LUND UNIVERSITY

Reglerteori

Åström, Karl Johan

1968

Document Version:
Förlagets slutgiltiga version

Link to publication

Citation for published version (APA):
Åström, K. J. (1968). *Reglerteori*. Almqvist & Wiksell.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

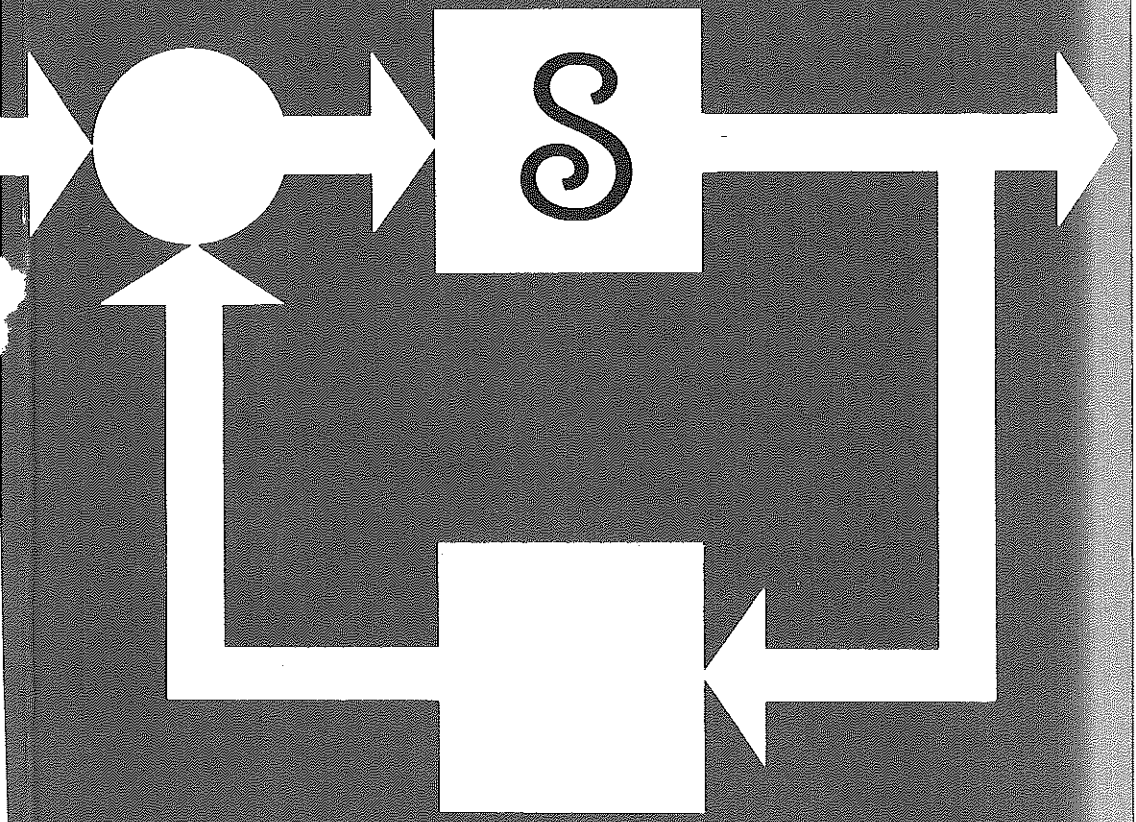
Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Reglerteori



Karl Johan Åström

Almqvist & Wiksell

Reglerteori

Reglertekniken är en ung vetenskap som får allt större betydelse genom den ständigt stigande automatiseringstakten. Dess utveckling har i hög grad påverkats av datateknikens framsteg. Processreglering med datorer är t. ex. ett mycket aktuellt tillämpningsområde.

I REGLERTEORI utvecklas den teori som ligger till grund för analys och konstruktion av reglersystem.

Ur innehållet:

- **Dynamiska system. Teorin utvecklas parallellt för insignal–utsignal-beskrivningar och tillståndsmodeller.**
- **Begreppen styrbarhet och observerbarhet behandlas ingående. Analysen utnyttjas genomgående för att förstå vad som händer med formellt manipulerande av överföringsfunktioner.**
- **Återkopplingsprincipen behandlas i ett kapitel. I detta presenteras intuitiva resonemang som sedan undersöks med matematisk analys.**
- **Stabilitetsbegrepp och stabilitetskriterier diskuteras ingående.**
- **Dimensionering av reglersystem ägnas stort utrymme. Specifikationer, reglerprinciper och reglerstrategier diskuteras översiktligt. Såväl klassiska frekvensanalytiska dimensioneringsmetoder som nyare metoder baserade på tillståndsåterkoppling och Kalmanfilter behandlas ingående.**

REGLERTEORI vänder sig till studerande vid tekniska högskolor och universitet, praktiskt verksamma ingenjörer, som vill färska upp sina kunskaper eller lära på nytt, och lärare vid gymnasieskolan.

Almqvist & Wiksell

ISBN 91-20-04438-0

Reglerteori

Karl Johan Åström

Almqvist & Wiksell

©
Karl Johan Åström
Almqvist & Wiksell Förlag AB, Stockholm 1968

Typografisk formgivning Dick Hallström

Mångfaldigande av innehållet i denna bok, helt eller delvis, är enligt lag om upphovsrätt av den 30 december 1960 förbjudet utan medgivande av förlaget, Almqvist & Wiksell Förlag AB, Stockholm. Förbudet avser såväl text som illustrationer och gäller varje form av mångfaldigande, genom tryckning, duplicering, stencilering, bandinspelning etc.

Första upplagen 1968
Andra omarbetade upplagan
Första tryckningen 1976
Andra tryckningen 1985

Graphic Systems AB
Göteborg 1985

Innehåll

Förord 9

1. Inledning 11

1. Vad är reglerteknik? 11
2. Historik och tillämpningar 14
   Översikt av olika tillämpningar 14
   Ångmaskinreglering 15
   Analys av reglersystem. Stabilitetsteori 17
   Autopiloter 18
   Återkopplade elektroniska förstärkare 20
   Konsolidering av teorin för servomekanismer 21
   Instrumenteringsteknik 21
   Processreglering 23
   Utvecklingen under de senaste åren 24
3. Referenser 26

2. Dynamiska system 31

1. Inledning 31
2. Karakterisering av dynamiska system 33
   Begreppet tillstånd 33
   Differentialekvationsmodeller 34
   Klassificering av dynamiska system 37
3. Linjära tidsinvarianta system 38
   Linjärisering 39
4. Lösning av systemekvationen 41
5. En funktionalekvation för exp At — kausalitetsprincipen 45
6. Byte av koordinater i tillståndsrummet 46
   Diagonalformen 46
   Kanonisk form 1 47
   Kanonisk form 2 51
7. Insignal–utsignalrelationer. Viktfunktionen 53
8. Laplacetransformen. Överföringsfunktionen 58
9. Grafisk representation av överföringsfunktionen 65
   Singularitetsdiagram 65
   Frekvenskurvor 66
   Bodediagram 69
   Bodes relationer. Minimum-fas-system 73

10. Grafisk representation av dynamiska system 77
Blockschema 78
Summation 78
Förgrening 79
Multiplikation med konstant 79
Dynamiska samband och integration 79
Signalflödesdiagram 82
11. Elementär stabilitetsteori för linjära tidsinvarianta system 83
Definitioner 83
Stabilitetskriterier i enkla fall 85
12. Orientering om den allmänna teorin för dynamiska system 87
Abstrakt definition av dynamiska system 87
13. Referenser 91

3. Styrbarhet och observerbarhet 94
1. Inledning 94
2. Ett exempel 95
3. Definitioner 97
4. Kriterier för styrbarhet 98
5. Kriterier för observerbarhet 102
6. Kalmans uppdelningssats 105
Uppdelning med avseende på styrbarhet 105
Uppdelning med avseende på observerbarhet 106
Uppdelning med avseende på styrbarhet och observerbarhet 107
7. Dualitet 110
8. Olika representationer av systemet $\mathbb{S}(A, B, C, D)$ 110
Tillståndsekvationer 110
Differentialekvationer som relaterar insignal till utsignal 111
Överföringsfunktion och viktfunktion 111
9. Konstruktion av en minimal realisation av en given överförings-
funktion 112
En observerbar realisation 113
En styrbar realisation 115
Flervariabla system 116
10. Referenser 118

4. Analys av återkopplade system 119
1. Inledning 119
2. Varvtalsreglering av en ångmaskin 120
Analys av det öppna systemet 121
Analys av centrifugalregulatorn 122
Analys av det återkopplade systemet 124

Proportionell och integrerande återkoppling 125
Bestämmning av lämpliga parametrar för PI-regulatorn 128
Överföringsfunktionen från belastning till vinkelhastighet. 130
Analys av centrifugalregulatorns dynamik. 131
Analys av ångmaskinreglering med proportionell återkoppling då regulatorns dynamik ej försummas 133
Sammanfattning och generalisering. 136
3. Kvalitativ analys av återkopplade system 137
Statisk noggrannhet. Felkoefficienter 139
Systemets förmåga att eliminera störningar. Styvhetskoefficienter. . 141
Känslighet för parametervariationer 143
4. Referenser 149

5. Stabilitet 150
1. Inledning 150
2. Stabilitet med avseende på störningar i begynnelsevillkor. 151
Definitioner 151
Linjära tidsinvarianta system. 153
Stabila system 154
3. Insignal–utsignalstabilitet 155
4. Översikt av metoder för stabilitetsanalys av linjära tidsinvarianta system 157
5. Rotortmetoden 158
Allmänt. 158
Användning av rotorten för analys av återkopplade system 159
Uppritning av rotorten 162
Funktionsteoretisk tolkning av rotorten 166
6. Routh–Hurwitzkriteriet. 166
Rouths algoritm 167
Hurwitzkriteriet 169
7. Funktionsteoretiska metoder och Nyquistkriteriet. 171
Argumentvariationsprincipen. 171
Nyquistteoremet 173
8. Konvergenshastighet. Praktisk stabilitet. 178
Motivering 178
Starkare villkor på den karakteristiska ekvationens rötter. 179
Förstärkningsmarginal och fasmarginal. 180
Resonanstoppens storlek M_p 181
Bestämning av det slutna systemets överföringsfunktion ur kretsöverföringen 182
Nicholsdiagrammet 184
9. Referenser 186

6. Syntes av reglersystem 188
1. Inledning 188
2. Specifikationer 191
Specifikationer i tidsplanet 192
Specifikationer i frekvensplanet 193
Specifikationer på överföringsfunktionen 195
Bandbreddskompromissen 202
Optimeringskriterier 204
Känslighet 206
Samband mellan olika systemspecifikationer 206
3. Reglerprinciper och regulatorstrukturer 208
PID regulatorn 209
Kaskadreglering 213
Framkoppling 214
En allmän regulatorstruktur som kombinerar framkoppling och återkoppling 215
Otto Smith-regulatorn 221
Statisk optimering och extremalsökande regulatorer 223
Dynamisk optimering 225
Varierande processdynamik — parameterstyrning och adaptiva regulatorer 226
4. Syntes av framkoppling och inversa system 229
Inverser av enkla linjära system 230
Inverser av flervariabla system 231
Inverser av system givna på tillståndsform 233
Inverser av olinjära system 235
Förkortning av poler och nollställen 236
Approximativa inverser 241
Sammanfattning 245
5. Syntes av system med given överföringsfunktion 245
Polplacering genom tillståndsåterkoppling 246
Rekonstruktion med hjälp av derivering 250
Rekonstruktion med dynamiska system 251
Analys av en regulator baserad på tillståndsåterkoppling och rekonstruktion 256
Polplacering genom utsignal-återkoppling 260
Jämförelse mellan metoderna för polplacering 270
Vad begränsar ett systems prestanda 271
6. Kompensering 274
Nyquist–Bode–Nichols kompenseringsmetod 275
Kompensering med rotortmetoden 284
Praktisk realisering av kompenseringsnät 292
7. Referenser 297

Register 300

Förord till andra upplagan

Reglerteknik är ett jämförelsevis nytt tillskott till ingenjörsutbildningen. Ordinarie högskolekurser i ämnet har förekommit först efter andra världskriget. Efter den sena starten har ämnet expanderat snabbt och reglerteknik ingår nu som en viktig del i nästan all ingenjörsutbildning. Högskolekurserna omfattar som regel den teori för servomekanismer som utvecklades under andra världskriget. Karakteristiskt för dessa kurser är att de i stor utsträckning baseras på teorin för Laplacetransformer och analys i frekvensplanet. Ofta är framställningen starkt formaliserad. Reglerteorin har emellertid förändrats genomgripande som ett resultat av de senaste årens forskning. De reglerproblem man ställts inför i samband med rymdteknologi och processreglering har i stor utsträckning stimulerat utvecklingen. Den har även starkt påverkats av datateknikens framsteg. Datorer används nu både för analys och syntes av reglersystem. Det finns s. k. interaktiva system för analys och syntes där en problemlösare kommunicerar med en dator med bildskärm. Datorn blir också allt vanligare som komponent i reglersystem.

Det är karakteristiskt för den nya utvecklingen att man använder analys såväl i tidsplanet som i frekvensplanet liksom att man i större utsträckning än tidigare utnyttjar kända matematiska resultat. Många nya grundläggande begrepp har införts och det är ej längre möjligt att följa med i tidskriftslitteraturen enbart med hjälp av förkunskaper om teorin för servomekanismer. I amerikansk litteratur presenteras de nya resultaten ofta under rubriker som »Modern Control Theory», »The State Space Approach», etc.

Denna lärobok omfattar samma problemkrets som den klassiska teorin för servomekanismer, nämligen teorin för linjära tidsinvarianta system, men teorin presenteras i såväl tids- som frekvensplanet. Vissa nya grundläggande begrepp såsom tillstånd, observerbarhet och styrbarhet införs från början och utnyttjas sedan för att bygga upp teorin. I läroboken presenteras metoder som gör det möjligt att få en överblick av ett problem med relativt enkla analytiska beräkningar. Dessa metoder kompletteras med andra som gör det möjligt att få precisa svar men som kräver beräkningar med dator.

Kap. 1 tjänar som inledning och där ges även en kort historik. Teorin för styrda dynamiska system presenteras i kap. 2. Begreppen styrbarhet och observerbarhet införs i kap. 3. En viktig konsekvens av dessa begrepp är Kalmans uppdelningssats, som säger att ett linjärt system kan delas upp i fyra delsystem utgående från styrbarhets- och observerbarhetsegenskaper. Systemets överföringsfunktion är entydigt bestämd av det delsystem som är både observerbart och styrbart. I kap. 4 behandlas analys av återkopplade system. En stor del av detta kapitel ägnas åt analys av en ångmaskin med centrifugalregulator. Denna del är helt fristående och kan läsas oberoende av boken i

övrigt. Stabilitetsteori presenteras i kap. 5. Syntes av linjära tidsinvarianta system presenteras slutligen i kap. 6. Där ges en översikt av olika reglerprinciper och regulatorstrukturer. Vidare behandlas klassiska metoder för dimensionering av servomekanismer med hjälp av frekvenskurvor och rotortdiagram. Systematiska metoder för att dimensionera framkoppling och återkoppling som ger ett slutet system med en given överföringsfunktion behandlas också. Syntesproblemet diskuteras med utgångspunkt såväl från interna beskrivningar (tillståndsekvationer) som från externa beskrivningar (överföringsfunktioner). Detta ger alternativa lösningsmetoder, olika algoritmer och en insyn i för- och nackdelar med olika systembeskrivningar.

Läroboken förutsätter vissa förkunskaper i matematik, bl. a. linjär algebra, analys, ordinära differentialekvationer, teorin för analytiska funktioner och Laplacetransformation, i en omfattning som motsvarar kurserna i matematik och numerisk analys vid de tekniska högskolorna. Denna upplaga skiljer sig från den tidigare därigenom att kapitel 3 har drastiskt omarbetats och att kapitel 6 som behandlar syntes är nyskrivet. I de övriga kapitlen har smärre ändringar gjorts. Ändringarna har motiverats av erfarenheter som gjorts i samband med högskoleundervisning, vidareutbildning och lösning av praktiska reglerproblem.

Ekvationer, figurer, satser m. m. har numrerats separat inom varje kapitel med en sifferkombination av typen (2.6) där siffran 2 anger det avsnitt där ekvationen förekommer och 6 är ett löpande nummer. Vid enstaka tillfällen är det nödvändigt att referera resultat i ett annat kapitel. I sådana fall skrivs kapitelnumret följt av kolon framför sifferkombinationen. Uttrycket »ekvation (2.6)» refererar således till ekvation (2.6) i det aktuella kapitlet medan uttrycket »ekvation (3:2.6)» avser ekvation (2.6) i kap. 3. Kapitel och avsnittsnummer har dessutom angivits på varje sida.

Vid revisionen har jag haft god hjälp av vänner och medarbetare. Benämningen »tyst» för icke observerbar har föreslagits av professor Lars Erik Zachrisson. Värdefulla förslag till ändringar har givits av tekn. dr Gunnar Bengtsson, tekn. dr Bo Leden, tekn. lic. Sture Lindahl och universitetslektorerna Gustaf Olsson och Björn Wittenmark. Olika versioner av manuskriptet har också lästs av civilingenjörerna Bo Egart, Per Olov Gutman och Lars Pernebo, tekn. dr Torkel Glad, tekn. dr Per Hagander och tekn. lic. Johan Wieslander. Detta har resulterat i att flera oklarheter uppdagats. Jag vill rikta ett varmt tack till min hustru Bia som bistått med inspiration, uppmuntran, formuleringar, ordval och korrekturläsning. Manuskriptet till de nya delarna har skrivits av Eva Dagnegård, och hon har också granskat korrekturet. De nya figurerna har ritats av Britt Marie Carlsson.

Jag vill slutligen tacka fil. dr Karl-Åke Kärnell och redaktör Britt-Marie Gustafsson i Almqvist & Wiksell Förlag AB för ett mycket gott samarbete.

Lund april 1976

Karl Johan Åström

1. Inledning

1. Vad är reglerteknik?

Begreppet reglerteknik kan definieras så att det omfattar varje rationell metod för att styra system. Motivet för reglertekniken är att få ett system att operera mer pålitligt, mer exakt eller mer ekonomiskt trots störningar från systemets omgivning. Ordet system får här tas mycket allmänt. Det kan t. ex. vara en servomotor, en luftvärnskanon, en månraket, ett ångkraftverk, en atomreaktor, en kemisk process, en gatukorsning, en biologisk, ekonomisk, administrativ eller psykologisk process. Historiskt har ofta väsentliga bidrag till reglertekniken vuxit fram ur lösningen till specifika teknologiska problem. Det har sedan visat sig att såväl begreppsbildning som teknik för problemlösning varit oberoende av de specifika teknologiska detaljerna. Metoder som utvecklats för styrning av rymdraketer har t. ex. med framgång kunnat tillämpas på styrning av pappersmaskiner. Metoder för analys och syntes av servomekanismer har funnit användning vid studium av ekonomiska och administrativa processer. Det är typiskt för reglertekniken att den spänner över mycket vida tillämpningsområden. Vad som binder vitt skilda områden samman är att i samtliga fall är de system, som skall styras *dynamiska* och det finns ett *syfte* för att styra processerna.

Signifikant för reglertekniska problem är att alla innehåller följande element: *system*, *omgivning* och *syfte* (fig. 1.1). Systemet karakteriseras av *styrvariabler* eller *insignaler* över vilka vi förfogar samt *mätsignaler* eller *utsignaler* med vars hjälp vi får information om systemet. Insignalerna kan t. ex. vara inställning av ventiler, kommandosignaler till elektriska servon, avfyrande av styrraketer, fattande av beslut etc. Mätsignalerna kan vara signaler från instrument, kontakttillslag etc. Systemet störs på grund av inverkan från omgivningen. För att förenkla teorin brukar man ofta anta att omgivningens verkan på systemet är enkelriktad, dvs. att systemet ej påverkar omgivningen. Detta antagande är, som vi skall se i exempel i det följande, inte alltid korrekt.

Det reglertekniska grundproblemet kan formuleras på följande sätt: *Bestäm med ledning av mätsignalerna en styrsignal sådan att systemets syfte uppfylls trots omgivningens inverkan.*

Problemet omfattar både bestämning av det funktionella sambandet och konstruktion av en apparat som realiserar den önskade relationen mellan insignal och utsignal. Ofta finns även ett flertal andra viktiga underproblem, t. ex.: Vad vinner vi genom att skaffa mer information? Lönar det sig att installera ytterligare mätinstrument? Lönar det sig att vänta med beslutet för att göra en marknadsundersökning?

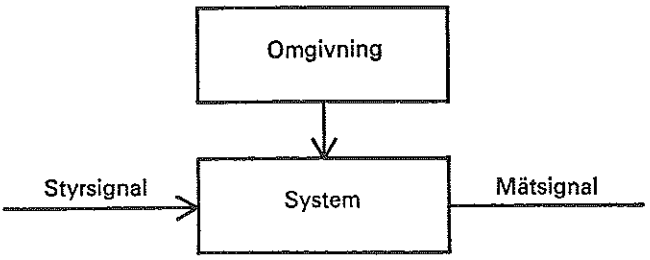


Fig. 1.1. Principschema för ett styrt system.

Man skiljer mellan *öppen styrning* (programstyrning, open-loop control, programming) och *sluten styrning* (återkoppling, closed loop control, feedback control). Vid öppen styrning bestäms styrvariabelns tidsförlopp en gång för alla och det ändras sedan ej under systemets operation. Vid sluten styrning får styrvariabeln bero av mätsignalerna. I det följande skall vi väsentligen studera sluten styrning. Först ges några exempel.

EXEMPEL 1.1. Uppskjutning av satellit

Problem 1. Uppskjutning av satellit i bana runt jorden.

Syfte: Sänd upp en noskon med största möjliga vikt.
Styrvariabler: Dragkraft, anfallsvinkel, stegutlösning.

Lösningen av detta problem resulterar i tidsprogram för dragkraft, anfallsvinkel och stegutlösning. Observera att i problemställningen finns inga mätsignaler. De styrsignaler som beräknas är enbart bestämda av initialdata, således en öppen styrning.

Problem 2. Vi antar att problem 1 är löst och skall nu undersöka hur raketens styrsystem skall konstrueras för att raketen verkligen skall följa det program som uppgjorts.

Syfte: Raketen skall följa den bana som beräknats under problem 1.
Omgivning: Vindbyar, oregelbunden dragkraft orsakad av inhomogeniteter i bränslet etc.
Styrvariabler: Luft- och strålroder, stegutlösning.
Mätsignaler: Markradar som mäter position och hastighet, gyron och accelerometrar i raketen.

Detta är ett typexempel på *sluten* styrning. Inställningen av rodren är beroende av de mätsignaler som erhålls.

EXEMPEL 1.2. Varvtalsreglering av ångmaskin

Ångmaskinen skall hålla konstant varvtal oberoende av variationer i ångtryck och belastning.

Syfte: Konstant varvtal.
Omgivning: Ångtryck och belastning.
Styrsignaler: Ångventil.
Mätsignaler: Vinkelhastighet t. ex. från centrifugalgivare.

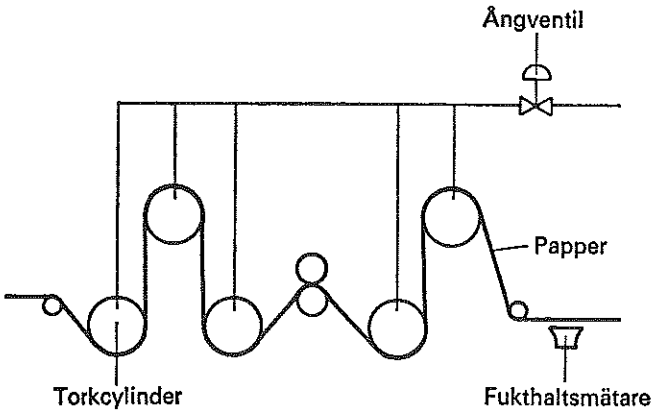


Fig. 1.2. Förenklad bild av torkpartiet i en pappersmaskin.

EXEMPEL 1.3. Fukthaltsreglering i pappersmaskin

I fig. 1.2 visas en förenklad bild av torkpartiet i en pappersmaskin. Det våta pappret (fukthalt 30–60 %) kommer in till vänster i bilden. Pappret får löpa runt ett antal torkcylindrar som värms med ånga. Det torra pappret kommer ut till höger. Det är önskvärt att det utgående pappret har konstant fukthalt.

Syfte: Utgående papprets fukthalt konstant.
Omgivning: Inkommande papprets fukthalt, trycket på ångledningen, ventilation av torkpartiet.
Styrvariabel: Ångtryck i torkcylindrarna.
Mätsignal: Fukthalt t. ex. från kapacitiv fukthaltsgivare.

EXEMPEL 1.4. Lagerhållning

Betrakta ett mellanlager som tar emot delar från ett huvudförråd (eller fabrik) och som levererar delar till köpare. Problemet för lagerinnehavaren är att hålla ett så litet lager som möjligt och att se till att han alltid kan leverera till köparen. Lagerinnehavaren har att ta hänsyn till att priset från huvudleverantören varierar med kvantiteten, leveranstider etc.

Syfte: Maximera vinsten.
Omgivning: Ordertillgång. Variationer i priser från huvudleverantören.
Styrvariabler: Beställning från huvudlager. Annonskampanjer, reklam etc. för att påverka marknaden.
Mätsignaler: Försäljningsstatistik, marknadsundersökningar etc.

EXEMPEL 1.5. Gatukorsning

Betrakta en korsning mellan två enkelriktade gator med trafikljus (fig. 1.3). Problemet i detta fall är hur man skall variera röd–grönperioderna för att få så stora trafikflöden som möjligt.

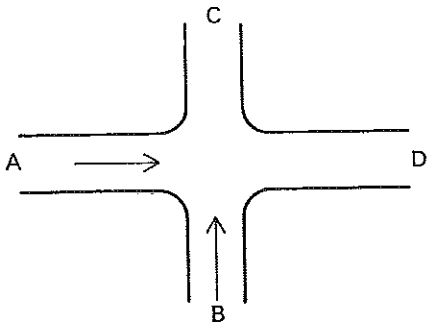


Fig. 1.3. Korsning mellan två enkelriktade gator.

Syfte: Maximera trafikflödet.
Omgivning: Inkommande trafikflöden vid A och B. Eventuella köer vid C och D.
Styrvariabler: Röd–grönperioder hos trafikljusen.
Mätsignaler: Information om trafikflöden från trafikräkningar. Uppgift om kölängd och aktuella trafikflöden från givare i vägbanan.

Detta är ett typiskt fall där det ej är lämpligt att isolera systemet från dess omgivning såsom gjorts i exemplet, ty om långa köer byggs upp vid A och B så kommer detta naturligtvis att påverka närliggande korsning. Antagandet om en enkelriktad inverkan från omgivning till system är således ej tillämpligt. Botemedlet är då att utöka systemet till att omfatta även närliggande gatukorsningar.

2. Historik och tillämpningar

Översikt av olika tillämpningar

Reglerteknikens tillämpningar har en lång historia. Den tidigast kända tillämpningen är ett system för konstbevattning som babylonierna begagnade omkring år 2100 f. Kr. Det finns även tidiga exempel på användning av reglersystem för tidmätning. Vattenklockor var i bruk redan omkring år 500 e. Kr. och Christian Huygens beskrev 1673 en pendelklocka där reglerteknikens principer begagnades. Holländaren Meikle uppfann år 1750 ett reglersystem för automatisk inriktning av väderkvarnar. Reglersystem fick vidsträckt användning först i och med industrialismens genombrott. Utvecklingen av reglertekniken har i stort sett följt industrins utveckling. I dag tillämpas reglerteknik inom en rad vitt skilda områden, t. ex.

Energialstring och energidistribution. Reglering av ångmaskiner, turbiner, vattenkraftverk, termiska kraftverk och atomkraftverk. System för samkörning av olika kraftverk, frekvensreglering och uppstartning.

Farkoststyrning. Autopiloter för flygplan, robotar, båtar, helikoptrar. Navigeringssystem. Automatiska landningssystem.

Militära vapensystem. Inriktning av kanoner och radarantenner. Målsökning. Bombsikten.

Kommunikation. Återkopplade förstärkare. Kretsar för faslåsning och automatik. Stabiliserade spänningsaggregat.

Rymdteknologi. System för satellituppskjutning. Raketstyrning. Månlandning och återfärd.

Instrumentering. Servomekanismer, instrumentservon och skrivare. Mätvärdesomvandlare.

Verkstadsindustri. Automatikutrustningar. Numeriskt styrda verktygsmaskiner.

Processindustri. Reglering av pappersmaskiner och kokare. Reglering och styrning av valsverk och stålugnar, cementugnar, destillationskolonner, värmeväxlare, indunstare, kemiska reaktorer etc.

Medicinsk teknik. Styrning av proteser. Modeller av fysiologiska och biologiska reglerförlopp.

Miljöteknologi. Luftkonditionering och uppvärmning.

I det följande skall vi redogöra för några av dessa tillämpningar i anslutning till den historiska utvecklingen.

Ångmaskinreglering

Watts centrifugalregulator från år 1788 för konstanthållning av en ångmaskins varvtal trots variationer i belastning och ångtryck är ett välkänt reglersystem. En förenklad skiss av Watts regulator visas i fig. 2.1.

För att beskriva centrifugalregulatorns egenskaper införs begreppet förstärkning som förhållandet mellan ändring i ventilläge och ändring av maskinaxelns vinkelhastighet vid en liten ändring av vinkelhastigheten. Regulatorns förstärkning kan påverkas genom att hävstångssystemet ändras. Genom empiriska studier fann man de resultat, som illustreras i fig. 2.2.

Den bestående förändringen i vinkelhastigheten efter en konstant belastningsförändring befanns vara omvänt proportionell mot regulatorns förstärkning. Ur denna synpunkt föreföll det således att vara gynnsamt med en så hög förstärkning som möjligt. Vid konstruktion av regulatorer med höga förstärkningar stötte man på konstruktiva problem. För att få tillräcklig kraft att vrida ventilen, måste kulorna göras tunga. Härigenom erhölls tröghet i regulatorn och man fann att vinkelhastigheten efter en belastningsändring ej svängde

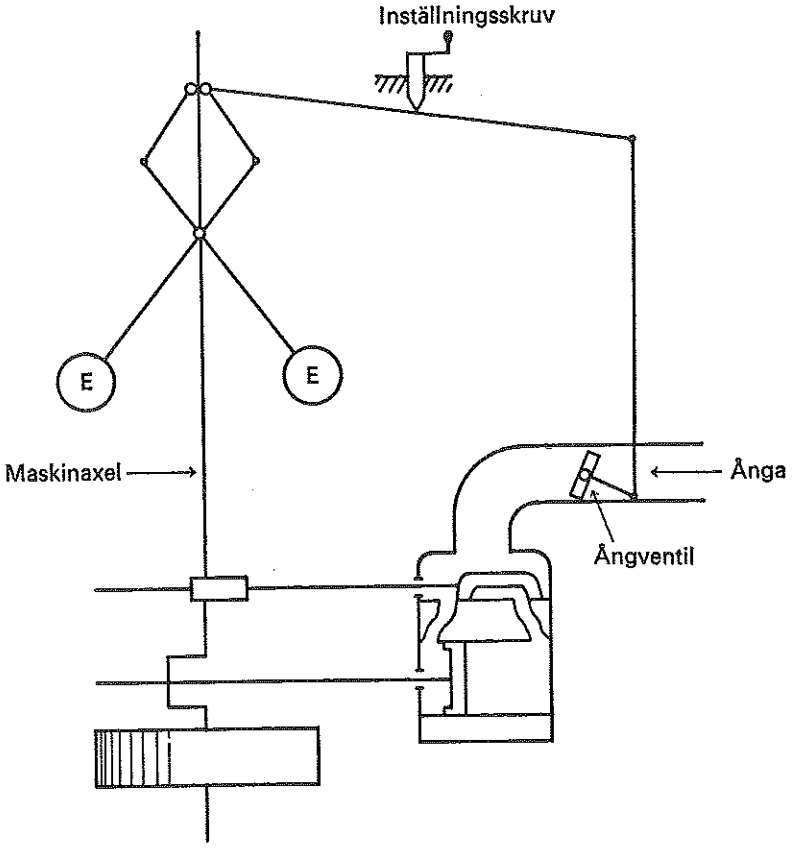


Fig. 2.1. Centrifugalregulator. Reglerkretsen fungerar på följande sätt: Om maskinens varvtal ökar så rör sig kulorna (E) ut från axeln på grund av centrifugalaccelerationen. Genom hävstångsanordningen minskar ångventilens öppning, och maskinens hastighet avtar.

in sig mot ett stationärt värde utan oscillerade med växande amplitud. Systemet blev således instabilt.

Kompromissen mellan stabilitet och förmåga att eliminera störningar är en karakteristisk egenskap för reglersystem som vi många gånger skall återkomma till. Med en regulator enligt fig. 2.1 kan man aldrig helt undvika ett bestående fel i varvtal efter en ändring i belastningen. För att komma över de svårigheter som var förenade med regulatorn i fig. 2.1 prövades många olika konstruktioner, av vilka de flesta nu är helt ointressanta. Fig. 2.3 visar en regulator som har många fördelar i jämförelse med den enkla regulatorn i fig. 2.1.

Denna regulator är intressant från två synpunkter. Dels har inverkan av trögheten hos kulorna i regulatorn eliminerats, dels är ventilens *hastighet* (ej dess läge) proportionell mot differensen mellan pendelns och maskinens rotationshastighet. Vid en hastighetsförändring hos spindeln kommer kulorna på grund av sin tröghet att bibehålla sin hastighet och ventilens läge ändras

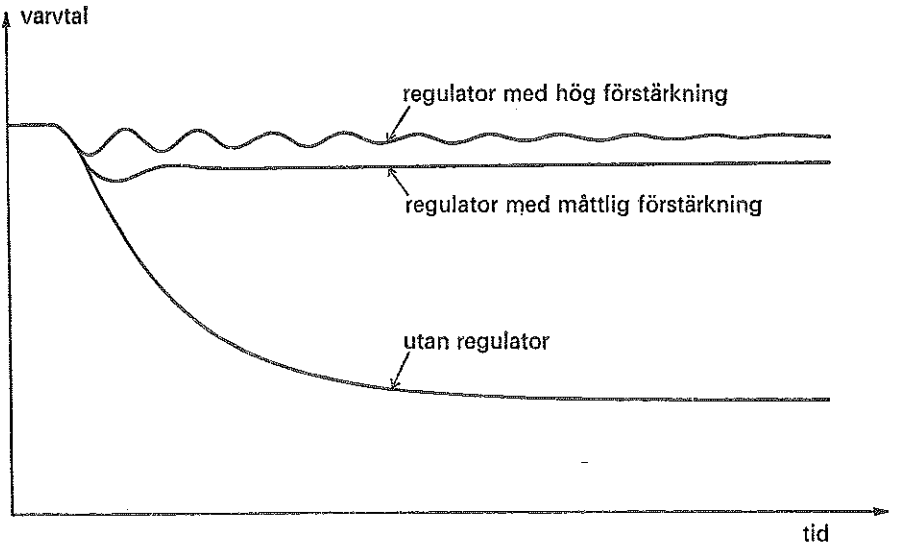


Fig. 2.2. Varvtalsändring vid en belastningsökning för ångmaskin med och utan regulator.

således omedelbart. Experimentella undersökningar visade att med en regulator enligt fig. 2.3 kommer maskinens hastighet i stationärt tillstånd alltid att återgå till referensvärdet efter en stegstörning i belastningen. En regulator enligt fig. 2.3 kallades därför »governor» medan regulatorn i fig. 2.1 kallades »moderator». Regulatorn i fig. 2.1 ger en förändring i ångventilens läge som är proportionell mot avvikelsen i varvtal (proportionell eller P-reglering). Regulatorn i fig. 2.3 ger en förändring i ångventilens läge som är proportionell mot tidsintegralen av varvtalsavvikelsen (integrerande- eller I-reglering).

Analys av reglersystem. Stabilitetsteori

Den teoretiska analysen av reglersystem har i regel kommit efter apparatbyggandet. De tidigast kända arbetena författade av Maxwell (1868) och Vyshnegradskii (1876) behandlade ångmaskinreglering. Maxwell utgick i sin analys från de experimentella resultat som fanns tillgängliga för centrifugalregulatorer. Han visade analytiskt att stationära avvikelser i många fall kunde undvikas genom att man lät styrsignalen vara proportionell mot integralen av regleravvikelsen. Maxwell angav även kriterier för systemets stabilitet och refererar praktiska experiment som verifierar hans slutsatser.

Stabilitetsproblemet, som i enkla fall behandlats redan av Maxwell, kan ofta reduceras till följande algebraiska problem: Vilka villkor skall ställas på koefficienterna $a_0, a_1, \ldots, a_n$ för att samtliga rötter till ekvationen

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0$$

skall ha negativ realdel?

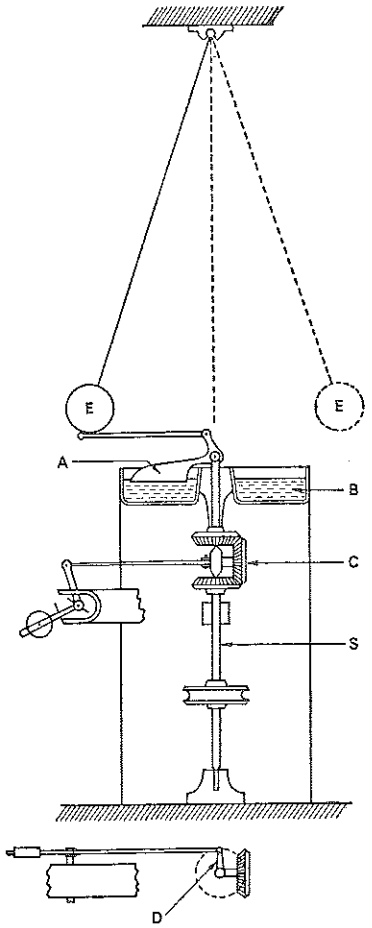


Fig. 2.3. Sofistikerad centrifugalregulator enligt Siemens, Proc. Roy. Soc. *156* (1868), s. 659. Anordningen med vingen (*A*) och vätskebehållaren (*B*) syftar till att driva pendeln med konstant vinkelhastighet. Med hjälp av differentialen (*C*) ges armen (*D*) en vinkelhastighet som är proportionell mot differensen mellan spindelns (*S*) och kulornas (*E*) hastighet. Rörelsen hos armen (*D*) överförs sedan direkt till ångventilen.

Detta föreslogs av Stodola i samband med studier av ångturbiner. Problemet löstes oberoende av Routh och Hurwitz. Stabilitetsproblemet har också studerats från andra utgångspunkter. Lagranges klassiska arbeten om stabilitet hos mekaniska system (ett jämviktsläge är stabilt om potentiella energin har minimum) generaliserades och utvidgades av Lyapunov.

Autopiloter

Bröderna Wrights första lyckade flygning gjordes den 17 december 1903. En av orsakerna till att bröderna Wright lyckades var att de drastiskt bröt mot de gängse konstruktionerna. I stället för att göra ett stabilt flygplan gjorde de ett instabilt men lättmanövrerbart plan. Det faktum att planet var instabilt var naturligtvis en stor belastning för piloten som hela tiden måste ingripa med styråtgärder. Flygplanets manövrerbarhet medförde dock att piloten mycket snabbt kunde motverka de störningar som orsakades av vindbyar. Det låg nära till hands att försöka konstruera en automat för att underlätta

pilotens arbete. Detta gjordes mycket tidigt. Redan 1912 installerades en autopilot i en Glenn H. Curtis-flygbåt. Autopiloten använde gyroskop för att detektera flygplanets avvikelser i attityd. Signalerna från gyroskopen påverkade rodren så att flygplanet bibehöll sin attityd trots störningar. Autopiloten, som konstruerats av Dr. E. A. Sperry, gjorde en remarkabel flygning den 18 juni 1914, då Curtis-flygbåten med autopilot vann ett pris på 50 000 francs i sin egenskap av det stabilaste flygplan som konstruerats. Vi citerar ur Sperrys originalhandbok:

"The Sperry Aeroplane Stabiliser is a carefully-designed, substantial and thoroughly tested instrument which has for its purpose the conception and correction of the slightest departure of an aeroplane from any desired attitude of flight.

This apparatus first came into prominence when it entered the Concours par l'Union pour la Securité en Aeroplane en France in the summer of 1914, where the marvellous feats accomplished by it caused the judges to unhesitatingly award it the first prize. A picture taken by a moving camera the day of the contest (which incidentally has been preserved to this day) is shown on the preceding page. The mechanician is shown standing out on the wing, where he walked back and forth, generating a strong upsetting couple while the pilot, as can be seen, stood with both hands over his head, the equilibrium of the aeroplane being automatically maintained by the gyroscopic stabiliser. An interesting feature is the proximity of the ground at which the demonstrations were conducted.

The full value of such an apparatus can only be appreciated by those who have experienced the difficulties of flying under unfavourable atmospheric conditions. The pilot is not only relieved of the nervous and physical fatigue of maintaining equilibrium of his machine, but, his hands being free, is enabled to draw maps, drop bombs, etc., for which an observer has heretofore been required. The increased ability for carrying fuel and war munitions due to the elimination of the passenger, will at once be recognised as a most important factor."

Det konstruktiva dilemma, som består i att stabilitet och manövrerbarhet ställer diametralt olika krav, löses således för flygplan genom att man konstruerar ett manövrerbart flygplan och därefter gör detta stabilt genom att införa reglersystem. Man kan säga, att reglersystemet ger konstruktören en ytterligare frihetsgrad med vars hjälp han kan lösa besvärliga kompromisser. Utvecklingen av flygplan och autopiloter är ett av de tidiga exemplen på denna konstruktionsfilosofi. Inom kärnkrafttekniken finner vi ett aktuellt exempel på ett liknande betraktelsesätt. De atomreaktorer som konstrueras idag är inherent stabila. En ekonomisk analys visar emellertid att kostnaden för kärnkraft skulle kunna nedbringas avsevärt om man vid utformandet av reaktorhärden ej fordrar inherent stabilitet. Detta kräver emellertid att reaktorn stabiliseras med ett reglersystem som måste vara exceptionellt driftsäkert.

Efter de första experimenten har alltmer sofistikerade autopiloter konstruerats. År 1947 genomförde ett amerikanskt militärflygplan C-47 en helt automatisk flygning över Atlanten inkluderande såväl automatisk start som automatisk landning. Styrsystem för robotar och raketer utgör en naturlig vidareutveckling på autopiloter för flygplan. Styrsystemen för månraketen APOLLO och dess månlandare är några av de mest avancerade reglersystem som konstruerats.

Återkopplade elektroniska förstärkare

Utvecklingen av långdistanstelefonin 1910–1940 krävde elektroniska förstärkare av mycket hög kvalitet. Konstruktion av förstklassiga elektroniska förstärkare var teknologiskt mycket svårt och kunde lösas först genom ett väsentligt bidrag av reglertekniken: Blacks uppfinning av den återkopplade förstärkaren. Vi illustrerar med ett utdrag ur ett föredrag av Bode, *Feedback — The History of an Idea*, Proc. Symp. on Active Networks and Feedback systems, Brooklyn Polytechnical Institute 1960.

"Most of you with hi-fi systems are no doubt proud of the quality of your audio amplifiers, but I doubt whether many of you would care to listen to the sound after the signal had gone in succession through several dozen or several hundred even of your own fine amplifiers. There is a 'tyranny of numbers', as my reliability friends say, which makes it necessary for the individual components of the system to become qualitatively better as the system as a whole becomes quantitatively more ambitious.

The causes of distortion were of various sorts. They included power supply noises, variations in gain, and so on. The dominant problem, however, was the intermodulation due to the slight nonlinearity in the characteristics of the last tube. Various efforts were made to improve this situation, by the selection of tubes, by careful biasing, by the use of matched tubes in push-pull to provide compensating characteristics, and so on. Until Blacks invention, however, nothing made a radical improvement in the situation."

Black visade att om man byggde en förstärkare med mycket hög förstärkning och återkopplade denna, så erhöll man en förstärkare som var mycket stabil och okänslig för parametervariationer. Forskningen på återkopplade förstärkare, som givit många bidrag till reglertekniken, utfördes väsentligen på Bell Telephone Laboratories. Bland de mest framträdande forskarna kan vi nämna Bode, Nyquist, MacCall och McMillan.

Liksom vid analys av ångturbiner och centrifugalregulatorer fann man att stabilitetsproblemet är ett av de viktigaste problemen i samband med konstruktion av återkopplade förstärkare. Nyquist-kriteriet för att avgöra stabilitet av ett återkopplat system är ett väsentligt resultat. Den teknik som utvecklades för analys av återkopplade förstärkare var baserade på transformmetoder (Laplace och Fourier) och funktionsteori.

Konsolidering av teorin för servomekanismer

Under andra världskriget insattes stora resurser på utveckling av vapensystem, bl. a. omfattande konstruktion av servomekanismer för styrning av båtar, flygplan, kanoner och radarantenner. En stor del av utvecklingsarbetet i USA utfördes vid tre laboratorier vid Massachussets Institute of Technology (MIT), Radiation Laboratory, Servomechanisms Laboratory och Instrumentation Laboratory. Teorin för servomekanismer utvecklades, man byggde främst vidare på de metoder, baserade på Laplacetransformering och funktionsteori, som framgångsrikt hade använts vid konstruktion av återkopplade förstärkare. Procedurer för att dimensionera servomekanismer utvecklades. Dessa procedurer utgick från specifikation av systemets snabbhet, statiska fel och känslighet för parametervariationer med kvantitativa värden. Man angav sedan metoder med vars hjälp man kunde dimensionera reglersystem som uppfyller sådana specifikationer, s. k. *kompenseringsmetoder*.

Många nya komponenter utvecklades även. Detta medförde att det var tekniskt möjligt att praktiskt realisera alltmer avancerade system. Speciellt utvecklades elektromekaniska komponenter såsom elgoner, servomotorer, tachometergeneratorer, gyroskop, accelerometrar, elektromagnetiska givare, *E*-transformatorer, microsyn etc. Tillkomsten av likspänningsförstärkare med mycket låg drift, t. ex. Goldberg-förstärkare, gav nya möjligheter att på ett enkelt sätt praktiskt realisera komplicerade överföringsfunktioner. Den ledde även till utvecklandet av analogimaskiner, som dels användes som räknehjälpmedel vid systemsyntes, dels som komponent i reglersystemen.

Som en följd av den intensiva utvecklingen under andra världskriget framstod reglertekniken efter kriget som en självständig vetenskap. Kurser i reglerteknik väsentligen omfattande teorin för servomekanismer, infördes som obligatoriska ämnen vid de flesta tekniska högskolor, ett flertal goda läroböcker skrevs, bl. a. av James, Nichols & Phillips, Chestnut & Mayer och Tsien.

Det reglertekniska syntesproblemet är ej helt tillfredsställande löst inom teorin för servomekanismer. De kvantitativa mått som används för att formulera systemets syfte, leder sällan till entydiga lösningar. Detta medför att de vanliga syntesprocedurerna i servomekanismteorin innehåller många tämligen godtyckliga val. Dessutom kan syntesmetoderna endast med svårighet tillämpas på system, som har flera insignaler och flera utsignaler.

Instrumenteringsteknik

Inom instrumenteringstekniken finns många reglertekniska tillämpningar. Genom att bygga instrumenten som servosystem kan man uppnå många fördelar. Vi illustrerar detta med att diskutera en accelerometer. En enkel accelerometer består av en fjäderupphängd massa (fig. 2.4).

Låt x beteckna accelerometerns utsignal, dvs. massans avvikelse från referensläget. Om massan är m, fjäderkonstanten k och dämpningskoefficienten d erhålls följande rörelseekvation för massan

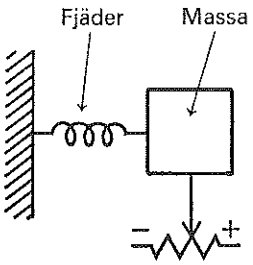


Fig. 2.4. Enkel accelerometer.

$$m\frac{d^2x}{dt^2} + d\frac{dx}{dt} + kx = ma \tag{2.1}$$

där a är accelerometerhöljets acceleration. I stationärt tillstånd erhålls utslaget

$$x^\circ = (m/k)a^\circ$$

Accelerometerns känslighet är således

$$S = m/k$$

Svängningsekvationen (2.1) har vidare egenfrekvensen

$$\omega = \sqrt{k/m} = \sqrt{1/S} \tag{2.2}$$

Om accelerometern skall kunna registrera snabba förlopp är det således önskvärt att den har hög egenfrekvens ω. Enligt ekvation (2.2) innebär detta emellertid att accelerometern samtidigt får låg känslighet. Kraven på hög känslighet och hög egenfrekvens är således motstridiga. Man skulle kunna komma ifrån detta dilemma genom att göra en okänslig accelerometer och förse den med lägesomvandlare, som omvandlar massans utslag i en spänning, följd av en förstärkare. Man kan emellertid få ett betydligt bättre instrument med en konstruktion enligt fig. 2.5.

I detta instrument lagras massan mycket lättrörligt. Massan förses med en strömspole med vars hjälp en kraft kan påverka massan, en lägesgivare

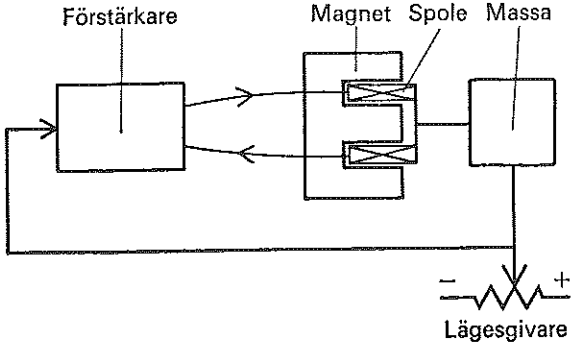


Fig. 2.5. Accelerometer konstruerad enligt kraftbalansprincipen.

och en förstärkare. Systemet utformas så att kraften från strömspolen alltid återför massan till nolläget. Då höljet accelererar kommer således en ström av sådan styrka att sändas genom spolen att den av accelerationen orsakade tröghetskraften balanseras. Strömmen genom strömspolen är således direkt proportionell mot accelerationen. I detta fall kan ett mycket känsligt system med hög egenfrekvens erhållas. Känsligheten bestäms väsentligen av lagringen, medan bandbredden bestäms av servosystemet. Den grundläggande idé som illustrerats här återfinns i ett flertal olika instrument såsom potentiometerskrivare, operationsförstärkare, mätvärdesomvandlare och givare.

Processreglering

Automatisering av industriella processer har fortgått ända sedan industrialismens genombrott. Varvtalsreglering av ångmaskiner är ett tidigt exempel. Proportionella och integrerande regulatorer av ungefär samma typ som diskuterades i samband med ångmaskinreglering, infördes tidigt för reglering av många typer av processer. Regulatorerna byggdes ofta samman med mätgivare. Dessa regulatorer var i stor utsträckning specialtillverkade för speciella processer. Det fanns spridda ansatser till analys av sådana regulatorer under mitten av 1930-talet, men processreglering var på denna tid mer en konst än en vetenskap. Processerna dimensionerades oftast från rent statiska betraktelser. Reglering och instrumentering betraktades som helt separata frågor. Regulatorerna infördes efter det att processerna tagits i bruk. Sambandet mellan processdynamik och regulator beaktades ej. I slutet av 1940-talet började man tillämpa teorin för servomekanismer på processreglering. Detta orsakade bl. a. att man började inse sammanhanget mellan processens och regulatorns egenskaper och betrakta process och regulator som en enhet. Man började intressera sig för processernas dynamiska egenskaper vid projekteringen av system. Vid tillämpning av teorin för servomekanismer stötte man emellertid på flera svårigheter.

- Processernas dynamiska egenskaper var ej kända.
- Det fanns ej instrument för att mäta de storheter som är av intresse. Speciellt saknades instrument för att mäta sammansättningar och koncentrationer.
- Teorin för servomekanismer kan endast med svårighet tillämpas på system med flera insignaler och flera utsignaler.

De flesta industriella processer har ett stort antal insignaler och utsignaler. Till att börja med installerades endast någon enstaka regulator men så småningom blev ambitionen allt större och man önskade reglera allt flera processvariabler. Man fann att det var omöjligt att göra alla reglerkretsar snabba på grund av interferens mellan de olika reglerkretsarna. Inställningen av en regulator i en reglerkrets kunde påverka stabiliten i andra kretsar. Olika

regulatorer kunde ibland motverka varandra. I teorin för servomekanismer är systemets syfte ofta att reproducera en viss signal. Man fann även att detta ej passade in på många situationer inom processindustrin. I vissa fall kunde dock avsevärda framsteg göras. Framför allt förde det ökade intresset för processdynamik det goda med sig, att man började dimensionera processer med tanke på att de skulle kunna regleras. Man gjorde även erfarenheter av samma typ som inom flyg- och instrumenteringstekniken, nämligen att regleringen ger en ytterligare frihet åt projektören.

Utvecklingen under de senaste åren

Vid försök att tillämpa teorin för servomekanismer på processreglering och även på många reglerproblem som uppträder inom rymdteknologin fann man att teorin ej var tillfyllest. Bland de största nackdelarna var bristen på entydighet och att man i allmänhet endast kunde behandla system med en insignal och en utsignal.

Som en direkt följd av dessa nackdelar sökte man andra metoder för att lösa syntesproblemet. Ett väsentligt bidrag gjorde bl. a. Wiener och Phillips i början av 1940-talet då de föreslog att systemets syfte skulle definieras med en skalär funktion, *kriteriefunktionen, vinstfunktionen* eller *förlustfunktionen,* som t. ex. anger den ekonomiska vinsten eller förlusten med systemets operation. Syntesproblemet kan då formuleras som ett extremalproblem: bestäm en styrlag sådan att kriteriefunktionen minimeras eller maximeras. På detta sätt kunde den rika floran av heuristiska kriterier elimineras. Man får en exakt definition av systemets syfte och lägger grunden till systematiska syntesmetoder. Efter 1950 skedde en snabb utveckling. Omformuleringen av syntesproblemet medförde att variationskalkylens kraftfulla resultat kunde utnyttjas. De speciella reglertekniska problemen ledde även till en vidareutveckling av den klassiska variationskalkylen. Bellman utvecklade *Dynamisk Programmering* och Pontryagin bidrog med *Maximumprincipen.* En väsentlig fördel med optimeringsteorin är att man får veta vad som överhuvudtaget kan åstadkommas och vad som begränsar systemets prestanda. Detta ger en måttstock med vars hjälp man kan bedöma olika teknologiska approximationer.

Wiener lämnade även ett annat väsentligt bidrag då han föreslog, att systemets omgivning skulle karakteriseras med hjälp av störningar som beskrivs som stokastiska processer. Detta har i kombination med variationskalkylen givit upphov till en rad mycket fruktbara problemställningar, vilka i sin tur har givit nya syntesmetoder.

Utvecklingen har i stor utsträckning påverkats av datorteknikens framsteg. Datorerna har dels gjort det möjligt att genomföra de beräkningar som krävs för analys och simulering av komplicerade system, dels givit en praktisk möjlighet att förverkliga även mycket komplexa styrsystem. Den första datorn avsedd för industriell processreglering tillverkades och installerades år 1959 av det amerikanska företaget Bunker-Ramo. Utvecklingen har sedan gått raskt

framåt. Många datorfabrikanter levererar nu speciella s. k. *processdatorer* för reglering av industriella processer. I slutet av år 1975 fanns t. ex. 130 processdatorer installerade enbart i Sveriges pappersindustri.

De första reglersystemen byggdes och brukades långt innan de analyserades. Maxwells analys kom t. ex. ca 80 år efter det att James Watt m. fl. framgångsrikt hade byggt regulatorer. I dag är situationen i många fall den omvända. Konstruktion av processreglersystem, styrsystem för raketer och moderna vapen bygger i stor utsträckning på analys och simulering, som kan vara mycket omfattande. Utvecklingen inom reglerteorin har också gått mycket snabbt. Det finns nu en rad syntesmetoder, som ännu ej har prövats praktiskt.

Tillämpningsfältet för reglerteknik har även ökat avsevärt under de senaste tio åren. De ständigt ökande tillämpningarna har närmat reglertekniken till många andra ämnen. Vid lösning av praktiska processregleringsproblem har det t. ex. visat sig vara mycket svårt att separera processreglering och produktionsplanering. Reglertekniken har härigenom fått en anknytning till operationsanalys. Utnyttjandet av datamaskiner har även givit anknytning till automatateori och informationsbehandling. Den praktiska utformningen av ett processreglersystem är t. ex. starkt beroende av de programmeringssystem som finns tillgängliga för processdatorer.

Under de senaste åren har man även kunnat märka en utvecklingslinje där reglertekniken tillämpas på problem inom medicin, biologi, ekonomi och sociologi. Vi illustrerar detta med att citera ur förordet till en bok, *The Application of Control Theory to Physiological Systems*, W. B. Saunders, Philadelphia, 1966 författad av H. T. Milhorn, professor i fysiologi och biofysik.

"The physiologist should ask himself, 'How may I, as a physiologist, be benefited by a knowledge of control theory?' The answer to this question should become quite evident upon reading the next few paragraphs.

It has been known for many years that homeostatic mechanisms maintain a constant internal environment. For instance, the blood pH level, water and electrolyte concentrations, body temperature, blood P_{CO_2}, arterial pressure, and many other physiological variables all are regulated at fairly constant values. What are the mechanisms which regulate variables such as these and how do they operate?

Many diseased states are simply abnormalities of a particular physiological control system. For instance, let us consider the following example: A patient comes into a neurological clinic. He is not paralyzed and he can move both of his legs when he is given the order. Yet he walks with an uncertain gait, with eyes downcast on the ground and his legs. He starts each step with a kick, throwing each leg in front of him in succession. If blindfolded, he cannot stand and totters to the ground. What is the matter with him?

This example involves an abnormality (diseased state) of a particular control system. This patient suffers from tabes dorsalis. The part of the spinal cord

which receives sensations has been damaged or destroyed by syphilis. Since his central nervous system is unable to receive the messages from the receptors in his joints and feet which normally inform him of the position and state of motion of his legs, he must rely upon his eyes and the balancing organs of his inner ear for information concerning his posture. In other words, he has lost part of his proprioceptive sense. In this case, the feedback loop has been broken by the lesion so that his brain does not know what his legs are doing and, hence, cannot correct the errors in leg position and motion that occur unless another feedback loop is substituted for the disrupted one."

Många av de icke tekniska tillämpningarna har sin grund i Norbert Wieners arbeten om cybernetik i början av 1940-talet. Den amerikanske matematikern Richard Bellman, som givit många väsentliga bidrag till reglertekniken, skriver i en uppsats, Control Theory, i Scientific American, september 1964:

"Norbert Wiener, the eminent mathematician who died last winter, formulated the provocative idea that it should be possible to develop a unified theory of feedback control applicable both to living organisms and machines. To express this idea he coined the term 'cybernetics'. It was his hope, shared by others that techniques used so successfully in control engineering could be applied to biomedical problems (for example the design of artificial human organs) and also that research into neurophysiology might provide valuable clues in the design and study of communication systems, computers and more general control systems of all kinds. But as mathematicians, physiologists and engineers explore the subtle difficulties of dealing with large-scale systems—living and nonliving—of different degrees of complexity, it seems less and less likely that any single 'cybernetic' theory will serve all purposes. Nonetheless, for those who want to understand both modern science and modern society, there is no better place to start than control theory." (Ur Richard Bellman, Control Theory. © September 1964 Scientific American, Inc.)

3. Referenser

Flera av de tidiga grundläggande arbetena bl. a. av Maxwell, Nyquist, Bode m. fl. har samlats i

Bellman, R. & Kalaba, R. (utg.), *Selected Papers on Mathematical Trends in Control Theory*, Dover, New York 1964.

