{KT_3}{T_1T_2} \frac{s + \frac{1}{T_3}}{(s - \lambda_1)(s - \lambda_2)(s - \lambda_3)} \end{aligned} \quad (4.144)$$

Vi har nå fått et 3. ordens system som har tre poler og et nullpunkt. Nullpunktet ligger fast i posisjonen $-1/T_3$ mens polene er avhengige av verdien til forsterkningen, K .

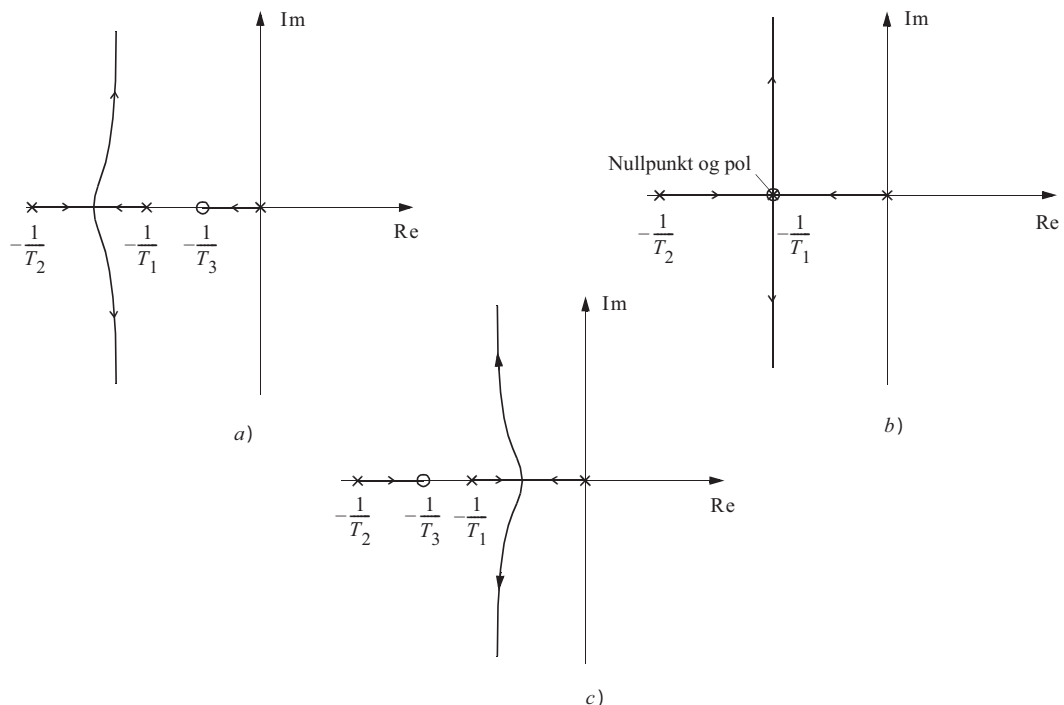
Figur 4.38 viser rotkurvene for tre forskjellige verdier av T_3 . De begynner som i de tre foregående eksemplene i verdier som er polene for det "åpne" systemet, $\lambda_1 = 0$, $\lambda_2 = -1/T_1$ og $\lambda_3 = -1/T_2$.

Figur 4.38 a) viser at når $T_2 < T_1 < T_3$ vil polen som ligger i origo bevege seg inn mot nullpunktet. De to andre polene vil møtes på den reelle aksen og bevege seg over i det komplekse planet. Vi ser at polene bøyer av i den andre retningen i forhold til i figur 4.38 c).

Figur 4.38 b) viser at når $T_2 < T_1 = T_3$ og K økes, går de to polene som ligger på hver sin side av nullpunktet først mot hverandre langs den reelle aksen, inntil de møtes i nullpunktet. Dersom K økes ytterligere, vil to polene bli komplekse og bevege seg langs en rett linje parallell med den imaginære aksen. I dette tilfellet ser vi at leddet $(1 + T_3s)$ i telleren kansellerer leddet $(1 + T_1s)$ i nevneren slik at vi sitter igjen med et 2. ordens system.

Figur 4.38 c) viser at når $T_2 < T_3 < T_1$, vil polen som ligger i $-1/T_2$ bevege seg inn mot nullpunktet. De andre to polene vil møtes på den reelle aksen og bevege seg over i det komplekse planet.

Figur 4.38
Rotkurver for
a) $T_2 < T_1 < T_3$
b) $T_2 < T_1 = T_3$
c) $T_2 < T_3 < T_1$



Vi kan nå gjøre noen generelle betraktninger om rotkurvenes forløp: Om vi tenker oss at vi beskuer det komplekse planet fra stor avstand, vil det åpne systemets poler og nullpunkter alle ligge i en “klynge” nær origo, og det er her polene starter når $K = 0$. Når K øker mot ∞ , vil rotkurvene nærme seg asymptoter som er forskjellige, avhengig av antallet poler N_p og nullpunkter N_n .

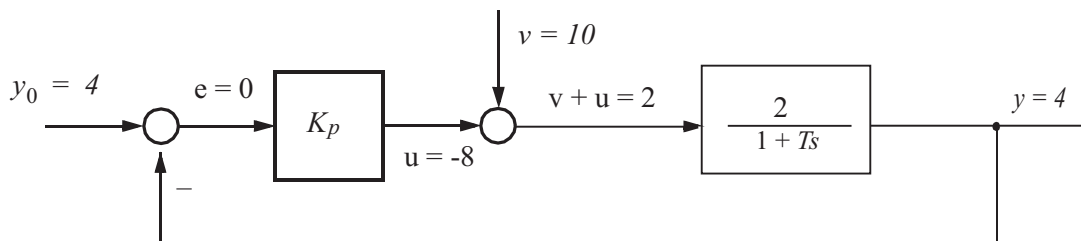
1. Med $N_p - N_n = 1$ får vi en asymptote som faller sammen med den negative reelle aksene.
2. Med $N_p - N_n = 2$ får vi to vertikale asymptoter som går parallelt med den imaginære aksene (asymptotene deler planet i to like deler).
3. Med $N_p - N_n = 3$ får vi tre asymptoter, en som faller sammen med den negative reelle aksene og to som stråler ut mot høyre med vinkler $\pm 60^\circ$ (asymptotene deler planet i tre like deler).
4. Med $N_p - N_n = 4$ får vi fire asymptoter som stråler ut med vinkler $\pm 45^\circ$ og $\pm 135^\circ$ (asymptotene deler planet i fire like deler).

4.7.6 Intuitiv introduksjon til PI-regulator

Til nå har vi bare sett på systemer med P-regulator (altså en regulator som er en rein forsterkning K_p). Vi skal nå introdusere Proporsjonal pluss Integralregulatoren (PI-regulatoren), som vi kommer grundig tilbake til seinere. Men vi introduserer den her fordi den har en svært nyttig egenskap som proporsjonalregulatoren (P-regulatoren) ikke har, og som kan forstås intuitivt uten matematiske utlegninger.

Se figur 4.39. Prosessen $\frac{2}{1+Ts}$ skal først reguleres med en proporsjonalregulator med forsterkning K_p . Prosessen utsettes for en forstyrrelse v som er et sprang med amplitude = 10. Referansen er et sprang med amplitude = 4. Vi tenker oss at det har gått så mye tid at alle variable har nådd konstante verdier, og disse er indikert i figuren. Utgangen y er nå blitt lik referansen. Summen av pådrag og forstyrrelse er $u + v = 2$, og med 2 i telleren i prosessens transfers-funksjon (prosessen har statisk forsterkning = 2), får vi at utgangen må bli 4 når de variable har nådd konstante verdier.

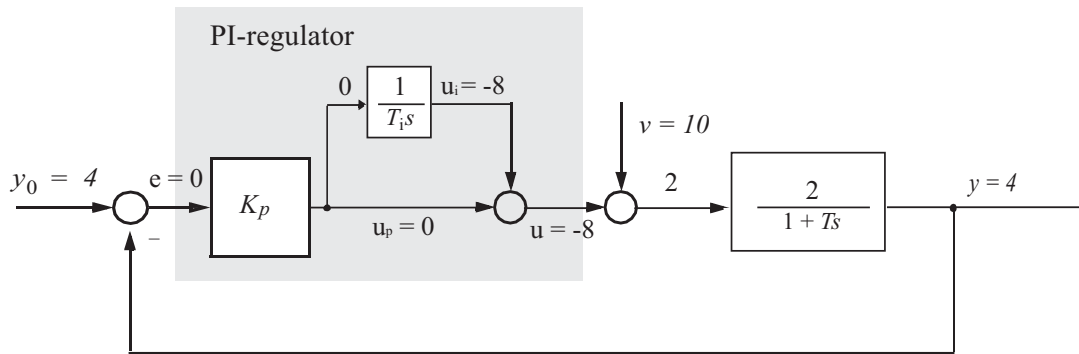
Figur 4.39
P-regulator med $e = 0$, noe som i praksis er umulig.



Men det settet av konstante verdier som er indikert i figuren kan ikke oppnås, for $u = K_p e$ kan ikke være $\neq 0$ når $e = 0$. For å i det hele tatt få ut en $u \neq 0$ må $e \neq 0$, dvs. vi ender opp med et konstant og uønsket avvik.

Derfor prøver vi med en ny struktur:

Figur 4.40



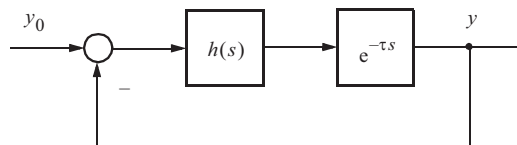
Den øvre greina med integratoren vil gi fra seg et konstant signal $\neq 0$ selv om $e = 0$. Nå er det mulig å motvirke en konstant $v \neq 0$ perfekt. Regulatoren er av type PI:

$$h_r(s) = K_p \left(1 + \frac{1}{T_i s} \right) = K_p \frac{1 + T_i s}{T_i s} \quad (4.145)$$

4.7.7 Transferfunksjoner med tidsforsinkelse

Alle eksemplene i avsnitt 4.7 tok for seg systemer som kunne beskrives med ordinære differensiallikninger. Vi skal nå se på et tilfelle der transferfunksjonen ikke er en rasjonal funksjon, og der bruk av ordinær invers Laplacetransformasjon gir et uoversiktlig resultat. Systemet er vist i figur 4.41.

Figur 4.41
Blokkdiagram



Det forestiller et monovariabelt system med negativ tilbakekopling, der det inni tilbakekoplingen finnes to transferfunksjoner i serie. Den ene er en generell rasjonal funksjon $h(s)$ hvor vi tenker oss at regulatoren er innlemmet, og den andre er transferfunksjonen for en **transportforsinkelse (tidsforsinkelse)**, $e^{-\tau s}$. Denne situasjonen forekommer i mange praktiske tilfeller. Reduserer vi blokkdiagrammet, finner vi

$$\frac{y}{y_0}(s) = \frac{h(s)e^{-\tau s}}{1 + h(s)e^{-\tau s}} \quad (4.146)$$

På grunn av leddet $e^{-\tau s}$ i nevneren, kan vi ikke finne en algebraisk løsning. Men (4.146) kan rekkeutvikles ved å dividere telleren med nevneren, hvilket gir resultatet

$$\frac{y}{y_0}(s) = h(s)e^{-\tau s} - h^2(s)e^{-2\tau s} + h^3(s)e^{-3\tau s} - \dots - (-1)^i h^i(s)e^{-i\tau s} + \dots \quad (4.147)$$

La oss først undersøke det tilfellet at $h(s) = K$ og referansen $y_0(s)$ er et enhetssprang ved tiden $t = 0$, som gir $y_0(s) = 1/s$. Dette gir

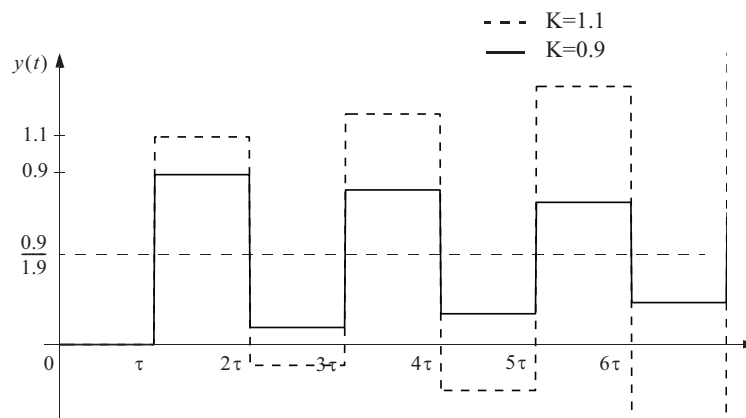
$$\begin{aligned} y(s) &= \frac{1}{s} (Ke^{-\tau s} - K^2e^{-2\tau s} + K^3e^{-3\tau s} - \dots - (-1)^i K^i e^{-i\tau s} + \dots) \\ &= \frac{K}{s} e^{-\tau s} - \frac{K^2}{s} e^{-2\tau s} + \frac{K^3}{s} e^{-3\tau s} - \dots - \frac{(-1)^i K^i}{s} e^{-i\tau s} + \dots \end{aligned} \quad (4.148)$$

Ved bruk av regel 5 i appendiks B får vi følgende resultat ved invers Laplacetransformasjon av (4.148)

$$y(t) = K\mu(t - \tau) - K^2\mu(t - 2\tau) + K^3\mu(t - 3\tau) - \dots - (-1)^i K^i \mu(t - i\tau) + \dots$$

Figur 4.42 viser $y(t)$ når $h(s) = K = 0.9$ og $h(s) = K = 1.1$. Vi kan tenke oss at dette eksemplet representerer et reguleringsystem der $h(s)$ er en såkalt *proporsjonalregulator* med en forsterkning som er lik 0.9 eller 1.1, og der prosessen er en ren transportforsinkelse. Det finnes mange slike prosesser (for eksempel massetransport i et rør). Figur 4.42 viser at kurven går mot den stasjonære verdien $0.9/(1 + 0.9)$, når tiden går mot uendelig og $K = 0.9$. Denne verdien er funnet ved hjelp av (4.146) og sluttverditteoremet (4.22).

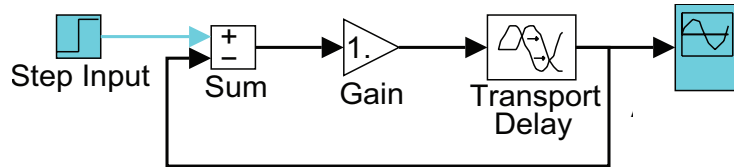
Figur 4.42
Sprangrespons når
 $h(s) = 0.9$ og
 $h(s) = 1.1$



Med $K = 1.1$ ser vi at den firkantede oscillasjonen øker i amplitude. Systemet er nå blitt *ustabilt*. Stabilitet skal vi studere nærmere i kapittel 8.

Figur 4.42 er generert ved hjelp av **Simulink**. Simulink kjøres under MATLAB og har et grafisk brukergrensesnitt. Figur 4.43 viser hvordan systemet ser ut i Simulink.

Figur 4.43
Simulink-diagram



Nå skal vi i stedet benytte en såkalt **integralregulator** i systemet i figur 4.41.

$$h(s) = \frac{1}{T_i s} \quad (4.149)$$

Referansen er fortsatt et enhetsprang ved tiden $t = 0$. Vi innfører dette i (4.147)

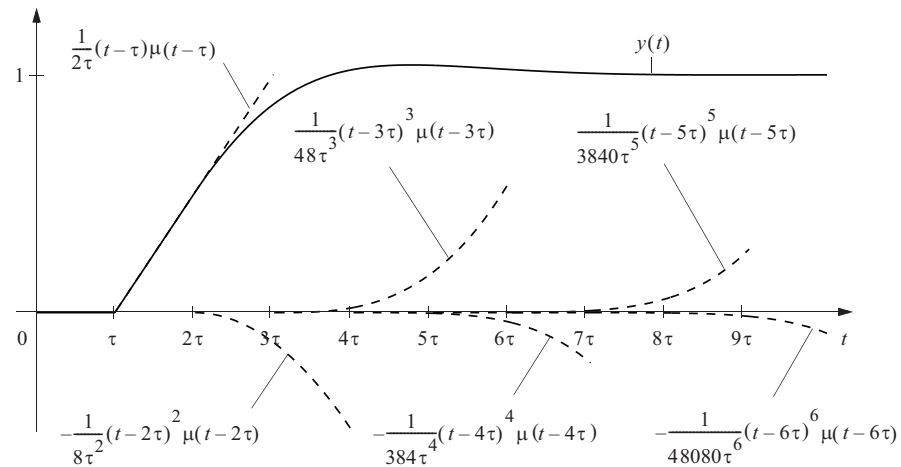
$$\begin{aligned} y(s) &= \frac{1}{s} \left(\frac{1}{T_i s} e^{-\tau s} - \left(\frac{1}{T_i s} \right)^2 e^{-2\tau s} + \left(\frac{1}{T_i s} \right)^3 e^{-3\tau s} - \dots - (-1)^j \left(\frac{1}{T_i s} \right)^j e^{-j\tau s} + \dots \right) \\ &= \frac{1}{T_i s^2} e^{-\tau s} - \frac{1}{T_i^2 s^3} e^{-2\tau s} + \frac{1}{T_i^3 s^4} e^{-3\tau s} - \dots - (-1)^j \frac{1}{T_i^j s^{j+1}} e^{-j\tau s} + \dots \end{aligned} \quad (4.150)$$

Ved bruk av regel 4 og 5 i appendiks B får vi følgende resultat ved invers Laplacetransformasjon av (4.150)

$$\begin{aligned} y(t) &= \frac{(t - \tau)}{T_i} \mu(t - \tau) - \frac{(t - 2\tau)^2}{2! T_i^2} \mu(t - 2\tau) + \frac{(t - 3\tau)^3}{3! T_i^3} \mu(t - 3\tau) - \dots \\ &\quad - (-1)^j \frac{(t - j\tau)^j}{j! T_i^j} \mu(t - j\tau) + \dots \end{aligned} \quad (4.151)$$

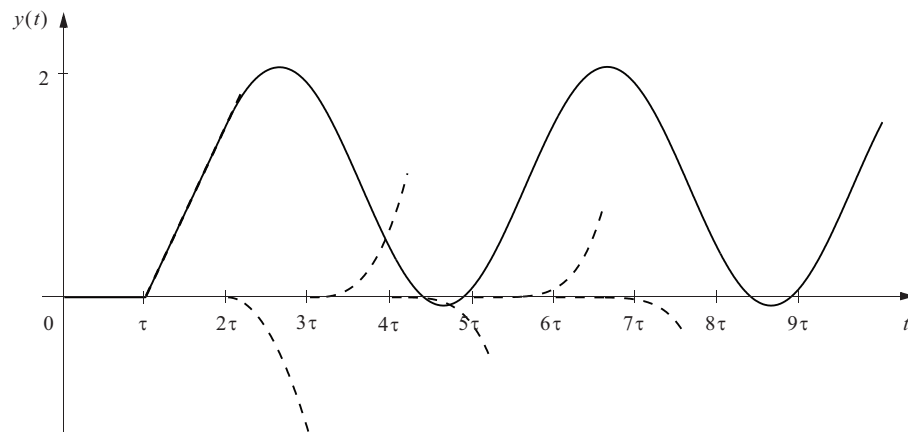
Figur 4.44 viser responsen $y(t)$ når $T_i = 2\tau$ (som siden skal vise seg å være en rimelig dimensjoneringsverdi). De stiplede kurvene viser de enkelte leddene i (4.151).

Figur 4.44
Sprangrespons når
 $h(s) = \frac{1}{2\tau s}$



Hvis vi senker T_i til $T_i = 2\tau/\pi$, som skal vise seg å være grensen for stabilitet (se kapittel 8), får vi sprangresponsen som er vist i figur 4.45.

Figur 4.45
Sprangrespons når
 $h(s) = \frac{\tau}{2\tau s}$



Transferfunksjonen $e^{-\tau s}$ kan erstattes med en rasjonal approksimasjon. Denne angrepsmåten skal studeres nærmere i forbindelse med stabilitetsanalyse i kapittel 8, men det er nyttig å møte problemstillingen allerede nå.

Den enkleste rasjonale approksimasjonen til en **transportforsinkelse** er

$$e^{-\tau s} = \frac{e^{-\frac{\tau}{2}s}}{e^{+\frac{\tau}{2}s}} \cong \frac{1 - \frac{\tau}{2}s}{1 + \frac{\tau}{2}s} \quad (4.152)$$

der vi i teller og nevner bare har tatt med førstegradsleddet i en Taylor-rekke. Dette kalles en **1. ordens Padé-approksimasjon**.

Vi setter nå inn (4.152) i nevneren (ikke i telleren, se forklaring nedenfor) i (4.146) og får, for det tilfellet at $h(s) = 1/(T_i s)$,

$$\frac{y}{y_0}(s) \approx \frac{h(s)e^{-\tau s}}{1 + h(s) \left(\frac{1 - \frac{\tau}{2}s}{\frac{1}{\tau}} \right)} = \frac{\left(1 + \frac{\tau}{2}s\right) e^{-\tau s}}{1 + \left(T_i - \frac{\tau}{2}\right)s + T_i \frac{\tau}{2}s^2} = \frac{1 + \frac{\tau}{2}s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} e^{-\tau s} \quad (4.153)$$

der

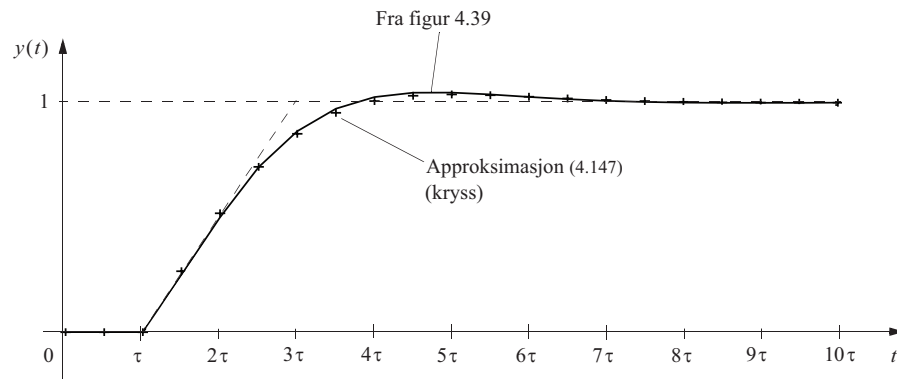
$$\omega_0 = \sqrt{\frac{2}{T_i \tau}} \quad (4.154)$$

og

$$\zeta = \sqrt{\frac{T_i}{2\tau}} - \sqrt{\frac{\tau}{8T_i}} \quad (4.155)$$

Bruker vi som foran $T_i = 2\tau$, får vi $\omega_0 = 1/\tau$ og $\zeta = 0.75$. Figur 4.46 viser sprangresponsen til approksimasjonen (4.153) sammenliknet med det vi fant i figur 4.44. Figuren viser at dette er en meget god approksimasjon til det eksakte svaret som vi har rekkeutviklingen av i (4.147).

Figur 4.46
Sprangrespons
ved rekkeutvikling,
og ved approksi-
masjonen (4.149)



Poenget med å bruke en slik approksimasjon er at man kan finne den algebraiske løsningen av sprangresponsen til

$$\frac{y}{y_0}(s) = \frac{1 + \frac{\tau}{2}s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} e^{-\tau s} \quad (4.156)$$

Dette gjøres ved først å se bort fra forsinkelsen $e^{-\tau s}$, og så forskyve løsningen τ tidsenheter mot høyre etterpå. Vi ser altså at en tidsforsinkelse *i serie* med et rasjonalt uttrykk ikke hindrer oss i å finne en algebraisk løsning, så denne tidsforsinkelsen trenger vi ikke å substituere for. Men i den opprinnelige transferfunksjonen,

$$\frac{y}{y_0}(s) = \frac{h(s)e^{-\tau s}}{1 + h(s)e^{-\tau s}} \quad (4.157)$$

inngår tidsforsinkelsen også i *nevneren*, og den må man substituere bort for å få et rasjonalt uttrykk, som jo er forutsetningen for å finne en algebraisk løsning.

Styrbarhet og observerbarhet

5.1 Innledning 175

5.2 Styrbarhet 176

5.3 Observerbarhet 179

5.4 Styrbarhet og observerbarhet i Laplacetransformerte systemer 182

5.5 Funksjonell styrbarhet (FS) av monovariabel lineære systemer 190

5.1 INNLEDNING

Styrbarhet og *observerbarhet* er to begreper av sentral betydning for dynamiske systemers oppførsel og deres bruk i reguleringsteknisk sammenheng.

Styrbarhet har i korte trekk å gjøre med hvorvidt et system ved hjelp av tilgjengelige pådrag over et endelig tidsintervall kan bringes til en vilkårlig gitt tilstand.

Observerbarhet har å gjøre med om man ved hjelp av tilgjengelige målinger over et endelig tidsintervall kan bestemme systemets starttilstand.

Det finnes en omfattende litteratur om strengt matematiske sider ved disse og beslektede begreper. Vi skal her gi en noe forenklet framstilling av problemstillingene som gjør det mulig å bruke resultatene uten at alle matematiske detaljer behandles. Det henvises også til appendiks C.

Vi vil i det følgende behandle systemer som kan beskrives med likningene

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx}\end{aligned}\tag{5.1}$$

5.2 STYRBARHET

Definisjon: Et system av typen (5.1) karakterisert ved tilstandsvektoren $\mathbf{x}(t)$ er styrtbart dersom det for alle tilstander $\mathbf{x}(t_0)$ og en vilkårlig sluttid $t_1 > t_0$ er mulig å finne en pådragsvektor $\mathbf{u}(t)$ i intervallet t_0 til t_1 som bringer systemet til en vilkårlig valgt tilstand $\mathbf{x}(t_1)$.

Vi skal presentere to forskjellige metoder for å påvise om et system er styrtbart.

Den første metoden forutsetter at systemet har lineært uavhengige egenvektorer. Den er basert på at systemet (5.1) bringes på diagonalform ved hjelp av den lineære transformasjonen $\mathbf{x} = \mathbf{M}\mathbf{q}$.

Da får vi

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{\Lambda}\mathbf{q} + \mathbf{M}^{-1}\mathbf{B}\mathbf{u} = \mathbf{\Lambda}\mathbf{q} + \mathbf{H}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{M}\mathbf{q} = \mathbf{F}\mathbf{q}\end{aligned}\tag{5.2}$$

Vi skriver første linje av (5.2) på komponentform

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{q}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_2 & & & & 0 \\ \cdot & & \cdot & & & 0 \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \lambda_n \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ \cdot \\ q_n \end{bmatrix} + \begin{bmatrix} h_{11} & h_{12} & \cdot & \cdot & h_{1r} \\ h_{21} & h_{22} & \cdot & \cdot & h_{2r} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ h_{n1} & h_{n2} & \cdot & \cdot & h_{nr} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_r \end{bmatrix}\tag{5.3}$$

(5.3) viser at dersom alle elementene h_{ij} i en *rad* av matrisen \mathbf{H} er lik null, vil pådragsvektoren ikke få noen innflytelse på den tilsvarende tilstandsvariabelen q_i . Denne tilstanden er derfor ikke styrtbar.

En nødvendig og tilstrekkelig betingelse for styrbarhet av et system med lineært uavhengige egenvektorer er at ingen rad i matrisen \mathbf{H} har alle elementer lik 0

For systemer med lineært *avhengige* egenvektorer som blant annet opptrer ved sammenfallende egenverdier, vil kriteriet ovenfor ikke være brukbart, fordi \mathbf{M} ikke kan inverteres.

Dessuten er det enklere å bruke et kriterium som i tillegg til ikke å forutsette noe om egenverdiene/vektorene, også bare trenger parametrene i den opprinnelige systembeskrivelsen (5.1).

Vi danner **styrbarhetsmatrisen**

$$\mathbf{Q}_s = [\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}] \quad (5.4)$$

Systemet er styrbart hvis og bare hvis rang $\mathbf{Q}_s = n$

Se appendiks C for bevis av dette.

Rang av en matrise er lik p dersom det finnes minst en underdeterminant av orden p forskjellig fra null, mens alle høyere ordens underdeterminanter er lik null.

EKSEMPEL 5.1: Å avgjøre om et system er styrbart

Som et eksempel på undersøkelse av styrbarhet, velger vi systemet

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u$$

Dette systemet har distinkte egenverdier. Vi finner en egenvektormatrise

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

og dens inverse

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$$

Vi finner

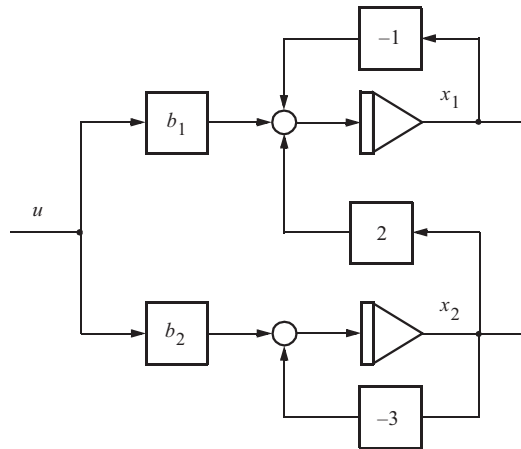
$$\mathbf{H} = \mathbf{M}^{-1}\mathbf{B} = \begin{bmatrix} b_1 + b_2 \\ -b_2 \end{bmatrix}$$

For dette systemet finner vi derfor at betingelsen for styrbarhet blir at

$$b_1 + b_2 \neq 0 \quad \text{og} \quad b_2 \neq 0$$

Systemet er illustrert i figur 5.1. Vi ser umiddelbart at b_2 må være forskjellig fra null for at pådraget i det hele tatt skal påvirke tilstanden x_2 . At vi også må forlange at $b_1 + b_2 \neq 0$, er imidlertid ikke så opplagt. Hvis vi velger akkurat $b_1 = -b_2$, vil den kanoniske (diagonaliser-te) tilstandsvariabelen q_1 ikke være påvirket av pådraget. Systemet er da ikke styrbart.

Figur 5.1
 Blokkdiagram
 for systemet



Vi kan også undersøke systemet i figur 5.1 ved hjelp av styrbarhetsmatrisen som er gitt i (5.4). Vi finner

$$\mathbf{Q}_s = \begin{bmatrix} b_1 & -b_1 + 2b_2 \\ b_2 & -3b_2 \end{bmatrix}$$

Rang $\mathbf{Q}_s = 2$ dersom $\mathbf{Q}_s \neq 0$, og dette oppstår når $(b_1 + b_2)b_2 \neq 0$. Dette gir nøyaktig de samme betingelsene som vi har funnet ovenfor.

■

5.3 OBSERVERBARHET

Definisjon: Et system av typen (5.1) er observerbart dersom det med kjennskap til systemets målevektor \mathbf{y} og pådragsvektor \mathbf{u} i et vilkårlig tidsrom (t_0, t_1) der $t_0 < t_1$, er mulig å bestemme alle komponentene i start-tilstandsvektoren $\mathbf{x}(t_0)$

Vi skal se på to typer kriterier for observerbarhet, slik det også ble gjort for styrbarhet.

Dersom systemet har lineært uavhengige egenvektorer, kan vi som foran utføre en diagonalisering og får en beskrivelse som gitt i (5.2). Vi er særlig interessert i den siste linjen $\mathbf{y} = \mathbf{F}\mathbf{q}$ som på komponentform kan skrives

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \cdot & \cdot & \cdot & f_{1n} \\ f_{21} & f_{22} & \cdot & \cdot & \cdot & f_{2n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ f_{m1} & f_{m2} & \cdot & \cdot & \cdot & f_{mn} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ \cdot \\ q_n \end{bmatrix} \quad (5.5)$$

Vi ser av (5.5) at dersom en (eller flere) av *kolonnene* i matrisen \mathbf{F} har alle elementer $f_{ij} = 0$, vil det ikke være noen informasjon om den (eller de) tilsvarende tilstandsvariabelen q_j i målevektoren.

At det ikke finnes noen kolonne i matrisen \mathbf{F} der alle elementene er lik null, er en nødvendig og tilstrekkelig betingelse for observerbarhet for et system med lineært uavhengige egenvektorer.

I likhet med det som er beskrevet under styrbarhet, er det behov for et kriterium for bestemmelse av observerbarhet dersom systemet ikke har lineært uavhengige egenvektorer. Et slikt kriterium har også den fordel at vi slipper å bringe systemet over på diagonalform.

Beviset for observerbarhetskriteriet er relativt likt det som er gitt for styrbarhetskriteriet i appendiks C. Men her vil vi nøye oss med en skisse til et bevis:

Vi benytter vi oss av (3.22)

$$\mathbf{x}(t_1) = e^{\mathbf{A}(t_1-t_0)}\mathbf{x}(t_0) + \int_{t_0}^{t_1} e^{\mathbf{A}(t_1-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau \quad (5.6)$$

Siden det er $\mathbf{x}(t_0)$ som skal bestemmes, og $\mathbf{u}(\tau)$ etter forutsetningen er kjent, vil integralet på høyresiden representere et kjent ledd. Vi kan derfor uten tap av generell gyldighet gå ut fra at $\mathbf{u}(t) = 0$. Videre kan vi gå ut fra at $t_0 = 0$. Vi har da ved et vilkårlig tidspunkt $t_i > 0$

$$\mathbf{y}(t_i) = \mathbf{C}e^{\mathbf{A}t_i}\mathbf{x}(0) \quad (5.7)$$

Rekkeutvikling gir

$$\mathbf{y}(t_i) = \mathbf{C} \left(\mathbf{I} + \mathbf{A}t_i + \frac{1}{2!}\mathbf{A}^2t_i^2 + \frac{1}{3!}\mathbf{A}^3t_i^3 + \dots \right) \mathbf{x}(0) \quad (5.8)$$

der vi ser at målevektoren ved et vilkårlig tidspunkt t_i er bestemt av en lineær kombinasjon av elementene

$$\mathbf{C}, \mathbf{C}\mathbf{A}, \mathbf{C}\mathbf{A}^2, \dots, \mathbf{C}\mathbf{A}^{n-1}, \mathbf{C}\mathbf{A}^n, \mathbf{C}\mathbf{A}^{n+1}, \dots$$

Ved å anta at \mathbf{y} er tilgjengelig ved forskjellige $t = t_i$, kan vi få forskjellige lineære kombinasjoner. Ved bruk av Cayley-Hamiltons teorem kan vi vise at alle elementer $\mathbf{C}\mathbf{A}^n, \mathbf{C}\mathbf{A}^{n+1}, \dots$ kan uttrykkes som en lineær kombinasjon av lavere ordens ledd $\mathbf{C}\mathbf{A}^k, k \leq n-1$. Dette fører fram til **observerbarhetsmatrisen**

$$\mathbf{Q}_o = [\mathbf{C}^T, (\mathbf{C}\mathbf{A})^T, \dots, (\mathbf{C}\mathbf{A}^{n-1})^T] \quad (5.9)$$

og observerbarhetskriteriet:

Systemet er observerbart hvis og bare hvis rang $\mathbf{Q}_o = n$

EKSEMPEL 5.2: Å avgjøre om et system er observerbart

Som et eksempel på undersøkelse av observerbarhet, kan vi ta det samme systemet som er undersøkt i eksempel 5.1,

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u$$

der målingen er karakterisert ved

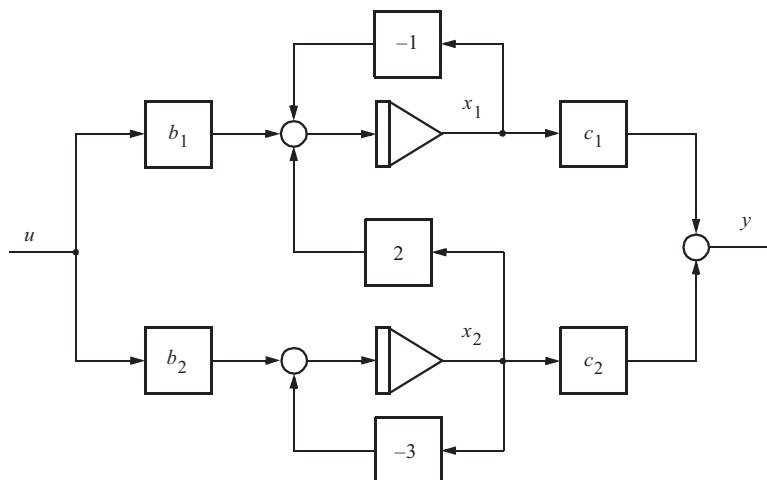
$$y = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \mathbf{x}$$

Vi benytter egenvektormatrisen og dens inverse fra eksempel 5.1 og finner ved innsetting i (5.2)

$$\mathbf{F} = \mathbf{CM} = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} c_1 & c_1 - c_2 \end{bmatrix}$$

Ifølge det første kriteriet for observerbarhet finner vi da at systemet er observerbart så sant $c_1 \neq 0$ og $c_1 - c_2 \neq 0$. Systemet er vist i figur 5.2.

Figur 5.2
Blokdiagram
for systemet



Vi ser med en gang at betingelsen $c_1 \neq 0$ er rimelig, fordi målingen ellers ikke vil inneholde opplysninger om det som foregår i den øvre delen av systemet. Betingelsen $c_1 - c_2 \neq 0$ er imidlertid ikke så opplagt.

Vi kan også benytte oss av kriteriet gitt i (5.9) og finner

$$\mathbf{Q}_o = \begin{bmatrix} c_1 & -c_1 \\ c_2 & 2c_1 - 3c_2 \end{bmatrix}$$

Rang $\mathbf{Q}_o = 2$ dersom $\det \mathbf{Q}_o \neq 0$, og dette får vi dersom $c_1(c_1 - c_2) \neq 0$. Kravet blir dermed det samme som vi har funnet ovenfor.

■

5.4 STYRBARHET OG OBSERVERBARHET I LAPLACETRANSFORMERTE SYSTEMER

I avsnitt 5.2 og 5.3 har vi utledet kriterier for styrbarhet og observerbarhet for lineære systemer med distinkte egenverdier, ved å studere den diagonaliserte tilstandsromformen som er gitt i (5.2)

$$\begin{aligned}\dot{\mathbf{q}} &= \Lambda \mathbf{q} + \mathbf{M}^{-1} \mathbf{B} \mathbf{u} = \Lambda \mathbf{q} + \mathbf{H} \mathbf{u} \\ \mathbf{y} &= \mathbf{C} \mathbf{M} \mathbf{q} = \mathbf{F} \mathbf{q}\end{aligned}\tag{5.10}$$

Vi ordner nå tilstandsvariablene i \mathbf{q} i fire grupper karakterisert ved undervektorer $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4$ slik at

$$\mathbf{q} = [\mathbf{q}_1^T, \mathbf{q}_2^T, \mathbf{q}_3^T, \mathbf{q}_4^T]^T$$

og slik at

- \mathbf{q}_1 representerer de styrbare og observerbare modi
- \mathbf{q}_2 representerer de ikke styrbare, men observerbare modi
- \mathbf{q}_3 representerer de styrbare, men ikke observerbare modi
- \mathbf{q}_4 representerer de ikke styrbare og ikke observerbare modi

Vi kan da skrive (5.10) på følgende form

$$\begin{aligned}\begin{bmatrix} \dot{\mathbf{q}}_1 \\ \dot{\mathbf{q}}_2 \\ \dot{\mathbf{q}}_3 \\ \dot{\mathbf{q}}_4 \end{bmatrix} &= \begin{bmatrix} \Lambda_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Lambda_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Lambda_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \Lambda_4 \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ \mathbf{q}_3 \\ \mathbf{q}_4 \end{bmatrix} + \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \mathbf{H}_3 \\ \mathbf{H}_4 \end{bmatrix} \mathbf{u} \\ \mathbf{y} &= \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 & \mathbf{F}_4 \end{bmatrix} \mathbf{q}\end{aligned}\tag{5.11}$$

der $\mathbf{H}_2 = \mathbf{0}$, $\mathbf{H}_4 = \mathbf{0}$, $\mathbf{F}_3 = \mathbf{0}$ og $\mathbf{F}_4 = \mathbf{0}$. Vi kan lage et blokkdiagram av innholdet i (5.11). Et slikt er vist i figur 5.3 der oppdelingen i fire undersystemer tydelig framgår.

Vi ser også av (5.13) at ordenstallet av transferfunksjonene som inngår i transfermatrisen, nemlig antallet poler, angir antall styrbare og observerbare tilstander. Transferfunksjonenes ordenstall vil alltid være mindre eller høyst lik det totale systemets ordenstall. I tilfelle det er mindre, vil det derfor være en eller flere faktorer i transferfunksjonenes teller og nevner *som kan forkortes bort*. Et eksempel på dette er vist i det følgende.

EKSEMPEL 5.3: Betraktning av transfermatrisen

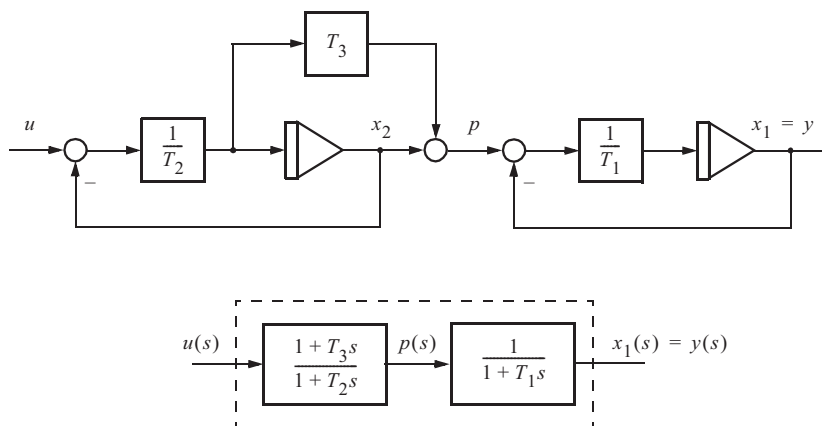
Figur 5.4 viser øverst et elementært blokkdiagram for et 2. ordens monovariabelt system som kan beskrives med en matematisk modell

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u \\ y &= \mathbf{c}^T \mathbf{x}\end{aligned}\tag{5.14}$$

der

$$\mathbf{A} = \begin{bmatrix} 1 & T_2 - T_3 \\ -\frac{1}{T_1} & \frac{T_1 T_2}{1} \\ 0 & -\frac{1}{T_2} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} T_3 \\ \frac{T_1 T_2}{1} \\ \frac{1}{T_2} \end{bmatrix}, \quad \mathbf{c}^T = \begin{bmatrix} 1 & 0 \end{bmatrix}\tag{5.15}$$

Figur 5.4
2. ordens mono-
variabelt system



I dette tilfellet er det lettest å undersøke styrbarhet og observerbarhet ved å finne rang av henholdsvis \mathbf{Q}_s og \mathbf{Q}_o .

Vi har at (se avsnitt 5.2 og 5.3)

$$\mathbf{Q}_s = [\mathbf{B}, \mathbf{AB}] = \begin{bmatrix} \frac{T_3}{T_1 T_2} & \frac{T_1(T_2 - T_3) - T_2 T_3}{T_1^2 T_2^2} \\ \frac{1}{T_2} & -\frac{1}{T_2^2} \end{bmatrix}$$

og

$$\mathbf{Q}_o = [\mathbf{C}^T, \mathbf{A}^T \mathbf{C}^T] = \begin{bmatrix} 1 & -\frac{1}{T_1} \\ 0 & \frac{T_2 - T_3}{T_1 T_2} \end{bmatrix}$$

Vi finner at rang $\mathbf{Q}_s = 2$ dersom

$$\frac{T_3}{T_1} \neq 1$$

og rang $\mathbf{Q}_s = 1$ dersom

$$\frac{T_3}{T_1} = 1$$

Systemet er derfor bare styrbart dersom $T_3/T_1 \neq 1$.

Tilsvarende finner vi for observerbarhet at rang $\mathbf{Q}_o = 2$ hvis

$$\frac{T_2}{T_3} \neq 1$$

og rang $\mathbf{Q}_o = 1$ dersom

$$\frac{T_2}{T_3} = 1$$

Dette betyr at begge systemets tilstander er observerbare bare dersom

$$\frac{T_2}{T_3} \neq 1$$

Transferfunksjonen mellom $\mathbf{u}(s)$ og $\mathbf{y}(s)$ finnes ved å erstatte integratorene i figur 5.4 med operatoren $(1/s)$ og deretter redusere blokkdiagrammet ved hjelp av reglene i figur 4.6, eller ved å benytte seg av formelen for transfermatrisen.

$$\mathbf{H}_u(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \quad (5.16)$$

Siden systemet er monovariabelt, gir dette transferfunksjonen

$$\begin{aligned} h(s) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s + \frac{1}{T_1} & -\frac{T_2 - T_3}{T_1 T_2} \\ 0 & s + \frac{1}{T_2} \end{bmatrix} \begin{bmatrix} \frac{T_3}{T_1 T_2} \\ \frac{1}{T_2} \end{bmatrix} \\ &= \frac{1 + T_3 s}{(1 + T_1 s)(1 + T_2 s)} \end{aligned} \quad (5.17)$$

Et blokkdiagram som viser hvordan denne transferfunksjonens elementer forekommer, er vist nederst i figur 5.4. Hvis

$$\frac{T_3}{T_1} = 1$$

skal systemet i følge det foregående ikke være styrbart. Vi ser av (5.17) at dette medfører at faktoren i telleren kan forkortes mot en av faktorene i nevneren. *Transferfunksjonen* er dermed blitt av 1. orden (en pol), til tross for at *systemet* jo er av annen orden.

Følgelig står vi overfor et tilfelle der et helt spesielt valg av systemparametre gjør systemet ikke-styrbart. For alle andre parameterverdier vil systemet være styrbart og derfor ha en transferfunksjon med to poler og et nullpunkt.

Vi kan gjerne kalle dette for **parametrisk ikke-styrbart** (PIS). Alternativet vil være at ikke-styrbart oppstår for (praktisk talt) alle parametre og er gitt av systemets *struktur*. Dette kan vi kalle **strukturell ikke-styrbart** (SIS).

Dersom

$$\frac{T_3}{T_2} = 1$$

vet vi av det foregående at systemet ikke er observerbart. Dette medfører ifølge (5.17) at telleren kan forkortes mot den andre faktoren i nevneren. Transferfunksjonen blir igjen av 1. orden, men får en annen pol (tidskonstant) enn i det foregående tilfellet. Tilsvarende det som er sagt ovenfor, ser vi at vi her står overfor et tilfelle av **parametrisk ikke-observerbarhet**

(PIO) fordi alle andre verdier enn

$$\frac{T_3}{T_2} = 1$$

vil gjøre at systemet blir observerbart og at transferfunksjonen vil få to poler og et nullpunkt.

Strukturell ikke-observerbarhet (SIO) vil vi ha hvis det er strukturen av systemet som er bestemmende.

■

EKSEMPEL 5.4: Elektrisk krets

Vi tar utgangspunkt i kretsen i figur 2.17 og antar at strømmen på inngangen måles og at pådraget er gitt av spenningen på inngangen. Fra (2.60) - (2.62) gir dette

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & a_{32} & a_{33} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_{11} \\ 0 \\ 0 \end{bmatrix} \quad (5.18)$$

$$\mathbf{c}^T = [1 \quad 0 \quad 0]$$

Vi antar for enkelhets skyld verdien 1 på alle R, L og C , noe som gir

$$\mathbf{A} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (5.19)$$

Styrbarhetsmatrisen er gitt av

$$\mathbf{Q}_s = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.20)$$

og observerbarhetsmatrisen er gitt av

$$\mathbf{Q}_o = \begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 2 & -1 \end{bmatrix} \quad (5.21)$$

Systemet er styrbart og observerbart. Det betyr at alle modi vil inngå i transferfunksjonen mellom inngangen og utgangen. Transferfunksjonen er gitt av

$$h(s) = \frac{1}{s^3 + 3s^2 + 3s + 1} \quad (5.22)$$

med tre poler i -1 . Dette er ekvivalent med egenverdiene til \mathbf{A} da

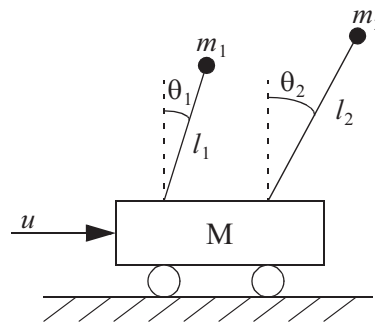
$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (5.23)$$

Alle modi inngår altså i transferfunksjonen.



EKSEMPEL 5.5: Vogn med inverterte pendler

Figur 5.5
Vogn med to
inverterte pendler



Vi betrakter en vogn med to inverterte pendler, som vist i figur 5.5. Massen til stengene av lengde l_1 og l_2 antas ubetydelig i forhold til m_1 og m_2 . Massen av vogna er M . Den antas å rulle uten friksjon, og utsettes for en kraft u , som er vårt pådrag. Reguleringsproblemet er nå å styre systemet slik at dette åpenbart ustabile systemet inntar en posisjon hvor begge pendlene står vertikalt, dvs. balansekunst. Vi bruker Newtons andre lov for å finne en modell av systemet, og forutsetter at vinkelutslagene θ_1 og θ_2 er så små at vi kan sette $\sin \theta = \theta$. Dette gir

$$\begin{aligned} M\dot{v} &= -m_1g\theta_1 - m_2g\theta_2 + u \\ m_1(\dot{v} + l_1\ddot{\theta}_1) &= m_1g\theta_1 \\ m_2(\dot{v} + l_2\ddot{\theta}_2) &= m_2g\theta_2 \end{aligned}$$

Her er v vognas hastighet, og g er som vanlig tyngdens akselerasjon. I den videre utledning antar vi for enkelhets skyld at $m_1 = m_2 = 1[\text{kg}]$ og $M = 10m_1$.

Vi velger nå tilstandene $x_1 = \theta_1$, $\dot{x}_1 = \dot{\theta}_1$, $x_3 = \theta_1 - \theta_2$ og $x_4 = \dot{\theta}_1 - \dot{\theta}_2$. Tilstandsrommodellen blir da

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

hvor

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1.2\alpha_1 & 0 & -0.1\alpha_1 & 0 \\ 0 & 0 & 0 & 1 \\ 1.2(\alpha_1 - \alpha_2) & 0 & \alpha_2 - 0.1(\alpha_1 - \alpha_2) & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ \beta_1 \\ 0 \\ \beta_1 - \beta_2 \end{bmatrix}$$

Her har vi innført $\alpha_1 = g/l_1$, $\alpha_2 = g/l_2$, $\beta_1 = -0.1/l_1$, $\beta_2 = -0.1/l_2$.

Styrbarhetsmatrisen for systemet, jf. (5.4), blir

$$\mathbf{Q}_s = [\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}]$$

Utrekning av determinanten gir

$$|\mathbf{Q}_s| = \frac{(0.011)^2 g^2 (l_1 - l_2)^2}{l_1^4 l_2^4}$$

Dette forteller oss at systemet er styrbart hvis og bare hvis $l_1 \neq l_2$. Når $l_1 = l_2$ har ikke \mathbf{Q}_s full rang (i dette tilfelle 4), og systemet er altså ikke styrbart. Vi har at $\alpha_1 = \alpha_2$ og $\beta_1 = \beta_2$, og dermed blir

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1.2\alpha_1 & 0 & -0.1\alpha_1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \alpha_1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ \beta_1 \\ 0 \\ 0 \end{bmatrix}$$

Vi ser umiddelbart at pådraget ikke kan influere på tilstandene x_3 og x_4 , verken direkte eller indirekte.

Ikke-styrbarheten for systemet med $l_1 = l_2$ kan vi også fatte intuitivt. Anta en startsituasjon hvor den ene av de to like lange pendler er i ferd med å falle til venstre og den andre til høyre: Det er umulig å "hente inn" den ene først ved å bevege vogna, uten å forverre posisjonen til den andre. Hvis derimot pendlene har forskjellig lengde, kan man "hente inn" den korteste først, for så å fange opp den lengste etterpå.

Så til systemets observerbarhet: Vi antar at den eneste målinga er av vinkelen $x_1 = \theta_1$. Vi har

$$y = \mathbf{c}^T \mathbf{x}, \quad \text{hvor } \mathbf{c}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

Observerbarhetsmatrisen blir

$$\mathbf{Q}_o = [\mathbf{c} \ (\mathbf{c}^T \mathbf{A})^T \ (\mathbf{c}^T \mathbf{A}^2)^T \ (\mathbf{c}^T \mathbf{A}^3)^T]$$

Determinanten blir

$$|\mathbf{Q}_o| = 0.01 \frac{g^2}{l_1^2}$$

Systemet er alltid observerbart. Dette betyr at sjøl om vi bare måler den ene vinkelen, kan vi alltid, via den matematiske modellen av systemet, finne verdien av den andre vinkelen til enhver tid.

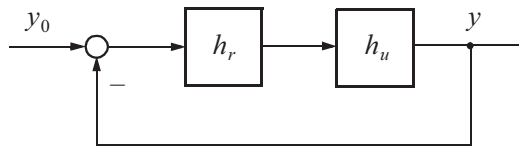
■

5.5 FUNKSJONELL STYRBARHET (FS) AV MONOVARIABLE LINEÆRE SYSTEMER

Definisjonene av *styrbarhet* og *observerbarhet* som er presentert i avsnitt 5.2 og 5.3 gjelder systemets *tilstand* $\mathbf{x}(t)$. Det ville derfor vært mest korrekt om vi hadde brukt betegnelsene *tilstandsstyrbarhet* (TS) og *tilstandsobserverbarhet* (TO).

Men betingelsene for å oppnå TS og TO er *for strenge* hvis vi bare vil vite om et regulerings-system som vist i figur 5.6 kan realiseres med formål å styre utgangen y mot en referanse y_0 . Vi forlanger altså *ikke* å kunne styre hele tilstandsvektoren $\mathbf{x}(t)$.

Figur 5.6
Monovariabel
tilbakekoplet sløyfe



En tilstrekkelig betingelse for denne forenklede problemstillingen betegnes “*funksjonell styrbarhet*” (FS) og er gitt av:

En lineær monovariabel reguleringsløyfe vil være FS dersom $h_u(s) \neq 0$ for alle s unntatt i systems nullpunkter.

Det forholder seg annerledes dersom prosessen er multivariabel, fordi vi da har mange transferfunksjoner ($h_{ij}(s)$) som forbinder inngangene med utgangene. Dette mer generelle problemet behandles i avsnitt 10.7.

EKSEMPEL 5.6: Styrbarhet/funksjonell styrbarhet

I avsnitt 5.4 er det påvist at prosessen gitt i figur 5.4 ikke er (*tilstands*)styrbar (TS) dersom $T_3/T_1 = 1$. Dette betyr at prosessens transferfunksjon har et nullpunkt som kansellerer (opphøyer) en pol. Men systemet er likevel *funksjonelt styrbart* (FS) fordi $h_u(s) \neq 0$ for alle s unntatt i systemets nullpunkt $s = -1/T_3$,

$$h_u(s) = \frac{1 + T_3s}{(1 + T_1s)(1 + T_2s)} \neq 0 \quad \text{for alle } s \neq -\frac{1}{T_3}$$

Å nøye seg med bare å kreve FS i stedet for det mer restriktive TS, er akseptabelt bare hvis den modus som forkortes bort når vi ikke har TS, er stabil, dvs. vi må kreve at $1/T_3 = 1/T_1$ ligger i venstre halvplan. Hvorfor?



Frekvensanalyse

6.1 Innledning og motivering 193

6.2 Utledning av frekvensrespons 195

6.3 Grafisk representasjon av frekvensrespons for monovariabel systemer 201

6.4 Asymptotisk amplitude/fase/frekvens-diagram (Bodediagram) 209

6.5 Skissering av tidsrespons på basis av frekvensrespons 224

6.1 INNLEDNING OG MOTIVERING

Vi skal i dette kapitlet studere **frekvensanalyse**. Frekvensanalysen muliggjør bruk av grafiske analysemetoder og gir derfor en hurtig oversikt over en rekke avgjørende systemegenskaper. La oss først forklare begrepet frekvensanalyse ved et innledende eksempel.

EKSEMPEL 6.1: Sinusformet variasjon i ankerspenning på likestrømsmotor

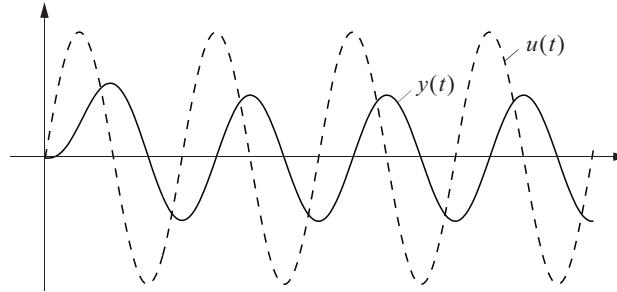
Vi tar for oss likestrømsmotoren med transferfunksjon som i (4.135). Der ble denne (vi forutsetter åpen sløyfe – ingen tilbakekopling) modellert med transferfunksjon fra ankerspenning u til vinkelhastighet $x_1 = y$:

$$\frac{y}{u}(s) = \frac{1}{K_v(1 + T_a s)(1 + T_m s)} \quad (6.1)$$

Figur 6.1 viser tidsresponsen til vinkelhastigheten når ankerspenningen varieres sinusformet. Ankerspenningen er $u(t) = u_0 \sin(\omega_0 t)$. u_0 kalles **amplitude** og ω **vinkelfrekvens**. Vinkelfrekvens er gitt av $\omega = 2\pi/T$, hvor T er periodetiden til sinussignalet.

Ved å studere figuren ser vi at tidsresponsen $y(t)$ inneholder en **transient** komponent som dør ut, slik at $y(t)$ etterhvert får en rein sinusform. Denne sinusfunksjonen har en annen amplitude enn inngangen $u(t)$. Vi sier at den er **amplitudeforsterket** sammenliknet med inngangen (denne forsterkninga kan godt være < 1). Videre er $y(t)$ forskjøvet sammenliknet med inngangen. Vi sier at den er **faseforskjøvet**.

Figur 6.1
Vinkelhastighet for rotor $y(t)$ ved sinusformet ankerspenning $u(t)$.



Dersom amplituden på inngangen endres, vil dette endre amplituden på utgangen. Forholdet mellom amplitudene vil være uendret. Videre vil faseforskyvningen mellom $y(t)$ og $u(t)$ forbli uendret.

Dersom amplituden på $u(t)$ er som i figur 6.1 samtidig som vinkelfrekvensen øker, vil amplituden til $y(t)$ avta. Samtidig vil faseforskyvningen øke. Denne typen analyse kalles **frekvensanalyse**, og metoder for dette presenteres i dette kapitlet.



Vi skal gi en mer formalisert beskrivelse av frekvensanalysen.

Et periodisk signal $u(t)$ som ved $t = 0$ settes på et lineært system med impulsresponsen $h(t)$, vil resultere i et signal $y(t)$ på utgangen som inneholder en *transient* og en *stasjonær* respons $y(t) = y_t(t) + y_s(t)$. Med bakgrunn i dette kan frekvensresponsen til et system defineres.

Frekvensresponsen til et system er defineres som den **stasjonære responsen** $y_s(t)$ på utgangen til et lineært system, når inngangssignalet er en sinussvingning.

De stasjonære signalene internt i, og på utgangen av, et lineært system som påtrykkes en sinussvingning, vil også være sinussvingninger med samme frekvens, men forskjøvet i fase og amplitude i forhold til sinussvingningen på inngangen

Merk at frekvensrespons-begrepet bare gjelder for *lineære systemer*.

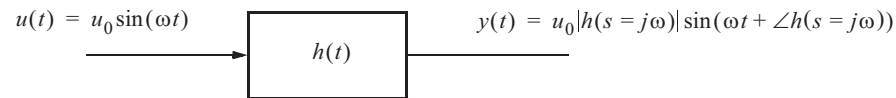
Vi har et inngangssignal med amplitude u_0 og vinkelfrekvens ω .

$$u(t) = u_0 \sin(\omega t) \quad (6.2)$$

som påtrykkes systemet i figur 6.2. Den stasjonære responsen på utgangen blir da

$$y_s(t) = u_0 |h(s = j\omega)| \sin(\omega t + \angle h(s = j\omega)) \quad (6.3)$$

Figur 6.2
Inngangs- og
utgangssignal for
system med
impulsrespons
 $h(t)$



Amplitudeforskyvningen til systemet er gitt av $u_0 |h(s = j\omega)|$ og faseforskyvningen er gitt av $\angle h(s = j\omega)$.

Alle stasjonære periodiske inngangssignaler kan uttrykkes som en sum av forskjellige frekvenskomponenter ved hjelp av **Fouriertransformasjonen**¹. Dersom inngangssignalet på et vilkårlig lineært system er sammensatt av mange frekvenskomponenter (sinuser), vil utgangssignalet inneholde de samme frekvenskomponentene (sinussvingningene), men endret i amplitude og forskjøvet i fase. Dette følger av superposisjonsprinsippet.

6.2 UTLEDNING AV FREKVENSRSPONS

Vi skal i dette avsnittet først bevise det som ble påstått i innledningen: at utgangen av et lineært system som påtrykkes en sinus, er en faseforskjøvet sinus.

Vi tar utgangspunkt i et lineært og monovariabelt system beskrevet med en transferfunksjon $h(s)$. Dette systemet påvirkes av et skalart pådrag $u(t)$, og observeres gjennom en skalar måling $y(t)$.

Vi lar pådraget være en sinussvingning.

$$u(t) = \begin{cases} u_0 \sin \omega t & \text{for } t > 0 \\ 0 & \text{for } t < 0 \end{cases} \quad (6.4)$$

Laplacestransformasjon av (6.4) gir

$$u(s) = \frac{u_0 \omega}{s^2 + \omega^2} \quad (6.5)$$

¹Fouriertransformasjonen får vi ved å sette $s = j\omega$ i Laplacestransformasjonen,

$$f(j\omega) = \int_0^{\infty} f(t) e^{-j\omega t} dt = f(s)|_{s=j\omega}. \text{ I matematiske lærebøker, i motsetning til i signalbehandlings- og}$$

kybernetikk-miljøet, defineres den ofte med konstanten $\frac{1}{\sqrt{2\pi}}$ foran integralet. Dette er en uvesentlig forskjell.

Dette pådraget gir følgende respons dersom det påtrykkes systemet med transferfunksjon $h(s)$

$$y(s) = h(s) \frac{u_0 \omega}{s^2 + \omega^2} \quad (6.6)$$

Vi skal bestemme tidsforløpet $y(t)$ ved invers Laplacetransformasjon av (6.6) og benytter (4.41)

$$y(t) = \frac{1}{2\pi j} \int_{\alpha-j\infty}^{\alpha+j\infty} h(s) \frac{u_0 \omega}{s^2 + \omega^2} e^{ts} ds = \frac{1}{2\pi j} \int_{\alpha-j\infty}^{\alpha+j\infty} y(s) ds \quad (6.7)$$

$$= \operatorname{res}_{s=j\omega} [y(s)] + \operatorname{res}_{s=-j\omega} [y(s)] + \sum_{i=1}^n \operatorname{res}_{s=s_i} [y(s)]$$

Residuenes beregnes i integrandens singulariteter, og det antas at transferfunksjonen $h(s)$ har n distinkte singulariteter (dette er ikke en nødvendig forutsetning, men vi gjør den for å forenkle framstillinga). Beregner vi residuenes som angitt i (4.39), finner vi

$$y(t) = \frac{h(j\omega)u_0\omega e^{j\omega t}}{2j\omega} + \frac{h(-j\omega)u_0\omega e^{-j\omega t}}{-2j\omega} + \sum_{i=1}^n \left. \frac{(s-s_i)h(s)u_0\omega e^{ts}}{s^2 + \omega^2} \right|_{s=s_i} \quad (6.8)$$

De to første leddene i (6.8) og det tredje leddet representerer henholdsvis den *stasjonære* og den *transiente* delen av responsen.

Vi er nå spesielt interessert i å finne ut hvordan responsen $y(t)$ blir når $t \rightarrow \infty$. Vi forutsetter at alle singularitetene (=polene) $s = s_i$ av $h(s)$ ligger i venstre halvplan¹, det vil si $\operatorname{Re}(s_i) < 0$. Alle leddene under summetegnet i (6.8) vil gå mot null når $t \rightarrow \infty$, fordi $e^{ts_i} \rightarrow 0$, når $t \rightarrow \infty$.

Vi har derfor bare igjen følgende uttrykk

$$y_s(t) = y(t) = u_0 \left(h(j\omega) \frac{e^{j\omega t}}{2j} - h(-j\omega) \frac{e^{-j\omega t}}{2j} \right) \quad (6.9)$$

$h(j\omega)$ er en vektor med lengde $|h(j\omega)|$ og vinkel i forhold til den reelle akse $\angle h(j\omega)$.

¹Noe annet ville være meningsløst. For at et system overhodet skal ha en (stasjonær) frekvensrespons, må det være asymptotisk stabilt (se kapittel 8), ellers vil systemet svinge seg ut til uendelig.

Vi har at

$$\begin{aligned} h(j\omega) &= |h(j\omega)|e^{j\angle h(j\omega)} \\ h(-j\omega) &= |h(j\omega)|e^{-j\angle h(j\omega)} \end{aligned} \quad (6.10)$$

fordi nullpunkter og poler i transferfunksjoner for fysiske systemer alltid ligger symmetrisk om den reelle akse. Det blir lettere å skjønne (6.10) etter å ha studert eksemplene 6.2 og 6.3.

Vi kan nå skrive (6.9) slik

$$\begin{aligned} y_s(t) &= u_0 |h(j\omega)| \frac{e^{j(\omega t + \angle h(j\omega))} - e^{-j(\omega t + \angle h(j\omega))}}{2j} \\ &= u_0 |h(j\omega)| \sin(\omega t + \angle h(j\omega)) \end{aligned} \quad (6.11)$$

Vi ser at responsen $y_s(t)$ vil være en sinusformet svingning med samme frekvens som pådraget og med en amplitude $y_0 = u_0 |h(j\omega)|$, men faseforskjøvet i forhold til den påtrykte sinussvingningen med en vinkel lik $\angle h(j\omega)$.

▲

- Funksjonen $h(j\omega)$ kalles systemets **frekvensrespons**.
- Forholdet $y_0/u_0 = |h(j\omega)|$ kalles **amplitudeforholdet**.
- $\angle h(j\omega)$ kalles *faseforskyvningen* eller **fasevinkelen**.

Utleddningen viser at vi kan finne ut hvordan et monovariabelt system vil reagere på en stasjonær sinussvingning, ved å betrakte transferfunksjonen. Vi lar da den komplekse variabelen s i transferfunksjonen $h(s)$ anta den imaginære verdien $j\omega$, der ω er sinussvingningens frekvens.

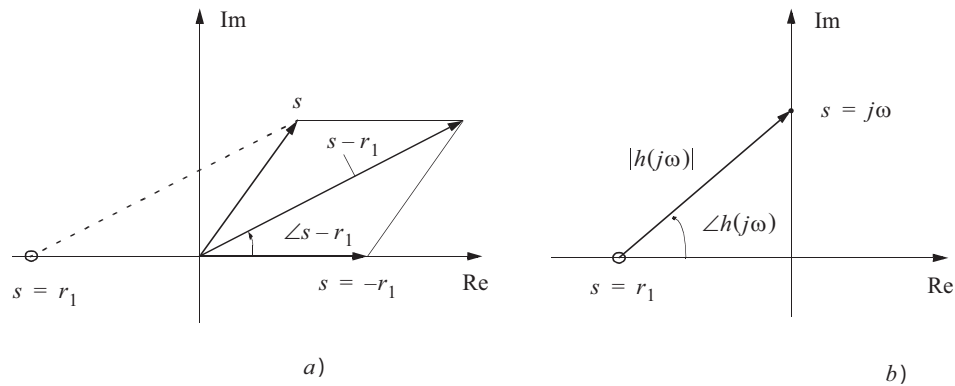
EKSEMPEL 6.2: Frekvensresponsen til en transferfunksjon med et nullpunkt

Vi ser først på et system som har transferfunksjonen $h(s) = s - r_1$, der r_1 er en reell konstant mindre enn null. Figur 6.3 viser at systemet har et nullpunkt i $s = r_1$.

s kan ha en vilkårlig kompleks verdi. Vi kan finne vektoren $s - r_1$ som en geometrisk sum av to vektorer. Den resulterende vektoren kunne vi også ha funnet ved å trekke vektoren fra nullpunktet fram til den aktuelle verdien av s som antydnet stiple.

For bestemmelse av frekvensresponsen velger vi nå $s = j\omega$ og finner som skissert i figur 6.3 b) frekvensresponsens tallverdi og vinkel ved hjelp av vektoren fra nullpunktet til det aktuelle punktet på den imaginære akse som angir frekvensen ω .

Figur 6.3
Vektoriell framstilling av $h(s)$



Positive vinkler regnes som vanlig i retning mot urviserens bevegelse slik at vinkelen som er antydnet i figur 6.3 b), er positiv. Vi ser umiddelbart hva som skjer dersom vi forandrer frekvensen fra null og oppover langs den imaginære aksen. Lengden av vektoren vil da øke fra en verdi $|h(j\omega)| = r_1$ ved $\omega = 0$ mot uendelig, og $\angle h(j\omega)$ vil øke fra null mot 90° ($\pi/2$ radianer). Den fysiske betydningen av dette er at hvis pådraget har konstant amplitude, vil utgangen (målingen) øke mot uendelig amplitude når frekvensen øker. Dette er ikke mulig i fysiske systemer. Transferfunksjoner for fysiske systemer inneholder aldri bare nullpunkter, men har minst like mange (som oftest flere) poler som nullpunkter. (En slik transferfunksjon kalles **proper**).

■

EKSEMPEL 6.3: Frekvensresponsen til en transferfunksjon med en pol

Vi velger nå et system med transferfunksjonen $h(s) = 1/(s - \lambda_1)$. Siden vi har at

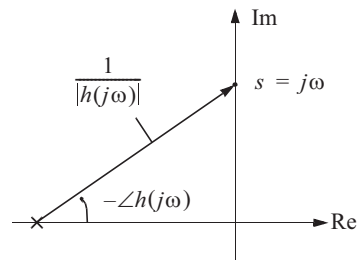
$$s - \lambda_1 = |s - \lambda_1|e^{j\angle(s - \lambda_1)}$$

kan vi skrive

$$h(s) = \frac{1}{s - \lambda_1} = \frac{1}{|s - \lambda_1|e^{j\angle(s - \lambda_1)}} = \frac{1}{|s - \lambda_1|}e^{-j\angle(s - \lambda_1)} \quad (6.12)$$

Figur 6.4 viser at vi kan få et inntrykk av frekvensresponsens tallverdi og vinkel ved å trekke en vektor fra transferfunksjonens pol ($s = \lambda_1$) til det aktuelle punktet på den imaginære aksen som angir frekvensen ω .

Figur 6.4
Vektoriell framstilling av $h(j\omega)$



Likning (6.12) viser at vektoren som framkommer, har en lengde $1/|h(j\omega)|$ og danner en vinkel med den reelle akse som er lik $-\angle h(j\omega)$. Vi kan også i dette tilfellet se hva som skjer med frekvensresponsen når frekvensen endres. Når ω økes fra null mot uendelig, vil frekvensresponsens tallverdi avta fra $|h(j\omega)| = 1/|\lambda_1|$ mot null, og $\angle h(j\omega)$ vil avta fra null til -90° ($-\pi/2$ radianer).

■

Et monovariabelt systems transferfunksjon vil kunne bestemmes entydig dersom frekvensresponsen $h(j\omega)$ er kjent (i tallverdi og vinkel) for alle verdier $s = j\omega$ mellom $-j\infty$ og $j\infty$. Siden vi imidlertid har at $|h(-j\omega)| = |h(j\omega)|$ og $\angle h(-j\omega) = -\angle h(j\omega)$, er det tilstrekkelig å kjenne frekvensresponsen for alle frekvenser mellom null og uendelig, det vil si langs den positive imaginære akse. Dette tilsvarer den fysiske situasjonen når systemet påtrykkes positive frekvenser (negative frekvenser har ingen fysisk mening).

Et analytisk uttrykk for frekvensresponsen får vi ved å sette $s = j\omega$ inn i transferfunksjonen. Antar vi at transferfunksjonen er en **rasjonal funksjon** som kan skrives

$$h(s) = \frac{K(1 + \tilde{T}_1 s)(1 + \tilde{T}_2 s) \dots (1 + \tilde{T}_m s)}{(1 + T_1 s)(1 + T_2 s) \dots (1 + T_n s)} \quad (6.13)$$

finner vi frekvensresponsens *tallverdi*

$$\begin{aligned} |h(j\omega)| &= \left. \frac{K|1 + \tilde{T}_1 s| |1 + \tilde{T}_2 s| \dots |1 + \tilde{T}_m s|}{|1 + T_1 s| |1 + T_2 s| \dots |1 + T_n s|} \right|_{s=j\omega} \\ &= K \sqrt{\frac{(1 + (\omega \tilde{T}_1)^2)(1 + (\omega \tilde{T}_2)^2) \dots (1 + (\omega \tilde{T}_m)^2)}{(1 + (\omega T_1)^2)(1 + (\omega T_2)^2) \dots (1 + (\omega T_n)^2)}} \end{aligned} \quad (6.14)$$

Så til frekvensresponsens *vinkel*: Den vil bli summen av vinkelbidragene for hvert enkelt element i telleren i transferfunksjonen minus summen av bidragene fra hvert enkelt ledd i nevneren

$$\begin{aligned} \angle h(j\omega) &= \arctan \omega \tilde{T}_1 + \arctan \omega \tilde{T}_2 + \dots + \arctan \omega \tilde{T}_m \\ &\quad - (\arctan \omega T_1 + \arctan \omega T_2 + \dots + \arctan \omega T_n) \end{aligned} \quad (6.15)$$

Ofte vil en finne at transferfunksjonen for et dynamisk system ikke er en rasjonal funksjon. Et vanlig tilfelle er for eksempel at systemet inneholder en eller annen form for transportforsinkelse (tidsforsinkelse). Dette medfører at det i transferfunksjonen vil inngå et multiplikativt ledd av typen $e^{-\tau s}$, der τ er **transport(tids-)forsinkelsen**. Frekvensresponsen av dette leddet vil være karakterisert ved

$$|e^{-j\omega\tau}| = 1 \quad , \quad \angle e^{-j\omega\tau} = -\omega\tau \quad (6.16)$$

6.2.1 Frekvensrespons av multivariable systemer

På basis av utledningen av transfermatriser for multivariable systemer i avsnitt 4.6.1 kan vi også finne **frekvensresponsmatriser** $\mathbf{H}(j\omega)$. En frekvensresponsmatrise framkommer ved å sette $s = j\omega$ inn i transfermatrisen $\mathbf{H}(s)$. Betydningen av en frekvensresponsmatrise blir bare en samling av individuelle frekvensresponsfunksjoner dersom hvert enkelt element i pådragsvektoren er uavhengig av elementene i responsvektoren (målevektoren). Hver enkelt respons $y_i(t)$ vil som kjent være en sum av enkeltresponser forårsaket av de enkelte elementene i pådragsvektoren. Dersom alle elementene i pådragsvektoren er sinussvingninger med samme frekvens og med kjente fasevinkler relativt til en fast vinkelreferanse, vil alle elementene i responsvektoren være sinussvingninger der både amplitude og fase kan bestemmes.

Dersom en frekvensresponsmatrise (transfermatrise) inngår som ledd i et tilbakekoplet multivariabelt system, skal vi i kapittel 10 se at det kan utledes en del nyttige resultater.

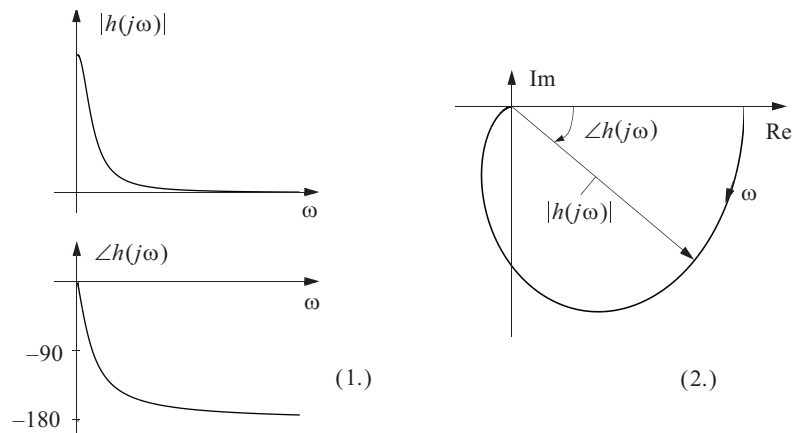
6.3 GRAFISK REPRESENTASJON AV FREKVENSRSPONS FOR MONOVARIABLE SYSTEMER

Vi har to hovedformer for grafisk representasjon av den komplekse størrelsen $h(j\omega)$:

1. Tallverdien $|h(j\omega)|$ og vinkelen $\angle h(j\omega)$ framstilles hver for seg som funksjoner av ω . Figur 6.5, del 1, viser dette.
2. Kurven for endepunktet av vektoren $h(j\omega)$ i det komplekse plan framstilles med ω som parameter, som vist i figur 6.5, del 2 (dette kalles et **Nyquistdiagram**, et **polar-diagram**, eller en **stedkurve**).

Figur 6.5

- 1) $|h(j\omega)|$ og $\angle h(j\omega)$ som funksjon av ω
- 2) Polar stedkurve for $|h(j\omega)|$



Representasjonsform 2 brukes mye, mens 1 er sjelden brukt. Vi introduserer nå to modifiserte representasjoner basert på del 1 og 2 ovenfor, vist i hhv. del 1.a og 2.a i figur 6.7 nedenfor.

Formen 1.a kalles **Bodediagram** eller **amplitude/fase/frekvens-diagram** (= **AFF-diagram**) og blir svært mye brukt. I dette diagrammet bruker vi *desibelskalaen*¹. Vi skal forklare den i tilknytning til diagrammet 1.a, som har følgende egenskaper: Det består, som i punkt 1 ovenfor, av to kurver med felles abscisseakse; et **amplitudediagram** og et **fasediagram**. Enheten på abscisseaksen (den horisontale aksene) er nå $\lg \omega$, der ω er frekvensen i [rad/tidsenhet]. Fra nå av vil ω alltid være i logaritmisk målestokk i denne boka.

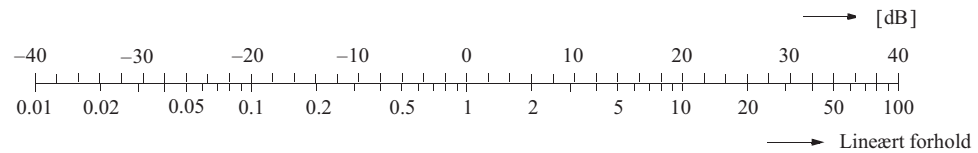
- I amplitudediagrammet tegnes absoluttverdien av $h(j\omega)$ (= “amplituden til $h(j\omega)$ ”) som funksjon av frekvensen. Enheten på ordinataksen er også logaritmisk, nærmere bestemt i enheten **desibel**, definert ved

$$|h(j\omega)|[\text{dB}] = 20 \lg |h(j\omega)|, \text{ der } \lg \text{ angir } 10\text{-logaritmen, } \lg x = \log_{10} x$$

¹Desibelskalaen er hentet fra telekommunikasjonsteknikken.

Figur 6.6 viser en omregningskala mellom lineære mål og dB-mål.

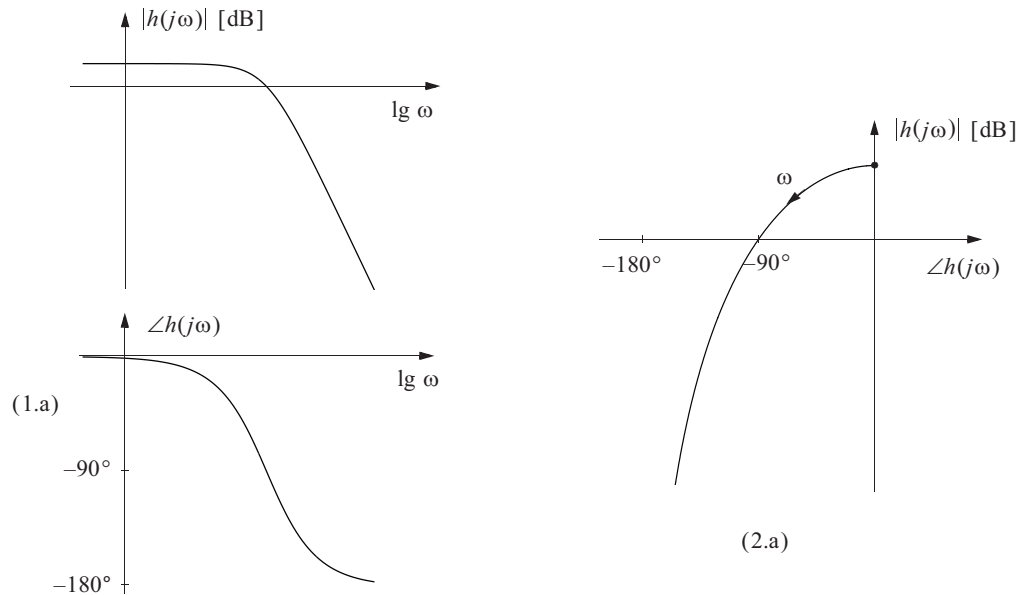
Figur 6.6
Desibelskalaen



- I den nedre delen av Bodediagrammet, dvs. fase-diagrammet, framstilles $\angle h(j\omega)$ som funksjon av samme frekvensskala, $\lg(\omega)$. Enheten på ordinataksen er grader eller radianer, som regel det første.

Figur 6.7

1.a Bodediagram
2.a AF-diagram



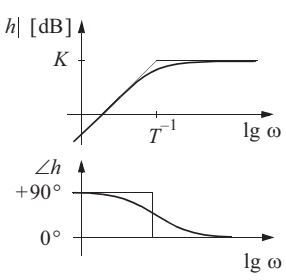
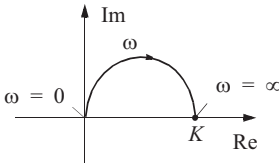
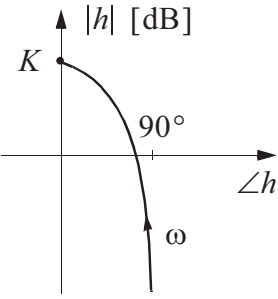
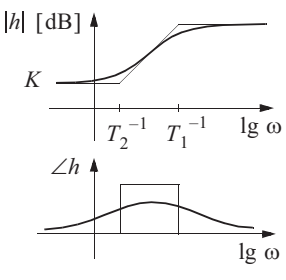
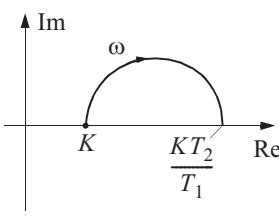
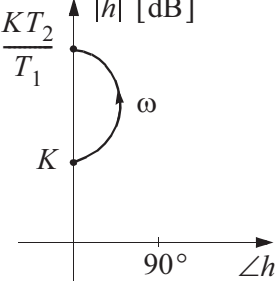
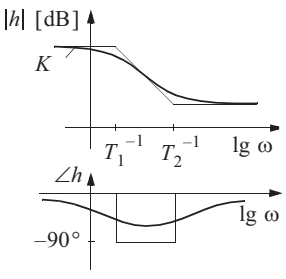
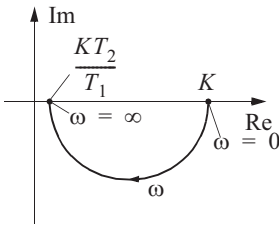
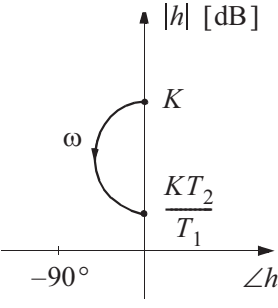
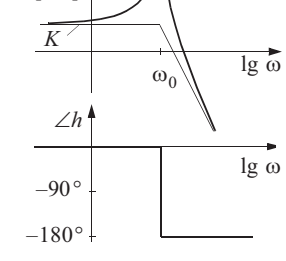
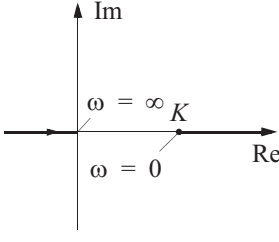
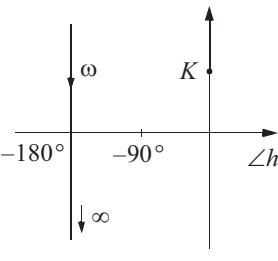
Nå til en variant av Nyquistdiagrammet fra figur 6.5, del 2. Denne er vist i figur 6.7 del 2.a:

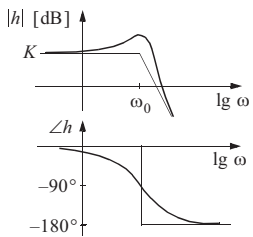
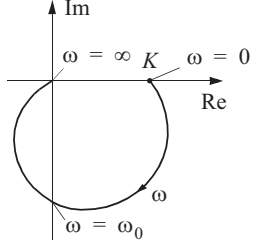
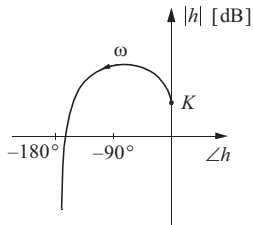
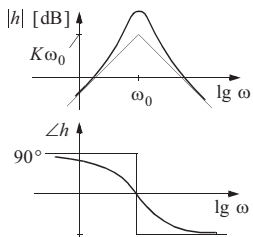
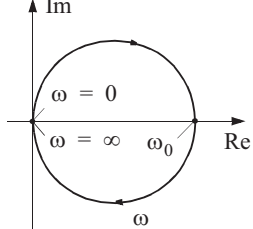
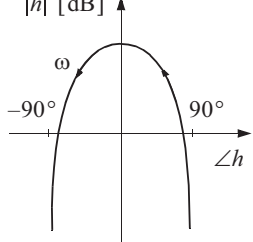
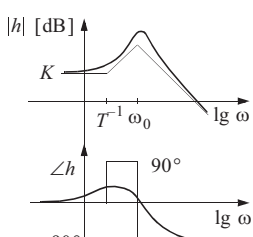
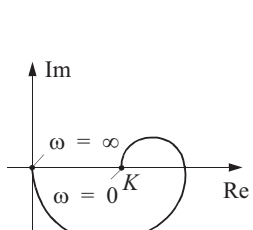
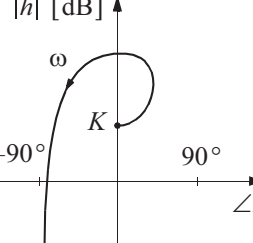
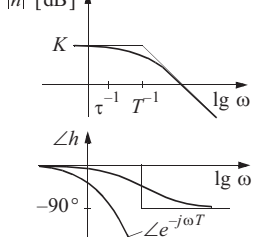
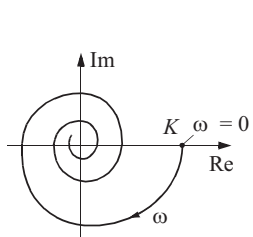
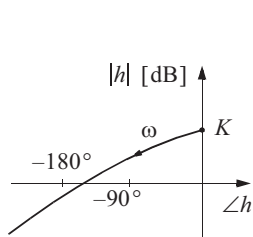
- $|h(j\omega)|$ [dB] framstilles som funksjon av $\angle h(j\omega)$ i et kartesisk koordinatsystem med ω som parameter. Dette kalles **amplitude/fase-diagram (AF-diagram)**.

En del eksempler på frekvensrespons—representert ved sine Bode-, Nyquist, og AF-diagrammer, er vist i tabell 6.1.

TABELL 6.1 Diagrammer for vanlig forekommende transferfunksjoner

Transfer-funksjon $h(s)$	AFF-(Bode-) diagram	Nyquist-diagram	AF-diagram
K			
$K \frac{1}{Ts}$			
$K \frac{1+Ts}{Ts}$			
$K \frac{1}{1+Ts}$			

Transfer-funksjon $h(s)$	AFF-(Bode-) diagram	Nyquist-diagram	AF-diagram
$K \frac{T s}{1 + T s}$			
$K \frac{1 + T_2 s}{1 + T_1 s} \quad T_2 > T_1$			
$K \frac{1 + T_2 s}{1 + T_1 s} \quad T_1 > T_2$			
$\frac{K}{1 + \left(\frac{s}{\omega_0}\right)^2}$			

Transfer-funksjon $h(s)$	AFF-(Bode-)diagram	Nyquist-diagram	AF-diagram
$\frac{K}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$			
$\frac{Ks}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$			
$K \frac{1 + Ts}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$			
$\frac{Ke^{-\tau s}}{1 + Ts}$			

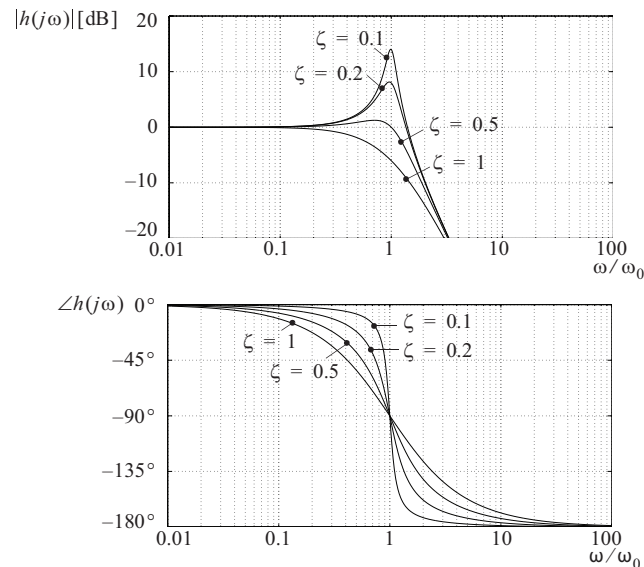
EKSEMPEL 6.4: AFF-diagram for et resonansledd

Vi skal se nærmere på resonansleddet i tabell 6.1, med transferfunksjon

$$h(s) = \frac{K}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$$

AFF-diagrammet i figur 6.8 viser at vi får forskjellige frekvensresponskurver, avhengig av verdien av ζ . (Merk at benevnningen langs abscisseaksen er ω/ω_0 , ikke $\lg(\omega/\omega_0)$, siden verdiene nå er merket av i logaritmisk skala langs denne aksen).

Figur 6.8
AFF-diagram for
resonansledd



Når ζ har en liten tallverdi, vil amplitudeforholdet $|h(j\omega)|$ oppvise en typisk resonans. Vi er interessert i å finne den maksimale verdien av dette forholdet og ved hvilken frekvens det oppstår. Det kan utledes at

$$|h(j\omega)|_{\max} = \frac{K}{2\zeta\sqrt{1-\zeta^2}} \quad \text{for} \quad \frac{\omega}{\omega_0} = \sqrt{1-2\zeta^2}$$

når $\zeta \leq 1/\sqrt{2}$, og

$$|h(j\omega)|_{\max} = K \quad \text{for} \quad \frac{\omega}{\omega_0} = 0$$

når $\zeta \geq 1/\sqrt{2}$.

Resonanstoppen (dersom en slik finnes) vil derfor alltid komme ved en frekvens som er noe lavere enn $\omega = \omega_0$. Dersom systemet er svakt dempet ($\zeta \ll 1$), finner vi at

$$|h(j\omega)|_{\max} \approx \frac{K}{2\zeta} \quad \text{for} \quad \frac{\omega}{\omega_0} \approx 1$$

Av den polare framstillingen (form 2) i tabell 6.1 framgår det at

$$\angle h(j\omega) = -\frac{\pi}{2} \quad \text{for} \quad \omega = \omega_0$$

uansett hvilken verdi ζ har.

Vi skal komme tilbake til resonansleddet i avsnitt 6.4 når vi behandler asymptotiske AFF-diagrammer.



EKSEMPEL 6.5: Nyquist-diagram for en transferfunksjon med pol i origo

Det er et spesielt problem med grafisk framstilling av frekvensresponsen av transferfunksjoner som har poler på den imaginære akse. I slike tilfeller vil innsetting av $s = j\omega$ lede til uttrykk som går mot uendelig, fordi transferfunksjonen ikke er analytisk i sine poler (singulariteter). Vi tar som eksempel for oss transferfunksjonen

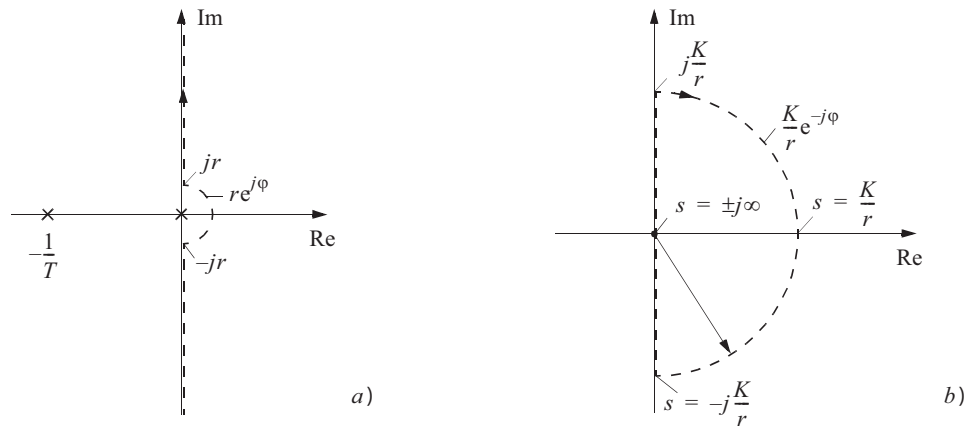
$$h(s) = \frac{K}{s(1+Ts)}$$

som har poler i henholdsvis $s = 0$ og $s = -1/T$. Vi kan undersøke denne funksjonens verdi langs en bane som gitt i figur 6.9 a). Denne banen består av den imaginære akse bortsett fra omkring origo der den er en liten halvsirkel i høyre halvplan. Halvsirkelens radius er r og kan gjøres så liten vi ønsker. Lar vi radien i halvsirkelen bli liten nok, ser vi at vektoren fra polen i $s = -1/T$ ikke vil forandre størrelse når vi lar s gjennomløpe halvsirkelen. Vektoren fra polen i origo vil imidlertid forandre vinkel fra $\pi/2$ til $-\pi/2$ når s gjennomløper halvsirkelen fra $-jr$ til $+jr$.

Den polare stedkurven for funksjonen K/s alene vil derfor bli som vist i figur 6.9 b) når s antar alle verdier langs den banen som er beskrevet ovenfor. Den lille halvsirkelen omkring origo i figur 6.9 a) resulterer altså i en stedkurve formet som en stor halvsirkel (i høyre halvplan) med radius $1/r$ som vist i figur 6.9 b).

Figur 6.9

- a) Bane for s som omslutter en pol i origo
 b) Stedkurve for K/s

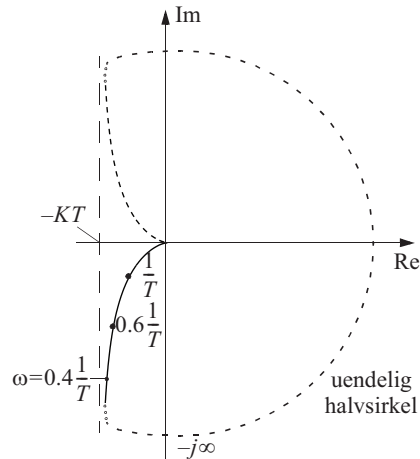


Tar vi også med leddet $1/(1 + Ts)$, vil stedkurven for hele transferfunksjonen $h(s)$ bli som vist i figur 6.10 der vi har latt $r \rightarrow 0$. Den reelle del av vektoren fra origo i figur 6.10 er

$$\operatorname{Re}(h(j\omega)) = \operatorname{Re}\left(\frac{K}{j\omega(1 + j\omega T)}\right) = \frac{-KT}{1 + (\omega T)^2} \rightarrow \frac{-KT}{\omega \rightarrow 0} \quad (6.17)$$

Figur 6.10

Stedkurve for $h(j\omega)$



Stedkurven i figur 6.10 tangerer altså en asymptote i venstre halvplan parallelt med den imaginære aksen i en avstand KT fra denne. Lar vi ω anta verdier fra 0 til $+\infty$, vil vektoren fra origo forandre lengde fra ∞ til 0 og vinkelen vil variere fra $-\pi/2 = -90^\circ$ til $-\pi = -180^\circ$. Dette framgår av figur 6.10.

■

6.4 ASYMPTOTISK AMPLITUDE/FASE/FREKVENSDIAGRAM (BODEDIAGRAM)

En transferfunksjon kan bestå av forsterkning, integratorledd eller derivatorledd, tidskonstantledd, resonansledd, nullpunktsledd og tidsforsinkelsesledd.

$$h(j\omega) = \frac{K(1 + j\omega\tilde{T}_1)\dots e^{-j\omega\tau}}{(j\omega)^n(1 + j\omega T_1)\dots \left(1 - \left(\frac{\omega}{\omega_0}\right)^2 + j2\zeta\frac{\omega}{\omega_0}\right)\dots} \quad (6.18)$$

Siden vi i amplitudediagrammet bruker [dB], dvs. $20\lg|h(j\omega)|$, kan vi summere bidragene fra de enkelte leddene som inngår i telleren til $h(j\omega)$, og trekke fra bidragene i nevneren,

$$\begin{aligned} |h(j\omega)|[\text{dB}] = & 20\left(\lg|K| + \lg|1 + j\omega\tilde{T}_1| + \dots + \lg|e^{-j\omega\tau}| \right. \\ & \left. - n\lg|j\omega| - \lg|1 + j\omega T_1| - \dots - \lg\left|1 - \left(\frac{\omega}{\omega_0}\right)^2 + j2\zeta\frac{\omega}{\omega_0}\right| - \dots\right) \end{aligned} \quad (6.19)$$

Likning (6.15) viser at vi også kan summere/subtrahere bidragene fra de enkelte leddene $h(j\omega)$ i for å finne fasen $\angle h(j\omega)$.

$$\begin{aligned} \angle h(j\omega) = & \angle K + \angle(1 + j\omega\tilde{T}_1) + \dots + \angle e^{-j\omega\tau} \\ & - \angle(j\omega)^n - \angle(1 + j\omega T_1) - \dots - \angle\left(1 - \left(\frac{\omega}{\omega_0}\right)^2 + j2\zeta\frac{\omega}{\omega_0}\right) - \dots \end{aligned} \quad (6.20)$$

De asymptotiske bidragene til $|h(j\omega)|$ [dB] og $\angle h(j\omega)$ fra de forskjellige typene ledd er:

Forsterkning:

Amplitudebidraget blir $20\lg K$ [dB].

Vinkelbidraget (vinkelen med den reelle aksene i det komplekse plan) $\angle K$ blir null som vist i figur 6.11 a). Siden K er uavhengig av ω , blir Nyquist-“kurven” i dette tilfellet bare et punkt.

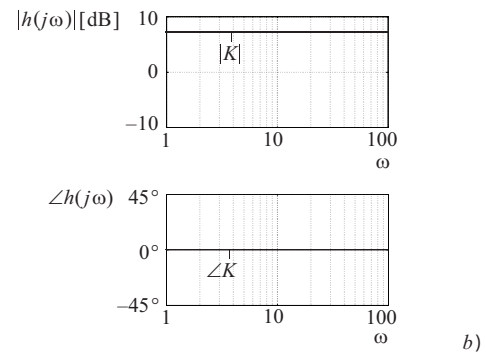
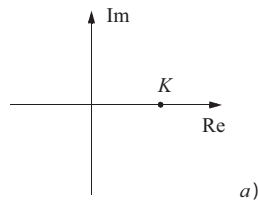
Figur 6.11 b) viser det trivielle AFF-diagrammet for forsterkningen K .

Integratorledd:

Amplituden til et integratorledd (som finnes i nevneren til $h(j\omega)$) blir et enkelt uttrykk, ω^{-1} [dB] = $-20\lg \omega$. Årsaken til dette er at $j\omega$ er rent imaginær i det komplekse plan, som vist i figur 6.12 a). Amplituden avtar med 20 [dB] pr. **dekade** (se figur 6.12 b)). En dekadere representerer en avstand langs den logaritmiske frekvensaksen svarende til en 10-dobling av frekvensen. I Bode-diagrammer bruker vi følgende notasjon for en asymptotes **stigningstall**: -20 [dB]/dekade indikeres med (-1), og -40 [dB]/dekade med (-2) osv.

Figur 6.11

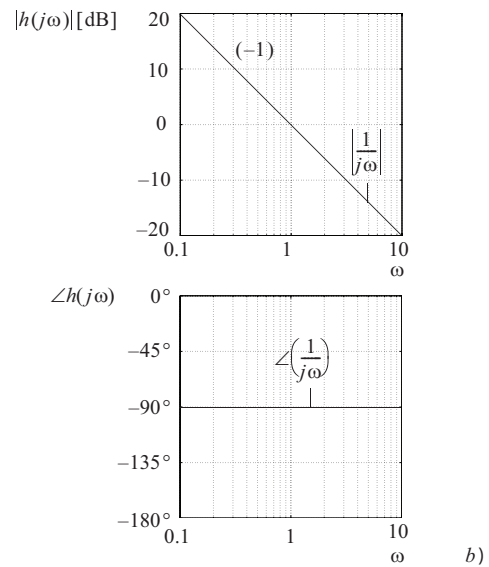
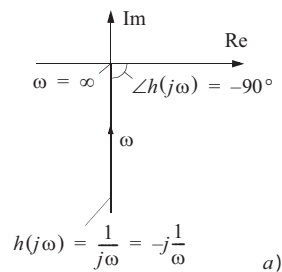
- a) Nyquist-diagram
for forsterkning
- b) AFF-diagram
for forsterkning



Vinkelen til integratorleddet med den reelle aksen i det komplekse plan (fasen), blir -90° som vist i figur 6.12 a).

Figur 6.12

- a) Nyquist-diagram
for integrator
- b) AFF-diagram
for integrator



Derivatorledd:

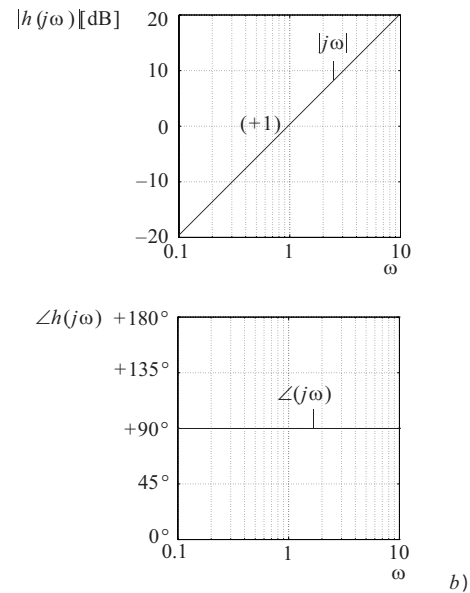
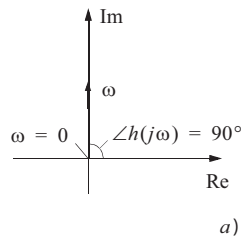
En integrator og en derivator oppveier hverandre i et AFF-diagram (slik de også gjør i en transferfunksjon).

Amplitudebidraget til et derivatorledd (som finnes i telleren til $h(j\omega)$) blir $20\lg \omega$ og gir et stigningstall (+1) i AFF-diagrammet (se figur 6.13 b)).

Fasen til derivatorleddet blir $+90^\circ$ som vist i figur 6.13 a).

Figur 6.13

- a) Nyquist-diagram
for derivator
b) AFF-diagram
for derivator



Tidskonstantledd (reell pol):

Amplitudebidraget til et tidskonstantledd er:

$$\left| \frac{1}{1 + jT_1 \omega} \right| [\text{dB}] = \frac{1}{\sqrt{1 + (T_1 \omega)^2}} [\text{dB}]$$

$$= -20 \lg \sqrt{1 + (T_1 \omega)^2} \begin{cases} \approx 0 [\text{dB}] & , \omega \ll 1/T_1 \\ \approx -20 \lg \sqrt{2} \approx -3 [\text{dB}] & , \omega = 1/T_1 \\ \approx -20 \lg \omega - 20 \lg T_1 & , \omega \gg 1/T_1 \end{cases} \quad (6.21)$$

Likning (6.21) viser at når tidskonstantleddet framstilles som en funksjon av $\lg \omega$, vil vi for svært store og svært små verdier av ω relativt til $1/T_1$ få tilnærmet rette linjer som krysser hverandre i $\omega = 1/T_1$. Linjene er **asymptoter**. Stigningstallet blir 0 for $\omega < 1/T_1$ og (-1) for $\omega \geq 1/T_1$. Dette er vist i figur 6.14 b).

Vinkelen til tidskonstantleddet med den reelle aksene i det komplekse plan blir

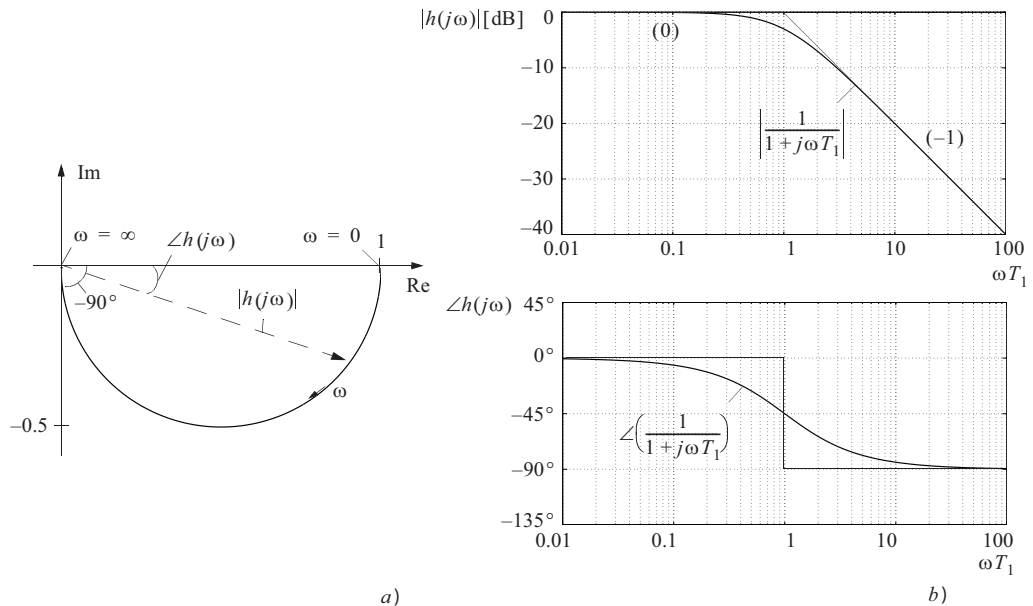
$$\angle \frac{1}{1 + jT_1 \omega} = -\arctan T_1 \omega \begin{cases} \approx 0 & , \omega \ll 1/T_1 \\ \approx -\frac{\pi}{4} = -45^\circ & , \omega = 1/T_1 \\ \approx -\frac{\pi}{2} = -90^\circ & , \omega \gg 1/T_1 \end{cases} \quad (6.22)$$

Hvis vi benytter de to grenseverdiene 0 og $-\pi/2$ som er gitt i (6.22), vil asymptotene bli som vist i figur 6.14 b).

Figur 6.14 a) viser at frekvensresponsen $h(j\omega)$ blir en halvsirkel i det imaginære plan når ω endrer seg mellom $\omega = 0$ og $\omega = \infty$. For $\omega = 1/T_1$ ligger vektorens endepunkt på det laveste stedet på halvsirkelen. Dette tilsvarer at $\angle h(j\omega) = -45^\circ$. Figuren viser også at $\angle h(j\omega)$ aldri blir større enn -90° noe som bekrefter (6.22).

Figur 6.14

- a) Nyquist-diagram for tidskonstantledd
b) AFF-diagram for tidskonstantledd



Resonansledd:

Amplitudebidraget til et resonansledd er:

$$\left| \frac{1}{1 - \left(\frac{\omega}{\omega_0}\right)^2 + j2\zeta \frac{\omega}{\omega_0}} \right| [\text{dB}] = \frac{1}{\sqrt{\left(1 - \left(\frac{\omega}{\omega_0}\right)^2\right)^2 + \left(2\zeta \frac{\omega}{\omega_0}\right)^2}} [\text{dB}] \quad (6.23)$$

$$= -20 \lg \sqrt{\left(1 - \left(\frac{\omega}{\omega_0}\right)^2\right)^2 + \left(2\zeta \frac{\omega}{\omega_0}\right)^2} \begin{cases} \approx 0 [\text{dB}] & , \omega \ll \omega_0 \\ = -20 \lg 2\zeta & , \omega = \omega_0 \\ \approx -40 \lg \omega + 40 \lg \omega_0 & , \omega \gg \omega_0 \end{cases}$$

Likning (6.23) viser at stigningstallet for asymptoten blir 0 for $\omega < \omega_0$ og (-2) for $\omega \geq \omega_0$. Figur 6.15 b) illustrerer dette.

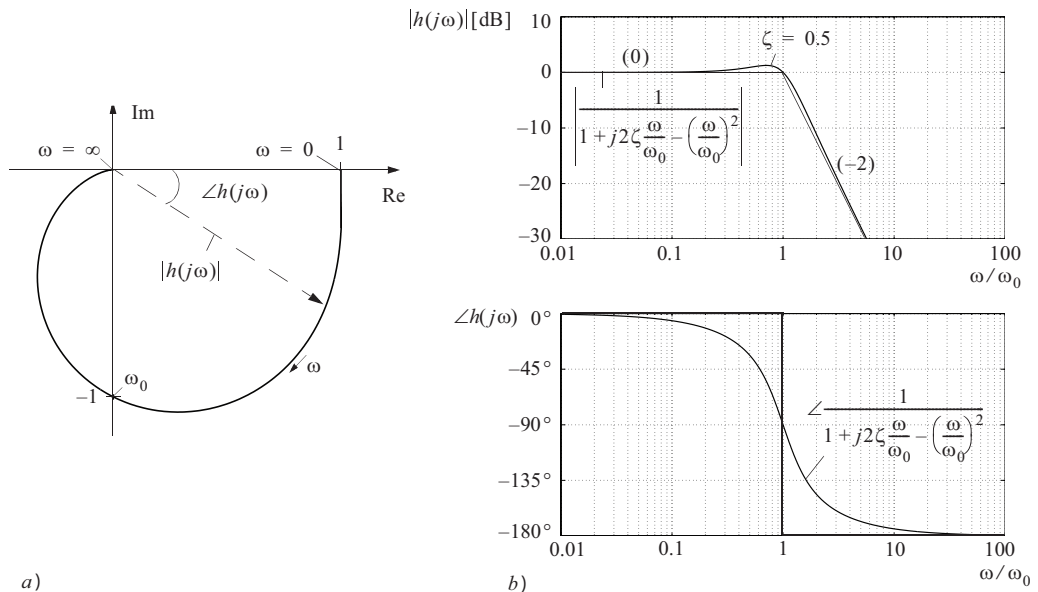
Fasen til resonansleddet blir

$$\angle \frac{1}{1 - \left(\frac{\omega}{\omega_0}\right)^2 + j2\zeta \frac{\omega}{\omega_0}} = -\arctan \frac{2\zeta \frac{\omega}{\omega_0}}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \begin{cases} \approx 0 & , \omega \ll \omega_0 \\ = -\frac{\pi}{2} = -90^\circ & , \omega = \omega_0 \\ \approx -\pi = -180^\circ & , \omega \gg \omega_0 \end{cases} \quad (6.24)$$

Hvis vi benytter de to grenseverdiene 0 og $-\pi$ som gitt i (6.24), vil asymptotene bli som vist i figur 6.15 b). Figur 6.15 a) viser også at $\angle h(j\omega)$ aldri blir mer negativ enn -180° .

Figur 6.15

- a) Nyquist-diagram for resonansleddet
b) AFF-diagram for resonansleddet



Nullpunktsledd:

Amplitudebidraget til et nullpunktsledd er:

$$|1 + j\tilde{T}_1 \omega| [\text{dB}] = 20 \lg \sqrt{1 + (\omega \tilde{T}_1)^2} \begin{cases} \approx 0 & , \omega \ll 1/\tilde{T}_1 \\ = 20 \lg \sqrt{2} & , \omega = 1/\tilde{T}_1 \\ \approx 20 \lg \omega + 20 \lg T_1 & , \omega \gg 1/\tilde{T}_1 \end{cases} \quad (6.25)$$

Likning (6.25) viser at stigningstallet blir 0 for $\omega < 1/\tilde{T}_1$ og (+1) for $\omega \geq 1/\tilde{T}_1$. Figur 6.16 b) viser dette.

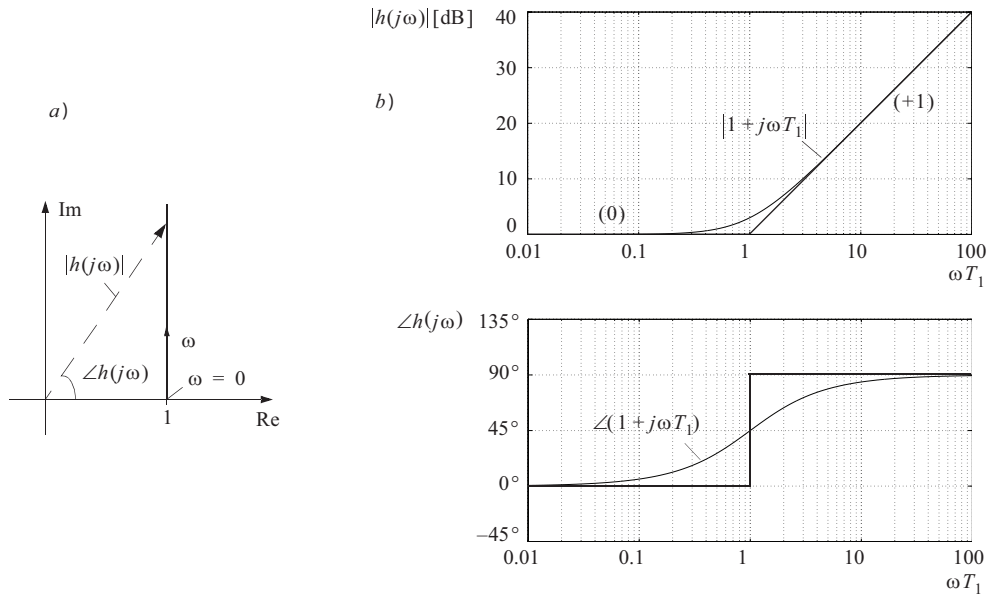
Vinkelen til nullpunktsleddet med den reelle aksen i det komplekse plan blir

$$\angle(1 + j\tilde{T}_1\omega) = \arctan \tilde{T}_1\omega \quad (6.26)$$

Dette resultatet er bortsett fra fortegnet naturlig nok identisk med (6.22). Grenseverdiene blir 0 og $\pi/2$, slik at asymptotene blir som vist i figur 6.16 b). Figur 6.16 a) bekrefter også at $\angle h(j\omega)$ aldri blir større enn $+90^\circ$.

Figur 6.16

- a) Nyquist-diagram for nullpunktsledd
b) AFF-diagram for nullpunktsledd



Tidsforsinkelsesledd:

Amplitudebidraget til et tidsforsinkelsesledd blir ifølge (6.16) en konstant faktor lik 1:

$$|e^{-j\omega\tau}|[\text{dB}] = 20\lg 1 = 0[\text{dB}] \quad (6.27)$$

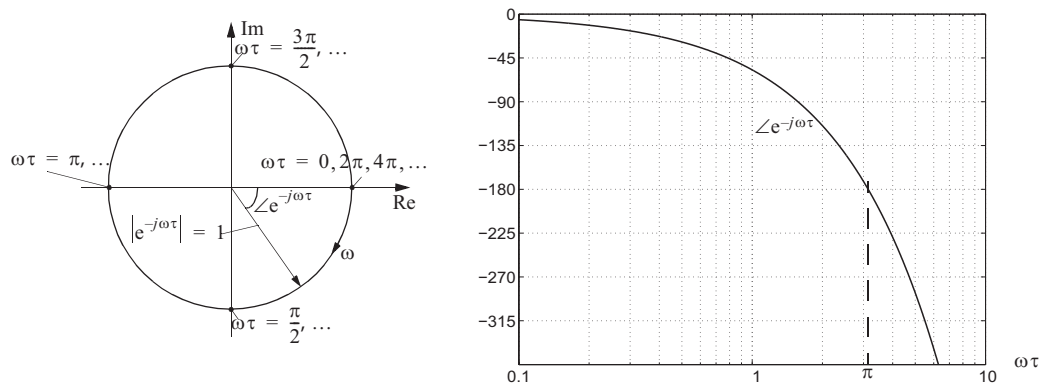
Dette er rimelig, for signalet forsinkes bare. Amplituden blir da uendret uansett frekvens.

Fasen til et tidsforsinkelsesledd er

$$e^{-j\omega\tau} = -\omega\tau \quad (6.28)$$

Det finnes ingen asymptoter for fasekarakteristikken når abscisseaksen er logaritmisk. Med logaritmisk frekvensakse får faseforløpet form som en negativ eksponentialfunksjon. Se figur 6.17, høyre del.

Figur 6.17
Nyquist-diagram
og faseforløp i
Bodediagram
for $e^{-j\omega\tau}$



6.4.1 Prosedyre for tegning av asymptotisk AFF-diagram

Her følger en generell prosedyre som alltid kan brukes:

1. Transferfunksjonen bringes over på formen (6.18). Dersom vi har transferfunksjonen

$$h(s) = \frac{K(\dots)}{(\dots)(a+s)}$$

må den omformes til

$$h(s) = \frac{K'(\dots)}{(\dots)(1+T's)}$$

der $K' = K/a$ og $T' = 1/a$. Dette må gjentas like mange ganger som vi har 1.ordens ledd i $h(s)$. Dersom vi har resonansledd

$$h(s) = \frac{K(\dots)}{(\dots)(c+bs+as^2)}$$

må vi først omforme dette til

$$h(s) = \frac{K'(\dots)}{(\dots) \left(1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0} \right)^2 \right)}$$

der $K' = K/c$ og $\omega_0 = \sqrt{c/a}$. Poenget er altså at alle ledd i nevner (og i teller!) må ha konstantledd lik 1, det vil si at hvert ledd er lik 1 for $s = j0$. Forsterkning i alle ledd “samles” da i en felles K' , fra nå av bare kalt K .

2. Begynn alltid til venstre i diagrammet, ved de laveste frekvensene ($\omega \ll 1$). (Vi symboliserer “liten ω ” med å skrive $\omega \ll 1$.) Ved slike frekvenser kan vi se bort fra virkningen av nullpunkts-, tidskonstant- og resonansledd. Videre prosedyre blir:

Tilfelle a): Hvis vi har q integrasjoner i $h(s)$, $q > 0$, har vi

$$h(j\omega)_{\omega \ll 1} \approx \frac{K}{(j\omega)^q}, \quad |h(j\omega)|_{\omega \ll 1} \approx \frac{K}{\omega^q}, \quad \angle h(j\omega)_{\omega \ll 1} \approx (-q) \cdot 90^\circ$$

q kan selvsagt være lik 0, da er vi over i Tilfelle b) nedenfor.

Amplitudeforløp

Asymptoten vil ved lave frekvenser ha helning $-q$ og skjære 0-dB-linjen i $\omega = K^{1/q}$. Dermed kan vi fastlegge venstre del av $|h(j\omega)|_{as}$ og 0-dB-linjen (subskript as betyr “asymptotisk verdi av”. Se forøvrig figur 6.18).

Faseforløp

Dette er enklere. Asymptoten er horisontal og starter til venstre i $(-q) \cdot 90^\circ$.

Tilfelle b): Ingen rene integrasjoner i $h(s)$

Amplitudeforløp: Da har vi $|h(j\omega)|_{\omega \ll 1} = K$, dvs. horisontal asymptote ved lave frekvenser, i K [dB] over 0-dB-linja.

Faseforløpet starter nå til venstre med en horisontal asymptote i 0° .

3. Videre amplitudeforløp

Asymptotene skal utgjøre en sammenhengende serie av linjestykker. Ta utgangspunkt i den venstre asymptoten som allerede er tegnet, og finn den minste **knekkfrekvensen**.

Knekkfrekvensene er der den asymptotiske kurven for et ledd i en transferfunksjon knekker. De er som vist ovenfor den inverse av tidskonstanten i nullpunkts- og tidskonstantledd, og frekvensen ω_0 i et resonansledd.

Finn helningen p på asymptoten til leddet du betrakter. $p = 1$ for nullpunktsledd, -1 for tidskonstantledd og -2 for resonansledd. Tegn et nytt linjestykke fra den minste knekkfrekvensen fram til den *etterfølgende* knekkfrekvensen med en helning som er lik forrige asymptotes helning $+ p$.

Sett “minste knekkfrekvens” = etterfølgende knekkfrekvens og gjenta prosedyren for neste ledd, osv.

Videre faseforløp

Asymptotene er ikke sammenhengende, men er horisontale, og springer opp eller ned ved knekkfrekvensen, med $p \cdot 90^\circ$ målt fra forrige asymptote.

For faseforløpet gjelder unntaksregelen at et nullpunktsledd av typen $1 - T_i s$ gir 90° knekk *ned*, ikke opp.

EKSEMPEL 6.6: Asymptotisk Bodediagram for en transferfunksjon med flere ledd

Vi ønsker å tegne et asymptotisk amplitudediagram og asymptotisk fasediagram for frekvensresponsen til transferfunksjonen

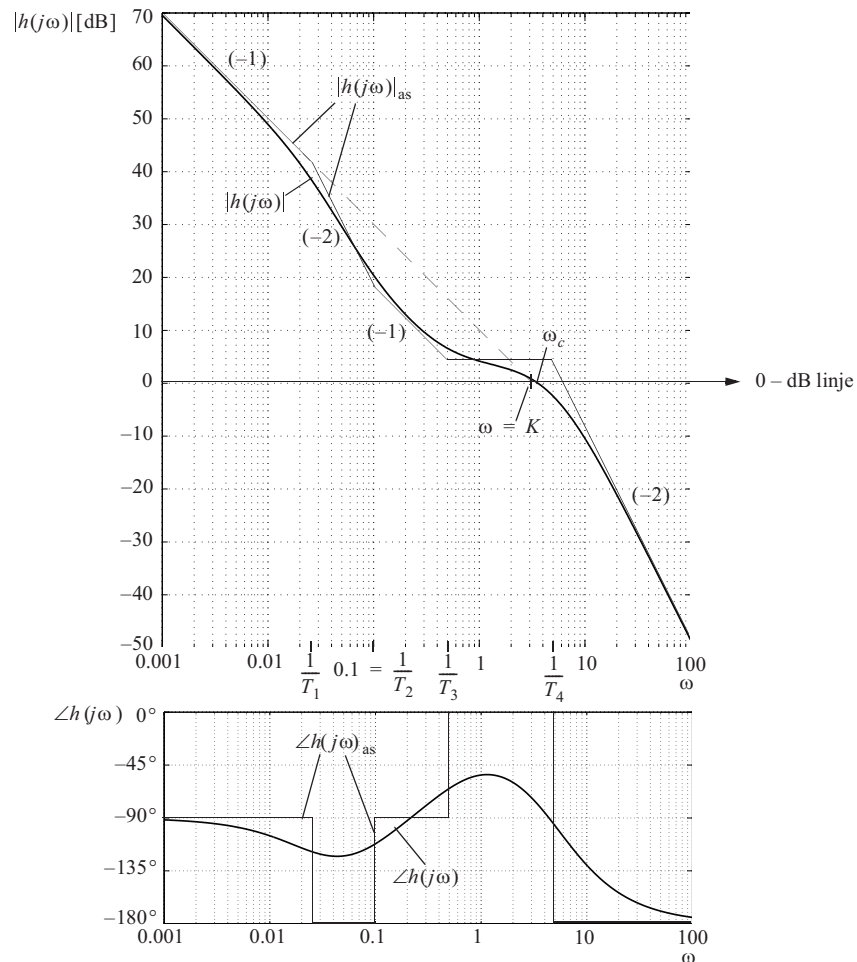
$$h(s) = \frac{K(1 + T_2s)(1 + T_3s)}{s(1 + T_1s)(1 + T_4s)^2} \quad (6.29)$$

der $K = 3$, $T_1 = 40$, $T_2 = 10$, $T_3 = 2$, $T_4 = 0.2$

Vi finner

$$\begin{aligned} \lg|h(j\omega)| = & \lg K + \lg|1 + j\omega T_2| + \lg|1 + j\omega T_3| \\ & - \lg \omega - \lg|1 + j\omega T_1| - 2\lg|1 + j\omega T_4| \end{aligned} \quad (6.30)$$

Figur 6.18
Bodediagram for $h(s)$, eksakt diagram, og asymptotisk diagram



Den totale asymptotiske karakteristikken består av summen av asymptotiske enkeltbidrag av den typen som er beskrevet foran. For transferfunksjonen i (6.29) får vi derfor et resultat som vist i figur 6.18.

Vi følger den oppgitte prosedyren: Vi har $q = 1$ og dermed $|h(j\omega)|_{\omega \ll 1} \approx K/\omega$. Med $K = 3$ får vi $\omega_{c,as} = 3$. Asymptoten har stigningstall (-1), dvs. -20dB/dekade. Dermed kan vi tegne den inn, se figur 6.18.

Ved frekvensen $\omega = 1/T_1 = 0.025$ i figuren vil karakteristikken “knekke av” til vinkelkoeffisient (-2) på grunn av polen i $s = -1/T_1$. Ved $\omega = 1/T_2 = 0.10$ vil den asymptotiske karakteristikken rette seg ut igjen til vinkelkoeffisient (-1) på grunn av nullpunktet i $s = -1/T_2$. Ved $\omega = 1/T_3 = 0.5$ vil vi få en ytterligere utretting til vinkelkoeffisient (0) på grunn av nullpunktet i $s = -1/T_3$. Den doble polen i $s = -1/T_4 = 5$ vil bevirke at den asymptotiske karakteristikken “knekker av” til en (-2) vinkelkoeffisient for $\omega > 1/T_4$. Den virkelige amplitudekarakteristikken vil “sno seg” mellom asymptotene som vist med kraftig strek i figur 6.18. Den asymptotiske karakteristikken avviker mest fra den virkelige i nærheten av knekkpunktene, og spesielt i punktet $\omega = 1/T_4$ der vi har en feil som i verste fall vil være lik 0.3 (i logaritmisk skala) svarende til 6 [dB].

Den asymptotiske og virkelige fasekarakteristikken for frekvensresponsen av transferfunksjonen i (6.29) er vist nederst i figur 6.18. Vi merker oss at både den asymptotiske og virkelige karakteristikken ved meget lave frekvenser begynner ved vinkelen $-\pi/2$ som er vinkelbidraget til en pol i origo (K/s). Forøvrig ser vi at den asymptotiske fasekarakteristikken gjør sprang ved de samme frekvensene som den asymptotiske amplitudekarakteristikk endrer vinkelkoeffisient.

Den asymptotiske fasekarakteristikken ved hver enkelt frekvens er proporsjonal med vinkelkoeffisienten til den asymptotiske amplitudekarakteristikken for systemer av denne arten.

$$\angle h(j\omega)_{as} = \frac{\pi}{2} \frac{d(\lg |h(j\omega)|_{as})}{d(\lg \omega)} = \frac{\pi}{2} k(\omega)_{as} \quad (6.31)$$

der $k(\omega)_{as}$ er vinkelkoeffisienten for den asymptotiske amplitudekarakteristikken.

AFF-diagrammet i figur 6.18 kan genereres ved bruk av MATLAB-koden i figur 6.19.

Figur 6.19
MATLAB-kode for generering av det “skreddersydde” AFF-diagrammet som går igjen i denne boka

```

1 K = 3; T4 = 0.2; T3 = 2; T2 = 10; T1 = 40;
2 t0 = K*conv([T2 1],[T3 1]); ← Transferfunksjonens teller
3 n0 = conv([T1 1 0],[T4^2 2*T4 1]); ← Transferfunksjonens nevner
4 w = logspace(-3,2,1000);
5
6 [amp0, fase0] = bode(t0,n0,w); ← Beregner amplituden og fasen som funksjon av frekvensen
7 semilogx(w, 20*log10(amp0), 'LineWidth', 1.2), grid; ← Plotter amplituden som funksjon av frekvensen
8 set(get(1, 'CurrentAxes'), 'YTick', [-50 -40 -30 -20 -10 0 10 20 30 40 50 60 70])
9 set(get(1, 'CurrentAxes'), 'YLim', [-50 70]) ← Bestemmer punktene som avmerkes på y-aksen
10 pause
11
12 semilogx(w, fase0, 'LineWidth', 2), grid ← Plotter fasen som funksjon av frekvensen
13 set(get(1, 'CurrentAxes'), 'YTick', [-180 -135 -90 -45 0])
14 set(get(1, 'CurrentAxes'), 'YLim', [-180 0]) ← Begrenser y-aksen

```

MATLAB har også et “standard” (etter vår mening dårligere) Bodediagram som genereres som vist i figur 6.20.

Figur 6.20
MATLAB-kode for generering av “standard” AFF-diagram

```

1 K = 3; T4 = 0.2; T3 = 2; T2 = 10; T1 = 40;
2 t0 = K*conv([T2 1],[T3 1]); ← Transferfunksjonens teller
3 n0 = conv([T1 1 0],[T4^2 2*T4 1]); ← Transferfunksjonens nevner
4 w = logspace(-3,2,1000);
5
6 bode(t0,n0,w) ← Beregner og plotter amplituden og fasen som funksjon av frekvensen

```

■

6.4.2 Ikke-minimum-fase-systemer

Likning (6.31) antyder at for visse typer systemer er det en entydig sammenheng mellom amplitudekarakteristikken og fasekarakteristikken. H. W. Bode (1945) har påvist at blant alle systemer som har en og samme amplitudekarakteristikk $|h(j\omega)|$, finnes det ett som har den minst negative fasekarakteristikken $\angle h(j\omega)$. Dette systemet kalles et **minimum-fase-system**, og oppfyller (6.31). Nødvendige og tilstrekkelige betingelser for at et stabilt system skal være et minimum-fase-system, er at alle nullpunktene og polene til transferfunksjonen ligger i venstre halvplan, og at $h(s)$ ikke inneholder en tidsforsinkelse.

En lang rekke fysiske fenomener leder til transferfunksjoner som medfører at fasekarakteristikken i sin helhet ligger lavere (mer negativ) enn den et minimum-fase-system har. Et slikt system kalles et **ikke-minimum-fase-system** (IMF-system).

Prosesser med ikke-minimum-fase-egenskaper er vanskeligere å regulere enn minimum-fase-prosesser på grunn av den større negative faseforskyvingen. Dette vil framgå av senere kapitler.

En type prosesser som er ikke-minimum-fase, er slike som inneholder en “**transportforsinkelse**” (tidsforsinkelse), som kan beskrive fysisk transport av stoff i et rør eller på et transportbånd eller transmisjon av signaler ved radiobølger eller lydbølger. Transferfunksjonen for en transportforsinkelse er $h(s) = e^{-\tau s}$. Frekvensresponsen for denne blir

$$|h(j\omega)| = 1 \quad \text{og} \quad \angle h(j\omega) = -\omega\tau$$

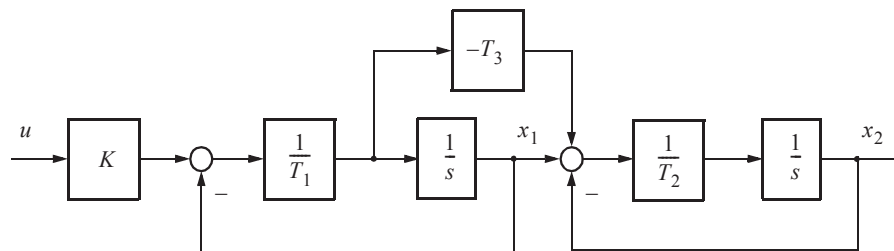
og er allerede omtalt, se figur 6.17. Dette er den type IMF-prosess som er vanskeligst å regulere.

En annen type ikke-minimum-fase-prosess oppstår når vi har et eller flere nullpunktsledd av typen $(1 - T_i s)$ i transferfunksjonen, dvs. **nullpunkt i høyre halvplan**¹. Da gjelder ikke (6.31) lenger, fordi $\angle h_{as}$ nå knekker *ned* -90° ved $\omega = 1/T_i$, selv om $|h|_{as}$ knekker *opp* ved samme frekvens.

EKSEMPEL 6.7: System med nullpunkt i høyre halvplan

Figur 6.21 viser Laplacetransformert blokkdiagram for et slikt system.

Figur 6.21
Blokkdiagram for system med nullpunkt i høyre halvplan



Vi har her to seriekoblede 1. ordens “prosesser” og i tillegg har vi en “negativ foroverkopling” av raten (den deriverte) av tilstandsvariabelen x_1 . Vi kan sette opp transferfunksjonen direkte ved

$$x_2(s) = \frac{1}{1 + T_2 s} (x_1(s) - T_3 s x_1(s)) = \frac{1}{(1 + T_2 s)(1 + T_1 s)} (1 - T_3 s) K u(s)$$

¹Dette kalles også (noe misvisende) for et “negativt nullpunkt”.

Dette gir

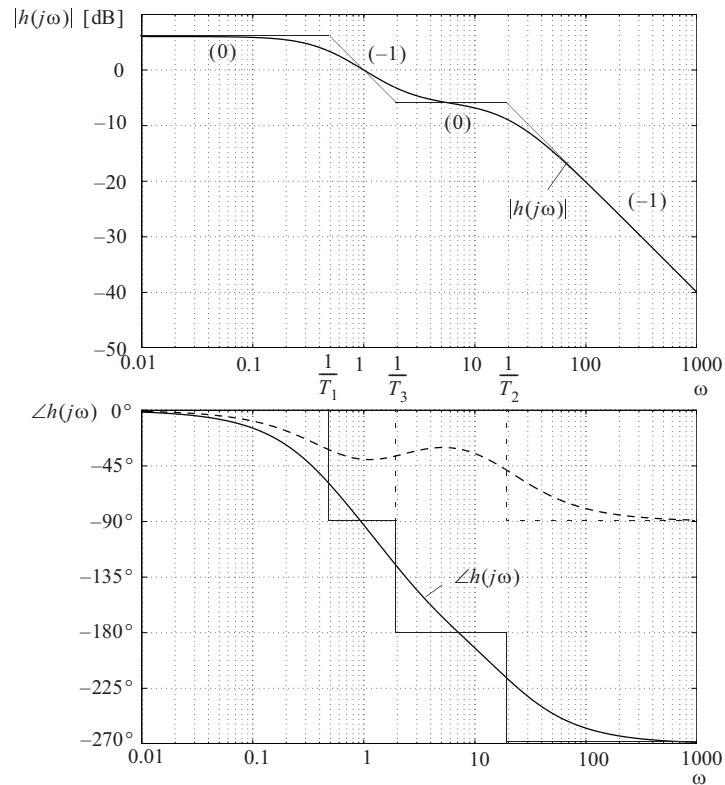
$$\begin{aligned} \frac{x_2(s)}{u(s)} = h(s) &= \frac{K(1 - T_3s)}{(1 + T_1s)(1 + T_2s)}, \text{ der } K = 2, T_1 = 2, T_2 = 1/20 \text{ og } T_3 = 1/2 \\ &= \frac{2(1 - 0,5s)}{(1 + 2s)(1 + 0,05s)} \end{aligned} \quad (6.32)$$

Figur 6.22 viser AFF-diagrammet for (6.32) når $T_1 > T_3 > T_2$.

Karakteristisk for dette AFF-diagrammet er at nullpunktet i høyre halvplan gitt av $s = +1/T_3$ medfører et fasebidrag på -90° . Med et tilsvarende nullpunkt i venstre halvplan ville vi fått et positivt fasebidrag på $+90^\circ$ som vist stiplet i figur 6.22.

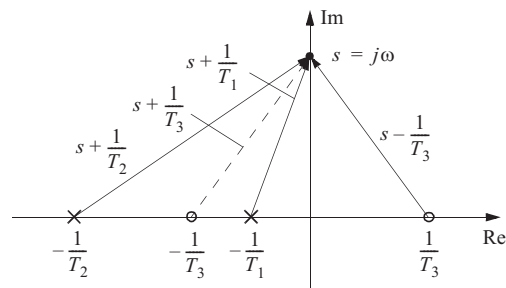
Figur 6.23 viser de vektorer som $h(j\omega)$ fra (6.32) består av. Vi ser at vektorlengdene er uavhengige av hvorvidt nullpunkter eller poler ligger i høyre eller venstre halvplan, mens fasevinkelen er negativ for en vektor med utgangspunkt i høyre halvplan.

Figur 6.22
AFF-diagram
for (6.32)

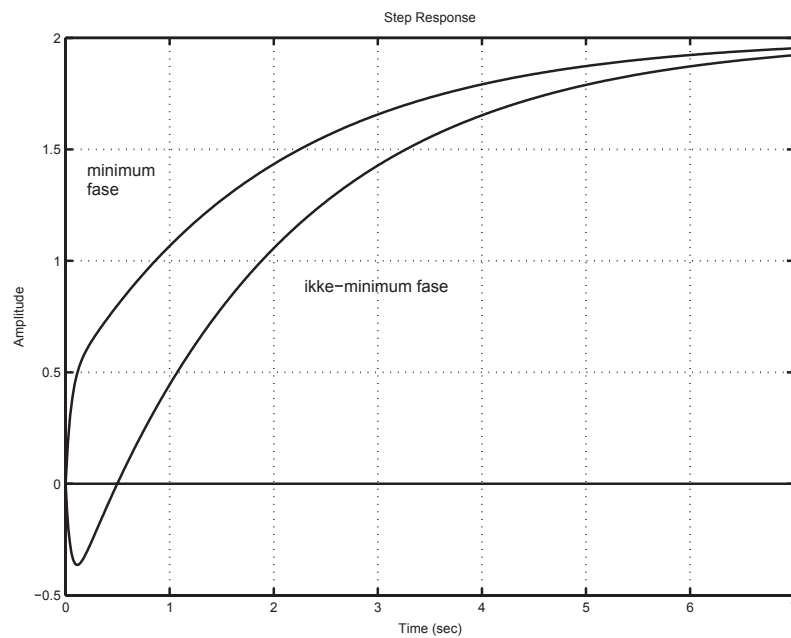


Figur 6.23

Vektorene til $h(j\omega)$ i det komplekse plan

**Figur 6.24**

Sprangrespons eksempel 6.7

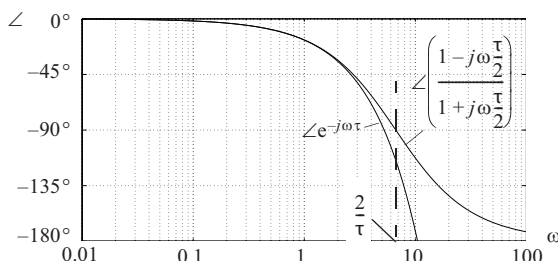


La oss nå betrakte en 1. ordens rasjonal tilnærming til en transportforsinkelse, se også (4.152). Vi ser at tilnærminga inneholder et nullpunkt i høyre halvplan, noe som indikerer slektskapet mellom de to formene for ikke-minimum-fase-ledd.

$$h(s) = e^{-\tau s} \approx \frac{1 - \frac{\tau}{2}s}{1 + \frac{\tau}{2}s} \quad (6.33)$$

Amplitudeforløpet for approksimasjonen svarer eksakt til amplitudeforløpet for $e^{\tau s}$ (hvorfor?). Faseforløpet for denne enkleste rasjonale tilnærmingen er vist i figur 6.25. Bedre tilnærminger behandles i eksemplene 8.10 - 8.12

Figur 6.25
AFF-diagram
for (6.33)



Ikke-minimum-fase-prosesser er uheldige i tilbakekoblede systemer fordi den økte negative fasevinkelen medfører at reguleringssystemet må gjøres langsomt for å unngå ustabilitet. Dette skal vi omtale mer i etterfølgende kapitler.

Et annet forhold er at såkalte “inverterende regulatorer” ikke kan realiseres for ikke-minimum-fase prosesser. En “inverterende regulator” har en transferfunksjon som inneholder den inverse av prosessens transferfunksjon. I tilfellet av en transportforsinkelse må regulatoren inneholde et ledd av typen $h_u^{-1}(s) = e^{\tau s}$. Dette betyr at regulatoren må utføre *prediksjon*, det vil si den *må gi respons før den er eksitert*. Dette kan ikke realiseres i en fysisk regulator.

Til slutt nevnes at regulering v.h.a. datamaskin, hvor målingene må tases (eng.: “sampling”), medfører en ikke-minimum-fase-virkning i reguleringssløyfen (se avsnitt 11.2.2).

6.5 SKISSERING AV TIDSRESPONS PÅ BASIS AV FREKVENSPONS

Ofte vil vi være interessert i å skissere en typisk tidsrespons, for eksempel en sprangrespons, ut fra kunnskaper om et lineært systems frekvensrespons.

Frekvensresponsen kan tenkes å være bestemt eksperimentelt. Vi påtrykker da et sinusformet pådrag og registrerer amplitudeforholdet og faseforskyvningen til responsen i forhold til pådraget. Dette gjentas for en rekke forskjellige frekvenser i et frekvensområde som er betydningsfullt i forhold til systemets dynamiske parametre.

En annen og kanskje langt viktigere situasjon er at frekvensresponsen er bestemt teoretisk (gjerne grafisk) på basis av kjennskap til frekvensresponsen for enkeltdeler og framfor alt for systemer der tilbakekopling inngår (se kapittel 7).

Den metoden som utvikles her har to trinn:

1. Finn fram til en transferfunksjon som med rimelig grad av nøyaktighet vil gi en frekvensrespons lik den som foreligger.
2. Finn systemets tidsrespons (for eksempel sprangrespons) ut fra kjennskap til denne transferfunksjonen.

Siden vi ikke vet om det systemet hvis frekvensrespons vi kjenner er et minimum-fase-system, må vi være nøye med å benytte både den informasjonen som ligger i amplitudekarakteristikken og i fasekarakteristikken.

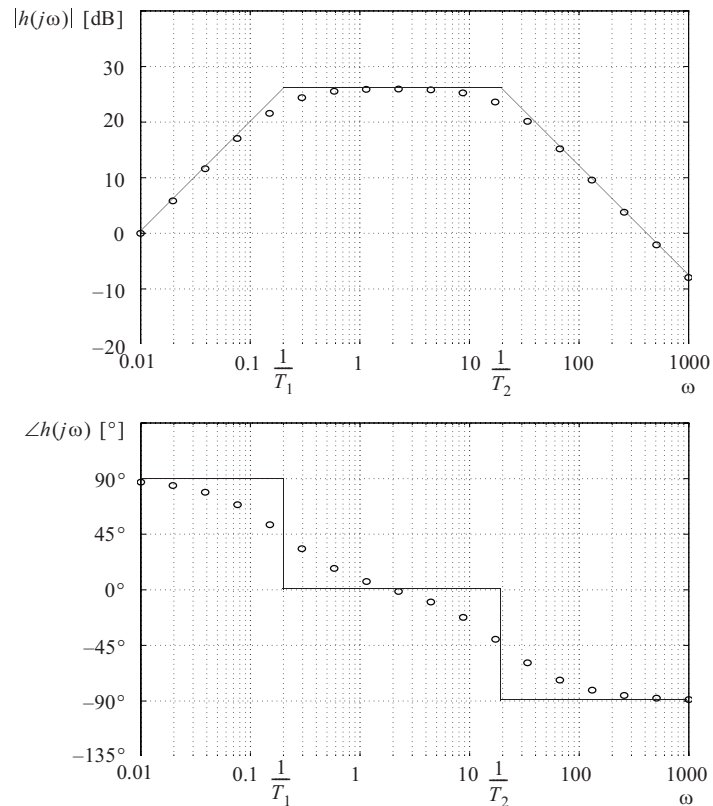
Dersom systemet kan beskrives med en rasjonal transferfunksjon (minimum-fase eller ikke-minimum-fase), skulle vi ved hjelp av asymptotiske karakteristikk og endel kurvetilpasning relativt lett finne en transferfunksjon ut fra frekvensresponsen. Inneholder systemet derimot fysiske fenomener som ikke lar seg beskrive med en rasjonal transferfunksjon, for eksempel en transportforsinkelse ($e^{-\tau s}$), vil vi som oftest ikke lykkes med å få noen god tilpasning. Dette gjelder særlig ved høye frekvenser. Det er da fordelaktig å bygge opp den "syntetiske" transferfunksjonen som produktet av en rasjonal transferfunksjon og en ikke rasjonal funksjon av den type som kan forventes, for eksempel $e^{-\tau s}$.

Hvilken nøyaktighet vi trenger i tilpasningen av frekvensresponsen av den "syntetiske" transferfunksjonen og den gitte frekvensresponsen avhenger av mange forhold. Som oftest ville vi være fornøyd med en relativt lav nøyaktighet når hensikten bare er å få en hurtig oversikt over de vesentligste egenskapene til tidsresponsen. Vi skal imidlertid være oppmerksom på at kravet til tilpasningsnøyaktighet ved de forskjellige frekvensene er helt forskjellig avhengig av hvilken form det aktuelle pådragets tidsforløp har (for eksempel sprang).

EKSEMPEL 6.8: Skissering av tidsrespons 1

Vi ser først på det tilfellet som er illustrert i figur 6.26.

Figur 6.26
Amplitude og fase
for noen frekvenser
og asymptotisk
tilnærming



Amplitudekarakteristikken og fasekarakteristikken for den gitte frekvensresponsen ($h(j\omega)$) er angitt med små sirkler for en lang rekke frekvenser. Vi legger merke til at fasekarakteristikken ved lave frekvenser later til å starte ved ca. $+90^\circ$, og at den ved høye frekvenser ser ut til å ende opp ved -90° . Tegner vi inn to asymptoter i amplitudediagrammet med henholdsvis $+1$ vinkelkoeffisient tilsvarende $+20$ [dB/dek] og -1 vinkelkoeffisient tilsvarende -20 [dB/dek], vil den gitte amplitudekurven nærme seg de to asymptotene ved henholdsvis lav og høy frekvens. Det er derfor rimelig å anta at vi her står overfor et relativt enkelt tilfelle med et system av minimum-fase-type.

Vi forsøker oss med en asymptotisk tilnærming på amplitudekurven, og ser at vi vil få fram to knekkpunkter ved frekvensene $\omega = 1/T_1$ og $\omega = 1/T_2$. En kontroll av fasekarakteristikken viser at den gitte fasevinkelen ved de to knekkfrekvensene er henholdsvis ca. $+45^\circ$ og -45° . Siden de to knekkfrekvensene ligger ca. 2 dekadere fra hverandre, vil dette resultatet være rimelig.

En “syntetisk” transferfunksjon som vil ha en frekvensrespons med asymptotiske karakteristikk, som skissert i figur 6.26, vil være

$$h_1(s) = \frac{KT_1s}{(1 + T_1s)(1 + T_2s)} \quad (6.34)$$

der $K = 20 = 26$ [dB], $T_1 = 1/(0.2) = 5$ og $T_2 = 1/20 = 0.050$.

Når vi skal tilpasse asymptotene for den “syntetiske” transferfunksjonens frekvensrespons til den gitte frekvensresponsen, er det viktig å ta hensyn til at de forskjellige knekkpunktene på kurvene innvirker på hverandre (dette framgår av appendiks D som tar for seg korreksjonskalaer).

Transferfunksjonen i (6.34) med de gitte tallverdiene har asymptotisk frekvensrespons som inntegnet med streker i figur 6.26. Den stemmer godt overens med de gitte dataene (sirklene).

Siden vi er kommet fram til transferfunksjonen $h_1(s)$, kan vi bestemme tidsresponsen av et eller annet pådrag ved hjelp av invers Laplacetransformasjon. Ofte er vi interessert i sprangresponsen og velger derfor $u(s) = 1/s$

$$y_1(s) = h_1(s)u(s) = \frac{KT_1}{(1 + T_1s)(1 + T_2s)} \quad (6.35)$$

Inverstransformen $y_1(t)$ av dette uttrykket finner vi i dette tilfellet lett ved delbrøksoppspaltning og deretter invers transform av hvert enkelt ledd.

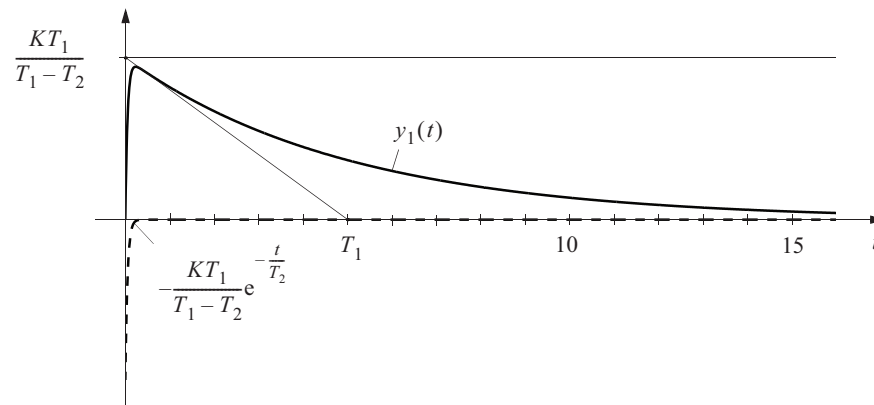
$$y_1(s) = \frac{KT_1^2}{1 + T_1s} - \frac{KT_1T_2}{1 + T_2s} \quad (6.36)$$

som gir

$$y_1(t) = \frac{KT_1}{T_1 - T_2} e^{-\frac{t}{T_1}} - \frac{KT_1}{T_1 - T_2} e^{-\frac{t}{T_2}} \quad (6.37)$$

Innholdet av (6.37) er vist i figur 6.27.

Figur 6.27
Tidsresponsen
 $y_1(t)$



Vi legger merke til at responsen $y_1(t)$ til å begynne med arter seg som om systemet bare skulle hatt en reell pol i $s = 1/T_2$. Etter hvert kommer imidlertid virkningen av mekanismene som gjør at den asymptotiske amplitudekarakteristikken går med en $+1$ vinkelkoeffisient fra 0 frekvens og opp til $\omega = 1/T_1$. Dette medfører at vi ikke får noen stasjonær respons, det vil si at sprangresponsen går mot 0 når t øker mot uendelig. “Halen” på sprangresponsen er karakterisert ved tidskonstanten T_1 .

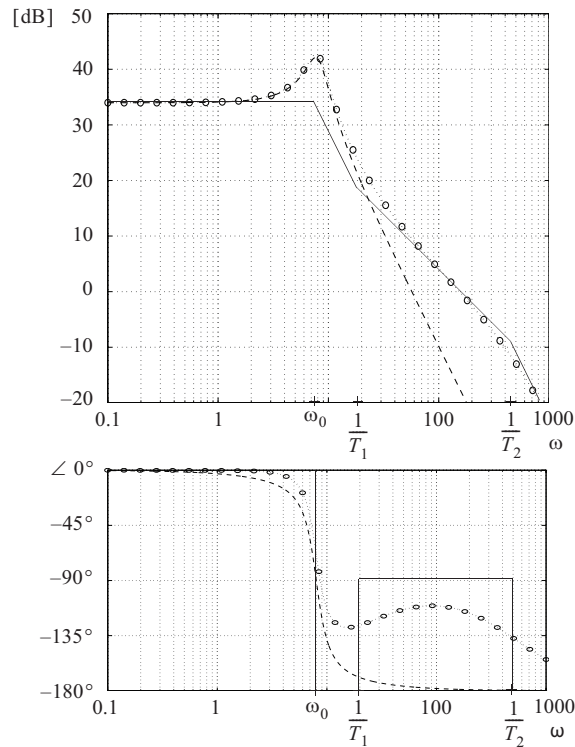
■

EKSEMPEL 6.9: Skissering av tidsrespons 2

Figur 6.28 viser med sirkler amplituden og fasen til en gitt frekvensrespons. Vi står her, så langt diagrammet viser, overfor et system som i alle fall ikke har noe nullpunkt i origo slik som systemet i eksempel 6.8. Ved lave frekvenser har vi en forsterkning på 34 [dB] tilsvarende 50. Det er videre tydelig at vi omkring frekvensen $\omega = \omega_0$ har en resonans, hvilket skulle tyde på at vi har to kompleks konjugerte poler. Dersom systemet er et minimum-fase-system, skulle vi med bare to kompleks konjugerte poler få en fasekarakteristikk som går mot -180° , slik som vist for eksempel i figur 6.15. Amplitudekarakteristikken skulle dessuten nærme seg en asymptote med vinkelkoeffisient -2 . Her ser det ut til at den ved frekvenser høyere enn resonansfrekvensen nærmer seg mer en vinkelkoeffisient -1 til å begynne med, for så senere å bøye mer nedover.

Dette tyder på at vi i tillegg til de to kompleks konjugerte polene også har et nullpunkt (i venstre halvplan) ved en frekvens $\omega = 1/T_1$ som ligger noe høyere enn ω_0 . Videre er det tydelig at vi har en ekstra pol som kommer inn ved frekvensen $\omega_2 = 1/T_2$. Hvorvidt vi har flere poler er det ikke mulig å si på grunn av at vi mangler data for høyere frekvenser enn vist i figur 6.28.

Figur 6.28
Amplitude og fase
for noen frekvenser
og asymptotisk
tilnærming



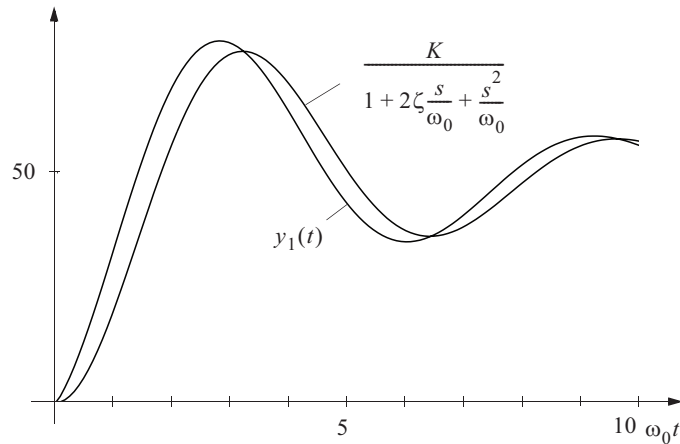
Resultatet av disse betraktningene leder til følgende “syntetiske” transferfunksjon

$$h_1(s) = \frac{K(1 + T_1 s)}{\left(1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2\right) (1 + T_2 s)} \quad (6.38)$$

Vanskeligheten er nå å bestemme konstantene i (6.38). Vi kjenner allerede til at $K \approx 50$. Vi går så som en første tilnærming ut fra at den udempede resonansfrekvensen ω_0 omtrent sammenfaller med resonanstoppen i figur 6.28 og finner $\omega_0 = 8$. Amplitudekurven viser at resonanstoppen ligger ca. 10 [dB] høyere enn det horisontale partiet. Anslår vi nå at nullpunktet i ω_1 kan bidra til å løfte karakteristikken i $\omega = \omega_0$ med omtrent 2 [dB], blir resonanstoppen ca. lik 8 [dB], hvilket tilsvarer $\zeta \approx 0.2$ (se for eksempel figur 6.8). De to kompleks konjugerte polenes fasekarakteristikk er stiplet i figur 6.28 nederst. Herav kan vi se hvor nullpunktet må legges, ved å finne frekvensen der differansen mellom de to fasekurvene er 45° . Dette ser vi opptrer omtrent ved frekvensen $\omega_1 = 20$. Frekvensen ω_2 finner vi omtrent der hvor den virkelige fasekarakteristikken har verdien -135° . Dette er ved ca. $\omega_2 = 500$.

Den asymptotiske amplitudekarakteristikken og fasekarakteristikken for frekvensresponsen til $h_1(s)$ med de funne tallverdiene er vist heltrukket i figur 6.28. Den virkelige frekvensresponsen til $h_1(s)$ er vist prikket i figuren. Vi ser at overensstemmelsen ved de gitte punktene er god.

Figur 6.29
Sprangrespons av
 $y_1(t)$



Sprangresponsen til den “syntetiske” transferfunksjon i (6.38) vil bli av samme art som de i figur 4.36, når vi ser bort fra polen i $s = 500$. Innvirkningen av denne polen vil bare bli en liten avrunding ved $t \approx 0$. For å finne innsvingningsforløpets form i detalj, kan vi benytte oss av resultatene i figur 4.20. Vi finner da at sprangresponsen må få en form som vist i figur 6.29.

De viktigste egenskapene ved denne sprangresponsen er størrelsen og tidspunktet for den første toppen, oscillasjonsfrekvensen og dempingen av oscillasjonene.

■

EKSEMPEL 6.10: Skissering av tidsrespons 3

Figur 6.30 viser amplituden og fasen for en gitt frekvensrespons med sirkler.

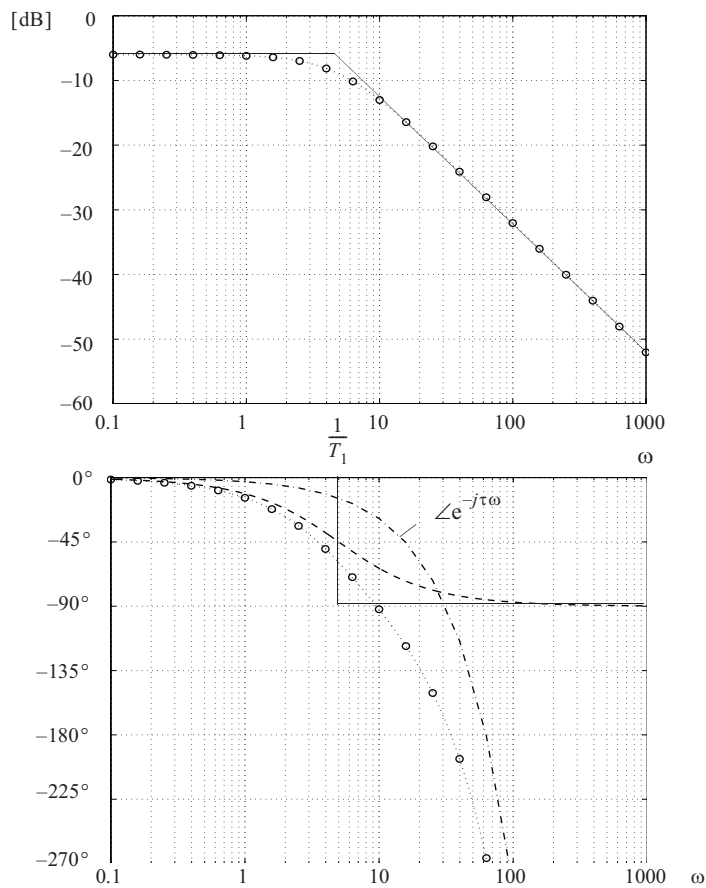
Forsøk med asymptoter på amplitudekarakteristikken antyder at asymptotene med vinkelkoeffisientene 0 og -1 gir god overensstemmelse, og vi finner et knekkpunkt ved $\omega = 1/T_1 = 5$. Dette kunne tyde på at systemet har en reell pol med tidskonstant $T_1 = 0.2$. Fasekarakteristikken tyder imidlertid på at vi her ikke står overfor et minimum-fase-system fordi fasevinkelen er atskillig mer negativ enn det som tilsvarer en reell pol. Vi antar derfor som en første tilnærming at den “syntetiske” transferfunksjonen er av typen

$$h_1(s) = \frac{Ke^{-\tau s}}{1 + T_1 s} \quad (6.39)$$

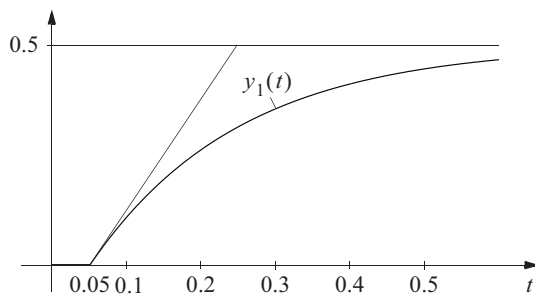
Av den horisontale asymptoten i amplitudediagrammet finner vi $K \approx 0.5 = -6$ [dB]. Tegner vi den tilhørende fasekarakteristikken til polen med tidskonstant $T_1 = 0.2$, som vist stiplet nederst i figur 6.30, ser vi at det manglende negative fasebidraget omtrent kan framskaffes ved å velge $\tau = 1/\omega_2 \approx 1/20$. Dette bidraget er vist strekpunkttert nederst i figur 6.30. De resulterende, virkelige amplitude- og fasekarakteristikkene for dette “syntetiske” systemet er vist prikket. Vi ser at overensstemmelsen er god.

Med de funne tallverdiene kan vi nå finne sprangresponsen av transferfunksjonen i (6.39). Denne blir som vist i figur 6.31.

Figur 6.30
Amplitude og fase
for noen frekvenser
og asymptotisk
tilnærming



Figur 6.31
Sprangresponsen
 $y_1(t)$



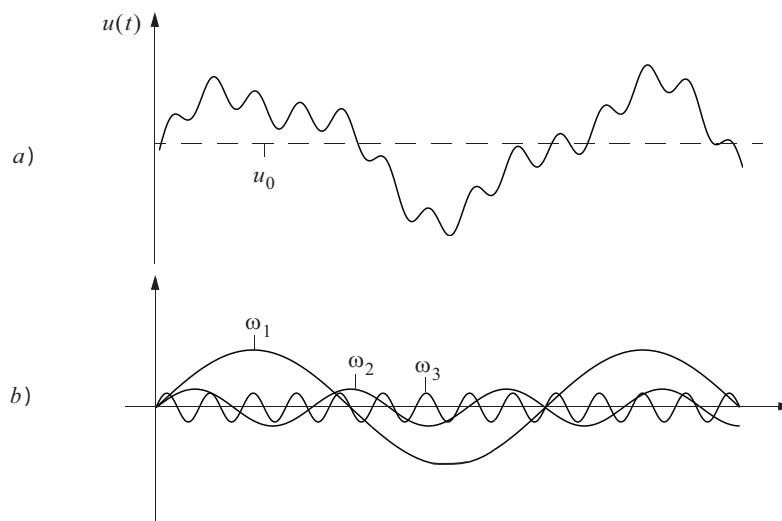
6.5.1 Effektivverdien av et lineært systems respons på et stasjonært “tilfeldig” pådrag

Mange fysiske systemer, for eksempel i industrielle prosesser, utsettes for påvirkninger som har en tilsynelatende “tilfeldig” eller “stokastisk” natur. En detaljert behandling av slike problemer hører ikke hjemme i denne teksten, men en tilnærmet oversikt over hva vi kan forvente av resultater kan fåes ved å benytte teknikken som beskrives i det følgende.

Vi antar at det foreligger en registrering av et typisk utsnitt av det “tilfeldige” pådraget (eller forstyrrelsen) som virker på prosessen, og vi antar at en slik registrering ser ut som vist øverst i figur 6.32.

Figur 6.32

- a) Stokastisk pådrag
 b) Tilnærmet signal bestående av en rekke frekvenser



Det ser her ut som om vi har en variasjon omkring en middelvei. Vi tenker oss nå at vi har klart å finne en rekke av sinussvingninger med forskjellig frekvens, amplitude og fase som til sammen gir en rimelig tilnærming til det foreliggende tidsforløpet. Vi skal da være oppmerksom på at om vi hadde tatt et annet utsnitt av det samme tidsforløpet, ville vi kanskje komme fram til en litt annen fordeling av frekvenser, amplituder og fasevinkler på de rene sinussvingningene. Det er derfor ikke svært nøye med presisjonen i bestemmelse av sinussvingningene. Svaret kan i alle fall ikke bli annet enn tilnærmet.

Vi antar derfor at vi har funnet fram til følgende tilnærming

$$u(t) \approx u_0 + u_1 \sin(\omega_1 t + \varphi_1) + u_2 \sin(\omega_2 t + \varphi_2) + \dots + u_m \sin(\omega_m t + \varphi_m) \quad (6.40)$$

der den høyeste frekvensen vi finner det nødvendig å ta med i beskrivelsen, er ω_m . Vi legger merke til at frekvensene $\omega_1, \omega_2, \dots$ ikke behøver å være **harmoniske frekvenser** (hele multiplum av en grunnfrekvens) slik resultatet vil bli dersom vi foretar en Fourier-rekkeutvikling av utsnittet ovenfor. En Fourier-rekkeutvikling vil imidlertid være en måte å komme fram til (6.40) på.

Dersom $u(t)$ påtrykkes et lineært system med kjent frekvensrespons $h(j\omega)$, kan vi direkte bestemme den stasjonære responsen $y(t)$ som

$$y(t) = u_0|h(j0)| + u_1|h(j\omega_1)|\sin(\omega_1 t + \varphi_1 + \angle h(j\omega_1)) + \dots + u_m|h(j\omega_m)|\sin(\omega_m t + \varphi_m + \angle h(j\omega_m)) \quad (6.41)$$

Likning (6.41) gjelder fordi responsen av de forskjellige svingningene i pådraget kan superponeres i et lineært system.

Nå er det imidlertid forbundet med atskillig arbeid å regne ut summen av alle leddene i (6.41) for eventuelt å tegne opp $y(t)$ i detalj. Dette interesserer imidlertid neppe i særlig grad fordi vi som sagt foran har tatt et vilkårlig utsnitt av pådraget. Det som derimot kunne interessere oss måtte være å bestemme **effektivverdien** av responsens svingning omkring sin middelvei. Betegner vi middelveien med symbolet \bar{y} , finner vi

$$\bar{y}(t) = u_0|h(j0)| \quad (6.42)$$

Effektivverdien av den resterende del av (6.41) betegner vi Δy_{eff} og vi har

$$\begin{aligned} \Delta y_{eff} &= \sqrt{(y(t) - \bar{y}(t))^2} \\ &= \sqrt{1/2(u_1^2|h(j\omega_1)|^2 + u_2^2|h(j\omega_2)|^2 + \dots + u_m^2|h(j\omega_m)|^2)} \end{aligned} \quad (6.43)$$

Likning (6.43) viser at vi for å bestemme effektivverdien av responsens svingning omkring sin middelvei, bare behøver å kjenne amplitudene av de forskjellige svingningene i pådraget og frekvensresponsens tallverdi ved disse svingningenes frekvenser. De innbyrdes fasevinklene mellom svingningene i pådraget og frekvensresponsens fasevinkel kommer ikke inn. Hadde vi i tillegg til effektivverdien av responsen også bestemt effektivverdien av pådragets svingning omkring sin middelvei, ville vi ha fått

$$\begin{aligned} \Delta u_{eff} &= \sqrt{(u(t) - \bar{u}(t))^2} \\ &= \sqrt{1/2(u_1^2 + u_2^2 + \dots + u_m^2)} \end{aligned} \quad (6.44)$$

På basis av (6.43) og (6.44) kunne vi nå kanskje tillate oss å definere en størrelse med betegnelsen “**effektivverdiforholdet**” analogt til størrelsen “*amplitudeforholdet*” slik:

$$h_{eff} = \frac{\Delta y_{eff}}{\Delta u_{eff}} = \sqrt{\frac{u_1^2 |h_1(j\omega_1)|^2 + \dots + u_m^2 |h_m(j\omega_m)|^2}{u_1^2 + \dots + u_m^2}} \quad (6.45)$$

Vi ser at h_{eff} er en funksjon av fordelingen av de forskjellige frekvenskomponentenes amplituder i pådraget og frekvensresponsens tallverdi ved de forskjellige frekvensene. Den er med andre ord både signal- og systemavhengig. h_{eff} er et mål for spredning av responsen omkring sin middelvei. Den sier imidlertid ingen ting om frekvensinnholdet eller tidsforløpet av responsen. Det gjør derimot de enkelte amplitudene som inngår i hvert ledd i (6.41). Et eksempel vil illustrere disse forholdene.

EKSEMPEL 6.11: Sammenheng i amplitude mellom pådrag og respons for visse frekvenser

Figur 6.33 viser tallverdien av frekvensresponsen til et system som påvirkes av et pådrag der svingeamplitudene ved de forskjellige frekvensene ω_1 , ω_2 , ω_3 og ω_4 , som er funnet å være dominerende, er angitt ved desibelverdier relativt en gitt referanse.

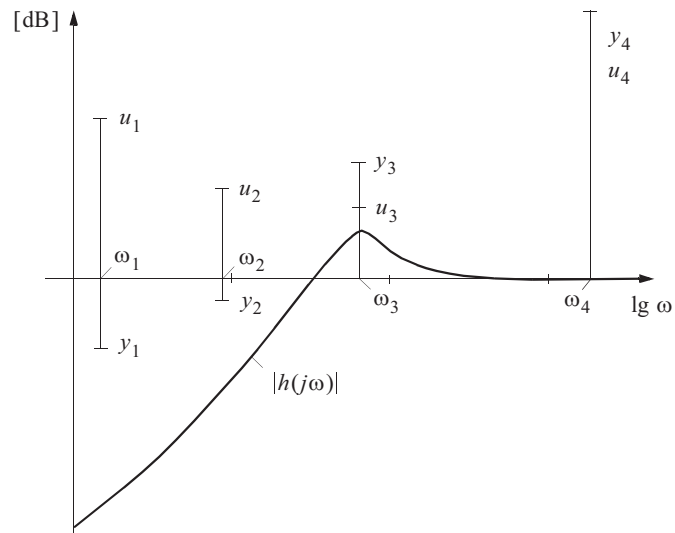
Vi ser at pådraget u har den største amplituden ved frekvensen ω_4 og den minste ved ω_3 .

Når amplituden til pådragssvingningen i [dB] *adderer* til verdien av $|h(j\omega)|$ [dB] ved en enkelt frekvens, framkommer amplituden (i [dB]) av responsens svingninger for den samme frekvensen.

Amplitudene til responsen for de gitte frekvensene er gitt betegnelsene y_1 , y_2 , y_3 og y_4 . De er inntegnet i figur 6.33 og inngår også i (6.43).

Vi legger merke til det viktige forholdet at dersom systemet har en utpreget resonans som gir en høy topp i frekvensresponsens amplitude, og pådraget har en komponent som ligger i nærheten av resonansfrekvensen, vil responsen kunne ha en stor amplitude ved denne frekvensen selv om amplituden til pådraget er liten.

Figur 6.33
 $u(\omega_i)$ og $y(\omega_i)$ for
dominerende frek-
venser i pådraget
og $|h(j\omega)|$ [dB]



Tilbakekoplede systemer

7.1 Innledning og motivering 237

7.2 Tilbakekopling i monovariabel systemer 240

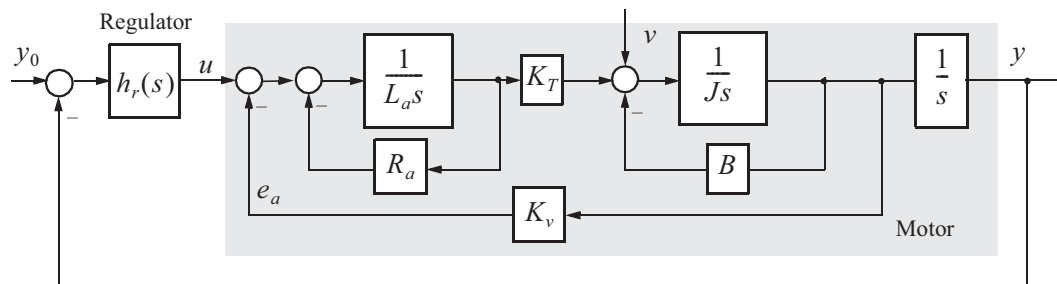
7.3 Følsomhet 260

7.1 INNLEDNING OG MOTIVERING

Tilbakekopling er av sentral betydning i reguleringsteknikken, men den forekommer også i naturlige systemer. Vi har allerede møtt fenomenet i beskrivelsen av matematiske modeller for dynamiske systemer, og vi skal benytte likestrømmotoren fra tidligere for å illustrere dette.

EKSEMPEL 7.1: Likestrømmotor og tilbakekopling

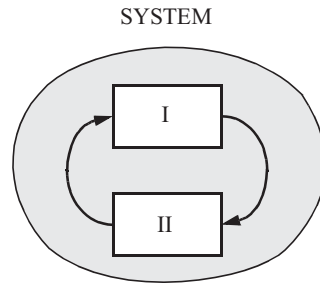
Figur 7.1
Blokkdiagram for
likestrømmotor
med regulerings-
sløyfe



Figur 7.1 viser likestrømmotoren med posisjonsregulering og en vilkårlig regulator $h_r(s)$. Pådraget er ankerspenning u og målingen er vinkelposisjon y . Dette reguleringsystemet har fire tilbakekoplinger. Tre av dem er interne i selve prosessen (motoren), og kan kalles “naturlige” tilbakekoplinger. For eksempel medfører vinkelhastigheten via en tilbakekopling K_v en motindusert spenning e_a som kommer til fradrag fra ankerspenningen.

Den ene ytterste tilbakekoplingen er derimot “kunstig”, i den forstand at den er etablert utenfor motoren for å realisere et egnet reguleringsystem.

Figur 7.2
Gjensidig påvirkning i et system



Felles for begge typer tilbakekoplinger er at de innebærer *gjensidige påvirkninger mellom enheter i et system*. Dette er illustrert i figur 7.2, hvor to vilkårlige blokker (enheter) i et vilkårlig system er framhevet, sammen med piler som danner en lukket sløyfe. Et sluttet kretsløp på grunn av slik gjensidig påvirkning kalles en **tilbakekoplingsløyfe**. En “kunstig” tilbakekopplingsløyfe kalles også en **reguleringsløyfe**.

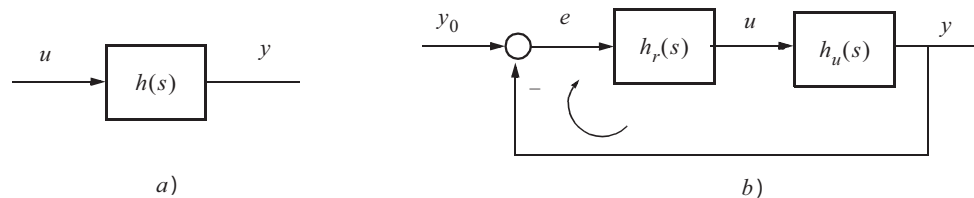


Et system som ikke inneholder reguleringsløyfer, kalles et **åpent system**. Figur 7.3 a) viser transferfunksjonen for en fysisk prosess, som er eksempel på et åpent system.

Et system som inneholder en eller flere reguleringsløyfer, kalles et **lukket system**, se figur 7.3 b). Vi snakker også om en **lukket sløyfe**, eller reguleringsløyfe. Reguleringsløyfen er realisert ved at utgangssignalet y subtraheres fra den ønskede utgangsverdien (referansen) y_0 . Dette gir et avvikssignal e . Avvikssignalet benyttes av en regulator til å beregne pådraget u .

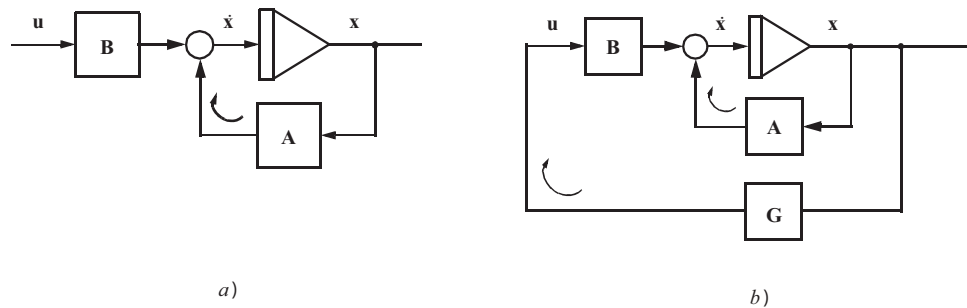
Vi definerer nå **sløyfetransferfunksjonen** $h_0(s)$ for et tilbakekoplet system. Den består av produktet av alle transferfunksjonene langs den lukkede sløyfen. Hvis vi bryter tilbakekoplingen, har vi et åpent system. Da bruker vi begrepet “det åpne system” om sløyfetransferfunksjonen $\frac{y}{e}(s) = h_0(s)$, dvs. vi tar med regulatoren. Da er altså $h_0(s) = h_u(s)h_r(s)$, jfr. figur 7.3 b).

Figur 7.3
Åpen og lukket sløyfe



For å komplettere framstillingen skal vi også beskrive tilbakekopling med utgangspunkt i tilstandsrommodellen. Den matematiske modellen $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$, se også figur 7.4 a), benyttes i svært mange sammenhenger. Dette viser at tilbakekopling forekommer i mange prosesser, i det modellen gir en “naturlig” tilbakekopling fra vektoren \mathbf{x} til vektoren $\dot{\mathbf{x}}$. Den fysiske betydningen av dette er at modellens tilstandsvariable, beskrevet ved \mathbf{x} , er med på å bestemme med hvilken hastighet \mathbf{x} varierer.

Figur 7.4
Elementært
blokkdiagram



Dersom pådraget genereres som en lineær transformasjon av \mathbf{x} ($\mathbf{u} = \mathbf{G}\mathbf{x}$), der \mathbf{G} er en konstant matrise, vil vi som vist i figur 7.4 b) få en ny og “kunstig” tilbakekopling i tillegg til den “naturlige” tilbakekoplingen i systemet. Antar vi at $\mathbf{u} = \mathbf{0}$ i figur 7.4 a), finner vi at dette systemet er beskrevet av differensiallikningen

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \quad (7.1)$$

mens systemet i figur 7.4 b) er beskrevet av

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{G}\mathbf{x} = (\mathbf{A} + \mathbf{B}\mathbf{G})\mathbf{x} = \tilde{\mathbf{A}}\mathbf{x} \quad (7.2)$$

Sammenlikner vi (7.1) og (7.2), ser vi at den “kunstige” tilbakekoplingen \mathbf{G} fra tilstandsvektoren \mathbf{x} til pådragsvektoren \mathbf{u} har modifisert systemets karakteristiske matrise fra å være \mathbf{A} for det åpne systemet til å være $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{B}\mathbf{G}$ for det lukkede systemet. Regulatoren endrer derfor systemets dynamiske egenskaper. Hensikten med regulering er å endre et systems egenskaper i en ønsket retning. *Fra nå av vil, hvis ikke annet er sagt, tilbakekopling være ensbetydende med “kunstig” tilbakekopling.*

I dette kapitlet skal vi se på en del vesentlige egenskaper ved tilbakekopling i monovariabel systemer. Dette er systemer med ett (skalart) pådrag og en (skalart) prosessutgang. Vi skal utvikle en enkel grafisk teknikk som gjør det mulig, på basis av frekvensanalyse, å bestemme egenskapene til tilbakekoblede monovariabel systemer.

7.2 TILBAKEKOPLING I MONOVARIABLE SYSTEMER

Vi tar utgangspunkt i en monovariabel dynamisk prosess beskrevet med en matematisk modell¹

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u + \mathbf{e}v \\ y &= \mathbf{c}^T \mathbf{x}\end{aligned}\quad (7.3)$$

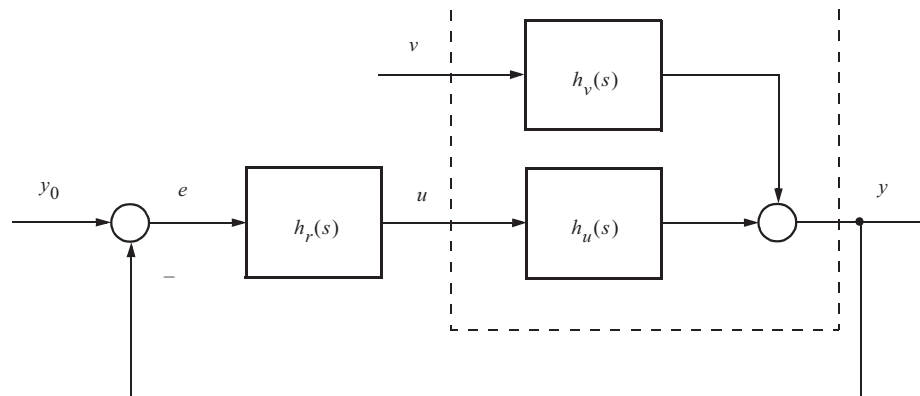
Vi legger merke til at pådraget, forstyrrelsen og målingen er skalare størrelser, mens tilstanden er en vektor. Ved Laplacetransformasjon finner vi

$$\begin{aligned}y &= \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b}u(s) + \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{e}v(s) \\ &= h_u(s)u(s) + h_v(s)v(s)\end{aligned}\quad (7.4)$$

Likning (7.4) viser at vi har en transferfunksjon $h_u(s)$ fra pådraget $u(s)$ til målingen $y(s)$, mens vi har en transferfunksjon $h_v(s)$ fra forstyrrelsen $v(s)$ til målingen.

Med utgangspunkt i systemet beskrevet i (7.4) kan vi nå konstruere et tilbakekoplet system med en struktur som vi etter hvert skal vise har mange gode egenskaper. Dette systemet er vist i figur 7.5.

Figur 7.5
Et monovariabelt tilbakekoplet system med forstyrrelse



¹Merk her at \mathbf{e} er en $n \times 1$ -vektor, et spesialtilfelle av den matrisen \mathbf{E} som vi bruker når v ikke er skalar, men en vektor \mathbf{v} . Vektoren \mathbf{e} må ikke forveksles med symbolet for reguleringsavvik.

Med dette systemet kan vi oppnå en rekke forskjellige egenskaper:

- Ved hensiktsmessig valg av **regulatortransferfunksjonen** $h_r(s)$ kan prosessens utgang (målingen y) følge en vilkårlig varierende referanse y_0 med et avvik e som ikke overskrider spesifiserte grenser i forskjellige situasjoner. En ønsker da også at y skal være mest mulig uavhengig av forstyrrelsene v . Denne problemstillingen forekommer i såkalte **følgesystemer** eller **servomekanismer**.
- Ved hensiktsmessig valg av $h_r(s)$ kan en når y_0 antas å være konstant, oppnå at y blir særlig lite influert av forstyrrelsen v . Dette er den vanlige problemstilling i **prosesserregulering**.
- Ved hensiktsmessig valg av $h_r(s)$ kan en prosess som i seg selv er ustabil, gjøres stabil. En ustabil prosess har egenverdier (poler) i høyre halvplan. Denne situasjonen forekommer både i følgesystemer og ved prosesserregulering.

Vi setter inn følgende i (7.4)

$$u(s) = h_r(s)(y_0(s) - y(s)) \quad (7.5)$$

og løser med hensyn på $y(s)$.

$$y(s) = \frac{h_u(s)h_r(s)}{1 + h_u(s)h_r(s)}y_0(s) + \frac{1}{1 + h_u(s)h_r(s)}h_v(s)v(s) \quad (7.6)$$

I (7.6) ser vi at produktet $h_0(s) = h_u(s)h_r(s)$ forekommer flere steder. Denne transferfunksjonen, som utgjør produktet av alle transferfunksjonene rundt tilbakekopplingsløyfen, kalles som allerede nevnt for **sløyfetransferfunksjon**. (I litteraturen finner en ofte at denne betegnes med $A(s)$, men dette kan medføre forveksling med \mathbf{A} som vi bruker som matrise-symbol i denne boka).

7.2.1 Følgeforhold. Avviksforhold. Reguleringsgrad

Vi ser av (7.6) at responsen til systemet består av summen av to bidrag, ett som skyldes referansen $y_0(s)$ og ett som skyldes forstyrrelsen $v(s)$. Det er hensiktsmessig å betrakte disse to bidragene hver for seg, noe vi kan tillate oss i lineære systemer (og så eventuelt superponere virkningen av dem hvis det er ønskelig). Dette innebærer å sette $v(s) = 0$ når vi studerer responsen fra $y_0(s)$, og $y_0(s) = 0$ når vi studerer responsen fra $v(s)$.

Betrakter vi bare første del av responsen i (7.6) skriver vi

$$y(s) = \frac{h_0(s)}{1 + h_0(s)}y_0(s) = N(s)h_0(s)y_0(s) = M(s)y_0(s) \quad (7.7)$$

Vi har her innført to nye funksjoner,

$$M(s) = \frac{h_0(s)}{1 + h_0(s)} = \frac{y}{y_0}(s) = \text{følgeforholdet} \quad (7.8)$$

$$N(s) = \frac{1}{1 + h_0(s)} = \frac{e}{y_0}(s) = \text{avviksforholdet} \quad (7.9)$$

Disse to størrelsene¹ spiller en sentral rolle i beskrivelsen av monovariabelt reguleringsystemer. Følgeforholdet $M(s)$ uttrykker hvor godt vi lykkes i å få utgangen y til å følge referansen y_0 . Når $M(s) = 1$, har vi perfekt følging. Likning (7.8) viser at dette tilnærmet oppnås når $|h_0(s)| \gg 1$.

Detaljene i avviket mellom y og y_0 ser vi best av avviksforholdet $N(s)$. Likning (7.9) viser at vi kan oppnå $N(s) \approx 0$ og dermed ha perfekt følging, ved å sørge for at $|h_0(s)| \gg 1$.

Det siste leddet i (7.6) uttrykker responsen av det tilbakekoplede systemet på prosessens forstyrrelser. Vi kan skrive

$$y(s) = \frac{1}{1 + h_0(s)} h_v(s) v(s) = N(s) h_v(s) v(s) \quad (7.10)$$

Vi antar i dette tilfellet at $y_0(s) = 0$.

Dersom vi ikke hadde hatt tilbakekopling rundt prosessen, ville responsen i y på grunn av v blitt

$$y(s)_{\text{uten}} = h_v(s) v(s) \quad (7.11)$$

(Indekset “uten” betyr: Uten tilbakekopling.)

(7.11) kan sees direkte fra figur 7.5. Benytter vi (7.10), finner vi

$$y(s)_{\text{med}} = N(s) y(s)_{\text{uten}} \quad , \quad \frac{y(s)_{\text{med}}}{y(s)_{\text{uten}}} = N(s) \quad (7.12)$$

(Indekset “med” betyr: Med tilbakekopling.)

Størrelsen $N(s) = 1/(1 + h_0(s))$ uttrykker direkte hvilken forbedring vi oppnår med hensyn på undertrykkelse av forstyrrelsene ved å bruke tilbakekopling.

¹Vi bruker stor M og N her sjøl om disse størrelsene er skalare, fordi dette er innarbeidet i internasjonal litteratur. Men ellers i denne boka er store bokstaver reservert for matriser.

Av denne grunn er det innført betegnelsen

$$N(s) = \text{reguleringsgrad}$$

Dette kan uttrykkes slik: Transferfunksjonen mellom en vilkårlig påvirkning og en vilkårlig respons i et tilbakekoplet system er lik transferfunksjonen *uten* tilbakekopling (mellom påvirkningen og responsen) multiplisert med størrelsen $N(s)$

Det kan være nyttig i fortsettelsen å merke seg at $N(s) + M(s) = 1$.

I avsnitt 7.3 vil vi bruke $N(s)$ i enda en betydning, som et mål for det som kalles *relativ følsomhet*.

EKSEMPEL 7.2: Følgeforholdet $M(s)$ til et system

Vi skal nå ta utgangspunkt i den regulering av likestrømsmotoren som ble diskutert i eksempel 4.12. Systemet er som vist i figur 7.5 med

$$h_u(s) = \frac{1}{s(1 + Ts)}$$

$$h_r(s) = K$$

Følgeforholdet i henhold til (7.8) blir

$$M(s) = \frac{1}{1 + \frac{1}{K}s + \frac{T}{K}s^2} = \frac{1}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2}$$

der

$$\omega_0 = \sqrt{\frac{K}{T}}$$

$$\zeta = \frac{1}{2\sqrt{TK}}$$

Vi ser at responsen y vil følge med referansen så sant de framtrepende frekvenskomponentene i y_0 er lavere enn ω_0 . Dersom vi ønsker at den relative dempningen skal være $\zeta = 0.5$, må vi velge $K = 1/T$ som gir $\omega_0 = 1/T$.

Avviksforholdet og reguleringsgraden er gitt av

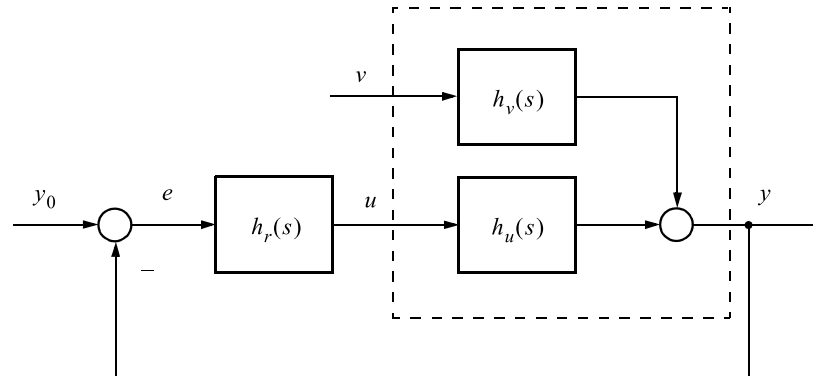
$$N(s) = \frac{1}{1 + h_u(s)h_r(s)} = \frac{1}{K} \frac{s(1 + Ts)}{1 + \frac{1}{K}s + \frac{T}{K}s^2}$$

Vi legger merke til at $N(s)$ har et nullpunkt i origo. Dette er en viktig egenskap som medfører at reguleringsavviket under stasjonære forhold (dvs. konstant referanse) vil bli redusert til null. Dette kommer vi grundig tilbake til i kapittel 9.



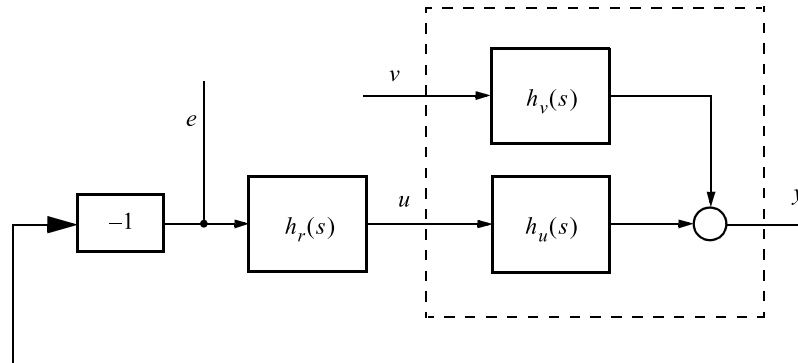
7.2.2 Å finne transferfunksjon fra v til e

Figur 7.6
Standard struktur
for monovariabelt
system



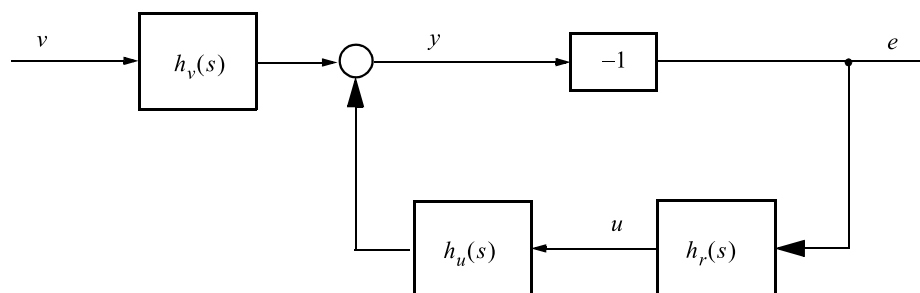
Setter $y_0 = 0$ (lov p.g.a. superposisjonsprinsippet), og betrakter e som utgang:

Figur 7.7



Ved å forlenge noen piler og korte litt ned på andre, ved å rette ut noen piler og å bøye andre, kan dette tegnes fullstendig ekvivalent som

Figur 7.8



Vi får

$$\begin{aligned} \frac{e}{v}(s) &= h_v(s) \frac{(-1)}{1 - (-1)h_u(s)h_r(s)} = -h_v(s) \frac{1}{1 + h_u(s)h_r(s)} = \\ &= -h_v(s) \frac{1}{1 + h_0(s)} = -h_v(s)N(s) \end{aligned}$$

7.2.3 Frekvensanalyse av monovariabele tilbakekoblede systemer

De begrepene som ble introdusert i avsnitt 7.2.1, egner seg godt til frekvensanalytisk beskrivelse av tilbakekoblede systemers egenskaper. Uttrykkene som er utledet der, gir frekvensresponsen til det tilbakekoblede systemet. Den fysiske tolkningen av frekvensresponsen er gitt i kapittel 6. På basis av frekvensresponsen kan vi avlede hvordan det tilbakekoblede systemet vil oppføre seg når det utsettes for periodiske påvirkninger, tilfeldige påvirkninger som kan tilnærmes med en sum av periodiske påvirkninger (avsnitt 6.5.1), eller transiente påvirkninger. Vi kan danne et tilnærmet uttrykk for systemets transferfunksjon ut fra frekvensresponsen, og finner deretter tidsresponsen.

Størrelsen $N(s)$, som ble kalt **reguleringsgrad** og **avviksforhold** i forrige avsnitt, har sin ekvivalent, $N(j\omega) = |N(j\omega)|e^{j\angle N(j\omega)}$, i frekvensplanet. Reguleringsgraden spiller en sentral rolle i frekvensanalyse av tilbakekoblede systemer. Den kan bestemmes grafisk med enkle virkemidler når frekvensresponsen målt rundt tilbakekoplingsløyfen $h_0(j\omega)$ er kjent.

Vi konstaterer først at

$$\begin{aligned} N(j\omega) &\approx \frac{1}{h_0(j\omega)} && \text{når } |h_0(j\omega)| \gg 1 \\ N(j\omega) &\approx 1 && \text{når } |h_0(j\omega)| \ll 1 \end{aligned} \quad (7.13)$$

Dette medfører at

$$\begin{aligned} |N(j\omega)|[\text{dB}] &\approx -|h_0(j\omega)|[\text{dB}] && \text{når } |h_0(j\omega)|[\text{dB}] \gg 0 \\ |N(j\omega)|[\text{dB}] &\approx 0 && \text{når } |h_0(j\omega)|[\text{dB}] \ll 0 \end{aligned} \quad (7.14)$$

Videre finner vi

$$\begin{aligned}\angle N(j\omega) &\approx -\angle h_0(j\omega) && \text{når } |h_0(j\omega)|[\text{dB}] \gg 0 \\ \angle N(j\omega) &\approx 0 && \text{når } |h_0(j\omega)|[\text{dB}] \ll 0\end{aligned}\quad (7.15)$$

(7.14) og (7.15) egner seg svært godt til skissering av en asymptotisk approksimasjon til $N(j\omega)$ når asymptotiske amplitude/fase/frekvens-karakteristikker er gitt for $h_0(j\omega)$.

Tilsvarende finner vi for **følgeforholdet** $M(j\omega)$. Vi har

$$M(j\omega) = \frac{h_0(j\omega)}{1 + h_0(j\omega)} \quad (7.16)$$

som gir

$$\begin{aligned}|M(j\omega)| &\approx 1 && \text{når } |h_0(j\omega)|[\text{dB}] \gg 0 \\ |M(j\omega)| &\approx |h_0(j\omega)| && \text{når } |h_0(j\omega)|[\text{dB}] \ll 0\end{aligned}\quad (7.17)$$

Videre har vi at

$$\begin{aligned}\angle M(j\omega) &\approx 0 && \text{når } |h_0(j\omega)|[\text{dB}] \gg 0 \\ \angle M(j\omega) &\approx \angle h_0(j\omega) && \text{når } |h_0(j\omega)|[\text{dB}] \ll 0\end{aligned}\quad (7.18)$$

EKSEMPEL 7.3: Asymptotisk AFF-diagram for $N(j\omega)$

La oss nå ta utgangspunkt i modellen av likestrømsmotor med regulering i eksempel 4.15. Sløyfetransferfunksjonen er

$$h_0(s) = \frac{K(1 + T_3s)}{s(1 + T_1s)(1 + T_2s)} \quad (7.19)$$

$$\text{der } K = 40, T_1 = 0.5, T_2 = 1/30, T_3 = 1/6$$

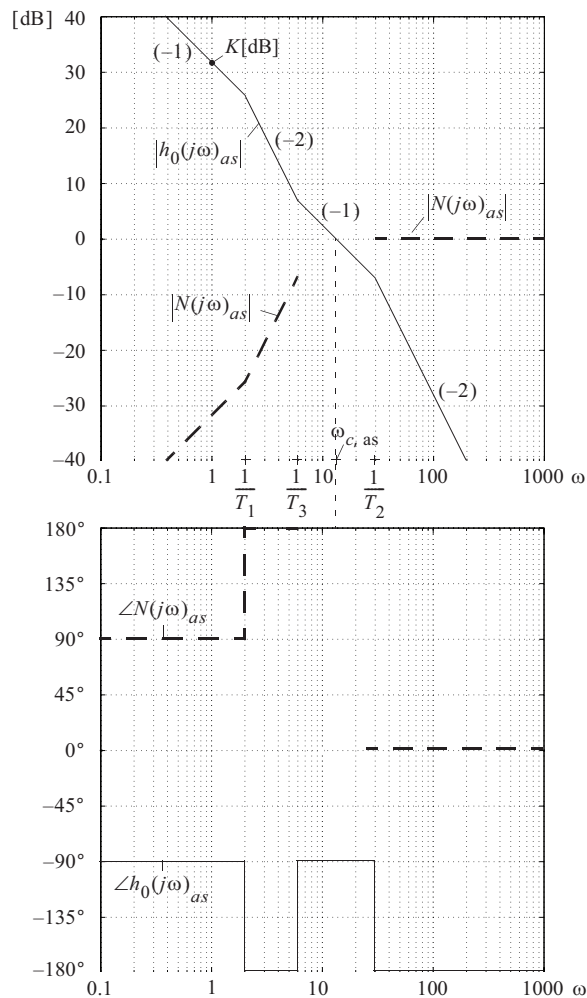
En asymptotisk amplitudekarakteristikk for $|h_0(j\omega)|$ er tegnet øverst i figuren, og den asymptotiske amplitudekarakteristikken for $|N(j\omega)|$ er avledet på basis av (7.14). Vi ser at frekvensen der asymptoten $|h_0(j\omega)|_{as} = 0$ [dB], er betegnet med $\omega_{c,as}$. Denne frekvensen kalles den **asymptotiske kryssfrekvensen**. Vi ser også at 0-dB-linjen karakteriserer overgangen mellom de to tilfellene, som er gitt av (7.14). Ved lave frekvenser, der $|h_0(j\omega)|$ har en stor verdi, vil $|N(j\omega)|$ være gitt av speilbildet til $|h_0(j\omega)|$ omkring absisseaksen. Ved høye frekvenser der $|h_0(j\omega)|$ er liten, vil $|N(j\omega)| = 0$ [dB].

Dette betyr at det tilbakekoplede systemet er effektivt som reguleringssystem for frekvenser som er lavere enn $\omega_{c,as}$. Systemet vil undertrykke eksterne forstyrrelser og medføre lite følgeavvik når $|N(j\omega)| < 0$ [dB].

Vi ser hvordan $\angle N(j\omega)_{as}$ avledes av $\angle h_0(j\omega)_{as}$ ved bruk av (7.15) nederst i figur 7.9.

(Nær $\omega_{c,as}$ vil asymptotene for $|N(j\omega)|$ og $\angle N(j\omega)_{as}$ avvike noe fra de tilsvarende speilbilder. Derfor er forløpene i dette området utelatt i figuren.)

Figur 7.9
Skissering av asymptotisk Bode-diagram for $N(j\omega)$



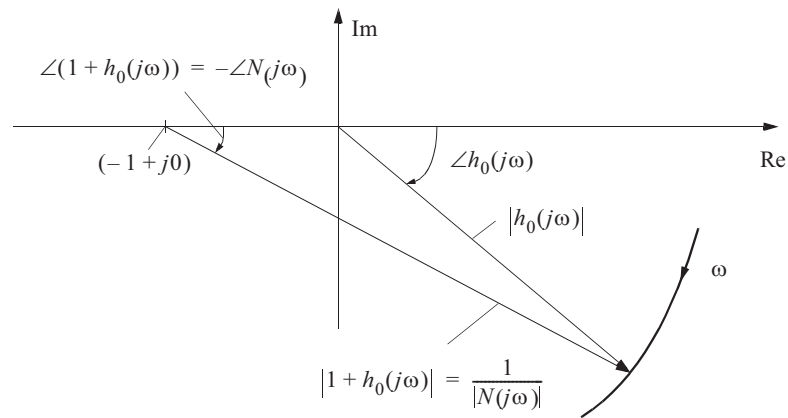
■

Likning (7.14) og (7.15) er bare egnet til å gi et noenlunde nøyaktig uttrykk for det tilbakekoblede systemets egenskaper når $|h_0(j\omega)|$ [dB] $\ll 0$ ($|h_0(j\omega)| \ll 1$) og når $|h_0(j\omega)|$ [dB] $\gg 0$ ($|h_0(j\omega)| \gg 1$). Vi må derfor finne en metode for å bestemme $N(j\omega)$ når $|h_0(j\omega)|$ [dB] ≈ 0 , dvs. i nærheten av $\omega_{c,as}$, den asymptotiske kryssfrekvensen.

Vi benytter oss da i første rekke av den polare representasjonen av $h_0(j\omega)$. Figur 7.10 viser en del av en polar stedkurve (Nyquistkurve) for frekvensresponsen til h_0 i et vilkårlig system. Vektoren $h_0(j\omega)$ er angitt fra origo til et punkt på stedkurven.

Figur 7.10

Den polare stedkurven for frekvensresponsen til et vilkårlig system

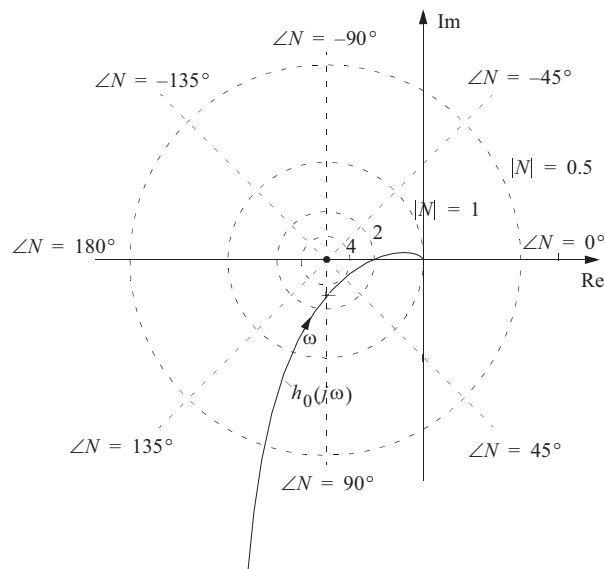


Trekkes en vektor fra punktet $(-1 + j0)$ til det samme punktet på stedkurven, ser vi at denne vil representere uttrykket $1 + h_0(j\omega) = 1/N(j\omega)$.

Vi kan derfor kartlegge det komplekse planet med kurver som angir de stedene endepunktet til vektoren $h_0(j\omega)$ må ligge, for at bestemte konstante verdier $|N(j\omega)|$ og $\angle N(j\omega)$ skal oppnås. Slike “isokurver” eller koter er vist i figur 7.11.

Figur 7.11

Isokurver



Kotene for konstante verdier av $|N(j\omega)|$ er konsentriske sirkler med sentrum i punktet $(-1 + j0)$ og med radius lik $1/|N(j\omega)|$. Kotene for konstant $\angle N(j\omega)$ er rettlinjete stråler ut fra punktet $(-1 + j0)$ som vist. Ved hjelp av dette “ruteneett” bestående av sirkler og stråler, kan vi altså lese av $N(j\omega)$ når vi har tegnet inn grafen for $h_0(j\omega)$. (I figur 7.11 har vi plottet inn $h_0(j\omega)$ for $h_0(s) = \frac{1}{s(s+1)^2}$).

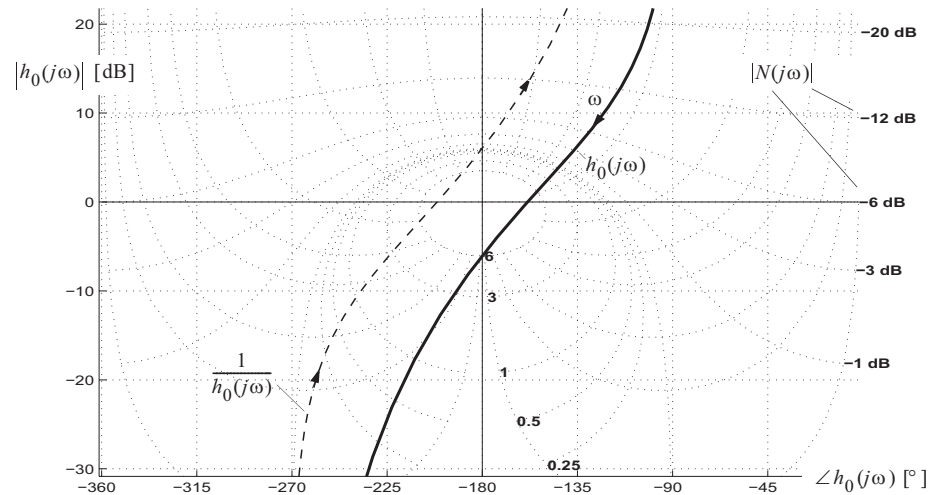
Har vi et slikt N-kote diagram, kan vi tegne inn den polare stedkurven for vektoren $h_0(j\omega)$ som vist heltrukket i figur 7.11 og så lese av $N(j\omega)$ for forskjellige verdier av frekvensen ω (som er en parameter langs stedkurven).

7.2.4 Nicholsdiagram

Vi kan få en mer anvendelig form av N-kote-diagrammet i figur 7.11 ved å bruke den representasjonsformen som i figur 6.7 er vist i delfigur 2.a. Vi framstiller da $|h_0(j\omega)|$ [dB] som funksjon av $\angle h_0(j\omega)$ med ω som parameter. Dette ser vi gir en kartesisk avbildning av det polare diagrammet i figur 7.11. Overfører vi grafen for $h_0(j\omega)$, og “rute-nettet” av sirkler og stråler fra figur 7.11 til den nye avbildningen, får vi et resultat som vist i figur 7.12.

Punktet $(-1 + j0)$ i figur 7.11 er i figur 7.12 avbildet i punktet $\angle h_0(j\omega) = -180^\circ$ og $|h_0(j\omega)| = 0$ [dB]. Kotene for $|N(j\omega)| = \text{konstant}$ [dB] er nå blitt sterkt deformerte “sirkler” og kotene for $\angle |N(j\omega)| = \text{konstant}$ ($^\circ$) er blitt krumme “stråler”. Dette spesielle kotediagrammet kalles et **Nicholsdiagram**.

Figur 7.12
Nicholsdiagram



Ved grafisk skissering av Nicholsdiagram for $h_0(j\omega)$ kan vi gå fram på følgende måte: Tegn AFF-diagram (Bodediagram). Overfør kurven fra AFF-diagrammet til Nicholsdiagrammet. Metoden er forklart i appendiks D.

Vi kan også gjøre det ved å regne ut $\angle h_0(j\omega)$ og $|h_0(j\omega)|$ [dB] for forskjellige ω og avmerke disse verdiene som punkter i Nicholsdiagrammet, der $\angle h_0(j\omega)$ er verdien langs absisseaksen og $|h_0(j\omega)|$ [dB] er verdien langs ordinataksen. Siden ω er en parameter i Nicholsdiagrammet og vi ikke har noen akse for ω , kan vi angi verdier for ω langs stedkurven h_0 . MATLAB gjør alt dette via funksjonene `nichols()` og `ngrid()`.

Det opprinnelige Nicholsdiagrammet som ble utviklet under den andre verdenskrig, framstilte $M(j\omega)$, se figur 7.13. I amerikansk litteratur og i MATLAB er det denne versjonen som brukes. Sammenhengen mellom N og M i Nicholsdiagram kan vi utlede på følgende måte:

$$N(j\omega) = \frac{1}{1+h_0(j\omega)} = \frac{1}{h_0(j\omega)} \frac{1}{1+\frac{1}{h_0(j\omega)}} \quad (7.20)$$

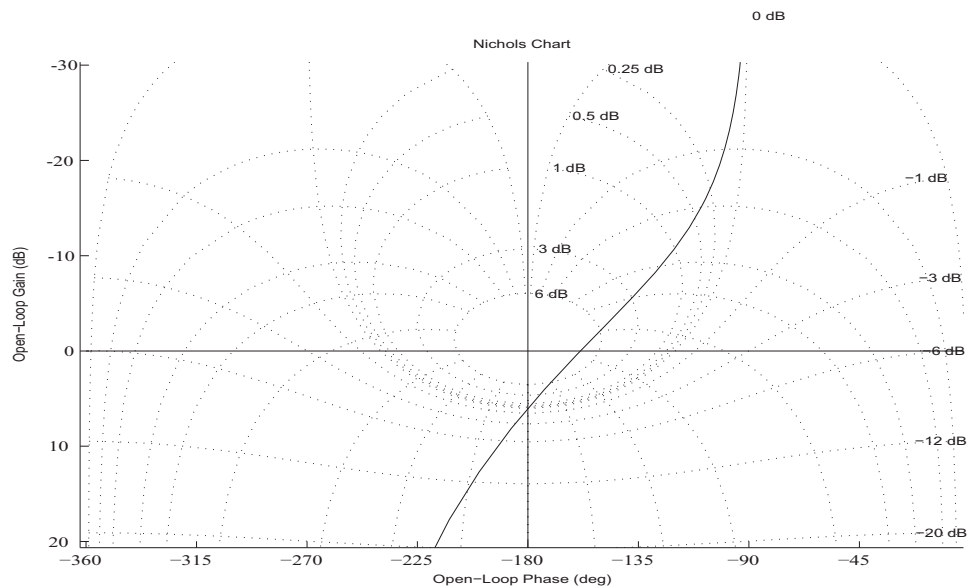
Hvis vi innfører $\frac{1}{h_0(j\omega)} = \tilde{h}_0(j\omega)$, har vi at $|\tilde{h}_0| \text{ [dB]} = -|h_0| \text{ [dB]}$ og $\angle\tilde{h}_0 = -\angle h_0$, dvs. en "speiling" av grafen rundt punktet ($\angle h_0 = -180^\circ$, $|h_0| = 0 \text{ [dB]}$).

Men uttrykket $M(j\omega)$ kan skrives

$$M(j\omega) = \frac{h_0(j\omega)}{1+h_0(j\omega)} = \frac{1}{1+\frac{1}{h_0(j\omega)}} = \frac{1}{1+\tilde{h}_0(j\omega)} \quad (7.21)$$

$M(j\omega)$ kan derfor avleses i figur 7.12 ved først å speile grafen for h_0 rundt punktet (-180° , 0 [dB]), se den stiplede kurven til venstre i figur 7.12.

Figur 7.13
Koter for $M(j\omega)$



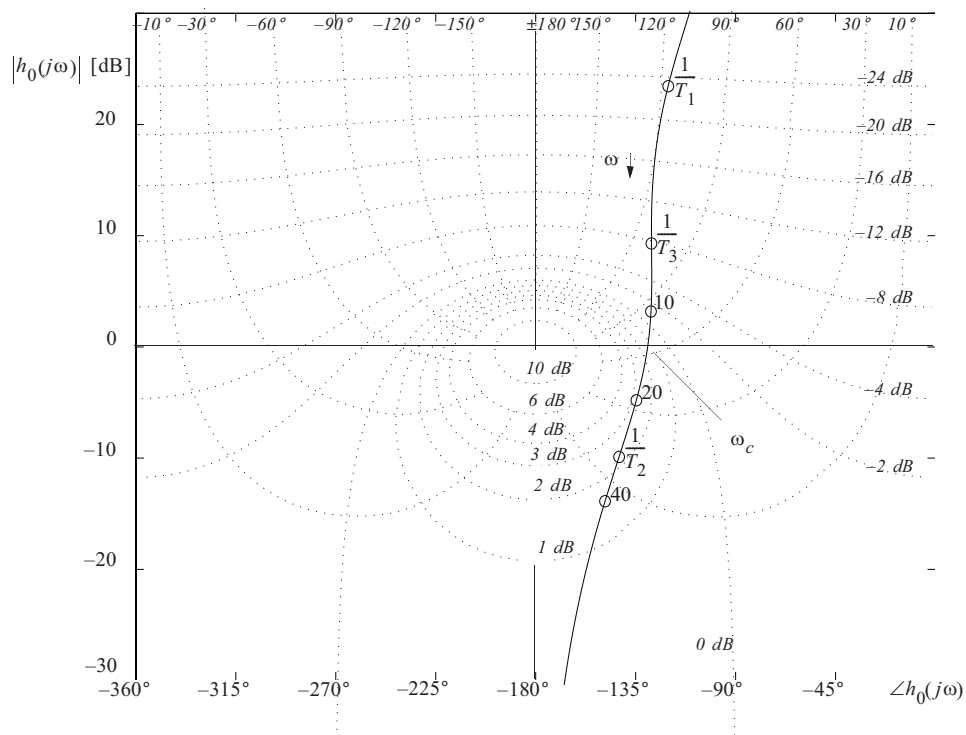
Oppsummering: Inntegnes $h_0(j\omega)$ i Nicholsdiagrammet i figur 7.12, kan vi ta ut $|N(j\omega)|$ og $\angle N(j\omega)$ som funksjoner av ω . Ønsker vi å lese av M i stedet kan dette gjøres i N -diagrammet i figur 7.12. Vi dreier først grafen for $h_0(j\omega)$ 180 grader om midtpunktet ($-180^\circ, 0[\text{dB}]$), og leser så av skjæringspunkter med koter i N -diagrammet. Dette vil da være verdier for $|M(j\omega)|$ og $\angle M(j\omega)$. Alternativt kan vi bruke et M -diagram direkte som vist i figur 7.13.

Det mest vanlige i dag er imidlertid å la MATLAB, på basis av h_0 , plote N og M direkte i Bodediagram, uten å gå veien om Nicholsdiagram. Vi har likevel lagt vekt på Nicholsdiagrammet fordi det gir innsikt i sammenhengen mellom $h_0(j\omega)$ og N, M for forskjellige valg av regulatorer, innsikt som ikke framkommer like lett i Bode-diagram. Dette skal vi studere i et eksempel:

EKSEMPEL 7.4: Bruk av Nicholsdiagram

Vi tar igjen utgangspunkt i regulering av likestrømsmotoren som diskutert i eksempel 4.15 og 7.3. Bruk av asymptotisk amplitude/fase/frekvens-diagram er vist i figur 7.9. Overført til et Nicholsdiagram¹ vil $h_0(j\omega)$ se ut som vist i figur 7.14.

Figur 7.14
Nicholsdiagram for $h_0(j\omega)$ fra eksempel 4.15 og 7.4

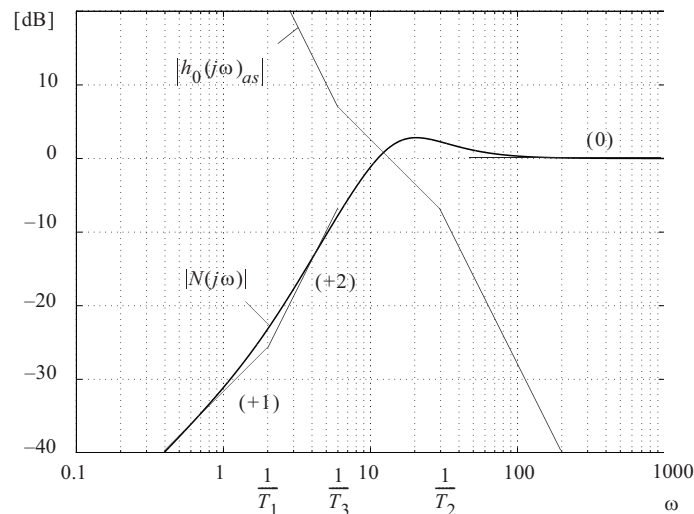


¹Her har vi også merket av $\angle N(j\omega)$ (kursiv øverst i figuren). Men $\angle N(j\omega)$ brukes sjelden.

Fasevinkelen for $h_0(j\omega)$ i nærheten av ω_c er ca. -130° . Dette gir en forholdsvis lav maksimalverdi av $|N(j\omega)|$. Siden $h_0(j\omega)$ avhenger av regulatoren vi velger, kan vi i stor grad bestemme formen på $h_0(j\omega)$. I Nichols-diagrammet i figur 7.14 er $h_0(j\omega)$ formet slik at grafen “smyger seg utenom” den uendelig høye “fjelltoppen” som $|N|$ har i punktet $(-180^\circ, 0[\text{dB}])$. Hvis vi tenker på Nicholsdiagrammet som et vanlig geografisk kart, svarer kotene for konstant $|N|$ til angivelse av høyde i terrenget. Vi starter ved en lav frekvens (øverst i diagrammet) og går “oppoverbakke”, så litt opp i “fjellsida” med “fjellet” på vår høyre side, og deretter noe nedover igjen og ut på ei flat “høyslette” med nivå 0 dB. Ved å forme $h_0(j\omega)$ slik at den går mest mulig utenom “fjellet”, får man en lav resonanstopp i $|N(j\omega)|$. Dette er et direkte mål for det regulerte systemets stabilitet, noe vi kommer til i kapittel 8.

Bodediagrammet for $|N(j\omega)|$ vil svare til et vertikalt snitt langs vår rute i “terrenget” (jfr. høydeprofilen for et langrennsløp). Dette Bodediagrammet med asymptoter er vist i figur 7.15. MATLAB-kode for $N(j\omega)$ og $M(j\omega)$ i Bode-diagram er vist i figur 7.16. $|M(j\omega)|$ og $\angle M(j\omega)$ blir som vist i figur 7.17.

Figur 7.15
 $|N(j\omega)|$ i Bode-
diagram



Figur 7.16
MATLAB-kode for
 $N(j\omega)$ og $M(j\omega)$
i Bode-diagram

```

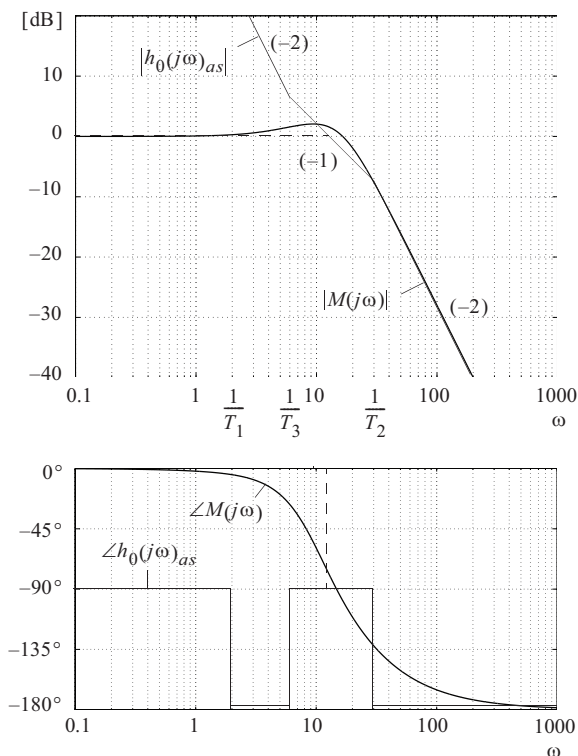
1 K = 40; T1 = 0.5; T2 = 1/30; T3 = 1/6;
2 t0 = K*[T3 1];
3 n0 = conv([T1 1 0], [T2 1]);
4 h0 = tf(t0,n0);
5 N = feedback(1,h0);
6 bode(N);
7 M = feedback(h0,1);
8 bode(M);

```

(reduserer blokkdiagrammet)

(MATLAB plottet dessverre ikke asymptoter)

Figur 7.17
 $|M(j\omega)|$ i Bode-
 diagram

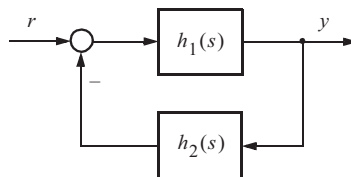


■

7.2.5 System med dynamisk tilbakekopling kontra enhetstilbakekopling

Figur 7.18 viser et tilbakekoplet system med en struktur som forekommer en del. Det spesielle er nå at vi i tillegg til $h_1(s)$ også har en transferfunksjon $h_2(s)$ i *tilbakekoplingsgrenen*, mens vi til nå har forutsatt **enhetstilbakekopling**, dvs. $h_2(s) = 1$. Transferfunksjonene $h_1(s)$ og $h_2(s)$ i henholdsvis forovergrenen og tilbakekoplingsgrenen av systemet, kan hver for seg igjen være resultatet av en rekke seriekoblede blokker. Vi betrakter overføringen mellom inngangsstørrelsen r og utgangsstørrelsen (målingen) y .

Figur 7.18
 System med dyna-
 misk tilbakekopling



Denne blir

$$\begin{aligned}\frac{y}{r}(s) &= \frac{h_1(s)}{1 + h_1(s)h_2(s)} = \frac{h_0(s)}{1 + h_0(s)} \frac{1}{h_2(s)} \\ &= M(s) \frac{1}{h_2(s)}\end{aligned}\quad (7.22)$$

Vi har som før

$$h_0(s) = h_1(s)h_2(s) = \text{sløyfetransferfunksjon}$$

Av (7.22) ser vi at

$$\begin{aligned}\frac{y}{r}(s) &\approx \frac{1}{h_2(s)} && \text{når } |h_0(s)| \gg 1 \\ \frac{y}{r}(s) &\approx h_1(s) && \text{når } |h_0(s)| \ll 1\end{aligned}\quad (7.23)$$

Likning (7.23) uttrykker at når tilbakekoplingsløyfen er effektiv ($|h_0(s)| \gg 1$), vil overføringen mellom r og y være omtrent lik *den inverse av tilbakekoplingen*. Den avhenger altså ikke av $h_1(s)$ i dette tilfellet. Dette resultatet er viktig og benyttes i avsnitt 9.5 under kompensasjon av reguleringsystemer med intern tilbakekopling. Når tilbakekoplingen ikke er effektiv ($|h_0(s)| \ll 1$), vil overføringen være tilnærmet lik transferfunksjonen i *forovergrenen*.

Vi merker oss at nå den svært vanlige enhetstilbakekoplingen, $h_2(s) = 1$, kan betraktes som et spesialtilfelle av dynamisk tilbakekopling. Den gir

$$\frac{y}{r}(s) \approx 1 \quad \text{når } |h_0(s)| \gg 1$$

EKSEMPEL 7.5: Tidskonstant i tilbakekoplingsgrenen

Anta at

$$h_1(s) = K \text{ og } h_2(s) = \frac{1}{1 + Ts}$$

Vi får da

$$\frac{y}{r}(s) = \frac{\frac{K}{1+K}}{1 + \frac{1+K}{1+K}s} (1 + Ts)$$

Det første leddet blir tilnærmet lik 1 når $K \gg 1$ og s har en rimelig verdi. Vi har dermed oppnådd at overføringen er blitt tilnærmet lik det inverse av tilbakekoplingen.

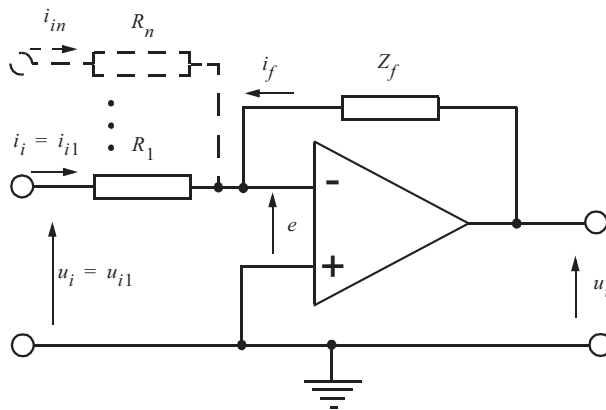
Dette utnyttes i elektronikken der man med forskjellige tilbakekoplinger rundt forsterkere (operasjonsforsterkere) med meget høy forsterkning kan oppnå forskjellige ønskede frekvensforløp.



EKSEMPEL 7.6: Operasjonsforsterker

Denne er sentral i kretsteknikken (se f.eks. Nielsson, Riedel: "Electric Circuits"). Vi skal her ta for oss operasjonsforsterkeren som tilbakekoplet system. Operasjonsforsterkeren er en svakstrøms signalforsterker og brukes i elektroniske kretser. Fysisk foreligger den som en mikrobrikke med bein, for påmontering på kretskort. Figur 7.19 viser et sterkt forenklet diagram med en typisk (for våre formål) oppkopling av operasjonsforsterker.

Figur 7.19
Operasjonsforsterker med tilbakekopling



Forsterkeren har to innganger. Den inverterende (-) brukes; dette er nødvendig for stabilitet av tilbakekoplingen. Den ikke-inverterende (+) forbindes til jord. Operasjonsforsterkeren har meget stor forsterkning $|A_0| \approx \infty$ i åpen sløyfe. Dette betyr $e \approx 0$. Den har også meget stor inngangsimpedans $|Z_i| \approx \infty$. Denne forutsetningen betyr at strømmen inn på den negative inngangen kan settes ≈ 0 .

Da får vi

$$i_i = -i_f \Rightarrow \frac{u_i}{R_1} = -\frac{u_u}{Z_f} \Rightarrow u_u = -\frac{Z_f}{R_1} u_i \quad (7.24)$$

Som antydnet stiplet i figur 7.19 kan vi også ha flere signaler inn på inngangen. Da vil operasjonsforsterkeren fungere som **summasjonsforsterker**. Vi har da

$$i_{i1} + \dots + i_{in} = -i_f \Rightarrow \frac{u_{i1}}{R_1} + \dots + \frac{u_{in}}{R_{1n}} = -\frac{u_u}{Z_f} \Rightarrow u_u = -Z_f \left(\frac{u_{i1}}{R_1} + \dots + \frac{u_{in}}{R_{1n}} \right) \quad (7.25)$$

$-u_u$ blir altså en veiet sum av inngangsspenningene.

Vi har til nå forutsatt $|A_0| \approx \infty$. La oss nå i stedet sette

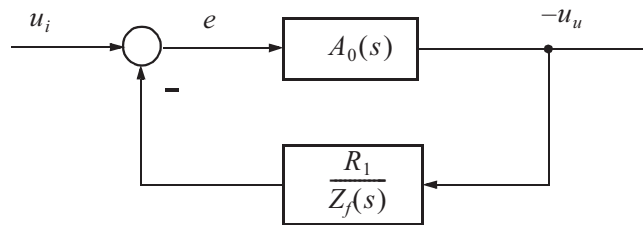
$$A_0(s) = \frac{K}{1 + Ts}, \text{ der } K > 0 \text{ er meget stor, og } T \text{ meget liten} \quad (7.26)$$

Dette er en mer realistisk modell av operasjonsforsterkeren *uten* tilbakekopling. Vi forutsetter nå ikke lenger $e \approx 0$, men kan fortsatt forutsette at strømmen inn på operasjonsforsterkeren er ≈ 0 . Med tilbakekopling som i figur 7.19, har vi

$$e = u_i - R_1 i_i = u_i - R_1 \left(-\frac{u_u}{Z_f} \right) = u_i + \frac{R_1}{Z_f} u_u \quad (7.27)$$

Vi kan nå sette opp blokkdiagrammet i figur 7.20

Figur 7.20
Blokkdiagram for operasjonsforsterker med tilbakekopling



Transferfunksjonen fra u_i til $-u_u$ (merk fortegn!) blir

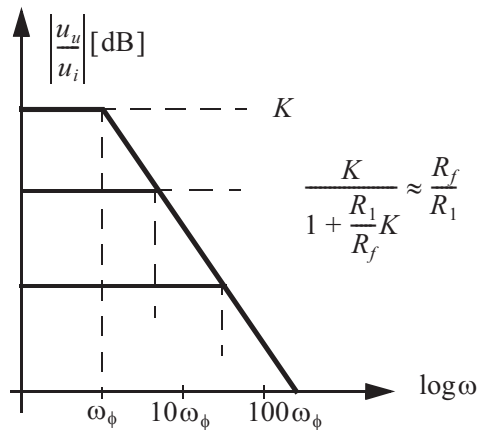
$$\frac{-u_u}{u_i}(s) = \frac{A_0(s)}{1 + \frac{R_1}{Z_f(s)} A_0(s)} = \frac{\frac{R_1}{Z_f(s)} A_0(s)}{1 + \frac{R_1}{Z_f(s)} A_0(s)} \cdot \frac{1}{\frac{R_1}{Z_f(s)}} = M(s) \frac{1}{h_2(s)} \approx \frac{Z_f(s)}{R_1} \quad (7.28)$$

Nest siste ledd svarer til notasjonen fra (7.22) og (7.23). Siste ledd gjelder når $|A_0 R_1 / Z_f| \approx \infty$; jfr. (7.24). Når dette gjelder, kan vi ved passende valg av $Z_f(s)$ og R_1 (der R_1 kan erstattes av en ikke-ohmsk impedans $Z_1(s)$ om ønskelig), oppnå en spesifisert frekvensrespons for operasjonsforsterkeren med tilbakekopling. Vi kan blant annet lage analoge regulatorer og filtre.

Hvis vi antar en rent ohmsk tilbakekopling $Z_f(s) = R_f$, ser vi i figur 7.21 hvordan båndbredden for operasjonsforsterkeren *med* tilbakekopling, øker med synkende R_f , mens forsterkningen avtar.

Den øverste asymptoten (kraftig strek) viser situasjonen uten tilbakekopling. Frekvensen ω_ϕ kalles operasjonsforsterkerens *3 dB-frekvens*. Over $\omega_\phi = 1/T$ faller den asymptotiske frekvensgangen med 20dB/dekade, jfr. (7.26).

Figur 7.21
Operasjonsforsterkerens frekvensrespons ved forskjellige ohmske tilbakekoplinger



■

EKSEMPEL 7.7: Dynamisk tilbakekopling i motorstyring

Virkningen av en dynamisk tilbakekopling for et reguleringssystem kan med fordel studeres ved hjelp av Bodediagrammer. Vi tar utgangspunkt i motorstyringen på figur 7.22 hvor transferfunksjonen fra pådrag til posisjon er (vi forenkler likestrømsmotorens dynamikk sterkt):

$$\frac{y(s)}{u(s)} = h_1(s) = \frac{K}{s}$$

Operasjonsforsterkerens tilbakekopling er rent ohmsk. Den brukes i dette tilfellet bare som summasjonsforsterker. Total forsterkning over operasjons- og kraftforsterker er valgt lik 1.

Anta at vi ønsker å lage et system som har transferfunksjon

$$\frac{y}{r}(s) = K_p \frac{1 + T_i s}{T_i s}$$

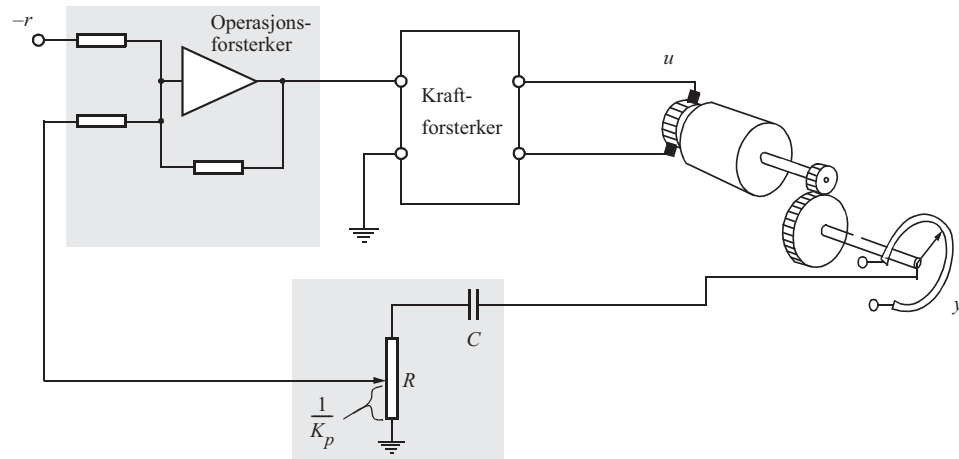
Dette oppnås ved å bruke en tilbakekopling som vist nederst i figur 7.15.

Forutsatt at vi kan få til $|h_0(s)| \gg 1$, kan vi da velge

$$h_2(s) = \frac{1}{K_p} \frac{T_i s}{1 + T_i s}$$

Dette realiseres av RC-leddet nederst i figur 7.22.

Figur 7.22
Elektrisk motor
med elektrisk for-
sterker $RC = T_i$



Da får vi

$$h_0(s) = \frac{KT_i}{K_p} \frac{1}{1 + T_i s}$$

Figur 7.23 viser et Bodediagram for $h_0(j\omega)$, $h_1(j\omega)$, $h_2(j\omega)$ og $\frac{y}{r}(j\omega)$.

Vi legger merke til at

$$\frac{y}{r}(j\omega) = K_p \frac{(1 + j\omega T_i)}{j\omega T_i} M(j\omega)$$

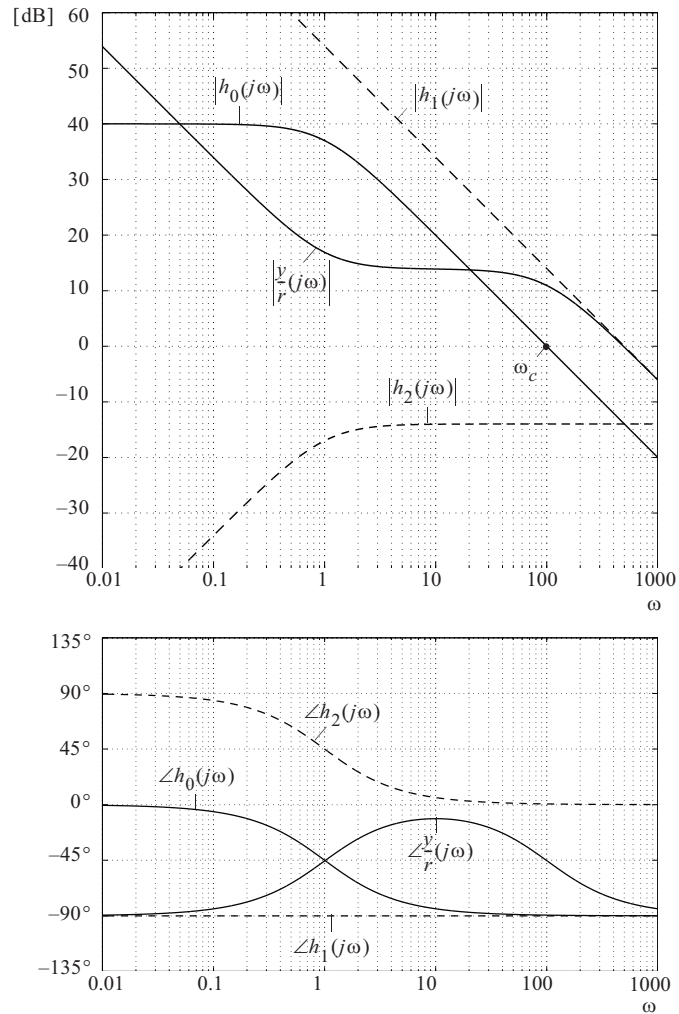
$|M(j\omega)| \approx 1$ opp til $\omega_c = (K/K_p)$.

Målsettingen

$$\frac{y}{r}(s) = K_p \frac{1 + T_i s}{T_i s}$$

er derfor bare tilfredsstilt opp til ω_c .

Figur 7.23
Bodediagram for
 $h_0(j\omega)$, $h_1(j\omega)$,
 $h_2(j\omega)$ og $\frac{y}{r}(j\omega)$



7.3 FØLSOMHET

Foran har vi tolket $N(s)$ som *avviksforhold* og *reguleringsgrad*. En tredje tolkning av $N(s)$ angår systemets **relative følsomhet** overfor parametervariasjoner. Vi ser på størrelsen

$$S_{\alpha}^{h_0} = \frac{\frac{dh_0(s)}{h_0(s)}}{\frac{d\alpha}{\alpha}} = \frac{\alpha}{h_0(s)} \frac{dh_0(s)}{d\alpha} \quad (7.29)$$

der α er en **parameter** (for eksempel en tidskonstant eller en forsterkning) som innvirker på $h_0(s)$. $S_{\alpha}^{h_0}$ kalles *den relative følsomheten* til h_0 med hensyn på α . I tilbakekoblede systemer er vi ofte interessert i å finne den relative følsomheten

$$S_{\alpha}^M = \frac{\frac{dM(s)}{M(s)}}{\frac{d\alpha}{\alpha}} = \frac{\alpha}{M(s)} \frac{dM(s)}{d\alpha} \quad (7.30)$$

Kjerneregelen for derivasjon gir

$$\frac{dM(s)}{d\alpha} = \frac{dM(s)}{dh_0(s)} \frac{dh_0(s)}{d\alpha} = \frac{1}{(1+h_0(s))^2} \frac{dh_0(s)}{d\alpha} \quad (7.31)$$

som innsatt i (7.30) gir

$$S_{\alpha}^M = \frac{\alpha}{h_0(s)} \frac{1}{(1+h_0(s))^2} \frac{dh_0(s)}{d\alpha} = N(s) S_{\alpha}^{h_0} \quad (7.32)$$

Den relative følsomheten til følgeforsholdet er derfor lik den relative følsomheten til sløyfe-transferfunksjonen multiplisert med $N(s)$. Siden tallverdien av $N(s)$ skal være et lite tall, vil den relative følsomheten til “det lukkede systemet” være mindre enn den relative følsomheten til “det åpne systemet”.

Med andre ord gjør tilbakekopling at innvirkningen av feil i anslag for prosessparametre eller forandring av prosessparametre over tid, blir redusert.

EKSEMPEL 7.8: S_α^M med hensyn på K og T for det samme systemet

Vi betrakter det samme systemet som i eksempel 7.2, og søker den relative følsomheten til følgeforholdet med hensyn på K og T gitt ved (7.32) og (7.29). Vi ser først på parameteren K ($\alpha = K$). Likning (7.29) gir

$$S_K^{h_0} = \frac{K}{\frac{K}{s(1+Ts)}} \frac{1}{s(1+Ts)} = 1$$

Den relative følsomhet til følgeforholdet med hensyn på forsterkningen K blir derfor

$$S_K^M = N(s)$$

Siden

$$N(s) = \frac{1}{K} \frac{s(1+Ts)}{1 + \frac{1}{K}s + \frac{T}{K}s^2} \rightarrow \frac{s}{K} \Big|_{|s| \rightarrow 0}$$

betyr dette at $S_K^M \rightarrow 0$ for lave frekvenser.

Betrakter vi i stedet følsomheten med hensyn på T , finner vi

$$S_T^{h_0} = \frac{T}{\frac{K}{s(1+Ts)}} \frac{-Ks^2}{s^2(1+Ts)^2} = -\frac{Ts}{1+Ts}$$

Den relative følsomhet til følgeforholdet med hensyn på T blir derfor

$$S_K^M = -N(s) \frac{Ts}{1+Ts} \rightarrow -\frac{Ts^2}{K} \Big|_{|s| \rightarrow 0}$$

Dette betyr at følsomheten til det lukkede systemet overfor endringer i tidskonstanten T blir enda mindre ved lave frekvenser, enn følsomheten med hensyn på forsterkningen K .



Stabilitet

- 8.1 Innledning og motivering 263
- 8.2 Definisjoner på stabilitet 264
- 8.3 Algebraiske kriterier for stabilitet 270
- 8.4 Nyquists stabilitetskriterium 284

8.1 INNLEDNING OG MOTIVERING

I det forrige kapitlet diskuterte vi tilbakekoplingsløyfen og viste hvordan en regulator kan benyttes til å endre et systems egenskaper, målt ved faktorer som følgeforhold og reguleringsgrad. En meget viktig egenskap som skal behandles i dette kapitlet, er *stabilitet* av tilbakekoplingsløyfen. En *ustabil* tilbakekopplingsløyfe gir en tidsrespons som “vokser over alle grenser”. Stabilitet er derfor nødvendig for at et reguleringsystem overhodet skal være anvendbart. Det er viktig å kunne forutsi om et system som er under planlegging, vil bli stabilt eller ustabil. Dersom beregninger viser at systemet blir ustabil, må vi vite hva som skal gjøres for å oppnå stabilitet. Vi ønsker også å beregne hvilke marginer som foreligger, før et stabilt system blir ustabil. Vi skal benytte et kjent eksempel til å illustrere problemstillingen før vi behandler stabilitet mer formelt.

EKSEMPEL 8.1: Likestrømsmotor og stabilitet

Likestrømsmotoren har vært benyttet i en rekke eksempler. I eksempel 4.13 studerte vi posisjonsregulering av motoren med en enkel proporsjonalregulator som vist i figur 4.29. Polene til det lukkede systemet ble beregnet for ulike positive verdier av forsterkningen K . Dette er illustrert i figur 4.30. Dersom

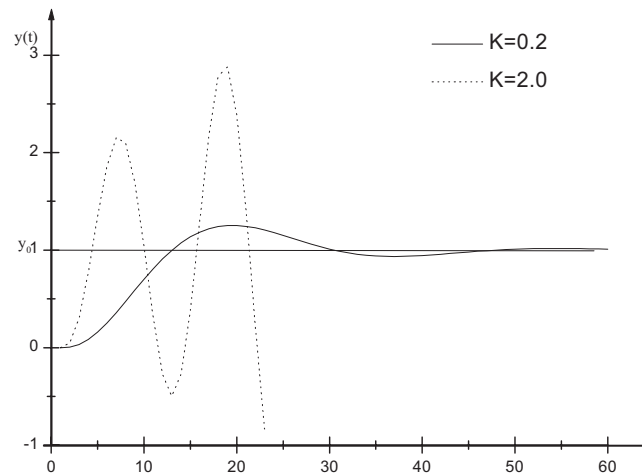
$$K > \frac{T_1 + T_2}{T_1 T_2}, \quad (8.1)$$

beveger to av systemets tre poler seg over i høyre halvplan, det vil si at de får en positiv realdel. Dette betyr at tidsresponsen til systemet vil gå mot uendelig: Vi får en oscillerende tidsrespons med en amplitude som “øker over alle grenser” som vist med den stiplede linjen i figur 8.1 for $K = 2.0$.

Systemet er ustabil¹ for denne verdien av K . Stabilitetsgrensen gitt av (8.1) er 1.2. Tidsresponsen for $K = 0.2$ er vist med heltrukken linje. Regulerings-sløyfen er stabil for denne verdien av K .

Figur 8.1

Tidsrespons ved settpunktendring for $K = 0.2$ og $K = 2.0$, $T_1 = 5$ og $T_2 = 1$



8.2 DEFINISJONER PÅ STABILITET

Det finnes en omfattende matematisk teori for stabilitet. Teorien blir mer komplisert for *ulineære* dynamiske systemer. Stabilitetsbegrepet er der langt mer mangfoldig enn for lineære systemer, fordi vi får mange former for dynamikk som har elementer av *både* ustabilitet og stabilitet i seg. Det henvises til for eksempel Khalil (1996) og Vidyasagar (1993).

I denne boka betrakter vi utelukkende stabilitet for tidsinvariante lineære systemer (“LTI-systemer”).

Stabilitet defineres i hovedsak på to måter:

- Inngangs-/utgangsstabilitet
- Stabilitet i tilstandsrommet

¹Men i praksis vil et system alltid ha noen mekanismer som begrenser økningen, for eksempel at en ventil når sin maksimale åpning eller at en motor når sin maksimale effekt. Systemet er ulineært ved store utslag.

8.2.1 Inngangs-/utgangs stabilitet

Vi skal først definere inngangs-/utgangs stabilitet, og tar utgangspunkt i et tilbakekoplet system $h(s)$ med inngang (referanse) y_0 og utgang y , slik at vi har $y(s) = h(s)y_0(s)$. Se figur 8.2, venstre del. Da har vi følgende definisjoner:

Systemet er **asymptotisk inngangs-/utgangsstabilt** dersom $y \rightarrow 0$ når $t \rightarrow \infty$, når y_0 har endelig varighet og endelig amplitude.

Systemet er **inngangs-/utgangsstabilt** dersom $|y| < \infty$ i alle tidspunkt $t \geq 0$, når y_0 har endelig varighet og endelig amplitude. Dette kalles også et **marginalt stabilt** system.

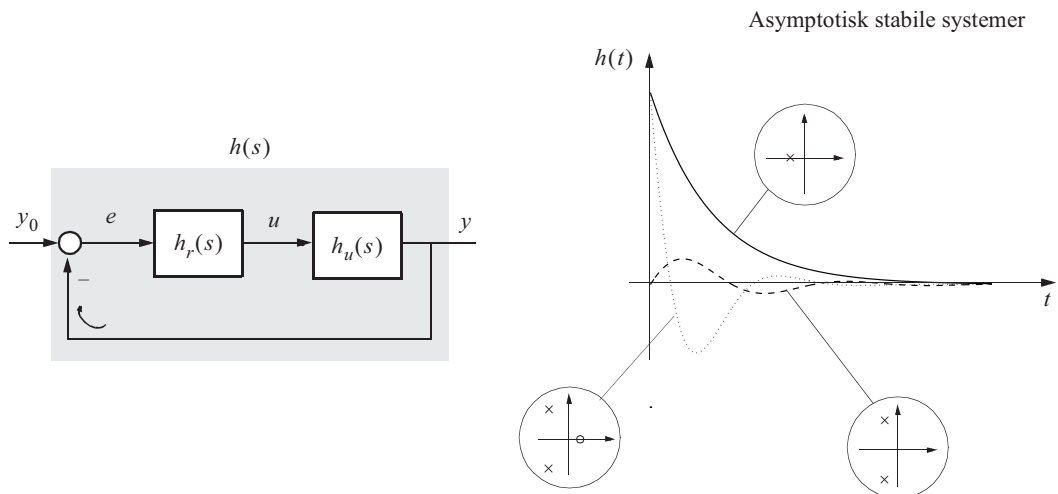
Dersom systemet ikke er stabilt, er det **ustabilt**

Marginal stabilitet for lineære systemer er et grensetilfelle som bare inntreffer for *ett eksakt sett* av systemparametre. Vi interesserer oss derfor primært for asymptotisk stabilitet kontra ustabilitet, som inntreffer for parameterverdier innenfor avgrensede *områder*. Av denne grunn sier vi også ofte bare “stabilitet” når vi mener asymptotisk stabilitet.

Vi kan definere asymptotisk stabilitet, marginal stabilitet og ustabilitet ved å betrakte *impulsresponsen* $h(t)$ til et system. Hvis $y_0(t) = \delta(t)$, blir $y(t) = h(t)$. Selv om en impulsfunksjon $\delta(t)$ ikke har endelig amplitude slik som nevnt i definisjonen ovenfor, har den infinitesimal kort varighet. Derfor kan impulsresponsen brukes til å klassifisere et lineært system med hensyn på dets inngangs-/utgangs stabilitet.

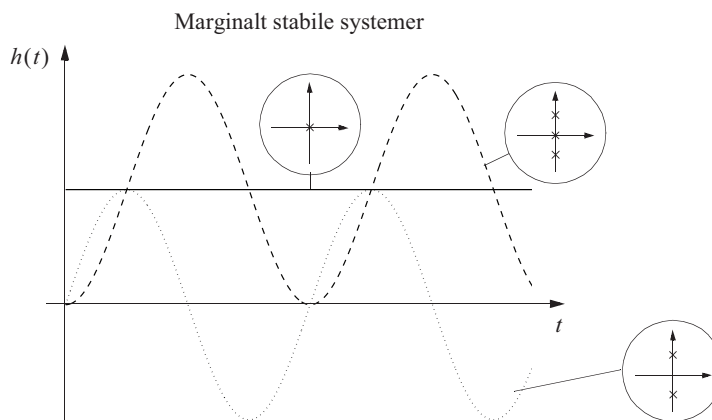
Figurene 8.2-8.4 viser eksempler på impulsresponsen, poler og nullpunkter til henholdsvis asymptotisk stabile systemer, marginalt stabile systemer og ustabile systemer.

Figur 8.2
Tilbakekoplet system, og dets impulsrespons for tre tilfeller av asymptotisk stabilitet

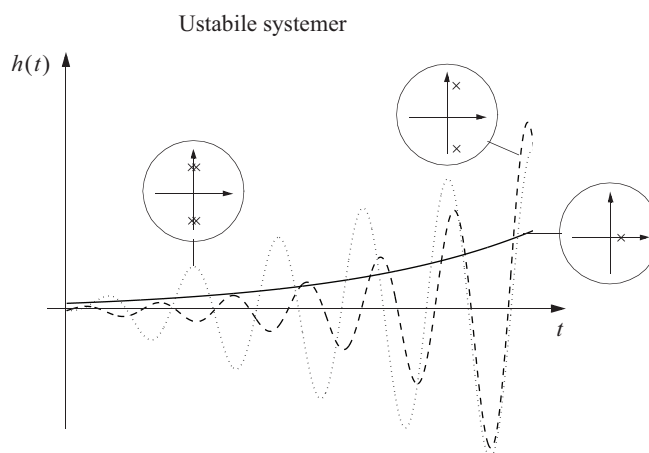


Figur 8.3

Impulsrespons,
marginalt stabile
systemer

**Figur 8.4**

Impulsrespons,
ustabile systemer



I figurene ovenfor er også vist poler (kryss) og nullpunkter (sirkler) i det lukkede systems transferfunksjon $h(s)$. Vi vet fra tidligere kapitler at impulsresponsens forløp når $t \rightarrow \infty$, avhenger av polenes plassering. Fordelen med å bruke polenes plassering som kriterium, er at vi da ikke bare kan *definere* stabilitet, men også *undersøke* den for et gitt system. Vi har:

Systemet er asymptotisk stabilt dersom $\text{Re}(\lambda_i) < 0$ for alle poler λ_i i $h(s)$.

Systemet er marginalt stabilt dersom polene er slik at en eller flere $\text{Re}(\lambda_i) = 0$, men ingen av disse er multiple.

Hvis en eller flere $\text{Re}(\lambda_i) > 0$, er systemet ustabil.

Dersom en pol ligger i origo, får vi et marginalt stabilt system. Et sett av to kompleks konjugerte poler på den imaginære akse tilsvarer en oscillasjon med konstant frekvens og amplitude. Dette er også et marginalt stabilt system.

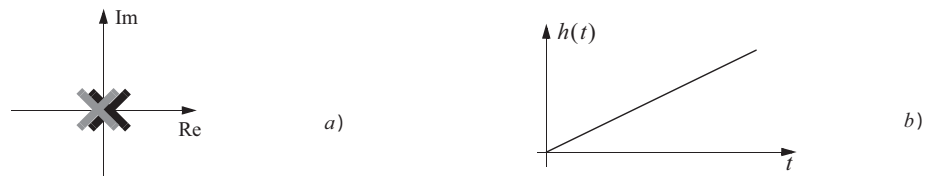
Nullpunktene i det lukkede systems transferfunksjon har ingen betydning for stabiliteten, bare for det transiente tidsforløpet.

EKSEMPEL 8.2: Dobbelintegrator

Vi tar for oss et system med to poler i origo som vist i figur 8.5 a). Impulsresponsen er

$$h(t) = L^{-1}\left(\frac{1}{s^2}\right) = \mu_2(t) = \text{enhetsrampefunksjon}$$

Figur 8.5
a) Plassering av egenverdiene
b) Impulsresponsen



Impulsresponsen i figur 8.5 b) viser at systemet er *ustabilt* og *ikke marginalt stabilt*, selv om egenverdiene ligger på den imaginære akse.

■

8.2.2 Stabilitet i tilstandsrommet

Stabilitet i tilstandsrommet er knyttet til oppførselen av tilstanden \mathbf{x} til et system, i stedet for prosessutgangen y . Systemet til venstre i figur 8.2 kan uttrykkes på tilstandsromform,

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}y_0 \\ y &= \mathbf{c}^T \mathbf{x}\end{aligned}\tag{8.2}$$

Merk her at systemet er monovariabelt, og at inngangsvariabelen er y_0 .

Systemet i (8.2) er **asymptotisk stabilt** i tilstandsrommet dersom $\mathbf{x} \rightarrow \mathbf{0}$ når $t \rightarrow \infty$, etter å ha vært påvirket av en eksitasjon (y_0) med endelig varighet og endelig amplitude.

Systemet er **stabilt** i tilstandsrommet dersom $|\mathbf{x}| < \infty$ i alle tidpunkt $t \geq 0$, når y_0 etter å ha vært påvirket av en eksitasjon (y_0) med endelig varighet og endelig amplitude. Dette kalles også et **marginalt stabilt** system.

Dersom systemet ikke er asymptotisk eller marginalt stabilt, er det **ustabilt**

Denne stabilitetsdefinisjonen har det til felles med den første i forrige avsnitt, at den baserer seg på responsen til systemet, og dermed bare er egnet til å definere, ikke undersøke, stabilitet. Vi kan undersøke stabilitet i tilstandsrommet ved å finne egenverdiene til systemet (8.2):

Systemet i (8.2) er asymptotisk stabilt dersom $\operatorname{Re}(\lambda_i) < 0$ for alle egenverdier λ_i i \mathbf{A} .

Systemet er marginalt stabilt dersom egenverdiene er slik at $\operatorname{Re}(\lambda_i) \leq 0$, og ingen egenverdier med $\operatorname{Re}(\lambda_i) = 0$ er multiple.

Hvis én eller flere $\operatorname{Re}(\lambda_i) > 0$, er systemet ustabil.

Vi merker oss at siden poler i en transferfunksjon svarer til egenverdier i en tilstandsrommodell av samme system, så er dette samme kriterium som i forrige avsnitt. Men for at alle egenverdier skal framkomme som poler i transferfunksjonen, må systemet være styrbart og observerbart. Mer om det lenger nede.

Vi skal sjekke kravene til stabilitet i tilstandsrommet med utgangspunkt i (8.2). Benytter vi oss av det som er funnet i avsnitt 3.3, finner vi følgende løsning av (8.2), når vi antar $y_0(t) = 0$ for alle $t > t_0$, men ikke for $t \leq t_0$:

$$\mathbf{x}(t) = \mathbf{M}e^{\mathbf{A}(t-t_0)}\mathbf{M}^{-1}\mathbf{x}(t_0) \quad (8.3)$$

$\mathbf{x}(t_0)$ er tilstanden til systemet umiddelbart etter at inngangssignalet $y_0(t)$ har opphørt å påvirke det. Likning (8.3) gir

$$\mathbf{x}(t) \xrightarrow[t \rightarrow \infty]{} 0 \text{ når alle } \operatorname{Re}(\lambda_i) < 0 \quad (8.4)$$

Dette tilsvarer asymptotisk stabilitet.

En fordel med å undersøke stabilitet i tilstandsrommet, i motsetning til inngangs-/utgangsstabilitet, er at man da også kan ta for seg **autonome** systemer, som ikke har noe inngangssignal,

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \quad (8.5)$$

Et slikt system var modellen for kapitalakkumulasjon i eksempel 2.3, som ble ustabil for visse parameterverdier.

Systemet (8.2) er monovariabelt. Anta nå i stedet et multivariabelt system,

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}_0 \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \end{aligned} \quad (8.6)$$

Vil dette bety noe for hvordan vi undersøker stabilitet?

Nei, stabilitet i tilstandsrom er fullstendig bestemt av egenskapene til \mathbf{A} . Derfor er det likegyldig hvilket inngangs-/utgangsforhold vi måtte velge å betrakte i et multivariabelt system; *stabilitet er en systemegenskap, og vil derfor være den samme uansett.*

Som nevnt ovenfor, svarer poler og egenverdier helt til hverandre hvis systemet er styrbart og observerbart. Men fra kapittel 5 vet vi at hvis dette ikke er oppfylt, vil transferfunksjonen bli av lavere orden enn tilstandsrommodellen, fordi ledd i telleren forkortes mot ledd i nevneren; poler kanselleres av nullpunkter. Dette skal vi se på i et eksempel:

EKSEMPEL 8.3: Inngangs-utgangsstabilitet kontra stabilitet i tilstandsrommet

I eksempel 5.3 studerte vi blant annet styrbarhet og observerbarhet av dynamiske systemer med utgangspunkt i systemet i figur 5.4. Analysen viste at systemet ikke er observerbart for $T_2 = T_3$. I dette tilfellet blir (5.15).

$$\dot{\mathbf{x}} = \begin{bmatrix} -\frac{1}{T_1} & 0 \\ 0 & -\frac{1}{T_2} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \frac{1}{T_1} \\ \frac{1}{T_2} \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}$$

Anta at dette systemet reguleres ved $u = K(y_0 - y) = K(y_0 - x_1)$, $K > 0$

Dette systemet er alltid inngangs-/utgangsstabil. Systemet er imidlertid ustabil i tilstandsrommet dersom $T_2 < 0$. I dette tilfelle vil $x_2 \rightarrow \infty$ når $t \rightarrow \infty$. Men dette registreres ikke på utgangen, da y ikke påvirkes av x_2 .

I figur 5.4 er transferfunksjonen for systemet avbildet. Når systemet ikke er observerbart, kanselleres telleren $(1 + T_3s)$ mot nevneren $(1 + T_2s)$. Når $T_2 < 0$, kanselleres altså en pol i høyre halvplan. Dette er årsaken til at ustabiliteten “ikke synes” med hensyn på inngangs-/utgangsstabilitet.

Tilsvarende vil skje hvis vi i stedet har $T_1 = T_3$ og $T_1 < 0$. Da er systemet ustabil og ikke styrbart. Den ustabile tilstanden x_1 kan ikke påvirkes fra u . Men transferfunksjonen blir stabil og av 1. orden, fordi telleren $(1 + T_3s)$ kanselleres mot nevneren $(1 + T_1s)$.



For å oppsummere: Stabilitet i tilstandsrommet vil alltid medføre inngangs- /utgangsstabilitet. Det motsatte er ikke alltid tilfellet.

I fortsettelsen av dette kapitlet benytter vi begrepene stabilitet både om inngangs-/utgangsstabilitet og stabilitet i tilstandsrommet. Det vil framgå av sammenhengen hvilke type stabilitet som diskuteres.

8.3 ALGEBRAISKE KRITERIER FOR STABILITET

Bestemmelse av egenverdier (poler) for dynamiske systemer medfører som kjent løsning av systemets karakteristiske likning eller bestemmelse av røttene i systemets karakteristiske polynom.

Systemet i (8.2) har **karakteristisk likning**

$$|\mathbf{A} - \lambda \mathbf{I}| = 0 \quad (8.7)$$

Det karakteristiske polynom framkommer ved utvikling av determinanten

$$\lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 = 0 \quad (8.8)$$

Vi trenger ikke å finne løsningen av (8.8) for å undersøke om systemet blir stabilt. Det er nok å finne betingelsene for at $\text{Re } \lambda < 0$. Det finnes en rekke hensiktsmessige metoder for å finne betingelsene som koeffisientene i det karakteristiske polynom (8.8) må tilfredsstille, for at systemet skal bli stabilt. Disse metodene er basert på samme teoretiske grunnlag, men har litt forskjellige bruksegenskaper. De mest anvendte metodene er *Rouths kriterium* og *Hurwitz's kriterium*.

8.3.1 Rouths kriterium

Rouths kriterium (algoritme) tar sitt utgangspunkt i det karakteristiske polynom

$$\alpha_n\lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 = 0 \quad (8.9)$$

der vi for fullstendighetens skyld har innført en koeffisient α_n foran leddet med høyeste potens.

Følgende kan sies om polynomet i (8.9):

1. Antallet røtter er lik polynomets grad n .
2. Fordi polynomets koeffisienter er reelle, forekommer røttene dersom de er komplekse, i kompleks konjugerte par.
3. En nødvendig, men ikke tilstrekkelig betingelse for at røttene skal ligge i venstre halvplan ($\text{Re } \lambda < 0$), er at alle polynomets koeffisienter har samme fortegn.
4. Hvis ikke alle koeffisientene i polynomet har samme fortegn, eller dersom en eller flere av koeffisientene er lik null, vil en eller flere av røttene ligge i høyre halvplan eller på den imaginære akse.

Dersom den nødvendige betingelsen i punkt 3 ovenfor er tilfredsstillt, kan vi undersøke polynomet ved hjelp av Rouths algoritme. Denne består i å bygge opp en koeffisienttabell (n er i eksemplet et oddetall)

$$\begin{array}{cccc}
 \alpha_n & \alpha_{n-2} & \cdots & \alpha_1 \\
 \alpha_{n-1} & \alpha_{n-3} & \cdots & \alpha_0 \\
 \beta_{n-1} & \beta_{n-3} & \cdots & \\
 \zeta_{n-1} & \zeta_{n-3} & \cdots & \\
 \cdot & & & \\
 \cdot & & & \\
 \cdot & & & \\
 \sigma_{n-1} & \sigma_{n-3} & & \\
 \eta_{n-1} & & & \\
 \omega_{n-1} & & &
 \end{array} \tag{8.10}$$

De nye koeffisientene i talltabellen framkommer etter følgende regler:

$$\beta_{n-1} = \frac{\begin{array}{cc} \alpha_n & \alpha_{n-2} \\ \alpha_{n-1} & \alpha_{n-3} \end{array}}{\alpha_{n-1}} = \frac{\alpha_{n-1}\alpha_{n-2} - \alpha_n\alpha_{n-3}}{\alpha_{n-1}} \tag{8.11}$$

$$\beta_{n-3} = \frac{\begin{array}{cc} \alpha_n & \alpha_{n-4} \\ \alpha_{n-1} & \alpha_{n-5} \end{array}}{\alpha_{n-1}} = \frac{\alpha_{n-1}\alpha_{n-4} - \alpha_n\alpha_{n-5}}{\alpha_{n-1}}$$

OSV.

Koeffisientene i 4. rad i talltabellen framkommer på følgende måte

$$\zeta_{n-1} = \frac{\begin{array}{cc} \alpha_{n-1} & \alpha_{n-3} \\ \beta_{n-1} & \beta_{n-3} \end{array}}{\beta_{n-1}} = \frac{\beta_{n-1}\alpha_{n-3} - \alpha_{n-1}\beta_{n-3}}{\beta_{n-1}} \tag{8.12}$$

$$\zeta_{n-3} = \frac{\begin{array}{cc} \alpha_{n-1} & \alpha_{n-5} \\ \beta_{n-1} & \beta_{n-5} \end{array}}{\beta_{n-1}} = \frac{\beta_{n-1}\alpha_{n-5} - \alpha_{n-1}\beta_{n-5}}{\beta_{n-1}}$$

OSV.

Legg merke til at koeffisientene i hver rad fra nummer tre og nedover framkommer ved å benytte de to ovenforliggende radene.

Talltabellen antar form av en trekant, der det til slutt bare er to koeffisienter igjen, for eksempel η_{n-1} og ω_{n-1} . Rouths kriterium uttrykker nå:

5. En nødvendig og tilstrekkelig betingelse for at røttene til polynomet i (8.9) skal ligge i venstre halvplan, er at alle koeffisientene i venstre kolonne i Rouths talltabell har samme fortegn. Dersom en eller flere av koeffisientene i venstre kolonne har forskjellig fortegn fra de andre, vil antallet røtter i høyre halvplan kunne bestemmes ut fra antall fortegnsskift som inntreffer ved gjennomløp ovenfra og nedover.

EKSEMPEL 8.4: Rouths kriterium brukt på et 3. ordens system

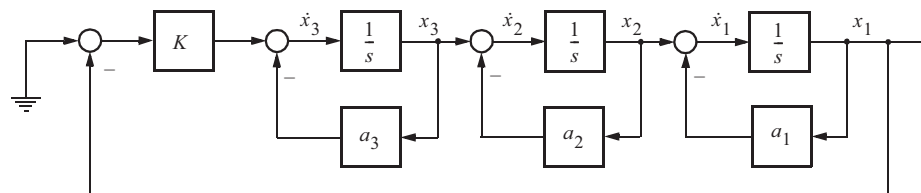
Figur 8.6 viser et elementært blokkdiagram for en 3. ordens prosess med proporsjonalregulering. Differensiallikningene for dette systemet er

$$\begin{aligned}\dot{x}_1 &= -a_1x_1 + x_2 \\ \dot{x}_2 &= -a_2x_2 + x_3 \\ \dot{x}_3 &= -Kx_1 + a_3x_3\end{aligned}\quad (8.13)$$

Systemmatrisen er som følger

$$\mathbf{A} = \begin{bmatrix} -a_1 & 1 & 0 \\ 0 & -a_2 & 1 \\ -K & 0 & -a_3 \end{bmatrix}\quad (8.14)$$

Figur 8.6
Blokkdiagram for systemet



Den karakteristiske likningen for (8.14) blir $|\mathbf{A} - \lambda\mathbf{I}| = 0$ som gir

$$\lambda^3 + (a_1 + a_2 + a_3)\lambda^2 + (a_1a_2 + a_1 + a_3 + a_2a_3)\lambda + a_1a_2a_3 + K = 0\quad (8.15)$$

Vi velger først å undersøke (8.15) ved hjelp av Rouths kriterium.

Da må vi først forlange at alle koeffisientene i (8.15) har samme fortegn, det vil si

$$\begin{aligned} a_1 + a_2 + a_3 &> 0 \\ a_1 a_2 + a_1 a_3 + a_2 + a_3 &> 0 \\ a_1 a_2 a_3 + K &> 0 \end{aligned} \quad (8.16)$$

Likning (8.16) inneholder nødvendige betingelser for stabilitet. Forutsatt at (8.16) er tilfredsstillt, bruker vi Rouths algoritme og finner koeffisienttabellen

$$\begin{array}{r} 1 \\ a_1 + a_2 + a_3 \\ a_1 a_2 + a_1 a_3 + a_2 a_3 - \frac{a_1 a_2 a_3 + K}{a_1 + a_2 + a_3} \\ a_1 a_2 a_3 + K \end{array} \quad , \quad \begin{array}{r} a_1 a_2 + a_1 a_3 + a_2 a_3 \\ a_1 a_2 a_3 + K \end{array} \quad (8.17)$$

Alle koeffisientene i første kolonne må ha samme fortegn ifølge Rouths kriterium. I tillegg til betingelsene i (8.16) får vi derfor

$$-a_1 a_2 a_3 < K(a_1 a_2 + a_1 a_3 + a_2 a_3)(a_1 + a_2 + a_3) - a_1 a_2 a_3 \quad (8.18)$$

Vi får et særtilfelle dersom $a_1 = a_2 = a_3 = a > 0$. Denne verdien tilfredsstiller de to første kravene i (8.16), og kravet i (8.18) vil da bli

$$-a^3 < K < 8a^3 \quad (8.19)$$

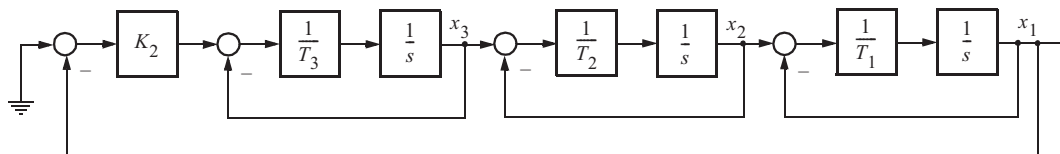
Med $a = 1$ får vi $-1 < K < 8$.

■

EKSEMPEL 8.5: Omforming av 3. ordens system og bruk av Rouths kriterium

Systemet vi har studert i eksempel 8.4 kan representeres som vist i figur 8.7.

Figur 8.7
Blokkdiagram for systemet



Differensiallikningene for dette systemet vil bli

$$\begin{aligned} \dot{x}_1 &= \frac{1}{T_1}(-x_1 + x_2) \\ \dot{x}_2 &= \frac{1}{T_2}(-x_2 + x_3) \\ \dot{x}_3 &= \frac{1}{T_3}(-K_2 x_1 - x_3) \end{aligned} \quad (8.20)$$

og den karakteristiske likningen blir

$$T_1 T_2 T_3 \lambda^3 + (T_1 T_2 + T_1 T_3 + T_2 T_3) \lambda^2 + (T_1 + T_2 + T_3) \lambda + 1 + K_2 = 0 \quad (8.21)$$

Bruker vi Rouths kriterium på (8.21), finner vi

$$-1 < K_2 < \left(\frac{1}{T_1} + \frac{1}{T_2} + \frac{1}{T_3} \right) (T_1 + T_2 + T_3) - 1 \quad (8.22)$$

Dette svaret tilsvare det som er funnet i (8.18) dersom $a_1 = T_1$, $a_2 = T_2$, $a_3 = T_3$ og $K_2 = K/(a_1 a_2 a_3)$. Det er interessant å konstatere at 8 er den minste verdien høyresiden i (8.22) kan få. Denne verdien oppnås når $T_1 = T_2 = T_3$. Den øvre grensen til K_2 vil bli større enn 8 for alle andre verdier av T_1 , T_2 og T_3 .

■

8.3.2 Hurwitz's kriterium

Som for Rouths kriterium gjelder det at en *nødvendig*, men *ikke tilstrekkelig*, betingelse for at røttene i det karakteristiske polynom skal ligge i venstre halvplan ($\text{Re } \lambda < 0$), er at alle koeffisientene har samme fortegn. Det kan forøvrig vises at Rouths og Hurwitz's kriterier er ekvivalente.

Hurwitz's kriterium tar utgangspunkt i koeffisientene til polynomet i (8.9) og uttrykker at en *nødvendig* og *tilstrekkelig* betingelse for at røttene til polynomet skal ligge i venstre halvplan, er at alle determinantene $|\Delta_1|$, $|\Delta_2|$, ..., $|\Delta_{n-1}|$ er positive. Δ_k er matriser som er markert i følgende koeffisienttabell

$$\begin{array}{l}
 \Delta_1 \begin{array}{|c|c|c|c|c|} \hline \alpha_{n-1} & \alpha_{n-3} & \alpha_{n-5} & \alpha_{n-7} & \alpha_{n-9} \\ \hline \alpha_n & \alpha_{n-2} & \alpha_{n-4} & \alpha_{n-6} & \alpha_{n-8} \\ \hline 0 & \alpha_{n-1} & \alpha_{n-3} & \alpha_{n-5} & \alpha_{n-7} \\ \hline 0 & \alpha_n & \alpha_{n-2} & \alpha_{n-4} & \alpha_{n-6} \\ \hline 0 & 0 & \alpha_{n-1} & \alpha_{n-3} & \alpha_{n-5} \\ \hline 0 & 0 & \alpha_n & \alpha_{n-2} & \alpha_{n-4} \\ \hline 0 & 0 & 0 & \alpha_{n-1} & \alpha_{n-3} \\ \hline 0 & 0 & 0 & \alpha_n & \alpha_{n-2} \\ \hline \end{array} \cdot \cdot \cdot \\
 \Delta_2 \cdot \cdot \cdot \\
 \Delta_3 \cdot \cdot \cdot \\
 \Delta_4 \cdot \cdot \cdot \\
 \Delta_5 \cdot \cdot \cdot
 \end{array} \quad (8.23)$$

I tabellen skal vi alltid ha $n - k \geq 0$. Dette er fordi den laveste indeksen i (8.9) er 0. Vi skal derfor sette inn $\alpha_{n-k} = 0$ ved verdier av k som gir $n - k < 0$.

Vi skal illustrere bruken av dette algebraiske kriteriet med en del konkrete eksempler.

EKSEMPEL 8.6: Hurwitz's kriterium

Vi bruker nå Hurwitz's kriterium på systemet i eksempel 8.4 og finner Hurwitz-matrisen (8.23) ved innsetting av koeffisientene fra (8.15):

$$\Delta_1 \begin{array}{|c} a_1 + a_2 + a_3 \\ \hline 1 \end{array} \quad \begin{array}{|c} a_1 a_2 a_3 + K \\ a_1 a_2 + a_1 a_3 + a_2 a_3 \\ \hline \end{array} \quad (8.24)$$

I tillegg til at koeffisientene i (8.15) skal ha samme fortegn som angitt i (8.16), må vi ifølge Hurwitz's kriterium ha at

$$\begin{aligned} |\Delta_1| &= a_1 + a_2 + a_3 > 0 \\ |\Delta_2| &= (a_1 + a_2 + a_3)(a_1 a_2 + a_1 a_3 + a_2 a_3) - (a_1 a_2 a_3 + K) > 0 \end{aligned} \quad (8.25)$$

Disse betingelsene leder til det samme svaret som vi har funnet i (8.18).

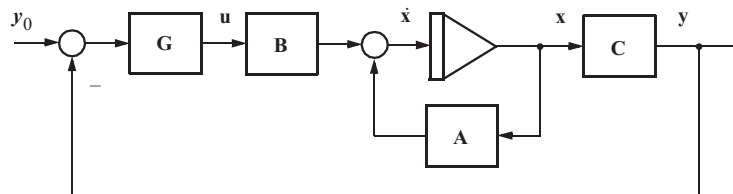


Rouths kriterium og Hurwitz's kriterium egner seg godt til å undersøke stabilitet av multivariable lineære systemer beskrevet på tilstandsromform. Et eksempel vil illustrere dette.

EKSEMPEL 8.7: Bruk av Rouths og Hurwitz's kriterium på et multivariabelt system

Det multivariable reguleringsystemet i figur 8.8 som har en prosess med to pådrag, to tilstander og to målinger antas å være karakterisert ved følgende matriser:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} -3 & 4 \\ 0 & -5 \end{bmatrix}, & \mathbf{B} &= \begin{bmatrix} 2 & -1 \\ 0 & 4 \end{bmatrix} \\ \mathbf{C} &= \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \end{aligned} \quad (8.26)$$



Figur 8.8
Blokkdiagram for
det multivariable
systemet

Reguleringssystemet antas å være bygget opp som to uavhengige proporsjonalregulatorer hvilket medfører at regulatormatrisen \mathbf{G} vil bli en diagonalmatrise

$$\mathbf{G} = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix}$$

Vi skal finne ut hvilke verdier g_1 og g_2 må ha for at det lukkede systemet skal være stabilt. Dersom $\mathbf{y}_0 = \mathbf{0}$, finner vi

$$\dot{\mathbf{x}}(\mathbf{A} - \mathbf{BGC})\mathbf{x} = \tilde{\mathbf{A}}\mathbf{x} \quad (8.27)$$

Den karakteristiske likningen for $\tilde{\mathbf{A}}$ blir

$$\lambda^2 + ((3 + 4g_1) + (5 + 4g_2))\lambda + (3 + 4g_1)(5 + 4g_2) = 0$$

Bruk av Rouths kriterium eller Hurwitz's kriterium gir følgende krav til stabilitet:

$$\begin{aligned} g_1 + g_2 &> -2 \\ (3 + 4g_1)(5 + 4g_2) &> 0 \end{aligned}$$

som igjen reduseres til

$$g_1 > -\frac{3}{4} \quad \text{og} \quad g_2 > -\frac{5}{4}$$

■

8.3.3 Algebraiske kriterier for et system gitt ved sin transferfunksjon

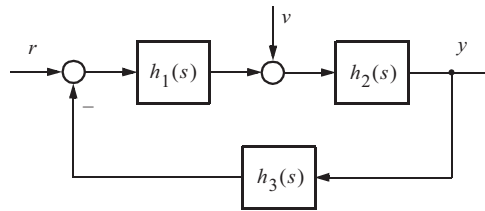
I utledningene foran gikk vi ut fra at systemet var karakterisert ved differensiallikninger på tilstandsromform. I kapittel 4 har vi imidlertid vist at den karakteristiske likningen for et dynamisk system også framkommer når vi beskriver systemet med transferfunksjoner (transfermatriser) som er et resultat av Laplace-transformasjon av differensiallikningene. Vi vet at et monovariabelt system beskrevet av n ordinære lineære differensiallikninger vil være karakterisert ved en transferfunksjon av typen

$$h(s) = \frac{\rho_{n-1}s^{n-1} + \dots + \rho_1s + \rho_0}{s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_1s + \alpha_0} \quad (8.28)$$

der nevneren utgjøres av det karakteristiske polynomet for systemet. Kjenner vi derfor systemets transferfunksjon, kan vi finne betingelsene for stabilitet ved å benytte enten Rouths kriterium eller Hurwitz's kriterium på polynomet i transferfunksjonens nevner.

Det knytter seg særlig interesse til stabilitet av tilbakekoblede systemer. Vi skal se på det monovariabelt tilbakekoblede systemet vist i figur 8.9.

Figur 8.9
Blokkdiagram for monovariabelt tilbakekoplede system



Dette tilbakekoblede systemet¹ kan beskrives med (fra nå av er avhengighet av s underforstått, så vi skriver h_i istedet for $h_i(s)$)

$$\begin{aligned} \frac{y}{r}(s) &= \frac{h_1 h_2}{1 + h_1 h_2 h_3} = \frac{h_1 h_2}{1 + h_0} \\ &= N(s) h_1 h_2 \end{aligned} \quad (8.29)$$

og

$$\frac{y}{v}(s) = \frac{h_2}{1 + h_0} = N(s) h_2 \quad (8.30)$$

Plasseringen til systemets poler (egenverdier) som er gitt ved røttene til polynomene i de to transferfunksjonenes nevnerer, er bestemmende for stabilitet. Antar vi at transferfunksjonene er rasjonale funksjoner betegnet med

$$h_i = \frac{t_i(s)}{n_i(s)}$$

der $t_i(s)$ og $n_i(s)$ er polynomer i s , finner vi ved innsetting i (8.29) og (8.30)

$$\frac{y}{r}(s) = \frac{n_3(s) t_1(s) t_2(s)}{n_1(s) n_2(s) n_3(s) + t_1(s) t_2(s) t_3(s)} = \frac{n_3(s) t_1(s) t_2(s)}{n_0(s) + t_0(s)} \quad (8.31)$$

og

$$\frac{y}{v}(s) = \frac{n_1(s) n_3(s) t_2(s)}{n_1(s) n_2(s) n_3(s) + t_1(s) t_2(s) t_3(s)} = \frac{n_1(s) n_3(s) t_2(s)}{n_0(s) + t_0(s)} \quad (8.32)$$

¹Merk at inngangen r her ikke kan betraktes som en referanse, siden vi ikke har enhetstilbakekopling. Men dette spiller ingen rolle for det vi drøfter i dette avsnittet; at stabilitet ikke avhenger av hvilke innganger og utganger vi betrakter.

Vi ser her at de to transferfunksjonene har samme polynom i nevneren. Polynomets røtter angir det lukkede systemets poler. Dette bekrefter at stabilitet av et lineært system er en systemegenskap som er uavhengig av hvilke påvirkninger og responser som studeres.

Betingelsene for stabilitet av et tilbakekoplet system kan nå bestemmes ved hjelp av **Rouths kriterium** eller **Hurwitz's kriterium** anvendt på polynomet $n_0(s) + t_0(s)$ der de to leddene er gitt av telleren og nevneren i sløyfetransferfunksjonen,

$$h_0 = \frac{t_0(s)}{n_0(s)}$$

Vi vil illustrere bruken av dette resultatet med et eksempel.

EKSEMPEL 8.8: Rouths kriterium og sløyfetransferfunksjonen 1

Vi velger det samme systemet som tidligere er undersøkt i eksempel 8.5 og vist i figur 8.7. Sløyfetransferfunksjonen for dette systemet er

$$h_0(s) = \frac{K}{(1 + T_1s)(1 + T_2s)(1 + T_3s)} \quad (8.33)$$

som gir

$$n_0(s) + t_0(s) = T_1T_2T_3s^3 + (T_1T_2 + T_1T_3 + T_2T_3)s^2 + (T_1 + T_2 + T_3)s + 1 + K$$

Dette polynomet er lik det vi fant i (8.21). Vi vil således få de samme betingelsene for stabilitet. Disse er angitt i (8.22).



EKSEMPEL 8.9: Rouths kriterium og sløyfetransferfunksjonen 2

I eksempel 4.13 undersøkte vi systemet som er vist i figur 4.29. Oppgaven var posisjonsregulering av en likestrømsmotor med proporsjonalregulator. Et problem i den forbindelse var å finne hvilken verdi av forsterkningen K som medførte at polene i det lukkede systemet beveget seg over fra venstre halvplan til høyre halvplan. Med andre ord, hvilken verdi av K gir stabilitetsgrensen? Figur 4.26 gir

$$h_0(s) = \frac{K}{s(1 + T_1s)(1 + T_2s)} \quad (8.34)$$

som leder til polynomet

$$n_0(s) + t_0(s) = T_1T_2s^3 + (T_1 + T_2)s^2 + s + K$$

Rouths kriterium brukt på dette polynomet gir følgende talltabell

$$\begin{array}{r} T_1 T_2 \quad \quad \quad , \quad 1 \\ T_1 + T_2 \quad \quad \quad , \quad K \\ \frac{T_1 + T_2 - K T_1 T_2}{T_1 + T_2} \\ K \end{array}$$

Alle elementene i først kolonne av Rouths talltabell skal ha samme fortegn. Antar vi at både T_1 og T_2 er positive, blir betingelsen for stabilitet

$$0 < K < \frac{T_1 + T_2}{T_1 T_2} \quad (8.35)$$

Den øvre grensen i dette uttrykket angir verdien der polene i det lukkede systemet beveger seg over fra venstre til høyre halvplan som vist i figur 4.30.



8.3.4 Algebraiske kriterier ved transportforsinkelse i transferfunksjonen

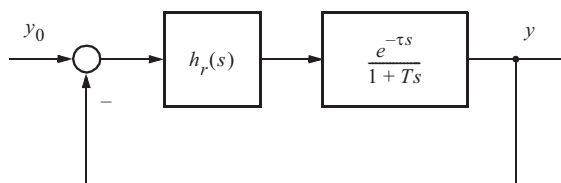
Mange fysiske prosesser inneholder fenomener som gjør at de får ledd av typen $e^{-\tau s}$ i sine transferfunksjoner. Årsaken til en slik “**transportforsinkelse**” kan være en veldefinert tidsforsinkelse i signalforplantningen slik tilfellet er ved materialtransport i rørledning eller på belter, lydtransmisjon i vann og liknende. Ofte har vi fysiske prosesser som med akseptabel tilnærming kan representeres ved transferfunksjonen

$$h(s) = \frac{e^{-\tau s}}{1 + T s} \quad (8.36)$$

Dersom en prosess av denne arten inngår i et reguleringsystem sammen med en regulator, som vist i figur 8.10, kan det være aktuelt å bestemme verdiene på regulatorparametrene som definerer grensen for systemets stabilitet.

Sprangrespons av tilbakekoblede systemer med transportforsinkelse er tidligere behandlet i avsnitt 4.7.7.

Figur 8.10
 Blokkdiagram for
 tilbakekoplet sys-
 tem med transport-
 forsinkelse



Transferfunksjonen i (8.36) er imidlertid ikke en rasjonal funksjon, og vi kan derfor ikke uten videre benytte oss av de algebraiske kriteriene som foreligger. En måte å omgå problemet på er å utvikle en rasjonal approksimasjon til funksjonen $e^{-\tau s}$. Vi kan gjøre dette på flere måter.

Den vanligste måten er som følger:

$$e^{-\tau s} = \frac{e^{-\frac{\tau}{2}s}}{e^{\frac{\tau}{2}s}} = \frac{1 - \frac{\tau}{2}s + \frac{1}{2!} \left(\frac{\tau}{2}s\right)^2 - \dots}{1 + \frac{\tau}{2}s + \frac{1}{2!} \left(\frac{\tau}{2}s\right)^2 + \dots} \quad (8.37)$$

$$\approx \frac{1 - \frac{\tau}{2}s + \frac{\tau^2}{8}s^2}{1 + \frac{\tau}{2}s + \frac{\tau^2}{8}s^2}$$

Vi har her avsluttet potensrekkeutviklingen av eksponentialfunksjonen med 2. ordens-leddet. Vi kunne ha avsluttet med 1. ordens-leddet og fått en dårligere approksimasjon slik det ble gjort i avsnitt 4.7.7, eller vi kunne ha tatt med flere ledd i rekkeutviklingen og fått en bedre approksimasjon. En annen approksimasjon som ofte har sine fordeler, er

$$e^{-\tau s} \approx \frac{\left(1 - \frac{\tau}{2n}s\right)^n}{\left(1 + \frac{\tau}{2n}s\right)^n} \quad (8.38)$$

Benytter vi $n = 2$ finner vi

$$e^{-\tau s} \approx \frac{1 - \frac{\tau}{2}s + \frac{\tau^2}{16}s^2}{1 + \frac{\tau}{2}s + \frac{\tau^2}{16}s^2} \quad (8.39)$$

Vi ser at (8.39) og (8.37) har mye til felles. Alle disse approksimasjonene gir eksakt overensstemmelse med *amplitudeforløpet* til $e^{-j\omega\tau} = 1$. Men *faseforløpet* kan de bare gjengi tilnærmet, dårligere jo høyere frekvens og jo færre ledd som tas med i rekkeutviklingen. Se eksempel 8.12. En sammenlikning av egenskapene til forskjellige rasjonale approksimasjoner til $e^{-\tau s}$ er også gitt i Balchen (1990).

Vi kan nå anvende slike approksimasjoner til bestemmelse av stabilitetsgrensen for noen eksempler.

EKSEMPEL 8.10: Stabilitetsgrensen ved proporsjonalregulering av en prosess med tidskonstant og transportforsinkelse

Vi betrakter først et system av en form som vist i figur 8.10, med proporsjonalregulator, dvs. $h_r(s) = K$. Vi benytter approksimasjon (8.37).

$$h_0(s) = K \frac{e^{-\tau s}}{1 + Ts} \approx K \frac{1 - \frac{\tau}{2}s + \frac{\tau^2}{8}s^2}{1 + \frac{\tau}{2}s + \frac{\tau^2}{8}s^2} \frac{1}{1 + Ts} \quad (8.40)$$

og

$$n_0(s) + t_0(s) = \frac{\tau^2}{8} T s^3 + \left(\frac{\tau}{2} T + \frac{\tau^2}{8} (1 + K) \right) s^2 + \left(T + \frac{\tau}{2} (1 - K) \right) s + 1 + K$$

Benytter vi Rouths kriterium på dette polynomet, finner vi følgende betingelser for stabilitet

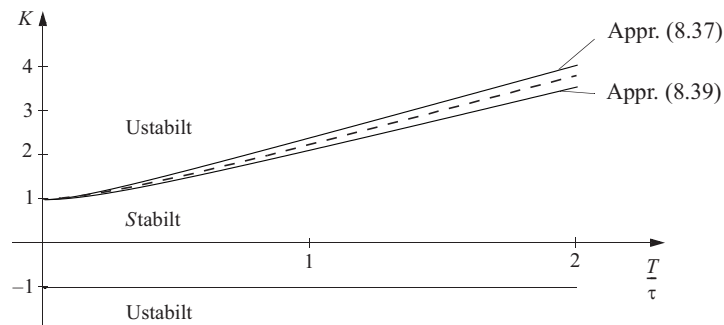
$$-1 < K < -2 \left(\frac{T}{\tau} \right) + \sqrt{12 \left(\frac{T}{\tau} \right)^2 + 4 \frac{T}{\tau} + 1} \quad (8.41)$$

Hadde vi benyttet approksimasjonen fra (8.39), ville vi fått følgende stabilitetsgrense

$$-1 < K < -4 \left(\frac{T}{\tau} \right) + \sqrt{32 \left(\frac{T}{\tau} \right)^2 + 8 \frac{T}{\tau} + 1} \quad (8.42)$$

Figur 8.11 viser stabilitetsgrensene for K som funksjon av T/τ for de to approksimasjonene. Det eksakte svaret ligger i mellom som vist stiplet.