Denna bok är mycket läsvärd och även billig.

Utmärkta redogörelser av den servomekanismteori som utvecklades vid MIT under andra världskriget har givits i

James, H. M., Nichols, N. B. & Phillips, R. S., *Theory of Servomechanisms*, McGraw-Hill, New York 1947

Brown, G. S. & Campbell, D. P., *Principles of Servomechanisms*, Wiley, New York 1948.

Nichols var ledare för Servomechanisms Group vid Radiation Laboratory och Brown var chef för Servomechanisms Laboratory.

En lärobok med något fylligare framställning av samma stoff är

Chestnut, H. & Mayer, R. W., *Servomechanisms and Regulating System Design*, Vol. 1, Wiley, New York 1959.

Denna bok har även en del II som behandlar olinjära servomekanismer.

Läroböckerna

Truxal, J. G., *Automatic Feedback Control System Synthesis*, McGraw-Hill, New York 1955

Tsien, H. S., *Engineering Cybernetics*, McGraw-Hill, New York 1954

är av senare datum. Truxals bok är att betrakta som standardverk för servomekanismteorin baserad på Laplacetransform. Tsien, som 1954 var professor vid California Institute of Technology, var mycket framsynt. I hans bok finns många antydningar till hur den fortsatta reglerteorin kommer att utvecklas. Tsien är för närvarande en av männen bakom Kinas ansträngningar att framställa ballistiska missiler. I Truxals bok finns bl. a. ett tidigt försök att utveckla systematiska syntesmetoder baserat på idéer från nätverkssyntes.

Ett tidigt försök att formulera syntesproblemet som ett optimeringsproblem presenteras i

Newton, G. C., Gould, L. A. & Kaiser, J. F., *Analytical Design of Linear Feedback Controls*, Wiley, New York 1957.

De metoder som presenteras där har emellertid begränsad användning.

Mer systematisk framställning av optimeringsteorin har givits i originalbidragen

Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V. & Mishchenko, E. F., *The Mathematical Theory of Optimal Processes*, Wiley, New York 1962

Bellman, R., *Adaptive Control Processes — A Guided Tour*, Princeton University Press, Princeton, New Jersey 1961.

Pontryagins bok är klar men ej helt lättläst. Bellmans bok är lätt att läsa. Den innehåller många idéer, men analysen är sällan fullständigt genomförd.

En detaljerad och välskriven framställning av optimeringsteorin har givits i

Athans, M. & Falb, P., *Optimal Control*, McGraw-Hill, New York 1966.

Som exempel på något nyare läroböcker kan nämnas

D'Azzo, J. J. & Houpis, C. H., *Feedback Control System Analysis and Synthesis*, McGraw-Hill, New York 1960

Gille, J. C., Pelegrin, M. J. & Decaulne, P., *Feedback Control Systems*, McGraw-Hill, New York 1959

Shinners, S. M., *Control Systems Design*, Wiley, New York 1964

Dorf, R. C., *Modern Control Systems*, Addison-Wesley, Reading, Mass. 1967.

Dessa böcker är direkta avläggare av James m. fl. och Truxals böcker.

På grund av teoriutvecklingen under 1960-talet fick läroböckerna i reglerteknik en annan inriktning därigenom att teorin ej baseras på frekvensanalys utan i stor utsträckning på teorin för differentialekvationer. Det brukar i amerikansk litteratur kallas »the State Space Approach». Typiska exempel på sådana läroböcker är

Derusso, P. M., Roy, J. R. & Close, C. M., *State Variables for Engineers*, Wiley, New York 1965.

Elgerd, O., *Control Systems Theory*, McGraw-Hill, New York 1967.

Ogota, K., *State Space Analysis of Control Systems*, Prentice-Hall, New Jersey 1967.

Schultz, D. G. & Melsa, J. M., *State Functions and Linear Control Systems*, McGraw-Hill, New York 1967.

Zadeh, L. A. & Desoer, C. A., *Linear System Theory*, McGraw-Hill, New York 1963.

De frekvensanalytiska metoderna upplever för närvarande en renässans då de utvidgats till flervariabla system. En tankeväckande presentation finns i

Rosenbrock, H. H., *State-space and Multivariable Theory*, Nelson, London 1970.

I den tidigare upplagan av föreliggande lärobok

Åström, K. J., *Reglerteori*, Almqvist & Wiksell, Stockholm 1968,

gjordes ett försök att ge en enhetlig presentation baserad både på frekvensanalys och differentialekvationer.

Den tyska läroboken

Oppelt, W., *Kleines Handbuch Technischer Regelvorgänge*, Verlag Chemie-GMBH, Weinheim 1964

innehåller många exempel från tillämpningarna.

Följande böcker beskriver reglerkomponenter

von Hamos, L. och Attebo, G., *Regleringsteknik*, Svenska Bokförlaget Bonniers, Lund 1964

Solheim, O. A., *Instrumenteringsteknikk*, Tapir, Trondheim 1966

Ahrendt, W. R. & Savant, C. J., *Servomechanism Practice*, McGraw-Hill, New York 1960.

Hydrauliska och pneumatiska servosystem behandlas i

Blackburn, J. F., Reethof, G. & Shearer, J. L. (utg.), *Fluid Power Control*, The MIT Press, Cambridge, Mass. 1960.

Tillämpningar på reglering av ångsystem beskrivs i

Profos, P., *Die Regelung von Dampfanlagen*, Springer Verlag, Berlin 1962.

Reglering av atomkraftverk behandlas i

Harrer, J. M., *Nuclear Reactor Control Engineering*, D. Van Nostrand, Princeton, New Jersey 1963.

En trevlig framställning av autopiloter och närliggande problem har gjorts i

Blakelock, J. H., *Automatic Control of Aircraft and Missiles*, Wiley, New York 1965.

I denna bok finns bl. a. en läsvärd redogörelse för adaptiva autopiloter.

Reglering av industriella processer behandlas t. ex. i

Buckley, P. S., *Techniques of Process Control*, Wiley, New York 1964

Campbell, D. P., *Process Dynamics*, Wiley, New York 1958.

Eckman, D. P., *Principles of Industrial Process Control*, Wiley, New York 1945.

Eckman, D. P., *Automatic Process Control*, Wiley, New York 1958.

Harriott, P., *Process Control*, McGraw-Hill, New York 1964.

Shinskey, F. G., *Process-Control Systems*, McGraw-Hill, New York 1967.

En översiktlig framställning av datamaskinstyrning av industriella processer har givits i

Savas, E. S., *Computer Control of Industrial Processes*, McGraw-Hill, New York 1965.

Shinskey, F. G., *Process-Control Systems*, McGraw-Hill, New York 1967.

Tillämpningar av reglerteknik på ekonomiska och administrativa processer behandlas i

Forrester, J. W., *Industrial Dynamics*, MIT Press, Cambridge, Mass. 1961

Harrison, T. J. (red.), *Handbook of Industrial Control Computers*, Wiley, New York 1972.

Holt, C. C. m. fl., *Planning Production, Inventories and Work Force*, Prentice Hall, New York 1960.

Forrester har med åren blivit allt djärvare i sina försök att använda reglerteori på alltmer komplicerade processer. I boken

Forrester, J. W., *Urban Dynamics*, MIT Press, Cambridge, Mass. 1969,

undersöker han problem som storstadsförslumning och i boken

Forrester, J. W., *World Dynamics*, Wright-Allen Press, Cambridge, Mass. 1971,

förutsäger han världens undergång genom överbefolkning, svält och miljöförstöring. Forrester beskriver världen som ett system av fem ordinära differentialekvationer, vilket kan synas förenklat i överkant. Genom skicklig marknadsföring har boken dock givit Forrester en unik ställning som en av våra främsta domedagspredikanter.

Medicinska och fysiologiska problem behandlas i

Brown, J. & Gann, D., *Engineering Principles in Physiology*, Academic Press, New York 1973.

Brown, J., Jacobs, J. & Stark, L., *Biomedical Engineering*, Davis, Philadelphia, Penn. 1971.

Clynes, M. & Milsum, J. H., *Biomedical Engineering Systems*, McGraw-Hill, New York 1970.

Goel, N. S., Maitra, S. C. & Montroll, E. W., *Nonlinear Models of Interacting Populations*, American Physical Society, Academic Press, New York 1971.

Milhorn, Jr., H. T., *The Application of Control Theory to Physiological Systems*, W. B. Saunders, London 1966

Milsum, J. H., *Biological Control Systems Analysis*, McGraw-Hill, New York 1966.

Stark, L., *Neurological Control Systems*, Plenum Press, New York 1968.

Riggs, D. S., *Control Theory and Physiological Feedback Mechanisms*, Williams & Wilkins, Baltimore, Maryland 1970.

Cybernetik kan lämpligen studeras i Wieners originalarbete

Wiener, N., *Cybernetics—or Control and Communication in the Animal and the Machine*, The MIT Press, Cambridge, Mass. 1961.

International Federation of Automatic Control (IFAC) organiserar kongresser om olika delområden varje år och en stor kongress vart tredje år som täcker allt från teori till tillämpningar. Kongressreferaten som publiceras regelbundet är mycket bra källor för den som vill följa med i utvecklingen.

De stora amerikanska ingenjörsorganisationerna

The American Institute of Aeronautics and Astronautics (AIAA)
The American Institute of Chemical Engineers (AIChE)
The American Society of Mechanical Engineers (ASME)
The Institute of Electrical and Electronics Engineers (IEEE)
The Instrument Society of America (ISA)

organiserar gemensamt en årlig konferens Joint Automatic Control Conference (JACC). Föredragen från denna konferens brukar publiceras i de olika ingenjörsorganisationernas tidskrifter

AIAA Journal; Chem. Engr. Progr.; Trans. ASME; ASME Journal on Basic Engineering; IEEE Transaction on Automatic Control; IEEE Transaction on Industry and General Applications; IEEE Transaction on Aerospace and Electronic Systems;
ISA Journal—Instrumentation—Systems—Automatic Control.

Bland övriga tidskrifter kan nämnas

Automation and Remote Control; Regelungstechnik; SIAM Journal on Control; Automatica; International Journal on Control; Information and Control; Control Engineering; Control.

2. Dynamiska system

1. Inledning

I detta kapitel presenteras den matematiska teori med vars hjälp vi skall beskriva de system som skall styras. Vi skall utveckla metoder, som gör det möjligt att matematiskt beskriva vad som händer med ett system då dess styrvariabler förändras. Den enklast tänkbara modellen vore att ange sambandet mellan processvariabeln x och styrvariabeln u med ett funktionssamband

$$x(t) = f(u(t))$$

Ett system som kan beskrivas på detta sätt kallas ett *statiskt* system och karakteriseras alltså av att processvariabelns värde vid tiden t beror enbart av styrvariabelns värde vid tiden t. De flesta system kan ej beskrivas på detta sätt, ty det är vanligt att processvariabelns värde vid tidpunkten t är beroende av insignalens förlopp i ett intervall (t_0, t). Sådana system kallas *dynamiska*. Den teori för dynamiska system som skall presenteras i detta kapitel är ej specifik för någon speciell typ av system, t. ex. elektriskt, mekaniskt, pneumatiskt eller termiskt. Det har visat sig att teorin kan tillämpas för att beskriva en rad mycket olika system. Detta är ett av skälen till att reglerteorin fått så vidsträckt användning. Teorin för dynamiska system har sina rötter i den klassiska mekaniken (Galilei, Newton, Laplace, Lagrange, Hamilton, Jacobi, Poincaré m. fl.). Teorin utvecklades först för att matematiskt beskriva och analysera himlakropparnas rörelse. På ett tidigt stadium leddes man till att utveckla integral- och differentialkalkylen, och man beskrev planeternas rörelser med hjälp av differentialekvationer. Det har sedan visat sig att många fysikaliska förlopp kan beskrivas med hjälp av differentialekvationer. Vi erinrar t. ex. om lagarna för konservering av massa, energi och rörelsemängd. Man fann tidigt att de ekvationer som erhölls kunde lösas analytiskt endast i vissa specialfall, och man försökte då dra slutsatser om rörelsen utan att först integrera rörelseekvationerna. Detta ledde så småningom till s. k. kvalitativa metoder (Hill, Poincaré, Lefschetz, Nemytskii, Stepanov, Carathéodory m. fl.). Via arbeten av bl. a. den välkände amerikanske matematikern Birkhoff har teorin för dynamiska system givits en fast form och är nu en självständig disciplin. Teorin för dynamiska system har fått ett nytt idéinnehåll genom impulser från elektrotekniken (Nyquist, Bode, Guillemin m. fl.). Vid studiet av planetrörelser var man väsentligen intresserad av att

analysera hur planeterna rör sig. Möjligheterna att utifrån påverka rörelsen var mycket små. Inom elektrotekniken finns helt andra problemställningar, bl. a. av den orsaken att systemen kan påverkas utifrån, och därför utvecklades metoder för att beskriva hur systemets utsignaler beror av systemets insignaler. Då den tekniska användningen ofta motiverades av att man önskade omforma signaler på givet sätt med hjälp av dynamiska system (filter), utvecklades även konstruktiva metoder för detta. Grovt generaliserat kan man säga att man inom mekaniken är mest intresserad av hur systemets tillstånd förändras med tiden, interna beskrivningar, medan man inom elektrotekniken huvudsakligen är intresserad av hur systemets utsignal beror av dess insignal, externa beskrivningar eller insignal–utsignal beskrivningar. Även den matematiska tekniken skiljer sig. Inom mekaniken betonas differentialekvationer, medan elektrotekniken huvudsakligen begagnar transformteori.

De tidigaste reglertekniska undersökningarna (Maxwell m. fl.) genomfördes uteslutande med hjälp av differentialekvationer. Detta synsätt var förhärskande ungefär fram till 1935. Man fann att det var stora problem förenade med analys av system som beskrivs av differentialekvationer av hög ordning (the curse of dimensionality). Under tidsperioden 1935–1945 tog man intryck av utvecklingen inom elektrotekniken och började tillämpa transformmetoder. Härigenom undvek man bl. a. den svårighet som var förenad med höga ordningstal. Vi citerar Bode: "As a matter of idle curiosity, I once counted to find what the order of the set of equations in an amplifier I had just designed, would have been, if I had worked with the differential equations directly. It turned out to be 55." Då reglertekniken konsoliderades under åren 1940–1950 användes nästan uteslutande externa beskrivningar och transformteori, något som återspeglar sig i dagens läroböcker. Under åren efter 1955 har man emellertid åter börjat använda interna beskrivningar och differentialekvationsteori på ett mycket fruktbart sätt ("The State Space Approach"). Det har bl. a. visat sig att det är det lämpliga tillvägagångssättet för att bygga upp en systematisk teori. Genom att kombinera idéer från mekanik och elektroteknik har man inom reglertekniken erhållit grunddragen till en teori för styrda dynamiska system. I detta kapitel skall vi redogöra för denna teori.

I avsnitt 2 införs begreppet dynamiskt system heuristiskt. Vi anger några olika sätt att beskriva dynamiska system matematiskt, diskuterar klassificering av dynamiska system och reglertekniska problemställningar i anslutning till dessa: identifiering, analys och syntes. Specialfallet linjära tidsinvarianta system behandlas i avsnitt 3. I de följande avsnitten diskuterar vi några elementära egenskaper hos de linjära systemen och olika sätt att representera sådana. Således införs viktfunktionen i avsnitt 7. I avsnitt 8 införs transformmetoder och vi definierar där överföringsfunktionen. Några olika sätt att representera överföringsfunktionen grafiskt, bl. a. singularitetsdiagram, Nyquist–Bodediagram, anges i avsnitt 9. Vidare införs begreppet minimum-fas-system. Grafisk representation av dynamiska system med blockschema och signalflödesdiagram behandlas i avsnitt 10. Avsnitt 11 ger en rudimentär

stabilitetsteori för linjära tidsinvarianta system. Den behövs bl. a. i kap. 4. En mer fullständig stabilitetsteori presenteras sedan i kap. 5. I avsnitt 12 ger vi slutligen en orientering om den abstrakta teorin för dynamiska system. Här ges bl. a. en formell definition av begreppet tillstånd. Trots att vi i detta kapitel ger ett antal olika sätt att karakterisera dynamiska system, kan vi ej göra någon jämförelse mellan de olika karakteriseringarna då detta kräver vidare teori. En sådan jämförelse presenteras i stället i kap. 3.

2. Karakterisering av dynamiska system

Begreppet tillstånd

De grundläggande idéerna för att karakterisera dynamiska system kan återföras på relationen orsak–verkan eller den s. k. kausalitetsprincipen. Tankegångarna går tillbaka till Newtons mekanik. Enligt denna kan rörelsen hos ett system av partiklar entydigt bestämmas för all framtid, $t \geqslant t_0$, om partiklarnas lägen och hastigheter är kända vid tiden t_0. Det är vidare irrelevant för partikelsystemets framtida rörelse hur partiklarna uppnått sina hastigheter och lägen. De krafter som påverkar partikelsystemet i framtiden har ej heller någon inverkan på partiklarnas rörelser i nuläget. För partikelsystemet definieras systemets tillstånd som samtliga partiklars positioner och hastigheter. Vi kan intuitivt säga att tillståndet är den minsta möjliga information om systemets förhistoria, som behövs för att exakt förutsäga systemets framtida rörelse. Vi skall nu närmare utveckla dessa idéer heuristiskt. I avsnitt 12 skall vi återkomma med formella definitioner.

Betrakta ett fysikaliskt system. Anta att systemets tillstånd vid tidpunkten t kan beskrivas med en ändlig uppsättning storheter $x_1(t)$, $x_2(t)$, ..., $x_n(t)$. Dessa storheter skall vi kalla *tillståndsvariabler* och vi skall ofta betrakta dem som komponenter av en vektor $x(t)$, *tillståndsvektorn*. I allmänhet kommer tillståndsvektorn att vara mångdimensionell, t. ex. vid studium av mekaniska system är tillståndsvariablerna positioner och hastigheter, vid passiva elektriska nätverk kan tillståndsvariablerna vara laddningar på kondensatorer resp. flöden i induktanser, vid kemiska system kan tillståndsvariablerna vara koncentrationer av olika konstituenter i reaktionskärl och behållare. I vissa stokastiska problem kan tillståndsvariablerna även vara sannolikheten att en process är i ett visst tillstånd. I reglertekniken skall vi studera reglerade eller styrda system. Dessa karakteriseras förutom av tillståndsvariablerna av *styrvariablerna* eller *instorheterna* $u_1(t)$, $u_2(t)$, ..., $u_r(t)$ samt *mätvariablerna, utstorheterna* eller *observablerna* $y_1(t)$, $y_2(t)$, ..., $y_p(t)$. Styrvariablerna är de variabler med vars hjälp vi kan påverka systemet, och mätvariablerna är de variabler som kan observeras. Med insignaler och utsignaler avses tidsfunktionerna $t \mapsto u(t)$ $t \mapsto y(t)$.

Differentialekvationsmodeller

När vi skall beskriva ett dynamiskt system ställs vi omedelbart inför följande problem: Hur beror systemets tillstånd av tiden och av styrvariablerna? Hur beror observablerna av systemets tillstånd och av styrvariablerna? Vi skall anta att tillståndsvariablerna är kontinuerligt deriverbara funktioner av tiden. Tillståndsförändringen kan då anges med hjälp av tillståndsvektorns tidsderivata dx/dt. Då x är systemets tillstånd och systemets framtida utveckling är oberoende av hur det aktuella tillståndet uppnåtts, måste dx/dt således vara en funktion av det aktuella tillståndet $x(t)$ och styrvariabelns aktuella värde $u(t)$. Systemet kan alltså beskrivas med hjälp av en ordinär differentialekvation.

$$\frac{dx_i(t)}{dt} = f_i(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), \ldots, u_r(t), t), \quad i = 1, 2, \ldots, n$$

$$x_i(0) = a_i \tag{2.1}$$

där a_i, $i = 1, \ldots, n$ är systemets initialtillstånd. Vidare antas att observablerna är funktioner av systemets tillstånd och styrvariablernas värden, dvs

$$y_i(t) = g_i(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), \ldots, u_r(t), t), \quad i = 1, 2, \ldots, p \tag{2.2}$$

För att minska skrivarbetet är det ofta bekvämt att införa vektorbeteckningar och vi skriver då ekvationerna (2.1) och (2.2) på följande sätt

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

$$x(0) = a$$

$$y(t) = g(x(t), u(t), t) \tag{2.3}$$

där vektorerna $x(t)$, $u(t)$, $y(t)$ uppfattas som kolonnvektorer.

En ekvation av typen (2.1) brukar benämnas *styrd differentialekvation* för att understryka att lösningarna kan påverkas genom val av styrvariablerna $u(t)$. Observera att även om vi primärt är intresserade av dynamiska system så omfattar ekvation (2.3) även statiska system. För sådana har vi $x(t) =$ konstant och systemet kan således beskrivas enbart med hjälp av ekvation (2.2).

EXEMPEL 2.1. Endimensionell partikel eller dubbelintegrator

Betrakta en partikel med enhetsmassa som rör sig längs en koordinataxel. Låt styrvariabeln u vara den kraft som påverkar partikeln och låt utsignalen eller observabeln y vara partikelns position. Rörelseekvationen ger

$$\frac{d^2y}{dt^2} = u \tag{2.4}$$

I detta fall vet vi från mekaniken att systemets tillstånd kan anges med två storheter. Vi väljer dessa att vara partikelns position x_1 och hastighet x_2. Systemekvationen kan då skrivas

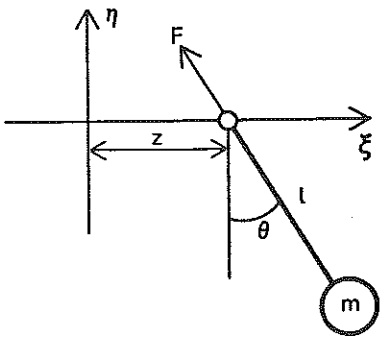


Fig. 2.1. Matematisk pendel.

$$\frac{dx_1}{dt} = x_2$$

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

$$y = x_1 \tag{2.5}$$

Vi observerar även att utsignalen $y(t)$ är insignalens dubbelintegral. Ekvation (2.4) beskriver således ett dynamiskt system där utsignalen är en dubbelintegration av insignalen.

EXEMPEL 2.2. Pendeln eller oscillatorn

Betrakta en matematisk pendel (fig. 2.1) med massan m och längden l i ett homogent gravitationsfält med tyngdaccelerationen g. Låt pendelns upphängningspunkt vara fritt rörlig längs en rät linje som är vinkelrät mot tyngdaccelerationen. Låt styrvariabeln vara pivotpunktens acceleration och låt observabeln vara pendelns utslagsvinkel $y = \theta$.

Om z får beteckna pivotpunktens ξ-koordinat, finner vi att massans koordinater blir

$$\xi = z + l \sin\theta$$

$$\eta = -l \cos\theta$$

och vi får

$$\ddot{\xi} = \ddot{z} + l\ddot{\theta} \cos\theta - l\dot{\theta}^2 \sin\theta$$

$$\ddot{\eta} = l\ddot{\theta} \sin\theta + l\dot{\theta}^2 \cos\theta$$

Om F får beteckna den kraft som verkar på pendeln i pivotpunkten erhålls följande rörelseekvationer

$$m\ddot{\xi} = -F \sin\theta$$

$$m\ddot{\eta} = F \cos\theta - mg$$

Elimineras F ur dessa ekvationer erhålls

$$l\ddot{\theta}+g\sin\theta+\ddot{z}\cos\theta=0 \tag{2.6}$$

Införs tillståndsvariablerna

$$x_1=\theta$$

$$x_2=\dot{\theta}$$

styrvariabeln $u=-\ddot{z}$ och observabeln $y=\theta$ erhålls

$$\frac{dx_1}{dt}=x_2$$

$$\frac{dx_2}{dt}=-\frac{g}{l}\sin x_1+\frac{u}{l}\cos x_1$$

$$y=x_1 \tag{2.7}$$

Vi antar vidare att enheterna väljs så att ekvationens alla koefficienter är ett och vi finner då

$$\frac{dx_1}{dt}=x_2$$

$$\frac{dx_2}{dt}=-\sin x_1+u\cos x_1$$

$$y=x_1$$

Innan vi fortsätter skall vi göra en utvikning för att påtala några begränsningar hos den uppställda systemmodellen. I många fall är det ej rimligt att förutsätta att antalet tillståndsvariabler är ändligt. T. ex. vid studium av värmeprocesser är det ofta nödvändigt att beskriva processen med hjälp av en partiell differentialekvation. Man kan då erhålla en beskrivning analog med ekvation (2.3), men tillståndsvektorn kommer att ha (uppräkneligt) oändligt många komponenter svarande t. ex. mot koefficienterna i en serieutveckling efter ortogonala funktioner. Flera system som man har intresse av att studera, har även tidsfördröjningar, orsakade t. ex. av att material transporteras på band eller i ledningar. I sådana fall erhålls primärt ekvationer av typen

$$\frac{dx}{dt}=f(x(t),x(t-T)) \tag{2.8}$$

eller

$$\frac{dx}{dt}=f(x(t),x(t-t_1),\ldots,x(t-t_k),u(t),u(t-t_1),\ldots,u(t-t_l),t) \tag{2.9}$$

s. k. differential-differensekvationer. Observera att $x(t)$ i ekvation (2.8) ej kan vara systemets tillstånd. För att lösa ekvation (2.8) i intervallet (t, t_1) räcker

det ej med att känna $x(t)$ utan det är uppenbarligen nödvändigt att känna värdena av x i hela intervallet $(t-T, t)$. I detta fall är systemets tillstånd således en funktion definierad på ett ändligt intervall på reella linjen.

I andra fall kan det vara nödvändigt att tillgripa ännu mer komplicerade ekvationer för att med tillräcklig noggrannhet beskriva ett system. Som exempel kan nämnas differential-integralekvationer

$$\frac{dx}{dt} = f\left(x(t), \int_{-\infty}^{t} x(s)\, dF_1(s,t), \ldots, \int_{-\infty}^{t} x(s)\, dF_k(s,t), t\right) \tag{2.10}$$

och integralekvationer

$$x(t) = f\left(\int_{-\infty}^{t} x(s)\, dF_1(s,t), \ldots, \int_{-\infty}^{t} x(s)\, dF_k(s,t), t\right) \tag{2.11}$$

Ekvationer av detta slag förekommer vid beskrivning av stokastiska system, vissa biologiska system, system med vågutbredning etc. För de system som bestäms av ekvationerna (2.10) och (2.11) är $x(t)$ ej heller systemens tillstånd, ty för att beskriva systemens framtida utveckling räcker det ej med att känna $x(t)$ utan funktionen x måste vara känd på hela intervallet $(-\infty, t)$. Vi kommer i fortsättningen ej att behandla sådana komplicerade system utan avrundar utvikningen med att påpeka att de komplicerade systemen ofta kan skrivas på formen (2.3) genom att man inför tillräckligt (oändligt) många tillståndsvariabler.

System som beskrivs av differensekvationer

$$x(t+\Delta) = F(x(t), u(t), t) \quad t = \Delta, 2\Delta, 3\Delta, \ldots$$

$$x(0) = a$$

$$y(t) = H(x(t), u(t), t)$$

kallas inom reglertekniken *tidsdiskreta* eller *samplade system*. Samplade system tjänstgör som approximationer till kontinuerliga system och uppträder dessutom naturligt när datorer används. De har genom datorernas tillkomst fått mycket stor betydelse.

Vi skall i denna bok med få undantag endast behandla system som kan beskrivas med ordinära differentialekvationer av typ (2.3).

Klassificering av dynamiska system

Vi skall nu sammanfattningsvis ange några allmänt förekommande klassificeringar av dynamiska system. Klassificeringar kan göras från många synpunkter, t. ex. antal insignaler och utsignaler, antal tillståndsvariabler eller komplexiteten hos systemekvationerna.

Ett system är *enkelt* om det endast har en insignal och en utsignal. Om

systemet har flera insignaler eller flera utsignaler kallas det *flervariabelt*. Ett system är av *n:te ordningen* om tillståndsvektorn har n komponenter. Ett system är *linjärt* om funktionerna $f(x, u, t)$ och $g(x, u, t)$ är linjära i x och u. Ett system är *tidsinvariant* eller *stationärt* om $f(x, u, t)$ och $g(x, u, t)$ ej beror explicit av tiden t.

Den endimensionella partikeln i exempel 2.1 är ett enkelt, linjärt och tidsinvariant system av andra ordningen. Pendeln i exempel 2.2 är ett enkelt, olinjärt och tidsinvariant system av andra ordningen.

I denna bok skall vi med några få undantag studera endast linjära tidsinvarianta system. För linjära system finns en systematisk teori, som är jämförelsevis lättillgänglig. Den linjära teorin används i praktiken vanligen på följande sätt. En linjär modell bestäms genom att de olinjära systemekvationerna approximeras med en linjär modell i närheten av en jämviktslösning, s. k. *linjärisering*. En styrlag bestäms ur den linjära modellen. Denna styrlag används sedan till det verkliga olinjära systemet. Det kan ifrågasättas om detta förfarande är korrekt. Ett partiellt svar på denna fråga ges av en viktig sats av Lyapunov och Poincaré, som säger att om det linjäriserade systemet är stabilt med den erhållna styrlagen så är med svaga restriktioner även det olinjära systemet stabilt.

3. Linjära tidsinvarianta system

Med ett linjärt tidsinvariant system $\mathcal{S}$ avses ett system där tillståndsvektorns tidsderivata dx/dt är en linjär funktion av tillståndet x och styrsignalen u och utsignalen y en linjär funktion av tillståndet x och styrvariabeln u. Ett linjärt tidsinvariant system $\mathcal{S} = \mathcal{S}(A, B, C, D)$ kan således karakteriseras med ekvationerna

$$\frac{dx}{dt} = Ax + Bu$$
$$y = Cx + Du \tag{3.1}$$

där x är en n-dimensionell tillståndsvektor, u är en r-dimensionell vektor av insignaler, y är en p-dimensionell vektor av utsignaler, A, B, C och D konstanta matriser, dvs.

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

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nr} \end{bmatrix}$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & & & \\ c_{p1} & c_{p2} & \cdots & c_{pn} \end{bmatrix} \qquad D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1r} \\ d_{21} & d_{22} & \cdots & d_{2r} \\ \vdots & & & \\ d_{p1} & d_{p2} & \cdots & d_{pr} \end{bmatrix}$$

EXEMPEL 3.1

Den endimensionella partikeln i exempel (2.1) är ett linjärt tidsinvariant system. Ekvation (2.5) kan skrivas på formen (3.1). Vi finner

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]\, x$$

Linjärisering

De linjära ekvationerna erhålls ofta på följande sätt. Ur fysikaliska grundekvationer som uttrycker balans av massa, energi och rörelsemängd uppställs de differentialekvationer som beskriver systemet. Tillståndsvariabler införs och ekvationerna skrivs på normalformen för tidsinvarianta system, dvs.

$$\frac{dx}{dt} = f(x, u)$$

$$y = g(x, u) \tag{3.2}$$

där vektorerna x, u och y antas ha dimensionerna n, r och p. Jämför med ekvationerna (2.1) och (2.2). Med en *stationär lösning* eller en arbetspunkt menas konstanta vektorer u^0 och x^0 sådana att

$$f(x^0, u^0) = 0 \tag{3.3}$$

Observera att det ej behöver finnas stationära lösningar. Det kan också finnas många stationära lösningar.

Med linjärisering menas att de olinjära ekvationerna (3.2) approximeras med linjära ekvationer i närheten av en arbetspunkt. Linjärisering är mycket användbart inom reglertekniken. De linjära ekvationerna är mycket lättare att analysera än motsvarande olinjära ekvationer. Eftersom meningen med reglersystemen är att se till att processvariablerna ligger så nära vissa föreskrivna värden som möjligt så är de linjäriserade modellerna ofta mycket goda approximationer.

För att linjärisera ekvationerna utgår vi således från en stationär lösning (x^0, u^0). Inför avvikelserna

$$\Delta x = x - x^0 \qquad \Delta u = u - u^0 \qquad \Delta y = y - y^0 = y - g(x^0, u^0)$$

Om funktionerna f och g är två gånger kontinuerliga deriverbara kan de utvecklas i Taylorserier och vi finner

$$\frac{d\Delta x}{dt} = f(x^0+\Delta x, u^0+\Delta u) = f(x^0, u^0) + f_x(x^0, u^0)\Delta x + f_u(x^0, u^0)\Delta u + O((\Delta x)^2, (\Delta u)^2) \tag{3.4}$$

$$\Delta y = y - g(x^0, u^0) = g_x(x^0, u^0)\Delta x + g_u(x^0, u^0)\Delta u + O((\Delta x)^2, (\Delta u)^2)$$

där

$$f_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}, \quad f_u = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} & \cdots & \frac{\partial f_1}{\partial u_p} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} & \cdots & \frac{\partial f_2}{\partial u_p} \\ \vdots & & & \\ \frac{\partial f_n}{\partial u_1} & \frac{\partial f_n}{\partial u_2} & \cdots & \frac{\partial f_n}{\partial u_p} \end{bmatrix}$$

och g_x och g_u definieras analogt. Observera att $f(x^0, u^0)$ är noll eftersom (x^0, u^0) är en jämviktslösning. Om termerna av högre ordning i (3.4) försummas finner vi således att (3.4) är ett linjärt system $S(A, B, C, D)$ om vi identifierar

$$A = f_x(x^0, u^0), \quad B = f_u(x^0, u^0), \quad C = g_x(x^0, u^0), \quad D = g_u(x^0, u^0)$$

EXEMPEL 3.2

Pendeln i exempel 2.2 är ett olinjärt system. Om vi antar att systemet utför små rörelser kring ett jämviktsläge så kan vi emellertid linjärisera ekvationerna. Om vi linjäriserar kring jämviktsläget $u = x_1 = 0$ erhålls följande system

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]\, x$$

Observera dock att koefficienterna i den linjära approximationen beror av den punkt kring vilken vi linjäriserar. Om vi i stället linjäriserar kring $\theta = \theta_0$ och inför tillståndsvariablerna

$$x_1 = \theta - \theta_0$$

$$x_2 = \dot{\theta}$$

erhålls

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ -1/\cos\theta_0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ \cos\theta_0 \end{bmatrix} (u - \tan\theta_0)$$

$$y = [1 \quad 0]\, x$$

Väljs $\theta = \pi$ erhålls t. ex.

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u$$

$$y = [1 \quad 0]\, x$$

Detta system brukar kallas den *inverterade pendeln.*

I exempel 2.2 var styrvariabeln pivotpunktens acceleration. Om vi i stället som styrvariabel väljer pivotpunktens koordinat lyder systemekvationen (2.6)

$$l\ddot{\theta} + g \sin \theta + \ddot{u} \cos \theta = 0$$

Om vi antar $g = l = 1$, linjäriserar denna ekvation kring jämviktsläget $\theta = 0$ och inför tillståndsvariablerna

$$x_1 = \theta + u$$

$$x_2 = \dot{\theta} + \dot{u}$$

finner vi att systemet kan beskrivas på följande sätt

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]\, x - u$$

Vilken av systemrepresentationerna som är relevant i ett praktiskt fall beror naturligtvis av förutsättningarna. Det fall då pivotpunktens acceleration är insignal förutsätter att styrsignalen levereras via en kraftgivare som styr pivotpunktens acceleration. Det fall då pivotpunktens position är insignal förutsätter att pivotpunktens läge styrs av ett positionsservo.

4. Lösning av systemekvationen

Ekvation (3.1) kan i allmänhet ej lösas explicit med hjälp av elementära funktioner. Lösningen kan däremot skrivas på en kompakt form som är mycket användbar för uppskattningar och som dessutom är väl lämpad för numeriska beräkningar. Denna form skall härledas i detta avsnitt.

Då ekvation (3.1) är linjär vet vi, att lösningen kan skrivas som en summa av en partikulärlösning och den allmänna lösningen till den homogena ekvationen. Vi skall först betrakta den homogena ekvationen

$$\frac{dx}{dt} = Ax \tag{4.1}$$

med initialvillkoret

$$x(0) = a \tag{4.2}$$

Lösningen till ekvation (4.1) med initialvillkoret (4.2) kan skrivas

$$x(t) = e^{At}a \tag{4.3}$$

där exponentialfunktionen av en matris definieras med hjälp av följande serieutveckling

$$e^B = \exp B = I + B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \ldots + \frac{1}{n!}B^n + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!}B^n \tag{4.4}$$

Denna serie konvergerar i norm för alla matriser B, ty varje term i serien domineras till sitt belopp av $\|B\|^n/n!$, där $\|B\|$ betecknar en tillordnad *matrisnorm*. (Se t. ex. *Formelsamling i Numerisk Analys*, Studentlitteratur, Lund 1967, s. 20.) Enligt satsen om absolut konvergens är då serien (4.4) konvergent. Serieutvecklingen (4.4) är också väl lämpad för numerisk beräkning av exp B. För att visa att (4.3) verkligen är en lösning till differentialekvationen (4.1) begagnar vi serieutvecklingen

$$e^{At} = I + At + \frac{1}{2!}A^2t^2 + \ldots + \frac{1}{n!}A^nt^n + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!}A^nt^n$$

Derivering med avseende på t ger

$$\frac{d}{dt}(e^{At}) = A + A^2t + \frac{1}{2!}A^3t^2 + \ldots + \frac{1}{(n-1)!}A^nt^{n-1} + \ldots$$

$$= A\left\{I + At + \frac{1}{2!}A^2t^2 + \ldots + \frac{1}{(n-1)!}A^{n-1}t^{n-1} + \ldots\right\} = Ae^{At}$$

Den deriverade serien är likformigt konvergent och vi finner även att funktionen exp At satisfierar differentialekvationen

$$\frac{dX}{dt} = AX \tag{4.5}$$

med initialvillkoret

$$X(0) = I \tag{4.6}$$

där $X(t)$ är en $n \times n$-matris. Genom att multiplicera från höger med a finner vi således att funktionen

$$x(t) = e^{At}a$$

satisfierar den homogena differentialekvationen (4.1) med initialvillkoret (4.2).

Betrakta ett system vars rörelse kan beskrivas med hjälp av differentialekvationen (4.1). Ekvation (4.3) innebär att tillståndsvektorns värde vid tidpunkten t är en linjär transformation av initialtillståndet a. Att integrera (4.1) över ett tidsintervall $(0, t)$ är således ekvivalent med att multiplicera initialtillståndet med exp At från vänster.

Innan vi löser den homogena ekvationen skall vi visa att funktionen exp At har egenskapen

$$e^{A(s+t)} = e^{As}e^{At} \tag{4.7}$$

Om vi använder serieutvecklingen (4.4) och det faktum att summationsordningen får bytas i en absolutkonvergent serie erhålls

$$e^{As}\,e^{At} = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k s^k\right)\left(\sum_{l=0}^{\infty} \frac{1}{l!} A^l t^l\right) = \sum_{n=0}^{\infty} A^n \left(\sum_{k=0}^{n} \frac{1}{k!} \frac{1}{(n-k)!} s^k t^{n-k}\right)$$

$$= \sum_{n=0}^{\infty} A^n \frac{(s+t)^n}{n!} = e^{A(s+t)}$$

ty

$$(s+t)^n = \sum_{k=0}^{n} \binom{n}{k} s^k t^{n-k} = \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} s^k t^{n-k}$$

och påståendet är bevisat.

Insätts $s = -t$ i ekvation (4.7) erhålls

$$e^{At}e^{-At} = I \tag{4.7'}$$

dvs. matrisen exp At är aldrig singulär och dess invers är given av exp $(-At)$. Detta är analogin till den skalära exponentialfunktionens egenskap att alltid vara positiv.

Då vi erhållit en lösning till den homogena ekvationen kan vi nu lösa den inhomogena ekvationen t. ex. med hjälp av Lagranges metod med variation av konstanten. Vi ansätter en lösning till (3.1) på formen

$$x(t) = e^{At}c(t)$$

Insätts denna i ekvation (3.1) erhålls

$$\frac{dx}{dt} = Ae^{At}c(t) + e^{At}\frac{dc(t)}{dt} = Ax + Bu$$

Men enligt (4.7') gäller att $(\exp At)^{-1} = \exp -At$. Vi finner då

$$\frac{dc(t)}{dt} = [e^{At}]^{-1}Bu = e^{-At}Bu$$

dvs.

$$c(t) = c(0) + \int_0^t e^{-As}Bu(s)\,ds$$

Lösningen till systemekvation (3.1) kan således skrivas

$$x(t) = e^{At}a + \int_0^t e^{A(t-s)}Bu(s)\,ds \tag{4.8}$$

Vi finner således att lösningen är en summa av två termer. Den första beror på initialtillståndet a, och den andra beror av insignalen $u(t)$.

Vi skall ge några exempel där systemekvationerna kan lösas explicit.

EXEMPEL 4.1

Betrakta den endimensionella partikeln som behandlades i exempel 3.1. Den kan beskrivas med ekvationen

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]\, x$$

Vi finner

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad A^2 = \ldots \; A^n = \ldots = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Definitionen (4.4) på exponentialmatrisen ger direkt

$$\exp At = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

Vidare gäller

$$(\exp At)\, B = \begin{bmatrix} t \\ 1 \end{bmatrix}$$

Ekvation (4.8) ger således

$$x(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} x(0) + \int_0^t \begin{bmatrix} t-s \\ 1 \end{bmatrix} u(s)\, ds$$

EXEMPEL 4.2

Betrakta den linjäriserade pendeln i exempel 3.2. För denna gäller

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]\, x$$

Vi finner

$$A^2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = -I$$

$$A^3 = -A$$

$$A^4 = I$$

$$\vdots$$

$$A^{2n} = (-1)^n I$$

$$A^{2n+1} = (-1)^n A$$

Det följer nu av ekvation (4.4) att

$$[e^{At}]_{11} = [e^{At}]_{22} = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{(2n)!} = \cos t$$

$$[e^{At}]_{12} = -[e^{At}]_{21} = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!} = \sin t$$

Sammanfattningsvis erhålls

$$\exp At = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$$

och

$$x(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} x(0) + \int_0^t \begin{bmatrix} \sin (t-s) \\ \cos (t-s) \end{bmatrix} u(s)\, ds$$

5. En funktionalekvation för exp At — kausalitetsprincipen

Den skalära exponentialfunktionen $f(t) = \exp at$ satisfierar funktionalekvationen

$$f(t+s) = f(t)\, f(s) \tag{5.1}$$

Vi visade i föregående avsnitt att matrisfunktionen $f(t) = \exp At$, där A är en $n \times n$-matris, satisfierar samma ekvation, dvs. att

$$e^{A(s+t)} = e^{As} e^{At} \tag{5.2}$$

Även omvändningen gäller. Funktionen exp At är således den enda kontinuerliga matrisfunktion som satisfierar (5.1). Jämför exempel 12.1.

Den härledning av funktionalekvationen (5.2) som gavs i föregående avsnitt är en verifikation. För att få en inblick i vad relationen innebär skall vi ge en fysikalisk tolkning till ekvation (5.2). Betrakta differentialekvationen

$$\frac{dX}{dt} = AX \qquad X(0) = I \tag{5.3}$$

Vi skall nu integrera denna ekvation över intervallet $(0, t+s)$ på två olika sätt. Om vi direkt integrerar över hela intervallet finner vi lösningen

$$X(t+s) = e^{A(t+s)} = e^{A(s+t)} \tag{5.4}$$

Låt oss nu först integrera över intervallet $(0, t)$. Vi finner då

$$X(t) = e^{At}$$

Om vi nu med $X(t)$ som initialvärde integrerar över $(t, t+s)$ finner vi

$$X(t+s) = e^{As}e^{At} \tag{5.5}$$

Då differentialekvationen (5.3) har en entydig lösning finner vi således att högerleden i ekvationerna (5.4) och (5.5) är lika, dvs.

$$e^{A(t+s)} = e^{A(s+t)} = e^{As}e^{At}$$

Om vi betraktar systemets rörelse över tidsintervallet $(t, t+s)$ finner vi således att om $X(t)$ är känd så kan denna rörelse exakt anges. Observera att då $X(t)$ är känd är det irrelevant för den framtida rörelsen hur $X(t)$ uppnåddes. Vi kan alltså säga att $X(t)$ är den minsta information om systemets förhistoria, som krävs för att förutsäga den framtida rörelsen. Enligt resonemanget i början av avsnitt 2 är $X(t)$ då systemets tillstånd, och funktionalekvationen (5.1) är således en matematisk formulering av *kausalitetsprincipen*. Generaliseringar av ekvation (5.1) ligger även till grund för abstrakta definitioner av kausalitetsbegreppet. Jämför avsnitt 12.

6. Byte av koordinater i tillståndsrummet

Det gäller för ett dynamiskt system att insignalen u och utsignalen y är entydiga fysikaliska storheter. Tillståndsvariabeln x är däremot ej entydig utan beror av det koordinatsystem vi valt för att representera systemet. För att inse detta inför vi nya koordinater z med hjälp av den linjära transformationen T

$$z = Tx \tag{6.1}$$

Vi finner då att ekvation (3.1) kan skrivas

$$\begin{aligned} \frac{dz}{dt} &= TAT^{-1}z + TBu = \tilde{A}z + \tilde{B}u \\ y &= CT^{-1}z + Du = \tilde{C}z + Du \end{aligned} \tag{6.2}$$

Matriserna A, B och C beror således på det koordinatsystem vi valt för att representera systemets tillstånd. Man kan då fråga sig om det finns koordinatsystem så att systemekvationerna får en speciellt enkel form. I det följande skall vi ange sådana koordinatsystem och motsvarande standardformer för systemekvationerna. Vi skall i hela avsnittet förutsätta att systemet har en insignal och en utsignal.

Diagonalformen

Studera först det fall då matrisen A har olika egenvärden. Det finns då en matris T sådan att

$$TAT^{-1} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

De transformerade tillståndsvariablerna Z_i kallas *moder*. Om systemet endast har en insignal och en utsignal kan systemekvationerna skrivas

$$\frac{dz_1}{dt} = \lambda_1 z_1 + b_1 u$$

$$\frac{dz_2}{dt} = \lambda_2 z_2 + b_2 u$$

$$\vdots$$

$$\frac{dz_n}{dt} = \lambda_n z_n + b_n u$$

$$y = c_1 z_1 + c_2 z_2 + \dots + c_n z_n + du \tag{6.3}$$

Ekvation (6.3) kallas *diagonalformen* för ett linjärt system med en insignal och en utsignal. Om vi endast vill ange sambandet mellan insignal och utsignal så finns i denna representation redundanta parametrar. Vi kan t. ex. dividera b_i med en faktor och multiplicera c_i med samma faktor och fortfarande få samma insignal–utsignalrelation. För enkla system av *n:te* ordningen är insignal–utsignalrelationen given av $2n+1$ parametrar. I ovanstående exempel kan vi t. ex. fixera koefficienterna $b_1 = b_2 \dots = b_n = 1$.

Diagonalformen existerar endast i det fall då matrisen A kan diagonaliseras. Detta går t. ex. alltid om A har skilda egenvärden.

Den karakteristiska ekvationen är invariant vid en transformation TAT^{-1}. Det gäller

$$\begin{aligned} \det[\lambda I - TAT^{-1}] &= \det T[\lambda T^{-1}T - A]T^{-1} = \det T[\lambda I - A]T^{-1} \\ &= \det T \det T^{-1} \det[\lambda I - A] = \det[\lambda I - A] \end{aligned}$$

Den karakteristiska ekvationens koefficienter och funktioner av dessa är också de enda invarianterna vid transformationen TAT^{-1}.

Denna invariansegenskap spelar en viktig roll i de transformationer som behandlas i följande avsnitt.

Kanonisk form 1 (Den observerbara kanoniska formen)

Betrakta ett system $S(A, B, C, D)$ med en insignal och en utsignal. Anta att matrisen A har den karakteristiska ekvationen

$$\det[\lambda I - A] = \lambda^n + a_1\lambda^{n-1} + \dots + a_n = 0 \tag{6.4}$$

och att matrisen

$$F = \begin{bmatrix} CA^{n-1} \\ CA^{n-2} \\ \vdots \\ C \end{bmatrix} \tag{6.5}$$

ej är singulär. Vi skall senare i kap. 3 ge en fysikalisk tolkning till detta villkor. Observera att C är en vektor då systemet endast har en insignal och en utsignal. Låt f beteckna en godtycklig kolonnvektor sådan att matrisen

$$T^{-1} = [A^{n-1}f \; A^{n-2}f \; \dots \; f] \tag{6.6}$$

ej är singulär. Då gäller

$$T[A^{n-1}f \; A^{n-2}f \; \dots \; f] = I \tag{6.7}$$

Betrakta systemet $\mathcal{S}(A, B, C, D)$ och inför nya koordinater i tillståndsrummet genom transformationen $z = Tx$. Vi finner då

$$\tilde{A} = TAT^{-1} = T[A^n f \; A^{n-1}f \; \dots \; Af] \tag{6.8}$$

Men enligt (6.7) gäller

$$TA^{n-1}f = e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad TA^{n-2}f = e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; \dots, \; Tf = e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Enligt Cayley–Hamiltons teorem (varje kvadratisk matris satisfierar sin egen karakteristiska ekvation) gäller

$$A^n = -a_1 A^{n-1} - a_2 A^{n-2} - \dots - a_n I$$

och vi finner således

$$TA^n f = -a_1 TA^{n-1}f - a_2 TA^{n-2}f - \dots - a_n Tf = -a_1 e_1 - a_2 e_2 - \dots - a_n e_n$$

Ekvation (6.8) ger då

$$\tilde{A} = TAT^{-1} = \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ -a_{n-1} & 0 & 0 & \dots & 1 \\ -a_n & 0 & 0 & \dots & 0 \end{bmatrix} \tag{6.9}$$

Den transformerade »A-matrisen» har alltså den karakteristiska ekvationens koefficienter som element. En matris av detta slag kallas i engelsk litteratur *companion matrix*.

Vi finner således att för varje f, sådan att T^{-1} definierad av (6.6) är reguljär,

erhålls en transformationsmatris som transformerar A på formen (6.9). Vi väljer nu f sådan att

$$CT^{-1}=[CA^{n-1}f\ CA^{n-2}f\ \dots\ Cf]=[1\ \ 0\ \dots\ 0]=e_1^T$$

dvs.

$$f=\begin{bmatrix}CA^{n-1}\\CA^{n-2}\\\vdots\\C\end{bmatrix}^{-1}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}=F^{-1}e_1 \tag{6.10}$$

Detta är alltid möjligt, ty vi har förutsatt att F är reguljär. Det återstår nu att visa, att om det f som ges av (6.10) insätts i (6.6), erhålls en reguljär matris.

Vi visar först följande identitet

$$A^rF^{-1}=F^{-1}\bar{A}^r \tag{6.11}$$

där $\bar{A}=\bar{A}^T$. Matrisen $\bar{A}$ definierades av ekvation (6.9).

Vi erhåller följande resultat

$$FA^r=\begin{bmatrix}CA^{n-1}\\CA^{n-2}\\CA^{n-3}\\\vdots\\C\end{bmatrix}A^r=\begin{bmatrix}CA^n\\CA^{n-1}\\CA^{n-2}\\\vdots\\CA\end{bmatrix}A^{r-1}=\begin{bmatrix}-a_1 & -a_2 & \dots & -a_{n-1} & -a_n\\1 & 0 & \dots & 0 & 0\\0 & 1 & \dots & 0 & 0\\\vdots & & & & \\0 & 0 & \dots & 1 & 0\end{bmatrix}\begin{bmatrix}CA^{n-1}\\CA^{n-2}\\CA^{n-3}\\\vdots\\C\end{bmatrix}A^{r-1}$$

Den första raden i matrisekvationen erhålls ur Cayley–Hamiltons ekvation. De övriga raderna är rena identiteter. Vi finner således

$$FA^r=\bar{A}FA^{r-1}=\bar{A}(\bar{A}FA^{r-2})=\bar{A}^rF$$

Multiplikation av ovanstående formel med F^{-1} från höger och vänster ger nu (6.11).

Genom att utnyttja ekvation (6.10) kan transformationsmatrisen (6.6) skrivas

$$T^{-1}=[A^{n-1}F^{-1}e_1\ \ A^{n-2}F^{-1}e_1\ \dots\ F^{-1}e_1]$$

Identiteten (6.11) ger vidare

$$T^{-1}=F^{-1}[\bar{A}^{n-1}e_1\ \ \bar{A}^{n-2}e_1\ \dots\ e_1]=F^{-1}H$$

där

$$H=\begin{bmatrix}h_{n-1} & h_{n-2} & \dots & h_1 & 1\\h_{n-2} & h_{n-3} & \dots & 1 & 0\\\vdots & & & & \\h_1 & 1 & \dots & 0 & 0\\1 & 0 & \dots & 0 & 0\end{bmatrix} \tag{6.12}$$

och

$$h_1 = -a_1$$

$$h_r = -a_r - \sum_{i=1}^{r-1} a_i h_{r-i}, \quad r > 1 \tag{6.13}$$

Matrisen H har determinanten ett och är således reguljär. Matrisen T^{-1} är då reguljär om F är reguljär. Då matrisen H är triangulär kan den inverteras direkt. Vi finner efter enkla algebraiska räkningar

$$H^{-1} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & a_1 \\ \vdots & & & & & \\ 0 & 1 & a_1 & \dots & a_{n-3} & a_{n-2} \\ 1 & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} \end{bmatrix} \tag{6.14}$$

Vi erhåller vidare följande explicita uttryck för transformationsmatrisen

$$T = H^{-1}F = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & a_1 \\ \vdots & & & & \\ 0 & 1 & \dots & a_{n-3} & a_{n-2} \\ 1 & a_1 & \dots & a_{n-2} & a_{n-1} \end{bmatrix} \begin{bmatrix} CA^{n-1} \\ CA^{n-2} \\ \vdots \\ CA \\ C \end{bmatrix} = \begin{bmatrix} C \\ CA + a_1 C \\ \vdots \\ CA^{n-1} + a_1 CA^{n-2} + \dots + a_{n-1} C \end{bmatrix} \tag{6.15}$$

Om t_i får beteckna T:s i:e rad erhålls följande rekursiva ekvationer för att bestämma matrisen T.

$$t_1 = C$$

$$t_{i+1} = t_i A + a_i C \tag{6.16}$$

Vi sammanfattar resultatet som

SATS 6.1

Anta att matrisen

$$F = \begin{bmatrix} CA^{n-1} \\ CA^{n-2} \\ \vdots \\ C \end{bmatrix}$$

är reguljär. Genom byte av koordinater i tillståndsrummet med transformationen $z = Tx$ där

$$T = \begin{bmatrix} C \\ CA + a_1 C \\ \vdots \\ CA^{n-1} + a_1 CA^{n-2} + \dots + a_{n-1} C \end{bmatrix}$$

kan systemet representeras som

$$\frac{dz}{dt}=\begin{bmatrix}-a_1 & 1 & 0 & \dots & 0\\ -a_2 & 0 & 1 & \dots & 0\\ \vdots & & & & \\ -a_{n-1} & 0 & 0 & \dots & 1\\ -a_n & 0 & 0 & \dots & 0\end{bmatrix} z+\begin{bmatrix}CB\\ CAB+a_1CB\\ \vdots\\ CA^{n-2}B+a_1CA^{n-3}B+\dots+a_{n-2}CB\\ CA^{n-1}B+a_1CA^{n-2}B+\dots+a_{n-1}CB\end{bmatrix} u$$

$$y=[1\quad 0\quad 0\ \dots\ 0]\ z+Du$$

Kanonisk form 2 (Den styrbara kanoniska formen)

Vi skall nu konstruera en transformation som mycket liknar den nyss gjorda transformationen. Anta att systemet $\mathcal{S}(A, B, C, D)$ har en insignal och en utsignal samt att matrisen

$$G=[A^{n-1}B\ A^{n-2}B\ \dots\ B] \tag{6.17}$$

ej är singulär. Vi skall senare i kap. 3 ge en fysikalisk tolkning till detta villkor. Låt g beteckna en godtycklig radvektor sådan att matrisen

$$S=\begin{bmatrix}gA^{n-1}\\ gA^{n-2}\\ \vdots\\ g\end{bmatrix} \tag{6.18}$$

ej är singulär. Betrakta systemet $\mathcal{S}(A, B, C, D)$ och inför nya koordinater i tillståndsrummet genom transformationen $z=Sx$. Vi finner

$$SAS^{-1}=\begin{bmatrix}gA^{n}\\ gA^{n-1}\\ \vdots\\ gA\end{bmatrix} S^{-1} \tag{6.19}$$

Men då $SS^{-1}=I$ erhålls ur ekvation (6.18)

$$gA^{n-1}S^{-1}=[1\quad 0\ \dots\ 0]=v_1$$
$$gA^{n-2}S^{-1}=[0\quad 1\ \dots\ 0]=v_2$$
$$\vdots$$
$$gS^{-1}\qquad =[0\quad 0\ \dots\ 1]=v_n$$

Med användande av Cayley–Hamiltons teorem finner vi

$$SAS^{-1}=\begin{bmatrix}-a_1 & -a_2 & -a_3 & \dots & -a_{n-1} & -a_n\\ 1 & 0 & 0 & \dots & 0 & 0\\ 0 & 1 & 0 & \dots & 0 & 0\\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & 1 & 0\end{bmatrix}=\bar{A} \tag{6.20}$$

Den transformerade »A-matrisen» har karakteristiska ekvationens koefficienter som element. Även matrisen (6.20) kallas i engelskspråkig litteratur *companion matrix*. Vi kan fortfarande förfoga över g och väljer då denna vektor så att

$$SB = \begin{bmatrix} gA^{n-1}B \\ gA^{n-2}B \\ \vdots \\ gB \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

dvs.

$$g = [1 \quad 0 \ldots 0] \quad [A^{n-1}B \quad A^{n-2}B \ldots B]^{-1} = v_1 G^{-1} \tag{6.21}$$

Vi skall nu ge en explicit framställning av transformationsmatrisen S samtidigt som vi skall visa att S är reguljär under förutsättning att G är reguljär. Vi finner helt analogt med analysen av den observerbara kanoniska formen att

$$G^{-1}A^r = \tilde{A}^r G^{-1} \tag{6.22}$$

där $\tilde{A} = \tilde{A}^T$. Jämför ekvation (6.9).