Figur 8.11
Stabilitetsgrensene for K som funksjon av T/τ



EKSEMPEL 8.11: Stabilitetsgrensen ved integralregulering av en prosess med bare transportforsinkelse

Dersom prosessdelen i systemet i figur 8.10 bare hadde hatt transferfunksjonen $e^{-\tau s}$, ville det være god grunn til å velge en regulator med ren integralvirkning,

$$h_r(s) = \frac{1}{T_i s}$$

I så fall ville vi ha fått

$$h_0(s) = \frac{1}{T_i s} e^{-\tau s}$$

Vi kan nå bestemme stabilitetsgrensen for dette systemet ved å bruke approksimasjonene i (8.39) og (8.37),

$$n_0(s) + t_0(s) = \frac{\tau^2}{8} T_i s^3 + \left(\frac{\tau}{2} T_i + \frac{\tau^2}{8} \right) s^2 + \left(T_i - \frac{\tau}{2} \right) s + 1$$

og

$$n_0(s) + t_0(s) = \frac{\tau^2}{16} T_i s^3 + \left(\frac{\tau}{2} T_i + \frac{\tau^2}{16} \right) s^2 + \left(T_i - \frac{\tau}{2} \right) s + 1$$

Bruk av et av de to algebraiske kriteriene på disse to polynomene gir følgende resultat for stabilitetsgrensene:

- Approksimasjon (8.37):

$$0 < \frac{\tau}{T_i} < \frac{4}{1 + \sqrt{3}} \approx 1.46$$

- Approsimasjon (8.39):

$$0 < \frac{\tau}{T_i} < \frac{4}{1 + \sqrt{2}} \approx 1.66$$

Vi skal senere se at det er særlig enkelt å bestemme det eksakte svaret i dette enkle tilfellet. Det eksakte svaret er

$$0 < \frac{\tau}{T_i} < \frac{\pi}{2} \approx 1.57 \quad (8.43)$$

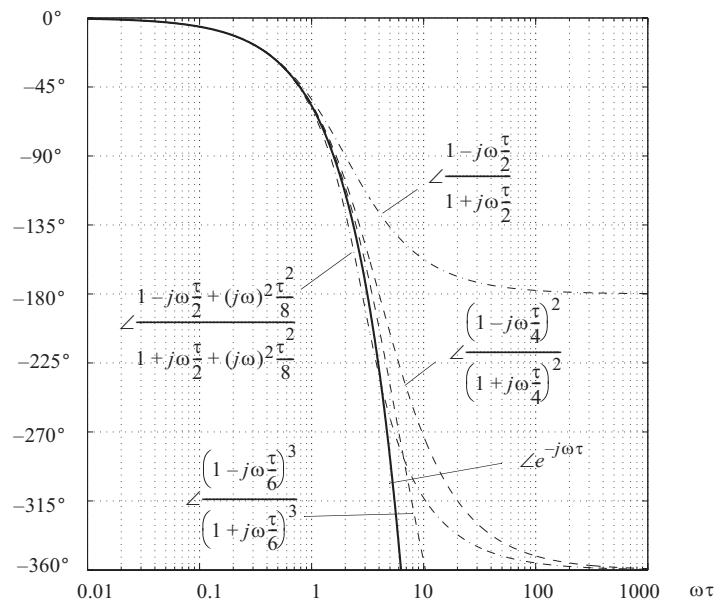
Dette viser at begge approksimasjonene ovenfor ga relativt gode resultater.



EKSEMPEL 8.12: Sammenlikning av $e^{-\tau s}$ og rasjonale approksimasjoner

En måte å sammenlikne de forskjellige approksimasjonene til $e^{-\tau s}$ på er å se på deres frekvensrespons. Tallverdien av frekvensresponsene til approksimasjonene vil være eksakt riktige, nemlig lik 1 for alle frekvenser, mens vinkelbidragene bare vil være tilnærmet. Figur 8.12 viser vinkelbidraget for frekvensresponsen til den virkelige funksjonen og en del approksimasjoner av typene (8.37) og (8.38).

Figur 8.12
Frekvensrespons
for eksponential-
funksjonen og
dens tilnærmelser



8.4 NYQUISTS STABILITETSKRITERIUM

8.4.1 Utleddning av Nyquists stabilitetskriterium

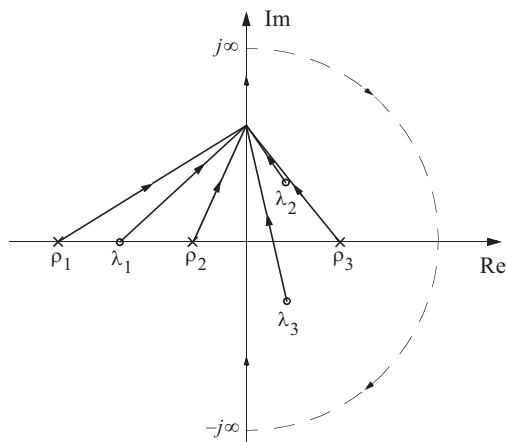
Et av de vesentligste bidrag til den regulerings-tekniske teori kom allerede i 1932 i Bell System Tech. Journal i et arbeid av H. Nyquist med tittelen "Regeneration Theory". Før vi starter gjennomgangen av Nyquists stabilitetskriterium presiseres det at dette kriteriet benyttes for å analysere inngangs- /utgangsstabilitet.

Nyquist tar utgangspunkt i formuleringer av den form som er gitt i (8.29) og (8.30) der det framgår at plasseringen av nullpunktene i uttrykket $1 + h_0(s)$ er avgjørende for stabilitet, siden disse blir poler i det lukkede systemet. Vi har

$$1 + h_0(s) = 1 + \frac{t_0(s)}{n_0(s)} = \frac{n_0(s) + t_0(s)}{n_0(s)} \quad (8.44)$$

der vi konstaterer at polene i det lukkede systemet framkommer som nullpunkter i $\frac{n_0(s) + t_0(s)}{n_0(s)}$, og polene til $1 + h_0(s)$ er identiske med polene til $h_0(s)$.

Figur 8.13
Poler og nullpunkters plassering i det komplekse plan



Dersom vi tegner inn alle nullpunktene og polene til $1 + h_0(s)$ i det komplekse plan, slik som antydning ved et eksempel i figur 8.13, kan vi finne en metode for å detektere hvor mange av nullpunktene til $1 + h_0(s)$ som ligger i høyre halvplan. Når vi betegner nullpunktene med λ_1 , λ_2 og λ_3 og polene med ρ_1 , ρ_2 og ρ_3 , kan vi trekke vektorer fra alle disse punktene til et vilkårlig punkt som angir den frie variabelen s .

Med de betegnelsene som er brukt i figur 8.13, finner vi da

$$1 + h_0(s) = \frac{n_0(s) + t_0(s)}{n_0(s)} = c \frac{(s - \lambda_1)(s - \lambda_2)(s - \lambda_3)}{(s - \rho_1)(s - \rho_2)(s - \rho_3)} \quad (8.45)$$

der c er en konstant. Vi merker oss at det alltid vil være like mange nullpunkter som poler i $1 + h_0(s)$.

I figur 8.13 har vi antatt at to av nullpunktene ligger i høyre halvplan hvilket betyr at det lukkede systemet er ustabil. Videre har vi antatt at en av polene til $1 + h_0(s)$ ligger i høyre halvplan. Dette betyr at det åpne systemet (uten tilbakekopling) er ustabil.

Tenker vi oss at den frie variabelen s gjennomløper alle verdier langs den imaginære akse fra $-j\infty$ til $+j\infty$ og langs en uendelig stor halvsirkel rundt hele høyre halvplan, som antydnet i figur 8.13, ser vi at de vektorene som kommer fra punkter i høyre halvplan, hver for seg vil gjennomløpe en hel omdreining mens de vektorene som kommer fra punkter i venstre halvplan, bare vil svinge opp og ned og ikke utføre noen hel omdreining.

Dette kan vi benytte til å finne det totale vinkelbidraget for $1 + h_0(s)$. Det blir summen av vinkelbidragene fra nullpunktene og polene i (8.45), når s gjennomløper alle verdier langs den imaginære akse fra $-j\infty$ til $+j\infty$ og langs en uendelig stor halvsirkel rundt hele høyre halvdel av det komplekse plan:

- Nullpunktene, λ , i høyre halvplan gir negative vinkelbidrag, fordi vektorene dreier en omdreining med urviseren (som er negativ vinkelretning).
- Polene, ρ , i høyre halvplan gir positive vinkelbidrag fordi de befinner seg i nevneren i uttrykket $1 + h_0(s)$, og beveger også seg en omdreining med urviseren.
- Polene og nullpunktene i venstre halvplan gir ikke noe vinkelbidrag siden disse vektorene beveger seg opp og ned og ender opp der bevegelsen startet.

Vektoren $1 + h_0(s)$ vil derfor få et totalt vinkelbidrag som er

$$\Delta\angle(1 + h_0(s)) = -2\pi(N_n - N_p) \quad (8.46)$$

der

$$\begin{aligned} N_n &= \text{antall poler i høyre halvplan for det lukkede system} \\ &= \text{antall røtter i h.h.p. av } n_0(s) + t_0(s) \\ &= \text{antall nullpunkter i h.h.p. av } 1 + h_0(s) \end{aligned}$$

$$\begin{aligned} N_p &= \text{antall poler i høyre halvplan for det åpne system} \\ &= \text{antall røtter i h.h.p. av } n_0(s) \\ &= \text{antall poler i h.h.p. av } 1 + h_0(s) \end{aligned}$$

Siden betingelsen for stabilitet av det lukkede systemet er at $N_n = 0$, kan (8.46) brukes til å formulere det **generelle Nyquists stabilitetskriterium**:

For at et lukket system skal være stabilt, må vektoren $1 + h_0(s)$ få en vinkeldreining lik $2\pi N_p$ når s gjennomløper en kontur rundt hele høyre halvplan med urviseren.

Et spesialtilfelle oppstår når det *åpne* systemet er stabilt, dvs. $N_p = 0$. Vi kan da formulere det **forenklede Nyquists stabilitetskriterium** som følger:

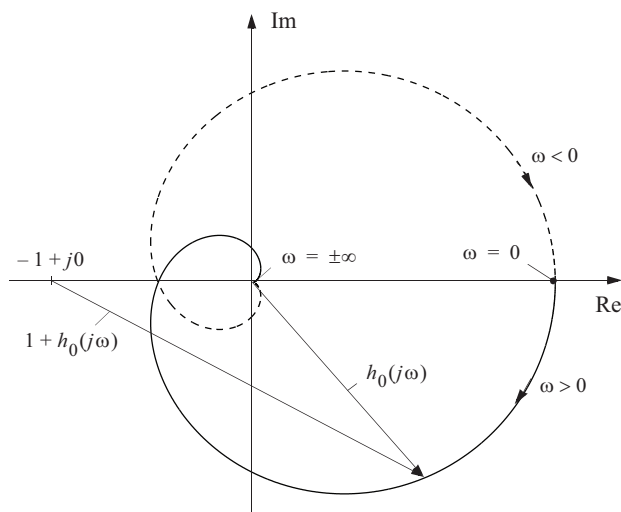
For at et system som er åpent stabilt, skal forbli stabilt når det lukkes, må vektoren $1 + h_0(s)$ få null vinkeldreining når s gjennomløper en kontur rundt hele høyre halvplan med urviseren.

Dette er det samme som å kreve at punktet $(-1, 0)$ ligger på *utsida* av stedkurven til h_0

Den viktigste fordelen med Nyquists stabilitetskriterium er at det er basert på grafisk tolkning. Det gir derfor muligheter for bestemmelse av stabilitetsforhold også i systemer beskrevet med eksperimentelt opptatte dynamiske karakteristikk (frekvensresponser) i stedet for analytiske uttrykk.

Nyquists stabilitetskriterium baserer seg på å la den frie variabelen s anta alle verdier langs den imaginære aksene og rundt hele det høyre halvplan. Fysisk realiserbare systemer er **strengt propre**; de har den egenskapen at $|h_0(s)| \rightarrow 0$ når $|s| \rightarrow \infty$. Dette medfører at vi bare vil være interessert i det som skjer når s antar verdier langs den imaginære aksene, dvs. $s = j\omega$. Vi kan altså se bort fra den reelle delen av s og utlede stabilitetsforholdene ved å betrakte frekvensresponser $h_0(j\omega)$. Vi kan også benytte oss av at vi alltid har $h_0(j\omega) = h_0^*(-j\omega)$, dvs. at den øvre delen av Nyquistkurven (den polare stedkurven) blir speilbildet av den nedre, som vi finner ved å la ω i $h_0(j\omega)$ variere fra 0 til ∞ .

Figur 8.14
Polar stedkurve for
 $h_0(j\omega)$



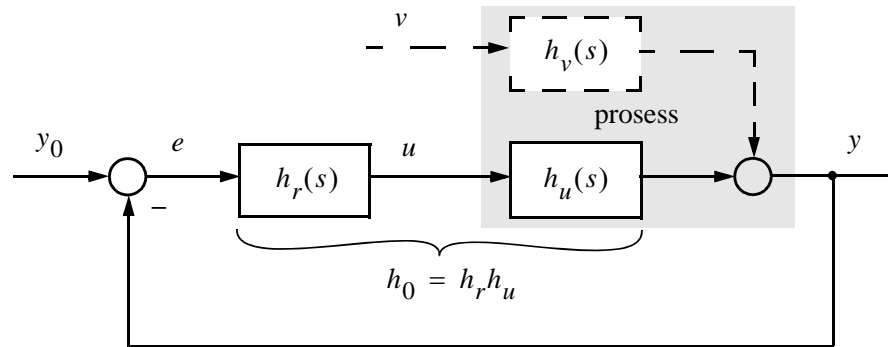
På basis av det som er beskrevet i avsnitt 7.2 (se for eksempel figur 7.10), har vi i figur 8.14 tegnet den polare stedkurven for $h_0(j\omega)$ for et monovariabelt, tilbakekoplet system med alle polene til $h_0(s)$ i venstre halvplan.

Vi legger spesielt merke til at $|h_0(j\omega)| \rightarrow 0$ når $|\omega| \rightarrow \infty$. Figuren viser også vektoren $1 + h_0(j\omega)$ fra punktet $(-1 + j0)$ til stedkurven. Lar vi nå s anta alle verdier langs den imaginære aksene ($s = j\omega$) og rundt høyre halvplan, vil vektorenes endepunkter bevege seg langs den epleformede stedkurven i retning med urviseren. Vektoren $h_0(j\omega)$ vil foreta to hele omdreininger med urviseren i dette tilfellet, mens vektoren $1 + h_0(j\omega)$ bare vil pendle opp og ned ca. 60° i hver retning uten å foreta noen hel omdreining.

8.4.2 Nyquist-prosedyre – oppsummert

Figur 8.15

Prosess



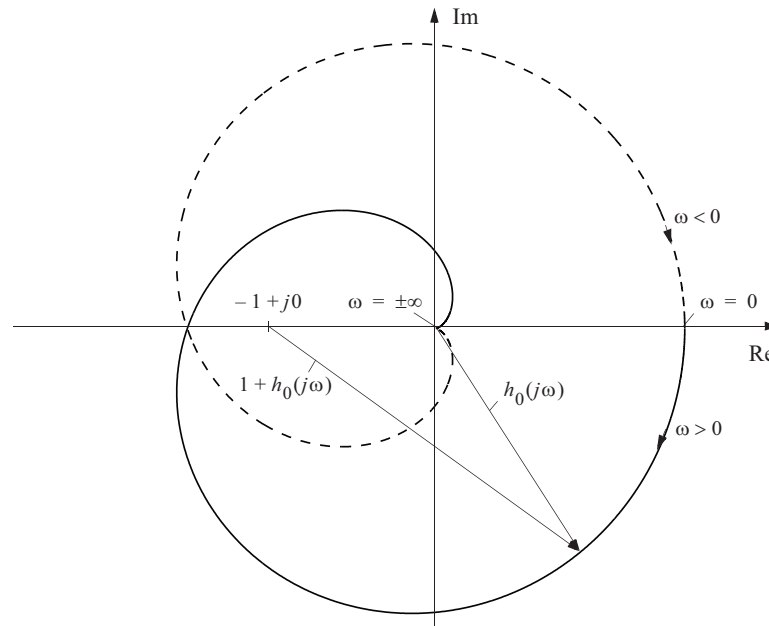
1. Gitt strukturen i figur 8.15. Øvre grein med forstyrrelse v er uten betydning ved stabilitetsanalysen.
2. Du må først finne antall poler N_p i h.h.p. (dvs. ustabile poler) for h_0 . Hvis $N_p > 0$, er h_0 åpent ustabil, ellers åpent stabil.
3. Plott polardiagrammet for h_0 med valgt regulator og forsterkning.
4. Tell hvor mange ganger vektoren $1 + h_0$ roterer (netto) når vektorens endepunkt gjennomløper grafen. Vinkelen $\Delta\angle(1 + h_0)$ blir 2π ganger dette heltallet (som kan være negativt, avhengig av dreieretning).
5. Antall poler i h.h.p. for det lukkede system, N_n , blir da: $N_n = N_p - \frac{\Delta\angle(1+h_0)}{2\pi}$. Hvis $N_n > 0$, er det lukkede systemet ustabil.
Sagt på en annen måte: Systemet er stabilt hvis $\Delta\angle(1 + h_0) = 2\pi N_p$.

8.4.3 Nyquists kriterium for åpent stabilt system

Siden systemet i figur 8.14 var forutsatt å ha sine poler av $h_0(s)$ i venstre halvplan, vil kravet til stabilitet for det lukkede systemet være at $\Delta\angle(1 + h_0(j\omega)) = 0$ slik som tilfellet er i figur 8.14. Dette tilbakekoblede systemet er derfor stabilt.

Figur 8.16 viser en tilsvarende stedkurve for et system med litt andre parametre enn det som er vist i figur 8.14.

Figur 8.16
Stedkurve for
 $h_0(j\omega)$



I dette tilfellet vil vektoren $1 + h_0(s)$ oppføre seg annerledes når s gjennomløper alle verdier langs den imaginære aksene og rundt høyre halvplan. Vi finner at $\Delta\angle(1 + h_0(j\omega)) = -4\pi$, siden vektoren $1 + h_0(j\omega)$ foretar to hele omdreininger med urviseren, når s gjennomløper alle verdier langs den imaginære aksene. Dersom systemet som forutsatt har alle poler av

$h_0(s)$ i venstre halvplan, det vil si $N_p = 0$, medfører dette ifølge (8.46) at $N_n = 2$. Dette betyr at $1 + h_0(s)$ har to nullpunkter i høyre halvplan. Det lukkede systemet er derfor *ustabilt* i følge Nyquists forenklete kriterium.

Dette prinsippet skal illustreres med noen eksempler.

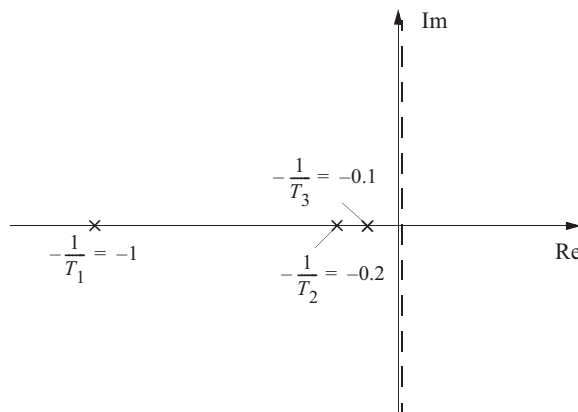
EKSEMPEL 8.13: Åpent stabilt system med poler i venstre halvplan

Vi undersøker først det samme systemet som er brukt i eksempel 8.5 og eksempel 8.8:

$$h_0(s) = \frac{K}{(1 + T_1 s)(1 + T_2 s)(1 + T_3 s)} \quad (8.47)$$

Velger vi $T_1 = 1$, $T_2 = 5$ og $T_3 = 10$ vil polene til $h_0(j\omega)$ få en plassering som vist i figur 8.17. Figur 8.18 viser stedkurvene til $h_0(j\omega)$ når $K = 10$, $K = 19.8$ og $K = 30$.

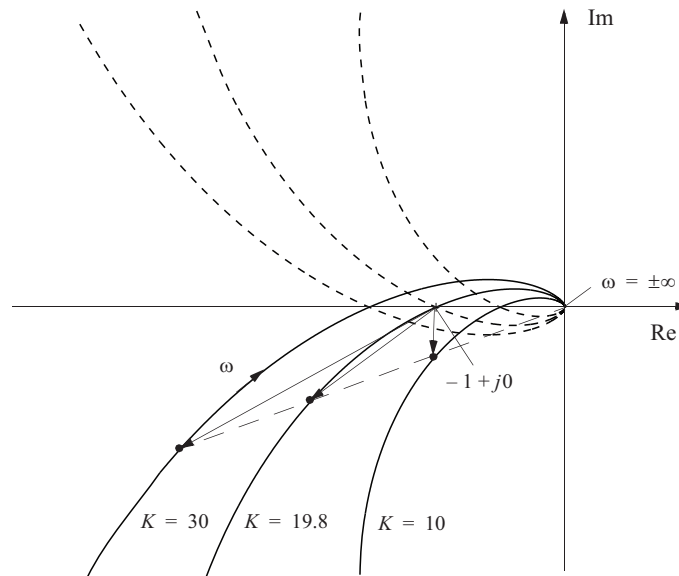
Figur 8.17
Poler og nullpunkter til $h_0(j\omega)$ i det komplekse plan



Figur 8.18 viser stedkurvene heltrukket for positive ω . Stedkurvene for negative ω er stiplet. Vektoren $1 + h_0(j\omega)$ er antydnet for en bestemt frekvens for hver av de tre verdiene av K . Når $K = 10$, ser vi at systemet vil være stabilt. Dette er fordi systemet er “åpent stabilt”, og fordi vektoren $1 + h_0(s)$ ikke får noen netto vinkelendring når s gjennomløper alle verdier rundt høyre halvplan. For $K = 30$ vil derimot vektoren $1 + h_0(j\omega)$ få et netto vinkelbidrag på -4π , hvilket betyr at systemet er ustabilt for denne verdien av K . Den kritiske verdien av K er 19.8 som er den verdien der stedkurven til $h_0(j\omega)$ akkurat går gjennom punktet $(-1 + j0)$. Dette resultatet kan kontrolleres mot (8.22) som gir samme svar.

Figur 8.18

Stedkurve for
 $h_0(j\omega)$ når
 $K = 10$,
 $K = 19.8$ og
 $K = 30$



■

EKSEMPEL 8.14: Åpent stabilt system med pol på den imaginære akse

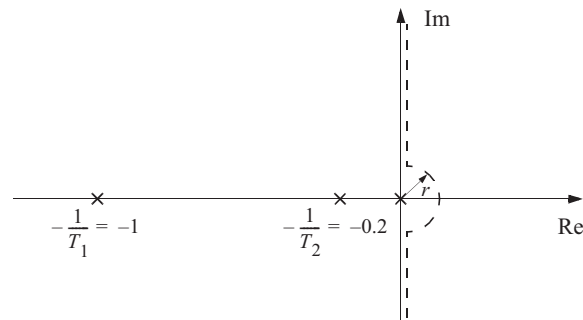
Vi undersøker nå det samme systemet som i eksempel 4.13 og eksempel 8.9. Vi har

$$h_0(s) = \frac{K}{s(1 + T_1 s)(1 + T_2 s)} \quad (8.48)$$

Figur 8.19 viser plasseringen av polene til $h_0(s)$ for det tilfellet at $T_1 = 1$ og $T_2 = 5$.

Figur 8.19

Polene og nullpunktene til $h_0(s)$ i det komplekse plan



Her har vi det spesielle tilfellet at en av polene ligger på den imaginære aksene. Siden den imaginære aksene utgjør randen av venstre halvplan, lar vi s anta alle verdier langs den

imaginære aksen bortsett fra i nærheten av origo der vi lar den følge en liten sirkel med radius r inn i høyre halvplan. Lar vi $r \rightarrow 0$, vil dette medføre at stedkurven for $h_0(j\omega)$ vil gjennomløpe en tilsvarende uendelig stor halvsirkel rundt høyre halvplan, se eksempel 6.5. Vi ønsker først å finne realverdien av vektoren $h_0(j\omega)$ når $\omega \rightarrow 0$.

$$\begin{aligned} h_0(j\omega) &= \frac{K}{j\omega(1+T_1j\omega)(1+T_2j\omega)} \\ &= \frac{K(1-T_1j\omega)(1-T_2j\omega)}{j\omega(1+T_1j\omega)(1+T_2j\omega)(1-T_1j\omega)(1-T_2j\omega)} \\ &= \frac{K(1-(T_1+T_2)j\omega-T_1T_2\omega^2)}{j\omega(1+(T_1\omega)^2)(1+(T_2\omega)^2)} \end{aligned}$$

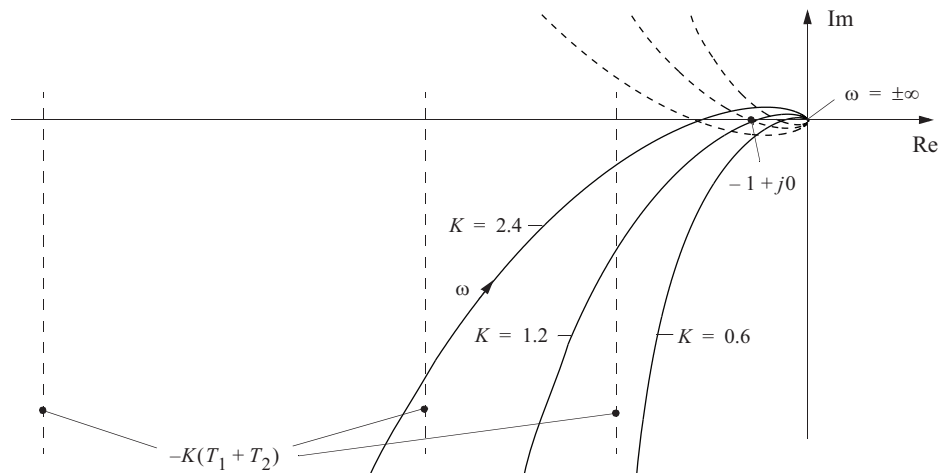
Vi dividerer med $j\omega$ i teller og nevner og ser at det første og siste leddet i telleren blir imaginært. Lar vi så $\omega \rightarrow 0$, blir resultatet som følger

$$\operatorname{Re} h_0(j\omega) \underset{\omega \rightarrow 0}{=} -K(T_1+T_2) \quad (8.49)$$

Dette medfører stedkurver av $h_0(j\omega)$ for positive ω som antydnet i figur 8.20.

Figur 8.20

Stedkurve for $h_0(j\omega)$ når $K = 2.4$, $K = 1.2$ og $K = 0.6$



Herav er det lett å utlede at systemet vil være stabilt for $K = 0.6$, ustabilt for $K = 2.4$ og akkurat på grensen til stabilitet for $K = 1.2$. Likning (8.35) i eksempel 8.9 viser at systemet er stabilt for

$$0 < K < \frac{T_1+T_2}{T_1T_2}$$

som gir en stabilitetsgrense på $K = 1.2$ for dette systemet. Den nedre grensen i denne ulikheten kunne vi også ha bestemt ved hjelp av Nyquists kriterium, fordi stedkurven med negativ K ville omslutte punktet $(-1 + j0)$ uansett verdiene av T_1 og T_2 .

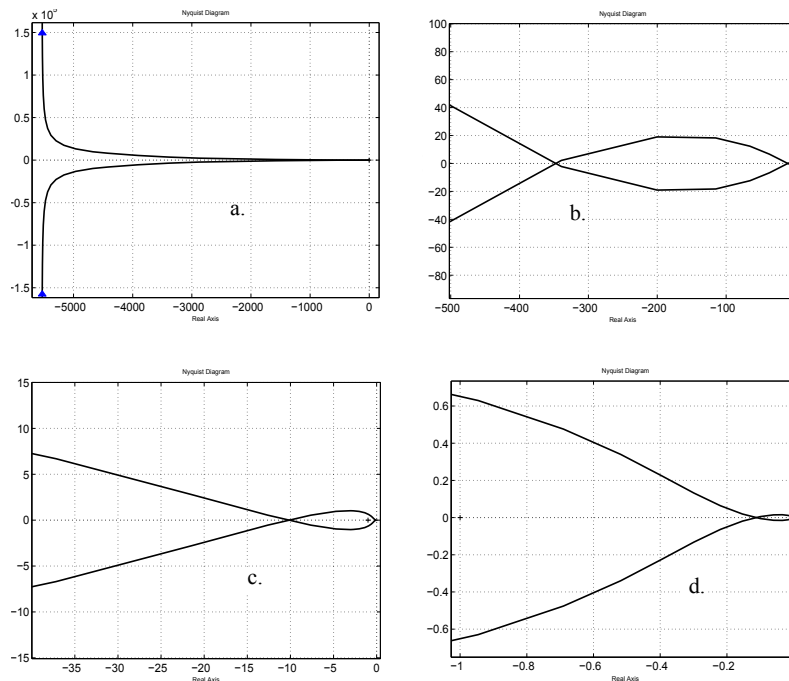


8.4.4 Det logaritmiske Nyquist-diagram

Vi skal her presentere et alternativt Nyquist-diagram. I dette er amplituden på radiusvektor i logaritmisk skala [dB], mens vinkelen på vektor er den samme som i det ordinære diagrammet. Dette *logaritmiske Nyquistdiagrammet* er spesielt fordelaktig når $h_0(s)$ har poler på den imaginære akse. Dette forklares best gjennom et eksempel. Figur 8.21 viser det ordinære Nyquistdiagrammet for transferfunksjonen

$$h_0(s) = \frac{K_p(1+3s)(1+2s)}{s(1+50s)(1+10s)(1+0.5s)(1+0.1s)}, \text{ med } K_p = 200 \quad (8.50)$$

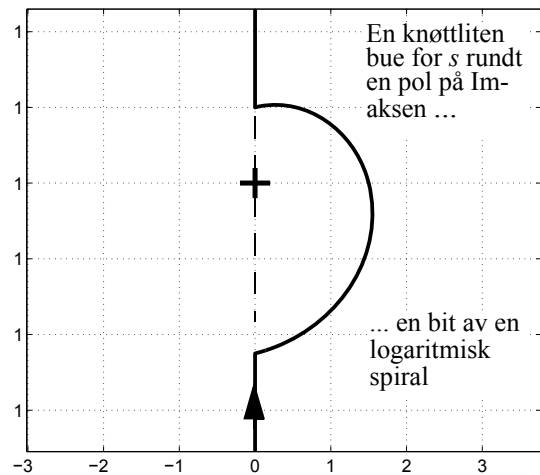
Figur 8.21



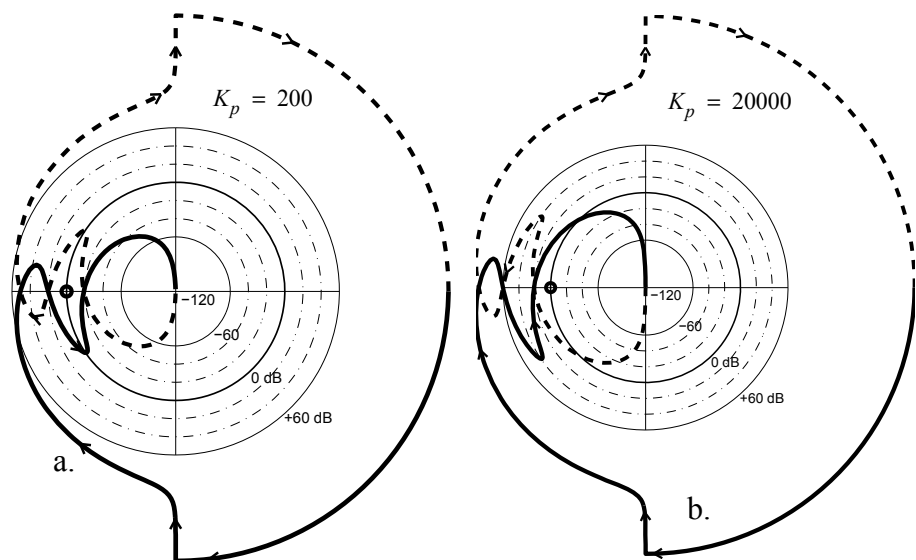
Dette systemet er åpent stabilt; det er ingen poler i høyre halvplan, vi har $N_p = 0$. Nyquists kriterium forteller oss da at grafen skal gå på innsida av det kritiske punkt -1 . Systemet har en pol i origo, og ellers tidskonstanter med stor spennvidde (den siste egenskapen gjør at slike systemer ofte kalles "stive", og det kan noen ganger være krevende å simulere dem numerisk). MATLAB-funksjonen `nyquist(h_0)` gir resultatet som er vist i figur 8.21. Det er den samme grafen, vist i fire forskjellige skalaer, hvor vi zoomer inn etter hvert. Oppe til venstre (delfigur a) må vi starte med svært stor verdier langs aksene, på grunn av polen i

origo. Etter hvert som vi zoomer inn, kommer det nye løkker fram på grafene, og disse skyldes at tidskonstantene gjør seg gjeldende etter tur. Totalbildet er likevel uklart: Er det lukkede system stabilt; går grafen på innsida av punktet -1 ? Ved å zoome inn og ut, og ta notater, kan man finne ut at det er stabilt med $K_p = 200$. Men stabiliteten er *betinget*: Ved å starte med en tilstrekkelig liten K_p , og prøve økende verdier, oppdager man at det lukkede system går gjennom stadiene stabilt, ustabil, stabilt, ustabil. Men dette er altså ikke lett å finne ut ved hjelp av et ordinært Nyquist-diagram.

Figur 8.22



Figur 8.23



Vi betrakter nå i stedet figur 8.23: her plotter vi radiusvektor ikke som $|h_0(j\omega)|$, men som $20\log_{10}|h_0(j\omega)|$, altså i [dB]. Alle verdier mindre enn 10^{-6} blir satt til $10^{-6} = -120$ dB,

som er plassert i sentrum i diagrammet. Fasevinkelen $\angle h_0(s)$ er den samme som i det vanlige Nyquistdiagrammet. Banen til s nær polen til $h_0(s)$ velges som en bit av en meget liten logaritmisk spiral inn i høyre halvplan, indikert på figur 8.22. Dette fører til en stor korresponderende bue¹ inn i h.h.p. for $h_0(j\omega)$, som vist i figur 8.23.

Vi ser nå at vi kan ta ut all informasjon for å sjekke stabiliteten av det lukkede system fra figur 8.23 a. Zooming ut og inn, med inspeksjon av grafer i forskjellig skala, er unødvendig. Vi ser at det lukkede systemet er stabilt. Og forsterknings- og fasemarginer (definert seinere i dette kapitlet, se avsnitt 8.4.9) kan hentes ut av figuren, bortsett fra at forsterkningsmarginen vil bli i dB. Figur 8.23 b viser dette **nyqlog**-diagrammet for 40 dB høyere forsterkning. Det lukkede system er nå ustabil.

MATLAB-funksjonen **nyqlog**(.) kan finnes og lastes ned via *Mathworks file exchange*. Ytterligere dokumentasjon finner man i Andresen, T. "A Logarithmic-amplitude Polar Diagram", *Modeling, Identification and Control*, vol. 22 no. 2, 2001.

8.4.5 Nyquists kriterium for åpent ustabil system

Et åpent ustabil system har poler i høyre halvplan, $N_p \neq 0$. For at det lukkede systemet da skal bli stabilt, må $\Delta\angle(1 + h_0(s)) = 2\pi N_p$.

Dette innebærer at Nyquistkurven *skal omslutte* det kritiske punkt N_p ganger i positiv vinkelretning når ω går fra $-\infty$ til $+\infty$. Å øke regulatorforsterkningen kan noen ganger være tilstrekkelig for å oppnå dette, som vi skal se i et eksempel:

EKSEMPEL 8.15: Åpent ustabil system – kjemisk reaktor

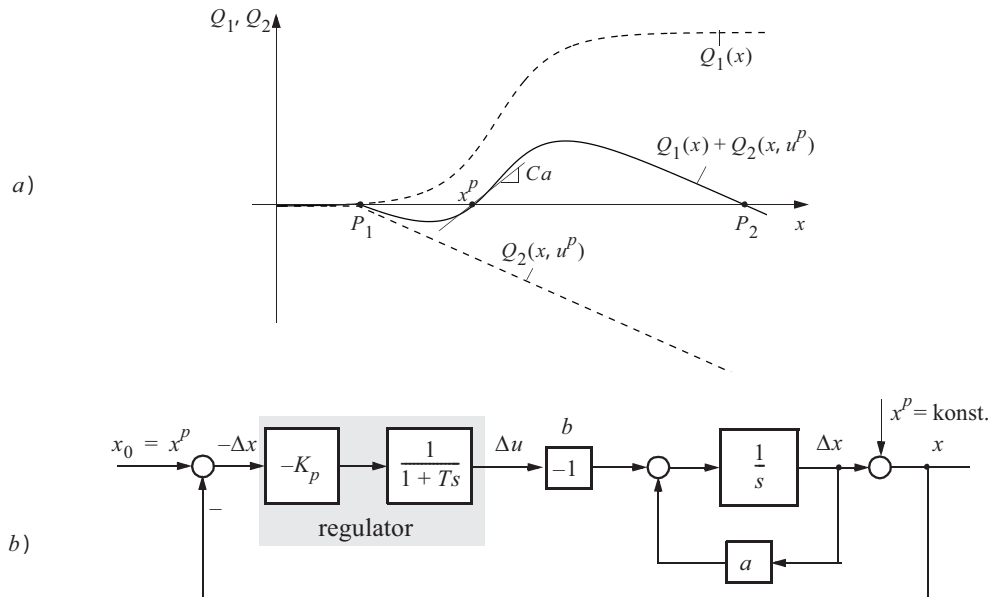
Enkelte prosesser som skal styres med automatisk regulering, vil under visse betingelser kunne være åpent ustabile. Dette er for eksempel tilfellet med en eksoterm reaksjon der den utviklede varmen i et visst temperaturområde øker sterkere med økende temperatur enn den bortførte varmen ved naturlig kjøling, slik tilfellet var for systemet fra eksempel 3.6 – som vi nå skal studere stabilitetsegenskapene til. I figur 8.24 a) er vist kjølekarakteristikken fra reaktoren. Vi baserer oss på den lineariserte modellen av systemet for små endringer rundt et arbeidspunkt, x^p , u^p .

Dersom temperaturen er noe lavere enn x^p , vil det være et overskudd av kjøling med det resultat at temperaturen må synke inntil likevekt oppnås i det punktet som er antydnet med P_1 . Dersom temperaturen i reaktoren er noe høyere enn x^p , er det et overskudd av reaksjonsvarme med det resultat at temperaturen i reaktoren må øke inntil likevekt oppstår ved temperaturen som er antydnet ved punktet P_2 . Ønsker vi at reaktoren skal holdes på en temperatur mellom de to stabile yttergrensene, må vi anordne en forsert og automatisk regulert kjøling slik at den resulterende kjølekarakteristikken går brattere enn reaktorkarakteristikken.

¹Denne buen blir en bit av en såkalt *Arkimedes*-spiral. Poenget med ikke å benytte små halvsirkler for s her, er å unngå at man får buer for $h_0(j\omega)$ som faller oppå hverandre. Da blir det vanskelig å se antall omdreininger for vektoren $1 + h_0(j\omega)$.

Figur 8.24 b) viser det lineariserte systemet¹ med regulator. Den lineariserte prosess har en positiv tilbakekopling a rundt en integrator. Vinkelkoeffisienten til funksjonen $Q_1 + Q_2$ i punktet x^p i figur 8.24 a), er Ca . For enkelhets skyld antar vi varmekapasiteten $C = 1$, og koeffisienten b fra (3.89). Regulatoren er en proporsjonalregulator $-K_p$. Transferfunksjonen $1/(1 + Ts)$ kommer (uønsket) i tillegg til i regulatoren fordi det er en viss dynamikk (treghet) forbundet med endringen av kjøleeffekten.

Figur 8.24
a) Kjølekaraktéristikk for reaktor
b) Blokkdiagram for linearisert reaktor med regulering



¹Den lineariserte strukturen basert på små utslag rundt et arbeidspunkt, bruker vi for å analysere egenskapene til den valgte regulatorløsning. Men vi implementerer reguleringsstrukturen som vist i figur 3.16 – vi setter på det totale pådrag u , ikke Δu

Det negative fortegnet i regulatoren er nødvendig fordi b er < 0 . Sagt på en annen måte: Hvis reaktoren blir for kald ($-\Delta x > 0$), må kjølinga *reduseres*, dvs. $\Delta u < 0$.

Det åpne systemets transferfunksjon blir

$$h_0(s) = \frac{K_p}{(1 + Ts)(s - a)} \quad (8.51)$$

Betingelser for stabilitet av dette lineariserte systemet kan finnes ved bruk av Rouths kriterium

$$n_0(s) + t_0(s) = Ts^2 + (1 - aT)s + K_p - a \quad (8.52)$$

Herav finner vi umiddelbart betingelsene for stabilitet:

$$T < \frac{1}{a} \quad , \quad K_p > a$$

Det er med andre ord en øvre grense for tidskonstanten T til oppbygging av kjøleeffekten som kan tolereres, og forsterkningen i regulatoren K_p må være større enn en viss verdi.

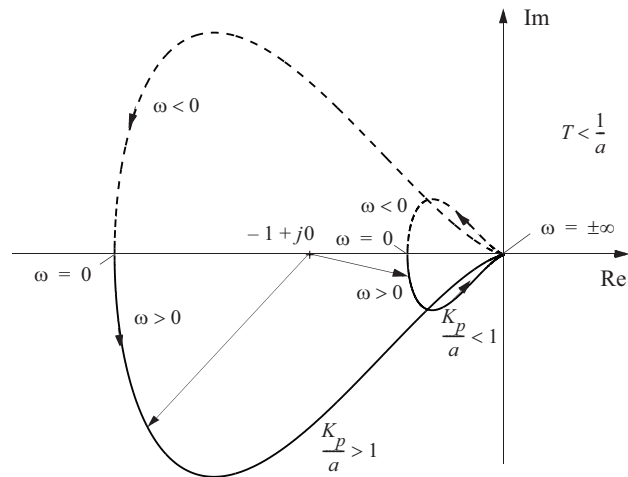
Vi skal nå benytte Nyquists kriterium til å bestemme de samme grenseverdiene. Figur 8.25 viser polare stedkurver for $h_0(j\omega)$ når $T < 1/a$ og for de to tilfellene $K_p/a > 1$ og $K_p/a < 1$. Siden $h_0(s)$ for $a > 0$ vil ha 1 pol i høyre halvplan, vil betingelsen for stabilitet av det lukkede systemet ifølge (8.46) være at

$$\Delta\angle(1 + h_0(s)) = 2\pi N_p = 2\pi \quad (8.53)$$

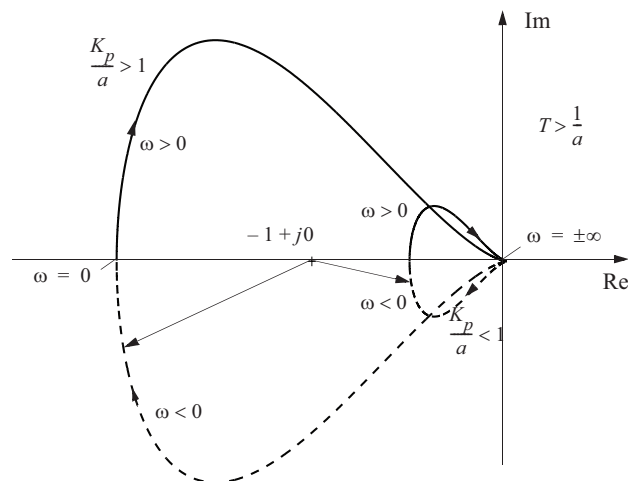
Trekker vi vektorer fra punktet $(-1 + j0)$ til stedkurvene i figur 8.25, ser vi at dette kravet vil bli oppfylt når $K_p/a > 1$. Vi har altså et tilfelle der stedkurven faktisk skal omslutte punktet $(-1 + j0)$ for at det lukkede systemet skal bli stabilt.

Figur 8.25

Stedkurve for $h_0(j\omega)$ når $T < 1/a$ og for de to tilfellene $K_p/a > 1$ og $K_p/a < 1$

**Figur 8.26**

Stedkurve for $h_0(j\omega)$ når $T > 1/a$ og for de to tilfellene $K_p/a > 1$ og $K_p/a < 1$



Dersom $T > 1/a$ ville stedkurvene for systemet med de tilsvarende betingelsene bli som vist i figur 8.26. Rotasjonsretningen for vektoren $1 + h_0(j\omega)$ når $K_p/a > 1$ vil nå være med urviseren. Dette medfører at vinkelbidraget blir $\Delta\angle(1 + h_0(s)) = -2\pi$, hvilket ikke tilfredsstillers (8.46). For $T > 1/a$ er det derfor ikke mulig å finne noen verdi av K_p som gjør systemet stabilt.

De refleksjoner som i dette tilfellet måtte gjøres for å bestemme stabilitetsbetingelsene ved hjelp av Nyquists kriterium, er mer omfattende enn ved bruk av Rouths kriterium. Hadde systemet hatt en mer komplisert dynamikk, ville det ikke nødvendigvis vært slik.



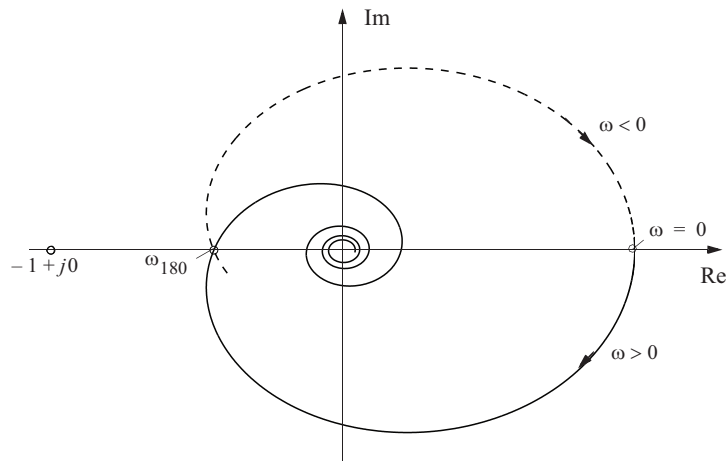
8.4.6 Nyquists stabilitetskriterium ved ikke-rasjonale funksjoner

Nyquists kriterium har, fordi det er grafisk basert, en vesentlig fordel framfor de algebraiske kriteriene ved undersøkelse av stabilitet for systemer som inneholder ikke-rasjonale funksjoner. En prosess i et reguleringssystem som inneholder en transportforsinkelse, er et slikt tilfelle, f.eks.

$$h_0(s) = \frac{Ke^{-\tau s}}{1 + Ts} \quad (8.54)$$

Det kan formelt påvises at Nyquists stabilitetskriterium også gjelder for slike tilfeller. Stedkurven til $h_0(j\omega)$ er vist i figur 8.27. Siden det åpne systemet er stabilt, må vi forlange at punktet $(-1 + j0)$ blir liggende utenfor stedkurven. Vi bruker gjerne betegnelsen ω_{180} på verdien av ω som gjør at $\angle h_0(j\omega) = -180^\circ$. Kravet til stabilitet blir da at $|h_0(j\omega_{180})| < 1$ og $h_0(0) = K > -1$.

Figur 8.27
Stedkurve for
 $h_0(j\omega)$



Også for systemer med høyt ordenstall vil Nyquists kriterium ofte være enklere i bruk enn de algebraiske kriteriene.

8.4.7 Bruk av amplitude/fase-diagram til stabilitetsundersøkelse

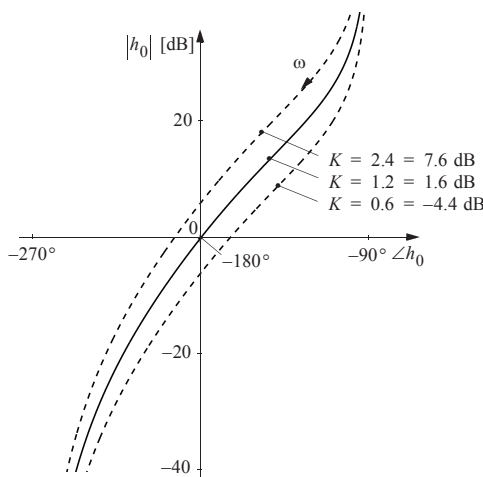
I enkelte tilfeller er det mer hensiktsmessig å undersøke stabilitetsforholdene ved framstilling av stedkurver i amplitude/fase-diagrammer i stedet for i Nyquist(polar-)diagrammer, i det minste for åpent stabile systemer. *AF-diagrammet svarer til Nichols-diagrammet uten det krumme rutenettet.* Det “kritiske punktet” $(-1 + j0)$ som stedkurven må gå på innsiden av i Nyquistdiagrammet, er i AF-diagrammet punktet $(0 \text{ [dB]}, -180^\circ)$. Kravet om å gå på innsiden av det kritiske punkt, svarer i AF-diagrammet til at $h_0(j\omega)$ må passere til venstre for $(0 \text{ [dB]}, -180^\circ)$ med økende frekvens. Se figur 8.28 nedenfor. Og jo lenger vekk $h_0(j\omega)$ er fra $(0 \text{ [dB]}, -180^\circ)$, jo mer stabilt er systemet. Nå har vi fått forklaringa på hvorfor det er så viktig å få $h_0(j\omega)$ til å “smyge seg utenom” den uendelig høye “fjelltoppen” til $N(j\omega)$ i punktet $(0 \text{ [dB]}, -180^\circ)$, som vi snakket om i forbindelse med Nicholsdiagrammet i eksempel 7.4. Videre innser vi nå at nærhet til stabilitetsgrensen innebærer en høy resonanstopp i $|N(j\omega)|$ og $|M(j\omega)|$. Hvis vi velger en regulator som gir en lav resonanstopp, vet vi med andre ord at systemet blir stabilt med god margin.

EKSEMPEL 8.16: Stabilitetsundersøkelse av åpent stabilt system i AF-diagram

Vi skal illustrere bruk av et amplitude/fase-diagram til stabilitetsundersøkelse, ved igjen å ta utgangspunkt i likestrømsmotoren i eksempel 4.13. Dette systemet har stedkurver som vist i figur 8.20. Disse er nå overført til amplitude/fase-diagrammet i figur 8.28.

Dersom verdien av K endres fra $K = 0.6$ til $K = 1.2$ og $K = 2.4$, vil stedkurven beholde sin form, men forskyves oppover. Når $K = 1.2$ går stedkurven akkurat gjennom punktet $(0 \text{ [dB]}, -180^\circ)$, og denne situasjon representerer grensen mellom et stabilt og et ustabilt system. Stabilitetsbetingelsen blir etter dette $K < 1.2$. Tidligere har vi funnet at vi også må forlange at $K > 0$. Denne betingelsen er det ikke fullt så lett å se i et AF-diagram.

Figur 8.28
Stedkurve for $h_0(j\omega)$ i amplitude/fase-diagram



I figur 8.28 har vi tegnet tre stedkurver som har samme form, for å finne den verdien av K som gir stabilitetsgrensen. Alternativt kunne vi hatt *en* kurve og forskjøvet 0-dB-linjen (aksekorset) i forhold til den.

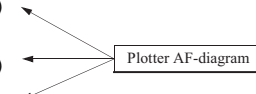
AF-diagrammet i figur 8.28 kan genereres ved bruk av MATLAB-koden i figur 8.29. Figur 8.28 bør sammenliknes med figur 8.20 og likning (8.35). Dette skulle kunne gi en forståelse for sammenhengen mellom tre ulike metoder for stabilitetsundersøkelser av et regulert system.

Figur 8.29
MATLAB-kode for
generering av AF-
diagram

```

1 K1 = 0.6; K2 = 1.2; K3 = 2.4;
2 T1 = 1; T2 = 5;
3 t1 = K1; t2 = K2; t3 = K3;
4 n0 = [T1*T2 (T1 + T2) 1 0];
5 w = logspace(-3, 3, 2000);
6
7 nichols(t1,n0,w)
8 hold on
9 nichols(t2,n0,w)
10 hold on
11 nichols(t3,n0,w), axis([-270 -90 -40 40])
12 hold off

```



■

EKSEMPEL 8.17: Stabilitetsundersøkelse av “betinget stabilt” system i AF-diagram

Gitt systemet

$$h_0(s) = \frac{K(1 + T_2s)(1 + T_3s)}{s^2(1 + T_1s)(1 + T_4s)(1 + T_5s)} \quad (8.55)$$

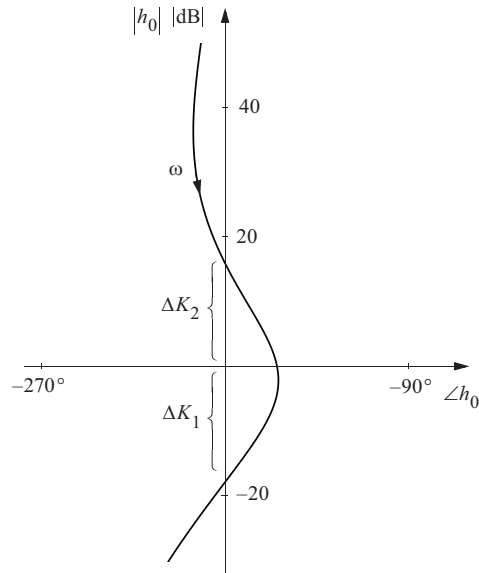
der $T_1 > T_2 > T_3 > T_4 > T_5 > 0$. Systemet er har ingen poler i høyre halvplan. Vi legger spesielt merke til at $h_0(s)$ har en dobbel pol i origo. Stedkurven for $h_0(j\omega)$ inntegnet i et amplitude/fase-diagram, vil for en bestemt verdi av K kunne bli som i figur 8.30.

For det viste tilfellet ser vi at det kritiske punktet (0 [dB], -180°) ligger til venstre for stedkurven $h_0(s)$. Et tilbakekoplet system med denne formen vil derfor være stabilt. Økes derimot forsterkningsfaktoren K med en størrelse ΔK_1 [dB], vil systemet komme til stabilitetsgrensen.

Dette tilsvarende den situasjonen som er vist i figur 8.28. I dette tilfellet ser vi imidlertid at systemet også vil komme til en stabilitetsgrense dersom forsterkningsfaktoren K reduseres, med ΔK_2 [dB]. Dette er noe nytt. Systemet er stabilt dersom

$$K_2 < K < K_1$$

Figur 8.30
Stedkurve for $h_0(j\omega)$ i AF-diagram



Vi har

$$K_1 \text{ [dB]} - K_2 \text{ [dB]} = \Delta K_1 \text{ [dB]} + \Delta K_2 \text{ [dB]}$$

eller i lineære enheter

$$\frac{K_1}{K_2} = \Delta K_1 \Delta K_2$$

Et system av denne arten kalles ofte **betinget stabilt**. Årsaken til navnet er at det også kan bli ustabil med *avtagende* positiv forsterkning i tilbakekoplingsløyfen. Dette kan i mange tilfeller være et problem dersom vi har å gjøre med komponenter som mister forsterkning på grunn av slitasje, elding, temperatur eller liknende. Samme fenomen oppstår også når forsterkningen i sløyfen avtar på grunn av ulineære effekter som “metning” og “dødsone”. Dette problemet behandles i kapittel 12.



8.4.8 Bode-Nyquist-kriteriet

Dersom det åpne systemet er stabilt, det vil si at $h_0(s)$ ikke har poler i høyre halvplan, kan vi med fordel bruke Bodediagram til undersøkelse av stabilitetsforhold i stedet for polardigram. Vi introduserer metoden med et eksempel.

EKSEMPEL 8.18: Bode-Nyquist-kriteriet på regulering av likestrømsmotor

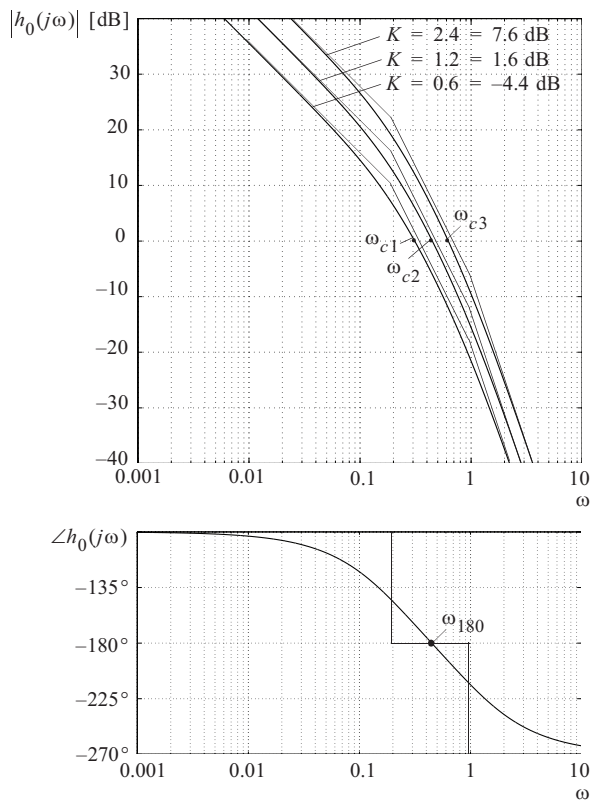
Vi benytter likestrømsmotor-eksempelet som er gjennomgått tidligere i eksemplene 8.14 og 8.16. Amplitude/fase/frekvens-diagrammer for dette systemet er vist i figur 8.31 for de tre verdiene

$$K = 0.6 = -4.4 \text{ [dB]}$$

$$K = 1.2 = 1.6 \text{ [dB]}$$

$$K = 2.4 = 7.6 \text{ [dB]}$$

Figur 8.31
Amplitude/fase/
frekvens-diagram
for systemet



Siden det bare er forsterkningen K som er endret, vil kurven for fasevinkelen $\angle h_0(j\omega)$ være den samme i alle de tre tilfellene. Vi bemerker at de kurvene som er vist i figur 8.31 inneholder eksakt samme informasjon som de som er vist i figur 8.20. For de tre forskjellige tilfellene får vi som avmerket tre forskjellige kryssfrekvenser ω_{c1} , ω_{c2} og ω_{c3} . Vi merker oss også den frekvensen der $\angle h_0(j\omega) = -180^\circ$. Denne betegnes som før ω_{180} .

Anvendelse av Nyquists kriterium på det foreliggende tilfellet vil gi følgende betingelser for stabilitet:

$$\angle h_0(j\omega_c) > -180^\circ \text{ og } |h_0(j\omega_{180})| < 1 = 0 \text{ [dB]} \quad (8.56)$$

Denne anvendelsen av Nyquists stabilitetskriterium er meget hensiktsmessig i praktisk bruk og vil bli ytterligere avklart ved bruk av to nye begreper som beskrives i neste avsnitt.



8.4.9 Fasemargin. Forsterkningsmargin

Den polare stedkurven for et vilkårlig åpent stabilt system, og tilsvarende amplitude/fase-diagram og amplitude/fase/frekvens-diagram er vist i figur 8.32.

Figur 8.32 a) viser en sirkel i det komplekse plan med sentrum i origo og radius lik 1. Den aktuelle stedkurven skjærer denne sirkelen i punktet P og den negative reelle aksens i punktet Q .

Strålen fra origo gjennom punktet P danner vinkelen ψ med den negative reelle aksens. Punktet Q ligger i avstanden $1/(\Delta K)$ fra origo. Vi ser at størrelsene ψ og ΔK kan brukes som et mål for hvor nær stedkurven kommer det kritiske punktet $(-1 + j0)$. Dersom $1/(\Delta K) > 1$ vil stedkurven komme på feil side av det kritiske punktet, og ΔK må derfor være større enn 1 (0 [dB]) for å oppnå stabilitet. Dersom $\psi < 0$, vil også stedkurven komme på feil side av det kritiske punktet, og ψ må derfor være større enn null for å oppnå stabilitet.

Vi har følgende betegnelser

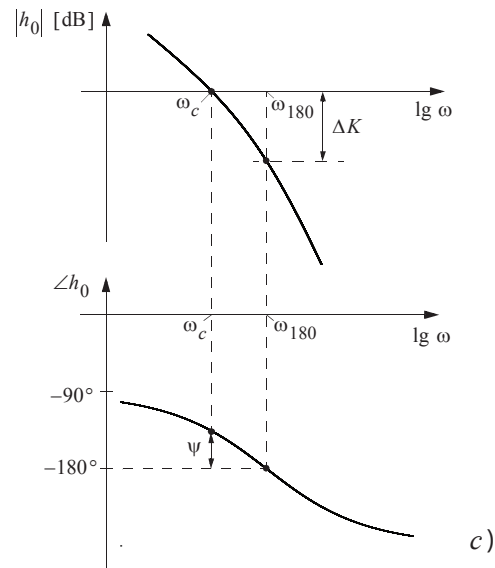
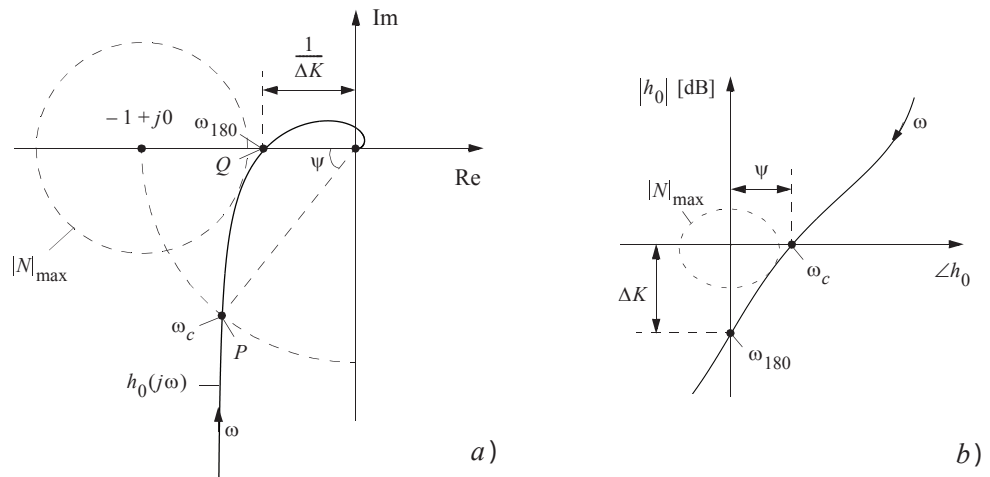
$$\begin{aligned} \psi &= \text{fasemargin} \\ \Delta K &= \text{forsterkningsmargin} \end{aligned}$$

De samme størrelsene er identifisert i figur 8.32 b) og c). Frekvensene ω_c og ω_{180} er også angitt for de tre framstillingsformene.

Figur 8.32

Fasemargin og forsterkningsmargin i

- a) Nyquist-diagram
 b) AF-diagram
 c) AFF-diagram



Innfører vi definisjonene

$$\psi = \angle h_0(j\omega_c) - (-180^\circ)$$

$$\frac{1}{\Delta K} = |h_0(j\omega_{180})| \quad , \quad \Delta K = \frac{1}{|h_0(j\omega_{180})|} \quad (8.57)$$

finder vi at betingelsene for stabilitet som er utledet i forrige avsnitt, kan uttrykkes som

$$\psi > 0 \quad , \quad \Delta K > 1 = 0 \text{ [dB]} \quad (8.58)$$

Fasemarginen ψ er et uttrykk for den ekstra negative fasedreiningen til $h_0(j\omega)$ som kan tåles ved $|h_0(j\omega)| = 1 = 0 \text{ [dB]}$, før systemet blir ustabil. Forsterkningsmarginen ΔK er et mål for forsterkningsøkningen som kan tåles ved $\omega = \omega_{180}$, før systemet blir ustabil.

Både systemets faseforskyvning og forsterkning kan tenkes å endre seg på grunn av parameterendringer i systemet som har sin årsak i ytre betingelser; slitasje og liknende. Det er derfor hensiktsmessig å spesifisere en viss stabilitetsmargin ved å bestemme minimale verdier for ψ og ΔK . Slike verdier kan for eksempel være

$$\begin{aligned} \psi &\geq 45^\circ \quad \text{og} \\ \Delta K &\geq 2 = 6 \text{ [dB]} \end{aligned}$$

Begrepene fasemargin og forsterkningsmargin er spesialtilfeller av et mer generelt begrep, **parametermargin**. En slik størrelse ville være et mål for hvor mye en vilkårlig parameter i systemet kan endre seg før systemet blir ustabil. Det er imidlertid som oftest vanskelig å bestemme marginen for vilkårlige parametre. Fasemargin og forsterkningsmargin er lette å identifisere og anvende. Et samlenavn for *fase- og forsterkningsmargin* er **stabilitetsmargin**.

Fasemargin og forsterkningsmargin angår egentlig bare forholdet til absolutt stabilitet. Figur 8.32 a) og b) viser koten for $|N(j\omega)|_{\max}$. Den største verdien av $|N(j\omega)|$ opptrer når $|1 + h_0(j\omega)|$ er minst fordi $|N(j\omega)| = 1/|1 + h_0(j\omega)|$. Figurene indikerer imidlertid at vi ved et hensiktsmessig valg av ψ og ΔK kan få en viss styring med hvilken verdi av $|N(j\omega)|_{\max}$ som vil oppstå. Dette avhenger av hvilken form stedkurven har mellom punktene P og Q . Det viser seg at systemer som kan beskrives med rasjonale funksjoner, og som ikke oppviser resonanser i frekvensområdet mellom P og Q , ofte vil ha en rimelig entydig sammenheng mellom størrelsen av paret $(\psi, \Delta K)$ og størrelsen $|N(j\omega)|_{\max}$.

Selvsagt kan vi unngå bruk av størrelsen ψ og ΔK og dimensjonere systemet både med hensyn til stabilitet og frekvensrespons ved hjelp av $N(j\omega)$, men denne størrelsen framgår ikke direkte av et Bodediagram. Størrelsene ψ og ΔK er derfor framfor alt nyttige i forbindelse med stabilitetsundersøkelse og syntese ved hjelp av Bodediagrammer.

EKSEMPEL 8.19: Fasemargin og forsterkningsmargin i stabilitetsundersøkelse

Vannkraftverk inneholder flere ulike reguleringssystemer slik at frekvensen og spenningen på den leverte elektriske effekten er tilfredsstillende. **Turbinregulatoren** styrer via en eller flere ventiler vannstrømmen inn på en turbin. Vi antar at turbinen tilføres vann over en *lang rørledning*. Transferfunksjonen mellom bevegelsen av pådraget (åpningen av ledeskovlene ved Francis-turbin) og det utviklede dreiemomentet kan da tilnærmet uttrykkes ved

$$h_1(s) = \frac{K(1 - T_r s)}{1 + 0.5T_r s} \quad (8.59)$$

T_r er her en konstant som er avhengig av rørdimensjonene og elastisiteten i vannet. Vi har fått et nullpunkt i høyre halvplan (dette refereres ofte noe villedende til som et “negativt” nullpunkt). Den fysiske betydningen av dette er at når pådraget plutselig økes, vil det utviklede dreiemomentet først plutselig avta for siden å stige mot sin nye stasjonære verdi.

Antas turbinen å være tapsfri, vil transferfunksjonen mellom netto dreiemoment (differansen mellom det utviklede dreiemomentet og belastningsmomentet) og hastighet bli

$$h_2(s) = \frac{1}{Is}$$

der I er aggregatets (turbin + generator) treghetsmoment.

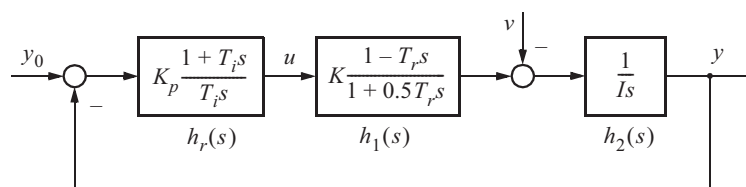
$h_0 = h_r h_1 h_2$ for dette systemet vil bli ikke-minimum-fase (IMF) på grunn av leddet $1 - T_r s$ i h_1 .

Vi skal nå anta at turbinregulatoren har transferfunksjon

$$h_r(s) = K_p \frac{1 + T_i s}{T_i s} \quad (8.60)$$

(Dette er en såkalt PI-regulator som blir nærmere presentert i avsnitt 9.3.1. Dette er en forsmak på en annen regulator enn den rene proporsjonal- (P-)regulatoren vi stort sett har anvendt til nå.)

Figur 8.33 viser et blokkdiagram for det komplette reguleringssystemet.



Figur 8.33
Blokkdiagram for turbin med regulator

(Denne modellen er ikke helt realistisk fordi en virkelig turbinregulator må ha minst en ekstra pol i sin transferfunksjon som skyldes pådragsmotoren. Men forenklingen har liten betydning for de punktene vi skal diskutere her).

Vi begynner med proposjonalregulator K_p . En vanlig prosedyre er å starte med ”kritisk” forsterkning K_{pk} = den forsterkningen som bringer systemet på stabilitetsgrensa. Da vil $\omega_c = \omega_{180}$. Ut fra symmetribetraktninger på $|h_0|_p$ og $\angle h_{op}$ må 0-dB-linja da ligge som vist i figur 8.34. Men vi må nå finne K_{pk} som svarer til denne posisjonen for 0-dB-linja: For $\omega \ll 1$ er $|h_0|_p \approx \frac{K_{pk}}{\omega}$. Det betyr at asymptoten til venstre skjærer 0-dB-linja i $\omega = K_{pk}$. Men denne ω er også lik $\frac{1}{T_r}$, da vi her knekker opp til horisontal asymptote $\Rightarrow \underline{\underline{K_{pk} = \frac{1}{T_r}}}$.

Mer generelt: Når man skal tegne Bode-diagram kan man for så vidt velge K_p fritt; man må da bare passe på å plassere den motsvarende 0-dB-linja i riktig posisjon. Etterpå kan man endre diagrammet til andre verdier av K_p ved å forskyve 0-dB-linja opp (mindre K_p) eller ned (større K_p). Husk at $\angle h_0$ (heldigvis) ikke endrer seg når K_p endres.

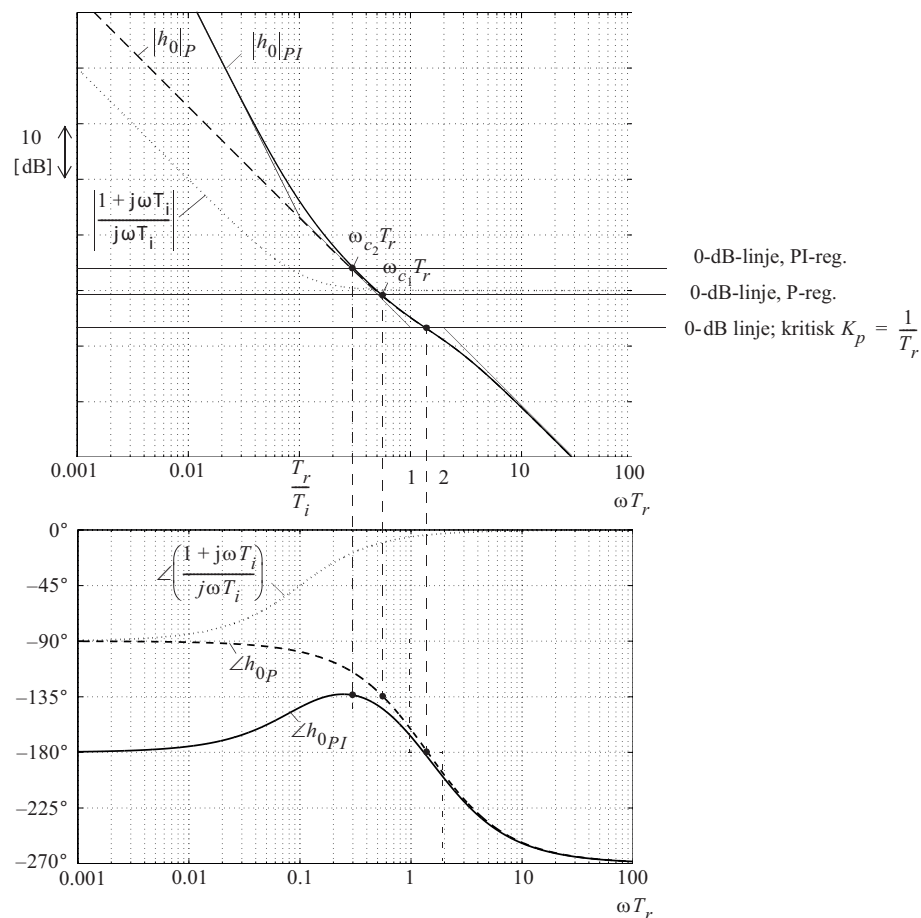
Vi antar for enkelhets skyld at $K/I = 1$ og tegner Bodediagram for $h_1(j\omega)h_2(j\omega)$. Resultatet blir som vist i figur 8.34. Vi legger spesielt merke til at nullpunktet i høyre halvplan gir en negativ fasedreining som nærmer seg -90° ved økende frekvens. Resultatet er at vi får en total fasevinkel som nærmer seg -270° for turbin med rørledning.

Antar vi først at $T_i = \infty$, vil regulatoren bli en ren proporsjonalregulator. Vi velger $K_p < 1/T_r$ slik at $\omega_c < \omega_{180}$, for at sløyfen med

$$h_0(s) = K_p \frac{1 + T_i s}{T_i s} \frac{1 - T_r s}{1 + 0.5 T_r s} \frac{1}{s}$$

skal tilfredsstillere kravene Bode-Nyquist-kriteriet stiller for stabilitet.

Figur 8.34
Bodediagram for
system med P-
regulator



Ved å dimensjonere systemet ut fra prinsippet med fasemargin og forsterkningsmargin og velge $\psi \geq 45^\circ$ og $\Delta K \geq 2 = 6$ [dB], finner vi at det er nødvendig å bruke $K_p T_r \approx 0.45 = -7$ [dB]. Hadde vi lagt 0-dB-linjene slik at den foreskrevne forsterkningsmarginen akkurat ble tilfredsstilt, ville vi ikke fått tilstrekkelig fasemargin. I andre tilfeller kan dette forholdet være omvendt.

Vi vil nå innføre PI-regulatoren. For vilkårlig verdi av T_i vil uttrykket

$$\frac{1 + j\omega T_i}{j\omega T_i}$$

få et amplitude- og faseforløp som vist i figur 8.34. Hvis $1/T_i \ll \omega_{c2}$ vil det negative fasebidraget på grunn av PI-virkningen bli lite i nærheten av ω_{c2} . Vi kan da superponere de to karakteristikkdelene og det er ikke nødvendig å endre den 0-dB-linjen som er bestemt for tilfellet med P-regulator. Velger vi derimot å forskyve punktet $1/T_i$ til høyre, vil dette som oftest medføre at vi må korrigere plasseringen av 0-dB-linjen for

å tilfredsstill den foreskrevne stabilitetsmarginen. For å oppnå hurtig nullstilling av statistisk avvik er det ønskelig å ha T_i minst mulig. En rimelig innstilling ut fra hensyn som her tas vil da være $T_i = 10T_r$ som gir karakteristikk som vist i figur 8.34. Denne verdien av T_i gjør at vi akkurat kan oppnå $\psi = 45^\circ$ ved å velge $K_p T_r = 0.3 = -11$ [dB]. Den nye 0-dB-linjen er vist i figuren. Vi kan derfor oppsummere de funne regulatorparametrene:

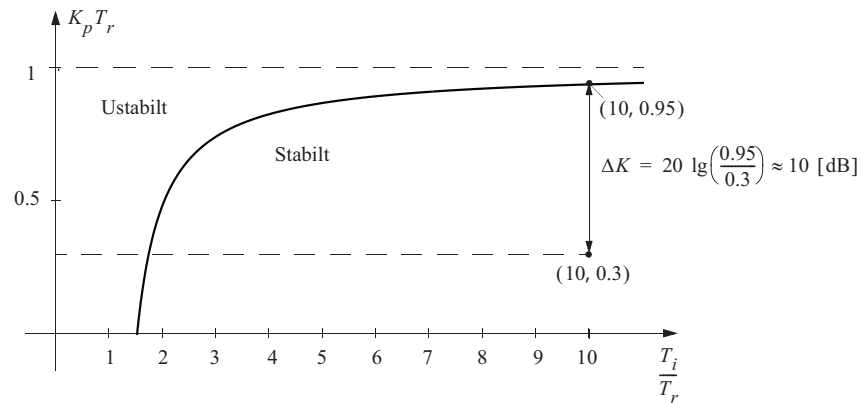
$$\begin{aligned} \text{P-regulator, } K_p T_r &= 0.45 = -7 \text{ [dB]}, & \omega_{c1} T_r &= 0.55 \\ \text{PI-regulator, } K_p T_r &= 0.3 = -11 \text{ [dB]}, & \omega_{c2} T_r &= 0.3, T_i = 10T_r \end{aligned}$$

Til sammenlikning kan vi benytte et av de algebraiske kriteriene på systemet i figur 8.33. Vi finner da stabilitetsbetingelsene:

$$\begin{aligned} 0 < K_p < \frac{1}{T_r} \\ T_i > \frac{1.5 - K_p T_r}{1 - K_p T_r} T_r \end{aligned} \tag{8.61}$$

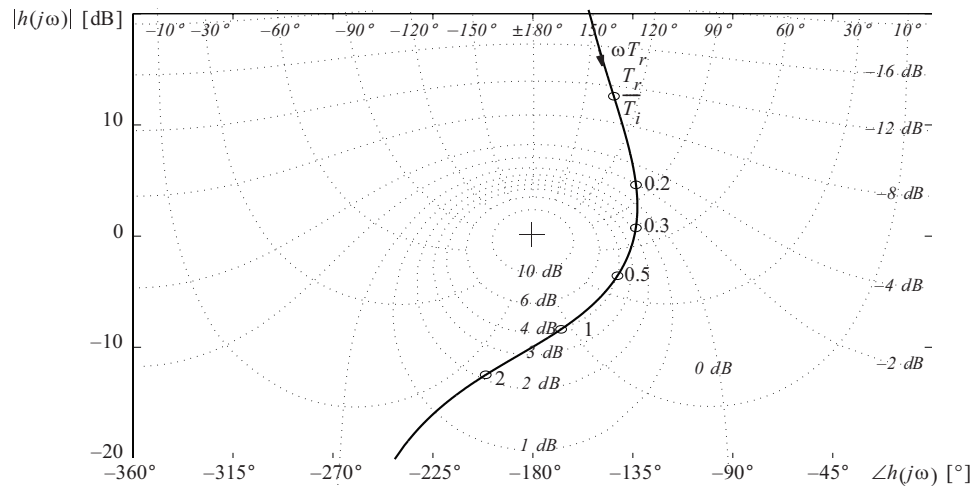
Disse betingelsene er framstilt grafisk i dimensjonsløs form i figur 8.35. I parameterområdet for stabilt system har vi også inntegnet et punkt som gjelder det valget av K_p og T_i vi gjorde foran. Det framgår her at vi har fått en økt forsterkningsmargin i forhold til det vi opprinnelig spesifiserte, nemlig $\Delta K = 2 = 6$ [dB]. Ifølge figur 8.35 har vi nå en forsterkningsmargin $\Delta K \approx 3.1 \approx 10$ [dB].

Figur 8.35
Grafisk framstilling
av stabilitets-
betingelser



Til slutt i dette eksemplet kan det være av interesse å se resultatet av bruken av begrepene fasemargin og forsterkningsmargin med hensyn på systemets reguleringsgrad $N(j\omega)$. Figur 8.36 viser de endelige karakteristikkene fra 8.34 i et Nicholsdiagram. Vi ser at vi har oppnådd $|N(j\omega)|_{\max} = 4.5$ [dB]. I dette tilfellet gir bruk av de gitte kravene til fasemargin og forsterkningsmargin en ekstra lav resonanstopp i reguleringsgraden, fordi grafen til $h_0(j\omega)$ krummer seg så sterkt i nærheten av det kritiske punkt. Men også generelt kan man si at krav til både fase- og forsterkningsmargin er nokså ekvivalent med krav om lav resonanstopp $|N(j\omega)|_{\max}$.

Figur 8.36
Nicholsdiagram



EKSEMPEL 8.20: Stabilitet av åpent ustabile systemer

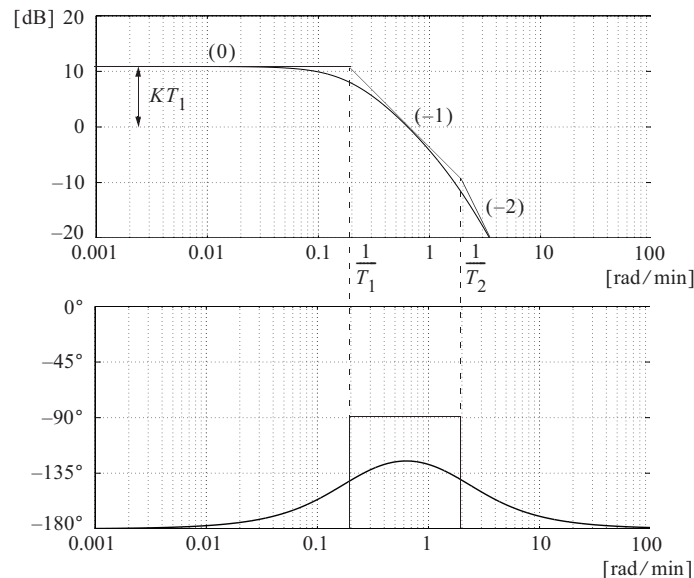
Bodediagrammer kan også benyttes til å undersøke stabilitetsforholdene av tilbakeløpede systemer som er åpent ustabile. En illustrasjon av dette får vi ved å undersøke reaktoren som ble introdusert i eksempel 3.6 og videre diskutert i eksempel 8.15.

Vi går ut fra

$$h_0(s) = \frac{K}{(1 + T_2s)(s - a)} = \frac{KT_1}{(1 + T_2s)(-1 + T_1s)} \quad (8.62)$$

Vi velger $T_1 = 5$ [min], $T_2 = 0.1T_1$ og $KT_1 = 3.5$. Figur 8.37 viser Bodediagrammet for systemet.

Figur 8.37
Bodediagram for
 $h_0(j\omega)$

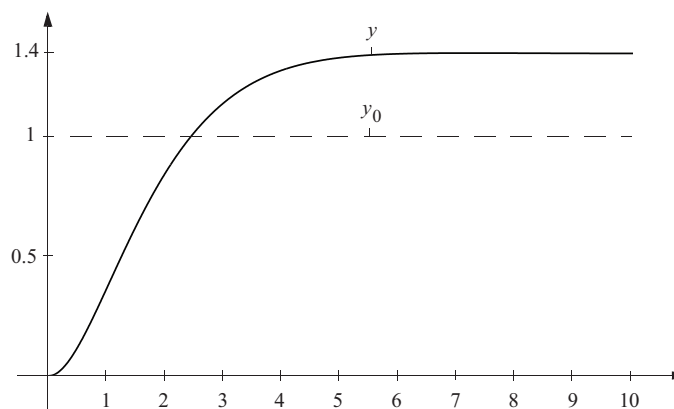


Polen i høyre halvplan medfører at fasekurven starter fra -180° i stedet for fra null, som er det vanlige. Betingelsen for stabilitet er som før at vi har positiv fasemargin

og positiv forsterkningsmargin, og det vil vi oppnå dersom $KT_1 > 1 = 0$ [dB] og $T_2 < T_1$. Dette er det samme resultatet som vi har funnet tidligere. En fordel med framstilling i Bode-diagram er at vi kan undersøke langt mer kompliserte systemer, for eksempel bruk av andre mer sammensatte regulatorer, enn det som er lett å håndtere ved hjelp av polare diagrammer av den typen som er vist i figur 8.25. Men når vi undersøker systemer som er åpent ustabile, bør vi alltid først plote $h(j\omega)$ i polar-diagram og bruke det generelle Nyquist-kriteriet som er beskrevet i avsnitt 8.4.1.

Sprangresponsen fra θ_0 til θ av det tilbakekoblede systemet er vist i figur 8.38.

Figur 8.38
Sprangrespons



Legg merke til at systemet har en stasjonær forsterkning som er større enn 1. For å finne årsaken til dette ser vi på følgefórhødet

$$\frac{\theta}{\theta_0}(s) = \frac{h_0(s)}{1 + h_0(s)} = \frac{KT_1}{KT_1 - 1} \frac{1}{1 + \frac{T_1 - T_2}{KT_1 - 1}s + \frac{T_1 T_2}{KT_1 - 1}s^2} \quad (8.63)$$

Dersom vi bruker sluttverditeoremet (4.22) på (8.63) og setter inn tallverdier, får vi et stasjonært avvik på 0.4 ved enhetssprang.

Vi kunne ha brukt en *PI-regulator* (som vi kommer tilbake til i kapittel 9) for å fjerne det stasjonære avviket.



Syntesemetoder for lineære reguleringsystemer

- 9.1 Innledning og motivering 313**
- 9.2 Systemspesifikasjoner 315**
- 9.3 Regulatorsyntese ved seriekompensasjon 327**
- 9.4 Når trenger vi integrerende regulator i en reguleringsløyfe? 386**
- 9.5 Kompensasjon med intern tilbakekopling 391**
- 9.6 Foroverkopling i monovariabel systemer 406**
- 9.7 Estimatorbasert regulering 428**
- 9.8 Parametrisk optimering av systemer med gitt struktur 436**
- 9.9 Robust regulering 453**
- 9.10 Sluttverdiregulering 459**

9.1 INNLEDNING OG MOTIVERING

Et reguleringsystem konstrueres og installeres for å gi en prosess ønskede egenskaper. Vi har tidligere i boka sett på egenskaper som stabilitet av reguleringsløyfer, følgeforhold og reguleringsgrad. I dette kapitlet skal vi systematisere konstruksjonsprosessen ved først å diskutere hvordan egenskapene til et reguleringsystem spesifiseres. Deretter beskrives viktige forhold i konstruksjonsprosessen med fokus på valg av regulatorstype. Tema som berøres er seriekompensasjon (avsnitt 9.3), bruk av integralvirkning (avsnitt 9.4), intern tilbakekopling (avsnitt 9.5), foroverkopling (avsnitt 9.6), estimatorbasert regulering (avsnitt 9.7) og optimalisering (avsnitt 9.8).

Konstruksjonsprosessen vil alltid avhenge av det aktuelle problemet som behandles. Reguleringsystemene for henholdsvis dynamisk posisjonering av et skip og for prosessering av olje og gass på en offshore-plattform vil for eksempel være vesensforskjellige. De inneholder alltid løsninger som er tilpasset lokale forhold. For å

formidle bredden i problemstillinger inneholder derfor avsnitt 9.3-9.8 et stort antall eksempler med rot i virkeligheten. Men eksemplene er idealiserte slik at det er mulig å beskrive dem på et begrenset antall sider.

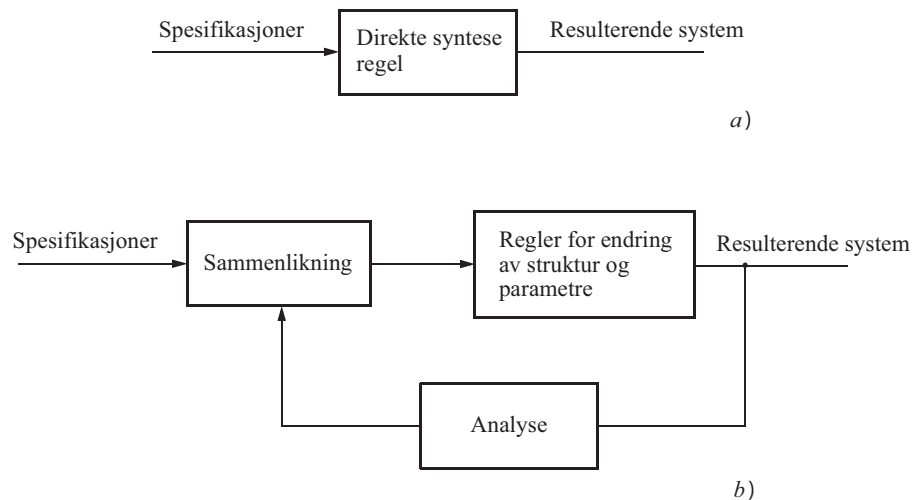
Konstruksjonen av et reguleringsystem som skal tilfredsstillende et på forhånd gitt sett av krav (spesifikasjoner), ut fra gitte systemelementer, kalles ofte for **syntese**. Dette er en av ingeniørens viktigste oppgaver.

Synteseoppgaven vil i de fleste tilfeller være svært omfattende og kreve kunnskaper innenfor en rekke fagfelter. *Prosessteknisk* innsikt (marinteknisk, maskinteknisk, kjemiteknisk, metallurgisk og liknende) vil være av avgjørende betydning for syntese av et regulerings teknisk system. Videre vil innsikt i *instrumenteringsteknikk* og *datateknikk* være viktig for realisering av systemet, og *driftsteknisk* og *økonomisk* innsikt vil være av betydning for planlegging og drift.

Når vi i dette kapitlet skal omtale syntese av reguleringsystemer, må det derfor presiseres at det er en liten del av den totale oppgaven som drøftes. Instrumentering og datateknisk implementering av reguleringsalgoritmer er for eksempel ikke behandlet.

Syntese av et system kan foregå på ulike måter. **Direkte syntese** og **iterativ syntese** som illustrert i figur 9.1, er to viktige metoder.

Figur 9.1
a) Direkte syntese
b) Iterativ syntese



Ved *direkte syntese* foreligger et sett synteseregler, som med utgangspunkt i gitte spesifikasjoner direkte leder til bestemmelse av systemets struktur og parametre. Ved *iterativ syntese* foreligger vanligvis noen ulike forslag til løsninger. Disse analyseres ved hjelp av tilgjengelige metoder. Resultatet av analysen sammenliknes med de gitte systemspesifikasjonene. Dersom resultatet ikke er tilfredsstillende, foretas det korreksjoner i struktur og parametre. Den iterative syntesen innebærer derfor en form

for tilbakekopling i konstruksjonsprosessen. Den vanligste syntesemetode i industriell virksomhet er den iterative metoden.

Dette kapitlet avsluttes med en kort behandling av robust reguleringsteori (avsnitt 9.9) og sluttverdiregulering (avsnitt 9.10).

9.2 SYSTEMSPESIFIKASJONER

Måten systemspesifikasjoner presenteres på avhenger gjerne av hvilken type reguleringssystem som betraktes. Et regulert (styrt) system vil som oftest vurderes ut fra:

- a) sin evne til å motvirke uønskede prosessforstyrrelser
- b) sin evne til å følge med varierende referanser eller bevegelige mål
- c) ufølsomhet overfor variasjoner i prosessparametre
- d) (viktigst!) evne til stabilisering av en ustabil prosess

Flere slike krav til egenskaper vil ofte forekomme samtidig.

Vi skal betrakte presentasjon av spesifikasjoner for **prosessregulering** der det stort sett er egenskapene a), c) og d) ovenfor som er vesentlige, og **følgesystemer** (**servomekanismer**, **fartøystyring** og liknende), der det stort sett er egenskapene b), c) og d) som er avgjørende.

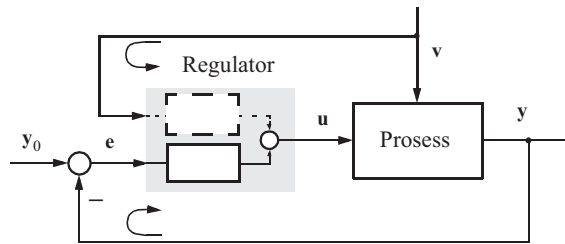
9.2.1 Spesifikasjon for prosessregulering

Vi skal her i hovedsak ta for oss de krav som stilles til prosessreguleringssystemer, men mange av dem vil også gjelde ved følgeregulering, som ellers er spesielt behandlet i avsnitt 9.2.2. Den konvensjonelle problemstillingen ved prosessregulering er konstruksjon av en regulator som skal generere et sett av *pådragsvariable* (\mathbf{u}), slik at et sett av målevariable (\mathbf{y}) til enhver tid er minst mulig forskjellig fra et sett av *referansestørrelser* (\mathbf{y}_0), når prosessen utsettes for normale *forstyrrelser* (\mathbf{v}).

Ved prosessregulering vil som oftest referansene (\mathbf{y}_0) ansees som konstante eller langsomt varierende. Det forutsettes at *målevariablene* (\mathbf{y}) er representative for *egenskapene* i prosessen som ønskes behersket.

Figur 9.2 viser en ganske generell struktur som løser det gitte problemet, der “regulatoren” påvirkes av avvikene (\mathbf{e}) mellom referansene og de tilhørende målingene og av prosessforstyrrelsene som måtte være målbare. Vi ser at reguleringssystemet utfører to funksjoner, en *tilbakekopling* fra (\mathbf{y}) gjennom regulatoren til \mathbf{u} og en *foroverkopling* fra \mathbf{v} gjennom regulatoren til \mathbf{u} .

Figur 9.2
System med
forstyrrelse



Hensikten med tilbakekoplingen er å gjøre $e = y_0 - y$ så liten som mulig, og hensikten med foroverkoplingen er å generere en direkte motvirkning av forstyrrelsens innflytelse på prosessen.

Vi skal først konsentrere oss om spesifikasjon av den delen av reguleringsystemet som utgjøres av tilbakekoplingen. Vi venter med foroverkoplingen til avsnitt 9.6.

Det teoretiske grunnlaget som er gitt i de foregående kapitlene, og særlig kapittel 7, gir gode verktøy for utvikling av enkle og effektive spesifikasjoner for prosessreguleringsystemer basert på tilbakekopling. Dette gjelder framfor alt monovariabel systemer.

Følgende egenskaper er av størst betydning ved spesifikasjon av et *prosessregulerings-system*:

- **Statisk nøyaktighet** er et mål for en regulert størrelses relative avvik fra sin referanseverdi når prosessen utsettes for en konstant forstyrrelse (belastning).
- **Dynamisk nøyaktighet** er et mål for systemets evne til å holde regulerte variable på sine referanseverdier når prosessen utsettes for dynamiske påvirkninger av en spesifisert karakter.
- **Stabilitetsmargin** er et mål for hvor mye essensielle parametre i systemet kan endres før stabilitetsgrensen nås.

Alle disse egenskapene og en rekke andre som senere skal berøres, kan gis en hensiktsmessig spesifikasjon ved hjelp av begrepene som er introdusert i kapittel 7. Dette gjelder framfor alt størrelsen $N(s)$ = reguleringsgrad. Denne størrelsen er en transferfunksjon i monovariabel systemer, mens den i multivariabel systemer er en transfermatrise. I det følgende skal vi bare se på spesifikasjon av monovariabel reguleringsystemer ved hjelp av størrelsen $N(s)$ ($N(j\omega)$). Multivariabel systemer diskuteres i kapittel 10.

Statisk nøyaktighet

Statisk nøyaktighet kan vi finne av (7.12) når vi antar at $y_0 = 0$. Statiske forhold reflekteres ved å la frekvensen $s = j\omega$ gå mot null, og vi finner derfor

$$\frac{e(s=0)_{\text{med}}}{y(s=0)_{\text{uten}}} = -N(s=0) \quad (9.1)$$

der indeksene “med” og “uten” henviser til henholdsvis med og uten regulering.

Likning (9.1) viser at den statiske nøyaktigheten er lik den statiske reguleringsgraden $N(0)$. Hvis vi som eksempel ønsker at det statiske reguleringsavviket skal være mindre enn 1% av variasjonen vi får i y når prosessen uten regulering påvirkes av aktuelle konstante forstyrrelser, må vi velge $N(0) < 0.01 = -40$ [dB].

Dynamisk nøyaktighet

Dynamisk nøyaktighet angir tilsvarende statisk nøyaktighet reguleringsavviket når systemet utsettes for dynamiske påvirkninger. Særlig enkelt ser vi dette når vi antar sinusformede forstyrrelser med varierende frekvens. Vi finner da at det relative reguleringsavviket blir

$$\left| \frac{e(j\omega)_{\text{med}}}{y(j\omega)_{\text{uten}}} \right| = |N(j\omega)| \quad (9.2)$$

Dette betyr at vi ved å spesifisere verdien av $|N(j\omega)|$ ved forskjellige verdier av ω , kan bestemme systemets evne til å undertrykke forstyrrelsers virkning ved disse frekvensene.

Dersom vi ville spesifisere detaljer i tidsforløpet av reguleringsavviket etter en kjent forstyrrelse, kunne vi ta utgangspunkt i uttrykket

$$e(t) = L^{-1}(e(s)) = L^{-1}(-N(s)y(s)_{\text{uten}}) \quad (9.3)$$

I denne forbindelse er det nyttig å ha for øyet de forenklede metodene for bestemmelse av sammenhengen mellom frekvensrespons og tidsrespons som er beskrevet i avsnitt 6.5.

Stabilitetsmargin

Stabilitetsmargin kan, som beskrevet i avsnitt 8.4.9, spesifiseres ved hjelp av størrelsene **fasemargin** og **forsterkningsmargin**. Dette er hensiktsmessig spesielt i den innledende fasen under dimensjonering av systemet.

Vi kan også spesifisere $|N(j\omega)|_{\text{max}}$, fordi vi derved automatisk oppnår bestemte verdier på fasemargin og forsterkningsmargin. Velger vi for eksempel å spesifisere $|N(j\omega)|_{\text{max}} = 6$ [dB], finner vi ved å studere et Nicholsdiagram at vi vil oppnå ihvertfall $\psi > 30^\circ$ og $\Delta K > 6$ [dB]. I de fleste tilfellene blir marginene en del større enn dette. Spesifikasjon av $|N(j\omega)|_{\text{max}}$ er også ønskelig ut fra hensyn til den dynamiske nøyaktigheten dersom systemet utsettes for påvirkninger med sterkt innslag av frekvenser i nærheten av resonanstoppen. Dette er tidligere illustrert i avsnitt 6.5.1.

Det er i mange tilfeller lett å innføre en spesifikasjon om variasjonsmarginer for spesifikke parametre som vi vet kan endre seg. Det beste er da som oftest å bruke et av de algebraiske stabilitetskriteriene og uttrykke stabilitetsbetingelsene med hensyn på de aktuelle parametrene. Dette ble gjort i eksempel 8.19.

Båndbredde

I tillegg til de tre hovedegenskapene nevnt ovenfor, kan en lang rekke andre egenskaper med fordel spesifiseres ved oppbygging av et reguleringsystem. Systemets båndbredde er et eksempel på en slik egenskap. Båndbredden til et tilbakekoplet reguleringsystem er frekvensområdet hvor tilbakekoplingen er effektiv, det vil si der $|N(j\omega)| < 1$ (ekvivalent: hvor $M(j\omega) \approx 1$). Båndbredden vil i de fleste systemer kunne ekvivaleres med frekvensområdet under **kryssfrekvensen** ω_c , dvs. frekvenser der $|h_0(j\omega)| > 1$.

I mange praktiske systemer ønsker vi at båndbredden skal være maksimal. Dette medfører at systemet får størst mulig reaksjonshurtighet. Ved prosessregulering vil dette si at forstyrrelser undertrykkes i et størst mulig frekvensområde. Dette er imidlertid ikke ønskelig i alle slags systemer. Stor båndbredde på reguleringen vil ofte kunne medføre at bevegelige pådragsorganer (ledeapparater i vannturbiner og liknende) blir utsatt for unødig slitasje. Det kan derfor være interessant å spesifisere en “optimal” båndbredde. Metodene som er beskrevet i foregående kapitler, gir mange muligheter for å ta hensyn til slike forhold. Dette gjelder for eksempel metoden som er beskrevet i avsnitt 6.5.1.

Belastningsstivhet

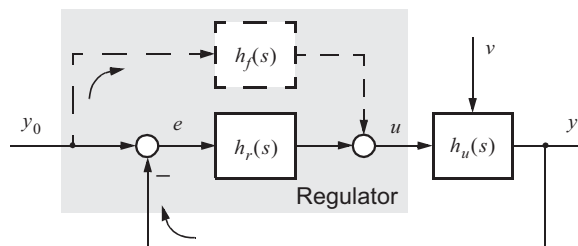
Systemets belastningsstivhet er en annen systemegenskap som ofte spesifiseres. Denne egenskapen er beslektet med statisk nøyaktighet som ble beskrevet foran. Mest typisk er dette ved turtallsregulering av kraftmaskiner som skal samkjøres, for eksempel over et elektrisk nett. Det er da ofte nødvendig å spesifisere den såkalte **statikken** som er et uttrykk for *turtallssenknningen* som funksjon av *belastningsmomentet* (belastningseffekten). Dette forholdet vil bli behandlet senere i forbindelse med et eksempel. Det er med andre ord ikke alltid ønskelig at systemets stivhet skal være maksimal. En ønsker snarere en spesifisert “ettergivenhet”.

9.2.2 Spesifikasjon for følgesystem (servomekanismer)

Et følgesystem har som primær oppgave å få en prosess (maskin, fartøy og liknende) til å endre tilstand som spesifisert ved en tidsvarierende referanse. Motorstyringen i eksempel 7.8 er et eksempel på et følgesystem. Figur 9.3 antyder at dette skulle kunne oppnås ved en kombinasjon av tilbakekopling og foroverkopling. Vi skal i dette avsnittet konsentrere oss om de viktigste trekkene ved spesifikasjon av monovariabelt tilbakekoplede følgesystemer. Spesifikasjon av foroverkoplingen vil vi komme tilbake til i avsnitt 9.6.

Figur 9.3

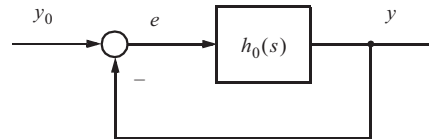
Et monovariabelt følgesystem med tilbakekopling og foroverkopling



I et følgesystem er vi framfor alt opptatt av systemets evne til å følge en referanseendring y_0 . Systemets evne til å undertrykke ytre forstyrrelser er imidlertid også essensiell og er behandlet i det foregående avsnittet.

I mange følgesystemer vil det være et dynamisk element i tilbakekoplingsgrenen som vist i figur 7.18. Analyse av dette systemet viser imidlertid som angitt i (7.22) at det vesentlige er å bestemme størrelsen $M(s)$, som er transferfunksjonen for et tilbakekoplet system med enhetstilbakekopling. Vi skal derfor i det følgende konsentrere oss om systemet som er vist i figur 9.4.

Figur 9.4
Et monovariabelt
tilbakekoplet
system



Et slikt følgesystems evne til å følge referansen y_0 kan beskrives med uttrykkene

$$\frac{y}{y_0}(s) = \frac{h_0(s)}{1 + h_0(s)} = M(s) = \text{følgforhold} \quad (9.4)$$

eller

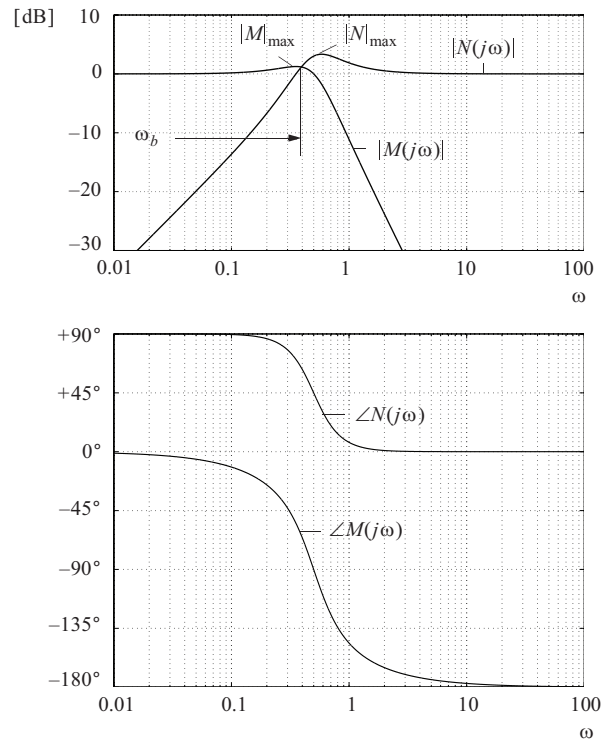
$$\frac{e}{y_0}(s) = \frac{1}{1 + h_0(s)} = N(s) = \text{avviksforhold} \quad (9.5)$$

De to uttrykkene egner seg til beskrivelse av systemet i forskjellige frekvensområder. Når $|h_0(s)| \gg 1$, vil $|M(s)| \approx 1$ og $|N(s)| \approx 1/|h_0(s)|$. Likning (9.5) egner seg da best til å beskrive systemets egenskaper, fordi vi får detaljert innsikt i hvordan systemets avvik forholder seg.

Når $|h_0(s)| \ll 1$, vil vi ha $|N(s)| \approx 1$ og $|M(s)| \approx |h_0(s)|$. Da egner (9.4) seg best til å beskrive systemets egenskaper.

Ved hjelp av frekvensanalytiske betraktninger er det lett å gi spesifikasjoner for et følgesystems egenskaper ut fra (9.4) og (9.5) når systemet er utsatt for periodiske referanseendringer eller endringer som med rimelig nøyaktighet kan beskrives som sammensatt av periodiske endringer, som diskutert i avsnitt 6.5.1. Vi kan da for eksempel beskrive den dynamiske nøyaktigheten ved å spesifisere verdier av $|N(j\omega)|$ ved forskjellige frekvenser som er aktuelle i referansesignalet. Frekvensen $\omega = 0$ representerer en konstant referanseverdi, og dersom en ønsker at systemet skal stille seg inn med null avvik når $y_0 = \text{konstant}$, må vi forlange at $|N(0)| = 0 = -\infty$ [dB].

Figur 9.5
Typiske diagrammer for $|M(j\omega)|$, $\angle M(j\omega)$, $|N(j\omega)|$ og $\angle N(j\omega)$ som funksjoner av $\lg \omega$



Figur 9.5 viser typiske diagrammer for $|M(j\omega)|$, $\angle M(j\omega)$, $|N(j\omega)|$ og $\angle N(j\omega)$ som funksjoner av $\lg \omega$, for et følgesystem. Frekvensanalytiske størrelser som kan være aktuelle å spesifisere, er:

- 1) **Båndbredden** ω_b som er det frekvensområdet hvor tilbakekoplingen er effektiv, det vil si der hvor $M(j\omega) \approx 1$.
- 2) $|M(j\omega)|_{\max}$ og/eller $|N(j\omega)|_{\max}$
- 3) $|N(j\omega)|$ og/eller $|M(j\omega)|$ for aktuelle frekvenser.
- 4) Formen på $h_0(s)$ ($h_0(j\omega)$) for å sikre systemets evne til å følge visse tidsforløp.

Et følgesystems evne til å følge visse tidsforløp

Ofte er det aktuelt å spesifisere et følgesystems evne til å følge visse gitte tidsforløp, som en sprangfunksjon $y_0(t) = \mu_1(t)$, en rampefunksjon $y_0(t) = \mu_2(t)$ eller en parabelfunksjon $y_0(t) = \mu_3(t)$. Disse funksjonene er beskrevet i tabell 4.1.

Vi er nå interessert i å finne hvordan følgesystemets avvik $e(t)$ (**følgeavviket**) utvikler seg når $t \rightarrow \infty$ for hver av disse tre typene referansesignaler.

Vi vet at

$$e(s) = N(s)y_0(s) = \frac{1}{1+h_0(s)}y_0(s) \quad (9.6)$$

og benytter oss av **sluttverditeoremet** som sier at

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} se(s) \quad (9.7)$$

Vi antar at det åpne følgesystemets transferfunksjon er

$$h_0(s) = \frac{K(1 + a_1s + a_2s^2 + \dots)}{s^p(1 + b_1s + b_2s^2 + \dots)} \quad (9.8)$$

og forutsetter som vist i (9.8), at det er p integrasjoner i den åpne sløyfen og et antall nullpunkter og poler som gir polynomene i teller og nevner.

Med bakgrunn i tabell 4.1 innfører vi

$$y_0(s) = q! \frac{1}{s^{q+1}} \quad (9.9)$$

der $q = 0$ tilsvarer $y_0(t) = \mu_1(t)$, $q = 1$ tilsvarer $y_0(t) = \mu_2(t)$ (dvs. rampe) og $q = 2$ tilsvarer $y_0(t) = \mu_3(t)$ (dvs. parabel).

Vi innfører (9.8) og (9.9) i (9.6)

$$e(s) = \frac{q!}{s^{q+1}} \frac{s^p(1 + b_1s + b_2s^2 + \dots)}{s^p(1 + b_1s + b_2s^2 + \dots) + K(1 + a_1s + a_2s^2 + \dots)} \quad (9.10)$$

Bruk av (9.7) på (9.10) gir

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} \frac{s^{p-q}}{s^p + K} q! \quad (9.11)$$

Resultatet er vist i tabell 9.1 og i figur 9.6

TABELL 9.1 Det stasjonære avviket, $\lim_{t \rightarrow \infty} e(t)$

Referanse	Systemtype		
	$p = 0$	$p = 1$	$p = 2$
Sprang $q = 0$	$\frac{1}{1 + K}$	0	0
Rampe $q = 1$	∞	$\frac{1}{K}$	0
Parabel $q = 2$	∞	∞	$\frac{2}{K}$

Tabell 9.1 viser at

- Vi må ha minst en integrasjon, altså $p \geq 1$, for å oppnå null stasjonært avvik når referansesignalet er et sprang. Dersom $p = 0$ (ingen integrasjoner i sløyfen), vil det stasjonære avviket bli proporsjonalt med faktoren $1/(1 + K)$ og med sprangets størrelse.
- Dersom referansesignalet er en rampefunksjon, må $p \geq 2$ for å oppnå null stasjonært avvik. Det stasjonære avvik vil bli proporsjonalt med $1/K$ og med rampens steilhet dersom $p = 1$. Et system med $p = 0$ vil ikke klare å følge med et rampesignal og vil få et stadig økende avvik.
- Dersom referansesignalet er en parabelfunksjon, vil vi selv med $p = 2$ få et stasjonært avvik som er proporsjonalt med $2/K$ og med parabelens akselerasjon. Dersom $p < 2$, vil avviket øke mot uendelig.

Resultatet i det foregående egner seg til spesifikasjon av følgesystemer på to måter.

For det første kan vi bestemme hvor mange integrasjoner (p) som trengs for å oppnå den ønskede stasjonære responsen på forskjellige typer referansesignaler.

Vi kan dessuten finne grunnlag for å spesifisere verdien av K ut fra størrelsen på det stasjonære avviket som tolereres. Mest typisk er dette for det tilfellet at $p = 1$ og $q = 1$. Dette tilfellet forekommer ofte i følgesystemer der $y_0(t)$ og $y(t)$ er posisjoner (eller vinkler) og der det finnes en eller annen form for motor i sløyfen. En vil da ha $p \geq 1$. Et referansesignal i form av en rampe benyttes dersom vi ønsker at følgesystemet (servomekanismen) skal følge et mål som beveger seg med konstant hastighet. Dersom $p = 1$, vil vi da få en konstant etterslepning som er omvendt proporsjonal med K , som vist i figur 9.6 b). Null følgeavvik oppnås når $p \geq 2$.

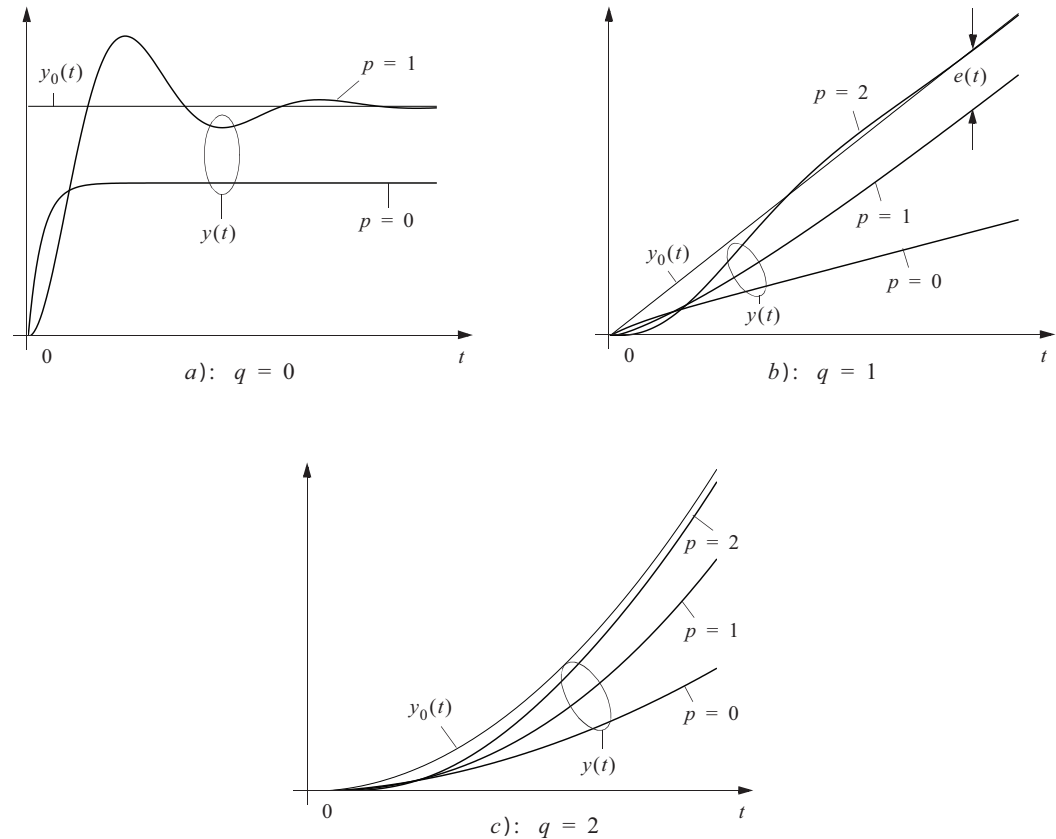
Figur 9.6

Respons når referansen er henholdsvis

a) Sprang

b) Rampe

c) Parabel

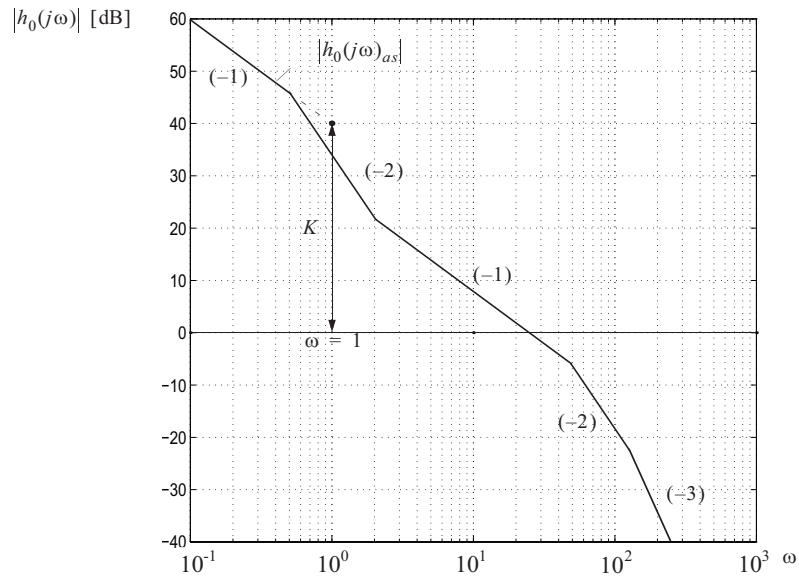


Figur 9.7 viser et asymptotisk amplitude-diagram for $|h_0(j\omega)|$ for en servomekanisme med $p = 1$. Vi ser der at størrelsen K framkommer ved skjæring mellom forlengelsen av asymptotene ved lave frekvenser og ordinaten gjennom $\omega = 1$. Denne størrelsen er altså et direkte mål for det stasjonære avviket når systemet påtrykkes en rampefunksjon og er vanlig som utgangspunkt for spesifikasjon av systemet.

I dette eksemplet finner $K = 100 = 40$ [dB] vi som gir en stasjonær ettersleping (som vist i figur 9.6 b)) på $e(t \rightarrow \infty) = 1/K = 0.01$.

I det foregående tok vi utgangspunkt i at referansesignalet var spesielt enkle tidsforløp. Vi skal nå se at en ut fra følgesystemets åpen-sløyfe-transferfunksjon $h_0(s)$ kan bestemme såkalte *avvikskoeffisienter* som i visse forbindelser kan være hensiktsmessige for spesifikasjon av systemegenskaper.

Figur 9.7
Asymptotisk amplitude-diagram for $|h_0(j\omega)|$



Vi antar at $h_0(s)$ har den form som er gitt i (9.8) og finner

$$\begin{aligned} N(s) &= \frac{s^p(1 + b_1s + b_2s^2 + \dots)}{s^p(1 + b_1s + b_2s^2 + \dots) + K(1 + a_1s + a_2s^2 + \dots)} \\ &= \frac{\tilde{b}_0 + \tilde{b}_1s + \tilde{b}_2s^2 + \dots}{\tilde{a}_0 + \tilde{a}_1s + \tilde{a}_2s^2 + \dots} \end{aligned} \quad (9.12)$$

Vi dividerer telleren med nevneren i likning (9.12) og finner

$$N(s) = e_0 + e_1s + e_2s^2 + \dots \quad (9.13)$$

der

$$\begin{aligned} e_0 &= \frac{\tilde{b}_0}{\tilde{a}_0} \\ e_1 &= \left(\tilde{b}_1 - \frac{\tilde{a}_1\tilde{b}_0}{\tilde{a}_0} \right) \frac{1}{\tilde{a}_0} \\ e_2 &= \left(\left(\tilde{b}_2 - \frac{\tilde{a}_2\tilde{b}_0}{\tilde{a}_0} \right) - \frac{\tilde{a}_1}{\tilde{a}_0} \left(\tilde{b}_1 - \frac{\tilde{a}_1\tilde{b}_0}{\tilde{a}_0} \right) \right) \frac{1}{\tilde{a}_0} \\ &\vdots \\ &\vdots \end{aligned}$$

Siden $\frac{e}{y_0}(s) = N(s)$, finner vi ved bruk av (9.13)

$$e(t) = e_0 y_0(t) + e_1 \frac{dy_0(t)}{dt} + e_2 \frac{d^2 y_0(t)}{dt^2} + \dots \quad (9.14)$$

Koeffisientene e_0 , e_1 og e_2 kalles **avvikskoeffisienter** og forteller oss altså hvordan følgesystemets avvik genereres av referansesignalet og dets deriverte.

Rekken i (9.13) er uendelig lang, og hvis vi bare betrakter noen få av de første leddene, kan betydelige feil oppstå. Vi får et inntrykk av dette ved å se på frekvensresponsen $|N(j\omega)|$ for et enkelt system

$$h_0(s) = \frac{K}{s(1+Ts)} \quad (9.15)$$

som gir

$$N(s) = \frac{s(1+Ts)}{K+s+Ts^2} \quad (9.16)$$

Det foregående resultatet gir for dette systemet

$$e(t) = 0 \cdot y_0(t) + \frac{1}{K} \dot{y}_0(t) + \frac{TK-1}{K^2} \ddot{y}_0(t) + \dots \quad (9.17)$$

Figur 9.8 viser $|N(j\omega)|$ sammen med frekvensresponsen for de tre første leddene i (9.17) når $K = 2/T$. Vi ser at overensstemmelsen mellom de to kurvene bare er brukbar ved de aller laveste frekvenser, det vil si når de høyere deriverte enn første deriverte av $y_0(t)$ bare antar svært små verdier. Metoden med avvikskoeffisienter vil derfor i praksis bare være brukbar når referansesignalets høyere deriverte er svært små eller dersom en tar med mange ledd i rekkeutviklingen.

Dersom et system av en type som vist i figur 9.4, er dimensjonert til å oppføre seg bra når referansen er en *rampefunksjon*, kan det godt tenkes at systemet ikke vil gi særlig gunstig respons når referansen er et *sprang*.

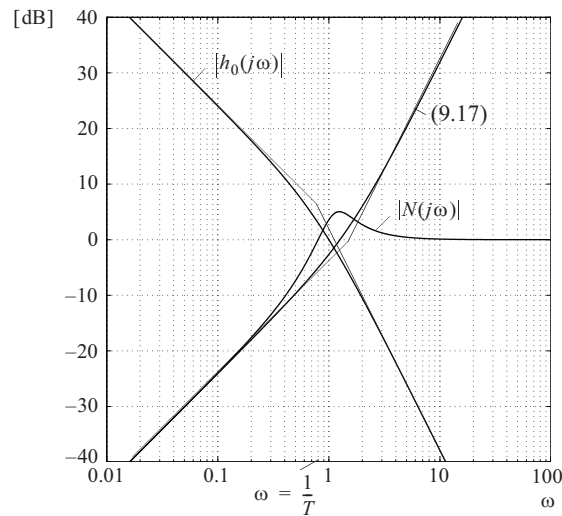
Et eksempel på dette finnes dersom vi ut fra tabell 9.1 velger å utføre systemet med to integrasjoner ($p = 2$), fordi vi ønsker at det skal kunne følge en rampefunksjon med null stasjonært avvik. Når dette systemet utsettes for et sprang i referansen, er responsen ikke særlig gunstig. Dette ser vi lettest ved å beregne

$$\int_0^{\infty} e(t) dt$$

Figur 9.8

$|N(j\omega)|$ og frekvensresponsen til de tre første ledene i (9.17) når

$$K = \frac{2}{T}$$



Vi bruker da følgende nyttige resultat

$$\int_0^{\infty} e(t) dt = \lim_{t \rightarrow \infty} \left(\int_0^t e(t) dt \right) = \lim_{s \rightarrow 0} \left(s \frac{e(s)}{s} \right) = e(s=0) \quad (9.18)$$

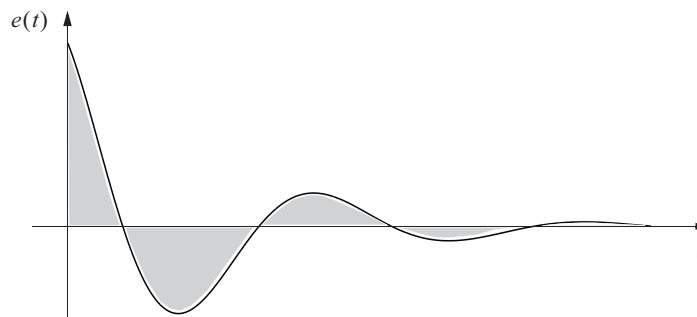
der vi har benyttet oss av både sluttverditeoremet og Laplace-transformen av det ubestemte tidsintegralet. Setter vi inn $h_0(s)$ som i (9.8), og $y_0(s)$ som i (9.9), med $p = 2$ og $q = 0$, finner vi

$$\int_0^{\infty} e(t) dt = e(s=0) = 0$$

Dette betyr at vi må få et innsvingningsforløp av avviket $e(t)$ med et stort oversving i negativ retning, fordi arealet under de positive delene skal utliknes av et tilsvarende areal under de negative delene, som vist i figur 9.9.

Figur 9.9

$e(t)$ som funksjon av tiden med $p = 2$ og $q = 0$



En slik løsning er selvsagt ikke gunstig. En enkel fysisk forklaring på dette er at den ene av de to integratorene i sløyfen vil bli “ladet” idet avviket går positivt og må “utlades” igjen for at systemet skal komme til ro. Derfor vil vi gjerne spesifisere at en av integratorene skal *utkoples*, så lenge avviket er over en bestemt liten tallverdi. med denne modifikasjonen vil vi få en **ulinear regulator**.

9.3 REGULATORSYNTese VED SERIEKOMPENSASJON

Den foregående analysen av tilbakekoblede reguleringsystemers funksjon og de metodene som benyttes for å spesifisere ønskede egenskaper til systemene, viser at den åpne sløyfens transferfunksjon $h_0(s)$ spiller en helt avgjørende rolle for resultatet. I de enkleste tilfellene vet vi at $h_0(s)$ oppstår som produktet av transferfunksjonene for prosessdelen $h_u(s)$ og regulatoren $h_r(s)$, altså $h_0(s) = h_u(s)h_r(s)$.

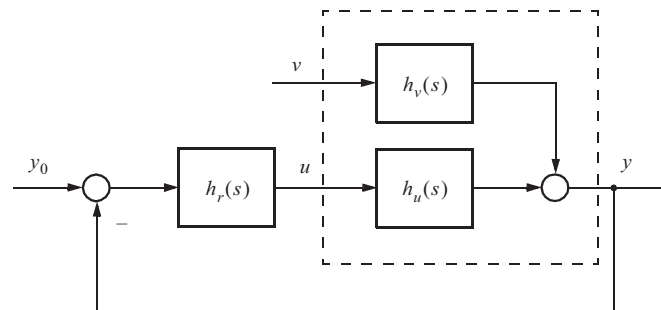
Dersom vi ønsker en bestemt form på $h_0(s)$ for å tilfredsstillе gitte spesifikasjoner og $h_u(s)$ er gitt, må vi finne fram til en regulator-transferfunksjon $h_r(s)$ som er realistisk og dessuten gjør at $h_0(s)$ tilfredstiller spesifikasjonene. Vi har ved å gjøre dette allerede valgt strukturen på løsningen. Den refereres ofte til som **seriekompensasjon** fordi kompensasjonselementet (regulatoren) står i serie med selve prosessen.

En annen form for kompensasjon benytter seg av interne (gjerne dynamiske) tilbakekoplinger rundt prosessen fra en eller flere “hjelpemålinger” i tillegg til den ene målingen som vanligvis betraktes i et monovariabelt system. Denne formen for kompensasjon kalles ofte **kompensasjon med intern tilbakekopling**. Denne behandles i avsnitt 9.5.

I det følgende skal vi behandle seriekompensasjon.

Analytisk seriekompensasjon av et monovariabelt tilbakekøpelt system består i å dimensjonere det dynamiske elementet som står i serie med prosessen, og som vi tidligere har referert til som en regulator (se figur 9.10).

Figur 9.10
Tilbakekøpelt system med forstyrrelse



Dersom $h_0(s)$ er gitt ut fra spesifikasjoner og $h_u(s)$ er gitt (fordi prosessen er gitt), skulle en i prinsippet kunne bestemme regulatoren som

$$h_r(s) = \frac{h_0(s)}{h_u(s)} \quad (9.19)$$

Siden $1/h_u(s)$ inngår, sier vi da at $h_r(s)$ finnes ved *inversjon* av $h_u(s)$. Dette svaret forutsetter imidlertid at en har vært i stand til å *spesifisere en fornuftig form på $h_0(s)$* slik at resultatet blir realiserbart.

Dersom vi velger $h_r(s)$ som en rasjonal transferfunksjon, vil polynomet i telleren høyst kunne være av samme grad som polynomet i nevneren, for at transferfunksjonen skal representere et realiserbart system. Dette setter bestemte restriksjoner på hva vi kan tillate oss å foreskrive for $h_0(s)$.

Et annet forhold er at en uheldig spesifisering av $h_0(s)$ kan gi en unødig komplisert $h_r(s)$, som ikke nødvendigvis gir noe bedre totalsystem enn en enklere og kanskje billigere form.

Hvis $h_u(s)$ inneholder en transportforsinkelse, har vi et problem med å realisere (9.19) fordi inversjon av $h_u(s)$ vil kreve at regulatoren tilføres *framtidige* (dvs. ennå ikke inntrufne) verdier av reguleringsavviket – den utfører *prediksjon*.

Hvis $h_u(s)$ inneholder nullpunkter i høyre halvplan blir regulatoren ustabil, fordi $h_r(s)$ får tilsvarende pol(er). I slike tilfeller er eksakt inversjon ikke mulig. Likevel kan den ideelle men ubrukelige $h_r(s)$ man da får av (9.19), brukes som utgangspunkt for å finne en mer realisérbar $h_r(s)$.

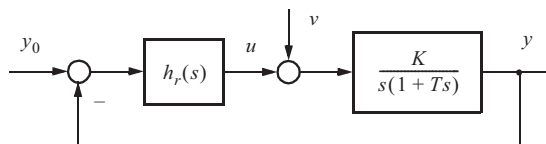
Vi skal med noen eksempler se på hvordan regulatoren kan utformes for noen viktige systemtyper.

EKSEMPEL 9.1: *Prosess med en ren integrator og én tidskonstant*

Vi skal betrakte en prosess $h_u(s) = \frac{K}{s(1+Ts)}$ (9.20)

Et blokkdiagram for denne prosessen med seriekompensasjon er vist i figur 9.11.

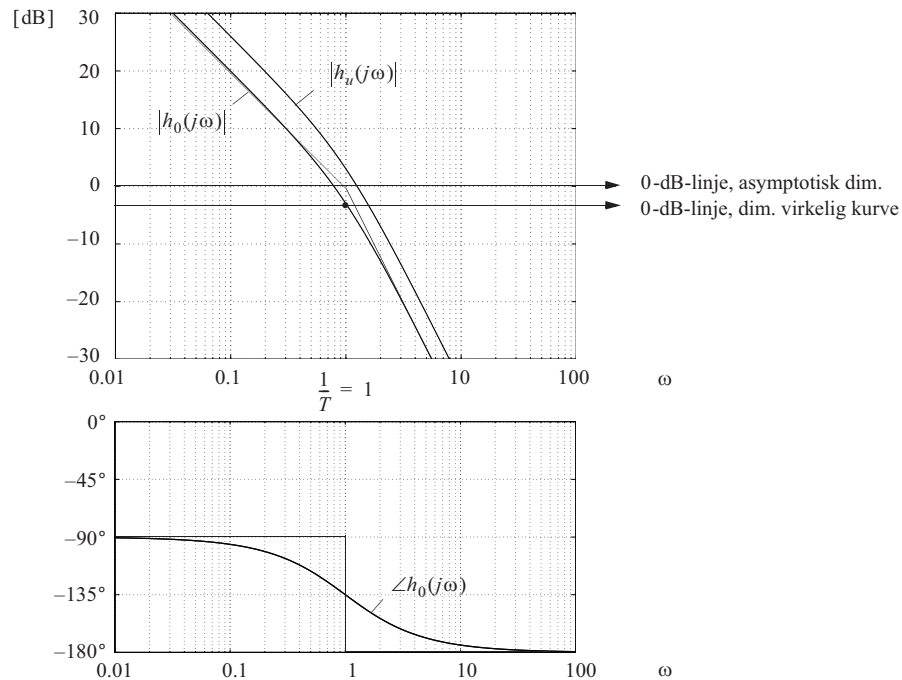
I dette tilfellet bestemmer vi oss for å bruke en regulator som bare er en forsterker. Det eneste vi forlanger er at systemet skal ha en rimelig stabilitetsmargin. Vi spesifiserer derfor en fasemargin $\psi \geq 45^\circ$ og setter *ingen krav til systemets båndbredde*.



Figur 9.11
Prosess med regu-
lator og tilbake-
kopling

Setter vi $T = 1$ og $K = 2$, får prosessen $h_u(s)$ et AFF-diagram som vist i figur 9.12.

Figur 9.12
AFF-diagram



Figuren viser at prosessen ikke trenger noen fasekorrigerende regulator for å få den spesifiserte fasemarginen, når kryssfrekvensen $\omega_c \leq 1/T$. Vi kan derfor benytte en proporsjonalregulator $h_r(s) = K_p$ og justerer denne for å oppnå en ønsket kryssfrekvens. Vi skal først dimensjonere K_p grovt ved å bruke asymptoter, så mer presis ved å bruke eksakte kurver.

Dimensjonering ved bruk av asymptotisk karakteristik

Det asymptotiske amplitudediagrammet i figur 9.12 viser at prosessens tidskonstant T ikke virker inn på $|h_0(j\omega)|_{as}$ for frekvenser mindre enn knekkfrekvensen. Vi finner derfor K_p ved hjelp av uttrykket

$$|h_0(j\omega_c)| \approx \frac{K_p K}{\omega_c} = K_p K T = 1 = 0 \text{ [dB]}$$

Dette gir forsterkningen

$$K_p = \frac{1}{KT} \quad (9.21)$$

Dimensjonering ved bruk av virkelig karakteristikk

Forsterkningen i (9.21) er noe lav fordi den er funnet med utgangspunkt i den asymptotiske kurven og ikke den virkelige.

Figur 9.12 viser at den virkelige amplitudekurven har verdien -3 [dB] ved frekvensen der $\angle h_0(j\omega) = 135^\circ$. For å oppnå $\psi = 45^\circ$ må amplitudekurven krysse 0-dB-linjen ved denne frekvensen. Vi plusser derfor 3 [dB] på forsterkningen fra (9.21). Dette gjøres ved å multiplisere med $\sqrt{2}$ siden $\sqrt{2} = 3$ [dB].

Forsterkningen funnet ved hjelp av de virkelige karakteristikkene blir da

$$K_p = \sqrt{2} \frac{1}{KT} \quad (9.22)$$

Det er viktig å vite at vi ikke trenger å gå veien om asymptotisk dimensjonering for å dimensjonere en forsterkning ut fra virkelige karakteristikk. Det rekkes å påplusse det antallet desibel $|h_u(j\omega)|$ må heves med for at den skal krysse 0-dB-linjen ved frekvensen som gir riktig fasemargin (forsterkningen virker ikke inn på fasekarakteristikken).

Figur 9.12 viser at systemet med proporsjonalregulator har uendelig stor forsterkningsmargin ($\Delta K = \infty$).

■

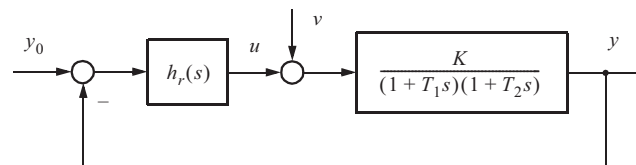
EKSEMPEL 9.2: Prosess med to tidskonstanter

Vi skal betrakte en modell som har framkommet i flere tidligere eksempler.

$$h_u(s) = \frac{K}{(1 + T_1s)(1 + T_2s)} \quad (9.23)$$

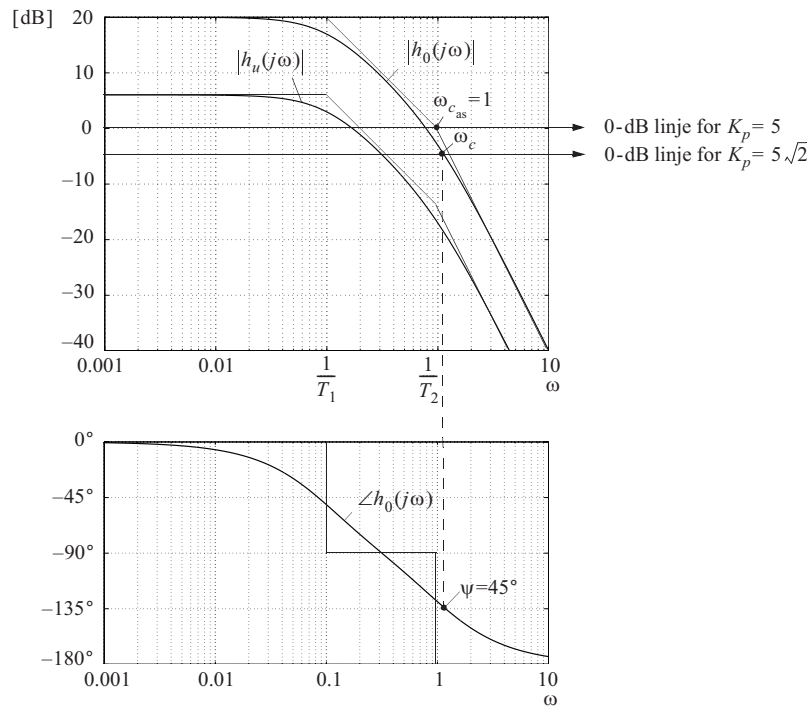
der $T_1 \gg T_2$. Et blokkdiagram for denne prosessen er vist i figur 9.13.

Figur 9.13
Prosess med regulator og tilbakekopling



Dersom vi setter $T_1 = 10$, $T_2 = 1$ og $K = 2$, får prosessen $h_u(s)$ et AFF-diagram som vist i figur 9.14.

Figur 9.14
AFF-diagram



Vi spesifiserer ikke noe krav til statisk nøyaktighet ($N(0)$), men ønsker størst mulig båndbredde og krever at fasemarginen $\psi > 45^\circ$.

Fasekarakteristikken for $h_0(j\omega)$ viser at prosessen ikke trenger noen fasekorreksjon for å få den spesifiserte fasemarginen, dersom $\omega_c \leq 1/T_2$. Vi benytter derfor en proporsjonalregulator K_p og må justere denne for å oppnå en ønsket kryssfrekvens.

Basert på de asymptotiske karakteristikkene i figur 9.14 kan vi bestemme K_p tilnærmet. Figur 9.14 viser at $|h_0|$ er $K_p K$ [dB] ved lave frekvenser. Siden asymptoten til amplituden faller med stigningstallet -1 mellom de to knekkfrekvensene, må $K_p K$ [dB] være lik avstanden mellom de to knekkpunktene $1/T_1$ og $1/T_2$ i diagrammet. Vi vet imidlertid at

$$\lg\left(\frac{1}{T_2}\right) - \lg\left(\frac{1}{T_1}\right) = \lg\left(\frac{T_1}{T_2}\right)$$

Dette gir (når vi vet at 20 [dB] tilsvarer 1 dekade)

$$\frac{20 \lg(K_p K)}{20} = \lg\left(\frac{T_1}{T_2}\right)$$

Forsterkningen blir da

$$K_p = \frac{T_1}{KT_2} \quad (9.24)$$

Denne verdien av K_p er altså bestemt ut fra en asymptotisk betraktning. Den stiplede linjen i figur 9.14 viser at vi ved å bruke de virkelige AFF-karakteristikkene kunne ha øket forsterkningen med ca. 3 [dB]. Dette svarer igjen til å multiplisere forsterkningen med en faktor på $\sqrt{2}$.

$$K_p \approx \sqrt{2} \frac{T_1}{KT_2} = 7 \quad (9.25)$$

Figur 9.14 viser at vi har uendelig forsterkningsmargin ($\Delta K = \infty$).

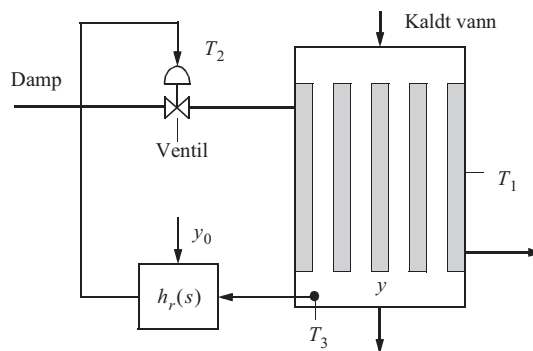
De verdiene av K_p som vi har funnet, er brukbare til å sikre stabilitet i det tilbakekoppledte systemet. Men ytelsen av systemet, for eksempel med hensyn til reguleringsgrad $N(j\omega)$, har vi ikke tatt hensyn til. Forsterkningen fra (9.25) gir for eksempel $N(0) = 1/(1 + 7 \cdot 2) = 0.07$. Dette svarer til ca. 7% reguleringsavvik, som ofte vil være for dårlig ytelse. Men det er hva vi kan oppnå med en proporsjonalregulator.

■

EKSEMPEL 9.3: Regulering av varmeveksler for varmtvannsforsyning

Vi betrakter den regulerede prosessen som er vist i figur 9.15, med blokkdiagram i figur 9.16. Denne typen reguleringsystem er aktuelt i f.eks. et fjernvarmeanlegg.

Figur 9.15
Varmtvannsforsyning som bruker damp til oppvarming



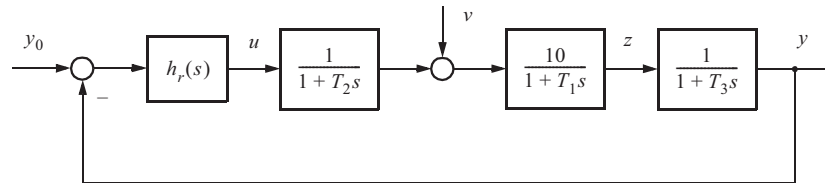
Prosessen består av en varmeveksler der damp som tilføres primærsiden gjennom en ventil skal varme opp kaldt vann på sekundærsiden. Selve varmeveksleren har en tidskonstant T_1 og ventilen har en tidskonstant T_2 . Temperaturen i det oppvarmede vannet måles ved hjelp av et måleelement som har en tidskonstant T_3 . Med god tilnærming kan vi anta at transferfunksjonen for denne prosessen fra inngangen på

ventilen til utgangen fra måleelementet vil være

$$h_u(s) = \frac{K}{(1 + T_1s)(1 + T_2s)(1 + T_3s)} \quad (9.26)$$

I (9.26) antar vi at $T_1 = 10$ [s], $T_2 = 2$ [s], $T_3 = 0.5$ [s] og $K = 10$.

Figur 9.16
Prosess med regu-
lator og tilbake-
kopling



Vi antar videre at den ønskede temperaturen y_0 til vannet stort sett er konstant, og at systemets oppgave for det meste er å undertrykke virkningen av forstyrrelser på prosessen (for eksempel endringer i temperaturen i det kalde vannet som kommer inn til varmeveksleren).

Følgende sett med spesifikasjoner antas å foreligge som beskrevet i avsnitt 9.2.1:

- $|N(0)| \leq -40$ [dB]
- $|N(j\omega_1)| \leq -20$ [dB] for $\omega_1 = 0.2$ [rad/s]
- $|N(j\omega)|_{\max} \leq +6$ [dB]
- $\omega_c =$ båndbredde = størst mulig

Spesifikasjon a) uttrykker som kjent systemets evne til å motvirke konstante forstyrrelser og vil i det foreliggende tilfellet medføre at systemets avvik vil være begrenset til 1% av utslaget vi ville ha fått i målingen, dersom det ikke var noen regulering.

Spesifikasjon b) er et mål for systemets evne til å undertrykke dynamiske forstyrrelser, for eksempel endringer i kaldtvannsstrømmen i frekvensområdet omkring $\omega = 0.2$ [rad/s]. Slike forstyrrelser vil bli redusert til omtrent 10% av det vi får uten regulering.

Spesifikasjon c) er et mål for en rimelig stabilitetsmargin og for at transiente innsvingningsforløp har en akseptabel demping.

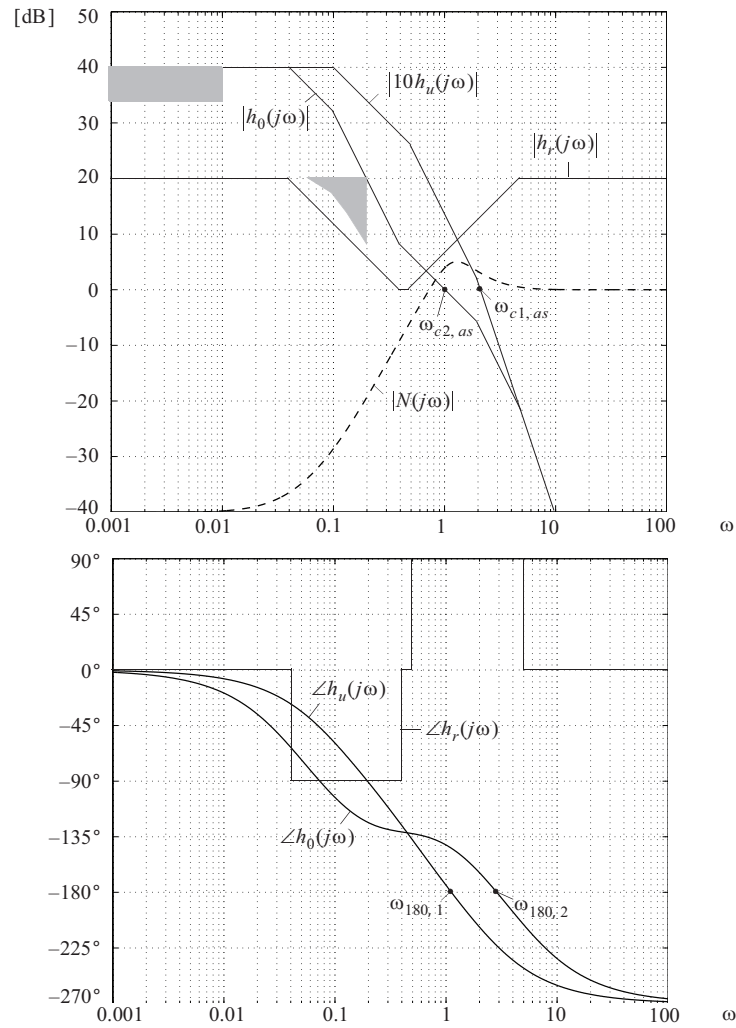
Spesifikasjon d) uttrykker bare at systemet skal regulere ut forstyrrelser så hurtig som mulig. Det forutsettes at det er tilstrekkelig effekt til disposisjon gjennom pådragsorganet til å gjøre dette ved de aktuelle formene for forstyrrelser.

Synteseoppgaven kan angripes på mange forskjellige måter. Vi vil i første omgang benytte oss av asymptotiske amplitude/fase/frekvens-diagrammer.

Siden oppgaven som antydnet i (9.19) er å finne en fornuftig $h_0(s)$ (eller $h_0(j\omega)$), kan vi forsøke å konstruere en $|h_0(j\omega)|$ som i tur og orden tilfredsstillter spesifikasjonene a) til d).

I det følgende refererer vi til figur 9.17.

Figur 9.17
AFF-diagram
for prosessen



Vi antar først at vi skal kunne klare oss med et kompensasjonselement (en regulator) som bare er en konstant (det vil si en proporsjonal regulator). Siden spesifikasjon a) krever at

$$|N(0)| = \left| \frac{1}{1 + h_0(0)} \right| \leq \frac{1}{100}$$

må vi kreve $|h_0(0)| > 100 = +40$ [dB]. Figur 9.17 viser derfor det asymptotiske amplitude-diagrammet for transferfunksjonen $10h_u(s)$. Denne funksjonen vil ved lave frekvenser ha en tallverdi lik 100 og spesifikasjon a) vil være tilfredsstilt dersom $|h_r(j\omega)| = 10$ ved $\omega = 0$.

Spesifikasjon b) som uttrykker

$$|N(j\omega_1)| = \left| \frac{1}{1 + h_0(j\omega_1)} \right| \leq \frac{1}{10}$$

medfører at vi må velge $|h_0(j\omega_1)| \geq 10 = +20$ [dB] ved $\omega_1 = 0.2$. Dette kravet er avmerket i figur 9.17 ved et hjørne som begrenser den endelige karakteristikken for $|h_0(j\omega)|$. Vi ser at karakteristikken for $10h_u(j\omega)$ ligger godt over denne spesifikasjonsverdien. Vi har derfor med denne funksjonen klart å tilfredsstille to av kravene.

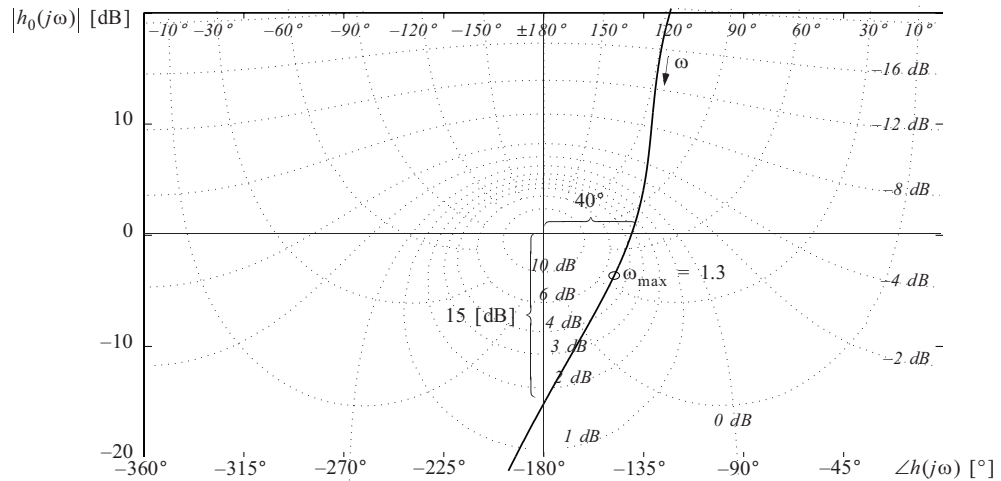
Går vi imidlertid videre i listen over spesifikasjoner, ser vi at spesifikasjon c) på langt nær tilfredsstilles av den foreløpige løsningen, fordi systemet faktisk er ustabil. Dette framkommer ved å sammenholde fasekarakteristikken $\angle h_u(j\omega)$ nederst i figur 9.17 med amplitudekarakteristikken for $10h_u(j\omega)$. Vi ser her at fasevinkelen ved systemets kryssfrekvens vil være omtrent -210° , det vil si at $\omega_{180} < \omega_{c,1}$. Noe må derfor gjøres for å oppnå tilstrekkelig forsterkningsmargin og fasemargin.

Dette kan vi få til ved å velge en asymptotisk karakteristik for $h_0(j\omega)$ som har et stykke med vinkelkoeffisient lik (-1) i frekvensområdet der den krysser 0-dB-linjen. En slik karakteristik er inntegnet øverst i figur 9.17. Den er framkommet på følgende måte:

- Vi har forutsatt at den høyfrekvente delen av $|h_0(j\omega)|$ skal falle sammen med $|10h_u(j\omega)|$. Dette er på ingen måte en generell regel, og vil avhenge av om det finnes nok effekt til disposisjon i pådraget til å drive kryssfrekvensen for det endelige systemet høyere ved aktuelle forstyrrelsesamplituder. (Dette problemet drøftes i kapittel 12.)
- Vi må la den asymptotiske karakteristikken for $|h_0(j\omega)|$ falle med en vinkelkoeffisient på minst (-2) i frekvensområdet omkring $\omega = 0.2$. Vi lar asymptoten tangere den angitte begrensingsverdien ved 20 [dB]. Knekkpunktene i den asymptotiske karakteristikken for $|h_0(j\omega)|$ ved $\omega = 0.1$ og ved $\omega = 2$ har vi lagt inn for å få en enkel transferfunksjon for regulatoren.
- Det resulterende stykket av den asymptotiske karakteristikken som har en vinkelkoeffisient lik (-1) , går fra $\omega = 0.4$ til $\omega = 2$. Den tilhørende fasekarakteristikken $\angle h_0(j\omega)$ er vist nederst i figur 9.17. Vi ser at området der $|h_0(j\omega)|$ krysser 0-dB-linjen, ved $\omega \approx 1$, har en atskillig mindre negativt faseforløp enn prosessen alene hadde. Det steile forløpet av den asymptotiske amplitudekarakteristikken omkring $\omega = 0.2$ medfører på den andre side at fasevinkelen i dette området blir mer negativ enn for prosessen alene.

Et Nicholsdiagram for $h_0(j\omega)$ er vist i figur 9.18.

Figur 9.18
Nicholsdiagram
for $h_0(j\omega)$



Vi ser at $|N(j\omega)|_{\max} \approx 5$ [dB], og denne maksimalverdien opptrer ved $\omega_{\max} \approx 1.3$. Vi merker oss også at systemet har fått en fasemargin $\psi \approx 40^\circ$ og en forsterkningsmargin $\Delta K \approx 15$ [dB].

Den resulterende sløyfetransferfunksjonen avledet av asymptotene i figur 9.17 blir

$$h_0(s) = \frac{100(1 + 2.5s)}{(1 + 25s)(1 + 10s)(1 + 0.5s)(1 + 0.2s)} \quad (9.27)$$

Benytter vi oss av (9.19) og (9.26), finner vi regulatoren

$$h_r(s) = \frac{h_0(s)}{h_u(s)} = \frac{10(1 + 2.5s)(1 + 2s)}{(1 + 25s)(1 + 0.2s)} \quad (9.28)$$

Grunnen til at vi tidligere har valgt knekkpunktene ved $\omega = 0.04$ og $\omega = 2$ ser vi nå. Dette valget har gitt en enkel form på regulatorens transferfunksjon uten at formen på $h_0(s)$ har blitt særlig ugunstig.

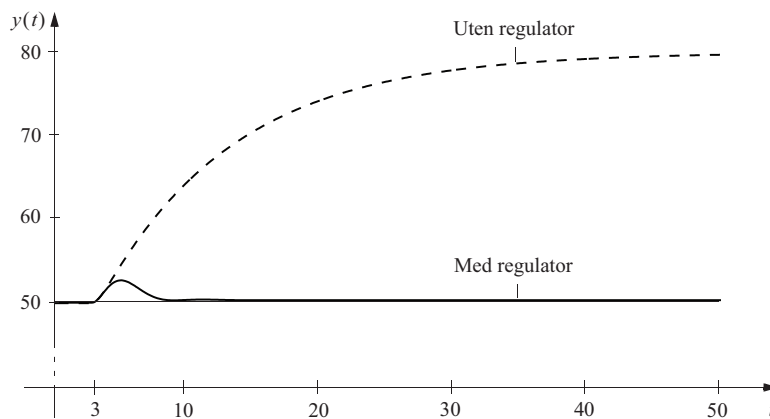
Figur 9.17 viser også asymptotiske karakteristikker for amplituden og fasen til regulatoren. Amplitudekarakteristikken regnet fra lave frekvenser har følgende sekvens av vinkelkoeffisienter: 0, -1, 0, +1, 0.

Regulatoren utfører to vesentlige funksjoner: en tilnærmet integralvirkning i området med vinkelkoeffisient lik (-1) og en tilnærmet derivatvirkning i området med vinkelkoeffisient lik (+1).

Regulatoren gir positiv fasekorleksjon ved frekvenser høyere enn $\omega = 0.45$. Dette var nødvendig for å forbedre stabilitetsforholdene. For frekvenser lavere enn $\omega = 0.45$ gir regulatoren negativ fasekorleksjon. Denne er et resultat av at amplitudekarakteristikken i dette frekvensområdet faller sterkt med økende frekvens fordi vi ønsker høy forsterkning ved lave frekvenser. Denne regulatorkarakteristikken skal vi nedenfor betegne med *proporsjonal + begrenset integral + begrenset derivat*.

Hvordan dette systemet oppfører seg i tidsplanet kan undersøkes ved simulering av blokk-diagrammet som er vist i figur 9.16. Vi antar at en sprangvis forstyrrelse v av størrelse 3 kommer inn som vist, og registrerer responsen i den målte utgangstemperaturen (y) uten og med regulering som vist i figur 9.19. Den regulatoren vi har valgt i (9.28) klarer å bringe $y(t)$ nesten tilbake til $y_0 = 50$. Feilen vil bare være ca. 0.3.

Figur 9.19
Respons på utgangen med og uten regulator ved sprang i forstyrrelsen på 3 ved tiden $t = 3$

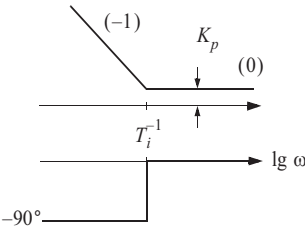
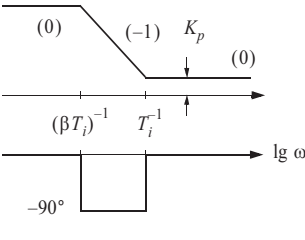
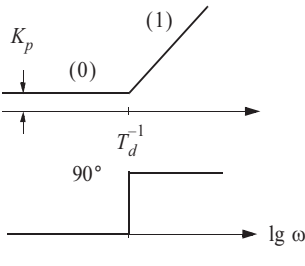
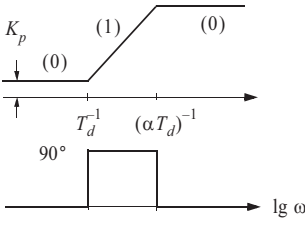
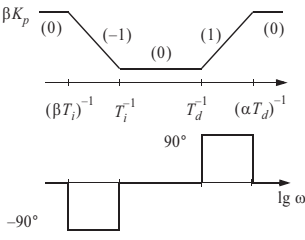


Vanlige seriekompensasjonselementer

Flere forskjellige virkninger kombineres i kompensasjonselementet som ble brukt i eksempel 9.3. Vi kan lage kompensasjonselementer med svært kompliserte transferfunksjoner, men ofte vil de vesentligste effektene kunne oppnås med relativt enkle kombinasjoner. Tabell 9.2 viser de vanligst forekommende seriekompensasjonselementene. De to første linjene inneholder elementer med *proporsjonal pluss integralvirkning*. Den første har ideell integralvirkning mens det andre har begrenset forsterkning ved lave frekvenser. De to neste linjene viser elementer med *proporsjonal pluss derivatvirkning*. Den siste av disse har begrenset forsterkning på derivatvirkningen. Til slutt vises kombinasjonen *proporsjonal pluss begrenset integral pluss begrenset derivatvirkning* som ble anvendt i eksempel 9.3¹

¹I engelskspråklig litteratur finner vi ofte kompensasjonselementet “proporsjonal + begrenset integral” betegnet med “lag-network” og “proporsjonal + begrenset derivat” med “lead-network”. Kombinasjonen av disse betegnes “lag/lead-network”.

TABELL 9.2 Vanlig forekommende regulatorer

Asymptotisk amplitude/ fase/frekvens-diagram	Transferfunksjon	Betegnelse Kode
	$K_p \frac{1 + T_i s}{T_i s}$	Proporsjonal + integral (-1, 0)
	$\beta K_p \frac{1 + T_i s}{1 + \beta T_i s}$ $\beta > 1$	Proporsjonal + begrenset integral (0, -1, 0)
	$K_p (1 + T_d s)$	Proporsjonal + derivat (0, 1)
	$K_p \frac{1 + T_d s}{1 + \alpha T_d s}$ $\alpha < 1$	Proporsjonal + begrenset derivat (0, 1, 0)
	$\beta K_p \frac{1 + T_i s}{1 + \beta T_i s} \frac{1 + T_d s}{1 + \alpha T_d s}$ $\beta > 1, \alpha < 1$	Proporsjonal + begrenset integral + begrenset derivat (0, -1, 0, 1, 0)

EKSEMPEL 9.4: Proporsjonal pluss begrenset integral-regulator for varmeveksler

I eksempel 9.3 begynte vi utformingen av seriekompensasjonselementet ved først å tilfredsstille spesifikasjonene ved lave frekvenser. I mange tilfeller vil vi gjøre omtrent det motsatte idet vi tenker oss at kompensasjonselementet antar form av en proporsjonalregulator der forsterkningen økes gradvis inntil vi når den største verdien som kan tolereres ut fra spesifikasjon til stabilitetsmarginen ($|N(j\omega)|_{\max}$). Båndbredden som derved oppnås, regnes som det resulterende systemets båndbredde. Dersom de øvrige spesifikasjonene til for eksempel verdier av $|N(j\omega)|$ ved forskjellige lavere frekvenser (særlig ved $\omega = 0$) ikke er tilfredsstilt, vil vi som oftest kunne klare oss med et seriekompensasjonselement av typen *proporsjonal pluss begrenset integral* som vist i andre linje i tabell 9.2.

Vi kan benytte samme prosess som i eksempel 9.3 med transferfunksjon gitt ved likning (9.26). Vi beholder spesifikasjonene:

- a) $|N(0)| \leq -40$ [dB]
- b) $|N(j\omega)|_{\max} \leq 6$ [dB]
- c) $\omega_c =$ båndbredde = størst mulig

Spesifikasjon b) fra eksempel 9.3 kan vi ikke tilfredsstille fordi vi allerede har bestemt oss for hvilken type kompensasjonselement vi vil bruke (proporsjonal + begrenset integral). Figurene 9.17 og 9.18 viser at med en ren proporsjonalregulator må forsterkningen i sløyfen være omtrent 18 [dB], for å oppnå en fasemargin $\psi = 45^\circ$ som omtrent svarer til $|N(j\omega)|_{\max}$. Siden prosessen allerede har $K = 10 = 20$ [dB], må kompensasjonselementet ha en forsterkning mindre enn 1, nemlig $K_p = 0.79 = -2$ [dB].

Vi har sagt at vi vil velge kompensasjonselementet

$$h_r(s) = \beta K_p \frac{1 + T_i s}{1 + \beta T_i s} \quad (9.29)$$

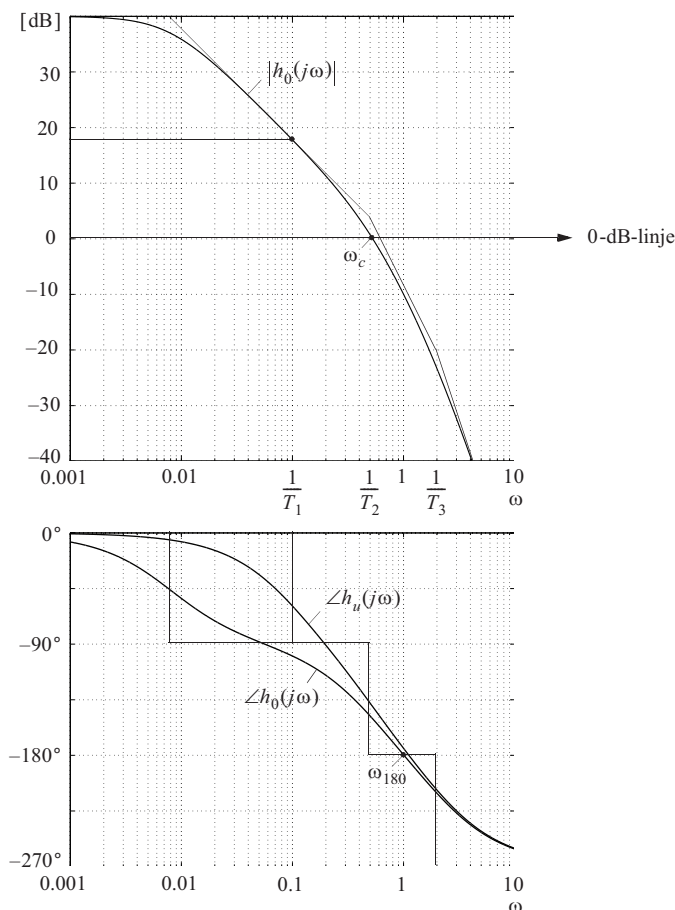
Figur 9.20 viser at $T_i = T_1 = 10$ vil være en rimelig verdi av parameteren T_i . Siden sløyfens forsterkning ved lave frekvenser skal være $|h_0(0)| = 100$, må vi velge $\beta K_p = 10$ og $\beta T_i = 126$. Dette gir

$$h_r(s) = 10 \frac{1 + 10s}{1 + 126s} \quad (9.30)$$

og

$$h_0(s) = \frac{100}{(1 + 126s)(1 + 2s)(1 + 0.5s)} \quad (9.31)$$

Figur 9.20
Amplitude/
fase/
frekvens-
diagram



Dette systemet har en båndbredde som er omtrent halvparten av den vi tidligere er kommet fram til i (9.27), og som er vist i figur 9.17. Systemet har også nå betydelig mindre evne til å undertrykke dynamiske forstyrrelser. Det nye systemet vil få $|N(j\omega)| = -20$ [dB] ved $\omega = 0.07$, mens systemet i figur 9.17 har $|N(j\omega)| = -20$ [dB] ved $\omega = 0.2$. Den vesentligste forskjellen på de to systemene er at vi i det første hadde begrenset derivatvirkning i sløyfen i tillegg til proporsjonal + begrenset integral.

Vi kunne ha fått et noe bedre resultat med hensyn til responsen ved midlere frekvenser ved for eksempel å ha valgt et kompensasjonselement som har følgende sekvens av asymptote vinkelkoeffisienter regnet fra lave frekvenser og oppover: (0, -1, -2, -1, 0). Et slikt element kan oppnås ved å kople to elementer av den type som er vist i andre linje i tabell 9.2, i serie (kaskade). Dette blir ofte gjort for å få et sterkere fall i den resulterende sløyfens asymptotiske amplitudekarakteristikk.



EKSEMPEL 9.5: PD-regulering av en prosess med to integrasjoner

Et særlig enkelt eksempel vil illustrere betydningen av bruk av proporsjonal pluss begrenset derivatvirkning for å oppnå positiv fasekompensasjon.

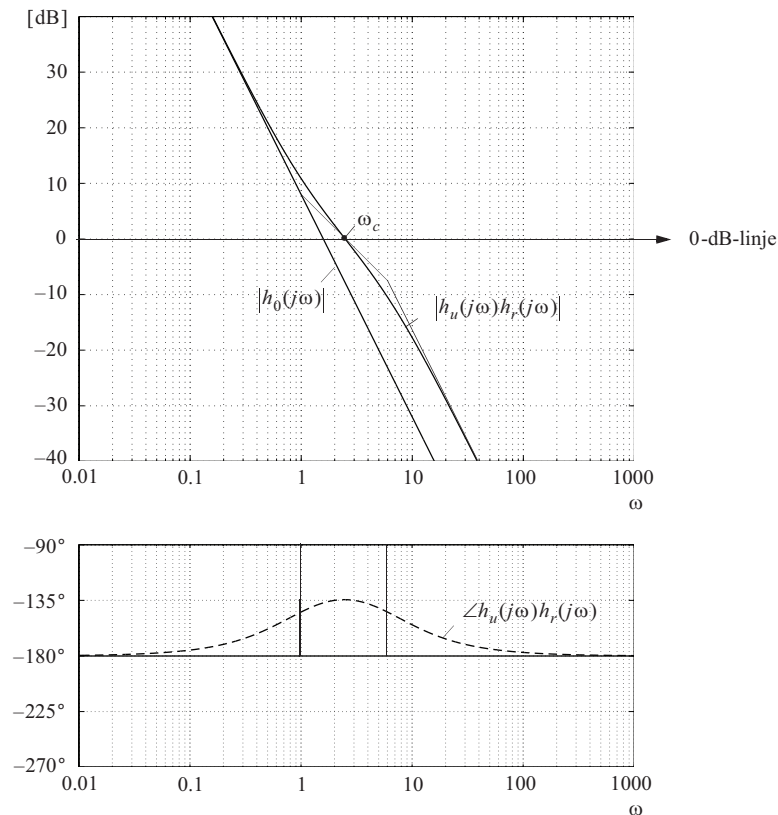
Vi antar at prosessens transferfunksjon er

$$h_u(s) = \frac{K}{s^2} \quad (9.32)$$

En slik transferfunksjon vil vi tilnærmet (ved lave hastigheter) få mellom propellskyvkraften og posisjonen av en farkost i vann. Dersom vi lager et reguleringssystem som i figur 9.10 for denne prosessen og først antar at $h_r(s) = K_p$, finner vi

$$h_0(s) = \frac{K_p K}{s^2} \quad (9.33)$$

Figur 9.21
Amplitude/fase/frekvens-diagram for prosessen



Dette tilbakekoblede systemet er spesielt fordi det for alle verdier K_p og K vil være akkurat på grensen av ustabilitet, det vil si at dets poler ligger på den imaginære akse fordi $\angle h_0(j\omega) = -180^\circ$ ved alle frekvenser (null fasemargin). Figur 9.21 viser amplitude/fase/frekvens-diagram for tilfellet i (9.33).

Skal vi oppnå positiv fasemargin, må vi innføre et kompensasjonselement som gir positivt fasebidrag i et visst frekvensområde. Dette er antydnet stiplet i figur 9.21. Regulatoren som kan utføre denne formen for kompensasjon, er av typen *proporsjonal pluss begrenset derivatvirkning* og er angitt i fjerde linje i tabell 9.2. Transferfunksjonen for et slikt element vil bli

$$h_r(s) = K_p \frac{1 + T_d s}{1 + \alpha T_d s} \quad (9.34)$$

Figur 9.21 viser at den positive fasekorleksjonen som dette elementet produserer, når sitt maksimum ved en frekvens som er den **geometriske middelveiden** av de to knekkfrekvensene,

$$\omega = \sqrt{\omega_1 \cdot \omega_2} = \sqrt{\frac{1}{T_d} \cdot \frac{1}{\alpha T_d}} = \sqrt{\frac{1}{\alpha T_d^2}} = \frac{1}{T_d} \sqrt{\frac{1}{\alpha}}$$

Den **maksimale fasekorleksjonen** $\angle h_r(j\omega)_{\max}$ er avhengig av frekvensavstanden mellom knekkfrekvensene, altså av forholdet

$$\frac{T_d}{\alpha T_d} = \frac{1}{\alpha}$$

Denne maksimale fasekorleksjonen er framstilt grafisk i figur 9.22. Figuren viser at dersom vi for eksempel ønsker å korrigere fasen slik at fasemarginen $\psi = 45^\circ = \angle h_r(j\omega)_{\max}$, må kompensasjonselementet ha

$$\frac{T_d}{\alpha T_d} = \frac{1}{\alpha} \approx 6$$

Ønsker vi også å spesifisere hvilken verdi kryssfrequensen ω_c skal ha og at den maksimale fasekorleksjonen skal opptre ved denne frekvensen (hvilket gir omtrent de gunstigste forholdene for $|N(j\omega)|_{\max}$), kan det vises at

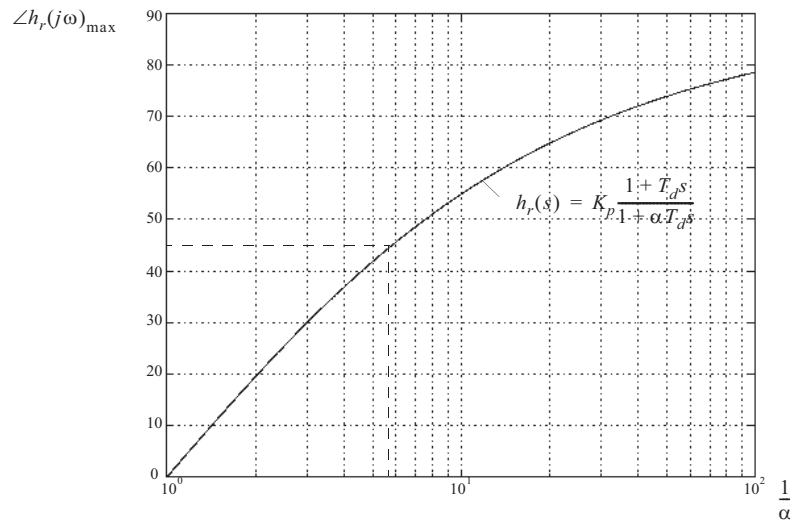
$$K_p K = \frac{\omega_c^2}{\sqrt{\frac{T_d}{\alpha T_d}}} = \omega_c^2 \sqrt{\alpha} \quad (9.35)$$

$$T_d = \sqrt{\frac{T_d}{\alpha T_d}} \frac{1}{\omega_c} = \frac{1}{\omega_c \sqrt{\alpha}} \quad (9.36)$$

$$\alpha T_d = \frac{1}{\sqrt{\frac{T_d}{\alpha T_d}}} \frac{1}{\omega_c} = \frac{\sqrt{\alpha}}{\omega_c} \quad (9.37)$$

Ovenfor har vi tenkt oss at systemets båndbredde ω_c skulle være spesifisert, og vi har påpekt at systemets evne til å reagere ved høye frekvenser er avhengig av at det finnes tilstrekkelig effekt tilgjengelig til at systemets komponenter faktisk kan reagere med ønskede amplituder ved disse frekvensene. En mer detaljert undersøkelse av slike begrensende forhold vil bli gitt i kapittel 12.

Figur 9.22
Maksimal fasekorreksjon som funksjon av α^{-1}



EKSEMPEL 9.6: Integralregulering av en ren transportforsinkelse

Dersom prosessen er dominert av en transportforsinkelse, slik at prosessens transferfunksjonen tilnærmet er gitt ved

$$h_u(s) = K e^{-\tau s} \quad (9.38)$$

er det liten mulighet til å få korrigert den sterkt negative fasevinkelen

$$\angle h_u(j\omega) = -\omega\tau \quad (9.39)$$

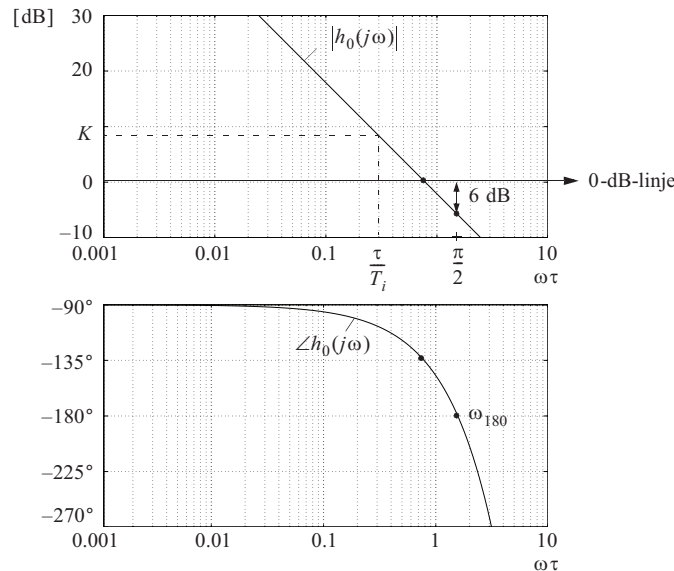
ved hjelp av begrenset derivatvirkning som gjort foran. Grunnen er at derivatvirkningen vil gi stigende amplitudekarakteristikk for hele sløyfen fordi prosessen selv bidrar med konstant forsterkning ved alle frekvenser (se likning (9.38)). En *proporsjonal + derivatregulator* vil derfor gi dårligere stabilitetsforhold enn en ren proporsjonalregulator som kan ha maksimal forsterkning på 1, før stabilitetsgrensen nås. Ønsker vi derfor i dette tilfellet et regulerings-system som har stor statisk nøyaktighet, for eksempel $|N(0)| = 0$, er det best å velge en *ren integralregulator*

$$h_r(s) = \frac{1}{T_i s} \quad (9.40)$$

Figur 9.23 viser amplitude/fase/frekvens-diagram for

$$h_0(s) = \frac{K}{T_i s} e^{-\tau s} \quad (9.41)$$

Figur 9.23
AFF-diagram for
 $h_0(s)$



Siden en ren integralregulator bidrar med vinkelen -90° ved alle frekvenser og vinkelbidraget fra prosessen er gitt av (9.39), finner vi at

$$\angle h_0(j\omega) = -\frac{\pi}{2} - \omega\tau \quad (9.42)$$

Det knytter seg særlig interesse til frekvensen ω_{180} der $\angle h_0(j\omega) = -\pi = -180^\circ$. Vi finner av (9.42)

$$\omega_{180} = \frac{\pi}{2\tau} \quad (9.43)$$

Ønsker vi å dimensjonere systemet ut fra fasemargin $\psi \geq 45^\circ$ og forsterkningsmargin $\Delta K \geq 6$ dB, finner vi av figur 9.23 å måtte velge

$$K \left(\frac{\tau}{T_i} \right) = \frac{\pi}{4} \text{ dvs. } T_i = \frac{4\tau K}{\pi} \quad (9.44)$$

Stabilitetsgrensen nås når

$$K \left(\frac{\tau}{T_i} \right) = \frac{\pi}{2} \quad (9.45)$$

hvilket stemmer med det som er funnet i (8.43) i eksempel 8.11.



9.3.1 Seriekompensasjon ved hjelp av standardregulatorer for prosessregulering

I det foregående har vi til stadighet henvist til visse standardregulatorer som brukes ved prosessregulering. Disse kan betraktes som seriekompensasjonselementer. De er hensiktsmessige ved oppbygging av prosessreguleringssystemer fordi de har parametre som kan innstilles innenfor vide grenser, og de gir som oftest et tilfredsstillende resultat. Standardregulatorene er ikke alltid fysiske enheter, men kan være algoritmer (beregningsregler, transferfunksjoner) som realiseres i en datamaskin. De omtalte seriekompensasjonselementene forekommer vanligvis i fire forskjellige former:

- a) **Proporsjonalregulatorer (P-regulator)** som bare er en forsterker med innstillbar forsterkning.

Den ideelle P-regulatoren virker ikke inn på sløvfens fase.

På grunn av den fysiske oppbyggingen av en slik forsterkning (gjerne innbefattet effektorganet som utgjør pådraget), vil vi bare ha konstant forsterkning innenfor et begrenset frekvensbånd, idet uunngåelige poler vil gjøre at forsterkningen synker ved økende frekvens. Dette frekvensbåndet vil imidlertid ofte være vesentlig større enn båndbredden til det ferdige reguleringssystemet, slik at P-regulatoren ofte kan regnes å ha konstant forsterkning til frekvenser høyere enn kryssfrekvensen ω_c .

- b) Proporsjonal + integral-regulator (**PI-regulator**) som ideelt sett skal ha en karakteristikk som vist i første linje i tabell 9.2.

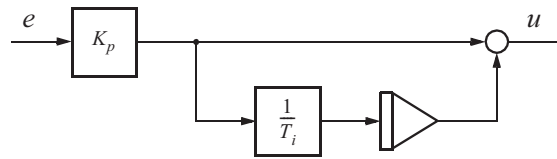
Hensikten med I-virkningen er å gi uendelig forsterkning ved $\omega = 0$ for å oppnå null stasjonært avvik.

PI-regulatoren gir sløyfens fase et negativt bidrag (nær -90° ved lave frekvenser).

Ved høye frekvenser er PI-regulatoren en P-regulator.

Figur 9.24 viser et blokkdiagram for PI-regulatoren.

Figur 9.24
PI-regulator



Transferfunksjonen for en ideell PI-regulator har formen

$$h_r(s) = K_p \left(1 + \frac{1}{T_i s} \right) = K_p \frac{1 + T_i s}{T_i s} \quad (9.46)$$

der K_p = proporsjonalkonstant og T_i = integraltid. Enkelte ganger ønsker vi ikke en ideell integrasjon, men heller en karakteristikk som vist i andre linje i tabell 9.2, som er en proporsjonal + begrenset integral-regulator.

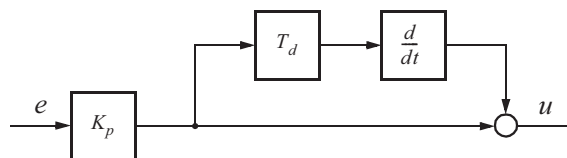
- c) En proporsjonal + derivat-regulator (**PD-regulator**) skal ideelt sett ha en karakteristikk som vist i tredje linje i tabell 9.2. I praksis blir karakteristikken gjerne som vist i fjerde linje i tabellen.

PD-regulatoren gir sløyfens fase et løft ved høye frekvenser. Båndbredden ω_c for den lukkede sløyfen kan derved bli noe høyere enn i en P-regulert sløyfe. Høy båndbredde gir hurtig regulering.

Ved lave frekvenser er PD-regulatoren en P-regulator.

Figur 9.25 viser et blokkdiagram for en PD-regulator.

Figur 9.25
PD-regulator



PD-regulatoren i figuren har transferfunksjonen

$$h_r(s) = K_p(1 + T_d s) \quad (9.47)$$

der T_d = derivattid.

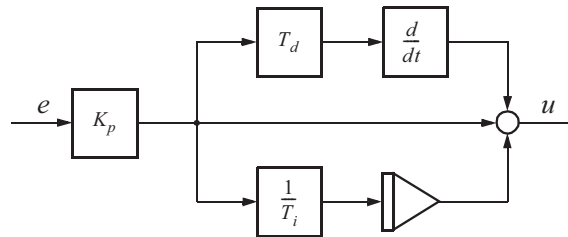
Derivatvirkningen vil som oftest gjøres begrenset som vist i fjerde linje i tabell 9.2, fordi for stor forsterkning ved høye frekvenser vil generere uønsket støy.

- d) Den ideelle proporsjonal + integral + derivat-regulator (**PID-regulator**) har en karakteristikk som er en kombinasjon av første og tredje linje i tabell 9.2. Proporsjonal + begrenset integral + begrenset derivat-regulatoren som er vist i femte linje i tabellen, er en implementasjon som er benyttet i mange praktiske systemer.

Med PID-regulator kan vi ofte oppnå både lite stasjonært avvik og høy båndbredde.

Figur 9.26 viser et blokkdiagram for en ideell PID-regulator.

Figur 9.26
Ideell PID-regulator



Transferfunksjonen for en ideell PID-regulator skrives gjerne slik

$$\begin{aligned} h_r(s) &= K_p \left(1 + \frac{1}{T_i s} + T_d s \right) = K_p \frac{1 + T_i s + T_i T_d s^2}{T_i s} \\ &\approx K_p \frac{(1 + T_i s)(1 + T_d s)}{T_i s} \end{aligned} \quad (9.48)$$

Tilnærmelsen i (9.48) er gyldig når $T_d \ll T_i$. Denne regulatortypen er meget populær fordi den omfatter alle funksjonene vi vanligvis ønsker, og fordi de tre parametrene (K_p , T_i og T_d) kan justeres innenfor vide grenser.

Ved bruk av slike regulatorer i prosessregulering må vi justere inn de forskjellige parametrene for å tilfredsstille spesifikasjonene som er gitt. Hvorvidt vi bruker en P-regulator eller en av de mer sammensatte regulatorene, avhenger av hvor strenge kravene er og hvor mye forsinkelse (poler eller transportforsinkelse) det finnes i prosessen.

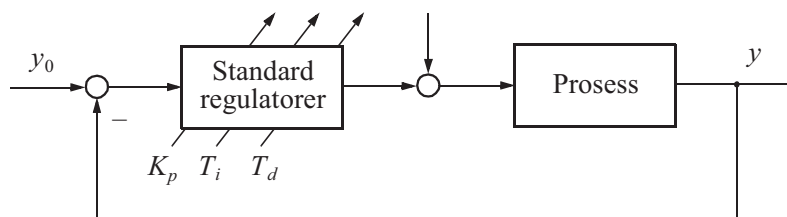
Se også appendiks E som inneholder en fabel som illustrerer den fysiske påvirkningen standardregulatorene har på en enkel prosess.

9.3.2 Eksperimentell innstilling av standardregulatorer for prosessregulering – Ziegler-Nichols metode

“**Ziegler-Nichols metode**” (Ziegler and Nichols 1942) er en enkel metode som kan brukes til eksperimentell innstilling av regulatorer, dvs. uten at man er avhengig av en prosessmodell. Mens prosessen er i drift stiller man (forsiktig) på regulatoren etter en prosedyre beskrevet nedenfor. Metoden forutsetter (sjølsagt) at reguleringsystemet i utgangspunktet er stabilt. Den forutsetter også at prosessens frekvensrespons har amplitude- og fasekarakteristikker som begge er monotont fallende ved økende frekvens.

Siden metoden baserer seg på systematiske eksperimenter utført på det ferdig monterte systemet der alle de virkelige komponentene inngår, vil de data vi får være ganske realistiske. (En metode som derimot baserer seg på individuell bestemmelse av de forskjellige systemkomponentenes dynamiske egenskaper, vil derimot lett kunne føre til at noe overses eller at gjensidige påvirkninger ikke kommer fram.)

Figur 9.27
Ziegler-Nichols
metode



Ziegler-Nichols metode kan i korthet beskrives som følger: Først settes $T_i \approx \infty$ og $T_d = 0$, da er regulatoren en ren P-regulator. Forsterkningen K_p økes gradvis, samtidig som små forstyrrelser påtrykkes prosessen inntil man oppnår en stående svingning (eller i det minste en svakt dempet svingning) på prosessens utgang (y). Vi noterer denne verdien av proporsjonalkonstanten som kalles **kritisk forsterkning**, K_{pk} . Likeledes noteres den stående svingningens periode

$$T_k = \frac{1}{f_k} = \frac{2\pi}{\omega_{180}} \quad (9.49)$$

Det som i realiteten skjer, er at systemet selv genererer en svingning, og det foretas en frekvensanalyse av den lukkede sløyfen i nærheten av frekvensen ω_{180} .

P-regulator: En gunstig innstilling hvis vi kan nøye oss med P-regulator vil da være

$$K_p = 0.5K_{pk} = K_{pk} [\text{dB}] - 6 [\text{dB}] \quad (9.50)$$

Den valgte innstillingen tilsvarer en forsterkningsmargin $\Delta K = 6 [\text{dB}]$. Denne innstillingen vil som oftest være akseptabel når systemet har relativt sterkt økende negativ fasevinkel, fordi forsterkningsmarginen da vil være utslagsgivende.

PI-regulator: En gunstig parameterkombinasjon vil være

$$K_p = 0.45K_{pk} = K_{pk} [\text{dB}] - 7 [\text{dB}] \quad (9.51)$$

$$T_i = 0.85T_k \approx 5 \frac{1}{\omega_{180}} \quad (9.52)$$

K_p for PI-regulator velges nå lavere enn for P-regulator p.g.a. det negative fasebidraget til integraldelen.

PID-regulator: Her velger vi

$$K_p = 0.6K_{pk} = K_{pk} [\text{dB}] - 4.5 [\text{dB}] \quad (9.53)$$

$$T_i = 0.5T_k \approx 3 \frac{1}{\omega_{180}} \quad (9.54)$$

$$T_d = 0.12T_k \approx 0.75 \frac{1}{\omega_{180}} \quad (9.55)$$

K_p kan nå velges litt større enn i (9.50) p.g.a. faseløftet som gis av derivatdelen. Vi legger merke til at ved PI-regulator vil integralvirkningen komme inn ved frekvenser lavere enn $0.2\omega_{180}$. Ved PID-regulator vil integralvirkningen komme inn ved frekvenser lavere enn $0.3\omega_{180}$, mens derivatvirkningen først kommer inn ved frekvenser høyere enn $1.3\omega_{180}$.

Det presiseres at denne eksperimentelle innstillingsteknikken ikke er universell, men den kan, eventuelt med en viss justering av de numeriske verdier som metoden

gir, benyttes i mange praktiske tilfeller. Vi oppsummer Ziegler-Nichols' regler i tabellform:

Regulator	K_p	T_i	T_d
P	$0.5K_{pk}$	∞	0
PI	$0.45K_{pk}$	$0.85T_k$	0
PID	$0.6K_{pk}$	$0.5T_k$	$0.12T_k$

Et naturlig spørsmål er nå: Hvorfor svinger systemet med ω_{180} med Ziegler-Nichols?

Ved å sette $h_r = K_p$ og øke K_p til K_{pk} , kommer systemet i en stående svingning med periodetid T_k . Kall svingefrekvensen i rad/tidsenhet ω_k . Da er

$$\omega_k = \frac{2\pi}{T_k} \Leftrightarrow T_k = \frac{2\pi}{\omega_k}$$

Vi skal vise at $\omega_k = \omega_{180}$:

Når det lukkede system er på stabilitetsgrensa, er $\omega_c = \omega_{180}$, dvs.

$$\angle h_0(\pm j\omega_c) = -180^\circ \Rightarrow h_0(\pm j\omega_{180}) = -1$$

Da er $1 + h_0(\pm j\omega_{180}) = 0$. Siden polene i det lukkede system kan finnes som røttene i nevneren til $N(s) = \frac{1}{1 + h_0(s)}$, må $\pm j\omega_{180}$ være to av polene i det lukkede system.

Den stående svingningen er gitt av disse to polene, for eventuelle øvrige poler er i v.h.p. Ergo er $\omega_k = \omega_{180}$.

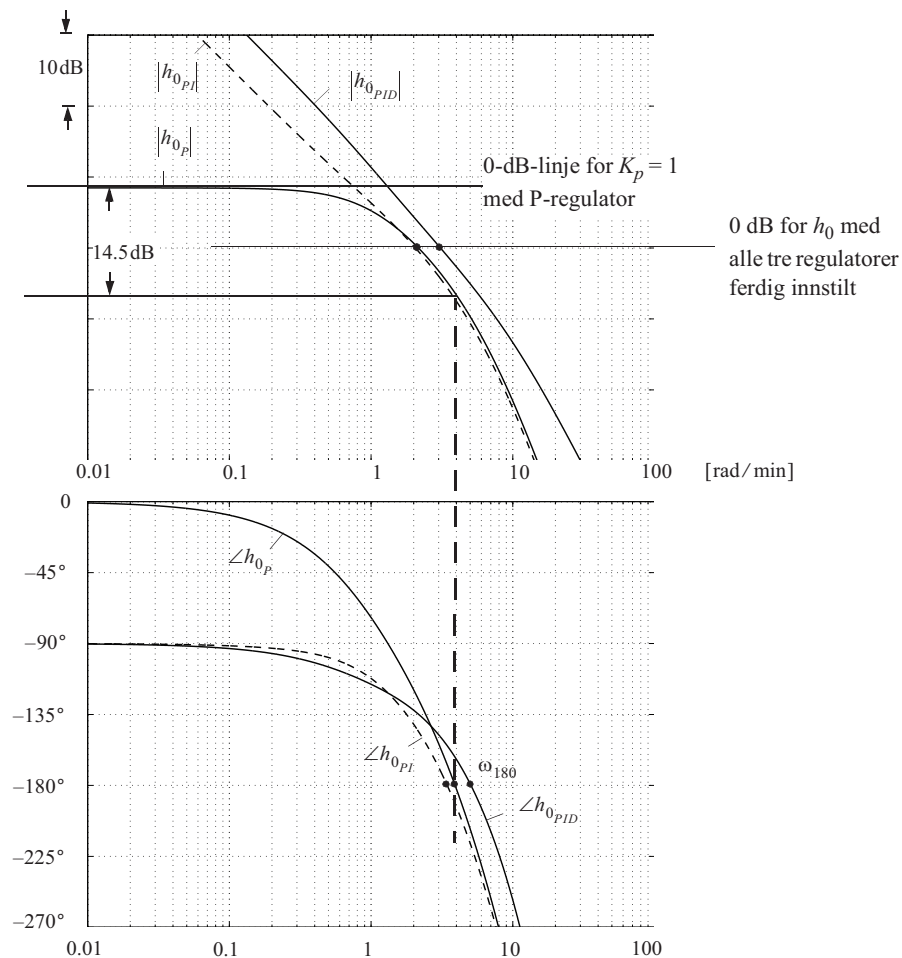
EKSEMPEL 9.7: Bruk av Ziegler-Nichols metode

Ziegler-Nichols metode er som forklart primært beregnet på å justere inn regulatoren for en fysisk prosess som man ikke har en modell av. Men den kan også brukes på en matematisk modell av en prosess: Gitt en prosess med transferfunksjonen

$$h_u(s) = \frac{Ke^{-\tau s}}{(1 + T_1 s)(1 + T_2 s)(1 + T_3 s)} \quad (9.56)$$

der $K = 1$, $T_1 = 0.1$ [min], $T_2 = 1.0$ [min] og $T_3 = \tau = 0.2$ [min].

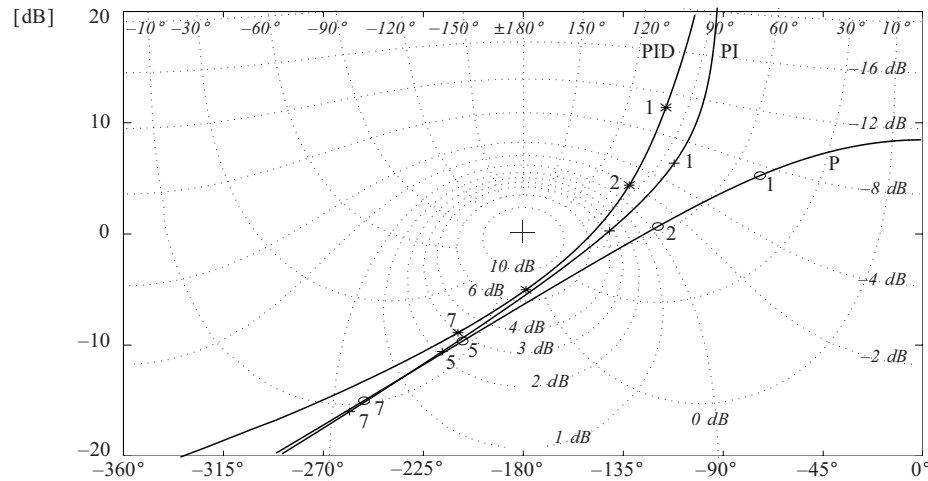
Figur 9.28
AFF-diagram for prosessen med P-regulator, PI-regulator og PID-regulator



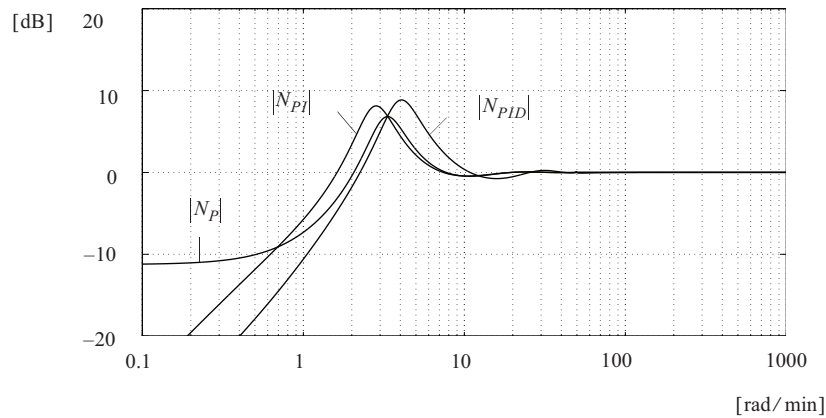
Vi kan nå bestemme K_{pk} og ω_{180} fra Bodediagrammet i figur 9.28, som viser amplitude- og fasediagrammer for reguleringsløyfen med de tre typene regulatorer innstilt etter Ziegler-Nichols metode. Figuren viser at $\omega_{180} = 4$ [rad/min]. Dette tilsvarer at systemet ville svinge med en periode $T \approx 1.6$ [min].

Vi leser også av $K_{pk} \approx 14.5$ [dB]. De resulterende frekvensresponsdiagrammer når regulatorparametre innstilles på grunnlag av K_{pk} og T_k , er overført til Nicholsdiagram i figur 9.29. De resulterende $|N(j\omega)|$ -diagrammer er vist i figur 9.30.

Figur 9.29
Nicholsdiagram for
prosess med P-
regulator, PI-regu-
lator og PID-
regulator



Figur 9.30
 $|N(j\omega)|$ -
diagram



Følgende resultater oppnås:

For P-regulator:

$$K_p = 8.5 \text{ [dB]} = 2.65$$

$$|N(0)| = -11.3 \text{ [dB]}, |N(j\omega)|_{\max} = 8 \text{ [dB]}$$

For PI-regulator:

$$K_p = 7.5 \text{ [dB]} = 2.35$$

$$T_i = 5 \cdot \frac{1}{4} = 1.25 \text{ [min]}$$

$$|N(0)| = -\infty \text{ [dB]} = 0, |N(j\omega)|_{\max} = 9 \text{ [dB]}$$

For PID-regulator:

$$K_p = 10 \text{ [dB]} = 3.16$$

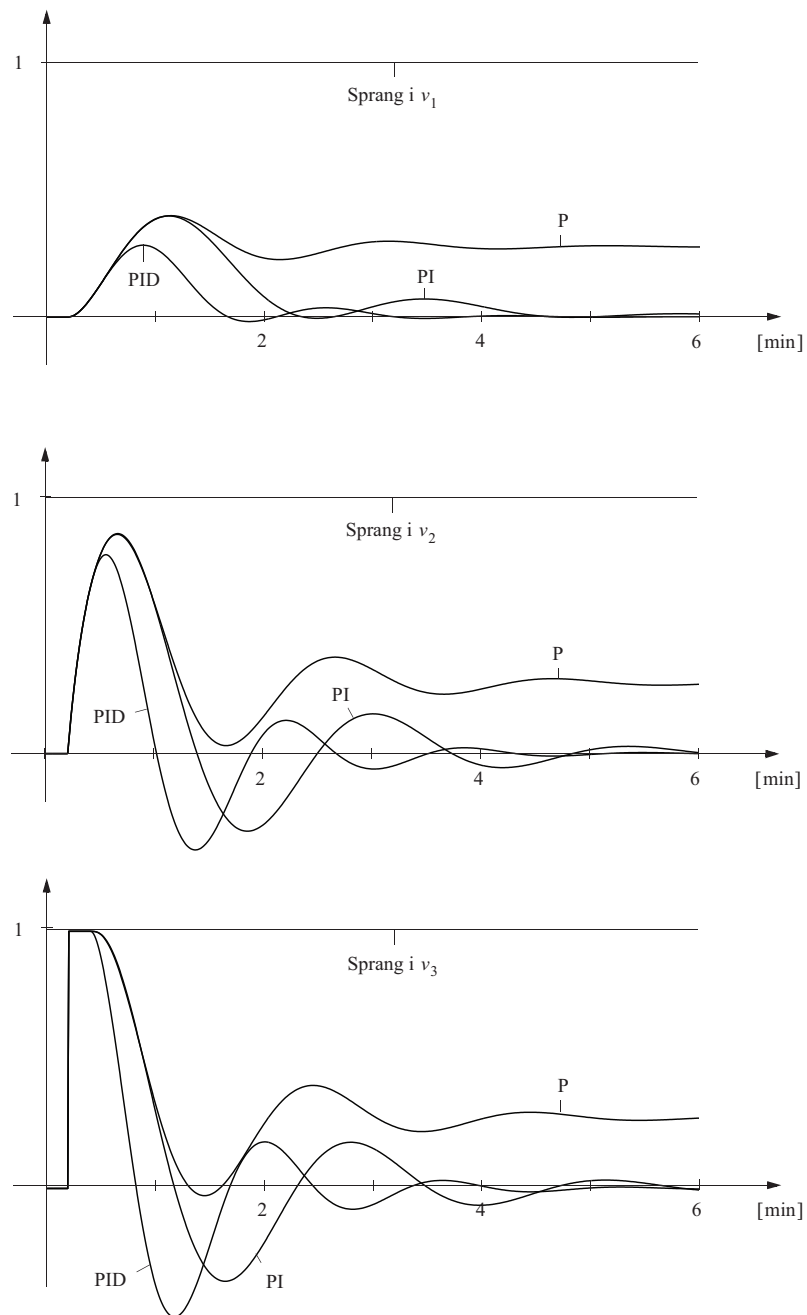
$$T_i = 3 \cdot \frac{1}{4} = 0.75 \text{ [min]}$$

$$T_d = 0.75 \cdot \frac{1}{4} = 0.19 \text{ [min]}$$

$$|N(0)| = -\infty \text{ [dB]} = 0, |N(j\omega)|_{\max} = 9 \text{ [dB]}$$

For å kontrollere hvordan dette systemet oppfører seg i tidsplanet, er det gjennomført en simulering der det antas at prosessens transferfunksjon framkommer som et resultat av fire seriekoblede dynamiske elementer, som vist i figur 9.32. En tenker seg også at prosessen kan utsettes for sprangvise forstyrrelser, som kan angripe på tre forskjellige steder. Disse betegnes med v_1 , v_2 og v_3 . Prosessens utgang (målingen) y registreres i hvert tilfelle for de tre regulatortypene. Resultatet er vist i figur 9.31.

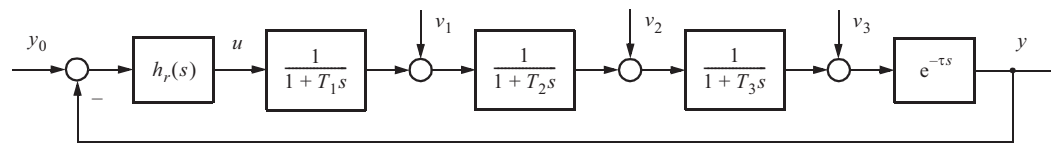
Figur 9.31
Respons ved
sprangvise forstyrrelser på tre
forskjellige steder



Responsten avhenger svært av hvor forstyrrelsen angriper. Når forstyrrelsen angriper “tidlig” i prosessen (det vil si nær pådraget), vil den arte seg som en “myk” forstyrrelse sett fra målingen. Uansett forstyrrelsens angrepssted gir P-regulatoren et stasjonært avvik mens både PI-regulatoren og PID-regulatoren får avviket til å gå mot null. PID-regulatoren gir den hurtigste responsten i alle tilfellene.

Man kan diskutere hvorvidt de valgte innstillingene er de beste med den foreliggende strukturen. I enkelte tilfeller vil vi nok foretrekke et system som er litt mindre oscillatorisk. Det svarer til innstilling av systemet slik at $|N(j\omega)|_{\max}$ blir litt lavere, for eksempel ca. 6 [dB] som tidligere antydnet.

Figur 9.32
Prosess med
regulator og
tilbakekopling



9.3.3 Eksperimentell innstilling av PI-regulator – SIMC-metoden

Her følger en oppskrift på åpen-sløyfe-tuning av PI-regulator. Metoden som her presenteres er SIMC (SIMC – *Skogestad's Internal Model Control tuning rules*). Metoden baseres på følgende vis:

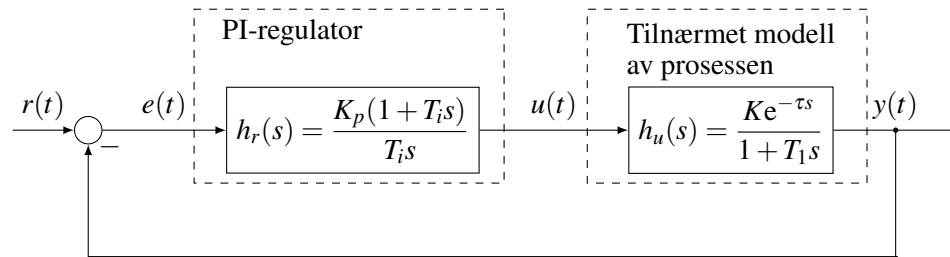
1. Ta utgangspunkt i en enkel sprangrespons i åpen sløyfe.
2. Tilpass en 1.ordens modell med tidsforsinkelse til responsten.
3. Spesifiser ønsket lukket sløyfe sprangrespons.
4. Beregn PI-parametrene fra en enkel formel

$$K_p = \frac{T_1}{K} \frac{1}{(\tau + T_L)} \quad (9.57)$$

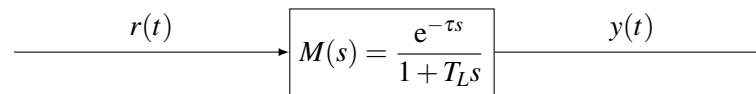
$$T_i = \min(T_1, 4(\tau + T_L))$$

der T_L er ønsket tidskonstant for lukket system fra r til y , se figur 9.34.

Figur 9.33
PI-
reguleringsløyfe

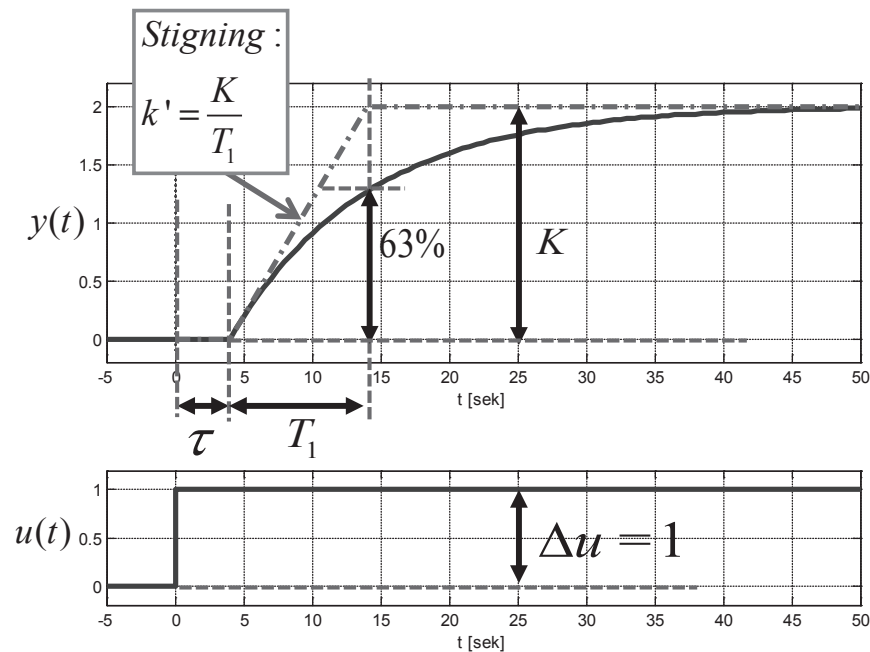


Figur 9.34
Ønsket lukket
sløyferespons.
Vi har valgt T_L .



Velger her ønsket lukket sløyferespons på tilsvarende form som $h_u(s)$. Vi kommer ikke utenom tidsforsinkelsen τ , men kan velge en ønsket tidsrespons gitt ved T_L .

Figur 9.35
Sprangresponsen
for 1. ordens modell
med tidsforsinkelse.



Figur 9.36 viser sprangresponsen for 1. ordens modell med tidsforsinkelse. Følgende parametre i prosessmodellen

$$h_u(s) = \frac{K e^{-\tau s}}{1 + T_1 s}$$

er brukt: $K = 2$, $T_1 = 10$, $\tau = 4$ og $\Delta u = 1$.

Steg 1: Tilpass den enkle 1.ordens modellen til en virkelig sprangrespons

1. Anslå først¹ stigningen like etter tidsforsinkelsen:

$$k' = \frac{dy/dt}{\Delta u} = \frac{K}{T_1}$$

Plott gjerne responsen på papir og bruk linjal og trekk tangenten langs den bratteste delen av responsen etter $y(t)$ begynner å stige.

2. Finn tidsforsinkelsen τ ut fra hvor denne tangenten skjærer tidsaksen ved $y(0)$.
3. Gjør et anslag på tidskonstanten T_1 (der $y(t)$ passerer 63% av fullt utslag eller der asymptoten (tangenten) når fullt utslag i $y(t)$). Om man vil finne K , bruk $K = k'T_1$

PS: Vær obs på å ta med rett størrelse på Δu når man anslår k' eller K fra $y(t)$!

Steg 2: Bruk følgende PI-parametre

$$K_p = \frac{T_1}{K} \frac{1}{(\tau + T_L)} = \frac{1}{k'} \frac{1}{(\tau + T_L)} \quad (9.58)$$

$$T_i = \min(T_1, 4(\tau + T_L))$$

For moderate tidskonstanter T_1 ser vi at integraltiden T_i velges slik at nullpunktet i PI-regulatoren kansellerer polen i prosessen, altså

$$T_i = T_1$$

Om valg av ønsket lukket sløyferespons

Juster aggressiviteten til regulatoren med valget av lukket sløyfe tidskonstant: T_L

$T_L = \tau$	Normalt rask innstilling, fasemargin ca 60°
$T_L = 0.3\tau$	Aggressiv innstilling, fasemargin ca 45°
$T_L > 2\tau$	Moderat/langsom innstilling, fasemargin $> 60^\circ$

For $T_L > 2\tau$ får vi ikke oversving, og lukket sløyferespons stemmer bra med

$$M(s) = \frac{e^{-\tau s}}{1 + T_L s}$$

¹Hvis man når en tydelig stasjonærverdi kan man selvsagt finne K først og trenger ikke k'

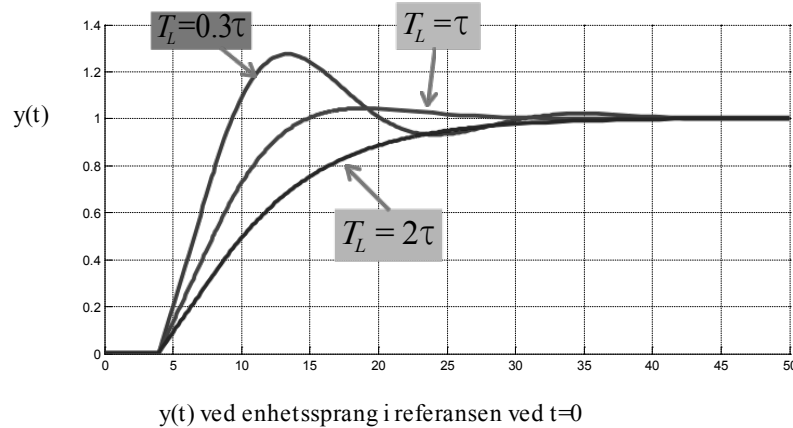
$$K = \frac{\Delta y}{\Delta u} = \frac{y(\infty) - y(0)}{u(\infty) - u(0)}$$

For lavere T_L vil vi få mer oversving etter hvor aggressiv (lav) verdi som velges.

Steg 3: Testing i lukket sløyfe

- Test om regulatoren fungerer som forventet
- Prøv ut noen verdier av T_L
- Hvis det ikke fungerer som forventet, se på modelltilpasningen igjen

Figur 9.36
Lukket sløyfe-
respons for
ulike T_L/τ .



Spesialtilfelle: Integrerende prosess med tidsforsinkelse

Anta først

$$h_u(s) = \frac{K e^{-\tau s}}{1 + T_1 s} = \frac{K}{T_1} \frac{e^{-\tau s}}{\frac{1}{T_1} + s} = k' \frac{e^{-\tau s}}{\frac{1}{T_1} + s}$$

Anta så at T_1 øker mens vi beholder samme stigningsrate k' , da får vi følgende prosess:

$$T_1 \rightarrow \infty \Rightarrow h_u(s) = k' \frac{e^{-\tau s}}{s}$$

SIMC gir da enkelt

$$K_p = \frac{1}{k'} \frac{1}{(\tau + T_L)}$$

$$T_i = 4(\tau + T_L)$$

Når tidskonstanten i åpen sløyfe er stor vil responsen i praksis ligne mer på en integrerende prosess, og følgende minimum seleksjon er derfor innført:

$$T_i = \min(T_1, 4(\tau + T_L)) \quad (9.59)$$

For integrerende prosesser velger man av og til enda lengre integraltid, eller en ren P-regulator for å unngå langsomme oscillasjoner på grunn av to integratorer i sløyfen.

Merk: En PI-regulator går mot en P-regulator i praksis når T_i velges stor ($T_i \gg 4(\tau + T_L)$).

Legg merke til at K_p er omvendt proporsjonal med åpen sløyfe forsterkning K eller stigning k' . K_p reduseres jo lengre tidsforsinkelse vi har eller jo langsommere vi ønsker responsen. Når $T_i = T_1$ i PI-regulatoren får vi følgende enkle sløyfetransferfunksjon

$$h_0(s) = h_r(s)h_u(s) = \left(\frac{T_1}{K} \frac{1}{(\tau + T_L)} \frac{(1 + T_1 s)}{T_1 s} \right) \left(\frac{K e^{-\tau s}}{(1 + T_1 s)s} \right) = \frac{e^{-\tau s}}{(\tau + T_L)s}$$

SIMC antar at man kan finne en tilpasning til en modell som er av 1. orden + tidsforsinkelse. Kan en da anta at en slik modell vil kunne representere alle prosesser? Generelt er svaret nei, men det er veldig mange prosesser der denne tilnærmingen vil kunne fungere. Nøkkelen er at det er fase og forsterkning for sløyfetransferfunksjonen i området der kryssfrekvensen kommer til å ligge som bør stemme med den egentlige prosessen. Avvik for lave og høye frekvenser har mindre betydning. Dermed kan vi si at denne enkle metoden vil kunne fungere for mange prosesser, men det er noen prosesser der forenklingen blir for grov.

Når vil da SIMC-metoden få problemer? Noen eksempler er

- Oscillatorisk åpen sløyfe prosess kan være vanskelig å håndtere.
- Langsomme nullpunkter (i venstre halvplan) kan gi til dels spesielle responser med stort oversving. Det kan da være utfordrende å anslå parametre i en 1. ordens modell.
- Ustabile prosesser har ikke fornuftig sprangrespons, og metoden kan ikke brukes.
- Høyere ordens prosesser kan noen ganger kreve mer avansert regulering enn PI.
- Fenomener som ulineariteter kan gjøre det vanskelig å tune enkle PI-regulatorer, for både settpunkt og forstyrrelser kan påvirke en linearisert modell.
- I multivariable systemer kan interaksjoner i prosessen gjøre at innstillinger i andre regulatorsløyfer påvirker prosessmodellen som hver enkelt regulator ser.

9.3.4 Standardregulatorer for et utvalg av prosesser. Parametervalg.

Tabell 9.3 gir en summarisk oversikt over rimelige parametervalg for forskjellige standardregulatorer brukt sammen med noen vanlige prosesstransferfunksjoner. Valgene må oppfattes som rimelige forslag. Andre regulatorparametre kan også gi gode resultater. Ved utledningene er følgende prinsipper brukt:

- Relativ demping i sløyfen $\zeta \approx 0.5$
- Forsterkningsmargin $\Delta K \approx 2 \approx 6$ [dB]
- Fasemargin $\psi \approx 45^\circ$
- Ziegler-Nichols metode

For noen av regulatorene er det anført $\omega_c = \omega_{bb} = \text{\textcircled{X}}$ = ønsket båndbredde. Dette skyldes at de oppgitte dataene ikke er tilstrekkelige til å bestemme båndbredden (kryssfrekvensen) ω_c . Andre spesifikasjoner, for eksempel effektbetraktninger (se kapittel 12), må til for å finne svaret.

De stedene i tabellen som inneholder X, er uaktuelle.

TABELL 9.3 Oversikt over passende regulatorparametre

		P	PD ⁽¹⁾
	$h_u(s)$	$h_r(s) = K_p$	$h_r(s) = K_p(1 + T_d s)$
1.1	$\frac{K}{s}$	$K_p = \frac{1}{K} \omega_c$ $\omega_c = \text{øbb}$	X
1.2	$\frac{K}{1 + Ts}$	$K_p = \frac{1}{K} T \omega_c$ $\omega_c = \text{øbb}$	X
2.1	$\frac{K}{s^2}$	X	$K_p = \frac{1}{K} \omega_c$ $T_d = \frac{1}{\omega_c}$ $\omega_c = \text{øbb}$
2.2	$\frac{K}{s(1 + Ts)}$	$K_p = \frac{1}{KT}$	$K_p = \frac{1}{K} \omega_c$ $T_d = T$ $\omega_c = \text{øbb}$
2.3	$\frac{K}{(1 + T_1 s)(1 + T_2 s)}$ $T_1 > T_2$	$K_p = \frac{1}{K} \left(\frac{T_1}{T_2} + 1 + \frac{T_2}{T_1} \right)$ $K_p \approx \frac{1}{K} \frac{T_1}{T_2}$ når $T_1 \gg T_2$	$K_p = \frac{1}{K} T_1 \omega_c$ $T_d = T_2$ $\omega_c = \text{øbb}$
3.1	$\frac{K}{s^2(1 + Ts)}$	X	$K_p = \frac{1}{K} \frac{\alpha \sqrt{\alpha}}{T^2}$ $T_d = \frac{1}{\alpha} T$, $\alpha \approx \frac{1}{10}$
3.2	$\frac{K}{(1 + T_1 s)(1 + T_2 s)}$ $T_1 > T_2$	$K_p = \frac{1}{2K} \frac{T_1 + T_2}{T_1 T_2}$	$K_p = \frac{1}{KT_2}$ $T_d = T_1$

	PD ⁽²⁾	PI	PID
	$h_r(s) = K_p \frac{1 + T_d s}{1 + \alpha T_d s}$	$h_r(s) = K_p \frac{1 + T_i s}{T_i s}$	$h_r(s) = K_p \frac{(1 + T_i s)(1 + T_d s)}{T_i s}$
1.1	X	$K_p = \frac{1}{K} \omega_c$ $T_i = \frac{5}{\omega_c}$ $\omega_c = \text{øbb}$	X
1.2	X	$K_p = \frac{1}{K} T \omega_c$ $T_i = T$ $\omega_c = \text{øbb}$	X
2.1	$K_p = \frac{1}{K} \sqrt{\alpha} \omega_c^2$ $T_d = \frac{1}{\omega_c \sqrt{\alpha}}$ $\omega_c = \text{øbb}$	X	X
2.2	$K_p = \frac{1}{K} \frac{1}{\alpha T}$ $T_d = T$	$K_p = \frac{1}{KT \sqrt{\beta}}$ $T_i = \beta T, \beta \approx 10$	$K_p = \frac{1}{K} \frac{T}{T_d} \omega_c$ $T_i = T$ $\omega_c = \text{øbb}$
2.3	$K_p = \frac{1}{K} \frac{T_1}{\alpha T_2}$ $T_d = T_2$	$K_p = \frac{1}{K} \frac{T_1}{T_2}$ $T_i = T_1$	$K_p = \frac{1}{K} T_i \omega_c$ $T_i = T_1, T_d = T_2$ $\omega_c = \text{øbb}$
3.1	$K_p = \frac{1}{K} \frac{0.05}{T^2}$ $T_d = \frac{1}{\alpha} T, \alpha = \frac{1}{20}$	X	X
3.2	$K_p = \frac{1}{2K} \frac{T_1 + 10T_2}{T_1 T_2}$ $T_d = T_1, \alpha = \frac{1}{10}$	$K_p = \frac{1}{3K} \frac{T_1 + T_2}{T_1 T_2}$ $T_i = 5\sqrt{T_1 T_2}$	$K_p = \frac{1}{K} \frac{1}{\sqrt{10} T_2}$ $T_i = 10T_2$ $T_d = T_1$

		P	PD⁽¹⁾
	$h_u(s)$	$h_r(s) = K_p$	$h_r(s) = K_p(1 + T_d s)$
3.3	$\frac{K}{(1 + T_1 s)(1 + T_2 s)(1 + T_3 s)}$ $T_1 > T_2 > T_3$	$K_p = \frac{1}{K} \frac{T_1}{T_2}$	$K_p = \frac{1}{K} \frac{T_1}{T_3}$ $T_d = T_2$
4.1	$\frac{K}{s} e^{-\tau s}$	$K_p = \frac{1}{K} \frac{\pi}{4\tau}$	X
4.2	$\frac{K}{(1 + Ts)} e^{-\tau s}$	$K_p = \frac{1}{K} \cdot$ $\left(\frac{1}{2} \sqrt{12 \left(\frac{T}{\tau}\right)^2 + 4 \frac{T}{\tau} + 1} - \frac{T}{\tau} \right)$ $\approx \frac{1}{K} \frac{T + \tau/2}{\tau}$	X
5.1	$\frac{K}{s^2} e^{-\tau s}$	X	$K_p = \frac{1}{K} \frac{0.1}{\tau^2}$ $T_d = 10\tau$
5.2	$\frac{K}{s(1 + Ts)} e^{-\tau s}$	$K_p = \frac{1}{K\tau} \frac{1 + 2\frac{T}{\tau}}{1 + 4\frac{T}{\tau}} \leq \frac{1}{KT}$	$K_p = \frac{1}{K} \frac{\pi}{4\tau}$ $T_d = T$
5.3	$\frac{K}{(1 + Ts)(1 + T_2 s)} e^{-\tau s}$ $T_1 > T_2 \gg \tau$	$K_p = \frac{1}{K} \frac{T_1 + T_2}{2\tau} \cdot$ $\frac{4T_1 T_2 + (T_1 + T_2)2\tau + \tau^2}{4T_1 T_2 + (T_1 + T_2)\tau}$ $\approx \frac{1}{K} \frac{T_1 + T_2}{2\tau}$	$K_p = \frac{1}{K} \cdot$ $\left(\frac{1}{2} \sqrt{12 \left(\frac{T}{\tau}\right)^2 + 4 \frac{T}{\tau} + 1} - \frac{T}{\tau} \right)$ $T_d = T_2$

	PD ⁽²⁾	PI	PID
	$h_r(s) = K_p \frac{1 + T_d s}{1 + \alpha T_d s}$	$h_r(s) = K_p \frac{1 + T_i s}{T_i s}$	$h_r(s) = K_p \frac{(1 + T_i s)(1 + T_d s)}{T_i s}$
3.3	$K_p \approx \frac{1}{K} \frac{T_1}{T_2}$ $T_d = T_2$	$K_p = \frac{1}{K} \frac{T_1}{T_2}$ $T_i = T_1$	$K_p = \frac{1}{K} \frac{T_i}{T_3}$ $T_i = T_1, T_d = T_2$
4.1	X	$K_p \approx \frac{1}{K} \frac{1}{2\tau}$ $T_i \approx 10\tau$	X
4.2	X	$K_p = \frac{1}{K} \frac{\pi T_i}{4\tau}$ $T_i = T\tau$	X
5.1	$K_p = \frac{1}{K} \frac{0.05}{\tau^2}$ $T_d = 10\tau$	X	X
5.2	$K_p = \frac{1}{K\tau} \frac{1 + 2\alpha \frac{T}{\tau}}{1 + 4\alpha \frac{T}{\tau}} \leq \frac{1}{K} \frac{1}{\alpha T}$ $T_d = T$	$K_p = \frac{1}{KT\sqrt{\beta}}$ $T_i = \beta\tau, \beta \approx 20$	$K_p = \frac{1}{K\tau}$ $T_i = \beta\tau, \beta \approx 20$ $T_d = T, \tau < T$
5.3	$K_p = \frac{5}{K} \frac{T_1}{T_2}$ $T_d = T_2$ $\alpha = \frac{1}{20}$	$K_p = \frac{1}{K} \frac{T_1}{\tau} \frac{(2T_2 + \tau)}{(4T_2 + \tau)}$ $T_i = T_1$	$K_p = \frac{1}{K} \frac{\pi T_1}{4\tau}$ $T_i = T_1$ $T_d = T_2$

9.3.5 Seriekompensasjon av følgesystemer (servomekanismer)

Seriekompensasjon er et viktig virkemiddel ved oppbygning av tilbakekoblede følgesystemer som skal tilfredsstillende gitte spesifikasjoner. Dette gjelder framfor alt i de enkleste og billigste systemtypene. I mer kompliserte systemer bruker vi gjerne en kombinasjon av seriekompensasjon og kompensasjon med intern tilbakekopling som skal beskrives i neste avsnitt.

Seriekompensasjon av et følgesystem vil stort sett følge de samme prinsippene som beskrevet til nå i dette kapitlet. Det finnes imidlertid visse forskjeller det kan være verd å legge merke til.

I mange tilfeller vil “prosessen” i et følgesystem være en eller annen form for motorisk enhet som gir en bevegelse (translasjon eller rotasjon). I så fall vil det være en integrasjon mellom enhetens hastighet og dens posisjon (vinkel). Dette har betydning fordi følgesystemets dynamiske egenskaper er avhengig av hvor mange integrasjoner som finnes i sløyfen, som beskrevet i avsnitt 9.2.2 (tabell 9.1). Dersom det “naturlige” antallet integrasjoner i prosessen er tilstrekkelig ut fra spesifikasjonene, vil det ikke være nødvendig å tilføre nye integrasjoner i regulatoren. Dersom spesifikasjonene fordrer to integrasjoner i sløyfen og prosessen bare har en, må imidlertid kompensasjonselementet bidra med en ekstra integrasjon.

Seriekompensasjon av følgesystemer basert på frekvensanalytiske spesifikasjoner skiller seg fra det som er beskrevet i foregående avsnittet bare ved at vi også gjerne undersøker størrelsen

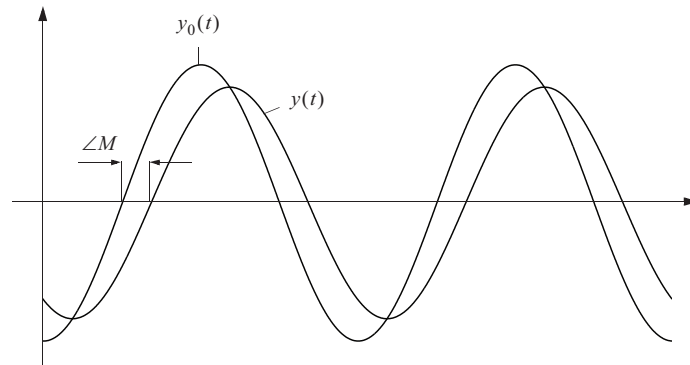
$$\frac{y}{y_0}(j\omega) = M(j\omega)$$

Det er vesentlig for et følgesystem at resonanstoppen i $|M(j\omega)|$ ikke blir for høy, og vi pleier ofte å spesifisere $|M(j\omega)|_{\max} \leq 3$ [dB]. Spesifikasjonen $|N(j\omega)|_{\max} \leq 5 - 6$ [dB] vil imidlertid lede til omtrent det samme resultatet. Et annet forhold er at det for mange anvendelser ikke bare er viktig at amplitudeforholdet $|M(j\omega)|$ skal være mest mulig lik 1, men også at $\angle M(j\omega)$ skal være minst mulig opp til høyest mulig frekvens.

$\angle M(j\omega)$ er et direkte mål for etterslepingen som $y(t)$ har i forhold til $y_0(t)$ når referansen er sinusformet, som illustrert i figur 9.37

Etterslepingen i tid finner vi lik

$$\Delta t = \frac{\angle M(j\omega)}{\omega} \quad (9.60)$$

Figur 9.37Betydning av
 $\angle M(j\omega)$ 

Antar vi eksempelvis at

$$M(s) = \frac{1}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \quad (9.61)$$

finner vi av (6.24) at

$$\angle M(j\omega) \approx -\arctan \frac{2\zeta \frac{\omega}{\omega_0}}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \quad (9.62)$$

For $\omega/\omega_0 \leq 0.3$ og dessuten $\zeta \leq 1$ finner vi

$$\angle M(j\omega) \approx -2\zeta \frac{\omega}{\omega_0} \quad (9.63)$$

Likning (9.63) og (9.60) gir da

$$\Delta t \approx -2\zeta \frac{1}{\omega_0} \quad (9.64)$$

Etterslepingen (forsinkelsen) ved frekvenser noe under den udedempede resonansfrekvensen er altså uavhengig av frekvensen på svingningen, men proporsjonal med den relative dempningsfaktoren. Eksempelvis vil et system med $\omega_0 = 10$ [rad/s] og $\zeta = 0.4$ (som gir $|M(j\omega)|_{\max} \approx 3$ [dB]) ha

$$\Delta t \approx -2 \cdot 0.4 \cdot \frac{1}{10} = 0.08 \text{ [s]}$$

I avsnitt 9.2.2 undersøkte vi et følgesystems evne til å følge typiske tidsforløp som sprangfunksjon, rampefunksjon eller parabelfunksjon. Det stasjonære avviket forårsaket av slike referansefunksjoner framgår av tabell 9.1. Dersom vi ønsker at systemets stasjonære avvik skal være lik null etter et sprang i referansen, må antallet integrasjoner (p) i sløyfen minst være lik 1.

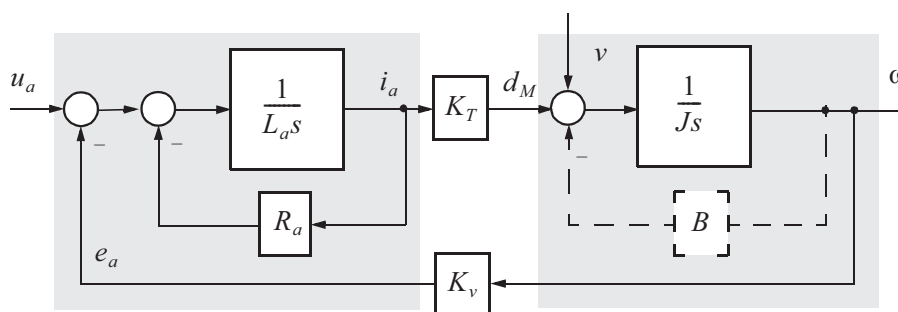
Figur 9.9 viser at innsvingningen vil få en uheldig oscillasjon ved valg av $p = 2$. Konklusjonen blir derfor at vi med hensyn på et sprang i referansen bør velge $p = 1$. Et system med $p = 1$ vil som det framgår av figur 9.6, få en viss ettersleping når det påtrykkes en rampefunksjon som referanse. Det stasjonære følgeavviket vil være proporsjonalt med faktoren $1/K$ der K er forsterkningen i sløyfen som framgår av (9.8) og eksemplet i figur 9.8.

Disse prinsippene skal illustreres ved hjelp av noen eksempler.

EKSEMPEL 9.8: Styring av likestrømsmotor

Vi skal i dette eksemplet se på dimensjonering av en servomekanisme hvor den vesentligste komponenten er en likestrømsmotor. Blokkdiagrammet for likestrømsmotoren (tegnet på grunnlag av likningene (2.69) og (2.70) i eksempel 2.8) blir

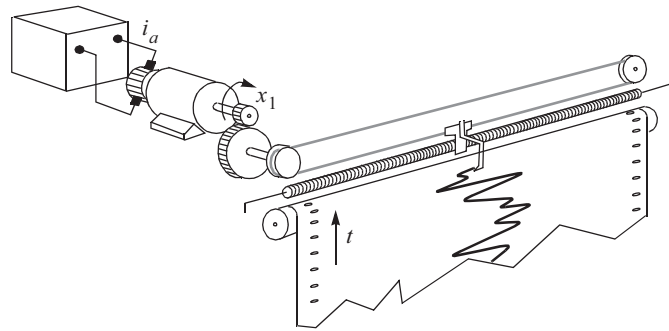
Figur 9.38
Blokkdiagram for
likestrømsmotor



De elektriske og mekaniske “undersystemene” er antydnet med skraverte rektangler. Fra nå av ser vi bort fra friksjon og setter $B = 0$

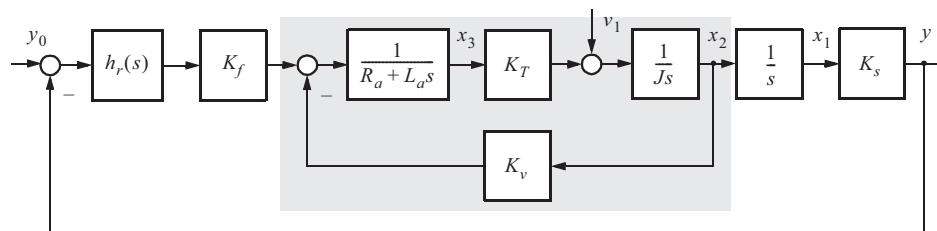
Likestrømsmotoren brukes til *styring av pennen i en potensiometerskriver*. Dette eksemplet har mange likhetstrekk med eksempel 7.8. En liten ankerstyrt likestrømsmotor med konstant felt inngår, som vist i figur 9.39. Motoren driver en penn som kan tegne kurver på et papir, over en mekanisk utveksling. Skyvekontakten til et potensiometer er festet til pennen. Spenningen over potensiometeret sammenliknes med et innkommende referansesignal som skal reproduseres som en kurve på papiret. For relativt små likestrømsmotorer kan vi få kontinuerlig virkende elektroniske kraftforsterkere som kan drive motoren direkte i ankerkretsen. (For større motorer i andre servo-anvendelser er det nødvendig å bruke enten thyristorer eller krafttransistorer for modulert effektilførsel eller en roterende generator).

Figur 9.39
Potensiometer-
skriver



Figur 9.40 viser et blokkdiagram der de forskjellige komponentene og fenomenen inngår. Den elektroniske kraftforsterkeren antas å ha en spenningsforsterkning lik K_f .

Figur 9.40
Blokkdiagram for
det totale systemet



Det skraverte rektangelet svarer til blokkdiagrammet i figur 9.38. Vi bruker nå notasjonen x for tilstand, y for måling, og v for forstyrrelse. Ankerstrømmen i motoren ($i_a = x_3$) gir gjennom momentkonstanten K_T opphav til et dreiemoment som sammen med et ytre belastningsmoment v_1 virker på motorens ekvivalente treghetsmoment J og gir en akselerasjon, som igjen resulterer i en rotasjonshastighet x_2 . Integralet av denne hastigheten gir motorens vinkelposisjon x_1 . Merk at vi i dette tilfelle ikke har *hastighetsregulering*, men *posisjonsregulering*, derfor den ekstra integratoren til høyre i blokkdiagrammet.

Overføringen mellom rotasjonen til motoren og bevegelsen av pennen, måles ved hjelp av spenningen som plukkes opp på potensiometerets skyvekontakt. Potensiometerkonstanten er K_s [volt/rad]. Transferfunksjonen fra forsterkerens inngang til potensiometerets spenning blir

$$\begin{aligned}
 h_u(s) &= K_f \frac{\frac{K_T}{(R_a + L_a s) J s} \frac{1}{s}}{\frac{K_v K_T}{(R_a + L_a s) J s}} K_s \\
 &= K_f \frac{1}{K_v \left(1 + \frac{J R_a}{K_v K_T} s + \frac{J L_a}{K_v K_T} s^2 \right)} \frac{1}{s} K_s \approx \frac{\frac{K_f K_s}{K_v}}{s(1 + T_a s)(1 + T_m s)}
 \end{aligned} \tag{9.65}$$

der

$$T_a = \frac{L_a}{R_a} = \text{ankekretsens elektriske tidskonstant}$$

og

$$T_m = \frac{JR_a}{K_v K_T} = \text{motorens mekaniske tidskonstant}$$

Vi vil ofte finne at $T_a \ll T_m$, slik at vi kan sette $T_a \approx 0$.

En servomekanisme for det formålet som her er skissert, skal framfor alt reproducere vilkårlige kurveforløp med minst mulig dynamisk feil.

Følgende fysiske parametre antas gitt:

K_f = kraftforsterkerens spenningsforsterkning	= 50[V/V]
R_a = ankekretsens motstand	= 40[ohm]
L_a = ankekretsens induktivitet	= 0.2[H]
K_v = motorens hastighetskonstant	= 0.1 $\left[\frac{\text{V}}{\text{rad/s}} \right]$
K_T = motorens momentkonstant	= 0.1 $\left[\frac{\text{Nm}}{\text{A}} \right]$
J = treghetsmomentet referert til motoraksel	= $1.25 \cdot 10^{-5}$ [kgm ²]
K_s = potensiometerkonstanten	= 0.5[V/rad]
$y_{0, \max}$ = skriverens spenningsområde	= ± 10 [V]

Med disse parametrene får vi av (9.65)

$$h_u(s) = \frac{250}{s(1 + 0.005s)(1 + 0.05s)} \quad (9.66)$$

En rimelig spesifisering til dette systemet angår etterslepingen når inngangssignalet $y_0(t)$ er en rampefunksjon. Vi spesifiserer at skriverens fulle spenningsområde (20 [V]) skal gjennomløpes i løpet av 1 sekund og at den stasjonære etterslepingen da ikke skal være større enn 0,2% av spenningsområdet, altså 0.04 [V]. Bruk av resultatet i (9.11) og tabell 9.1 gir da

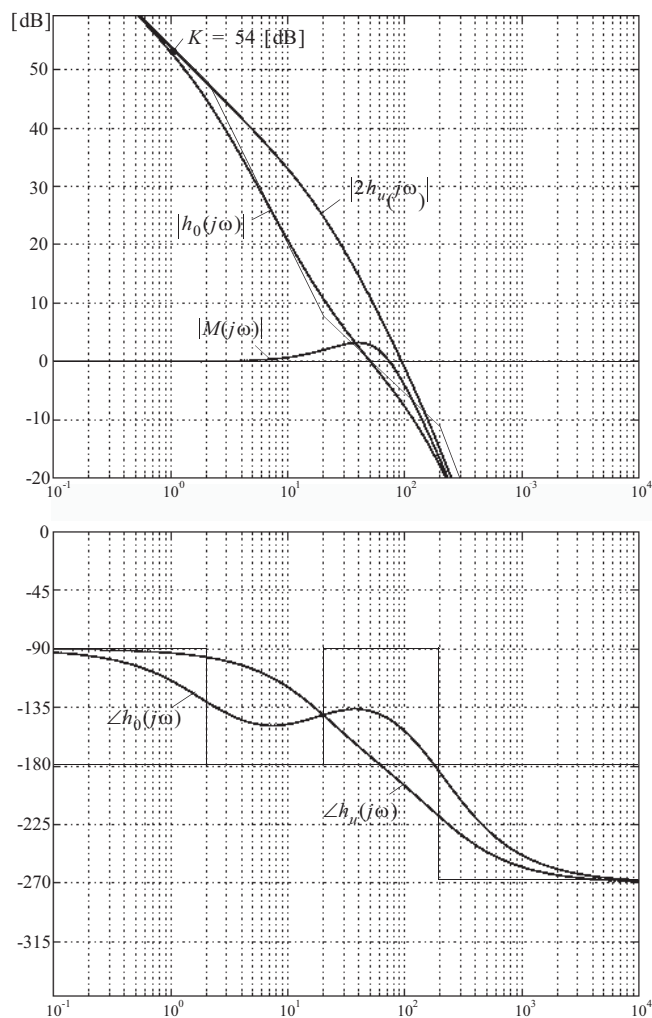
$$\frac{20}{K} = 0.04 \quad (9.67)$$

der K er den endelige sløyfens forsterkning under forutsetning av at sløyfen har én integrasjon ($p = 1$). Dette gir $K = 500 = 54$ [dB].

Siden vi også er interessert i at dette systemet skal oppvise en fornuftig sprangrespons uten vesentlig oversving, vil vi *ikke* spesifisere at systemet skal ha $p = 2$ integrasjoner i sløyfen. Med to integrasjoner kunne vi oppnådd perfekt stasjonær følgning av en rampefunksjon, men ville fått betydelig oversving i sprangresponsen som vist i figur 9.9.

Vi kan nå tegne et asymptotisk amplitude-diagram for sløyfen i figur 9.40 når bare kravet $K = 500$ skal tilfredstilles. Dette oppnås ved å velge et kompensasjonselement som bare er en konstant forsterkning $h_r^{(1)}(s) = 2$. Resultatet er vist i figur 9.41. Av dette diagrammet ser vi at systemet ikke er akseptabelt med denne enkle kompensatoren på grunn av ustabilitet ($\omega_c > \omega_{180}$). Vi må derfor innføre frekvensavhengige elementer i kompensatoren for blant annet å oppnå stabilitet.

Figur 9.41
AFF-diagram
for sløyfen



Vi kan i prinsippet oppnå stabilitet i systemet ved å innføre en forholdsvis sterk positiv fasekorreksjon i frekvensområdet 200 – 500 [rad/s]. Dette medfører imidlertid at vi må ha to begrensede derivatvirkninger i kaskade for å oppnå tilstrekkelig fasemargin. En mer realistisk løsning er å bruke en kombinasjon av proporsjonal + begrenset integralvirkning og proporsjonal + begrenset derivatvirkning som vist i figur 9.41. Kryssfrekvensen vil derved bli $\omega_c = 50$ [rad/s]. Skulle vi oppnådd en vesentlig høyere båndbredde enn dette, måtte vi ha valgt en mekanisk konstruksjon som tillot høyere båndbredde, for eksempel en hurtigere motor. Kompensasjonselementet som er benyttet for det korrigerede forløp i figur 9.41, har formen

$$h_r(s) = \frac{2(1 + 0.05s)^2}{(1 + 0.5s)(1 + 0.005s)} \quad (9.68)$$

og den resulterende sløyfetransferfunksjonen blir

$$h_0(s) = \frac{500(1 + 0.05s)}{s(1 + 0.5s)(1 + 0.005s)^2} \quad (9.69)$$

Det lukkede systemets frekvensrespons $M(j\omega)$ bestemmes ved bruk av Nicholsdiagram eller plottes ved hjelp av MATLAB og er vist i figur 9.41.

Den lukkede sløyfens transferfunksjon blir

$$M(s) = \frac{h_0(s)}{1 + h_0(s)} = \frac{1 + 0.05s}{1 + 0.052s + 0.00102s^2 + 1.01 \cdot 10^{-5}s^3 + 2.5 \cdot 10^{-8}s^4} \quad (9.70)$$

Knekkfrekvensen til den doble polen i $h_0(s)$ er en dekadere høyere enn den nest høyeste knekkfrekvensen. Virkningen av den doble polen er derfor tilnærmet uavhengig av de andre polene og nullpunktene i transferfunksjonen. Vi forsøker derfor følgende tilnærming for $M(s)$

$$M(s) \approx \frac{1 + 0.05s}{\left(1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2\right) (1 + 0.005s)^2} \quad (9.71)$$

Vi multipliserer ut nevneren i (9.71) og finner parametrene ζ og ω_0 , slik at leddene av fjerde og første grad stemmer overens med (9.70). Det lukkede systemets transferfunksjon vil nå tilnærmet være lik

$$M(s) \approx \frac{1 + 0.05s}{\left(1 + 2 \cdot 0.664 \frac{s}{31.6} + \left(\frac{s}{31.6}\right)^2\right) (1 + 0.005s)^2} \quad (9.72)$$

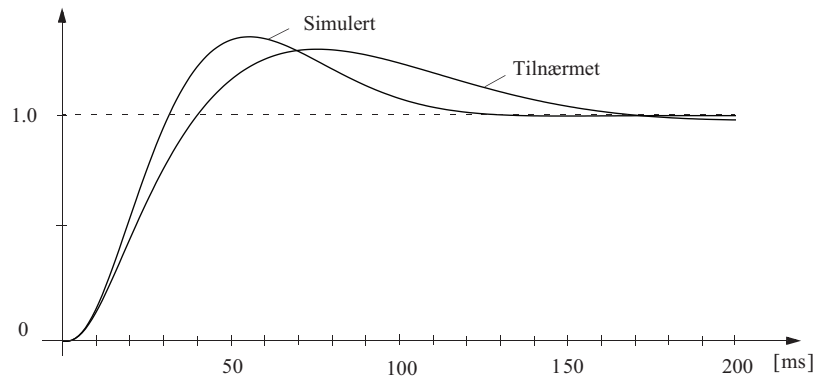
Ved å bruke MATLAB kan vi finne det eksakte uttrykket for $M(s)$ som blir

$$M(s) \approx \frac{1 + 0.05s}{\left(1 + 2 \cdot 0.675 \frac{s}{61.4} + \left(\frac{s}{61.4}\right)^2\right) (1 + 0.00355s)(1 + 0.0265s)} \quad (9.73)$$

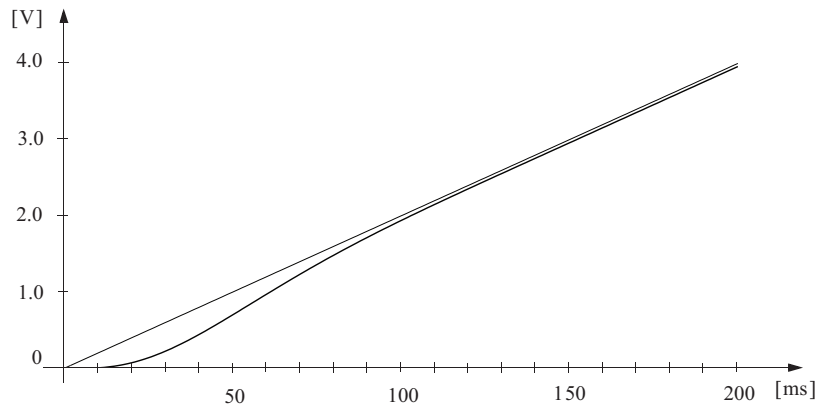
Det er tydelig at tilnærmingen i (9.72) har ganske stor feil i resonansfrekvensen og i de to høyfrekvente polene.

Likning (9.72) kan brukes for skissering av en tilnærmet sprangrespons av systemet som vist i figur 9.42. Figuren viser også den egentlige sprangresponsen som finnes ved simulering. Hovedsaken er at systemet har en stigetid etter et sprang som ligger i området 40 [ms].

Figur 9.42
Sprangrespons



Figur 9.43
Respons ved påtrykking av en rampe



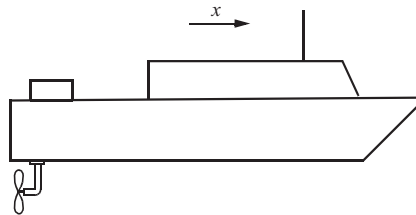
Etterslepingen (med en rampefunksjon $20[V/s]$ som referanse) vil som vist i figur 9.43 til å begynne med komme opp i ca. $0.3 [V]$, men etter ca. $0.1 [s]$ være redusert til ca. $0.04 [V]$ som spesifisert. Dersom dette ikke er tilfredsstillende, er det som tidligere sagt mulig å presse den endelige sløyfens frekvensrespons noe mer mot høyre med bruk av sterkere derivatvirkning eller aller helst ved å bruke en hurtigere motor.

■

EKSEMPEL 9.9: Dynamisk posisjonering av farkoster

Et eksempel på problemet betinget stabilitet som er nevnt i eksempel 8.17, har vi ved såkalt “dynamisk posisjonering” av farkoster i sjøen som vist i figur 9.44. Dette området er behandlet i stor detalj i Fossen (1994)

Figur 9.44
Dynamisk posisjonering av farkoster



En propell (thruster) drives av en elektrisk motor og gir skyvkraft til å posisjonere båten i horisontal retning. Når hastigheten til båten er liten, vil den hydrauliske dempingen være liten og transferfunksjonen mellom propell-skyvkraft og posisjon være

$$h_u(s) = \frac{1}{ms^2} \quad (9.74)$$

der $m =$ farkostens masse.

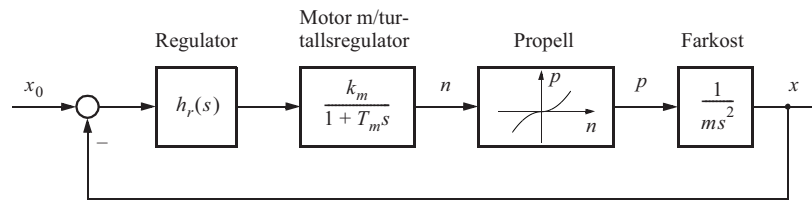
En propell vil gi en skyvkraft $p = k_t |n|n$, der $n =$ propellens turtall og $k_t =$ konstant. Dette betyr at skyvekraften er svært liten når turtallet er lite (ulineært). På grunn av de to integrasjonene i $h_u(s)$ vil vi velge proporsjonal+begrenset derivat-regulatoren

$$h_r(s) = K_p \frac{1 + T_d s}{1 + \alpha T_d s}$$

der $\alpha \approx 0.1$ er en rimelig verdi.

Dersom motoren utstyres med *turtallsregulering*, vil vi få en reguleringsløyfe som vist i figur 9.45

Figur 9.45
System med
turtallsregulering



Hadde propellen hatt en karakteristikk som ga proporsjonalitet k_t mellom turtall n og skyvkraft p , ville vi fått en reguleringsløyfe som fungerte bra med sløyfetransferfunksjonen

$$h_0(s) = K_p \frac{1 + T_d s}{1 + \alpha T_d s} \frac{k_m}{1 + T_m s} k_t^{(1)} \frac{1}{m s^2} \quad (9.75)$$

der

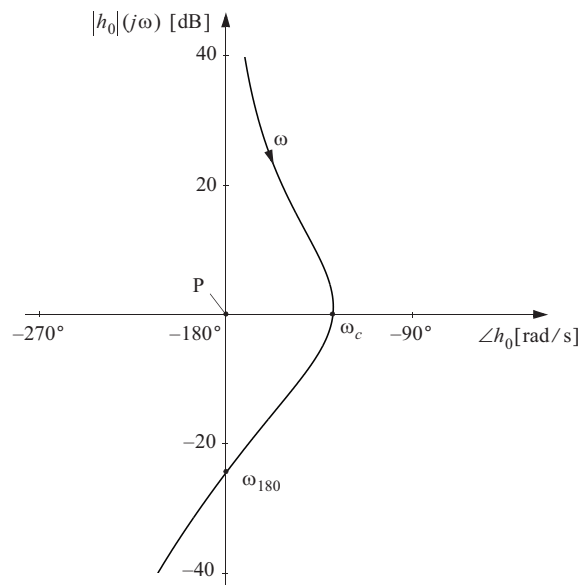
$$K_p = \frac{\sqrt{10}}{2500} \cdot \frac{m}{T_m^2 k_m k_t^{(1)}}$$

$$T_d = 50 T_m \text{ og } \alpha = 0.1$$

ville gitt en akseptabel stabilitetsmargin.

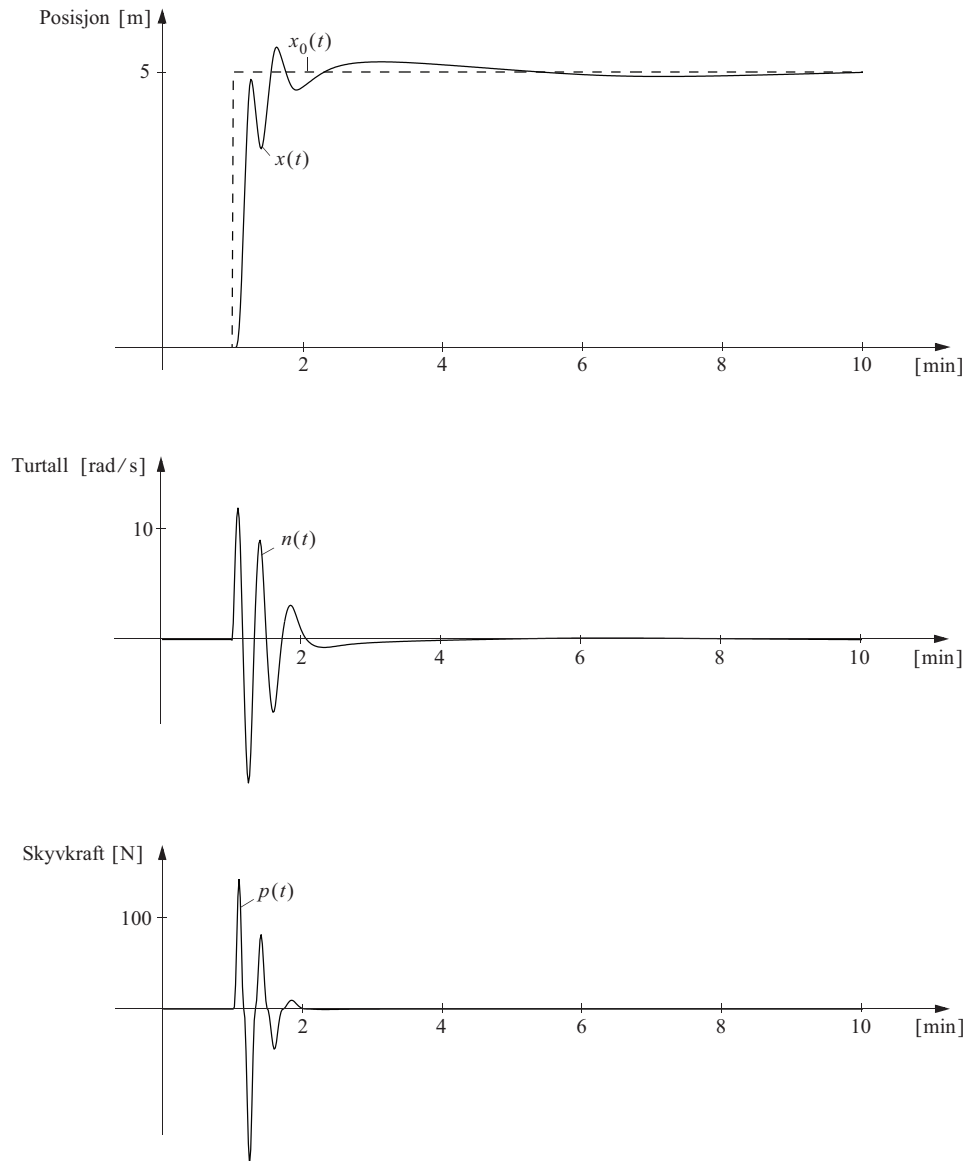
Men siden $p = k_t |n| n$ er en ulineær funksjon i n , vil “forsterkningen” assosiert med propellen gå mot null når turtallet blir lite.

Figur 9.46
A/F-diagram
for $h_0(j\omega)$



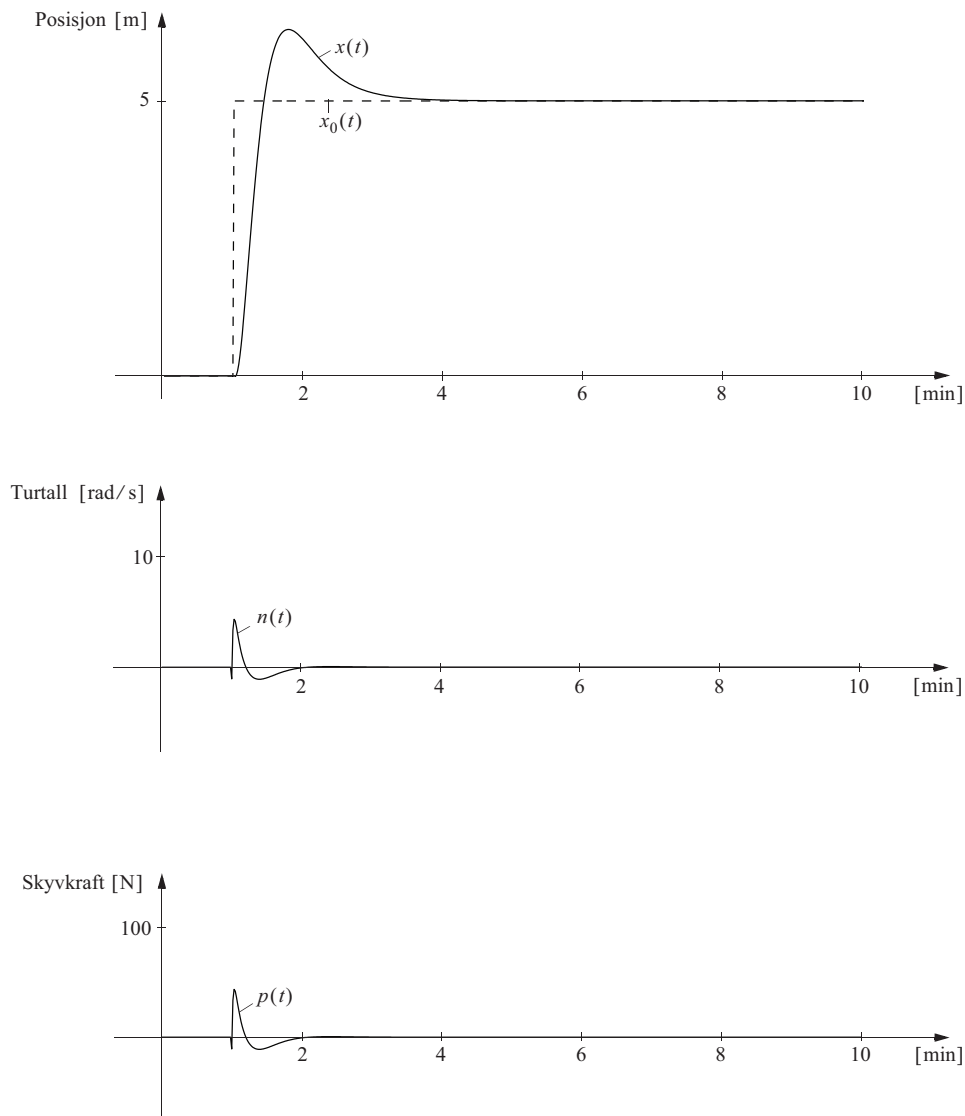
Figur 9.46 viser et amplitude/fase-diagram for $h_0(j\omega)$ dersom “forsterkningen” k_t i propellen er konstant. Siden denne “forsterkningen” i virkeligheten er $k_t|n|$, vil stedkurven for $h_0(j\omega)$ forskyves nedover i forhold til det kritiske punkt P (0 [dB], -180°). Dette er ekvivalent med å si at det kritiske punkt P beveger seg oppover og nærmer seg stedkurven.

Figur 9.47
Horisontal posisjon, turtall av propell og skyvkraft ved sprang i posisjonsreferansen, ulineært system



Dette betyr at dempningen i systemet blir mindre og mindre med det resultat at systemet vil begynne å svinge, men med avtagende amplitude. Figur 9.47 illustrerer dette der farkosten gis et sprang i horisontal posisjonsreferanse. Responsen i posisjon $x(t)$, skyvkraft $p(t)$ og propellerturtall $n(t)$ er vist i figuren. Figur 9.48 viser det samme med lineær karakteristik for propellen.

Figur 9.48
Lineært system



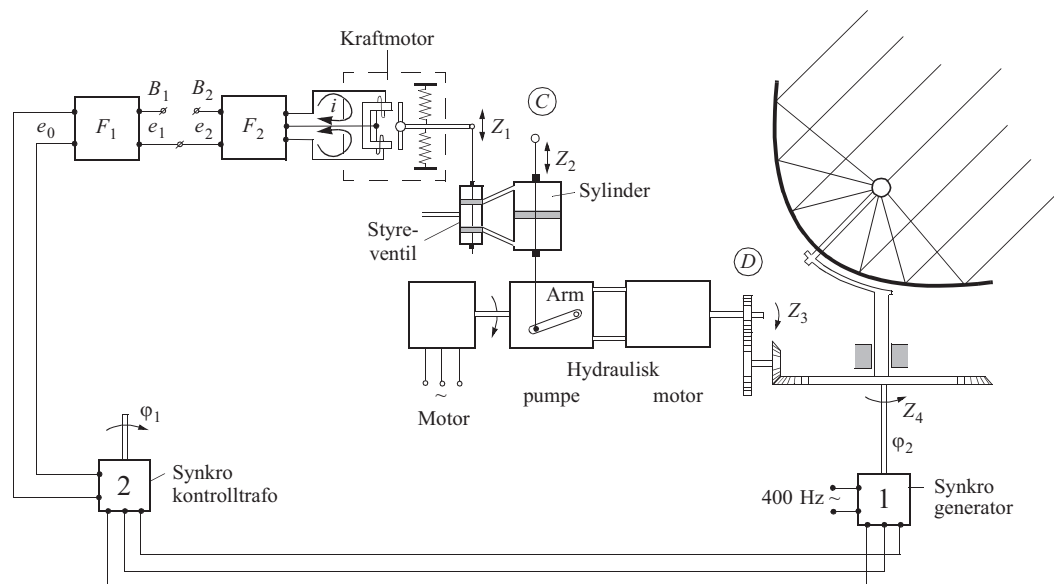
EKSEMPEL 9.10: Styring av en antenne

Figur 9.49 viser skjematisk de viktigste elementene som inngår i en elektro-hydraulisk servomekanisme for drift av en tung antenne-installasjon. Dette er et eksempel som prinsipielt er likt med motorstyringseksemplet 7.7. Problemet er å dreie drivakselen til en antenne slik at dens vinkel φ_2 til enhver tid blir mest mulig lik en referansevinkel φ_1 . Denne foreligger også som en akseldreining. Systemet er typisk ved antenner som skal følge satellitter for mottaking av radiosignaler. For ikke å miste følsomhet i antennen som har et svært smalt strålingsdiagram, er det viktig at antennen peker nøyaktig mot satellitten. Satellittens posisjon er beregnet av en datamaskin som ved hjelp av en instrument-servomekanisme driver referanseakselen, for det systemet som behandles her.

Antennen er utstyrt med to forholdsvis like servomekanismer, en for sidevinkel og en for høydevinkel. Bare den første av disse behandles her.

Da antennen har stort treghetsmoment, store friksjoner og er utsatt for stor vindpåkjenning, trengs det stor motoreffekt for å dreie den. En hydraulisk motor som drives med olje fra en pumpe, er derfor valgt. Mellom motor og utgangsaksel finnes en giroversetning for å få gunstig tilpasning mellom motorens normale turtall og antennens normale rotasjonshastighet.

Figur 9.49
Elektrohydraulisk
servomekanisme
som driver
en antenne



Pumpen, som er en stempelpumpe drevet av en elektrisk motor med konstant hastighet, er av en slik konstruksjon at stempelslaget kan justeres kontinuerlig. Oljestrømmen som pumpen leverer til motoren, kan varieres fra null til maksimal verdi og kan dessuten skifte retning. Motorens hastighet vil dermed kunne varieres

kontinuerlig i begge retninger. En justering av stemplenes slaglengde foregår ved hjelp av en hydraulisk sylinder som beveger en solid arm, tilgjengelig på pumpens utside. Fordi friksjonskrefter og reaksjonskrefter fra pumpen er store, er det nødvendig å bruke en forholdsvis kraftig hydraulisk sylinder til å bevege denne armen. Sylindren styres i sin tur fra en liten hydraulisk styreventil som får trykkolje fra en egen pumpe.

For bevegelse av styreventilen finnes en elektromagnetisk kraftmotor (magnet) som er i stand til å gi små, men hurtige bevegelser av styreventilens sleide. Gjennom kraftmotorens spole drives det strøm fra en likestrømsforsterker.

Alle disse forskjellige elementene er nødvendige for å oppnå det ønskede effektnivået, og for at enheten skal kunne koples sammen med det elektriske vinkelmåleelementet. Dette vinkelmåleelementet, en såkalt **synkrokjede**, består av to motorliknende enheter, synkrogenerator (merket 1) og kontrolltransformator (merket 2). Synkrogeneratoren mates av et vekselspenningsnett (her antydnet 400 Hz). Ut av kontrolltransformatoren fås da en vekselspenning med frekvens lik nettfrekvensen og med amplitude proporsjonal med sinus til vinkelavviket. Dette kan uttrykkes slik

$$e_0(t) = K_7 \sin(\varphi_1 - \varphi_2) \sin \omega_n t \quad (9.76)$$

der

$$\begin{aligned} K_7 &= \text{synkrokonstant} \\ \varphi_1 - \varphi_2 &= \text{vinkelavvik} \\ \omega_n = 2\pi f_n &= \text{nettfrekvens} \end{aligned}$$

Vinkelavviket foreligger altså som et modulert signal, og for relativt små vinkelavvik kan man tilnærmet skrive

$$e_0(t) = K_7(\varphi_1 - \varphi_2) \sin \omega_n t \quad (9.77)$$

To elektroniske forsterkere F_1 og F_2 følger etter synkrokjeden. F_1 er en vekselspenningsforsterker med demodulator (fasefølsom likeretter) som omdanner det modulerede signalet til et varierende likespenningssignal. F_2 er en likestrømforsterker som driver strømmen i kraftmotoren. Mellom de to forsterkerne er det i figur 9.49 et åpent rom hvor det er mulighet for å plassere et seriekompensasjonselement. Senere, under behandling av kompensasjon med intern tilbakekopling, skal vi benytte dette punktet for tilkopling av tilbakekoplingssignalet.

Fordi synkroelementenes nettfrekvens er mye høyere enn den kryssfrekvensen som kan oppnås i denne servomekanismen (se senere), kan vi med god tilnærming anta at vinkel-målingen er kontinuerlig i stedet for modulert. Signalet som kommer ut av demodulatoren, antas derfor til enhver tid å være proporsjonalt med vinkelavviket.

Under henvisning til figur 9.49 innføres følgende betegnelser:

$$\begin{aligned}
 K_1 &= \text{forsterkning i vekselspenningsforsterkeren} \\
 (F_1)_{\text{m/demodulator}} &= \frac{e_1}{e_0} \left[\frac{\text{V}}{\text{V}} \right] \\
 K_2 &= \text{forsterkning i likestrømsforsterkeren } (F_2) &= \frac{i}{e_0} \left[\frac{\text{mA}}{\text{V}} \right] \\
 z_1 &= \text{bevegelse i styresleiden} &[\text{mm}] \\
 z_2 &= \text{bevegelse av pumpepådraget} &[\text{mm}] \\
 z_3 &= \text{motorhastigheten} &[\text{rad/s}] \\
 z_4 &= \text{antennehastigheten} &[\text{rad/s}]
 \end{aligned}$$

Følgende transferfunksjoner antas kjent for de gitte elementene

$$\frac{z_1}{i}(s) = \frac{K_3}{1 + T_1 s} \quad (9.78)$$

der

$$\begin{aligned}
 K_3 &= 0.1 \left[\frac{\text{mm}}{\text{mA}} \right] \\
 T_1 &= 0.02 [\text{s}]
 \end{aligned}$$

$$\frac{z_2}{z_1}(s) = \frac{K_4}{s} \quad (9.79)$$

der

$$K_4 = 100 \left[\frac{\text{mm/s}}{\text{mm}} \right]$$

$$\frac{z_3}{z_2}(s) = \frac{K_5}{(1 + T_2 s)(1 + T_3 s)} \quad (9.80)$$

der

$$K_5 = 1.0 \left[\frac{\text{rad/s}}{\text{mm}} \right]$$

$$T_2 = 0.2[\text{s}]$$

$$T_3 = 0.05[\text{s}]$$

$$\frac{z_4}{z_3} = K_6 = \frac{1}{20} = 0.05 \left[\frac{\text{rad}}{\text{rad}} \right] \quad (9.81)$$

$$\frac{\varphi_2}{z_4}(s) = \frac{1}{s} \quad (9.82)$$

$$\frac{e_0}{\varphi_1 - \varphi_2} = \frac{e_0}{\Delta\varphi} = K_7 = 50 \left[\frac{\text{V}}{\text{rad}} \right] \quad (9.83)$$

Dersom sløyfen lukkes ved å forbinde de to punktene B_1 og B_2 mellom de to forsterkerne, vil vi finne den ukompenserte sløyfens transferfunksjon

$$h_0^{(0)}(s) = \frac{25K_1K_2}{s^2(1+0.02s)(1+0.2s)(1+0.05s)} \quad (9.84)$$

Vi ser at denne transferfunksjonen har fem poler, hvorav to ligger i origo.

Dette systemet vil helt sikkert være ustabilt hvis vi ikke innfører frekvensavhengig kompensasjon.

Følgende krav stilles nå til denne servomekanismen:

- a) For at antennen skal følge et mål som beveger seg med konstant hastighet, uten stasjonært vinkelavvik, fordres det at $h_0(s)$ har to integrasjoner (to poler i origo).
- b) Rimelig stabilitetsmargin og relativ dempning (for eksempel $|N(j\omega)|_{\max} = 6$ [dB]).
- c) Størst mulig båndbredde.

Konstruksjon av elektrisk element for seriekompensasjon

Seriekompensasjon kan vi få til ved å skyte inn et elektrisk nettverk mellom de to forsterkerne. Dette nettverket må være av typen “proporsjonal + begrenset derivatvirkning” for å få et positivt fasebidrag omkring den endelige kryssfrekvensen fordi den ukompenserte sløyfen har større negativ fasevinkel enn -180° ved alle frekvenser.

Antar vi til å begynne med

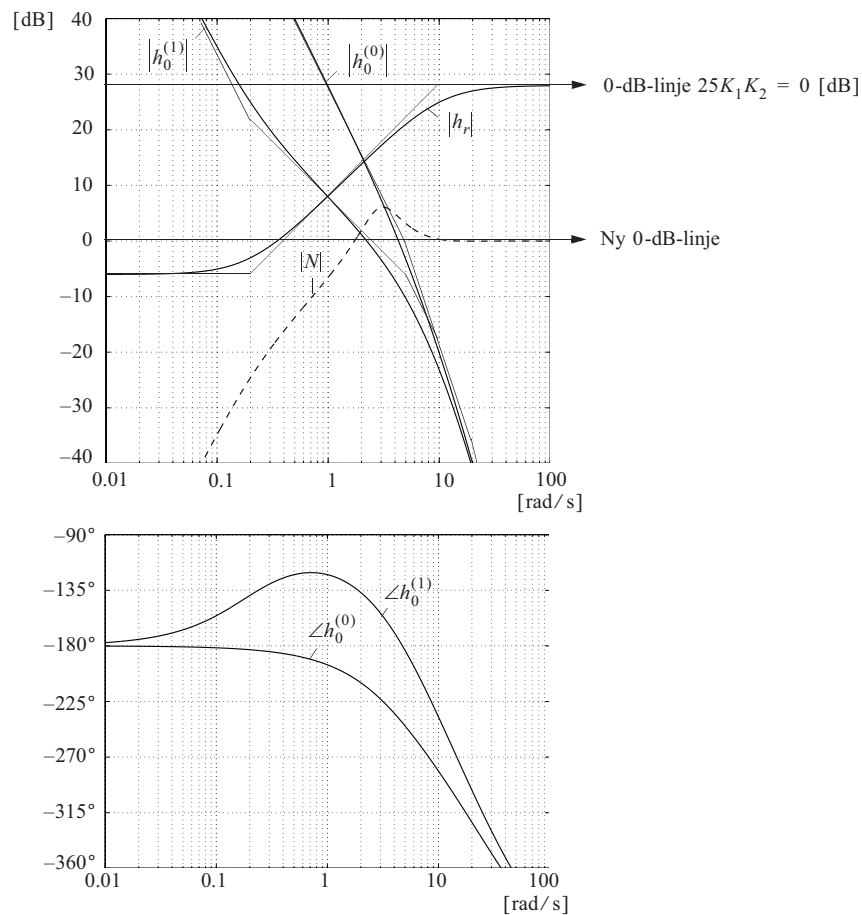
$$25K_1K_2 = 1 = 0 \text{ [dB]}$$

kan vi tegne diagrammer for $|h_0^0(j\omega)|$ og $\angle h_0^0(j\omega)$ som vist i figur 9.50. Fordi den asymptotiske amplitudekarakteristikken faller med vinkelkoeffisient (-3) for frekvenser over $\omega = 5$ [rad/s], kan vi bare regne med å få nytte av en derivatkompensasjon dersom denne i det vesentlige legges inn ved frekvenser lavere enn $\omega = 5$ [rad/s]. Et passivt seriekompensasjonselement med transferfunksjon

$$h_r(s) = \frac{1}{50} \frac{1 + 5s}{1 + 0.1s} \quad (9.85)$$

vil gi karakteristikker for den korrigerte sløyfe $|h_0^1(j\omega)|$ og $\angle h_0^1(j\omega)$ som vist i figur 9.50. Senker vi 0-dB-linjen med 28 [dB], som vist i figuren (det vil si $25K_1K_2 = 25 = 28$ [dB]), får vi et $|N(j\omega)|$ -diagram som vist streket, og som har en maksimal verdi omtrent lik 6 [dB].

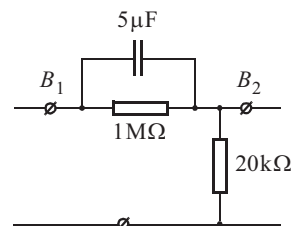
Figur 9.50
AFF-diagram for
systemet med og
uten regulator



Vi legger merke til at med den valgte løsningen har $|h_0^1(j\omega)|$ vinkelkoeffisient -1 i et forholdsvis bredt frekvensområde under kryssfrequensen. Dette gir faseløft over et tilsvarende frekvensområde, og skulle kanskje tyde på at vi har valgt en unødig kraftig derivat-virkning (dvs. over et unødig stort frekvensområde). Det er imidlertid en grunn til dette valget: Et system av denne typen som har to integrasjoner, kan bli mer oscillatorisk når forsterkningen i sløyfen avtar. På grunn av eldning av komponenter som kan gi større friksjon, økt viskositet til oljen i det hydrauliske systemet ved lave temperaturer og liknende, må vi regne med at 0-dB-linjen kan bli forskjøvet minst 6 [dB] oppover (forsterkningen avtar til det halve). Vi ønsker at systemet skal ha en rimelig stabilitetsmargin selv om dette inntreffer.

Et elektrisk nettverk som gir den ønskede seriekompensasjon er vist i figur 9.51.

Figur 9.51
Regulatoren implementert med elektriske komponenter

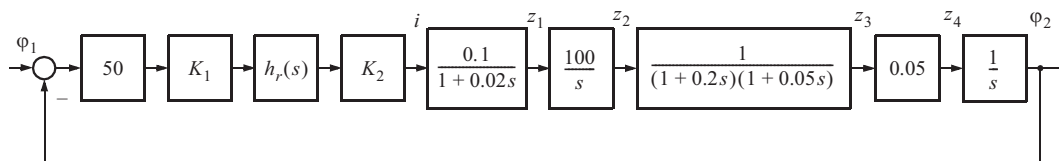


Komponentverdiene som er valgt, er ikke de eneste mulige. Ved valg av nettverkets impedansnivå må hensyn tas til utgangsimpedansen i demodulatoren i forsterkeren F_1 og inngangsimpedansen i den etterfølgende likestrømsforsterkeren F_2 .

Vi har ovenfor funnet $K_1 K_2 = 1 = 0$ [dB]. Hvordan vi fordeler forsterkningen mellom de to forsterkerne, har prinsipielt ingen betydning og vil bare være gitt av konstruktive hensyn.

Figur 9.52 viser et blokkdiagram for det ferdige seriekompenserte systemet der transferfunksjonene slik de er benyttet foran, vil framgå. Bortsett fra kompensatoren skal det samme blokkdiagrammet benyttes senere når det samme systemet skal kompenseres ved hjelp av intern tilbakekopling.

Figur 9.52
Prosessen med regulator



Den valgte løsningen med seriekompensasjon er ikke den beste løsningen selv om den er enkel. Med de løsningene som skal demonstreres senere, vil vi kunne oppnå større båndbredde for systemet, mindre følsomhet overfor parametervariasjoner og større evne til demping av forstyrrelser, for eksempel ytre momenter. Systemet blir da imidlertid mer komplisert.

■

EKSEMPEL 9.11: Frekvensanalyse av biltrafikk

På en lang og rett bilvei kjører en lang kø av biler med konstant nominell hastighet (v). Følgende antas:

- Alle biler og sjåførere har samme atferd.
- Bil nummer i har posisjon $x_i(t)$ ($x_i(s)$) og hastighet $v_i(t)$ ($v_i(s)$).
- Hver sjåfør ser bilen foran seg og bestemmer visuelt posisjonsavviket (avstanden) $\Delta x_i = x_{i-1} - x_i$ til den.
- Sjåføren styrer bilens akselerasjon ($a_i(t)$) og dermed hastigheten ($v_i(t)$) og posisjonen ($x_i(t)$) med en reaksjon som svarer til en PD-regulator, men med en forsinkelse τ_i slik at

$$\frac{a_i}{\Delta x_i}(s) = K_p(1 + T_d s)e^{-\tau s}$$

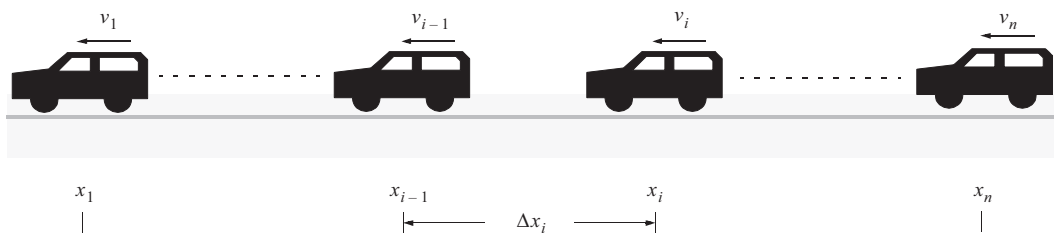
$$\frac{v_i}{a_i}(s) = \frac{1}{s} \text{ og}$$

$$\frac{x_i}{v_i}(s) = \frac{1}{s}$$

- Hver sjåfør forsøker å holde en avstand Δx° til bilen foran.

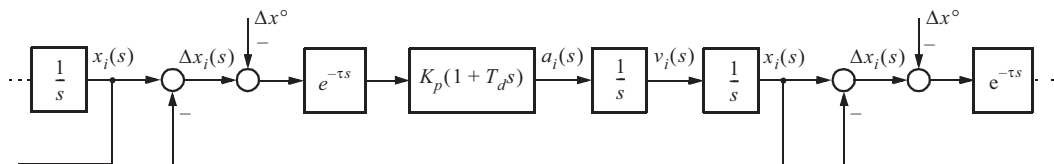
Figur 9.53 viser et bilde av denne bilkøen

Figur 9.53
En bilkø



Et blokkdiagram som beskriver bilkøen er vist i figur 9.54.

Figur 9.54
Blokkdiagram
for bilkøen



Den første bilen i køen (som ikke har noen bil foran seg) antas å kjøre med hastighetsregulering.

Transferfunksjonen fra en bils posisjon til den neste bakenfor blir

$$\frac{x_i}{x_{i-1}}(s) = \frac{h_0(s)}{1 + h_0(s)} = M(s)$$

Det samme gjelder bilenes hastighet

$$\frac{v_i(s)}{v_{i-1}(s)} = M(s)$$

Transferfunksjonen fra den første bilens hastighet $v_1(s)$ til bil nummer n sin hastighet $v_n(s)$ er

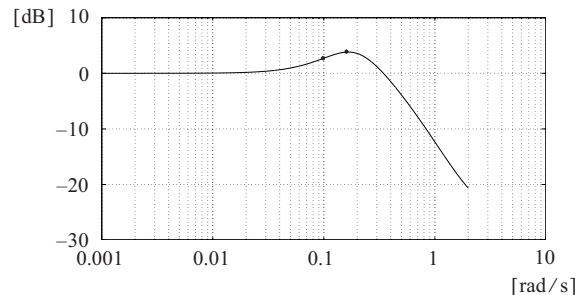
$$\frac{v_n}{v_1}(s) = M^{n-1}(s)$$

Det er interessant å undersøke $|M(j\omega)|$ og $|M(j\omega)|^{n-1}$ i frekvensplanet. Ved bruk av dB-skalaen får vi,

$$\left| \frac{v_n}{v_1}(j\omega) \right| = (n-1)|M(j\omega)| \text{ [dB]}$$

Figur 9.55 viser $|M(j\omega)|$ når vi har $\tau = 2$ [s], $T_d = 10$ [s] og $K_p = 0.02$ [N/m].

Figur 9.55
Amplitudekarakteristikk for $M(j\omega)$



$|M(j\omega)|$ har en resonans ved $\omega_r = 0.17$ [rad/s] og $|M(j\omega_r)| = 3.8$ [dB].

Det betyr at dersom det påtrykkes en hastighetsendring (perturbasjon) med denne frekvensen omkring den nominelle kjørehastigheten (v°), vil denne endringen *forsterkes* bakover i køen med 3.8 [dB] for hver bil. Ved andre frekvenser som er mindre eller større enn ω_r vil det også foregå en slik forsterkning, men med en mindre forsterkningsfaktor.

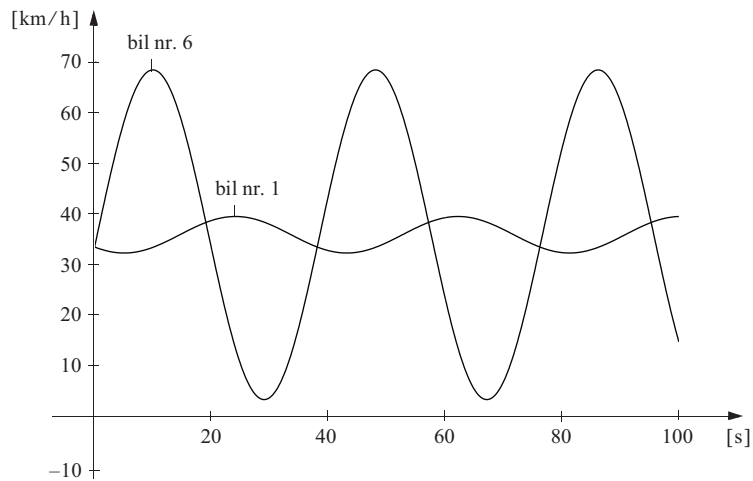
Vi antar at den nominelle kjørehastigheten er $v^\circ = 36 \text{ [km/h]} = 10 \text{ [m/s]}$ og at den første bilen gis en hastighetsendring med amplitude lik 10% av v° (det vil si $\Delta v^\circ = 1 \text{ [m/s]}$) ved frekvensen $\omega_1 = \omega_r = 0.17 \text{ [rad/s]}$.

Siden hver bil forsterker denne hastighetsendringen med 3.8 [dB], vil bil nummer 6 få en hastighetsendring mellom ca. 0 [km/h] og ca. 72 [km/h] fordi

$$n - 1 = \frac{20 \text{ [dB]}}{3.8 \text{ [dB]}} = 5.3$$

Figur 9.56 viser dette.

Figur 9.56
Forplantning av
hastighetsendring
i bilkøen



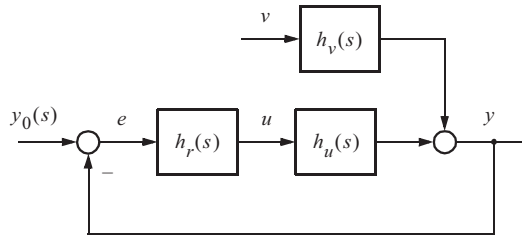
Hadde hastighetsendringen i den første bilen kommet med en frekvens $\omega_2 = 0.1 \text{ [rad/s]}$ som gir $|M(j\omega_2)| = 2.7 \text{ [dB]}$, ville vi funnet $n - 1 = 20/2.7 = 7.4$. Bil nummer 8 ville altså fått en hastighetsendring mellom ca. 0 [km/h] og ca. 72 [km/h].



9.4 NÅR TRENGER VI INTEGRERENDE REGULATOR I EN REGULERINGSSLØYFE?

Figur 9.57 viser en monovariabel reguleringsløyfe som gjelder både for prosessregulering og et følgesystem. Spørsmålet vi nå stiller er: *Når trenger vi en integrerende regulator, det vil si en I-, PI- eller PID-regulator?*

Figur 9.57
Monovariabel
reguleringsløyfe



Det finner vi ut ved å analysere responsen til systemet i figur 9.57 etter forskjellige eksitasjoner i de to driftsmodi, prosessregulering og følgesystem.

Anta at vi ønsker $e(t)$ når enten $v(t) =$ sprangfunksjon eller $y_0(t) =$ sprangfunksjon.

Vi finner ved Laplacetransformasjon

$$e(s) = e_{y_0}(s) + e_v(s) = N(s)y_0(s) - N(s)h_v(s)v(s) \quad (9.86)$$

der

$$N(s) = \frac{1}{1 + h_0(s)}, \quad h_0(s) = h_u(s)h_r(s)$$

Sluttverditeoremet (4.22) sier

$$\lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} s e(s)$$

Dersom $h_0(s)$ inneholder minst en integrasjon enten i regulatoren $h_r(s)$ eller i prosessen $h_u(s)$, vil bidraget fra det første leddet i (9.86) bli

$$\lim_{t \rightarrow \infty} e_{y_0}(t) = \lim_{s \rightarrow 0} s e_{y_0}(s) = \lim_{s \rightarrow 0} s \frac{1}{1 + \frac{t_0(s)}{s^p n_0(s)}} \frac{y_{00}}{s} = 0 \quad (9.87)$$

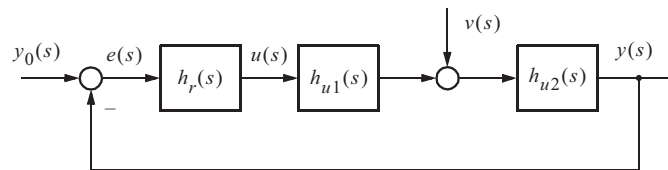
der $t_0(s)$ og $n_0(s)$ er polynomer i s , $p \geq 1$ og y_{00} er størrelsen på spranget i $y_0(t)$.

Følgeresponsen vil altså få null stasjonært avvik $e_{y_0}(t \rightarrow \infty)$, forutsatt at sløyfetransferfunksjonen $h_0(s)$ har minst en integrasjon ($p \geq 1$) uansett om denne finnes i regulatoren eller i prosessen (se tabell 9.1).

Det andre leddet i (9.86) inneholder faktoren $h_v(s)$. Dersom denne faktoren ikke inneholder noen integrasjon, vil resultatet bli det samme som for følgeresponsen ovenfor. Forstyrrelsesresponsen $e_v(t)$ etter et sprang i $v(t)$ vil da gå mot null når $t \rightarrow \infty$, forutsatt at sløyfetransferfunksjonen $h_0(s)$ har minst en integrasjon ($p \geq 1$), uansett om denne finnes i regulatoren eller i prosessen.

Det vil være annerledes dersom $h_v(s)$ inneholder en integrasjon (eller flere). Da må $h_0(s)$ ha minst to (eller flere) integrasjoner. Dersom en integrasjon finnes i prosessen $h_u(s)$, må det også legges en integrasjon i regulatoren $h_r(s)$.

Figur 9.58
Tilbakekoplet system med forstyrrelse



Ofte vil det være en tett sammenheng mellom transferfunksjonene $h_u(s)$ og $h_v(s)$. Dersom prosessens struktur er som vist i figur 9.58 $h_v(s) = h_{u2}(s)$, kan vi avlede en viktig men enkel konklusjon:

For at forstyrrelsesresponsen $e_v(t)$ etter et sprang i $v(t)$ skal gå mot null når $t \rightarrow \infty$, må det finnes en integrasjon foran angrepspunktet for forstyrrelsen, det vil si enten i $h_{u1}(s)$ eller i $h_r(s)$. Hvis integrasjonen ikke finnes i $h_{u1}(s)$ må den altså legges i regulatoren $h_r(s)$.

Integrerende regulatorer finnes i flere former:

- I-regulator, $h_r(s) = \frac{1}{T_i s}$
- PI-regulator, $h_r(s) = K_p \frac{1 + T_i s}{T_i s}$
- PID-regulator, $h_r(s) = K_p \frac{(1 + T_i s)(1 + T_d s)}{T_i s}$
- PI(begrenset)D-regulator, $h_r(s) = K_p \frac{(1 + T_i s)(1 + T_d s)}{T_i s(1 + \alpha T_d s)}$

Hvilken av disse formene som skal velges, avhenger av hva vi vil oppnå med hensyn til båndbredde og stabilitet og av prosessens dynamikk $h_u(s)$.

Vi skal gjøre det samme med e_v som gjøres med e_{y_0} i (9.87):

Vi definerer $h_v(s) = \frac{t_v}{h_v s^q}$, hvor $t_v(s)$ og $n_v(s)$ ikke har noen rene integratorer, og hvor q er et heltall, $q \geq 0$. Fra (9.86) har vi

$$\begin{aligned} e_v(s) &= -N(s)h_v(s)v(s) = -\frac{s^p n_0}{s^p n_0 + t_0} \cdot \frac{t_v}{n_v s^q} \cdot \frac{v_0}{s} \\ \Rightarrow \lim_{t \rightarrow \infty} e_v(t) &= \lim_{s \rightarrow 0} s \left[-\frac{s^p n_0}{s^p n_0 + t_0} \cdot \frac{t_v}{n_v s^q} \cdot \frac{v_0}{s} \right] \\ &= -v_0 \left(\lim_{s \rightarrow 0} \left[\frac{s^{p-q} n_0}{s^p n_0 + t_0} \cdot \frac{t_v}{n_v} \right] \right) = \begin{cases} 0, & p > q \\ -\infty < -v_0 \left(\frac{n_0}{t_0} \cdot \frac{t_v}{n_v} \right) \Big|_{s=0} < 0, & p = q \\ -\infty, & p < q \end{cases} \quad (9.88) \end{aligned}$$

Dette er en generalisering av det som står i de to nest øverste avsnittene over figur 9.58.

9.4.1 Sammenhengen mellom sluttverditheorem og frekvensrespons

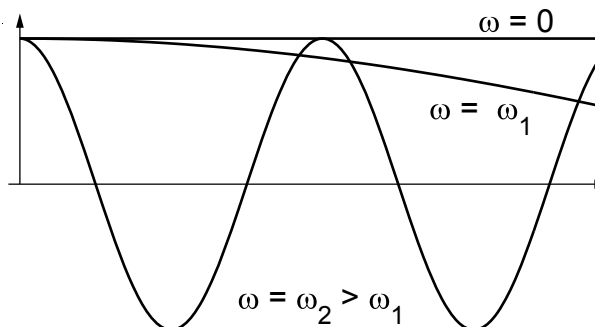
Gitt systemet $h(s)$ med pådrag $u(s) = \frac{1}{s}$ (sprang). Vi antar at $h(s)$ har alle poler i venstre halvplan, dvs. $h(s)$ er stabil og dermed har en frekvensrespons (jfr. avsnitt 6.2). Vi søker responsen $y(t = \infty)$. Sluttverditheorem (4.22) gir

$$\begin{aligned} y(t = \infty) &= \lim_{s \rightarrow 0} s [h(s)u(s)] = \lim_{s \rightarrow 0} s [h(s) \frac{1}{s}] \\ &= \lim_{s \rightarrow 0} h(s) = \underline{h(0)} \end{aligned}$$

La oss nå se om vi kan få samme resultat med en frekvensanalyse-betraktning:

Et sprang $\mu_1(t) \xrightarrow{\mathcal{L}} \frac{1}{s}$ kan betraktes som et grensetilfelle av $\cos \omega t \xrightarrow{\mathcal{L}} \frac{s}{s^2 + \omega^2}$ når frekvensen $\omega \rightarrow 0$:

Figur 9.59
Lav- og
høyfrekvent co-
sinus, blir sprang
når $\omega = 0$



Hvis $u(t) = \cos \omega t$ blir utgangssignalet en cosinus med amplitude $|h(j\omega)|$ og fase $\angle h(j\omega)$. Og når $\omega = 0$ (da blir $\cos \omega t$ et sprang $\mu_1(t)$), blir amplituden $|h(j0)| = |h(0)| = h(0)$. Siste overgang fordi alle ledd med s i $h(s)$ bortfaller, og da er

$$|h(0)| = h(0)$$

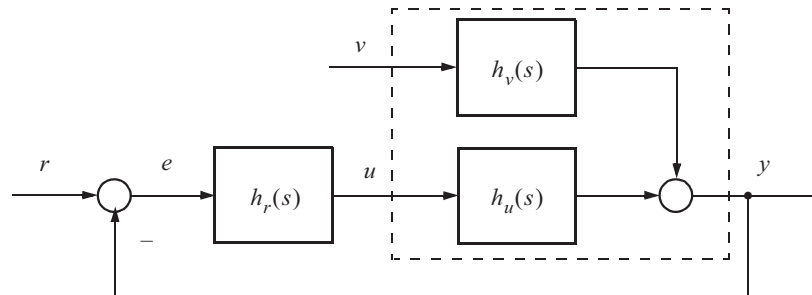
En frekvensanalytisk betraktning gir samme resultat som sluttverditeoremet.

(P.S: $\angle h(0) = 0$ fordi $h(j0)$ blir reell)

EKSEMPEL 9.12: Stasjonært avvik – utvidelse av eksempel 9.5

Gitt den generelle strukturen

Figur 9.60



I avsnitt 7.2.2 manipulerte vi blokkdiagrammet for å finne $\frac{e}{v}(s)$:

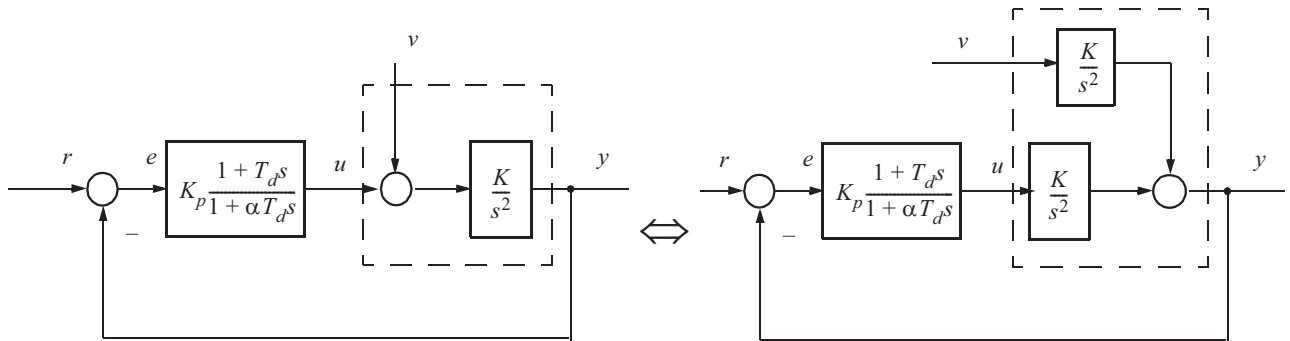
$$\frac{e}{v}(s) = h_v(s) \frac{(-1)}{1 - (-1)h_u(s)h_r(s)} = -h_v(s) \frac{1}{1 + h_0(s)} = -h_v(s)N(s)$$

Vi har også at

$$\frac{e}{r}(s) = \frac{1}{1 + h_u(s)h_r(s)} = N(s)$$

Vi utvider nå eksempel 9.5 til å inkludere en prosessforstyrrelse v , som kommer inn som vist under. Vi kan for eksempel tenke oss at v og u er kraft eller dreiemoment i et mekanisk system, og y er da posisjon.