Vi finner

$$A^r G = A^{r-1}[A^n B \; A^{n-1}B \ldots AB] = A^{r-1}[A^{n-1}B \; A^{n-2}B \ldots B] \begin{bmatrix} -a_1 & 1 & 0 \ldots 0 \\ -a_2 & 0 & 1 \ldots 0 \\ \vdots & & \\ -a_{n-1} & 0 & 0 \ldots 1 \\ -a_n & 0 & 0 \ldots 0 \end{bmatrix}$$

dvs.

$$A^r G = A^{r-1} G \tilde{A} = G \tilde{A}^r$$

Multiplikation med G^{-1} från höger och vänster ger nu (6.22). Utnyttjas (6.21) kan transformationsmatrisen S given av (6.18) skrivas

$$S = \begin{bmatrix} gA^{n-1} \\ gA^{n-2} \\ \vdots \\ g \end{bmatrix} = \begin{bmatrix} v_1 G^{-1} A^{n-1} \\ v_1 G^{-1} A^{n-2} \\ \vdots \\ v_1 G^{-1} \end{bmatrix} = \begin{bmatrix} v_1 \tilde{A}^{n-1} \\ v_1 \tilde{A}^{n-2} \\ \vdots \\ v_1 \end{bmatrix} G^{-1} = HG^{-1}$$

där matrisen H ges av (6.12). Matrisen H har determinanten ett och är således reguljär. Matrisen S är alltså reguljär om G är reguljär. Då matrisen H har en enkel invers, jämför (6.14), erhålls följande explicita uttryck på transformationsmatrisens invers

$$S^{-1} = GH^{-1} = [A^{n-1}B \quad A^{n-2}B \ldots B] \begin{bmatrix} 0 & 0 \ldots 0 & 1 \\ 0 & 0 \ldots 1 & a_1 \\ \vdots & & \\ 0 & 1 \ldots a_{n-3} & a_{n-2} \\ 1 & a_1 \ldots a_{n-2} & a_{n-1} \end{bmatrix}$$

dvs.

$$S^{-1} = [B, \; AB + a_1 B, \; \ldots, \; A^{n-1}B + a_1 A^{n-2}B + \ldots + a_{n-1}B] \tag{6.23}$$

Om s_i får beteckna den i:e kolonnen i S^{-1} erhålls följande rekursiva samband

$$s_1 = B$$

$$s_{i+1} = As_i + a_i B \tag{6.24}$$

Vi sammanfattar resultatet som

SATS 6.2

Anta att matrisen

$$G = [A^{n-1}B \quad A^{n-2}B \ \dots \ B]$$

är reguljär. Genom byte av koordinater i tillståndsrummet $z = Sx$ med

$$S^{-1} = [B, AB + a_1 B, \dots, A^{n-1}B + a_1 A^{n-2}B + \dots + a_{n-1}B]$$

transformeras det linjära tidsinvarianta systemet $\mathcal{S}(A, B, C, D)$ till

$$\frac{dz}{dt} = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u$$

$$y = [CB, CAB + a_1 CB, \dots, CA^{n-1}B + a_1 CA^{n-2}B + \dots + a_{n-1}CB]z + Du.$$

7. Insignal–utsignalrelationer. Viktfunktionen

I avsnitten 3, 4, 5 och 6 behandlades *interna* beskrivningar av det linjära systemet $\mathcal{S}(A, B, C, D)$. Speciellt fann vi i avsnitt 6 att matriserna A, B och C beror av det koordinatsystem som valts för att beskriva systemet. Vi skall nu behandla *externa* beskrivningar av systemet, dvs. beskrivningar som direkt anger en relation mellan insignalen och utsignalen. Formeln (4.8) ger

$$y(t) = Cx(t) + Du(t) = Ce^{A(t-t_0)}x(t_0) + Du(t) + C\int_{t_0}^{t} e^{A(t-s)}Bu(s)\,ds \tag{7.1}$$

Om vi nu antar att initialtillståndet $x(t_0)$ är noll samt att $t_0 = -\infty$, dvs. insignalen har påverkat systemet under intervallet $(-\infty, t)$, erhålls

$$y(t) = Du(t) + \int_{-\infty}^{t} Ce^{A(t-s)}Bu(s)\,ds \tag{7.2}$$

Denna formel är oberoende av det koordinatsystem som valts för att beskriva tillståndet. För att visa detta inför vi nya koordinater

$z = Tx$

och vi finner då

$$y(t) = CT^{-1}z(t) + Du(t) = Du(t) + CT^{-1}\int_{-\infty}^{t} e^{TAT^{-1}(t-s)}\, TBu(s)\, ds$$

Men

$$\exp TAT^{-1} = I + TAT^{-1} + \frac{1}{2!}\, TAT^{-1}TAT^{-1} + \ldots \frac{1}{n!}\,(TAT^{-1})^n + \ldots$$

$$= T\left[I + A + \frac{1}{2!}\, A^2 + \ldots + \frac{1}{n!}\, A^n + \ldots\right] T^{-1}$$

$$= T(\exp A)\, T^{-1}$$

och vi får således

$$y(t) = Du(t) + C\int_{-\infty}^{t} e^{A(t-s)}\, Bu(s)\, ds$$

Låt oss först anta att $D = 0$, dvs. att insignalen ej påverkar utsignalen direkt. Ekvation (7.2) innebär då att utsignalens värde vid tidpunkten t är ett viktat medelvärde av insignalens värden på intervallet $(-\infty, t)$. Insignalens värde vid tiden t_1 har vikten $C \exp A(t-t_1)B$. Funktionen

$$h(t) = Ce^{At}B \quad t \geqslant 0$$

kallas därför *viktfunktion* (weighting function). I det allmänna fallet då $D \neq 0$ finner vi att (7.2) formellt kan skrivas

$$y(t) = \int_{-\infty}^{t} [D\delta(t-s) + Ce^{A(t-s)}B]\, u(s)\, ds$$

där $\delta(t)$ är Dirac's deltafunktion.

DEFINITION 7.1

Med viktfunktionen för ett linjärt tidsinvariant system $\mathbb{S}(A, B, C, D)$ avses

$$h(t) = D\delta(t) + Ce^{At}B \quad t \geqslant 0 \tag{7.3}$$

Vi finner således från (7.1) och (7.3) att under förutsättning att initialtillståndet är noll kan insignal–utsignalrelationen skrivas

$$y(t) = \int_{-\infty}^{t} h(t-s)\, u(s)\, ds = \int_{0}^{\infty} h(s)\, u(t-s)\, ds \tag{7.4}$$

Vi skall slutligen ge en fysikalisk tolkning till viktfunktionen. Om vi antar att insignalen är en impuls vid tidpunkten $t_0 < t$, så blir utsignalen enligt (7.4)

$$y(t) = \int_{-\infty}^{t} h(t-s)\, \delta(s-t_0)\, ds = h(t-t_0)$$

Viktfunktionen är således systemets utsignal då insignalen är en impuls applicerad vid tiden $t_0 = 0$. Detta ger en konstruktiv metod för att bestämma viktfunktionen för ett system genom att applicera en puls på systemet och registrera utsignalen. Det är även orsaken till att viktfunktionen ibland kallas *impulssvar*.

Vi kan även utgå från egenskapen (7.4) och definiera viktfunktionen för linjära tidsinvarianta system på följande sätt.

DEFINITION 7.2

Viktfunktionen för ett linjärt tidsinvariant system är en funktion $h(t)$ sådan att följande relation gäller

$$y(t) = \int_{-\infty}^{t} h(t-s)\, u(s)\, ds$$

där u är systemets insignal och y systemets utsignal då initialtillståndet är noll.

Observera att viktfunktionen endast har definierats för icke negativa argument. Om vi önskar utsträcka definitionen även till negativa argument, är det med den ovan givna fysikaliska tolkningen naturligt att sätta $h(t) = 0$ för $t < 0$, ty annars skulle systemet kunna ge en utsignal innan insignalen applicerats, vilket strider mot kausalitetsprincipen. Med hjälp av viktfunktionen kan vi direkt ange sambandet mellan insignal och utsignal för ett system utan att specificera ett koordinatsystem i tillståndsrummet. Vi skall nu visa att vi för ett linjärt tidsinvariant system $\mathbb{S}(A, B, C, D)$ kan ange ett direkt samband mellan insignal och utsignal med hjälp av differentialoperatorer. Utgå ifrån

$$\frac{dx}{dt} = Ax + Bu$$

$$y = Cx + Du$$

Bilda

$$x = x$$

$$\frac{dx}{dt} = Ax + Bu$$

$$\frac{d^2x}{dt^2} = A\frac{dx}{dt} + B\frac{du}{dt} = A^2x + ABu + B\frac{du}{dt}$$

$$\vdots$$

$$\frac{d^nx}{dt^n} = A^nx + A^{n-1}Bu + A^{n-2}B\frac{du}{dt} + \ldots + B\frac{d^{n-1}u}{dt^{n-1}}$$

Låt matrisen A ha den karakteristiska ekvationen

$$\det[\lambda I - A] = \lambda^n + a_1\lambda^{n-1} + \ldots + a_n = 0$$

Multiplicera ovanstående ekvationer i ordning med $a_n, a_{n-1}, \ldots, a_1, 1$ och addera. Termerna i högerledet som innehåller x försvinner då på grund av Cayley–Hamiltons teorem och vi finner

$$\frac{d^n x}{dt^n} + a_1 \frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_n x = B\frac{d^{n-1}u}{dt^{n-1}} + (AB + a_1 B)\frac{d^{n-2}u}{dt^{n-2}} + \ldots$$
$$+ (A^{n-1}B + a_1 A^{n-2}B + \ldots + a_{n-1}B)\,u$$

Inför nu differentialoperatorn $p = \dfrac{d}{dt}$ och vi finner att insignal–utsignalrelationen kan skrivas

$$y(t) = \left\{D + \frac{B_1 p^{n-1} + B_2 p^{n-2} + \ldots + B_n}{p^n + a_1 p^{n-1} + \ldots + a_n \cdot 1}\right\} u(t) \tag{7.5}$$

där $B_1, B_2, \ldots, B_n$ är matriser som definieras av

$$B_1 = CB$$
$$B_2 = CAB + a_1 CB$$
$$\vdots$$
$$B_n = CA^{n-1}B + a_1 CA^{n-2}B + \ldots + a_{n-1}CB$$

Observera att representationen (7.5) är oberoende av koordinatsystemet i tillståndsrummet. Vid transformationen $z = Tx$ övergår systemet $\mathcal{S}(A, B, C, D)$ i $\mathcal{S}(TAT^{-1}, TB, CT^{-1}, D)$. Som vi tidigare har funnit i avsnitt 6 är den karakteristiska ekvationens koefficienter invarianta. Vidare övergår vid transformationen uttrycket CA^rB i följande uttryck

$$(CT^{-1})(TAT^{-1})^r TB = (CT^{-1})(TA^rT^{-1})TB = CA^rB$$

Vi finner således att även matriserna B_i är invarianta vid transformationen.

EXEMPEL 7.1

För den endimensionella partikeln, som behandlats i exempel 3.1 och exempel 4.1, gäller

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\quad C = [1 \quad 0],\quad D = 0$$

Viktfunktionen är således

$$h(t) = C\,e^{At}B = [1 \quad 0]\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = t$$

Insignal–utsignalrelationen kan således skrivas

$$y(t) = \int_{-\infty}^{t} (t - s)\,u(s)\,ds \tag{7.6}$$

För att ange insignal–utsignalrelationen som en differentialekvation bildar vi först A:s karakteristiska polynom

$$\det [\lambda I - A] = \lambda^2$$

Vidare gäller

$$B_1 = CB = [1 \quad 0] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

$$B_2 = CAB + a_1 CB = [1 \quad 0] \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0 = 1$$

Vi finner således

$$y = \frac{1}{p^2} u$$

eller det välbekanta

$$\frac{d^2y}{dt^2} = u \tag{7.7}$$

Detta kan även erhållas genom direkt derivation av (7.6).

EXEMPEL 7.2

För pendeln, som behandlats i exemplen 3.2 och 4.2, gäller om pivotpunktens koordinat väljs som insignal

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

Viktfunktionen blir

$$H(t) = D\delta(t) + Ce^{At}B = -\delta(t) + [1 \quad 0] \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\delta(t) + \sin t$$

Insignal–utsignalrelationen kan alltså skrivas

$$y(t) = \int_{-\infty}^{t} \sin (t-s)\, u(s)\, ds - u(t) \tag{7.8}$$

För att erhålla differentialekvationssambandet mellan insignal och utsignal bildar vi

$$\det [\lambda I - A] = \lambda^2 + 1$$

Vidare gäller

$$B_1 = CB = 0$$

$$B_2 = CAB + a_1 CB = [1 \quad 0] \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0 = 1$$

Vi finner således

$$y(t) = \left[-1 + \frac{1}{1+p^2} \right] u(t) = -\frac{p^2}{1+p^2}\, u(t)$$

eller

$$\frac{d^2y}{dt^2} + y = -\frac{d^2u}{dt^2} \tag{7.9}$$

Alternativt deriveras (7.8) två gånger

$$\frac{dy}{dt} = \int_{-\infty}^{t} \cos(t-s)\, u(s)\, ds - \frac{du}{dt}$$

$$\frac{d^2y}{dt^2} = u(t) - \int_{-\infty}^{t} \sin(t-s)\, u(s)\, ds - \frac{d^2u}{dt^2} = -y(t) - \frac{d^2u}{dt^2} \tag{7.10}$$

vilket ger samma resultat.

8. Laplacetransformen. Överföringsfunktionen

Laplacetransformen $\mathcal{L}f$ av en tidsfunktion f definieras som

$$(\mathcal{L}f)(s) = F(s) = \int_{0-}^{\infty} e^{-st} f(t)\, dt, \quad \text{Re } s > s_0 \tag{8.1}$$

om integralen existerar. Talet s_0 kallas konvergensabscissan och s är en komplex variabel. Med ovanstående formel är Laplacetransformen således endast definierad i ett halvplan, Re $s > s_0$. För värden på s som ej ligger i halvplanet Re $s > s_0$ definieras funktionen genom analytisk fortsättning om så är möjligt.

Laplacetransformen är således en linjär integraltransformation från reella funktioner $f(t)$, originalfunktioner, definierade på $(0, \infty)$ till en komplex funktion $F(s)$ av en komplex variabel s, bildfunktioner. Det är vanligt att beteckna originalfunktioner med gemena bokstäver $f(t)$, och bildfunktioner med versaler $F(s)$. I reglerteknisk litteratur förekommer ett allmänt utbrett (miss-) bruk genom att man använder samma symbol f för såväl bildfunktion som originalfunktion. Man skiljer då på funktionen och dess transform genom deras argument t (för originalfunktion) resp. s (för bildfunktion). Detta symbolmissbruk leder sällan till svårigheter men bör undvikas.

Det är ofta fördelaktigt att använda Laplacetransformen vid lösning av initialvärdesproblem för differentialekvationer (ordinära såväl som partiella). Skälet till detta är att följande samband råder

$$\mathcal{L}\frac{df}{dt} = sF(s) - f(0) \tag{8.2}$$

Derivation av en tidsfunktion motsvarar således en algebraisk operation (multiplikation med s) på bildfunktionen. Detta medför att Laplacetransformen kan användas för att bygga upp en algebraisk formalism för att lösa

systemekvationerna (3.1). Med hjälp av Laplacetransformen kan vi även införa det viktiga begreppet överföringsfunktion. En annan fördel med Laplacetransformen är att den öppnar vägen för användning av många av funktionsteorins mycket kraftfulla resultat på systemteorin. För en utförlig behandling av Laplacetransformen hänvisas till läroböcker i matematik.

Det är två svårigheter förenade med Laplacetransformen.

- För att få en bekvämt användbar kalkyl är det ej tillräckligt att låta $f(t)$ vara en vanlig funktion utan $f(t)$ bör tolkas som en distribution eller en generaliserad funktion. (Det är vanligt i reglertekniska läroböcker att undvika denna svårighet genom att räkna formellt med Dirac's deltafunktion.)
- Den inversa transformen given genom omvändningsformeln

$$f(t) = (\mathcal{L}^{-1}F)(t) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{st}F(s)\,ds \tag{8.3}$$

är ej en begränsad operator. Detta betyder att en transform som är »liten» i någon mening kan ha en mycket »stor» tidsfunktion. Som exempel kan vi betrakta följande par av transformer

$$f(t) = \frac{1}{\sqrt{a}}\sin at, \quad F(s) = \frac{\sqrt{a}}{s^2+a^2}$$

där tidsfunktionen f blir mycket stor trots att transformen är mycket liten för små a.

Några ord om beteckningar är även på sin plats. Det är vanligt att använda såväl s som p för att beteckna transformens argument. Vi har valt att följa det amerikanska bruket, dvs. s. Då variabeln t betecknar en tidsvariabel så följer att variabeln s har dimensionen invers tid eller frekvens. Av ekvation (8.1) följer att originalfunktion och bildfunktion har olika dimensioner. Detta medför att man i stället för Laplacetransformen ibland använder transformen

$$\hat{F}(p) = p\int_0^\infty e^{-pt}f(t)\,dt$$

den s. k. p-multiplicerade Laplacetransformen. Denna transformation bevarar uppenbarligen originalfunktionens dimension. I reglerteknisk litteratur används uteslutande Laplacetransformer.

För att lösa initialvärdesproblemet för systemekvationerna (3.1) måste vi först definiera Laplacetransformationen av en vektorvärd funktion. Vi gör detta genom att transformera komponentvis och finner då att ekvation (8.1) gäller även om $f(t)$ är en vektor (eller en matris).

Låt nu $X(s)$, $U(s)$ och $Y(s)$ beteckna Laplacetransformerna av $x(t)$, $u(t)$ och $y(t)$. Vi finner då med hjälp av (8.2) att Laplacetransformen av differentialekvationen (3.1) kan skrivas

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

Vi finner således

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

$$Y(s) = C[sI - A]^{-1}x(0) + C[sI - A]^{-1}BU(s) + DU(s) \tag{8.4}$$

En jämförelse med (7.1) ger som resultat

$$\mathcal{L}e^{At} = [sI - A]^{-1} \tag{8.5}$$

$$\mathcal{L}\int_0^t e^{A(t-\tau)}Bu(\tau)\,d\tau = [sI - A]^{-1}BU(s) \tag{8.6}$$

Observera analogin med skalära funktioner

$$\mathcal{L}\,e^{\alpha t} = (s - \alpha)^{-1} = \frac{1}{s - \alpha}$$

$$\mathcal{L}\int_0^t e^{\alpha(t-\tau)}u(\tau)\,d\tau = \frac{1}{s - \alpha}\,U(s)$$

Vi skall nu införa begreppet *överföringsfunktion* (transfer function).

DEFINITION 8.1

Överföringsfunktionen för ett linjärt tidsinvariant system är en funktion $G(s)$ sådan att följande relation gäller

$$Y(s) = G(s)\,U(s)$$

där $U(s)$ är insignalens Laplacetransform och $Y(s)$ utsignalens Laplacetransform då systemets initialtillstånd är noll.

Om systemet endast har en insignal och en utsignal så är överföringsfunktionen således kvoten mellan utsignalens och insignalens Laplacetransformer, då det förutsatts att initialtillståndet sätts lika med noll vid beräkning av utsignalens Laplacetransform. För flervariabla system är överföringsfunktionen en matris. Följande samband råder mellan viktfunktionen (definition 7.2) och överföringsfunktionen.

SATS 8.1

Överföringsfunktionen är viktfunktionens Laplacetransform.

Bevis:
Definitionen på viktfunktionen ger

$$y(t) = \int_{-\infty}^{t} h(t - \tau)\,u(\tau)\,d\tau$$

För att kunna definiera insignalens Laplacetransform förutsätter vi att tidsskalan valts så att $u(t) = 0$ $t < 0$. Då gäller

$$y(t) = \int_0^t h(t - \tau)\,u(\tau)\,d\tau$$

Definitionen 7.2 på Laplacetransformen ger

$$Y(s) = \int_0^\infty e^{-st} y(t)\, dt = \int_0^\infty \left[\int_0^t e^{-st} h(t-\tau)\, u(\tau)\, d\tau \right] dt$$

Skriv om integranden som

$$e^{-st} h(t-\tau)\, u(\tau) = e^{-s(t-\tau)} h(t-\tau)\, e^{-s\tau} u(\tau)$$

och byt integrationsordningen. Observera ordningsföljden då exponentialfunktionerna är skalärer men h och u är matriser! Vi finner

$$Y(s) = \int_0^\infty \left[\int_0^\infty e^{-s(t-\tau)} h(t-\tau)\, dt \right] e^{-s\tau} u(\tau)\, d\tau$$

Betrakta nu uttrycket inom hakparanteser. Det gäller

$$J = \int_0^\infty e^{-s(t-\tau)} h(t-\tau)\, dt = \int_\tau^\infty e^{-s(t-\tau)} h(t-\tau)\, dt = \int_0^\infty e^{-st} h(t)\, dt$$

Den första likheten följer av att $h(t) = 0$ då $t < 0$, den andra erhålls genom variabelbyte. Vi finner således

$$Y(s) = \left(\int_0^\infty e^{-st} h(t)\, dt \right) U(s) = H(s)\, U(s)$$

Påståendet i satsen följer nu av definitionen på överföringsfunktionen.

För det linjära tidsinvarianta systemet $\mathcal{S}(A, B, C, D)$ gäller

$$\frac{dx}{dt} = Ax + Bu$$

$$y = Cx + Du$$

Laplacetransformation ger under förutsättning att initialtillståndet är noll

$$sX(s) = AX(s) + BU(s)$$

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

Elimination av $X(s)$ ger

$$Y(s) = \{D + C[sI - A]^{-1} B\}\, U(s)$$

Överföringsfunktionen för detta system är således

$$G(s) = D + C[sI - A]^{-1} B \tag{8.7}$$

Om matrisen $[sI - A]^{-1}$ utvärderas med Kramers regel blir elementen rationella funktioner. Täljarpolynomen har gradtal mindre än n. Nämnarpolynomet $\det[sI - A]$, som är gemensamt för alla element, har graden n och kallas *det karakteristiska polynomet*. Överföringsfunktionen är således en matris vars

element är rationella funktioner. Den är oberoende av val av koordinater i tillståndsrummet ty efter transformationen $z = Tx$ övergår (3.1) i

$$\frac{dz}{dt} = TAT^{-1}z + TBu$$

$$y = CT^{-1}z + Du$$

Överföringsfunktionen för detta system är

$$G(s) = D + CT^{-1}[sI - TAT^{-1}]^{-1}TB$$

Men

$$[sI - TAT^{-1}]^{-1} = [sTT^{-1} - TAT^{-1}]^{-1} = \{T[sI - A]T^{-1}\}^{-1} = T[sI - A]^{-1}T^{-1}$$

Vi finner således

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

och påståendet är visat. Såväl viktfunktion som överföringsfunktion ger alltså en representation av insignal–utsignalrelationen, som är oberoende av val av koordinatsystem i tillståndsrummet.

EXEMPEL 8.1

För den endimensionella partikeln, som tidigare behandlats i exemplen 3.1, 4.1 och 7.1, gäller

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1 \quad 0], \quad D = 0$$

Vi finner

$$[sI - A]^{-1} = \begin{bmatrix} s & -1 \\ 0 & s \end{bmatrix}^{-1} = \frac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix}$$

Överföringsfunktionen för detta system är således

$$G(s) = 0 + [1 \quad 0]\frac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{s^2}$$

Detta resultat kan även erhållas genom en direkt Laplacetransformering av insignal–utsignalrelationen i exempel 7.1, ekvation (7.7).

EXEMPEL 8.2

Betrakta det dynamiska system, som består av pendeln med pivotpunktens läge som insignal, vilket behandlats i exempel 3.2 och 7.2. För detta system gäller

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

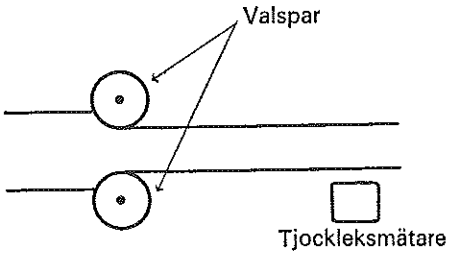


Fig. 8.1. Schematisk bild av bandvalsverk. Bandets tjocklek då det lämnar valsparet bestäms av dess tjocklek vid inträdet i valsparet, bandets hårdhet och valsparets ansättning. Av teknologiska skäl kan tjockleksmätaren ej placeras omedelbart intill valsparet.

Vi finner

$$[sI-A]^{-1}=\begin{bmatrix} s & -1 \\ 1 & s \end{bmatrix}^{-1}=\frac{1}{s^2+1}\begin{bmatrix} s & 1 \\ -1 & s \end{bmatrix}$$

Överföringsfunktionen för systemet blir således

$$G(s)=-1+\frac{1}{s^2+1}\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix} s & 1 \\ -1 & s \end{bmatrix}\begin{bmatrix}0 \\ 1\end{bmatrix}=-\frac{s^2}{s^2+1}$$

Detta kan även erhållas genom direkt Laplacetransformering av den differentialekvation (7.9) som relaterar insignal till utsignal.

Observera att definitionerna av överföringsfunktion och viktfunktion är giltiga även för oändligt dimensionella linjära system, dvs. för linjära tidsinvarianta system som ej kan beskrivas med $\mathbb{S}=\mathbb{S}(A, B, C, D)$ där A, B, C och D är matriser av ändlig ordning. Vi skall ge några exempel på detta.

EXEMPEL 8.3. Tidsfördröjning

I fig. 8.1 visas ett enkelt bandvalsverk.

Betrakta det dynamiska system vars insignal är valsparets ansättning u och vars utsignal y är tjockleksmätarens utslag. Anta att bandets hårdhet och dess tjocklek vid inträde i valsparet är konstant. Valsparets ansättning u kan direkt kalibreras i utgående tjocklek. Det dynamiska systemet kan då beskrivas med ekvationen

$$y(t)=u(t-T) \tag{8.8}$$

där T är den tid det tar för bandet att passera från valsarna till tjockleksmätaren. Systemets tillstånd vid tiden t är $\{u(s), t-T\leqslant s<t\}$, dvs. en funktion definierad på ett intervall av reella linjen. Laplacetransformation av (8.8) ger

$$Y(s)=\int_0^\infty e^{-st}y(t)\,dt=\int_0^\infty e^{-st}u(t-T)\,dt$$

$$=e^{-sT}\int_0^\infty e^{-s(t-T)}u(t-T)\,dt$$

$$=e^{-sT}\int_{-T}^0 e^{-st}u(t)\,dt+e^{-sT}U(s)$$

Om initialtillståndet är noll, dvs.

$u(t)=0$ då $-T\leqslant t<0$

erhålls

$Y(s)=e^{-sT}U(s)$

Definition på överföringsfunktionen ger

$$G(s)=\frac{Y(s)}{U(s)}=e^{-sT} \tag{8.9}$$

EXEMPEL 8.4. Värmeledning

Betrakta det dynamiska system som består av en rak homogen stång med täthet ϱ, specifikt värme C och värmeledningsförmåga k. Låt systemets insignal u vara temperaturen $u(t)$ i stavens vänstra ände. Låt $\vartheta(x,t)$ beteckna temperaturen vid tidpunkten t i en punkt på avståndet x från stavens vänstra ändpunkt samt låt utsignalen vara $\vartheta(x_0,t)$. Om vi antar att temperaturen är konstant i tvärsnitt vinkelräta mot stavens längdaxel gäller

$$\frac{\partial^2\vartheta(x,t)}{\partial x^2}-\frac{1}{a^2}\frac{\partial\vartheta(x,t)}{\partial t}=0,\quad a^2=\frac{k}{\varrho C} \tag{8.10}$$

$\vartheta(0,t)=u(t)$

Systemets tillstånd är temperaturfördelningen i staven, dvs. $\vartheta(x,t)$. För att bestämma överföringsfunktionen för systemet Laplacetransformerar vi ekvation (8.10). Inför

$$\theta(s,x)=\int_0^\infty e^{-st}\vartheta(x,t)\,dt$$

Vi finner då

$$\frac{\partial^2\theta(s,x)}{\partial x^2}-\frac{1}{a^2}(s\theta(s,x)-\vartheta(x,0))=0$$

där

$\theta(s,0)=U(s)$

och $\vartheta(x,0)$ är temperaturfördelningen i staven vid tidpunkten $t=0$, dvs. systemets initialstillstånd. Anta nu att initialtillståndet är noll och vi får då

$$\frac{\partial^2\theta}{\partial x^2}-\frac{s}{a^2}\theta=0$$

Denna ekvation har lösningen

$$\theta(s,x)=A\exp\sqrt{s}\,\frac{x}{a}+B\exp-\sqrt{s}\,\frac{x}{a}$$

Koefficienterna bestäms ur randvillkor. Överföringsfunktionen är ändlig då $x \to \infty$ endast om $A = 0$. Koefficienten B bestäms sedan ur

$$U(s) = \theta(s, 0) = B$$

Vi finner således

$$\theta(s, x) = U(s) \exp - \sqrt{s}\, \frac{x}{a} \tag{8.11}$$

Men utsignalen är temperaturen i punkten x_0, dvs.

$$Y(s) = U(s) \cdot \exp - \sqrt{sT}, \quad T = x_0^2 a^{-2} = x_0^2 \varrho\, C/k$$

Systemets överföringsfunktion är således

$$G(s) = \frac{Y(s)}{U(s)} = \exp - \sqrt{sT} \tag{8.12}$$

I detta fall erhålls alltså en transcendent funktion. Genom inverstransformering finner vi efter vissa räkningar följande uttryck för viktfunktionen

$$h(t) = \frac{1}{2\pi i} \int_\Gamma e^{st} G(s)\, ds = \frac{x}{2a\sqrt{\pi}}\, t^{-\frac{3}{2}} \exp - \frac{x^2}{4a^2 t} \tag{8.13}$$

9. Grafisk representation av överförings-funktionen

Överföringsfunktioner definierades i föregående avsnitt. Vi fann speciellt att överföringsfunktionen $G(s)$ för det linjära tidsinvarianta systemet $\mathcal{S}(A, B, C, D)$ är en matris vars element är brutna rationella funktioner av en komplex variabel. I detta avsnitt skall vi behandla några olika grafiska representationer för sådana funktioner. För att förenkla behandlingen skall vi uteslutande diskutera enkla system. För flervariabla system får vi genomföra en liknande diskussion för varje element av matrisen $G(s)$.

Singularitetsdiagram

En bruten rationell funktion är med undantag av en konstant faktor fullständigt karakteriserad av sina poler och nollställen. För att studera detta skriver vi överföringsfunktionen på följande form

$$G(s) = K\, \frac{(s - z_1)(s - z_2) \ldots (s - z_m)}{(s - p_1)(s - p_2) \ldots (s - p_n)} \tag{9.1}$$

där z_i är nollställen och p_i poler. Funktionen $G(s)$ är fullständigt bestämd om samtliga p_i och z_i samt talet K anges. Funktionen kan då representeras genom

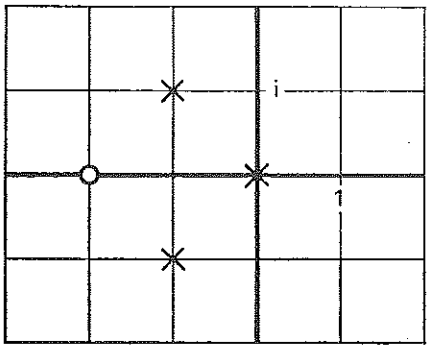


Fig. 9.1. Singularitetsdiagram för överföringsfunktionen

$$G(s) = 5\,\frac{s+2}{s(s^2+2s+2)}$$

att polerna p_i och nollställena z_i markeras i komplexa talplanet. Konstanten K anges separat. Det erhållna diagrammet kallas *singularitetsdiagram*. Då överföringsfunktionen har reella koefficienter uppträder komplexa singulariteter som komplexkonjugerade par och singularitetsdiagrammet är därför symmetriskt med avseende på reella axeln i s-planet (fig. 9.1).

Om överföringsfunktionen är specificerad med singularitetsdiagrammet, kan man med fördel utvärdera dess värden för speciella argument med en grafisk metod. Det gäller att

$$|G(s)| = K\,\frac{\prod|s-z_i|}{\prod|s-p_i|} \tag{9.2}$$

$$\arg G(s) = \Sigma \arg (s-z_i) - \Sigma \arg (s-p_i) \tag{9.3}$$

Såväl $|s-z_i|$ som $\arg (s-z_i)$ har enkla geometriska tolkningar, och $G(s)$ kan således utvärderas genom direkta mätningar i singularitetsdiagrammet (fig. 9.2). En speciell räknesticka, s. k. spirule, underlättar arbetet.

Frekvenskurvor

Vi har tidigare konstaterat att överföringsfunktionen är analytisk med undantag av vissa singularitetspunkter. En analytisk funktion är emellertid fullständigt bestämd av sina värden på en konvergerande punktföljd. Detta innebär att det är onödigt att ange överföringsfunktionens värden i hela s-planet. Det är tillräckligt att ange dess värden på en konvergerande punktföljd. Med hjälp av en Taylorserieutveckling kan man då definiera funktionen i en cirkulär omgivning till hopningspunkten. Sedan kan man med analytisk fortsättning utsträcka funktionens definitionsområde. Det är t. ex. vanligt att ange överföringsfunktionens värden på imaginära axeln i s-planet. Motivet till att välja funktionsvärden på imaginära axeln är att dessa har en trevlig fysikalisk tolkning. För att demonstrera detta betraktar vi systemet $\mathcal{S}(A, B, C, D)$. Låt insignalen vara en sinusfunktion

$$u = u^0 \sin \omega t = \mathrm{Im}\,\{u^0 e^{i\omega t}\}$$

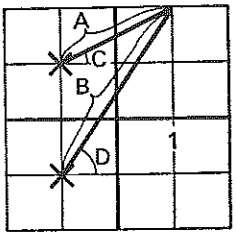


Fig. 9.2. Illustrerar beräkning av $G(1+2i)$ ur singularitetsdiagrammet för överföringsfunktionen

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

Beloppet $|G(1+2i)|$ erhålls som inverterade produkten av sträckorna A och B och argumentet arg $G(1+2i)$ erhålls som negativa summan av vinklarna C och D.

Systemekvationen (3.1) har då en partikulärlösning

$$x(t) = \text{Im}\,\{x^0 e^{i\omega t}\}, \quad y(t) = \text{Im}\,\{y^0 e^{i\omega t}\}$$

där x^0 är ett komplext tal givet av $i\omega x^0 = Ax^0 + Bu^0$ dvs.

$$x^0 = [i\omega I - A]^{-1} B u^0$$

Det har då förutsatts att matrisen A ej har något egenvärde med realdelen noll. Vi finner således

$$y^0 = \{D + C[i\omega I - A]^{-1} B\} u^0 = G(i\omega) u^0 = |G(i\omega)|\, u^0 e^{i \arg G(i\omega)}$$

dvs.

$$y(t) = \text{Im}\,\{|G(i\omega)|\, u^0 e^{i(\omega t + \arg G(i\omega))}\} = u^0 |G(i\omega)| \sin(\omega t + \arg G(i\omega)) \tag{9.4}$$

Om insignalen är en sinusfunktion med frekvensen ω finns det således en partikulärlösning till systemekvationerna sådan att även utsignalen blir en sinusfunktion med samma frekvens ω. Överföringsfunktionens belopp $|G(i\omega)|$ anger alltså kvoten mellan utsignalens och insignalens amplituder, och överföringsfunktionens argument arg $G(i\omega)$ anger utsignalens fasläge i förhållande till insignalen. På grund av detta har $G(i\omega)$ fått ett speciellt namn, *frekvensfunktionen*. Om man låter insignalen vara en sinusfunktion och mäter insignalens och utsignalens relativa faser och amplituder, har vi således en praktisk metod för bestämning av överföringsfunktioner. Denna metod kallas *frekvensanalys*.

Observera att i den analys som ledde fram till ekvation (9.4) bestämde vi endast en partikulärlösning. I det allmänna fallet tillkommer till ekvation (9.4) en term, som utgör allmänna lösningen till homogena delen. Vid praktisk frekvensanalys får man se till att denna tilläggsterm blir försumbar. Detta uppnås vanligen genom att insignalen får påverka systemet en viss tid innan mätningen utförs. Om matrisen A har samtliga egenvärden i vänstra halv-

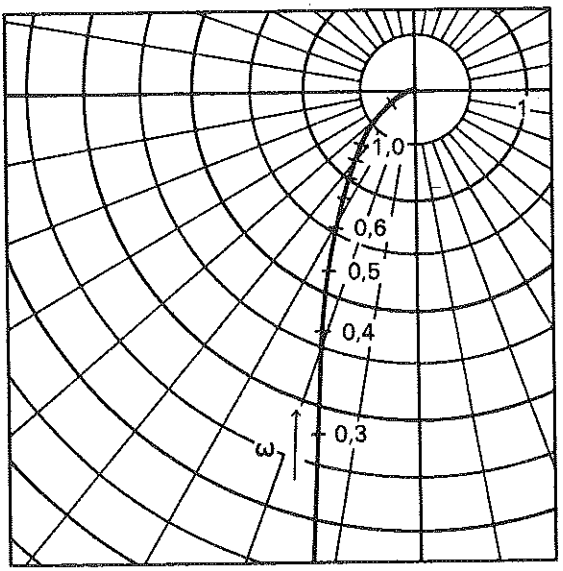


Fig. 9.3. Frekvenskurva för överföringsfunktionen $G(s) = \dfrac{1}{s(s+1)}$.

planet, konvergerar nämligen varje lösning till homogena ekvationen mot noll då t växer. Jämför avsnitt 11.

Med *frekvenskurvan, polära diagrammet* eller *Nyquistdiagrammet*, avses grafen av det komplexa talet $G(i\omega)$ (fig. 9.3). Låt alternativt w vara en komplex variabel $w = G(s)$. Frekvenskurvan är då avbildningen av imaginära axeln i s-planet på w-planet. Då överföringsfunktionens koefficienter är reella är frekvenskurvan symmetrisk med avseende på den reella w-axeln, och det är därför tillräckligt att endast avbilda den positiva halvaxeln $s = i\omega$, $\omega \geqslant 0$. Frekvenskurvan brukar parametriseras i ω. Frekvenskurvan har således en mycket åskådlig fysikalisk tolkning då den anger hur amplitud och fas av en sinusfunktion förändras vid passage genom systemet.

Uppritning av frekvenskurvor är i allmänhet tidsödande. Vid uppritandet har man ofta hjälp av att approximera $G(s)$ för stora och små $|s|$. I det fall som illustreras i fig. 9.3 erhålles t. ex. för små $|s|$

$$G(s) \sim \frac{1}{s}$$

Punkter $s = i\varepsilon$, $\varepsilon \ll 1$, avbildas således på en punkt $G \sim -i/\varepsilon$, dvs. nära negativa imaginära axeln och på stort avstånd från origo. För att bestämma i vilken kvadrant punkten ligger approximerar vi litet bättre

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

och finner då att bilden $s = i\varepsilon$, $\varepsilon > 0$, $G = -1 - i/\varepsilon$, ligger i tredje kvadranten. Analogt finner vi för stora $|s|$ att

$$G(s) \sim s^{-2}$$

Punkten $s = ir$, $r \gg 1$, avbildas på $G = -1/r^2$, dvs. på en punkt på negativa reella axeln. På detta sätt kan man snabbt skissa frekvenskurvans ungefärliga förlopp.

Bodediagram

Det är jämförelsevis tidsödande att rita upp frekvenskurvor. Man använder därför en annan grafisk representation av överföringsfunktionen, *logaritmiska frekvenskurvor* eller *Bodediagram*. Sådana diagram kan lätt uppritas även för mycket komplicerade överföringsfunktioner. Bodediagrammet består av två kurvor: *beloppkurvan* eller *amplitudkurvan* anger $\log|G(i\omega)|$ som funktion av ω, *argumentkurvan* eller *faskurvan* anger $\arg G(i\omega)$ som funktion av ω. Abscissa är i båda fallen frekvensen ω i logaritmisk skala. Beloppet anges ibland i *decibel* (dB), *decilog* (dl) eller *neper* (N) varvid $|G| = 10$ svarar mot 20 dB, 10 dl eller 2.303 N. För frekvensaxelns gradering används ofta begreppen *oktav* och *dekad*, motsvarande fördubbling respektive tiofaldigande.

Vi skall nu visa hur man ritar Bodediagram för system av ändlig ordning. För sådana system är överföringsfunktionen en rationell funktion. Uppritning av Bodediagram börjar med att man faktorerar överföringsfunktionen

$$G(s) = K\frac{(s-z_1)(s-z_2)\ldots(s-z_m)}{s^n(s-p_1)(s-p_2)\ldots(s-p_n)}$$

Komplexa poler och nollställen skrivs härvid som andra ordningens faktorer av typen

$$s^2 + 2\zeta\omega s + \omega^2$$

För ett system av ändlig ordning kan överföringsfunktionen således skrivas som en produkt av termer av typen

$$K$$

$$s^n$$

$$(1+sT)^n$$

$$(1+2\zeta sT+(sT)^2)^n$$

där n är ett positivt eller negativt heltal. Överföringsfunktionen G:s logaritm är en summa av logaritmer av sådana termer. Om vi kan rita Bodediagrammen för de enskilda termerna, kan vi sedan erhålla Bodediagrammet för överföringsfunktionen G genom att addera de enskilda termernas Bodediagram. Vi skall nu visa hur man ritar Bodediagram för de enkla termerna.

1. *Bodediagram för $G(s) = $ konstant*

Om $G(s) = K$, $K > 0$ gäller $\log|G(s)| = \log K$ och $\arg K = 0$. Bodediagrammet för en konstant består således av två horisontella räta linjer $\log G(i\omega) = \log K$ och $\arg G(i\omega) = 0$.

2. *Bodediagram för* $G(s)=s^n$

Det gäller att $\log|G(i\omega)|=n\log\omega$ och $\arg G(i\omega)=\frac{1}{2}n\pi$. Argumentkurvan är således en horisontell linje. Beloppkurvan är också en rät linje som har lutningen n om samma skala väljs på axlarna.

3. *Bodediagram för* $G(s)=(1+sT)^n$

Det gäller att $\log|G(i\omega)|=n\log\sqrt{1+\omega^2T^2}$ och $\arg G(i\omega)=n\operatorname{arctg}\omega T$. För $\omega T\ll 1$ fås $\log|G(i\omega)|\sim 0$ och $\arg G(i\omega)\sim 0$. För höga frekvenser, $\omega T\gg 1$ fås $\log|G(i\omega)|\sim n\log\omega T$ och $\arg G(i\omega)\sim\frac{1}{2}n\pi$. Vi finner således att beloppkurvan har två asymptoter

$\log|G(i\omega)|=0$ lågfrekvensasymptoten

$\log|G(i\omega)|=n\log\omega T$ högfrekvensasymptoten

Lågfrekvensasymptoten är horisontell och högfrekvensasymptoten har lutningen n. Dessa asymptoter skär varandra då $\omega=1/T$, den s. k. *brytfrekvensen*. Vidare finner vi att argumentkurvan har två horisontella asymptoter

$\arg G(i\omega)=0$ lågfrekvensasymptoten

$\arg G(i\omega)=n\frac{\pi}{2}$ högfrekvensasymptoten

Bodediagrammet för $n=1$ anges i fig. 9.4. Det framgår av figuren att beloppkurvans avvikelse från asymptoten är obetydlig utom i närheten av brytfrekvensen.

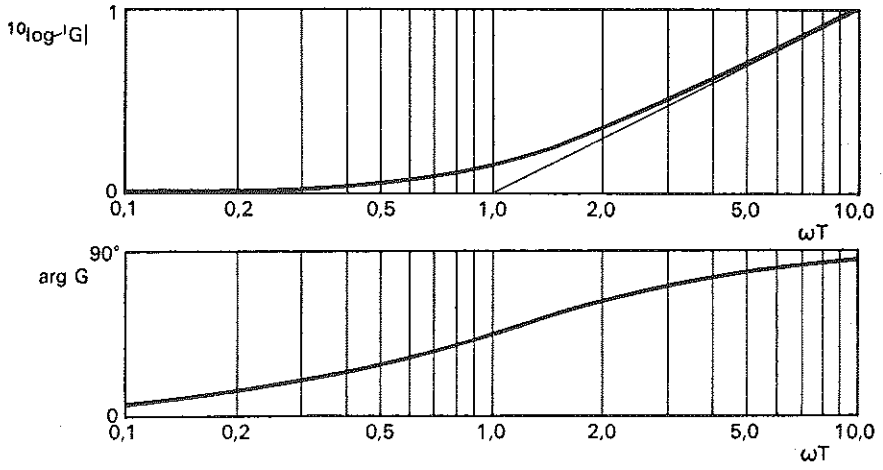
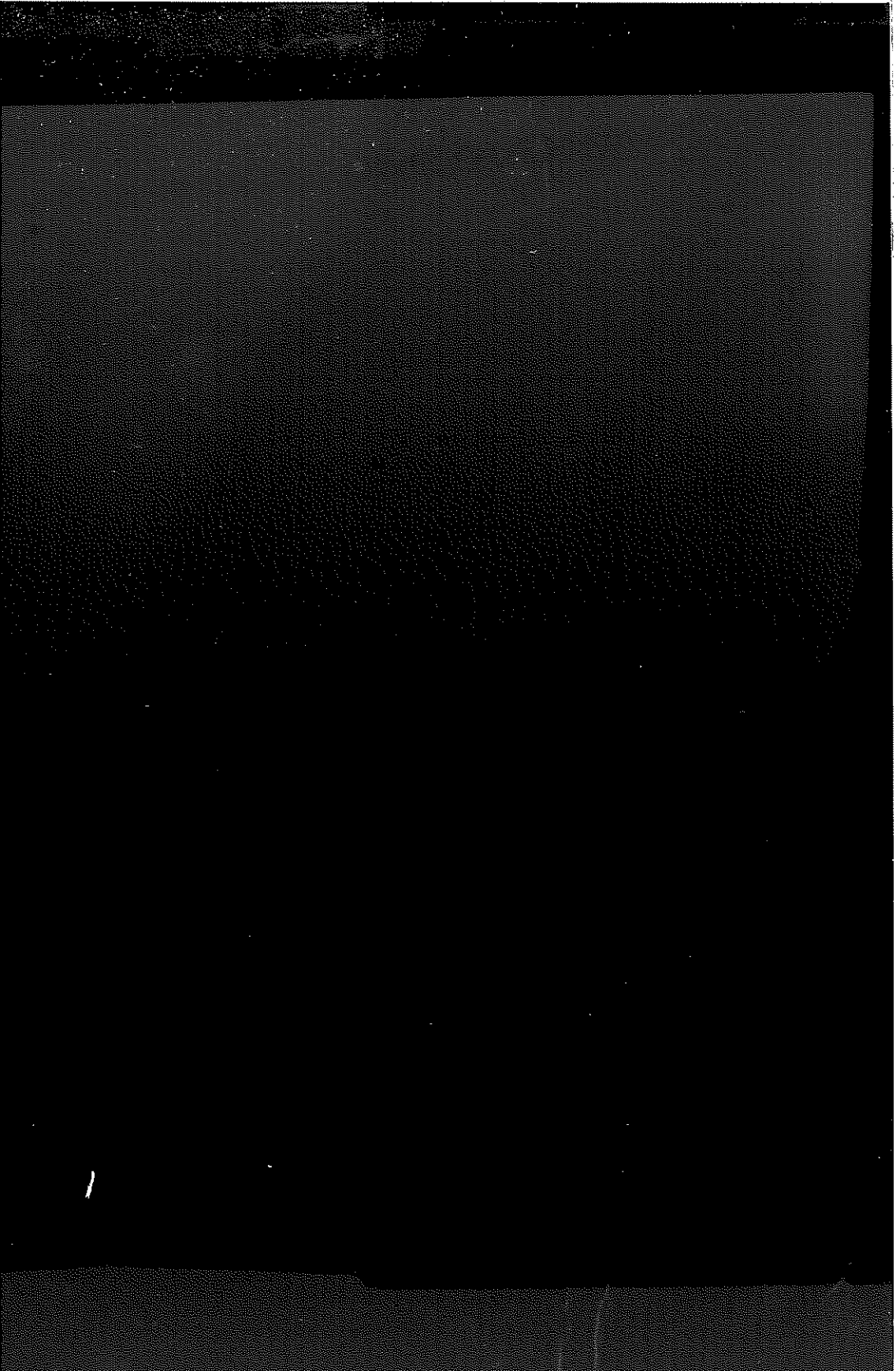


Fig. 9.4. Bodediagram för överföringsfunktionen $G(s)=1+sT$.



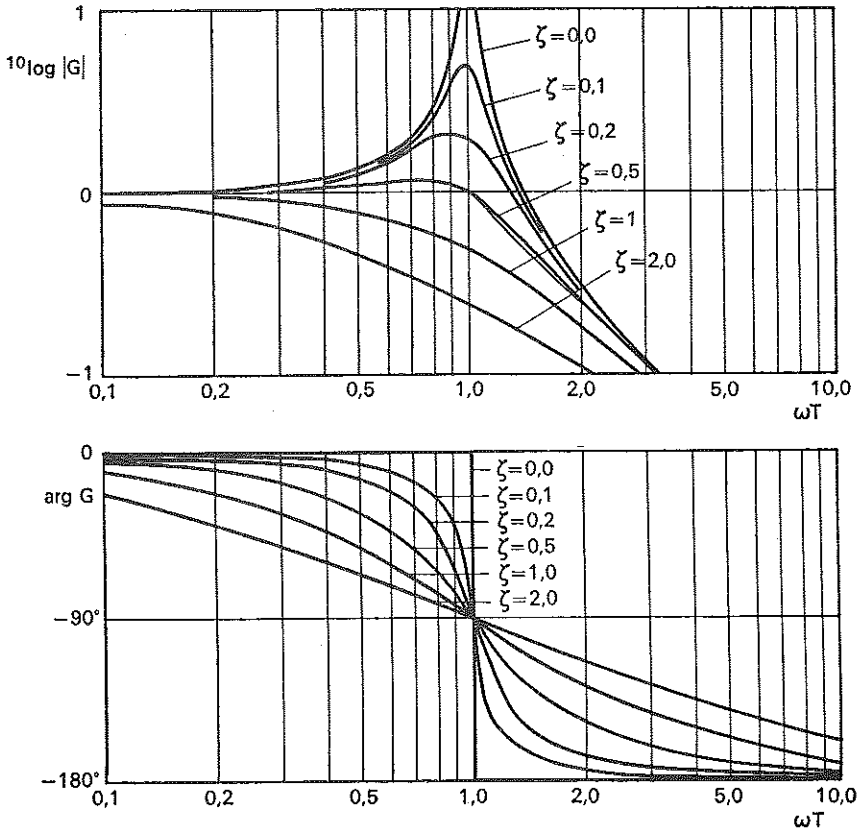


Fig. 9.5. Bodediagram för överföringsfunktionen $G(s) = (1 + 2\zeta sT + s^2T^2)^{-1}$.

4. *Bodediagram för* $G(s) = (1 + 2\zeta sT + s^2T^2)^n$

Det gäller att

$$\log|G(i\omega)| = \frac{n}{2}\log\{(1-\omega^2T^2)^2 + 4\zeta^2\omega^2T^2\}$$

$$\arg G(i\omega) = n \operatorname{arctg}\frac{2\zeta\omega T}{1-\omega^2T^2}$$

För $\omega T \ll 1$ fås $\log|G(i\omega)| \sim 0$ och $\arg G(i\omega) \sim 0$. För $\omega T \gg 1$ fås $\log|G(i\omega)| \sim 2n \log \omega T$ och $\arg G(i\omega) \sim n\pi$.

Beloppkurvan har således två asymptoter

$\log|G(i\omega)| = 0$ lågfrekvensasymptoten

$\log|G(i\omega)| = 2n \log \omega T$ högfrekvensasymptoten

Dessa asymptoter skär varandra då $\omega = 1/T$, den s. k. *brytfrekvensen*.

Lågfrekvensasymptoten är horisontell och högfrekvensasymptoten har lutningen $2n$. Argumentkurvan har två horisontella asymptoter

$\arg G(i\omega) = 0$ lågfrekvensasymptoten

$\arg G(i\omega) = n\pi$ högfrekvensasymptoten

Bodediagrammet för $n = -1$ framgår av fig. 9.5.

Bodediagrammet för en rationell överföringsfunktion kan erhållas genom att sätta samman diagrammen för de elementärfaktorer som har behandlats. Eftersom amplitudkurvan har asymptoter som är sammansatta av räta linjer är det lätt att skissa dess utseende. Vi illustrerar detta genom att skissa Bodediagrammet för överföringsfunktionen

$$G(s) = \frac{100(s+0{,}1)^2}{s^2(s^2+0{,}4s+0{,}25)(s+2)}$$

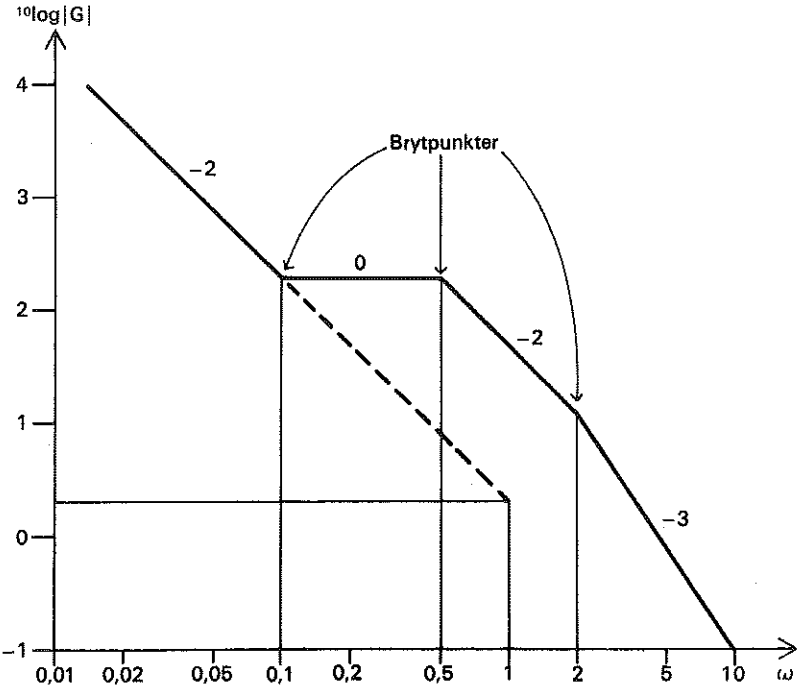


Fig. 9.6. Illustrerar konstruktion av Bodediagram för överföringsfunktionen

$$G(s) = \frac{2(1+10s)^2}{s^2\{1+0{,}8(2s)+(2s)^2\}(1+0{,}5s)}$$

Vi skriver först $G(s)$ på en standardform

$$G(s) = \frac{2(1+10s)^2}{s^2[(2s)^2 + 2 \cdot 0{,}4(2s) + 1](1+0{,}5s)}$$

Överföringsfunktionen $G(s)$ kan således erhållas genom multiplikation av följande elementärfaktorer

$$G_1(s) = 2 \cdot s^{-2}$$

$$G_2(s) = (1+10s)^2$$

$$G_3(s) = [1 + 2 \cdot 0{,}4(2s) + (2s)^2]^{-1}$$

$$G_4(s) = (1+0{,}5s)^{-1}$$

vilka har brytfrekvenserna $\omega = 0{,}1$, 0,5 och 2. Amplitudkurvans asymptot är sammansatt av segment av räta linjer mellan brytpunkterna. För låga frekvenser gäller $G(s) \approx 2 \cdot s^{-2}$. Lågfrekvensasymptoten har således lutningen -2. Vidare passerar asymptoten genom punkten $|G(i\omega)| = 2 (\log |G(i\omega)| = 0{,}301)$ för $\omega = 1$. Lågfrekvensasymptoten kan nu uppritas. Se fig. 9.6. Den lägsta brytfrekvensen $\omega = 0{,}1$ svarar mot ett dubbelt nollställe (faktorn G_2). Vid brytpunkten ökar asymptotens lutning med två enheter och asymptoten kan nu ritas i intervallet $0{,}1 \leqslant \omega < 0{,}5$. Se fig. 9.6. Den näst lägsta brytfrekvensen $\omega = 0{,}5$ svarar mot faktorn G_3 som har två komplexa poler. Vid denna brytfrekvens minskar asymptotens lutning med två enheter och asymptoten kan nu ritas i intervallet $0{,}5 \leqslant \omega < 2$. Se fig. 9.6. Den sista brytfrekvensen $\omega = 2$ svarar mot faktorn G_4 som har en enkel pol och asympotens lutning minskar då med en enhet till -3. Eftersom inga ytterligare brytpunkter finns kan asymptoterna nu kompletteras i hela intervallet $2 \leqslant \omega < \infty$. Se fig. 9.6. Den gjorda konstruktionen kan kontrolleras genom att observera att vid höga frekvenser gäller $|G(i\omega)| \approx 100 \cdot \omega^{-3}$. För höga frekvenser skall asymptoten således ha lutningen -3. Högfrekvens-asymptoten passerar också genom punkten $\omega = 10$, $|G(i\omega)| = 0{,}1$. Det är således lätt att rita amplitudkurvans asymptoter. Om man önskar känna amplitudkurvan noggrannare kan dess värden bestämmas genom grafisk addition av bidragen från elementarfaktorerna. Man kan för detta ändamål ha stor glädje av standardkurvor för elementarfaktorernas Bodediagram, fig. 9.4 och fig. 9.5. Faskurvan bestäms på analogt sätt.

Bodes relationer. Minimum-fas-system

Belopp och argumentkurvorna för överföringsfunktionen till ett dynamiskt system är beroende av varandra. Det går att visa att för ett system vars

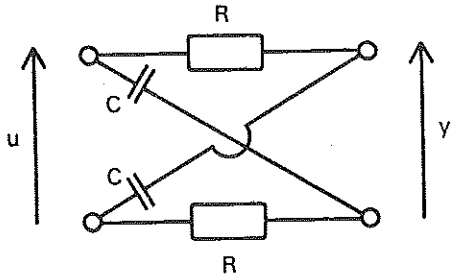


Fig. 9.7. Enkel *RC*-krets som ej är ett minimum-fas-system.

överföringsfunktion ej har några singulariteter i högra halvplanet och för vilken $(\log|G(i\omega)|)/(1+\omega^2)$ är integrerbar över $(0, \infty)$ gäller att

$$\arg G(i\omega_1) \leqslant \frac{2\omega_1}{\pi}\int_0^\infty \frac{\log|G(i\omega)| - \log|G(i\omega_1)|}{\omega^2-\omega_1^2}\,d\omega \qquad (9.5)$$

Olikheten innebär att om överföringsfunktionens beloppkurva är given, så ligger argumentkurvan alltid under en given kurva. Man kan vidare visa att det alltid finns en överföringsfunktion sådan att likhet erhålls i ovanstående relation. System med denna egenskap kallas *minimum-fas-system*. Av alla linjära system med given beloppkurva har således minimum-fas-systemen minsta möjliga fasförskjutning. Man kan visa att ett nödvändigt och tillräckligt villkor för att ett system skall vara minimum-fas är att överföringsfunktionens samtliga poler och nollställen ligger i vänstra halvplanet. För ett minimum-fas-system gäller således

$$\arg G(i\omega_1) = \frac{2\omega_1}{\pi}\int_0^\infty \frac{\log|G(i\omega)| - \log|G(i\omega_1)|}{\omega^2-\omega_1^2}\,d\omega \qquad (9.6)$$

$$\log|G(i\omega)| = \log|G(0)| - \frac{2\omega_1^2}{\pi}\int_0^\infty \frac{\omega^{-1}\arg G(i\omega) - \omega_1^{-1}\arg G(i\omega_1)}{\omega^2-\omega_1^2}\,d\omega \qquad (9.7)$$

Ekvationerna (9.6) och (9.7) vilka ger sambandet mellan amplitud- och faskurvor för ett minimum-fas-system kallas *Bodes relationer*. De kan härledas genom tillämpning av teorin för analytiska funktioner. Det följer av ekvation (9.6) att för ett minimum-fas-system motsvarar en asymptot till beloppkurvan med lutningen n, en asymptot till argumentkurvan som har ekvationen $\arg G(i\omega) = \frac{1}{2}n\pi$. Vi skall här ej analysera Bodes relationer närmare utan nöjer oss med att konstatera att ett system med en rationell överföringsfunktion är minimum fas om alla poler och nollställen ligger i vänstra halvplanet.

Begreppet minimum-fas infördes av Bode i samband med analys av återkopplade förstärkare. Det är mycket viktigt för reglertekniken ty det finns en gräns för reglersnabbheten hos system som *ej* är minimum-fas. Vi skall återkomma till detta i kapitel 6. Vi avslutar detta avsnitt med att ge några exempel på system som ej är minimum-fas.

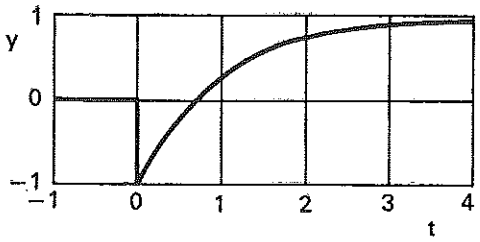


Fig. 9.8. Stegfunktionssvaret för överföringsfunktionen

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

EXEMPEL 9.1

Nätverket i fig. 9.7 har överföringsfunktionen

$$G(s) = \frac{1-sT}{1+sT}, \quad T = RC$$

Det gäller att

$$|G(i\omega)| = 1$$

$$\arg G(i\omega) = -2 \operatorname{arctg} \omega T$$

Systemet med överföringsfunktionen $G_1(s) = 1$ har uppenbarligen samma amplitudfunktion som $G(s)$ men $\arg G_1(i\omega) = 0$. Nätverket i fig. 9.7 är således ej minimum-fas. Systemets stegfunktionssvar framgår av fig. 9.8. Observera att utsignalen till en början går åt fel håll. Detta ger en indikation på att systemet kan vara svårt att styra.

EXEMPEL 9.2. Vattenturbinens överföringsfunktion

Betrakta det dynamiska system som utgörs av en vattenturbin enligt fig. 9.9. Låt systemets styrvariabel vara turbinpådraget graderat i den effektiva öppningsarean a. Låt utsignalen vara den maximalt utnyttjningsbara turbineffekten P definierad av

$$P = \frac{1}{2} \varrho v_{ut}^2 Q$$

där v_{ut} är utströmningshastigheten, ϱ tätheten och Q volymflödet. För att erhålla en matematisk modell för systemet tillämpas Bernoullis teorem på strömlinjen AB i fig. 9.9. Vi finner då

$$\int_A^B \frac{\partial \mathbf{v}}{\partial t} \cdot d\mathbf{r} + \frac{1}{2}(v_B^2 - v_A^2) + \Omega_B - \Omega_A + \int_A^B \frac{dp}{\varrho} = 0$$

där $\mathbf{v}$ betecknar hastighetsvektorn i strömningsfältet, $d\mathbf{r}$ linjeelementet längs strömlinjen, v hastighetens belopp, Ω tyngdkraftens potential, p trycket och ϱ tätheten.

Om vi antar att bassängen är mycket stor är hastigheten väsentligen skild från noll endast i anloppsröret. Anta vidare att tätheten är konstant. Vi finner då

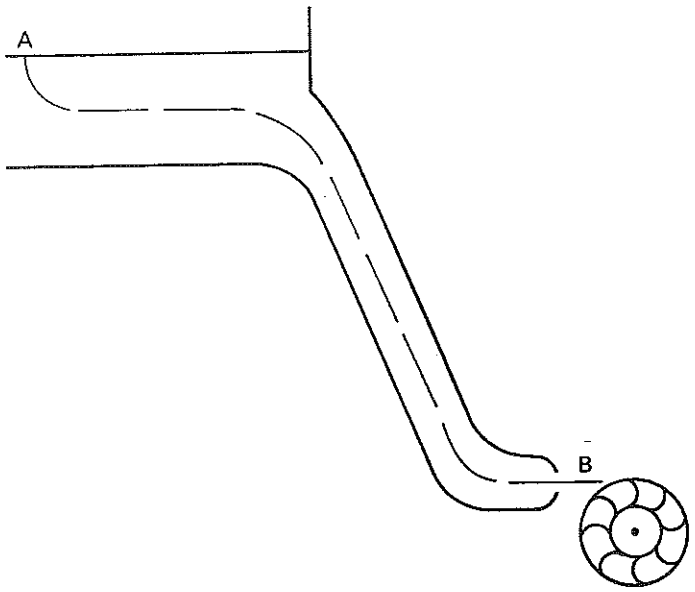


Fig. 9.9. Förenklad bild av vattenturbin. Samband mellan maximalt tillgänglig effekt och turbinpådrag representerar ett icke minimum-fas-system.

$$L\frac{dv}{dt}+\frac{1}{2}v_{ut}^2-gh=0$$

där v är strömningshastigheten i anloppstuben vilken antas ha konstant tvärsnitt. Kontinuitetsekvationen ger

$$v_{ut}a=vA$$

Vi finner

$$\frac{dv}{dt}=\frac{1}{L}gh-\frac{1}{2L}\left(\frac{A}{a}v\right)^2$$

Den maximalt uttagbara turbineffekten kan skrivas

$$P=\frac{1}{2}\varrho v_{ut}^2 Q=\frac{1}{2}\varrho a v_{ut}^3=\frac{1}{2}\varrho\frac{A^3v^3}{a^2}$$

Införs tillståndsvariabeln $x=v$, styrvariabeln $u=a/A$ och utsignalen $y=P$ så erhålls följande ekvationer för systemet

$$\frac{dx}{dt}=gh/L-x^2/(2Lu^2)$$

$$y=\varrho A x^3/(2u^2)$$

Vid ett givet turbinpådrag $u=u_0$ erhålls den stationära lösningen

$$x_0=u_0\sqrt{2gh};\quad y_0=P_0=\varrho A x_0^3/(2u_0^2)$$

Linjärisering av systemekvationerna kring den stationära lösningen $x(t)=x_0$ ger

$$\frac{d(x-x_0)}{dt} = -\frac{\sqrt{2gh}}{u_0 L}(x-x_0)+\frac{2gh}{u_0 L}(u-u_0)$$

$$y-P_0 = \frac{3P_0}{u_0\sqrt{2gh}}(x-x_0)-\frac{2P_0}{u_0}(u-u_0)$$

Efter Laplacetransformation av dessa ekvationer

$$Y(s)=\mathcal{L}[y(t)-y_0];\ U(s)=\mathcal{L}[u(t)-u_0]$$

erhålls följande överföringsfunktion för systemet

$$\frac{Y(s)}{U(s)} = \frac{P_0}{u_0}\frac{1-2u_0 sT}{1+u_0 sT} \tag{9.8}$$

där $T=L/\sqrt{2gh}$ betecknar den tid det tar för vattnet att rinna genom anloppsröret. Tidskonstanten T är av storleksordningen sekunder. Vid en plötslig ökning av ventilpådraget kommer således effekten momentant att minska för att sedan växa upp exponentiellt till ett högre värde. Den momentana minskningen är dubbelt så stor som den ökning som erhålls i stationärt tillstånd.

Det finns många andra tekniska system som ej är minimum-fas. Överföringsfunktionen e^{-sT} som representerar en transportfördröjning (jämför exempel 8.3) är ej minimum-fas. Då transportfördröjningar förekommer i många tekniska system förstår vi att förekomsten av icke minimum-fas-system är vanlig. Sambandet mellan höjd och höjdroderutslag i ett flygplan är ett annat exempel. Då höjdrodret vrids uppåt erhålls en kraft, som till att börja med accelererar flygplanet nedåt. Kraften ger emellertid även ett moment, som vrider flygplanet så att anfallsvinkeln ökar. Härvid ökar lyftkraften och flygplanet stiger. Relationen mellan domnivå och matarvattenflöde i en ångpanna liksom positionen för en backande bil är ytterligare exempel på icke minimum-fas-system. Vi återkommer senare i kap. 6 med en diskussion om svårigheterna att reglera icke minimum-fas-system.

10. Grafisk representation av dynamiska system

Det är vanligt i den reglertekniska litteraturen att dynamiska system representeras med grafiska symboler av olika slag. I det följande ges en kortfattad översikt av två sådana grafiska representationer, *blockschema* resp. *signalflödesdiagram*.

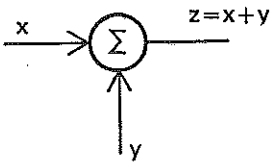


Fig. 10.1. Blockschema för addition.

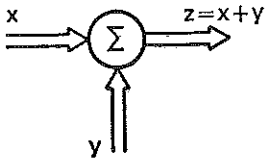


Fig. 10.2. Blockschema för addition av vektorer.

Blockschema

För representation av dynamiska system med blockschema införs symboler svarande mot addition, förgrening, integration och dynamiska eller statiska samband. Dessa ansluter sig nära till de operationer, dvs. addition, multiplikation med konstant och integration, som kan utföras av en analogimaskin. Ett blockschema kan därför som regel direkt läggas till grund för lösning av ett problem på analogimaskin.

Summation

För att representera summation av två variabler, dvs. $z=x+y$, används den symbol som visas i fig. 10.1. Om storheterna x och y är vektorer av samma dimension och man speciellt vill betona detta används en symbol enligt fig. 10.2. Om flera variabler summeras ritas flera pilar till summationssymbolen. I amerikansk litteratur förekommer även de symboler som visas i fig. 10.3, där ett tecken invid pilarna anger addition eller subtraktion. I denna bok skall vi genomgående använda rena additionssymboler enligt fig. 10.1 och 10.2. Om vi önskar subtrahera en variabel från en annan kan vi alltid göra detta genom att multiplicera den ena med -1 och addera.

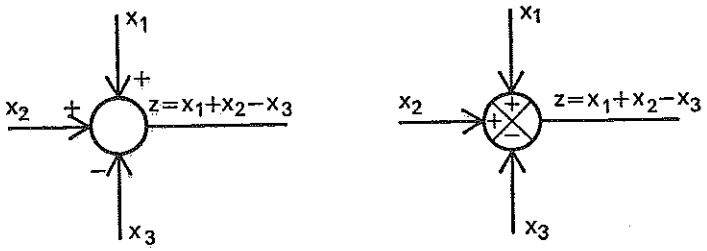


Fig. 10.3. Alternativa blockscheman för addition och subtraktion.

Förgrening

För att beteckna att en signal går flera vägar används de symboler som visas i fig. 10.4.

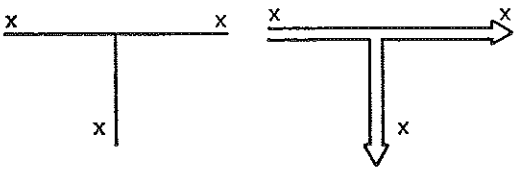


Fig. 10.4. Blockschema för förgrening.

Multiplikation med konstant

Den symbol som visas i fig. 10.5 anger att utsignalen y erhålls genom multiplikation av insignalen u med en konstant a. Observera att symbolen även anger ett kausalt samband som indikeras av pilens riktning. Utsignalen y orsakas således av insignalen u. Blocket i fig. 10.5 är alltså ej reversibelt.

Fig. 10.5. Blockschema för multiplikation med konstant.

Dynamiska samband och integration

För att beteckna att y är utsignalen till ett dynamiskt system $\mathcal{S}$ med insignal u och överföringsfunktion $G(s)$ används symbolerna i fig. 10.6. Specialfallet integration har en egen symbol (fig. 10.7). Om man speciellt vill markera att storheterna u och y är vektorer används symbolerna i fig. 10.8.

Vi illustrerar användningen av blockschemasymboler med några exempel.

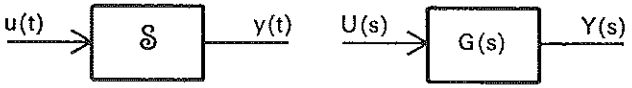


Fig. 10.6. Blockschema för dynamiska system.

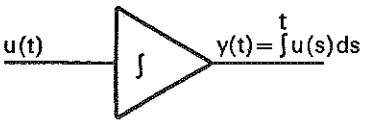


Fig. 10.7. Blockschema för integration.

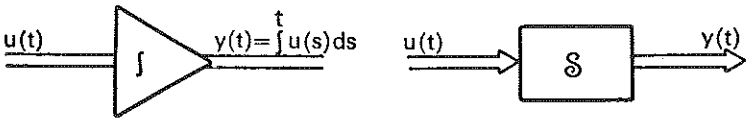


Fig. 10.8. Blockschema för dynamiska system då man speciellt vill markera att insignalerna och utsignalerna är vektorer.

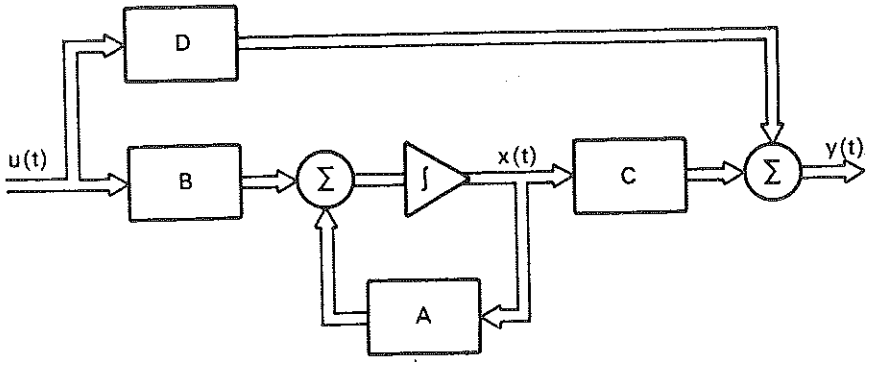


Fig. 10.9. Blockschema för systemet $\mathcal{S}(A, B, C, D)$.

EXEMPEL 10.1

Det linjära tidsinvarianta dynamiska systemet $\mathcal{S}(A, B, C, D)$ kan representeras med blockschemat i fig. 10.9, eller om endast insignal–utsignalrelationen är av intresse med blockschemat i fig. 10.6 med $G(s) = D + C[sI - A]^{-1}B$.

EXEMPEL 10.2. Systemet $\mathcal{S}(A, B, C, D)$ i observerbar kanonisk form

Systemet

$$\frac{dx}{dt} = \begin{bmatrix} -a_1 & 1 & 0 \dots 0 \\ -a_2 & 0 & 1 \dots 0 \\ \vdots & & \\ -a_{n-1} & 0 & 0 \dots 1 \\ -a_n & 0 & 0 \dots 0 \end{bmatrix} x + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} u$$

$$y = [1 \quad 0 \dots 0]\, x + Du$$

kan representeras med blockschemat i fig. 10.10.

EXEMPEL 10.3. Systemet $\mathcal{S}(A, B, C, D)$ i styrbar kanonisk form

Systemet

$$\frac{dx}{dt} = \begin{bmatrix} -a_1 & -a_2 \dots & -a_{n-1} & -a_n \\ 1 & 0 \dots & 0 & 0 \\ 0 & 1 \dots & 0 & 0 \\ \vdots & & & \\ 0 & 0 \dots & 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u$$

$$y = [b_1 \quad b_2 \dots b_n]\, x + Du$$

kan representeras med blockschemat i fig. 10.11.

Observera att systemen i exemplen 10.2 och 10.3 har samma överföringsfunktion, nämligen

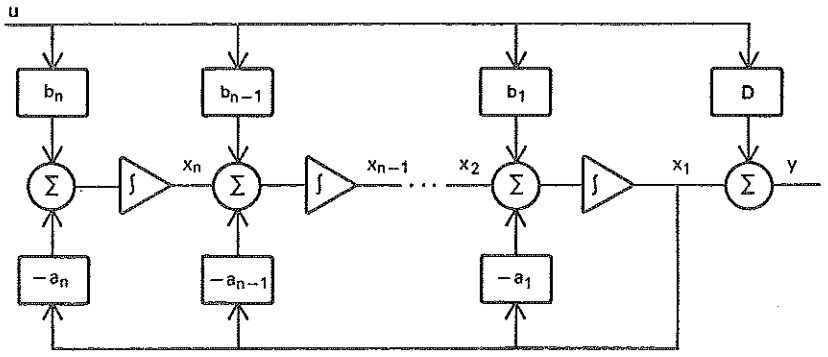


Fig. 10.10. Blockschema för systemet $\mathcal{S}(A, B, C, D)$ i observerbar kanonisk form. Systemet har överföringsfunktionen

$$G(s) = D + \frac{b_1 s^{n-1} + b_2 s^{n-2} + \cdots + b_n}{s^n + a_1 s^{n-1} + \cdots + a_n}$$

$$G(s) = D + \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + \ldots + a_n}$$

Lägg även märke till likheterna mellan fig. 10.10 och 10.11. Blockschemat i fig. 10.10 kan erhållas ur blockschemat i fig. 10.11 genom att pilarnas riktning omkastas, insignal byts mot utsignal i varje block, förgreningspunkter byts mot summationspunkter och summationspunkter mot förgreningspunkter.

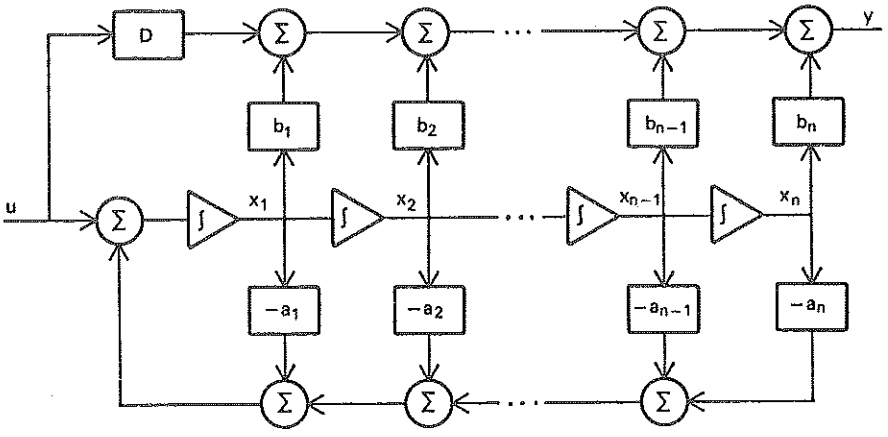


Fig. 10.11. Blockschema för det dynamiska systemet $\mathcal{S}(A, B, C, D)$ i styrbar kanonisk form. Systemet har överföringsfunktionen

$$G(s) = D + \frac{b_1 s^{n-1} + b_2 s^{n-2} + \cdots + b_n}{s^n + a_1 s^{n-1} + \cdots + a_n}$$

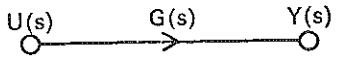


Fig. 10.12. Signalflödesdiagram för ett dynamiskt system med insignal u, utsignal y och överföringsfunktion $G(s)$.

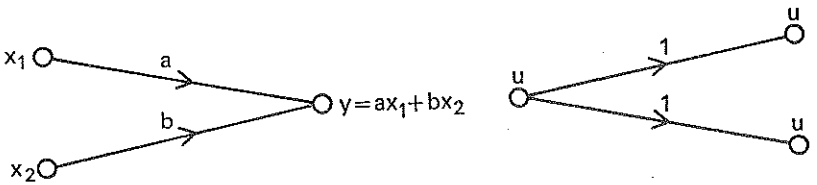


Fig. 10.13. Signalflödesdiagram för addition och förgrening.

Signalflödesdiagram

Blockschema kan ibland bli ganska otympliga. Man kan då i stället använda signalflödesdiagram. I ett sådant representeras systemvariablerna med *noder*. Variabler som påverkar varandra förbinds med en linje. För att ange kausalitet används en pil. Invid linjen anges det samband som råder mellan variablerna operatorform. Sambandet

$$Y(s) = G(s)\,U(s)$$

representeras således med symbolen i fig. 10.12. Observera skillnaderna med blockschema. I blockschema representeras dynamiska system med block och signalerna med streck. I signalflödesdiagram representeras dynamiska system med streck och variabler med prickar eller ringar. Addition och förgrening representeras enligt fig. 10.13.

Några exempel på signalflödesdiagram ges i fig. 10.14, fig. 10.15 och fig. 10.16.

Både för blockschema och signalflödesdiagram har en algebra utvecklats, så att man direkt kan manipulera diagrammen på olika sätt, t. ex. förenkla komplicerade signalflödesdiagram, finna samband mellan två godtyckliga variabler i ett diagram etc. Då alla dessa frågor även kan besvaras genom ett direkt studium av systemekvationerna, avstår vi här från att gå vidare in på dessa frågor utan hänvisar till litteraturen.

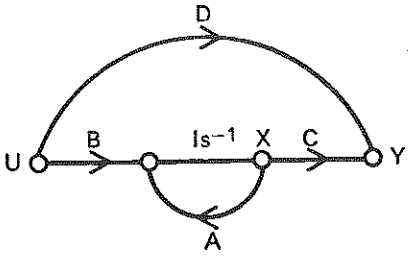


Fig. 10.14. Signalflödesdiagram för systemet $S(A, B, C, D)$. Jämför med blockschemat för samma system i fig. 10.9.

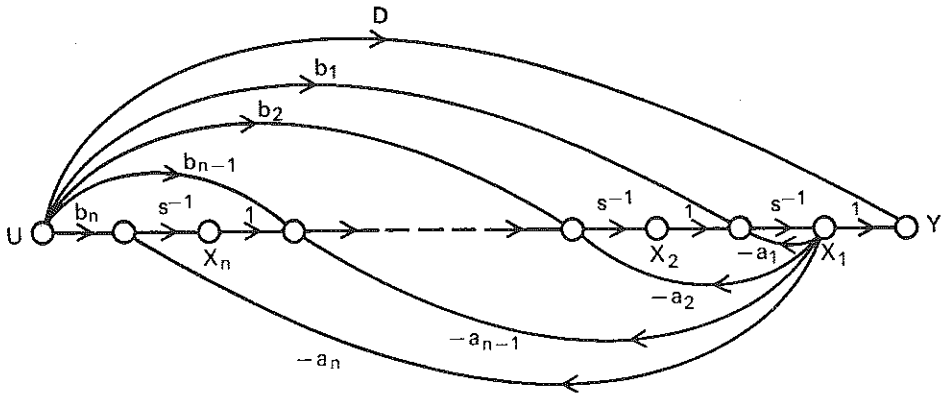


Fig. 10.15. Signalflödesdiagram för systemet $\mathcal{S}(A, B, C, D)$ i observerbar kanonisk form. Jämför med blockschemat för samma system i fig. 10.10.

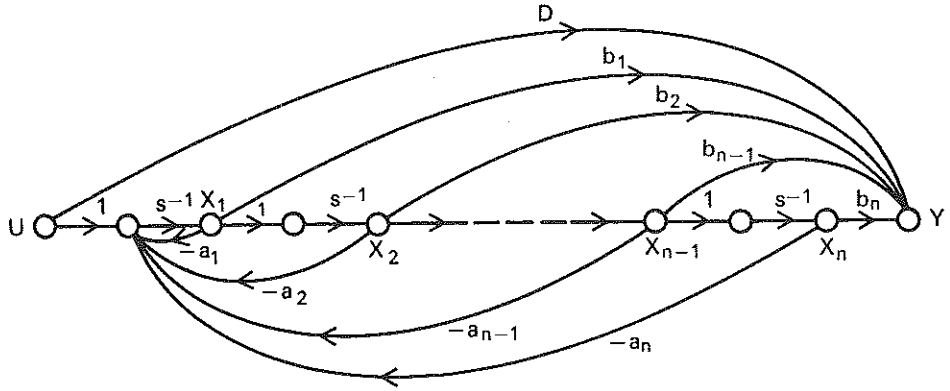


Fig. 10.16. Signalflödesdiagram för systemet $\mathcal{S}(A, B, C, D)$ i styrbar kanonisk form. Jämför med blockschemat för samma system i fig. 10.11.

11. Elementär stabilitetsteori för linjära tidsinvarianta system

Definitioner

Vid många tillämpningar är det väsentligt att veta under vilka villkor samtliga lösningar till den homogena ekvationen

$$\frac{dx}{dt} = Ax \tag{11.1}$$

konvergerar mot noll. Om så är fallet säges denna ekvation vara *stabil*. Annars är ekvationen *instabil*. Man säger även att det system som beskrivs av (11.1) är stabilt resp. instabilt. Vi skall senare i kap. 5 grundligt motivera detta.

Om matrisen A har skilda egenvärden, så finns alltid en reguljär matris T sådan att

$$\exp At = T^{-1}(\exp \Lambda t)T$$

där Λ är en diagonalmatris av A:s egenvärden. Då A har skilda egenvärden, finner vi således att ett nödvändigt och tillräckligt villkor för stabilitet är att alla egenvärden har negativ realdel. Vi skall nu visa att detta gäller även då A har multipla egenvärden.

För att visa detta använder vi det resultat från matristeorin som säger att en godtycklig matris A alltid kan transformeras på triangelform, dvs. det finns en reguljär matris T så att

$$TAT^{-1} = \begin{bmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & \lambda_2 & \dots & b_{2n} \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

där λ_i är A:s egenvärden.

Lösningen till ekvation (11.1) kan då skrivas

$$x(t) = T^{-1}z(t)$$

där

$$\frac{dz}{dt} = \begin{bmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & \lambda_2 & \dots & b_{2n} \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} z$$

I detta ekvationssystem kan vi integrera ekvationerna successivt från slutet. Vi finner då att om egenvärdena λ_i är skilda så blir z_k en summa av exponentialfunktionen $\exp(\lambda_i t)$. Varje sådan term kallas en *mod* till systemet. Om matrisen A däremot har multipla egenvärden så kan lösningen innehålla termer som är produkter av polynom och exponentialfunktioner. Ett egenvärde av multipliciteten k ger således upphov till termer av typen $P_{k-1}(t)\exp(\lambda_k t)$ där $P_l(t)$ är ett polynom av graden l eller lägre.

Alternativt kan vi härleda detta resultat ur Laplacetransformen

$$\mathcal{L}\{\exp At\} = [sI - A]^{-1} \tag{11.2}$$

Varje element i matrisen (11.2) har nämnaren $\det[sI - A]$, dvs. det karakteristiska polynomet till A. Med användande av inverstransformen finner vi då att enkla nollställen till den karakteristiska ekvationen ger upphov till termer av typen $\exp(\lambda_k t)$. Multipla nollställen till den karakteristiska ekvationen motsvaras av termer av typen

$$P_{k-1}(t)\, e^{\lambda_k t} \tag{11.3}$$

Varje lösning till (11.1) har således komponenter av typen (11.3). Då exponentialfunktionen växer snabbare än varje potens följer nu resultatet. Vi sammanfattar resultatet som

SATS 11.1

Ett nödvändigt och tillräckligt villkor för att ekvation (11.1) skall vara stabil är att samtliga egenvärden till A har negativ realdel. Varje komponent $x_i(t)$ av en lösning $x(t)$ till (11.1) har egenskapen

$$|x_i(t)| < K \cdot e^{\alpha t} \tag{11.4}$$

där

$$\alpha > \max_i \operatorname{Re} \lambda_i(A)$$

Det första påståendet har visats direkt och det andra påståendet följer av det gjorda resonemanget.

Stabilitetskriterier i enkla fall

Vi har infört ett stabilitetsbegrepp och genom sats 11.1 visat att stabilitetsundersökningen reduceras på ett rent algebraiskt problem, nämligen att undersöka om samtliga rötter till den karakteristiska ekvationen $\det [sI - A] = 0$ har negativ realdel. Vi skall nu ge stabilitetskriterier i enkla fall.

EXEMPEL 11.1. Första ordningens system

Betrakta ett första ordningens system. Den karakteristiska ekvationen

$$\det [sI - A] = 0 \tag{11.5}$$

kan då skrivas

$$s + a_1 = 0 \tag{11.6}$$

Stabilitetskriteriet är uppenbarligen $a_1 > 0$.

EXEMPEL 11.2. Andra ordningens system

Betrakta ett system av andra ordningen. Den karakteristiska ekvationen (11.5) kan skrivas

$$s^2 + a_1 s + a_2 = 0 \tag{11.7}$$

Vi finner genom direkt lösning att stabilitetsvillkoret är $a_1 > 0$ och $a_2 > 0$.

EXEMPEL 11.3. Tredje ordningens system

Betrakta ett system av tredje ordningen. Den karakteristiska ekvationen (11.5) kan skrivas

$$s^3 + a_1 s^2 + a_2 s + a_3 = 0 \tag{11.8}$$

Då en tredjegradsekvation med reella koefficienter alltid har en reell rot kan ekvationen alltid faktoreras som

$$(s+\alpha)(s^2+\beta s+\gamma)=0 \tag{11.9}$$

där

$$\begin{aligned} a_1&=\alpha+\beta \\ a_2&=\gamma+\alpha\beta \\ a_3&=\alpha\gamma \end{aligned} \tag{11.10}$$

Enligt exemplen 11.1 och 11.2 kan stabilitetsvillkoret för (11.9) skrivas som

$$\begin{aligned} \alpha&>0 \\ \beta&>0 \\ \gamma&>0 \end{aligned} \tag{11.11}$$

För att uttrycka stabilitetsvillkoret för (11.9) med hjälp av a_1, a_2 och a_3 behöver vi således endast undersöka hur området (11.11) i $\alpha\beta\gamma$-rummet avbildas i $a_1a_2a_3$-rummet vid transformationen (11.10).

Vi finner

$\alpha\beta$-planet $\gamma=0 \Rightarrow a_1=\alpha+\beta$, $a_2=\alpha\beta$ och $a_3=0$.

Avbildas således på a_1a_2-planet.

$\beta\gamma$-planet $\alpha=0 \Rightarrow a_1=\beta$, $a_2=\gamma$ och $a_3=0$.

Avbildas således på a_1a_2-planet.

$\alpha\gamma$-planet $\beta=0 \Rightarrow a_1=\alpha$, $a_2=\gamma$ och $a_3=\alpha\gamma$.

Avbildas således på ytan $a_3=a_1a_2$.

Den positiva oktanten i $\alpha\beta\gamma$-rummet vid transformationen (11.10) avbildas alltså på en volym i positiva oktanten i $a_1a_2a_3$-rummet karakteriserad av

$$\begin{aligned} &0<a_3<a_1a_2 \\ &a_1,\ a_2>0 \end{aligned} \tag{11.12}$$

och vi finner således att olikheten (11.12) är ett nödvändigt och tillräckligt villkor för att ekvationen (11.8) skall ha alla rötter i vänstra halvplanet.

12. Orientering om den allmänna teorin för dynamiska system

Vi har sett exempel på hur teorin för dynamiska system kan utvecklas i specialfallet linjära tidsinvarianta system av ändlig ordning. I detta avsnitt skall vi ge en mycket kortfattad introduktion till den allmänna teorin för dynamiska system. Den generella teorin är kanske ej så matnyttig för den praktiskt inriktade, men den ger ett allmänt betraktelsesätt, som gör det möjligt att se sammanhangen mellan problem som vid ett ytligt betraktande kan förefalla mycket olika. Den generella teorin kan även ge uppslag till lösningsmetoder för problem som ej tillhör standardkategorin. För att penetrera den generella teorin fordras kunskaper i matematik som går utöver fordringarna för godkänt i civilingenjörsexamen.

Abstrakt definition av dynamiska system

Ett dynamiskt system $\mathcal{S}$ är en matematisk struktur som definieras av följande axiom:

A1 Det är givet ett topologiskt rum $\mathcal{X}$, tillståndsrummet, och en mängd av tider $\mathcal{T}$, där $\mathcal{T}$ är en ordnad halvgrupp av reella tal.

A2 Det är givet ett topologiskt rum $\mathcal{W}$, av funktioner definierade på $\mathcal{T}$ som tar värden i det topologiska rummet $\mathcal{U}$. Vidare är givet ett topologiskt rum $\mathcal{Y}$, där utsignalerna tar sina värden. Ett insignalsegment $u_{[t_1, t_2)}$ är en avbildning $[t_1, t_2) \to u$, $u \in \mathcal{U}$ där $[t_1, t_2)$ är en delmängd av $\mathcal{T}$ som innehåller alla element som uppfyller ordningsrelationen $t_1 \leqslant t < t_2$.

A3 För varje tillstånd $x^0 \in \mathcal{X}$, varje tidpunkt $t_0 \in \mathcal{T}$ och varje insignalsegment $u_{[t_0, t)}$ är systemets tillstånd vid tidpunkten $t \in \mathcal{T}$ bestämd av övergångsfunktionen φ, vilken avbildar produktrummet $\mathcal{T} \times \mathcal{T} \times \mathcal{X} \times \mathcal{W}$ i tillståndsrummet $\mathcal{X}$, dvs.

$$x(t) = \varphi(t;\, t_0,\, x^0,\, u_{[t_0, t)})$$

Funktionen φ skall vara sådan att

$$x^0 = \varphi(t_0;\, t_0,\, x^0,\, u_{[t_0, t_0)}) \tag{12.1}$$

$$\varphi(t_2;\, t_0,\, x^0,\, u_{[t_0, t_1)}) = \varphi(t_2;\, t_1,\, \varphi(t_1;\, t_0,\, x^0,\, u_{[t_0, t_1)}),\, u_{[t_1, t_1)}) \tag{12.2}$$

Vid varje tidpunkt $t \in \mathcal{T}$, varje tillstånd $x \in \mathcal{X}$ och varje värde av insignalen $u(t) \in \mathcal{U}$ gäller att utsignalen y definieras av funktionen ψ som avbildar $\mathcal{X} \times \mathcal{U} \times \mathcal{T}$ på $\mathcal{Y}$, dvs.

$$y(t) = \psi(x,\, u(t),\, t)$$

A4 Funktionerna φ och ψ är kontinuerliga med avseende på de topologier, som definierats för $\mathcal{T}$, $\mathcal{X}$ och $\mathcal{W}$ och de därav inducerade topologierna i produktrummen.

Vi skall beteckna ett dynamiskt system $\mathcal{S}$ med sju-tipeln $\mathcal{S} = \mathcal{S}(\mathcal{X}, \mathcal{T}, \mathcal{U}, \mathcal{W}, \mathcal{Y}, \varphi, \psi)$. Mängden $\mathcal{T}$ är vanligen mängden av reella tal eller mängden av heltal, systemen kallas då *tidskontinuerliga* resp. *tidsdiskreta*.

Ett system säges vara *tidsinvariant* om $u(t) = \tilde{u}(t+\tau)$, för varje $\tau \in \mathcal{T}$ medför att

$$\varphi(t; t_0, x^0, u_{[t_0, t)}) = \varphi(t+\tau; t_0+\tau, x^0, \tilde{u}_{[t_0+\tau, t+\tau)}) \tag{12.3}$$

$$\psi(x, u, t) = \psi(x, u, t+\tau) \tag{12.4}$$

för varje $t \in \mathcal{T}$, $t_0 \in \mathcal{T}$, $x^0 \in \mathcal{X}$, $x \in \mathcal{X}$ och $u \in \mathcal{U}$.

Ett system $\mathcal{S}$ är *linjärt* om funktionerna φ och ψ är linjära i x och u. Ett system $\mathcal{S}$ är *ändligt dimensionellt* om rummen $\mathcal{X}$, $\mathcal{U}$ och $\mathcal{Y}$ har ändlig dimension. Ett system $\mathcal{S}$ är en *ändlig automaton* (finite automata, finite state machine) om rummen $\mathcal{X}$, $\mathcal{U}$ och $\mathcal{Y}$ är ändliga mängder och rummet $\mathcal{W}$ är mängden av heltal. Detta visar sammanhanget mellan automatateori och reglerteknik. En väsentlig skillnad är att i automatateori betonas algebraiska och kombinatoriska metoder medan reglerteorin betonar analytiska metoder.

EXEMPEL 12.1. Linjära system

För att illustrera axiomen skall vi undersöka deras innebörd i det specialfall då rummen $\mathcal{X}$, $\mathcal{U}$, $\mathcal{Y}$ och $\mathcal{T}$ är mängder av reella tal $\mathcal{R}$ och $\mathcal{W} = \mathcal{C}$ är mängden av kontinuerliga funktioner på reella talen. Vi skall vidare förutsätta att systemen är linjära. Den allmänna linjära funktionen från $\mathcal{R}$ till $\mathcal{R}$ är Ax. Enligt en sats av F. Riesz (se t. ex. F. Riesz & B. Sz.-Nagy, *Leçon d'Analyse Fonctionelle* s. 110) kan varje linjär funktion från $\mathcal{C}$ till $\mathcal{R}$ skrivas som en Stieltjesintegral. Vi finner då att funktionerna φ och ψ har formen

$$\varphi(t; t_0, x^0, u_{[t_0, t)}) = f(t; t_0)\, x^0 + \int_{t_0}^{t} dH(t, s)\, u(s) \tag{12.5}$$

$$\psi(x, u, t) = C(t)\, x + D(t)\, u \tag{12.6}$$

Använd nu egenskapen (12.2) i A3 och vi finner

$$\varphi(t; t_0, x^0, u_{[t_0, t)}) = \varphi(t; t_1, \varphi(t_1; t_0, x^0, u_{[t_0, t)}), u_{[t_1, t)})$$

$$= f(t; t_1) \left[f(t_1; t_0)\, x^0 + \int_{t_0}^{t_1} dH(t_1, s)\, u(s) \right] + \int_{t_1}^{t} dH(t, s)\, u(s)$$

Identifiering med (12.5) ger

$$f(t; t_0) = f(t; t_1) f(t_1; t_0) \tag{12.7}$$

$$\int_{t_0}^{t_1} f(t; t_1)\, dH(t_1, s)\, u(s) = \int_{t_0}^{t_1} dH(t, s)\, u(s) \tag{12.8}$$

Då (12.8) gäller för alla $u \in \mathcal{U}$ finner vi

$$dH(t, s) = f(t; t_1)\, dH(t_1, s) \tag{12.9}$$

Sätt $t_1 = s$ och vi erhåller

$$dH(t, s) = f(t; s)\,dH(s, s) = f(t; s)\,dH(s) \tag{12.10}$$

Insätts ekvation (12.5) i ekvation (12.1) finner vi således

$$x(t) = f(t; t_0)\,x^0 + \int_{t_0}^{t} f(t; s)\,dH(s)\,u(s) \tag{12.11}$$

Enligt (12.1) gäller vidare

$$f(t; t) = 1 \tag{12.12}$$

Det följer således av axiomen att ett endimensionellt linjärt system kan beskrivas med ekvationerna (12.6) och (12.11) där funktionen f satisfierar funktionalekvationen (12.7) med bivillkoret (12.12).

Vi skall nu göra den ytterligare förutsättningen att systemet är tidsinvariant. Det följer då av ekvation (12.3) att $f(t; t_0) = f(t+\tau; t_0+\tau)$ och $dH(s) = dH(s+\tau)$ dvs.

$$f(t; t_0) = f(t - t_0)$$

$$H(t) = Bt + \text{konstant}$$

Funktionalekvationen (12.7) med bivillkoret (12.12) reduceras då till

$$f(s+t) = f(s)\,f(t)$$

$$f(0) = 1 \tag{12.13}$$

Enligt axiom A4 är funktionen φ kontinuerlig i t. Funktionen f är då också kontinuerlig. Vi skall nu visa att f även är deriverbar. Vi observerar först att f är integrerbar över ett intervall $(0, h)$ och att

$$\lim_{h \downarrow 0} \frac{1}{h} \int_0^h f(t)\,dt = 1$$

För t litet men positivt gäller då att

$$k = \int_0^t f(s)\,ds > 0$$

Bilda nu

$$\frac{1}{h}\,[f(h) - f(0)] = \frac{1}{h}\,[f(h) - 1] = \frac{1}{hk}\,[f(h) - 1] \int_0^t f(s)\,ds$$

$$= \frac{1}{hk}\left[\int_0^t f(s)\,f(h)\,ds - \int_0^t f(s)\,ds\right] = \frac{1}{hk}\left[\int_0^t f(s+h)\,ds - \int_0^t f(s)\,ds\right]$$

$$= \frac{1}{hk}\left[\int_h^{t+h} f(s)\,ds - \int_0^t f(s)\,ds\right] = \frac{1}{hk}\left[\int_t^{t+h} f(s)\,ds - \int_0^h f(s)\,ds\right]$$

Men då f är kontinuerlig gäller

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} f(s)\, ds = f(t)$$

och vi finner således att gränsvärdet

$$f'(0) = \lim_{h \downarrow 0} \frac{1}{h} [f(h) - f(0)]$$

existerar och att

$$f'(0) = \frac{1}{k} [f(t) - 1]$$

dvs.

$$f(t) = 1 + kf'(0) = 1 + A \int_0^t f(s)\, ds \tag{12.14}$$

där $A = f'(0)$. Vidare gäller

$$f(t+h) - f(t) = [f(h) - 1] f(t)$$

Funktionen f är således deriverbar för alla t. Derivation av (12.14) ger

$$\frac{df}{dt} = Af \tag{12.15}$$

Vi finner således att för linjära, tidsinvarianta, endimensionella dynamiska system gäller

$$\frac{dx}{dt} = \frac{\partial f(t; t_0)}{\partial t} x_0 + \int_{t_0}^t \frac{\partial f(t; s)}{\partial t} Bu(s)\, ds + f(t; t)\, Bu(t)$$

dvs.

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) \tag{12.16}$$

Då systemet är tidsinvariant gäller vidare, enligt definition, ekvation (12.4). Talen C och D i ekvation (12.6) är då konstanter, dvs.

$$y = Cx + Du \tag{12.17}$$

Vi har således visat att för linjära tidsinvarianta system med en insignal och en utsignal är den abstrakta definitionen av dynamiska system ekvivalent med den speciella definition som gavs i avsnitt 3. Resonemanget kan omedelbart generaliseras till det fall då $\mathcal{R}$, $\mathcal{U}$ och $\mathcal{Y}$ är euklidiska rum av godtycklig dimension.

Observera att vi även har visat att en kontinuerlig funktion som satisfierar funktionalekvationen (12.13) är en exponentialfunktion. Jämför avsnitt 5.

13. Referenser

För ett vidare studium av dynamiska system rekommenderas i första hand någon av nedanstående läroböcker

Zadeh, L. A. & Desoer, C. A., *Linear Systems Theory—The State Space Approach*, McGraw-Hill, New York 1963

Porter, W. A., *Modern Foundations of Systems Engineering*, The Macmillan Company, New York 1966

Koenig, H. E. m. fl., *Analysis of Discrete Physical Systems*, McGraw-Hill, New York 1967.

Zadeh & Desoers bok är mycket läsvärd. Den ger en mycket grundlig framställning av ändligt dimensionella system. Porters bok är mer speciellt inriktad och innehåller flera rent matematiska avsnitt, t. ex. en översikt av fundamenta av funktionalanalysen. Både Zadehs och Porters böcker fordrar matematiska kunskaper, som delvis går utöver vad som krävs för godkänt i civilingenjörsexamen. Koenigs bok har starkare praktisk anknytning och kräver mindre matematiska förkunskaper. Den innehåller även en framställning om hur man ställer upp matematiska modeller för elektriska, hydrauliska och pneumatiska system. Uppställning av matematiska modeller för dynamiska system presenteras även utmärkt i

Shearer, L. L., Murphy, A. I. & Richardson, H. H., *Introduction to System Dynamics*, Addison-Wesley, Reading, Mass. 1967.

Dynamiska system som beskrivs av differential-differensekvationer behandlas i

Bellman R. & Cooke, K. L., *Differential-Difference Equations*, Academic Press, New York 1963.

System beskrivna av partiella differentialekvationer är ej tillnärmelsevis fullständigt genomarbetade. Vissa resultat finns i

Leondes, C. T. (red.), *Advances in Control Systems—Theory and Applications*, Vol. 1, kap. 3, Control of Distributed Parameter Systems, Academic Press, New York 1964.

För en detaljerad framställning av transformteorin och dess tillämpning på linjära system hänvisas till

Gardner, M. F. & Barnes, J. L., *Transients in Linear Systems*, Vol. 1 & 2, John Wiley, New York 1942

Chestnut, H. & Mayer, R. W., *Servomechanisms and Regulation Systems Design*, Vol. 1, John Wiley, New York 1951

Truxal, J. G., *Automatic Feedback Control System Synthesis*, McGraw-Hill, New York 1955.

Alla dessa läroböcker är omsorgsfullt skrivna och väl genomarbetade.

För teorin för Laplacetransformen hänvisas till standardverken

Widder, D. V., *The Laplace Transform*, Princeton University Press, Princeton, New Jersey 1941

Doetsch, G., *Handbuch der Laplace Transformation*, 3 delar, Basel 1950–1960.

Widder ger bl. a. fullständig genomgång av begynnelse- och slutvärdessatserna, s. k. Abel- och Taubersatser. En mycket trevlig framställning av Taubersatser har även givits i

Feller, W., *An Introduction to Probability Theory and its Applications*, Vol. II, kapitel 13 och 14, John Wiley, New York 1966.

För övningsexempel i Laplacetransform hänvisas till

Spiegel, M. R., *Theory and Problems of Laplace Transforms*, Schaum's Outline Series, New York 1965.

För att undvika formella manipulationer med dirac-funktioner, får man tillgripa distributionsteori, generaliserade funktioner eller Mikusinskis operatorkalkyl. En lättfattlig översikt har givits i

Beckenbach, E. F. (red.), *Modern Mathematics for the Engineer*, Second Series kap. 1, From Delta Functions to Distributions, A., McGraw-Hill, New York 1961.

I denna artikel finns referenser till originalarbeten av Schwartz, Lighthill och Mikusinski.

Vid praktiskt räknearbete med Laplacetransformer har man stor nytta av tabellverken

Magnus, W. & Oberhettinger, F., *Formulas and Theorems for the Special Functions of Mathematical Physics*, Chelsea, New York 1949

Erdelyi, A. (red.), *Tables of Integral Transforms*, Vol. I, McGraw-Hill, New York 1954.

Numerisk inversion av Laplacetransformer har behandlats av

Bellman, R. m. fl., *Numerical Inversion of the Laplace Transform*, Elsevier, New York 1966.

För teorin för ordinära differentialekvationer hänvisas till standardverk såsom

Coddington, E. A. & Levinson, N., *Theory of Ordinary Differential Equations*, McGraw-Hill, New York 1955

Lefschetz, S., *Differential Equations—Geometric Theory*, John Wiley, New York 1963.

Matristeori kan lämpligen studeras i

Gantmacher, F. R., *The Theory of Matrices*, Vol. 1 & 2, Chelsea, New York 1959.

För en härledning av Bodes relation och ytterligare behandling av icke minimum-fas system hänvisas till

Bode, H. W., *Network Analysis and Feedback Amplifier Design*, D. Van Nostrand, New Jersey 1945.

För den abstrakta teorin för dynamiska system som ej är styrda, refereras till

Birkhoff, G. D., *Dynamical Systems*, Am. Math. Soc. 1927

Nemytskii, V. V. & Stepanov, V. V., *Qualitative Theory of Differential Equations*, Princeton University Press, Princeton, N.J. 1960.

En diskussion om sambanden mellan reglerteori och automatateori finns i artikeln

Arbib, M. A., *A Common Framework for Automata Theory and Control Theory*, S.I.A.M. Journal on Control *3* (1965) 206–222.

Där ges även ytterligare referenser.

3. Styrbarhet och observerbarhet

1. Inledning

I detta kapitel skall vi behandla två fundamentala frågor i anslutning till styrda dynamiska system: Kan ett dynamiskt system alltid styras från ett givet tillstånd till ett annat? Om insignalens och utsignalens värden är kända i ett intervall $(0, T)$ kan då systemets tillstånd vid tidpunkten T bestämmas?

Svaren på dessa frågor anger i vilken utsträckning systemet går att styra samt vilken information utsignalerna ger om systemets tillstånd. I ett konkret fall kan vi t. ex. avgöra om antalet reglerventiler är tillräckligt för att styra en process eller om de tillgängliga instrumenten ger fullständig information om processens tillstånd. Det är förvånande att dessa fundamentala frågor formulerats och besvarats relativt sent. De villkor som delvis besvarar frågorna uppträdde först som förutsättningar för att vissa formella beräkningar skulle gå att genomföra. Jämför avsnitt 2:6. Problemen formulerades och löstes först av Kalman (1960), som även införde begreppen styrbarhet (controllability) och observerbarhet (observability). Ibland används ordet kontrollerbarhet i stället för styrbarhet.

Avsnitt 2 i detta kapitel behandlar ett enkelt exempel, som belyser frågeställningarna i ett specialfall. Formella definitioner av styrbarhet och observerbarhet ges i avsnitt 3. I de två följande avsnitten ges kriterier för styrbarhet och observerbarhet för linjära, tidsinvarianta dynamiska system. I avsnitt 6 ges ett viktigt resultat som visar att ett linjärt tidsinvariant system kan uppdelas i delsystem utgående från styrbarhets- och observerbarhetsegenskaper. Kriterierna för styrbarhet och observerbarhet är snarlika. I själva verket är styrbarhet och observerbarhet matematiska dualer. Detta visas i avsnitt 7. I avsnitt 8 ges en översikt av olika sätt att representera det dynamiska systemet $\mathcal{S}(A, B, C, D)$. Begreppen observerbarhet och styrbarhet gör det möjligt att klart inse skillnaden mellan att representera ett dynamiskt system med hjälp av tillståndsekvationer respektive överföringsfunktioner. Hur man skall konstruera en tillståndsekvation av lägsta möjliga ordningstal som har en given överföringsfunktion behandlas i avsnitt 9.

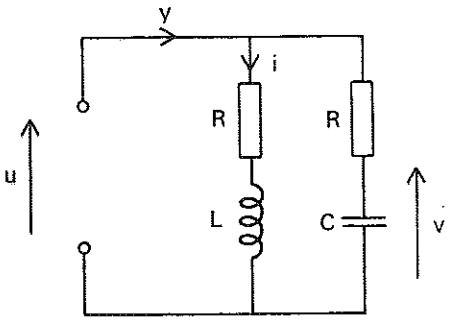


Fig. 2.1. Parallellresonanskrets.

2. Ett exempel

Betrakta det elektriska nät som visas i fig. 2.1.

Låt insignalen vara polspänningen u och utsignalen y strömmen i huvudgrenen. Anta vidare att $R=1$, $L=1$ och $C=1$. Inför strömmen genom induktansen, i, och spänningen över kapacitansen, v, och vi får följande ekvationer

$$u = Ri + L\frac{di}{dt}$$

$$u = v + RC\frac{dv}{dt}$$

$$y = i + C\frac{dv}{dt} = i - \frac{v}{R} + \frac{u}{R}$$

Inför tillståndsvariablerna

$$x_1 = i$$

$$x_2 = v$$

och systemekvationerna kan skrivas

$$\frac{dx_1}{dt} = -x_1 + u$$

$$\frac{dx_2}{dt} = -x_2 + u$$

$$y = x_1 - x_2 + u$$

eller

$$\frac{dx}{dt} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad -1]\, x + u$$

Vi finner således att systemet är ett stationärt dynamiskt system av andra ordningen med en insignal och en utsignal. För att närmare undersöka systemets egenskaper införs nya koordinater i tillståndsrummet med transformationen

$$z_1 = x_1 - x_2$$

$$z_2 = x_1 + x_2$$

Vi finner då

$$\frac{dz_1}{dt} = -z_1$$

$$\frac{dz_2}{dt} = -z_2 + 2u$$

$$y = z_1 + u$$

Ovanstående ekvationer visar att styrsignalen u kan påverka tillståndsvariabeln z_2 men ej variabeln z_1 samt att utsignalen endast påverkas av variabeln z_1. I detta speciella exempel är det således ej möjligt att styra systemet till ett tillstånd med $z_1 \neq 0$ med hjälp av insignalen u. Vi säger att variabeln z_2 är *styrbar* medan variabeln z_1 ej är styrbar. Vidare ger utsignalen ingen som helst information om variabeln z_2. Vi säger att variabeln z_1 är *observerbar* men att variabeln z_2 ej är observerbar.

Låt oss även beräkna överföringsfunktionen för systemet. Vi finner då

$$G(s) = \frac{Y(s)}{U(s)} = \frac{Z_1(s) + U(s)}{U(s)} = 1$$

ty enligt definition 2:8.1 är överföringsfunktionen kvoten mellan utsignalens och insignalens Laplacetransformer då initialtillståndet är noll vilket medför $Z_1(s) = 0$. Överföringsfunktionen kan även beräknas direkt ur fig. 2.1. Vi finner

$$\frac{Y(s)}{U(s)} = \frac{1}{R + sL} + \frac{1}{R + \frac{1}{sC}} = \frac{1 + 2RCs + LCs^2}{(R + sL)(1 + RCs)} = \left(\frac{1+s}{s+1}\right)^2 = 1$$

Då systemet är av andra ordningen skulle man vänta sig att överföringsfunktionen är en kvot mellan två andragradspolynom. Trots detta finner vi att överföringsfunktionen är en konstant. Den första härledningen visar att detta faktum är en omedelbar konsekvens av att variabeln z_1 är observerbar men ej styrbar och att variabeln z_2 är styrbar men ej observerbar. I den senare kalkylen yttrar sig effekten därigenom att två poler och två nollställen i överföringsfunktionen kan förkortas bort.

3. Definitioner

De begrepp som infördes heuristiskt i föregående avsnitt skall nu definieras. Betrakta ett linjärt dynamiskt system. Anta att systemets tillstånd är x^0 vid tiden t_0 och att styrsignalen u definierad i intervallet $(0, t_1)$ medför att $x(t_1) = x^1$. Styrsignalen u sägs då *överföra* eller *styra* tillståndet x^0 till x^1. Vi kan nu införa

DEFINITION 3.1. Styrbarhet

Ett *tillstånd* x^0 är *styrbart* (controllable) om det finns ett ändligt $t_1 > 0$ och en styrsignal u i intervallet $(0, t_1)$ som överför tillståndet x^0 till origo. Ett system är styrbart om varje tillstånd är styrbart.

Observera att det är väsentligt att förutsätta t_1 ändligt, annars skulle tillståndet z_1 i exemplet i avsnitt 2 bli styrbart, vilket ej är önskvärt.

Det skall senare visas att om ett linjärt tidsinvariant system är styrbart så kan det styras från ett godtyckligt tillstånd till varje annat tillstånd. För teoribildningen är det dock bekvämt att standardisera så att definitionen görs t. ex. för det fall då systemet styrs från ett givet tillstånd till origo som i definition 3.1. I litteraturen förekommer också begreppet *uppnåeligt* tillstånd (reachable state) som relaterar till den situation då systemet styrs från origo till ett föreskrivet tillstånd. För linjära tidsinvarianta system är mängderna av styrbara och uppnåeliga tillstånd lika. Detta är dock ej nödvändigtvis sant för andra klasser av system, t. ex. linjära tidsvariabla system.

Begreppet styrbarhet uttrycker således möjligheten att styra ett system från ett tillstånd till ett annat. Vi skall nu införa ett annat begrepp *observerbarhet*, som uttrycker möjligheten att bestämma tillstånden ur mätningar av systemets in- och utsignaler.

För ett linjärt tidsinvariant system gäller

$$y(t) = Cx(t) + Du(t)$$

Enligt ekvation (2:4.8) gäller

$$x(t) = e^{At}x^0 + \int_0^t e^{A(t-s)}Bu(s)\,ds$$

De förändringar i tillstånd och utsignal som orsakas av insignalen kan således alltid beräknas. För att avgöra om det är möjligt att ur in- och utsignaler bestämma systemets tillstånd kan det därför förutsättas att $u(t) = 0$.

Av tekniska skäl är det bekvämt att först definiera icke observerbarhet.

DEFINITION 3.2. Observerbarhet

Ett tillstånd $x^0 \neq 0$ sägs vara *icke observerbart* (unobservable) eller *tyst* om det finns ett ändligt $t_1 > 0$ så att $y(t) = 0$ för $0 \leqslant t \leqslant t_1$ då $x(0) = x^0$ och $u(t) = 0$ för $0 \leqslant t \leqslant t_1$ Ett *system* är observerbart om det ej finns några tysta tillstånd.

Definitionen återspeglar klart den intuitiva innebörden att de icke observerbara tillstånden ej kan bestämmas ur observationer av in- och utsignalerna.

Ett alternativ till ovanstående definition är att säga att ett tillstånd är *icke rekonstruerbart* (non reconstructible, not determinable) om det finns ett ändligt t_1 så att $x(t_1) \neq 0$ även om $y(t) = 0$ och $u(t) = 0$ för $0 \leqslant t \leqslant t_1$. Det skall senare visas att för linjära tidsinvarianta system så är mängderna av tysta och icke rekonstruerbara tillstånd lika. Detta är ej sant för andra klasser av system, t. ex. linjära tidsvariabla system.

Observera att om tillstånden x^1 och x^2 är tysta så gäller

$$0 = Ce^{At}x^1$$

$$0 = Ce^{At}x^2$$

Då gäller också

$$0 = Ce^{At}[\alpha x^1 + \beta x^2]$$

Tillståndet $\alpha x^1 + \beta x^2$ är således också tyst. De tysta eller icke observerbara tillstånden bildar alltså ett linjärt underrum. Detta är motivet till att man först definierar tysta eller icke observerbara tillstånd.

4. Kriterier för styrbarhet

Kriterier som anger om ett tillstånd är styrbart skall nu härledas. Betrakta ett linjärt tidsinvariant system $\mathcal{S}(A, B, C, D)$. Enligt ekvation (2:4.8) kan lösningen till systemekvationerna skrivas

$$x(t_1) = e^{At_1}\left[x^0 + \int_0^{t_1} e^{-As}Bu(s)\,ds\right]$$

Tillståndet x^0 är således styrbart om det finns en styrsignal u sådan att ekvationen

$$x^0 = -\int_0^{t_1} e^{-As}Bu(s)\,ds \tag{4.1}$$

har en lösning för ändligt t_1. Vi kan således först observera att styrbarheten endast beror av matriserna A och B. Vi observerar vidare att om tillstånden x^1 och x^2 är styrbara så finns styrsignaler u^1 och u^2 sådana att

$$x^1 = -\int_0^{t_1} e^{-As}Bu^1(s)\,ds$$

$$x^2 = -\int_0^{t_1} e^{-As}Bu^2(s)\,ds$$

Då gäller också

$$\alpha x^1 + \beta x^2 = -\int_0^{t_1} e^{-As} B[\alpha u^1(s) + \beta u^2(s)]\, ds$$

Tillståndet $\alpha x^1 + \beta x^2$ är alltså styrbart. Vi finner således att de styrbara tillstånden bildar ett linjärt underrum av tillståndsrummet. För att karakterisera detta linjära underrum observerar vi att ekvation (4.1) kan skrivas på följande sätt

$$x^0 = -\int_0^{t_1} e^{-As} B u(s)\, ds = -\sum_{k=0}^{n-1} A^k B \int_0^{t_1} \alpha_k(s)\, u(s)\, ds \tag{4.2}$$

där funktionerna α_k är linjärt oberoende på intervallet $(0, t_1)$. Detta följer av en sats om matrisfunktioner. Man kan intuitivt inse att resultatet är rimligt genom att använda serieutvecklingen för $\exp(-As)$. En sats av *Cayley* och *Hamilton* säger att varje matris satisfierar sin egen karakteristiska ekvation, dvs.

$$A^n + a_i A^{n-1} + \ldots + a_{n-1} A + a_n I = 0$$

där a_i är koefficienterna i matrisen A:s karakteristiska polynom, dvs.

$$\det[\lambda I - A] = \lambda^n + a_1 \lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$$

Genom upprepad användning av denna sats finner vi att varje potens A^k med $k \geqslant n$ kan uttryckas som en linjärkombination av I, A, ..., A^{n-1}.

Ekvationen (4.2) kan således endast lösas om vektorn x^0 kan erhållas som en linjärkombination av kolonnerna i matriserna

$$B \;\; AB \;\; A^2B \ldots A^{n-1}B$$

Sammanfattningsvis finner vi

Sats 4.1

De styrbara tillstånden ligger i det linjära underrum av tillståndsrummet som spänns upp av kolonnerna i matrisen

$$W_s = [B \;\; AB \;\; A^2B \ldots A^{n-1}B] \tag{4.3}$$

Systemet $\mathcal{S}(A, B, C, D)$ är styrbart om och endast om W_s har rangen n.

Matrisen W_s kallas *styrbarhetsmatris*.

I avsnitt 2:6 studerades byte av koordinater i tillståndsrummet. För att genomföra analysen förutsattes att matrisen G definierad av ekvation (2:6.17) var reguljär. Det följer nu av sats 4.1 att denna förutsättning innebär att systemet är styrbart, vilket motiverar namnet styrbar kanonisk form.

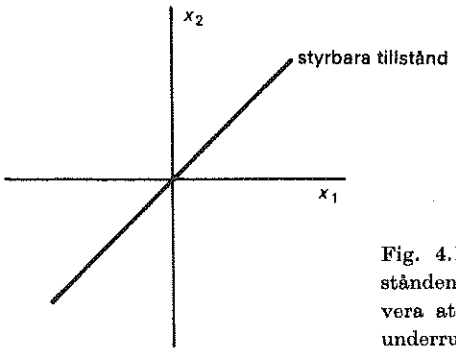


Fig. 4.1. Tillståndsrummet och de styrbara tillstånden för parallellresonanskretsen i fig. 2.1. Observera att de styrbara tillstånden bildar ett linjärt underrum.