Figur 9.61



Med $h_0 = t_0/n_0$, får vi $t_0 = K_p K(1 + T_d s)$ og $n_0 = (1 + \alpha T_d s)^2$. Vi har

$$N(s) = \frac{1}{1 + h_0(s)} = \frac{1}{1 + t_0/n_0} = \frac{n_0}{n_0 + t_0} = \frac{s^2(1 + \alpha T_d s)}{s^2(1 + \alpha T_d s) + K_p K(1 + T_d s)}$$

Anta nå at r og/eller v er et enhetssprang som kan betraktes som en cosinussvingning med frekvens ω som går mot 0. $\cos(\omega t)$ har Laplacetransform $s^2/(s^2 + \omega^2)$. Vi får

$$\left. \frac{e}{r}(s)r(s) \right|_{s=j0=0} = N(s)r(s) \Big|_{s=0} = 0, \text{ men}$$

$$\left. \frac{e}{v}(s)v(s) \right|_{s=0} = -h_v N(s)v(s) \Big|_{s=0} = -1/K_p$$

Vi merker oss at stasjonært avvik m.h.p. forstyrrelsen, i motsetning til referansen, ikke fjernes! Dette fordi forstyrrelsen angriper *foran* de to integratorene i prosessen, eller sagt på en annen måte: Det er integratorer i $h_v(s)$. Disse “oppveier” integratorene i $h_0(s)$.

Hvordan finne transferfunksjonen fra v til e ved hjelp av blokkdiagrammet er forklart i avsnitt 7.2.2.



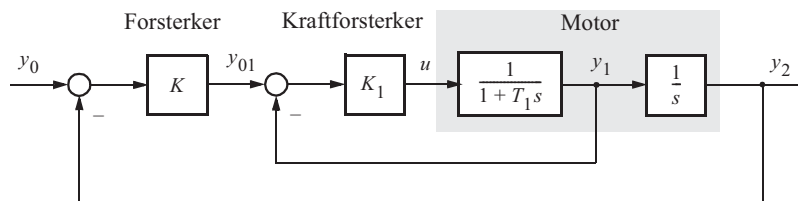
9.5 KOMPENSASJON MED INTERN TILBAKEKOPLING

I avsnitt 7.2.4 så vi at et monovariabelt tilbakekoplet systems egenskaper i betydelig grad er bestemt av tilbakekoplingen. Transferfunksjonen for det lukkede systemet opp til kryssfrekvensen er tilnærmet lik den inverse transferfunksjonen i tilbakekoplingsgrenen, det vil si tilnærmet lik 1 ved det viktige spesialtilfellet enhetstilbakekopling. Ved å etablere **intern(e) tilbakekopling(er)** i tillegg til den ene, ytre tilbakekopling som vi har studert til nå, kan man oppnå at deler av systemet får høyere båndbredde og bedre stabilitetsmarginer, noe som igjen gjør det mulig å oppnå tilsvarende forbedringer for det totale system.

EKSEMPEL 9.13: Likestrømsmotor- intern tilbakekopling

La oss igjen ta utgangspunkt i systemet i figur 4.26. Dersom vi innfører måling av rotasjonshastighet tillegg til vinkelposisjon for motoren, kan vi ha **intern tilbakekopling** fra rotasjonshastigheten. Dette gir følgende struktur, med to enhetstilbakekoplinger:

Figur 9.62
Blokkdiagram for enkel motormodell med P-regulator og intern tilbakekopling



Den fysiske prosess – motoren – er markert med skravert felt. Den har ett pådrag som før, men *to* målinger. Det som er utenfor det skraverte feltet, utgjør våre reguleringsanordninger. Vi har nå *to* enkle regulatorer – det er innført en forsterkning K_1 i en ekstra, indre sløyfe. K_1 er nå kraftforsterkeren som besørger pådraget u , mens K velges som en (“svakstrøms”) signalforsterker. Transferfunksjonen for den indre sløyfen, fra y_{01} til y_1 , er

$$h_1 = \frac{K_1}{T_1} \frac{1}{s + \frac{1 + K_1}{T_1}}$$

Transferfunksjonen $h_1(j\omega)$ er tilnærmet lik 1 for stor K_1 og $\omega \ll (1 + K_1)/T_1$. Dermed båndbredden for den ytre sløyfen er mye mindre enn $(1 + K_1)/T_1$, kan den ytre sløyfen designes under antakelsen $h_1 \approx 1$. Dette forenkler konstruksjonen av den ytre sløyfen. Videre gjør dette at man kan gi totalsystemet høyere båndbredde og/eller bedre stabilitetsmarginer enn hvis man bare hadde hatt en ytre sløyfe med seriekompensasjon.

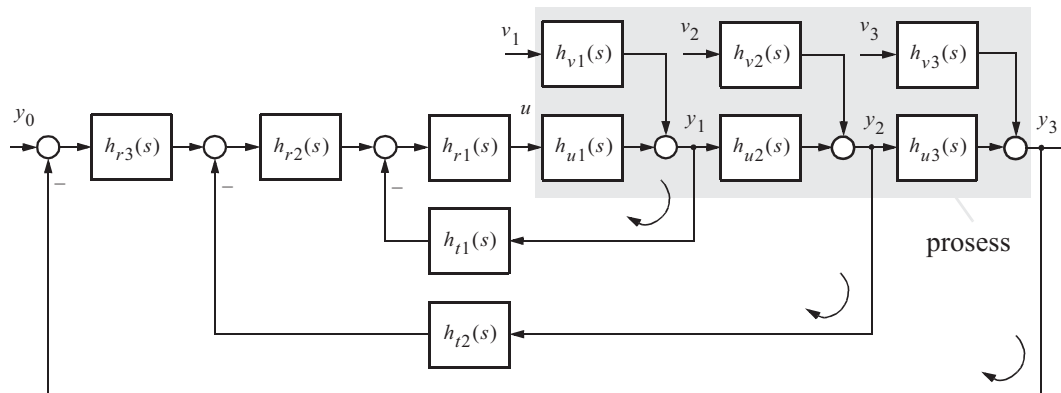
■

Vi skal nå presentere kompensasjon med intern tilbakekopling mer generelt. Vi begrenser oss til et system med to indre sløyfer. Figur 9.63 viser en struktur som omfatter en prosess, som i tillegg til den ordinære målingen y_3 på utgangen av prosessen, også har målinger “tidligere” i prosessen, y_1 og y_2 .

For denne prosessen kan vi skrive

$$h_u(s) = h_{u1}(s)h_{u2}(s)h_{u3}(s) \quad (9.89)$$

Figur 9.63
System med
interne tilbake-
koplinger



Fra de interne målingene y_1 og y_2 er det tilbakekoplinger henholdsvis gjennom transferfunksjonene $h_{t1}(s)$ og $h_{t2}(s)$. Den innerste av de interne tilbakekoplingene har en “regulator” med transferfunksjon $h_{r1}(s)$. Denne vil stort sett inneholde et forsterkende element, men kan også tillegges en viss funksjon av seriekompensasjon for å tilfredsstille krav til blant annet stabilitet av den innerste sløyfen.

Den neste interne sløyfen, som inneholder den innerste sløyfen som en komponent, har en “regulator” med transferfunksjon $h_{r2}(s)$. Denne vil stort sett også bestå av en forsterker, men med mulighet for seriekompensasjon.

Til slutt vil hovedsløyfen inneholde en “regulator” $h_{r3}(s)$ som i de fleste tilfeller er en forsterker.

På basis av det som er utledet i kapittel 7, og strukturen i figur 9.63, finner vi nå at hovedsløyfens transferfunksjon utgjøres av produktet av transferfunksjonen for det

siste avsnittet av prosessen, den ytterste undersløyfen og hovedsløyfens “regulator”. Vi finner derfor

$$h_0(s) = h_{u3}(s)M_2(s)\frac{1}{h_{r2}(s)}h_{r3}(s) \quad (9.90)$$

der $M_2(s)$ er den ytterste undersløyfens lukkede transferfunksjon med enhetstilbakekopling, gitt av

$$M_2(s) = \frac{h_{u2}(s)M_1(s)\frac{1}{h_{r1}(s)}h_{r2}(s)h_{r2}(s)}{1 + h_{u2}(s)M_1(s)\frac{1}{h_{r1}(s)}h_{r2}(s)h_{r2}(s)} \quad (9.91)$$

Transferfunksjonen $M_1(s)$ for den innerste undersløyfen med enhetstilbakekopling er gitt av

$$M_1(s) = \frac{h_{u1}(s)h_{r1}(s)h_{r1}(s)}{1 + h_{u1}(s)h_{r1}(s)h_{r1}(s)} \quad (9.92)$$

Likning (9.90) viser at transferfunksjonen $h_{r2}(s)$ er avgjørende for formen av hovedsløyfens transferfunksjon fordi $h_{u3}(s)$ er gitt, $M_2(s) \approx 1$ og $h_{r3}(s)$ skal fortrinnsvis bare være en konstant forsterkning. Vi legger merke til at transferfunksjonene $h_{u1}(s)$ og $h_{u2}(s)$ spiller en mer underordnet rolle i dannelsen av $h_0(s)$ fordi de bare inngår i $M_1(s)$ og $M_2(s)$.

Dersom vi ut fra spesifikasjonene til systemet og hensyn til realiserbarhet har fått fram en ønsket form for $h_0(s)$, kan vi av (9.90) finne hvilken form $h_{r2}(s)$ må ha. Vi kan dermed arbeide oss innover til den innerste sløyfen ved bruk av (9.91) og ut fra tilsvarende argumentasjon finne hvilken form $h_{r1}(s)$ må ha.

I de fleste tilfellene er det egentlig best å starte med konstruksjon av den innerste undersløyfen, men valget av transferfunksjoner i tilbakekoplingen i denne sløyfen er ofte avhengig av funksjonen den neste sløyfen skal ha.

De vesentligste fordelene og ulempene med løsningen som er vist i figur 9.56, sammenliknet med en ren seriekompensasjon i hovedsløyfen som beskrevet i avsnitt 9.3, er:

- Forstyrrelser som kommer inn “tidlig” i prosessen (v_1 og v_2) vil bli fanget opp av de interne tilbakekoplingene og raskt motvirket, fordi undersløyfen kan få høyere båndbredde enn hovedsløyfen.
- Systemet er mindre følsomt for varierende parametre og ulineariteter i prosessen enn et system med seriekompensasjon, fordi en undersløyfes transferfunksjon stort sett er bestemt av tilbakekopplingselementets transferfunksjon.

- I noen tilfeller er kompensasjon med intern tilbakekopling den eneste praktiske mulighet, for eksempel i elektriske systemer der signalene er modulerte.
- Et system med interne tilbakekoplinger vil som oftest være noe mer ressurskrevende å utvikle enn et system med bare seriekompensasjon.

Et eksempel vil illustrere det som her er utledet. Et viktig designprinsipp for intern tilbakekopling er at den indre sløyfen alltid har høyere båndbredde enn den ytre sløyfen.

EKSEMPEL 9.14: Intern tilbakekopling i servomekanisme for antennestyring

Vi benytter det samme systemet som tidligere er innført i eksempel 9.10, og som omhandler et følgesystem for drift av en antenne. Figur 9.49 antyder at det på to steder midt i prosessen er mulig å utføre målinger av viktige interne variable (tilstandsvariable) som egner seg for intern tilbakekopling. Disse punktene er merket med C og D og gjelder henholdsvis størrelsene $z_2 =$ stempelbevegelse og $z_3 =$ hastighet av den hydrauliske motoren.

Vi vil først forsøke om det er tilstrekkelig å bare benytte en intern tilbakekopling og begynner med å velge tilbakekopling fra størrelsen z_2 . Vi velger altså $y_1 = z_2$. Senere skal vi også utnytte tilbakekopling fra størrelsen z_3 . I eksempel 9.10 krevde vi at den åpne hovedsløyfen skulle ha *to integrasjoner*, for at systemet skulle kunne følge et mål med konstant hastighet uten posisjonsavvik. Figur 9.52 viser at det ligger en integrasjon rett foran størrelsen z_2 og en helt til høyre i diagrammet. Skal vi legge en tilbakekopling fra z_2 , må tilbakekoplingens transferfunksjon $h_{t1}(s)$ derfor ha et nullpunkt i origo. Denne interne tilbakekoplingen vil da bidra med en pol i origo som ønsket. Da transferfunksjonen mellom z_2 og φ_2 inneholder en pol i origo og to reelle poler, er det nødvendig at den interne tilbakekoplingen bidrar med proporsjonal+derivatvirkning for å oppnå stabilitet.

Det er derfor fornuftig å velge

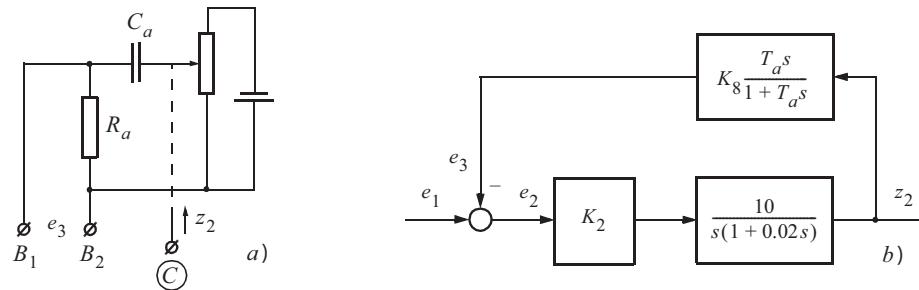
$$h_{t1}(s) = K_8 \frac{T_a s}{1 + T_a s} \quad (9.93)$$

Fordi vi foreløpig ikke har tatt standpunkt til hvilken forsterkning vi skal ha i vekselspenningsforsterkeren med demodulator (K_1), står vi noenlunde fritt ved valg av forsterkningen K_8 i (9.93). Vi velger derfor for enkelhets skyld $K_8 = [\text{V/mm}]$. Et elektromekanisk element som vil utføre den funksjon som er foreskrevet i (9.93), er antydnet i figur 9.64 a).

Elementet består av et potensiometer hvis skyver bevegges av stempelstangen (punkt C), og hvis lengde er lik den maksimale forskyvningen av stemplet. Over

Figur 9.64

- a) Implementasjon av h_{t1}
 b) Blokkdiagram over tilbakekopplingsløyfen



potensiometeret sitter en spenningskilde som er avpasset slik at spenningen på skyveren vil variere med 1 [V/mm]. Etter potensiometeret følger et RC-nettverk som gir transferfunksjonen i (9.93) med $T_a = R_a C_a$. Et lite blokkdiagram som gjelder for den nye overføringen mellom demodulatoren og sylinderposisjonen er vist i figur 9.64. Vi ønsker nå å drive forsterkningen K_2 i likestrømsforsterkeren så høyt at denne overføringen transferfunksjon i hovedsak er bestemt av tilbakekopplingselementet, og at den indre sløyfen får stor båndbredde. Ved å sammenlikne med det som ble utledet under seriekompensasjon av dette systemet, vet vi at verdien av T_a i (9.93) som vi må fram til, vil være i området $1 < T_a < 10$ [s]. For frekvenser høyere enn $\omega = 1$ [rad/s] vil derfor i alle fall tilbakekopplingselementet i figur 9.64 bli tilnærmet frekvensuavhengig ($h_{t1}(j\omega) \approx 1$). En rimelig verdi for K_2 vil derfor være den som gir en kryssfrekvens for undersløyfen lik

$$\omega_c \approx \frac{1}{T_1} = \frac{1}{0.02}$$

der T_1 er tidskonstanten for kraftmotoren. Vi finner da å måtte velge $K_2 = 5 = 14$ [dB]. Overføringen for undersløyfen blir derfor

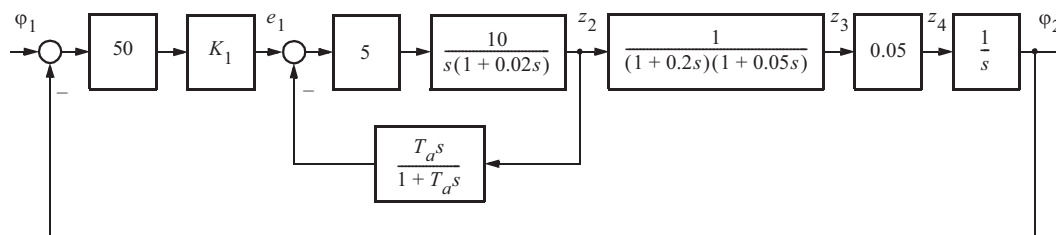
$$\frac{z_2}{e_1}(s) = \frac{1 + T_a s}{T_a s} M_1(s) \left[\frac{\text{mm}}{\text{V}} \right] \quad (9.94)$$

der $|M_1(j\omega)| \approx 1 = 0$ [dB] for alle frekvenser opp til $\omega = 50$ [rad/s]. Det resulterende systemet er vist i figur 9.65.

Kombineres (9.94) med de øvrige elementene i systemet og antar vi foreløpig $K_1 = 1 = 0$ [dB], kan vi tegne amplitude/fase/frekvens-diagrammet vist i figur 9.66.

Vi har her valgt $T_a = 5$ [s] som svarer til derivattiden i (9.85). Forlanger vi $|N(j\omega)|_{\max} \approx 6$ [dB] i hovedsløyfen (tilsvarer her ca. 40° fasemargin), kan vi legge inn en forsterkning $K_1 = 2 = 6$ [dB] i sløyfen. Den resulterende 0-dB-linjen er vist i figur 9.66.

Figur 9.65
System med en intern tilbakekopling

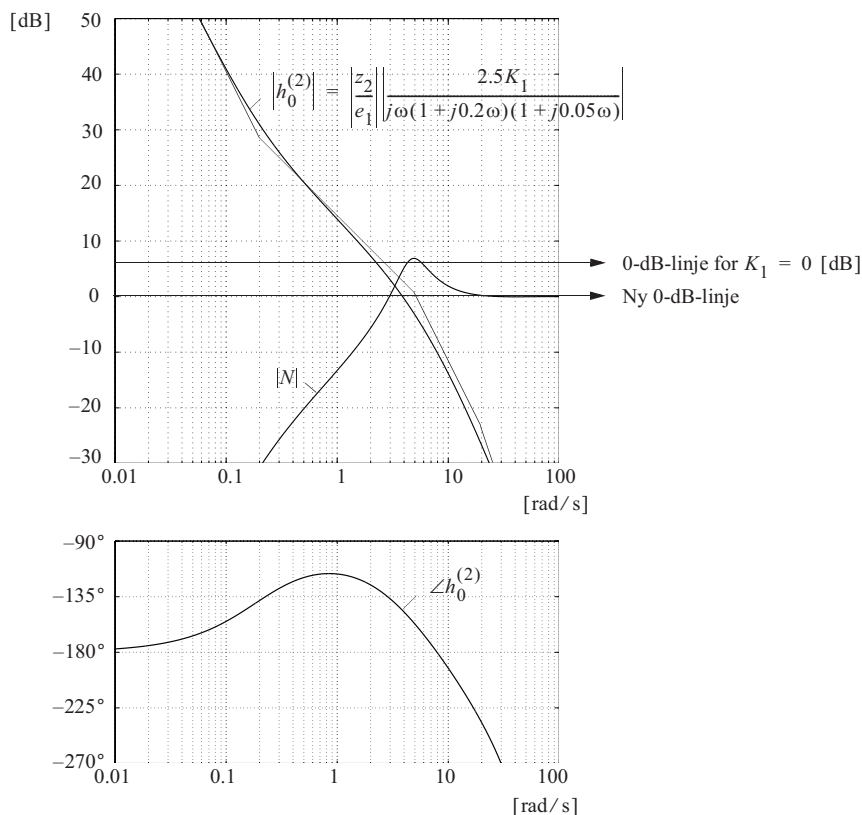


Med innsetting av tallverdiene som nå er bestemt, vil transferfunksjonen i hovedsløyfen bli

$$h_0(s) = \frac{1 + 5s}{s^2(1 + 0.2s)(1 + 0.05s)} M_1(s) \quad (9.95)$$

Figur 9.66 viser at vi med denne løsningen har fått en båndbredde i systemet som er omtrent dobbelt så stor som den vi oppnådde med seriekompensasjon, og som er vist i figur 9.50. Det vesentligste er imidlertid at systemet med intern tilbakekopling er bedre fordi uheldige virkninger av ulineariteter, støy og variabel forsterkning i likestrømsforsterker, kraftmotor, styreventil og sylinder nå er så godt som eliminert.

Figur 9.66
AFF-diagram



En forskjell ved den nye løsningen i forhold til seriekompensasjon er at vi nå må sørge for forsterkningen $K_1 K_2 = 10 = 20$ [dB] der vi ved seriekompensasjon fikk $K_1 K_2 = 1 = 0$ [dB].

Det finnes en rekke andre måter å anordne en tilsvarende tilbakekopling fra punktet C på, men i prinsippet må de fungere som antydnet foran. Et detaljert studium av hvor stor tidskonstanten T_a i tilbakekoplingen bør velges, vil nok avsløre at en noe mindre verdi enn den som her er valgt, kan benyttes. Resultatet av en mindre T_a vil være at den transiente etterslepningen av antennevinkelen $\varphi_2(t)$ i forhold til den jevnt økende referansevinkelen $\varphi_1(t) = kt$, vil bli av kortere varighet. Gjøres T_a for liten, vil imidlertid systemet bli oscillatorisk, som vist foran.

Vi kunne nå undersøke hvordan en tilbakekopling bare fra punktet D i figur 9.49 til punktene $B_1 - B_2$ ville arte seg. Dette skal ikke gjøres her. Vi skal derimot innføre en tilbakekopling fra punktet D i tillegg til den interne tilbakekoplingen fra punktet C som nettopp ble beskrevet.

Tilbakekoplingen fra punktet D dannes ved hjelp av et elektrisk likespenningstachometer tilkoplek motorakselen. Da vi fremdeles er interessert i å beholde begge integrasjonene i hovedsløyfen, må også denne tilbakekoplingen gjøres forbigående (det vil si at dens transferfunksjon må ha et nullpunkt i origo). Figur 9.67 viser hvordan dette kan gjøres. Spenningen fra tachometeret som er proporsjonal med motorhastigheten z_3 , tilføres et RC-nettverk. En del av spenningen over motstanden tas ut og føres tilbake i serie med den første interne tilbakekoplingen. Kalles den spenningen som tas ut mellom klemmene $B_1 - B_2$ for e_4 , finner vi

$$\frac{e_4}{z_3}(s) = h_{t2}(s) = K_t K_9 \frac{T_b s}{1 + T_b s} \quad (9.96)$$

der

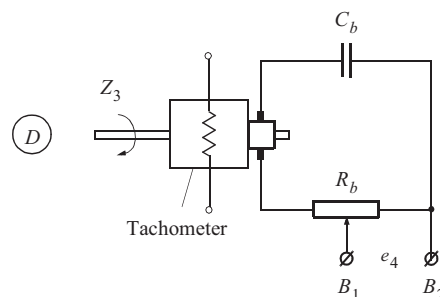
$$K_t = \text{tachometerkonstant} = 5 \left[\frac{\text{V}}{\text{rad/s}} \right]$$

$$K_9 = \text{potensiometerkonstant} \leq 1 \left[\frac{\text{V}}{\text{V}} \right]$$

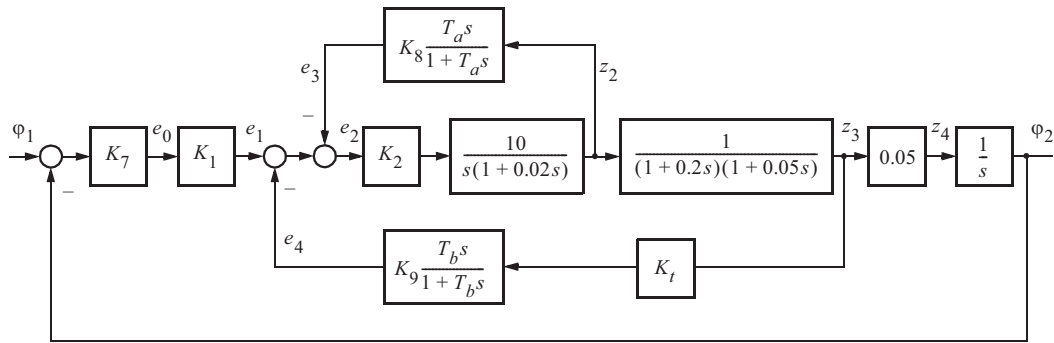
Et komplett blokkdiagram for hele systemet med to interne tilbakekoplinger er vist i figur 9.68. Signalene e_3 i figur 9.64 a) og e_4 i figur 9.67 seriekoples.

Figur 9.67

Implementasjon av tilbakekopling fra D

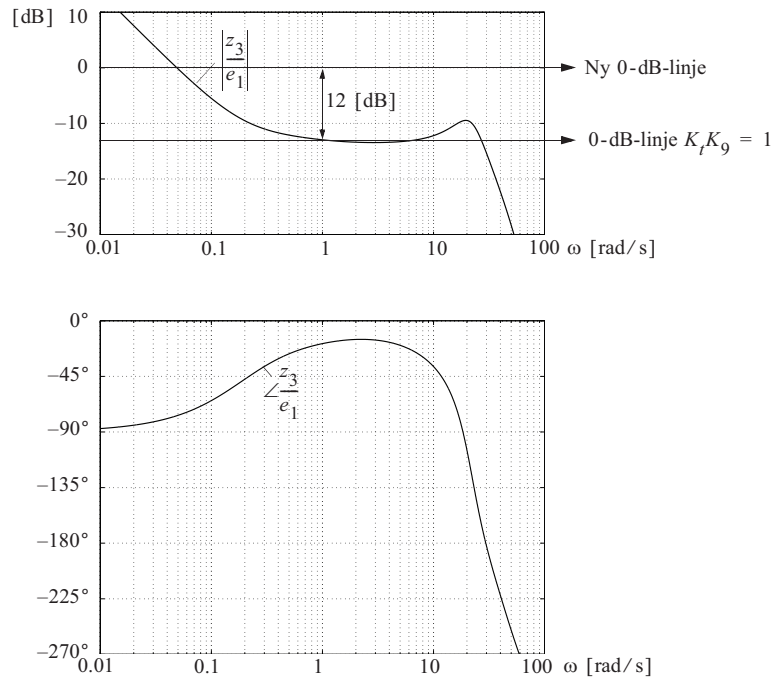


Figur 9.68
Systemet med to
interne tilbake-
koplinger



Fordi vi nå har fått en ny intern tilbakekopling, vil det være mulig å redusere tidskonstanten T_a i den opprinnelige interne tilbakekoplingen atskilleg. Hensikten med dette er å få så stor forsterkning som mulig i den nye interne sløyfen ved lave frekvenser. En rimelig verdi av T_a vil nå være $T_a = 1$ [s]. Tidskonstanten T_b i den nye tilbakekoplingen kan nå velges like stor som vi før hadde T_a , altså $T_b = 5$ [s]. Med et gitt tachometer kan forsterkningen i den nye interne sløyfen bare justeres med koeffisienten K_9 . Dette betyr at vi må sørge for at tachometerkonstanten K_t er stor nok. Vi justerer da K_9 til tachometersløyfen får størst mulig båndbredde, med for eksempel 45° fasemargin. Overføringen for kombinasjonen av de to interne sløyfene vil da bli som vist i figur 9.69.

Figur 9.69
AFF-diagram for
kombinasjonen av
de to interne
tilbakekoplingene

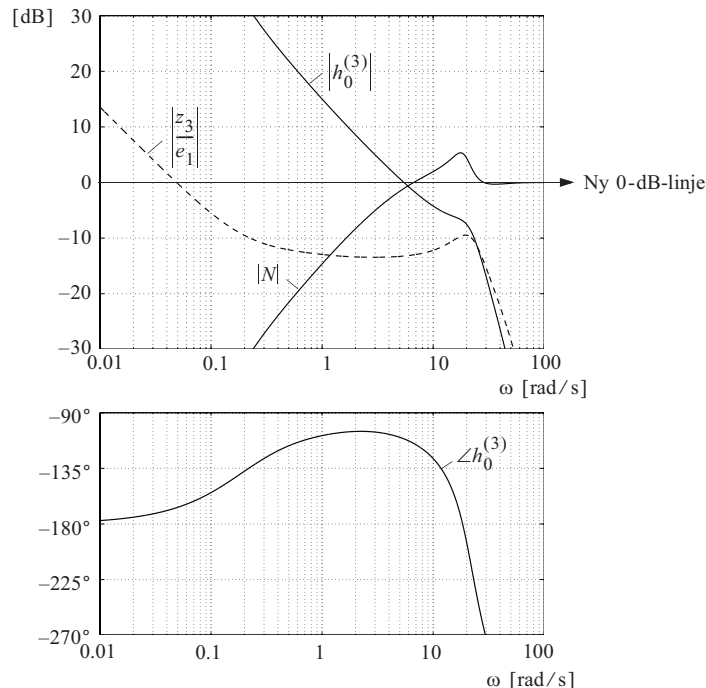


En rimelig verdi av forsterkningen i tilbakekoplingen blir da

$$K_7 K_9 = 4 \left[\frac{\text{V}}{\text{rad/s}} \right] = 12 \text{ [dB]} \quad (9.97)$$

Dette resultatet kombineres nå med de få resterende elementene i sløyfen (i det vesentlige en integrasjon), og vi får hovedsløyfens amplitude/fase/frekvens-diagram som vist i figur 9.70.

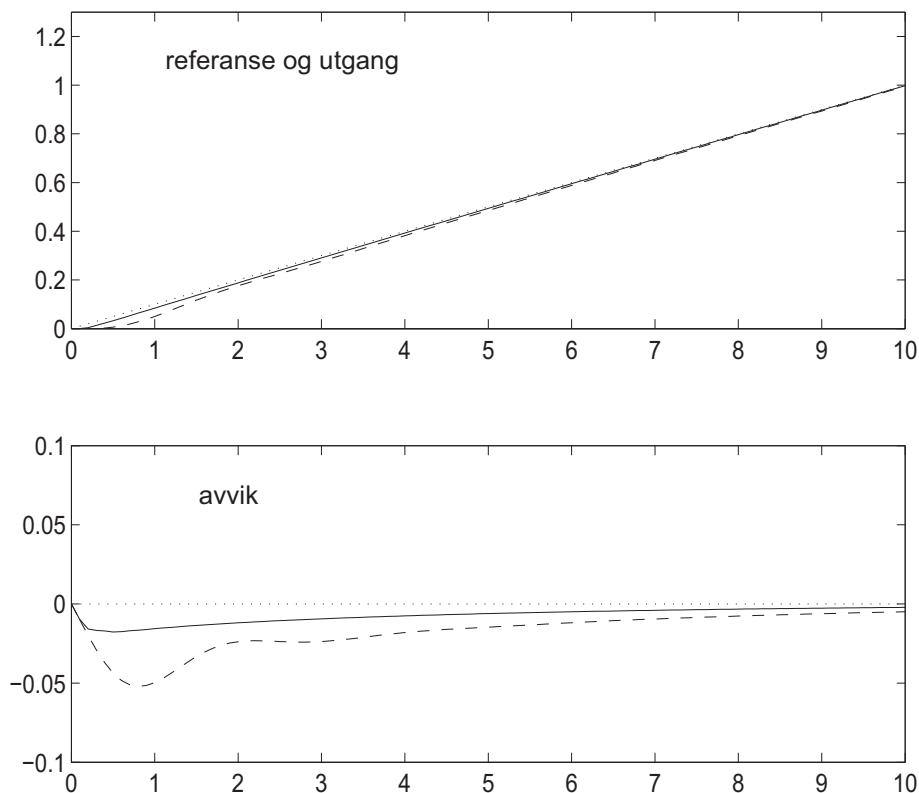
Figur 9.70
Hovedsløyfens
amplitude/fase/fre-
kvens-diagram



Vi ser at vi i alle fall kan legge inn en forsterkning $K_1 = 10 = 20$ [dB] uten at hovedsløyfens stabilitetsmargin kommer i fare. Vi oppnår derved en kryssfrekvens for hovedsløyfen $\omega_c \approx 6$ [rad/s] med $|N(j\omega)|_{\max} \approx 6$ [dB]. Dette er en vesentlig forbedring i systemets hurtighet i forhold til det vi kan oppnå med de enklere løsningene som er demonstrert tidligere. Dette systemet vil forøvrig ha den egenskapen at eventuelle forandringer i den hydrauliske transmisjonen (hydraulisk pumpe og motor) får mindre betydning, idet tachometersløyfen har gjort systemet mindre følsomt for slike endringer. Innføring av et tachometer i systemet medfører en viss fordyring, men i et tilfelle som dette vil det neppe være av særlig betydning idet de øvrige komponentene er mye mer kostbare. Vi ser også at vi med denne løsningen må sørge for en total forsterkning gitt av $K_1 K_2 = 50 = 34$ [dB].

La oss til slutt sjekke hva vår løsning, funnet ved syntese i frekvensplanet, fører til i tids-planet. Figur 9.71 viser responsen når referansen er en rampefunksjon. Som forventet går avviket mot null, fordi vi har sørget for å beholde de to integrasjonene mellom referanse og utgang. Vi ser også at løsningen med intern tilbakekopling (heltrukken linje) er raskere enn løsningen med bare seriekompensasjon (heltrukken linje). Dette kunne vi forvente p.g.a. høyere båndbredde med intern tilbakekopling.

Figur 9.71
Respons på
rampefunksjon



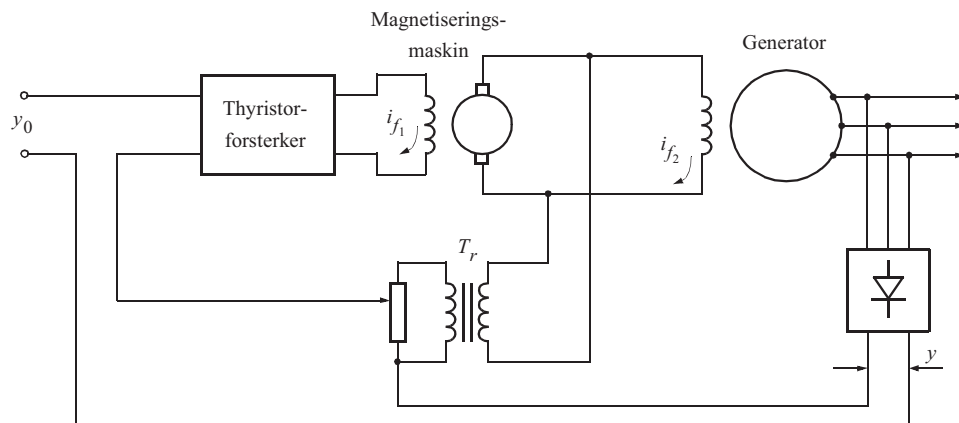
EKSEMPEL 9.15: Intern tilbakekopling ved spenningsregulering av vekselstrømsgeneratorer

Spenningsregulering av store vekselstrømsgeneratorer i et kraftnett er et annet typisk eksempel på bruk av intern tilbakekopling ved oppbygging av et reguleringsystem.

Hensikten med spenningsreguleringen er framfor alt å holde generatorens klemmespenning konstant ved varierende belastning, eventuelt at generatoren ved hjelp av spenningsreguleringsystemet får en utgangsimpedans som er hensiktsmessig og innstillbar. Generatorens utgangsimpedans er avgjørende for genereringen av reaktiv effekt i et samkjøringssystem, og for den enkelte forbruker er variasjonen i spenningen på nettet ofte av stor betydning.

Vekselstrømsgeneratorer for kraftnett leveres i dag med ytelser opp i flere hundre megawatt. Magnetiseringseffekten utgjør vanligvis et par prosent av generatoreffekten, det vil si at vi kan få magnetiseringseffekter på flere megawatt. For en del år tilbake var det nødvendig å benytte en egen likestrømsgenerator koplet til samme aksel som vekselstrømsgeneratoren for levering av magnetiseringseffekt. Utviklingen av moderne halvlederteknologi har imidlertid gjort det mulig å sløyfe den roterende magnetiseringsmaskinen og i stedet lage en statisk magnetiseringsenhet som gjør det hele i ett trinn. Det regulerings tekniske problemet blir derved mindre enn før, fordi reguleringsløyfen vil inneholde færre dynamiske elementer. Fordi det framdeles er aktuelt med magnetisering ved hjelp av en egen magnetiseringsmaskin, men framfor alt fordi reguleringsproblemet da er mer interessant, skal vi se på det systemet som er vist i figur 9.72.

Figur 9.72
Spenningsregulering av vekselstrømsgenerator i kraftnett



Til høyre i figuren er vekselstrømsgeneratoren som drives av en turbin. Den magnetiseres av en likestrømsgenerator som igjen magnetiseres fra en thyristorforsterker. Vekselstrømsgeneratorens tre-fase klemmespenninger måles, likerettes og sammenliknes med en likespenningsreferanse. Avviket mellom disse spenningene føres til

triggerkretsen for thyristorforsterkeren. Noe forenklet kan vi anta at transferfunksjonen fra thyristorforsterkerens inngang til feltstrømmen i magnetiseringsmaskinen i_{f1} er karakterisert ved en reell pol der tidskonstanten er T_1 . Transferfunksjonen fra i_{f1} til vekselstrømsgeneratorens magnetiseringsstrøm i_{f2} vil likeledes ha en reell pol med tidskonstant T_2 . Dersom vi tillater oss å neglisjere en rekke detaljfenomener i vekselstrømsgeneratoren, kan vi regne med at spenningen på klemmene er omtrent proporsjonal med feltstrømmen. Vi har med andre ord en sløyfe som inneholder to reelle poler. Forholdet mellom disse polenes tidskonstanter

$$\frac{T_2}{T_1}$$

vil bestemme hvor stor proporsjonal forsterkning vi kan legge inn i sløyfen.

Ved hjelp av en intern tilbakekopling over en spesielt bygget transformator (TR), som angitt i figur 9.72, kan vi nå forme de dynamiske egenskapene til sløyfen på en gunstig måte. Magnetiseringsmaskinen inklusive tilbakekoplingen kan gjerne betraktes som systemets “regulator”. Tilbakekoplingstransformatoren vil ha en transferfunksjon

$$h_{t1}(s) = \frac{K_3 T_3 s}{1 + T_3 s} \quad (9.98)$$

Transferfunksjonen for “regulatoren” blir derfor

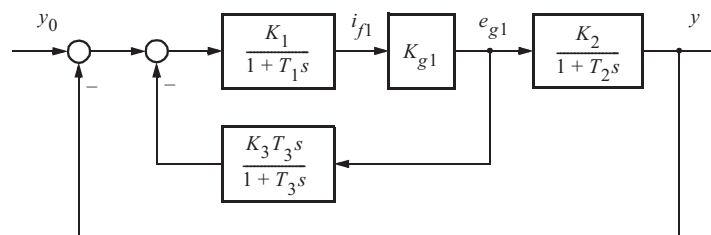
$$\begin{aligned} h_r(s) &= \frac{1}{h_{t1}(s)} M_1(s) \\ &= \frac{1}{K_3} \frac{1 + T_3 s}{T_3 s} M_1(s) \end{aligned} \quad (9.99)$$

$|M_1(j\omega)| \approx 1$ for frekvenser som er vesentlig høyere enn det endelige systemets kryssfrekvens. Den totale sløyfens transferfunksjon blir

$$h_0(s) = \frac{1}{K_3} \frac{1 + T_3 s}{T_3 s} \frac{K_2}{1 + T_2 s} M_1(s) \quad (9.100)$$

Et blokkdiagram er vist i figur 9.73.

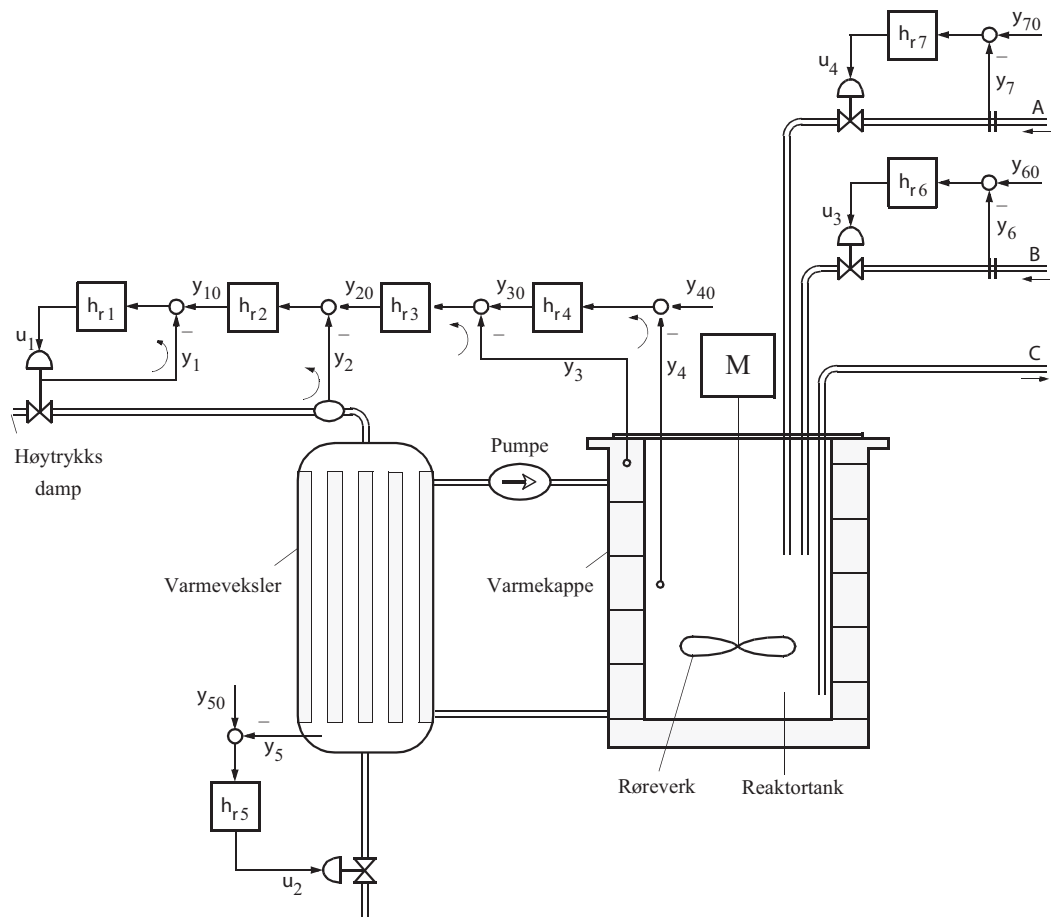
Figur 9.73
Vekselstrømsregulator med spenningsregulering



EKSEMPEL 9.16: Kjemisk reaktor med kaskaderegulering

Ved prosessregulering vil kompensasjon med intern tilbakekopling gjerne gis betegnelsen *kaskaderegulering*.

Figur 9.74
Kjemisk reaktor
med kaskade-
regulering



Figur 9.74 viser en ofte forekommende anvendelse av kaskaderegulering. En kjemisk reaktor består av en lukket tank som på toppen tilføres de to stoffene (A og B), som skal reagere med hverandre ved at de tilføres varme, og resulterer i stoffet C. Oppvarmingen av reaktoren skjer ved hjelp av vann som omslutter reaktortanken, og som varmes opp i en damp/vann-varmeveksler. I reaktortanken er det et kraftig motordrevet røreverk som sørger for god omrøring. Varmerveksleren tilføres mettet vanddamp fra et høytrykks dampnett. Prosessen er *endoterm*, det vil si den forbruker varme. Hadde den vært *eksoterm*, ville den produsert varme i overskudd og da kunne den fordret avkjøling i stedet for oppvarming.

Reaksjonen forløper best ved en bestemt temperatur som må reguleres presist. Dette oppnås ved å måle temperaturen y_4 i reaktortanken, og gjennom regulatoren h_{r4} styre referansen i et reguleringsystem som påvirker temperaturen i det varme vannet fra varmeveksleren idet det går inn i reaktorens varmekappe.

Damp/vann-varmevekslerens effektnivå bestemmes av temperaturen på vanddampen. I *mettet* vanddamp er det en helt presis sammenheng mellom damptrykket og damp-temperaturen (damptrykk 1 bar svarer til temperatur 100°C , 5 bar til 150°C , 16 bar til 200°C). Det er derfor fornuftig å styre damptrykket ved hjelp av et reguleringsystem som tar utgangspunkt i trykket y_2 og med regulatoren h_{r2} regulerer referansen for et lite reguleringsystem som sørger for en presis bevegelse av dampventilen y_1 .

Behovet for dette lille reguleringsystemet (ventilforstiller) dikteres av at det i en dampventil for høyt trykk vil kunne være stor friksjon på grunn av pakninger som kan gi stor **hysterese** i ventilbevegelsen. Hysterese er et ulineært fenomen som oppstår på grunn av slark i ventilen. Denne effekten finnes spesielt i mekaniske overføringer som tannhjul, ledd og liknende.

I tillegg til de fire regulatorene som inngår i kaskadereguleringen i figur 9.74, er det også vist tre andre reguleringsystemer på samme prosess. I bunnen av varmeveksleren finnes et lite reguleringsystem h_{r5} som måler nivået av vanddampkondensatet, og styrer ventilen i bunnen. Denne slipper ut vann etter hvert som det dannes ved kondensasjon av damp. Uten dette reguleringsystemet ville varmeveksleren enten gå tom for kondensat i bunnen og derved blåse damp tvers igjennom, eller den ville kunne gå full av vann og derved ikke virke tilfredsstillende.

Tilførselsledningen for de to stoffene (A og B) som skal reagere i reaktortanken, er utstyrt med strømningsregulering (regulatorene h_{r6} og h_{r7}) for å sikre at den kjemiske støkiometrien er tilfredsstillt (her vil det også være aktuelt med forholdsregulering som beskrives i figur 9.86).

Målingen y_3 i figur 9.74 kunne vært plassert på andre steder enn vist i figuren, for eksempel i *utløpet* av varmekappen i stedet for *innløpet*. Resultatet av dette ville vært en forskyvning av dynamikken i prosessen slik at det ble mer forsinkelse i sløyfe 3 og mindre forsinkelse i sløyfe 4 enn det som er vist i figur 9.74.



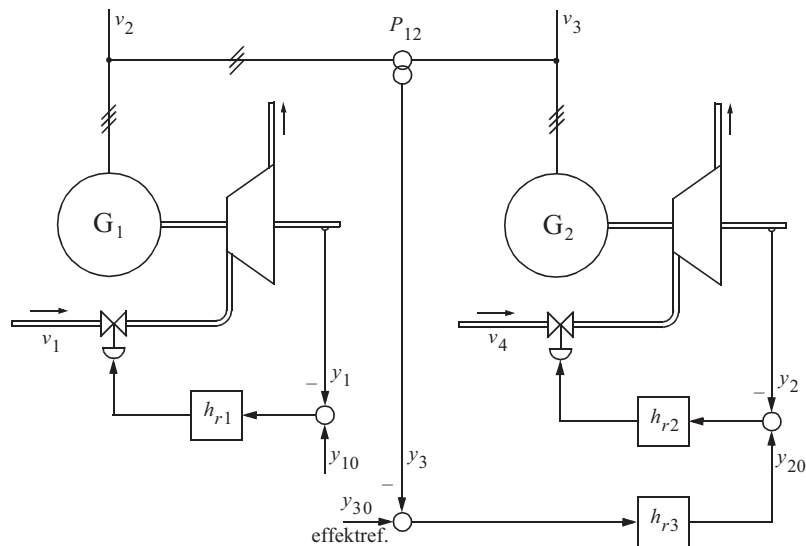
EKSEMPEL 9.17: Kaskaderegulering av turbiner

Figur 9.75 viser en annen typisk kaskaderegulering. I figuren er det to turbin-generatorsatser i forskjellige kraftstasjoner. Hver turbin er utstyrt med sin tur-tallsregulator (**turbinregulator**). Foruten å være tilkoplest sine lokale forbrukernett, er de to kraftstasjonene sammenbundet med en samkjøringslinje.

Den effekten som stasjonen til høyre leverer over samkjøringslinjen, skal nå reguleres. Effekten måles av elektrisk vei, og ved hjelp av en egen regulator h_{r3} påvirkes referansen for turbinregulatoren til høyre.

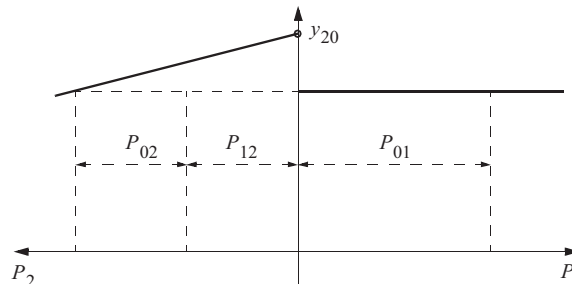
Siden de to generatorene er koplet sammen, må de gå med *samme frekvens*. Stilles det store krav til frekvensnøyaktigheten, kan det være at regulator h_{r1} utføres som en PI-regulator (isodrom-regulator). Regulatoren h_{r2} må i så fall være en P-regulator. Forstyrrelsene som kommer inn og nødvendiggjør de forskjellige regulatorene, er varierende last i de to nettene og varierende trykk foran turbinene (vanntrykk ved vantturbiner og damptrykk ved dampturbiner). I tillegg til de tre reguleringsystemene som her er antydnet, vil det også være en spenningsregulator på hver generator som imidlertid ikke har vesentlig innflytelse i denne forbindelse.

Figur 9.75
Regulering av
turbiner i sam-
kjøringsnett



Figur 9.76 viser et statisk diagram som gir sammenhengen mellom frekvensen i samkjøringsnettet ($y_1 = y_2$) og effekten P_i som hver generator leverer. Det er her antatt at stasjon 1 er frekvensbestemmende. h_{r1} er derfor en PI-regulator. Effekten som flyter fra nett 2 til nett 1 kalles P_{12} , og lokalforbruket i nett 2 er P_{02} . Turbinregulatoren i nett 2 er en P-regulator. Dette medfører at sammenhengen mellom frekvens og effekt er en fallende karakteristikk (statikk) som kan forskyves opp og ned ved hjelp av y_{02} .

Figur 9.76
Statisk diagram

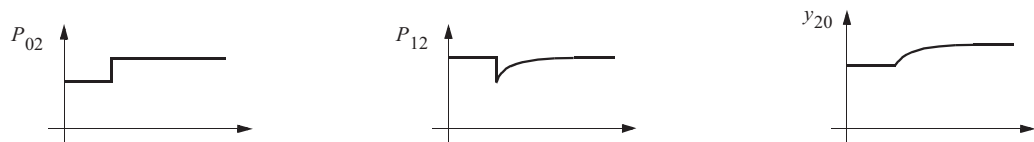


En typisk situasjon kan være som følger:

- Det er kontraktmessig bestemt at stasjon 2 skal levere effekten P_{12} til nett 1. Siden det er et lokalt forbruk P_{02} , må regulatoren h_{r3} forskyve referansen y_{02} som vist i figur 9.76.
- En endring i lokalforbruket P_{01} i nett 1 vil fanges opp av stasjon 1 ved endring av P_1 .
- Dersom det kommer en plutselig endring i lokalforbruket P_{02} , vil denne i første omgang fanges opp av stasjon 1, men derved synker den overførte effekten P_{12} slik at “statikken” må løftes ytterligere opp av h_{r3} . Figur 9.77 gir en skisse av denne situasjonen.

Figur 9.77

Resultat ved sprang i P_{02}



9.6 FOROVERKOPLING I MONOVARIABLE SYSTEMER

Vi har introdusert begrepet *foroverkopling* i avsnitt 1.4 og 9.2. Vi skal betrakte to forskjellige anvendelser av foroverkopling:

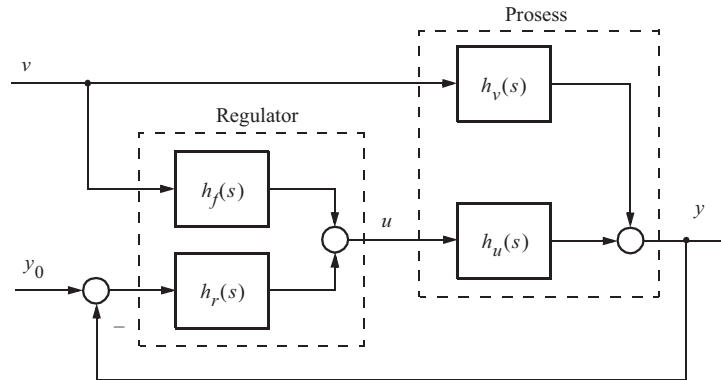
- *Foroverkopling i prosessreguleringssystemer:*
Prosessreguleringssystemer er systemer hvor referansen gjerne er konstant (betegnelsen settpunkt brukes ofte i stedet for referanse). Den primære oppgaven til reguleringssystemet er derfor å undertrykke forstyrrelsers innvirkning på prosessen. *Foroverkopling fra en målt forstyrrelse* kan være aktuelt for slike systemer. Et eksempel på foroverkopling i prosessregulering er såkalt *forholdsregulering*.
- *Foroverkopling i følgesystemer:*
Dersom en utgang i et system skal følge en referanse som varierer med tiden, benytter vi ofte betegnelsen følgesystem. Et følgesystems ideelle transferfunksjon vil derfor være lik 1. Med foroverkopling mellom referansen og pådraget forsøker vi å “oppeve” noe av dynamikken i prosessen.

Foroverkopling er en typisk *modellbasert* reguleringsform idet den forutsetter at vi har realisert en matematisk modell av prosessen og bruker denne til å regne ut hvordan pådraget må endres for å oppveie målbare ytre påvirkninger på prosessen.

9.6.1 Foroverkopling (fra forstyrrelsen) i prosessreguleringssystemer

Figur 9.78 viser hvordan foroverkopling kombinert med tilbakekopling kan utformes.

Figur 9.78
Tilbakekoplet system med foroverkopling fra forstyrrelsen



Det antas at forstyrrelsen på prosessen kan måles. Dette signalet benyttes, sammen med y og y_0 , til å beregne pådraget. Vi observerer videre at forstyrrelsen påvirker utgangen gjennom transferfunksjonen $h_v(s)$.

Hensikten med foroverkoplingen er å generere endringer i pådraget u som motvirker forstyrrelsens påvirkning av prosessen, slik at resultatet blir ingen eller minst mulig respons i y . Elimineringen av forstyrrelsers virkning kan bare oppnås for forstyrrelser som er målbare.

Antar vi i første omgang at tilbakekoplingen i figur 9.78 ikke er aktiv, det vil si at $h_r(s) = 0$, blir responsen til prosessen som følge av forstyrrelsene:

$$\begin{aligned} y(s) &= h_u(s)u(s) + h_v(s)v(s) \\ &= (h_u(s)h_f(s) + h_v(s))v(s) \end{aligned} \quad (9.101)$$

hvor foroverkoplingen er gitt av $h_f(s)$.

Ideell foroverkopling oppnås dersom uttrykket i (9.101) kan gjøres lik null. Vi finner da den ideelle foroverkoplingen

$$h_{fi}(s) = -\frac{h_v(s)}{h_u(s)} \quad (9.102)$$

(9.101) kan omskrives til

$$y(s) = \left(\frac{h_u(s)h_f(s)}{h_v(s)} + 1 \right) h_v(s)v(s) \quad (9.103)$$

Siden $h_v(s)v(s)$ er et mål for prosessens respons på forstyrrelsen uten noen form for foroverkopling eller tilbakekopling, vil uttrykket i parentes i (9.103) være et mål for forbedringen som oppnås ved hjelp av foroverkoplingen. Tilsvarende definisjonen i (7.12) av begrepet *reguleringsgrad* i et tilbakekoplet system kan vi nå innføre begrepet **foroverkoplingsgrad** som

$$\frac{y(s)_{\text{med}}}{y(s)_{\text{uten}}} = \frac{h_u(s)h_f(s)}{h_v(s)} + 1 = L(s) \quad (9.104)$$

Den **ideelle foroverkopplingsgraden** er $L(s) = 0$ og oppnås når

$$h_{f_i}(s) = -\frac{h_v(s)}{h_u(s)} \quad (9.105)$$

Vi definerer **statisk foroverkopling** som $h_f(0)$. Hvis setter $s = j\omega = 0$ inn i uttrykket (9.105) for den ideelle foroverkopling, får vi en konstant forsterkning

$$h_{f_i}(0) = -\frac{h_v(0)}{h_u(0)} = K_{f_i} \quad (9.106)$$

Denne vil undertrykke en konstant forstyrrelse (et sprang i v) perfekt (under forutsetning av at vår "lavfrekvente" prosessmodell gitt av $h_v(0)$ og $h_u(0)$ er riktig).

Dersom foroverkoplingen kombineres med en tilbakekopling som vist i figur 9.78, vil systemets totale evne til å undertrykke forstyrrelser være gitt av *produktet* av tilbakekoplingens *reguleringsgrad* $N(s)$ og *foroverkopplingsgraden* $L(s)$,

$$N(s)L(s) = \frac{1}{1 + h_u(s)h_r(s)} \left(\frac{h_u(s)h_f(s)}{h_v(s)} + 1 \right) \quad (9.107)$$

Det er derfor viktig at de to virkemidlene vi har (tilbakekopling og foroverkopling) utnyttes slik at de supplerer hverandre på en mest mulig effektiv måte. Betrakter vi systemets funksjon i frekvensplanet, vet vi at tilbakekoplingen blir ineffektiv når sløyfens forsterkning er mindre enn 1. Dette vil som oftest være tilfellet ved frekvenser høyere enn systemets kryssfrekvens ω_c .

Det er derfor framfor alt viktig å oppnå at $|L(j\omega)| < 1$ der $|N(j\omega)| \approx 1$, ved hjelp av foroverkoplingen.

Virkingen av foroverkoplingen vil ikke være av vesentlig betydning der tilbakekoplingen er effektiv. Størst interesse har derfor foroverkopling i systemer der vi på grunn av store forsinkelser i prosessen ("langsomme" poler eller transportforsinkelser) ikke kan oppnå høy båndbredde i tilbakekoplingen og derfor ikke hurtig nok motvirkning av forstyrrelsene. Dette er særlig problematisk i systemer med transportforsinkelse.

Ved frekvensanalytisk behandling av uttrykket i (9.104) kan det være fordelaktig å benytte seg av et Nicholsdiagram for bestemmelse av $|L(j\omega)|$. Vi skriver da

$$L(j\omega) = g_0(j\omega) + 1 \quad (9.108)$$

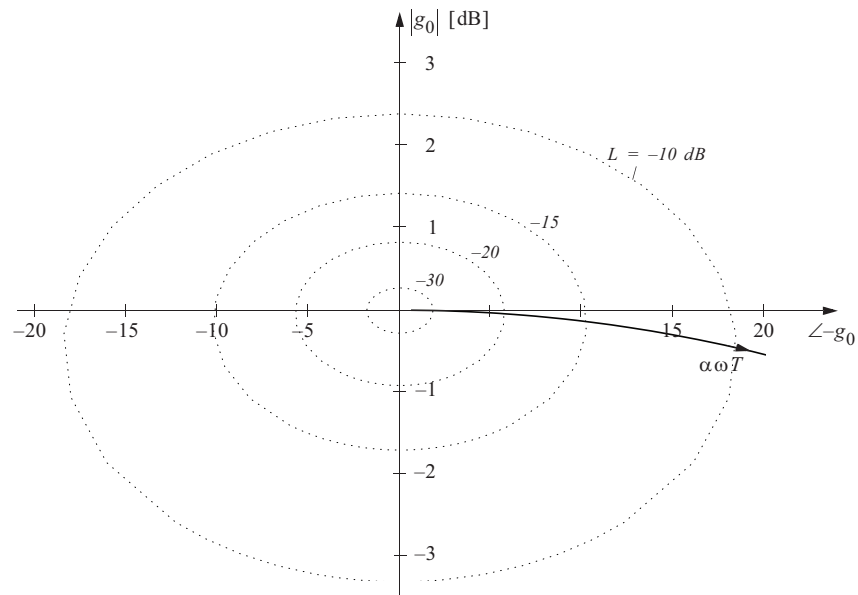
der

$$\begin{aligned} g_0(j\omega) &= \frac{h_u(j\omega)}{h_v(j\omega)} h_f(j\omega) \\ &= -\frac{h_f(j\omega)}{h_{f_i}(j\omega)} \end{aligned} \quad (9.109)$$

Vi har tidligere sett at $N(j\omega) = (1 + h_0(j\omega))^{-1}$ kan leses av fra koter i et Nicholsdiagram der $|h_0(j\omega)|$ tegnes som funksjon av $\angle h_0(j\omega)$. Vi ser at (9.108) likner på uttrykket for $N(j\omega)$, med den forskjell at uttrykket er invertert og at $g_0(j\omega)$ erstatter $h_0(j\omega)$ som “argument”. Vi kan derfor bestemme $L(j\omega)$ ved hjelp av et ordinært Nicholsdiagram når $g_0(j\omega)$ er gitt.

Figur 9.79 viser et forstørret utsnitt av et Nicholsdiagram som i dette tilfellet framstiller koter for bestemmelse av $|L(j\omega)|$, i et $|g_0(j\omega)|/\angle -g_0(j\omega)$ -diagram

Figur 9.79
Hvordan $|L(j\omega)|$
avhenger av den
ideelle foroverkop-
lingen, $g_0(j\omega)$



Ideell foroverkopling svarer til $|L(j\omega)| = 0 = -\infty$ [dB].

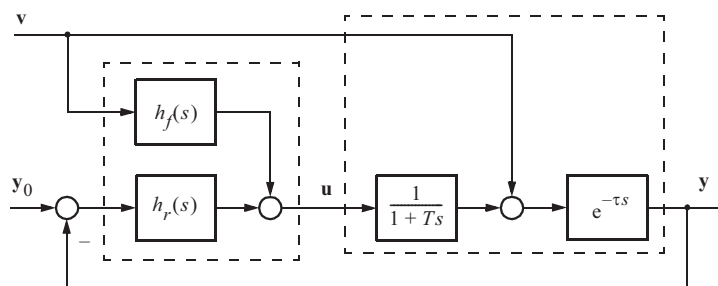
Vi ser at for å oppnå en foroverkopplingsgrad på for eksempel $|L(j\omega)| = -20$ [dB], må transferfunksjonen til foroverkoplingen $h_f(j\omega)$ stemme overens med den ideelle foroverkoplingen $h_{fi}(j\omega)$ innenfor omtrent 1 [dB] amplitude og omtrent 5° i vinkel. Dersom vi er fornøyd med en foroverkopplingsgrad $|L(j\omega)| = -10$ [dB], er kravet til “perfekthet” i foroverkoplingen mindre og ligger i området ca. 2.5 – 3.0 [dB] i amplitude og omtrent 15 – 18° i fasevinkel.

Noen eksempler vil illustrere de betraktninger og definisjoner som er gitt ovenfor.

EKSEMPEL 9.18: Foroverkopling i en prosess med tidskonstant og transportforsinkelse

Figur 9.80 viser et enkelt reguleringsystem omkring en prosess som mellom pådraget og målingen har en reell pol med tidskonstant T og en transportforsinkelse τ .

Figur 9.80
Enkelt system
med foroverkopling



En forstyrrelse v som antas målbar, virker inn etter polen. Vi skal finne løsningen generelt for vilkårlige verdier av T og τ . Først ser vi hva som kan oppnås med tilbakemeldingen og velger med bakgrunn i det som er beskrevet i avsnitt 9.3, en regulator av formen

$$h_r(s) = K_p \frac{1 + T_i s}{T_i s} \quad (9.110)$$

$T_i = T$ vil være et fornuftig valg av integraltid fordi det medfører forkorting av en pol og et nullpunkt. Vi ønsker å oppnå en forsterkningsmargin $\Delta K = 6$ [dB] som i dette tilfellet gir $|N(j\omega)|_{\max} \approx 6.5$ [dB]. I eksempel 9.6 valgte vi

$$\frac{K_p \tau}{T} = \frac{\pi}{4} \approx -2 \text{ [dB]} \quad (9.111)$$

Vi ønsker å finne kryssfrekvensen og benytter at

$$|h_0(j\omega)| = \frac{K_p}{T\omega} \quad (9.112)$$

Siden $|h_0(j\omega_{180})| = 1/2 = -6$ [dB] og $|h_0(j\omega_c)| = 1$, gir (9.112) $\omega_c = \omega_{180}/2$. Ligning (9.43) gir derfor den dimensjonsløse kryssfrekvensen $\omega_c\tau = \pi/4$. Frekvensen som $|N(j\omega)|_{\max}$ oppstår ved, ligger et sted mellom ω_c og ω_{180} som vist i figur 8.32 a). Dette vil tilsi at den aktuelle frekvensen er $\omega\tau \approx 1.3$ for vårt system.

Vi finner den ideelle foroverkoplingen for dette systemet ved å benytte (9.105)

$$h_{fi}(s) = -(1 + Ts) \quad (9.113)$$

Fordi en ubegrenset derivatvirkning ikke lar seg realisere i praksis, blant annet på grunn av uunngåelig høyfrekvent støy i målingen, vil det være realistisk med en foroverkopling av formen

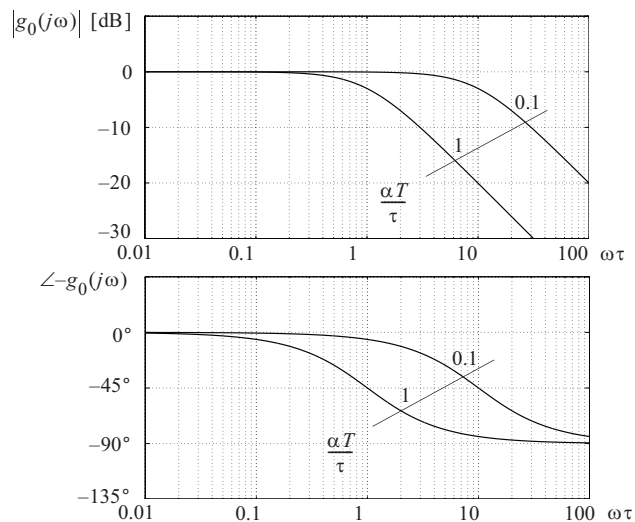
$$h_f(s) = -\frac{1 + Ts}{1 + \alpha Ts} \quad (9.114)$$

der $\alpha < 1$. Vi får da

$$g_0(s) = -\frac{h_f(s)}{h_{fi}(s)} = -\frac{1}{1 + \alpha Ts} = -\frac{1}{1 + \left(\alpha \frac{T}{\tau}\right) s\tau} \quad (9.115)$$

Figur 9.81 viser frekvensresponsen til dette uttrykket for forskjellige verdier av faktoren $\alpha T/\tau$. Dette resultatet er så overført til figur 9.79 for bestemmelse av $|L(j\omega)|$. Resultatet i form av kurver for $|N(j\omega)|$ og $|L(j\omega)|$ for forskjellige verdier av faktoren $\alpha T/\tau$ er vist i figur 9.82, der vi har brukt den dimensjonsløse frekvensen $\omega\tau$.

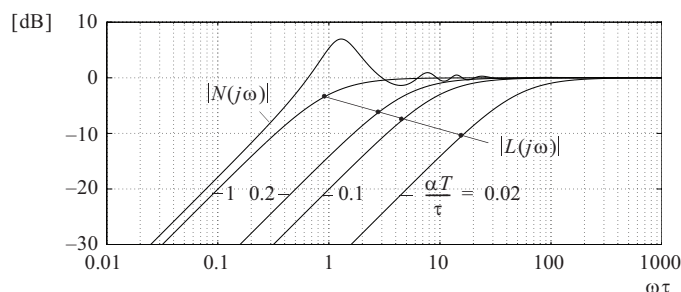
Figur 9.81
Frekvensrespons
for forskjellige
verdier av $\alpha T/\tau$



Siden det er produktet av $N(j\omega)$ og $L(j\omega)$ som bestemmer systemets evne til å undertrykke forstyrrelser, ser vi at vi ved valg av for eksempel $\alpha T/\tau = 0.1$ får en betraktelig forbedring.

For en gitt verdi av α ser vi at den forbedringen som oppnås ved bruk av foroverkopling i forhold til det som kan oppnås ved tilbakekopling, øker når transportforsinkelsen (τ) blir stor i forhold til tidskonstanten (T). Foroverkopling er derfor framfor alt nyttig i systemer med stor transportforsinkelse eller annen ikke-minimum-fase virkning.

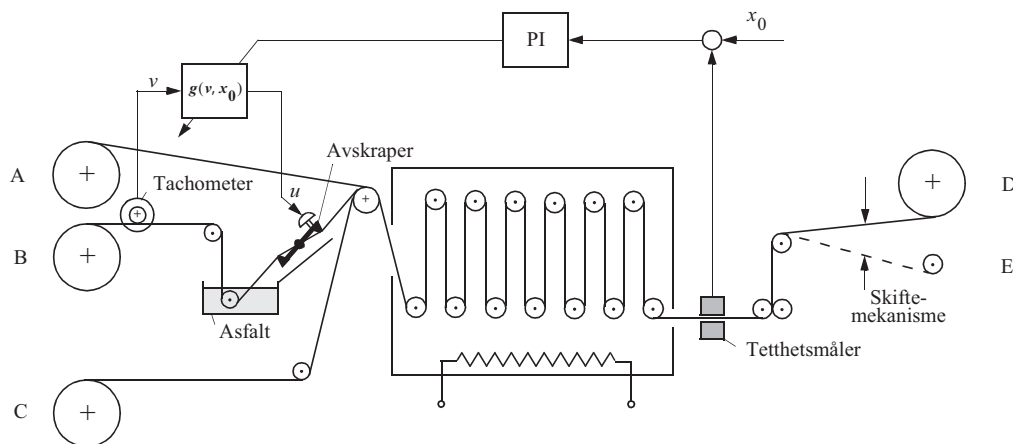
Figur 9.82
 $|N(j\omega)|$ og
 $|L(j\omega)|$ for
 forskjellige verdier
 av faktoren $\alpha T/\tau$



EKSEMPEL 9.19: Foroverkopling ved produksjon av impregnert papir

Figur 9.83 viser forenklet prinsippet for produksjon av asfaltimpregnert papir (for framstilling av sekker) der en legger sammen tre papirbånd (A, B, C). Det midterste papirbåndet (B) har passert gjennom et asfaltbad.

Figur 9.83
 Oversikt over
 produksjons-
 prosessen



Det tredobbelte, impregnerte papiret blir varmebehandlet kontinuerlig en viss tid og til slutt oppviklet på ruller (D, E). Kvaliteten på det ferdig impregnerte papiret er sterkt avhengig av impregneringsgraden (mengden av asfalt absorbert i det midterste papirlaget), og vi ønsker derfor å regulere denne. Impregneringsgraden kan måles ved hjelp av β -stråleabsorpsjon (tetthetsmåling), *men bare etter at papiret har passert varmebehandlingen*. Påleggingen av asfalt kan påvirkes ved hjelp av en avskrapningsmekanisme som befinner seg like etter asfaltbadet.

Maskinen har en normal kapasitet på 300 [m/min]. Vi vil forsøke å holde størst mulig gjennomsnittlig produksjonshastighet, samtidig som kravene til konstant impregneringsgrad tilfredsstilles.

En sterk begrensning ligger i skiftemekanismen som legger det ferdig impregnerte papiret over fra en oppviklingsenhet til den neste, når den første er full. Denne skiftemekanismen fungerer ikke tilfredsstillende ved større hastigheter enn 100 [m/min].

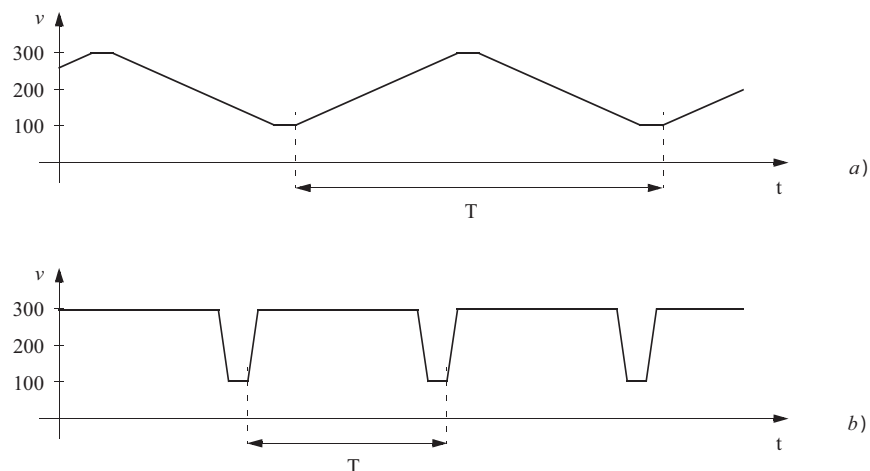
Benytter vi oss av en regulering av impregneringsgraden basert utelukkende på en tilbakekopling fra tetthetsmåleren over en regulator til avskrapningsmekanismen, vil det resulterende reguleringssystemet få meget liten båndbredde på grunn av den store transportforsinkelsen (oppholdstiden) i varmebehandlingsenheten. Resultatet er at vi for å holde kravene til impregneringsgrad ikke kan kjøre papirhastigheten hurtig ned fra 300 [m/min] til 100 [m/min] og opp igjen når skiftemekanismen skal i funksjon. Figur 9.84 a) viser hvordan vi er tvunget til å betjene papirhastigheten i et slikt tilfelle.

Figur 9.84

Produksjonshastigheten

a) uten foroverkopling

b) med foroverkopling



Midlere papirhastighet er altså vesentlig lavere enn 300 [m/min], hvilket vil si at produksjonstiden (T) for en rull ferdig papir er vesentlig større enn maskinens produksjonskapasitet skulle tilsi. Figur 9.84 b) viser hvordan vi ønsker å kjøre hastigheten ned og opp raskere, for derved å øke produksjonskapasiteten. Dette kan oppnås ved hjelp av foroverkopling.

Vi antar at asfaltbadet har temperaturregulering og at asfalten derfor har konstant temperatur og dermed konstant viskositet. Derved vil mengden asfalt som påføres papiret, bare være en funksjon av papirhastigheten v og avskrapningstrykket som er gitt ved stillingen av avskrapermekanismen u . Impregneringsgraden x vil da være gitt av

$$x = f(u, v) \quad (9.116)$$

Denne relasjonen inneholder altså ingen vesentlige dynamiske ledd, idet impregneringsgraden er bestemt umiddelbart etter at papiret forlater avskrapningen, selv om tetthetsmåleren ikke kan måle dette før lenge etterpå.

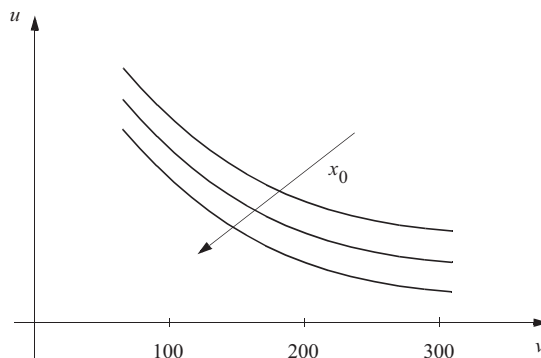
Funksjonen (9.116) kan bestemmes eksperimentelt ved å måle v , u og x under stasjonære forhold. Fordi vi ønsker en konstant impregneringsgrad, i det minste på hver papirkvalitet, kan vi finne den nødvendige foroverkopling fra v til u av

$$u = g(v, x_0) \quad (9.117)$$

der x_0 er den ønskede impregneringsgraden.

Figur 9.85 viser sammenhengen fra (9.117). Denne foroverkoplingen kan realiseres ved at en måler papirhastigheten ved hjelp av et elektrisk tachometer med en trinse som rider på papirbanen. Tachometerspenningen føres til et ulineært elektrisk nettverk, som gir ut en elektrisk spenning proporsjonal med den ønskede u . En servomekanisme (antydnet i figur 9.83 med et ventilsymbol) sørger for betjeningen av avskrapermekanismen. På grunn av variasjoner i asfaltkvalitet, papirporøsitet etc. vil vi ikke oppnå den beste presisjonen med et slikt arrangement. En tilbakekopling fra tetthetsmåleren, som jo bestemmer impregneringsgraden som i virkeligheten er oppnådd, vil nå kunne benyttes til å justere foroverkoplingen som antydnet i figur 9.83. Denne justeringen kan bestå enten i at den beregnede kurven i figur 9.85 forskyves parallelt med en av aksene eller "vrís". Regulatoren i tilbakekoplingen vil vi velge som en PI-regulator med integraltid omtrent lik "tidskonstanten" i avskrapermekanismen.

Figur 9.85
Foroverkopling
gitt ved (9.117)



Den løsningen som er beskrevet i dette eksemplet er typisk for en lang rekke industrielle prosesser. Den kan blant annet anvendes ved regulering av valsepalt, valsetrykk, strekk og hastighet i valseverk, basert på måling av tykkelsesvariasjoner i emnet (plate eller bånd) et stykke før dette kommer fram til valsene. En slik foroverkopling, kombinert med en regulering basert på tilbakekopling fra dimensjonsmåling etter valsene, vil kunne gi nøyaktig regulering både for hurtigere og langsomme variasjoner i emnets egenskaper.



Det er åpenbart av blant annet utledningen omkring figur 9.79 at kravet til overensstemmelse mellom den ideelle foroverkoplingen og den virkelige foroverkoplingen er ganske store dersom vi skal oppnå en vesentlig virkning. I mange prosesser vil prosessparametrene endre seg så mye at foroverkoplingen blir mistilpasset og den ønskede virkningen uteblir. Et alternativ er selvsagt å gjøre foroverkoplingen adaptiv (selvtilpassende), men en slik løsning blir mer komplisert enn det som her er skissert.

En fordel med foroverkopling framfor tilbakekopling er at det *ikke oppstår noe stabilitetsproblem*. Stabilitet av en lukket tilbakekoplingsløyfe er som kjent en vesentlig hindring for blant annet å oppnå stor båndbredde (dvs. hurtighet).

9.6.2 Forholdsregulering

En ofte forekommende form for reguleringsystem som godt kan betraktes som en anvendelse av foroverkopling, er såkalt *forholdsregulering*. Slik regulering kommer til anvendelse i tilfeller der to eller flere strømmer av faste stoffer, væsker eller gasser til enhver tid bør stå i et bestemt forhold til hverandre. Slik er det for eksempel ved regulering av olje - og lufttilførselen til en oljebrenner som hører til en dampkjel, eller ved dosering av syre for nøytralisering av en base i en kjemisk prosess.

EKSEMPEL 9.20: Bruk av forholdsregulering

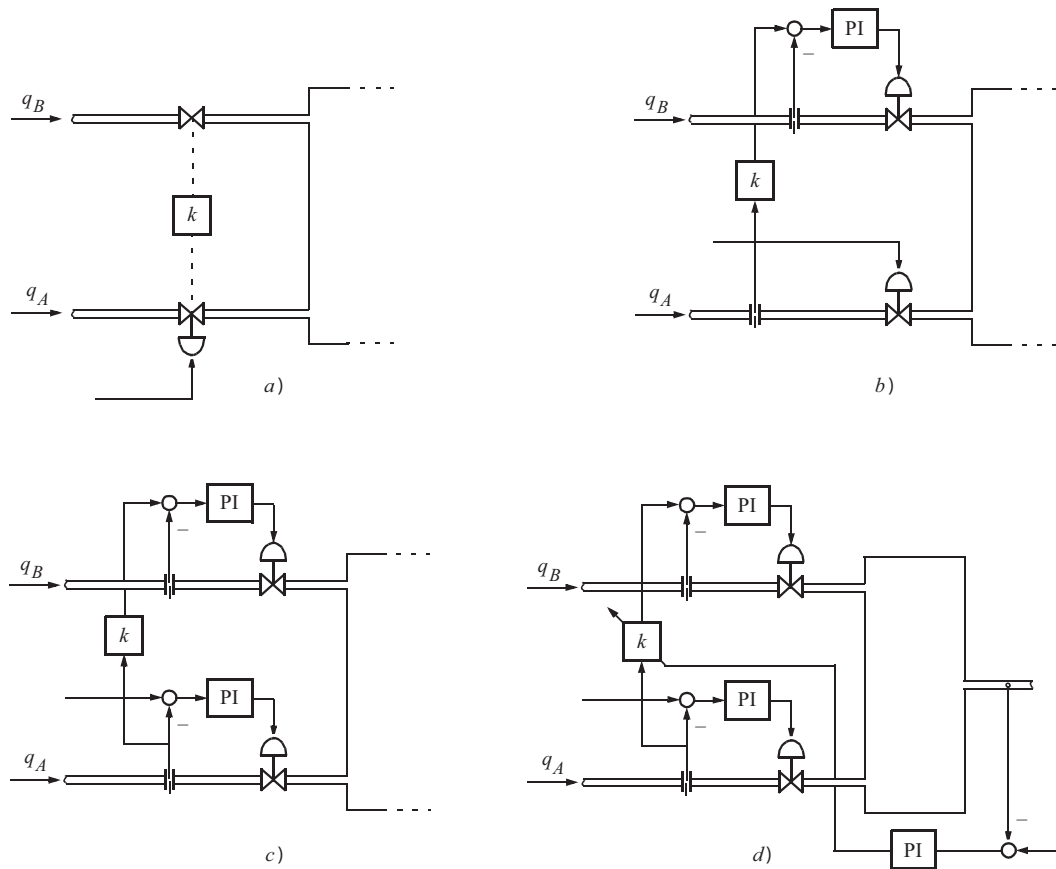
Figur 9.86 viser noen vanlig forekommende utforminger av forholdsregulering som f.eks. ofte benyttes for å regulere tilførselen av ulike stoffer til en reaktor av den typen som er beskrevet i eksempel 3.6. Problemet er at vi ønsker $q_B = kq_A$. Den enkleste løsningen, vist i figur 9.86 a), anvender to ventiler som er direkte mekaniske forbundet, og som åpner og stenger parallelt. For å oppnå et godt resultat med denne løsningen kreves det imidlertid høy presisjon ved framstilling av ventilene.

I figur 9.86 b) er en strømningsregulator innført i *B*-røret, og dennes referanse styres fra en måling av q_A .

Den sikreste løsningen har vi i figur 9.86 c) der begge strømmene er regulert individuelt, og referansen for strømningsregulatoren i rør *B* er styrt fra målingen av q_A . Referansen for strømningsregulatoren i *A*-røret antas å komme utenfra. Denne referansen vil ved fyringsregulering være den som angir pådraget.

Dersom forholdstallet k ikke alltid kan være konstant, men må justeres langsomt ut fra målinger senere i prosessen, vil vi kunne benytte et system som antydnet i figur 9.86 d). Ved fyringsregulering vil det for eksempel kunne være aktuelt å foreta gassanalyse i røykgassen for bestemmelse av CO/CO_2 -forholdet og gjennom en PI-regulator justere forholdstallet k . Derved kan en ideell forbrenning oppnås under sterkt varierende last og med forskjellige kvaliteter på brennstoffet.

Figur 9.86
Forholdsregulering



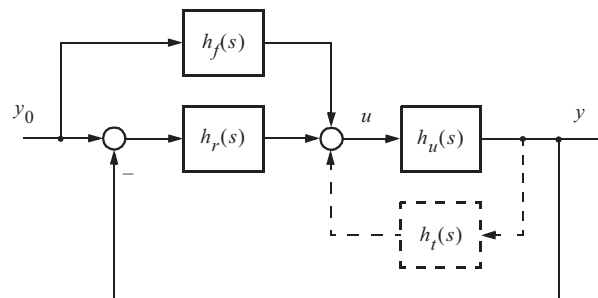
■

9.6.3 Foroverkopling (fra referansen) i følgesystemer

Det som er beskrevet foran for prosessreguleringssystemer, kan også gjøres gjeldende for følgesystemer. Foroverkoplingen vil i et følgesystem tas fra den størrelsen som er den viktigste påvirkningen på systemet, referansen y_0 . Det kan selvsagt også være aktuelt å nytte foroverkoplinger fra eventuelle forstyrrelser v på samme måte som beskrevet i foregående avsnitt.

Figur 9.87 viser en mulig struktur for foroverkopling i et monovariabelt følgesystem karakterisert ved en hovedsløyfe (tilbakekopling) med regulator $h_r(s)$. Det er også tatt med en blokk $h_t(s)$ som representerer en eventuell intern tilbakekopling, som er svært aktuelt i nettopp følgesystemer. Foroverkoplingen utgjøres av transferfunksjonen $h_f(s)$ mellom referansen y_0 og pådraget u .

Figur 9.87
Foroverkopling
fra referansen



Hensikten med den interne tilbakekoplingen kan, som beskrevet i avsnitt 9.5, være å modifisere prosessens transferfunksjon $h_u(s)$ slik at den blir gunstigere for hovedsløyfen. I det følgende vil vi analysere systemet uten og med slik tilbakekopling.

Figur 9.87 viser at siden et ideelt følgesystem tilsikter $y = y_0$, vil den ideelle foroverkoplingen, når vi ser bort fra virkningen av tilbakekoplingen, måtte bli

$$h_{f_i}(s) = \frac{1}{h_u(s)} \quad (9.118)$$

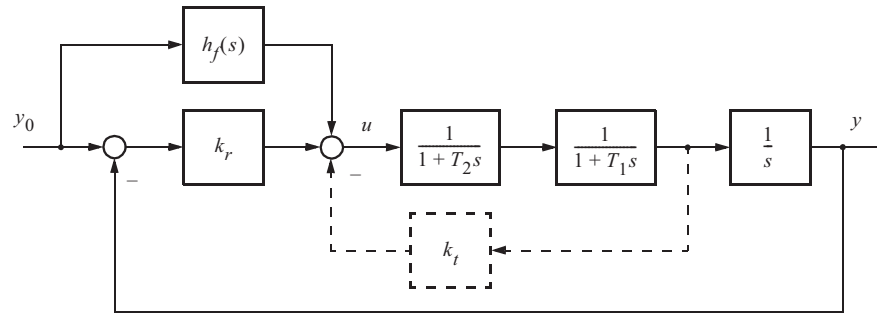
Denne foroverkoplingen kan selvsagt ikke realiseres i praksis, men den er et gunstig utgangspunkt for dannelse av tilnærmelser.

Det er hensiktsmessig å illustrere virkningen av foroverkopling på den typen følgesystemer som er mest aktuelle; servomekanismer.

EKSEMPEL 9.21: Foroverkopling i en servomekanisme

Figur 9.88 viser et blokkdiagram for en servomekanisme, der prosessen utgjøres av en integrasjon (pol i origo) forårsaket av integrasjonen mellom hastighet og posisjon, en pol med tidskonstant T_1 som i hovedsak skyldes motorens mekaniske treghet, og en pol med tidskonstant T_2 som i hovedsak skyldes motorens elektriske treghet. Penneskriveren i eksempel 9.8 er et eksempel på et slikt system.

Figur 9.88
Servomekanisme
med foroverkopling
fra referansen



Dersom y og y_0 er rotasjoner, er det forholdsvis lett å instrumentere deres deriverte ved hjelp av tachometergeneratorer. Det er derfor svært vanlig å utruste en slik servomekanisme med en intern tilbakekopling fra et tachometer. Denne tilbakekoplingen vil i enkelte tilfeller bare være en konstant k_t .

Vi skal sammenlikne resultatene som kan oppnås dersom regulatordelen i servomekanismens hovedsløyfe antas å være en konstant k_r , og:

1. Ingen tilbakekopling fra tachometeret og ingen foroverkopling
2. Tilbakekopling fra tachometeret k_t og ingen foroverkopling
3. Tilbakekopling fra tachometeret k_t og foroverkopling

Vi antar i det følgende at $T_2 \leq 0.1T_1$. Dette vil i praktiske tilfeller være forholdsvis realistisk. De tre tilfellene som er nevnt, blir da:

Tilfelle 1:

Siden $T_2 \leq 0.1T_1$, kan vi se bort i fra tidskonstanten T_2 ved dimensjonering av regulatorforsterkningen, k_r . Vi har da et system som svarer til det i eksempel 9.1.

Likning (9.21) gir $k_r = 1/T_1$. (Stabilitetsgrensen er $K_{rkrit} = \frac{T_1 + T_2}{T_1 T_2}$.)

$$M(s) = \frac{1}{1 + T_1 s + T_1(T_1 + T_2)s^2 + T_1^2 T_2 s^3} \quad (9.119)$$

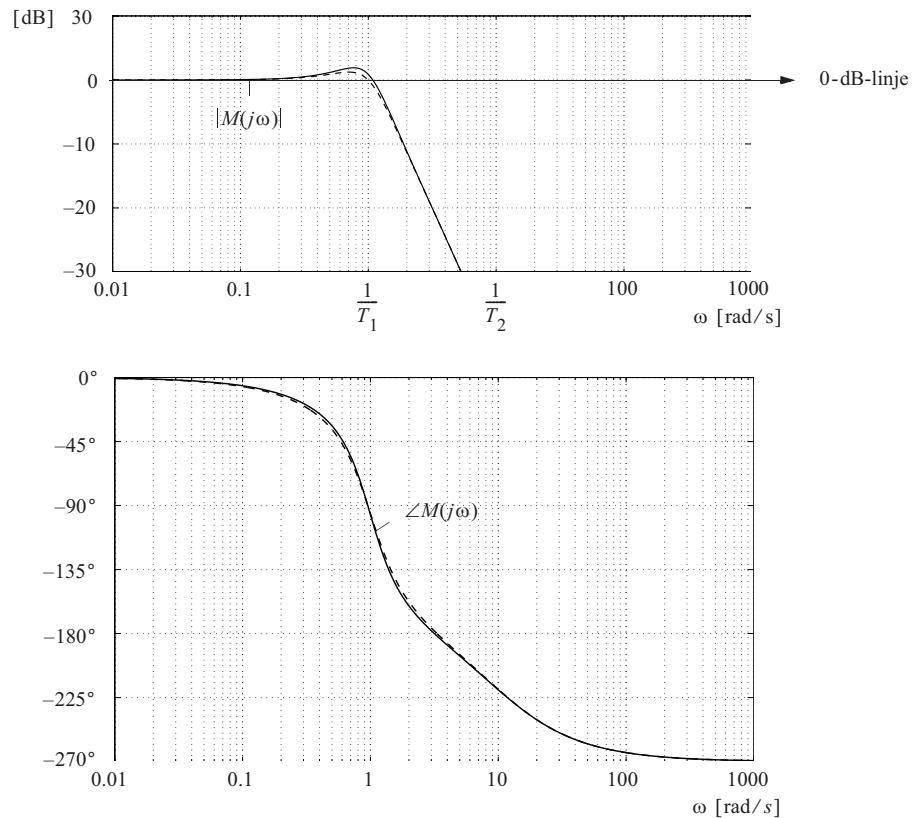
Vi kan finne en tilnærming til $M(s)$ siden $T_2 \leq 0.1T_1$ (slik vi gjorde i eksempel 9.8).

$$\frac{y}{y_0}(s) \approx \frac{1}{1 + 2\zeta \frac{s}{\omega_{01}} + \left(\frac{s}{\omega_{01}}\right)^2} \frac{1}{1 + T_2 s} \quad (9.120)$$

Vi multipliserer ut nevneren i (9.120) og justerer parameterene ω_{01} og ζ slik at de to leddene av høyeste orden blir lik de tilsvarende leddene i (9.119). De aktuelle parametrene blir da $\omega_{01} = 1/T_1$ og $\zeta = 0.5$. Følgesystemets båndbredde er blitt omtrent lik $\omega_{01} = 1/T_1$.

Figur 9.89 viser et AFF-diagram for $M(s)$ når $T_1 = 1$ og $T_2 = 0.1$. De heltrukne kurvene viser de virkelige karakteristikkene mens de stiplede kurvene gjengir tilnærmelsen.

Figur 9.89
AFF-diagram
for $M(s)$



Tilfelle 2:

Ved dimensjonering av den indre sløyfen har vi et system som svarer til det i eksempel 9.2. Likning (9.24) gir da

$$k_t = \frac{T_1}{T_2} \quad (9.121)$$

Den modifiserte prosesstransferfunksjonen (inkludert intern tilbakekopling k_t) blir, når vi benytter (9.121),

$$h_u'(s) = \frac{1}{s((1+T_1s)(1+T_2s)+k_t)} = \frac{1}{1+k_t} \frac{1}{s \left(1+T_2s + \frac{T_1T_2}{1+k_t} s^2 \right)} \quad (9.122)$$

Siden $T_2 \leq 0.1T_1$, vil $1/(1+k_t) \approx 1/k_t$. Vi kan derfor finne følgende tilnærming til $h_u'(s)$

$$h_u'(s) \approx \frac{1/k_t}{s(1+T_2sT_2^2s^2)} \quad (9.123)$$

Regulatortransferfunksjonen $h_r(s) = k_r$ benyttes nå i hovedsløyfen. Vi antar at den doble polen i $h_u'(s)$ kan betraktes separat, slik at $M(s)$ tilnærmet blir

$$\frac{y}{y_0}(s) \approx \frac{1}{1 + \frac{k_t}{k_r} s} \frac{1}{(1+T_2s+T_2^2s^2)} \quad (9.124)$$

Dersom

$$\frac{k_r}{k_t} < \frac{1}{T_2}$$

(det vil si for forholdsvis små verdier av k_r) er (9.124) en rimelig tilnærming, fordi polen i $-k_r/k_t$ ikke virker særlig inn på den doble polen. Øker vi k_r , vil dette systemet med en rimelig stabilitetsmargin kunne få en båndbredde omtrent lik $0.8(1/T_2)$. Vi legger merke til at stabilitetsgrensen nå er blitt

$$K_{rkrit} = \frac{T_1+T_2}{T_1T_2}(1+k_t)$$

som kan finnes ved bruk av Rouths kriterium.

Figur 9.90 viser et AFF-diagram for $h_u'(s)$ og $M'(s)$ når $T_1 = 1$ og $T_2 = 0.1$. De heltrukne kurvene viser de virkelige kurvene mens de stiplede kurvene gjengir tilnærmelsene.

Ved å sammenlikne figur 9.89 og 9.90 ser vi at bruk av intern tachometer tilbakekopling har gitt en økning i systemets båndbredde på nær en dekad.

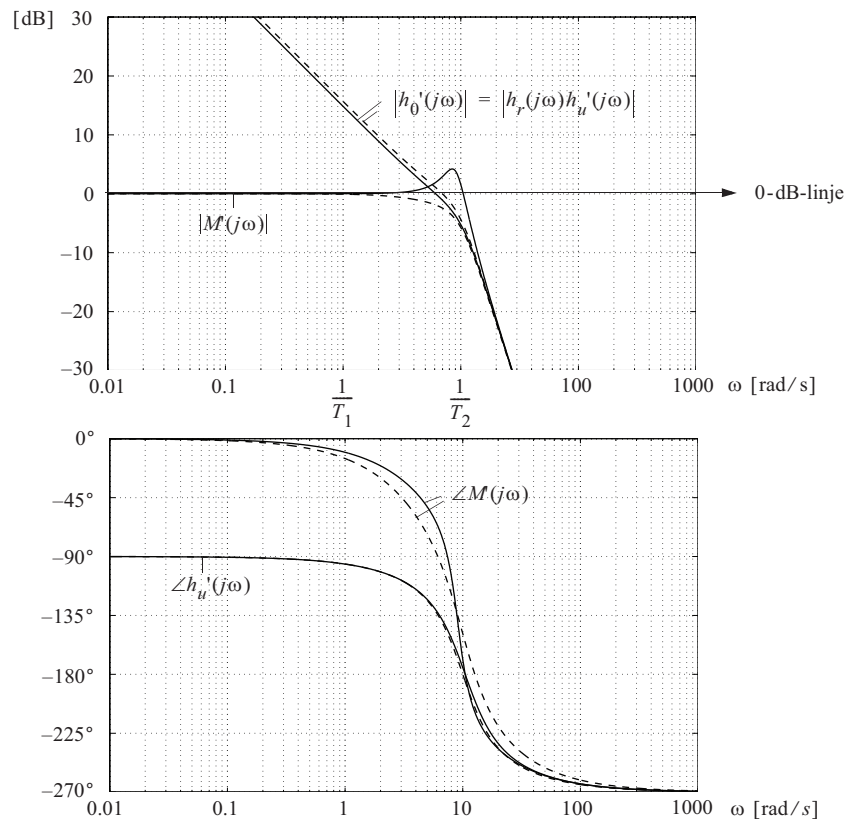
Tilfelle 3:

Kombinasjonen av intern tilbakekopling over faktoren k_t og foroverkopling med

$$h_f(s) = k_t s \quad (9.125)$$

skal vise seg å ha gunstige egenskaper. Foroverkoplingen i (9.125) betyr at vi, for eksempel ved hjelp av et tachometer koplet til referansebevegelsen y_0 , genererer et signal proporsjonalt med hastigheten av referansen \dot{y}_0 og med samme vekt k_t som vi har brukt i tilbakekoplingen fra den målte hastigheten på motoren.

Figur 9.90
AFF-diagram for
 $h_u'(s)$ og $M'(s)$



Analyserer vi sammenhengene i figur 9.87, finner vi

$$\begin{aligned}\frac{y}{y_0}(s) &= \frac{h_u(s)h_r(s)}{1+h_0(s)} + \frac{h_u(s)h_f(s)}{1+h_0(s)} \\ &= \left(1 + \frac{h_f(s)}{h_r(s)}\right)M(s)\end{aligned}\quad (9.126)$$

der

$$M(s) = \frac{h_0(s)}{1+h_0(s)} \quad (9.127)$$

som i dette tilfellet er gitt av (9.124). Setter vi den valgte foroverkoplingen fra (9.125) inn i (9.126), blir resultatet

$$\frac{y}{y_0}(s) \approx \frac{1}{1+T_2s+T_2^2s^2} \quad (9.128)$$

uansett hvilken verdi som velges for forsterkningen k_r , så sant denne ikke er for liten. Vi ser at det resulterende systemet får en båndbredde lik $1/T_2$ som er det beste vi hittil har oppnådd.

Vi kunne selvsagt ved bruk av (9.123) i (9.118) ha kommet fram til en foroverkopling som gir et enda gunstigere resultat, nemlig

$$\begin{aligned}h_f(s) &= k_t s(1+T_2s+T_2^2s^2) \\ &= k_t s + k_t T_2 s^2 + k_t T_2^2 s^3\end{aligned}\quad (9.129)$$

Denne foroverkoplingen betinger imidlertid i tillegg til hastighetsmålingen på referanseinngangen at vi også bestemmer akselerasjonen og akselerasjonens deriverte. Dette kan være noe vanskeligere i praksis.

■

Løsningen vi får ved å bruke tachometerforoverkopling kan også tolkes som et system med hastighetsstyring (følgesystemets hastighets-tilbakekopplingsløyfe har hastighetsreferanse). Posisjonssløyfen som kan ha en relativt lav forsterkning, brukes primært til å finjustere posisjonen i henhold til posisjonsreferansen. En slik betraktning leder oss til en mer enhetlig beskrivelse av kombinasjonen av intern tilbakekopling og foroverkopling i følgesystemer.

Vi tar da utgangspunkt i en dynamisk prosess beskrevet av

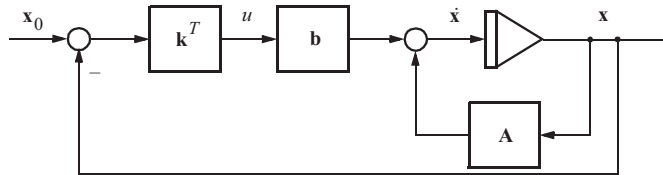
$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \quad (9.130)$$

For å gjøre sammenlikningen med det foregående enkel, antas prosessen å ha bare en pådragsvariabel. Alle tilstandsvariablene antas målbare, det vil si $y = \mathbf{x}$. Skal vi få denne prosessen til å følge en referanse, er det selvsagt rimelig at referansen er en vektor som refererer til hele tilstandsvektoren. Vi antar derfor en referansevektor \mathbf{x}_0 som angir det ønskede forløp av hver enkelt av prosessens tilstandsvariable. Avviket mellom referansevektoren og tilstandsvektoren ($\mathbf{x}_0 - \mathbf{x}$) kan nå brukes til å generere pådraget idet vi velger strategien

$$u = \mathbf{k}^T (\mathbf{x}_0 - \mathbf{x}) \quad (9.131)$$

der $\mathbf{k}^T = [k_1, k_2, \dots, k_n]$. Et følgesystem i henhold til dette er vist i figur 9.91.

Figur 9.91
Følgesystem med
tilbakekopling

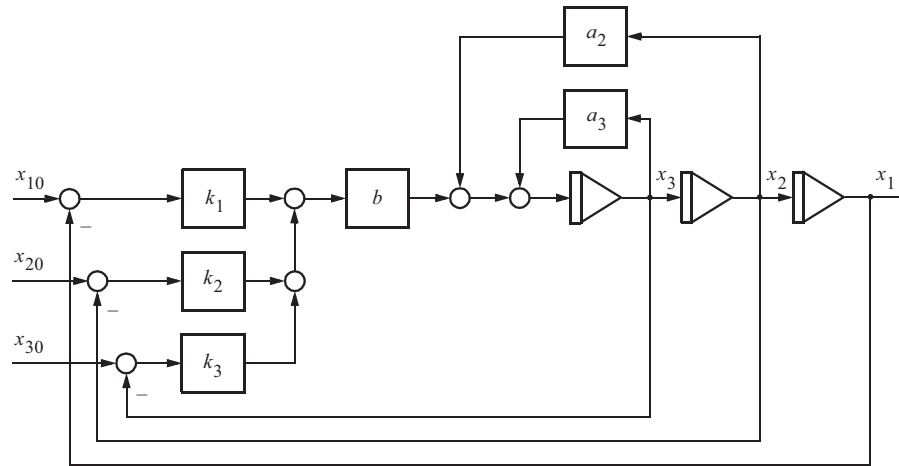


Tar vi som eksempel en prosess bestående av en likestrømmotor som driver en last av samme type som vist i figur 9.39, vil matrisene i (9.130) bli av formen

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & a_2 & a_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix}$$

Et elementært blokkdiagram for følgesystemet vil da bli som vist i figur 9.92.

Figur 9.92
Elementært blokk-
diagram for den
dynamiske
prosessen



Siden x_1 , x_2 og x_3 representerer henholdsvis posisjon, hastighet og akselerasjon av prosessen, ser vi at løsningen består av tilbakekoplingsløyper fra hver av disse størrelsene.

Siden x_{10} , x_{20} og x_{30} er referanser for de samme størrelsene, har vi realisert en kombinasjon av foroverkopling og tilbakekopling. Foroverkoplingsfenomenet ligger i at elementene i \mathbf{x}_0 alle kan avledes av hverandre.

Setter vi (9.131) inn i (9.130), finner vi

$$\begin{aligned}\dot{\mathbf{x}} &= (\mathbf{A} - \mathbf{b}\mathbf{k}^T)\mathbf{x} + \mathbf{b}\mathbf{k}^T \mathbf{x}_0 \\ &= \tilde{\mathbf{A}}\mathbf{x} + \tilde{\mathbf{B}}\mathbf{x}_0\end{aligned}\quad (9.132)$$

som er differensiallikningen for det lukkede følgesystemet. Matrisen $\tilde{\mathbf{A}}$ som karakteriserer det lukkede følgesystemet, vil for eksemplet i figur 9.92 bli

$$\begin{aligned}\tilde{\mathbf{A}} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -bk_1 & -bk_2 + a_2 & -bk_3 + a_3 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 \end{bmatrix}\end{aligned}\quad (9.133)$$

Vi ser altså at det med tilbakekoplingsmatrisen \mathbf{k}^T er mulig å korrigere selektivt elementene i den nederste linjen i $\tilde{\mathbf{A}}$. Systemer av en form som beskrives av (9.132) og (9.133) har vi tidligere behandlet i avsnitt 3.5. Vi fant da ut at konstantene α_0 , α_1 og α_2 utgjør koeffisientene i systemets karakteristiske likning, som i dette tilfellet vil være

$$\lambda^3 + \alpha_2\lambda^2 + \alpha_1\lambda + \alpha_0 = 0 \quad (9.134)$$

Følgeegenskapene til systemet i figur 9.92 kan vi undersøke ved Laplacetransformasjon av (9.132).

Vi finner da

$$\mathbf{x}(s) = (s\mathbf{I} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{B}}\mathbf{x}_0(s) \quad (9.135)$$

Innsetting av de aktuelle matrisene gir

$$\begin{aligned}\frac{\mathbf{x}}{\mathbf{x}_0}(s) &= \frac{1}{s^3 + \alpha_2s^2 + \alpha_1s + \alpha_0} \begin{bmatrix} s^2 + \alpha_2s + \alpha_1 & s + \alpha_2 & 1 \\ -\alpha_0 & s^2 + \alpha_2s & s \\ -\alpha_0s & -\alpha_1s - \alpha_0 & s^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ bk_1 & bk_2 & bk_3 \end{bmatrix} \\ &= \frac{b}{s^3 + \alpha_2s^2 + \alpha_1s + \alpha_0} \begin{bmatrix} k_1 & k_2 & k_3 \\ sk_1 & sk_2 & sk_3 \\ s^2k_1 & s^2k_2 & s^2k_3 \end{bmatrix}\end{aligned}\quad (9.136)$$

Med dette uttrykket kan vi i detalj studere alle mulige kombinasjoner av interne tilbakekoplinger, med og uten foroverkopling.

Kombinasjon 1:

Har vi for eksempel begge de interne tilbakekoplingene i tillegg til hovedtilbakekoplingen, men ingen foroverkopling fra hastighet og akselerasjon, må vi sette

$$\mathbf{x}_0(s) = \begin{bmatrix} x_{10}(s) \\ 0 \\ 0 \end{bmatrix} \quad (9.137)$$

Bruker vi denne i (9.136), finner vi

$$x_1(s) = \frac{bk_1}{s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0} x_{10}(s) \quad (9.138)$$

Siden $\alpha_0 = bk_1$, kan vi skrive dette som

$$x_1(s) = \frac{1}{\left(1 + \frac{s}{\omega_1}\right) \left(1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2\right)} x_{10}(s) \quad (9.139)$$

der ω_1 , ω_0 og ζ kan innstilles ved hjelp av tilbakekoplingene k_1 , k_2 og k_3 .

Kombinasjon 2:

Tar vi i bruk foroverkopling, i første rekke bare fra hastigheten av referansen, vil vi i (9.136) måtte benytte

$$\mathbf{x}_0(s) = \begin{bmatrix} x_{10}(s) \\ x_{20}(s) \\ 0 \end{bmatrix} = \begin{bmatrix} x_{10}(s) \\ s x_{10}(s) \\ 0 \end{bmatrix} = x_{10}(s) \begin{bmatrix} 1 \\ s \\ 0 \end{bmatrix} \quad (9.140)$$

Resultatet blir

$$x_1(s) = \frac{1 + \frac{k_2}{k_1} s}{\left(1 + \frac{s}{\omega_1}\right) \left(1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2\right)} x_{10}(s) \quad (9.141)$$

Vi legger merke til at vi ved hensiktsmessig valg av tilbakekoplingene k_1 , k_2 og k_3

kan forme transferfunksjonen i (9.141) slik at vi oppnår bedre følgning enn det vi kunne med (9.139).

Kombinasjon 3:

Tar vi til slutt med foroverkopling fra akselerasjonen av referansesignalet, finner vi

$$\mathbf{x}_0(s) = \begin{bmatrix} x_{10}(s) \\ x_{20}(s) \\ x_{30}(s) \end{bmatrix} = x_{10}(s) \begin{bmatrix} 1 \\ s \\ s^2 \end{bmatrix} \quad (9.142)$$

som brukt i (9.136) gir

$$x_1(s) = \frac{1 + \frac{k_2}{k_1}s + \frac{k_3}{k_1}s^2}{\left(1 + \frac{s}{\omega_1}\right) \left(1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2\right)} x_{10}(s) \quad (9.143)$$

Bruk av akselerasjonsforoverkopling gir derfor muligheter for ytterligere forbedring av følgeegenskapene idet det er kommet inn enda et nullpunkt i transferfunksjonen. Polene er de samme i alle de tre tilfellene.

9.6.4 Prediktiv foroverkopling

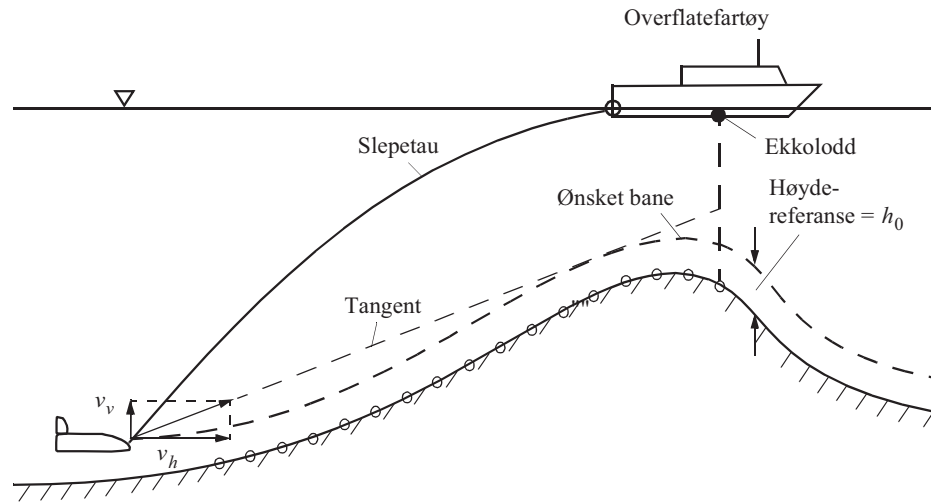
Foroverkopling er særlig viktig i tilfeller der pådraget er beheftet med forsinkelse eller begrensninger på andre måter slik at det lønner seg å utnytte kunnskap vi har om forstyrrelser som vi vet kommer. I følgesystemer er dette særlig aktuelt når det er mulig å prediktere framtidig forløp av referansen $y_0(t)$. Slik vil det være når sjåføren i en bil ser framover og i tide kan ta hensyn til en sving som kommer.

EKSEMPEL 9.22: Slepning og dybdestyring av undervannsfarkost

En tilsvarende problemstilling har en ved sleping av en undervannsfarkost bak et overflatefartøy. Undervannsfarkosten er utstyrt med automatisk regulering av vertikalposisjonen, på basis av en ekkoloddmåling av avstanden til bunnen (ca. 3 m), og vertikalstyring ved hjelp av en motordrevet propell (egentlig to vertikalthrustere, se figur 10.10). Dette systemet er skissert i figur 9.93.

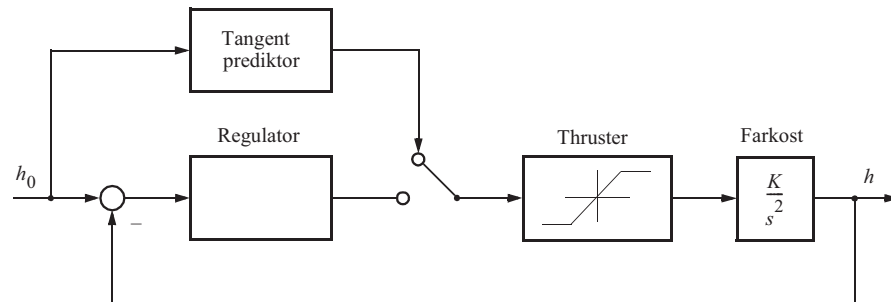
Problemet er at vertikalhastigheten av undervannsfarkosten er begrenset på grunn av begrenset motoreffekt. Det betyr at dersom det kommer en brå stigning i bunnprofilen, må vertikalbevegelsen starte tidligere jo mindre den maksimale vertikale hastigheten er for å unngå kollisjon med bunnen.

Figur 9.93
Sleping av
undervannsfarkost



Siden overflatefartøyet er utstyrt med ekkolodd, kjenner vi imidlertid det framtidige forløpet av bunnprofilen, som vist i figur 9.93. En rett linje kan trekkes som tangerer den ønskede banen, til enhver tid. Når vinkelkoeffisienten til denne tangenten er lik forholdet mellom den maksimale vertikalthastigheten $v_{v\max}$ og farkostens horisontalhastighet v_h , må vertikalthrusteren kjøres med maksimal hastighet for å unngå kollisjon med bunnen. Denne “predikator”-baserte løsningen er antydnet i figur 9.94.

Figur 9.94
“Predikator”-basert
regulering av dybden
til en farkost



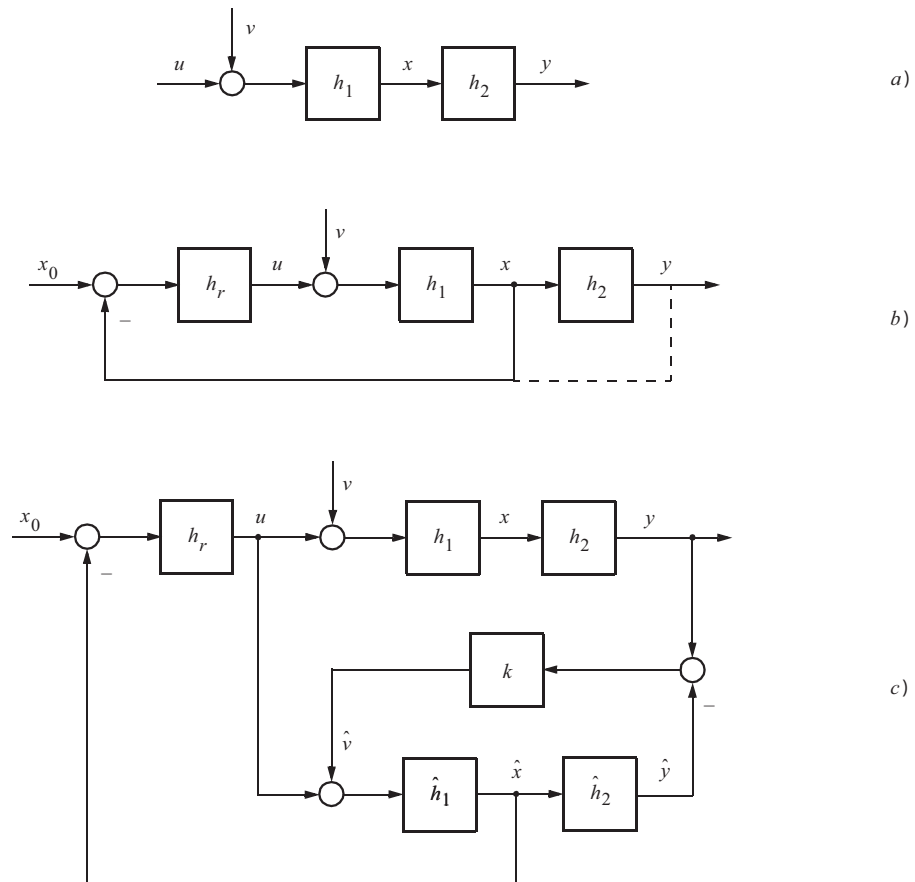
■

9.7 ESTIMATORBASERT REGULERING

9.7.1 Forenklet “tilstandsestimator” ved regulering av prosess med stor forsinkelse

Vi vil ofte finne at det i praksis er vanskelig å få arrangert målinger så “tidlig” i prosessen som ønskelig. Dette er illustrert i figur 9.95 a), der det er vist en monovariabel prosess som består av to kaskadekoblede systemdeler med transferfunksjonene $h_1(s)$ og $h_2(s)$.

Figur 9.95
Prosess med to
kaskadekoblede
systemdeler



$h_1(s)$ kan vi tenke oss er av minimum-fase-type (for eksempel $h_1(s) = K/(1 + Ts)$), mens $h_2(s)$ er av ikke-minimum-fase-type (for eksempel $h_2(s) = e^{-\tau s}$).

Vi forestiller oss for enkelhets skyld at alle prosessens forstyrrelser er representert ved en felles påvirkning v som virker inn på samme sted som pådraget u .

Den størrelsen vi egentlig ønsker å regulere er x som vist i figur 9.95 b), men siden x ikke er målbar, er vi henvist til å lage tilbakekoplingen som vist stippet fra målingen y . Dette vil ikke være noe problem dersom forsinkelsene i $h_2(s)$ er små. Er de derimot store, har vi et problem. Det kan i noen grad omgås ved å benytte strukturen som er vist i figur 9.95 c).

Her er det bygget en **estimator** som består av en matematisk modell av prosessen, og som arbeider i parallell med den virkelige prosessen. Modellen er drevet av pådragssignalet u og en størrelse \hat{v} som spiller rollen som et estimat (tilnærmede) av den virkelige prosessforstyrrelsen v . Ved hjelp av transferfunksjonene $h_1(s)$ og $h_2(s)$ i modellen genereres det nå estimater av størrelsene x og y , som gis betegnelsen henholdsvis \hat{x} og \hat{y} . Estimateret \hat{y} av den egentlige målingen subtraheres fra målingen y og differansen genererer gjennom elementet med transferfunksjon $k(s)$ estimateret \hat{v} av forstyrrelsen.

Den tilbakekoplingen fra x i figur 9.95 b) som ikke kan realiseres på grunn av at x ikke er målbar, erstattes nå med tilbakekopling fra estimateret \hat{x} som vist i figur 9.95 c).

Denne strukturen er en forenkling av en mer generell struktur som vi kan komme fram til for multivariable prosesser (se avsnitt 10.5) på basis av teori for *optimal tilstandsestimering* og *optimal regulering*. Vi skal her bare drøfte den forenklete monovariable løsningen.

I en undersøkelse av systemet i figur 9.95 c) antar vi først for enkelhets skyld at $\hat{h}_1(s) = h_1(s)$ og $\hat{h}_2(s) = h_2(s)$. Det vil si at vi er i stand til å modellere den fysiske prosessen perfekt. Senere skal vi se på innvirkningen av eventuelle forskjeller mellom modell og system.

Når modellen er perfekt, finner vi for systemet i figur 9.95 c)

$$x(s) = M_r(s)x_0(s) + h_1(s)(1 - M_r(s)M_e(s))v(s) \quad (9.144)$$

der

$$M_r(s) = \frac{h_r(s)h_1(s)}{1 + h_r(s)h_1(s)} = \begin{array}{l} \text{følgeforhold for regulert} \\ \text{minimum fase prosessdel} \end{array} \quad (9.145)$$

og

$$M_e(s) = \frac{k(s)h_1(s)h_2(s)}{1 + k(s)h_1(s)h_2(s)} = \text{følgeforhold for estimeringssløyfen} \quad (9.146)$$

Siden $h_1(s)$ er en minimum-fase-transferfunksjon, vil vi kunne realisere en reguleringsløyfe som beskrevet i (9.145), som har svært høy båndbredde (det vil si at $M_r(j\omega) \approx 1$ til meget høy frekvens). Dette medfører at (9.144) antar formen

$$x(s) = x_0(s) + h_1(s)N_e(s)v(s) \quad (9.147)$$

der

$$N_e(s) = \frac{1}{1 + k(s)h_1(s)h_2(s)} \quad (9.148)$$

Dette resultatet må sammenliknes med det vi ville ha fått med systemet basert på den stiplede tilbakekoplingen i figur 9.95 b). I denne tilbakekopplingsløyfen ville regulatoren måtte få nøyaktig den samme transferfunksjonen som den tilsvarende i estimeringsløyfen, nemlig

$$h_r(s) = k(s) \quad (9.149)$$

Med denne regulatoren innsatt ville vi ha fått

$$x(s) = \frac{1}{h_2(s)}M_e(s)x_0(s) + h_1(s)N_e(s)v(s) \quad (9.150)$$

Forskjellen mellom (9.147) og (9.150) er første leddet. Der ser vi at følgeresponsen er perfekt i den estimatorbaserte løsningen (figur 9.95 c)), mens den ordinære løsningen (figur 9.95 b) stiplet) er beheftet med transferfunksjonen $M_e(s)$, som på grunn av ikke-minimum-fase-funksjonen $h_2(s)$ vil ha meget lav båndbredde. Det siste leddet i de to løsningene $h_1(s)N_e(s)v(s)$ er identisk.

De to løsningene oppfører seg med andre ord likt med hensyn på responsen til forstyrrelsen v , men forskjellig med hensyn på x_0

Det har interesse å se mer i detalj på transferfunksjonen $M_e(s)$ for følgeforholdet for estimatoren og tilbakekoplingen stiplet i figur 9.95 b).

Går vi ut fra at $h_1(s) = K/(1 + Ts)$, $h_2(s) = e^{-\tau s}$ og $k(s) = K_p(1 + T_i s)/(T_i s)$ og velger $T_i = T$, finner vi for sløyfen

$$h_{0e}(s) = \frac{K_p K e^{-\tau s}}{T s} \quad (9.151)$$

ω_{180} for denne sløyfen vil bli $\omega_{180} = \pi/(2\tau)$. Fordi amplitudekarakteristikken faller med vinkelkoeffisient (-1) , kan vi velge å innstille parametrene slik at vi får en kryssfrekvens $\omega_c = \omega_{180}/2 = \pi/(4\tau)$. Siden sløyfeforsterkningen ved ω_c skal være lik 1, finner vi $|h_{0e}(j\omega)| = K_p K/(T\omega_c) = 1$ som gir $K_p K/T = \pi/(4\tau)$

Da finner vi

$$M_e(s) = \frac{h_{0e}(s)}{1 + h_{0e}(s)} = \frac{\frac{\pi e^{-\tau s}}{4\tau s}}{1 + \frac{\pi e^{-\tau s}}{4\tau s}} \quad (9.152)$$

I (9.152) kan vi sette inn approksimasjonen

$$e^{-\tau s} \approx \frac{1 - \frac{\tau}{2}s}{1 + \frac{\tau}{2}s} \quad (9.153)$$

for å få et tilnærmet uttrykk av $M_e(s)$ som en rasjonal transferfunksjon. Da finner vi

$$M_e(s) = \frac{1 - \frac{\tau}{2}s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \quad (9.154)$$

der

$$\omega_0 = \sqrt{\frac{\pi}{2}} \frac{1}{\tau} = 1.253 \frac{1}{\tau}$$

$$\zeta = \left(\frac{2}{\pi} - \frac{1}{4}\right) \sqrt{\frac{\pi}{2}} = 0.485$$

Vi kan nå komme tilbake til den opprinnelige formen ved å skrive (9.154) slik:

$$M_e(s) = \frac{1 + \frac{\tau}{2}s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} \frac{1 - \frac{\tau}{2}s}{1 + \frac{\tau}{2}s} = \frac{1 + \frac{\tau}{2}s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} e^{-\tau s} \quad (9.155)$$

Setter vi (9.155) inn i (9.150), finner vi for systemet i figur 9.95 b) (stiplet)

$$x(s) = \frac{1 + \frac{\tau}{2}s}{1 + 2\zeta \frac{s}{\omega_0} + \left(\frac{s}{\omega_0}\right)^2} x_0(s) + h_1(s) N_e(s) v(s)$$

$$\approx \frac{1}{1 + \frac{\tau}{\pi}s} x_0(s) + h_1(s) N_e(s) v(s) \quad (9.156)$$

Resultatet i (9.156) må sammenliknes med resultatet i (9.147).

Vi ser at forskjellen er dramatisk til fordel for den estimatorbaserte løsningen når vi ønsker en endring i referansen x_0 .

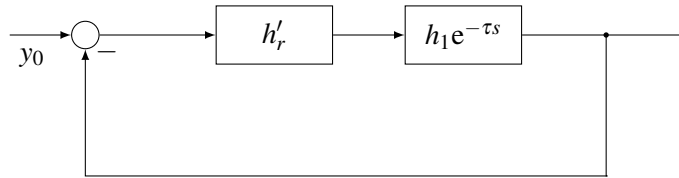
9.7.2 Estimatorbasert regulering – Otto-Smith-regulator

Gitt en prosess med en stor tidsforsinkelse τ . Vi skal nå finne en bedre løsning enn den vi fant i eksempel 9.6.

$$h_u(s) = h_1(s)e^{-\tau s} \quad (9.157)$$

Gitt en reguleringsløyfe med en regulator h'_r , se figur 9.96.

Figur 9.96



Fra figuren har vi:

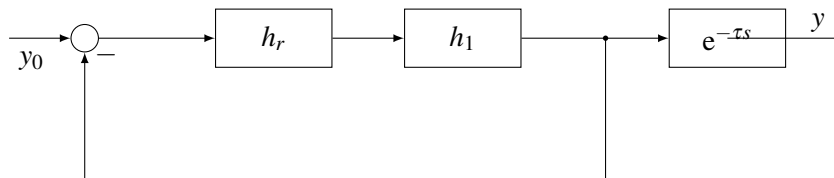
$$\frac{y}{y_0} = \frac{h'_r(s)h_1(s)e^{-\tau s}}{1 + h'_r(s)h_1(s)e^{-\tau s}} \quad (9.158)$$

Vi setter dette uttrykket lik

$$\frac{h_r(s)h_1(s)}{1 + h_r(s)h_1(s)} e^{-\tau s} \quad (9.159)$$

dvs, vi håper at det er mulig å finne en $h'_r(s)$ slik at vi kan behandle problemet som om transportforsinkelsen var utafor den lukkede sløyfen. Da kan vi velge $h_r(s)$ som en vanlig seriekompensatorer basert på strukturen Dvs. vi kan se bort fra $e^{-\tau s}$ i valget

Figur 9.97



av $h_r(s)$. Vi prøver altså

$$\frac{h'_r(s)h_1(s)e^{-\tau s}}{1 + h'_r(s)h_1(s)e^{-\tau s}} = \frac{h_r(s)h_1(s)}{1 + h_r(s)h_1(s)} e^{-\tau s} \quad (9.160)$$

stryker vi fellesledd på begge sider, får vi

$$\frac{h'_r(s)}{1 + h'_r(s)h_1(s)e^{-\tau s}} = \frac{h_r(s)}{1 + h_r(s)h_1(s)} \quad (9.161)$$

Hvis vi løser dette med hensyn på $h'_r(s)$:

$$h'_r(s)(1 + h_r(s)h_1(s)) = h_r(s)(1 + h'_r(s)h_1(s)e^{-\tau s}) \quad (9.162)$$

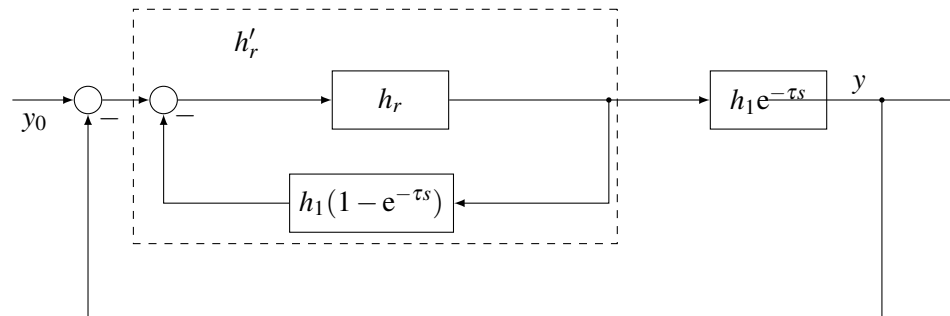
$$h'_r(s)(1 + h_r(s)h_1(s) - h_r(s)h_1(s)e^{-\tau s}) = h_r(s) \quad (9.163)$$

finner vi til slutt

$$h'_r(s) = \frac{h_r(s)}{1 + h_r(s)h_1(s)(1 - e^{-\tau s})} \quad (9.164)$$

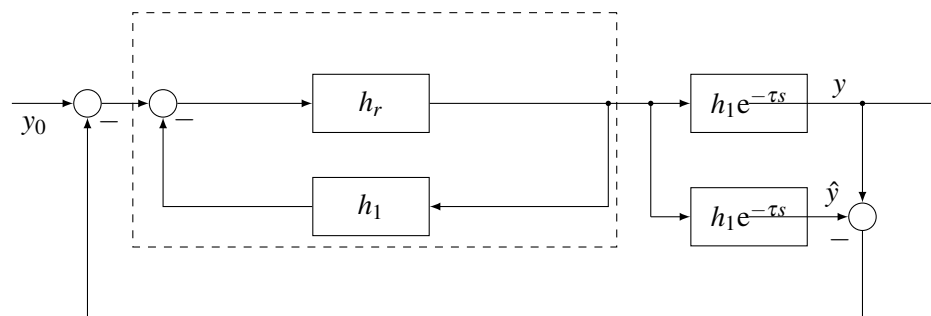
Dette svarer til en struktur som vist i Figur 9.98

Figur 9.98



Denne kan med litt omstrukturering omgjøres til systemet vist i Figur 9.99.

Figur 9.99

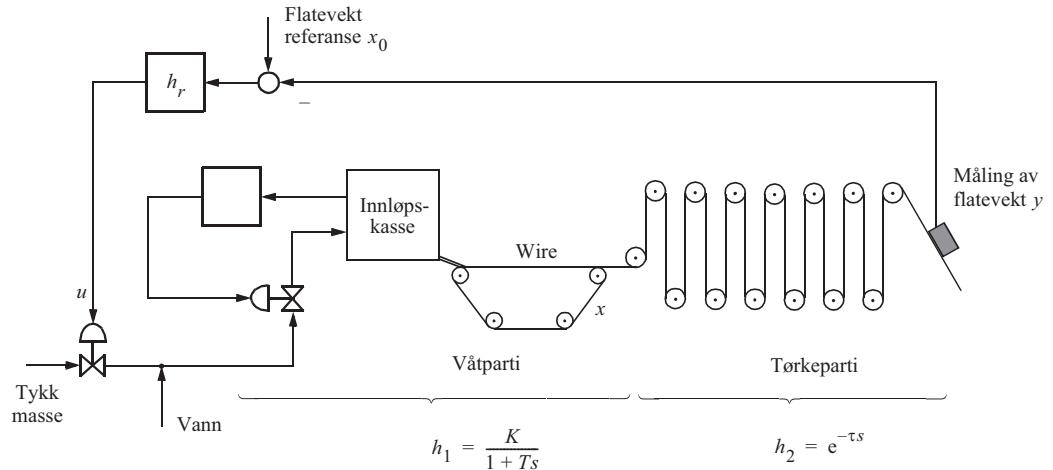


Ved perfekt overensstemmelse mellom y og \hat{y} som er estimert via en modell i parallell med prosessen $h_1(s)e^{-\tau s}$, ser vi at reguleringsproblemet forenkles til det vi ønsket; en seriekompensator $h_r(s)$ bare basert på prosessdelen $h_1(s)$, som ikke inneholder en transportforsinkelse! Figur 9.99 viser at vi har endt opp med et tilfelle av noe som det finnes mange varianter av i reguleringsteknikken: Estimatorbasert regulering. Denne forutsetter at den parallelle modellen ikke avviker for mye fra prosessen man skal styre. Estimatorbasert regulering realiseres oftest basert på tilstandsrommodeller, og ofte for multivariable prosesser.

EKSEMPEL 9.23: Regulering av fiberkonsentrasjonen i en papirmaskin

En problemstilling som illustrert i figur 9.95, har vi i en papirmaskin der en viktig oppgave er regulering av konsistensen (fiberkonsentrasjonen) i den svært fortynnede fibersuspensjonen som tilføres maskinens innløpskasse for å oppnå en spesifisert “flatevekt” av det endelige papiret. Dette er illustrert i figur 9.100.

Figur 9.100
Regulering av “flatevekten” til papir

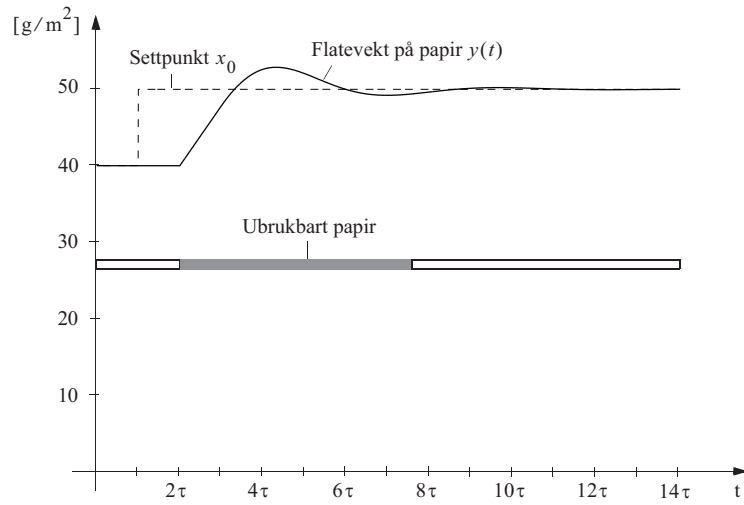


Papirmaskinens *våtparti* (tilførselspumpen, innløpskassen, første del av wirepartiet) kan gjerne representeres med en 1. ordens transferfunksjon ($h_1(s) = K/(1 + Ts)$) mens tørkepartiet (pressparti, tørkesylindre etc.) utgjør en stor transportforsinkelse ($h_2(s) = e^{-\tau s}$).

I moderne papirproduksjon ønskes en hyppig omstilling av papirkvaliteten (flatevekt, etc.) etter kundespesifikasjon. Et system som skissert i figur 9.100, med regulering av konsistensen for innløpskassen på basis av måling av flatevekten etter tørkepartiet, vil få svært lav båndbredde (ω_{180} lav). Derfor vil vi som følge av en omstilling av settpunktet (referansen) for eksempel for flatevekt få produsert store mengder ubrukt papir mens systemet svinger seg inn som vist i figur 9.101. Ved bruk av en Otto-Smith-regulator vil omstillingen foregå svært mye raskere som vist i figur 9.102. Dette resulterer i økt produktivitet for prosessen.

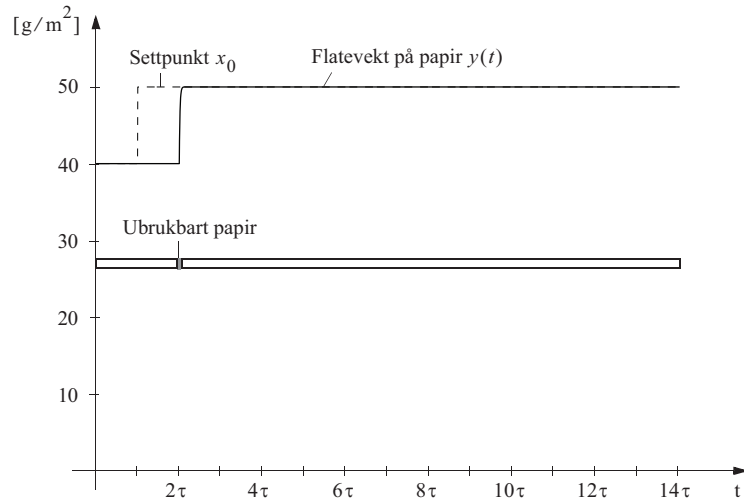
Figur 9.101

Endring av sett-
punktet for
flatevekt



Figur 9.102

Endring av sett-
punktet for flate-
vekt med
estimatorbasert
regulering



9.8 PARAMETRISK OPTIMERING AV SYSTEMER MED GITT STRUKTUR

Det finnes en betydelig teori for *optimal* dimensjonering av multivariable reguleringsystemer, og strukturen for en slik løsning blir antydnet i avsnitt 10.5. En beslektet, men forenklet problemstilling, får vi ved å gå ut fra en gitt *struktur* karakterisert ved et *sett parametre*, og så med en *optimaliseringsprosedyre* finne fram til de gunstigste verdiene av disse parametrene. For å få dette til må en definere en *objekt-funksjon* (*objektfunksjonal*) som gir et mål for systemets godhet. Det optimale sett av parametre er da det som gir en *ekstremalverdi* (maksimum eller minimum avhengig av formuleringen) for denne objektfunksjonen.

9.8.1 Kvadratisk objektfunksjon

Vi kan formulere **objektfunksjoner** på mange måter, men ikke alle egner seg like godt for analytisk behandling. Særlig interesse knytter det seg til objektfunksjoner som i det multivariable tilfellet har formen

$$L(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{P} \mathbf{u} \quad (9.165)$$

der \mathbf{Q} og \mathbf{P} er matriser.

I et skalart tilfelle vil en tilsvarende funksjon kunne anta formen

$$L(x, u) = \alpha x^2 + \beta u^2 \quad (9.166)$$

Likning (9.165) og (9.166) er skalare størrelser som gir et momentant mål for viktige prosessvariables utslag. Beregner vi integralet av dette momentane målet over et visst tidsintervall (for eksempel fra 0 til ∞) eller gjennomsnittsverdien, får vi det som betegnes med en **objektfunksjonal**. Et enkelt eksempel på en slik er

$$J = \int_0^{\infty} (\alpha x^2 + \beta u^2) dt \quad (9.167)$$

Objektfunksjonalen i (9.167) er særlig interessant, fordi vi kan bestemme verdien av integralet ved hjelp av Laplacetransformasjon. I avsnitt 4.2.1 finnes en rekke regler for bruk av Laplacetransformasjon. Vi skal i det følgende benytte oss av (4.14), (4.20) og (4.23).

Bruk av (4.14) og (4.23), gir

$$\int_0^{\infty} g(t) dt = \lim_{t \rightarrow \infty} \int_0^t g(t) dt = \lim_{s \rightarrow 0} s \frac{g(s)}{s} = g(s=0) \quad (9.168)$$

Bruk av dette resultatet sammen med (4.20), gir

$$\int_0^{\infty} f_1(t) f_2(t) dt = \frac{1}{2\pi j} \int_{\gamma-j\infty}^{\gamma+j\infty} f_1(-w) f_2(w) dw \quad (9.169)$$

For det spesielle tilfellet at vi i (9.169) har $f_1(t) = f_2(t) = f(t)$, finner vi

$$\int_0^{\infty} f(t)^2 dt = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} f(-w) f(w) dw = \frac{1}{\pi} \int_0^{\infty} |f(j\omega)|^2 d\omega \quad (9.170)$$

idet det er antatt at $\gamma = 0$. (Når venstre del av (9.170) eksisterer, kan vi tillate oss å sette $\gamma = 0$.)

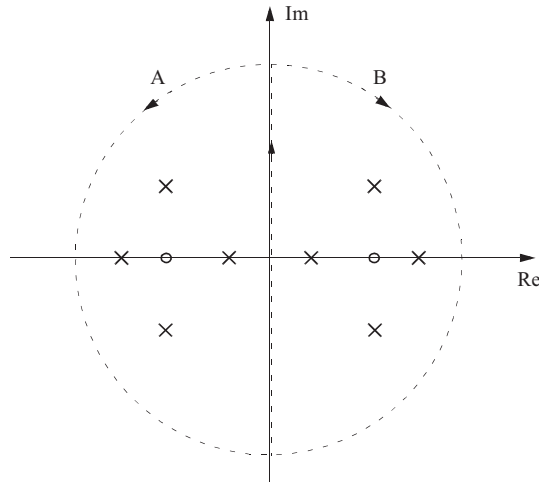
(9.170) kalles **Parsevals sats**.

Vi ser at venstre side i (9.170) er av samme form som (9.167). Vi kan derfor benytte oss av Laplacetransformen til de tidsforløp som inngår under objektfunksjonalens integral og har derved overført problemet fra å være en integrasjon i tidsplanet til å være en integrasjon langs hele den imaginære akse i det komplekse plan. Integralet i midten i likning (9.170) kan bestemmes ved hjelp av residusatsen (avsnitt 4.4.1).

Når $f(t)$ er kjent og dens Laplacetransform er en rasjonal funksjon med poler (singulariteter) i venstre halvplan, vil integranden til høyre i (9.170) være karakterisert

ved poler som vist i figur 9.103. Integralet langs den imaginære akse kan, dersom integrandens tallverdi går mot null når argumentet går mot uendelig, erstattes med et kurveintegral rundt hele venstre halvplan som antydnet med halvsirkel *A*, eller rundt hele høyre halvplan som antydnet med halvsirkel *B*. Uansett om vi velger integrasjonsbane *A* eller *B*, vil resultatet bli det samme.

Figur 9.103
Polene til integranden på høyre side i (9.170) når $f(s)$ er en rasjonal funksjon



Velger vi integrasjonsbane *A*, vil integralet bli lik summen av residuene i polene til $f(s)$. Velger vi integrasjonsbane *B*, vil verdien av integralet bli summen av residuene i polene til $f(-s)$.

Vi illustrerer dette med et enkelt eksempel.

EKSEMPEL 9.24: Beregning av tidsintegral

Dersom $f(t) = e^{at}$ og $a < 0$, finner vi

$$f(s)f(-s) = \frac{1}{(s-a)(-s-a)} \quad (9.171)$$

som gir

$$\int_0^{\infty} f^2(t) dt = \text{res } f(s)|_{s=a} = (s-a) \frac{1}{(s-a)(-s-a)} \Big|_{s=a} = -\frac{1}{2a} \quad (9.172)$$

Dette svaret kan kontrolleres ved å integrere tidsfunksjonen $f^2(t)$.
 ■

Ved analytisk behandling av reguleringsystemer vet vi imidlertid at det ofte vil være brysomt å bestemme polene for et lukket system hvis dette er av høyere orden enn tredje. Hvis $f(s)$ er en rasjonal funksjon, er vi imidlertid så heldige at det er mulig å uttrykke integralets verdi som funksjon av den rasjonale funksjonens polynomkoeffisienter. Denne metoden går som følger:

Vi antar

$$f(s) = \frac{p(s)}{q(s)} \quad (9.173)$$

der $p(s)$ og $q(s)$ er polynomer i s , og polynomet $p(s)$ er av en grad lavere enn polynomet $q(s)$, altså

$$p(s) = p_{n-1}s^{n-1} + p_{n-2}s^{n-2} + \dots + p_1s + p_0 \quad (9.174)$$

$$q(s) = q_n s^n + q_{n-1}s^{n-1} + \dots + q_1s + q_0 \quad (9.175)$$

Da kan integralet i (9.170) skrives

$$J_{(n)} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \frac{p(s)}{q(s)} \frac{p(-s)}{q(-s)} ds \quad (9.176)$$

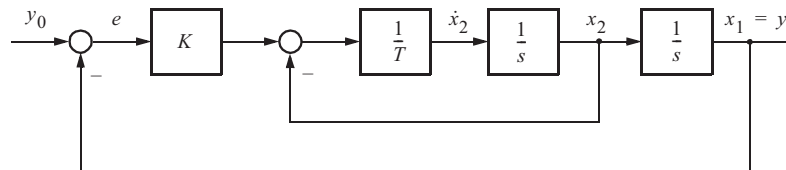
Indekset n står for graden til polynomet i nevneren.

Tabell 9.4 gir verdiene av integralet som funksjon av koeffisientene i de to polynomene opp til $n = 5$. Større tabeller finnes i Newton, Gould, Kaiser (1957).

Vi har nå verktøyet som trengs for analytisk optimering av et system med gitt struktur på basis av en objektfunksjonal av typen (9.167). Et eksempel vil demonstrere dette.

EKSEMPEL 9.25: Analytisk optimalisering av følgesystem

Figur 9.104 viser et blokkdiagram for et enkelt følgesystem der vi har valgt å benytte oss av seriekompensasjon med en forsterkning K . Tilbakekoplingen fra hastigheten x_2 utgjøres av den naturlige dempingen i systemet. Problemet er å finne den verdien av K som minimaliserer en hensiktsmessig objektfunksjonal når referansen $y_0(t)$ antar en realistisk form.