EXEMPEL 4.1

För nätverket i fig. 2.1 gäller

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

Vi finner således

$$W_s = [B \;\; AB] = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

Styrbarhetsmatrisen W_s har rangen 1. De styrbara tillstånden bildar således ett endimensionellt underrum. Detta underrum spänns upp av vektorn (1, 1). Se fig. 4.1. Jämför analysen i avsnitt 2.

Ett styrbart tillstånd definierades så att det var möjligt att styra från x^0 till origo. I många praktiska fall är det naturligtvis intressant att också styra till andra punkter. Följande resultat motiverar den mer inskränkta definitionen.

SATS 4.2

För ett linjärt tidsinvariant system går det att styra mellan två godtyckliga tillstånd i det styrbara underrummet.

Bevis:

Det följer av sats 4.1 att det går att styra från x^0 till origo. Det återstår således endast att visa att det går att styra från origo till tillståndet x^1. Om initialtillståndet är noll gäller enligt ekvation (2:4.8)

$$x^1 = x(t_1) = \int_0^{t_1} e^{A(t_1 - s)} B u(s)\, ds = \int_0^{t_1} e^{A\tau} B u(t_1 - \tau)\, d\tau \tag{4.4}$$

där sista likheten erhållits genom variabelbytet $\tau = t_1 - s$.

Samma resonemang som utnyttjades vid bevis av sats 4.1 visar nu att det finns en styrsignal som satisfierar (4.4) om x^1 tillhör det styrbara underrummet.

Några av det styrbara underrummets egenskaper skall nu analyseras. Vi undersöker först vad som händer vid byte av koordinater i tillståndsrummet. Inför

$$z = Tx$$

där T är en reguljär matris. Systemekvationen lyder då

$$\frac{dz}{dt} = TAT^{-1}z + TBu$$

Det transformerade systemet har styrbarhetsmatrisen

$$V_s = [TB \;\; (TAT^{-1})TB \;\ldots\; (TAT^{-1})^{n-1}TB]$$
$$= [TB \;\; TAB \;\ldots\; TA^{n-1}B] = TW_s$$

Om x^0 tillhör det underrum som spänns upp av kolonnerna i W_s så tillhör $z^0 = Tx^0$ det underrum som spänns upp av kolonnerna i TW_s. Men då $TW_s = V_s$ är z^0 styrbar. Vi finner således

SATS 4.3

Det styrbara underrummet är invariant under byte av koordinater i tillståndsrummet.

Vi har vidare

SATS 4.4

Det styrbara underrummet är invariant under transformationen A.

Bevis:

Låt x^0 vara styrbar. Vektorn x^0 är då en linjärkombination av kolonnerna i

$$B \; AB \ldots A^{n-1}B$$

vektorn Ax^0 är då en linjärkombination av

$$AB \; A^2B \ldots A^nB$$

Men enligt Cayley–Hamiltons sats är A^nB en linjärkombination av B, AB, ..., $A^{n-1}B$, och Ax^0 är således en linjärkombination av kolonnerna i W_s.

Kriteriet för styrbarhet blir speciellt enkelt för system som kan skrivas på diagonalform. Om matrisen A har skilda egenvärden finns enligt avsnitt 2.6 en transformation $z = Tx$ så att systemekvationerna kan skrivas som

$$\frac{dz}{dt} = \Lambda z + \beta u$$
$$y = \gamma z + Du \qquad (4.5)$$

där Λ är en diagonalmatris

$$\Lambda = \text{diag}\,[\lambda_1, \lambda_2, \ldots, \lambda_n]$$

Betrakta den i:te moden, dvs. tillståndsvektorns i:te komponent. Vi finner

$$z_i(t) = e^{\lambda_i t} z_i^0 + \int_0^t e^{\lambda_i (t-s)} \beta_i u(s)\, ds$$

där β_i är i:te raden av matrisen β. Enligt definitionen på styrbarhet är moden z_i^0 styrbar om det finns en styrsignal sådan att $z_i(t_1) = 0$. Detta är möjligt endast om det finns något element i β_i, som är skilt från noll. I detta fall spänns det styrbara underrummet upp av koordinataxlar i z-rummet. Vi sammanfattar resultatet som

SATS 4.5

Anta att samtliga egenvärden λ_i är skilda. Moden z_i är styrbar om och endast om något element i β_i (dvs. i:te raden av β-matrisen) är skilt från noll.

Följande exempel visar att det måste krävas att alla egenvärden är olika

$$\frac{dx}{dt} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

Avslutningsvis kan det betonas att begreppet styrbarhet uttrycker att ett system kan styras från ett tillstånd till ett annat på ändlig tid. Begreppet säger däremot ej något om att styrningen kan göras så att tillståndet eller utsignalerna följer föreskrivna banor. Detta kräver ytterligare villkor. Man kan t. ex. lätt övertyga sig om att dubbelintegratorn i exempel 3:3.1 är styrbar. Däremot kan i varje tidpunkt endast tillståndsvektorns andra komponent väljas fritt genom val av styrsignalens värde

5. Kriterier för observerbarhet

Betrakta ett linjärt tidsinvariant system $\mathcal{S}(A, B, C, D)$. Vi skall nu undersöka när det går att bestämma systemets tillstånd ur observation av insignaler och utsignaler. Som tidigare anmärkts i avsnitt 3 kan man förutsätta att $u = 0$. Vi skall nu karakterisera det tysta underrummet. Det följer av ekvation (2:4.8) att

$$y(t) = Ce^{At}x^0$$

Låt x^0 vara ett tyst tillstånd. Enligt definition 3.2 gäller då $y(t) = 0$ och eftersom y kan deriveras godtyckligt många gånger även

$$y(0) = Cx^0 = 0$$

$$y'(0) = CAx^0 = 0$$
$$\vdots$$
$$y^{(n-1)}(0) = CA^{n-1}x^0 = 0$$

Ytterligare derivationer ger inget nytt ty enligt Cayley–Hamiltons ekvation gäller

$$A^n + a_1 A^{n-1} + \ldots a_n I = 0$$

Vi finner således

$$y^{(k)} = -a_1 y^{(k-1)} - a_2 y^{(k-2)} - \ldots - a_n y^{(k-n)} \quad k \geqslant n$$

Samtliga högre derivator av y kan således uttryckas i linjärkombinationer av $y(0)$, $y'(0)$, ..., $y^{(n-1)}(0)$. Om y och de $n-1$ första derivatorna av y är noll så blir alltså även alla högre derivator noll. Vi finner således

SATS 5.1

Det tysta tillstånden ligger i nollrummet till matrisen

$$W_0 = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

Systemet $\mathcal{S}(A, B, C, D)$ är observerbart om och endast om matrisen W_0 har rangen n.

Matrisen W_0 kallas *observerbarhetsmatris*. Med nollrummet till en $m \times n$ matrix W_0 menas alla $x \in R^n$ sådana att $W_0 x = 0$.

Jämför avsnitt 2:6 som behandlar byte av koordinater i tillståndsrummet. För att genomföra analysen antogs där att matrisen F definierad av ekvation (2:6.5) var reguljär. Enligt sats 5.1 innebär detta emellertid att systemet är observerbart, vilket motiverar namnet observerbar kanonisk form.

EXEMPEL 5.1

För nätverket i fig. 2.1 gäller att

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

Vi finner

$$W_0 = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Matrisen W_0 har rangen 1. De icke observerbara tillstånden ligger således i ett endimensionellt underrum, vilket är ortogonalt mot vektorn $(1, -1)$. Det icke observerbara underrummet spänns således upp av vektorn $(1, 1)$ se fig. 5.1.

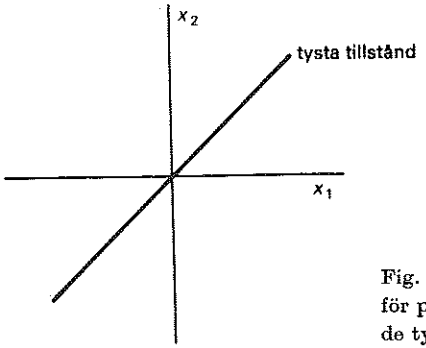


Fig. 5.1. Tillståndsrummet och de tysta tillstånden för parallellresonanskretsen i fig. 2.1. Observera att de tysta tillstånden bildar ett linjärt underrum.

Observera att i detta speciella fall överensstämmer det icke observerbara underrummet med det styrbara underrummet. Se fig. 4.1.

Några egenskaper hos det icke observerbara underrummet skall nu undersökas. Det gäller

SATS 5.2

Det tysta (icke observerbara) underrummet är invariant vid byte av koordinater i tillståndsrummet.

Bevis:

Inför nya koordinater genom transformationen

$$z = Tx$$

Det transformerade systemet får då observerbarhetsmatrisen

$$V_0 = \begin{bmatrix} CT^{-1} \\ CT^{-1}TAT^{-1} \\ \vdots \\ CT^{-1}(TAT^{-1})^{n-1} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} T^{-1} = W_0 T^{-1}$$

Antag att x^0 är icke observerbar då gäller $W_0 x^0 = 0$. För motsvarande transformerade tillstånd $z^0 = Tx^0$ gäller då $W_0 T^{-1} z^0 = 0$. Men enligt ovanstående ekvation innebär detta att z^0 är icke observerbar.

Vidare gäller

SATS 5.3

Det tysta (icke observerbara) underrummet är invariant under transformation med matrisen A.

Bevis:

Låt x^0 vara tyst. Då gäller

$$Cx^0 = 0$$

$$CAx^0 = 0$$

$$\vdots$$

$$CA^{n-1}x^0 = 0$$

Vi skall nu visa att Ax^0 också är tyst. Det följer av ovanstående ekvationer att

$$C(Ax^0) = 0$$

$$\vdots$$

$$CA^{n-2}(Ax^0) = 0$$

Det följer av Cayley–Hamiltons ekvation att

$$CA^n x^0 = -a_1 CA^{n-1}x^0 - \ldots - a^n Cx^0$$

Men högerledet försvinner eftersom x^0 är tyst och även Ax^0 är således tyst.

Observera att det följer av sats 5.3 att mängderna av icke observerbara och icke rekonstruerbara tillstånd är lika.

Kriteriet för observerbarhet blir speciellt enkelt för system som kan skrivas på diagonalform. Om matrisen A har skilda egenvärden kan systemekvationerna skrivas på formen (4.5). Detta ger omedelbart

SATS 5.4

Betrakta ett system på diagonalform med skilda egenvärden. Moden z_i är icke observerbar om alla element i i:e kolonnen av γ är noll. Systemet är observerbart om och endast om det i varje kolonn i γ finns åtminstone något element som är skilt från noll.

6. Kalmans uppdelningssats

Vi skall nu visa en konsekvens av begreppen styrbarhet och observerbarhet som är central för att förstå sambanden mellan olika representationer av dynamiska system. Resultatet innebär att ett linjärt system kan uppdelas i ett antal olika delsystem utgående från styrbarhets- och observerbarhetsegenskaper. Resultatet erhålls genom att göra en speciell transformation av koordinaterna i tillståndsrummet. Det centrala resultatet skall byggas upp successivt genom att man först löser två enklare problem.

Uppdelning med avseende på styrbarhet

Det visades i avsnitt 4 att de styrbara tillstånden bildade ett linjärt underrum av tillståndsrummet, som kallas det styrbara underrummet $\mathcal{X}_s$. Det visades också i sats 4.4 att $\mathcal{X}_s$ var oberoende av val av koordinater i tillstånds-

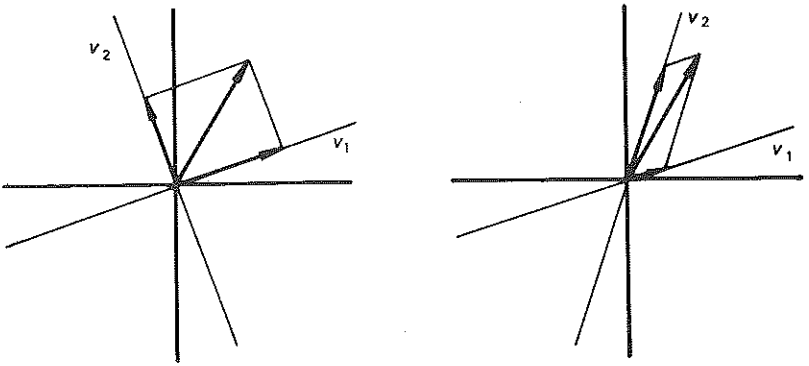


Fig. 6.1. Illustrerar att rummet R^2 kan skrivas som en direkt summa av det givna linjära underrummet V_1 och ett annat linjärt underrum V_2 på två olika sätt.

rummet. Vi skall nu ge ett val av koordinatsystem i vilket systemekvationerna får en speciell form.

Varje vektor x i tillståndsrummet kan skrivas som

$$x = x^1 + x^2$$

där x^1 ligger i det styrbara underrummet $\mathcal{X}_s$ och x^2 ligger i ett linjärt underrum $\mathcal{X}_{\bar{s}}$. Man säger att tillståndsrummet är en direkt summa av $\mathcal{X}_s$ och $\mathcal{X}_{\bar{s}}$ och skriver

$$\mathcal{X} = \mathcal{X}_s \oplus \mathcal{X}_{\bar{s}}$$

Det linjära rummet $\mathcal{X}_{\bar{s}}$ är ej entydigt såsom visas i fig. 6.1. Oavsett vilket val som görs så är dock varje tillstånd i $\mathcal{X}_{\bar{s}}$ icke styrbart och $\mathcal{X}_{\bar{s}}$ kallas därför det »icke styrbara underrummet». Observera att språkbruket är lätt missvisande då de icke styrbara tillstånden ej bildar ett linjärt underrum. Jämför fig. 4.1.

Välj nu en bas i tillståndsrummet som sammansätts av en bas för $\mathcal{X}_s$ och en bas för $\mathcal{X}_{\bar{s}}$. Med detta val av koordinater ges matriserna A och B av

$$A' = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix} \quad B' = \begin{bmatrix} B'_1 \\ 0 \end{bmatrix} \tag{6.1}$$

där koordinaterna svarande mot basen i $\mathcal{X}_s$ utgör första delen av tillståndsvektorn. Delmatrisen av A' i position 21 måste vara noll ty annars är det styrbara underrummet ej invariant under A' (sats 4.4). Vidare måste delmatrisen av B' i position 2 vara noll ty annars kunde tillstånden i $\mathcal{X}_{\bar{s}}$ vara styrbara.

Uppdelning med avseende på observerbarhet

På analogt sätt kan tillståndsrummet $\mathcal{X}$ också uppdelas som

$$\mathcal{X} = \mathcal{X}_o \oplus \mathcal{X}_{\bar{o}}$$

där $\mathcal{X}_{\bar{o}}$ är det entydiga linjära underrummet av icke observerbara tillstånd, och $\mathcal{X}_o$ är ett linjärt underrum som tillsammans med $\mathcal{X}_{\bar{o}}$ spänner upp tillståndsrummet $\mathcal{X}$. Observera att $\mathcal{X}_o$ ej är entydigt. Alla tillstånd i $\mathcal{X}_o$ är dock observerbara. Med något oegentligt språkbruk kan $\mathcal{X}_o$ kallas »det observerbara underrummet». De observerbara tillstånden bildar dock ej ett linjärt underrum. Se t. ex. fig. 5.1. Inför nu en bas i tillståndsrummet som är sammansatt av en bas i $\mathcal{X}_o$ och en bas i $\mathcal{X}_{\bar{o}}$. Med detta val av koordinater kan matriserna A och C skrivas som

$$A' = \begin{bmatrix} A'_{11} & 0 \\ A'_{21} & A'_{22} \end{bmatrix} \quad C' = [C'_1 \quad 0] \tag{6.2}$$

där tillståndsvektorns komponenter svarande mot $\mathcal{X}_0$ grupperats först. Delmatrisen av A' i position 12 måste vara noll ty enligt sats 5.3 är underrummet av tysta tillstånd $\mathcal{X}_{\bar{o}}$ invariant under A'. Vidare måste delmatrisen av C' i position 2 försvinna ty annars vore ej $\mathcal{X}_{\bar{o}}$ tyst.

Uppdelning med avseende på styrbarhet och observerbarhet

Eftersom skärningen mellan två linjära underrum är ett linjärt underrum så är

$$\mathcal{X}_{\bar{o}s} = \mathcal{X}_{\bar{o}} \cap \mathcal{X}_s$$

ett linjärt underrum. Inför nu underrummet $\mathcal{X}_{\bar{o}\bar{s}}$ vilket tillsammans med $\mathcal{X}_{\bar{o}s}$ spänner upp $\mathcal{X}_{\bar{o}}$, dvs.

$$\mathcal{X}_{\bar{o}} = \mathcal{X}_{\bar{o}s} \oplus \mathcal{X}_{\bar{o}\bar{s}}$$

Inför också det linjära underrummet $\mathcal{X}_{os}$ vilket tillsammans med $\mathcal{X}_{\bar{o}s}$ spänner upp $\mathcal{X}_s$, dvs.

$$\mathcal{X}_s = \mathcal{X}_{\bar{o}s} \oplus \mathcal{X}_{os}$$

Slutligen införs det linjära underrummet $\mathcal{X}_{o\bar{s}}$ vilket tillsammans med $\mathcal{X}_{os}\,\mathcal{X}_{\bar{o}s}$ och $\mathcal{X}_{\bar{o}\bar{s}}$ spänner upp hela tillståndsrummet $\mathcal{X}$, dvs.

$$\mathcal{X} = \mathcal{X}_{os} \oplus \mathcal{X}_{o\bar{s}} \oplus \mathcal{X}_{\bar{o}s} \oplus \mathcal{X}_{\bar{o}\bar{s}}$$

Observera att i denna uppdelning finns visst godtycke ty endast $\mathcal{X}_{\bar{o}s}$ är entydigt bestämd. Inför nu ett koordinatsystem för $\mathcal{X}$ som är sammansatt av baserna i de linjära rum som definierats ovan. Med detta val av koordinater får matriserna följande form

$$A' = \begin{bmatrix} A'_{11} & A'_{12} & 0 & 0 \\ 0 & A'_{22} & 0 & 0 \\ A'_{31} & A'_{32} & A'_{33} & A'_{34} \\ 0 & A'_{42} & 0 & A'_{41} \end{bmatrix} \quad B' = \begin{bmatrix} B'_1 \\ 0 \\ B'_3 \\ 0 \end{bmatrix}$$

$$C' = [C'_1 \quad C'_2 \quad 0 \quad 0] \tag{6.3}$$

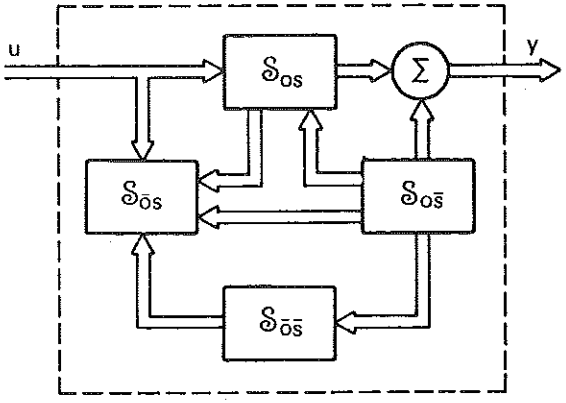


Fig. 6.2. Blockschema som illustrerar Kalmans uppdelning av ett godtyckligt system i delsystemen $\mathcal{S}_{os}$, $\mathcal{S}_{o\bar{s}}$, $\mathcal{S}_{\bar{o}s}$, $\mathcal{S}_{\bar{o}\bar{s}}$.

där tillståndsvektorns komponenter grupperats i ordning efter basvektorerna i rummet $\mathcal{X}_{os}$, $\mathcal{X}_{o\bar{s}}$, $\mathcal{X}_{\bar{o}s}$, $\mathcal{X}_{\bar{o}\bar{s}}$.

Delmatriserna av A' i positionerna 13, 14, 23 och 24 försvinner ty det icke observerbara underrummet $\mathcal{X}_{\bar{o}} = \mathcal{X}_{\bar{o}s} \oplus \mathcal{X}_{\bar{o}\bar{s}}$ är invariant under A' enligt sats 5.3. Delmatriserna av A' i positionerna 21, 23, 41 och 43 måste också vara noll ty det styrbara underrummet $\mathcal{X}_s = \mathcal{X}_{os} \oplus \mathcal{X}_{\bar{o}s}$ är invariant under A' enligt sats 4.4. Vidare måste delmatriserna av B' i positionerna 2 och 4 vara noll ty det styrbara underrummet sammansätts av $\mathcal{X}_{os}$ och $\mathcal{X}_{\bar{o}s}$ vilket motsvarar positionerna 1 och 3. På samma sätt är delmatriserna av C' i positionerna 3 och 4 noll ty det tysta delrummet motsvarar dessa positioner.

Enkla algebraiska räkningar ger vidare att systemets överföringsfunktion ges av

$$G(s) = C_2[sI - A_{11}]^{-1}B_1 + D \tag{6.4}$$

Sammanfattningsvis finner vi således

Sats 6.1. Kalmans uppdelningssats

Ett linjärt system $\mathcal{S}$ kan uppdelas i fyra delsystem $\mathcal{S}_{os}$, $\mathcal{S}_{o\bar{s}}$, $\mathcal{S}_{\bar{o}s}$ och $\mathcal{S}_{\bar{o}\bar{s}}$ där delsystemen har följande egenskaper

$\mathcal{S}_{os}$ observerbart och styrbart

$\mathcal{S}_{o\bar{s}}$ observerbart men inga styrbara tillstånd

$\mathcal{S}_{\bar{o}s}$ inga observerbara tillstånd men styrbart

$\mathcal{S}_{\bar{o}\bar{s}}$ inga observerbara och styrbara tillstånd

Systemets överföringsfunktion är entydigt bestämd av det observerbara och styrbara delsystemet och en eventuell direktterm.

Observera att uppdelningen ej är entydig. I fig. 6.2 ges ett blockschema som illustrerar uppdelningen.

Den till synes oskyldiga förutsättningen att initialtillståndet sätts lika med

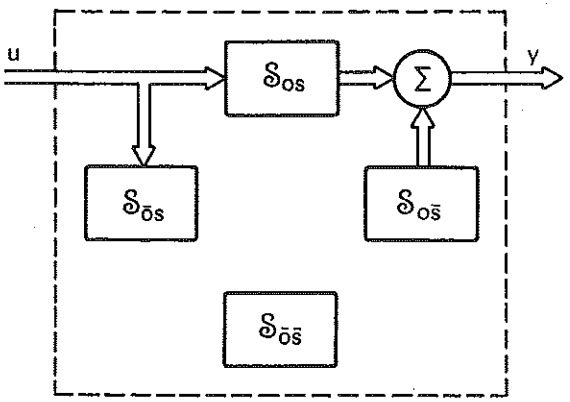


Fig. 6.3. Blockschema som illustrerar Kalmans uppdelning av ett system på diagonalform i delsystemen $\mathcal{S}_{os}$, $\mathcal{S}_{o\bar{s}}$, $\mathcal{S}_{\bar{o}s}$, $\mathcal{S}_{\bar{o}\bar{s}}$.

noll vid definition av överföringsfunktionen har således vittgående konsekvenser. Detta faktum har klargjorts först genom införandet av begreppen observerbarhet och styrbarhet. Det finns tyvärr i litteraturen ett flertal felaktigheter som har sin orsak i att man ej beaktat att överföringsfunktionen endast representerar det observerbara och styrbara delsystemet.

EXEMPEL 6.1

Betrakta parallellresonanskretsen i avsnitt 2. Efter det byte av koordinaterna som gjordes kan systemekvationerna skrivas

$$\frac{dz}{dt} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} z + \begin{bmatrix} 0 \\ 2 \end{bmatrix} u$$

$$y = [1 \quad 0]\, z + u$$

I detta fall saknas det delsystem som är observerbart och styrbart och det delsystem som varken är observerbart eller styrbart. Överföringsfunktionen bestäms enbart av den direkta kopplingen mellan u och y.

Tillståndet z_1 är observerbart men ej styrbart och tillståndet z_2 är styrbart men ej observerbart.

EXEMPEL 6.2

För system som kan diagonaliseras kan systemekvationerna skrivas på följande sätt

$$\frac{dz}{dt} = \begin{bmatrix} \Lambda_1 & 0 & 0 & 0 \\ 0 & \Lambda_2 & 0 & 0 \\ 0 & 0 & \Lambda_3 & 0 \\ 0 & 0 & 0 & \Lambda_4 \end{bmatrix} z + \begin{bmatrix} \beta_1 \\ 0 \\ \beta_2 \\ 0 \end{bmatrix} u$$

$$y = [\gamma_1 \quad \gamma_2 \quad 0 \quad 0]\, z + Du \tag{6.5}$$

där $\Lambda_1, \ldots, \Lambda_4$ är diagonalmatriser. Ty om moden z_i ej är styrbar så är enligt sats 4.5 alla element i raden β_i noll, och om moden z_i är tyst så är enligt sats 5.4 alla element i kolonnen γ_i noll. Vi finner således att systemet $\mathcal{S}(\Lambda, \beta, \gamma, 0)$ kan uppdelas i fyra delsystem $\mathcal{S}_{os}$, $\mathcal{S}_{o\bar{s}}$, $\mathcal{S}_{\bar{o}s}$ och $\mathcal{S}_{\bar{o}\bar{s}}$ svarande mot den uppdelning av tillståndsrummet, som gavs av sats 6.1. Ett blockschema för uppdelningen visas i fig. 6.3.

7. Dualitet

En jämförelse mellan sats 4.1 och sats 5.1 visar att kriterierna för styrbarhet och observerbarhet är mycket lika varandra. Betrakta t. ex. följande system

$$\frac{dx}{dt} = Ax + Bu \tag{7.1}$$

$$y = Cx$$

$$\frac{dx}{dt} = A^T x + C^T u \tag{7.2}$$

$$y = B^T x$$

Det följer omedelbart av sats 4.1 och sats 5.1 att om systemet (7.1) är styrbart så är (7.2) observerbart och om (7.1) är observerbart så är (7.2) styrbart. System med denna egenskap kallas *duala*. Systemen i exempel 2:10.2 och exempel 2;10.3 är duala.

8. Olika representationer av systemet $\mathcal{S}$ (A, B, C, D)

I kap. 2 visade vi att det linjära tidsinvarianta systemet $\mathcal{S}(A, B, C, D)$ kunde representeras på många olika sätt med tillståndsekvationer, differentialoperatorer, viktfunktion, överföringsfunktion samt flera olika grafiska representationer. I avsnitten 4 och 5 i detta kapitel har vi med hjälp av begreppen observerbarhet och styrbarhet närmare klargjort skillnaden mellan tillståndsekvationerna och överföringsfunktionen. Vi skall nu göra en sammanställning av de olika representationerna och samtidigt jämföra deras egenskaper.

Tillståndsekvationer

I avsnitt 2;3 definierades ett linjärt tidsinvariant dynamiskt system $\mathcal{S}$ med hjälp av tillståndsekvationen

$$\frac{dx}{dt} = Ax + Bu$$

$$y = Cx + Du \tag{8.1}$$

Vid härledning av systemekvationer direkt ur fysikaliska grundlagar leds man direkt till en representation av typ (8.1). Om tillståndsekvationerna är kända kan vidare systemet direkt simuleras på analogimaskin. Tillståndsekvationerna kallas därför även en *realisation* av systemet. Vid numeriska beräkningar på datamaskin är det lämpligt att utgå ifrån tillståndsekvationerna då de flesta integrationsalgoritmer förutsätter att ekvationerna är skrivna som ett system av första ordningens differentialekvationer. I vissa fall har tillståndsvariablerna fysikalisk tolkning som interna variabler i systemet. Observera emellertid att ekvationernas form beror av det koordinatsystem, som införts i tillståndsrummet. Ofta kan det vara bekvämt att införa en transformation så att ekvationerna får en trevlig form. Se avsnitt 2:6. En nackdel med representation med tillståndsekvationer är att det kan vara svårt att handskas med system av höga ordningstal.

Differentialekvationer som relaterar insignal till utsignal

Representationen med tillståndsekvationer beror av valet av koordinatsystem i tillståndsrummet. I avsnitt 2:7 visade vi att man kunde härleda en differentialekvation av n:e ordningen som relaterar insignal till utsignal

$$y(t) = \left[D + \frac{B_1 p^{n-1} + B_2 p^{n-2} + \ldots + B_n}{p^n + a_1 p^{n-1} + \ldots + a_n} \right] u(t)$$

$$p = \frac{d}{dt} \tag{8.2}$$

där matriserna D, B_1, ..., B_n och skalärerna a_1, ..., a_n är oberoende av koordinatsystemet i tillståndsrummet. Representationen (8.2) är ekvivalent med tillståndsekvationen, ty vi kan härleda (8.2) ur tillståndsekvationen och omvänt. Jämför avsnitt 2:7.

Överföringsfunktion och viktfunktion

Överföringsfunktionen och viktfunktionen är parameterfria representationer. Om viktfunktionen ges i form av en tabell så behöver systemets ordningstal ej anges explicit. Då överföringsfunktionen är viktfunktionens Laplacetransform, är representationerna ekvivalenta i den meningen att den ena kan direkt beräknas ur den andra och omvänt. Både viktfunktion och överföringsfunktion har trevliga fysikaliska tolkningar och det finns mätmetoder, impulssvarsanalys och frekvensanalys, med vars hjälp de kan bestämmas experimentellt. Överföringsfunktionen lämpar sig väl för uppskattningar av

olika slag. Då ett system representeras med överföringsfunktionen och dess grafiska representationer kan man någorlunda hanterbart analysera system av höga ordningar.

9. Konstruktion av en minimal realisation av en överföringsfunktion

Det framgår av analysen i avsnitten 4 och 5 att överföringsfunktionen endast representerar de observerbara och styrbara delsystemen av ett givet system. Detta medför att det ej är möjligt att skriva upp tillståndsekvationerna för ett system med given överföringsfunktion då man saknar information om delsystemen $\mathcal{S}_{\bar{o}s}$, $\mathcal{S}_{o\bar{s}}$ och $\mathcal{S}_{\bar{o}\bar{s}}$. Det är emellertid alltid möjligt att ange tillståndsekvationen för ett ekvivalent system som har den givna funktionen till överföringsfunktion. En sådan tillståndsekvation kallas en *realisation* av den givna överföringsfunktionen. Om tillståndsekvationen dessutom är av lägsta möjliga ordningstal sägs den vara en *minimal realisation* av den givna överföringsfunktionen. Om vi analyserar system som är specificerade med överföringsfunktioner och vill ange realisationer för dessa system, t. ex. för att simulera dem på analogimaskin eller för att utföra numeriska beräkningar, är det mycket viktigt att systemen representeras med minimala realisationer. Vi illustrerar detta med ett exempel.

EXEMPEL 9.1

Betrakta ett system med överföringsfunktionen

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

Följande dynamiska system är en realisation av denna överföringsfunktion

$$\frac{dx}{dt}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}x+\begin{bmatrix}1\\ 0\end{bmatrix}u$$

$$y=[1 \quad -1]\,x \tag{9.1}$$

ty detta system har överföringsfunktionen

$$G(s)=[1 \quad -1]\begin{bmatrix}s & -1\\ -1 & s\end{bmatrix}^{-1}\begin{bmatrix}1\\ 0\end{bmatrix}=\frac{1}{s+1}$$

Realisationen är uppenbarligen ej minimal, ty systemet (9.1) är av andra ordningen och $G(s)$ kan representeras av ett första ordningens system. Systemet (9.1) har således ett tillstånd som ej är observerbart och styrbart. För att demonstrera detta transformerar vi systemet på diagonalform. Inför $z=Tx$ där

$$T=\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix},\quad T^{-1}=\frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$$

Vi finner då

$$\frac{dz}{dt}=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}z+\begin{bmatrix}1\\ 1\end{bmatrix}u$$

$$y=[1\quad 0]\,z$$

Systemet har således två moder svarande mot egenvärdena -1 och $+1$. Den första moden är observerbar, styrbar och stabil. Den andra moden är styrbar, ej observerbar och instabil. Om systemet (9.1) simuleras på analogimaskin kommer således den instabila moden snabbt att divergera och förstärkarna blir överstyrda.

Det är illustrativt att i detalj gå igenom vad som händer om vi vill beräkna stegsvaret för överföringsfunktionen

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

genom att använda realisationen (9.1). Om $u=1$ har ekvation (9.1) lösningen

$$x_1(t)=\sinh t$$

$$x_2(t)=\cosh t-1$$

Utsignalen blir

$$y(t)=\sinh t-\cosh t+1=1-e^{-t}$$

Vi finner således att båda tillståndsvariablerna växer exponentiellt, men att utsignalen som är differensen mellan tillståndsvariablerna konvergerar mot ett konstant värde. Om ett system med överföringsfunktionen $1/(1+s)$ simuleras på analogimaskin kommer således båda tillståndsvariablerna att divergera. Förstärkarna blir överstyrda och simuleringen måste avbrytas. Ett liknande fenomen uppträder om vi i stället simulerar numeriskt. Då tillståndsvariablerna ökar exponentiellt uppnår de snabbt värden som är större än det största tal som kan representeras i en datamaskin. »Overflow» uppträder och beräkningen måste avbrytas. Om koefficienterna i realisationen (9.1) ej anges mycket noggrant, kan den instabila moden z_2 vara svagt kopplad till utsignalen och då z_2 uppnår stora värden kan även utsignalen påverkas.

Det valda exemplet är extremt, dels är det i det enkla fallet lätt att inse att realisationen ej är minimal, dels är det tillstånd som ej är observerbart instabilt. För system med en insignal och en utsignal är det lätt att konstruera minimala realisationer.

En observerbar realisation

Vi har i själva verket redan tidigare angivit en observerbar realisation av en

rationell överföringsfunktion. Jämför exempel 2:10.2, fig. 2:10.10 och fig. 2:10.15. För att visa konstruktivt hur realisationen erhålls studeras ett exempel

EXEMPEL 9.2

Vi skall ange en minimal realisation av överföringsfunktionen

$$G(s)=\frac{b_1 s^2+b_2 s+b_3}{s^3+a_1 s^2+a_2 s+a_3} \tag{9.2}$$

där vi förutsätter att täljare och nämnare ej har gemensamma faktorer. Vi söker således ett dynamiskt system av lägsta ordning som har följande insignal–utsignalrelation

$$Y(s)=\frac{b_1 s^2+b_2 s+b_3}{s^3+a_1 s^2+a_2 s+a_3}\ U(s)$$

Vi inför första tillståndsvariabeln $x_1=y$. Då gäller

$$s^3X_1+a_1 s^2X_1+a_2 sX_1+a_3X_1=b_1 s^2U+b_2 sU+b_3U$$

dvs.

$$sX_1=-a_1X_1+b_1U+\frac{1}{s^2}\left[-a_2 sX_1-a_3X_1+b_2 sU+b_3U\right]$$

Inför tillståndsvariabeln x_2 definierad av

$$X_2=\frac{1}{s^2}\left[-a_2 sX_1-a_3X_1+b_2 sU+b_3U\right]$$

dvs.

$$sX_2=-a_2X_1+b_2U+\frac{1}{s}\left[-a_3X_1+b_3U\right]$$

Inför tillståndsvariabeln X_3 definierad av

$$X_3=\frac{1}{s}\left[-a_3X_1+b_3U\right]$$

dvs.

$$sX_3=-a_3X_1+b_3U$$

Sammanfattningsvis finner vi således följande realisation

$$\frac{dx}{dt}=\begin{bmatrix}-a_1 & 1 & 0\\ -a_2 & 0 & 1\\ -a_3 & 0 & 0\end{bmatrix}x+\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}u$$

$$y=[1\quad 0\quad 0]\,x \tag{9.3}$$

I denna representation är alla tillstånd observerbara. Om polynomen $b_1 s^2+$

$b_2 s + b_3$ och $s^3 + a_1 s^2 + a_2 s + a_3$ ej har några gemensamma faktorer, så är systemet även styrbart. Under den givna förutsättningen är således systemet både observerbart och styrbart, alltså en minimal realisation. Vi observerar att systemet (9.3) är på den observerbara kanoniska formen. Jämför avsnitt 2:6.

Den konstruktion som gjordes i exemplet kan uppenbarligen också genomföras i det allmänna fallet.

En styrbar realisation

En styrbar realisation av en rationell överföringsfunktion visades i exempel 2:10.3. Realisationens blockschema och signalflödesdiagram visades i fig. 2:10.11 respektive fig. 2:10.16. Vi visar med hjälp av ett exempel hur realisationen kan konstrueras.

EXEMPEL 9.3

Betrakta överföringsfunktionen (9.2). Inför tillståndsvariablerna

$$X_1(s) = \frac{s^2}{s^3 + a_1 s^2 + a_2 s + a_3} U(s)$$

$$sX_2 = X_1$$

$$sX_3 = X_2$$

Vi finner då

$$s^3 X_1 + a_1 s^2 X_1 + a_2 s X_1 + a_3 X_1 = s^2 U(s)$$

dvs.

$$sX_1 = -a_1 X_1 - a_2 X_2 - a_3 X_3 + U$$

Systemekvationerna kan då skrivas

$$\frac{dx}{dt} = \begin{bmatrix} -a_1 & -a_2 & -a_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$$

$$y = [b_1 \quad b_2 \quad b_3]\, x \tag{9.4}$$

Vi observerar att i denna representation är samtliga tillstånd styrbara. Om polynomen $s^3 + a_1 s^2 + a_2 s + a_3$ och $b_1 s^2 + b_2 s + b_3$ ej har några gemensamma faktorer så är samtliga tillstånd också observerbara. Systemet är då en minimal realisation. Vi noterar vidare att systemet (9.4) är på den styrbara kanoniska formen. Jämför avsnitt 2:6.

Konstruktionen kan uppenbarligen genomföras för en överföringsfunktion av godtyckligt ordningstal.

Flervariabla system

Vi fann i föregående avsnitt att det var mycket lätt att konstruera minimala realisationer till en given överföringsfunktion för ett system med en insignal och en utsignal. I detta avsnitt skall vi genomföra denna konstruktion för ett flervariabelt system. För att förenkla analysen skall vi anta att den givna överföringsfunktionen har skilda poler. Vi kan då bestämma en realisation på diagonalform, dvs.

$$\frac{dz}{dt} = \Lambda z + \beta u$$

$$y = \gamma z + Du \tag{9.5}$$

Detta system har överföringsfunktionen

$$G(s) = \gamma[sI - \Lambda]^{-1}\beta + D = \sum_{i=1}^{n} \frac{\gamma_i \beta_i}{s - \lambda_i} + D$$

där γ_i är i:te kolonnen hos matrisen γ och β_i är i:te raden hos matrisen β. Observera att för moder som ej är observerbara eller styrbara försvinner produkten $\gamma_i\beta_i$ och summationen behöver därför endast utföras över index i som svarar mot observerbara och styrbara moder. Överföringsfunktionen $G(s)$ för ett system på diagonalform kan således alltid skrivas

$$G(s) = \sum_{i=1}^{n} \frac{K_i}{s - \lambda_i} + D \tag{9.6}$$

där matriserna K_i har rangen 1 och n är dimensionstalet för det styrbara och observerbara underrummet av tillståndsrummet.

Vi skall nu undersöka det omvända problemet, nämligen hur man skall konstruera en differentialekvation av lägsta ordning på diagonalform som har en given överföringsfunktion. Betrakta en överföringsfunktion $H(s)$ som har enkla poler $s = \lambda_i$, $i = 1, 2, \ldots, m$. Låt partialbråksutvecklingen av H vara

$$H(s) = \sum_{i=1}^{m} \frac{K_i}{s - \lambda_i} + D \tag{9.7}$$

där

$$D = \lim_{s \to \infty} G(s)$$

$$K_i = \lim_{s \to \lambda_i} (s - \lambda_i)\, G(s) \tag{9.8}$$

Låt r_i vara rangen av K_i. Då K_i har rangen r_i finns det r_i linjärt oberoende kolonner i K_i. Låt dessa kolonner vara e_{ji}, $j = 1, \ldots, r_i$. Varje kolonn i K_i är då en linjärkombination av e_{ji}, dvs.

$$K_i = E_i F_i$$

där $E_i = [e_{1i}, e_{2i}, \ldots, e_{r_i i}]$. Matriserna F_i bestäms på vanligt sätt genom

$$F_i = (E_i^T E_i)^{-1} E_i^T K_i$$

där inversen alltid existerar. Då matriserna E_i och F_i erhållits kan K_i uttryckas på följande sätt

$$K_i = \sum_{j=1}^{r_i} e_{ji} f_{ij}$$

där f_{ij} är j:te raden av matrisen F_i. Vi finner då

$$H(s) = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{e_{ji} f_{ij}}{s - \lambda_i} + D$$

Denna formel är identisk med (9.5) och överföringsfunktionen $H(s)$ kan således representeras med ett system på diagonalform där

$$\Lambda = \begin{bmatrix} \lambda_1 I_{r_1} & & & \\ & \lambda_2 I_{r_2} & & \\ & & \ddots & \\ & & & \lambda_m I_{r_m} \end{bmatrix}, \quad \beta = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}$$

$$\gamma = [E_1 \; E_2 \ldots E_m]$$

I_i är en enhetsmatris av ordningen i. Egenvärdet λ_i har multipliciteten r_i. Systemets ordning är

$$n = \sum_{i=1}^{m} r_i = \sum_{i=1}^{m} \text{rang } (K_i)$$

Vi sammanfattar resultatet som

SATS 9.1

Givet en överföringsfunktion $G(s)$ med enkla poler $s = \lambda_i$. Låt partialbråksutvecklingen vara

$$G(s) = \sum_{i=1}^{m} \frac{K_i}{s - \lambda_i} + D$$

Låt r_i vara rangen av K_i. Det finns då ett dynamiskt system av ordningen

$$n = \sum_{i=1}^{m} r_i$$

som har den givna överföringsfunktionen.

Vi visar med ett exempel att man lätt kan underskatta ordningstalet för ett system. Betrakta överföringsfunktionen

$$G(s) = \begin{bmatrix} \dfrac{1}{s+1} & \dfrac{2}{s+1} \\ \dfrac{-1}{(s+1)(s+2)} & \dfrac{1}{s+2} \end{bmatrix} = \frac{\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}}{s+1} + \frac{\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}}{s+2}$$

Vid hastigt påseende verkar det som om systemet har ordningen 2 men $r_1+r_2=2+1=3$. Således behövs det åtminstone ett tredje ordningens system för att realisera överföringsfunktionen.

Felbedömning av ordningstal är ett av de vanligaste felen vid analys av flervariabla system. Det finns tyvärr många fel av detta slag i den publicerade litteraturen.

10. Referenser

Begreppen observerbarhet och kontrollerbarhet har utvecklats av Kalman. Se t. ex.

Kalman, R. E., *On the General Theory of Control Systems*, in *Automatic and Remote Control*, Proceedings First IFAC Congress in Moscow, Butterworths, London 1961

Kalman, R. E., Ho, Y. C. & Narendra, K. S., *Controllability of Linear Dynamical Systems*, Contributions to Differential Equations *1* (1961) 189–213

Kalman, R. E., *Mathematical Description of Dynamical Systems*, S.I.A.M. Journal on Control Ser. A *1* (1963) 152–192.

En lättillgänglig framställning baserad på diagonalformen har givits av

Gilbert, E. G., *Controllability and Observability in Multivariable Control Systems*, S.I.A.M. Journal on Control Ser. A *1* (1963) 128–151.

Vår framställning av minimala realisationer i avsnitt 9 är baserad på Gilberts resultat. Dessa resultat har generaliserats till det fall då matrisen A har multipla egenvärden i

Kalman, R. E., *On Structural Properties of Linear Constant Multivariable Systems*, Automatic and Remote Control III. Proc. Third IFAC Congress, Inst. Mech. Engrs., Vol. 1.

En intressant vidareutveckling av begreppen styrbarhet och observerbarhet ges i

Wonham, W. M., *Linear Multivariable Control—a Geometric Approach*, Springer-Verlag, Berlin 1974.

Översiktliga presentationer av linjär algebra och matristeori ingår i kurserna matematik och numerisk analys. Lättlästa framställningar finns i

Björk, Å. & Dahlquist, G., *Numeriska metoder*, Gleerup, Lund 1969.

Fröberg, C.-E., *Lärobok i numerisk analys*. Andra upplagan. Svenska Bokförlaget/Bonniers, Stockholm 1967.

Cayley–Hamiltons sats beskrivs t. ex. på s. 58 i Fröbergs bok.

4. Analys av återkopplade system

1. Inledning

Vi skall nu närma oss det reglertekniska huvudproblemet — syntesen. För att få insyn i hur ett återkopplat system fungerar och vilka fördelar som kan uppnås med återkoppling, skall vi först analysera ett typiskt reglersystem. Resultaten från det speciella exemplet skall sedan generaliseras.

Betrakta principschemat för ett reglersystem i fig. 1.1. I kap. 2 och 3 utvecklades matematisk teknik för analys av dynamiska system. Vi skaffade oss således verktyg för att studera delsystemet I. Delsystemet II som representerar återkopplingen eller styrlagen, är i många praktiska fall också ett dynamiskt system. Vi observerar att om såväl system som styrlag är stationära dynamiska system, så kommer även det återkopplade systemet att vara ett stationärt dynamiskt system. Den teknik som utvecklades i kap. 2 och 3 kan således användas även för analys av återkopplade system.

I detta kapitel skall vi undersöka hur ett systems egenskaper påverkas vid återkoppling. Vi skall finna att återkopplingen ger många fördelar. Systemets känslighet för belastningsvariationer och variationer i systemparametrarna kan avsevärt nedbringas. Systemets förmåga att följa variationer i insignaler kan likaledes förbättras med hjälp av återkoppling.

Vi skall i stort sett följa den historiska utvecklingen och vi börjar i avsnitt 2 med att studera varvtalsreglering av en ångmaskin. Avsikten med analysen är att handgripligt demonstrera några av de viktigaste egenskaperna hos återkopplade system. För att icke i onödan tynga framställningen med algebraiska manipulationer förenklas problemet. De approximationer som görs, är i stort sett samma som gjordes i Maxwells och Vyshnegradskiis originalarbeten. Den förenklade analysen av ångmaskinregleringen visar att med hjälp av återkoppling kan fel, orsakade av störningar från systemets omgivning, göras mindre. Återkopplingen medför även att systemet ej är så känsligt för parametervariationer. Analysen visar att, i det fall då styrsignalen görs proportionell mot utsignalens avvikelse från det önskade referensvärdet (proportionell reglering), erhålls i stationärt tillstånd ett bestående fel i utsignalen. Detta kan avhjälpas om styrlagen modifieras så att utsignalen får en term, som är proportionell mot reglerfelets tidsintegral (proportionell och integrerande reglering). Vi demonstrerar även hur en alltför förenklad analys kan leda till

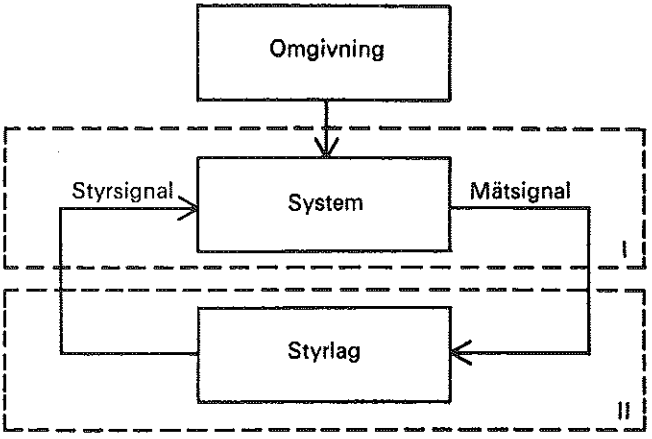


Fig. 1.1. Principschema för reglersystem.

felaktiga slutsatser. Om man försummar centrifugalregulatorns dynamik (dvs. antar att regulatorn efter en förändring i maskinens rotationshastighet omedelbart antar sitt nya jämviktsläge), så finner vi att det är gynnsamt att ha hög förstärkning i återföringskretsen. En mer detaljerad analys visar dock att en alltför hög förstärkning leder till instabilitet.

I avsnitt 3 visas hur de slutsater, som erhållits ur det enkla exemplet, kan generaliseras. Vi undersöker för en klass av system hur deras förmåga att följa styrsignaler och att eliminera störningar kan härledas ur överföringsfunktionerna för systemets delar. Ett av huvudresultaten är en enkel metod för bestämning av asymptotiska uttryck för reglerfelet då störningar och styrsignaler är polynom i tidsvariabeln. Denna analys leder till införandet av begreppen felkoefficienter och styvhetskoefficienter. Resultaten av analysen kan direkt användas för att i enkla fall avgöra huvurvida en regleruppgift går att lösa med proportionell reglering eller om det är nödvändigt att införa en eller flera integratorer i systemet. Vi undersöker vidare hur känsligheten för variationer i systemets parametrar påverkas av återkopplingen. Vi anknyter här till Blacks uppfinning av den återkopplade förstärkaren.

2. Varvtalsreglering av en ångmaskin

Vi skall nu betrakta det första reglertekniska system som analyserats matematiskt — ångmaskinen. Analysen är baserad på Maxwells och Vyshnegradskiis arbeten.

Det system vi betraktar är således en ångmaskin. Systemets insignal är ångventilens öppningsvinkel. Systemets utsignal är maskinens rotationshastighet. Systemets syfte är att maskinen skall rotera med konstant varvtal,

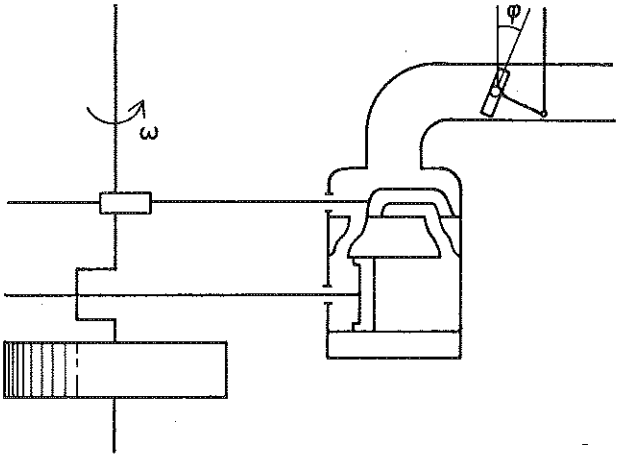


Fig. 2.1. Förenklad bild av ångmaskinsystemet (det öppna systemet). Insignalen är ångventilens öppningsvinkel φ. Utsignalen är maskinens vinkelhastighet ω.

trots störningar från omgivningen i form av belastningsvariationer och variationer i ångtryck. Dessutom är det önskvärt att man snabbt skall kunna ändra maskinens rotationshastighet från ett konstant värde till ett annat. I fig. 1:2.1 visades en enkel centrifugalregulator för konstanthållning av ångmaskinens varvtal. Vi hänvisar till detta kapitel för en redogörelse av empiriskt funna egenskaper hos centrifugalregulatorer. Vi skall nu analysera det reglersystem, som består av en ångmaskin med den centrifugalregulator som visades i fig. 1:2.1. Först analyserar vi ångmaskinens dynamiska egenskaper (det öppna systemet) och sedan skall vi studera hur systemets egenskaper ändras då regulatorn införs, dvs. då systemet återkopplas. Slutligen skall vi undersöka hur resultaten beror av regulatorns egenskaper. Härvid skall vi även diskutera de mer sofistikerade regulatorerna som omnämndes i kap. 1.

Analys av det öppna systemet

Vi skall först ställa upp ekvationer, som beskriver hur ångmaskinens varvtal beror av ångpådraget. Bektrakta det system vars insignal är ångventilens öppningsvinkel φ och vars utsignal är ångmaskinens vinkelhastighet ω (fig. 2.1).

Låt J beteckna tröghetsmomentet för ångmaskinens axel med svänghjul och belastning. Låt vidare M_B vara det moment varmed lasten påverkar maskinen samt M_D det drivande moment som erhålles från ångmaskinens cylinder. På motoraxeln verkar dessutom ett friktionsmoment M_F från lagringen. En momentekvation för svänghjulsaxeln ger

$$J\frac{d\omega}{dt} = M_D - M_B - M_F \tag{2.1}$$

Det drivande momentet M_D beror av ångventilens öppningsvinkel och dess utformning, ångans tryck och temperatur, axelns vridningsvinkel m. m. Funk-

tionssambandet är i allmänhet mycket komplicerat. I analysen skall vi förutsätta att svänghjulets tröghetsmoment är så stort att vi kan räkna med ett medelvärde av det drivande momentet över vridningsvinkeln 2π. Vi gör vidare det starkt förenklade antagandet att friktionsmomentet är en linjär funktion av vinkelhastigheten ω, dvs.

$$M_F = D\omega$$

Efter dessa förenklingar finner vi att ångmaskinen kan beskrivas med följande ekvation

$$J\frac{d\omega}{dt} + D\omega = M_D - M_B \tag{2.2}$$

där det drivande momentet M_D beror av insignalen φ. Denna förenklade ekvation användes även i Maxwells analys. Med de förenklade antagandena finner vi således att ångmaskinen är ett dynamiskt system av första ordningen. Systemet är olinjärt framförallt beroende på det samband som råder mellan det drivande momentets medelvärde över ett varv och ångventilens öppningsvinkel.

I stationärt tillstånd, dvs. för konstant ventilläge och därmed även konstant drivande moment och konstant belastning, roterar ångmaskinen med konstant varvtal

$$\omega^0 = (M_D^0 - M_B^0)/D \tag{2.3}$$

Vid en plötslig belastningsändring från ett konstant belastningsmoment till ett annat kommer rotationshastigheten att svänga in sig efter en exponentialfunktion med tidskonstanten J/D. I stationärt tillstånd erhålls en bestående förändring

$$\Delta\omega = -\Delta M_B/D \tag{2.4}$$

Ur denna synpunkt verkar det vara mycket fördelaktigt att ha en hög dämpningskoefficient. Dels erhålls en snabb insvängning, dels en liten avvikelse i vinkelhastigheten. En hög dämpningskoefficient D betyder emellertid att friktionen på motoraxeln är hög. Detta orsakar naturligtvis energidissipation, värmeutveckling och låg verkningsgrad, vilket ej är önskvärt.

Innan vi kan undersöka det återkopplade systemet, måste vi först undersöka regulatorns dynamiska egenskaper.

Analys av centrifugalregulatorn

En schematisk skiss av centrifugalregulatorn med hävarmssystem och ångventil visas i fig. 2.2.

Vi inför beteckningar enligt fig. 2.2. För att ej komplicera analysen antas att geometrin är sådan att ångventilens öppningsvinkel φ är lika med vin-

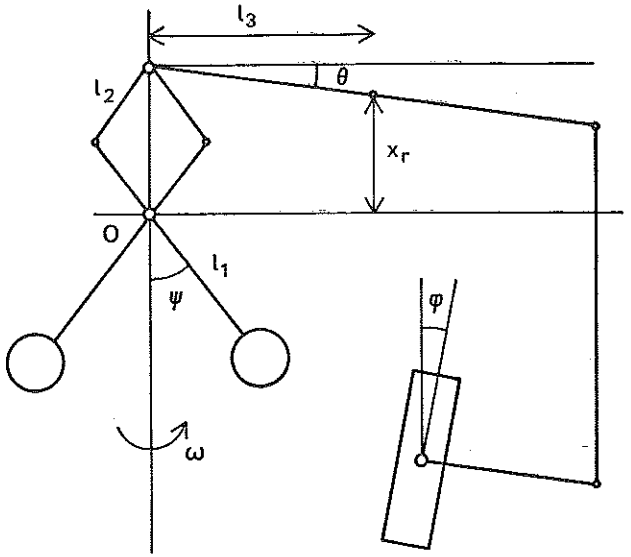


Fig. 2.2. Förenklad bild av centrifugalregulator med hävarmssystem och ångventil. Insignalen är spindelns vinkelhastighet ω. Utsignalen är ventilens inställningsvinkel φ.

keln θ. Vinkeln θ beror av läget x_r hos inställningsskruven och kulornas utslagsvinkel ψ. Det gäller att

$$\operatorname{tg} \varphi = \operatorname{tg} \theta = (2l_2 \cos \psi - x_r)/l_3 \tag{2.5}$$

Låt l_1 beteckna avståndet mellan pivotpunkten O och ena kulans masscentrum och låt g beteckna tyngdaccelerationen. Anta vidare att regulatorn är symmetrisk samt att hävarmarnas massa och friktion kan försummas. Då gäller i stationärt tillstånd följande samband mellan hävarmarnas utslagsvinkel ψ och axelns vinkelhastighet ω

$$g \sin \psi = \omega^2 l_1 \sin \psi \cos \psi$$

Regulatorns utslagsvinkel ges således av

$$\cos \psi = g/(\omega^2 l_1) \tag{2.6}$$

Sammanfattningsvis finner vi således följande relation för regulatorn, som anger hur ångventilens öppningsvinkel φ beror av vinkelhastigheten ω och regulatorinställningen x_r i stationärt tillstånd

$$\varphi = \varphi(\omega, x_r) = \operatorname{arctg} [2l_2 g/(\omega^2 l_1 l_3) - x_r/l_3] \tag{2.7}$$

Om vi nu gör det förenklade antagandet att regulatorns massa är försumbar så att det stationära tillståndet inställer sig momentant, finner vi således att regulatorn kan karakteriseras med ett rent statiskt samband (2.7). Vi skall senare analysera konsekvenserna av antagandet.

Analys av det återkopplade systemet

Vi skall nu analysera det slutna systemet ångmaskin med centrifugalregulator. Detta system kan beskrivas med ekvationerna

$$J\frac{d\omega}{dt}+D\omega=M_D(\varphi)-M_B \tag{2.8}$$

$$\varphi=\varphi(\omega, x_r) \tag{2.7}$$

Vi antar att regulatorinställningen (x_r) är konstant och vi skall undersöka hur varvtalet kommer att förändras vid små belastningsvariationer. Då regulatorinställningen är konstant kan vi sätta

$$\varphi=\varphi(\omega)$$

Anta nu att $M_B=M_B^0=$konstant. Maskinens varvtal antar då det konstanta värdet ω_0 givet av

$$D\omega_0=M_D(\varphi(\omega_0))-M_B^0 \tag{2.9}$$

För att linjärisera systemekvationerna kring jämviktslösningen $\omega=\omega_0$ införs

$$C=\left.\frac{dM_D}{d\varphi}\right|_{\varphi=\varphi(\omega_0)} \tag{2.10}$$

$$-K=\left.\frac{d\varphi(\omega)}{d\omega}\right|_{\omega=\omega_0}=-\frac{4l_2g\cos^2\varphi_0}{\omega_0^3 l_1 l_3} \tag{2.11}$$

där $\varphi_0=\varphi(\omega_0)$. Talet C anger således kvoten mellan ändring i drivande moment och ändring i ventilinställning. Talet K är regulatorns förstärkning. Det anger vid små förändringar kvoten mellan ändringen i ångventilens vridningsvinkel och ändringen i varvtal. Talet K är positivt, ty regulatorn fungerar så att ventilen stängs då vinkelhastigheten ökar (fig. 1:2.1). Vid små förändringar av ventilinställningen kring ett jämviktsläge φ_0 kan ångmaskinen således beskrivas med den linjäriserade ekvationen

$$J\frac{d}{dt}(\omega-\omega_0)+D(\omega-\omega_0)=C(\varphi-\varphi_0)-(M_B-M_B^0) \tag{2.12}$$

Vidare kan regulatorn beskrivas med

$$\varphi=\varphi_0-K(\omega-\omega_0) \tag{2.13}$$

där termer av andra och högre ordning har utelämnats. Insätt (2.13) i (2.12). Den linjäriserade ekvationen för det återkopplade systemet kan då skrivas

$$J\frac{d}{dt}(\omega-\omega_0)+(D+CK)(\omega-\omega_0)=-(M_B-M_B^0) \tag{2.14}$$

Om belastningsmomentet ändras från M_B^0 till $M_B^0+\Delta M_B$, så ändras vinkelhastigheten till $\omega_0+\Delta\omega$ där

$$\Delta\omega=-\Delta M_B/(D+CK)$$

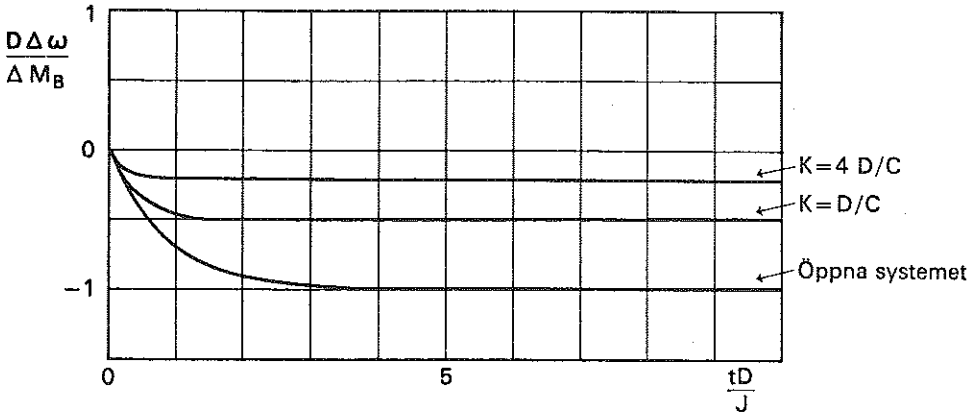


Fig. 2.3. Visar varvtalsförändringen $\Delta\omega$ vid en ökning av belastningen med ΔM_B för det öppna systemet och för det slutna systemet med två olika värden på förstärkningsparametern K.