Figur 9.104
Den dynamiske
prosessen

TABELL 9.4 Kvadratiske objektfunksjoner

Verdi av (9.176) for forskjellige n	
$J_{(1)}$	$\frac{p_0^2}{2q_0q_1}$
$J_{(2)}$	$\frac{p_1^2q_0 + p_0^2q_2}{2q_0q_1q_2}$
$J_{(3)}$	$\frac{p_2^2q_0q_1 + (p_1^2 - 2p_0p_2)q_0q_3 + p_0^2q_2q_3}{2q_0q_3(-q_0q_3 + q_1q_2)}$
$J_{(4)}$	$\frac{p_3^2(-q_0^2q_3 + q_0q_1q_2) + (p_2^2 - 2p_1p_3)q_0q_1q_4 + (p_1^2 - 2p_0p_2)q_0q_3q_4 + p_0^2(-q_1q_4^2 + q_2q_3q_4)}{2q_0q_4(-q_0q_3^2 - q_1^2q_4 + q_1q_2q_3)}$
$J_{(5)}$	$\frac{1}{2\Delta_5}(p_4^2m_0 + (p_3^2 - 2p_2p_4)m_1 + (p_2^2 - 2p_1p_3 + 2p_0p_4)m_2 + (p_1^2 - 2p_0p_2)m_3 + p_0^2m_4)$ <p>der</p> $m_0 = \frac{1}{q_5}(q_3m_1 - q_1m_2) \quad , \quad m_3 = \frac{1}{q_0}(q_2m_2 - q_4m_1)$ $m_1 = -q_0q_3 + q_1q_2 \quad , \quad m_4 = \frac{1}{q_0}(q_2m_3 - q_4m_2)$ $m_2 = -q_0q_5 + q_1q_4 \quad , \quad \Delta_5 = q_0(q_1m_4 - q_3m_3 + q_5m_2)$

Vi velger å studere objektfunksjonen

$$J = \int_0^{\infty} (\alpha e^2(t) + \beta \dot{x}_2^2(t)) dt \quad (9.177)$$

Det første leddet under integraltegnet er et mål for følgeavviket. Vi ser at positive avvik veier like mye som negative avvik. Videre ser vi at store avvik veies sterkere enn små avvik. Det andre leddet i integralet ivaretar et ønske om å begrense akselerasjonen $\dot{x}_2(t)$ til prosessen. Også her vil positive verdier veies like mye som negative verdier, og store akselerasjoner veies forholdsvis sterkere enn små akselerasjoner. Forholdet mellom vekten på følgeavviket og akselerasjonen er gitt av faktorene α og β , der vi gjerne kan velge $\alpha = 1$.

Grunnen til at vi gjerne vil ha med et ledd av typen $\dot{x}_2^2(t)$ eller en annen tilsvarende intern prosessstørrelse i objektfunksjonalen, er at denne gir et mål for den dynamiske påkjenningen på prosessen. Et eksempel på dette har en ved styring av en bil. Det er riktignok viktig at bilen følger veien med lite posisjonsavvik, men det er også viktig at sideakselerasjonene ikke blir for store fordi det ellers vil være ubehagelig for passasjerene. Vi finner derfor ved kjøring av bil å måtte foreta et kompromiss mellom sideavvik og akselerasjon og bør avpasse farten slik at totalresultatet blir akseptabelt.

For systemet i figur 9.104 finner vi

$$e(s) = N(s)y_0(s) \quad (9.178)$$

$$\dot{x}_2(s) = s^2 h_0(s)N(s)y_0(s) \quad (9.179)$$

der

$$N(s) = \frac{1}{1 + h_0(s)} \quad (9.180)$$

og

$$h_0(s) = \frac{K}{s(1 + Ts)} \quad (9.181)$$

Vi velger å optimere systemet for et referansesignal av typen

$$y_0(s) = \frac{y_{00}}{s(1 + T_0s)} \quad (9.182)$$

som er et mykt sprang med amplitude y_{00} og tidskonstant T_0 . Vi finner da

$$e(s) = \frac{T y_{00} s + y_{00}}{T T_0 s^3 + (T + T_0) s^2 + (T_0 K + 1) s + K} \quad (9.183)$$

$$\dot{x}_2(s) = \frac{K y_{00} s}{T T_0 s^3 + (T + T_0) s^2 + (T_0 K + 1) s + K} \quad (9.184)$$

Disse rasjonale funksjonene har begge samme polynom i nevneren, og polynomene i tellerne er av 1. orden. Det blir derfor enkelt å beregne integralet av typen (9.176) ved hjelp av tabell 9.4.

Vi finner

$$J_{(3)} = \frac{y_{00}^2}{2} \frac{T^2 K + (T + T_0) + \beta K^3}{K(K T_0^2 + T + T_0)} \quad (9.185)$$

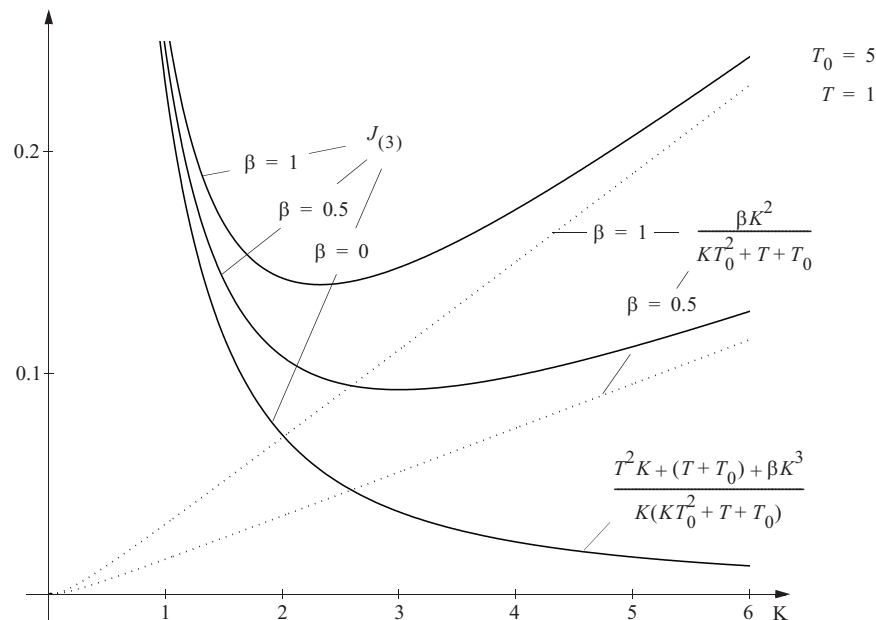
Objektfunksjonens verdi som funksjon av K , er vist i figur 9.105. Vi ser at forløpet ved små verdier av K blir omtrent omvendt proporsjonalt med K , mens forløpet ved store verdier av K blir omtrent proporsjonalt med K .

Vi vil få en minimal verdi av objektfunksjonen ved en middels stor verdi av K . Denne bestemmes av

$$\frac{\partial J_{(3)}}{\partial K} = 0 \quad (9.186)$$

som gir en fjerdegradslikning i K . Figur 9.105 viser tydelig at dersom vi ikke legger noen vekt på akselerasjonen, det vil si ved å velge $\beta = 0$, vil den optimale forsterkingen bli $K = \infty$. Hvordan det da går, har vi tidligere undersøkt i eksempel 4.12, og resultatet er vist i figur 4.27 og 4.28. Systemet vil bli sterkt oscillatorisk, men vil allikevel (bedømt ut i fra den valgte objektfunksjonen) bli det optimale. Figur 9.105 viser ellers at jo mer vekt vi legger på akselerasjonen, det vil si $\beta > 0$, jo lavere vil den optimale verdien av K bli.

Figur 9.105
Objektfunksjonen som funksjon av K



Eksemplet viser at det er ganske lett, både prinsipielt og med hensyn til numerisk beregning, å få et kvantitativt mål for objektfunksjonen i form av den type uttrykk som er gitt i (9.185). Det er en kurant oppgave å skrive et datamaskinprogram for utregning av slike uttrykk, selv for systemer av forholdsvis stor kompleksitet. Det som imidlertid ikke er så enkelt, er å løse likninger av typen (9.186) analytisk. For praktiske systemer vil det derfor være aktuelt å anvende en numerisk søkemetode for

å finne fram til optimale parametre. Det finnes en rekke slike metoder tilgjengelig.

I det foregående eksemplet var det bare en parameter (K) som skulle optimeres. I et praktisk tilfelle kan en tenke seg at flere parametre skal innstilles til sine optimale verdier. Disse parametrene kan samles til en *parametervektor*, for eksempel \mathbf{k} . Den optimale parametervektoren bestemmes da ved hjelp av gradienten til objektfunksjonen med hensyn på parametervektoren

$$\frac{\partial J}{\partial \mathbf{k}} = \left[\frac{\partial J}{\partial k_1} \quad \frac{\partial J}{\partial k_2} \quad \dots \quad \frac{\partial J}{\partial k_p} \right]^T = \mathbf{0} \quad (9.187)$$

9.8.2 Objektfunksjon basert på systemets frekvensrespons

I Balchen (1958) er det utledet en metode som kan ha endel fordeler framfor den som er demonstrert i avsnitt 9.8.1.

En objektfunksjonal svarende til den i (9.167) som kunne være interessant, er

$$J = \int_0^{\infty} |x(t)| dt \quad (9.188)$$

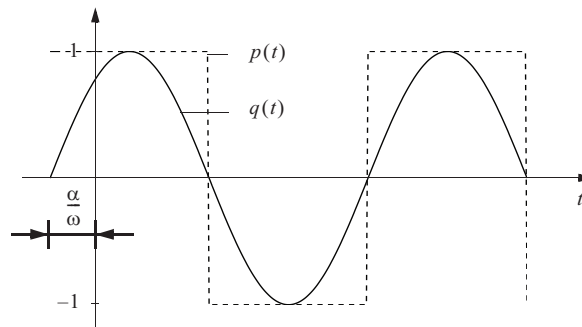
der $x(t)$ for eksempel kan være reguleringsavviket i et tilbakekoplet system som vist i figur 9.104, eller en annen avledet størrelse. Slik uttrykket står i (9.188), kan det ikke finnes noe analytisk svar på en enkel måte. Vi kan imidlertid skrive

$$|x(t)| = x(t) \frac{|x(t)|}{x(t)} = x(t) \operatorname{sgn} x(t) = x(t)p(t) \quad (9.189)$$

der funksjonen $p(t)$ vil være en “firkantbølge” med amplitude ± 1 som vist i figur 9.106.

Figur 9.106

Eksempel på
 $p(t)$ og $q(t)$



Dersom denne firkantbølgen er tilnærmet periodisk med konstant frekvens, er det rimelig å tilnærme den med en sinussvingning med samme frekvens, det vil si samme nullgjennomgangspunkter. Vi erstatter altså $p(t)$ med en ny funksjon

$$q(t) = \sin(\omega t + \alpha) \quad (9.190)$$

der ω er frekvensen og α fasevinkelen for den sinussvingningen som best tilnærmer den riktige firkantbølgen. Vi får da en ny objektfunksjonal definert slik

$$J(\omega, \alpha) = \int_0^{\infty} x(t) \sin(\omega t + \alpha) dt \quad (9.191)$$

Siden uttrykket i (9.191) er mindre eller i beste fall lik uttrykket i (9.188), vil vi søke den kombinasjonen av ω og α som gjør $J(\omega, \alpha) = \text{maksimum}$. Dette gir den sinussvingningen $q(t)$ som best tilnærmer $p(t)$.

Vi skriver nå

$$\sin(\omega t + \alpha) = \frac{e^{j\omega t} e^{j\alpha} - e^{-j\omega t} e^{-j\alpha}}{2j} \quad (9.192)$$

som innsatt i (9.191) gir

$$J(\omega, \alpha) = \frac{1}{2j} \left(e^{j\alpha} \int_0^{\infty} x(t) e^{j\omega t} dt - e^{-j\alpha} \int_0^{\infty} x(t) e^{-j\omega t} dt \right) \quad (9.193)$$

Da

$$\int_0^{\infty} g(t) dt = \lim_{s \rightarrow 0} g(s) \quad (9.194)$$

der

$$g(s) = L(g(t))$$

og

$$L(x(t) e^{\pm j\omega t}) = x(s \mp j\omega) \quad (9.195)$$

finner vi

$$J(\omega, \alpha) = \frac{1}{2j} (e^{j\alpha} x(-j\omega) - e^{-j\alpha} x(j\omega))$$

Vi innfører videre

$$x(\pm j\omega) = |x(j\omega)|e^{\pm\angle x(j\omega)}$$

som gir

$$J(\omega, \alpha) = |x(j\omega)| \sin(\alpha - \angle x(j\omega))$$

Denne funksjonen får sitt maksimum med hensyn på α når

$$\sin(\alpha - \angle x(j\omega)) = 1$$

og

$$|x(j\omega)| = \text{maksimum}$$

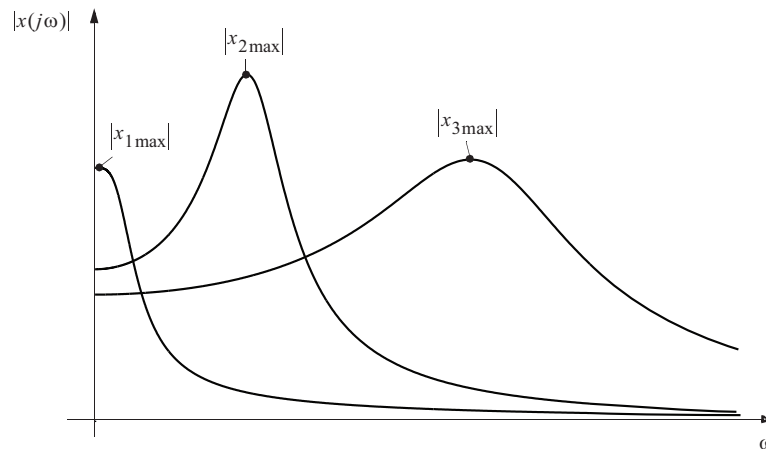
Funksjonen

$$x(j\omega) = F(x(t))$$

er Fourier-transformen av $x(t)$. Vi ser at den nye objektfunksjonen er uttrykt som maksimalverdien med hensyn på ω av Fourier-transformens tallverdi. Altså

$$J(\cdot) = \max_{\omega} |x(j\omega)| \quad (9.196)$$

Fourier-transformen av en tidsvariabel finnes direkte av dens Laplacetransform ved å erstatte variabelen s med $j\omega$. Figur 9.107 viser hvordan objektfunksjonalens verdi finnes som maksimalverdien av $|x(j\omega)|$.



Figur 9.107
Spektral representasjon av Fourier-transformen

En oppsummering av de forskjellige objektfunksjonene J_i som er definert hittil, vil gi

$$J_1 = \int_0^{\infty} x^2(t) dt = \frac{1}{\pi} \int_0^{\infty} |x(j\omega)|^2 d\omega \quad (9.197)$$

$$J_2 = \int_0^{\infty} |x(t)| dt \quad (9.198)$$

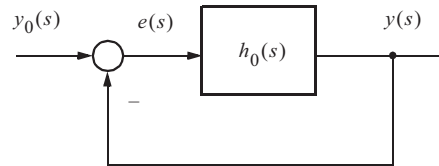
$$J_3 = \max_{\omega} |x(j\omega)| \quad (9.199)$$

En rekke andre typer objektfunksjonaler kan defineres, men de skal ikke behandles her.

Objektfunksjonen J_3 i (9.199) baserer seg som sagt på Fourier-transformen $x(j\omega)$ av respons-funksjonen $x(t)$. Men responsen av et system finnes av systemets transferfunksjon multiplisert med eksitasjonens Laplacetransform. Ser vi for eksempel på systemet i figur 9.108 og ønsker å optimalisere med hensyn på avviket ($e(s)$, $e(t)$) ved å endre parametre i $h_0(s)$, finner vi som kjent

$$e(s) = \frac{1}{1 + h_0(s)} y_0(s) = N(s) y_0(s) \quad (9.200)$$

Figur 9.108
Blokkdiagram for et monovariabelt system med enhets-tilbakekopling



Vi må nå velge hvilken form eksitasjonen $y_0(s)$ ($y_0(t)$) skal ha fordi resultatet av optimaliseringen uansett hvilken type objektfunksjon vi velger, vil være helt avhengig av hvilken eksitasjon vi har brukt. Vi kan for eksempel velge $y_0(s) = 1$ hvilket betyr at $y_0(t) = \delta(t) =$ enhetsimpuls.

I så fall vil vi for systemet i figur 9.108 få

$$J_3 = \max_{\omega} |e(j\omega)| = \max_{\omega} |N(j\omega)| \quad (9.201)$$

Dette resultatet er svært interessant fordi det direkte viser betydningen av kriteriet vi tidligere har brukt en rekke ganger når vi har tegnet stedkurven for i et Nichols-diagram og bestemt verdien av . Vi kan med andre ord si at å innstille et tilbakekoplet

systems parametre slik at vi oppnår et minimum av $|N(j\omega)|_{\max}$, svarer til å optimalisere systemet med hensyn på objektfunksjonen J_3 når systemeksitasjonen $y_0(t)$ er en enhetsimpuls.

Hadde vi valgt et enhetssprang $y_0(s) = 1/s$ i stedet for en enhetsimpuls, ville vi ha funnet de optimale parametre ut fra å minimalisere

$$J_3 = \max_{\omega} \left| \frac{1}{\omega} N(j\omega) \right| \quad (9.202)$$

Hvis vi har en parametervektor

$$\mathbf{k} = \begin{bmatrix} k_1 & k_2 & \cdots & k_p \end{bmatrix}^T$$

vil vi altså søke etter

$$\min_{\mathbf{k}} (\max_{\omega} |N(j\omega)y_0(j\omega)|) \quad (9.203)$$

EKSEMPEL 9.26: Minimalisering av avviket $e(t)$ ved bruk av J_3

Vi velger i systemet i figur 9.108 med sløyfetransferfunksjon

$$h_0(s) = \frac{1}{T_i s} e^{-\tau s} \quad (9.204)$$

Vi vil optimalisere dette systemet ved bruk av objektfunksjonen J_3 på avviket $e(t)$ når eksitasjonen $y_0(t)$ er et enhetssprang. Parameteren som skal justeres, er $k_1 = T_i/\tau$. Figur 9.109 viser sammenhengen mellom J_3 og T_i/τ der optimum tydelig framgår ved

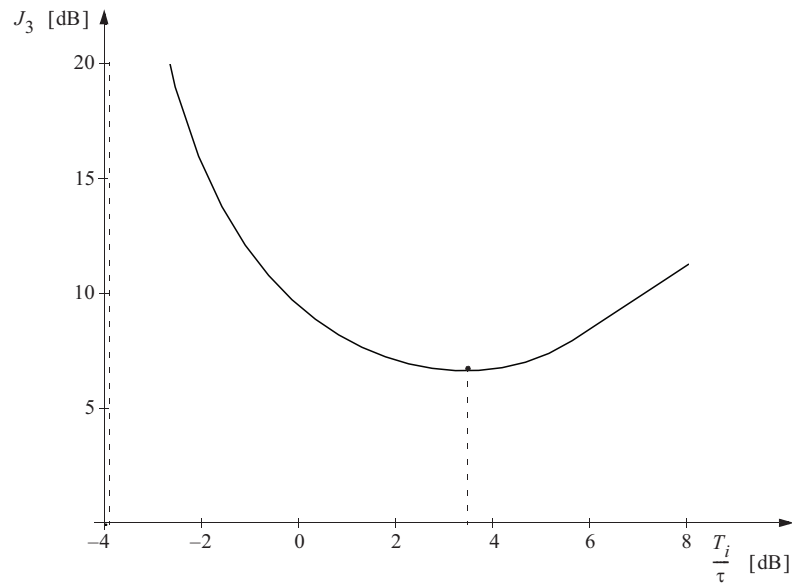
$$\frac{T_i}{\tau} = 3.5 \text{ [dB]} \approx 1.5$$

Dette resultatet er funnet ved hjelp av stedkurver i et Nicholsdiagram, og det er derfor hensiktsmessig å framstille resultatet med dB-skalaer. Stabilitetsgrensen for dette systemet er gitt av

$$\frac{T_i}{\tau} = \frac{2}{\pi} = -3.9 \text{ [dB]}$$

hvilket betyr at vi har funnet fram til et system med forsterkningsmargin $\Delta K = 7.4$ [dB]. Dette er rimelig.

Figur 9.109
Sammenhengen
mellom J_3
og T_i/τ



Etter at ovenstående metode var presentert i 1958, er den blitt generalisert til bruk for multivariable systemer der en frekvensresponsmatrise representeres med sin “største singularverdi” $\bar{\sigma}(j\omega)$. Objektfunksjonen blir da

$$J_3 = \max_{\omega} \bar{\sigma}(j\omega) \quad (9.205)$$

Dette kriteriet er i moderne litteratur gitt betegnelsen H_{∞} (uttales H-uendelig, eng. “H-infinity”).

Singularverdianalyse behandles i avsnitt 10.6.



9.8.3 Regulatorinnstilling ved numerisk optimalisering

Man kan bruke programvare for å stille inn regulatorparametre ”optimalt”. Vi skal studere en optimal innstilling av regulator for likestrømsmotoren gitt i eksempel 7.3 og 7.4. En rimelig transferfunksjon for likestrømsmotoren fra rotorspenningen u til vinkelposisjonen y , er gitt i likning (7.19),

$$h_p(s) = \frac{1}{s(1 + T_2s)(1 + T_1s)} \quad (9.206)$$

Vi velger å regulere likestrømsmotoren med en PD-regulator,

$$h_r(s) = K(1 + T_d s) \quad (9.207)$$

Systemet med regulator får da transferfunksjon

$$h_0(s) = K \frac{(1 + T_d s)}{s(1 + T_2 s)(1 + T_1 s)} \quad (9.208)$$

Fra eksempel 7.4 er $T_1 = 0.5$ og $T_2 = 1/30$, med valgte verdier $K = 40$ og $T_d = 1/6$ ($T_d = T_3$). Betrakt avviksforholdet

$$N(s) = \frac{1}{1 + h_0(s)}, \quad (9.209)$$

med følgefórhóldet

$$M(s) = \frac{h_0(s)}{1 + h_0(s)}, \quad (9.210)$$

og identiteten $M + N = 1$ for monovariablene systemer. Avviket $e(s)$ i et tilbakekoblet system er proporsjonalt med $N(s)$,

$$e(s) = r - y = N(s)r - N(s)h_v(s)v, \quad (9.211)$$

hvor r er referansen og h_v er transferfunksjonen fra forstyrrelsen v . Ideelt er det derfor ónskelig at $|N(j\omega)|$ er liten opp til hóyest mulig frekvens. Videre er det ónskelig at resonanstoppen i $|N(j\omega)|$ er liten, da dette betyr at det lukkede system har tilstrekkelig stabilitetsmargin¹. Fra absoluttverdien av h_0 ,

$$|h_0(j\omega)| = \frac{K}{\omega} \frac{|1 + jT_d\omega|}{|1 + jT_2\omega||1 + jT_1\omega|} = \frac{|t_0(j\omega)|}{|n_0(j\omega)|}, \quad (9.212)$$

kan absoluttverdien av $N(j\omega)$ utledes til

$$\begin{aligned} |N(j\omega)| &= \frac{1}{|1 + h_0(j\omega)|} = \frac{|n_0|}{|n_0 + t_0|} \\ &= \sqrt{\frac{\omega^2(1 + T_2^2\omega^2)(1 + T_1^2\omega^2)}{(K - (T_1 + T_2)\omega^2)^2 + (KT_d\omega + \omega(1 - T_1T_2\omega^2))^2}} \end{aligned} \quad (9.213)$$

La

$$\|N(j\omega)\|_\infty = \max_\omega |N(j\omega)|, \quad (9.214)$$

¹Dessverre kan ikke regulatoren tunes slik at $|N(j\omega)|$ holdes lav over alle frekvenser, fordi $|N(j\omega)|$ er fundamentalt begrenset av Bodes integral,

$$\int_0^\infty \ln |N(j\omega)| d\omega = 0,$$

for stabile systemer.

altså største verdi av $|N(j\omega)|$ over alle frekvenser. Det er ønskelig med lav absoluttverdi $|N(j\omega)|$ for å få lite oscillatorisk tidsrespons, og begrense den maksimale feilen $\max |e(j\omega)|$ (fra likning (9.211)), altså ønsker vi å minimere $\|N(j\omega)\|_\infty$. Men problemet med å minke $\|N(j\omega)\|_\infty$ er at båndbredden til systemet minker. Generelt er det ønskelig med høy båndbredde, da høy båndbredde gir mindre stigetid og hurtigere regulering, siden høyfrekvente signaler lettere passerer gjennom systemet. Båndbredden til et system kan defineres på flere vis. En vanlig definisjon, som i reguleringssteknikkboka, er frekvensområdet opp til kryssfrekvensen ω_c , dvs frekvensområdet hvor $h_0(\omega_c) \geq 1$. Båndbredden til et tilbakekoblet system kan også defineres i form av $N(j\omega)$ og $M(j\omega)$. Vi bruker følgende definisjon av båndbredden ω_B for $N(j\omega)$ Skogestad (2005):

Lukket sløyfe båndbredde ω_B for $N(j\omega)$ er frekvensen hvor $|N(j\omega)|$ først krysser $\frac{1}{\sqrt{2}} \approx -3\text{dB}$ nedenfra.

Båndbredden ω_B for $N(j\omega)$ vil som regel være nær ω_c , men litt mindre (dvs, $\omega_B < \omega_c$). Når vi snakker om båndbredde, bruker vi ω_B og ω_c litt om hverandre. Basert på diskusjonen over, er det tydelig at avveininger må gjøres i design av en regulator. I et forsøk på å finne optimale verdier av K og T_d i forhold til ytelseskrav for systemet, kan vi spesifisere et optimeringsproblem. For sløyfetransferfunksjon (9.208) ønsker vi å beholde virkningen av den rene integratoren for å fjerne stasjonært avvik. Hvis produktet KT_d blir for stort, opphører virkningen av integratoren $\frac{1}{s}$ (hvorfor?). Vi burde derfor ikke velge K og T_d slik at dette produktet blir for stort. En mulighet er å stille inn T_d slik at virkningen av den største tidskonstanten T_1 forkortes vekk, da denne representerer den tregeste dynamikken i systemet. Vi kan da kreve at $T_d \leq T_1$. Samtidig ønsker vi så stor båndbredde som mulig, men at $\|N(j\omega)\|_\infty$ er øvre begrenset av en konstant N^{max} . Dette gir oss et optimeringsproblem, som vi kan skrive slik

$$\max_{K, T_d} \quad \omega_B \quad (9.215a)$$

s.t.¹

$$\|N(j\omega)\|_\infty \leq N^{max} \quad (9.215b)$$

$$T_d \leq T_1 \quad (9.215c)$$

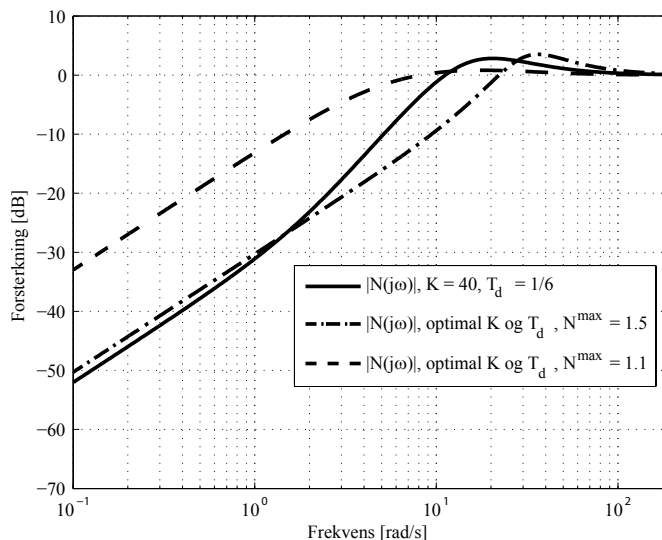
Et vanlig krav er at $N^{max} \leq 6$ dB. Optimeringsproblemet (9.215) kan implementeres i MATLAB og løses med innebygde numeriske algoritmer. Siden begrensningen (9.215b) er ulineær (gitt av likning (9.213) og (9.214)), har vi i dette tilfellet brukt funksjonen `fmincon` som løser generelle ulineære optimeringsproblemer. Figur 9.110 viser $|N(j\omega)|$ for optimerte verdier av K og T_d for henholdsvis $N^{max} = 1.5$ og

¹s.t. = subject to = "under betingelse av at", "med bibetingelser".

$N^{max} = 1.1$, sammenliknet mot de verdier som er brukt i eksempel 7.4. Begrensning

Figur 9.110

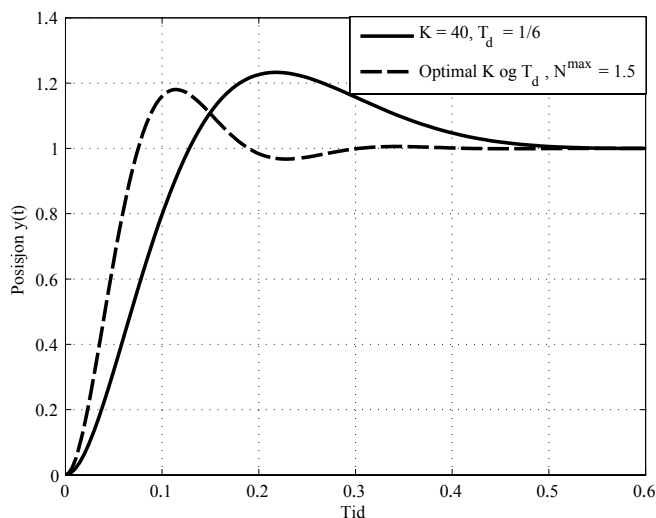
$|N(j\omega)|$ som funksjon av K og N^{max} .



gen (9.215c) fører i begge tilfeller til at $T_d = T_1$. Løsningen av optimeringsproblemet (9.215) for $N^{max} = 1.5$ gir $K = 32.7$ og $\omega_B = 17.4$ rad/s, og responsen av $|N(j\omega)|$ gir tydelig bedre båndbredde, men noe høyere resonanstopp enn verdiene brukt i eksempel 7.4. Dersom vi derimot velger $N^{max} = 1.1$, får vi $K = 4.46$ og $\omega_B = 3.88$ rad/s, altså mindre resonanstopp, men vesentlig dårligere båndbredde. En annen ulempe med å velge $N^{max} = 1.1$ er at $|N(j\omega)|$ blir høyere for lave frekvenser. Dette resulterer i at stasjonær forsterkning $|h_0(j\omega \rightarrow 0)|$ i systemet blir dårligere, og systemets evne til å motvirke (tilnærmet) konstante forstyrrelser blir dårligere.

Figur 9.111

Sprangrespons for eksempel 7.4 sammenliknet med optimerte verdier for K og T_d med $N^{max} = 1.5$.



Figur 9.111 sammenlikner sprangresponsen for eksempel 7.4 med det samme systemet med optimerte verdier for K og T_d , basert på $N^{max} = 1.5$. Den økte båndbredden for den optimerte regulatorstillingen gir tydelig mindre stigetid, kjappere respons og lavere oversvingning, men på bekostning av at sprangresponsen blir noe mer oscillatorisk, som svarer til en litt høyere resonanstopp i $|N(j\omega)|$.

Konklusjon:

Eksemplet viser at det er mulig å finne optimale verdier for regulatorparametere ved å løse et numerisk optimaliseringsproblem, men at disse parameterne fortsatt vil avhenge av hvilke (tildels motstridende) krav den som designer regulatoren setter til ytelsen av systemet.

9.9 ROBUST REGULERING

Under behandlingen av systemspesifikasjoner (se avsnitt 9.2) innførte vi noen enkle kriterier for “stabilitetsmargin”:

- Fasemargin ψ
- Forsterkningsmargin ΔK
- $|N(j\omega)|_{\max}$

Disse tre størrelsene, som alle kan avledes av reguleringsløyvens frekvensrespons $h_0(j\omega)$, gir effektive mål for det regulerede systemets robusthet.

Når vi analyserer et reguleringsystems stabilitetsforhold og ytelse benytter vi en *modell* av systemets “*prosess*” $h_u(s)$ og “*regulatoren*” $h_r(s)$. Til sammen utgjør disse ved seriekompensasjon sløyfetransferfunksjonen $h_0(s) = h_u(s)h_r(s)$. Prosessmodellen $h_u(s)$ uttrykker vår *nominelle* kunnskap om de mange systemelementene og vil alltid være gjenstand for *usikkerhet*. Typiske kilder til slik usikkerhet er:

- Fysiske fenomener som er utelatt i modelleringen (umodellerte fenomener)
- Ulineariteter forenklet ved linearisering
- Endringer i prosessparametrene på grunn av endringer i omgivelser (temperatur, hastighet) og arbeidspunkt

Usikkerhet i prosessmodellen kan representeres på ulike måter. Vi skal se på **multiplikativ perturbasjon** som er illustrert i figur 9.112. Sløyfetransferfunksjonen kan uttrykkes som

$$\tilde{h}_0(s) = h_0(s)(1 + l(s)) \quad (9.216)$$

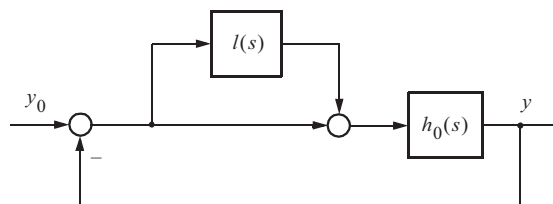
der

$h_0(s)$: **nominell sløyfetransferfunksjon** ($h_0(s) = h_u(s)h_r(s)$)

$l(s)$: **ustrukturert usikkerhet**

$\tilde{h}_0(s)$: resulterende sløyfetransferfunksjon

Figur 9.112
Multiplikativ
usikkerhet



Usikkerheten beskrives gjerne i frekvensplanet slik at

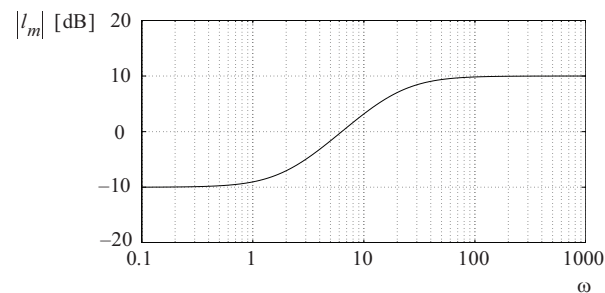
$$|l(j\omega)| < |l_m(\omega)| \quad (9.217)$$

hvor en typisk $|l_m(\omega)|$ er vist i figur 9.113. Dette betyr at *enhver* $l(j\omega)$ som tilfredsstiller (9.217) er tillatt. Merk at det er vanlig å skrive $l_m(\omega)$ i stedet for $l_m(j\omega)$, da l_m er en reell funksjon.

Robust regulering behandler en regulators evne til å regulere et system med usikkerhet til forskjell for nominell regulering som behandler en regulators evne til å regulere uten usikkerhet. **Robust stabilitet** vil si at et system er stabilt for alle tillatte $h_0(s)$. **Robust ytelse** vil si at et system har en ønsket ytelse for alle tillatte $h_0(s)$.

Figur 9.113 viser et typisk valg for $l_m(\omega)$. Den sier at usikkerheten er liten ved lave frekvenser, $\omega < 2$. Videre kan vi lese fra figuren at usikkerheten er helt dominerende for $\omega > 6$. En slik usikkerhet kan begrunnes som følger. Ved lave frekvenser er den nominelle modellen god da de fleste viktige fenomener inkluderes. Ved høye frekvenser er det typisk stor usikkerhet i den nominelle modellen på grunn av at høyfrekvente umodellerte fenomener er utelatt.

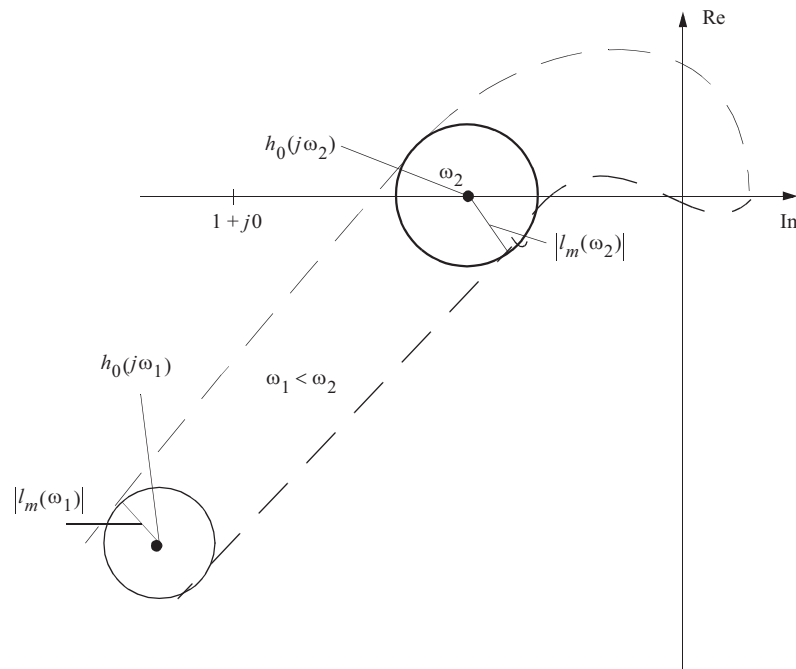
Figur 9.113
Multiplikativ
usikkerhet



Figur 9.114 viser et polart diagram (Nyquist-diagram) med et eksempel på en stabil frekvensrespons $h_0(j\omega)$ (det vil si $\frac{h_0(j\omega)}{1+h_0(j\omega)}$ er stabil). Videre er mulige perturbasjoner som følge av $l_m(\omega)$, angitt. Dette gir “usikkerhetssirkler” ved hver frekvens ω . Legg merke til at radiusen på sirkelen øker med økende frekvens.

Betingelsen for robust stabilitet av systemet i figur 9.114 er at alle tillatte $\tilde{h}_0(j\omega)$ skal ha det samme antallet omsirklinger av det kritiske punkt $(1+j0)$ som det nominelt stabile systemet.

Figur 9.114
 Frekvensrespons
 forbundet med
 usikkerhet framstilt
 i Nyquistdiagram



Systemet i figur 9.112 har et følgeforhold

$$\tilde{M}(s) = \frac{(1 + l(s))h_0(s)}{1 + (1 + l(s))h_0(s)} \quad (9.218)$$

Dette kan omskrives til

$$\tilde{M}(s) = \frac{l(s)M(s)}{1 + l(s)M(s)} \cdot \left(1 + \frac{1}{l(s)}\right) \quad (9.219)$$

der

$$M(s) = \frac{h_0(s)}{1 + h_0(s)} \quad (9.220)$$

er det nominelle systems følgeforhold. Vi ser av (9.219) at vi kan tolke omskrivingen som en ny ekvivalent tilbakekoplingsløyfe som har sløyfetransferfunksjon $l(s)M(s)$. For at denne sløyfen, som svarer til den opprinnelige, skal være stabil er det *tilstrekkelig* at¹

$$|l(j\omega)M(j\omega)| < 1 \quad (9.221)$$

¹Vi benytter her det såkalte ‘Small-gain’ teoremet, se for eksempel Skogestad og Postlethwaite (1996). Det forutsettes at $l(s)M(s)$ er åpen sløyfe stabil.

Dette er oppfylt dersom

$$|l_m(\omega)| \cdot |M(j\omega)| < 1 \quad (9.222)$$

eller

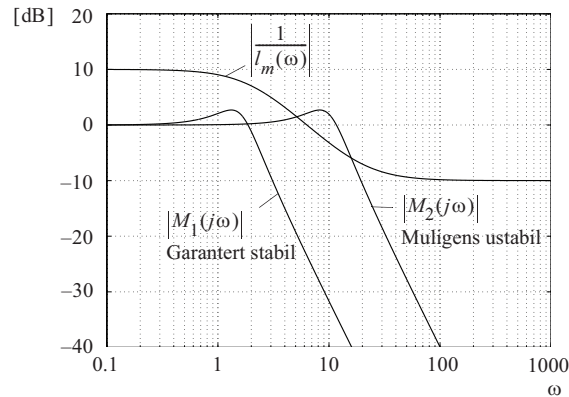
$$|M(j\omega)| < \frac{1}{|l_m(\omega)|} \quad (9.223)$$

Dette betyr at vi har robust stabilitet dersom det nominelle systemets følgeførlighet $|M(j\omega)|$ er mindre enn den inverse av den maksimale ustrukturerte usikkerheten $|l_m(\omega)|$.

Figur 9.115 viser et eksempel på resultatet i (9.223).

Figur 9.115

$|1/l_m(\omega)|$,
 $|\tilde{M}_1(j\omega)|$ og
 $|\tilde{M}_2(j\omega)|$ i
AFF-diagram



Vi tenker oss at M_1 er basert på en regulator h_{r1} og M_2 på en annen regulator h_{r2} hvor både h_{r1} og h_{r2} stabiliserer det nominelle system $h_0(j\omega)$. Usikkerheten er definert av $l_m(\omega)$ i figur 9.113.

$M_1(j\omega)$ er robust stabil med en gitt ustrukturert usikkerhet $l_m(\omega)$, mens et annet system $M_2(j\omega)$ med den samme usikkerheten ikke er robust stabil. Dette systemet kan altså bli ustabil.

Siden den ustrukturerte usikkerheten $l_m(\omega)$ normalt vil være økende med økende frekvens som illustrert i figur 9.113, finner vi at det kritiske frekvensområdet med hensyn til *stabilitet* vil være fra $M(j\omega)$'s resonansfrekvens og oppover. Når det gjelder systemets *ytelse*, har vi tidligere gitt spesifikasjoner til $|N(j\omega)|$ ved lave frekvenser og oppover til rundt $|N(j\omega)|_{\max}$. Men siden

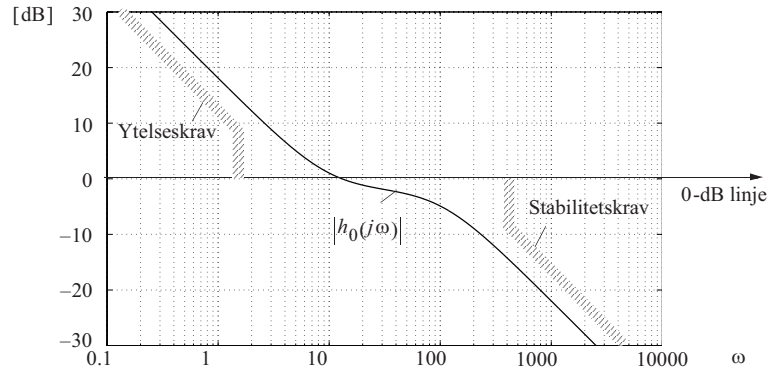
$$|N(j\omega)| \approx \left| \frac{1}{h_0(j\omega)} \right|$$

når $|h_0(j\omega)| \gg 1$, og

$$|M(j\omega)| \approx |h_0(j\omega)|$$

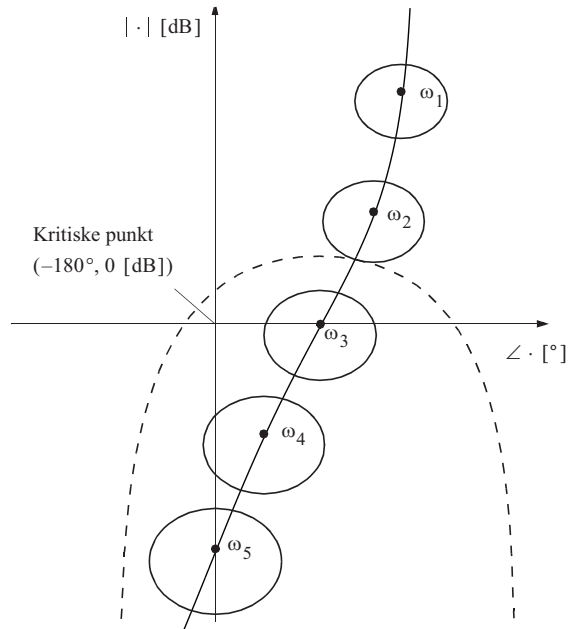
når $|h_0(j\omega)| \ll 1$, finner vi å kunne oppsummere hvilke krav som må stilles til $|h_0(j\omega)|$ ut fra både *ytelse* og *stabilitet* som vist i figur 9.116.

Figur 9.116
Krav som stilles til $|h_0(j\omega)|$ mhp. *ytelse* og *stabilitet*



En tolkning av Nyquistdiagrammet vist i figur 9.114 kan gis ved hjelp av en tilsvarende framstilling i et Nicholsdiagram som vist i figur 9.117.

Figur 9.117
Nicholsdiagram



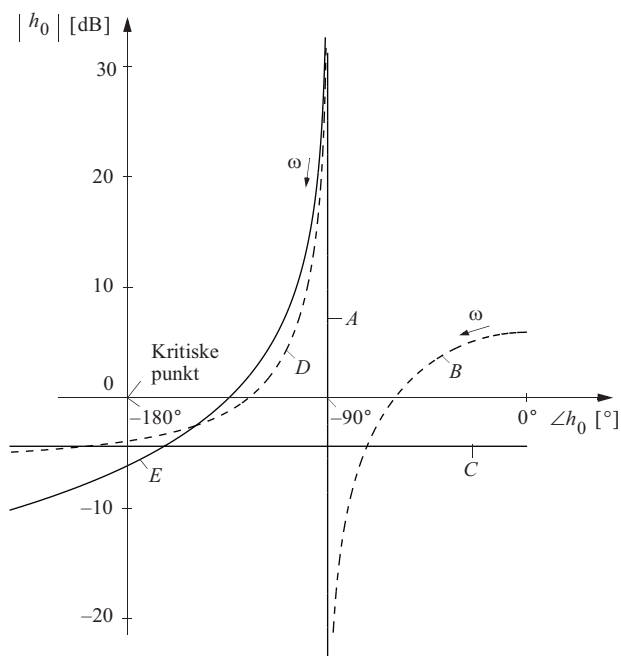
Det som i figur 9.114 er usikkerhetssirkler, vil i figur 9.117 være “deformerte sirkler” av samme form som finnes i Nicholsdiagrammets sentrum. Usikkerheten $l_m(\omega)$ er en relativ størrelse siden den multipliseres med $h_0(j\omega)$. Det er derfor riktig å angi $l_m(\omega)$ i [dB]. En usikkerhet på $|l_m| = -6$ [dB] svarer til en mulig variasjon i $|\tilde{h}_0(j\omega)|$ på $\pm 50\%$. En usikkerhet på $|l_m| = +20$ [dB] gir tilsvarende en mulig variasjon på $\pm 1000\%$.

I eksemplet i figur 9.117 er det angitt “usikkerhetssirkler” for hver frekvens svarende til det som er illustrert i figur 9.114. Av dette ser vi direkte betydningen av $|N(j\omega)|_{\max}$ -kriteriet som tidligere er innført. Men den nye tolkningen i figur 9.117 gir noe mer enn $|N(j\omega)|_{\max}$ -kriteriet. Nå kan vi nemlig se hva konsekvensen med hensyn til stabilitet vil være ved for eksempel en økning av usikkerheten $|l_m(\omega)|$ ved ω_3 til 0 [dB] (det vil si 100% usikkerhet) som vist stiplet i figur 9.117. Den stiplede konturen svarer til en $|N| = 0$ dB-kontur i et Nicholsdiagram. Vi ser at systemet nå kan bli ustabil fordi den stiplede konturen omslutter det kritiske punktet (0 [dB], -180°).

Det er nyttig å vite hvilke fenomener som gjør en tilbakekoplet sløyfe “robust” med hensyn på stabilitet. Som vi ser av usikkerhetssirklene i figur 9.117, har usikkerheten $l(j\omega)$ et bidrag både i amplitude $|l(j\omega)|$ og i fase $\angle l(j\omega)$. Hvilken virkning som er mest alvorlig, avhenger av det aktuelle systemet.

Figur 9.118 viser en del eksempler på stedkurver for $h_0(j\omega)$ i et amplitude/fase-diagram.

Figur 9.118
Eksempler på stedkurver for $h_0(j\omega)$ i AF-diagram



Stedkurven A som representerer $h_0(s) = 1/(Ts)$ (ren integrator), er robust med hensyn på endringer i amplitude. Den har en *amplitudemargin* på $\pm\infty$ [dB]. *Fasemarginen* er 90° .

Stedkurven C representerer en ren transportforsinkelse $h_0(s) = Ke^{-\tau s}$. Vi ser at den er helt ufølsom for usikkerhet i fasen $\angle l(j\omega)$, men svært følsom for amplitudeusikkerhet $|l(j\omega)|$. P-regulering av en ren transportforsinkelse er ikke særlig vellykket (som vist i kurve C). PI-regulering som vist i kurve D, er heller ikke særlig vellykket, men ren I-virkning (integral) som vist i kurve E, er nær det beste som kan oppnås når prosessen er en ren transportforsinkelse eller en annen ikke-minimum-fase transfer-funksjon.

De resultatene som er utledet foran, basert på en *ustrukturert usikkerhet* $l_m(s)$ gir gjerne *konservative* (pessimistiske) kriterier. Robust regulering er utførlig behandlet i Skogestad og Postlethwaite (1996).

9.10 SLUTTVERDIREGULERING

I dette avsnittet skal vi kort berøre et problem som er noe forskjellig fra de som er behandlet i foregående avsnitt. Problemet kan kort formuleres slik:

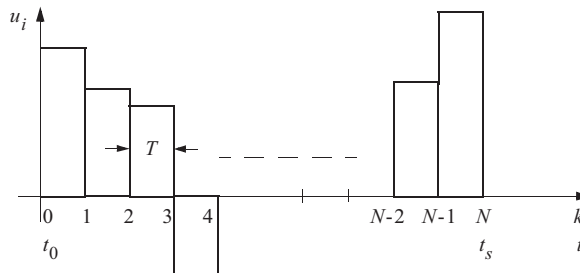
Hvordan skal en endre pådraget $\mathbf{u}(t)$ for at en prosess, karakterisert ved $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$, skal bringes fra tilstanden $\mathbf{x}(t_0)$ ved begynnelsestidspunktet t_0 til en gitt slutttilstand $\mathbf{x}(t_s)$ ved et gitt sluttidspunkt t_s ?

Det motsatte problemet, å bestemme slutttilstanden til systemet når pådragsvektoren er kjent, har vi behandlet tidligere. Vi vet at

$$\mathbf{x}(t_s) = e^{\mathbf{A}(t_s-t_0)}\mathbf{x}(t_0) + \int_{t_0}^{t_s} e^{\mathbf{A}(t_s-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau \quad (9.224)$$

En generell behandling av dette problemet hører ikke hjemme her, men vi skal se hvordan problemet kan gis en enkel løsning dersom vi tillater oss å anta at pådragsvektoren har en slik form at den er stykkevis konstant i tidsintervallet fra t_0 til t_s slik som vist for en av pådragsvariablene i figur 9.119

Figur 9.119
Stykkevis konstant pådragsvariabel



Vi har tidligere i avsnitt 3.3.2 utledet et uttrykk for hvordan tilstandsvektoren utvikler seg i et slikt diskret tilfelle. Kaller vi tidsintervallet i figur 9.119 for T og innfører $t - t_0 = kT$ og $t_s - t_0 = NT$, kan vi med bakgrunn i (3.38)(3.38) skrive

$$\mathbf{x}((k+1)T) = \mathbf{F}(T)\mathbf{x}(kT) + \mathbf{D}(T)\mathbf{u}(kT) \quad (9.225)$$

der $\mathbf{F}(T) = e^{\mathbf{A}T}$ = kjent og $\mathbf{D}(T) = \mathbf{A}^{-1}(e^{\mathbf{A}T} - \mathbf{I})\mathbf{B}$ = kjent. Vi kan nå benytte (9.225) en rekke ganger for å beskrive utviklingen av tilstandsvektoren og finner

$$\begin{aligned} \mathbf{x}(T) &= \mathbf{F}\mathbf{x}(0) + \mathbf{D}\mathbf{u}(0) \\ \mathbf{x}(2T) &= \mathbf{F}\mathbf{x}(T) + \mathbf{D}\mathbf{u}(T) = \mathbf{F}^2\mathbf{x}(0) + \mathbf{F}\mathbf{D}\mathbf{u}(0) + \mathbf{D}\mathbf{u}(T) \\ &\vdots \\ \mathbf{x}(NT) &= \mathbf{F}^N\mathbf{x}(0) + \mathbf{F}^{N-1}\mathbf{D}\mathbf{u}(0) + \dots + \mathbf{D}\mathbf{u}((N-1)T) \end{aligned} \quad (9.226)$$

Den siste av disse likningene uttrykker slutttilstanden $\mathbf{x}(NT)$ som funksjon av begynnelsestilstanden $\mathbf{x}(0)$ og pådragsvektoren i alle de små tidsintervallene.

Tenker vi oss alle de ukjente verdier av pådragsvektoren samlet i en stor vektor

$$\mathbf{U}_N = \begin{bmatrix} \mathbf{u}(0) \\ \mathbf{u}(T) \\ \vdots \\ \mathbf{u}((N-1)T) \end{bmatrix} \quad (9.227)$$

og innfører vi matrisen

$$\mathbf{C}_N = \begin{bmatrix} \mathbf{F}^{N-1}\mathbf{D} & \mathbf{F}^{N-2}\mathbf{D} & \dots & \mathbf{D} \end{bmatrix} = \text{kjent} \quad (9.228)$$

kan vi skrive siste linje i (9.226) som

$$\mathbf{x}(NT) = \mathbf{F}^N\mathbf{x}(0) + \mathbf{C}_N\mathbf{U}_N \quad (9.229)$$

Første ledd på høyre side i (9.229) utgjør bidraget til slutttilstanden av begynnelsestilstanden $\mathbf{x}(0)$, mens det andre leddet er bidraget fra pådragsvektoren. Vi legger merke til at matrisen \mathbf{C}_N er kjent og avhengig av hvor mange små intervaller N vi har inndelt tidsrommet i. Siden $\mathbf{x}(0)$ er kjent, kan vi like gjerne skrive (9.229) på formen

$$\delta\mathbf{x}(NT) = \mathbf{x}(NT) - \mathbf{F}^N\mathbf{x}(0) = \mathbf{C}_N\mathbf{U}_N \quad (9.230)$$

Likning (9.230) skal løses med hensyn på \mathbf{U}_N . Dette er mulig så sant

$$\text{rang } \mathbf{C}_N = n \quad (9.231)$$

der n er dimensjonen av tilstandsvektoren. Kravet i (9.231) overensstemmer med kravet til *styrbarhet* av systemet og er helt analogt til det som er presentert i (5.4) i kapittel 5. Det er selvsagt logisk at systemet må være styrbart for at vi skal kunne bestemme pådragsvektoren i de forskjellige tidsintervallene. Løsningen av (9.230) er ikke entydig. Det finnes en rekke \mathbf{U}_N som vil tilfredsstillе likningen. En løsningsmetode er å bruke en optimaliseringsprosedyre, som gir minimum av størrelsen $\mathbf{U}_N^T \mathbf{U}_N$. Dette er et mål for energien i pådraget. Løsningen vil ha formen

$$\mathbf{U}_N = \mathbf{C}_N^T (\mathbf{C}_N \mathbf{C}_N^T)^{-1} \delta \mathbf{x}(NT) \quad (9.232)$$

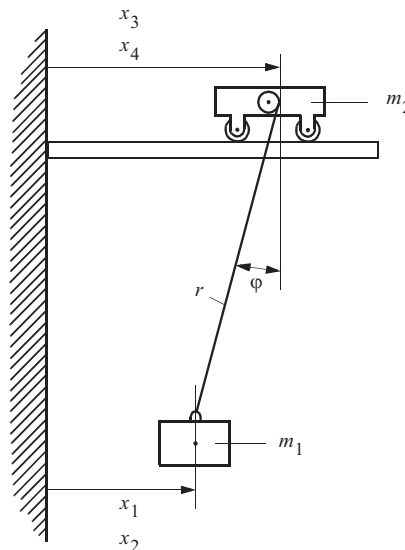
På grunn av (9.231) vil $|\mathbf{C}_N \mathbf{C}_N^T| = \det(\mathbf{C}_N \mathbf{C}_N^T) \neq 0$ og derfor eksisterer $(\mathbf{C}_N \mathbf{C}_N^T)^{-1}$. Vi kan overbevise oss om at (9.232) er riktig ved å premultiplisere på begge sider med \mathbf{C}_N . Vi får da direkte uttrykket i (9.230). Uttrykket $\mathbf{C}_N^T (\mathbf{C}_N \mathbf{C}_N^T)^{-1} = \mathbf{C}_N^+$ kalles ofte den **pseudoinverse** av \mathbf{C}_N .

Et eksempel vil illustrere bruken av resultatet i likning (9.232).

EKSEMPEL 9.27: Sluttverdiregulering av motorisert løpekatt

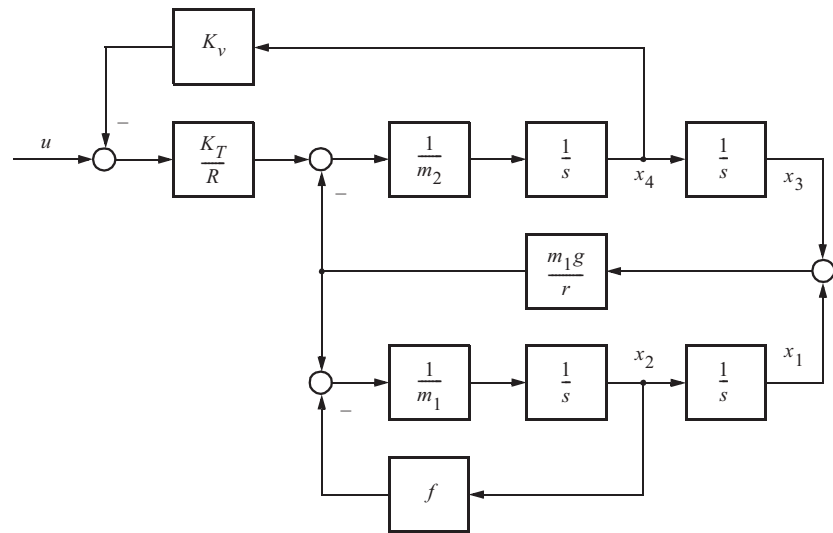
Figur 9.120 viser en forenklet skisse av en løpekatt med motordrift som skal forflytte en last som er festet til løpekatten via en wire.

Figur 9.120
Løpekatt med last



Vi kaller posisjonen og hastigheten til lasten i forhold til en fast vegg henholdsvis x_1 og x_2 . Løpekattens posisjon og hastighet i forhold til den samme veggen kalles henholdsvis x_3 og x_4 . Massene av last og løpekatt betegnes henholdsvis m_1 og m_2 , og r angir lengden av wiren. Figur 9.121 viser et elementært blokkdiagram som beskriver balansene i dette systemet når vinkelutslaget til wiren er lite, slik at $\tan \varphi \approx \sin \varphi$. Øverst til venstre i blokkdiagrammet finner vi spenningen u som tilføres løpekattens elektriske motor. Motoren er en likestrømsmotor med konstant felt der ankerets elektriske tidskonstant antas svært liten. Ankerstrømmen framkommer som netto ankerets spenning dividert med ankerkretsens motstand R . Ankerstrømmen multipliseres med momentkonstanten K_T og gir akselerasjonskraften som virker på løpekattens masse. Motorens induserte motspenning er gitt av løpekattens hastighet x_4 multiplisert med K_v . Mot den akselererende kraften fra motoren virker reaksjonskraften fra wiren.

Figur 9.121
Elementært blokk-
diagram for
systemet



Kraften i wiren virker på den andre side akselererende på massen til lasten som vist nederst i diagrammet. Det antas å være en viss luftdempning proporsjonal med hastigheten x_2 og en konstant f .

Kraften som virker på lasten og løpekatten er lik

$$m_1 g \tan \varphi \approx m_1 g \sin \varphi = m_1 g \frac{x_1 - x_3}{r} \quad (9.233)$$

som vist i diagrammet. Vi setter nå opp differensiallikningene for systemet ut fra figur 9.121.

$$\begin{aligned}
 \dot{x}_1 &= x_2 \\
 \dot{x}_2 &= -\frac{g}{r}x_1 - \frac{f}{m_1}x_2 + \frac{g}{r}x_3 \\
 \dot{x}_3 &= x_4 \\
 \dot{x}_4 &= \frac{m_1g}{m_2r}x_1 - \frac{m_1g}{m_2r}x_3 - \frac{K_vK_T}{Rm_2}x_4 + \frac{K_T}{Rm_2}u
 \end{aligned} \tag{9.234}$$

Vi setter inn tallverdier for et realistisk system med en elektrisk motor av størrelse ca. 5 kW, masse i størrelsesorden noen hundre kilo og wirelengde ca. 5 m og får en systemmodell av formen

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \tag{9.235}$$

der

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -2 & -0.0033 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0.857 & 0 & -0.857 & -6.04 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.093 \end{bmatrix}$$

En undersøkelse av egenverdiene til matrisen \mathbf{A} vil vise at dette systemet er sterkt oscillatorisk. Lasten vil overlatt til seg selv svinge med en periode ca. lik 4.4 [s]. Vi benytter oss nå av den metoden som er beskrevet i foregående avsnitt. Vi antar da at både vogn og last befinner seg i ro inne ved veggene til å begynne med. Dette gir

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

Så bestemmer vi at systemet skal bringes til sluttstanden $\mathbf{x}(NT) = \begin{bmatrix} 5 & 0 & 5 & 0 \end{bmatrix}^T$ ved hjelp av periodevis konstante pådrag. Dette vil si at vi ønsker å bringe både last

og vogn til ro 5 [m] ut fra veggen ved hjelp av N skritt i løpet av tiden NT . Problemet er undersøkt for $NT = 4$ [s] og med $N = 4$, $N = 8$ og $N = 16$.

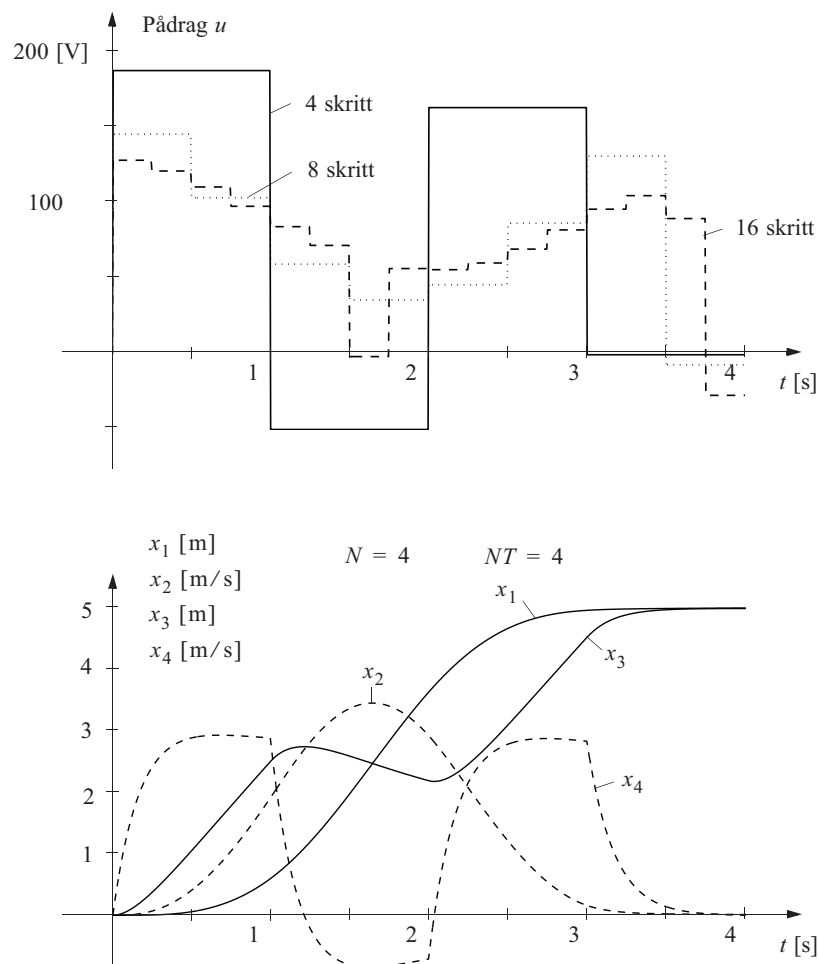
Resultatene er vist i figur 9.122. Øverst i figuren ser vi hvordan pådraget (det vil si spenningen) som tilføres motorens ankerkrets må endres i de tre tilfellene. Det er her karakteristisk at ved å velge flere skritt (16, 8 og 4), får vi mindre utslag i pådraget uttrykt ved en lavere verdi av den forbrukte pådragsenergien som er gitt av

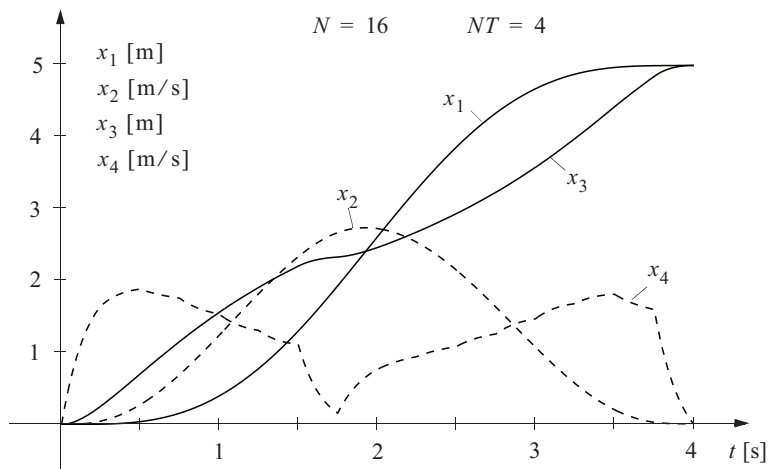
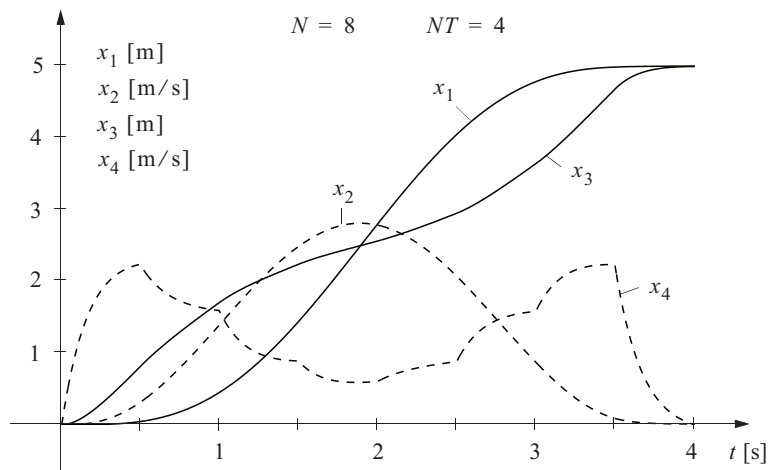
$$\mathbf{U}_N^T \mathbf{U}_N = \sum_{i=0}^{N-1} u^3(iT) \quad (9.236)$$

De tre nederste kurvesettene i figur 9.122 viser forløpet av de 4 tilstandsvariablene ved henholdsvis 4, 8 og 16 skritt. Vi ser at for $N = 4$ vil vognen en kort periode gå bakover. Det vil den ikke gjøre når $N = 8$ og 16.

Ut i fra det siste kurvesettet ($N = 16$) får vi en ide om hvordan $u(t)$ og $x_i(t)$ ville sett ut i det kontinuerlige tilfellet ($N \rightarrow \infty$).

Figur 9.122
Kurver som viser pådrag og tilstander som funksjon av tiden for forskjellige N





Multivariable systemer

10.1 Innledning og motivering 467

10.2 Stabilitet og respons 471

10.3 Seriekompensasjon av en multivariabel reguleringsløyfe 481

10.4 Foroverkopling i multivariable systemer 487

10.5 Optimalregulering og modellbasert regulering 488

10.6 Singulærverdianalyse av lineære multivariable dynamiske systemer 491

10.7 Funksjonell styrbarhet (FS) av multivariable systemer 497

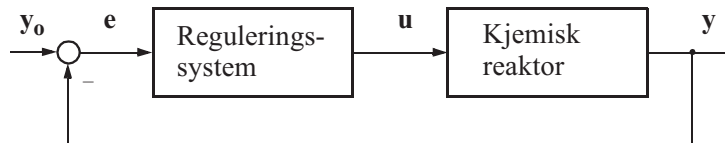
10.1 INNLEDNING OG MOTIVERING

Vi har tidligere i boka støtt på multivariable systemer. Et multivariabelt system er et system hvor antallet pådrag og/eller antallet prosessutganger er større enn én. Den kjemiske reaktoren i eksempel 9.16 og turbinene i eksempel 9.17 er eksempler på multivariable systemer. Dette kapitlet skal forstås som et oversiktskapittel da temaet multivariable systemer er meget omfattende. Vi skal i et innledende eksempel se nærmere på reaktoreksemplet.

EKSEMPEL 10.1: Multivariabel regulering av en kjemisk reaktor

Reguleringsystemet som er vist i figur 9.74, kan skjematisk beskrives som i figur 10.1, dersom følgende er oppfylt:

Figur 10.1
En skjematisk beskrivelse av den regulerede reaktoren i figur 9.74



- Dimensjonen til prosessutgangen er 7 og pådragsdimensjonen er 4.
- y_1, y_2, y_3, y_4 benyttes til å beregne u_1 (høytrykksdamp).
- y_5 benyttes til å beregne u_2 .
- y_6 benyttes til å beregne u_3 (ventil som styrer tilførsel av stoff B).
- y_7 benyttes til å beregne u_4 (ventil som styrer tilførsel av stoff A).

Det finnes selvsagt mange andre måter å kople prosessutgangene til pådragene på. Den valgte løsningen representerer en mulig reguleringsstruktur. Valg av reguleringsstruktur er en svært viktig del i enhver syntese av et multivariabelt reguleringsystem.

Vanligvis velges en enkel reguleringsstruktur, forutsatt at spesifikasjonene for det regulerede systemet kan tilfredstilles. Den enkleste strukturen får vi når reaktoren kan betraktes som en samling av uavhengige (ukoplede) monovariabel systemer. Da vil vi konstruere regulatorsystemet ved å bruke 4 monovariabel regulatorer. Reguleringsystemet i figur 9.74 består av 3 monovariabel regulatorer. Den 4. sløyfen blir monovariabel dersom den benytter *en* av prosessutgangene y_1, y_2, y_3, y_4 til å beregne u_1 . Referert til strukturen i figur 10.2 vil regulatoren $\mathbf{H}_r(s)$, som i dette tilfelle er en 4×7 -matrise (en **prosesstransfermatrise**) ha fire elementer forskjellige fra null.

I noen tilfeller er det viktig for ytelsen av et reguleringsystem at alle prosessutgangene kan påvirke alle pådragene. For reaktoreksemplet betyr dette at y_1, \dots, y_9 inngår i beregningen av u_1 , og de tre andre pådragene. Dette gir en meget kompleks reguleringsstruktur, og betyr at matrisen $\mathbf{H}_r(s)$ i figur 10.2 vil være full (den har ingen null-elementer).

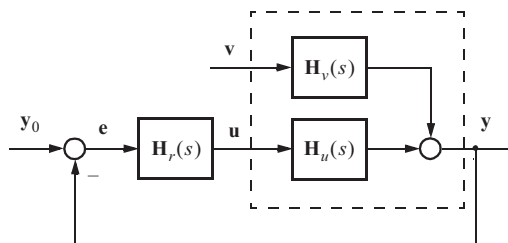
■

Eksemplet ovenfor indikerer at analyse og syntese av multivariable reguleringsystemer er vesentlig mer komplisert enn for monovariabel systemer. Det er derfor en stor fordel om en, i alle fall tilnærmet, kan betrakte et multivariabelt system som en samling av uavhengige (ukoplede) monovariabel systemer. Dette gjøres ofte ved konstruksjon av industrielle reguleringsystemer. Når dette ikke er mulig, må vi ha metoder som tar hensyn til krysskoplingene i prosessen. Det finnes et omfattende teoretisk grunnlag for dimensjonering av generelle multivariable systemer. Vi skal imidlertid her begrense oss til relativt elementære betraktninger.

Vi vil studere tilbakekopling i multivariable systemer på tilsvarende måte som i avsnitt 7.2. Analogt med figur 7.5 foreslår vi et tilbakekoplet system av en form som er vist i figur 10.2.

Figur 10.2

Et multivariabelt tilbakekoplet system med vektoriell forstyrrelse



Forskjellen mellom figur 7.5 og 10.2 er at størrelsene \mathbf{y} , \mathbf{y}_0 , \mathbf{e} og \mathbf{u} som forekommer i figur 10.2 er vektorer, mens de tilsvarende størrelsene i figur 7.5 var skalare. Størrelsene $\mathbf{H}_r(s)$, $\mathbf{H}_u(s)$ og $\mathbf{H}_v(s)$ er transfermatriser der elementene er transferfunksjoner. Vektorene \mathbf{y} , \mathbf{y}_0 og \mathbf{e} har samme dimensjon. Vi skal i dette avsnittet forutsette at også \mathbf{u} har den samme dimensjon, noe som ikke var tilfelle i eksempel 10.1. Dette betyr at $\mathbf{H}_r(s)$ og $\mathbf{H}_u(s)$ blir kvadratiske matriser. Vi setter ingen betingelser på dimensjonen til \mathbf{v} .

$\mathbf{H}_u(s)$ og $\mathbf{H}_v(s)$ kan beregnes fra en tilstandsrombeskrivelse.

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{v} \\ \mathbf{y} &= \mathbf{C}\mathbf{x}\end{aligned}\tag{10.1}$$

Laplacetransformasjon gir

$$\begin{aligned}\mathbf{y}(s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{u}(s) + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{v}(s) \\ &= \mathbf{H}_u(s)\mathbf{u}(s) + \mathbf{H}_v(s)\mathbf{v}(s)\end{aligned}\tag{10.2}$$

Innsatt regulatoren

$$\mathbf{u}(s) = \mathbf{H}_r(s)(\mathbf{y}_0(s) - \mathbf{y}(s))\tag{10.3}$$

i (10.2) gir

$$\begin{aligned}\mathbf{y}(s) &= (\mathbf{I} + \mathbf{H}_u(s)\mathbf{H}_r(s))^{-1}\mathbf{H}_u(s)\mathbf{H}_r(s)\mathbf{y}_0(s) \\ &\quad + (\mathbf{I} + \mathbf{H}_u(s)\mathbf{H}_r(s))^{-1}\mathbf{H}_v(s)\mathbf{v}(s)\end{aligned}\tag{10.4}$$

I det monovariabelt tilbakekoblede systemet behandlet i avsnitt 7.2 innførte vi sløyfe-transferfunksjonen $h_0(s)$ som produktet av alle transferfunksjoner rundt tilbakekoplingsløyfen. I dette produktet vil faktorenes orden være likegyldig. I et multivariabelt system kan vi innføre en transfermatrise for å karakterisere tilbakekoplingene. Denne vil være produktet av transfermatrisene til alle elementene i sløyfen. Men i dette tilfellet er ikke faktorenes orden likegyldig. Vi vil derfor få forskjellige **sløyfe-transfermatriser** avhengig av hvilke punkter i sløyfen vi betrakter.

Vi innfører betegnelsen $\mathbf{H}_0^y(s)$ = produktet av alle transfermatriser i tilbakekoplingsløyfen regnet fra \mathbf{y} mot signalretningen og får

$$\mathbf{H}_0^y(s) = \mathbf{H}_u(s)\mathbf{H}_r(s)\tag{10.5}$$

Videre innfører vi analogt til det som er gjort for det skalare tilfellet i avsnitt 7.2

$$\mathbf{N}^y(s) = (\mathbf{I} + \mathbf{H}_0^y(s))^{-1}\tag{10.6}$$

og da vil (10.4) anta formen

$$\begin{aligned}\mathbf{y}(s) &= \mathbf{N}^y(s)\mathbf{H}_u(s)\mathbf{H}_r(s)\mathbf{y}_0(s) + \mathbf{N}^y(s)\mathbf{H}_v(s)\mathbf{v}(s) \\ &= \mathbf{M}(s)\mathbf{y}_0(s) + \mathbf{N}^y(s)\mathbf{H}_v(s)\end{aligned}\quad (10.7)$$

som ledd for ledd tilsvarende uttrykket i (7.6).

Vi definerer størrelsen $\mathbf{N}^y(s)$ som det multivariable systemets **reguleringsgradmatrise** referert til vektoren \mathbf{y} . $\mathbf{M}(s)$ er systemets **følgeforholdsmatrise**.

La oss så studere oppførselen til vektoren \mathbf{u} istedenfor \mathbf{y} , i det tilbakekoplede systemet i figur 10.2.

Vi benytter et resultat fra avsnitt 7.2.1 som også gjelder for multivariable systemer: Transferfunksjonen mellom en vilkårlig påvirkning og en vilkårlig respons i et tilbakekoplede system er lik transferfunksjonen uten tilbakekopling (mellom påvirkningen og responsen) (for)multiplisert med størrelsen $\mathbf{N}^p(s)$ (p angir responspunktet).

$$\mathbf{u}(s) = \mathbf{N}^u(s)\mathbf{H}_r(s)\mathbf{y}_0(s) - \mathbf{N}^u(s)\mathbf{H}_r(s)\mathbf{H}_v(s)\mathbf{v}(s)\quad (10.8)$$

der vi tilsvarende definisjonene i (10.5) og (10.6) har

$$\mathbf{N}^u(s) = (\mathbf{I} + \mathbf{H}_0^u(s))^{-1}\quad (10.9)$$

og

$$\mathbf{H}_0^u(s) = \mathbf{H}_r(s)\mathbf{H}_u(s)\quad (10.10)$$

Vi merker oss her at rekkefølgen av faktorene i (10.5) og (10.10) er forskjellig.

Med litt matrisemanipulasjon kan det vises at

$$\mathbf{N}^u(s) = \mathbf{H}_u^{-1}(s)\mathbf{N}^y(s)\mathbf{H}_u(s)\quad (10.11)$$

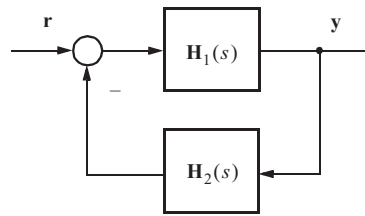
og

$$\mathbf{N}^y(s) = \mathbf{H}_u(s)\mathbf{N}^u(s)\mathbf{H}_u^{-1}(s)\quad (10.12)$$

Likning (10.11) kan være nyttig dersom vi har konstruert systemet slik at $\mathbf{N}^y(s)$ har en spesielt ønsket form og vi ønsker å finne hva $\mathbf{N}^u(s)$ blir.

Tilsvarende det som er beskrevet i avsnitt 7.2.4 kan det for et multivariabelt tilbakekoplede system være av interesse å finne det tilbakekoplede systemets dynamiske egenskaper når tilbakekoplingen er dynamisk, som vist i figur 10.3.

Figur 10.3
Multivariabelt system med dynamisk tilbakekopling



Vi finner at

$$\mathbf{y}(s) = (\mathbf{I} + \mathbf{H}_1(s)\mathbf{H}_2(s))^{-1}\mathbf{H}_1(s)\mathbf{r}(s) \quad (10.13)$$

Antar vi at \mathbf{H}_2^{-1} eksisterer (forutsetter at $\dim \mathbf{r} = \dim \mathbf{y}$), vil vi kunne skrive (10.13) slik

$$\begin{aligned} \mathbf{y}(s) &= [(\mathbf{I} + \mathbf{H}_1(s)\mathbf{H}_2(s))^{-1}\mathbf{H}_1(s)\mathbf{H}_2(s)]\mathbf{H}_2^{-1}(s)\mathbf{r}(s) \\ &= \mathbf{M}(s)\mathbf{H}_2^{-1}(s)\mathbf{r}(s) \end{aligned} \quad (10.14)$$

Dette uttrykket tilsvarer ledd for ledd det vi har funnet i (7.22). Vi ser at dersom

$$\mathbf{M}(s) \approx \mathbf{1}$$

vil vi oppnå at

$$\mathbf{y}(s) \approx \mathbf{H}_2^{-1}(s)\mathbf{r}(s) \quad (10.15)$$

hvilket tilsvarer det vi har funnet i (7.23).

10.2 STABILITET OG RESPONS

I de foregående avsnittene har vi stort sett bare behandlet stabilitet av monovariabile tilbakekoblede systemer. Det er imidlertid åpenbart at en med utgangspunkt i en tilstandsrom-beskrivelse også kan bruke de algebraiske kriteriene til bestemmelse av plasseringen av *multivariable systems egenverdier*, og dermed få stabilitetskriterier for slike systemer.

De metoder som tar utgangspunkt i systemkomponentenes transferfunksjoner er foreløpig begrenset til monovariabile systemer. I dette avsnittet skal vi gå videre og betrakte multivariable systemer med transfermatriser.

Vi benytter oss da av den notasjonen som er introdusert i avsnitt 10.1 og går ut fra en konfigurasjon som vist i figur 10.2.

Bestemmelse av de inverse matrisene i (10.6) og (10.9) leder til

$$\mathbf{N}^y(s) = \frac{\text{adj}(\mathbf{I} + \mathbf{H}_0^y(s))}{\det(\mathbf{I} + \mathbf{H}_0^y(s))} \quad (10.16)$$

og

$$\mathbf{N}^u(s) = \frac{\text{adj}(\mathbf{I} + \mathbf{H}_0^u(s))}{\det(\mathbf{I} + \mathbf{H}_0^u(s))} \quad (10.17)$$

Stabilitet av det tilbakekoblede systemet i figur 10.2 vil nå, tilsvarende det vi har funnet for monovariabel systemer, være gitt av formen på matrisen $\mathbf{N}^y(s)$ eller $\mathbf{N}^u(s)$. Systemet vil være stabilt dersom det polynomet i s som forekommer i nevneren i $\mathbf{N}^y(s)$ eller $\mathbf{N}^u(s)$, har sine røtter i venstre halvplan. Det kan vises at

$$\det(\mathbf{I} + \mathbf{H}_0^y(s)) = \det(\mathbf{I} + \mathbf{H}_0^u(s)) = \frac{P_2(s)}{P_0(s)} \quad (10.18)$$

der $P_2(s)$ og $P_0(s)$ er polynomer i s . $P_2(s)$ er det polynomet som skal ha røtter i venstre halvplan. $P_2(s)$ kan gjerne kalles *det lukkede multivariable systemets karakteristiske polynom*. Tilsvarende er $P_0(s)$ *det åpne multivariable systemets karakteristiske polynom*.

Disse begrepene forklares enklest ved å gå tilbake til monovariabel systemer og da spesielt (8.29) og (8.30) og forklaringen i forbindelse med (8.44).

Når polynomet $P_2(s)$ er bestemt med koeffisienter som avhenger av systemets parametre, kan vi benytte Rouths kriterium eller Hurwitz's kriterium til å finne betingelser for stabilitet av hele systemet.

Bruken av denne metoden er i prinsippet svært enkel, men vil for selv relativt enkle systemer ofte lede til forholdsvis store algebraiske uttrykk der vi lett kan miste oversikten. Som oftest vil det i systemer beskrevet med ordinære differensiallikninger (rasjonale transferfunksjoner) være lettere å bringe systemet på tilstandsform og så bestemme det totale multivariable systemets karakteristiske polynom ut fra uttrykket

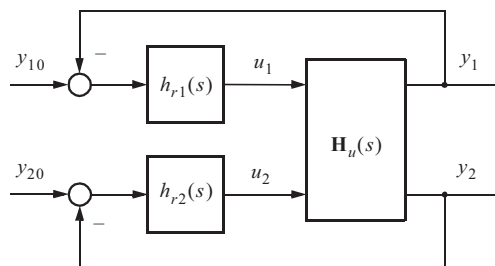
$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \quad (10.19)$$

10.2.1 Stabilitet og respons av et 2×2 multivariabelt system med diagonalregulator

Et særlig enkelt spesialtilfelle som det kan være verd å se på, oppstår dersom prosessen har to pådrag, to målinger og to separate regulatorer som vist i figur 10.4.

Figur 10.4

Blokkdiagram over prosess med to pådrag, to målinger og to separate regulatorer



Prosessmatrisen vil ha formen

$$\mathbf{H}_u(s) = \begin{bmatrix} h_{11}(s) & h_{12}(s) \\ h_{21}(s) & h_{22}(s) \end{bmatrix} \quad (10.20)$$

og regulatormatrisen

$$\mathbf{H}_r(s) = \begin{bmatrix} h_{r1}(s) & 0 \\ 0 & h_{r2}(s) \end{bmatrix} = \begin{bmatrix} h_1(s) & 0 \\ 0 & h_2(s) \end{bmatrix} \quad (10.21)$$

Sammenhengen mellom inngangene og utgangene i figur 10.4 kan vi skrive slik

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \bar{M}_{11} & \bar{M}_{12} \\ \bar{M}_{21} & \bar{M}_{22} \end{bmatrix} \begin{bmatrix} y_{10} \\ y_{20} \end{bmatrix} = \mathbf{M}\mathbf{y}_0 \quad (10.22)$$

$$\bar{\mathbf{M}}(s) = (\mathbf{I} + \mathbf{H}_0^y(s))^{-1} \mathbf{H}_0^y(s) \quad (10.23)$$

og

$$\mathbf{H}_0^y(s) = \mathbf{H}_u(s)\mathbf{H}_r(s) \quad (10.24)$$

Ved endel manipulasjon av likningene kan det vises at

$$\bar{M}_{11} = M_1 \frac{1 - YM_2}{1 - YM_1M_2} = M_1K_1 \quad (10.25)$$

$$\bar{M}_{22} = M_2 \frac{1 - YM_1}{1 - YM_1M_2} = M_2K_2 \quad (10.26)$$

$$\bar{M}_{12} = \frac{h_{12}}{h_{22}} \frac{N_1M_2}{1 - YM_1M_2} \quad (10.27)$$

$$\bar{M}_{21} = \frac{h_{21}}{h_{11}} \frac{N_2M_1}{1 - YM_1M_2} \quad (10.28)$$

der

$$Y = \frac{h_{12}h_{21}}{h_{11}h_{22}} \quad (10.29)$$

$$M_i = \frac{h_{ii}h_i}{1 + h_{ii}h_i} \quad (10.30)$$

og

$$N_i = \frac{1}{1 + h_{ii}h_i} \quad (10.31)$$

Størrelsen Y i (10.29) er et mål for graden av **krysskopling** i prosessen, mens M_i , $i = 1, 2$, er følgeforholdet vi ville fått i den øvre og nedre tilbakekoplingsløyfen i figur 10.4 dersom det ikke hadde vært noen krysskopling i prosessen ($Y = 0$).

Inspeksjon av (10.25) viser at følgeforholdet M_{11} for systemet *med* krysskopling i prosessen kan avledes av følgeforholdet M_1 *uten* krysskopling multiplisert med en korreksjonsfaktor K_1 .

Når vi skriver

$$\bar{M}_{11}(s) = M_1(s)K_1(s) \quad (10.32)$$

og

$$\bar{M}_{22}(s) = M_2(s)K_2(s) \quad (10.33)$$

ser vi at betingelsen for stabilitet av det totale systemet blir at ingen av funksjonene $M_1(s)$, $M_2(s)$, $K_1(s)$ og $K_2(s)$ kan ha poler i høyre halvplan.

Dette betyr at de systemer som ville oppstå hvis man ser bort fra krysskoplingen må være stabile, og dessuten må faktoren $(1 - Y(s)M_1(s)M_2(s))$ ikke ha nullpunkter i høyre halvplan. Vi ser at denne faktoren går igjen i alle uttrykkene (10.25)-(10.28).

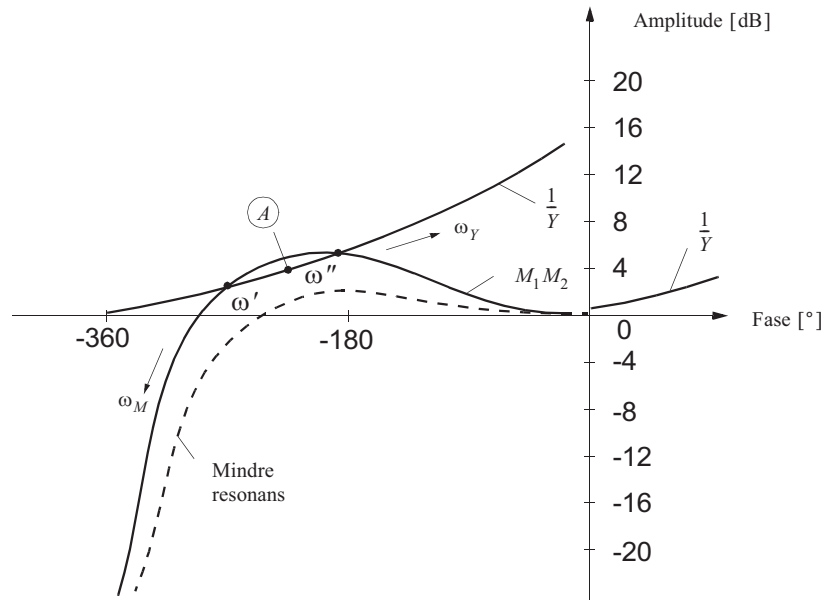
Siden faktoren $Y(s)$ i (10.29) er uavhengig av regulator-transferfunksjonene, kan vi med fordel uttrykke den nye stabilitetsbetingelsen ved at $(-1/Y(s) + M_1(s)M_2(s))$ ikke skal ha nullpunkter i høyre halvplan.

Denne stabilitetsbetingelsen kan vi undersøke ved hjelp av Nyquists stabilitetskriterium ved å sette $s = j\omega$. Størrelsen $1/Y(j\omega)$ blir da et bevegelig (frekvensavhengig) kritisk punkt svarende til punktet $(-1 + j0)$ i det ordinære Nyquist-kriteriet.

En hensiktsmessig tolkning av dette får en ved å bruke et amplitude/fase-diagram som vist i figur 10.5 der det er inntegnet stedkurver for $1/Y(j\omega)$ og produktet $M_1(j\omega)M_2(j\omega)$. Dersom det kritiske punktet hadde vært fast ($1/Y = \text{konstant}$) som antydnet ved punkt A i figuren, ville betingelsen for stabilitet vært at stedkurven for $M_1(j\omega)M_2(j\omega)$ ikke skulle omslutte dette punktet. Men når det kritiske punktet beveger seg som funksjon av ω , blir den korrekte tolkningen av stabilitetsbetingelsen at punktet ikke må “fanges” av stedkurven for M_1M_2 når ω varierer fra 0 til ∞ .

Systemet antydnet i figur 10.5 vil altså være ustabil dersom frekvensskalaen langs $1/Y(j\omega)$ er slik at det kritiske punktet blir “fanget”. For å oppnå stabilitet, må derfor stedkurven for $M_1(j\omega)M_2(j\omega)$ trykkes nedover som antydnet stiplet i figur 10.5. Det vil i dette tilfellet si at de to isolerte reguleringsløyfen, som oppstår når henholdsvis den nedre sløyfen og deretter den øvre sløyfen settes ut av funksjon, må justeres til større stabilitetsmargin (mindre resonans).

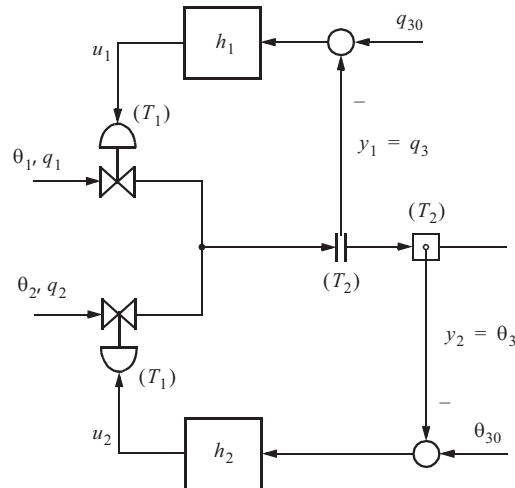
Figur 10.5
 $1/Y(j\omega)$ og
 $M_1(j\omega)M_2(j\omega)$
 i amplitude/fase-
 diagram



EKSEMPEL 10.2: Blanding av vannstrømmer

Et velkjent eksempel på et 2×2 multivariabelt system er det som er gitt figur 10.6. Figuren illustrerer et reguleringsystem for blanding av to vannstrømmer, en varm og en kald. Hensikten er å holde konstant utgangsstrøm q_3 og utgangstemperatur θ_3 .

Figur 10.6
System for regu-
ling av væskestrøm
og vanntemperatur



Når vi formulerer en enkel matematisk modell for dette systemet, antar vi at de to reguleringsventilene er lineære, har konstant trykkfall og drives av motorer med identiske tidskonstanter (T_1). Videre antas at målelementene for væskestrøm og temperatur er lineære og har like store tidskonstanter (T_2). Vi antar nå at systemet er i et stasjonært arbeidspunkt og at variablene gir små endringer omkring dette arbeidspunktet. Selve blandeprosessen antar vi er momentan, mens de dynamiske virkningene er konsentrert i ventilmotorene (T_1) og i målelementene (T_2). Den stasjonære massebalansen for systemet er relativt grei å sette opp. Vi finner da for væskestrømmen

$$q_3 = q_1 + q_2 \quad (10.34)$$

$$\frac{\partial q_3}{\partial q_1} = \frac{\partial q_3}{\partial q_2} = 1 \quad (10.35)$$

Energibalansen påvirkes kun av inngangs og utgangsstrømmene.

$$\theta_1 q_1 + \theta_2 q_2 = \theta_3 (q_1 + q_2) = \theta_3 q_3 \quad (10.36)$$

$$\theta_3 = \theta_1 \frac{q_1}{q_1 + q_2} + \theta_2 \frac{q_2}{q_1 + q_2} \quad (10.37)$$

Av (10.37) finner vi

$$\begin{aligned}\frac{\partial \theta_3}{\partial q_1} &= \frac{\theta_1}{(q_1 + q_2)^2} ((q_1 + q_2) - q_1) + \frac{\theta_2}{(q_1 + q_2)^2} (-q_2) \\ &= \frac{(\theta_1 - \theta_2)q_2}{(q_1 + q_2)^2} = \frac{(\theta_1 - \theta_2)q_2}{q_3^2}\end{aligned}\quad (10.38)$$

$$\frac{\partial \theta_3}{\partial q_2} = -\frac{(\theta_1 - \theta_2)q_1}{q_3^2}\quad (10.39)$$

Legg merke til at de partiellderiverte må beregnes fordi at de gir oss forsterkningen i arbeidspunktet. Med den notasjonen som er innført foran for systemets transferfunksjoner, finner vi da

$$h_{11}(s) = \frac{\partial q_3}{\partial q_1} \frac{1}{(1 + T_1 s)(1 + T_2 s)}\quad (10.40)$$

$$h_{21}(s) = \frac{\partial \theta_3}{\partial q_1} \frac{1}{(1 + T_1 s)(1 + T_2 s)}\quad (10.41)$$

$$h_{12}(s) = \frac{\partial q_3}{\partial q_2} \frac{1}{(1 + T_1 s)(1 + T_2 s)}\quad (10.42)$$

$$h_{22}(s) = \frac{\partial \theta_3}{\partial q_2} \frac{1}{(1 + T_1 s)(1 + T_2 s)}\quad (10.43)$$

hvor vi husker at $\frac{1}{1 + T_1 s}$ skriver seg fra dynamikken i ventilen og $\frac{1}{1 + T_2 s}$ skyldes dynamikken i måleinstrumentet. Som eksempel er $h_{21}(s) = \frac{y_2(s)}{u_1(s)}$ gitt av dynamikken i den øverste ventilen, dynamikken i temperaturmåleelementet og forsterkningen (10.38).

Som eksempel velger vi nå referanseverdiene (settpunktene) for reguleringssystemene

$$q_{30} = 20 \text{ [l/min]} \text{ og } \theta_{30} = 50 \text{ [}^\circ\text{C]}$$

og antar forsyningsvannets temperaturer

$$\theta_1 = 90 \text{ [}^\circ\text{C]} \text{ og } \theta_2 = 10 \text{ [}^\circ\text{C]}$$

Vi får da de nominelle verdiene av væskestrømmene

$$\bar{q}_1 = 10 \text{ [l/min]} \text{ og } \bar{q}_2 = 10 \text{ [l/min]}$$

Disse verdiene innsatt i (10.38) og (10.39) gir

$$\frac{\partial \theta_3}{\partial q_1} = \frac{80 \cdot 10}{400} = 2 \text{ og } \frac{\partial \theta_3}{\partial q_2} = -\frac{80 \cdot 10}{400} = -2$$

Uttrykkene i (10.38) og (10.39) kan omskrives ved bruk av uttrykket i (10.36) og vi finner

$$\frac{\partial \theta_3}{\partial q_1} = \frac{\theta_1 - \theta_{30}}{q_{30}} \quad (10.44)$$

og

$$\frac{\partial \theta_3}{\partial q_2} = -\frac{\theta_{30} - \theta_2}{q_{30}} \quad (10.45)$$

idet vi har satt inn referanseverdiene for utgangstemperatur og utgangsstrøm.

Vi kan nå innsette (10.35), (10.44) og (10.45) i (10.40)-(10.41) og ser at temperaturreguleringssløyfen vil få varierende forsterkning avhengig av settpunktene for de to sløyfene og temperaturene på inngangsvannet, mens strømningsreguleringssløyfen vil ha konstant forsterkning.

Ved innsetting av de valgte eksempelverdiene finner vi graden av krysskopling.

$$Y(s) = \frac{h_{12}(s)h_{21}(s)}{h_{11}(s)h_{22}(s)} = \frac{\theta_1 - \theta_{30}}{\theta_2 - \theta_{30}} = \frac{90 - 50}{10 - 50} = -1 \quad (10.46)$$

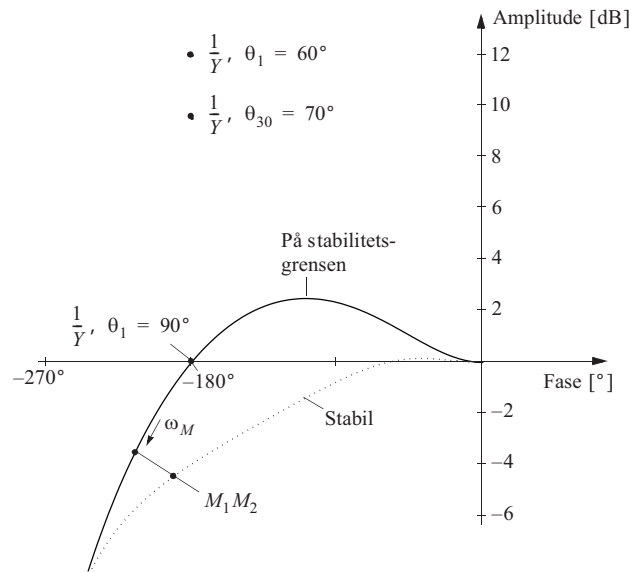
Vi antar nå at $T_1 > T_2$ og at de to reguleringssløyfene har PI-regulatorer som er innstilt med integraltid $T_i = T_1$ og forsterkning K_p som er justert slik at den relative dempningen i hver sløyfe blir $\zeta = 0.5$. Vi vil da få

$$M_1(s) = M_2(s) = \frac{1}{1 + T_2 s + T_2^2 s^2} \quad (10.47)$$

(10.46) og (10.47) gir som resultat stedkurvene vist i figur 10.7.

Vi ser av denne figuren at systemet *er akkurat på grensen til ustabilitet* fordi stedkurven $M_1(j\omega)M_2(j\omega)$ går gjennom det kritiske punkt $1/Y(j\omega) = -1$. Dette skyldes at de to reguleringssløyfene er innstilt til en relativ dempning ($\zeta = 0.5$) som er for lav. Figur 10.8 a) viser sprangresponsen til dette systemet. Vi kunne ha oppnådd stabilitet hvis regulatorenes parametre var justert slik at hver sløyfe hadde mer enn kritisk dempning ($\zeta > 1$). Dette oppnås ved å redusere forsterkningen i hver sløyfe med en faktor 4 eller mer. Båndbredden i reguleringssløyfen vil derved bli redusert til halvparten.

Figur 10.7
 $1/Y(j\omega)$ og
 $M_1(j\omega)M_2(j\omega)$
 i amplitude/fase-
 diagram



En annen årsak til instabiliteten er at de to sløyfene har identisk frekvensrespons og derved sine resonanstopper ved samme frekvens. Hvis en sløyfe hadde vært hurtigere enn den andre, ville ikke de to resonanstopperne blitt addert ved det kritiske området rundt $\angle M_1 M_2 = -180^\circ$. En av sløyfene (for eksempel M_2) innstilles til resonans ved en lavere frekvens (ved å senke regulatorforsterkningen med 6 [dB]). Den andre sløyfen gis sin resonans ved en høyere frekvens (ved å øke regulatorforsterkningen med 6 [dB]). Resultatet av dette er vist i den prikkede kurven i figur 10.7. Sprangresponsen er vist i figur 10.8 b).

Det “kritiske punktet” er $1/Y(s) = -1$ i figur 10.7 som utledet i likning (10.46). Men, som vi ser av likning (10.46), er punktet avhengig av temperaturene på forsyningsvannet og referanseverdien for temperaturen i blandingsvannet. Antar vi nå at vi beholder de samme betingelsene som i eksemplet ovenfor, bortsett fra temperaturen θ_1 i varmtvannsforsyningen, vil vi få

$$\frac{1}{Y(s)} = \frac{10 - 50}{\theta_1 - 50} = -\frac{40}{\theta_1 - 50} \quad (10.48)$$

Velger vi eksempelvis $\theta_1 = 60[^\circ\text{C}]$, får vi

$$\frac{1}{Y(s)} = -4 = (12 \text{ [dB]}, -180^\circ)$$

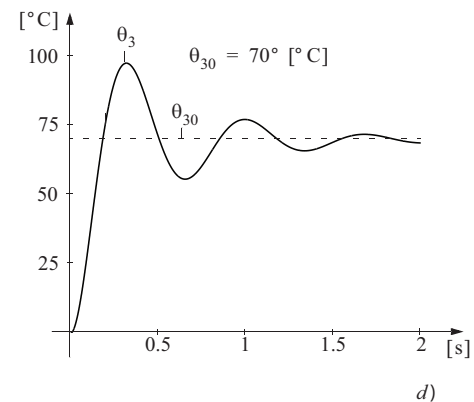
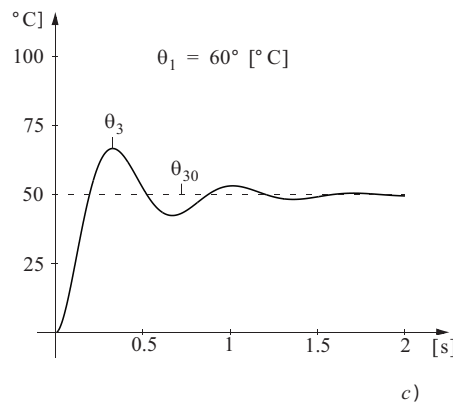
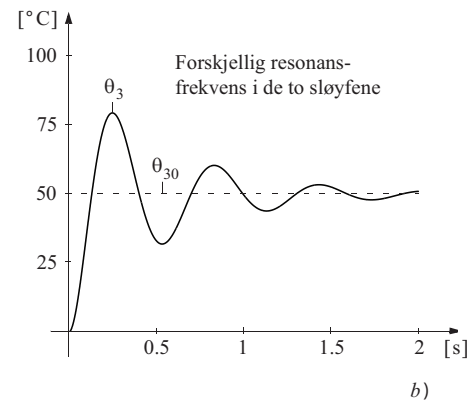
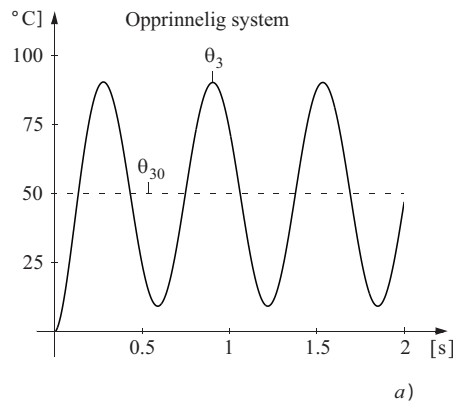
Dette bringer det kritiske punktet i figur 10.7 langt nok over $M_1 M_2$ -kurven til å sikre stabilitet. Rent fysisk betyr senkningen av temperaturen i varmtvannsforsyningen at det vil strømme mer vann gjennom varmtvannsventilen enn kaldtvannsventilen, hvilket medfører en dekopling av de to reguleringsløyfene. Sprangresponsen er vist i figur 10.8 c).

Et tilsvarende resultat ville vi ha fått som vist i figur 10.7 dersom vi med de opprinnelige parametrene hadde endret referanseverdien for blandingstemperaturen θ_{30} til 70 [$^{\circ}\text{C}$]. Figur 10.8 d) viser sprangresponsen for dette tilfellet. Men det er ikke særlig aktuelt å gjøre dette dersom det for eksempel dreier seg om et dusjbatteri.

Figur 10.8

Sprangresponser

- a) $1/Y(s) = -1$
- b) $1/Y(s) = -1$,
men forskjellig resonans-
frekvens i de
to sløyfene
- c) θ_1 endret slik at
 $1/Y(s) = -4$
- d) θ_{30} endret slik
at
 $1/Y(s) = -3$



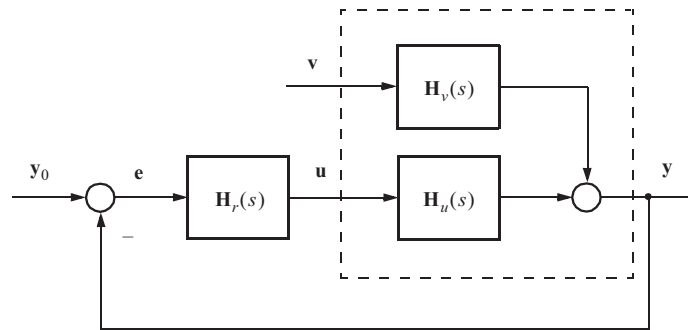
Den konklusjonen som kan trekkes ut fra dette eksemplet, er at betingelsene for stabilitet av et 2×2 -system er ganske sammensatte og at en grafisk løsning som vist i figur 10.7 er til stor hjelp for å forstå de indre mekanismene.

10.3 SERIEKOMPENSASJON AV EN MULTIVARIABEL REGULERINGSSLØYFE

Vi skal i dette avsnittet bare se på en begrenset del av den omfattende problemstillingen som seriekompensasjon av multivariable systemer er. Figur 10.9 viser systemet som skal undersøkes. Det er det samme som vist i figur 10.2.

Figur 10.9

Et multivariabelt tilbakekoplet system med vektoriell forstyrrelse



Proessen antas karakterisert ved transfermatrisene $\mathbf{H}_u(s)$ og $\mathbf{H}_v(s)$. Problemet er å bestemme seriekompensasjonsmatrisen (regulormatrisen) $\mathbf{H}_r(s)$ slik at det resulterende lukkede systemet får visse foreskrevne egenskaper.

Vi baserer oss på det som er utledet i avsnitt 10.1, og antar at vi er i stand til å spesifisere en fornuftig sløyfe-transfermatrise $\mathbf{H}_0^y(s)$ der

$$\mathbf{H}_0^y(s) = \mathbf{H}_u(s)\mathbf{H}_r(s) \quad (10.49)$$

Vi kan da bestemme seriekompensasjonsmatrisen av

$$\mathbf{H}_r(s) = \mathbf{H}_u^{-1}(s)\mathbf{H}_0^y(s) \quad (10.50)$$

Dette uttrykk tilsvare det vi tidligere har omtalt i (9.19).

Ved løsning av (10.50) er det først og fremst nødvendig at $\mathbf{H}_u(s)$ har en invers. Det er dessuten, i likhet med det som er diskutert for monovariabel systemer, nødvendig å ta hensyn til at det resulterende systemet skal være teoretisk og praktisk realiserbart, ved valget av $\mathbf{H}_0^y(s)$. Det må derfor gjøres en del kompromisser. Å foreta en suksessiv utvikling av $\mathbf{H}_r(s)$ på basis av et fornuftig valg av $\mathbf{H}_0^y(s)$ etterfulgt av nødvendige korreksjoner, er en primitiv, men brukbar metode. Problemet er i prinsippet det samme som det vi har ved monovariabel systemer.

Vi kan for eksempel ikke fjerne virkningen av en pol i prosessen uten at vi innfører derivatvirkning. Dette krever at det finnes stor nok effekt til å øke systemets hurtighet.

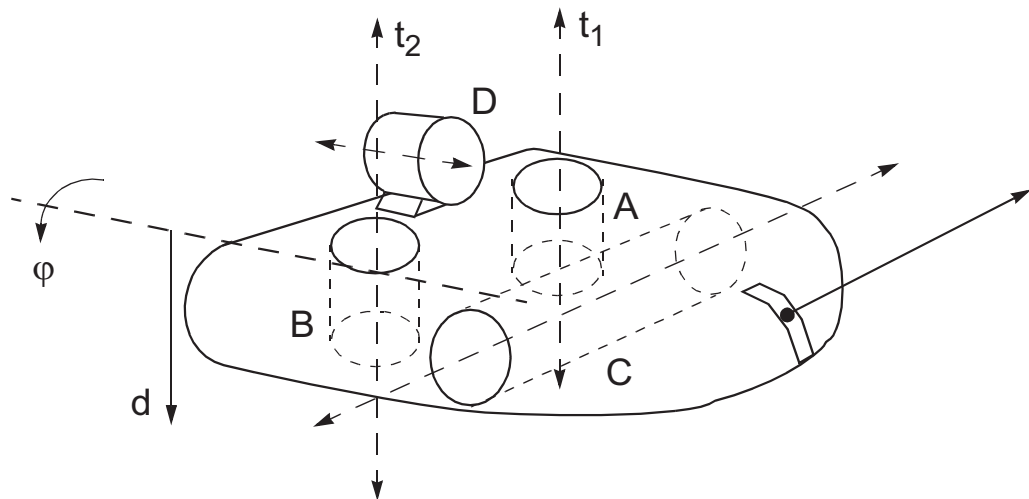
Ofte kan det være ønskelig ved regulering av en multivariabel prosess å få til en “**dekopling**” av de krysspåvirkninger som finnes i prosessen. Krysspåvirkningene består av transferfunksjonene i prosessens transfermatrise som ligger utenfor matrisens diagonal. Vi vil altså tilstrebe en korrigert prosesstransfermatrise som er en diagonalmatrise. Dette vil i noen tilfeller være relativt lett å realisere, om enn tilnærmet, ved hjelp av seriekompensasjon. Den resulterende dekoplete prosesstransfermatrisen kan så benyttes som utgangspunkt for valg av en regulormatrise som kan være en diagonalmatrise av ordinære regulatortransferfunksjoner. Resultatet skulle altså bli et antall uavhengige monovariabelt reguleringsløyper.

Vi skal illustrere det som her er beskrevet med et eksempel.

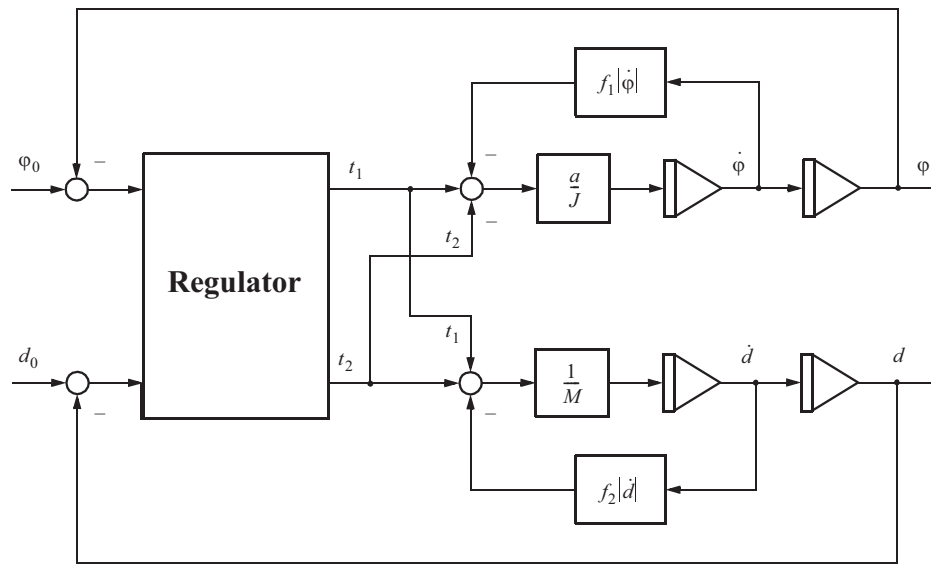
EKSEMPEL 10.3: Undervannsfarkosten DAHABU

Vi skal studere undervannsfarkosten DAHABU som brukes til registrering av data fra havbunnssubstratet (DAtainnsamling fra HAVBUUnnssubstrat), som er vist i figur 10.10. Farkosten slepes med en kabel bak et overflatefartøy og har automatisk regulering av rullvinkelen ϕ og dybden under havoverflaten eller høyden over havbunnen. Farkosten har fire elektrisk drevne thrustere (propeller) hvorav A og B sørger for rullvinkel- og dybdestyring, C sørger for sidestyling og D sørger for bevegelse bakover. Vi skal bare se på systemet for rullvinkel- og dybdestyring.

Figur 10.10
Undervannsfarkosten DAHABU



Figur 10.11
Elementært blokk-
diagram for
farkosten



Figur 10.11 viser et elementært blokkdiagram for farkostens respons på thrustkref-
tene t_1 og t_2 fra vertikalthrusterne, henholdsvis A og B. Avstanden mellom A og B
er $2a$ [m] og treghetsmomentet omkring rotasjonsaksen er J [kgm²]. Rotasjonsvinke-
len betegnes med ϕ [rad] og det hydrodynamiske dempningsmomentet som oppstår
ved rotasjonsbevegelsen, beskrives med $f_1|\dot{\phi}|\phi$ [Nm]. Differensiallikningen for ro-
tasjonsbevegelsen blir derved

$$\ddot{\phi} \frac{a}{J} (-f_1|\dot{\phi}|\phi + (t_1 - t_2)) \quad (10.51)$$

Vi ser at første ledd i denne likningen er kvadratisk i $\dot{\phi}$ og vil derfor være svært lite
når rotasjonshastigheten er liten. Vi vil for enkelhets skyld se bort fra dette leddet for
derved å gjøre likningen lineær og får

$$\ddot{\phi} = \frac{a}{J}(t_1 - t_2) \quad (10.52)$$

Tilsvarende betraktninger gjøres gjeldende for vertikalbevegelsen. Også mot denne
bevegelsen vil det være en hydrodynamisk dempningskraft lik $f_2|\dot{d}|d$ [N] der d [m]
er dybden (vertikalposisjonen) og \dot{d} [m/s] er vertikalhastigheten. Betegner vi massen
av farkosten med M [kg], får vi et forenklet uttrykk for vertikalakselerasjonen

$$\ddot{d} = \frac{1}{M}(t_1 + t_2) \quad (10.53)$$

Vi definerer systemets pådragsvektor \mathbf{u} og målevektor \mathbf{y} som

$$\mathbf{u} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \varphi \\ d \end{bmatrix} \quad (10.54)$$

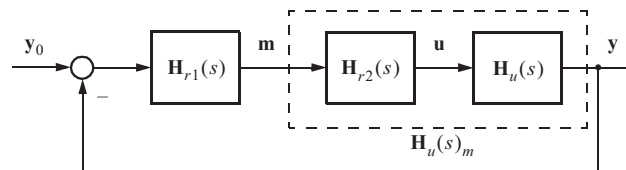
og finner farkostens transfermatrise av

$$\mathbf{y}(s) = \mathbf{H}_u(s)\mathbf{u}(s) \quad (10.55)$$

der

$$\mathbf{H}_u(s) = \begin{bmatrix} \frac{a}{Js^2} & -\frac{a}{Js^2} \\ \frac{1}{Ms^2} & \frac{1}{Ms^2} \end{bmatrix} \quad (10.56)$$

Figur 10.12
Diagonalisering
i multivariabelt
system



Av (10.56) ser vi at denne “prosessen” har meget sterk krysskopling. Vi velger derfor å realisere regulatortransfermatrisen i to trinn som vist i figur 10.12, der

$$\mathbf{H}_r(s) = \mathbf{H}_{r2}(s)\mathbf{H}_{r1}(s) \quad (10.57)$$

$\mathbf{H}_{r2}(s)$ har til formål å diagonalisere prosessen, og vi ønsker at \mathbf{H}_{r1} skal være en diagonalmatrise. Vi betegner den modifiserte (diagonaliserte) prosesstransfermatrisen $\mathbf{H}_u(s)_m$ og finner

$$\mathbf{H}_u(s)_m = \mathbf{H}_u(s)\mathbf{H}_{r2}(s) \quad (10.58)$$

som gir

$$\mathbf{H}_{r2}(s) = \mathbf{H}_u^{-1}(s)\mathbf{H}_u(s)_m \quad (10.59)$$

Vi kan nå spesifisere

$$\mathbf{H}_u(s)_m = \begin{bmatrix} \frac{a}{Js^2} & 0 \\ 0 & \frac{1}{Ms^2} \end{bmatrix} \quad (10.60)$$

og får av (10.59)

$$\mathbf{H}_{r2}(s) = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (10.61)$$

Med denne “krysskopplingsalgoritmen” er prosessen blitt diagonalisert og transferfunksjonene på diagonalen i (10.60) er “dobbelintegratorer”. Fra eksempel 9.5 vet vi at det riktige valget av regulator da er en proporsjonal pluss derivatregulator (PD-regulator), eller mer realistisk en proporsjonal pluss begrenset derivatregulator

$$\mathbf{H}_{r1}(s) = \begin{bmatrix} K_{p1} \frac{1 + T_{d1}s}{1 + \alpha_1 T_{d1}s} & 0 \\ 0 & K_{p2} \frac{1 + T_{d2}s}{1 + \alpha_2 T_{d2}s} \end{bmatrix} \quad (10.62)$$

Det vil ikke være urimelig å velge $\alpha_1 = \alpha_2 = 0.1$ og kryssfrekvensene i de to resulterende dekoplete reguleringsløyfer lik $\omega_{c1} = \omega_{c2} = \omega_c$ og da finner vi

$$K_{p1} = \frac{\omega_c^2 J \sqrt{0.1}}{a} \quad (10.63)$$

$$K_{p2} = \omega_c^2 M \sqrt{0.1} \quad (10.64)$$

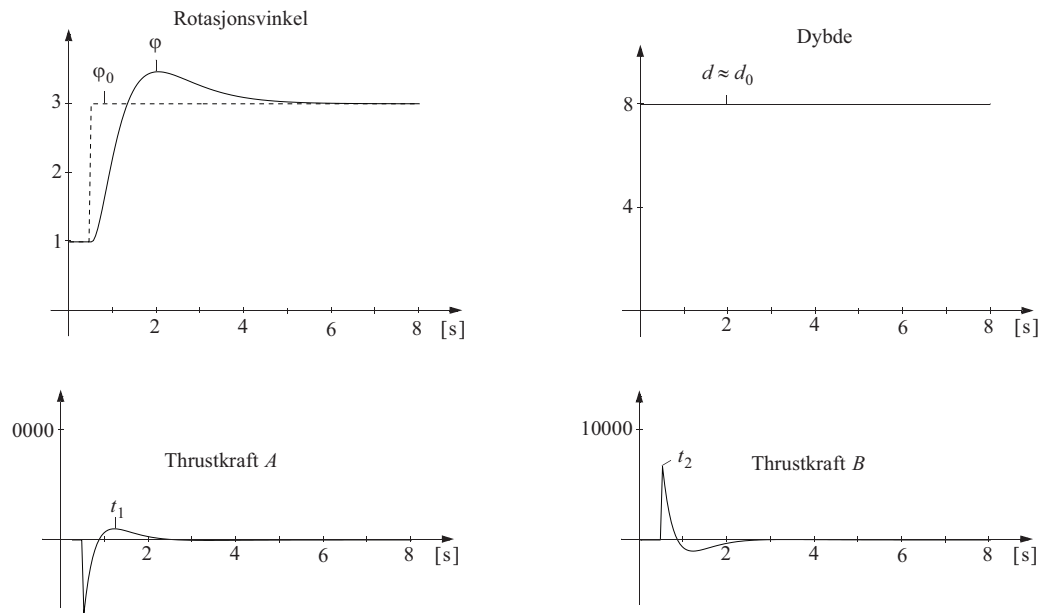
$$T_{d1} = T_{d2} = \frac{1}{\omega_c} \sqrt{10} \quad (10.65)$$

Det faktiske valg av ω_c vil være særlig influert av thrusternes kapasitet (motoreffekten). Dette problemet behandles i kapittel 12.

Figur 10.13 viser hvordan responsen blir i $\varphi(t)$ og $d(t)$ når referansen $\varphi_0(t)$ er et sprang, mens $d_0(t)$ holdes konstant. Det er dessuten vist hvordan pådragene $t_1(t)$ og $t_2(t)$ utvikler seg. Til sammenlikning viser figur 10.14 hvordan de samme responsene ville bli dersom vi hadde brukt en ren diagonal regulator (bare $\mathbf{H}_{r1}(s)$) uten “krysskopplingsalgoritmen” $\mathbf{H}_{r2}(s)$.

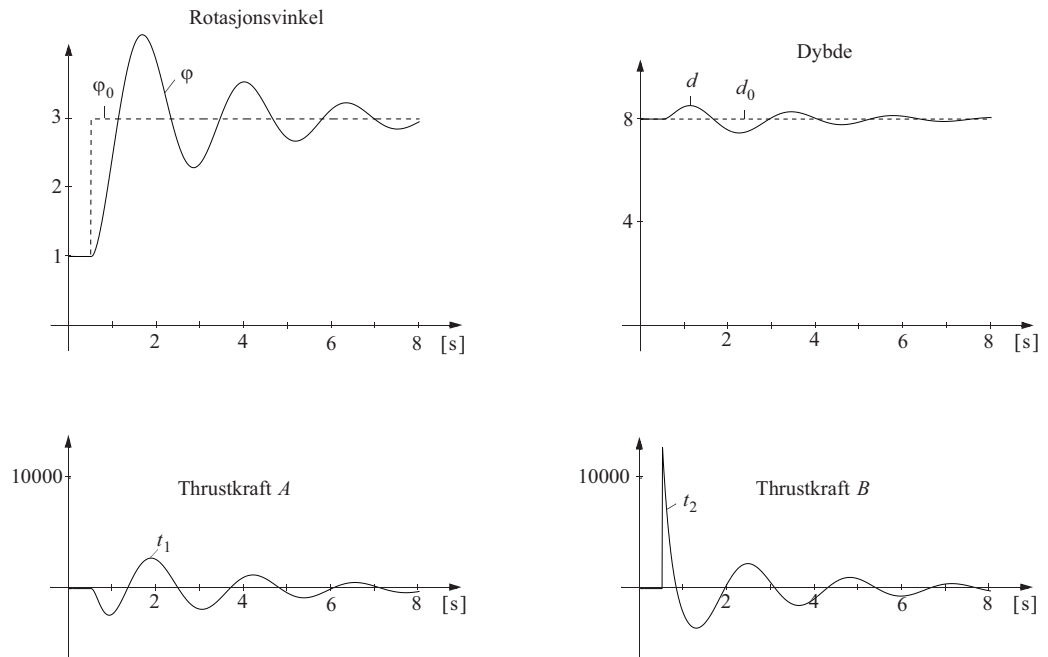
Figur 10.13

Rotasjonsvinkel $\varphi(t)$, dybde $d(t)$ og thrustkreftene $t_1(t)$ og $t_2(t)$



Figur 10.14

Rotasjonsvinkel $\varphi(t)$, dybde $d(t)$ og thrustkreftene $t_1(t)$ og $t_2(t)$ uten bruk av "krysskopplingsalgoritmen" $\mathbf{H}_{r2}(s)$

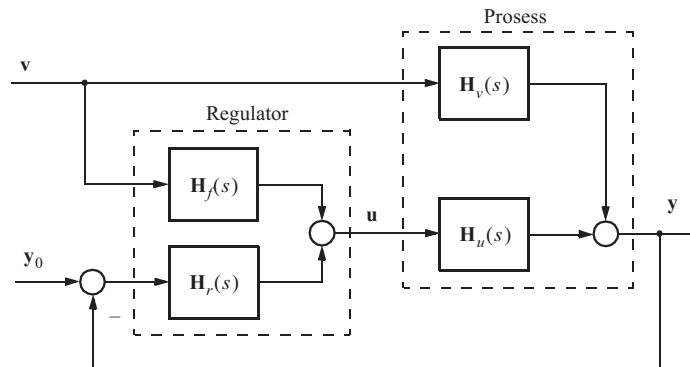


10.4 FOROVERKOPLING I MULTIVARIABLE SYSTEMER

I forbindelse med figur 9.78 introduserte vi foroverkopling for monovariabel systemer. Vi skal nå utvide denne teorien til å gjelde multivariable systemer.

Figur 10.15 viser hvordan foroverkopling kombinert med tilbakekopling kan utformes i et multivariabelt system for prosessregulering.

Figur 10.15
Multivariabelt tilbakekoplet system med foroverkopling fra forstyrrelsen



Det antas at noen eller alle av de mest vesentlige forstyrrelsene på prosessen kan måles. Dette er egentlig litt ulogisk fordi vi tidligere har definert \mathbf{y} som det sett av målinger som utføres på prosessen. Vi skal derfor oppfatte \mathbf{y} som det sett av prosessvariable som danner utgangspunkt for bedømmelse av prosessens driftstilstand, og som sammenliknes med referansevektoren \mathbf{y}_0 .

Hensikten med foroverkoplingen er å generere endringer i pådraget \mathbf{u} som motvirker forstyrrelsens påvirkning av prosessen, slik at resultatet blir ingen eller minst mulig respons i \mathbf{y} . Elimineringen av forstyrrelsers virkning kan bare oppnås for forstyrrelser som er målbare.

Antar vi i første omgang at tilbakekoplingen i figur 10.15 ikke er aktiv, blir responsen til prosessen som følge av forstyrrelsene:

$$\begin{aligned} \mathbf{y}(s) &= \mathbf{H}_u(s)\mathbf{u}(s) + \mathbf{H}_v(s)\mathbf{v}(s) \\ &= (\mathbf{H}_u(s)\mathbf{H}_f(s) + \mathbf{H}_v(s))\mathbf{v}(s) \end{aligned} \quad (10.66)$$

Vi forutsetter at foroverkoplingen utgjøres av transfermatrisen $\mathbf{H}_f(s)$.

Ideell foroverkopling oppnås dersom uttrykket i (10.66) kan gjøres lik null. Vi finner da den ideelle foroverkoplingen

$$\mathbf{H}_{f_i}(s) = -\mathbf{H}_u^{-1}(s)\mathbf{H}_v(s) \quad (10.67)$$

Matrisen $\mathbf{H}_u(s)$ må være inverterbar for at det skal eksistere en løsning av (10.67). En nødvendig betingelse for inverterbarhet er at $\mathbf{H}_u(s)$ er kvadratisk (det vil si at det er like mange pådrag som målinger). Det er derimot ikke satt noen betingelser til antallet forstyrrelser.

10.5 OPTIMALREGULERING OG MODELLBASERT REGULERING

En vesentlig del av den moderne reguleringsteorien (utviklet etter ca. 1960) dreier seg om syntese av multivariable reguleringssystemer. I denne forbindelse spiller tilstandsrombeskrivelse av systemer en viktig rolle hvor teorien for **optimalregulering** er blitt langt utviklet. Vi presenterte noe teori i avsnitt 9.8. Multivariable reguleringssystemers strukturer og parametre blir da bestemt ut fra en minimalisering av en *objektfunksjonal* som er et kvantitativt mål for systemets prestasjon.

En fullstendig behandling av dette emnet hører ikke hjemme i denne boka. Vi skal fra denne teorien bare gjengi strukturen av et av de viktigste resultatene. For en mer fullstendig behandling refereres til Lewis (1986) og Anderson og Moore (1989).

Et av resultatene sier at et lineært dynamisk system beskrevet med

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \quad (10.68)$$

som er brakt ut fra sin likevektstilstand $\mathbf{x} = \mathbf{0}$, bringes tilbake mot denne likevektstilstanden på en "optimal" måte ved å la pådraget \mathbf{u} bli generert ved en lineær transformasjon av tilstandsvektoren

$$\mathbf{u} = \mathbf{Gx} \quad (10.69)$$

I de enkleste tilfeller vil matrisen \mathbf{G} ha konstante elementer.

Dette resultatet fordrer at alle elementene i tilstandsvektoren må foreligge. Dersom målingene er karakterisert ved

$$\mathbf{y} = \mathbf{Cx} \quad (10.70)$$

der matrisen \mathbf{C} er kvadratisk og kan inverteres (det vil si $\det \mathbf{C} \neq 0$), vil vi kunne beregne

$$\mathbf{x} = \mathbf{C}^{-1}\mathbf{y} \quad (10.71)$$

Innsetting i (10.69) gir da

$$\mathbf{u} = \mathbf{GC}^{-1}\mathbf{y} \quad (10.72)$$

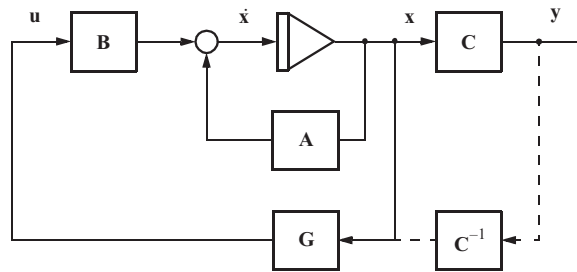
Dette resultatet er vist i figur 10.16 der den siste løsningen er antydnet stiplet.

Setter vi (10.69) inn i (10.68), får vi

$$\dot{\mathbf{x}} = (\mathbf{A} + \mathbf{BG})\mathbf{x} = \tilde{\mathbf{A}}\mathbf{x} \quad (10.73)$$

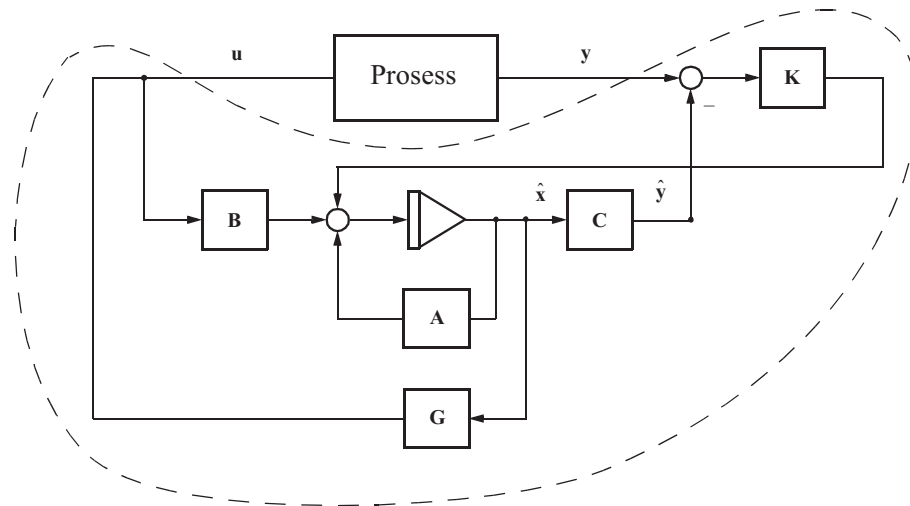
$\tilde{\mathbf{A}}$ er den matrisen som karakteriserer det resulterende multivariable tilbakekoblede systemet. Egenverdiene og egenvektorene til denne matrisen er med å bestemme hvordan tilstandsvektoren til det tilbakekoblede systemet skal svinge seg inn. Hvordan vi i detalj bestemmer matrisen \mathbf{G} for å oppnå en “optimal” innsvingning, skal vi som sagt ikke behandle her. Vi nøyer oss bare med å presisere at alle tilstandsvariable \mathbf{x} må være med for å generere pådragsvektoren \mathbf{u} .

Figur 10.16
Et multivariabelt tilbakekoplede system



Dersom ikke alle tilstandsvariablene er målbare eller matrisen \mathbf{C} ikke er inverterbar, må vi på en eller annen måte skaffe oss et *estimat* av tilstandsvektoren $\hat{\mathbf{x}}$. Dette kan gjøres ved hjelp av en struktur som er vist i figur 10.17, der en modell av prosessens dynamikk løper parallell med prosessen (en sanntidssimulator basert på (10.68) og (10.70)).

Figur 10.17
Et multivariabelt tilbakekoplede system med estimator

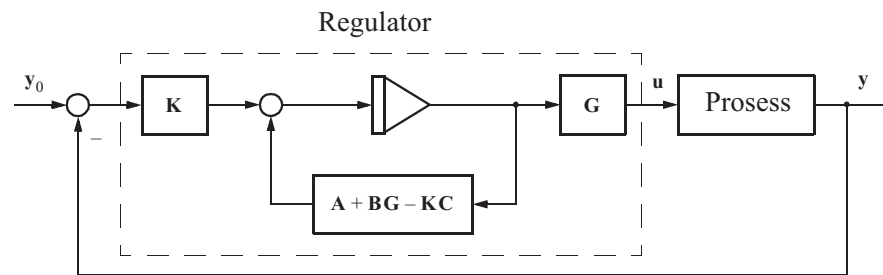


Ved hjelp av tilbakekopling gjennom matrisen \mathbf{K} blir denne modellen tvunget til å følge prosessen, og dermed genereres estimatet $\hat{\mathbf{x}}$. Dette kalles **tilstandsestimering**. Svært mye kan sies om dette emnet, men det hører heller ikke hjemme her. Vi kan for eksempel finne fram til en “optimal” tilstandsestimator, og det finnes en teori for bestemmelse av matrisen \mathbf{K} for slike tilfeller (Kalman-filter). En alternativ metode vil være å velge \mathbf{K} slik at egenverdiene til estimeringsløyfen, som er gitt av egenverdiene til $\mathbf{A} + \mathbf{K}\mathbf{C}$. Disse velges gjerne betydelig større (5-10 ganger) enn egenverdiene til $\mathbf{A} + \mathbf{B}\mathbf{G}$, som gir det tilbakekoblede systemets dynamikk med tilstandstilbakekopling. Vi kan nå realisere et multivariabelt reguleringssystem ved å ta en tilbakekopling gjennom matrisen \mathbf{G} fra $\hat{\mathbf{x}}$ istedenfor \mathbf{x} (som ikke er målbar), og under visse betingelser kan det vises at dette er den “optimale” løsningen.

Betrakter vi det sammensatte systemet som ligger innenfor den stiplede konturen i figur 10.17, ser vi at kombinasjonen av en tilstandsestimator og multivariabel tilbakekopling gjennom matrisen \mathbf{G} utgjør en *dynamisk multivariabel regulator* mellom \mathbf{y} og \mathbf{u} .

Hvis målevektoren \mathbf{y} er representativ for det vi ønsker å regulere (\mathbf{z} = egenskapsvektor og $\mathbf{z} = \mathbf{y}$), kan strukturen i figur 10.17 omformes som vist i figur 10.18.

Figur 10.18
En “optimal”
reguleringsstruktur



der regulatorens transfermatrise blir

$$\mathbf{H}_r(s) = \mathbf{G}(s\mathbf{I} - (\mathbf{A} + \mathbf{B}\mathbf{G} - \mathbf{K}\mathbf{C}))^{-1}\mathbf{K} \quad (10.74)$$

Vi ser at en prosess av orden n leder til en n -ordens regulator. Når matrisene \mathbf{A} , \mathbf{B} og \mathbf{C} er gitte, kan \mathbf{G} og \mathbf{K} bestemmes, for eksempel ved hjelp av såkalt LQG-teori, som er en kombinasjon av *optimal regulering* og *optimal tilstandsestimering*. MATLAB har beregningsrutiner for dette formålet.

10.6 SINGULÆRVERDIANALYSE AV LINEÆRE MULTIVARIABLE DYNAMISKE SYSTEMER

Et lineært multivariabelt dynamisk system kan som vi har sett, karakteriseres ved hjelp av systemets *transfermatrise* $\mathbf{H}(s)$ eller, ved innsetting av $s = j\omega$, med *frekvensresponsmatrisen* $\mathbf{H}(j\omega)$. Dersom $\mathbf{H}(s)$ har r pådrag som innganger og m målinger som utganger, vil denne matrisen ha $r \times m$ elementer som er *transferfunksjoner* $h_{ij}(s)$ (eller frekvensrespons $h_{ij}(j\omega)$). En slik representasjon er derfor svært omfangsrik.

Ofte vil det være tilstrekkelig for å få en rask oversikt over et multivariabelt systems ytelse å finne øvre og nedre grense av responsfunksjonene. For dette formålet er *singulærverdianalyse* et effektivt verktøy.

Vi skal illustrere metoden på en frekvensresponsmatrise $\mathbf{H}(j\omega)$ og antar i første omgang at den er kvadratisk ($m = r$), det vil si den har m^2 elementer. Vi finner $\mathbf{H}^*(j\omega)$ som er den transponerte og komplekse konjugerte av $\mathbf{H}(j\omega)$, og danner produktet

$$\mathbf{H}^*(j\omega)\mathbf{H}(j\omega)$$

Dette produktet er en skalar matrise som er en funksjon av ω , og har dimensjon $m \times m$. Denne matrisen kan karakteriseres ved sine m egenverdier $\lambda_i(\mathbf{H}^*\mathbf{H})$ og enda bedre med sine m singulærverdier

$$\sigma_i(\mathbf{H}(j\omega)) = \sqrt{\lambda_i(\mathbf{H}^*(j\omega)\mathbf{H}(j\omega))} \quad (10.75)$$

Alle $\sigma_i(j\omega)$ er positive reelle funksjoner av ω som kan sammenliknes med $|h_{ij}(j\omega)|$. Det kan vises at for alle $|h_{ij}(j\omega)|$ gjelder

$$|h_{ij}(j\omega)| \leq \bar{\sigma}(j\omega) \quad (10.76)$$

der $\bar{\sigma}(j\omega) = \bar{\sigma}(\mathbf{H}(j\omega))$ er den *største* singulærverdien til $\mathbf{H}(j\omega)$.

Når det gjelder den minste singulærverdien $\underline{\sigma}(j\omega) = \underline{\sigma}(\mathbf{H}(j\omega))$, har vi

$$\underline{\sigma}(j\omega) \leq \frac{\|\mathbf{H}(j\omega)\mathbf{u}(j\omega)\|_2}{\|\mathbf{u}(j\omega)\|_2} = \frac{\|\mathbf{y}(j\omega)\|_2}{\|\mathbf{u}(j\omega)\|_2}$$

der $\|\mathbf{u}(j\omega)\|_2 = \sqrt{u_1(-j\omega)u_1(j\omega) + \dots + u_r - j\omega u_r j\omega}$.

(Det bør bemerkes at

$$\underline{\sigma}(j\omega) \leq |h_{ij}(j\omega)|$$

ikke alltid er oppfylt, som illustrert med et eksempel lengre fram, i figur 10.21.)

Vi ser at det er lagt inn to negative foroverkoplinger for å illustrere virkningen av ikke-minimum-fase (IMF) fenomener.

Dette fører til at

$$\mathbf{A} = \begin{bmatrix} -2 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 5 & -4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (10.80)$$

$$\mathbf{C} = \begin{bmatrix} -a & (1+a) & a \\ 0 & -b & 1 \end{bmatrix}$$

Vi finner elementene i $\mathbf{H}_u(s)$ ved å bruke (10.79) eller direkte ut fra blokkdiagrammet i figur 10.19.

$$h_{11} = \frac{y_1}{u_1}(s) = \frac{4}{18} \frac{(1+0.25s)(1-as)}{(1+0.5s) \left(1 + 2 \cdot 0.833 \cdot \frac{s}{3} + \left(\frac{s}{3}\right)^2 \right)}$$

$$h_{12} = \frac{y_1}{u_2}(s) = -\frac{1}{9} \frac{1-as}{\left(1 + 2 \cdot 0.833 \cdot \frac{s}{3} + \left(\frac{s}{3}\right)^2 \right)}$$

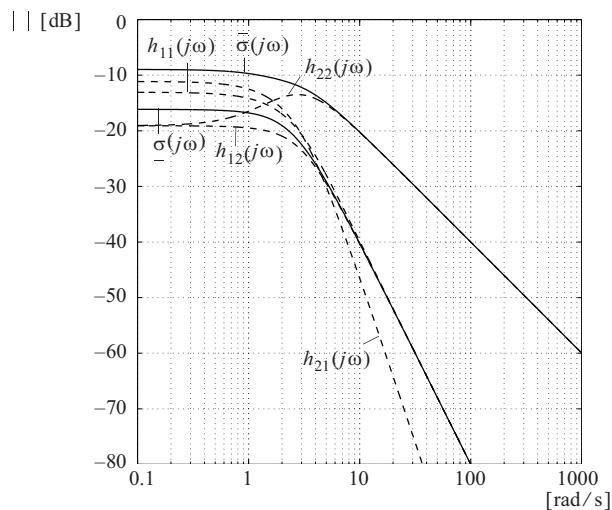
$$h_{21} = \frac{y_2}{u_1}(s) = \frac{5-4b}{18} \frac{\left(1 - \frac{b}{5-4b}s \right)}{(1+0.5s) \left(1 + 2 \cdot 0.833 \cdot \frac{s}{3} + \left(\frac{s}{3}\right)^2 \right)}$$

$$h_{22} = \frac{y_2}{u_2}(s) = \frac{1+b}{9} \frac{\left(1 + \frac{1}{1+b}s \right)}{\left(1 + 2 \cdot 0.833 \cdot \frac{s}{3} + \left(\frac{s}{3}\right)^2 \right)}$$

Frekvensresponsen $h_{ij}(j\omega)$ for disse når $a = 0$ og $b = 0$, er vist i figur 10.21 sammen med $\bar{\sigma}(j\omega)$ og $\underline{\sigma}(j\omega)$. Vi ser at $\bar{\sigma}(j\omega)$ gir en god representasjon av øvre grense for $|h_{ij}(j\omega)|$. Men vi får ingen opplysninger om fasevinklene $\angle h_{ij}(j\omega)$ som jo er av avgjørende betydning for systemets oppførsel.

Figur 10.21

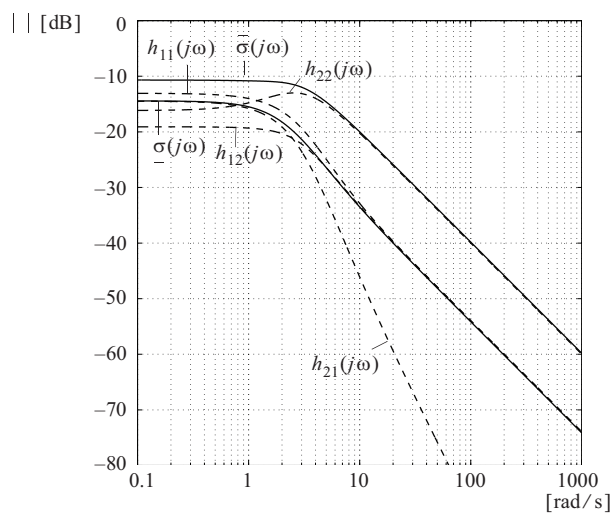
Frekvensresponsene
 $h_{ij}(j\omega)$ og singu-
 lærverdiene $\bar{\sigma}(j\omega)$
 og $\underline{\sigma}(j\omega)$ når
 $a = 0$ og $b = 0$



Et tilsvarende diagram er vist i figur 10.22 der $a = 0.2$ og $b = 0.4$.

Figur 10.22

Frekvensresponsene
 $h_{ij}(j\omega)$ og singu-
 lærverdiene $\bar{\sigma}(j\omega)$
 og $\underline{\sigma}(j\omega)$ når
 $a = 0.2$ og
 $b = 0.4$



Figur 10.22 kan genereres ved bruk av MATLAB-koden i figur 10.23.

Figur 10.23
MATLAB-kode

```

1 a = 0.2;
2 b = 0.4;
3 A = [-2, 0, 0; 1, -1, -1; 0, 5, -4];
4 B = [1, 0; 0, 0; 0, 1];
5 C = [-a, (1+a), a; 0, -b, 1];
6 D = [0, 0; 0, 0];
7 w = logspace(-1, 3, 1000);
8
9 [sv] = sigma(A,B,C,D,w); ← Beregner singularverdiene som funksjon av frekvensen
10 [amp0, fase0] = bode(A,B,C,D,1,w); ← Beregner frekvensrespons fra input 1
11 [amp1, fase1] = bode(A,B,C,D,1,w); ← Beregner frekvensrespons fra input 2
12
13 semilogx(w, 20*log10(sv), w, 20*log10(amp0(:,1)), '--', w, 20*log10(amp0(:,2)),
14 '--', w, 20*log10(amp1(:,1)), '--', w, 20*log10(amp1(:,2)), '--', 'linewidth'1.2)
15 grid
16 set(get(1,'CurrentAxes'),'YTick', [-80 -70 -60 -50 -40 -30 -20 -10 0])
17 set(get(1,'CurrentAxes'),'YLim', [-80 0])

```

Det er imidlertid mer interessant å undersøke størrelsen

$$\mathbf{N}^y(j\omega) = (\mathbf{I} + \mathbf{H}_0^y(j\omega))^{-1} = (\mathbf{I} + \mathbf{H}_u(j\omega)\mathbf{H}_r(j\omega))^{-1} \quad (10.81)$$

som gir et uttrykk for hvordan en tilbakekopling rundt prosessen $\mathbf{H}_u(j\omega)$ vil oppføre seg. Største singularverdi $\bar{\sigma}_N(j\omega)$ av denne vil derfor være en nyttig øvre skranke for $|\mathbf{N}^y(j\omega)|$.

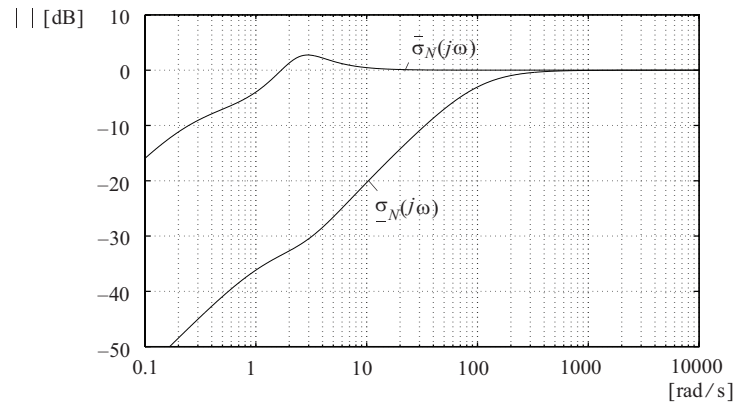
Figur 10.24 viser kurver for $\bar{\sigma}_N(j\omega)$ og $\underline{\sigma}_N(j\omega)$ når vi for prosessen i (10.80) har $a = 0$ og $b = 0$, og velger en regulator med transfermatrisen

$$\mathbf{H}_r(s) = \begin{bmatrix} K_{p1} \frac{1 + T_{i1}s}{T_{i1}s} & 0 \\ 0 & K_{p2} \frac{1 + T_{i2}s}{T_{i2}s} \end{bmatrix}$$

der $K_{p1} = 5$, $T_{i1} = 3$ [s], $K_{p2} = 100$ og $T_{i2} = 0.3$ [s].

Figur 10.24

$\bar{\sigma}_N(j\omega)$ og
 $\underline{\sigma}_N(j\omega)$ med
 $a = 0$ og $b = 0$

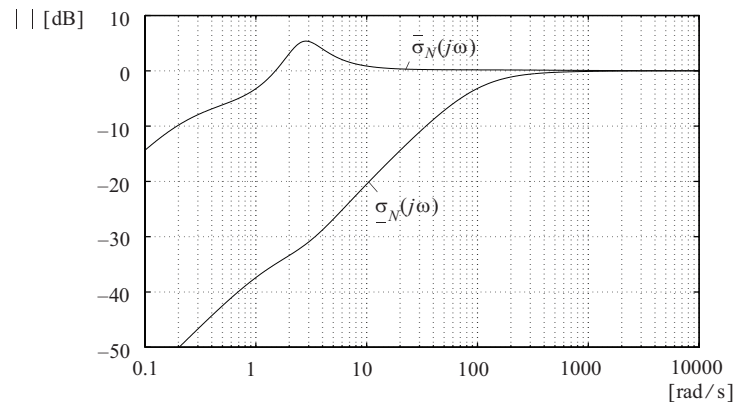


Vi ser at $\bar{\sigma}_N(j\omega)$ som er en øvre skranke for alle $|\mathbf{N}^y(j\omega)|$, gir et godt bilde av systemets ytelse (se avsnitt 9.9). $\underline{\sigma}_N(j\omega)$ vil i dette tilfellet være en nedre skranke for alle $|\mathbf{N}^y(j\omega)|$, men denne er av mindre nytte for oss.

Hvis vi så setter $a = 0.2$ og $b = 0.4$ og beholder den opprinnelige regulatoren, får vi følgende forløp for $\bar{\sigma}_N(j\omega)$ og $\underline{\sigma}_N(j\omega)$.

Figur 10.25

$\bar{\sigma}_N(j\omega)$ og
 $\underline{\sigma}_N(j\omega)$ med
 $a = 0.1$ og
 $b = 0.2$



Vi ser nå tydelig at resonanstoppen i $\bar{\sigma}_N(j\omega)$ er øket. Dette tyder på at stabilitetsmarginen er blitt dårligere.

Robust regulering slik som denne kort ble beskrevet i avsnitt 9.9 for monovariable systemer, kan utvides til multivariable systemer. Dette er detaljert behandlet i Skogestad og Postlethwaite (1996).

Figur 10.25 kan genereres ved bruk av MATLAB-koden i figur 10.26. Fremgangsmåten er annerledes enn ved generering av figur 10.22, siden vi ikke kan sette tilstandsrommatriser inn i *sigma*.

Figur 10.26
MATLAB-kode

```

1 a = 0.2; b = 0.4;
2 Kp1 = 5; Kp2 = 100;
3 Ti1 = 3; Ti2 = 0.3;
4 A = [-2, 0, 0; 1, -1, -1; 0, 5, -4];
5 B = [1, 0; 0, 0; 0, 1];
6 C = [-a, (1+a), a; 0, -b, 1];
7 D = [0, 0; 0, 0];
8 w = logspace(-1, 4, 1000);
9
10 sys_Hu = pck(A,B,C,D);
11 sys_hr1 = nd2sys(Kp1*[Ti1 1], [Ti1 0]);
12 sys_hr2 = nd2sys(Kp2*[Ti2 1], [Ti2 0]);
13 sys_Hr = daug(sys_hr1, sys_hr2);
14 sus_H0 = mmult(sys_Hu, sys_Hr);
15 sys_N = minv(madd(eye(2), sys_H0));
16 N_jw = frsp(sys_N, w);
17 sigma_N_w = vsvd(N_jw);
18 [sigma_N_dB] = vebe('20*log10', sigma_N_w);
19
20 vplot('liv,d', sigma_N_dB), grid

```

Lager systemmatrise for prosessen
 Lager systemmatrise for $H_{r,11}$
 Lager systemmatrise for $H_{r,22}$
 Danner regulatormatrise
 Systemmatrise for sløyfetransferfunksjonen
 Systemmatrise for avviksforholdet
 Frekvensrespons for avviksforholdet
 Beregner singularverdiene som funksjon av frekvensen
 Tar 20 \log_{10} av elementene i $\sigma_N(j\omega)$
 Plotter øvre og nedre singularverdi

I denne MATLAB-koden benyttes funksjoner fra μ -tools.



10.7 FUNKSJONELL STYRBARHET (FS) AV MULTIVARIABLE SYSTEMER

Kravene til (tilstands) styrbarhet er behandlet i kapittel 5. For et monovariabelt system er det lett å erkjenne at disse kravene er unødig strenge i forhold til det som kreves for å realisere et enkelt tilbakekoplet reguleringssystem, nemlig $h_u(s) \neq 0$ for alle s unntatt i systemets nullpunkter. (Dette er betegnet *funksjonell styrbarhet* (FS)).

I multivariable systemer er problemet betydelig mer komplisert fordi vi i slike systemer har mange forbindelser $h_{ij}(s)$ mellom inngangene (pådragene) og utgangene (målingene/egenskapene).

Problemet er å finne betingelser for at det er mulig å styre (påvirke) alle m utganger ($m = \dim \mathbf{y}$) av et system uavhengig av hverandre ved hjelp av systemets r innganger (pådrag) ($r \geq m$).

Rosenbrock (1970) har presentert følgende betingelser for funksjonell styrbarhet (functional controllability):

Et multivariabelt ($r \times m$) system $\mathbf{H}_u(s)$ ($r \geq m$) er funksjonelt styrbart (FS) dersom normal rang av $\mathbf{H}_u(s)$ er lik antall utganger m .

Normal rang av $\mathbf{H}_u(s)$ er rang $\mathbf{H}_u(s)$ ved alle verdier av s unntatt i et endelig antall punkter som er nullpunkter av $\mathbf{H}_u(s)$.

Funksjonell ikke-styrbarhet (FIS) har vi når et system ikke tilfredsstillers kravet til FS.

Vi kan oppsummere noen resultater knyttet til FIS (Skogestad og Postlethwaite 1996):

- For et system på tilstandsromform

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

gjelder at $\mathbf{H}_u(s)$ er FIS hvis

- rang $\mathbf{B} < m$ (inngangsreduisert)
- rang $\mathbf{C} < m$ (utgangsreduisert)
- rang $(s\mathbf{I} - \mathbf{A}) < m$ (færre tilstander enn utganger)
- En prosess $\mathbf{H}_u(s)$ er FIS hvis og bare hvis minst en singularverdi $\sigma_{\mathbf{H}_u}(j\omega) = 0$ for alle ω . Som et mål for systemets *grad* av FIS kan en, når $r \geq m$, bruke $\underline{\sigma}_{\mathbf{H}_u}(j\omega)$.
- Vanligvis er FIS en *strukturell* egenskap ved prosessen og er derfor ikke avhengig av spesielle parameterverdier i prosessmodellen.
- Et enkelt eksempel på FIS er en prosess der ingen av pådragene u_i har noen påvirkning på en bestemt utgang y_j (eller z_j).
- Et annet eksempel på FIS er når en prosess har færre pådrag enn utganger ($r < m$).

EKSEMPEL 10.5: FIS ved 2×2 system

2×2 -systemet som vi studerte i avsnitt 10.2, har prosesstransferfunksjonen

$$\mathbf{H}_u(s) = \begin{bmatrix} h_{11}(s) & h_{12}(s) \\ h_{21}(s) & h_{22}(s) \end{bmatrix} \quad (10.82)$$

rang $\mathbf{H}_u(s) = 2$ dersom

$$\begin{aligned} \det \mathbf{H}_u(s) &= h_{11}(s)h_{22}(s) - h_{12}(s)h_{21}(s) \\ &= h_{11}(s)h_{22}(s)[1 - Y(s)] \neq 0 \end{aligned} \quad (10.83)$$

$$\text{hvor } Y(s) = \frac{h_{12}(s)h_{21}(s)}{h_{11}(s)h_{22}(s)}$$

Systemet er FIS dersom rang $\mathbf{H}_u(s) = 1$. Dette oppstår når $Y(s) = 1$.

■

Diskret regulering av kontinuerlige systemer

11.1 Innledning og motivering 501

11.2 Analyse ved kontinuerlig tilnærming 505

11.3 Filtre og filtrering 512

11.4 Diskret analyse og syntese — Z-transformasjonen 518

11.5 Analyse av diskrete systemer ved W-transformasjon 534

11.6 Bruk av W-transformasjon ved syntese av diskret regulator 545

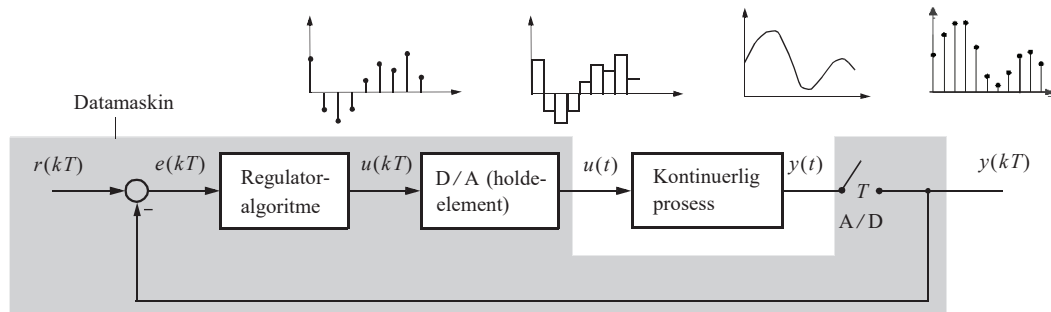
11.7 Tidsrespons av diskret system 548

11.8 Den multivariable W-transformasjonen 550

11.1 INNLEDNING OG MOTIVERING

Reguleringssystemer realiseres i dag ved hjelp av datamaskiner som i sin natur fungerer tidsdiskret. I de foregående kapitlene har vi gått ut fra at prosesser og reguleringsutstyr arbeider kontinuerlig, og at pådraget settes ut i det samme tidspunkt som målingen foretas. Men forsinkelser fra måling til pådrag vil framkomme av forskjellige årsaker. Den viktigste forsinkelsesfaktor skyldes *diskret regulering*. Det tas en måling av en eller flere kontinuerlige variable med jevne mellomrom; de **tastes** (norsk-engelsk: “**samples**”). Denne leses så inn i en datamaskin der beregning foregår, og omgjøres igjen til et kontinuerlig signal som sendes til et pådragsorgan. I denne kjeden får man en tidsforsinkelse som blir viktig når taste-tidsintervallet (**tastetid**, “samplingstid”) er forholdsvis stort i forhold til prosessens dynamikk. For å kunne analysere slike situasjoner, skal vi derfor utvide teorien til å omfatte kontinuerlige systemer med tasting og datamaskin i reguleringsløyfen. Figur 11.1 viser problemstillingen. En monovariabel og kontinuerlig prosess tastes med et tidsintervall T . Målingen gjøres om til et binært tall av en analog-til-digital (A/D) - omsetter, og blir inngangsvariabel til en regulator-algoritme i en datamaskin. Denne leverer resultatet til en D/A - omsetter, som også er et “**holdeelement**”. En slik elektronisk komponent sørger for at det binære tallet som den mottar fra reguleringsalgoritmen, omsettes til en motsvarende analog verdi. Denne settes ut til pådragsorganet og holdes konstant inntil neste tastetidspunkt, da en ny analog verdi settes ut.

Figur 11.1
Monovariabel kontinuerlig prosess med diskret regulator



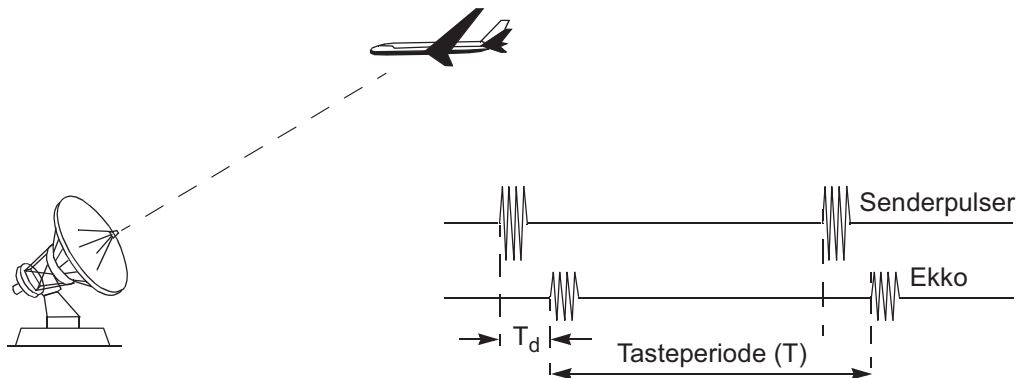
Vi har $y(kT) = y(t)|_{t=kT}$. Når **tastetida** T er gitt, skriver vi $y(kT)$ som $y[k]$. Hakeparentes $[]$ symboliserer **diskret tid** $k - 1, k, k + 1, \dots$. I det skraverte feltet er alle signal tidsdiskrete (tastede), ellers er de kontinuerlige.

Men før vi tar for oss regulering med datamaskin som vist i figur 11.1, skal vi nevne noen *andre* tilfeller hvor tasting av kontinuerlige signaler er nødvendig.

EKSEMPEL 11.1: Avstandsmåling ved bruk av radar

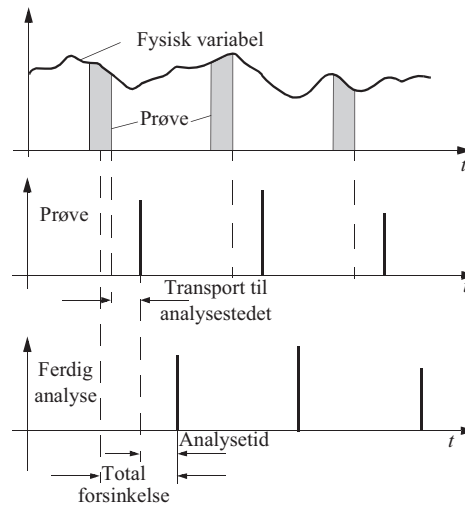
Ved bestemmelse av avstanden til et bevegelig mål ved hjelp av radar (figur 11.2), vil informasjon om målets avstand bare foreligge diskret hver gang det sendes ut en høyfrekvent puls gjennom antennen i retning av målet, og når dette reflekterer pulsen til antennen. Hvis antennen til enhver tid peker mot målet, vil **tastefrekvensen** være relativt høy, ca. 1000 pulser/s. Dersom antennen roterer eller beveger seg på en annen måte for å søke over et større område, vil det foreligge informasjon om et bestemt punkt bare hver gang antennen peker i den bestemte retningen, for eksempel en gang hvert tredje sekund. Dette fenomenet er selvsagt ikke tilsiktet, men det ligger i det benyttede radarprinsippets natur. Vi må derfor ta hensyn til det ved analyser der dette fenomenet kommer inn.

Figur 11.2
Avstandsmåling ved hjelp av radar



EKSEMPEL 11.2: Prøvetaking i kjemisk prosess

Figur 11.3
Forsinkelse som følge av transport til analysested og analysetid

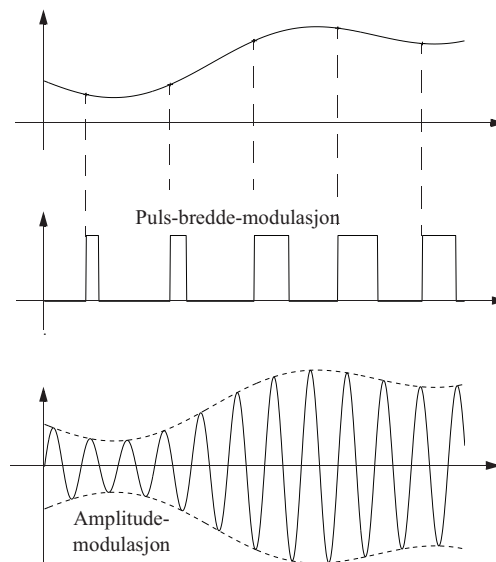


Mange måletekniske oppgaver, for eksempel kompliserte kjemiske analyser ved prosessregulering, lar seg ofte best utføre på diskret basis. Det foregår da en mekanisert prøvetaking (tasting) av mediet som skal analyseres. En kjemisk analyse, for eksempel ved titrering eller spektroskopi, foregår i et mekanisert apparat hvoretter resultatet framkommer etter noen tids forløp som vist i figur 11.3. I et slikt tilfelle har vi i høy grad en diskret angivelse, men den vil også være tidsforsinket (ekvivalent med transportforsinkelse i kontinuerlige systemer).

■

EKSEMPEL 11.3: Pulsbreddemodulasjon

Figur 11.4
Pulsbreddemodulasjon og amplitudemodulasjon



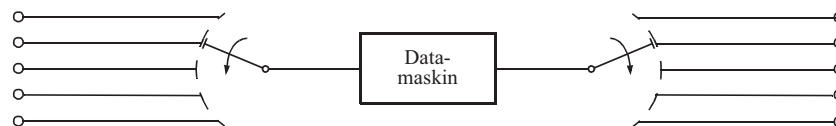
Flere av de vanligst forekommende modulasjonsmetoder som benyttes i måleapparater og for signalforsterkning, vil medføre en diskret angivelse av informasjonen. **Pulsbreddemodulasjon** (figur 11.4) vil for eksempel medføre det samme diskrete fenomen som radarantennen ovenfor.

(Men **amplitudemodulasjon** av en sinusformet bæreølge vil tilnærmet gi de samme fenomener som skal analyseres i det følgende.)

EKSEMPEL 11.4: Multiplekser

Figur 11.5 viser skjematisk en “multiplekser” i form av to synkroniserte velgere (som oftest elektroniske) som kan kople et stort antall kontinuerlige (analoge) signalkanaler til et databehandlingsutstyr for innlesing og utlesing. På inngangssiden vil det før datamaskinen være en **A/D-omsetter** (analog/digital-omsetter) og på utgangssiden vil det tilsvarende være en **D/A-omsetter**.

Figur 11.5
Inn- og utlesing til datamaskin



11.2 ANALYSE VED KONTINUERLIG TILNÆRMING

Fra nå av og i resten av kapittel 11 vil vi behandle reguleringstrukturer av den type som er skissert i figur 11.1. Den enkleste måten å angripe diskret regulering på, er å betrakte den diskrete regulatoren *som om den var kontinuerlig*, for da kan hele reguleringssystemet betraktes som kontinuerlig. Denne tilnærmede metoden går som regel bra hvis tastetida T er $\ll 1/|\lambda|_{\max}$; altså at den egenverdien som har størst tallverdi er bestemmende (det er den som representerer de hurtigste svingningene i prosessen). Men en bedre regel er gitt rett etter eksempel 11.5.

En diskret regulator er en rekursiv algoritme i en datamaskin. Vi skal nå utlede en regulatoralgoritme som svarer til en kontinuerlig PI-regulator. Siden en av de mest vanlige regulatorer i industriell sammenheng er PI-regulatoren, er det av interesse å utlede en diskret variant. Dessuten er prosedyren for å finne andre diskrete regulatoralgoritmer (PD, PID, begrenset PI og PID, m.fl.), tilsvarende.

11.2.1 Diskret PI-regulator

I det kontinuerlige tilfellet har vi for PI-regulatoren

$$h_r(s) = K_p \frac{1 + T_i s}{T_i s} \quad (11.1)$$

som i tidsplanet svarer til

$$u(t) = K_p \left(e(t) + \frac{1}{T_i} \int_0^t e(\sigma) d\sigma \right) \quad (11.2)$$

Vi skal utlede en diskret PI-regulator og starter med å dele integralet opp i to deler

$$u(t) = K_p \left(e(t) + \frac{1}{T_i} \int_0^{t-T} e(\sigma) d\sigma + \frac{1}{T_i} \int_{t-T}^t e(\sigma) d\sigma \right) \quad (11.3)$$

(11.2) gir

$$\frac{u(t)}{K_p} - e(t) = \frac{1}{T_i} \int_0^t e(\sigma) d\sigma \quad (11.4)$$

Andre ledd i parentesens i (11.3) kan derfor skrives som

$$\frac{u(t-T)}{K_p} - e(t-T)$$

mens siste ledd i (11.3) kan approksimeres med trapesintegrasjon som gir

$$\frac{1}{T_i} \left(\frac{T}{2} (e(t) + e(t - T)) \right)$$

Med den diskrete notasjonen $u(kT) = u[k]$ osv., får vi

$$\begin{aligned} u[k] &= K_p \left(e[k] + \left(\frac{u[k-1]}{K_p} - e[k-1] \right) + \frac{1}{T_i} \frac{T}{2} (e[k] + e[k-1]) \right) \\ &= u[k-1] + K_p \left(1 + \frac{T}{2T_i} \right) e[k] - K_p \left(1 - \frac{T}{2T_i} \right) e[k-1] \end{aligned} \quad (11.5)$$

eller

$$u[k] = u[k-1] + g_0 e[k] + g_1 e[k-1] \quad (11.6)$$

Dette er en rekursiv algoritme som kan realiseres i en datamaskin.

Vi skal i avsnitt 11.2.3 se hvordan vi kan lage vilkårlige diskrete regulatorer.

11.2.2 Frekvensanalyse, kontinuerlig tilnærming

Vi antar nå at T er så liten at (11.6) har samme egenskaper som en kontinuerlig PI-regulator, og velger K_p og T_i basert på en kontinuerlig syntesemetode, for eksempel at et minimumskrav til fase- og forsterkningsmargin skal oppfylles. Dette kan best forklares via et eksempel:

EKSEMPEL 11.5: Dimensjonering av kontinuerlig og diskret PI-regulator

Anta

$$h_u(s) = \frac{1}{(1 + T_1 s)(1 + T_2 s)}$$

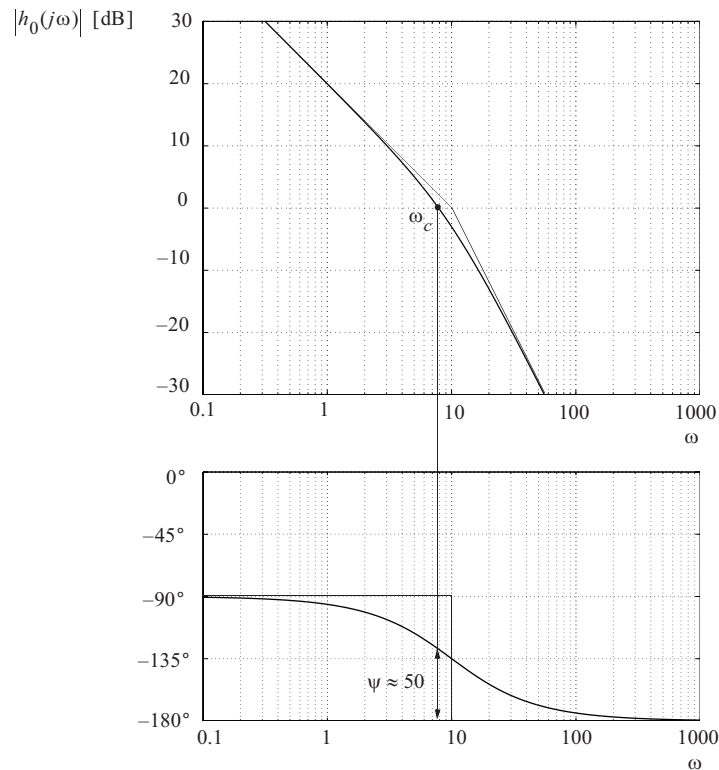
og tastetid $= T$ i regulatoren.

Videre antas $T_1 = 1$ og $T_2 = 0.1$. Vi later som om alt er kontinuerlig, og starter med å velge integraltid $T_i = T_1$, slik at sløyfetransferfunksjonen blir

$$h_0(s) = \frac{K_p}{T_1 s(1 + T_2 s)} = \frac{K_p}{s(1 + 0.1s)} \quad (11.7)$$

Av Bode-diagrammet i figur 11.6 ser vi at $K_p = 1/T_2 = 10$ gir $\psi \approx 50^\circ$ og $\Delta K = \infty$, noe som er tilfredsstillende.

Figur 11.6
Bodediagram for
prosessen $h_u(s)$
med kontinuerlig
regulator

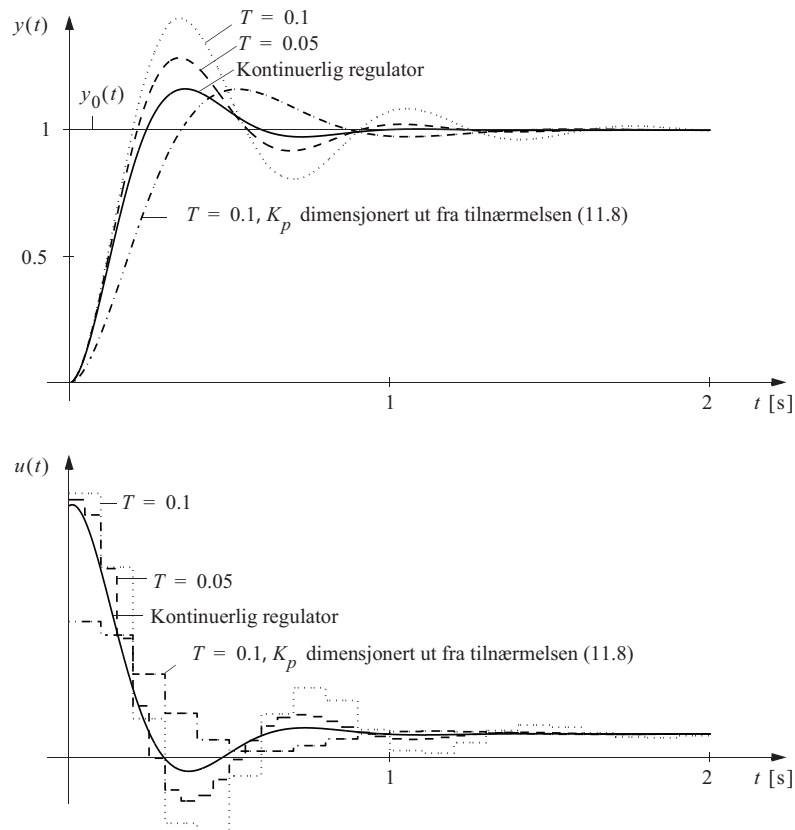


Sprangresponsene for $y(t)$ og $u(t)$ med kontinuerlig regulator og med denne regulatorinnstillingen er vist i figur 11.7 med heltrukne linjer.

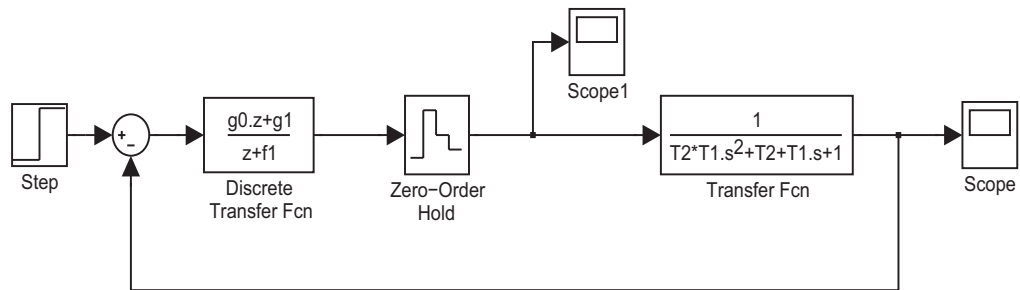
Men nå *er* regulatoren altså diskret. Figur 11.7 viser også den virkelige sprangresponsen når den kontinuerlige prosessen styres med de valgte verdiene for K_p og T_i , men nå med diskret PI-regulator. Responsen er vist for to alternative valg av tasterintervaller, $T = 0.05$ og $T = 0.1 = T_2$. Vi ser at det er risikabelt å bruke denne enkle tilnæringsmetoden når T økes så mye at den nærmer seg den minste tidskonstanten T_2 i prosessen.

Responsene i figur 11.7 med diskret regulator kan genereres med simulinkfilen i figur 11.8, men vi må huske å gi konstantene en verdi. Dette kan gjøres ved å lage en parameterfil som kjøres i Matlab.

Figur 11.7
Utgang og pådrag ved kontinuerlig og diskret regulator

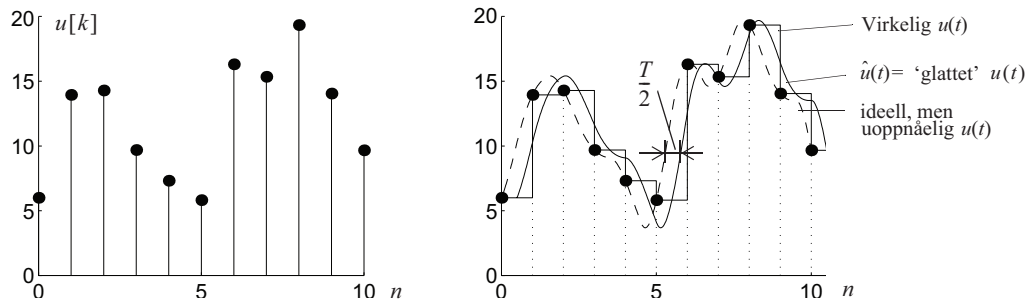


Figur 11.8
Simulinkvindu



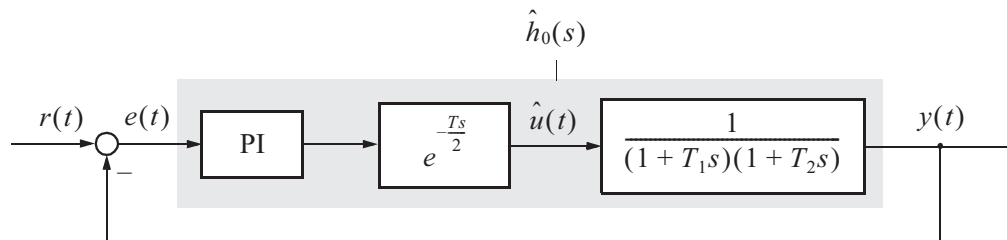
En bedre kontinuerlig tilnæringsmetode tar hensyn til den viktigste bivirkningen av holdelementet, nemlig en *tidsforsinkelse*. Som det framgår av figur 11.9, vil holdelementet tilnærmet innebære en tidsforsinkelse på halve tastetida.

Figur 11.9
Virkningen av
holdeelementet



Et tilnærmet ekvivalent kontinuerlig system blir da som vist i figur 11.10.

Figur 11.10
Det ekvivalente
kontinuerlige
system



Vi får

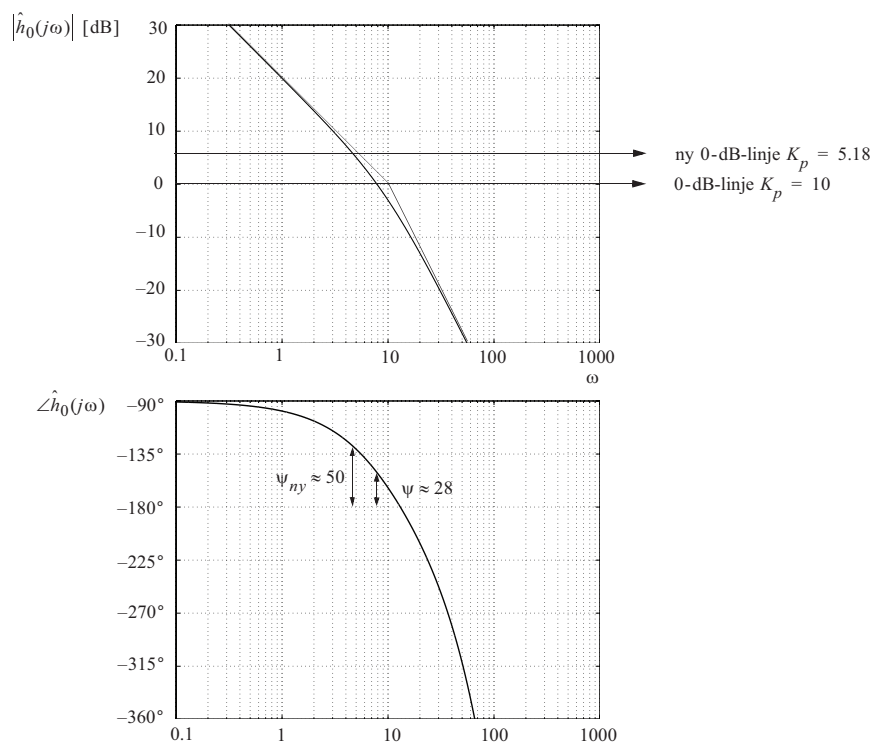
$$\hat{h}_0(s) = \frac{K_p e^{-\frac{T_s}{2}}}{s(1 + 0.1s)} \quad (11.8)$$

Figur 11.11 viser Bodediagram for $\hat{h}_0(s)$ med $T = 0.1$.

Vi må nå velge K_p mindre for å få rimelige stabilitetsmarginer. Den siste og fjerde responsen i figur 11.7 viser at selv for stor tastetid $T = 0.1$, får vi akseptabel respons med diskret regulator, når vi har tatt tilnærmet hensyn til holdeelementets virkning.

Når dette er sagt, ønsker vi adgang til et begrepsapparat og metoder som gjør det mulig å analysere diskret regulering av kontinuerlige systemer *eksakt*. De resterende avsnitt i kapittel 11 dreier seg om dette.

Figur 11.11
Bode-diagram
for $\hat{h}_0(s)$ med
 $T = 0.1$



Fra figurene 11.9-11.11 ser vi at tastetida T bør velges slik at fasebidraget til den syntetiske tidsforsinkelsen som vi putter inn for å approksimere virkninga av holdeelementet, ikke bør bli mer enn et par grader. Dette oppnås ved å kreve at $\omega_c T/2 < d\psi$, der $d\psi$ er akseptabel reduksjon av fasemargin (i radianer). Vi får altså

$$T < 2d\psi/\omega_c \quad (11.9)$$

11.2.3 Diskret regulator ved hjelp av tidsforskyvingsoperator

Vi har tidligere sett i avsnitt 4.5.3 hvordan vi kan bruke operatortotasjon i kontinuerlige systemer: derivasjon symboliseres med operatoren s . Dette tillater algebraiske manipulasjoner på differensiallikninger som om s var en multiplikativ konstant.

For diskrete systemer kan vi tilsvarende innføre **tidsforskyvingsoperatoren**, kalt z , slik at $zx[k]$ symboliserer $x[k+1]$, og $z^{-1}x[k]$ symboliserer $x[k-1]$.

Betrakt integrasjonen

$$x(t) = \int_{-\infty}^t u(\sigma) d\sigma \quad (11.10)$$

En tilnærming ved hjelp av trapesintegrasjon blir da, jfr. avsnitt 11.4.1,

$$x[k] = x[k-1] + \frac{T}{2}(u[k] + u[k-1]) \quad (11.11)$$

der T er tastetida. Vi kan nå benytte tidsforskyvingsoperatoren, og får

$$x[k] = z^{-1}x[k] + \frac{T}{2}(u[k] + z^{-1}u[k]) \quad (11.12)$$

Vi kan løse dette m.h.p. $x[k]$:

$$x[k] = \left(\frac{T}{2} \frac{1+z^{-1}}{1-z^{-1}} \right) u[k] = \left(\frac{T}{2} \frac{z+1}{z-1} \right) u[k] \quad (11.13)$$

(11.13) er altså trapes-tilnærmelsen til kontinuerlig integrasjon,

$$x(t) = \frac{1}{s}u(t) \quad (11.14)$$

s er her den kontinuerlige derivasjonsoperator. Den diskrete tilnærming til kontinuerlig derivasjon blir da

$$u[k] = \left(\frac{2}{T} \frac{z-1}{z+1} \right) x[k] \quad (11.15)$$

som svarer til $u(t) = sx(t)$. Dette resultatet kan vi bruke for å lage en diskret regulator med utgangspunkt i en vilkårlig kontinuerlig regulator. Vi erstatter da bare alle forekomster av s i den kontinuerlige regulatorens transferfunksjon med $\frac{2}{T} \frac{z-1}{z+1}$.

EKSEMPEL 11.6: Utleddning av diskret PI-regulator ved hjelp av tidsforskyvingsoperator

Inngangsvariabelen er avviket $e[k]$, utgangsvariabelen er pådraget $u[k]$. Vi erstatter først alle s i den kontinuerlige PI-regulatoren med $\frac{2}{T} \frac{z-1}{z+1}$, og får

$$u[k] = K_p \frac{1 + T_i \left(\frac{2}{T} \frac{z-1}{z+1} \right)}{T_i \left(\frac{2}{T} \frac{z-1}{z+1} \right)} \quad (11.16)$$

Vi multipliserer med $T(z+1)$ i teller og nevner, og får

$$u[k] = K_p \frac{T(z+1) + 2T_i(z-1)}{2T_i(z-1)} e[k] \quad (11.17)$$

Dette gir

$$\begin{aligned} 2T_i(z-1)u[k] &= K_p(T(z+1) + 2T_i(z-1))e[k] \Leftrightarrow \\ u[k+1] - u[k] &= \frac{K_p}{2T_i}(Te[k+1] + Te[k] + 2T_ie[k+1] - 2T_ie[k]) \Leftrightarrow \\ u[k+1] &= u[k] + K_p \left(\left(1 + \frac{T}{2T_i}\right) e[k+1] - \left(1 - \frac{T}{2T_i}\right) e[k] \right) \end{aligned} \quad (11.18)$$

som er samme resultat som det vi utledet i avsnitt 11.4.1. At vi har $u[k+1]$ på venstre side av likhetstegnet i (11.18) stedet for $u[k]$ slik som i (11.5) gjør ingen forskjell. Det er bare å dekrementere tida et skritt i alle ledd på begge sider av likhetstegnet i (11.18), hvis man foretrekker den ekvivalente formen (11.5).

■

Denne metoden kan anvendes uansett hvilken type kontinuerlig regulatorstruktur som er utgangspunktet.

Som vi ser, tillater operatoren z algebraiske manipulasjoner på **differenslikninger** (som er det samme som tidsdiskrete rekursive likninger) som om z var en multiplikativ konstant.

Merk også at vi i likning (11.17) har noe som minner om en transferfunksjon, med fri variabel z i stedet for s . Den diskrete transferfunksjon behandles i avsnitt 11.4.4.

11.3 FILTRE OG FILTRERING

Vi skal ta for oss fire slags analoge filtre: Lavpass, høypass, båndpass og båndstopp (også kalt “notch” = “hakk”-filter på engelsk). Først viser vi hvordan filtrene kan konstrueres, så forklarer vi i avsnitt 11.3.6 om hvordan/hvorfor de er nyttige i reguleringssystemer.

11.3.1 Analoge og diskrete filtre

Et analogt filter realiseres ved hjelp elektroniske komponenter, og virker på et kontinuerlig tidssignal. Man kan derfor også kalle det et analogt, elektronisk eller kontinuerlig filter. Et slikt filter beskriver vi ved hjelp av en rasjonal transferfunksjon $h(s)$.

Vi har også (tids)diskrete filtre, som realiseres som en algoritme i en datamaskin. En PI-regulator slik vi har presentert den i flere kapitler til nå, kan betraktes som et analogt (tids-kontinuerlig) filter, mens den tidsdiskrete utgaven introdusert i avsnitt 11.2.1 kan betraktes som et tilsvarende diskret filter. For å komme seg fra analog til diskret filter, kan man substituere for s i $h(s)$ som indikert i avsnitt 11.2.3, og så finne en rekursiv formel for det

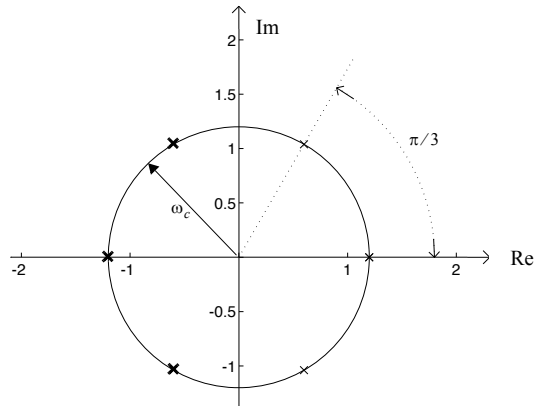
diskrete filteret som demonstrert i eksempel 11.6. Men diskrete filtre er ikke tema i denne boka.

11.3.2 Butterworth-filteret – et lavpassfilter

Dette er et av flere typer analoge lavpassfiltre. Det særegne med BW-filteret er at frekvensresponsen er maksimalt flat i passbåndet, og at frekvensresponsen er monotont fallende i stoppbåndet – det er ingen “bulker” noe sted i frekvensforløpet. Andre typer analoge filtre kan ha brattere overgang mellom pass og stoppområde, men da er prisen man betaler at det blir bulker i frekvensforløpet. Transferfunksjonen $h_L(s)$ (“L” = “lavpass”) for et BW-filter av orden n kan finnes ved en enkel geometrisk betraktning: de n polene ligger på hjørnene i en regulær mangekant med $2n$ sider, der mangekanten er plassert slik at den ligger symmetrisk om Im-aksen, og hvor avstanden fra origo og ut til hvert hjørne (= pol) er filterets knekkfrekvens ω_c . Halvparten av disse polene ligger i venstre halvplan, og de utgjør filterets n poler λ_i . Se figur 11.12 hvor $n = 3$ og mangekanten da blir en sekskant. (Hvorfor akkurat en slik beliggenhet av polene gir BW-filteret dets egenskaper, vises ikke her.)

Figur 11.12

Polene i et
BW-filter med
 $n = 3$ og $\omega_c = 1.2$



Transferfunksjonen for BW-filteret blir

$$h_L(s) = \frac{\omega_c^n}{(s - \lambda_1)(s - \lambda_2) \dots (s - \lambda_n)} \quad (11.19)$$

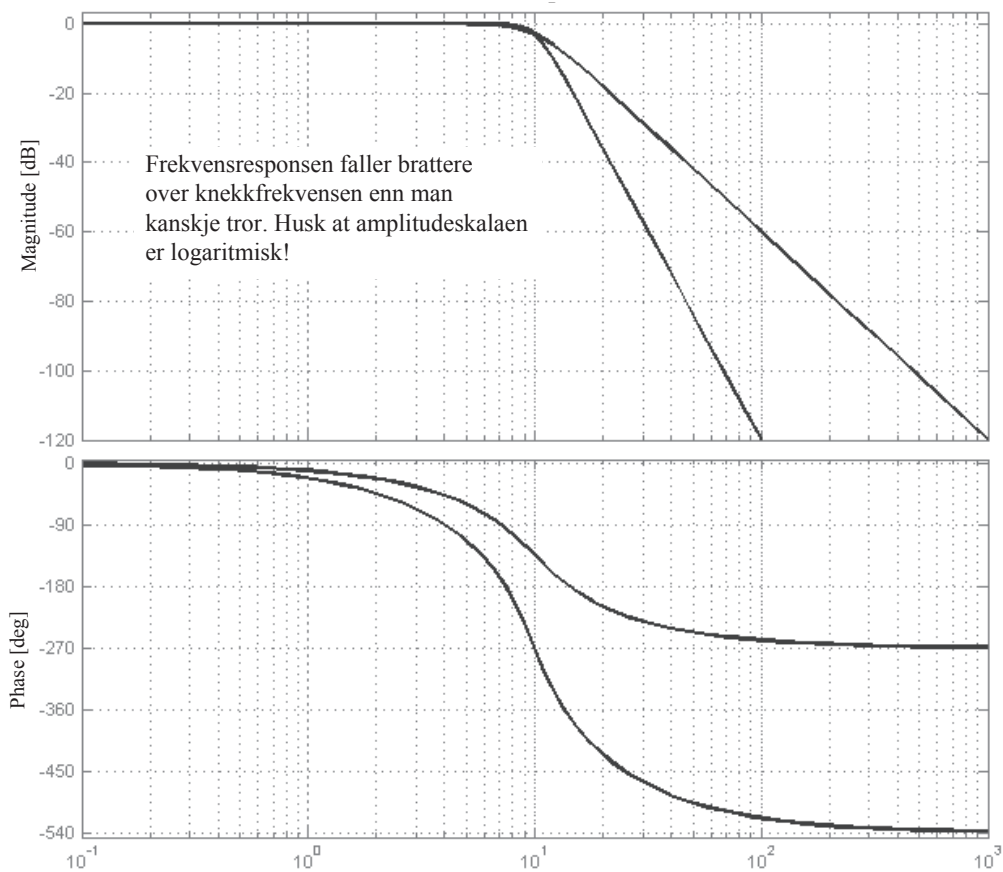
Den statiske forsterkinga for filteret er

$$h_L(s)|_{s=j0} = \frac{\omega_c^n}{(-\lambda_1)(-\lambda_2) \dots (-\lambda_n)} = 1 \quad (11.20)$$

Det siste følger av at alle polene ligger enten kompleks konjugert eller på den negative reelle akse uansett n , og at de alle har tallverdi ω_c . $h_L(j0) = 1$ er akkurat det vi krever av et lavpassfilter. Figur 11.13 viser Bode-diagram for BW-filteret med $n = 3$ og 6. Vi ignorerer faseforløpet her.

Figur 11.13

Frekvensrespons
for BW-filter med
 $n = 3, 6$ og $\omega_c = 10$



Merk at et førsteordens BW-filter, men $n = 1$, blir:

$$h_L(s) = \frac{\omega_c}{(s - \lambda_1)} = \frac{1/T}{s - (-1/T)} = \frac{1}{Ts + 1} \quad (11.21)$$

Når $n = 1$ er $\omega_c = 1/T$. Vi kjenner igjen den mye brukte første ordens transferfunksjonen. Den elektriske kretsen i eksempel 2.2 er et 1.ordens BW-filter. Se også figur 6.14 b). Men dette filteret har en frekvensrespons som faller bare 20dB/dekade i stopp-området, mens 3. og 6. ordens filtre faller med henholdsvis 60 og 120 dB/dekade, se figur 11.13. Vi ønsker brattest mulig overgang til stopp-båndet, og det er argumentet for å bruke høyere ordens filtre.

11.3.3 Høypassfilter fra Butterworth-filteret

Et høypassfilter kan vi få til ved å kreve at transferfunksjonen $h_H(s)$ (“H” = “høypass”) skal ha samme amplitudeforløp når ω går fra ω_c til ∞ , som $h_L(s)$ har når ω går fra ω_c til 0.

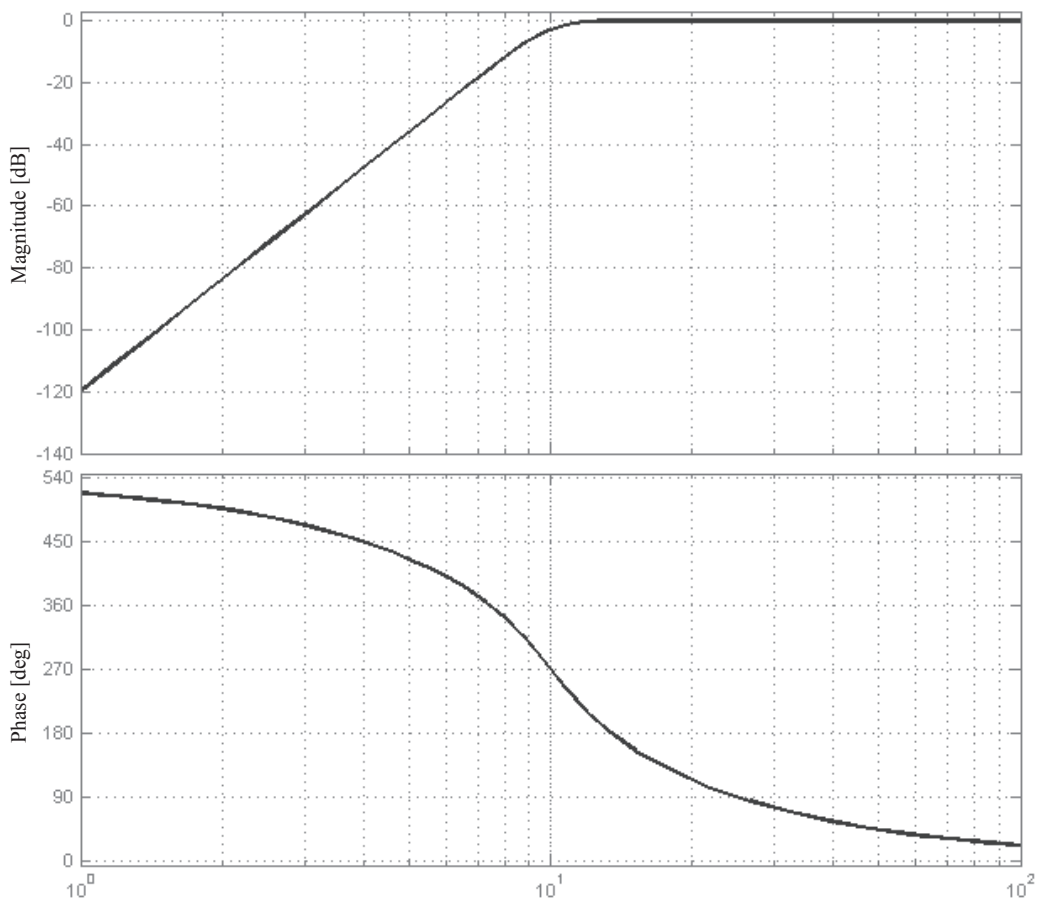
Det oppnår vi ved å erstatte alle s -er i (1.1) med ω_c^2/s :

$$\begin{aligned} h_H(s) &= \frac{\omega_c^n}{(\omega_c^2/s - \lambda_1)(\omega_c^2/s - \lambda_2)\dots(\omega_c^2/s - \lambda_n)} \\ &= \frac{s^n \omega_c^n}{(\omega_c^2 - s\lambda_1)(\omega_c^2 - s\lambda_2)\dots(\omega_c^2 - s\lambda_n)} \end{aligned} \quad (11.22)$$

Frekvensresponsen for et slikt høypassfilter med $n = 6$ er vist i figur 11.14:

Figur 11.14

Frekvensrespons for høypassfilter med $n = 6$ og $\omega_c = 10$



11.3.4 Flere lav-ordens filtre eller et med høy orden?

Nå er det slik at lavere ordens filtre er enklere å realisere enn direkte implementerte høyere ordens filtre. Det blir også mindre støy (altså “forurensing” som ligger oppå det korrekte signalet) ved flere lavere ordens filtre i serie. Det er derfor hensiktsmessig å koble f.eks. tre 2. ordens filtre i serie istedet for et direkte implementert 6. ordens filter.

11.3.5 Båndpass- og båndstopp-filtre

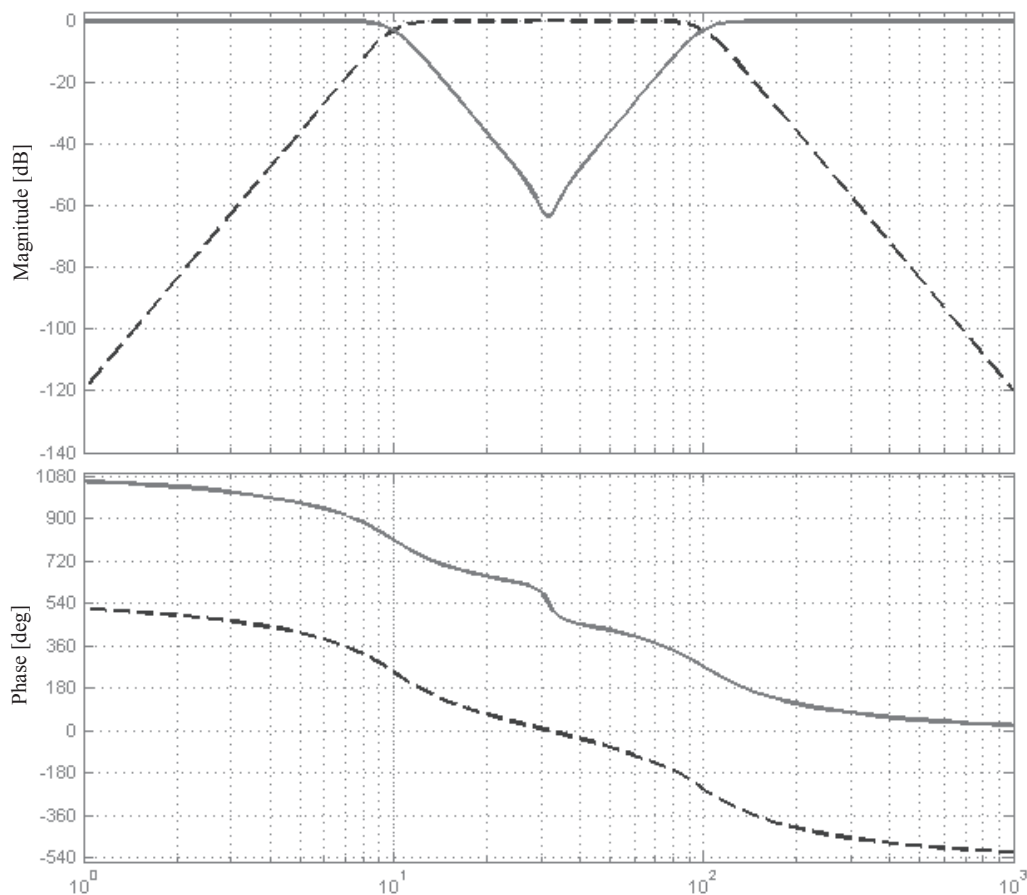
Slike kan man konstruere ved å kombinere lavpass- og høypass-filtre, som igjen er laget i følge oppskriftene ovenfor. Metoden forklares enklest med eksempler:

Vi skal først lage et **båndpassfilter** som slipper gjennom frekvensene mellom $\omega_L = 10$ og $\omega_H = 100$. Dette oppnås ved å koble et lavpassfilter med $\omega_c = 100$ i *serie* med et høypassfilter med $\omega_c = 10$. Merk den “motsatte” plasseringa av de to knekkfrekvensene.

Så til et **båndstoppfilter** som demper frekvensene mellom $\omega_L = 10$ og $\omega_H = 100$. Dette oppnås ved å koble et lavpassfilter med $\omega_c = 10$ i *parallel* med et høypassfilter med $\omega_c = 100$. Se figur 11.15, som viser båndpass (stiplet) båndstopp (heltrukket) for $n = 6$.

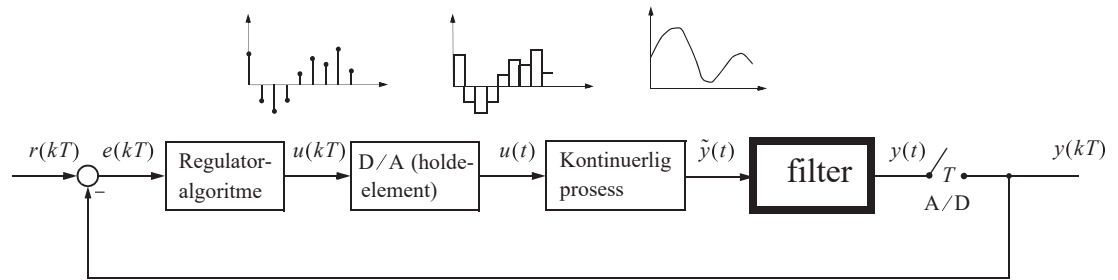
Figur 11.15

Frekvensresponser
for båndpass og
stopp-
filter med $n = 6$



11.3.6 Filtre i reguleringsstekniske anvendelser

Figur 11.16
Reguleringskrets
med filter



Figur 11.16 er en utvidelse av figur 11.1. Den ubehandlede målingen $\tilde{y}(t)$ fra den fysiske prosessen blir nå filtrert før den samples. Slikt kan være nyttig:

1. Når det er høyfrekvent *målestøy* overlappet målingen. Denne uttrykker ikke prosessens virkelige bevegelser og bør fjernes/dempes med et lavpassfilter. Heldigvis er målestøy ofte høyfrekvent i forhold til prosessens dynamikk, slik at filterets knekkfrekvens kan legges så høyt at dette ikke har forringende effekt på reguleringsløyfa.
2. Når det er bevegelser i prosessen man ikke ønsker å regulere bort, men ignorere. Et eksempel kan være bølger innafor et gitt frekvensområde, hvor man ikke ønsker å begrense bevegelsene i det aktuelle frekvensområde, men la fartøyet svinge med bølgene. Da kan et båndstoppfilter være nyttig.
3. Når målingen inneholder en konstant feil (“bias”), som man ikke ønsker å ha med som grunnlag for reguleringa. Da kan et høypassfilter være aktuelt. Dette er en problemstilling som ofte dukker opp i navigasjonssystemer.

Merk til slutt at *en hel reguleringsløyfe kan betraktes som et filter*. Anta at vi ønsker å fjerne stasjonært avvik ved hjelp av integralvirkning i regulatoren. Vi oppnår at transferfunksjonen

$$\frac{e}{v}(s) = -N(s)h_v(s) \quad (11.23)$$

får en frekvensrespons hvor $\left| \frac{e}{v}(0) \right| = 0$, dvs. null stasjonært avvik, og at $\left| \frac{e}{v} \right|$ er liten for lave frekvenser, men større etter hvert. Med andre ord så kan vi betrakte et reguleringsystem med passende integralvirkning som et høypassfilter som fjerner forstyrrelser som er konstante eller lavfrekvente.

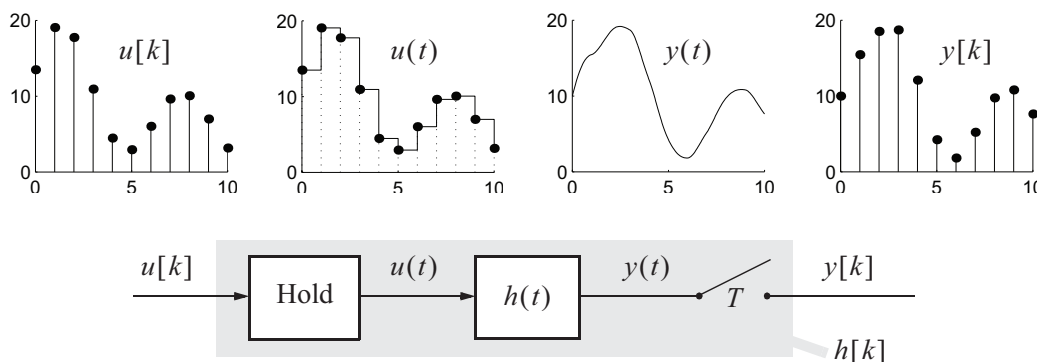
11.4 DISKRET ANALYSE OG SYNTSE – Z-TRANSFORMASJONEN

Vi har til nå omtalt metoder for å lage diskrete regulatorer som om de var kontinuerlige. Denne type tilnærming er ikke alltid tilstrekkelig. Nå skal vi presentere begreper og verktøy som trengs for å utføre matematisk eksakt analyse av kontinuerlige systemer som er diskretiserte.

11.4.1 Diskret signal og system

Figur 11.17 viser hvordan et kontinuerlig system kan betraktes som et diskret system, når virkningen av holdeelement og taster tas med: En tallfølge $u[k]$ på inngangen fører til en annen tallfølge $y[k]$ på utgangen. På samme måte som den kontinuerlige prosessen er gitt ved sin transferfunksjon $h(s)$ - eller ekvivalent - sin impulsrespons $h(t)$, så er det diskrete systemet gitt ved sin **impulsrespons** $h[k]$ (eller sin diskrete transferfunksjon $h(z)$, som vi kommer tilbake til i avsnitt 11.4.4).

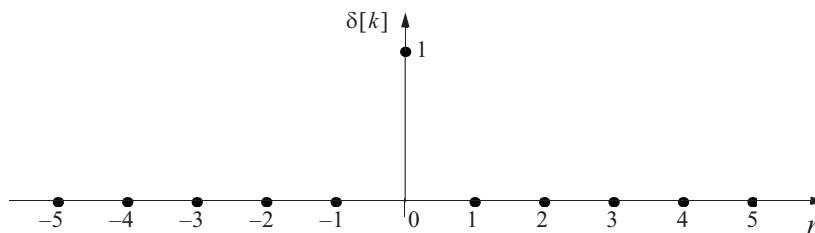
Figur 11.17
“Diskretisert”
kontinuerlig
LTI-system



Den diskrete **enhetspuls** $\delta[k]$, se også figur 11.18, er definert som:

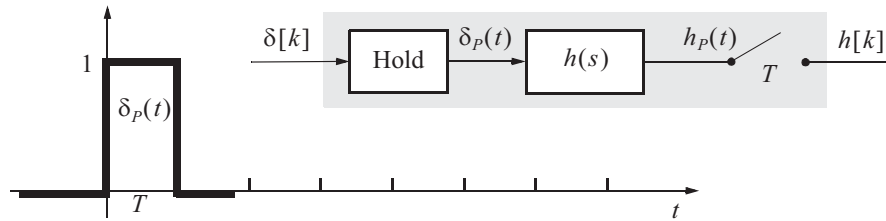
$$\delta[k] = \begin{cases} 1 & , k = 0 \\ 0 & , \text{ellers} \end{cases} \quad (11.24)$$

Figur 11.18
Diskret
enhetspuls $\delta[k]$



Hvis vi sender en diskret enhetspuls $\delta[k]$ inn på holdeelementet, vil dette igjen sende en firkantpuls $\delta_p[k]$ inn på det kontinuerlige systemet som vist på figur 11.19.

Figur 11.19
Firkantpuls fra
holdeelement



Den diskrete impulsresponsen $h[k]$ er da det tastede signal av utgangen $h_p(t)$ på den kontinuerlige prosessen. Et generelt diskret pådragsignal $u[k]$ kan betraktes som et tog av diskrete pulser. Hver slik pådragspuls gir opphav til en respons på utgangen. Vi kan dermed bruke superposisjonsprinsippet og skrive

$$\begin{aligned} y[n] &= u[0]h[n] + u[1]h[n-1] + \dots + u[n-1]h[1] + u[n]h[0] \\ &= \sum_{k=0}^n h[n-k]u[k] = h[n] * u[n] \end{aligned} \quad (11.25)$$

Likning (11.25) er **diskret folding**, og svarer til kontinuerlig folding som ble introdusert i avsnitt 2.4. Vi skal nå finne $h[k]$ når den kontinuerlige prosessen er monovariabel og gitt på tilstandsromformen

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u \\ \mathbf{y} &= \mathbf{c}^T \mathbf{x} \end{aligned} \quad (11.26)$$

Fra (3.34) har vi da

$$\begin{aligned} \mathbf{x}[k+1] &= \Phi(T)\mathbf{x}[k] + \tilde{\delta}(T)u[k] \\ y[k] &= \mathbf{c}^T \mathbf{x}[k] \end{aligned} \quad (11.27)$$

Hvor $\Phi = e^{\mathbf{A}T}$ og $\tilde{\delta} = \mathbf{A}^{-1}(\Phi(T) - \mathbf{I})\mathbf{b}$

Impulsresponsen $h[k]$ kan da finnes slik:

$$\begin{aligned} \mathbf{x}[0] &= \mathbf{0} \\ \mathbf{x}[1] &= \Phi\mathbf{x}[0] + \tilde{\delta} \cdot 1 = \tilde{\delta} \\ \mathbf{x}[2] &= \Phi\mathbf{x}[1] + \tilde{\delta}u[1] = \Phi\tilde{\delta} + \tilde{\delta} \cdot 0 = \Phi\tilde{\delta} \\ \mathbf{x}[3] &= \Phi^2\tilde{\delta} \\ &\vdots \\ \mathbf{x}[k] &= \Phi^{k-1}\tilde{\delta} \Rightarrow \\ h[k] &= \mathbf{c}^T \mathbf{x}[k] = \mathbf{c}^T \Phi^{k-1}\tilde{\delta} \end{aligned} \quad (11.28)$$

¹Akkurat her bruker vi tilde (~) for å unngå sammenblanding med “ $\delta[k]$ ” som jo representerer diskret enhetspuls.

Vi har dermed vist hvordan vi, når en kontinuerlig prosess (11.26) og tastetid T er gitt, kan finne den ekvivalente diskrete prosess, gitt ved sin impulsrespons $h[k]$.

(Hvis den kontinuerlige prosessen i stedet for (11.13) er gitt ved sin transferfunksjon $h(s)$, kan denne omgjøres til tilstandsrombeskrivelse ved for eksempel å bruke fasevariabel form, (4.82)).

11.4.2 z-transformasjon

Laplacetransformasjonen $f(s)$ for et kontinuerlig signal $f(t)$ er som kjent

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

For et diskret signal $f[k]$ kan vi innføre **Z-transformasjonen**¹ $f(z)$, der

$$f(z) = \sum_{k=0}^{\infty} f[k]z^{-k} \quad (11.29)$$

Z-transformasjonen er “tvillingen” til Laplace-transformasjonen for diskrete signaler og systemer. z er en kompleks variabel på samme måte som s . Vi snakker derfor om “z-planet”, analogt med bruken av “s-planet” for det kontinuerlige tilfellet.

Z-transformasjonen har en rekke nyttige egenskaper som svarer til slike som Laplace-transformasjonen har for det kontinuerlige tilfellet². Vi skal bare ta med de aller viktigste her:

1. Linearitet

$$Z(af_1[k] + bf_2[k]) = af_1(z) + bf_2(z) \quad (11.30)$$

2. Tidsforskyvning

$$Z(f[k+1]) = zf(z) - zf[k=0] \quad (11.31)$$

3. Reell folding

$$Z\left(\sum_{k=0}^n f_1[n-k]f_2[k]\right) = Z(f_1[n] * f_2[n]) = f_1[z]f_2[z] \quad (11.32)$$

¹Merk at vi også her bruker små bokstaver *både* i tidsplanet og i z-planet, dvs. $f[k]$ og $f(z)$, akkurat som med $f(t)$ og $f(s)$. Fordelen i det diskrete tilfellet er at parentesene hjelper oss å skille $f[k]$ og $f(z)$ fra hverandre.

²Egentlig er (11.29) den *ensidige* Z-transformasjonen fordi den bare gjelder for $0 \leq k < \infty$, ikke $k < 0$. Den svarer til den ensidige Laplace-transformasjonen som vi bruker i denne læreboka. Det finnes dobbeltsidige versjoner av både L- og Z-transformasjonene.

4. Sluttverdi

$$\lim_{k \rightarrow \infty} f[k] = \lim_{z \rightarrow 1} (z-1)f(z) \quad (11.33)$$

5. Begynnelsesverdi

$$f[0] = \lim_{z \rightarrow \infty} f(z) \quad (11.34)$$

Vi skal ikke føre bevis for alle disse egenskapene; (11.30) og (11.34) er for eksempel helt trivielle å bevise ut fra definisjonen (11.29).

Vi nøyer oss med to bevis:

Tidsforskyvning, (11.31):

$$\begin{aligned} Z(f[k+1]) &= \sum_{k=0}^{\infty} f[k+1]z^{-k} = \sum_{m=1}^{\infty} f[m]z^{-(m-1)} \\ &= \sum_{m=0}^{\infty} f[m]z^{-m+1} - f[0]z = zf(z) - zf[0] \end{aligned}$$

▲

Reell folding, (11.32):

$$\begin{aligned} Z\left(\sum_{k=0}^{\infty} f_1[n-k]f_2[k]\right) &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} f_1[n-k]f_2[k]\right) z^{-n} \\ &= \sum_{k=0}^{\infty} f_2[k] \left(\sum_{n=0}^{\infty} f_1[n-k]z^{-(n-k)}\right) z^{-k} \\ &= \sum_{k=0}^{\infty} f_2[k]z^{-k} \sum_{m=-k}^{\infty} f_1[m]z^{-m} \\ &= f_2(z) \sum_{m=0}^{\infty} f_1[m]z^{-m} = f_2[z]f_1[z] \end{aligned}$$

Her har vi kunnet endre på summasjonsgrensene siden både $f_1[n] = 0$ og $f_2[n] = 0$ for $n < 0$.

▲

11.4.3 Invers z-transformasjon

Den inverse z-transformasjonen¹ er

$$f[k] = \frac{1}{2\pi j} \oint_{|z|=r} f(z)z^{k-1} dz \quad (11.35)$$

Der integrasjonen er en sirkel som omslutter alle integrandens poler.

Bevis:

Vi setter inn (11.29) i høyresiden i (11.35), og er i mål hvis vi kan vise at den blir lik $f[k]$:

$$\begin{aligned} & \frac{1}{2\pi j} \oint \left(\sum_{n=0}^{\infty} f[n]z^{-n} dz \right) z^{k-1} dz = \sum_{n=0}^{\infty} \frac{1}{2\pi j} \oint f[n]z^{k-n-1} dz = \\ & \dots + \frac{1}{2\pi j} f[k-2] \oint z dz + \frac{1}{2\pi j} f[k-1] \oint dz + \frac{1}{2\pi j} f[k] \oint \frac{1}{z} dz + \frac{1}{2\pi j} f[k+1] \oint \frac{1}{z^2} dz + \dots \end{aligned}$$

Her er $f[i]$, $i = \dots, k, k+1, \dots$ konstante m.h.p. integralene, og derfor satt utenfor integraltegnene. Vi ser at alle ledd med $i < k$ er analytiske innenfor enhver sirkel $|z| = r$, og da blir disse leddene lik null. Alle ledd med $i > k$ får en pol med multiplisitet > 1 i integranden:

$$\frac{dz}{z^2}, \frac{dz}{z^3}, \dots$$

Disse integralene kan løses ved residuegning og blir

$$I_{i-k} = \frac{\partial^{i-k}}{\partial z^{i-k}} \left\{ \frac{z^{i-k} \cdot 1}{z^{i-k}} \right\} \Bigg|_{z=0} = \begin{cases} 1 & \text{for } i = k \\ 0 & \text{for } i > k \end{cases}$$

Dermed blir bare ett eneste ledd $\neq 0$, nemlig leddet $f[k]I_{k-k} = f[k]$.

▲

¹Merk likheten med invers Laplace-transformasjon, likning (4.9).

11.4.4 Transferfunksjon i z

Fra (11.27) og (11.31) har vi, når vi z-transformerer på begge sider:

$$\begin{aligned} z\mathbf{x}(z) - z\mathbf{x}[0] &= \Phi\mathbf{x}(z) + \tilde{\delta}u(z) \Rightarrow \\ \mathbf{x}(z) &= z(z\mathbf{I} - \Phi)^{-1}\mathbf{x}[0] + (z\mathbf{I} - \Phi)^{-1}\tilde{\delta}u(z) \end{aligned}$$

Hvis vi forutsetter $\mathbf{x}(0) = \mathbf{0}$, finner vi den **diskrete transferfunksjonen** $h(z)$, gitt ved

$$y(z) = h(z)u(z)$$

der¹

$$h(z) = \mathbf{c}^T (z\mathbf{I} - \Phi)^{-1} \tilde{\delta}$$

eller

$$\begin{aligned} h(z) &= \mathbf{c}^T \frac{\text{adj}(z\mathbf{I} - \Phi)}{\det(z\mathbf{I} - \Phi)} \tilde{\delta} = \frac{b_1 z^{n-1} + \dots + b_n}{z^n + a_1 z^{n-1} + \dots + a_n} \\ &= \frac{b_1 z^{-1} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} \end{aligned} \quad (11.36)$$

Vi har altså funnet et rasjonalt uttrykk for den diskrete transferfunksjonen $h(z)$. Den andre varianten brukes oftest, og den skriver vi som $h(z^{-1})$. Den inverse z-transformen til $h(z)$ er den diskrete impulsresponsen $h[k]$. Vi kan med andre ord finne $h(z)$ som z-transformen til $h[k]$, men det er vanligvis unødig tungvint.

Det er nå lett å finne løsningen i diskret tid, når pådraget $u[k]$ er kjent. Vi multipliserer ut den andre varianten av (11.36) og finner

$$(1 + a_1 z^{-1} + \dots + a_n z^{-n})y(z) = (b_1 z^{-1} + \dots + b_n z^{-n})u(z) \quad (11.37)$$

og derav

$$y(z) = -(a_1 z^{-1}y(z) + \dots + a_n z^{-n}y(z)) + b_1 z^{-1}u(z) + \dots + b_n z^{-n}u(z) \quad (11.38)$$

Vi benytter tidsforskyvingsegenskapen og inverstransformerer ledd for ledd, og får

$$y[k] = -a_1 y[k-1] - \dots - a_n y[k-n] + b_1 u[k-1] + \dots + b_n u[k-n] \quad (11.39)$$

¹Denne utledningen er analog til den for $h(s)$ i avsnitt 4.6.

Dette betyr at den nåværende verdien av prosessens utgang ($y[k]$) kan beskrives som en lineær kombinasjon av n tidligere verdier av utgangen og n tidligere verdier av pådraget. Uttrykket i (11.39) egner seg derfor godt for utregning ved hjelp av en datamaskin som lagrer $2n$ gamle verdier og multipliserer disse med tilsvarende koeffisienter.

Se ellers eksempel 11.13.

11.4.5 Stabilitet for et diskret system

Vi tar utgangspunkt i (11.27). For å studere stabilitet, er det tilstrekkelig å betrakte det autonome systemet

$$\mathbf{x}[k+1] = \Phi \mathbf{x}[k] \quad , \quad \mathbf{x}[0] = \mathbf{x}_0 \quad (11.40)$$

Fra (3.13) har vi (vi antar distinkte egenverdier i matrisen \mathbf{A}):

$$\Lambda = \mathbf{M}^{-1} \mathbf{A} \mathbf{M} = \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & & & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix} \quad (11.41)$$

Likning (3.17) og (3.21) gir videre

$$\mathbf{e}^{\Lambda T} = \mathbf{M}^{-1} \mathbf{e}^{\mathbf{A}T} \mathbf{M} = \mathbf{M}^{-1} \Phi(T) \mathbf{M} = \begin{bmatrix} e^{\lambda_1 T} & 0 & \cdot & \cdot & 0 \\ 0 & e^{\lambda_2 T} & & & 0 \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & e^{\lambda_n T} \end{bmatrix} \quad (11.42)$$

Egenvektormatrisen (\mathbf{M}) er den samme i begge tilfeller. Fra (11.42) har vi, når vi setter $\mathbf{x} = \mathbf{M}\mathbf{q}$:

$$\begin{aligned} \mathbf{M}\mathbf{q}[k+1] &= \Phi(T) \mathbf{M}\mathbf{q}[k] \Rightarrow \\ \mathbf{q}[k+1] &= \mathbf{e}^{\Lambda T} \mathbf{q}[k] \quad , \quad \mathbf{q}[0] = \mathbf{M}^{-1} \mathbf{x}_0 \end{aligned}$$

det vil si på komponentform

$$\begin{aligned} q_i[k+1] &= q_i[k] e^{\lambda_i T} \\ &= q_i[k] r_i e^{j\varphi_i} \end{aligned} \quad (11.43)$$

hvor $r_i = |e^{\lambda_i T}|$ og $\varphi_i = \angle e^{\lambda_i T}$.

Kravet for asymptotisk stabilitet blir i følge (11.42) og (11.43) at alle egenverdier ρ_i til Φ , $\rho_i = e^{\lambda_i T}$, må ha $|\rho_i| = r_i < 1$, det vil si egenverdiene til Φ må alle ligge innenfor enhets sirkelen.

Nevneren i transferfunksjonen $h(z)$ blir

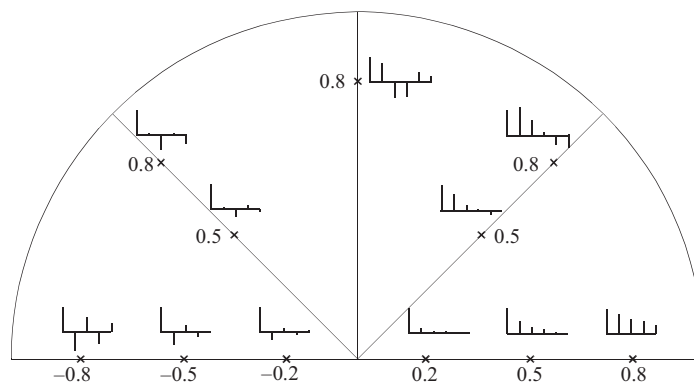
$$|z\mathbf{I} - \Phi| = (z - \rho_1) \dots (z - \rho_n) = (z - e^{\lambda_1 T}) \dots (z - e^{\lambda_n T}) \quad (11.44)$$

på samme måte som nevneren i $h(s)$ er

$$|s\mathbf{I} - \mathbf{A}| = (s - \lambda_1) \dots (s - \lambda_n) \quad (11.45)$$

Figur 11.20 viser innsvingningsforløpet for $q[k]$ avhengig av om man har en pol på den reelle akse eller to kompleks konjugerte poler symmetrisk plassert rundt den reelle akse. Bare øvre halvdel av z -planet er vist på grunn av symmetrien.

Figur 11.20
Innsvingningsforløp avhengig av egenverdiens plassering



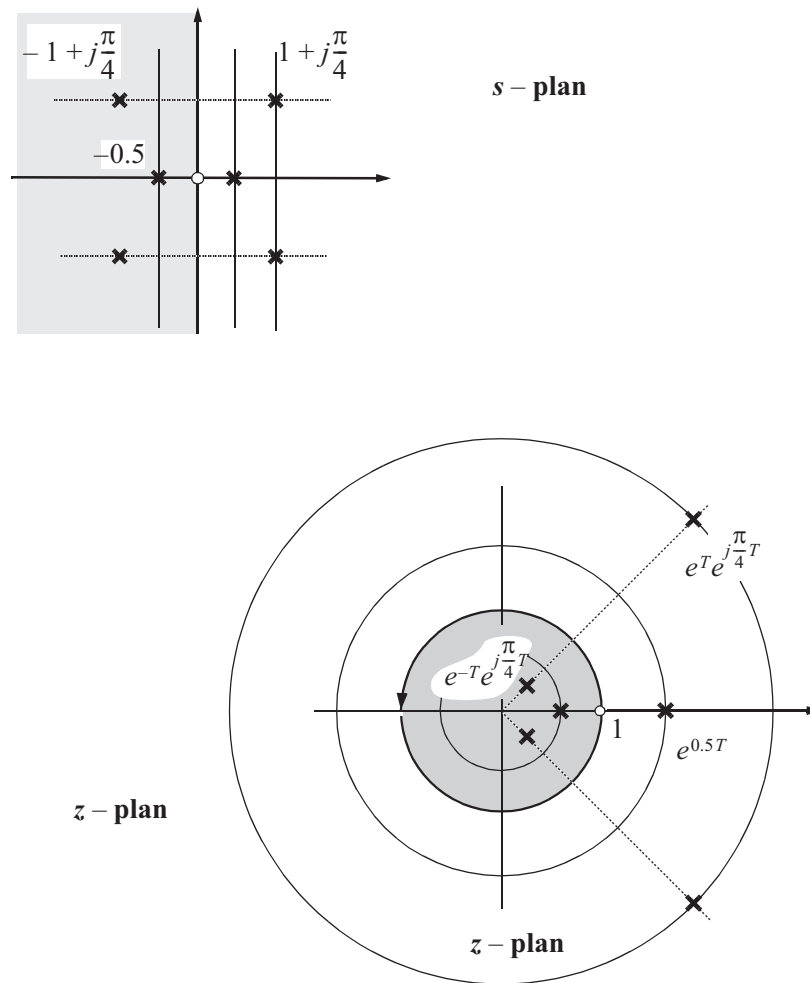
Hurtig innsvingning oppnås når egenverdien ligger nær origo, mens dempningen blir liten jo nærmere egenverdien kommer enhets sirkelen. Egenverdier i venstre del av enhets sirkelen gir den mest oscillerende innsvingningen.

Polene λ_i i det kontinuerlige systemet fører ved diskretisering til polene $\rho_i = e^{\lambda_i T}$ i z -planet. Matematisk kan vi uttrykke dette som at venstre halvplan i s -planet *avbildes* innenfor enhets sirkelen i z -planet, mens høyre halvplan avbildes utenfor. Se figur 11.21.

Linjer parallelle med den imaginære akse i s -planet svarer til konsentriske sirkler om origo i z -planet. Linjer parallelle med den reelle akse svarer til stråler ut fra origo. Merk at linjene i rutenettet fortsatt krysser hverandre med 90° vinkel unntatt i det singulære punktet origo. Dette er et eksempel på en **konform avbildning**, hvor vinkler forblir de samme ved avbildningen.

I figur 11.21 sammenliknes polene i s -planet og z -planet for en gitt T . Det er viktig å merke seg at polenes plassering i z -planet avhenger av tastetida T , for en gitt kontinuerlig prosess.

Figur 11.21
Poler og nullpunkt-
er i s -plan, og i z -
plan for det tilsva-
rende diskretiserte
system



11.4.6 Frekvensrespons for diskrete systemer

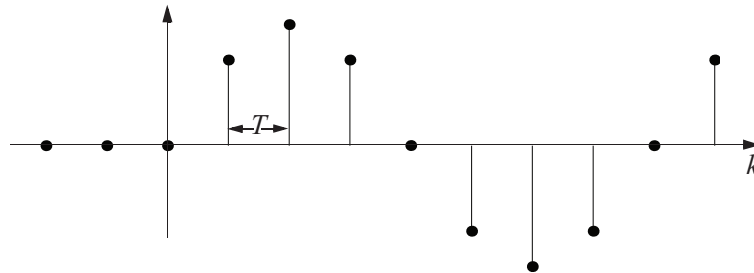
Anta at et diskretisert system er gitt ved sin diskrete transferfunksjon $h(z)$. Vi forutsetter nå at systemet er asymptotisk stabilt, det vil si at alle poler i $h(z)$ er innenfor enhets sirkelen. Vi skal nå utlede frekvensresponsen til systemet, og vil gjøre dette på en måte som er helt parallell til det kontinuerlige tilfellet (avsnitt 6.2).

Vi tenker oss at systemet utsettes for et signal

$$u[k] = \mu_1[k] \sin(\omega kT) \quad (11.46)$$

Figur 11.22 viser signalet¹. I figuren er $\omega = \frac{\pi}{4T}$, men ω kan ha enhver verdi $0 \leq \omega \leq \frac{\pi}{T}$.

Figur 11.22
Diskret signal



Z-transformasjon av (11.46) gir

$$\begin{aligned} u(z) &= \sum_{k=0}^{\infty} u[k] z^{-k} = \sum_{k=0}^{\infty} \frac{(e^{j\Omega k} - e^{-j\Omega k})}{2j} z^{-k} \\ &= \frac{1}{2j} \left(\sum_{k=0}^{\infty} (e^{j\Omega} z^{-1})^k - \sum_{k=0}^{\infty} (e^{-j\Omega} z^{-1})^k \right) \end{aligned} \quad (11.47)$$

Her har vi innført $\Omega = \omega T$. Ω er proporsjonal med ω , og er en skalert, **dimensjonsløs frekvens** som svarer til en skalert, **dimensjonsløs tastetid** $T = 1$. Ved å bruke Ω i stedet for ω , blir vår beskrivelse av diskrete signaler og systemer den samme, uansett fysisk tastetid. For å komme tilbake til fysisk frekvens, bruker vi sammenhengen $\omega = \Omega/T$.

De to leddene i (11.47) kan uttrykkes som to geometriske rekker, når vi forutsetter $|z| > 1$:

$$\begin{aligned} u(z) &= \frac{\frac{1}{2j}}{1 - e^{j\Omega} z^{-1}} - \frac{\frac{1}{2j}}{1 - e^{-j\Omega} z^{-1}} \\ &= \frac{(e^{j\Omega} - e^{-j\Omega})}{2j} z^{-1} \\ &= \frac{(\sin \Omega) z}{(1 - e^{j\Omega} z^{-1})(1 - e^{-j\Omega} z^{-1})} = \frac{(\sin \Omega) z}{(z - e^{j\Omega})(z - e^{-j\Omega})} \end{aligned} \quad (11.48)$$

¹ $\mu_1[k]$ er den **diskrete enhetsprangfunksjonen**. Her brukes den for å indikere $u[k] = 0$ for $k < 0$.

Fra (11.32) har vi $y(z) = h(z)u(z)$, og $y[k] = Z^{-1}(y(z))$.

$$y[k] = \frac{1}{2\pi j} \oint_{|z|=r} h(z)u(z)z^{k-1} dz = \sum \operatorname{res}(h(z)u(z)z^{k-1})|_{z=z_i} \quad (11.49)$$

der z_i er polene til $h(z)u(z)z^{k-1}$. Nå er vi interessert i responsen for “store” k , det vil si når de transiente komponentene har dødd ut, noe som vil skje fordi $h(z)$ er stabil. Da kan vi ignorere residuene som følger av polene i $h(z)$, og står igjen med de to residuene på grunn av polene i $u(z)$. Likning (11.48) gir da:

$$\begin{aligned} y[k] &= \operatorname{res}(h(z)u(z)z^{k-1})|_{z=e^{j\Omega}} + \operatorname{res}(h(z)u(z)z^{k-1})|_{z=e^{-j\Omega}} = \\ &= \left. \frac{(\sin \Omega)h(z)z^k}{(z - e^{-j\Omega})} \right|_{z=e^{j\Omega}} + \left. \frac{(\sin \Omega)h(z)z^k}{(z - e^{j\Omega})} \right|_{z=e^{-j\Omega}} = \frac{e^{j\Omega} - e^{-j\Omega}}{e^{j\Omega} - e^{-j\Omega}} h(e^{j\Omega})e^{j\Omega k} - \frac{e^{j\Omega} - e^{-j\Omega}}{e^{j\Omega} - e^{-j\Omega}} h(e^{-j\Omega})e^{-j\Omega k} \end{aligned}$$

Vi setter $h(e^{j\Omega}) = |h(e^{j\Omega})|e^{j\angle h(e^{j\Omega})}$ og benytter¹ at $h(e^{-j\Omega}) = h^*(e^{+j\Omega})$, det vil si $|h(e^{-j\Omega})| = |h(e^{j\Omega})|$ og $\angle h(e^{-j\Omega}) = -\angle h(e^{j\Omega})$:

$$y[k] = |h(e^{j\Omega})| \frac{e^{j(\Omega k + \angle h(e^{j\Omega}))} - e^{-j(\Omega k + \angle h(e^{j\Omega}))}}{2j} = |h(e^{j\Omega})| \sin(\Omega k + \angle h(e^{j\Omega})) \quad (11.50)$$

Vi ser altså at det diskrete utgangssignalet blir en sinus med samme frekvens, men med amplitude og faseforskyvning gitt av henholdsvis $|h(e^{j\Omega})|$ og $\angle h(e^{j\Omega})$.

Med andre ord: Vi finner frekvensresponsen for et diskret system ved å substituere $z = e^{j\Omega} = e^{j\omega T}$ i $h(z)$. Dette er analogt med at frekvensresponsen for et kontinuerlig system finnes ved å substituere $s = j\omega$ i $h(s)$.

EKSEMPEL 11.7: Frekvensresponsen for 1. ordens prosess

Gitt den kontinuerlige prosessen

$$h(s) = \frac{1}{1 + T_1 s} \quad \text{med } T_1 = 4.48 \quad (11.51)$$

¹ Skisse til bevis: Sett opp et rasjonalt uttrykk i z , erstatt alle z med $e^{j\Omega}$. Lag et annet uttrykk, hvor alle $e^{j\Omega}$ er erstattet med $(e^{j\Omega})^* = e^{-j\Omega}$. Polynomene i teller og nevner blir da konjugerte, og kvotienten mellom konjugert teller og nevner er også konjugert, dvs. de to uttrykkene blir konjugerte av hverandre.

Proessen¹ har fase- og amplitudeforløp som vist ved de stiplede linjene i figur 11.23 og 11.24. Figur 11.23 viser responsene i lineær skala langs begge aksene. Figur 11.24 er en mer gjenkjennelig versjon (Bodediagram) siden både frekvens og amplitude er i logaritmisk skala.

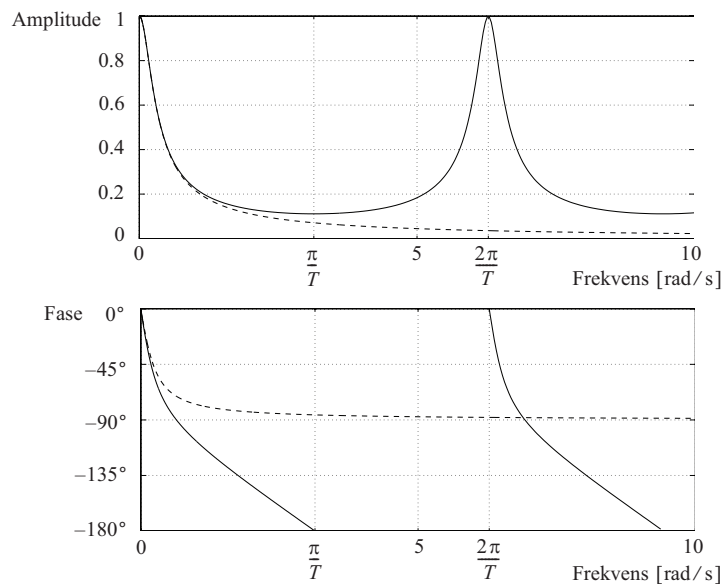
Vi diskretiserer prosessen med tastetid $T = 1$ og $T_1 = 4.48$ og får

$$h(z) = \frac{(1 - e^{-T/T_1})}{z - e^{-T/T_1}} = \frac{0.2}{z - 0.8} \quad (11.52)$$

Fase- og amplitudeforløpet for frekvensresponsen $h(z = e^{j\omega T})$ er vist med heltrukket linje i figur 11.23 og 11.24. Vi merker oss at overensstemmelsen mellom $h(s = j\omega)$ og $h(z = e^{j\omega T})$ er best ved lave frekvenser. Dette er rimelig, da innflytelsen til holde-elementet på dynamikken blir minst når inngangssignalet varierer langsomt.

Figur 11.23

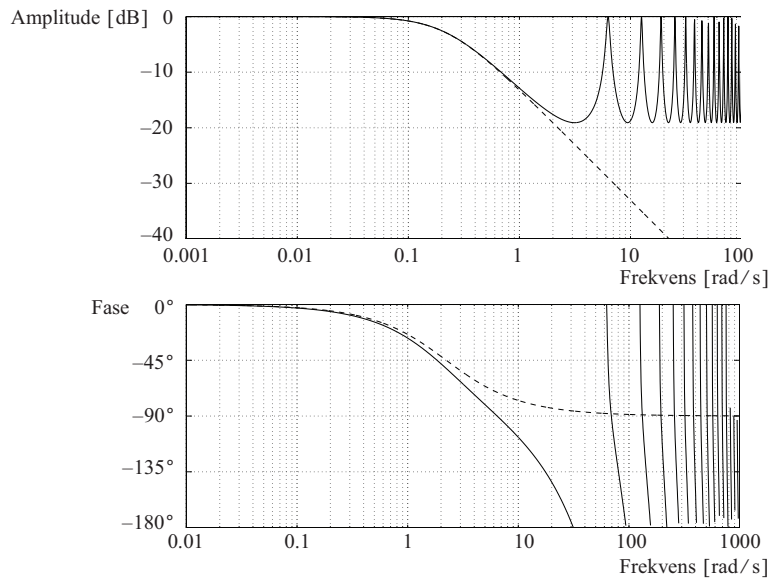
$h(s)|_{s=j\omega}$ og
 $h(z)|_{z=e^{j\omega T}}$
 med lineær
 frekvensskala



¹ $T_1 = 4.48$ og $T = 1$, valgt slik her for å få runde tall i $h(z)$, det gir $e^{-T/T_1} = 0.8$.

Den dramatiske forskjellen inntreffer ved høyere frekvenser. Vi ser at $h(e^{j\omega T})$ blir periodisk med periode $2\pi/T$. Dette sees tydeligst i figur 11.23 hvor frekvensaksen er lineær. Det ligger ikke noen ny informasjon om $h(e^{j\omega T})$ i området mellom $\omega = \pi/T$ og $\omega = 2\pi/T$, fordi $|h|$ er symmetrisk om $\omega = \pi/T$ og $\angle h$ er antisymmetrisk om $\omega = \pi/T$. Dette er en følge av samplingsteoremet, som behandles i avsnitt 11.4.7.

Figur 11.24
 $h(s)|_{s=j\omega}$ og
 $h(z)|_{z=e^{j\omega T}}$
 med logaritmisk
 frekvensskala



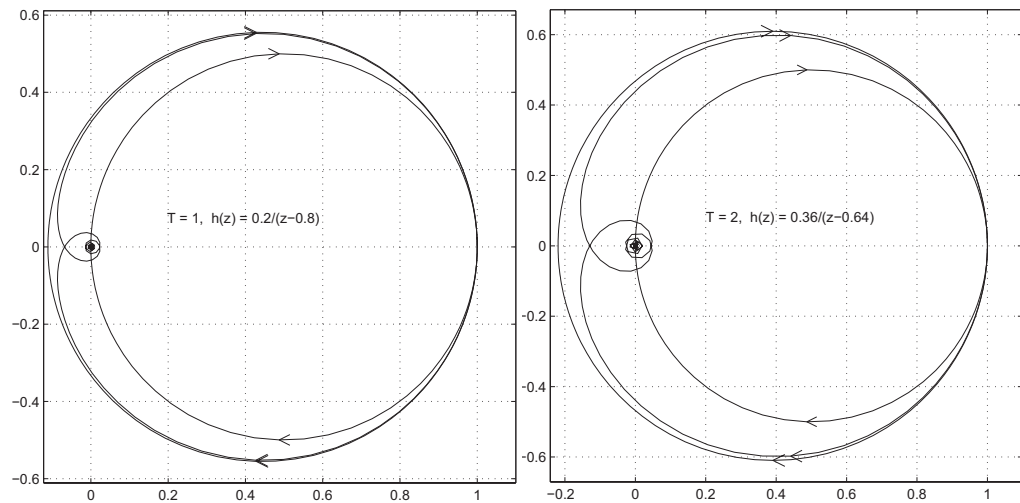
EKSEMPEL 11.8: Nyquists stabilitetskriterium for diskret system

Vi betrakter fortsatt prosessen (11.51). Figur 11.25 viser, for to verdier av T , Nyquist-diagrammet for den kontinuerlige prosessen (den innerste kurven), den kontinuerlige prosessen med en transportforsinkelse i serie lik $T/2$ (midterste kurve), og frekvensresponsen for det diskretiserte systemet (ytterste kurve). Den innerste og den ytterste kurven i venstre del av figur 11.25, er den polare representasjonen av hhv. det stiplede og heltrukne forløp i diagrammene i figurene 11.23 og 11.24. I høyre del av figur 11.25 har vi gått noe opp i tastetid.

Den midtre kurven i begge delfigurer viser hva som skjer når vi bruker metoden i avsnitt 11.2, å analysere som om systemet var kontinuerlig, men med en kunstig transportforsinkelse for å ta hensyn til tidsforsinkelsen som oppstår p.g.a. holdeelementet.

Vi kan undersøke stabiliteten for diskrete systemer med Nyquists stabilitetskriterium, akkurat som for kontinuerlige systemer.

Figur 11.25
Nyquist-
diagrammer for
kontinuerlig og
diskretisert
prosess, to
forskjellige
tastetider



Betrakt det aktuelle eksemplet: Hvis vi har proporsjonalregulering og *kontinuerlig* regulering, vet vi at dette systemet ikke kan bli ustabil. Dette Leser vi ut av Nyquist-diagrammet ved at den innerste kurven aldri kan omslutte det kritiske punkt -1 , uansett hvor stor forsterkning vi har. Kurven vil for økende K_p bare vokse ut mot høyre, men aldri gå inn i v.h.p.

Ser vi på forløpet med transportforsinkelse, får vi bekreftet det vi vet, at en tilstrekkelig stor K_p vil gi ustabilitet fordi det kritiske punkt da vil bli omsluttet.

De ytre kurvene¹ gir også åpenbart ustabilitet for stor nok K_p , og forsterkningsmarginen er noe mindre enn den vi får ved den tilnærmede analysen med transportforsinkelse. Høyre delfigur viser også at forsterkningsmarginen går ned når T økes, som forventet.

Den ytre kurven er frekvensforløpet av en *diskret* transferfunksjon,

$$h(z) = \frac{(1 - e^{-T/T_1})}{z - e^{-T/T_1}}, \text{ hvor } z = e^{j\omega T} \quad (11.53)$$

¹For dette 1. ordens systemet, vil Nyquistkurvene både for $h(s)$ og $h(z)$ bli *sirkler*. Sentrum for sirkelen for $h(z)$, vil ligge på den reelle akse, i $a/(1+a)$, der $a = e^{-T/T_1}$. Radius vil bli $1/(1+a)$. $h(s)$ og $h(z)$ er *konforme avbildninger* (inngår i kompleks funksjonsteori). Den spesielle klasse av konforme avbildninger som kan uttrykkes som $y = (az+b)/(cz+d)$, har den egenskap at rette linjer eller sirkler alltid avbildes som rette linjer eller sirkler. Dette forklarer sirklene både i det kontinuerlige og diskrete tilfelle i vårt eksempel. Men merk at dette bare gjelder for 1. ordens systemer. Det har heller ikke spesiell reguleringsteknisk betydning.

Hvordan har det seg at vi kan bruke Nyquists stabilitetskriterium også for denne?

– Den komplekse variable z beveger seg rundt enhets sirkelen i positiv dreieretning, og dette gir kurven for $h(z)$. Men enhets sirkelen er akkurat den banen som utgjør skillet mellom stabile og ustabile poler for et diskret system. Dermed spiller z langs enhets sirkelen ved stabilitetsundersøkelse av et diskret system, samme rolle som s gjør når den gjennomløper den imaginære akse ved stabilitetsundersøkelse av et kontinuerlig system. Av dette følger at Nyquists stabilitetskriterium kan brukes på samme måte for diskrete systemer, som det kan for kontinuerlige.



11.4.7 Tasteteoremet (“Samplingsteoremet”)

Anta at vi har et kontinuerlig sinusformet signal $x(t) = \sin \omega t$. Vi lager en diskret utgave av dette signalet ved å taste det ved tidspunktene $\dots, (k-1)T, kT, (k+1)T, \dots$

$$x[k] = \sin(\omega kT)$$

Tilfelle 1:

Anta at vi velger T slik at $\pi/T < \omega < 2\pi/T$. Vi definerer en ny frekvens $\bar{\omega} = 2\pi/T - \omega$. Da blir $0 < \bar{\omega} < \pi/T$. Vi har $\omega = 2\pi/T - \bar{\omega}$ og dermed

$$x[k] = \sin\left(\left(\frac{2\pi}{T} - \bar{\omega}\right)kT\right) = \sin(2\pi k - \bar{\omega}kT)$$

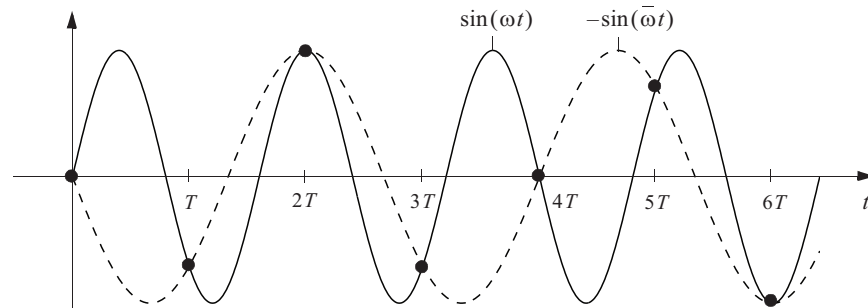
Da har vi

$$\begin{aligned} x[k] &= \sin(2\pi k) \cdot \cos(\bar{\omega}kT) - \sin(\bar{\omega}kT) \cdot \cos(2\pi k) \\ &= -\sin(\bar{\omega}kT) \end{aligned}$$

Med andre ord: Etter tasting vil en kontinuerlig funksjon $\sin(\omega kT)$ ikke være mulig å skille fra en annen tastet funksjon $-\sin(\bar{\omega}kT)$, når $\pi/T < \omega < 2\pi/T$. Dette er illustrert i figur 11.26, hvor $T = 1$, $\omega = 5\pi/4$ og $\bar{\omega} = 3\pi/4$.

Figur 11.26

To sinussignaler som er identiske i tastetidspunktene

**Tilfelle 2:**

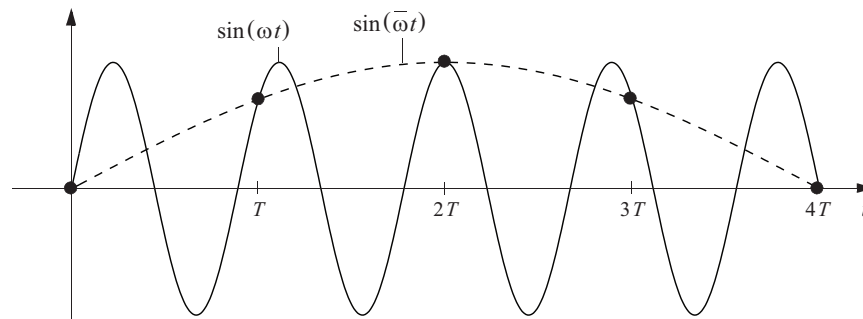
For $2\pi < \omega < 3\pi/T$ kan vi definere $\bar{\omega} = \omega - 2\pi/T$ slik at $\bar{\omega}$ igjen oppfyller $0 < \bar{\omega} < \pi/T$. Da har vi

$$x[k] = \sin\left(\left(\bar{\omega} + \frac{2\pi}{T}\right)kT\right) = \sin(\bar{\omega}kT + 2\pi k) = \sin(\bar{\omega}kT)$$

Nå er det etter tasting ikke mulig å skille mellom $\sin(\omega t)$ og $\sin(\bar{\omega} t)$. Se figur 11.27, hvor $T = 1$, $\omega = 9\pi/4$ og $\bar{\omega} = \pi/4$.

Figur 11.27

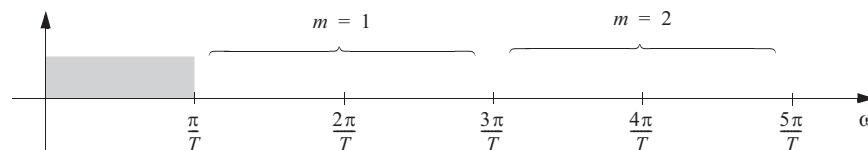
To andre sinus-signaler som er identiske i tastetidspunktene



Disse to tilfellene kan generaliseres til å gjelde alle frekvenser $(2m-1)\pi/T < \omega < 2m\pi/T$ og $2m\pi/T < \omega < (2m+1)\pi/T$, der $m = 1, 2, \dots$. Se figur 11.28.

Figur 11.28

Frekvensakse



Til nå har vi bare tatt for oss tasting av rene sinusformede signaler. Men i følge Fourier-analysen kan vi betrakte et vilkårlig signal som bestående av en uendelig sum av sinusfunksjoner med innbyrdes forskjellig amplitude og fase. Dermed kan vi konkludere med at:

Kontinuerlige signaler som inneholder frekvenskomponenter over π/T ikke kan representeres korrekt etter tasting, fordi slike frekvenskomponenter vil framstå som ikke-eksisterende (falske) frekvenskomponenter *under* π/T . Vi må altså passe på å velge T så liten at dette ikke skjer.

Dette er **tasteteoremet (samplingsteoremet)**. Frekvensen $\omega_s = 2\pi/T$ kalles **tastefrekvensen**. For at et kontinuerlig signal skal kunne representeres korrekt, må alle frekvenskomponenter i det kontinuerlige signalet være $< \omega_s/2$.

Tasteteoremet forklarer noe som skjer ved polenes forflytning i z -planet med økende T , for et diskretisert system. Betrakt figur 11.21, nedre del. Øk T til polene møtes på den negative reelle akse. Hva blir da T ? Har det noen mening å øke T ytterligere? Stemmer din konklusjon med tasteteoremet?

11.5 ANALYSE AV DISKRETE SYSTEMER VED w -TRANSFORMASJON

Vi innfører nå en ny variabel w slik at

$$z = \frac{1 + \frac{T}{2}w}{1 - \frac{T}{2}w}, \text{ det vil si } w = \frac{2}{T} \frac{z-1}{z+1} \quad (11.54)$$

1. Da vil enhver rasjonal transferfunksjon i z (teller og nevner er begge polynomer i z) bli en rasjonal funksjon i w (dette følger av (11.54)).
2. Videre vil w være i venstre halvplan når z er på innsida av enhetssirkelen. Enhetssirkelen i z -planet blir transformert til den imaginære akse i w -planet og innsiden og utsiden av enhetssirkelen blir transformert til henholdsvis venstre og høyre halvdel av w -planet (dette følger også av (11.54)).
3. w vil spille samme rollen som s i $h(s)$, når T er liten i forhold til prosessens dynamikk.

Egenskapene 1. og 2. ovenfor betyr at vi kan bruke all den teorien vi har lært blant annet om stabilitet i s -planet også i w -planet, f.eks. algebraiske stabilitetskriterier av typen Routh/Hurwitz. Egenskap 3. (som blir nærmere forklart og begrunnet nedenfor) betyr at vi kan bruke Bode-diagram for analyse og syntese. Kort sagt, w -transformasjonen er meget hensiktsmessig til analyse og syntese av diskrete reguleringssystemer, da vi kan benytte de samme transferfunksjons-baserte metodene som er utviklet for kontinuerlige systemer.

Vi skal nå begrunne egenskap 3. ovenfor, og søker frekvensresponsen til $h(w)$:

Fra (11.54), høyre del, får vi

$$w(\omega) = \left. \frac{2z-1}{Tz+1} \right|_{z=e^{j\omega T}} = \frac{2e^{j\omega T} - 1}{Te^{j\omega T} + 1} = \frac{2e^{j\omega T/2} - e^{-j\omega T/2}}{Te^{j\omega T/2} + e^{-j\omega T/2}} = j\frac{2}{T} \tan \frac{\omega T}{2} \quad (11.55)$$

Hvis vi nå definerer en ny variabel v slik at $w = jv$, får vi

$$v = \frac{2}{T} \tan \left(\frac{\omega T}{2} \right), \quad \text{dvs.} \quad \frac{vT}{2} = \tan \left(\frac{\omega T}{2} \right) \quad (11.56)$$

Av dette ser vi at frekvensresponsen til $h(w)$ framkommer ved å sette $w = jv$ (i analogi med $s = j\omega$ for $h(s)$). Av (11.56) ser vi også at $v \approx \omega$ når $1/T$ er stor i forhold til ω . Vi ser videre at når $\omega \rightarrow \pi/T$, så går $v \rightarrow \infty$. Dette er en konsekvens av tastingen og uttrykker at signalkomponenter med frekvens over halve tastefrekvensen ikke vil kunne reproduseres (jfr. tasteteoremet).

Konklusjonen blir da at vi kan framstille frekvensresponsen til et diskret system $h(w)$ ved å substituere $w = jv$, og at hjelpe“frekvensen” v vil være nær virkelig frekvens ω så lenge vi betrakter frekvenser godt under halve tastefrekvensen. Fordi (11.56) er strengt monoton, svarer det en entydig verdi av v til hver verdi av ω . Den eksakte sammenheng er gitt av (11.56). Vi kan bruke (asymptotiske) Bode-diagram til å skissere forløpet av amplitude og fase for $h(w)$, og den eneste forskjell fra kontinuerlige systemer blir at den sanne frekvensaksen blir strukket ut mer og mer, jo lenger vi kommer mot høyre, jfr. (11.56).

Merk at Nyquist-diagrammet (og -stabilitetskriteriet) ved bruk av $h(w)$ i stedet for $h(z)$ blir akkurat det samme. Sammenhengene (11.54) har ingen betydning for formen på Nyquist-diagrammet, siden frekvensen ikke er en eksplisitt variabel i et polardiagram.

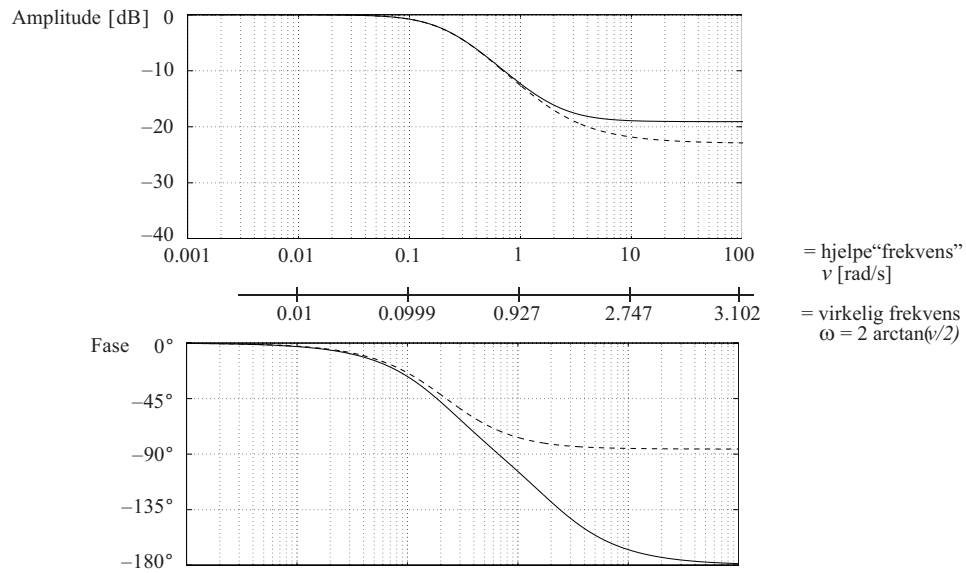
EKSEMPEL 11.9: “Frekvens”responsen for kontinuerlig og diskretisert prosess

Hvis vi bruker (11.54) til å substituere for z i (11.52), får vi

$$h(w) = \frac{1 - 0.5w}{1 + 4.5w} \quad (11.57)$$

Figur 11.29 viser frekvensforløpet til $h(w)$ med $w = jv$. Det er akkurat samme frekvensforløp som for (11.52) i figur 11.24. Forskjellen er bare at skalaen langs frekvensaksen nå er logaritmisk i v , ikke i ω . Virkelig frekvens ω er også vist. Vi ser hvordan overensstemmelsen er meget god mellom v og ω ved lave frekvenser, men at $v \rightarrow \infty$ når $\omega \rightarrow \pi/T$ (Vi har $T = 1$ i dette tilfellet).

Figur 11.29
 $h(s)$ og $h(w)$



For $h(w)$ med $w = jv$ (heltrukne kurver i figuren), vil vi kunne bruke asymptoter for å skissere forløp for amplitude og fase. Som i figur 11.24 er frekvensresponsen til den kontinuerlige prosess $h(s)$ (gitt i (11.51)) tegnet inn stiplet. I denne frekvensskalaen vil man *ikke* kunne bruke asymptoter for å skissere $h(s)$, bare $h(w)$.

Den vesentlige forskjell mellom $h(w)$ og $h(s)$ er leddet $(1 - 0.5w)$ i telleren. Dette skyldes holdeelementet, og er omtalt nedenfor.



Ved hjelp av metoden som er beskrevet i 11.4.1 og 11.4.4 for å finne $h(z)$, og deretter innsetting av (11.54), kan vi finne $h(w)$. Dette kan vi gjøre ved hjelp av MATLAB, eller ved hjelp av metoden som er forklart i tabell 11.1 nedenfor. Først til MATLAB, hvor vi fortsatt bruker (11.51), dvs. $h(s) = 1/(1 + T_1s)$ med tastetid $T = 1$, som eksempel:

Figur 11.30
MATLAB-kode for
generering av
 $h(w)$ fra $h(s)$,
gitt T

```

1 T = 1; T1 = -1/log(0.8);
2 num = 1;
3 den = [T1 1];
4 sys = tf(num,den);
5 [nz, dz] = c2dm(num,den,T);
6 [nw, dw] = d2cm(nz,dz,T,'tustin');
7 sysw = tf(nw,dw);

```

teller i $h(s)$

nevner i $h(s)$

Tustins metode innebærer å substituere venstre del av (11.53) for z

(“Tustins metode” er egentlig en metode for å lage et kontinuerlig system på grunnlag av et diskret, altså å “gå den andre veien”. MATLAB viser derfor den resulterende transferfunksjon som funksjon av s . Men det *er* altså $h(w)$ vi får ved å bruke `d2cm(...)` på denne måten.)

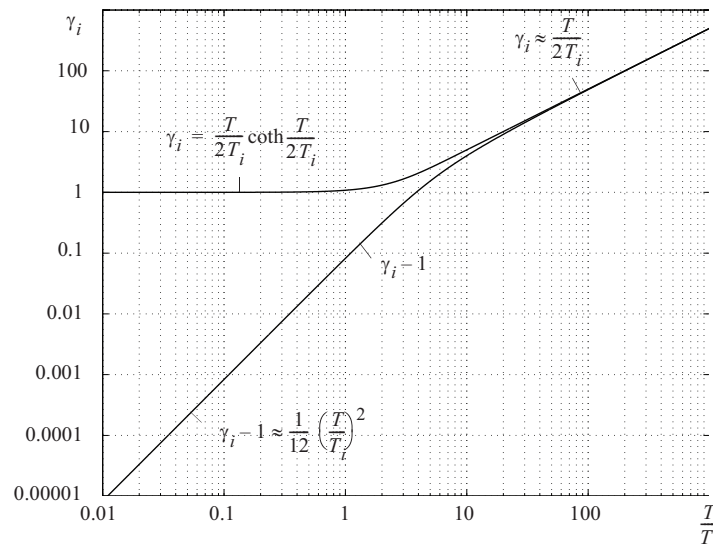
Så til den andre, “manuelle”, metoden: Tabell 11.1 viser par av Laplace-transformerte og w -transformerte transferfunksjoner for noen viktige monovariabel systemer. Husk at den w -transformerte gjelder for et system som inkluderer et holdeelement. Alle de w -transformerte transferfunksjonene inneholder faktoren $(1 - Tw/2)$. Denne faktoren skyldes holdeelementet, og kan forstås enklest ved å betrakte figur 11.10 med tilhørende tekst. Tabell 11.1 viser at w -transformasjonen er en lineær transformasjon.

I en rekke av uttrykkene i tabell 11.1 finnes en faktor γ_i som er gitt av

$$\gamma_i = \frac{T}{2T_i} \coth\left(\frac{T}{2T_i}\right) \quad (11.58)$$

Vi ser at $\gamma_i \rightarrow 1$ når $T/(2T_i) \rightarrow 0$, det vil si når tastetida T blir kort i forhold til den aktuelle tidskonstanten T_i . Figur 11.31 viser grafisk hvordan faktoren γ_i varierer med forholdet T/T_i . Når $T/T_i < 0.1$, kan vi med god nøyaktighet anta $\gamma_i = 1$. Dette gir en betydelig forenkling av de forskjellige uttrykkene i tabell 11.1.

Figur 11.31
 γ_i som funksjon
av T/T_i



Vi skal nå innføre regulator i w -planet. Vi studerer dette ved å betrakte en diskret PI-regulator, gitt ved (11.5). Hvis vi nå z -transformerer ledd for ledd i (11.5) ved å benytte tidsforskyvingsegenskapen, og deretter setter inn (11.54) for z , finner vi

$$h_r(w) = K_p \frac{1 + T_i w}{T_i w} \quad (11.59)$$

(11.59) har samme form som transferfunksjonen for en kontinuerlig PI-regulator. Denne, eller andre regulatorer, sammen med transferfunksjonen for den diskretiserte fysiske prosess $h(w)$, kan da brukes til å danne sløyfetransferfunksjonen $h_0(w) = h_r(w)h(w)$. Dermed kan vi bruke Nyquist, Nichols- og (asymptotiske) Bode-diagrammer til stabilitetsundersøkelse, analyse og syntese, helt tilsvarende det som gjøres for kontinuerlige systemer.

TABELL 11.1 w -transformerte transferfunksjoner*

	$h(s)$ (kont. prosess)	$h(w)$ (diskretisert kont. prosess med holdeelement)
	$h_1(s) + h_2(s)$	$h_1(w) + h_2(w)$
	$h_1(s)h_2(s)$	$h(w) \neq h_1(w)h_2(w)$
	$h(s)e^{-mTs}$ $m = \text{heltall}$	$h(w) \left(\frac{1 - \frac{T}{2}w}{1 + \frac{T}{2}w} \right)^m$
0.1	1	$\frac{1 - \frac{T}{2}w}{1 + \frac{T}{2}w}$ fra 1.2 når $T_1 \rightarrow 0$
0.2	$\frac{1 + T_2s}{T_2s}$	$\frac{\left(1 - \frac{T}{2}w\right) \left(1 + \left(T_2 + \frac{T}{2}\right)w\right)}{\left(1 + \frac{T}{2}w\right) T_2w}$
0.3	$\frac{1 + T_2s}{1 + T_1s}$	$\frac{\left(1 - \frac{T}{2}w\right) \left(1 + \left(\gamma_1 T_2 + \frac{T}{2} \left(1 - \frac{T_2}{T_1}\right)\right)w\right)}{\left(1 + \frac{T}{2}w\right) 1 + \gamma_1 T_1 w}$
0.4	$\frac{T_1s}{1 + T_1s}$	$\frac{\left(1 - \frac{T}{2}w\right) \left(\gamma_1 T_1 - \frac{T}{2}\right)w}{\left(1 + \frac{T}{2}w\right) 1 + \gamma_1 T_1 w}$
1.1	$\frac{1}{s}$	$\frac{1 - \frac{T}{2}w}{w}$
1.2	$\frac{1}{1 + T_1s}$	$\frac{1 - \frac{T}{2}w}{1 + \gamma_1 T_1 w}$

TABELL 11.1 w-transformerte transferfunksjoner*

	$h(s)$ (kont. prosess)	$h(w)$ (diskretisert kont. prosess med holdeelement)
1.3	$\frac{1 + T_2 s}{s^2}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + T_2 w)}{w^2}$
1.4	$\frac{T_1 s}{(1 + T_1 s)(1 + T_2 s)}$	$\frac{\left(1 - \frac{T}{2}w\right) \frac{(\gamma_1 T_1 - \gamma_2 T_2) T_1 w}{T_1 - T_2}}{(1 + \gamma_1 T_1 w)(1 + \gamma_2 T_2 w)}$
1.5	$\frac{1 + T_2 s}{s(1 + T_1 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + (T_2 + (\gamma_1 - 1)T_1)w)}{w(1 + \gamma_1 T_1 w)}$
1.6	$\frac{1 + T_3 s}{(1 + T_1 s)(1 + T_2 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + Bw)}{(1 + \gamma_1 T_1 w)(1 + \gamma_2 T_2 w)}$ $B = T_3 \frac{\gamma_1 T_1 - \gamma_2 T_2}{T_1 - T_2} + \frac{T_1 T_2}{T_1 - T_2} (\gamma_2 - \gamma_1)$
2.1	$\frac{1}{s^2}$	$\frac{1 - \frac{T}{2}w}{w^2}$
2.2	$\frac{1}{s(1 + T_1 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + (\gamma_1 - 1)T_1 w)}{w(1 + \gamma_1 T_1 w)}$
2.3	$\frac{1}{(1 + T_1 s)(1 + T_2 s)}$	$\frac{\left(1 - \frac{T}{2}w\right) \left(1 + \frac{T_1 T_2}{T_1 - T_2} (\gamma_2 - \gamma_1) w\right)}{(1 + \gamma_1 T_1 w)(1 + \gamma_2 T_2 w)}$
2.4	$\frac{1 + T_2 s}{s^2(1 + T_1 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + (T_2 + (\gamma_1 - 1)T_1)w - (\gamma_1 - 1)(T_1 - T_2)T_1 w^2)}{w^2(1 + \gamma_1 T_1 w)}$

TABELL 11.1 w -transformerte transferfunksjoner*

	$h(s)$ (kont. prosess)	$h(w)$ (diskretisert kont. prosess med holdeelement)
2.5	$\frac{1 + T_3 s}{s(1 + T_1 s)(1 + T_2 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + Bw + Cw^2)}{w(1 + \gamma_1 T_1 w)(1 + \gamma_2 T_2 w)}$ $B = T_3 + (\gamma_1 - 1)T_1 + (\gamma_2 - 1)T_2$ $C = \frac{T_1 T_2}{T_1 - T_2}(\gamma_1 \gamma_2 (T_1 - T_2) - \gamma_2 (T_1 - T_3) + \gamma_1 (T_2 - T_3))$
3.1	$\frac{1}{s^2(1 + T_1 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + (\gamma_1 - 1)T_1 w - (\gamma_1 - 1)T_1^2 w^2)}{w^2(1 + \gamma_1 T_1 w)}$ <p>som 2.4 med $T_2 = 0$</p>
3.2	$\frac{1}{s(1 + T_1 s)(1 + T_2 s)}$	som 2.5 med $T_3 = 0$
3.3	$\frac{1}{(1 + T_1 s)(1 + T_2 s)(1 + T_3 s)}$	$\frac{\left(1 - \frac{T}{2}w\right)(1 + Bw + Cw^2)}{(1 + \gamma_1 T_1 w)(1 + \gamma_2 T_2 w)(1 + \gamma_3 T_3 w)}$ $B = a_1(\gamma_2 T_2 + \gamma_3 T_3) + a_2(\gamma_1 T_1 + \gamma_3 T_3) + a_3(\gamma_1 T_1 + \gamma_2 T_2)$ $C = a_1 \gamma_2 T_2 \gamma_3 T_3 + a_2 \gamma_1 T_1 \gamma_3 T_3 + a_3 \gamma_1 T_1 \gamma_2 T_2$ $a_1 = \frac{1}{\left(1 - \frac{T_2}{T_1}\right)\left(1 - \frac{T_3}{T_1}\right)}$ $a_2 = \frac{1}{\left(1 - \frac{T_1}{T_2}\right)\left(1 - \frac{T_3}{T_2}\right)}$ $a_3 = \frac{1}{\left(1 - \frac{T_1}{T_3}\right)\left(1 - \frac{T_2}{T_3}\right)}$

*Første siffer i første kolonne angir relativ grad, det vil si ordenen til nevnerpolynommet minus ordenen til tellerpolynommet.

EKSEMPEL 11.10: Bruk av w -transformasjon på servomekanisme - stabilitet

Som eksempel på bruk av w -transformasjon skal vi først analysere en servomekanisme med diskret regulering. Servomekanismen består som vist i figur 11.32 av en motor som drives av en forsterker. Motoren trekker en last hvis posisjon (eller vinkel) måles av et måleelement (vi kan f.eks. tenke på dette som et reguleringsystem for en heisekran). Transferfunksjonen for den fysiske prosessen er gitt av

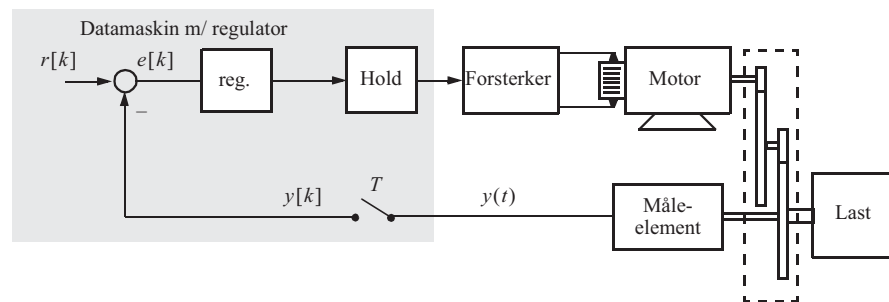
$$h(s) = \frac{1}{T_1 s(1 + T_2 s)} \quad (11.60)$$

der T_2 er tidskonstanten for motor med last. Vi tenker oss at regulatoralgoritmen er en ren forsterkning. For enkelhets skyld inkorporerer vi den i faktoren $1/T_1$ som også inneholder prosessens “fysiske forsterkning”. Dermed vil $1/T_1$ uttrykke den åpne sløyfes totale forsterkning og være justérbar.

Den w -transformerte av transferfunksjonen i (11.60) finnes i tabell 11.1 i linje 2.2 som

$$h_0(w) = \frac{\left(1 - \frac{T}{2}w\right)(1 + (\gamma_1 - 1)T_2w)}{T_1w(1 + \gamma_2T_2w)} \quad (11.61)$$

Figur 11.32
Servomekanisme



Vi har benyttet betegnelsen $h_0(w)$ fordi (11.61) utgjør den diskrete servomekanismens sløyfetransferfunksjon.

Denne sløyfes stabilitet kan nå undersøkes ved å se på uttrykket

$$1 + h_0(w) = \frac{\left(\gamma_2 T_1 T_2 - (\gamma_2 - 1) \frac{T T_2}{2}\right) w^2 + \left(T_1 + (\gamma_2 - 1) T_2 - \frac{T}{2}\right) w + 1}{T_1 w(1 + \gamma_2 T_2 w)} \quad (11.62)$$

Bruk av Rouths kriterium gir betingelsene for at (11.62) ikke skal ha nullpunkter i høyre halvplan

$$2\gamma_2 T_1 T_2 - (\gamma_2 - 1) T T_2 > 0 \quad (11.63)$$

$$2T_1 + 2(\gamma_2 - 1)T_2 - T > 0 \quad (11.64)$$

Innfører vi her de dimensjonsløse størrelsene $\alpha = T/T_2$ og $\beta = T_1/T_2$, finner vi ved bruk av (11.58) betingelsene for stabilitet.

$$\beta > \frac{\alpha}{2} - \frac{e^\alpha - 1}{e^\alpha + 1} \quad (11.65)$$

$$\beta > 1 - \frac{\alpha}{e^\alpha - 1} \quad (11.66)$$

Innholdet av (11.65) og (11.66) er vist i figur 11.33. Vi ser at når tasteperioden (T) er kort, vil vi ha stabilitet når $\beta > \alpha/2$ eller $T_1 > T/2$.

Når tasteperioden er lang, vil vi ha stabilitet når $\beta > \alpha/2 - 1$ eller $T_1 > T/2 - T_2$

Vi legger merke til at det kontinuerlige systemet karakterisert bare ved uttrykket i (11.60) vil være asymptotisk stabilt for alle verdier av T_1 , mens det diskrete systemet vil kunne bli ustabil ved stor tastetid. Som forventet blir grensa for høyeste forsterkning lavere når tastetida er stor. Dette skyldes økt fase-etterslep (tidsforsinkelse) p.g.a. holdeelementet, og kommer til uttrykk i w -plan-analysen i form av faktoren $1 - (T/2)w$ i telleren i (11.61). Denne faktoren gir et nullpunkt i høyre halvplan og representerer et "ikke-minimum-fase"-fenomen.

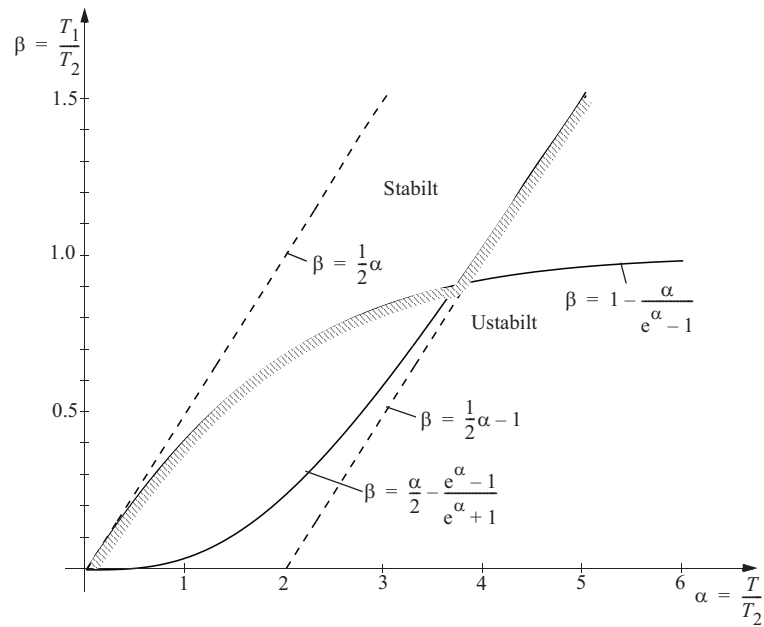
Vi kan, i likhet med det som gjøres for kontinuerlige systemer, studere "frekvensresponsen" av uttrykket i (11.61). Vi velger tallverdiene $T = T_2 = 1$. Da finner vi ved bruk av kurvene i figur 11.31

$$h_0(w) = \frac{(1 - 0.5w)(1 + 0.08w)}{T_1 w(1 + 1.08w)} \quad (11.67)$$

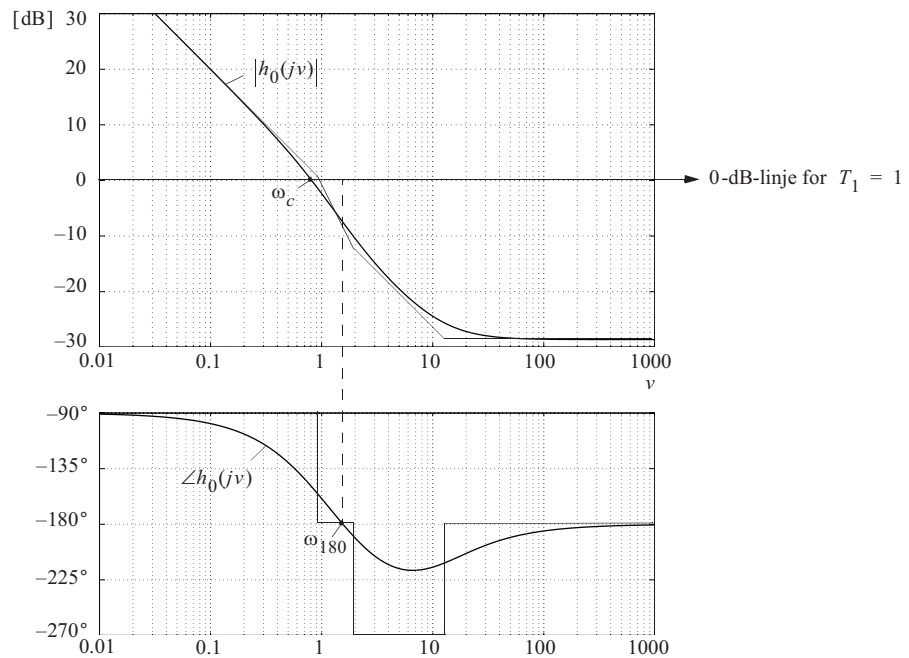
Innsatt $w = j\nu$ finner vi asymptotiske og eksakte "frekvens" karakteristikk som vist i figur 11.34, for $T_1 = 1$. Vi legger merke til at fasekurven går gjennom -180° og vi finner $\nu_{180} \approx 1.5$. Den maksimale forsterkningen systemet kan tåle, er ca. 7 [dB]. En nøyaktigere verdi finner vi i figur 11.33 som

$$\frac{1}{T_1} = \frac{1}{0.42} = 2.38 = 7.56 \text{ [dB]}$$

Figur 11.33
Verdier av β
som gir stabilitet/
ustabilitet



Figur 11.34
“Frekvens”karakteristikker
for $h_0(w)$



11.6 BRUK AV w -TRANSFORMASJON VED SYNTSE AV DISKRET REGULATOR

EKSEMPEL 11.11: Diskret proporsjonal pluss begrenset derivat regulator

Vi skal se på en kontinuerlig monovariabel prosess som består av to integratorer i serie, som et første eksempel på bruk av w -transformasjon ved syntese av en diskret regulator

$$h_u(s) = \frac{K}{s^2} \quad (11.68)$$

Ifølge linje 2.1 i tabell 11.1 vil den w -transformerte transferfunksjonen bli

$$h_u(w) = \frac{K \left(1 - \frac{T}{2}w\right)}{w^2} \quad (11.69)$$

En diskret regulator med transferfunksjon $h_r(w)$ vil nå sammen med uttrykket i (11.69) gi den diskrete sløyfetransferfunksjonen

$$h_0(w) = h_r(w)h_u(w) \quad (11.70)$$

Problemet blir altså som i det kontinuerlige tilfellet, å formulere en $h_r(w)$ som gir en fornuftig $h_0(w)$ ut fra hensynet til stabilitet og respons. Likning (11.69) viser at $h_r(w)$ må være av en type som gir positiv fase-korreksjon, for eksempel

$$h_r(w) = K_p \frac{1 + T_d w}{1 + \alpha T_d w} \quad (11.71)$$

som er en diskret “proporsjonal-pluss-begrenset-derivatregulator”. Før vi går videre, skal vi se på denne transferfunksjonens tolkning i z -planet ved innsetting av

$$w = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (11.72)$$

Vi finner

$$h_r(z) = K_p \frac{(2T_d + T) - (2T_d - T)z^{-1}}{(2\alpha T_d + T) - (2\alpha T_d - T)z^{-1}} \quad (11.73)$$

Vi velger $T = 2\alpha T_d$. Valget er ikke kritisk, poenget er bare at T må være “liten” i forhold til de øvrige tidskonstanter i $h_u(w)$. Med dette valget vil (11.71) gi

$$h_r(w) = K_p \frac{1 + T_d w}{1 + \frac{T}{2}w} \quad (11.74)$$

Likning (11.74) sammen med (11.69) gir nå den diskrete sløyfetransferfunksjonen

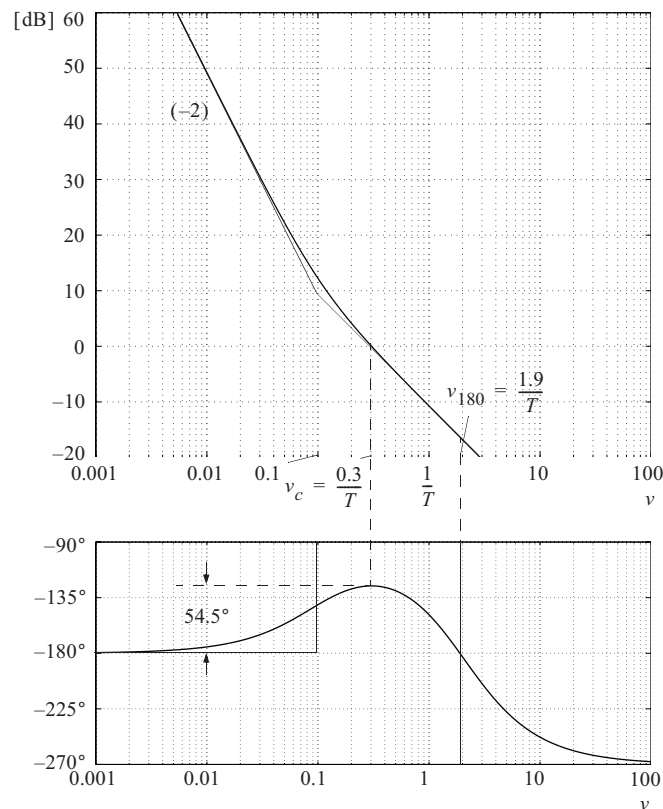
$$h_0(w) = K_p K \frac{1 + T_d w}{w^2} \frac{1 - \frac{T}{2} w}{1 + \frac{T}{2} w} \quad (11.75)$$

Vi kan velge å studere (11.75) ved hjelp av amplitude/fase/“frekvens”-diagrammer og får da et resultat som vist i figur 11.35, når $T = 1$. Velger vi for eksempel $T_d = 10T$, vil vi finne at fasemarginen blir størst ved “frekvensen” $\nu = 0.303/T = 3.03/T_d$.

Ved denne verdien av ν vil fasemarginen være ca. 54.5° . Velger vi å legge kryss-“frekvensen” slik at fasemarginen er maksimal, finner vi forsterkningen i systemet

$$K_p K = 0.0289 \frac{1}{T^2}$$

Figur 11.35
“Frekvens”karakteristikker for $h_0(w)$



EKSEMPEL 11.12: Diskret PID-regulator

Et noe mer komplisert eksempel hentes fra prosessregulering der vi antar en monovariabel og kontinuerlig prosess med transferfunksjon

$$h_u(s) = \frac{Ke^{-\tau s}}{(1 + T_1 s)(1 + T_2 s)} \quad (11.76)$$

ved hjelp av tabell 11.1 finner vi den w -transformerte transferfunksjonen

$$h_u(w) = \frac{K \left(1 - \frac{T}{2}w\right) \left(1 + \frac{T_1 T_2}{T_1 - T_2}(\gamma_2 - \gamma_1)w\right)}{(1 + \gamma_1 T_1 w)(1 + \gamma_2 T_2 w)} \left(\frac{1 - \frac{T}{2}w}{1 + \frac{T}{2}w}\right)^m \quad (11.77)$$

der $m = \tau/T$.

Vi ønsker å utforme en diskret PID-regulator med begrenset derivatvirkning som, jfr. (11.73), får transferfunksjonen

$$h_r(w) = K_p \frac{(1 + T_i w)(1 + T_d w)}{T_i w \left(1 + \frac{T}{2}w\right)} \quad (11.78)$$

Det vil være rimelig å velge $T_i = \gamma_1 T_1$ og $T_d = \gamma_2 T_2$. Da vil vi få

$$h_0(w) = K_p K \frac{\left(1 + \frac{T_1 T_2}{T_1 - T_2}(\gamma_2 - \gamma_1)w\right)}{T_i w} \left(\frac{1 - \frac{T}{2}w}{1 + \frac{T}{2}w}\right)^{m+1} \quad (11.79)$$

Antar vi at $T_1 \gg T_2$ og velger vi $T = T_2$, får vi

$$h_0(w) = K_p K \frac{(1 + 0.08 T w)}{T_i w} \left(\frac{1 - \frac{T}{2}w}{1 + \frac{T}{2}w}\right)^{m+1} \quad (11.80)$$

Faktoren $(1 + 0.08 T w)$ vil praktisk talt bli uten betydning fordi dens virkning vil komme et godt stykke over kryss“frekvensen”.

Vi sitter derfor igjen med en sløyfe hvis asymptotiske amplitudekarakteristikk faller med vinkelkoeffisienten (-1) innenfor hele det interessante området, fordi den siste

faktoren i (11.80) ikke bidrar til amplitudekarakteristikken. Den asymptotiske vinkelkarakteristikken starter ved -90° og foretar et sprang på $-(m+1)180^\circ$ ved "frekvensen" $\nu = 2/T$.

Dersom eksempelvis $\tau = 2T_2 = 2T$, det vil si $m = 2$, vil vi få $\angle h_0(j\nu) = -180^\circ$ ved $\nu_{180} = 0.536/T$. Ønsker vi et system med fasemargin 45° , finner vi $\angle h_0(j\nu) = -135^\circ$ ved $\nu_c = 0.263/T$. En slik innstilling av regulatorforsterkningen vil gi en forsterkningsmargin på $\Delta K = 2.04 = 6.2$ [dB].



11.7 TIDSRESPONS AV DISKRET SYSTEM

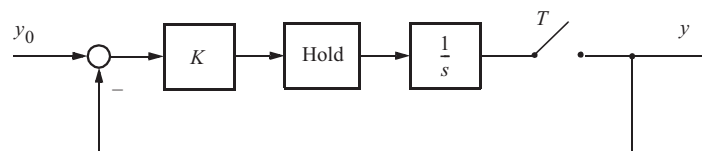
W -transformasjonen egner seg dårlig til å klarlegge hvordan et diskret regulerings-system oppfører seg i tidsplanet. Til dette er z -transformasjonen svært gunstig, som vi allerede har forklart i avsnitt 11.4.4. Når vi har syntetisert en regulator i w -planet, kan vi derfor med fordel gå tilbake til z -planet og via dette finne responsen i det diskrete tidsplanet.

EKSEMPEL 11.13: Et lukket systems respons i det diskrete tidsplan

Som et eksempel på dette skal vi se på det særdeles enkle tilfellet der den kontinuerlige prosessen som skal reguleres, er en ren integrator. Vi vil bruke en diskret proporsjonal regulator slik som vist i figur 11.36. Vi finner den w -transformerte transferfunksjonen for den diskrete åpne sløyfes transferfunksjon av linje 1.1 i tabell 11.1

$$h_0(w) = \frac{K \left(1 - \frac{T}{2}w\right)}{w} \quad (11.81)$$

Figur 11.36
Tilbakekoplet system med diskret proporsjonal regulator



Innsatt z -transformasjonen fra (11.54) finner vi

$$h_0(z) = \frac{K \left(1 - \frac{1 - z^{-1}}{1 + z^{-1}}\right)}{\frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}}} = KT \frac{z^{-1}}{1 - z^{-1}} \quad (11.82)$$

Transferfunksjonen for den lukkede sløyfen blir

$$\frac{y}{y_0}(z) = \frac{h_0(z)}{1 + h_0(z)} = \frac{KTz^{-1}}{1 - (1 - KT)z^{-1}} \quad (11.83)$$

I det diskrete tidsplanet kan (11.83) tolkes som

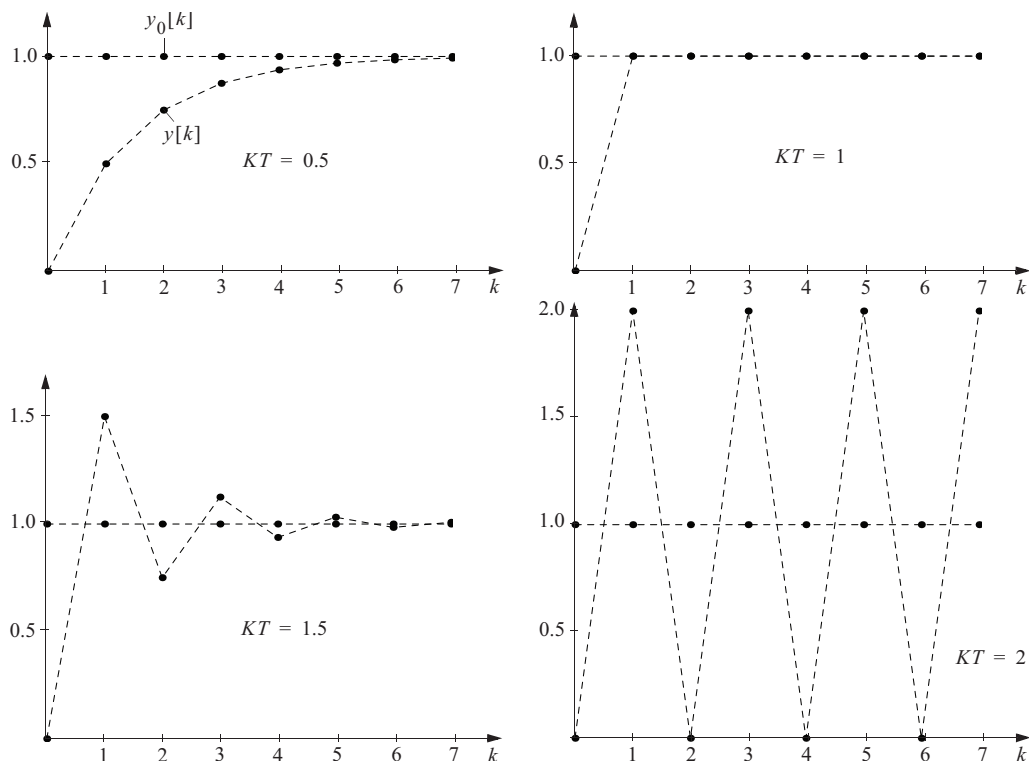
$$y[k] = (1 - KT)y[k - 1] + KTy_0[k - 1] \quad (11.84)$$

For en gitt verdi av produktet KT og en gitt sekvens av $y_0[k]$ er det da lett å finne sekvensen $y[k]$, ved å bruke den rekursive formelen (11.84) gjentatte ganger for $k = 1, 2, \dots$

Ved å sjekke polene i (11.83) finner vi at for å oppnå asymptotisk stabilitet, må vi ha $0 < KT < 2$. Figur 11.37 viser den diskrete responsen $y[k]$ for forskjellige verdier av KT når $y_0[k]$ er et diskret enhetsprang.