En jämförelse med analysen av det öppna systemet visar skillnaden mellan de öppna och slutna systemen. I det öppna systemet orsakar en belastningsvariation ΔM_B en varvtalsändring $\Delta\omega = -\Delta M_B/D$. Se ekvation (2.4). I fig. 2.3 visas varvtalsförändringen efter en ökning i belastningen. I det återkopplade systemet har känsligheten för belastningsvariation M_B således minskats med en faktor $D/(D+CK)$ jämfört med det öppna systemet. Efter en plötslig lastförändring svänger det återkopplade systemet in sig efter en exponentialfunktion med tidskonstanten $J/(D+CK)$, dvs. en faktor $(D+CK)/D$ snabbare än det öppna systemet.

Observera att vid konstant ventilinställning i det öppna systemet är ekvationerna för öppna och slutna systemen (2.2) resp. (2.14) identiska sånär som på värdet av dämpningsparametern. Om man experimenterar med det återkopplade systemet (2.14) genom att ändra belastningen och observera maskinens vinkelhastighet så uppför sig systemet på precis samma sätt som ett öppet system (2.2) med dämpningskoefficienten $D+CK$. Vi kan således säga att återkopplingen skenbart ökar maskinens dämpning. Då dämpningen införts artificiellt genom återkoppling, åtföljs den emellertid ej av den energidissipation, som erhålls vid ökad viskös friktion på motoraxeln. Återkopplingen gör det alltså möjligt att erhålla ett system som har alla fördelarna hos ett system med hög dämpning (snabb insvängning och litet fel i stationärt tillstånd) utan att ha dess stora nackdel (energidissipation).

Av ekvation (2.9) följer vidare att om regulatorns inställning (x_r) är konstant kommer maskinens hastighet i stationärt tillstånd att bero av belastningen. Det är således ej möjligt att kalibrera regulatorns inställning i varvtal.

Proportionell och integrerande återkoppling

I den enkla centrifugalregulatorn, fig. 1:2.1, är förändringen i styrsignalen (ångventilens vinkelläge) proportionell mot varvtalsändringen. En sådan re-

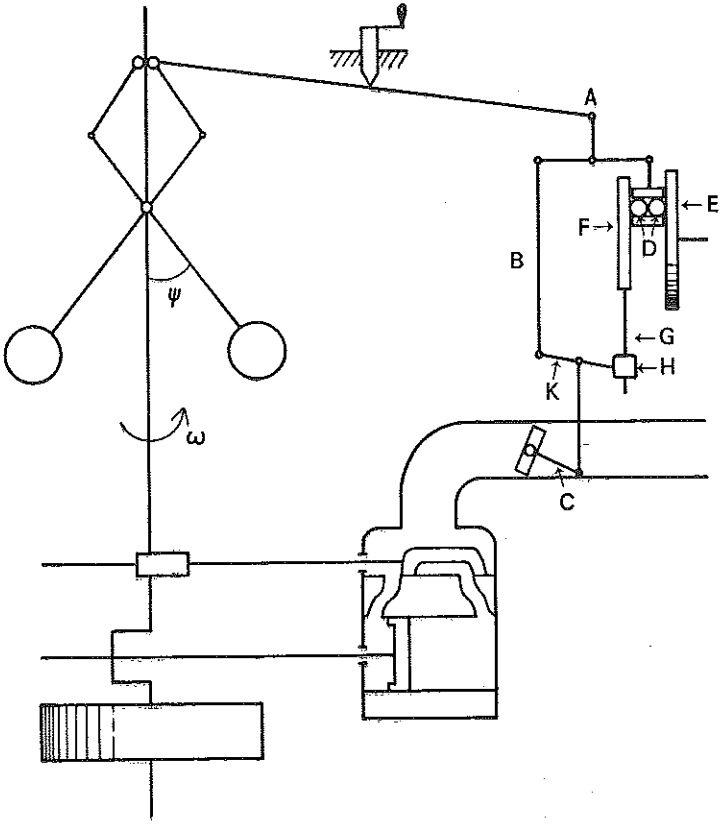


Fig. 2.4. Ångmaskin med PI-regulator. Då ångmaskinens hastighet ω ökar så ökas även kulornas utslagsvinkel ψ. Härvid lyfts punkten A via hävarmssystemet. Ventilarmen C höjs då genom kopplingen via länken B. Även kulorna D höjs då de är förbundna med A. Då D ligger an mot skivan E, som roterar med konstant hastighet, ökas kulornas hastighet proportionellt mot lägesförändringen för A. Kulorna D ligger an mot cylindern F, vilken är förbunden med skruven G som gängar i en mutter H fäst vid hävarmen K. Hävarmens högra ändpunkt får då en uppåtriktad hastighet, som är proportionell mot tidsintegralen av punkten A:s förskjutning. Ventilarmens läge blir då proportionell mot A:s läge och mot tidsintegralen av A:s läge.

gulator kallas därför *proportionell regulator* eller *P-regulator*. Analysen av ångmaskinregleringen visade att med denna regulator erhålls även i stationärt tillstånd en avvikelse i vinkelhastigheten efter en ändring av belastningen från ett konstant värde till ett annat. Man kan då fråga om det inte är möjligt att konstruera en regulator så beskaffad att en ändring av belastningen ej leder till någon bestående förändring av vinkelhastigheten. Sådana regulatorer konstruerades redan på 1800-talet av flera ångmaskinkonstruktörer. Den matematiska innebörden av dessa konstruktioner är att ventilutslaget görs proportionellt mot såväl reglerfelet $(\omega - \omega_0)$ som reglerfelets tidsintegral, dvs.

$$\varphi(\omega, t) - \varphi(\omega_0) = -K(\omega - \omega_0) - K_1 \int^t (\omega - \omega_0)\, ds \tag{2.15}$$

Den styrlag som beskrivs av ovanstående ekvation kallas *proportionell* och *integrerande*. Motsvarande regulator benämns ofta *PI-regulator*. Fig. 2.4 visar hur en *PI*-regulator kan realiseras praktiskt. I fig. 1:2.3 visas en annan regulator, som är rent integrerande.

Vi skall nu undersöka egenskaperna hos en ångmaskin, som är försedd med en proportionell och integrerande regulator. Ekvationen för regulatorn (2.15) kan skrivas

$$\frac{d\varphi}{dt} = -K\frac{d\omega}{dt} - K_1(\omega - \omega_0) \tag{2.16}$$

Vi har tidigare visat att ångmaskinen kan beskrivas med den linjäriserade ekvationen (2.12) dvs.

$$J\frac{d}{dt}(\omega - \omega_0) + D(\omega - \omega_0) = C(\varphi - \varphi_0) - (M_B - M_B^0) \tag{2.12}$$

För att beskriva det slutna systemet införs tillståndsvariablerna

$x_1 = \omega - \omega_0$

$x_2 = \varphi - \varphi_0$

Låt insignalen vara

$u = M_B - M_B^0$

och utsignalen

$y = \omega - \omega_0$

Kombineras ekvationerna (2.12) och (2.16) erhålls följande ekvation för det återkopplade systemet

$$\frac{dx_1}{dt} = -(D/J)\,x_1 + (C/J)\,x_2 - u/J$$

$$\frac{dx_2}{dt} = -K\frac{dx_1}{dt} - K_1 x_1 = -(K_1 - KD/J)\,x_1 - (KC/J)\,x_2 + (K/J)\,u \tag{2.17}$$

Det återkopplade systemet är således ett linjärt, tidsinvariant system $S(A, B, C, D)$ med

$$A = \begin{bmatrix} -D/J & C/J \\ -(K_1 - KD/J) & -KC/J \end{bmatrix},\quad B = \begin{bmatrix} -1/J \\ K/J \end{bmatrix},\quad C = [1 \quad 0],\ D = 0 \tag{2.18}$$

Insignal–utsignalrelationen erhålls direkt genom att man deriverar (2.12) och eliminerar $d\varphi/dt$ med hjälp av (2.16). Vi finner då

$$J\frac{d^2}{dt^2}(\omega - \omega_0) + (D + KC)\frac{d}{dt}(\omega - \omega_0) + K_1 C(\omega - \omega_0) = -\frac{d}{dt}(M_B - M_B^0) \tag{2.19}$$

Systemets karakteristiska ekvation är

$$\det [sI - A] = s^2 + s(D/J + KC/J) + K_1 C/J = 0 \tag{2.20}$$

Då samtliga koefficienter är positiva, följer av stabilitetsanalysen i avsnitt 2:11 att systemet är stabilt. För konstant värde på $M_B - M_B^0$ erhålls då en stationär lösning

$$x^0 = \begin{bmatrix} \omega - \omega_0 \\ \varphi - \varphi_0 \end{bmatrix} = -A^{-1}B(M_B - M_B^0) = \begin{bmatrix} 0 \\ (M_B - M_B^0)/C \end{bmatrix} \tag{2.21}$$

I stationärt tillstånd kommer vinkelhastigheten ω således alltid att överensstämma med referensvärdet ω_0 oberoende av belastningsvariationerna. Ventilinställningens stationära värde kommer emellertid att bero av belastningen. Ventilen kommer således att ställa in sig så att det drivande momentet M_D exakt balanserar belastningsmomentet M_B. Observera att detta följer direkt av formen på styrlagen (2.15) och att om det finns en stationär lösning $\omega = \text{konstant}$ så måste det gälla att $\omega = \omega_0$. Om $\omega = \text{konstant} \neq \omega_0$ så kommer enligt (2.15) ventilläget φ att växa linjärt med tiden. Ekvation (2.15) medför då att maskinens vinkelhastighet ej kan vara konstant och vi har fått en motsägelse.

En jämförelse mellan ekvationerna (2.13) och (2.15) visar att den anmärkningsvärda egenskapen erhålls på grund av att styrsignalen är en funktion av reglerfelets integral. Detta faktum påpekades redan av Maxwell, som gjorde en skarp distinktion mellan de båda styrlagarna (2.13) och (2.15). System som realiserade den första styrlagen kallades *moderatorer* och system som realiserade den andra styrlagen *regulatorer*.

Vid analys av ångmaskinen med en proportionell regulator fann vi att maskinens varvtal i stationärt tillstånd ej endast beror av regulatorns inställning utan även av belastningen. Med en proportionell och integrerande regulator är däremot varvtalet i stationärt tillstånd endast beroende av regulatorinställningen. Inställningsskruven för PI-regulatorn kan således direkt kalibreras i varvtal. Kalibreringen tillgår så att man för given inställning x_{r_i} (fig. 2.4) bestämmer vid vilken vinkelhastighet ω_i integratorn ger utsignalen noll. Inställningen x_{r_i} markeras sedan ω_i. Noggrannheten hos vinkelhastigheten för det slutna systemet i stationärt tillstånd bestäms således enbart av noggrannheten hos regulatorn med länksystem. Den är oberoende av ångmaskinens egenskaper i övrigt. Om vi betraktar regulatorns inställning som en insignal till det återkopplade systemet, kan vi således säga att med en PI-regulator överensstämmer utsignalen (maskinhastigheten) i stationärt tillstånd alltid med insignalen. Med en P-regulator däremot kan utsignalen avvika från insignalen även i stationärt tillstånd.

Bestämning av lämpliga parametrar för PI-regulatorn

Vi har således funnit att styrlagen (2.15) har en lämplig form. Om problemet vore att konstruera en lämplig regulator för en ångmaskinreglering återstår

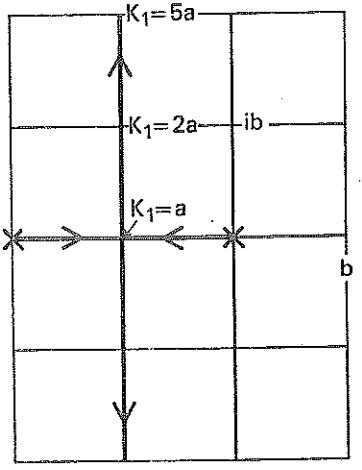


Fig. 2.5. Den karakteristiska ekvationens rötter som funktion av parametrarna K och K_1. Enheterna är $a = (D + KC)^2/4JC$ och $b = (D + KC)/2J$.

nu problemet att bestämma koefficienterna K och K_1. Vi löser detta problem analysvägen genom att undersöka hur det återkopplade systemets egenskaper varierar med K och K_1.

Systemets (2.17) karakteristiska ekvation har rötterna

$$s_{1,2} = -(D + KC)/(2J) \pm \sqrt{[D + KC/(2J)]^2 - K_1 C/J} \qquad (2.22)$$

Rötterna ligger således alltid symmetriskt med avseende på punkten $s = -(D + KC)/(2J)$. För $K_1 \leqslant a = (D + KC)^2/(4JC)$ är rötterna reella, för $K_1 > a$ är rötterna komplexa. Fig. 2.5 visar hur rötterna beror av parametrarna K och K_1. Då lösningen till systemekvationerna (2.17) sammansätts av linjära kombinationer av $\exp s_i t$ där s_i är den karakteristiska ekvationens rötter, ger

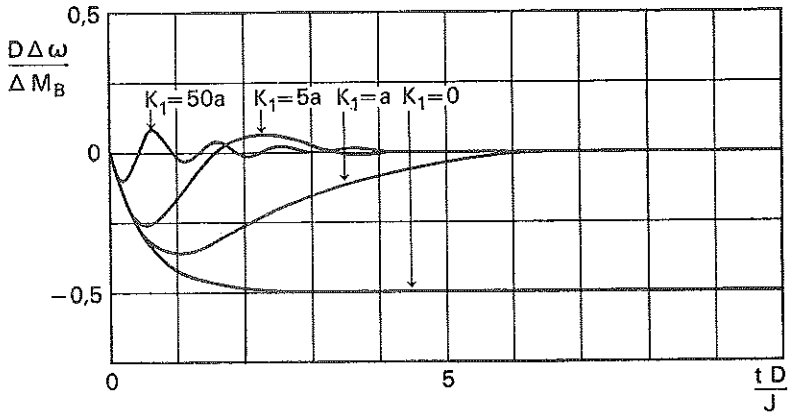


Fig. 2.6. Vinkelhastigheten $\Delta\omega$ efter en ökning av belastningen för en ångmaskin med proportionell och integrerande reglering ($u = Ke + K_1 \int e(t)\,dt$). Konstanten K har värdet $K = D/C$ och konstanten K_1 har givits fyra olika värden 0, a, $5a$ och $50a$ där $a = (D + CK)^2/4JC$.

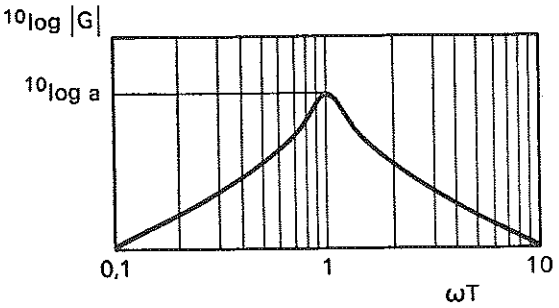


Fig. 2.7. Amplitudkurva för överföringsfunktion från störmoment till vinkelhastighet för ångmaskin med proportionell och integrerande reglering. Överföringsfunktionens maximum är $a = 1/(2\zeta\sqrt{K_1 CJ})$

fig. 2.5 ett kvalitativt uttryck på hur det återkopplade systemets egenskaper beror av parametrarna K och K_1.

Fig. 2.6 visar ändringen i vinkelhastighet efter en språngvis ökning av belastningen M_B. Vi finner att ett lämpligt värde på K_1 är $2a-5a$. Vidare finner vi att systemet svänger in sig snabbare ju högre värdet på K är. Systemet förefaller alltså att bli bättre och bättre ju högre värde på K vi väljer. Vi skall återkomma till denna fråga i nästa avsnitt.

Överföringsfunktionen från belastning till vinkelhastighet

Vi har hittills endast diskuterat ändringar i vinkelhastigheten orsakade av stegstörningar i belastningen. Vi skall nu kvalitativt bedöma hur andra typer av störningar påverkar systemet. Överföringsfunktionen från momentstörningar till vinkelhastighet är

$$G(s) = -\frac{s}{Js^2 + (D + KC)\,s + K_1 C}$$

Inför

$$T^2 = J/(K_1 C)$$

$$\zeta = (D + KC)/(2TK_1 C)$$

Överföringsfunktionen G kan då skrivas på standardformen

$$G(s) = -(K_1 CJ)^{-\frac{1}{2}} \frac{sT}{1 + 2\zeta sT + s^2 T^2}$$

Amplitudkurvan för denna överföringsfunktion visas i fig. 2.7. Det framgår av diagrammet att systemet är känsligast för momentstörningar med frekvens $\omega = 1/T = \sqrt{K_1 C/J}$, men att såväl högre som lägre frekvenser dämpas effektivt.

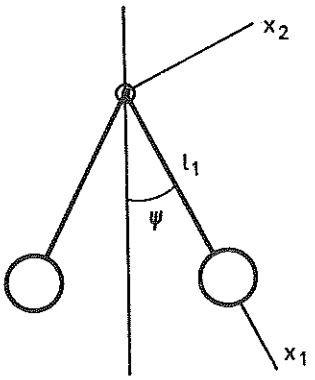


Fig. 2.8. Förenklad bild av centrifugalregulatorn med det koordinatsystem som används för analys av systemet. Koordinataxeln x_3 är vinkelrät mot x_1- och x_2-axlarna.

Analys av centrifugalregulatorns dynamik

Den förenklade analysen i föregående avsnitt gav bl. a. som resultat att, vid såväl proportionell som proportionell och integrerande reglering, felen blir mindre och insvängningen snabbare ju högre värde på parametern K som väljs. I verkligheten kommer dock stora värden på K att leda till att systemet blir instabilt. Diskrepansen mellan verkligheten och vår analys beror på de förenklade antaganden som gjorts i analysen. Vi skall nu göra en något förfinad analys, som ger en mer realistisk bild av verkligheten i detta avseende. Det finns många brister i den förenklade beskrivning av systemet, som ges av ekvationerna (2.12) och (2.13). Såväl modellen för ångmaskinen (2.12) som modellen för regulatorn (2.13) är mycket schematiserade. Vi skall här godtyckligt välja att göra en bättre beskrivning av regulatorn. Det finns två skäl till detta val. Dels är det lättare att analysera regulatorn, dels överensstämmer det med den historiska utvecklingen. Vår analys följer i stort sett Vyshnegradskiis arbete.

Den mest markanta bristen i modellen (2.13) av regulatorn är att vi försummat regulatorns dynamik. Vi har med andra ord antagit att vid en förändring i ångmaskinens varvtal från ett konstant värde till ett annat, så kommer centrifugalregulatorns kulor att omedelbart inta det nya jämviktsläget.

Fig. 2.8 visar en förenklad bild av den del av regulatorn, som innehåller pendelarmarna. För en schematiserad bild av hela regulatorn hänvisas till fig. 1:2.1.

Vi antar fortfarande att regulatorn är symmetrisk och att hävarmarnas massa kan försummas. Dessa antaganden är oväsentliga men förenklar räknearbetet. Vinkeln ψ är regulatorns utslagsvinkel, ω är regulatorns vinkelhastighet. Problemet är att finna det dynamiska sambandet mellan ω och ψ. För att beskriva regulatorns rörelse införs koordinatsystemet $x_1x_2x_3$ som antytts i fig. 2.8. Koordinatsystemets origo är fixt men koordinatsystemet är låst till pendeln och medföljer i dess rörelse. Koordinatsystemet har vinkelhastigheten

$$\Omega = -\omega \hat{x}_1 \cos \psi + \omega \hat{x}_2 \sin \psi + \frac{d\psi}{dt} \hat{x}_3 \tag{2.23}$$

där $\hat{x}_i$ betecknar enhetsvektorn längs x_i-axeln.

Om J får beteckna pendelns tröghetsmoment med avseende på axlarna x_2 och x_3 så blir pendelns rörelsemängdsmoment

$$\mathcal{M} = J\omega \sin \psi \hat{x}_2 + J \frac{d\psi}{dt} \hat{x}_3$$

Pendeln påverkas av två moment. Det ena beror av tyngdkraftens inverkan och det andra beror av den dämpning som erhålls på grund av pendelarmarnas lagring. Vi gör det förenklade antagandet att friktionsmomentet är proportionellt mot $d\psi/dt$. Det yttre moment som påverkar pendeln är således

$$-\left(\mathrm{mg}\, l_1 \sin \psi + bJ \frac{d\psi}{dt}\right) \hat{x}_3$$

Vidare erhålls följande uttryck på rörelsemängdsmomentets tidsderivata

$$\frac{d\mathcal{M}}{dt} = J\left(\frac{d\omega}{dt} \sin \psi + \omega \frac{d\psi}{dt} \cos \psi\right) \hat{x}_2 + J \frac{d^2\psi}{dt^2} \hat{x}_3 + J\omega \sin \psi \frac{d\hat{x}_2}{dt} + J \frac{d\psi}{dt} \frac{d\hat{x}_3}{dt} \tag{2.24}$$

Men $d\hat{x}_i/dt = \Omega \times \hat{x}_i$ där $\times$ betecknar kryssprodukt. Vi finner således följande ekvation för regulatorn

$$J \frac{d^2\psi}{dt^2} - J\omega^2 \sin \psi \cos \psi + bJ \frac{d\psi}{dt} + \mathrm{mg}\, l_1 \sin \psi = 0 \tag{2.25}$$

där

$$J = ml_1^2$$

Ekvation (2.25) har en stationär lösning som ges av

$$l_1 \omega^2 \cos \psi = g \tag{2.26}$$

vilket överensstämmer med ekvation (2.6), som erhölls ur en förenklad analys. Om vi vidare försummar trögheten i hävarmssystemet gäller enligt (2.5) följande relation mellan regulatorns utslagsvinkel ψ och ångventilens öppningsvinkel

$$\mathrm{tg}\, \varphi = (2l_2 \cos \psi - x_r)/l_3 \tag{2.6}$$

Vi finner således att om vi tar hänsyn till kulornas massa, så är centrifugalregulatorn ett dynamiskt system av andra ordningen. Insignalen är pendelns rotationshastighet ω och utsignalen är ångventilens öppningsvinkel φ. Om vi linjäriserar ekvation (2.25) kring den jämviktslösning ω_0, ψ_0 som ges av (2.26) och sedan utnyttjar det statiska sambandet (2.6) mellan kulornas utslagsvinkel ψ och ångventilens öppningsvinkel φ erhålls efter elementära räkningar

$$\frac{d^2}{dt^2}(\varphi-\varphi_0)+b\frac{d}{dt}(\varphi-\varphi_0)+a(\varphi-\varphi_0)=-aK(\omega-\omega_0) \tag{2.27}$$

där

$$a=g\sin^2\psi_0/(l_1\cos\psi_0) \tag{2.28}$$

och K är regulatorns statiska förstärkning som definieras av ekvation (2.11). Härledningen av ekvation (2.27) underlättas om vi observerar att det följer av ekvationerna (2.5) och (2.26) att $\varphi-\varphi_0=-K(\omega-\omega_0)$ i stationärt tillstånd.

Analys av ångmaskinreglering med proportionell återkoppling då regulatorns dynamik ej försummas

Vi skall nu undersöka hur analysen av en ångmaskin med P-regulator kommer att förändras om vi tar hänsyn till regulatorns dynamik. Om vi linjäriserar kring jämviktsläget, som ges av ekvation (2.9), finner vi att ångmaskinen kan beskrivas av ekvation (2.12), dvs.

$$J\frac{d}{dt}(\omega-\omega_0)+D(\omega-\omega_0)=C(\varphi-\varphi_0)-(M_B-M_B^0)$$

och att regulatorn kan beskrivas med ekvationen (2.27), dvs.

$$\frac{d^2}{dt^2}(\varphi-\varphi_0)+b\frac{d}{dt}(\varphi-\varphi_0)+a(\varphi-\varphi_0)=-aK(\omega-\omega_0)$$

För att beskriva det återkopplade systemet införs nu tillståndsvariablerna

$x_1=\omega-\omega_0$

$x_2=\varphi-\varphi_0=\varphi-\varphi(\omega_0)$

$x_3=dx_2/dt$

insignalen

$u=M_B-M_B^0$

och utsignalen

$y=x_1$

Det återkopplade systemet är ett linjärt, tidsinvariant system $\mathcal{S}(A, B, C, D)$ med

$$A=\begin{bmatrix}-D/J & C/J & 0\\ 0 & 0 & 1\\ -aK & -a & -b\end{bmatrix},\quad B=\begin{bmatrix}-1/J\\ 0\\ 0\end{bmatrix},\quad C=[1\ \ 0\ \ 0],\ D=0 \tag{2.29}$$

Systemet karakteristiska ekvation är

$$\det[sI-A]=(s+D/J)s(s+b)+aKC/J+a(s+D/J)=0 \tag{2.30}$$

dvs.

$$s^3 + a_1 s^2 + a_2 s + a_3 = 0$$

där

$$a_1 = b + D/J$$
$$a_2 = bD/J + a$$
$$a_3 = aD/J + aKC/J \qquad (2.31)$$

Det följer nu av exempel 2:11.3 att karakteristiska ekvationen (2.30) har samtliga rötter i vänstra halvplanet om

$$a_3 < a_1 a_2$$

Systemet är då stabilt. Utnyttjas (2.31) kan stabilitetsvillkoret skrivas

$$KC/D < b^2/a + bJ/D + bD/(aJ)$$

Vi finner således att stabilitetsvillkoret ger en övre gräns för regulatorns förstärkning K. För att kunna ge en fysikalisk tolkning till stabilitetsvillkoret införs parametrarna

$T = J/D$ ångmaskinens tidskonstant

$\beta = \sqrt{a}$ regulatorns egenfrekvens

$\zeta = \dfrac{b}{2\beta}$ regulatorns dämpningskoefficient (i typiska fall är $0 < \zeta < 1$).

Stabilitetsvillkoret kan då skrivas

$$KC/D < 2\zeta[2\zeta + \beta T + (\beta T)^{-1}] \qquad (2.32)$$

Om stabilitetsvillkoret är uppfyllt och om belastningsmomentet ändras från jämviktsvärdet M_B^0 till det konstanta värdet $M_B^0 + \Delta M_B$, ändras det dynamiska systemets tillstånd från $x = 0$ till det konstanta värdet x^0 givet av

$$Ax^0 + B\Delta M_B = 0$$

dvs.

$$x^0 = \begin{bmatrix} \omega - \omega_0 \\ \varphi - \varphi_0 \\ d\varphi/dt \end{bmatrix} = \begin{bmatrix} \Delta M_B/(D + KC) \\ -\Delta M_B K/(D + KC) \\ 0 \end{bmatrix}$$

Ändringen i vinkelhastighet blir således densamma som i det förenklade fallet.

Vi finner således att den förenklade analys som baseras på antagandet att regulatorns dynamik är försumbar kan ge helt felaktiga resultat. Den förenklade analysen visar således att det är gynnsamt att välja regulatorns

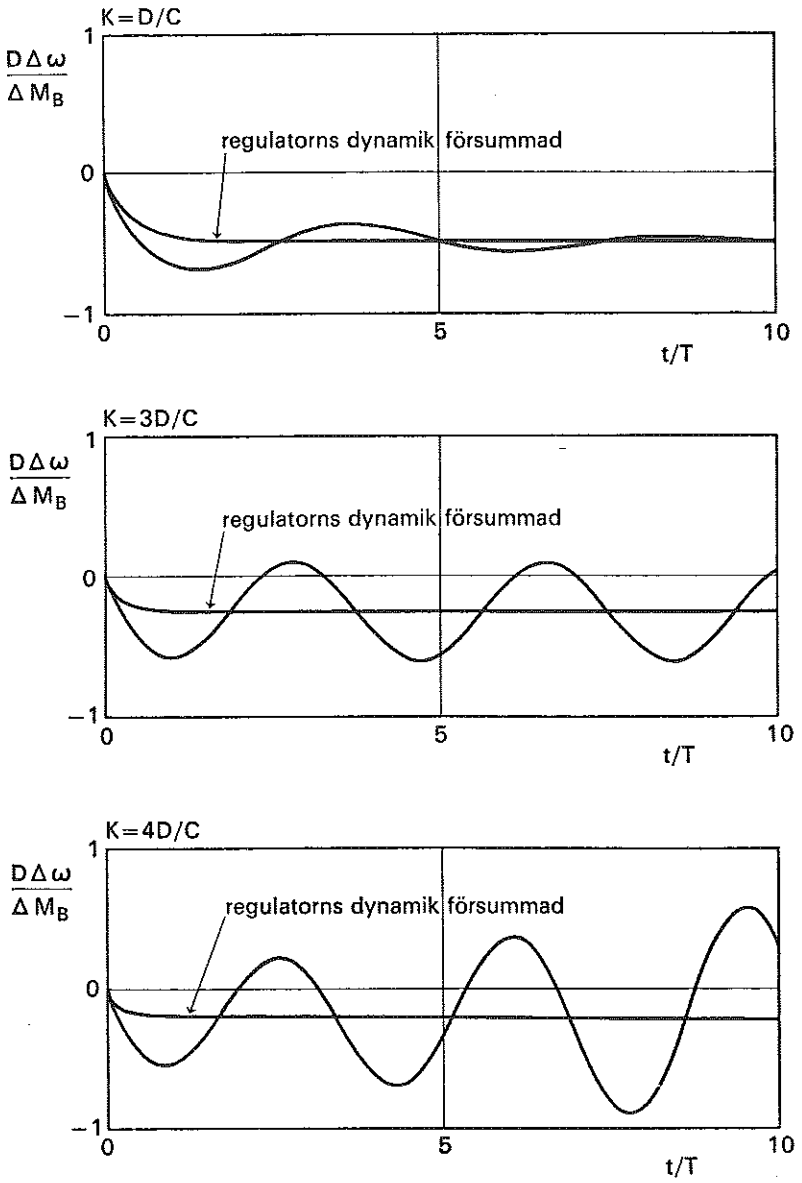


Fig. 2.9. Visar vinkelhastighetsförändringen $\Delta\omega$ efter en ökning av belastningen för en ångmaskin med proportionell återkoppling med olika värden på regulatorns förstärkning. Regulatorns parametrar har valts så att $\beta T = 1$ och $\zeta = 0{,}5$. Figuren visar även vad som händer då regulatorns dynamik försummas i analysen.

förstärkning så stor som möjligt. Om regulatorns dynamik medtas finner vi emellertid att det finns en övre gräns för regulatorns förstärkning som ges av ekvation (2.32). Om denna gräns överskrids, blir systemet instabilt. Jämför fig. 2.9.

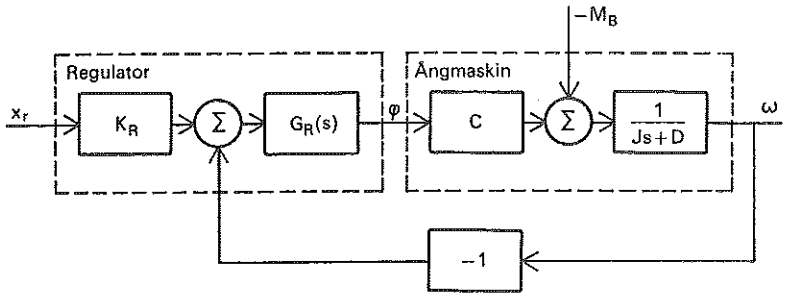


Fig. 2.10. Blockschema för ångmaskin med regulator. Ångmaskinblocket har två insignaler: ångventilens öppningsvinkel φ och belastningsmomentet M_B, och en utsignal: maskinens vinkelhastighet ω. Regulatorblocket har två insignaler: regulatorinställning x_r och vinkelhastighet ω, och en utsignal: ångventilens öppningsvinkel φ.

Sammanfattning och generalisering

Sammanfattningsvis finner vi således

- Med proportionell reglering erhålls ett bestående fel i vinkelhastighet efter en belastningsändring. Felet avtar då regulatorns förstärkning ökar. Regulatorns inställningsskruv (x_r i fig. 2.2) kan ej kalibreras i vinkelhastighet då maskinens varvtal i stationärt tillstånd beror av belastningen.

- Med proportionell och integrerande reglering erhålls inget bestående fel efter en belastningsändring. Regulatorns inställningsskruv kan graderas i vinkelhastighet. Noggrannheten i kalibreringen beror endast av precisionen i regulatorns mekaniska delar.

- Starkt idealiserade matematiska modeller leder lätt till felaktiga slutsater. Den förenklade modell där regulatorns dynamik försummats indikerar t. ex. att det skulle vara gynnsamt med en regulator, som har mycket hög förstärkning. En noggrannare matematisk modell, som tar hänsyn till regulatorns dynamik, visar dock att höga förstärkningar medför att systemet blir instabilt.

I fig. 2.10 visas ett blockschema för ångmaskin med regulator.

De olika fall vi analyserat motsvarar olika överföringsfunktioner $G_R(s)$ för regulatorn. Det gäller

Proportionell reglering då regulatorns tröghet försummas

$$G_R(s) = K$$

Proportionell och integrerande reglering då regulatorns tröghet försummas

$$G_R(s) = K + K_1/s$$

Proportionell reglering då regulatorns tröghet ej försummas

$$G_R(s) = \frac{aK}{s^2 + bs + a}$$

Det vore nu naturligt att även undersöka fallet med PI-reglering då regulatorns tröghet ej är försumbar. Vi genomför ej analysen utan nöjer oss med att ange att även i detta fall måste förstärkningsparametern K vara mindre än ett givet värde för att systemet skall vara stabilt.

Låt oss slutligen påpeka att de enda egenskaper hos ångmaskinen vi utnyttjade för analysen var att den kunde beskrivas som ett linjärt dynamiskt system av första ordningen, som karakteriseras av ekvationen

$$T\frac{dx}{dt} + x = K_u u + K_v v$$

där x är utsignal, u styrvariabel och v störning. Då många andra tekniska system kan beskrivas approximativt med dessa ekvationer, är den genomförda analysen även giltig för många andra system, t. ex. vid hastighetsreglering av explosionsmotorer, ankarstyrda likströmsmotorer, turbiner, temperaturreglering av enkel ugn och nivåreglering av tank med hjälp av utflödesventil.

3. Kvalitativ analys av återkopplade system

Ur den förenklade analysen av ångmaskinregleringen erhålls följande slutsatser:

- Genom en proportionell återkoppling kan det fel i rotationshastigheten, som orsakas av tryck och belastningsvariationer, minskas. Om styrlagen förses med en term proportionell mot felets tidsintegral kan felet i stationärt tillstånd helt elimineras.
- Genom återkoppling kan systemets egenskaper i viss utsträckning göras oberoende av variationer i systemets parametrar. Det återkopplade systemets noggrannhet i stationärt tillstånd bestäms väsentligen av mätinstrumentets precision medan det öppna systemets noggrannhet kritiskt beror av samtliga komponenters precision. Även de komponenter som skall ge stora krafter, måste hos det öppna systemet vara noggranna.

Vi skall nu undersöka i vilken utsträckning dessa slutsater gäller för andra typer av återkopplade system. Vi studerar en klass av system vars blockschema visas i fig. 3.1.

Systemets insignal betecknas med u, dess utsignal med y. Variabeln v betecknar en störning och e är reglerfelet. Vi förutsätter att systemets syfte är att utsignalen skall reproducera insignalen så väl som möjligt trots störningens

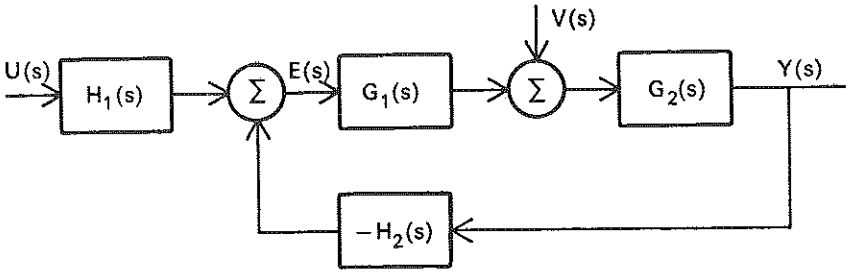


Fig. 3.1. Blockschema för reglersystem med en insignal u och en störsignal v.

inverkan. Överföringsfunktionerna $G_1(s)$, $G_2(s)$ och $H_1(s)$, $H_2(s)$ motsvarar såväl systemet som de styrlagar som införts för att få ett system med önskade egenskaper. Överföringsfunktionen

$$G_0(s) = G_1(s)\,G_2(s)\,H_2(s) \tag{3.1}$$

kallas *kretsöverföring* (loop transfer function).

Vi skall nu för systemet i fig. 3.1 undersöka hur dess

- förmåga att följa styrsignaler
- förmåga att eliminera störningar
- känslighet för parametervariationer

beror av överföringsfunktionerna $G_1(s)$, $G_2(s)$, $H_1(s)$ och $H_2(s)$.

Från fig. 3.1 erhålls följande samband mellan u, v, y och e

$$Y(s) = \frac{G_1(s)\,G_2(s)\,H_1(s)}{1 + G_1(s)\,G_2(s)\,H_2(s)}\,U(s) + \frac{G_2(s)}{1 + G_1(s)\,G_2(s)\,H_2(s)}\,V(s) \tag{3.2}$$

$$E(s) = H_1(s)\,U(s) - H_2(s)\,Y(s)$$

Vi finner således att om beloppet av kretsöverföringen $G_0(s)$ är mycket större än 1 så gäller approximativt

$$Y(s) \sim \frac{H_1(s)}{H_2(s)}\,U(s) + \frac{1}{G_1(s)\,H_2(s)}\,V(s) \tag{3.3}$$

Signaltransmissionen från u till y bestäms således väsentligen av överföringsfunktionerna $H_1(s)$ och $H_2(s)$ och är i stort sett oberoende av kretsöverföringen $G_0(s)$. Observera att om $H_1(s) = 1$, så är överföringsfunktionen från u till y approximativt $1/H_2(s)$, dvs. inversen av överföringsfunktionen i återkopplingskretsen. Det följer av (3.3) att om $|G_1(s)H_2(s)| \gg 1$, dvs. om det väsentliga bidraget till kretsförstärkningen $G_0(s)$ kommer från $G_1(s)H_2(s)$, så kommer störningen v endast att ha obetydlig inverkan på utsignalen y. Jämför 3.1. Då en integrator införs i systemet (jämför exemplet med ångmaskin-

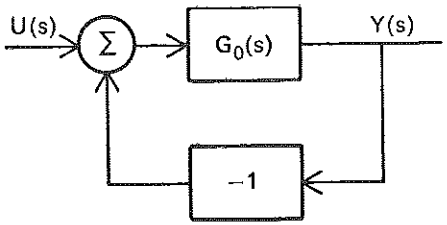


Fig. 3.2. Blockschema för enkelt slutet system med kretsöverföringen $G_0(s)$.

reglering) ökas G_1 mycket kraftigt vid låga frekvenser och inverkan från långsamma störningar minskas avsevärt.

I följande avsnitt skall vi göra en kvantitativ analys baserad på ovanstående heuristiska resonemang. Då systemet är linjärt gäller superpositionsprincipen och vi kan således separat studera inverkan av insignaler och störningar på systemet.

Statisk noggrannhet. Felkoefficienter

Vi skall först undersöka systemets förmåga att följa en klass av insignaler. Det förutsätts att störningarna är noll. Vidare antas att i fig. 3.1 är $H_1(s) = H_2(s) = 1$. Systemet har således det blockschema som visas i fig. 3.2

Vi antar att det öppna systemet har överföringsfunktionen

$$G_0(s) = \frac{k_n(1 + b_1 s + b_2 s^2 + \ldots + b_p s^p)}{s^n(1 + a_1 s + a_2 s^2 + \ldots + a_q s^q)} = \frac{k_n B(s)}{s^n A(s)} \tag{3.4}$$

Det *slutna* systemets överföringsfunktion är då

$$\frac{Y(s)}{U(s)} = G(s) = \frac{G_0(s)}{1 + G_0(s)} = \frac{k_n B(s)}{k_n B(s) + s^n A(s)}$$

Observera att det kan vara lätt att förväxla de öppna och slutna systemens överföringsfunktioner. Vi skall därför indicera det öppna systemets överföringsfunktion med index 0.

Om $n > 0$ har det öppna systemets överföringsfunktion således en singularitet i origo. Talet n kallas systemets *typnummer* och motsvarar antalet integrationer i det öppna systemet. Vanliga värden på n är 0, 1 och 2. Talet k_n kallas systemets *kretsförstärkning*. Observera att $G_0(s)$ ofta antar mycket stora värden för små s medan $G(s)$ oftast antar värden nära 1 för små s. Vi förutsätter vidare att koefficienterna a_i och b_i är sådana att funktionen $1 + G_0(s)$ har alla nollställen i halvplanet Re $s < -\alpha$ där $\alpha > 0$. Samtliga poler till det slutna systemet ligger då till vänster om den räta linjen Re $s = -\alpha$.

Vi skall nu undersöka reglerfelet $e = u - y$, dvs. differensen mellan utsignal och insignal då insignalen u är polynom av n:te ordningen

$$u(t) = u_0 + u_1 t + u_2 t^2 + \ldots + u_n t^n$$

Reglerfelets Laplacetransform $E(s)$ är given av

$$E(s) = \frac{1}{1 + G_0(s)}\, U(s)$$

Då $u(t)$ är ett polynom av n:te graden i t är dess Laplacetransform $U(s)$ analytisk i hela planet utom i origo där funktionen har en pol av ordningen $n+1$. Funktionen $E(s)$ har således poler i origo och poler till vänster om linjen $\operatorname{Re} s = -\alpha$. Vi skall nu göra en asymptotisk uppskattning av reglerfelet $e(t)$. För att göra detta skall vi skriva $E(s)$ som en summa av två termer, där den ena har alla poler i origo och den andra har alla poler till vänster om linjen $\operatorname{Re} s = -\alpha$.

Inför koefficienterna e_i definierade av

$$\frac{1}{1 + G_0(s)} = e_0 + e_1 s + e_2 s^2 + \ldots + e_n s^n + \ldots \tag{3.5}$$

Serien konvergerar åtminstone för $|s| < \alpha$, ty vi hade förutsatt att funktionen $1 + G_0(s)$ ej har några nollställen i halvplanet $\operatorname{Re} s \geqslant -\alpha$.

Funktionen $E(s)$ kan då skrivas på följande sätt:

$$E(s) = [e_0 + e_1 s + e_2 s^2 + \ldots + e_n s^n]\, U(s) + \left[\frac{1}{1 + G_0(s)} - e_0 - e_1 s - e_2 s^2 - \ldots - e_n s^n\right] U(s) \tag{3.6}$$

Genom att använda formeln för Laplacetransformen av derivatan av en tidsfunktion (ekvation 2:8.2) kan vi direkt erhålla den tidsfunktion, som svarar mot den första termen i högerledet i (3.6). Vi finner

$$\mathcal{L}^{-1}(e_0 + e_1 s + \ldots + e_n s^n)\, U(s) = e_0 u(t) + e_1 \frac{du(t)}{dt} + \ldots + e_n \frac{d^n u(t)}{dt^n}, \quad t > 0$$

Vi observerar vidare att den andra termen i högerledet av ekvation (3.6) ej har några singulariteter i halvplanet $\operatorname{Re} s \geqslant -\alpha$, ty

$$[1 + G_0(s)]^{-1} - e_0 - e_1 s - \ldots - e_n s^n = \underset{s \to 0}{0}\,(s^{n+1})$$

$$U(s) = \underset{s \to 0}{0}\,(s^{-n-1})$$

där $F(s) = \underset{s \to 0}{0}\,(s^m)$ betecknar att $F(s)/s^m$ är begränsad då $s \to 0$.

Enligt sats 2:11.1 gäller då att motsvarande tidsfunktion är begränsad av $Ke^{-\alpha t}$. Vi finner således

$$\left| e(t) - e_0 u(t) - e_1 \frac{du(t)}{dt} - \ldots - e_n \frac{d^n u(t)}{dt^n} \right| < Ke^{-\alpha t}, \quad t > 0 \tag{3.7}$$

Vilket leder till följande asymptotiska uttryck för reglerfelet e

$$e(t) \sim e_0 u(t) + e_1 \frac{du(t)}{dt} + \ldots + e_n \frac{d^n u(t)}{dt^n} \tag{3.8}$$

Uppskattningen (3.7) visar hur snabbt det asymptotiska uttrycket uppnås. Ekvation (3.8) visar att felet är en summa av termer som är proportionella mot styrsignalen och dess derivator. Proportionalitetsfaktorerna utgörs av talen e_i. Talen e_i kallas därför *felkoefficienter* (error coefficients).

Vi sammanfattar resultatet som

SATS 3.1

Om överföringsfunktionen $[1+G_0(s)]^{-1}$ ej har några poler i halvplanet Re $s \geqslant -\alpha$ och om insignalen u är ett polynom av n:e graden så gäller följande olikhet för reglerfelet e

$$\left| e(t) - e_0 u(t) - e_1 \frac{du(t)}{dt} - \ldots - e_n \frac{d^n u(t)}{dt^n} \right| < K e^{-\alpha t}, \quad t > 0 \tag{3.9}$$

där talen e_n är felkoefficienterna, dvs. koefficienterna för s^n i Maclaurinutvecklingen av $1 - G(s) = [1+G_0(s)]^{-1}$.

Ekvation (3.9) visar speciellt att om $e_0 = e_1 = \ldots = e_n = 0$, så går felet vid en insignal i form av ett polynom av n:te ordningen mot noll då t växer. Vi finner vidare att för ett system med kretsöverföringen (3.4) gäller

$$\frac{1}{1+G_0(s)} = \frac{s^n A(s)}{s^n A(s) + k_n B(s)}$$

Vi finner således följande samband mellan felkoefficienterna och talen k_n.

1. $n = 0 \quad e_0 = \dfrac{1}{1+k_0}$

2. $n > 0 \quad e_0 = \ldots = e_{n-1} = 0, \; e_n = \dfrac{1}{k_n}$

De första talen k_n har därför fått speciella namn. k_0 kallas *statisk konstant* (position constant, step constant), k_1 kallas *hastighetskonstant* (velocity constant) och k_2 kallas *accelerationskonstant* (acceleration constant). Detta kommer sig av att det stationära felet efter en stegstörning h är $h/(1+k_0)$. För ett system med $k_0 = \infty$ gäller vidare att det stationära felets storlek efter en rampstörning vt är v/k_1 etc. Med hjälp av felkoefficienterna kan vi således direkt uppskatta reglerfelen vid enkla insignaler. Den genomförda analysen visar även hur följfelen kan minskas genom att man inför integrationer i det öppna systemet.

Systemets förmåga att eliminera störningar. Styvhetskoefficienter

I föregående avsnitt analyserades systemets följsamhet för insignaler under förutsättningen att inga störningar verkar på systemet. Vi skall nu undersöka systemets känslighet för störningar då inga insignaler påverkar systemet. Genom superposition kan sedan de båda resultaten kombineras.

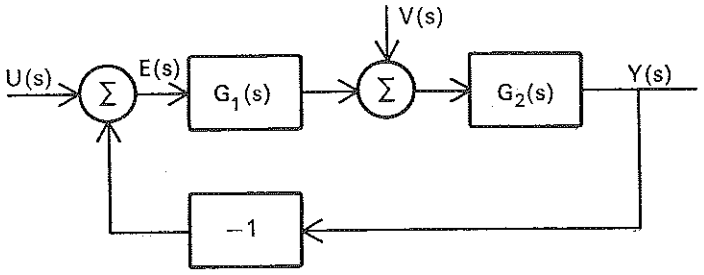


Fig. 3.3.

Betrakta det system vars blockschema framgår av fig. 3.3.

Låt $v(t)$ beteckna störningen och $e(t)$ reglerfelet. Anta att insignalen är noll. Vi finner

$$E(s) = -\frac{G_2(s)}{1+G_1(s)\,G_2(s)}\,V(s) = -\frac{1}{G_1(s)}\,\frac{G_0(s)}{1+G_0(s)}\,V(s)$$

där

$$G_0(s) = G_1(s)\,G_2(s)$$

För ett reproducerande system skall gälla att kretsöverföringens belopp $|G_0(s)|$ är stort inom insignalens frekvensband. I vilken utsträckning störningar dämpas beror då på var i systemet störningen angriper. Om $G_1 = 1$, dvs. om störningen inträder i systemet på samma sätt som insignalen, så kommer störningen att dämpas endast i ringa utsträckning. Om däremot $G_2(s) = 1$, dvs. störningen påverkar utsignalen direkt, så blir störningens dämpning mycket effektivare.

För att kvantitativt uppskatta störningarnas inverkan kan vi förfara på samma sätt som vid analysen av systemets följsamhet. Vi finner då

SATS 3.2

Om överföringsfunktionen $G_2(s)[1+G_1(s)\,G_2(s)]^{-1}$ ej har några poler i halvplanet $\mathrm{Re}\,s \geqslant -\alpha$, $\alpha > 0$ och om v är ett polynom av n:te graden så gäller följande olikhet för reglerfelet e

$$\left| e(t) - c_0 v(t) - c_1 \frac{dv(t)}{dt} - \ldots - c_n \frac{dv^n(t)}{dt^n} \right| < Ke^{-\alpha t} \tag{3.10}$$

där K är en konstant och talen c_n är koefficienterna för s^n i Maclaurinutvecklingen av funktionen $-G_2(s)[1+G_1(s)\,G_2(s)]^{-1}$.

Koefficienterna i serieutvecklingen

$$-\frac{G_2(s)}{1+G_1(s)\,G_2(s)} = c_0 + c_1 s + c_2 s^2 + \ldots$$

anger således hur systemet kan motstå inverkan av störningar, som kan beskrivas med polynom. Talen c_i kallas *styvhetskoefficienter*. Observera att

om $G_2(s)=1$, dvs. om störningen påverkar utsignalen direkt, så gäller $c_i=-e_i$ där e_i är felkoefficienterna.

EXEMPEL 3.1. Varvtalsreglering av en ångmaskin

Vid proportionell reglering av en ångmaskin gäller

$$G_1(s)=KC$$

$$G_2(s)=\frac{1}{Js+D}$$

Jämför fig. 2.10. Vi har således följande uttryck på överföringsfunktionen från störningen v till felet e

$$\frac{-G_2(s)}{1+G_1(s)\,G_2(s)}=\frac{-1}{CK+D+Js}=\frac{-1}{D+CK}+s\,\frac{J}{(D+CK)^2}-\cdots$$

Vi får

$$c_0=-(D+CK)^{-1}$$

$$c_1=J(D+CK)^{-2}$$

och finner omedelbart att ett konstant störmoment ΔM_B i stationärt tillstånd ger ett fel i vinkelhastigheten

$$\Delta\omega=c_0\Delta M_B=-\Delta M_B/(D+CK)$$

Jämför avsnitt 2. I fallet proportionell och integrerande återkoppling gäller i stället

$$G_1(s)=C(Ks+K_1)/s$$

och vi finner följande uttryck på överföringsfunktionen från störningen v till felet e

$$\frac{-G_2(s)}{1+G_1(s)\,G_2(s)}=\frac{-s}{CK_1+s(D+CK)+Js^2}=-\frac{s}{CK_1}+\frac{D+CK}{(CK_1)^2}\,s^2+\cdots$$

Vi får

$$c_0=0$$

$$c_1=-(CK_1)^{-1}$$

$$\vdots$$

Ett konstant störmoment ger alltså ej något fel i maskinens hastighet i stationärt tillstånd.

Känslighet för parametervariationer

Vid den förenklade analysen av ångmaskinreglering fann vi att det återkopplade systemet hade en skenbar dämpningskoefficient $D+CK$, där K var regulatorns

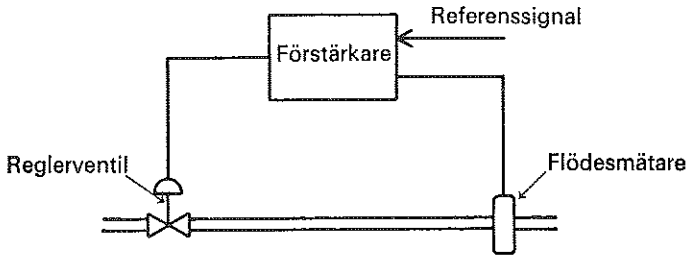


Fig. 3.4. Reglerkrets bestående av flödesmätare, förstärkare och reglerventil.

förstärkning. Om K är stor så dominerar termen CK och variationer i D kommer ej nämnvärt att påverka det slutna systemets egenskaper. Återkopplingen kan således medföra att det slutna systemet blir mindre känsligt för variationer i systemparametrarna. Detta är ibland det enda skälet till att återkoppling användes. Vi har tidigare i avsnitt 1:2 berört återkopplingens betydelse för långdistanstelefonin. I detta sammanhang är problemet att konstruera en elektronisk förstärkare, som har extremt konstant förstärkning trots variationer i de ingående komponenternas parametrar. Detta problem löstes som tidigare nämnts av H. S. Black genom uppfinningen av den återkopplade förstärkaren. Vi citerar »By building an amplifier whose gain is deliberately made, say 40 decibels higher than necessary (10 000 fold excess on energy basis) and then feeding the output back on the input in such a way as to throw away the excess gain, it has been found possible to effect extraordinary improvement in constancy of amplification and freedom from nonlinearity.» Vi illustrerar effekten med ett exempel.

EXEMPEL 3.2

En reglerventil har olinjär karakteristik mellan ventilläge u och flöde y som anges av funktionen

$$y = g(u) = u^2 \quad 0{,}1 \leqslant u \leqslant 1$$

Vid små förändringar i ventilläget kan vi definiera ventilens förstärkning som

$$\frac{dy}{du} = g'(u) = 2u$$

Om reglerventilen används i en reglerkrets där ventilläget kan variera mellan gränserna 0,1 och 1,0, så kommer ventilens förstärkning således att variera mellan värdena $g'(0{,}1) = 0{,}2$ och $g'(1) = 2$.

En förstärkningsvariation med en faktor 10 kan i vissa fall medföra mycket stora olägenheter. För att eliminera dessa olägenheter kan ventilen införas i ett återkopplat system på det sätt som illustreras i fig. 3.4.

Man inför således en flödesmätare. Signalen från denna y subtraheras från en referenssignal u_r, differensen $u_r - y$ förstärks. Den förstärkta signalen styr

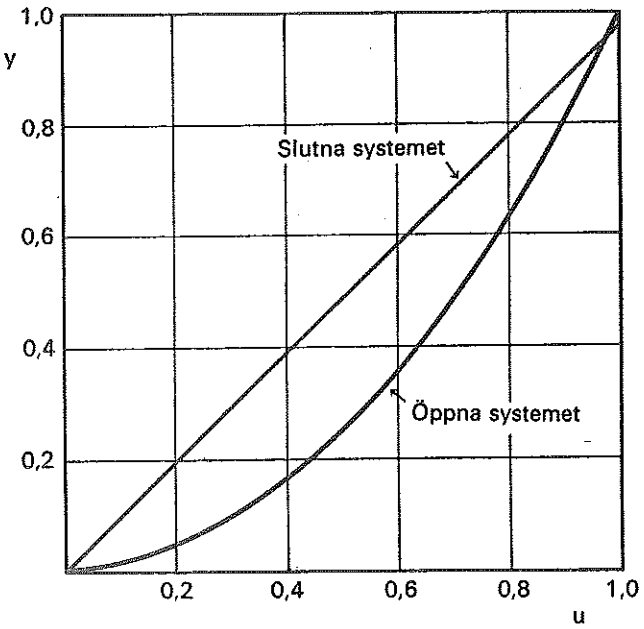


Fig. 3.5. Insignal–utsignalrelationen för öppna och slutna systemen. Det slutna systemet har $k = 50$. Jämför linjariteten för öppna och slutna systemen.

ventilens läge med hjälp av en servomotor. Vi skall nu visa att förstärkningen för det återkopplade systemet kan hållas mycket nära konstant trots variationerna i reglerventilens förstärkning. Anta att den krets, som består av förstärkare och motor, är ett linjärt statiskt system, dvs. att kvoten mellan ändring i reglerventilens läge och ändring i flödesmätarens signal är konstant k. Det återkopplade systemet kan då beskrivas med ekvationerna

$$e = u_r - y$$

$$y = g(ke)$$

Elimination av e ger följande samband mellan insignalen u_r och utsignalen y

$$u_r = y + \frac{1}{k}\, g^{-1}(y) = y + \frac{1}{k}\sqrt{y}$$

där g^{-1} betecknar den till g inversa funktionen. Det slutna systemets förstärkning är

$$G = \frac{dy}{du_r} = \frac{2k\sqrt{y}}{1 + 2k\sqrt{y}} = \frac{2ku}{1 + 2ku}$$

Genom att välja k tillräckligt stort kan det slutna systemets förstärkning fås att variera mycket litet trots att reglerventilens förstärkning varierar med en

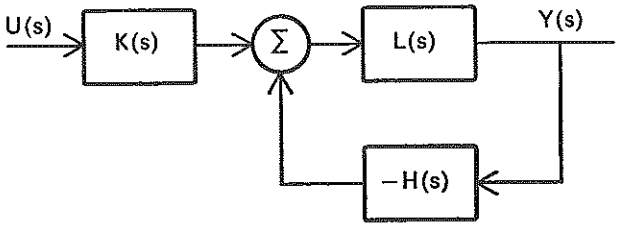


Fig. 3.6.

faktor 10. För $k=50$ varierar det slutna systemets förstärkning mellan gränserna

$$\frac{10}{11} \leqslant G \leqslant \frac{100}{101}$$

Det slutna systemets linjaritet är även mycket god. Med $k=50$ är den största avvikelsen från det linjära sambandet $y=u_r$ given av

$$\max |u_r - y| = \max_y \frac{1}{k} \sqrt{y} = \frac{1}{k} = 0{,}02$$

I fig. 3.5 visas insignal–utsignalrelationen för det öppna och det återkopplade systemet.

Vi skall nu undersöka hur återkopplingen påverkar systemets känslighet för parametervariationer i mer generella fall. Betrakta det system vars konfiguration framgår av blockschemat i fig. 3.6.

Överföringsfunktionen från insignal till utsignal är

$$G(s) = \frac{Y(s)}{U(s)} = \frac{K(s)\,L(s)}{1 + L(s)\,H(s)}$$

För att undersöka hur en systemparameter påverkar den totala överföringsfunktionen $G(s)$ införs begreppet *känslighet* (sensitivity). Med systemets känslighet S_A för variationer i ett delsystem A avses

$$S_A = \frac{dG}{G}\,\frac{A}{dA} = \frac{d \log G}{d \log A}$$

Känsligheten är således en överföringsfunktion. Den anger den relativa variationen i systemets överföringsfunktion vid en förändring av någon av de i systemet ingående överföringsfunktionerna. Känsligheten är ett tekniskt mycket viktigt begrepp då den indikerar vilka krav, som måste ställas på de i systemet ingående komponenterna.

Ur ovanstående ekvation erhålls följande uttryck på de olika *känsligheterna*

$$S_K = \frac{d \log G}{d \log K} = \frac{d \log K}{d \log K} = 1$$

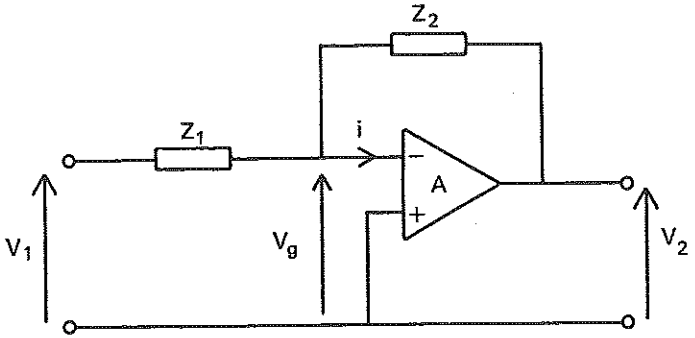


Fig. 3.7. Schematiskt diagram över återkopplad operationsförstärkare.

$$S_L = \frac{d \log G}{d \log L} = \frac{d \log L - d \log (1+LH)}{d \log L} = \frac{1}{1+LH}$$

$$S_H = \frac{d \log G}{d \log H} = -\frac{d \log (1+LH)}{d \log H} = -\frac{LH}{1+LH}$$

Vi finner således att känsligheten S_K alltid är 1. I de frekvensband där kretsförstärkningen är stor ($|LH| \gg 1$) gäller vidare att känsligheten S_L är liten medan känsligheten S_H är nära ett. De komponenter som ingår i L behöver således ej tillverkas med stor noggrannhet utan det är tillfyllest om de komponenter som ingår i K och H är noggranna.

EXEMPEL 3.3. Operationsförstärkaren

Operationsförstärkaren är en direktkopplad likspänningsförstärkare med mycket hög förstärkning. Operationsförstärkaren används ofta för att realisera överföringsfunktioner. Vid dessa tillämpningar kopplas förstärkaren på det sätt som visas i fig. 3.7.

Överföringsfunktionen från förstärkarens inspänning V_g till dess utspänning V_2 är $-A$. Vi skall visa att man med denna koppling kan erhålla en överföringsfunktion, som är mycket stabil även om operationsförstärkarens parametrar skulle variera. Vi skall först skriva upp de ekvationer som beskriver kopplingen i fig. 3.7.

Om vi antar att strömmen i är försumbar gäller

$$(V_1 - V_g)/Z_1 = (V_g - V_2)/Z_2$$

Detta ger efter omskrivning

$$(Z_1 + Z_2)\, V_g = Z_2 V_1 + Z_1 V_2 \tag{3.11}$$

Vidare har förstärkaren förstärkningen $-A$, dvs.

$$V_2 = -A V_g \tag{3.12}$$

Dessa ekvationer kan representeras med blockschemat i fig. 3.8.

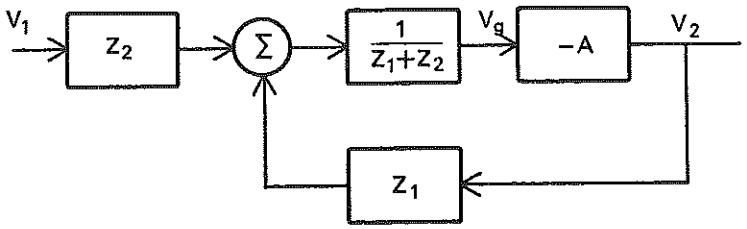


Fig. 3.8. Blockschema för operationsförstärkaren i fig. 3.7.

Ekvationerna (3.11) och (3.12) ger

$$\frac{V_2(s)}{V_1(s)} = -\frac{Z_2(s)}{Z_1(s)} \cdot \frac{1}{1 + Z_2(s)\, Z_1^{-1}(s)\, A^{-1}(s) + A^{-1}(s)} = -G(s) \qquad (3.13)$$

Om överföringsfunktionen A är tillräckligt stor så är uppenbarligen det slutna systemets överföringsfunktion väsentligen $-Z_2/Z_1$.

Av uttrycket (3.13) följer t. ex. att om Z_1 och Z_2 är rena resistanser så är överföringsfunktionen $G(s)$ frekvensoberoende. Om Z_1 är en resistans och Z_2 är kapacitans blir överföringsfunktionen en ren integration. Detta är det vanliga sättet att realisera integrationer i en analogimaskin.

Observera att i fig. 3.8 motsvarar Z_1 och Z_2 passiva komponenter t. ex. resistanser, kapacitanser etc., medan A motsvarar aktiva komponenter, såsom transistorer, rör etc. Det slutna systemets överföringsfunktion bestäms således av de passiva komponenterna. Dessa kan utföras med mycket god stabilitet och temperaturstabiliseras. De aktiva komponenterna är däremot mycket mer komplicerade och har ej så god stabilitet.

Vi skall nu med hjälp av känslighetsbegreppet analysera hur det slutna systemets överföringsfunktion påverkas av variationer i överföringsfunktionen $A(s)$, dvs. variationer i de aktiva komponenterna.

Anta t. ex. att vi vill göra en förstärkare med förstärkningen 100 ggr med hjälp av en operationsförstärkare med $A(s) = \text{konst} = A = 10^5$ där A kan variera med 10 %. Vi väljer då $Z_2(s)/Z_1(s) = 100$. Överföringsfunktionen för det slutna systemet blir då

$$G(s) = 100\, \frac{1}{1 + \dfrac{101}{A}}$$

Känsligheten hos $G(s)$ för variationer i A är då

$$S_A = \frac{d \log G}{d \log A} \approx \frac{101}{A} = 0{,}00101$$

Det slutna systemets överföringsfunktion ändras således endast med 0,01 %, trots att det öppna systemets överföringsfunktion ändras med 10 %.

4. Referenser

Exemplet med ångmaskinreglering är baserat på

Maxwell, J. C., *On Governors*, Proc. Roy. Soc. London *16* (1868), s. 270–283

Vyshnegradskii, J., *Sur la Théorie Générale des Régulateurs*, Compt. Rend. Acad. Sci. Paris *83* (1876), s. 318–321

Vyshnegradskii, J., *Über direkt wirkende Regulatoren*, Der Civilingenieur *23* (1877), s. 95–132.

En liknande analys av turbinreglering återfinns i

Stodola, A., *Über die Regulierung von Turbinen*, Schweiz Bau-Zeitung *22* (1893), s. 113–117, *24* (1894), s. 108–112 och 115–117.

En trevlig framställning av hur känslighet för parametervariationer kan minskas med återkoppling har givits i

Black, H. S., *Stabilized Feedback Amplifiers*, Bell System Techn. Jour. *13* (1934), s. 1–18.

Felkoefficienterna introducerades av Nichols med flera. Se t. ex.

James, H. M., Nichols, N. B. & Phillips, R. S., *Theory of Servomechanisms*, McGraw-Hill, New York 1947.

En grundlig analys av återkopplade elektroniska förstärkare finns i boken

Hammond, P. H., *Feedback Theory and its Application*, The English Universities Press, London 1958.

I denna bok finns många ytterligare referenser.

En redogörelse för felkoefficienterna och deras användning återfinns i nästan alla läroböcker i reglerteknik. Se t. ex. böckerna av Truxal och Chestnut & Meyer som nämndes i avsnitt 2:13.

5. Stabilitet

1. Inledning

Stabilitetsteori har varit ett centralt tema i reglerteorins utveckling. Till en början utvecklades kriterier för att avgöra om ett system är stabilt eller instabilt. Det var sedan naturligt att ange konstruktiva metoder för att modifiera instabila system så att de blir stabila. Många syntesmetoder kan således direkt återföras på metoder för stabilitetsanalys.

Stabilitetsproblem uppträder ej endast i samband med analys av återkopplade system utan även i många andra sammanhang. Vid tillverkning av dynamiska system är det således naturligt att fråga om små avvikelser i systemets koefficienter kan leda till stora förändringar i systemets egenskaper. En närbesläktad fråga är om systemets egenskaper kan påverkas drastiskt genom små störningar. Stabilitetsteorin har således många beröringspunkter med känslighetsanalysen i avsnitt 4:3.

Vi har tidigare i avsnitt 2:11 infört begreppet stabilitet för linjära tidsinvarianta dynamiska system. Stabilitetskriterier för system av låg ordning har även angivits. I detta kapitel skall vi utveckla stabilitetsbegreppet. Vi skall börja med att studera olinjära system, som beskrivs av ekvation (2:2.3) och vi skall införa olika stabilitetsbegrepp i anslutning till dessa. Vi skall således ej börja med att diskutera linjära tidsinvarianta system. Det är ej svårare att göra definitionerna generella och vi vinner även i klarhet. Vid angivande av explicita stabilitetskriterier skall vi dock begränsa oss till specialfallet linjära tidsinvarianta system. Det visar sig att vi i det allmänna fallet ej kan säga att ett system är stabilt. Vi får i stället göra uttalanden om speciella lösningars stabilitet. För linjära tidsinvarianta system gäller dock, att om en speciell lösning är stabil, så är även varje annan lösning stabil. I detta specialfall kan vi således säga att systemet är stabilt. Några olika stabilitetsbegrepp, bl. a. stabilitet, asymptotisk stabilitet och insignal–utsignalstabilitet, införs i avsnitten 2 och 3. Sambandet mellan dessa och det stabilitetsbegrepp som införts i avsnitt 2:11 behandlas även. Avsnitt 4 ger en översikt av metoder för stabilitetsanalys av linjära tidsinvarianta system: rotortmetoden, Routh-Hurwitzkriteriet och Nyquistteoremet. Dessa metoder behandlas detaljerat i avsnitten 5, 6 och 7. I avsnitt 8 behandlas slutligen begreppet praktisk stabilitet.

2. Stabilitet med avseende på störningar i begynnelsevillkor

Betrakta ett system som kan beskrivas med differentialekvationen

$$\frac{dz}{dt} = g(z, t) \tag{2.1}$$

med initialvillkoret

$$z(t_0) = z^0 \tag{2.2}$$

Anta att funktionen $g(z, t)$ är sådan att ekvation (2.1) har en entydig lösning för $t_0 \leqslant t < \infty$. Den entydiga lösningen betecknas med $z(t)$ eller $z(z^0, t_0, t)$, om vi speciellt vill understryka lösningens beroende av initialvillkoret, och kallas den *ostörda* rörelsen. Det är vanligt i stabilitetsteorin att använda orden trajektorie, rörelse och lösning synonymt. Vi skall nu undersöka s. k. *störda* lösningar till ekvation (2.1). Störningar kan uppkomma på flera olika sätt, t. ex. genom att systemet (2.1) ges ett annat initialvärde än (2.2). Avvikelser i hastighet och position vid startraketens brinnslut vid satellituppskjutning är ett typiskt exempel på detta. En annan typ av störning består i att högerledet i ekvation (2.1) ändras så att systemet beskrivs av

$$\frac{dz}{dt} = g(z, t) + h(z, t) \tag{2.3}$$

där funktionen h är liten i någon specificerad mening. Detta kan uppkomma t. ex. på grund av krafter, som försummats vid uppställning av systemekvationerna eller genom att vissa parametrar varierar på grund av toleranser vid tillverkningen. Beroende på störningarnas karaktär kan vi få en rad olika stabilitetsbegrepp. I det följande skall vi undersöka störningar i ekvationens begynnelsevillkor. Orsaken till detta är, att det fall då störningarna uppträder som tilläggstermer i systemekvationens högerled, alltid kan reduceras till fallet med störningar i begynnelsevillkor. För linjära tidsinvarianta system skall detta visas i avsnitt 3. För det allmänna fallet hänvisas till läroböcker om olinjära system.

Definitioner

Låt $\zeta(t)$ vara en störd lösning till ekvation (2.1), som erhållits genom att begynnelsevillkoret (2.2) ändrats. Inför nu skillnaden mellan den störda och den ostörda lösningen

$$x(t) = \zeta(t) - z(z^0, t_0, t) \tag{2.4}$$

Se fig. 2.1. Då $\zeta(t)$ satisfierar ekvation (2.1) finner vi

$$\frac{dx}{dt} = g(\zeta, t) - g(z, t) = g(z + x, t) - g(z, t) = f(x, t) \tag{2.5}$$

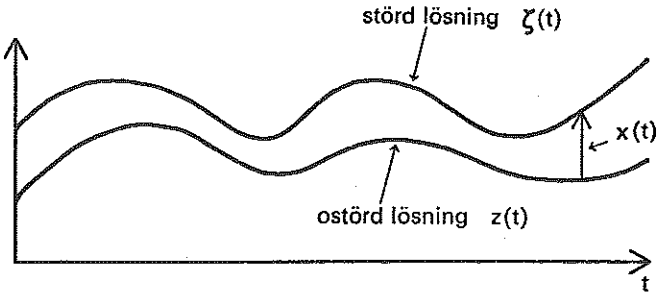


Fig. 2.1. Illustration av störda och ostörda lösningar.

ty $z(z^0, t_0, t)$ är en känd funktion av tiden. Observera att även om funktionen $g(x, t)$ är oberoende av tidsvariabeln t så kan funktionen $f(x, t)$ vara beroende av t.

Den ostörda lösningen motsvarar $x(t)=0$ och det följer av (2.5) att funktionen $f(x, t)$ har egenskapen

$$f(0, t)=0$$

Detta innebär att $x(0, t_0, t)=0$ för $t>t_0$. Om den störda lösningen överensstämmer med den ostörda lösningen vid tidpunkten t_0, så överensstämmer lösningarna således för alla $t \geqslant t_0$.

I stabilitetsteorin är det brukligt att alltid göra transformationen (2.4) och studera stabiliteten av nollösningen $x(0, t_0, t)$ till ekvationen

$$\frac{dx}{dt}=f(x, t)$$

$$x(t_0)=0 \tag{2.6}$$

Vi skall även anta att funktionen $f(x, t)$ är sådan att ekvation (2.6) har en entydig lösning i hela intervallet $t_0 \leqslant t<\infty$ för varje x^0 som ligger tillräckligt nära origo.

Vi kan nu införa ett stabilitetsbegrepp.

Definition 2.1. Stabilitet — Lyapunov

Nollösningen $x(t)=0$ till (2.6) är *stabil* (vid $t=t_0$) om för givet $\varepsilon>0$ det finns ett $\delta=\delta(\varepsilon, t_0)>0$ så att för alla lösningar med initialvillkor $\|x^0\|<\delta$, det gäller att $\|x(x^0, t_0, t)\|<\varepsilon$ för alla $t \geqslant t_0$. En lösning som ej är stabil kallas *instabil*.

I definitionen betecknar $\|x\|$ normen av vektorn x. Definitionen innebär således att nollösningen är stabil om varje störd lösning kan fås att ligga godtyckligt nära nollösningen genom att initialtillståndet väljs tillräckligt nära noll. Observera att vi endast definierar stabilitet för en specifik lösning och ej för ett system, och att stabilitet är ett kontinuitetsbegrepp.

Det är instruktivt att betrakta ett icke-trivialt fall av instabilitet.

EXEMPEL 2.1. Matematisk pendel

Betrakta en matematisk pendel. Låt oss undersöka stabiliteten hos den rörelse som är definierad av initialvillkoren $\phi(t_0)=\alpha$, $\dot{\phi}(t_0)=0$ där $\phi(t)$ är pendelns utslagsvinkel. Vi skall visa att denna rörelse är instabil. Betrakta därför en störd rörelse definierad av $\phi(t_0)=\beta$, $\dot{\phi}(t_0)=0$. Det är bekant från mekaniken att svängningens period beror av initialamplituden. Perioden ökar med ökande amplitud. Se t. ex. Faxén, Dynamik, del II, 2 uppl., s. 129 ff. Vinkelskillnaden mellan den störda och den ostörda rörelsen kommer således åtminstone vid någon tidpunkt att vara större än α, oavsett hur liten differensen $\alpha-\beta$ är, och den ostörda rörelsen är alltså instabil.

Ett skarpare stabilitetsbegrepp ges av följande

DEFINITION 2.2. Asymptotisk stabilitet — Lyapunov

Nollösningen $x(t)=0$ till ekvation (2.6) är *asymptotiskt stabil* (för $t=t_0$) om lösningen är stabil och det finns $\delta>0$ så att för alla lösningar med $\|x^0\|<\delta$ gäller att

$$\|x(x^0, t_0, t)\| \to 0, \text{ då } t-t_0 \to \infty.$$

Asymptotisk stabilitet innebär således förutom stabilitet att lösningen konvergerar mot noll då $t\to\infty$.

I definition 2.1 är i allmänhet $\delta=\delta(\varepsilon, t_0)$ en funktion av initialtidpunkten t_0. Om man i definition 2.1 till varje $\varepsilon>0$ kan finna ett tal δ som är oberoende av t_0 kallas nollösningen *likformigt stabil*. Analogt kan vi även tala om *likformig asymptotisk stabilitet*.

Linjära tidsinvarianta system

Vi skall nu tillämpa definitionerna 2.1 och 2.2 på det linjära tidsinvarianta systemet

$$\frac{dx}{dt}=Ax \tag{2.7}$$

Vi särskiljer tre fall:

A:s alla egenvärden har negativ realdel. Vi finner omedelbart att om samtliga egenvärden till matrisen A har negativ realdel så är nollösningen asymptotiskt stabil. I definition 2.1 finner vi vidare att $\delta(\varepsilon, t_0)$ kan väljas oberoende av t_0. Nollösningen är således likformigt stabil. Vidare finner vi enligt definition 2.2 att lösningen är asymptotiskt stabil. Det gäller speciellt

$$\|x(x_0, t_0, t)\| < K\cdot e^{-\alpha(t-t_0)} \tag{2.8}$$

där $-\alpha>\max \operatorname{Re}\lambda_i$ och λ_i är A:s egenvärden (sats 2:11.1).
Den störda lösningen konvergerar i detta fall mycket snabbt (exponentiellt) mot den ostörda lösningen.

Något av A:s egenvärden har positiv realdel. Nollösningen är då instabil.

A:s alla egenvärden är icke positiva och åtminstone något egenvärde har realdelen noll. Nollösningen kan då vara såväl stabil som instabil. Vi nöjer oss med att ge några exempel.

$$A=\begin{bmatrix}0&0\\0&0\end{bmatrix} \qquad \exp At=\begin{bmatrix}1&0\\0&1\end{bmatrix} \qquad \text{stabil}$$

$$A=\begin{bmatrix}0&1\\0&0\end{bmatrix} \qquad \exp At=\begin{bmatrix}1&t\\0&1\end{bmatrix} \qquad \text{instabil}$$

$$A=\begin{bmatrix}0&1\\-1&0\end{bmatrix} \qquad \exp At=\begin{bmatrix}\cos t&\sin t\\-\sin t&\cos t\end{bmatrix} \qquad \text{stabil}$$

$$A=\begin{bmatrix}0&1&0\\-1&0&1\\0&0&0\end{bmatrix} \qquad \exp At=\begin{bmatrix}\cos t&\sin t&1-\cos t\\-\sin t&\cos t&\sin t\\0&0&1\end{bmatrix} \qquad \text{stabil}$$

$$A=\begin{bmatrix}0&1&0\\-1&0&0\\1&0&0\end{bmatrix} \qquad \exp At=\begin{bmatrix}\cos t&\sin t&0\\-\sin t&\cos t&0\\\sin t&1-\cos t&1\end{bmatrix} \qquad \text{stabil}$$

Stabila system

Betrakta en lösning till (2.7) som ej är en nollösning, säg

$$\frac{dz}{dt}=Az$$

$$z(0)=a$$

För att undersöka om denna lösning är stabil studerar vi en störd lösning

$$\frac{d\zeta}{dt}=A\zeta$$

$$\zeta(0)=\alpha$$

Inför nu enligt (2.4)

$$x(t)=\zeta(t)-z(t)$$

Vi finner då

$$\frac{dx}{dt}=\frac{d\zeta}{dt}-\frac{dz}{dt}=A(\zeta-z)=Ax$$

$$x(0)=\alpha-a$$

För linjära tidsinvarianta system gäller således att om nollösningen är stabil (asymptotiskt stabil eller instabil) så är varje annan lösning stabil (asympto-

tiskt stabil eller instabil). *För linjära tidsinvarianta system är stabilitet således en egenskap hos systemet* och ej en egenskap som hänför sig till en speciell lösning. Systemet (2.7) är asymptotiskt stabilt om och endast om A:s alla egenvärden har negativ realdel.

3. Insignal–utsignalstabilitet

I föregående avsnitt definierades stabilitet med avseende på störningar i systemets initialvillkor. Vi nämnde även att man kunde definiera andra stabilitetsbegrepp genom att införa störningar i ekvationens högerled. I detta avsnitt skall vi genomföra detta för linjära tidsinvarianta system.

DEFINITION 3.1. Insignal–utsignalstabilitet

Ett linjärt system är *insignal–utsignalstabilt* (*i–u*-stabilt) om en begränsad insignal ger en begränsad utsignal för varje initialtillstånd.

Vi skall nu undersöka några av de samband som finns mellan de olika stabilitetsbegreppen. Vi har

SATS 3.1

Betrakta det linjära tidsinvarianta systemet

$$\frac{dx}{dt} = Ax + Bu \tag{3.1}$$

där x är systemets utsignal. Om samtliga egenvärden till A har negativ realdel så är systemet insignal–utsignalstabilt.

Bevis:

Ekvation (3.1) har lösningen

$$x(t) = e^{A(t-t_0)}x^0 + \int_{t_0}^{t} e^{A(t-s)}Bu(s)\,ds \tag{3.2}$$

Jämför ekvation (2:4.8).

Då gäller

$$\|x(t)\| \leqslant \|e^{A(t-t_0)}\|\,\|x^0\| + \int_{t_0}^{t} \|e^{A(t-s)}\|\,\|B\|\,\|u(s)\|\,ds \tag{3.3}$$

där $\|A\|$ betecknar den med vektornormen $\|x\|$ associerade matrisnormen.

Då samtliga egenvärden till matrisen A har negativ realdel, säg $\mathrm{Re}\,\lambda(A) < -\alpha$, gäller enligt sats 2:11.1

$$\|e^{At}\| < K \cdot e^{-\alpha t}, \quad \alpha > 0$$

Vi finner således

$$\|x(t)\| < Ke^{-\alpha(t-t_0)}\|x^0\| + \int_{t_0}^{t} Ke^{-\alpha(t-s)}\|B\|\|u(s)\|\,ds$$

$$< Ke^{-\alpha(t-t_0)}\|x^0\| + \frac{K}{\alpha}\{1-e^{-\alpha(t-t_0)}\}\|B\| \max_{t_0\leqslant s\leqslant t}\|u(s)\| \tag{3.4}$$

Om insignalen är begränsad, dvs.

$$\max_s \|u(s)\| < K_1$$

så är även utsignalen begränsad och satsen är därmed bevisad.

Sats 3.1 innebär således att för linjära tidsinvarianta system medför asymptotisk stabilitet av nollösningen att systemet är insignal–utsignalstabilt. Vi skall med hjälp av ett motexempel visa att omvändningen av denna sats ej är sann.

EXEMPEL 3.1

Betrakta det system som beskrivs av

$$\frac{dx}{dt} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t) \tag{3.5}$$

där utsignalen antas vara tillståndsvektorn. Ovanstående ekvation har lösningen

$$x(t) = \begin{bmatrix} x_1^0 \cos(t-t_0) + x_2^0 \sin(t-t_0) \\ -x_1^0 \sin(t-t_0) + x_2^0 \cos(t-t_0) \\ x_3^0 e^{t_0-t} + \int_{t_0}^{t} e^{s-t}u(s)\,ds \end{bmatrix} \tag{3.6}$$

Detta system är uppenbart insignal–utsignalstabilt. Det ostörda systemets nollösning ($u(t)=0$, $x(t)=0$) är dock ej asymptotiskt stabilt, ty tillstånden x_1 och x_2 konvergerar ej mot noll då t växer. Om tillstånden x_1 och x_2 kunde påverkas av insignalen u så skulle x_1 och x_2 kunna ges mycket stora värden även med en liten insignal, t. ex. genom att man väljer u periodisk med perioden 2π. I exemplet är emellertid x_1 och x_2 ej styrbara.

Vi skall vidare med hjälp av ett motexempel visa att motsvarigheten till sats 3.1 ej gäller för tidsvariabla linjära system.

EXEMPEL 3.2

Betrakta det system som beskrivs av ekvationen

$$\frac{dx}{dt} = -\frac{1}{2t}x + u(t) \tag{3.7}$$

Det fria systemet

$$\frac{dx}{dt} = -\frac{1}{2t}\,x$$

har lösningen

$$x(t) = \sqrt{t_0/t}\; x^0$$

Nollösningen är således asymptotiskt stabil men ej *likformigt* asymptotiskt stabil. Vi finner t. ex. genom användning av Lagranges metod med variation av konstanten att den störda ekvationen (3.7) har lösningen

$$x(t) = \sqrt{t_0/t}\; x^0 + \int_{t_0}^{t} \sqrt{s/t}\; u(s)\, ds$$

Anta nu $u(t) = a =$ konstant. Vi finner då

$$x(t) = x^0 \sqrt{t_0/t} + (2/3)\, at - (2/3)\, at_0 \sqrt{t_0/t}$$

Utsignalen är således ej begränsad och systemet är ej insignal–utsignalstabilt.

4. Översikt av metoder för stabilitetsanalys av linjära tidsinvarianta system

Vi skall i de avsnitt som följer utveckla systematiska metoder för att undersöka om ett dynamiskt system är stabilt. Först måste vi bestämma vilket stabilitetsbegrepp som skall användas. Från teknisk synpunkt är det rimligt att kräva åtminstone insignal–utsignalstabilitet, ty vi vill gardera oss mot att små störningar leder till att systemets lösningar förändras drastiskt. För att undvika sådana fall som representeras av exempel 3.1 skall vi kräva *asymptotisk stabilitet* i Lyapunovs mening. När vi i fortsättningen talar om stabilitet avses således asymptotisk stabilitet enligt definition 2.2 om ej annat sägs. Observera att detta val är starkt betingat av förutsättningen om tidsinvarianta system. Exempel 3.2 visar att för tidsvariabla system medför asymptotisk stabilitet ej insignal–utsignalstabilitet. Observera även att den valda definitionen 2.2 kan appliceras direkt på flervariabla system, medan en undersökning av insignal–utsignalstabilitet kräver undersökning av alla möjliga par av insignaler och utsignaler.

Vi fann i avsnitt 2 att asymptotisk stabilitet var ekvivalent med villkoret att samtliga egenvärden till systemmatrisen A har negativ realdel. Det stabilitetsbegrepp vi skall använda oss av överesstämmer alltså med det begrepp som infördes i avsnitt 2:11.

För att undersöka om ett system är stabilt behöver vi således metoder som avgör om den karakteristiska ekvationens rötter ligger i vänstra halvplanet. Vi kan indela metoderna i två kategorier, *direkta* respektive *indirekta* metoder.

Stabilitetskriterier baserade på direkta metoder utgår från en direkt lösning av den karakteristiska ekvationen. I de indirekta metoderna avgörs i stället stabiliteten utan att den karakteristiska ekvationen löses.

Vad som hittills sagts gäller för linjära tidsinvarianta dynamiska system i allmänhet. För återkopplade system finns ytterligare en synpunkt att beakta. Det är bekvämt att kunna avgöra stabiliteten med hjälp av det öppna systemets egenskaper (avsnitt 4:3). Stabilitetskriterierna bör således formuleras som villkor på kretsöverföringen. Ett annat skäl för detta är att kretsöverföringen ofta direkt motsvarar systemets fysikaliska komponenter. En vanlig metod att modifiera ett reglersystem är att införa ett filter efter felsignalgivaren. Detta är ekvivalent med att kretsöverföringen multipliceras med filtrets överföringsfunktion. Vid en sådan förändring ändras kretsöverföringen på ett enkelt sätt (multiplikativt), medan ändringen i det återkopplade systemets överföringsfunktion är betydligt mer komplicerad.

Följande metoder skall undersökas:

Direkta metoder

- Numerisk lösning av den karakteristiska ekvationen
- Rotortmetoden

Indirekta metoder

- Routh–Hurwitzkriteriet
- Argumentvariationsprincipen och Nyquistkriteriet

En numerisk lösning av den karakteristiska ekvationen kan endast göras för hand med system av låg ordning, t. ex. fjärde och femtegradsekvationer. Om man har tillgång till en datamaskin så är det naturligtvis möjligt att undersöka högre gradtal med hjälp av standardalgoritmer. En direkt lösning ger sällan insyn i hur den karakteristiska ekvationens rötter beror av systemets parametrar. Detta är en mycket stor nackdel då man i reglertekniska problem ofta har intresse av att veta lösningen för flera parametervärden. Man vill t. ex. veta hur karakteristiska ekvationens rötter varierar med förstärkningen i återkopplingskretsen.

Låt oss slutligen nämna att det finns en generell metod — Lyapunovs andra metod — som kan användas även för olinjära system. En behandling av denna går emellertid utanför ramen av denna bok.

5. Rotortmetoden

Allmänt

Rotortmetoden (the root locus method) som introducerades av Evans 1948, är en elegant grafisk metod att lösa algebraiska ekvationer, t. ex.

$$A(s) = s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_n = 0 \tag{5.1}$$

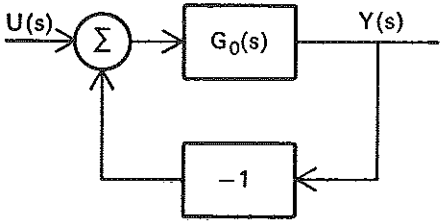


Fig. 5.1. Blockschema för enkelt återkopplat system.

Metoden baseras på *inbäddning* (imbedding) och går ut på att man, i stället för att lösa ekvationen för fixa värden på parametrarna, löser ekvationen för samtliga värden på någon parameter (eller parameterkombination). Vi kan t. ex. skriva ekvationen (5.1) som

$$A(s)=P(s)+Q(s)=0 \tag{5.2}$$

där (detta är endast ett *exempel* på uppdelning)

$$P(s)=s^n+a_1s^{n-1}+\ldots+a_ms^{n-m}$$

$$Q(s)=a_{m+1}s^{n-m-1}+a_{m+2}s^{n-m-2}+\ldots+a_n$$

I stället för att lösa ekvationen

$$A(s)=P(s)+1\cdot Q(s)=0 \tag{5.3}$$

löser vi ekvationen

$$A_K(s)=P(s)+K\cdot Q(s)=0 \tag{5.4}$$

för positiva reella värden på parametern K. Vi kan således säga att det speciella problemet ($K=1$) bäddats in i ett allmännare problem (positiva reella K). Därav namnet inbäddning. Det anmärkningsvärda är att det allmännare problemet är lättare att lösa.

Vi skall nu införa

DEFINITION 5.1

Med *rotorten* för ekvation (5.4) med avseende på parametern K avses den geometrisk orten för ekvationens rötter då parametern K varierar över intervallet $(0, \infty)$.

För att lösa ekvationen (5.1) med hjälp av rotortmetoden ritar vi således rotorten för ekvation (5.4). Genom att sedan i den erhållna figuren avläsa rötternas läge för $K=1$ har vi fått en lösning till det ursprungliga problemet.

Användning av rotorten för analys av återkopplade system

Rotortmetoden är mycket användbar då den ger en god överblick av hur den karakteristiska ekvationens rötter varierar med någon systemparameter. Vi illustrerar med ett exempel.

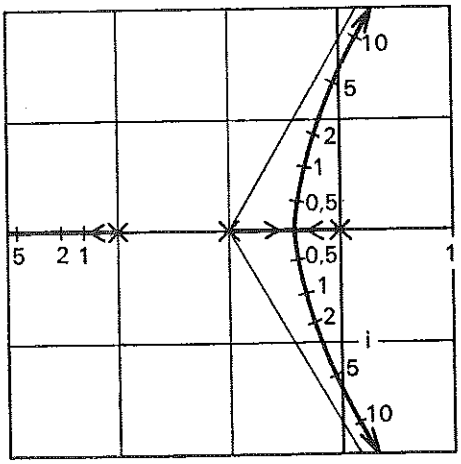


Fig. 5.2. Rotorten för ekvationen $s^3 + 3s^2 + 2s + K = 0$ då parametern K antar positiva reella värden.

EXEMPEL 5.1

Betrakta det reglersystem vars blockschema visas i fig. 5.1. Låt kretsöverföringen vara

$$G_0(s) = KQ(s)/P(s)$$

där K betecknar förstärkningsparametern. Studera hur det slutna systemets egenskaper varierar med förstärkningsparametern K.

Överföringsfunktionen från insignal till utsignal (»det slutna systemets överföringsfunktion») är

$$\frac{Y(s)}{U(s)} = \frac{G_0(s)}{1 + G_0(s)} = \frac{KQ(s)}{P(s) + KQ(s)} \tag{5.5}$$

Systemet har följande karakteristiska ekvation

$$P(s) + KQ(s) = 0 \tag{5.6}$$

Vi finner att det slutna systemets överföringsfunktion (5.5) har samma nollställen som det öppna systemet, men att det slutna systemets poler är givna av den karakteristiska ekvationens (5.6) rötter. Genom att rita rotorten för den karakteristiska ekvationen (5.6) får vi således veta hur polerna till det slutna systemets överföringsfunktion beror av förstärkningsparametern K. Med hjälp av den karakteristiska ekvationens rotort kan vi kvalitativt bedöma det slutna systemets egenskaper. Vi illustrerar detta genom att specifikt anta

$$G_0(s) = K\frac{Q(s)}{P(s)} = \frac{K}{s(s+1)(s+2)} \tag{5.7}$$

Det öppna systemet har således inga nollställen och enkelpoler i $s = 0$, $s = -1$ och $s = -2$. Den karakteristiska ekvationen lyder

$$P(s) + KQ(s) = s^3 + 3s^2 + 2s + K = 0 \tag{5.8}$$

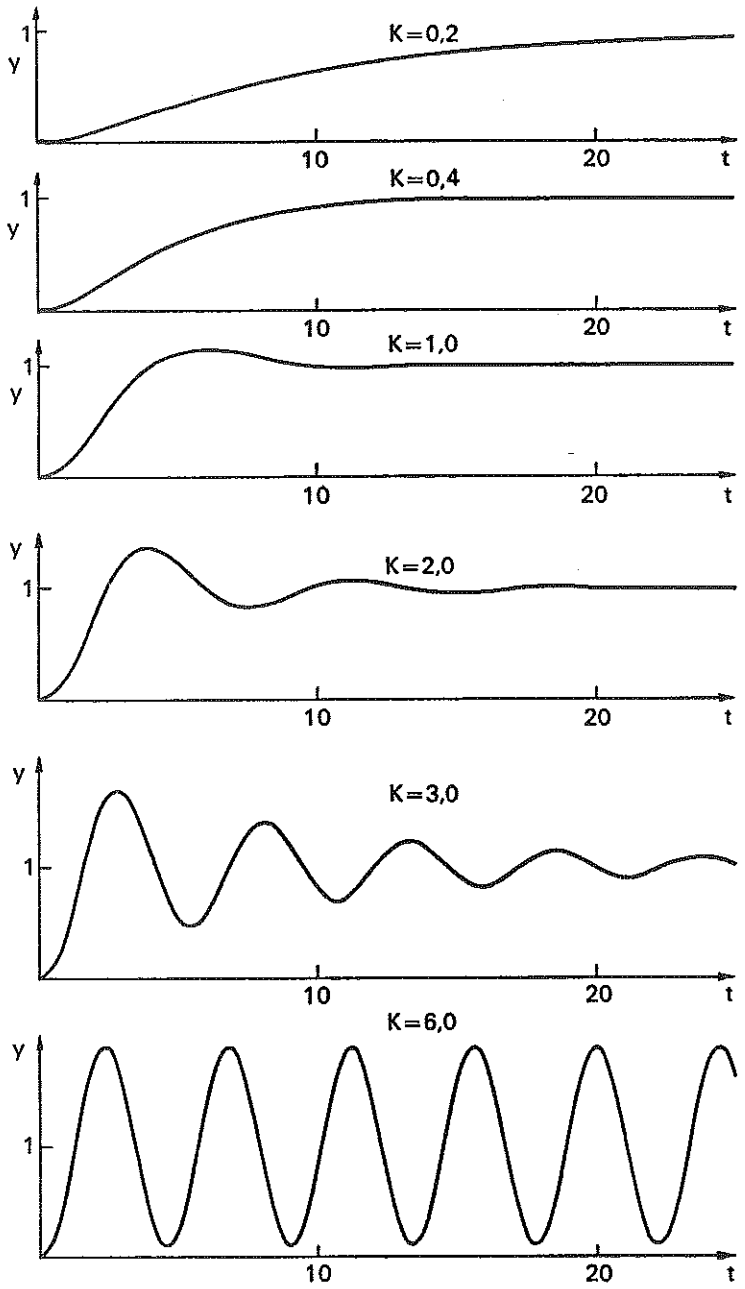


Fig. 5.3. Stegfunktionssvar för enkelt återkopplat system med olika värden på kretsförstärkningen K. Systemets kretsöverföring är

$$G_0 = \frac{K}{s(s+1)(s+2)}$$

Fig. 5.2 visar rotorten för denna ekvation. Med hjälp av rotortkurvan i fig. 5.2 kan vi kvalitativt bedöma det återkopplade systemets egenskaper för olika värden på förstärkningsparametern K. Vid låga värden, $K<0{,}4$ har den karakteristiska ekvationen (5.8) endast reella rötter. Det slutna systemets viktfunktion är då sammansatt av tre exponentialfunktioner med reella exponenter, där den ena exponenten är $> -0{,}5\,t$. Då $K>0{,}4$ har den karakteristiska ekvationen två komplexa rötter och systemet är oscillativt. Med växande värden på K blir systemet allt svagare dämpat och för $K=6$ blir systemet instabilt. I fig. 5.3 visas systemets stegfunktionssvar för olika värden på förstärkningsparametern K. Lämpliga värden är $K=0{,}8$–$1{,}5$.

Uppritning av rotorten

Rotorten definierades som geometriska orten för rötterna till ekvation

$$P(s)+KQ(s)=0 \tag{5.4}$$

då $0\leqslant K<\infty$. I denna ekvation är $P(s)$ och $Q(s)$ polynom av den komplexa variabeln s. För att endast behöva arbeta med reella storheter kan vi skriva ekvationen (5.4) som

$$\arg Q(s)-\arg P(s)=\pi+2k\pi,\ k=0,\ 1,\ 2,\ \ldots \tag{5.9}$$

$$K=|P(s)|/|Q(s)| \tag{5.10}$$

Ekvation (5.9) kallas *argumentvillkoret* och ekvation (5.10) kallas *beloppvillkoret*. Alla punkter s som satisfierar argumentvillkoret ligger således på rotorten. Beloppvillkoret är användbart för att bestämma värdet av K i en punkt på rotorten. För att rita rotorten begagnar man sig av ekvationerna (5.4) och (5.9). För att parametrisera rotorten med hjälp av K-värden används sedan (5.10). Uppritning av rotorten förenklas av följande iakttagelser

- *Symmetri.* Då polynomen $P(s)$ och $Q(s)$ har reella koefficienter är rotorten symmetrisk med avseende på den reella axeln.
- *Antal grenar.* Enligt algebrans fundamentalsats har en n:te gradsekvation n rötter. Rotorten har således flera kurvgrenar. Antalet grenar är lika med det största av gradtalen för $P(s)$ och $Q(s)$.
- *Start- och ändpunkter.* Genom att låta $K\to 0$ och $K\to\infty$ i ekvation (5.4) finner vi att nollställena till polynomen $P(s)$ och $Q(s)$ ligger på rotorten. Rotortkurvorna börjar således i P:s nollställen ($K=0$) och slutar i Q:s nollställen ($K=\infty$) eller i oändligheten.
- *Nära start- och ändpunkter.* I närheten av nollställen till P eller Q erhålls rotorten lätt ur argumentvillkoret. För att demonstrera detta skriver vi polynomen som

  $$P(s)=\alpha_1(s-p_1)(s-p_2)\ \ldots\ (s-p_n)$$

  $$Q(s)=\beta_1(s-z_1)(s-z_2)\ \ldots\ (s-z_m)$$

där multipla nollställen skrivs som multipla faktorer. P-polynomets nollställen p_i ligger på rotorten. För att undersöka rotortens utseende i närheten av $s=p_i$ användes argumentvillkoret (5.9) och vi får

$$\arg\,(s-p_i)=\pi+\sum_{j=1}^{m}\arg\,(s-z_j)-\sum_{\substack{j=1\\ j\neq i}}^{n}\arg\,(s-p_j)+2k\pi$$

för

$$|s-p_i|\ll\min\{\min_j|p_i-z_j|,\ \min_{j\neq i}|p_i-p_j|\}$$

kan argumentvillkoret då skrivas (observera den geometriska tolkningen)

$$\arg\,(s-p_i)\sim\pi+\sum_{j=1}^{m}\arg\,(p_i-z_j)-\sum_{\substack{j=1\\ j\neq i}}^{n}\arg\,(p_i-p_j)+2k\pi \tag{5.11}$$

Speciellt finner vi att delar av reella axeln satisfierar argumentvillkoret. Till höger om varje punkt på rotorten måste det finnas ett udda antal nollställen till $P(s)$ och $Q(s)$.

● *Asymptoter.* Vi kan även ur argumentvillkoret lätt erhålla rotortens utseende för stora värden på $|s|$. Polynomen $P(s)$ och $Q(s)$ skrivs

$$P(s)=s^n+a_1s^{n-1}+\ldots+a_n$$

$$Q(s)=s^m+b_1s^{m-1}+\ldots+b_m$$

Det är ingen förlust i allmängiltighet att anta att koefficienterna för högsta ordningens termer är 1, ty en annan faktor kan alltid inkluderas i K. Argumentvillkoret kan nu skrivas

$$\arg\frac{s^m+b_1s^{m-1}+\ldots+b_m}{s^n+a_1s^{n-1}+\ldots+a_n}=\pi+2k\pi,\ k=0,1,2,\ldots$$

Antas t. ex. $n>m$ erhålls

$$\arg\,[s^{n-m}+(a_1-b_1)s^{n-m-1}+\ldots]=\pi+2k\pi,\ k=0,1,2,\ldots$$

För stora $|s|$, dvs. då $|s|\gg\max\{|p_i|,|z_i|\}$, gäller

$$\arg\,[s^{n-m}+(a_1-b_1)s^{n-m-1}+\ldots]=$$

$$=\arg\left[\left(s+\frac{a_1-b_1}{n-m}\right)^{n-m}+\alpha_2s^{n-m-2}+\ldots\right]$$

$$=\arg\left(s+\frac{a_1-b_1}{n-m}\right)^{n-m}+\arg\left[1+\frac{\beta_2}{s^2}+\ldots\right]$$

$$\sim\arg\left(s+\frac{a_1-b_1}{n-m}\right)^{n-m}+\arg\,1$$

Argumentvillkoret lyder då

$$\arg\left(s+\frac{a_1-b_1}{n-m}\right)\sim\frac{\pi}{n-m}+2k\frac{\pi}{n-m},\ k=1,2,\ldots \tag{5.12}$$

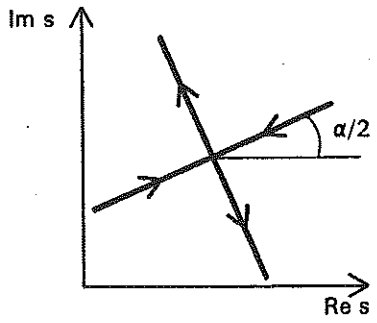


Fig. 5.4. Rotortens utseende i närheten av en dubbelrot.

För stora värden på $|s|$ har rotorten således rätlinjiga asymptoter, som är symmetriskt orienterade.

Samband mellan rötter och koefficienter ger

$$a_1 = -\sum_{i=1}^{n} p_i \text{ och } b_1 = -\sum_{i=1}^{m} z_i$$

Asymptoternas skärningspunkt är

$$s_1 = -\frac{a_1 - b_1}{n-m} = \frac{1}{n-m}\left(\sum_{i=1}^{n} p_i - \sum_{i=1}^{m} z_i\right) \tag{5.13}$$

Detta kan således tolkas fysikaliskt som masscentrum till den massfördelning som har positiva enhetsmassor i nollställena till $P(s)$ och negativa enhetsmassor i nollställena till $Q(s)$.

● *Multipelrötter.* Låt s_0 vara en punkt på rotorten svarande mot parametervärdet K_0. Utvecklas den karakteristiska ekvationen (5.4) i Taylorserie erhålls

$$P(s_0) + K_0 Q(s_0) + (s-s_0)[P'(s_0) + K_0 Q'(s_0)] + \tfrac{1}{2}(s-s_0)^2[P''(s_0) + K_0 Q''(s_0)] + \ldots$$
$$+ (K-K_0)Q(s_0) + (K-K_0)(s-s_0)Q'(s_0) + \ldots = 0 \tag{5.14}$$

De första två termerna försvinner, ty punkten s_0 är en punkt på rotorten svarande mot K_0. Om $P'(s_0) + K_0 Q'(s_0) \neq 0$ gäller således i närheten av s_0 följande approximativa ekvation för rotorten

$$s = s_0 - (K-K_0)Q(s_0)[P'(s_0) + K_0 Q'(s_0)]^{-1} \tag{5.15}$$

Om däremot s_0 är en dubbelrot så försvinner även den tredje termen i ekvation (5.14) .I närheten av s_0 gäller då approximativt följande ekvation för rotorten

$$(s-s_0)^2 + 2(K-K_0)Q(s_0)[P''(s_0) + K_0 Q''(s_0)]^{-1} = 0 \tag{5.16}$$

dvs.

$$\arg(s-s_0) = \tfrac{1}{2}\arg(K_0 - K) + \alpha/2 + k\pi$$

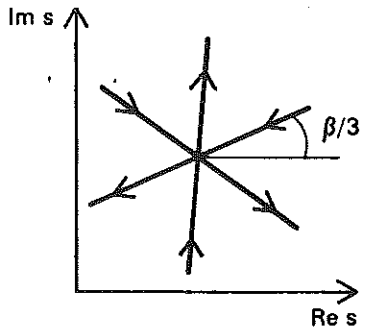


Fig. 5.5. Rotortens utseende i närheten av en trippelrot.

där

$$\alpha = \arg Q(s_0) - \arg [P''(s_0) + K_0 Q''(s_0)] \tag{5.17}$$

eller

$$\arg (s - s_0) = \begin{cases} \alpha/2 + k\pi & K < K_0 \\ (\alpha + \pi)/2 + k\pi & K > K_0 \end{cases}, \quad k = 1, 2, \ldots \tag{5.18}$$

I närheten av en dubbelrot har rotorten således den form som visas i fig. 5.4. På samma sätt finner vi att i närheten av en trippelrot har rotorten den form som visas i fig. 5.5.

EXEMPEL 5.2

Skissa rotorten för ekvationen

$$s(s+1)(s+2) + K = 0$$

Då ekvationen är av tredje graden har rotorten tre grenar. Dessa börjar i $s = 0$, $s = -1$ och $s = -2$. Det följer vidare av ekvation (5.11) att i närheten av $s = 0$ gäller arg $s \sim \pi$, i närheten av $s = -1$ gäller arg $(s+1) \sim 0$ och i närheten av $s = -2$ gäller arg $(s+2) \sim \pi$. Vi verifierar sedan med hjälp av argumentvillkoret (5.9) att intervallen $(-\infty, -2)$ och $(-1, 0)$ tillhör rotorten. På intervallet $(-1, 0)$ möts två kurvgrenar. Där måste då finnas en dubbelrot till den karakteristiska ekvationen. Den karakteristiska ekvationens derivata har då en enkelrot vilket medför att

$$\frac{d}{ds}\{s(s+1)(s+2) + K\} = 3s^2 + 6s + 2 = 0$$

Denna ekvation har lösningen $s = -1 \underset{(-)}{\overset{+}{}} \sqrt{3}/3$. Beloppvillkoret (5.10) ger $K = 2\sqrt{3}/9 = 0{,}386$.

Vi finner vidare att talet α definierat av ekvation (5.17) är noll. Rotortens utseende i närheten av dubbelroten framgår då av fig. 5.4.

Vi undersöker nu asymptoterna. För stora värden på $|s|$ gäller enligt (5.12)

$$\arg (s+1) \sim \pi/3 + 2k\pi/3, \quad k = 0, 1, 2$$

Rotorten har således för stora $|s|$ tre asymptoter som skär varandra i $s=-1$. Vi kan nu skissera rotorten och erhåller då något som liknar fig. 5.2. För att bestämma skärningen med imaginära axeln införes $s=i\omega$ och vi får

$$i\omega(2-\omega^2)+(K-3\omega^2)=0$$

Denna ekvation har lösningen $\omega=\pm\sqrt{2}$ och $K=6$.

Det är efter viss övning mycket lätt att skissa rotortkurvor även för system av hög ordning. En skiss är ofta tillräcklig. Att rita en rotort noggrant är mycket tidsödande. För detta ändamål finns emellertid utmärkta datamaskinprogram. Några av dessa program är anmärkningsvärda därigenom att de ej baseras på algoritmer för lösning av algebraiska ekvationer. I stället undersöker man systematiskt ett givet punktgitter i s-planet och testar om argumentvillkoret är uppfyllt. För punkter som satisfierar argumentvillkoret beräknas sedan på vanligt sätt parametern K ur beloppvillkoret.

Funktionsteoretisk tolkning av rotorten

Betrakta avbildningen $s\to G=Q(s)/P(s)$. För att representera denna avbildning en-entydigt måste vi ha ett antal Riemann-ytor i s-planet, lika många som det största av gradtalen av $P(s)$ och $Q(s)$. Rotorten som definieras av (5.9) dvs.

$$\arg G=\arg Q(s)-\arg P(s)=\pi+2k\pi \tag{5.19}$$

är således bilden i s-planet av negativa reella axeln i G-planet. Vi får således en kurvgren för varje Riemannyta.

Funktionen $\log|G(s)|$ är en potentialfunktion. Kurvorna

$$\log|G(s)|=\text{konstant} \tag{5.20}$$

kan således tolkas som ekvipotentialytor. Kurvskaran

$$\arg G(s)=\text{konstant} \tag{5.21}$$

är ortogonal mot kurvskaran (5.20) och representerar således strömlinjer. Rotorten definierad av (5.19) kan alltså tolkas som en speciell strömlinje.

6. Routh–Hurwitzkriteriet

Vi skall nu övergå till de indirekta metoderna för stabilitetsundersökning. För linjära stationära system kan, som visats i avsnitt 2, stabilitetsproblemet reduceras på det rent algebraiska problemet genom att man undersöker huruvida en algebraisk ekvation har alla rötter i det vänstra halvplanet. Betydelsen av detta algebraiska problem för reglertekniken underströks i såväl

Maxwells som Vyshnegradskiis tidiga arbeten. I själva verket var problemet löst av Hermite redan år 1856, men hans resultat var ej på en sådan form att de direkt kunde tillämpas. År 1875 publicerade den engelske matematikern Routh en algoritm för problemets lösning. Problemet löstes oberoende av Hurwitz år 1895. Hurwitz fick problemet av Stodola, som ovetande om de tidigare resultaten formulerat problemet i samband med analys av ångturbinreglering. Det fundamentala resultatet är nu känt som Routh–Hurwitz' olikheter. Ett ytterligare bidrag lämnades av de franska matematikerna Liénard och Chipart (1914), som visade att Routh–Hurwitz' olikheter är beroende och att flera av dessa kan utelämnas.

Rouths algoritm

Betrakta ekvationen

$$F(z)=a_0z^n+b_0z^{n-1}+a_1z^{n-2}+b_1z^{n-3}+\ldots=0 \tag{6.1}$$

Det säregna valet av beteckningar för ekvationens koefficienter skall senare visa sig vara praktiskt.

Anta att koefficienterna a_i, b_i är reella och att a_0 är positiv. Vi skall nu undersöka vilka villkor som skall ställas på koefficienterna a_i, b_i för att samtliga rötter till ekvationen skall ha negativ realdel. Då koefficienterna är reella vet vi att ekvationens rötter antingen är reella eller parvis komplexkonjugerade. Ekvationens vänsterled faktoreras då i termer av typen $(s+\alpha)$ eller $s^2+\beta s+\gamma$. Om samtliga rötter skall ligga i vänstra halvplanet så måste det alltså gälla att $\alpha>0$, $\beta>0$ och $\gamma>0$. Vi får då omedelbart det nödvändiga villkoret att alla koefficienter måste vara positiva. Ty en produkt av polynom med positiva koefficienter är ett polynom med positiva koefficienter. Villkoret att samtliga koefficienter är positiva är dock ej tillräckligt för att alla rötter skall ha negativ realdel. För att erhålla ett tillräckligt villkor förfar vi på följande sätt.

Inför tablån

$$\begin{array}{cccc} a_0 & a_1 & a_2 & \ldots \\ b_0 & b_1 & b_2 & \ldots \\ c_0 & c_1 & c_2 & \ldots \\ d_0 & d_1 & d_2 & \ldots \\ \vdots & & & \end{array} \tag{6.2}$$

där

$$\begin{aligned} c_0&=a_1-a_0b_1/b_0 \\ c_1&=a_2-a_0b_2/b_0 \\ &\vdots \\ d_0&=b_1-b_0c_1/c_0 \\ d_1&=b_2-b_0c_2/c_0 \\ &\vdots \end{aligned} \tag{6.3}$$

Varje rad i tablån erhålls således ur de två föregående. Övre raden minskas med undre raden multiplicerad med ett tal sådant att första differensen försvinner. Den nya raden erhålls genom att det första försvinnande elementet utelämnas. Om vid subtraktionen något element fattas utfylls tablån med nollor. Proceduren upprepas tills $(n+1)$ rader erhållits. Vi visar nedan hur tablån byggs upp i fallen $n=4$ och $n=5$.

$n=4$			$n=5$		
x	x	x	x	x	x
x	x	0	x	x	x
x	x		x	x	0
x	0		x	x	
x			x	0	
			x		

I ovanstående uppställning betecknar x ett tal och 0 betecknar de nollor som införts vid tablåns uppbyggnad.

SATS 6.1. Routh

Antalet teckenväxlingar hos sviten $a_0, b_0, c_0, d_0 \ldots$ är lika med det antal nollställen som polynomet $F(z)$ har i högra halvplanet $\operatorname{Re} z > 0$. Alla nollställen till polynomet $F(z)$ ligger i vänstra halvplanet om samtliga tal $a_0, b_0, c_0, d_0, \ldots$ är positiva.

För bevis av denna sats hänvisas till någon lärobok i matristeori, t. ex. Gantmacher, *The Theory of Matrices*, Vol. II, kap. 15.

Vi ser att den givna regeln för att bygga upp Rouths tablå bryter samman om någon av koefficienterna $a_0, b_0, c_0, \ldots$ försvinner. Om vi endast skall avgöra om polynomet har alla rötter i vänstra halvplanet är svaret givet (nej). Om vi däremot skall bestämma hur många rötter som ligger i högra halvplanet måste vi ha kompletterande regler. Sådana regler har givits av Routh. Vi säger att ekvationen är *reguljär* om alla tal $a_0, b_0, c_0, d_0, \ldots$ är skilda från noll.

EXEMPEL 6.1

Ett system har den karakteristiska ekvationen

$$z^5 + 2z^4 + 10z^3 + 30z^2 + 100z + 360 = 0$$

Är systemet stabilt? Vi bildar Rouths tablå

1	10	100
(2	30	360)
1	15	180
(−5	−80)	
−1	−16	
−1	180	
(−196	0)	
−1	0	
180		

Observera att vi kan multiplicera en rad med ett godtyckligt, positivt tal utan att ändra antalet teckenväxlingar. Detta är användbart för att förenkla räknearbetet. Då en ny rad erhållits på detta sätt har den rad ur vilken den nya raden bildats satts inom parantes i tablån ovan.

I första kolonnen finns två teckenväxlingar, ekvationen har således två rötter i högra halvplanet och systemet är då instabilt.

Vi skall slutligen ge en regel som kan vara ett stöd för att memorera Rouths tablå. Om vi inför polynomen

$$A(z)=a_0z^n+a_1z^{n-2}+\dots$$

$$B(z)=b_0z^{n-1}+b_1z^{n-3}+\dots$$

$$C(z)=c_0z^{n-2}+c_1z^{n-3}+\dots$$

$$D(z)=d_0z^{n-3}+d_1z^{n-5}+\dots$$

där koefficienterna c_i och d_i är definierade av (6.3), så finner vi att den rationella funktionen $A(z)/B(z)$ kan skrivas på följande sätt

$$\frac{A(z)}{B(z)}=\frac{a_0}{b_0}z+\frac{C(z)}{B(z)}=\frac{a_0}{b_0}z+\frac{1}{\dfrac{B(z)}{C(z)}}$$

$$=\frac{a_0}{b_0}z+\frac{1}{\dfrac{b_0}{c_0}z+\dfrac{D(z)}{C(z)}}=\frac{a_0}{b_0}z+\frac{1}{\dfrac{b_0}{c_0}z+\dfrac{1}{\dfrac{C(z)}{D(z)}}}$$

$$=\frac{a_0}{b_0}z+\cfrac{1}{\dfrac{b_0}{c_0}z+\cfrac{1}{\dfrac{c_0}{d_0}z+\dots}}$$

Koefficienterna $a_0, b_0, c_0, d_0, \dots$ kan således erhållas genom att man utvecklar den rationella funktionen $A(z)/B(z)$ i kedjebråk.

Hurwitzkriteriet

Rouths tablå är utmärkt för att avgöra om en ekvation är stabil eller ej. Om ekvationens koefficienter beror på en parameter, t. ex. en förstärkningskoefficient, och man skall ange för vilka parametervärden systemet är stabilt, är det bekvämt att ha ett stabilitetskriterium som explicit beror av ekvationens koefficienter. Betrakta ekvationen

$$F(z)=a_0z^n+a_1z^{n-1}+a_2z^{n-2}+\dots+a_n=0 \tag{6.4}$$

där koefficienterna a_i är reella och a_0 är positiv.

Ur ekvationens koefficienter a_i bildar vi de s. k. Hurwitzdeterminanterna

$$\Delta_1 = a_1$$

$$\Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}$$

$$\Delta_3 = \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}$$

$$\vdots$$

$$\Delta_k = \begin{vmatrix} a_1 & a_3 & a_5 & \dots & a_{2k-1} \\ a_0 & a_2 & a_4 & \dots & a_{2k-2} \\ 0 & a_1 & a_3 & \dots & a_{2k-3} \\ \vdots & & & & \vdots \\ & & & & a_k \end{vmatrix} \quad k = 1, 2, \dots, n \tag{6.5}$$

I varje rad ökar således index med två enheter. I kolonnerna avtar index successivt med en enhet. Termen a_m sättes till noll om m är större än n eller negativ.

Vi kan nu formulera Hurwitz' teorem.

SATS 6.2. Hurwitz

Låt ekvation (6.4) ha reella koefficienter. Anta att $a_0 > 0$. Ett nödvändigt och tillräckligt villkor för att alla rötter till ekvation (6.4) har negativ realdel är att alla determinanter $\Delta_1, \Delta_2, \dots, \Delta_n$ är positiva.