Verdiene som ble funnet, gjelder bare i tastetidspunktene, men i dette tilfellet er den kontinuerlige responsen $y(t)$ ganske lett å finne. Dette fordi $y(t)$ er integralet av det trappeformede (p.g.a. holdeelementet) pådraget $u(t)$. $y(t)$ må da bestå av rette sammenhengende linjestykker. Disse er markert stiplet i figuren.

Figur 11.37
Diskret sprang-
respons



11.8 DEN MULTIVARIABLE W-TRANSFORMASJONEN

Gitt et multivariabelt LTI-system med tilstandsrom-representasjonen

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x}\end{aligned}\tag{11.85}$$

Vi har

$$\begin{aligned}\Phi &= e^{\mathbf{A}T} \\ \Delta(T) &= \mathbf{A}^{-1}(\Phi(T) - \mathbf{I})\mathbf{B}\end{aligned}\tag{11.86}$$

Transfermatrisen $\mathbf{H}(z)$ blir

$$\mathbf{H}(z) = \mathbf{C}(z\mathbf{I} - \Phi)^{-1}\Delta\tag{11.87}$$

Ved å substituere (11.54) for z , får vi

$$\mathbf{H}(w) = \mathbf{C} \left(\begin{array}{c} 1 + \frac{T}{2}w \\ \frac{1 - \frac{T}{2}w}{T}\mathbf{I} - \Phi \\ 1 - \frac{T}{2}w \end{array} \right)^{-1} \Delta\tag{11.88}$$

Videre utvikling av (11.88) gir

$$\mathbf{H}(w) = \left(1 - \frac{T}{2}w\right) \mathbf{C} \left(\frac{T}{2}(\mathbf{I} + \Phi)w + (\mathbf{I} - \Phi) \right)^{-1} \Delta\tag{11.89}$$

Innfører vi nå

$$\Phi = e^{\mathbf{A}T} = e^{\mathbf{A}\frac{T}{2}} e^{\mathbf{A}\frac{T}{2}}$$

og

$$\Delta = \mathbf{A}^{-1}(e^{\mathbf{A}T} - \mathbf{I})\mathbf{B}$$

finner vi etter en del matrisemanipulasjon

$$\mathbf{H}(w) = \left(1 - \frac{T}{2}w\right) \mathbf{C}(w\mathbf{I} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{B}}\tag{11.90}$$

der

$$\tilde{\mathbf{A}} = \mathbf{A} \left(\frac{\mathbf{A}T}{2} \right)^{-1} \tanh \left(\frac{\mathbf{A}T}{2} \right)\tag{11.91}$$

$$\tilde{\mathbf{B}} = \left(\frac{\mathbf{A}T}{2} \right)^{-1} \tanh \left(\frac{\mathbf{A}T}{2} \right) \mathbf{B}\tag{11.92}$$

Rekkeutvikling av disse uttrykkene gir

$$\tilde{\mathbf{A}} = \mathbf{A} \left(\mathbf{I} - \frac{1}{12}(\mathbf{AT})^2 + \frac{1}{120}(\mathbf{AT})^4 - \frac{1}{1186}(\mathbf{AT})^6 + \frac{1}{11706}(\mathbf{AT})^8 - \dots \right) \quad (11.93)$$

$$\tilde{\mathbf{B}} = \left(\mathbf{I} - \frac{1}{12}(\mathbf{AT})^2 + \frac{1}{120}(\mathbf{AT})^4 - \dots \right) \mathbf{B} \quad (11.94)$$

Rekken som forekommer i disse to uttrykkene har som vi ser meget hurtig konvergens, hvilket betyr at vi som oftest vil kunne klare oss med få ledd. Når $\mathbf{AT} \rightarrow 0$, ser vi at $\tilde{\mathbf{A}} \rightarrow \mathbf{A}$ og $\tilde{\mathbf{B}} \rightarrow \mathbf{B}$ hvilket medfører

$$\mathbf{H}(w) \rightarrow \left(1 - \frac{T}{2}w \right) \mathbf{C}(w\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \quad (11.95)$$

Resultatet i (11.90) kan illustreres med et eksempel.

EKSEMPEL 11.14: w -transformasjon av monovariabelt kontinuerlig system

Vi ser på et monovariabelt kontinuerlig system karakterisert ved

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{1}{T_1} \end{bmatrix}, \quad \mathbf{B} = \mathbf{b} = \begin{bmatrix} 0 \\ \frac{1}{T_1} \end{bmatrix}, \quad \mathbf{C} = \mathbf{c}^T = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

som gir

$$\mathbf{H}(s) = h(s) = \frac{1}{s(1 + T_1s)}$$

Velger vi $T/T_1 \leq 1$ (forholdsvis grov tasting), vil vi ifølge (11.93) og (11.94) kunne skrive

$$\tilde{\mathbf{A}} \cong \mathbf{A} \left(\mathbf{I} - \frac{1}{12} \begin{bmatrix} 0 & T \\ 0 & -\frac{T}{T_1} \end{bmatrix}^2 \right) = \begin{bmatrix} 0 & 1 - \frac{1}{12} \left(\frac{T}{T_1} \right)^2 \\ 0 & -\frac{1}{T_1} \left(1 - \frac{1}{12} \left(\frac{T}{T_1} \right)^2 \right) \end{bmatrix}$$

$$\tilde{\mathbf{B}} \cong \begin{bmatrix} \frac{1}{12} \left(\frac{T}{T_1} \right)^2 \\ \frac{1}{T_1} \left(1 - \frac{1}{12} \left(\frac{T}{T_1} \right)^2 \right) \end{bmatrix}$$

$$\mathbf{H}(w) = \left(1 - \frac{T}{2}w\right) \begin{bmatrix} 1 & 0 \end{bmatrix} \frac{\text{adj}(w\mathbf{I} - \tilde{\mathbf{A}})}{\det(w\mathbf{I} - \tilde{\mathbf{A}})} \begin{bmatrix} \frac{1}{12} \left(\frac{T}{T_1}\right)^2 \\ \frac{1}{T_1} \left(1 - \frac{1}{12} \left(\frac{T}{T_1}\right)^2\right) \end{bmatrix}$$

som gir

$$\begin{aligned} \mathbf{H}(w) = h(w) &= \frac{\left(1 - \frac{T}{2}w\right) \left(1 + \frac{\frac{1}{12} \left(\frac{T}{T_1}\right)^2}{1 - \frac{1}{12} \left(\frac{T}{T_1}\right)^2} T_1 w\right)}{w \left(1 + \frac{1}{1 - \frac{1}{12} \left(\frac{T}{T_1}\right)^2} T_1 w\right)} \\ &\approx \frac{\left(1 - \frac{T}{2}w\right) \left(1 + \frac{1}{12} \left(\frac{T}{T_1}\right)^2 T_1 w\right)}{w \left(1 + \left(\frac{1}{12} \left(\frac{T}{T_1}\right)^2\right) T_1 w\right)} \end{aligned}$$

Dette resultatet kan kontrolleres mot tabell 11.1, linje 2.2 og det tilnærmede uttrykket for som er påført nederst i figur 11.31.



Begrensninger i lineære reguleringsystemer

12.1 Innledning og motivering 553

12.2 Metning. Amplitudebegrensning 555

12.3 Effektbruk i reguleringsystemer 561

12.4 Bestemmelse av den realiserbare båndbredden av et regulerings-system 564

12.5 Pådragsmetning i systemer med integrerende regulator 566

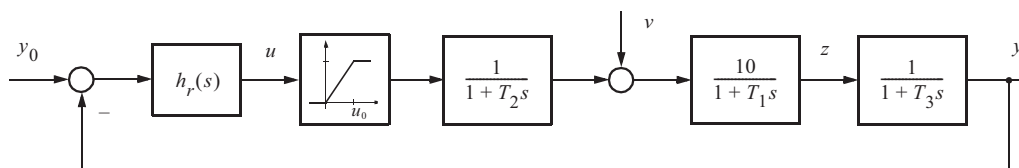
12.1 INNLEDNING OG MOTIVERING

I de foregående kapitlene er det en forutsetning at alle elementene i reguleringsløyfen er lineære. Vi husker for eksempel at Laplace-modeller og frekvensanalysen baserer seg på antakelsen om at systemer er lineære. Men alle fysiske systemer blir ulineære hvis man bare betrakter store nok utsving vekk fra systemets arbeidspunkt. La oss starte med et enkelt eksempel som viser nødvendigheten av å inkludere ulineære fenomener i flere sammenhenger.

EKSEMPEL 12.1: Ulineariteter som følge av begrensning i pådragsorgan

Temperaturreguleringen i figur 9.16 baserer seg på varmeveksleren i figur 9.15. I figur 12.1 har vi inkludert en ekstra blokk med en metningsfunksjon. Den angir at ventilåpningen er lik 0 når pådraget er negativt ($u < 0$), og at ventilåpningen når en maksimal åpning når $u = u_0$. Dette betyr at det reelle pådraget inn på prosessen vil være forskjellig fra det pådraget som regulatoren beregner, når $u < 0$ eller når $u > u_0$. Dersom pådraget alltid holder seg i området $0 < u < u_0$, vil reguleringsløyfen være lineær og all den foregående teorien kan benyttes. Dersom pådraget når sin begrensning, må reguleringsløyfen behandles som et ulineært reguleringsystem. Teori for å analysere denne typen systemer er ikke et tema i denne boka.

Figur 12.1
Regulatorsløyfe
med begrensning
i pådragsorganet



I eksemplet ovenfor er det pådragsorganet som gjør at reguleringsløyfen blir ulineær. En annen viktig årsak til at reguleringsløyfen blir ulineær er at prosessen i seg selv er ulineær.

Alle virkelige prosesser er ulineære. Et eksempel på dette er reaktoren som modelleres i eksempel 3.6. Den lineariserte modellen avhenger av arbeidspunktet, i dette tilfellet pådrags-nivå, og temperatur i reaktoren. Industrielt er det vanlig å håndtere dette ved **parameterstyring**. Dette betyr at en eller flere regulatorparametre, for eksempel forsterkningen i en PID-regulator, endrer seg med prosessens arbeidspunkt. Dette blir dermed en ulineær regulator. Parameterstyring vil ikke bli behandlet videre i dette kapitlet idet fokus er ulineariteter som følge av begrensning i pådragsorganet. Et annet eksempel på en ulineær prosess er vist i figur 9.45 hvor propellen har en ulineær karakteristikk.

Begrensninger forekommer i mange forskjellige former i reguleringsystemer. Vi skal nå diskutere begrensning i pådragsorganet nærmere. Dette fenomenet benevnes vanligvis **metning** og er vist i figur 12.1.

Dersom et element drives ut over sin yteevne, vil dette innvirke på hele systemets oppførsel. En motor som har en påstemplet ytelse på 1 [kW], kan ikke yte 2 [kW] i lengre tid. Det er derimot ikke umulig at den i kortere perioder er i stand til å prestere langt mer enn sin merkeeffekt, fordi det ofte er den termiske påkjenningen som setter grensen for langtidsytelsen. Store termiske kapasiteter vil som oftest sørge for at kortvarige overbelastninger ikke gir store temperaturstigninger i motoren.

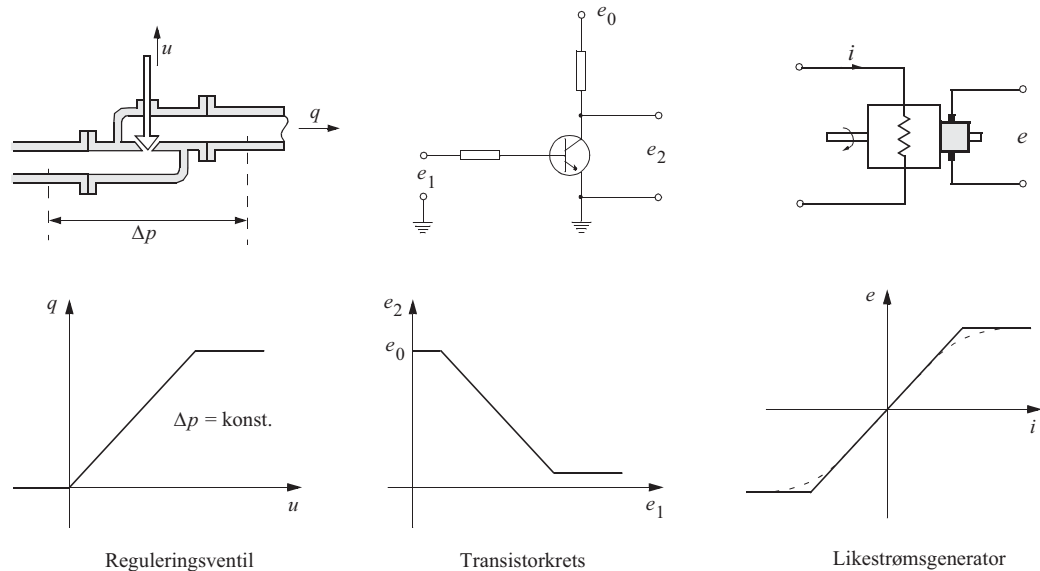
Det forholder seg annerledes dersom motoren eller et annet element på grunn av effektbegrensninger (*metning*) ikke er i stand til å yte mer enn sin merkeeffekt (pådrag). Det er tilfellet med en reguleringsventil som ikke kan gi mer enn full åpning i den rørledningen den er montert. Likeledes vil en elektronisk forsterker bare kunne levere strøm innenfor et begrenset område. Utenfor dette området “går den i metning”. Dersom et eller flere elementer i et reguleringsystem går i metning mens de utfører sine reguleringsfunksjoner, er det klart at reguleringsforløpet må bli annerledes enn om de hadde vært lineære (ubegrensede). Systemet kan da ikke analyseres ut fra de lineære metodene som er beskrevet foran, men det er ikke nødvendigvis slik at systemet er ubrukbart fordi elementer går i metning en gang i mellom. Det er snarere slik at systemet er litt for rikelig dimensjonert, dersom ingen av elementene noen gang når sine metningsverdier under utførelse av sine reguleringsfunksjoner.

I dette kapitlet skal vi se på hvordan vi kan konstatere om noen av systemets elementer når sine metningsverdier og hvor stor effekt elementene leverer.

12.2 METNING. AMPLITUDEBEGRENSNING

Figur 12.2 viser statiske karakteristikker for noen typiske elementer fra regulerings-teknikken. De er alle karakterisert ved at de har en øvre og nedre begrensningsverdi, mer eller mindre skarpt definert. Slike elementer eller fysikalske relasjoner kan forekomme på de forskjelligste steder i reguleringsystemer.

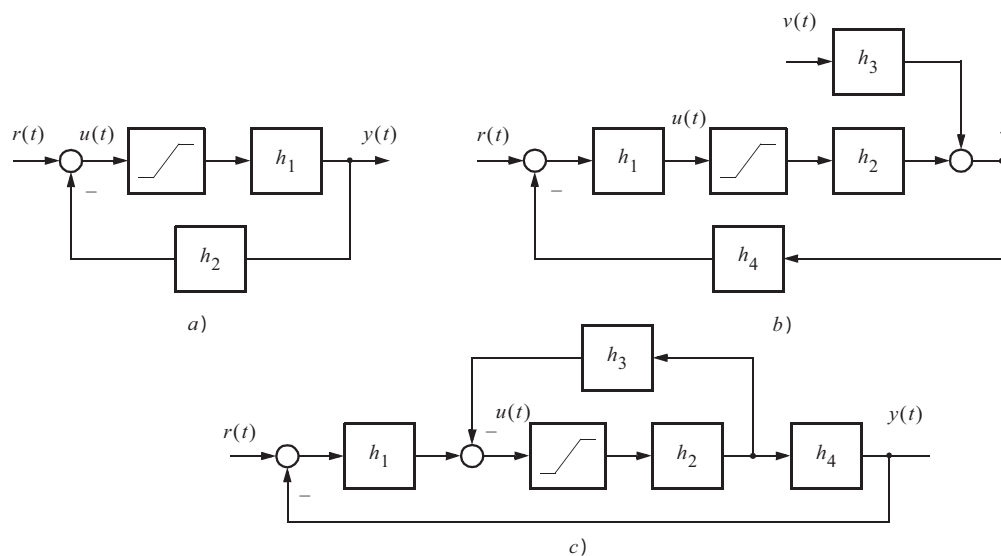
Figur 12.2
Elementer som
kan gå i metning



Figur 12.3 viser noen enkle eksempler på systemer med et element som kan gå i metning, men som forøvrig kan beskrives med lineære relasjoner (transferfunksjoner $h(s)$). De mest betydningsfulle inngangs- og utgangsvariable er også antydnet. Ved hjelp av analysemetoder for lineære systemer kan vi nå lett finne ut betingelsene for at metning *ikke skal inntre*. Vi kan derimot ikke uten videre finne ut hva som skjer dersom metning inntre.

Kjenner vi formen på inngangsvariabelen til hele systemet, kan vi i alle fall i enkle tilfeller finne formen på variabelen u foran det elementet der metningen inntre. Et vesentlig forhold er at de variables størrelse (amplitude) overalt i lineære systemer vil være proporsjonale. Har vi funnet formen av u for en størrelse av inngangsvariabelen, kan vi direkte finne den amplituden av denne variabelen som vil medføre metning. I tilfeller der slik metning medfører dårligere ytelser av systemet, har vi dermed fått en viktig opplysning.

Figur 12.3
Systemer med et
element som
kan gå i metning



12.2.1 Undersøkelse av metning i tidsplanet

Som kjent kan vi i et lineært system prinsipielt finne responsen hvor som helst, forutsatt at inngangsvariablene er gitt. Selv om vi for å konstatere metning bare trenger å finne maksimalverdien av signalet umiddelbart foran metningselementet, kan det imidlertid innebære svært tidkrevende beregninger dersom disse foretas i tidsplanet.

Antar vi at det elementet som kan gå i metning, har en forsterkning lik 1 innenfor sitt lineære område, finner vi for de tre systemene som er skissert i figur 12.3:

$$\text{System a)} \quad \frac{u}{r}(s) = N(s) \quad \text{der } N = \frac{1}{1 + h_1 h_2} \quad (12.1)$$

$$\text{System b)} \quad \frac{u}{v}(s) = -h_3 h_4 h_1 N \quad \text{der } N = \frac{1}{1 + h_1 h_2 h_4} \quad (12.2)$$

$$\text{System c)} \quad \frac{u}{r}(s) = h_1 N_1 N_2 \quad \text{der } N_1 = \frac{1}{1 + h_1 h_4 \frac{h_2}{1 + h_2 h_3}} \quad (12.3)$$

$$\text{og } N_2 = \frac{1}{1 + h_2 h_3}$$

$u(t)$ finnes ved inverstransformasjon av $u(s)$ i disse uttrykkene. Dette illustreres ved et eksempel, der vi velger system a) i figur 12.3. Vi antar da gitt:

$$r(t) = r_0 \left(1 - e^{-\frac{t}{T_0}}\right), \text{ det vil si } r(s) = \frac{r_0}{s(1 + T_0s)}$$

$$h_1(s) = \frac{1}{T_1s}$$

$$h_2(s) = 1$$

Da blir

$$u(s) = \frac{r_0T_1}{(1 + T_0s)(1 + T_1s)} \quad (12.4)$$

$$u(t) = \frac{r_0T_1}{T_1 - T_0} \left(e^{-\frac{t}{T_1}} - e^{-\frac{t}{T_0}}\right) \quad (12.5)$$

Vi bestemmer nå $\frac{du}{dt} = 0$ og finner

$$u_{\max} = r_0 \left(\frac{T_0}{T_1}\right)^{\frac{T_0}{T_1 - T_0}} \quad (12.6)$$

ved

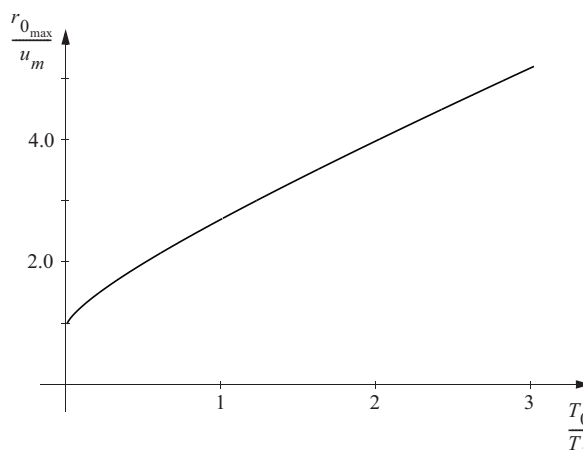
$$t = \frac{T_1T_0}{T_0 - T_1} \ln \frac{T_0}{T_1} \quad (12.7)$$

Når vi kaller metningsgrensen av u for u_m , finner vi den maksimale størrelsen av r_0 som ikke gir metning i systemet:

$$r_{0\max} = u_m \left(\frac{T_0}{T_1}\right)^{\frac{T_0}{T_1 - T_0}} \quad (12.8)$$

$r_{0\max}/u_m$ fra (12.8) er vist grafisk i figur 12.4. Denne figuren viser tydelig at når inngangssignalet blir “mykere”, det vil si T_0 større relativt til systemets tidskonstant T_1 , kan det tåles en større verdi av r_0 uten at systemet går i metning.

Figur 12.4
 $r_{0\max}/u_m$ som
 funksjon av T_0/T_1



Vi ser av dette enkle eksemplet at de beregninger som må utføres, er forholdsvis omfattende. For et mer realistisk tilfelle ville det knapt nok være mulig å få fram et hensiktsmessig svar ved analytiske metoder og uten bruk av datamaskin. Det kan derfor være behov for å benytte andre betraktningmåter.

12.2.2 Undersøkelse av metning i frekvensplanet

Dersom $r(t)$ kan beskrives med følgende uttrykk

$$r(t) = r_1 \sin \omega_1 t + r_2 \sin(\omega_2 t + \varphi_2) + \dots \quad (12.9)$$

finnes $u(t)$ for system a) i figur 12.3 på følgende måte

$$u(t) = r_1 |N(j\omega_1)| \sin(\omega_1 t + \angle N(j\omega_1)) + r_2 |N(j\omega_2)| \sin(\omega_2 t + \varphi_2 + \angle N(j\omega_2)) + \dots \quad (12.10)$$

Dersom vi i (12.9) har $\omega_n = n\omega$, vil $r(t)$ være periodisk med grunnfrekvens ω_1 . Tilsvarende uttrykk som (12.10) finnes for de andre systemene i figur 12.3.

Størrelsene $|N(j\omega)|$ og $\angle N(j\omega)$ tas ut fra systemets frekvensdiagrammer. Det oppstår imidlertid et visst problem med bestemmelse av u_{\max} , fordi de forskjellige frekvenskomponentene har forskjellige fasevinkler. En god tilnærming til den riktige verdien kan vi få ved å regne ut effektivverdien u_{eff} av $u(t)$ og så anslå u_{\max} ut fra denne.

Vi har tidligere i avsnitt 6.5.1 utledet hvordan effektivverdien kan finnes, og får

$$u_{\text{eff}} = \frac{1}{\sqrt{2}} \sqrt{(r_1^2 |N(j\omega_1)|^2 + r_2^2 |N(j\omega_2)|^2 + \dots)} \quad (12.11)$$

Avhengig av amplituden og fasen til de forskjellige frekvenskomponentene, vil vi nå kunne få forskjellige forhold mellom maksimalverdi og effektivverdi. I et virkelig signal som har en tilfeldig natur, vil det ikke ha særlig mening å angi dette forholdet med stor presisjon. Det vesentlige er som nevnt at systemet *ikke går i metning en stor del av tiden*. At systemet går i metning et lite øyeblikk en gang i mellom vil vanligvis ikke være særlig skadelig. Vi kan derfor som oftest anslå forholdet mellom maksimalverdi og effektivverdi med tilstrekkelig grad av nøyaktighet, når vi vet at

$$\frac{u_{\max}}{u_{\text{eff}}} = \sqrt{3} \quad \text{for} \quad u(t) = \text{trekantkurve}$$

$$\frac{u_{\max}}{u_{\text{eff}}} = \sqrt{2} \quad \text{for} \quad u(t) = \text{sinuskurve}$$

$$\frac{u_{\max}}{u_{\text{eff}}} = 1 \quad \text{for} \quad u(t) = \text{firkantkurve}$$

Hvis vi i et foreliggende tilfelle anslår forholdstallet (“formfaktoren”) lik α , finner vi derfor $u_{\max} = \alpha u_{\text{eff}}$. Denne verdien benytter vi så i et uttrykk helt tilsvarende det som er utledet i (12.8) til bestemmelse av maksimalt tillatelig amplitude av inngangssignalet sine komponenter.

EKSEMPEL 12.2: Retningsstabilisering av plattform i rullende skip

For det spesielle tilfellet at systemets inngangssignal er rent sinusformet, vil forholdene bli enklere. Slik er det eksempelvis ved retningsstabilisering av en plattform ombord i et skip som ruller, når det antas at rullebevegelsen er nær sinusformet. Et slikt retningsstabiliseringssystem vil gjerne ha et blokkdiagram som vist i figur 12.3 c). Da vil vi ha

$$h_1(s) = K_1 = \text{konstant}$$

$$h_2(s) = \frac{K_2}{(1 + T_1s)(1 + T_2s)} = \text{(forsterker + motor)}$$

$$h_3(s) = \frac{K_3T_3s}{1 + T_3s} = \text{(intern tilbakekopling)}$$

$$h_4(s) = \frac{K_4}{s} = \text{(motor + gir)}$$

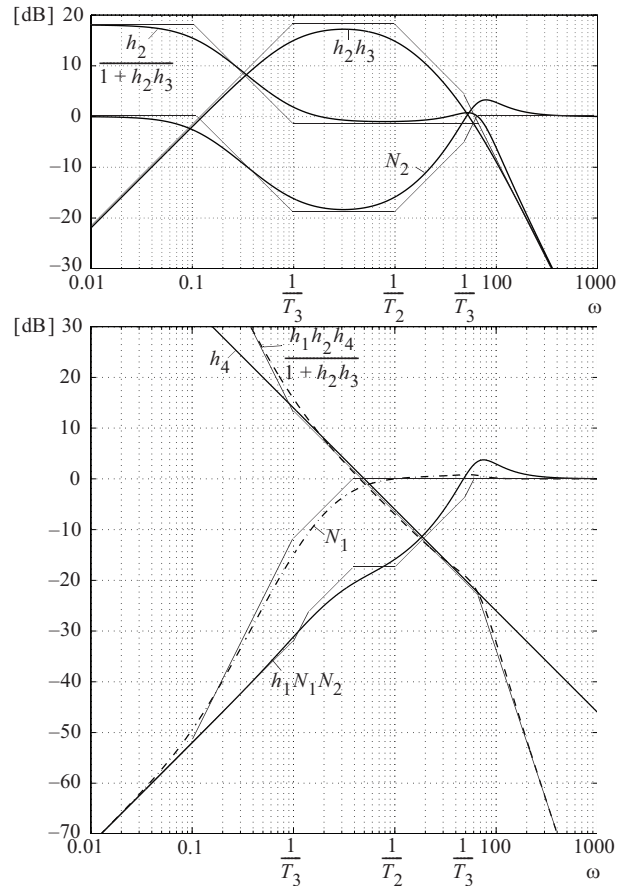
Når vi igjen antar at inngangssignalet sine amplitude er r_1 og metningsnivået u_m , vil forholdet

$$\frac{r_{1\max}}{u_m} = \left| \frac{1}{h_1N_1N_2} \right|$$

for dette systemet finnes av den asymptotiske karakteristikken for $|h_1 N_1 N_2|$ som vist i figur 12.5.

Av det foregående skulle det framgå at det er relativt lett å finne (tilnærmet) om et system som er under dimensjonering vil kunne beherske de signalamplituder som forekommer, uten at noe element går i metning. Vi ser at systemet gjerne kan ha en rekke metningselementer plassert på forskjellige steder, uten at det forandrer undersøkelsesmetoden. *Det element som først går i metning vil bli avgjørende for hele systemet*

Figur 12.5
Amplitude/fre-
kvens-diagram



12.3 EFFEKTBRUK I REGULERINGSSYSTEMER

Den momentane effekten ($P(t)$ [Watt]) som formidles av kraftorganene i et reguleringsystem vil, avhengig av arten, kunne uttrykkes ved:

$$\text{Elektrisk effekt: } P(t) = e(t)i(t) \quad \text{der} \quad \begin{array}{l} e = \text{spenning [Volt]} \\ i = \text{strøm [Ampere]} \end{array}$$

$$\text{Mekanisk effekt: } P(t) = d(t)n(t) \quad \text{der} \quad \begin{array}{l} d = \text{dreiemoment [Nm]} \\ n = \text{hastighet [rad/s]} \end{array}$$

$$\text{Hydraulisk effekt: } P(t) = p(t)q(t) \quad \text{der} \quad \begin{array}{l} p = \text{trykk [N/m}^2 = 10^5 \text{ bar]} \\ q = \text{strømning [m}^3/\text{s]} \end{array}$$

Variablene (spenning, strøm og liknende) i disse likningene er avhengige av inngangssignalene som påvirker systemet. I et lineært system med kjente parametre kan vi finne disse variablenes Laplacetransformer hver for seg. Laplacetransformen av effekten finner vi av uttrykket i (4.20). I tilfelle av for eksempel elektrisk effekt, finner vi

$$P(s) = \frac{1}{2\pi j} \int_{\Gamma} e(s-w)i(w)dw \quad (12.12)$$

Γ er her en integrasjonsbane som foreskrives avhengig av singularitetene i $e(s-w)$ og $i(w)$.

Det er mulig å finne en løsning av (12.12) i de enkleste tilfeller. Når $P(s)$ er funnet, kan $P(t)$ bestemmes ved invers Laplacetransformasjon. Av formen på denne kan vi så utlede viktige dimensjoneringsstørrelser for kraftorganene.

I uttrykkene for effekt legger vi merke til at siden begge de to størrelsene som inngår i produktformen vil øke proporsjonalt med økningen av et vilkårlig inngangssignal på systemet, vil effekten øke med kvadratet av inngangssignalets størrelse. Påtrykker vi et dobbelt så stort signal, vil altså effektbehovet bli fire ganger så stort.

Når det gjelder beregning av effektbehovet er det også ønskelig med en forenklet teknikk. Vi kan utlede en slik ved hjelp av frekvensanalytiske betraktninger.

Dersom inngangsvariabelen til et lineært system kan beskrives med en sum av sinusvingninger (om enn tilnærmet), slik som gjort i (12.9), vil de to komponentene som inngår i effektlikningene (e og i), også kunne beskrives med liknende summer der de samme frekvenskomponentene forekommer.

Vi innfører nå frekvensresponsene:

$$\begin{aligned} g_1(j\omega) &= \frac{e}{r}(j\omega) \\ g_2(j\omega) &= \frac{i}{r}(j\omega) \end{aligned} \quad (12.13)$$

og finner

$$\begin{aligned} e(t) &= r_1 |g_1(j\omega)| \sin(\omega_1 t + \angle g_1(j\omega_1)) + \dots \\ i(t) &= r_1 |g_2(j\omega)| \sin(\omega_1 t + \angle g_2(j\omega_1)) + \dots \end{aligned} \quad (12.14)$$

Uttrykket for den *momentane effekten* som framkommer ved å multiplisere de to uttrykkene i (12.14) med hverandre, vil selvsagt bli ganske komplisert hvis vi har mange forskjellige frekvenskomponenter. Men dersom vi tillater oss å *midle over en tid* som er mer enn en periode av den laveste frekvensen, finner vi at den totale effekten som formidles av systemet, kan uttrykkes ved

$$P_{\text{tot}} = \frac{1}{2} (r_1^2 |g_{11}| |g_{21}| + r_2^2 |g_{12}| |g_{22}| + r_3^2 |g_{13}| |g_{23}| + \dots) \quad (12.15)$$

der $g_{1j} = g_1(j\omega_j)$ og $g_{2j} = g_2(j\omega_j)$

Likning (12.15) framkommer fordi komponenter med forskjellig frekvens i e og i ikke har effekt med hensyn på hverandre. Denne totaleffekten består av en *aktiv* og en *reaktiv* komponent. Den **aktive effekten** er den som “koster noe”, og finnes av uttrykket:

$$\begin{aligned} P_a &= \frac{1}{2} (r_1^2 |g_{11}| |g_{21}| \cos(\angle g_{11} - \angle g_{21}) + \\ &\quad r_2^2 |g_{12}| |g_{22}| \cos(\angle g_{12} - \angle g_{22}) + \dots) \end{aligned} \quad (12.16)$$

Den **reaktive effekten** sirkulerer i systemet. Maskineriet i systemet må imidlertid dimensjoneres for å yte totaleffekten i (12.15). Størrelsene i (12.15) og (12.16) kan alle finnes av systemets amplitude/frekvens-karakteristikker.

EKSEMPEL 12.3: Effektbruk i servomekanisme

Vi skal se på en servomekanisme skissert i blokkdiagrammet i figur 12.6. d og n er henholdsvis dreiemoment og hastighet av motoren. $h_1(s)$ og $h_2(s)$ er transferfunksjoner for signal- og effektførsterkende elementer, og $h_3(s)$ er en stabiliserende tilbakekopling. J er de roterende massenes treghetsmoment referert til motorakselen, og k er giroversettingen

$$g_1(s) = \frac{d}{r}(s) = \frac{y}{r}(s) \frac{1}{k} J s^2 = M(s) \frac{1}{k} J s^2 \quad (12.17)$$

$$g_2(s) = \frac{n}{r}(s) = \frac{y}{r}(s) \frac{1}{k} s = M(s) \frac{1}{k} s \quad (12.18)$$

For det tilfellet at

$$r(t) = r_1 \sin(\omega_1 t)$$

finner vi av (12.15)

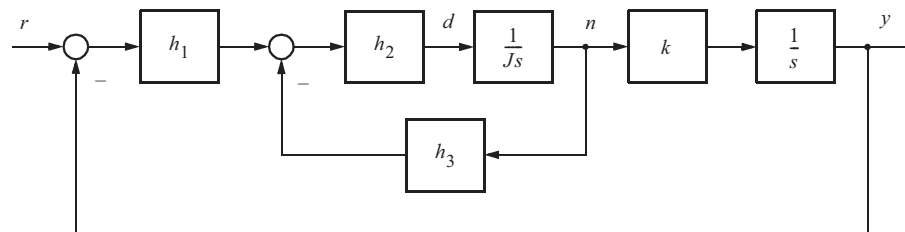
$$P_{\text{tot}} = \frac{1}{2} r_1^2 |g_{11}| |g_{21}| = \frac{J}{2} \left(\frac{r_1}{k} \right)^2 |M(j\omega_1)|^2 \omega^3 \quad (12.19)$$

Fordi $(\angle g_{11} - \angle g_{21}) = -\pi/2$ for alle ω , ser vi av (12.16) at

$$P_a = \text{aktiv effekt} = 0$$

Dette betyr bare at vi ikke har regnet med noe forbruk av effekten. Den oscillerer bare mellom effektkilden og systemets svingmasse. Selv om den aktive effekten er lik null, må selvsagt motoren dimensjoneres for å kunne håndtere denne effekttransporten.

Figur 12.6
Blokkdiagram for servomekanisme



Av (12.19) kan vi trekke følgende viktige konklusjoner:

Den totale effekten som formidles av motoren, er proporsjonal med:

- De roterende massers treghetsmoment (J).
- Kvadratet av inngangssvingningens amplitude, r_1^2 , det vil si at dersom vi ønsker dobbelt så stor svingeamplitude, fordres det fire ganger så stor motor.
- Kvadratet av følgeforholdet ($|M(j\omega)|^2$). Da $|M| \approx 1$ for $\omega < \omega_c$, betyr dette at effekten for svingefrekvenser lavere enn ω_c er tilnærmet uavhengig av systemparametrene.
- Tredje potens av svingefrekvensen, ω^3 , det vil si at hvis vi øker svingefrekvensen for $\omega < \omega_c$ til det dobbelte, vil effekten bli åtte ganger så stor. Ønsker vi å øke systemets båndbredde (ω_c) til det dobbelte, fordres det åtte ganger så stor motor. En større motor vil dessuten få større treghetsmoment, slik at forholdet blir enda verre.



Disse enkle betraktningene lærer oss at den effiktmessige dimensjoneringen av elementene i et reguleringsystem er av største viktighet. Alt for mange reguleringsystemer dimensjoneres for svakt, med den følge at de ikke tilfredsstillende rimelige spesifikasjoner.

12.4 BESTEMMELSE AV DEN REALISERBARE BÅNDBREDDEN ω_c AV ET REGULERINGSYSTEM

De konklusjonene som er trukket i avsnittene foran, om sammenhengen mellom metning/effektbegrensning og oppnåelig båndbredde, kan reformuleres på forskjellige måter. Vi tar for enkelthets skyld utgangspunkt i systemet i figur 12.6 og ser først på det tilfellet at motorens dreiemoment (d) har en metningsgrense (d_m).

Av (12.17) finner vi

$$d(s) = M(s) \frac{1}{k} J s^2 r(s) \quad (12.20)$$

Når $r(t) = r_1 \sin(\omega_1 t)$ og når

$$\begin{aligned} |M(j\omega)| &\approx 1 \text{ for } \omega < \omega_c \text{ og} \\ |M(j\omega)| &< \left(\frac{\omega_c}{\omega}\right)^2 \text{ for } \omega > \omega_c \end{aligned} \quad (12.21)$$

vil den maksimale frekvensen (ω_c) (båndbredden) som tillates uten at **dreiemomentgrensen** (d_m) nås, bli

$$\omega_{cd} = \left(\frac{d_m k}{J r_1}\right)^{1/2} \quad (12.22)$$

Tilsvarende finner vi når motoren er valgt, at den maksimale amplituden (r_1) av inngangssignalet som kan tillates, er

$$r_{1\max} = \frac{d_m k}{J \omega^2} \quad (12.23)$$

Dreiemomentgrensen (d_m) vil vanligvis være gitt og forårsaket av den kraftforsterkeren som driver motoren, og ikke av motoren selv. Dersom kraftforsterkeren er rikelig dimensjonert, kan det tenkes at det er effektbegrensningen for motoren som setter grensen.

Utfra (12.19) finner vi den maksimale frekvensen (ω_c) (båndbredden) som tillates uten at motorens **effektgrense** (P_m) nås, av

$$\omega_{c_p} = \left(\frac{2P_m k^2}{J r_1^2} \right)^{1/3} \quad (12.24)$$

Tilsvarende finner vi at når motoren er valgt, vil den maksimale amplituden (r_1) som kan tillates, være

$$r_{1\max} = \left(\frac{2P_m k^2}{J \omega^3} \right)^{1/2} \quad (12.25)$$

Ovenstående resultater (12.22)-(12.25) kan etterprøves for en konkret elektrisk motor (type INLAND TT5301A) som er karakterisert ved

$$P_m = 4000 \quad [\text{watt}]$$

$$d_m = 136 \quad [\text{Nm}]$$

$$J = 2 \cdot J_{m0} = 80 \cdot 10^{-3} \quad [\text{kgm}^2]$$

$$J_{m0} = \text{motortreghetsmoment}$$

Anta at giroversettingen er

$$k = \frac{1}{100}$$

og ønsket maksimal amplitude

$$r_1 = 1$$

Da gir (12.22)

$$\omega_{c_d} = \left(\frac{136}{100 \cdot 80 \cdot 10^{-3}} \right)^{1/2} = 4.12 \text{ [rad/s]} \quad (12.26)$$

og (12.24)

$$\omega_{c_p} = \left(\frac{2 \cdot 4000}{10^4 \cdot 80 \cdot 10^{-3}} \right)^{1/3} = 2.15 \text{ [rad/s]} \quad (12.27)$$

Når disse to svarene er litt forskjellige, er det bare en indikasjon på at metoden er tilnærmet, men vi får en god pekepinn om hva som kan oppnås.

Hadde vi antatt en giroversetting $k = 1/10$ ville vi fått

$$\omega_{c_d} = 13 \text{ [rad/s]} \quad (12.28)$$

og

$$\omega_{c_p} = 10 \text{ [rad/s]} \quad (12.29)$$

Resultater som ovenstående vil være helt avgjørende for dimensjonering av regulatorparametre og/eller knekkfrekvensen i seriekompensatoren som ble diskutert i kapittel 9.

12.5 PÅDRAGSMETNING I SYSTEMER MED INTEGRERENDE REGULATOR

Alle pådragsorganer viser i praksis “metning” i en eller annen form. En ventil kan for eksempel ikke levere mer væske enn svarende til full åpning og kan ikke mer enn stenges.

Når et pådragsorgan med metning styres av en regulator med integrerende virkning (I-regulator, PI-regulator eller PID-regulator), må vi sørge for at integratoren (integralvirkningen) ikke fortsetter å bygge seg opp selv om det etterfølgende elementet (ventilen) er gått i metning. En slik “**integrator-overlading**” (eng: integrator wind-up) vil det kunne ta lang tid å lade ut igjen.

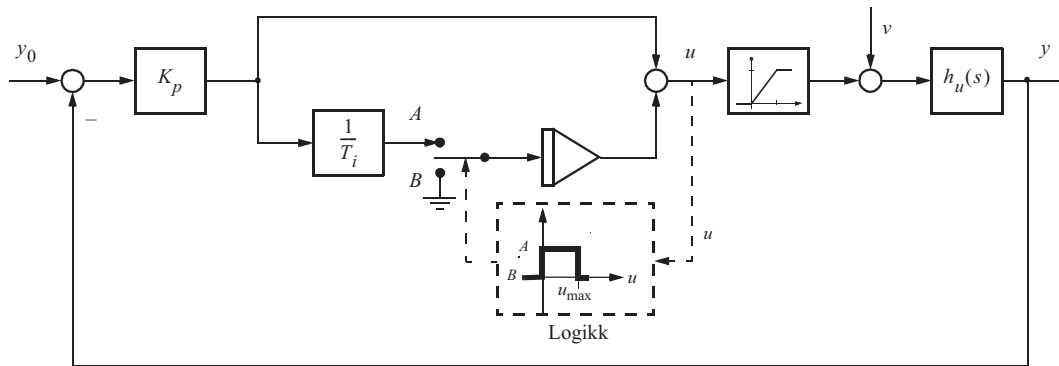
Dette problemet kan løses på flere måter. Alle kommersielt tilgjengelige reguleringsystemer er utrustet med en “**anti-overlader**” (eng.: **anti wind-up**) i en eller annen form.

Den mest vanlige løsningen benytter seg av en logisk krets (eller algoritme) som stopper integrasjonen og holder integratorens utgangsverdi, når regulatorutgangen overskrider pådragsorganets metningsgrenser (øvre og nedre).

EKSEMPEL 12.4: PI-regulator med "anti-overlader"

Figur 12.7 viser et blokkdiagram som beskriver denne metoden i et system med en PI-regulator som regulerer en prosess gjennom en ventil. Ventilen har metningsgrensene $u_{\max} = +1$ og $u_{\min} = 0$.

Figur 12.7
System med
anti-overlader



Vi antar at prosessen har transferfunksjonen

$$h_u(s) = \frac{K}{1 + Ts} e^{-\tau s} \quad (12.30)$$

Når den lineære regulatoren velges som

$$h_r(s) = K_p \frac{1 + T_i s}{T_i s}$$

med $T_i = T$, kan vi benytte forsterkningen som ble funnet i eksempel 9.6. Se også eksempel 9.18.

$$K_p = \frac{1}{K} \frac{T_i \pi}{\tau 4}$$

Med eksempelvis: $K = 1$, $T = 5$ og $\tau = 2$ gir dette

$$T_i = 5, \quad K_p = \frac{5 \pi}{2 \cdot 4} = 1.963$$

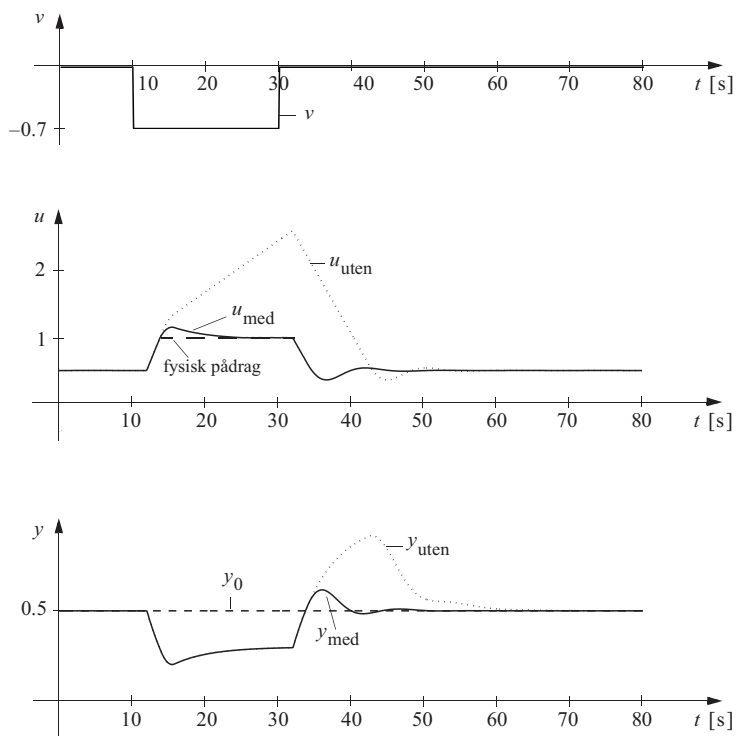
$$\omega_{180} = \frac{\pi}{4} = 0.785$$

$$\omega_c = \frac{\pi}{8} = 0.393$$

Vi antar videre at $y_0 = 0.5 =$ konstant og at forstyrrelsen v er en firkantpuls fra 0 til -0.7 av varighet $T_0 = 20$ [s].

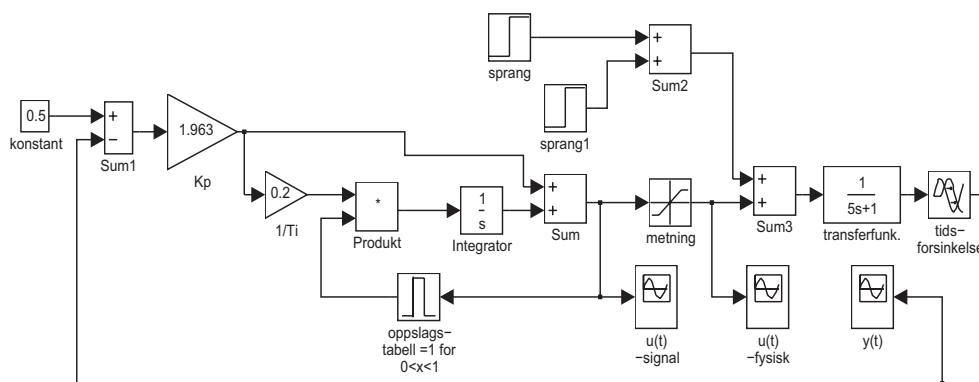
Figur 12.8 viser tidsresponsen av forskjellige variable i systemet med og uten anti-overlader-koplingen i bruk.

Figur 12.8
Respons i u og
utgang y ved
sprang i v med og
uten anti-overlader



Problemet oppstår altså *etter* at forstyrrelsen er over, når integratoren feilaktig er “opp-ladet”. Responsen i figur 12.8 er generert med simulink-filen i figur 12.9.

Figur 12.9
Simulinkvindu



Appendiks

Matrisealgebra

A.1 Lineære transformasjoner. Matriser 572

A.2 Derivasjon av matriser og vektorer 580

Dette appendiks skal gi en kort innføring i matrisealgebra.

En mengde av tall som er ordnet i rekker og kolonner

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdot & \cdot & a_{mn} \end{bmatrix} \quad (\text{A.1})$$

kalles for en matrise. Tallene i matrisen, a_{ij} , kalles for elementer i matrisen. Den første indeksen, i , indikerer nummeret på raden som elementet befinner seg i, og den andre indeksen, j , indikerer kolonnen.

Vi kan også symbolisere matrisen som vist til høyre i (A.1), dvs. som en “mengde av elementer a_{ij} ”. Dette vil vi benytte senere.

Vi kaller en matrise som har m rader og n kolonner for en $m \times n$ -matrise.

Dersom antallet kolonner i en matrise er 1, $n = 1$, har vi en kolonnevektor. Dersom antallet rader i en matrise er 1, $m = 1$, har vi en radvektor.

A.1 LINEÆRE TRANSFORMASJONER. MATRISER

Dersom to sett av tall x_1, x_2, \dots, x_n og y_1, y_2, \dots, y_m er avhengig av hverandre på en slik måte at hvert element i det ene settet er en lineær kombinasjon av elementene i det andre settet, sier vi at det ene settet er avledet av det andre ved en lineær transformasjon. Vi kan da ha

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\cdot \\ &\cdot \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{aligned} \tag{A.2}$$

Man kan fremstille likning (A.2) som to kolonne vektorer og en matrise

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{A.3}$$

$$\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_m \end{bmatrix}^T = m\text{-vektor}$$

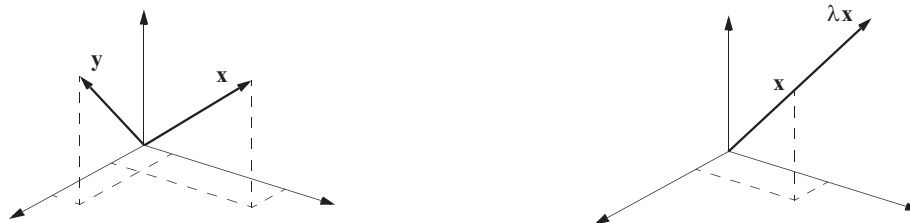
$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T = n\text{-vektor}$$

der T betyr transponert.

A.1.1 Egenverdier og egenvektorer

Vi skal til å begynne med betrakte lineære transformasjoner der vektorene \mathbf{y} og \mathbf{x} har samme dimensjon, n . \mathbf{A} blir da en kvadratisk $n \times n$ -matrise, og \mathbf{y} og \mathbf{x} vil være vektorer i det samme vektorrommet. Likning (A.3) uttrykker hvordan en vektor \mathbf{x} kan transformeres over til en vektor \mathbf{y} i det samme n -dimensjonale rommet. Dette er illustrert med et 3-dimensjonalt tilfelle i A.1.

Figur A.1
Vektorer i et tredimensjonalt rom



Et spesielt tilfelle er når \mathbf{y} er et skalart multiplum av \mathbf{x} . Vi har da

$$\mathbf{y} = \mathbf{Ax} = \lambda \mathbf{x} \quad (\text{A.4})$$

der λ er en skalar.

Vi kan omformulere (A.4) til

$$\lambda \mathbf{x} - \mathbf{Ax} = (\lambda \mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0} \quad (\text{A.5})$$

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & 1 \end{bmatrix} = \text{identitetsmatrisen}$$

(A.4) er et sett homogene likninger med komponentene til \mathbf{x} som ukjente. For at en ikke-triviell løsning (i.e. alle komponentene er null) av dette systemet skal eksistere, må determinanten til systemet være lik null,

$$\det(\mathbf{A} - \lambda \mathbf{I}) = |\mathbf{A} - \lambda \mathbf{I}| = 0 \quad (\text{A.6})$$

Fullt utskrevet vil (A.6) anta formen

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & a_{nn} - \lambda \end{vmatrix} \quad (\text{A.7})$$

Venstresiden av (A.7) er et polynom av grad n i λ . Vi kaller dette polynomet for det *karakteristiske polynomet* til \mathbf{A} . Likning (A.6) kalles den *karakteristiske likningen* til \mathbf{A} .

En vektor $\mathbf{x} = \mathbf{m}$ som tilfredsstillers (A.4) kalles en *egenvektor* og den tilhørende skalare størrelsen, λ_i , kalles en *egenverdi*. En $n \times n$ -matrise som den i (A.4) vil ha n egenvektorer og n egenverdier.

En $n \times n$ -matrise har minst 1 og høyst n forskjellige egenverdier, fordi egenverdiene er røttene til det karakteristiske polynomet. Likning (A.7) viser at egenverdiene vil være lik elementene på diagonalen for en diagonalmatrise ($a_{ij} = 0$, $i \neq j$).

EKSEMPEL A.1: Egenverdiene til en matrise

En 2×2 -matrise

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

vil ved innsetting i (A.7) og utregning få en **karakteristisk likning**

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0$$

som har løsningsene

$$\lambda_1 = \frac{a_{11} + a_{22} + \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2}$$

$$\lambda_2 = \frac{a_{11} + a_{22} - \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2}$$

■

Egenverdiene kan være reelle eller kompleks konjugerte. Dersom et systems orden er større enn 2, er det mest hensiktsmessig å bestemme egenverdiene ved hjelp av en datamaskin (f.eks. i MATLAB).

Når egenverdiene er bestemt, kan vi finne et tilsvarende antall egenvektorer ved å sette de forskjellige egenverdiene inn i (A.4). Siden determinanten for likningssystemet er lik null, kan vi ikke finne alle komponentene i \mathbf{m}_i . Vi kan imidlertid velge en av elementene i \mathbf{m}_i og bestemme de andre ut i fra denne. Dette skyldes at alle vektorer med samme retning som en egenvektor, vil være egenvektorer til \mathbf{A} . Egenvektorenes lengde er derfor ubestemte. Vi bestemmer ikke egenvektorene til en matrise \mathbf{A} , men et sett med egenvektorer.

Hvis egenverdiene til et sett egenvektorer er forskjellige, er egenvektorene **lineært uavhengige**. At vektorer er lineært uavhengig, betyr at en eller flere av vektorene ikke kan dannes ved en lineær kombinasjon av de øvrige. Dette skal vise seg å ha betydning senere. Egenvektorer kan i likhet med egenverdier ha reelle og komplekse komponenter. En egenvektor med komplekse komponenter kan vanskelig visualiseres i et enkelt aksesystem som i figur A.1 siden det trenges en ny dimensjon for å skille reelle og imaginære deler.

A.1.2 Matriseoperasjoner

Med matriser kan man utføre en lang rekke operasjoner som svarer til de som utføres med skalare størrelser. En kan f.eks. utføre multiplikasjon av to matriser. Rekkefølgen av matrisene er helt avgjørende for resultatet. Dette er ikke tilfelle for skalare størrelser. I dette avsnittet skal 2 matriseoperasjoner diskuteres, *transponering* og *inversjon*.

En matrise

$$\mathbf{A} = \{a_{ij}\}$$

er karakterisert ved sitt generelle element a_{ij} i rad i og kolonne j . Den **transponerte av matrisen**

$$\mathbf{A}^T = \{a_{ji}\}$$

er karakterisert ved at rader og kolonner er byttet om i forhold til den opprinnelig matrisen.

Når \mathbf{A} , \mathbf{B} , \mathbf{C} er vilkårlige matriser har vi $(\mathbf{ABC})^T = \mathbf{C}^T \mathbf{B}^T \mathbf{A}^T$.

Den **inverse matrisen** til en kvadratisk matrise \mathbf{A} er definert som den matrisen vi må multiplisere \mathbf{A} med for å få identitetsmatrisen. Vi betegner den inverse matrisen med \mathbf{A}^{-1} ,

$$\mathbf{A}^{-1} \mathbf{A} = \mathbf{I} \quad (\text{A.8})$$

eller

$$\mathbf{A} \mathbf{A}^{-1} = \mathbf{I} \quad (\text{A.9})$$

Den inverse av en matrise \mathbf{A} er gitt ved

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{|\mathbf{A}|} = \frac{\text{adj } \mathbf{A}}{\det \mathbf{A}} \quad (\text{A.10})$$

$$\text{adj } \mathbf{A} = \{\mathbf{A}^{ij}\}^T \quad (\text{A.11})$$

$$\mathbf{A}^{ij} = (-1)^{i+j} \mathbf{U}_{ij} = \text{kofaktoren til } a_{ij} \text{ i } \mathbf{A}$$

$$\mathbf{U}_{ij} = \text{underdeterminanten til } \mathbf{A}$$

Underdeterminanten til \mathbf{A} fremkommer ved å stryke rad nr. i og kolonne nr. j i determinanten til \mathbf{A} .

EKSEMPEL A.2: Den inverse av en matrise

Vi skal finne den inverse av

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

For å finne $\text{adj } \mathbf{A}$ må vi først beregne kofaktorene til \mathbf{A}

$$\begin{aligned} \mathbf{A}^{11} &= (-1)^2(2 \cdot 2 - 1 \cdot 1) = 3 \\ \mathbf{A}^{12} &= (-1)^3(1 \cdot 2 - 1 \cdot 1) = -1 \\ \mathbf{A}^{13} &= (-1)^4(1 \cdot 1 - 2 \cdot 1) = -1 \\ \mathbf{A}^{21} &= (-1)^3(0 \cdot 2 - 2 \cdot 1) = 2 \\ \mathbf{A}^{22} &= (-1)^4(1 \cdot 2 - 2 \cdot 1) = 0 \\ \mathbf{A}^{23} &= (-1)^5(1 \cdot 1 - 0 \cdot 1) = -1 \\ \mathbf{A}^{31} &= (-1)^4(0 \cdot 1 - 2 \cdot 2) = -4 \\ \mathbf{A}^{32} &= (-1)^5(1 \cdot 1 - 2 \cdot 1) = 1 \\ \mathbf{A}^{33} &= (-1)^6(1 \cdot 2 - 0 \cdot 1) = 2 \end{aligned} \tag{A.12}$$

Vi setter deretter sammen $\text{adj } \mathbf{A}$ i henhold til (A.11)

$$\text{adj } \mathbf{A} = \{\mathbf{A}^{ij}\}^T = \begin{bmatrix} 3 & 2 & -4 \\ -1 & 0 & 1 \\ -1 & -1 & 2 \end{bmatrix} \tag{A.13}$$

Deretter finner vi determinanten til \mathbf{A}

$$\det \mathbf{A} = 1 \cdot (2 \cdot 2 - 1 \cdot 1) - 0 \cdot (1 \cdot 2 - 1 \cdot 1) + 2 \cdot (1 \cdot 1 - 2 \cdot 1) = 1 \tag{A.14}$$

Likning (A.10), (A.13) og (A.14) gir

$$\mathbf{A}^{-1} = \{\mathbf{A}^{ij}\}^T = \begin{bmatrix} 3 & 2 & -4 \\ -1 & 0 & 1 \\ -1 & -1 & 2 \end{bmatrix}$$

■

Av (A.10) fremgår det at betingelsen for at den inverse matrisen skal eksistere, er at $|\mathbf{A}| \neq 0$. En matrise som har denne egenskapen, kalles **ikke-singulær** eller **regulær**. Dersom $|\mathbf{A}| = 0$ kalles matrisen **singulær**.

For de ikke-singulære matrisene \mathbf{A} , \mathbf{B} , \mathbf{C} gjelder $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

A.1.3 Similaritetstransformasjon. Diagonalisering

\mathbf{A} undergår en *similaritetstransformasjon* dersom \mathbf{S} er en ikke-singulær $n \times n$ -matrise, og en ny matrise \mathbf{B} dannes fra \mathbf{A} ved transformasjonen

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} \quad (\text{A.15})$$

Vi er interessert i hvordan det går med egenverdiene og egenvektorene til en matrise som følge av similaritetstransformasjonen. Vi kaller en egenvektor til \mathbf{A} for \mathbf{m}_A og en egenvektor til \mathbf{B} for \mathbf{m}_B .

$$\mathbf{B}\mathbf{m}_B = \lambda_B\mathbf{m}_B \quad (\text{A.16})$$

Innsetting av (A.15) gir

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{m}_B = \lambda_B\mathbf{m}_B$$

Vi formultipliserer med \mathbf{S}

$$\mathbf{A}(\mathbf{S}\mathbf{m}_B) = \lambda_B(\mathbf{S}\mathbf{m}_B)$$

Vi sammenligner dette med

$$\mathbf{A}\mathbf{m}_A = \lambda_A\mathbf{m}_A \quad (\text{A.17})$$

Resultatet blir at

$$\lambda_A = \lambda_B \quad (\text{A.18})$$

og

$$\mathbf{m}_A = \mathbf{S}\mathbf{m}_B \quad (\text{A.19})$$

\mathbf{A} og \mathbf{B} har altså de samme egenverdiene mens egenvektorene er gitt av en lineær transformasjon over matrisen \mathbf{S} .

En spesiell similaritetstransformasjon transformerer \mathbf{A} til en diagonalmatrise.

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{M} = \mathbf{\Lambda} \quad (\text{A.20})$$

$\mathbf{\Lambda}$ er en diagonalmatrise. Diagonalleddene i $\mathbf{\Lambda}$ er egenverdiene til \mathbf{A} .

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix} \quad (\text{A.21})$$

For å finne \mathbf{M} formultipliserer vi (A.20) med \mathbf{M} og får

$$\mathbf{A}\mathbf{M} = \mathbf{M}\mathbf{\Lambda} \quad (\text{A.22})$$

Utvikling av høyresiden gir

$$\mathbf{M}\mathbf{\Lambda} = \begin{bmatrix} m_{11} & m_{12} & \cdot & \cdot & m_{1n} \\ m_{21} & m_{22} & \cdot & \cdot & m_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{n1} & m_{n2} & \cdot & \cdot & m_{nn} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix} \quad (\text{A.23})$$

$$= \begin{bmatrix} m_{11}\lambda_1 & m_{12}\lambda_2 & \cdot & \cdot & m_{1n}\lambda_n \\ m_{21}\lambda_1 & m_{22}\lambda_2 & \cdot & \cdot & m_{2n}\lambda_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{n1}\lambda_1 & m_{n2}\lambda_2 & \cdot & \cdot & m_{nn}\lambda_n \end{bmatrix}$$

Vi innfører vektorer for å representere kolonnene i \mathbf{M}

$$\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n] \quad (\text{A.24})$$

(A.22) og (A.23) kan nå skrives slik

$$\mathbf{A} [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n] = [\lambda_1\mathbf{m}_1, \lambda_2\mathbf{m}_2, \dots, \lambda_n\mathbf{m}_n] \quad (\text{A.25})$$

Dette viser at

$$\mathbf{A}\mathbf{m}_i = \lambda_i\mathbf{m}_i \quad (\text{A.26})$$

Likning (A.26) viser at kolonnene i \mathbf{M} er egenvektorer til \mathbf{A} .

Dersom en matrise \mathbf{A} er ikke-singulær og ikke har sammenfallende egenverdier slik at egenvektorene ikke er lineært avhengige, vil $|\mathbf{M}| \neq 0$ slik at \mathbf{M}^{-1} eksisterer. Det er alltid mulig å finne en similaritetstransformasjon som transformerer \mathbf{A} til en diagonalmatrise dersom disse vilkårene er oppfylt.

EKSEMPEL A.3: Egenvektorene og egenverdiene til en matrise

Vi tar utgangspunkt i matrisen

$$\mathbf{A} = \begin{bmatrix} -7 & 4 & -4 \\ 4 & -7 & -2 \\ -4 & 4 & -13 \end{bmatrix}$$

Vi vil bestemme matrisens egenverdier og deretter et sett av egenvektorer. Likning (A.7) gir

$$\lambda^3 + 27\lambda^2 + 207\lambda + 405 = 0$$

Røttene til denne likningen som er lettest å finne ved hjelp av datamaskin (f.eks. ved hjelp av MATLAB), er $\lambda_1 = -3$, $\lambda_2 = -9$ og $\lambda_3 = -15$. Likning (A.26) benyttes for å bestemme de tre egenvektorene. Egenvektoren \mathbf{m}_1 bestemmes på følgende måte

$$\begin{aligned} -7m_{11} + 4m_{21} - 4m_{31} &= -3m_{11} \\ 4m_{11} - 7m_{21} - 2m_{31} &= -3m_{21} \\ -4m_{11} + 4m_{21} - 13m_{31} &= -3m_{31} \end{aligned} \quad (\text{A.27})$$

Vi kan velge ett av elementene i \mathbf{m}_1 fritt fordi egenvektoren kan ha vilkårlig lengde. Vi velger $m_{11} = 1$. Den første likningen i (A.27) gir $m_{21} - m_{31} = 1$. Ved også å se på den andre likningen i (A.27), kommer vi fram til $m_{21} = 1$ og $m_{31} = 0$. Dersom vi gjør det samme for de to øvrige egenvektorene, finner vi til slutt en **egenvektormatrise**

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

Vi vil nå bestemme den inverse av egenvektormatrisen ved hjelp av (A.10).

$$\mathbf{M}^{-1} = \frac{\text{adj } \mathbf{M}}{\det \mathbf{M}} = \frac{1}{3} \begin{bmatrix} 2 & 1 & -1 \\ -2 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Vi kan kontrollere resultatet ved å benytte (A.20)

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{M} = \frac{1}{3} \begin{bmatrix} 2 & 1 & -1 \\ -2 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -7 & 4 & -4 \\ 4 & -7 & -2 \\ -4 & 4 & -13 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & -15 \end{bmatrix} \quad (\text{A.28})$$

Vi ser at den siste matrisen er lik egenverdimatrisen til \mathbf{A} .



A.2 DERIVASJON AV MATRISER OG VEKTORER

Den deriverte av en tidsvarierende matrise $\mathbf{X}(t)$ er per definisjon matrisen av deriverte elementer:

$$\dot{\mathbf{X}}(t) = \{\dot{x}_{ij}(t)\} \quad (\text{A.29})$$

Den deriverte av en vektor (som kan betraktes som en $(n \times 1)$ “matrise” blir da

$$\dot{\mathbf{x}}(t) = \{\dot{x}_i(t)\} \quad (\text{A.30})$$

Den deriverte av en konstant matrise ganger en tidsfunksjon er:

$$\frac{d}{dt}[f(t)\mathbf{A}] = \frac{d}{dt}\{f(t)a_{ij}\} = \dot{f}(t)\mathbf{A} \quad (\text{A.31})$$

Den deriverte av et produkt av to matriser bestående av tidsfunksjoner $\mathbf{X}(t)\mathbf{Y}(t)$ er:

$$\frac{d(\mathbf{X}(t)\mathbf{Y}(t))}{dt} = \mathbf{X}(t)\dot{\mathbf{Y}}(t) + \dot{\mathbf{X}}(t)\mathbf{Y}(t) \quad (\text{A.32})$$

Vi merker oss at regelen svarer til den vi kjenner fra det skalare tilfellet. Det eneste vi må passe på, er multiplikasjonsrekkefølgen.

Bevis av (A.32):

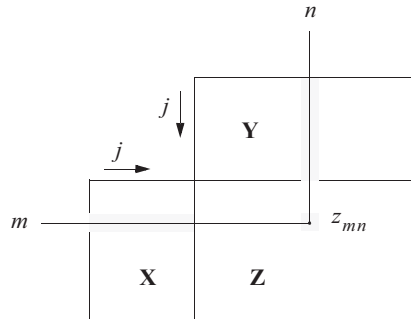
Vi definerer $\mathbf{Z}(t) = \mathbf{X}(t)\mathbf{Y}(t)$.

Vi har

$$\mathbf{Z}(t) = \{z_{mn}\} = \left\{ \sum_j x_{mj} y_{jn} \right\} \quad (\text{A.33})$$

(A.33) innser vi ved å studere figur A.2.

Figur A.2
Matrise-
multiplikasjon



Vi deriverer (A.33):

$$\begin{aligned} \dot{\mathbf{Z}}(t) = \{\dot{z}_{mn}\} &= \left\{ \frac{d}{dt} \left(\sum_j x_{mj} y_{jn} \right) \right\} = \left\{ \sum_j (x_{mj} \dot{y}_{jn} + \dot{x}_{mj} y_{jn}) \right\} \\ &= \left\{ \sum_j x_{mj} \dot{y}_{jn} \right\} + \left\{ \sum_j \dot{x}_{mj} y_{jn} \right\} = \mathbf{X}\dot{\mathbf{Y}} + \dot{\mathbf{X}}\mathbf{Y} \end{aligned} \quad (\text{A.34})$$

▲

(A.32) kan brukes ved utledning av (3.22) slik som skissert i fotnoten på samme side. Dette overlater vi til den interesserte leser.

Laplaceformasjons-par

TABELL B.1 Laplaceformasjons-par

	$f(s)$	$f(t)$	
1.	1	$\delta(t)$, enhetspuls ved $t = 0$
2.	$\frac{1}{s}$	$\mu_1(t)$, enhetssprang ved $t = 0$
3.	$\frac{1}{s^2}$	$\mu_2(t) = t\mu_1(t)$, enhetsrampe
4.	$\frac{1}{s^n}$	$\frac{1}{(n-1)!}t^{n-1}\mu_1(t)$, n er et positivt heltall
5.	$\frac{1}{s}e^{-as}$	$\mu_1(t-a)$, enhetssprang i $t = a$
6.	$\frac{1}{s}(1 - e^{-as})$	$\mu_1(t) - \mu_1(t-a)$, rektangulær puls
7.	$\frac{1}{s+a}$	e^{-at}	
8.	$\frac{1}{s(s+a)}$	$\frac{1}{a}(1 - e^{-at})$	
9.	$\frac{1}{(s+a)(s+b)}$	$\frac{1}{b-a}(e^{-at} - e^{-bt})$	
10.	$\frac{1}{(s+a)^2 + b^2}$	$\frac{1}{b}e^{-at} \sin(bt)$	
11.	$\frac{1}{s^2 + 2\zeta\omega_0s + \omega_0^2}$	$\frac{1}{\omega_0\sqrt{1-\zeta^2}}e^{-\zeta\omega_0t} \sin(\sqrt{1-\zeta^2}\omega_0t)$	

TABELL B.1 Laplacetransformasjons-par

	$f(s)$	$f(t)$
12.	$\frac{\omega}{s^2 + \omega^2}$	$\sin(\omega t)$
13.	$\frac{s}{(s+a)(s+b)}$	$\frac{1}{a-b}(ae^{-at} - be^{-bt})$
14.	$\frac{s+\alpha}{(s+a)(s+b)}$	$\frac{1}{b-a}((\alpha-a)e^{-at} - (\alpha-b)e^{-bt})$
15.	$\frac{s+a}{(s+a)^2 + b^2}$	$e^{-at} \cos(bt)$
16.	$\frac{s+\alpha}{(s+a)^2 + b^2}$	$\frac{\sqrt{(\alpha-a)^2 + b^2}}{b} e^{-at} \sin(bt + \varphi)$, $\varphi = \arctan\left(\frac{b}{\alpha-a}\right)$
17.	$\frac{s}{s^2 + \omega^2}$	$\cos(\omega t)$
18.	$\frac{s+\alpha}{s^2 + \omega^2}$	$\frac{\sqrt{\alpha^2 + \omega^2}}{\omega} \sin(\omega t + \varphi)$, $\varphi = \arctan\left(\frac{\omega}{\alpha}\right)$
19.	$\frac{s \sin \phi + \omega \cos \phi}{s^2 + \omega^2}$	$\sin(\omega t + \phi)$
20.	$\frac{1}{s(s+a)(s+b)}$	$\frac{1}{ab} \left(1 - \frac{b}{b-a} e^{-at} + \frac{a}{b-a} e^{-bt}\right)$
21.	$\frac{1}{(s+a)(s+b)(s+c)}$	$\frac{e^{-at}}{(b-a)(c-a)} + \frac{e^{-bt}}{(c-b)(a-b)} + \frac{e^{-ct}}{(a-c)(b-c)}$
22.	$\frac{1}{s((s+a)^2 + b^2)}$	$\frac{1}{a^2 + b^2} + \frac{1}{b\sqrt{a^2 + b^2}} e^{-at} \sin(bt - \varphi)$, $\varphi = \arctan\left(\frac{b}{-a}\right)$
23.	$\frac{1}{s(s^2 + 2\zeta\omega_0 s + \omega_0^2)}$	$\frac{1}{\omega_0^2} - \frac{1}{\omega_0^2 \sqrt{1 - \zeta^2}} e^{-\zeta\omega_0 t} \cos(\sqrt{1 - \zeta^2}\omega_0 t - \varphi)$, $\varphi = \arcsin \zeta$
24.	$\frac{1}{(s+c)((s+a)^2 + b^2)}$	$\frac{e^{-ct}}{(c-a)^2 + b^2} + \frac{e^{-at} \sin(bt - \varphi)}{b\sqrt{(c-a)^2 + b^2}}$, $\varphi = \arctan\left(\frac{b}{c-a}\right)$

TABELL B.1 Laplacetransformasjons-par

	$f(s)$	$f(t)$
25.	$\frac{1}{s(s^2 + \omega^2)}$	$\frac{1}{\omega^2}(1 - \cos(\omega t))$
26.	$\frac{1}{(s+a)(s^2 + \omega^2)}$	$\frac{e^{-at}}{a^2 + \omega^2} + \frac{1}{\omega\sqrt{a^2 + \omega^2}} \sin(\omega t - \varphi)$, $\varphi = \arctan\left(\frac{\omega}{a}\right)$
27.	$\frac{1}{s(s+a)^2}$	$\frac{1}{a^2}(1 - e^{-at} - ate^{-at})$
28.	$\frac{1}{s^2(s+a)}$	$\frac{1}{a^2}(at - 1 + e^{-at})$
29.	$\frac{s + \alpha}{s(s+a)(s+b)}$	$\frac{1}{ab} \left(\alpha - \frac{b(\alpha - a)}{b - a} e^{-at} + \frac{a(\alpha - b)}{b - a} e^{-bt} \right)$
30.	$\frac{s + \alpha}{(s+a)(s+b)(s+c)}$	$\frac{(\alpha - a)e^{-at}}{(b - a)(c - a)} + \frac{(\alpha - b)e^{-bt}}{(c - b)(a - b)} + \frac{(\alpha - c)e^{-ct}}{(a - c)(b - c)}$
31.	$\frac{s + \alpha}{s((s+a)^2 + b^2)}$	$\frac{\alpha}{a^2 + b^2} + \frac{1}{b} \sqrt{\frac{(\alpha - a)^2 + b^2}{a^2 + b^2}} e^{-at} \sin(bt + \varphi)$ $\varphi = \arctan\left(\frac{b}{\alpha - a}\right) - \arctan\left(\frac{b}{-a}\right)$
32.	$\frac{s + \alpha}{s(s^2 + \omega^2)}$	$\frac{\alpha}{\omega^2} - \frac{\sqrt{\alpha^2 + \omega^2}}{\omega^2} \cos(\omega t + \varphi)$, $\varphi = \arctan\left(\frac{\omega}{\alpha}\right)$
33.	$\frac{s + \alpha}{s(s+a)^2}$	$\frac{1}{a^2}(\alpha - \alpha e^{-at} + a(a - \alpha)te^{-at})$
34.	$\frac{s^2 + \alpha_1 s + \alpha_0}{s(s+a)(s+b)}$	$\frac{\alpha_0}{ab} + \frac{a^2 - \alpha_1 a + \alpha_0}{a(a-b)} e^{-at} - \frac{b^2 - \alpha_1 b + \alpha_0}{b(a-b)} e^{-bt}$
35.	$\frac{s^2 + \alpha_1 s + \alpha_0}{s((s+a)^2 + b^2)}$	$\frac{\alpha_0}{a^2 + b^2} + \frac{\sqrt{(a^2 - b^2 - \alpha_1 a + \alpha_0)^2 + b^2(\alpha_1 - 2a)^2}}{b\sqrt{a^2 + b^2}} e^{-at} \sin(bt + \varphi)$ $\varphi = \arctan\left(\frac{b(\alpha_1 - 2a)}{a^2 - b^2 - \alpha_1 a + \alpha_0}\right) - \arctan\left(\frac{b}{-a}\right)$

TABELL B.1 Laplacetransformasjons-par

	$f(s)$	$f(t)$
36.	$\frac{1}{s(s+c)((s+a)^2+b^2)}$	$\frac{1}{c(a^2+b^2)} - \frac{e^{-ct}}{c((c-a)^2+b^2)} + \frac{e^{-at} \sin(bt-\varphi)}{b\sqrt{a^2+b^2}\sqrt{(c-a)^2+b^2}}$ $\varphi = \arctan\left(\frac{b}{-a}\right) + \arctan\left(\frac{b}{c-a}\right)$
37.	$\frac{1}{(s^2+\omega^2)((s+a)^2+b^2)}$	$\frac{(1/\omega)\sin(\omega t + \varphi_1) + (1/b)e^{-at}\sin(bt + \varphi_2)}{\sqrt{4a^2\omega^2 + (a^2+b^2-\omega^2)^2}}$ $\varphi_1 = \arctan\left(\frac{-2a\omega}{a^2+b^2-\omega^2}\right) \quad \varphi_2 = \arctan\left(\frac{2ab}{a^2-b^2+\omega^2}\right)$
38.	$\frac{s+\alpha}{s^2((s+a)^2+b^2)}$	$\frac{1}{c^2}\left(\alpha t + 1 - \frac{2\alpha a}{c^2}\right) + \frac{\sqrt{(\alpha-a)^2+b^2}}{bc^2}e^{-at}\sin(bt+\varphi)$ $c^2 = a^2 + b^2$ $\varphi = 2\arctan\left(\frac{b}{a}\right) + \arctan\left(\frac{b}{\alpha-a}\right)$
39.	$\frac{s+\alpha}{s(s+c)((s+a)^2+b^2)}$	$\frac{\alpha}{c(a^2+b^2)} + \frac{(c-\alpha)e^{-ct}}{c((c-a)^2+b^2)}$ $+ \frac{\sqrt{(\alpha-a)^2+b^2}}{b\sqrt{a^2+b^2}\sqrt{(c-a)^2+b^2}}e^{-at}\sin(bt+\varphi)$ $\varphi = \arctan\left(\frac{b}{\alpha-a}\right) - \arctan\left(\frac{b}{-a}\right) - \arctan\left(\frac{b}{c-a}\right)$
40.	$\frac{s+\alpha}{(s^2+\omega^2)((s+a)^2+b^2)}$	$\frac{1}{\omega}\sqrt{\frac{\alpha^2+\omega^2}{c}}\sin(\omega t + \varphi_1) + \frac{1}{b}\sqrt{\frac{(\alpha-a)^2+b^2}{c}}e^{-at}\sin(bt + \varphi_2)$ $c = (2a\omega)^2 + (a^2+b^2-\omega^2)^2$ $\varphi_1 = \arctan\left(\frac{\omega}{\alpha}\right) - \arctan\left(\frac{2a\omega}{a^2+b^2+\omega^2}\right)$ $\varphi_2 = \arctan\left(\frac{b}{\alpha-a}\right) + \arctan\left(\frac{2ab}{a^2-b^2+\omega^2}\right)$

TABELL B.1 Laplacetransformasjons-par

	$f(s)$	$f(t)$
41.	$\frac{s^2 + \alpha_1 s + \alpha_0}{s^2(s+a)(s+b)}$	$\frac{\alpha_1 + \alpha_0 t}{ab} - \frac{\alpha_0(a+b)}{(ab)^2} - \frac{1}{a-b} \left(1 - \frac{\alpha_1}{a} + \frac{\alpha_0}{a^2}\right) e^{-at}$ $- \frac{1}{a-b} \left(1 - \frac{\alpha_1}{b} + \frac{\alpha_0}{b^2}\right) e^{-bt}$
42.	$\frac{1}{(s+a)^n}$	$\frac{1}{(n-1)!} t^{n-1} e^{-at}$, n er et positivt heltall
43.	$e^{\left(\frac{\sigma^2 s^2}{2}\right)}$	$\frac{1}{\sqrt{2\pi} \cdot \sigma} e^{\left(\frac{-t^2}{2\sigma^2}\right)}$

Utledning av styrbarhetskriteriet

Gitt systemet

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx}\end{aligned}\tag{C.1}$$

Systemet sies å være styrbart hvis det finnes en pådragsvektor $\mathbf{u}(\tau)$, $0 \leq \tau \leq t$ som er slik at systemet bringes fra en vilkårlig starttilstand $\mathbf{x}(0)$ til en vilkårlig slutttilstand $\mathbf{x}(t)$ der $t > 0$, men ellers vilkårlig.

Kriteriet for styrbarhet er som følger: Vi danner *styrbarhetsmatrisen*

$$\mathbf{Q}_s = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}\tag{C.2}$$

som har dimensjon $n \times nr$.

Systemet (C.1) er styrbart hvis og bare hvis $\text{rang}(\mathbf{Q}_s) = n$

Før vi starter med beviset, forenkler vi problemet litt. Den generelle løsningen av (C.1) er

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{Bu}(\tau)d\tau\tag{C.3}$$

Vi definerer $\mathbf{x}'(t) = \mathbf{x}(t) - e^{\mathbf{A}t}\mathbf{x}(0)$ og får

$$\mathbf{x}'(t) = \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{Bu}(\tau)d\tau\tag{C.4}$$

Siden $\mathbf{x}'(t)$ alltid kan avledes av $\mathbf{x}(t)$ og $\mathbf{x}(0)$, kan vi definere styrbarhet ut fra (C.4): Å kunne bringe systemet til en vilkårlig slutttilstand når man starter i *nulltilstanden*. Fra nå av skriver vi $\mathbf{x}(t)$ (dropper “merket”) og fører beviset med utgangspunkt i (C.4).

Bevis:

Vi deler intervallet $(0, t)$ i n like deler

$$\delta = \frac{t}{n} \quad (\text{C.5})$$

Innenfor hvert intervall forutsettes pådraget å være konstant. Vi har en sekvens av pådrag u_0, \dots, u_{n-1} . Vi skriver $\mathbf{x}(t)|_{t=i\delta} = \mathbf{x}(i\delta) = \mathbf{x}[i]$, og får med utgangspunkt i (C.4):

$$\mathbf{x}[1] = \int_0^{\delta} e^{\mathbf{A}(\delta-\tau)} \mathbf{B} d\tau \mathbf{u}(0) = \mathbf{D} \mathbf{u}(0) \quad (\text{C.6})$$

Her er

$$\mathbf{D} = \mathbf{A}^{-1} (e^{\mathbf{A}\delta} - \mathbf{I}) \mathbf{B} = \left(\delta \mathbf{I} + \frac{\mathbf{A}\delta^2}{2!} + \frac{\mathbf{A}^2\delta^3}{3!} + \dots \right) \mathbf{B} \quad (\text{C.7})$$

Gjentatt bruk av (C.4) for $i = 2, 3, \dots, n$ gir

$$\begin{aligned} \mathbf{x}[2] &= e^{\mathbf{A}\delta} \mathbf{x}[1] + \mathbf{D} \mathbf{u}[1] = e^{\mathbf{A}\delta} \mathbf{D} \mathbf{u}[0] + \mathbf{D} \mathbf{u}[1] \\ \mathbf{x}[3] &= e^{2\mathbf{A}\delta} \mathbf{D} \mathbf{u}[0] + e^{\mathbf{A}\delta} \mathbf{D} \mathbf{u}[1] + \mathbf{D} \mathbf{u}[2] \\ &\vdots \\ \mathbf{x}[n] &= e^{(n-1)\mathbf{A}\delta} \mathbf{D} \mathbf{u}[0] + \dots + e^{\mathbf{A}\delta} \mathbf{D} \mathbf{u}[n-2] + \mathbf{D} \mathbf{u}[n-1] \end{aligned} \quad (\text{C.8})$$

eller

$$\mathbf{x}(t) = \sum_{i=0}^{n-1} e^{(n-1-i)\mathbf{A}\delta} \mathbf{D} \mathbf{u}[i] \quad (\text{C.9})$$

Problemet er altså å finne ut hva som kreves for at det skal eksistere en pådragssekvens $\mathbf{u}[i]$, $i = 0, \dots, n-1$ slik at (C.9) oppfylles.

Før vi går videre på dette, har vi behov for å omforme uttrykket for $e^{\mathbf{A}\tau}$

$$e^{\mathbf{A}\tau} = \mathbf{I} + \mathbf{A}\tau + \frac{\mathbf{A}^2\tau^2}{2!} + \frac{\mathbf{A}^3\tau^3}{3!} + \dots \quad (\text{C.10})$$

Vi benytter Cayley-Hamiltons teorem, (3.49), som gjør det mulig å erstatte alle \mathbf{A}^j , $j > n - 1$ i (C.10) med funksjoner av \mathbf{A}^j , $j = 0, \dots, n - 1$. Dermed kan (C.10) omformes til

$$e^{\mathbf{A}\tau} = \gamma_0(\tau)\mathbf{I} + \dots + \gamma_{n-1}(\tau)\mathbf{A}^{n-1} \quad (\text{C.11})$$

Vi ser at koeffisientene γ_i blir tidsfunksjoner. Ut fra et tilsvarende resonnement kan vi omforme (C.7), slik at vi får

$$\mathbf{D} = (\rho_0(\tau)\mathbf{I} + \dots + \rho_{n-1}(\tau)\mathbf{A}^{n-1})\mathbf{B} \quad (\text{C.12})$$

Hvis vi nå vender tilbake til (C.9), ser vi at den inneholder ledd av typen $e^{k\mathbf{A}\delta}\mathbf{D}$; $k = 0, \dots, n - 1$. Benytter vi (C.11) og (C.12) på hvert slikt ledd, og Cayley-Hamiltons teorem for å substituere for alle \mathbf{A}^j , $j > n - 1$, så vil (C.9) kunne skrives som

$$\mathbf{x}(t) = \sum_{i=0}^{n-1} (\varepsilon_{1, i+1}\mathbf{I} + \varepsilon_{2, i+1}\mathbf{A} + \dots + \varepsilon_{n, i+1}\mathbf{A}^{n-1})\mathbf{B}\mathbf{u}[i] \quad (\text{C.13})$$

Her er koeffisientene ε konstanter som følger på grunn av at tidsintervallet $\delta = \text{konstant} = t/n$.

Vi kan bytte summasjonsrekkefølgen i (C.13) og får

$$\mathbf{x}(t) = \sum_{k=0}^{n-1} \left(\sum_{i=0}^{n-1} \varepsilon_{k+1, i+1} \mathbf{A}^k \mathbf{B} \mathbf{u}[i] \right) \quad (\text{C.14})$$

som igjen kan skrives

$$\mathbf{x}(t) = \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{B} \mathbf{v}[k] \quad (\text{C.15})$$

der

$$\mathbf{v}[k] = \sum_{i=0}^{n-1} \varepsilon_{k+1, i+1} \mathbf{u}[i] \quad (\text{C.16})$$

eller skrevet ut:

$$\begin{aligned}
 \mathbf{v}[0] &= \varepsilon_{1,1} \mathbf{u}[0] + \varepsilon_{1,2} \mathbf{u}[1] + \dots + \varepsilon_{1,n} \mathbf{u}[n-1] \\
 &\cdot \\
 &\cdot \\
 \mathbf{v}[n-1] &= \varepsilon_{n,1} \mathbf{u}[0] + \varepsilon_{n,2} \mathbf{u}[1] + \dots + \varepsilon_{n,n} \mathbf{u}[n-1]
 \end{aligned} \tag{C.17}$$

Vi skal komme tilbake til (C.17), men betrakter nå (C.15):

Den kan formuleres som et produkt av en matrise og en lang kolonnevektor:

$$\begin{array}{c}
 \left. \begin{array}{c} \mathbf{v}_0 \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_{n-1} \end{array} \right\} n \cdot r = \tilde{\mathbf{v}} \\
 \\
 \left. \begin{array}{c} \mathbf{A} \quad \mathbf{AB} \quad \dots \quad \mathbf{A}^{n-1} \mathbf{B} \end{array} \right\} n \\
 \left. \begin{array}{c} \mathbf{A} \quad \mathbf{AB} \quad \dots \quad \mathbf{A}^{n-1} \mathbf{B} \end{array} \right\} n \cdot r \quad \leftarrow \mathbf{x}(t)
 \end{array} \tag{C.18}$$

Vi gjenkjenner styrbarhetsmatrisen \mathbf{Q}_s , og definerer dessuten

$$\tilde{\mathbf{v}} = \begin{bmatrix} \mathbf{v}_1 \\ \cdot \\ \cdot \\ \mathbf{v}_n \end{bmatrix}$$

Da blir (C.18)

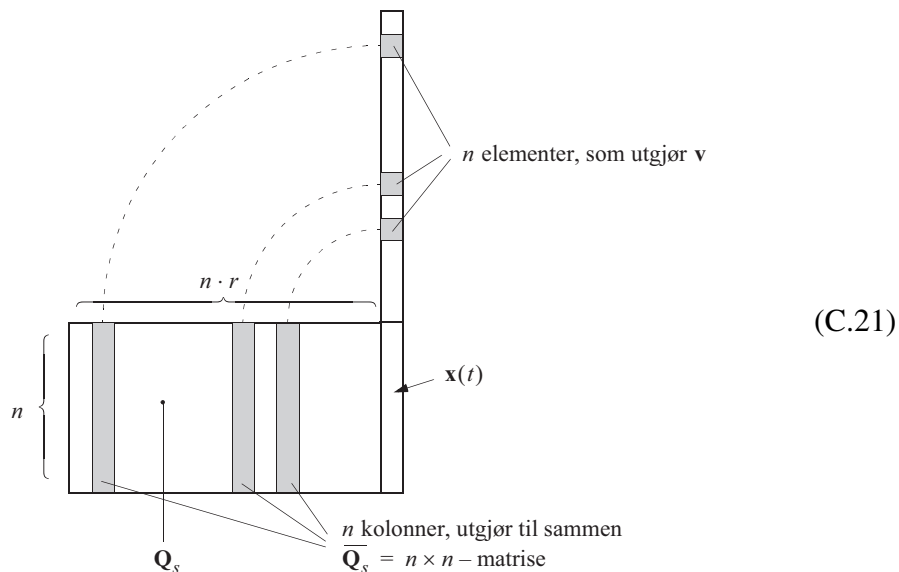
$$\mathbf{x}(t) = \mathbf{Q}_s \tilde{\mathbf{v}} \tag{C.19}$$

For at det overhodet skal være mulig å finne en $\tilde{\mathbf{v}}$ som gir $\mathbf{x}(t)$, kreves det at n kolonner i \mathbf{Q}_s er lineært uavhengige. Disse kolonnene samles i en kvadratisk matrise \mathbf{Q}_s som da blir inverterbar.

Dermed kan vi finne en $\tilde{\mathbf{v}}$ fra likningen

$$\tilde{\mathbf{v}} = \bar{\mathbf{Q}}_s^{-1} \mathbf{x}(t) \quad (\text{C.20})$$

$\tilde{\mathbf{v}}$ dannes av elementene til $\bar{\mathbf{v}}$, som er plassert i posisjoner som svarer til de n kolonnevektorene som er ekstrahert fra \mathbf{Q}_s , og av nuller ellers:



Vi har derved vist at rang $\mathbf{Q}_s = n$ er *nødvendig* for styrbarhet.

Så til *tilstrekkelig*-delen av beviset: Vi vet nå at vi kan finne et sett vektorer $\mathbf{v}_0, \dots, \mathbf{v}_{n-1}$ som gir $\mathbf{x}(t)$, når rang $\mathbf{Q}_s = n$. Men vårt pådrag er $\mathbf{u}_0, \dots, \mathbf{u}_{n-1}$.

Formel (C.17) forteller oss at hvis matrisen

$$\mathbf{E} = \begin{bmatrix} \mathcal{E}_{11} & \cdot & \cdot & \cdot & \mathcal{E}_{1n} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \mathcal{E}_{n1} & \cdot & \cdot & \cdot & \mathcal{E}_{nn} \end{bmatrix} \quad (\text{C.22})$$

er inverterbar, $\mathbf{G} = \mathbf{E}^{-1}$, så blir

$$\begin{aligned}\mathbf{u}[0] &= g_{11}\mathbf{v}[0] + \dots + g_{1n}\mathbf{v}[n-1] \\ &\cdot \\ &\cdot \\ \mathbf{u}[n] &= g_{n1}\mathbf{v}[0] + \dots + g_{nn}\mathbf{v}[n-1]\end{aligned}\tag{C.23}$$

Her fører vi ikke bevis for at \mathbf{E} alltid er inverterbar, men dette kan vises. Vi vil alltid kunne finne et sett av pådrag \mathbf{u}_i , $i = 1, \dots, k-1$ som gir oss $\mathbf{v}_i(t)$ og dermed $\mathbf{x}(t)$.



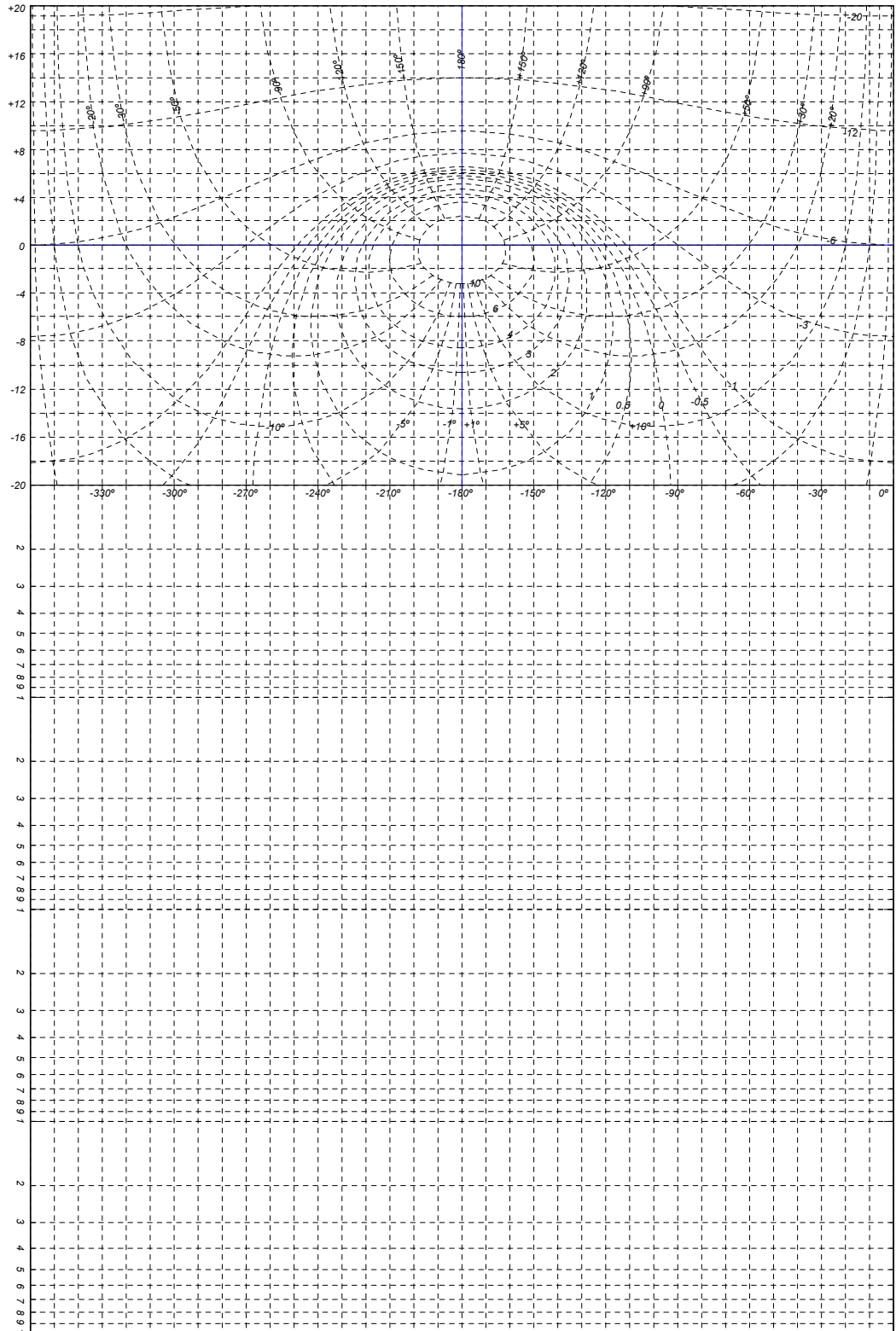
Manuell skissering av Nicholsdiagram ved hjelp av Bode-diagrammer

Når amplitude- og fasediagrammer er tegnet for elementene i et tilbakekoplet system, kan man manuelt tegne den tilsvarende stedkurven i et Nicholsdiagram¹ (dB-vinkel-diagram). Et diagrampapir er vist i figur D.1. Det består av en semilogaritmisk diagramdel til høyre, og et Nicholsdiagram i riktig målestokk til venstre. Dette diagram-papiret er spesielt framstilt for å lette overføringen av data fra amplitude/fase/frekvens-diagrammer til Nicholsdiagrammer. Det benyttes to diagrampapirer. På det ene tegnes amplitudekarakteristikken. Den semilogaritmiske delen har ca. 4 frekvensdeka-der, og langs ordinataksen er det plass til 72 [dB]. På det andre papiret tegnes fase-karakteristikken med samme frekvensskala, og med faseskala som angitt på papiret (0° øverst, -180° midt på, -360° nederst).

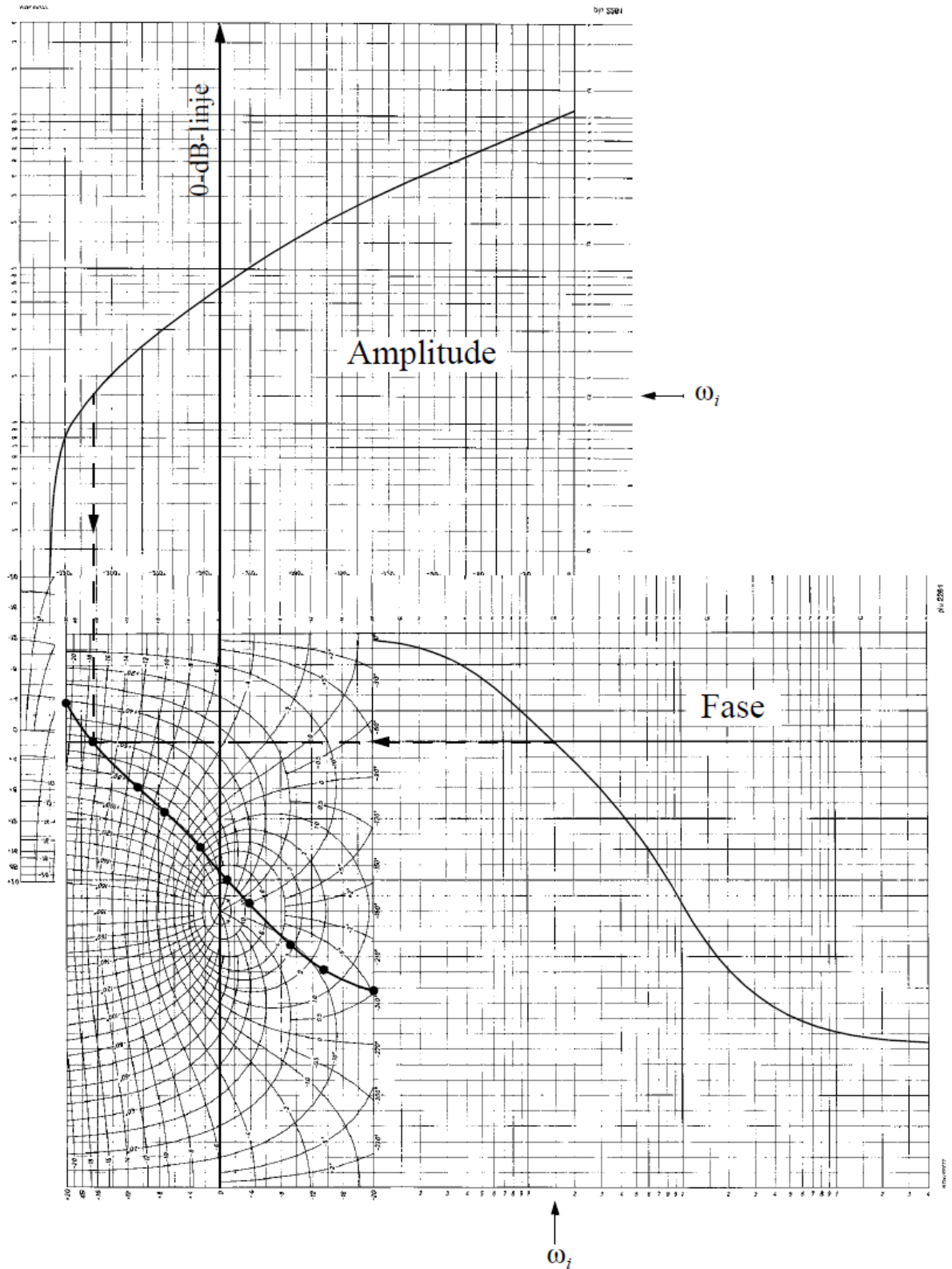
Overføringen av amplitude- og fasedata foregår som antydnet i figur D.2, idet papiret med fasekurven dreies 90° og legges oppå amplitude- diagrampapiret som vist. Det finnes flere andre måter å gjøre denne overføringen på, men en av fordelene med denne metoden er at det ikke er nødvendig å foreta en avlesning av amplituden og fasen (i dB og grader), siden overføringen foregår helt grafisk.

¹Dette kan selvsagt også gjøres i MATLAB, via kommandoen `nichols()`

Figur D.1
Diagrampapir



Figur D.2
Tegning av
Nicholsdiagram



Brønnmesteren

BRØNNMESTEREN OG HANS 3 SØNNER - EN FABEL FOR GODTFOLK OM REGULERINGSPRINSIPPER

Modifisert på basis av en opprinnelig dansk fabel formidlet av dosent Erling Sonne, Danmarks Ingeniør Akademi, Lyngby

Det var en gang en gammel brønnmester, som følte seg mett av dage og gjerne ville tilbringe sin livsaften i ro og fred. Brønnmesteren hadde 3 sønner, Peter, Ivan og David, i landsbyen gjerne kalt PID-brødrene. Og alle tre ville gjerne overta farens stilling ved dennes avgang fra tjenesten.

Brønnmesterens arbeid besto i at han skulle regulere tilstrømningen av vann til en stor høytliggende vannbeholder (vanntårn) fra landsbyens store pumpe som var drevet av en vindmølle. Han skulle holde vannstanden i beholderen konstant, slik at det alltid ville være samme trykk i landsbyens vannledninger, uansett om forbruket var stort eller lite.

Sønnene hadde hver sin mening om hvordan den riktige regulering skulle være - sønner er jo ofte i deres egne øyne ikke så lite klokere enn deres fedre. Den gamle brønnmester vil gjerne være rettferdig overfor sine sønner, og han besluttet derfor å sette dem på en praktisk prøve. Hver av sønnene skulle i en viss tid få regulere vannstanden i beholderen, og den som klarte oppgaven best, ville han anbefale byens borgermester å ansette i brønnmesterstillingen.

For å gi sine sønner et par gode råd med på veien, ga brønnmesteren noen retningslinjer for hva der etter hans skjønn skulle kreves av en god regulering:

1. Vannstanden i beholderen skal holdes så nær som mulig den riktige høyden, uansett forbruksvariasjon.
2. Hvis vannstanden av en eller annen grunn er forskjellig fra den ønskede, skal den føres tilbake til det riktige nivå i et passende tempo. Det må ikke være for hurtig, for det sliter på ventilen og kan medføre pendlinger i vannstanden. Men tempoet må heller ikke være for langsomt slik at vannstandens avvik holder seg unødig lenge.

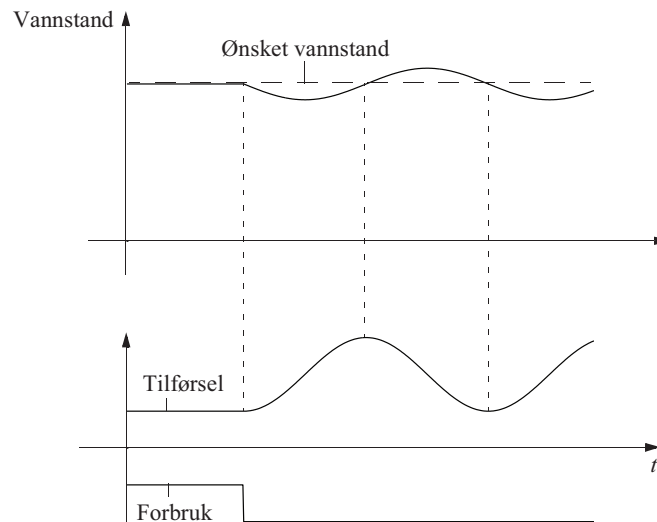
Så begynte prøvene, og den eldste av sønnene, Ivan, var først ute. Han var en grundig og samvittighetsfull type som ikke var altfor intelligent, men arbeidet langsomt, jevnt og rolig.

Mens Ivan stod ved ventilen, oppstod plutselig en stigning i forbruk av vann, og vannstanden i beholderen begynte å falle. Han resonnererte da som så: Det skal ikke mangle på flid fra min side; Jeg fortsetter å regulere, inntil vannstanden igjen er helt korrekt.

Ivan begynte langsomt å øke vanntilstrømningen ved å dreie ventilen med en hastighet som var proporsjonal med vannstandsavviket. Han fortsatte å dreie opp ventilen selv om han etterhvert så at vannstanden nå ikke falt mer, men var begynt å stige. Etterhvert som vannstanden nærmet seg det riktige nivået, dreide han langsommere på ventilen, men han sluttet først med å øke vanntilstrømningen, da vannstanden hadde nådd den korrekte høyden. Til sin forbauselse så han da at vannstanden fortsatte å stige, og nå ble han derfor nødt til å dreie ventilen den andre veien for å lukke den. Han vedble å dreie ventilen så lenge vannstanden var over den ønskede høyden og med en hastighet proporsjonal med vannstandsavviket. Vannstanden kom derfor ikke bare ned på sin riktige verdi igjen, men fortsatte til et godt stykke under denne, og Ivan måtte derfor begynne forfra igjen med å åpne ventilen.

I langsomme svingninger over- og underskred vannstanden hele tiden den riktige verdien, og Ivan fikk ikke en rolig stund under sitt arbeid, idet han hele tiden var nødt til å regulere for de endringer som han til en viss grad selv var skyld i. Så lenge det eksisterte et avvik fra den ønskede vannstanden, mente Ivan, at han skulle dreie på ventilen, og han kunne derfor aldri nå en stabil tilstand. Figur E.1 viser hvordan vannstanden forandret seg under Ivans regulering. Ivan arbeidet som en ren *integralregulator*.

Figur E.1
Respons og pådrag
ved fall i forstyrrelsen
(=forbruket)
ved
integralregulering



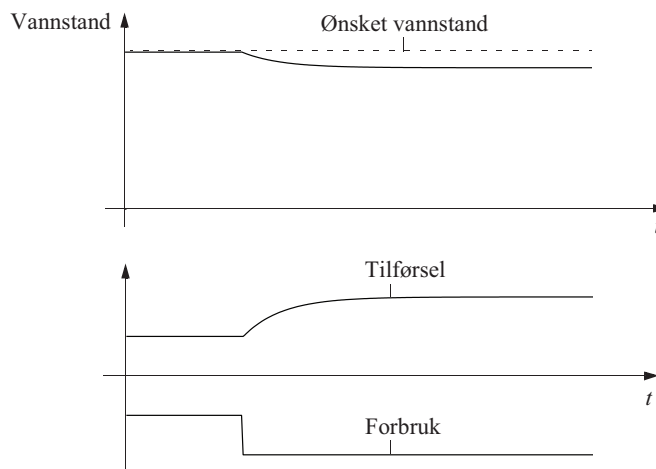
Om ettermiddagen fikk Ivan avløsning av sin bror, Peter. Peter var en meget beregnende fyr og hadde i forveien nøyaktig regnet ut hvordan han ville angripe regule-

ringsproblemet. Hvis vannstanden var blitt 10 cm for lav, kunne han bremse avviket ved å åpne sitt ventilhåndtak 5 omdreininger og derved utlikne det økte forbruket. Peters regulering virket ikke dårlig i praksis. Han var virkelig i stand til å utlikne det økte forbruket ved å dreie ventilen et antall omdreininger *proporsjonalt med vannstandsavviket*. Når han på denne måten hadde sørget for at vanntilførselen var like stor som landsbyens forbruk, bekymret han seg ikke noe om at vannstanden faktisk var blitt noe for lav. Han sa til seg selv, at det varte jo ikke så lenge, før vannforbruket ville falle igjen, og så ville vannstanden gå tilbake på sitt riktige nivå. Da vannstanden en tid senere endret seg slik at den var 6 cm for lav, altså en endring på 4 cm i forhold til det han hadde før (10 cm), dreide han ventilhåndtaket 2 omdreininger i lukkeretningen og stabiliserte dermed vannstanden igjen. Peter regulerte hurtig og sikkert, men nøyaktigheten av vannstanden tok han seg ikke av.

Peter arbeidet som en *proporsjonal regulator*. Figur E.2 viser hvordan vannstanden varierte under Peters regulering, da vannforbruket plutselig steg første gang.

Figur E.2

Respons og pådrag ved sprang i forstyrrelsen (=forbruket) ved proporsjonalregulering



Om kvelden overtok David, den yngste av sønnene, reguleringen. David var en overordentlig energisk fyr, men noe komplisert i sin psyke. Han gikk først igang med arbeidet etter at han hadde tenkt nøye i gjennom hvordan han ville angripe situasjonen. Han mente, at hans metode var langt bedre enn noen av hans brødres metoder. Det gjelder, mente han, å stoppe vannstandsendringer så fort som mulig.

Han ville derfor iaktta den *hastighet* hvormed vannet steg eller sank og *innstille ventilen proporsjonalt med dette*. Hvis vannstanden f.eks. plutselig begynte å falle 10 cm pr. minutt, ville han lynhurtig skru ventilen 10 omdreininger opp, 20 omdreininger opp ved 20 cm pr. minutt fall osv.

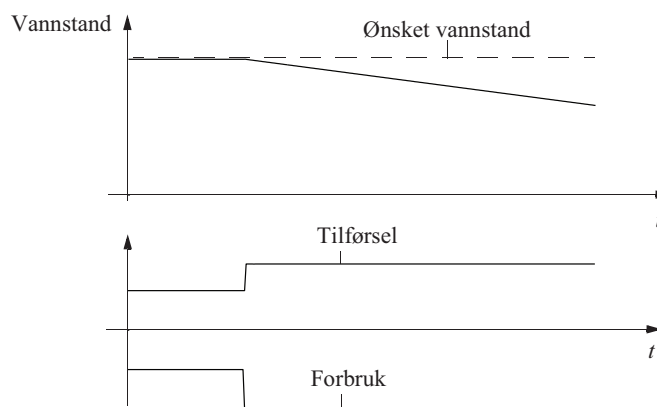
Da David kom ut til brønnen, inntraff plutselig en forbruksstigning, slik at vannet sank hurtig i beholderen. Han begynte derfor hurtig å skru ventilen opp, som han

hadde tenkt ut på forhånd. Derved ble hastigheten som vannspeilet sank med, redusert. Da David hadde arbeidet meget hurtig, var vannspeilet bare sunket ganske lite siden forbruksstigningen inntraff. Imidlertid måtte det være noe galt med Davids reguleringsmetode, for riktignok var det hurtige fallet av vannspeilet stanset, men vannstanden sank stadig, om enn ganske langsomt. Det brydde David seg ikke om, for etter hans prinsipper skulle det være overensstemmelse mellom ventilstilling og vannstandsending og det var der. Litt senere kom der et fall i vannforbruket, og David fikk det igjen travelt, denne gang med å skru ventilen til. Vannstandsendingen ble igjen nesten stoppet.

Likegyldig hvor meget David slet med regulering av ventilen, maktet han verken å holde vannstanden konstant eller i nærheten av den korrekte verdi. David oppførte seg som en *derivatregulator*. Figur E.3 viser vannstandsendingen ved en plutselig forbruksending.

Figur E.3

Respons og pådrag ved sprang i forstyrrelsen (=forbruket) ved derivatregulering



Det ble etter dette ytterst vanskelig for byens borgermester å treffe en beslutning. Ivan, som arbeidet som en integral regulator, brukte lang tid på å få vann-nivået riktig, men fikk det ikke til. Han genererte derimot en stående svingning. Hvis vannstanden først var kommet i svingninger, var det umulig å få stanset den.

Peter som arbeidet som en proporsjonal regulator, oppnådde faktisk det motsatte. Ved hans metode ble det ikke noen store avvik i vann-nivået, men det riktige nivået kom han ikke til. Det ble diskutert om Ivan eller Peter var den beste, men man kom ikke til noe resultat.

David gjorde det vanskelig for seg selv. Hver av hans håndbevegelser var energiske og hurtige, nesten som piskeslag. Med hans arbeide var borgermesteren og landsbyens innbyggere minst tilfreds, idet han tross alle sine anstrengelser fjernet vannstanden lengre og lengre fra den riktige.

Siden ingen av de tre brødrene hadde demonstrert de samme egenskapene som den

gamle kloke brønnmesteren, besluttet borgermesteren at han ville prøve å la dem regulere sammen. Forsøket var lett å arrangere da de fikk satt inn to ekstra ventiler i parallell med den første mellom pumpen og vannbeholderen. Dermed fikk de hver sin ventil å regulere på. Forsøket ble fastsatt til å finne sted i den følgende uken.

Imidlertid ble David syk, og Peter og Ivan prøvde alene sammen.

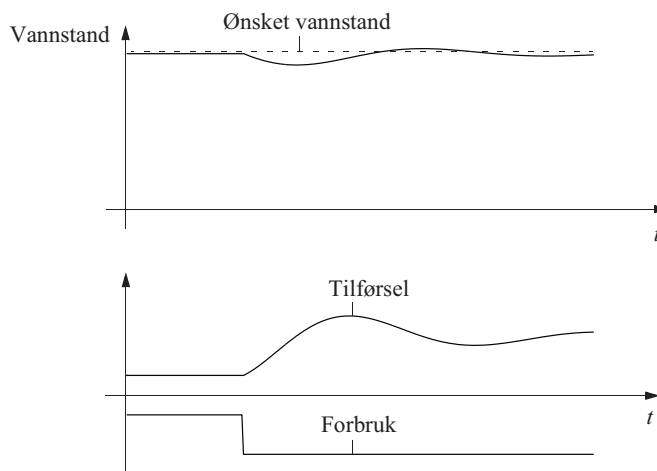
Mens de sto klar med sine ventiler, kom det en forbruksøkning som resulterte i at vannstanden sank. Peter åpnet sin ventil proporsjonalt med vannstandsavviket og sørget dermed for at vannstanden ikke sank ytterligere. Ivan åpnet langsomt og rolig sin ventil med en hastighet proporsjonal med vannstandsavviket og holdt på med dette inntil vannstanden var kommet på plass. Det dreide seg for Ivan denne gangen, siden han arbeidet sammen med Peter, bare om å gjøre Peters arbeid ferdig og korrigere for hans unøyaktighet.

Figur E.4 viser hvordan Peter og Ivan regulerer hvis byens vannforbruk plutselig stiger. Denne reguleringsform kalles *proporsjonal pluss integral-regulering* eller *PI-regulering*.

Alt i alt var alle godt tilfredse med Peters og Ivans samarbeid, men man syntes David måtte få en sjanse til å vise hva han dugde til når han arbeidet sammen med de andre.

Figur E.4

Respons og pådrag ved sprang i forstyrrelsen (=forbruket) ved proporsjonal+integralregulering



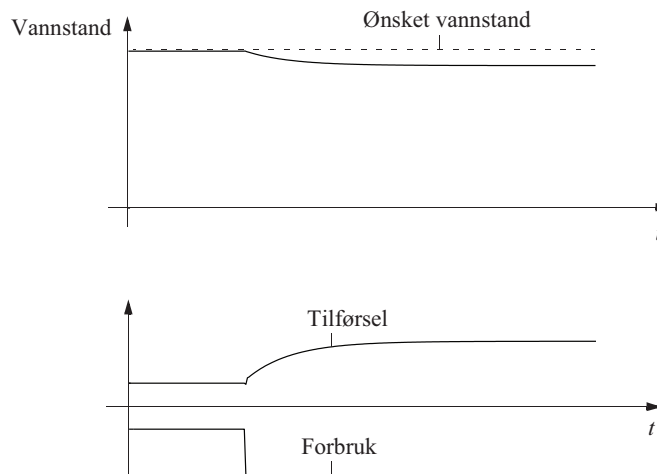
Et par dager etter var David kommet på beina igjen, men til gjengjeld hadde Ivan lagt seg syk og man lot derfor David og Peter prøve sammen.

Da de startet, begynte vannstanden å falle. Peter regulerte som vanlig med en proporsjonal åpning av sin ventil. David åpnet opp sin ventil svært hurtig og øket vanntilstrømmingen vesentlig. Da vannstanden deretter falt langsommere, sluttet han å åpne sin ventil mere. Han sa til seg selv, at Peter jo alltid kom baketter, og at hans egen hurtige regulering var nødvendig for å bremse vannstandssenkingen hurtig nok. Da-

vid mente liksom tidligere at det var viktigere å regulere vannstandens fallhastighet eller stigningshastighet enn den absolutte vannstanden. Peter hadde med sin rolige og besindige åpning av ventilen bremsset fallet, og den sammenlagte virkningen av deres anstrengelser gjorde at vannstanden sank, men den falt mye langsommere enn da Peter regulerte alene. Figur E.5 viser hvordan Peter og David sammen regulerer vanntilstrømningen om vannforbruket plutselig stiger. Denne reguleringsformen kalles *proporsjonal pluss derivat-regulering* eller PD-regulering.

Figur E.5

Respons og pådrag ved sprang i forstyrrelsen (=forbruket) ved proporsjonal+derivatregulering



Endelig lyktes det en dag å få alle tre brødrene i gang med å regulere ventilene sammen.

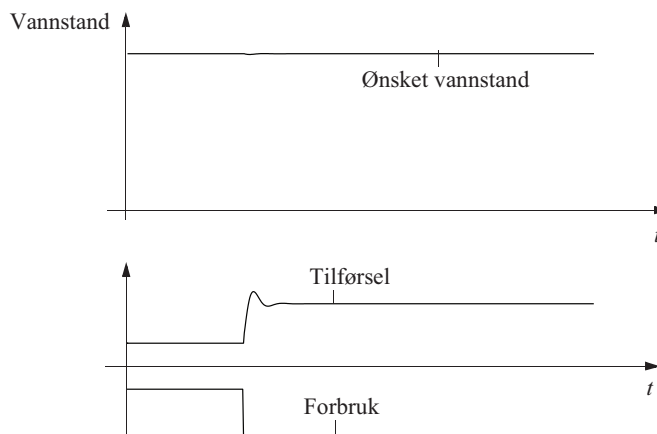
Vannstanden begynte å falle som den hadde gjort ved de tidligere forsøkene. Peter åpnet sin ventil proporsjonalt med vannstandsendingen som han pleide å gjøre. Ivan tok seg som vanlig god tid og dreide sin ventil med en hastighet som var proporsjonal med vannstandsavviket. Men før de to brødrene hadde fått endret sine ventiler noe særlig, hadde David åpnet sin ventil langt opp. Derved hadde David med sitt hurtige inngrep forhindret et stort avvik fra den ønskede vannstanden. Peter hadde kompensert for det økte forbruket og Ivan hadde fjernet det avviket som ble igjen.

Figur E.6 viser hvordan de tre brødrene i fellesskap regulerer vanntilstrømningen når landsbyens forbruk plutselig øker. Til sammen arbeider de tre som en *proporsjonal pluss integral pluss derivat-regulator* eller PID-regulator. De utgjorde den perfekte kombinasjon.

Til belønning for deres mønsterverdige prestasjon innstilte borgermesteren dem alle tre til vervet som brønnmester og byrådet ansatte dem alle. De hadde jo tydelig demonstrert, at bare ved at de tre arbeidet sammen var man i stand til å få en perfekt regulering.

Figur E.6

Respons og pådrag ved sprang i forstyrrelsen (=forbruket) ved PID-regulering



Hele byen gledet seg over at det var kommet en løsning på det viktige problemet, og gleden fikk først et skår da neste regnskap fra byrådet viste at det nå var tre brønnmestere på lønningslisten i stedet for bare en slik det hittil hadde vært.

Etterskrift til den danske fabel.

Etter at de tre PID-brødrene hadde utført sin funksjon i omtrent ett år, skjedde det noe underlig. Til en av landsbyens bondegårder var det kommet en ny dreng. Det var en ung gutt ved navn Ola fra nabolandet Norge. Han ble snart kjent som en ganske gløgg gutt og hadde det vi kaller "*sunt bondevett*". En dag fikk han lyst til å besøke de tre brødrene mens de arbeidet ved vanntårnet. Han spurte dem hva de holdt på med og de forklarte ham at jobben deres var å regulere ventilene slik at vannstanden holdt seg konstant. Da han spurte om hvorfor vannstanden forandret seg, ble han forklart at det skyldtes at vannforbruket til landsbyen varierte svært.

Ola tenkte litt på dette og så gikk det et lys opp for ham. Han ville helst ikke fortelle de tre brødrene hva han hadde funnet ut fordi de var jo ekspertene og ansatt som brønnmestere. Men han spurte dem om han også kunne få en ventil til å prøve seg på. Og det mente brødrene at han jo lett kunne få. Det skulle de sørge for.

Så spurte Ola om han kunne få se det røret som førte vann ut av beholderen til landsbyen. Det fikk han. Han la øret inn til røret for han tenkte han kunne kanskje høre når vannet rant igjennom. Det kunne han. Han tenkte at hvis han kunne høre hvor mye vann som gikk gjennom røret, kunne han bare åpne sin ventil tilsvarende og da ville det renne inn like mye vann i beholderen som landsbybefolkningen tok ut av beholderen.

Neste morgen hadde brønnmesterbrødrene installert en ny ventil, og sa til Ola at nå kunne han forsøke seg. Det viste det seg at nivåendringen i beholderen når vannforbruket endret seg, ble svært liten slik at brønnmesterbrødrenes arbeide ble sterkt

reduisert.

Det Ola hadde laget var en *foroverkopling* fra forstyrrelsen (vannforbruket) til pådraget (vanntilførselen).

Da borgermesteren hørte om Ola Nordmann som hadde sunt bondevett, tenkte han som så at nå hadde han funnet løsningen hvis han ansatte Ola istedenfor de tre brødrene. Da kunne han spare mange penger. Han spurte Ola om han kunne tenke seg å overta jobben som brønnmester, men Ola ble svært betenkt. Han skjønnte jo at den målingen han gjorde av vannstrømmen ut av beholderen ved å høre på lyden i røret, ikke var særlig nøyaktig og derfor kunne det bli en differanse mellom vannstrømmen ut av beholderen og den som han regulerte inn i beholderen med ventilen. Derfor tenkte han som så at jeg må nok i alle fall ha hjelp av en av brødrene til å sikre at ikke vann-nivået endrer seg langsomt.

Da borgermesteren fikk høre dette, ble han lei seg og innså at han likevel kanskje ikke kunne tjene noe særlig ved å ansette Ola. Men han roste Ola for den gode innsatsen.

Bibliografi

- [1] Ackermann, J. Der Entwurf Linearer Regelungssysteme im Zustandsraum, Regelungstech. Prozess-Datenverarb., vol 7, pp. 297-300 1972.
- [2] Anderson, B. D. O. and Moore, J. B. Optimal Control. Linear Quadratic Methods, Prentice Hall. 1989.
- [3] Anderson, B. D. O. and Moore, J. B. Linear Optimal Control, Prentice Hall, Englewood Cliffs, NJ 1971.
- [4] Balchen, J. G. A Performance Index for Feedback Control Systems Based on the Fourier Transform of the Control Deviation. Acta Polytechnica Scandinavia, No. 1, 1958.
- [5] Balchen, J. G. Rational transfer function approximations to transport delay. Modeling, Identification and Control, Vol. 11, No. 3, 127-140 1990.
- [6] Balchen, J. G. The stability of multivariable control systems. Modeling, Identification and Control, Vol. 11, No. 2, 97-108 1990.
- [7] Bélanger, P. R. Control Engineering - a modern approach, Saunders College Publishing 1995.
- [8] Bode, H. W. Network analysis and Feedback Amplifier Design, D. Van Nostrand Co., Inc., New York 1945.
- [9] Bryson, A. E. and Ho, Y. C. Applied Optimal Control, Waltham, MA 1969.
- [10] Campbell, D. P. Process Dynamics, John Wiley, New York. 1958.
- [11] Cheney, W. Numerical Mathematics and Computing, Brooks/Cole Publishing Company 1994.
- [12] Dorf, R. C. and Bishop, R. H. Modern Control Systems, Pearson 2011.
- [13] Doyle, J. C. and Stein, G. Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis. IEEE Trans. Autom. Control, vol. AC-26. no. 1, pp. 4-16, feb 1981.

- [14] Evans, W. R. Graphical Analysis of Control Systems. Trans. Am. Inst. Electr. Eng., vol 67, pp. 547-551 1948.
- [15] Fossen, T. I. Guidance and Control of Ocean Vehicles, Wiley, Chichester 1994.
- [16] Franklin, G. F., Powell, J. D. & Emami-Naeini, A. Feedback Control of Dynamic Systems, Addison Wesley Publishing Company, Inc. 2010.
- [17] Goodwin, G. C., Graebe, F. S. & Salgado, M. E. Control System Design, Prentice-Hall 2001.
- [18] James, H. M., Nichols, N. B. & Philips, R. S. Theory of Servomechanisms. Radiation Lab. Series, 25 McGraw-Hill, New York 1947.
- [19] Kailath, T. Linear Systems, Prentice-Hall, Englewood Cliffs, NJ 1980.
- [20] Kalman, R. E. On the General Theory of Control Systems. in Proc. 1st Int. Congr. Autom. Control, pp. 481-493, Moscow. 1960.
- [21] Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. J. Basic Eng., vol 85, pp. 34-45 1960.
- [22] Khalil, H. K. Nonlinear Systems, Prentice Hall, Upper Saddle, N. J. 1996.
- [23] Ljung, L. Reglerteori. Moderna analys- og syntesmetoder, Studentlitteratur Lund, Sweden 1981.
- [24] Ljung, L. System Identification: Theory for the user, Prentice-Hall, Englewood Cliffs, NJ 1987.
- [25] Lewis, F. Optimal Control, John Wiley & Sons 1986.
- [26] Newton, G. C., Gould, L. A. & Kaiser, J. F. Analytical Design of Linear Feedback Controls, John Wiley, New York 1957.
- [27] Nise, N. Control Systems Engineering, The Benjamin/Cummings Publishing Company, Inc. 1995.
- [28] Nyquist, H. Regeneration Theory Bell System Tech. Journal, 1932.
- [29] Ogata, K. Modern Control Engineering, Prentice-Hall, Inc. 2002.
- [30] Oppenheim, A. V., Willsky, A. S. & Young, I. T. Signals and System, Prentice-Hall, Inc. 1983.
- [31] Rosenbrock, H. H. State-space and Multivariable Theory, Nelson, London 1970.
- [32] Shahian, B. & Hassul, M. Control System Design Using MATLAB, Prentice-Hall, Inc. 1993.

-
- [33] Skogestad, S. & Postlethwaite, I. Multivariable Feedback Control. Analysis and Design, John Wiley & Sons, New York 1996.
- [34] Smith, O. J. M. Feedback Control Systems, McGraw-Hill, New York 1958.
- [35] Strang, G. Linear Algebra and Its Applications, Thomson Learning 1988.
- [36] Truxal, J. G. Control System Synthesis, McGraw-Hill, New York 1955.
- [37] Vidyasagar, M. Nonlinear Systems Analysis, Prentice Hall, Englewood Cliffs, NJ 1993.
- [38] Wiener, N. Cybernetics or Control and Communication in the Animal and the Machine, MIT Press, Cambridge, Mass. 1948.
- [39] Zadeh, L. A. & Desoer, C. A. Linear System Theory - The State Space Approach, McGraw-Hill, New York 1963.
- [40] Ziegler, J. G. and Nichols, N. B. Optimum settings for automatic controllers, Transactions of the A.S.M.E 64: 759-768 1942.
- [41] Åström, K. J. Reglerteori, Almqvist & Wiksell, Stockholm 1968.
- [42] Åström, K. J. Introduction to Stochastic Control Theory, Academic Press, New York 1970.

Tidsskrifter som omhandler regulerings- tekniske emner

American Control Conference (ACC) Proceedings

Automatisierungstechnik, Oldenburg Verlag, München

Automatica, Pergamon Press, Ltd., England

European Control Conference (ECC) Proceedings

International Federation of Automatic Control (IFAC) Proceedings

Institute of Electrical and Electronics Engineers (IEEE), USA:
Transactions on Automatic Control (AC)
Transactions on Control Systems
Conference on Decision and Control (CDC) Proceedings
Control Systems

International Journal of Systems Science, London: Taylor and Francis, Ltd.

Journal of Process Control, Elsevier Science, UK

Modeling, Identification and Control (MIC), Oslo: Publ. Norges forskningsråd

Stikkordsliste

A

A/D-omsetter	504
absolutt dempningsfaktor	144, 149
AF-diagram	202
AFF-diagram	201
aktiv effekt	564
amplitude	193
amplitude/fase-diagram	202
amplitude/fase/frekvens-diagram	201
amplitudediagram	201
amplitudeforhold	197
amplitudeforsterkning	193
amplitudemodulasjon	504
analytisk funksjon	115
antennestyling	377, 394
anti wind-up	568
anti-overlader	568
arbeidspunkt	99, 102
asymptote	211
asymptotisk AFF-diagram	209
asymptotisk kryssfrekvens	246
asymptotisk stabilitet	265, 267, 268
asynkronmotor	61
autonom	41, 80, 268
autopilot	23
avvik	14
avviksforhold	242, 245, 319
avvikskoeffisienter	325

B

båndbredde	318, 320
båndbredde, realiserbar	566
båndpassfilter	516

båndstoppfilter	516
begrenset derivatvirkning	371
begrenset integralvirkning	371
begynnelsesverditeoremet	111
begynnelsesverditeoremet, diskret	521
belastningsstivhet	318
betinget stabilt	301
blokkdiagram	122
blokkdiagram, elementært	71, 122
blokkdiagram, reduksjon av	122
Bode-Nyquist-kriteriet	302
Bodediagram	201

C

Cauchys integralsats	110
Cayley-Hamiltons teorem, transisjonsmatrisen	88

D

D/A-omsetter	504
dekade	209
dekopling	482
deltafunksjon	48
derivasjon	110
derivasjonsoperator	128
derivatvirkning	371
derivatvirkning, begrenset	371
desibel	201
destillat	25
diagonaliserte variable	94, 135
differensiallikning, høyereordens	65
differensiallikninger	40
differensiallikninger, drevet	40
differensiallikninger, førsteordens	40

differenslikning 512
 dimensjonsløs frekvens 527
 dimensjonsløs tastetid 527
 Dirac-puls 48
 direkte syntese 314
 diskret enhetspuls 518
 diskret enhetssprangfunksjon 527
 diskret foldning 519
 diskret impulsrespons 518
 diskret PI-regulator 505, 506
 diskret regulator 545
 diskret regulering 501
 diskret sluttverditeorem 521
 diskret system, stabilitet 524
 diskret system, tidsrespons 548
 diskret tid 502
 diskret transferfunksjon 523
 diskrete systemer, frekvensanalyse
 527
 diskretiseringsintervall 69
 dreiemomentgrensen 566
 dynamisk nøyaktighet 316
 dynamisk tilbakekopling 253
 dynamiske systemer 12

E

effekt, aktiv 564
 effekt, reaktiv 564
 effektbruk 563
 effektgrense 567
 effektivverdi 233
 effektivverdiforholdet 234
 egenskap 13
 egenskapsvektor 14
 egenvektor 574
 egenvektormatrise 77, 581
 egenverdi 574
 egenverdimatrise 78
 eksoterm reaksjon 102
 elektrisk balanse 58, 61
 elektrisk tidskonstant 155

elementært blokkdiagram 71, 122
 energibalanse 54
 enhetspuls 518
 enhetssprang 18
 enhetssprang, diskret 527
 enhetstilbakekopling 253, 391
 estimator 429
 estimatorbasert regulering .. 428, 432
 estimering 28

F

føde 25
 fartøystyring 315
 fasediagram 201
 faseforskyvning 144, 193
 fasemargin 303, 317
 fasevariable 67, 95, 136
 fasevinkel 197
 filtere 512
 filtre, analog 512
 filtre, båndpass 516
 filtre, båndstopp 516
 filtre, Butterworth 513
 filtre, diskre 512
 filtrering 512
 foldingsintegral 42
 foldning, diskret 519
 fordelte parametre 40
 forholdsregulering 415
 foroverkopling 15, 406
 foroverkopling, følgesystemer ... 417
 foroverkopling, predikativ 426
 foroverkopling, prosessregulerings-
 systemer 407
 foroverkoplingsgrad 408
 foroverkoplingsgrad, ideell 408
 forsterkning kritisk 348
 forsterkningsmargin 303, 317
 forstyrrelse 13
 Fourier-transformasjon 195
 frekvensanalyse 194

frekvensrespons 194, 197
 frekvensrespons, grafisk representa-
 sjon 201
 frekvensresponsmatrise 200
 funksjonell ikke-styrbarhet, multiva-
 riabel 498
 funksjonell styrbarhet 190
 funksjonell styrbarhet, multivariabel
 497
 fysiologisk system 28
 fysiske balanser 50
 følgeforhold 242
 følgesystem 318
 følgesystemer, foroverkopling ... 417
 følgesystemer, seriekompensasjon 365
 følgeavvik 320
 følgeforhold 246, 319
 følgeforholdsmatrise 470
 følgesystem 241, 315

G

geometriske middelvei 342

H

høyereordens differensiallikninger 65
 harmoniske frekvenser 233
 holdeelement 501
 Hurwitz's kriterium 274, 278
 hysteresis 404

I

ideell foroverkopling 407, 487
 ideell foroverkoplingsgrad 408
 ikke-dynamisk prosess 18
 ikke-minimum-fase-system 220
 ikke-singulær matrise 579
 impedans 120
 impulsfunksjon 48
 impulsrespons 48
 impulsrespons, diskret 518
 industriell automatisering 30
 inngang(svariabel) 13

inngangs-/utgangsstabilt system . 265
 integralregulator 170
 integralvirkning 371
 integralvirkning, begrenset 371
 integrasjon 110
 integrator-overlading 568
 integrerende regulator 386
 intern tilbakekopling 391
 intern tilbakekopling, multivar. system
 488
 invers Laplacetransformasjon ... 109,
 115
 invers z-transformasjon 522
 invertert matrise 577
 irrasjonal transferfunksjon 168
 iterativ syntese 314

K

kanonisk form 94
 kanonisk form 3 97
 karakteristisk likning 270, 576
 knekkfrekvens 216
 kompensasjon med intern tilbakekop-
 ling 327
 kompleks folding 111
 kompleks konjugert 139
 kompleks translasjon 111
 kompleks multiplikasjon 111
 konform avbildning 526
 konsentrerte parametre 40
 konvolusjon 49
 kraftbalanse 54
 kritisk forsterkning 348
 kryssfrekvens 318
 kryssfrekvens, asymptotisk 246
 krysskopling 474
 kybernetikk 8

L

laplacetransformasjon 65
 Laplacetransformasjon, invers ... 109,
 115

- laplacetransformasjons-par 585
 likeløpsregulering 23
 likestrømsmotor 61, 367
 likevektstilstand 102
 lineær tidsveing 111
 lineære transformasjoner 574
 linearisering 53
 linearitet 110
 lineær uavhengighet 576
 lineært system 46
 LTI-systemer 41
 lukket sløyfe 238
 lukket system 157, 238
- M**
- maksimal fasekorreksjon 342
 marginal stabilitet 265, 268
 massebalanse 51
 matrise, ikke-singulær 579
 matrise, invertert 577
 matrise, regulær 579
 matrise, singulær 579
 matrise, transponert 577
 matrisealgebra 573
 matriseoperasjoner 577
 mekanisk tidskonstant 155
 memoryless process 18
 metning 556, 560
 minimum-fase-system 219
 modalmatrise) 130
 modell av 1. orden 50
 modellreduksjon 154
 modulasjon 504
 modus 79
 momentan prosess 18
 momentbalanse 51, 63
 monovariabel 15
 monovariebelt system 91
 multiplerer 504
 multiplikativ perturbasjon 453
 multiplert pol 116
- multivar. system, intern tilbakekopling
 488
 multivariabel 15
 multivariabel w-transformasjon .. 550
 multivariabelt system 91
 multivariable systemer 91, 467
 multivariable systemer, stabilitet . 471
 målematrise 93
 målestøy 13, 91
 målevektor 13
 målinger 13
- N**
- Nicholsdiagram 249
 Nicholsdiagram, skissering 597
 nominell sløyfetransferfunksjon . 453
 nullpunkt i høyre halvplan 220
 nullpunkter 135, 138, 164
 numerisk optimalisering 448
 Nyquist-prosedyre 287
 Nyquistdiagram 201
 Nyquistdiagram, logaritmisk 292
 Nyquists stabilitetskriterium, forenklet
 286
 Nyquists stabilitetskriterium, generelt
 285
 Nyquists stabilitetskriterium, åpent
 stabilt system 288
 Nyquists stabilitetskriterium, åpent
 ustabilt system 294
 nøyaktighet, dynamisk 316
 nøyaktighet, statisk 316
- O**
- objektfunksjonal 437
 objektfunksjoner 436
 Observerbarhet 175
 observerbarhet, parametrisk ikke- 186
 observerbarhet, strukturell ikke- . 187
 observerbarhetsmatrisen 180
 optimalregulering, multivariabel . 488
 optimering, parametrisk 436

orden 15
ordinære differensiallikninger 40
oscillatorisk 21
Otto-Smith-regulator 432
overføringsfunksjon 119

P

P-regulator 157, 345, 349
Padé-approximasjon 172
parameter 260
parametermargin 305
parameterstyring 556
parametrisk ikke-observerbarhet . 186
parametrisk ikke-styrbarhet 186
parametrisk optimering 436
Parsevals sats 437
partielle differensiallikninger 40
PD-regulator 161, 346
PI-regulator 167, 346, 349, 355
PI-regulator, diskret 505, 506
PID-regulator 347, 349
polardiagram 201
poler 116, 135, 138
predikativ foroverkopling 426
produktivitet 31
proper 198
proporsjonalregulator 345
prosess 12
prosessforstyrrelser 91
prosessindustri 24
prosessregulering 241, 315
prosessreguleringssystemer, forover-
kopling 407
prosesstransfermatrise 468
prosesstøy 91
pseudoinverse 461
pulsbreddemodulasjon 504
pumpesymbol 52
pådrag 13
pådragsmatrise 93
pådragsvektor 13

R

rasjonal funksjon 138, 199
reaktiv effekt 564
realiserbar båndbredde 566
reell folding 111
reell folding, diskret 520
reell multiplikasjon 111
reell translasjon 111
referanse 14
reflux 25
regulator 15
regulatortransferfunksjon 241
reguleringsgrad 243, 245
reguleringsgradmatrise 470
reguleringsobjekt 12
reguleringsløyfe 238
reguleringsystem 12
reguleringsteknikk 8
reguleringsteknikk, historie 9
regulær matrise 579
relativ dempningsfaktor 144, 149
relativ følsomhet 260
residuesatsen 116
residuregning 110
residusats 141
resiprok basisvektor 131
resolventmatrisen 130
robust regulering 453, 454
robust stabilitet 454
robust ytelse 454
root locus 158
rotkurver 158
Rouths kriterium 278

S

sampling 501
samplingsteoremet 532, 534
seriekompensasjon, multiv. system 481
seriekompensasjon 327
seriekompensasjon, følgesystemer 365
servomekanisme 241, 315

- servomekanisme, seriekompensasjon
 365
 settpunkt 14
 SIMC-metoden 355
 similaritetstransformasjon 579
 Simulink 170
 singulærverdianalyse 491
 singulariteter 135
 singulær matrise 579
 singulært punkt 102
 sluttverdiregulering 459
 sluttverditeoremet 111, 127, 321
 sløyfetransferfunksjon 238, 241
 sløyfetransfermatrise 469
 sosialt system 28
 spenningsregulering 401
 spesifikasjon 314
 sprang 44
 sprangrespons 18, 147
 stabilitet i tilstandsrommet 267
 stabilitet, algebraiske kriterier ... 270
 stabilitet, asymptotisk . 265, 267, 268
 stabilitet, marginal 265, 267, 268
 stabilitet, multivariable systemer . 471
 stabilitetsmargin 305, 316, 317
 standardregulatorer 345
 stasjonær respons 194
 stasjonær modell 39
 statikk 318
 statisk foroverkopling 408
 statisk nøyaktighet 316
 statisk prosess 18
 stedkurven 201
 stigningstall 209
 stoffbalanse 51
 strengt proper 286
 strengt proper transferfunksjon .. 117
 strukturell ikke-observerbarhet .. 187
 strukturell ikke-styrbarhet 186
 Styrbarhet 175
 styrbarhet, funksjonell 497
 styrbarhet, parametrisk ikke- 186
 styrbarhet, strukturell ikke- 186
 styrbarhetsmatrisen 177
 summasjonsforsterker 255
 superposisjonsprinsippet 46
 svingefrekvens 144, 149
 symbol styrbar ventil 25
 synkrokjede 378
 syntese 314
 syntese, direkte 314
 syntese, iterativ 314
 system 12
 systemmatrise 93
- T**
- tachometer 61
 tastefrekvens 502
 tastefrekvensen 534
 tasteteoremet 532, 534
 tastetid 501, 502
 tasting 501
 teknisk kybernetikk 8
 tidsforsinkelse 111, 168
 tidsforskyving 510
 tidsforskyvingsoperator 510
 tidsinvariant 41
 tidskonstant 43
 tidsmålestokk 110
 tidsrespons, skissering 224
 tidsvariant 40
 tidsvarierende differensiallikninger 68
 tilbakekoplet system 237
 tilbakekopling 15, 39
 tilbakekopling, dynamisk 253
 tilbakekopling, monovariable syste-
 mer 240
 tilbakekoplingsløyfe 238
 tilstand 12
 tilstandsestimator 428
 tilstandsestimering 490
 tilstandsrom 75

- tilstandsrommodell 74, 75
 tilstandsvariabel 66, 75
 tilstandsvektor 12
 transferfunksjon 119
 transferfunksjon, strengt proper .. 117
 transferfunksjoner, benevninger .. 128
 transfermatrise 120
 transient 193
 transisjonsfunksjonen 69
 transisjonsmatrise 80
 transisjonsmatrisen 79
 transisjonsmatrisen, Cayley-Hamiltons
 teorem 88
 transponert matrise 577
 transportforsinkelse ... 168, 171, 220,
 279
 transportforsinkelse, frekvensrespons
 200
 transportforsinkelse, integralregule-
 ring 343
 turbinregulator 306, 404
- U**
 udempet resonansfrekvens .. 144, 149
 ulineær differensiallikning 40
 ulineær regulator 327
 ustabilitet 21
- ustabilt system 267
 ustrukturert usikkerhet 453
 utgang(svariabel) 13
- V**
 vannkraftturbin 26
 varmekapasitet 54
 varmeovergangstall 54
 vektordifferensiallikninger 75
 ventilsymbol 25
 vinkelfrekvens 193
- W**
 w-transformasjon 534
 w-transformasjon, multivariabel . 550
- Z**
 z-transformasjon 520
 z-transformasjon, invers 522
 Ziegler-Nichols metode 348, 350
- Ø**
 økonomisk system 28
- Å**
 åpent system 238
 åpent ustabilt system, nyquists stabili-
 tetskriterium 294