Beviset av satsen som baseras på funktionsteorin återges ej. Vi observerar att

$$\Delta_n = a_n \Delta_{n-1} \tag{6.6}$$

EXEMPEL 6.2

För en tredjegradsekvation ($n = 3$) erhålls följande villkor ur Hurwitz' teorem

$$a_1 > 0, \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, a_3 > 0$$

Stabilitetsvillkoret lyder således (jämför exempel 2:11.3)
$a_1 > 0,\ a_1 a_2 > a_0 a_3,\ a_3 > 0$

EXEMPEL 6.3

För en fjärdegradsekvation ($n = 4$) erhålls följande villkor

$$a_1 > 0, \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} > 0, a_4 > 0$$

Hurwitzkriteriet ger följande olikheter

Ur ekvationens koefficienter a_i bildar vi de s. k. Hurwitzdeterminanterna

$$\Delta_1 = a_1$$

$$\Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}$$

$$\Delta_3 = \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}$$

$$\vdots$$

$$\Delta_k = \begin{vmatrix} a_1 & a_3 & a_5 & \dots & a_{2k-1} \\ a_0 & a_2 & a_4 & \dots & a_{2k-2} \\ 0 & a_1 & a_3 & \dots & a_{2k-3} \\ \vdots & & & & \vdots \\ & & & & a_k \end{vmatrix} \quad k = 1, 2, \dots, n \tag{6.5}$$

I varje rad ökar således index med två enheter. I kolonnerna avtar index successivt med en enhet. Termen a_m sättes till noll om m är större än n eller negativ.

Vi kan nu formulera Hurwitz' teorem.

SATS 6.2. Hurwitz

Låt ekvation (6.4) ha reella koefficienter. Anta att $a_0 > 0$. Ett nödvändigt och tillräckligt villkor för att alla rötter till ekvation (6.4) har negativ realdel är att alla determinanter $\Delta_1, \Delta_2, \dots, \Delta_n$ är positiva.

Beviset av satsen som baseras på funktionsteorin återges ej. Vi observerar att

$$\Delta_n = a_n \Delta_{n-1} \tag{6.6}$$

EXEMPEL 6.2

För en tredjegradsekvation ($n = 3$) erhålls följande villkor ur Hurwitz' teorem

$$a_1 > 0, \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, a_3 > 0$$

Stabilitetsvillkoret lyder således (jämför exempel 2:11.3)
$a_1 > 0,\ a_1 a_2 > a_0 a_3,\ a_3 > 0$

EXEMPEL 6.3

För en fjärdegradsekvation ($n = 4$) erhålls följande villkor

$$a_1 > 0, \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} > 0, a_4 > 0$$

Hurwitzkriteriet ger följande olikheter

$$a_1 > 0 \tag{6.7}$$

$$a_1 a_2 - a_0 a_3 > 0 \tag{6.8}$$

$$a_3(a_1 a_2 - a_0 a_3) - a_1^2 a_4 > 0 \tag{6.9}$$

$$a_4 > 0 \tag{6.10}$$

Då alla koefficienter är positiva finner vi att villkor (6.8) följer av villkor (6.9). Stabilitetsvillkoret kan således reduceras till

$$a_i > 0 \quad i = 0, \ldots, 4$$

$$a_1 a_2 a_3 - a_0 a_3^2 - a_1^2 a_4 > 0$$

Villkoret (6.8) som erhålls ur Hurwitz' sats är alltså onödigt. Lienard och Chipart har observerat detta, generaliserat och erhållit följande resultat.

SATS 6.3. Liénard och Chipart

Låt ekvation (6.4) ha reella koefficienter. Anta att $a_0 > 0$. Ett nödvändigt och tillräckligt villkor för att samtliga rötter till ekvation (6.4) har negativ realdel är att endera av följande villkor är uppfyllda

1. $a_n > 0, a_{n-2} > 0, \ldots, \Delta_1 > 0, \Delta_3 > 0, \ldots$
2. $a_n > 0, a_{n-2} > 0, \ldots, \Delta_2 > 0, \Delta_4 > 0, \ldots$
3. $a_n > 0, a_{n-1} > 0, a_{n-3} > 0, \ldots, \Delta_1 > 0, \Delta_3 > 0$
4. $a_n > 0, a_{n-1} > 0, a_{n-3} > 0, \ldots, \Delta_2 > 0, \Delta_4 > 0$

Vi visar inte heller denna sats, men observerar att satsen för en ekvation med positiva koefficienter innebär att alla Hurwitzdeterminanter av udda ordning är positiva om de av jämn ordning är positiva och omvänt. Det är således tillräckligt att undersöka tecknet på endera de udda eller de jämna Hurwitzdeterminanterna.

7. Funktionsteoretiska metoder och Nyquistkriteriet

Argumentvariationsprincipen

Det finns flera resultat från funktionsteorin som är användbara när man skall undersöka hur många poler och nollställen en meromorf funktion har i ett givet område. Vi har följande fundamentala resultat.

SATS 7.1

Låt funktionen $f(z)$ vara analytisk i ett slutet område D och på områdets rand C utom i ett ändligt antal poler och nollställen i D. Det gäller då

$$\frac{1}{2\pi}\Delta_C \arg f(z) = \frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz = N - P \tag{7.1}$$

där N är antalet nollställen och P antalet poler i D. Poler respektive nollställen av ordningen m räknas m gånger. Vidare betecknar $\Delta_C \arg f(z)$ argumentvariationen av funktionen $f(z)$ då kurvan C genomlöps i *positiv* riktning.

Bevis:

Anta att $z=a$ är ett nollställe av ordningen m. I närheten av punkten z gäller då

$$f(z) = (z-a)^m g(z)$$

där $g(z)$ är analytisk och skild från noll. Då gäller

$$\frac{f'(z)}{f(z)} = \frac{m}{z-a} + \frac{g'(z)}{g(z)}$$

Den sista termen är analytisk för $z=a$ och funktionen $f'(z)/f(z)$ har således en enkelpol i $z=a$ med residue m. Summan av residuerna vid nollställen är således N. Analogt finner vi att summan av residuerna vid polerna är $-P$. (Ändra tecknet på m i ovanstående resonemang.) Vidare gäller

$$\frac{d}{dz}\log f(z) = \frac{f'(z)}{f(z)}$$

dvs.

$$\int_C f'(z)/f(z)\,dz = \Delta_C \log f(z)$$

där Δ_C betecknar variationen av funktionen längs konturen C. Vidare gäller

$$\log f(z) = \log|f(z)| + i \arg f(z)$$

Då $|f(z)|$ är en entydig funktion gäller således för en sluten kontur

$$\Delta_C \log f(z) = i\,\Delta_C \arg f(z)$$

och satsen är bevisad.

Med hjälp av sats 7.1 kan man nu visa flera satser, som är användbara för att undersöka hur många rötter en algebraisk ekvation har i ett givet område, t. ex. Rouchés sats. För formulering och bevis av dessa hänvisas till någon lärobok i funktionsteori t. ex. Brinck & Persson, *Analytiska funktioner*, Studentlitteratur, Lund 1967, s. 181. För att använda sats 7.1 måste vi utvärdera argumentvariationen för den funktion vars nollställen vi vill undersöka.

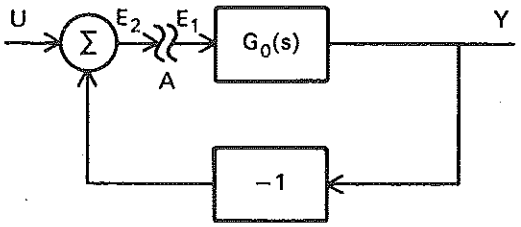


Fig. 7.1. Blockschema för enkelt återkopplat system.

Ett praktiskt sätt att göra detta är att undersöka hur kurvan C i z-planet avbildas på $f(z)$-planet. Argumentvariationen av funktionen $f(z)$ anges av det antal varv som bilden av C omsluter origo i $f(z)$-planet.

Nyquistteoremet

Då Nyquists teorem är en tillämpning av argumentvariationsprincipen (sats 7.1) på ett specialfall kan hans insats förefalla obetydlig. Så är dock ej fallet. Nyquist var nämligen den förste som insåg sambandet mellan stabiliteten för återkopplade system (elektronrörsförstärkare) och funktionsteori.

För att förstå bakgrunden till Nyquistteoremet skall vi studera ett enkelt återkopplat system enligt fig. 7.1. Före Nyquistteoremets tillkomst brukade man använda följande heuristiska resonemang för att avgöra om systemet är stabilt. Anta att insignalen $u(t) \equiv 0$ och att återkopplingskretsen bryts vid A såsom indikerats i fig. 7.1. Anta vidare att e_1 är en sinussignal med amplituden 1 och frekvensen ω. Då det gäller att $E_2(s) = -G_0(s)\,E_1(s)$ blir signalen e_2 i *stationärt tillstånd* en sinusfunktion med amplituden $|G_0(i\omega)|$ och fasläget $-180^0 + \arg G_0(i\omega)$ relativt signalen e_1. Om signalen e_1:s frekvens varierar så kommer således det relativa fasläget mellan e_1 och e_2 att ändras. Välj nu, om så är möjligt, signalen e_1:s frekvens så att e_1 och e_2 ligger i fas. Om signalerna e_1 och e_2 då har samma amplitud kan systemet kopplas ihop vid A och en självsvängning erhålls. Om e_2 har lägre amplitud än e_1 borde systemet vid sammankopplingen vara stabilt och om e_2 har högre amplitud än e_1 borde systemet vid sammankoppling vara instabilt. Om detta resonemang vore riktigt skulle man således kunna avgöra systemets stabilitet genom att undersöka systemets kretsöverföring, t. ex. genom att rita frekvenskurvan. Jämför avsnitt 2:9. Det återkopplade systemet skulle således vara stabilt om frekvenskurvan skär reella axeln i G_0-planet till höger om punkten -1 och instabilt om frekvenskurvan skär reella axeln i G_0-planet till vänster om punkten -1. Tyvärr är det genomförda resonemanget ej korrekt. Orsaken till detta utreddes klart av Nyquist. Vi citerar ur hans artikel 'The Regeneration Theory' i Oldenburger, R. (utg.), Frequency Response, Mac Millan, New York 1956:

»Mr. Black proposed a negative feedback repeater and proved by tests that it possessed the advantages which he had predicted for it. In particular, its gain was constant to a high degree, and it was linear enough so that spurious

signals caused by the interaction of the various channels could be kept within permissible limits. For best results, the feedback factor, the quantity usually known as $\mu\beta(\equiv -G_0(i\omega))$, had to be numerically much larger than unity. The possibility of stability with a feedback factor greater than unity was puzzling. Granted that the factor is negative it was not obvious how that would help. If the factor was minus 10, the effect of one round trip around the feedback loop is to change the mangitude of an original current from, say, 1 to -10. After a second trip around the loop the current becomes 100, and so forth. The totality looks much like a divergent series and it was not clear how such a succession of ever-increasing components could add to something finite and so stable as experience had shown. The missing part in this argument is that the numbers that describe the successive components 1, -10, 100, and so on, represent the steady state, whereas at any finite time many of the components have not yet reached the steady state and some of them, which are destined to become very large, have barely reached perceptible magnitude. My calculations were principally concerned with replacing the indefinite divergent series referred to by a series which gives the actual value attained at a specific time t. The series thus obtained is convergent instead of divergent and, moreover, converges to values in agreement with the experimental findings.

This explains how I came to undertake the work. It should perhaps be explained also how it comes to be so detailed. In the course of the calculations, the facts with which the term conditional stability have come to be associated, became apparent. One aspect of this is that it is possible to have a feedback loop which is stable and can be made unstable by increasing the loop loss. This seemed a very surprising result and appeared to require that all the steps be examined and set forth in full detail.»

Vi skall nu visa Nyquistteoremet för det enkla återkopplade systemet i fig. 7.1. Vi använder ej Nyquists originalbevis utan stöder oss i stället på argumentvariationsprincipen vilket är enklare. För systemet i fig. 7.1 gäller att överföringsfunktionen från insignal till utsignal är

$$Y(s) = \frac{G_0(s)}{1+G_0(s)}\,U(s) \tag{7.2}$$

Systemets karakteristiska ekvation har samma rötter som ekvationen

$$1+G_0(s)=0 \tag{7.3}$$

Då vi önskar formulera resultatet med hjälp av kretsöverföringen $G_0(s)$ skall vi utgå ifrån (7.3) i stället för systemets karakteristiska ekvation.

För att avgöra om den ostörda ($u(t)=0$) lösningen ($y(t)=0$) är asymptotisk stabil, skall vi alltså undersöka om ekvationen (7.3) har alla rötter i vänstra halvplanet. Vi undersöker då argumentvariationen för funktionen $1+G_0(s)$ då s genomlöper en kontur C som omsluter högra halvplanet $\mathrm{Re}\, s>0$.

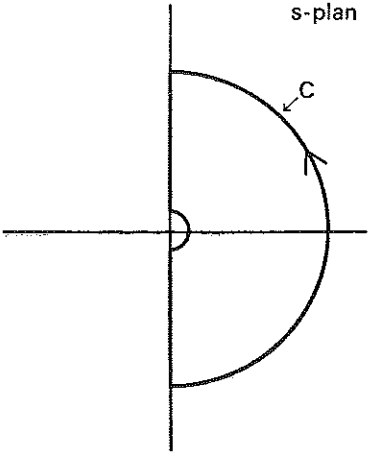


Fig. 7.2. Konturen C som används vid formulering av Nyquist-teoremet.

Det är brukligt att som kontur C välja den kontur som visas i fig. 7.2. Den består av de positiva och negativa imaginära halvaxlarna, två halvcirklar med origo som medelpunkt, en liten (radie r) och en stor (radie R). För att undersöka argumentvariationen hos funktionen $1+G_0(s)$ ritas bilden av C i G_0-planet. Sedan detta gjorts görs en gränsövergång $r\to 0$ och $R\to\infty$.

Vi kan nu formulera Nyquistteoremet.

SATS 7.2. Nyquist

Om $G_0(s)$ ej har några poler i det område i s-planet som omslutes av C, och om bilden av C i G_0-planet ej omsluter punkten -1 så har den karakteristiska ekvationens alla rötter negativ realdel.

Bilden av positiva imaginära halvaxeln är frekvenskurvan och brukar ritas heldragen. Jämför avsnitt 2:9. Observera att då konturen C genomlöpes i positiv riktning rör sig punkten G_0 längs frekvenskurvan från $\omega=\infty$ till $\omega=0$.

Observera att vi med hjälp av Nyquistteoremet kan avgöra det slutna systemets stabilitet enbart med hjälp av kretsöverföringen. Det är även lätt att inse hur kretsöverföringen skall modifieras för att ett instabilt system skall bli stabilt. Dessa egenskaper är de viktigaste motiven för Nyquistteoremets användbarhet.

Det bör även noteras att Nyquistteoremet förutsätter att $G_0(s)$ ej har poler i det område i s-planet som omsluts av C. Om G_0 har poler i högra halvplanet rekommenderas direkt användning av argumentvariationsprincipen såsom illustreras i exempel 7.2.

En väsentlig fördel med Nyquistteoremet är att det ger direkt anvisning om hur ett instabilt system skall modifieras för att det skall bli stabilt.

EXEMPEL 7.1

Betrakta ett reglersystem enligt fig. 7.1 med kretsöverföringen

$$G_0(s)=\frac{K}{s(s+1)^2} \tag{7.4}$$

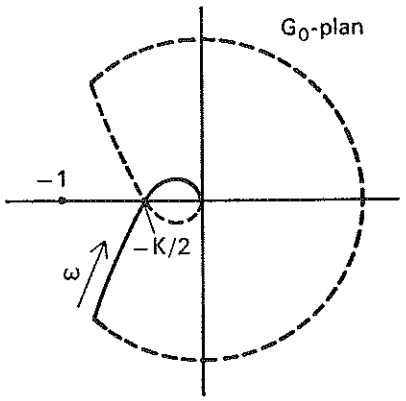


Fig. 7.3. Bilden av konturen C i fig. 7.2 vid avbildningen

$$G_0(s) = \frac{K}{s(s+1)^2}$$

Frekvenskurvan för kretsöverföringen har markerats med heldragen linje.

Hur stor får förstärkningsparametern K högst vara för att det slutna systemet skall vara stabilt?

För att använda Nyquistteoremet skall vi först undersöka hur konturen C i fig. 7.2 avbildas på G_0-planet vid transformationen

$$G_0 = G_0(s) = \frac{K}{s(s+1)^2}$$

Den »lilla halvcirkeln»

$$s = re^{i\theta},\ -\pi/2 \leqslant \theta \leqslant \pi/2,\ r \ll 1$$

i s-planet avbildas approximativt på halvcirkeln

$$G_0 = \frac{K}{r} e^{-i\theta}$$

dvs. en halvcirkel som omsluter högra halvplanet. Den »stora halvcirkeln» i s-planet

$$s = Re^{i\theta},\ -\pi/2 \leqslant \theta \leqslant \pi/2,\ R \gg 1$$

avbildas på origo i G_0-planet. Används det resonemang som beskrevs i avsnitt 2:9 finner vi att frekvenskurvan börjar ($\omega = 0$) i tredje kvadranten och att den slutar ($\omega = \infty$) i andra kvadranten.

Vidare finner vi att frekvenskurvan skär reella axeln då

$$\arg G_0(i\omega) = -\arg(i\omega) - 2\arg(1+i\omega) = -\pi$$

dvs. för $\omega = 1$.

För detta ω-värde gäller att $|G_0(i\omega)| = K/2$. Bilden av konturen C framgår av fig. 7.3. Vi finner att bilden av konturen C ej omsluter punkten -1 om $K < 2$.

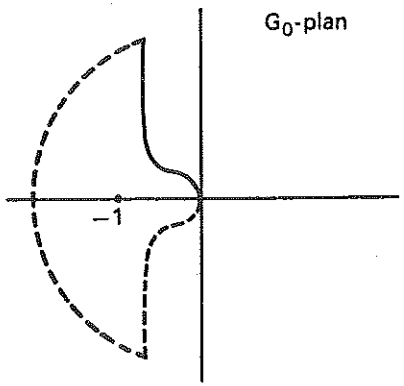


Fig. 7.4. Bilden av konturen C i fig. 7.2 vid avbildningen

$$G_0(s) = \frac{K}{s(s-1)(s+5)}$$

Frekvenskurvan för överföringsfunktionen $G_0(s)$ har markerats med heldragen linje.

EXEMPEL 7.2

Betrakta ett system enligt fig. 7.1 med

$$G_0(s) = \frac{K}{s(s-1)(s+5)}$$

Ange för vilka värden på förstärkningsparametern systemet är stabilt.

I detta fall har G_0 således en pol i högra halvplanet. Nyquistteoremet kan då ej användas utan vi utnyttjar i stället argumentvariationsprincipen direkt. Vi undersöker först bilden av konturen C i fig. 7.2 som omsluter högra halvplanet. Den »lilla halvcirkeln» i s-planet

$$s = re^{i\theta},\ -\pi/2 \leqslant \theta \leqslant \pi/2,\ r \ll 1$$

avbildas approximativt på halvcirkeln

$$G_0 = -K/(5s) = -Ke^{-i\theta}/(5r)$$

dvs. en halvcirkel som omsluter vänstra halvplanet. Den »stora halvcirkeln» i s-planet

$$s = Re^{i\theta},\ -\pi/2 \leqslant \theta \leqslant \pi/2,\ R \gg 1$$

avbildas på origo i G_0-planet. Vidare finner vi att frekvenskurvan börjar ($\omega = 0+$) i andra kvadranten och slutar ($\omega = +\infty$) i andra kvadranten. Villkoret för att frekvenskurvan skall skära negativa reella axeln i G_0-planet lyder

$$\arg(i\omega) + \arg(i\omega - 1) + \arg(i\omega + 5) = \pi$$

dvs.

$$\operatorname{arctg}\omega - \operatorname{arctg}(\omega/5) = \pi/2,\quad 0 < \omega < \infty$$

Denna ekvation har ej någon lösning och frekvenskurvan förlöper alltså helt i andra kvadranten. Bilden av konturen C framgår av fig. 7.4. Vi finner att

$\Delta_C \arg [1+G_0(s)] = +2\pi$

då konturen C i fig. 7.2 genomlöpes i positiv riktning. Argumentvariationsprincipen ger då

$N-P=1$

Då funktionen $G_0(s)$ har en pol i högra halvplanet gäller således $P=1$ dvs. $N=2$. Systemets karakteristiska ekvation har således två rötter i högra halvplanet och systemet är då instabilt för alla värden på K.

8. Konvergenshastighet. Praktisk stabilitet

Motivering

Begreppet asymptotisk stabilitet, definition 2.2, som är grunden för vår stabilitetsanalys innebär endast att den störda lösningen konvergerar mot nolllösningen. De stabilitetskriterier som garanterar asymptotisk stabilitet säger således ingenting om hur snabbt den störda lösningen konvergerar mot nolllösningen. Vi ger ett exempel som illustrerar att det ej är tillräckligt att kräva asymptotisk stabilitet.

EXEMPEL 8.1

Betrakta det reglersystem som behandlades i exempel 5.1. Systemet har den karakteristiska ekvationen

$$s^3+3s^2+2s+K=0 \tag{8.1}$$

Vi finner, t. ex. med hjälp av Hurwitz' kriterium sats 6.2, att systemet är stabilt då $K<6$, men instabilt då $K\geqslant 6$. Vi skall nu undersöka systemets egenskaper då förstärkningsparametern K antar värden i närheten av stabilitetsgränsen. Vi identifierar (8.1) med uttrycket

$$(s+\alpha)(s^2+2\zeta\omega s+\omega^2) \tag{8.2}$$

Med $K=6-\varepsilon$ erhålls då termer av andra och högre ordning i ε försummas

$$\begin{aligned}\alpha &= 3-\varepsilon/11\\ \omega &= \sqrt{2}-3\sqrt{2}\,\varepsilon/44\\ \zeta &= \varepsilon\sqrt{2}/44\end{aligned} \tag{8.3}$$

För små värden på $\varepsilon>0$ har den karakteristiska ekvationen två komplexa rötter, som ligger mycket nära den imaginära axeln. I fig. 8.1 visas stegfunktionssvaret för system med $K=5{,}0$ och $K=5{,}9$. Även om systemet är stabilt för dessa värden på förstärkningsparametern, så är det fullständigt oanvänd-

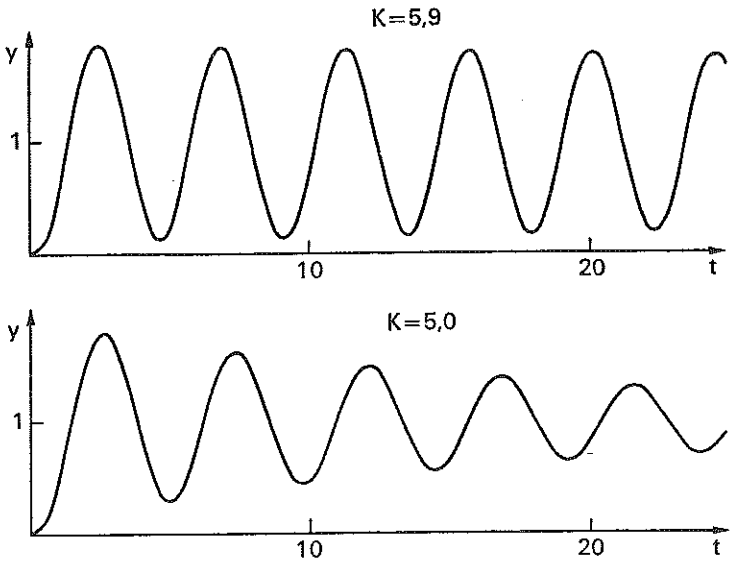


Fig. 8.1. Stegfunktionssvar för enkelt slutet reglersystem med kretsöverföringen

$$G_0(s) = \frac{K}{s(s+1)(s+2)}$$

och $K = 5{,}9$ respektive $K = 5{,}0$.

bart som reglersystem. Vi kan t. ex. tänka oss att systemet representerar en potentiometerskrivare. Vidare finner vi att redan mycket små variationer i systemets parametrar kan leda till att systemet blir instabilt.

Exemplet visar klart att det ej är tillräckligt att kräva enbart asymptotisk stabilitet om ett reglersystem skall fungera tillfredsställande. I de flesta praktiska tillämpningar måste man åtminstone ha kunskap om konvergenshastigheten. Dessa frågeställningar leder naturligt fram till problemet om hur ett reglersystems syfte skall specificeras. En ingående behandling av detta problem ges i kap. 6. I föreliggande avsnitt skall vi som en förberedelse diskutera några stabilitetsbegrepp, som är mer restriktiva än asymptotisk stabilitet.

Starkare villkor på den karakteristiska ekvationens rötter

Ett naturligt sätt att införa restriktivare stabilitetsbegrepp vore att fordra att

$$\|x(t)\| < Ke^{-\alpha t}, \quad \alpha > 0 \tag{8.4}$$

Detta är ekvivalent med att kräva att den karakteristiska ekvationens rötter ligger i halvplanet $\operatorname{Re} s < -\alpha$. Jämför sats 2:11.1. Ett alternativ till (8.4) vore att fordra att den karakteristiska ekvationens rötter skall ligga i sektorn $\pi/2 + \alpha < \arg s < 3\pi/2 - \alpha$. Kriterier av denna typ kan lätt undersökas med de direkta metoderna, t. ex. rotortmetoden, där rötternas värden anges explicit.

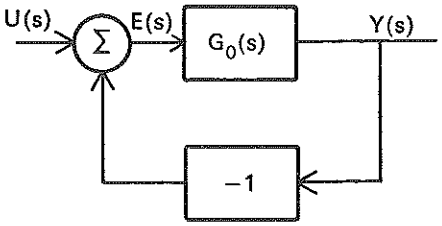


Fig. 8.2. Blockschema för enkelt återkopplat system.

Det är också möjligt att erhålla kriterier för dessa mer restriktiva krav genom modifiering av Hurwitz' teorem eller med argumentvariationsprincipen. Då kriterier som erhålls på detta sätt är mycket komplicerade, används de sällan i praktiken. Man har i stället infört andra heuristiska kriterier som är lättare att evaluera. Samtliga kriterier har konstruerats för att analysera ett enkelt återkopplat system, vars blockschema framgår av fig. 8.2. Kriterierna är baserade på analys av kretsöverföringen $G_0(s)$.

Förstärkningsmarginal och fasmarginal

Begreppen förstärkningsmarginal och fasmarginal är båda relaterade till känslighetsanalys. Man undersöker således hur mycket vissa systemparametrar kan ändras innan systemet blir instabilt. Låt $G_0(s)$ vara kretsöverföringen, dvs. det öppna systemets överföringsfunktion. Låt vidare ω_0 vara det minsta positiva tal sådant att $\arg G_0(i\omega_0)=\pi$, och funktionen $\arg G_0(i\omega)$ är avtagande för $\omega=\omega_0$. Med *förstärkningsmarginalen* eller amplitudmarginalen avses då talet

$$A_m=1/|G_0(i\omega_0)| \tag{8.5}$$

Låt vidare ω_c vara det minsta positiva tal sådant att $|G_0(i\omega_c)|=1$. Med *fasmarginalen* avses talet

$$\varphi_m=\pi+\arg G_0(i\omega_c) \tag{8.6}$$

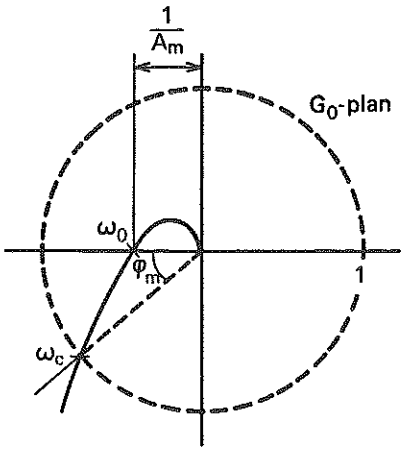


Fig. 8.3. Illustrerar bestämning av fasmarginal φ_m och förstärkningsmarginal A_m i Nyquistdiagrammet.

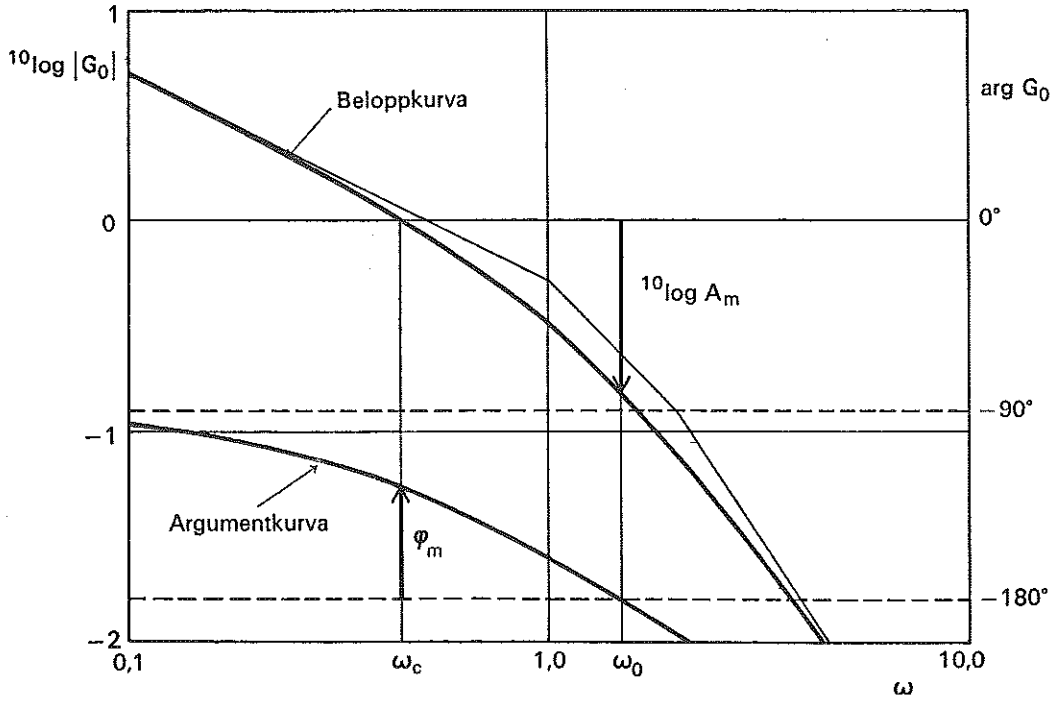


Fig. 8.4. Illustrerar bestämning av fasmarginal φ_m och amplitudmarginal A_m i Bodediagrammet. I figuren visas Bodediagrammet för överföringsfunktionen

$$G_0(s) = \frac{K}{s(s+1)(s+2)}$$

Vi finner $\varphi_m = 54°$ och $^{10}\log A_m = 0{,}78$ dvs. $A_m = 6$.

Förstärkningsmarginalen anger således hur många gånger förstärkningen i den öppna kretsen kan ökas innan systemet blir instabilt. Om förstärkningsmarginalen är mindre än ett är systemet instabilt. Fasmarginalen anger analogt hur mycket extra fasförskjutning som kan införas i det öppna systemet innan det slutna systemet blir instabilt. Om fasmarginalen är negativ är systemet instabilt. Man har empiriskt funnit lämpliga värden för amplitudmarginal och fasmarginal för olika typer av system. För vanliga servomekanismer har följande värden visat sig ge rimliga system $A_m = 2$–6 och $\varphi_m = 30°$–$60°$. Både fasmarginal och amplitudmarginal kan direkt utläsas ur Nyquist- eller Bodediagrammen för det öppna systemet (fig. 8.3 och 8.4).

Resonanstoppens storlek M_p

Den logaritmiska frekvenskurvan för reglersystemet i exempel 5.1 med $K = 5{,}0$ visas i fig. 8.5.

Det starkt oscillerande stegsvaret vi fann i fig. 8.1 har tydligen nära sam-

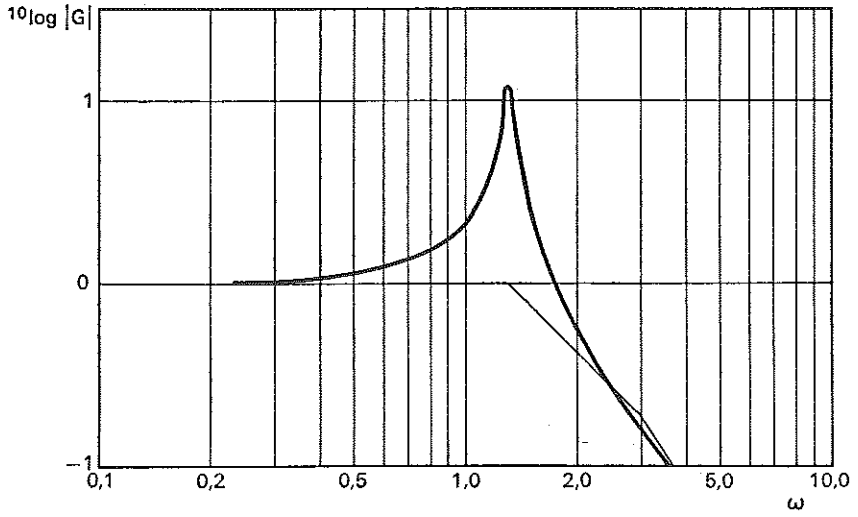


Fig. 8.5. Amplitudkurva för reglersystemet i exempel 5.1 med $K = 5{,}0$. Jämför med stegsvaret för samma system i fig. 8.1.

band med den höga toppen i det slutna systemets resonanskurva. Detta är utgångspunkten för ett annat kvalitativt stabilitetsbegrepp, nämligen resonanstoppens storlek

$$M_p = \max_{\omega} |G(i\omega)| = \max_{\omega} \left| \frac{G_0(i\omega)}{1 + G_0(i\omega)} \right| \tag{8.7}$$

Man har empiriskt funnit att för vanliga servomekanismer är $M_p = 1{,}3$ (1,0–1,5) ett lämpligt värde. För att bekvämt kunna använda resonanstoppens höjd som kriterium fordras metoder, som gör det möjligt att på ett enkelt sätt bestämma M_p ur kunskap om systemets kretsöverföring. I det följande skall vi ange sådana metoder.

Bestämning av det slutna systemets överföringsfunktion ur kretsöverföringen

Det slutna systemet i fig. 8.2 har överföringsfunktionen

$$G(s) = \frac{G_0(s)}{1 + G_0(s)} \tag{8.8}$$

där $G_0(s)$ är det öppna systemets överföringsfunktion. Om överföringsfunktionen G_0 är given genom sin frekvenskurva $G_0(i\omega)$ kan det slutna systemets frekvenskurva lätt bestämmas med hjälp av en grafisk konstruktion (fig. 8.6).

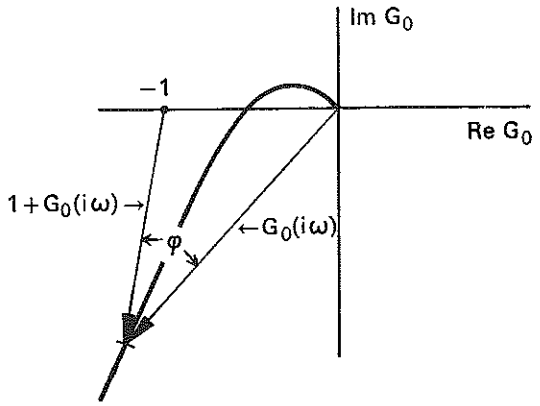


Fig. 8.6. Illustrerar hur det slutna systemets frekvenskurva $G(i\omega)$ kan konstrueras ur det öppna systemets frekvenskurva $G_0(i\omega)$. De komplexa talen $G_0(i\omega)$ och $1+G_0(i\omega)$ kan direkt utläsas i diagrammet. Beloppet $|G(i\omega)|$ erhålls som kvoten mellan $|G_0(i\omega)|$ och $|1+G_0(i\omega)|$. Argumentet arg $G(i\omega)$ är lika med negativa värdet av vinkeln φ.

Den grafiska konstruktionen underlättas om vi ritar följande kurvskaror i G_0-planet

$$|G_0/(1+G_0)| = M = \text{konstant } (M\text{-cirklar}) \tag{8.9}$$

$$\arg [G_0/(1+G_0)] = \varphi = \text{konstant } (\varphi\text{-cirklar}) \tag{8.10}$$

Ett G_0-plan med dessa kurvskaror inritade kallas Halldiagram (fig. 8.7). Man kan visa att båda kurvskarorna är cirkelskaror som skär varandra under räta vinklar. Kurvskaran (8.9) är de harmoniska cirklarna till punkterna $G_0=0$ och $G_0=-1$. En kurva i denna skara har medelpunkten i

$$G_0 = -M^2/(M^2-1)$$

och radien

$$R = |M/(M^2-1)|$$

En kurva i den ortogonala skaran (8.10) har medelpunkten i

$$G_0 = -\frac{1}{2} + \frac{i}{2\,\mathrm{tg}\,(\varphi/2)}$$

och radien

$$R = \frac{1}{2\sin(\varphi/2)}$$

För att bestämma det slutna systemets överföringsfunktion $G=G_0/(1+G_0)$ ur kretsöverföringen G_0 ritas således först kretsöverföringens frekvenskurva. Frekvenskurvan graderas i ω på vanligt sätt. Genom att interpolera i Halldiagrammets M- och φ-kurvor kan man sedan för en given punkt på frekvenskurvan bestämma belopp och argument för G.

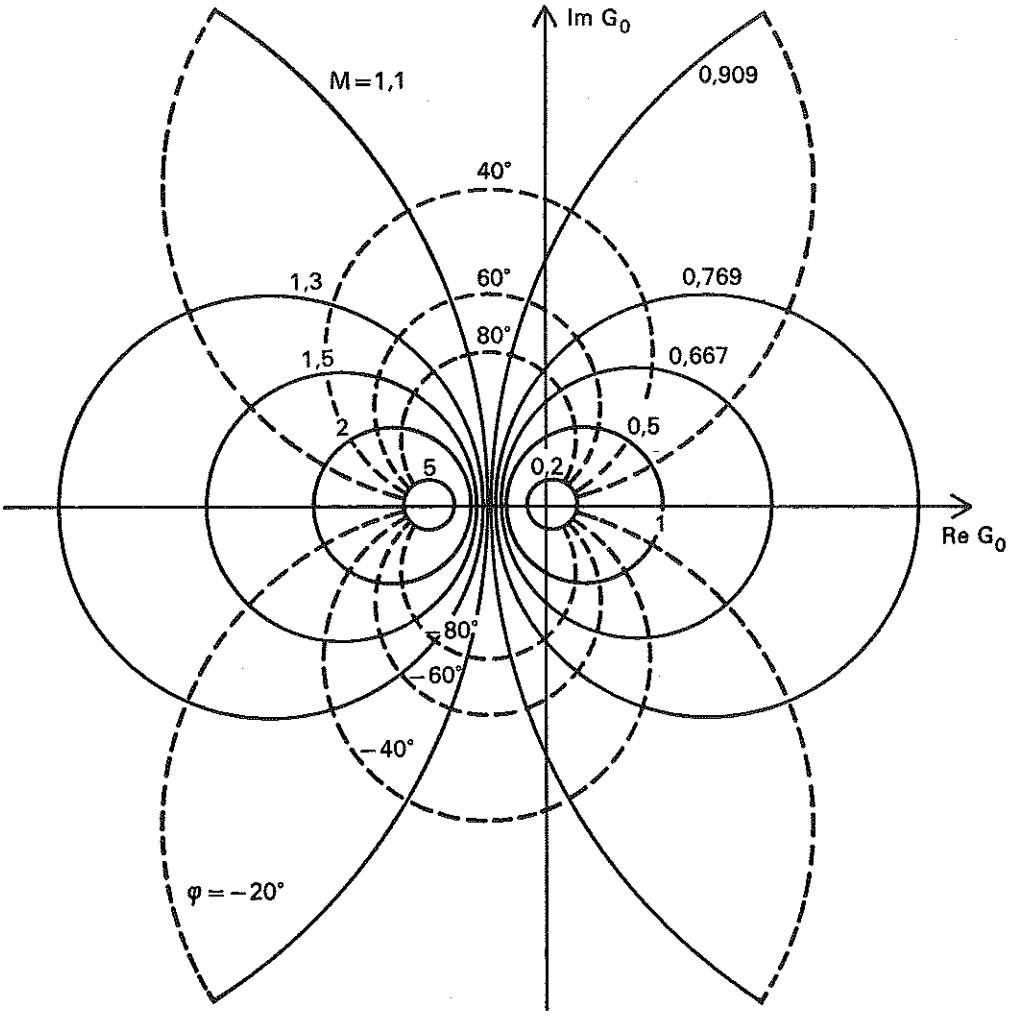


Fig. 8.7. Halldiagram, dvs. G_0-plan med kurvskarorna $|G_0/(1+G_0)| = M$ och arg $G_0/(1+G_0) = \varphi$. Jämför ekvipotential- och fältlinjer från en källa och sänka i planet.

Resonanstoppens höjd M_p, definierad av ekvation (8.7), kan direkt bestämmas ur Halldiagrammet som det största M-värde som antas för någon punkt på frekvenskurvan. Detta värde antas i en punkt där frekvenskurvan tangerar en M-cirkel.

Nicholsdiagrammet

Då det är tidsödande att rita frekvenskurvor används Halldiagrammet sällan vid praktiskt arbete. I stället används Nicholsdiagrammet, som är den kon-

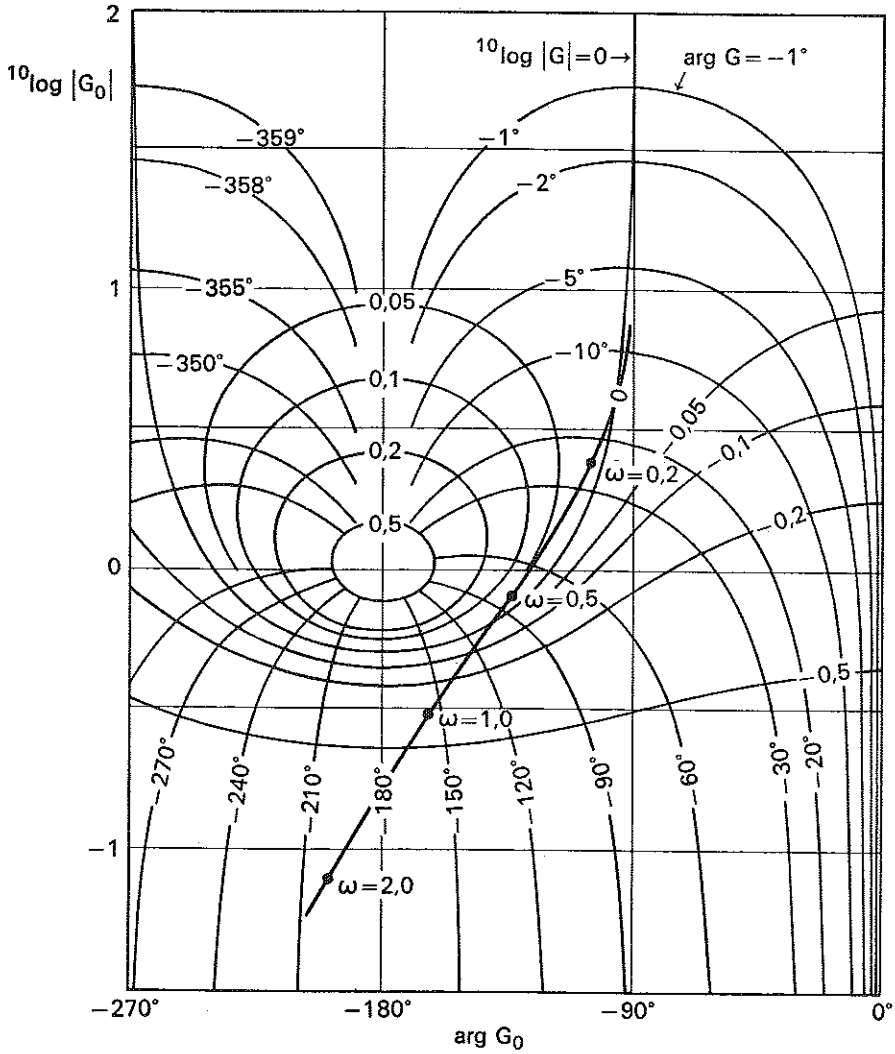


Fig. 8.8. I Nicholsdiagrammet ritas frekvenskurvan för det öppna systemets överföringsfunktion G_0, med arg $G_0(i\omega)$ som abscissa och $^{10}\log|G_0(i\omega)|$ som ordinata. I diagrammet finns kurvskarorna $G_0/(1+G_0)=M$ och arg $G_0/(1+G_0)=\varphi$ inritade. Dessa kurvskaror har graderats med 10-logaritmen för M-värdet och φ-värdet i grader. Det slutna systemets överföringsfunktion $G=G_0/(1+G_0)$ kan bestämmas ur kretsöverföringen G_0 med hjälp av kurvskarorna. Resonanstoppen M_P kan direkt utläsas ur diagrammet.

forma avbildningen av Halldiagrammet vid transformationen $G_0 \rightarrow \log G_0 = \log|G_0| + i \arg G_0$. Frekvenskurvans transformation i $\log G_0$-planet kan nämligen lätt erhållas ur Bodediagrammet för G_0. I Bodediagrammet kan mätas direkt $\log|G_0(i\omega)|$ och arg $G_0(i\omega)$. Genom att för varje ω avsätta $\log|G_0(i\omega)|$ och arg $G_0(i\omega)$ i ett rätvinkligt koordinatsystem erhålls således en punkt på frekvenskurvans bild i $\log G_0$-planet.

M- och φ-cirklarna, definierade av ekvationerna (8.9) och (8.10), blir i Nicholsdiagrammet konforma avbildningar av cirkelskarorna i Halldiagrammet. Nicholsdiagram ritas ofta med andra enheter än naturliga logaritmer och radianer. Kurvskarorna (8.9) och (8.10) bildar då ej räta vinklar med varandra. Fig. 8.8 visar ett sådant diagram.

Nicholsdiagrammet används på samma sätt som Halldiagrammet. Med dess hjälp kan man således lätt bestämma det slutna systemets överföringsfunktion G ur kretsöverföringen G_0. För att göra detta ritas först frekvenskurvans avbildning i Nicholsdiagrammet. Genom att interpolera i det kroklinjiga koordinatsystemet kan man sedan för varje punkt på frekvenskurvans bild bestämma belopp och argument för det komplexa talet $G_0/(1+G_0)$. Med hjälp av kurvskaran $|G_0/(1+G_0)| = M$ kan man även avläsa resonanstoppens höjd M_p som det största M-värdet som antas för någon punkt på frekvenskurvans avbildning. Detta värde uppnås naturligtvis i någon punkt där frekvenskurvans bild tangerar en M-kurva.

I Nicholsdiagrammet i fig. 8.8 har vi ritat bilden av frekvenskurvan för ett system med kretsöverföringen

$$G_0(s) = \frac{K}{s(s+1)(s+2)}, \quad K = 1$$

För detta system erhålls $^{10}\log M_p = 0{,}05$, dvs. $M_p = 1{,}1$.

Det slutna systemets bandbredd ω_B kan vidare erhållas som den frekvens för vilken frekvenskurvan skär kurvan $|G_0/(1+G_0)| = 0{,}707$. För systemet i fig. 8.8 erhålls $\omega_B = 0{,}8$.

Då en multiplikation av G_0 med en konstant K motsvaras av en addition av $\log K$ och $\log G_0$, finner vi att vid en ändring av kretsförstärkningen translateras frekvenskurvans avbildning i Nicholsdiagrammet vertikalt. Det är således lätt att se hur en ändring av det öppna systemets förstärkning påverkar det slutna systemets överföringsfunktion. För systemet i fig. 8.8 finner vi t. ex. att om vi tillåter $^{10}\log M_p = 0{,}1$ ($M_p = 1{,}3$ så kan kretsförstärkningen ökas med $^{10}\log K = 0{,}1$, dvs. $K = 1{,}3$. Systemets bandbredd ökar härvid till $\omega_B = 0{,}9$.

9. Referenser

En elementär framställning av Lyapunovs resultat har givits av

LaSalle, J. & Lefschetz, S., *Stability by Liapunov's Direct Method*, Academic Press, New York 1961.

Vi hänvisar till denna bok för ytterligare referenser.

En utförlig behandling av rotortmetoden har givits av dess skapare

Evans, W. R., *Control-System Dynamics*, McGraw-Hill, New York 1954.

Rotortmetoden diskuteras även i de flesta läroböcker i reglerteknik, t. ex. de tidigare nämnda av Truxal, Chestnut & Mayer och Elgerd.

Bevis av Rouths och Hurwitz teorem återfinns i

Gantmacher, F. R., *Theory of Matrices*, Vol. II, kap. 15.

Nyquistteoremet studeras lämpligen i originalarbetet

Nyquist, H., *Regeneration Theory*, Bell. System Technical Journal *II* (1932), s. 126–147.

Artikeln finns reproducerad i

Bellman, R. & Kalaba, R., *Mathematical Trends in Control Theory*, Dover, New York 1964.

Komplettering av kunskaper i funktionsteori kan t. ex. göras med hjälp av någon av

Ahlfors, L. V., *Complex Analysis*, McGraw-Hill, New York 1956

Brinck, I. & Persson, A. *Elementär teori för analytiska funktioner*, Studentlitteratur, Lund 1967

Titchmarsh, E. C., *The Theory of Functions*, Oxford University Press, London 1939.

En utförlig behandling av Halldiagrammen och Nicholsdiagram finns i Chestnut & Mayers lärobok.

6. Syntes av reglersystem

1. Inledning

Lösningen av ett reglerproblem kan uppdelas i följande moment:

- Specifikation av problemet
- Modellbygge och identifiering
- Analys
- Simulering
- Syntes
- Förverkligande
- Drifttagning och trimning

För att ange specifikationerna bör man göra klart för sig vilka krav som systemet bör uppfylla. Detta kan t. ex. innebära att ange systemets förmåga, att eliminera störningar och att följa kommandosignaler.

Modellbyggandet avser att ge en matematisk modell av den process som skall regleras och av de störningar som påverkar processen. Modellen kan erhållas genom att tillämpa grundläggande naturlagar såsom massbalans, energibalans och rörelsemängdsbalans. Om den process som skall regleras finns tillgänglig kan man också bestämma modellerna genom att experimentera med processen. Experimenten utförs genom att insignalen varieras och motsvarande variationer i utsignalen registreras. Bestämning av dynamiska modeller genom analys av experimentella insignal-utsignal-data kallas *identifiering*. Överföringsfunktionen kan t. ex. bestämmas genom att variera insignalen sinusformigt och mäta amplitud- och fasrelationerna mellan insignal och utsignal (frekvensanalys). Viktfunktionen kan t. ex. bestämmas genom att registrera utsignalen då insignalen är en kort puls (transientanalys). I många fall används flera modeller av varierande komplexitet från enkla modeller av låg ordning till mycket komplicerade modeller som representerar reglerobjektets dynamiska egenskaper med stor detaljrikedom. Man eftersträvar också att göra modeller som beskriver systemets omgivning genom att karakterisera störningar och mätfel. Några av de modeller som används presenterades i kapitel 2.

Analysen ger en överblick av modellens egenskaper såsom stabilitet, styrbarhet, observerbarhet, känslighet, icke-minimum-fas-egenskaper m. m. Vid dessa moment används de verktyg som presenterades i kapitel 3, 4 och 5.

Analysen kompletteras ofta med simulering, då många problem ej kan lösas analytiskt.

Syntesen består i att matematiskt karakterisera regulatorns funktion. Man skall beskriva hur styrvariabeln skall genereras från de tillgängliga mätningarna. Simulering används ofta för att verifiera att den erhållna styrlagen ger ett slutet system med önskade egenskaper.

Med förverkligande, även kallat realisering, implementering eller apparatisering, avses att konstruera, tillverka och montera en apparat som utför den signalbehandling som angivits av syntesen. Rent tekniskt består reglersystemen ofta av komponenter av analogityp, dvs. återkopplade operationsförstärkare, eller datorer med analog-digital och digital-analog omvandlare.

Med drifttagning och trimning avses att installera regulatorn och att ställa in den så att den fungerar enligt specifikationerna.

I praktiken kan de olika momenten av problemlösandet ej genomföras steg för steg i den ordning som angivits ovan. Problemlösandet är i stället en iterativ process, som sällan är helt systematisk och logisk. Slutmålet är att få en praktiskt fungerande regulator, som uppfyller de tekniska specifikationerna, som är billig, pålitlig och lätt att trimma in, modifiera och underhålla. Detta mål kan uppnås på många olika sätt och ofta måste många tekniskt-ekonomiska avvägningar göras. Det finns med andra ord utmärkta tillfällen för konstruktören av ett reglersystem att visa sin ingenjörsmässiga förmåga. Man bör också lägga märke till att det ingår i reglerteknikerns uppgifter att undersöka om reglerproblemen kan lösas genom ändringar av processen, införandet av nya mätsignaler och nya styrsignaler.

Reglerproblemen kan ha mycket olika karaktär. Det arbete som fordras för att lösa ett reglerproblem kan vara mycket varierande främst beroende av om det finns tidigare erfarenheter eller om det gäller en ny tillämpning. Reglering av nivå i behållare förekommer allmänt i processindustrin. För att lösa en sådan regleruppgift installeras vanligen en proportionell och integrerande standardregulator. Regulatorns parametrar inställs sedan empiriskt med ledning av tumregler. Man vrider på rattarna, observerar processens uppförande och modifierar inställningen till dess önskat resultat erhålls. Vid nya tillämpningar t. ex. vid lösning av styrproblem för nya vapensystem, mjuklandning på månen, avancerad processreglering etc. är projekteringsarbetet mycket omfattande. Det kan i omfattning motsvara tiotals manår eller mer. Det är vanligt att systemet simuleras antingen genom att man bygger en skalmodell eller genom att man använder analogimaskin eller dator. Eventuellt kan en kombination göras därigenom att verkliga komponenter används tillsammans med en analogimaskin eller dator.

En annan faktor av stor betydelse är om man fritt kan välja systemets alla komponenter eller om man är bunden att använda vissa givna komponenter. Vid nykonstruktion har man mycket stora möjligheter, framförallt bör samspelet mellan reglerutrustning och övriga komponenter i systemet beaktas. Jämför diskussionen i kap. 1:2 om autopiloter och accelerometrar.

Vid automatisering av redan befintlig utrustning är man i stor utsträckning bunden av de dynamiska egenskaperna hos den existerande utrustningen.

Detta kapitel behandlar syntes. Vi antar att det system som skall styras är givet. Problemet är hur insignalen u skall beräknas ur den information om systemet som erhålls ur utsignalen y för att systemets syfte skall uppnås. För att lösa detta problem är det uppenbarligen väsentligt att karakterisera systemets omgivning och specificera dess syfte. I reglerteknikens begynnelse lade man huvudsakligen kvalitativa synpunkter på systemets syfte och man var belåten med de fördelar, som återkopplingen gav i sig själv. Syftet var närmast att åstadkomma ett stabilt återkopplat system med små statiska fel. Man var inte heller så noga med att definiera systemets omgivning, dvs. de störningar som påverkar systemet. De analytiska verktyg som användes var väsentligen stabilitetskriterier. Kvantitativa uttryck för snabbhet, stabilitet och noggrannhet var sedan en naturlig vidareutveckling. Vi har redan i tidigare avsnitt behandlat sådana uttryck, t. ex. förstärkningsmarginal, fasmarginal, M_p-värde, felkoefficienter, styvhetskoefficienter. De specifikationer som erhålls på detta sätt leder endast i undantagsfall till en entydig styrlag. Som en direkt följd av detta utvecklades ett stort antal syntesprocedurer. Dessa syntesprocedurer har alla följande karaktär:

- Välj styrlagens form (systemets struktur).
- Justera kretsförstärkningen så att stabiliteten blir acceptabel.
- Undersök om reglerfelet är tillfredsställande vid typiska störningar med hjälp av analys eller simulering på analogimaskin eller dator.
- Upprepa proceduren från början och fortsätt tills specifikationerna är uppfyllda.

Då specifikationerna i allmänhet ej ger en entydig styrlag måste godtyckliga val göras i flera av stegen.

Ovanstående procedur kan synas mycket grov. Det har emellertid visat sig att en bra ingenjör med hjälp av dessa riktlinjer och viss träning kan dimensionera servosystem som fungerar tillfredsställande. De flesta servosystem, som är i bruk idag, är hopkomna på detta sätt. De heuristiska syntesprocedurerna har emellertid betydande nackdelar. Det är omöjligt att med den angivna proceduren avgöra om problemet överhuvudtaget går att lösa. Specifikationerna kan också vara inbördes motstridande. Syntesmetoderna kan endast med stor svårighet tillämpas på system som ej är enkla och tidsinvarianta. Man vet aldrig om väsentligt bättre prestanda kan erhållas med en annan struktur. Man har försökt att komma ifrån dessa svårigheter genom att ange en skalär prestandafunktion som t. ex. motsvarar den ekonomiska vinsten. På detta sätt kan man få klarhet i vilka prestanda som kan uppnås och var de fundamentala begränsningarna finns. Den grundläggande svårigheten är att det sällan går att ange ett systems prestanda med ett enda tal.

Syntes av ett reglersystem kommer därför alltid att göras genom en kombination av analytiska och erfarenhetsmässiga synpunkter.

Tillgången på datorer med bildskärmar och dialogprogramvara har givit drastiska möjligheter att förenkla syntesarbetet. Människans intuition och överblick kan då kombineras med datorns förmåga att snabbt utföra beräkningar. Det finns f. n. vid flera universitet (t. ex. i Lund) programsystem för datorstödd dimensionering av reglersystem. Det är t. ex. möjligt att via tangentbord ge beskrivningar av system och att ändra parametrar i en systembeskrivning. Man kan sedan med hjälp av kommandon generera Bode-diagram och Nyquist-kurvor på en bildskärm. Med hjälp av andra kommandon kan man t. ex. begära undersökning av styrbarhet och observerbarhet. Det finns också möjligheter att via tangentbordet definiera specifikationer och via kommandon begära syntes av regulatorer baserade på olika syntesmetoder. Program för datorstödd syntes av reglersystem är ett mycket kraftfullt hjälpmedel för konstruktören.

I avsnitt 2 i detta kapitel ges en översikt av olika sätt att specificera reglersystem. Här behandlas såväl heuristiska specifikationer i tids- och frekvensplanet som optimeringskriterier. Avsnitt 3 ger en översikt av olika reglerprinciper och reglerstrategier. Metoder för att syntetisera ett urval av reglerprinciperna behandlas i följande avsnitt. Syntes av framkoppling diskuteras i avsnitt 4. Detta är intimt förknippat med problemet att beräkna inversa system. Avsnitt 5 behandlar två olika metoder att dimensionera en regulator som ger en föreskriven överföringsfunktion. Om ett system med en insignal är styrbart så kan det slutna systemets poler ges godtyckliga värden genom en återkoppling från samtliga tillståndsvariabler. Om tillståndsvariablerna ej kan mätas direkt så kan de rekonstrueras ur mätningar av in- och utsignaler. Olika metoder att göra detta behandlas också i avsnitt 5. I avsnitt 6 slutligen diskuteras två metoder baserade på frekvenskurvor och rotortdiagram med vars hjälp ett systems egenskaper kan modifieras utan att det slutna systemets samtliga poler måste fastläggas.

2. Specifikationer

Att specificera ett system består i att ange dess syfte och omgivning. I vissa fall kan detta vara mycket lätt, t. ex. vid dimensionering av ett servosystem avsett att driva ett teleskop så att det följer en viss stjärna. Systemets syfte kan således vara att teleskopet skall följa stjärnans rörelse med ett vinkelfel som är mindre än en given tolerans. Systemets omgivning utgörs i detta fall av stjärnans rörelse, som är en given funktion av tiden. I detta speciella fall karakteriseras systemets omgivning med en enda funktion. I andra fall kan det vara svårare att specificera systemets omgivning,

t. ex. vid dimensionering av ett instrumentservo för en potentiometerskrivare eller vid dimensionering av ett servo för inriktning av en luftvärnskanon. I sådana fall är det många olika typer av insignaler som skall reproduceras och det kan vara mycket svårt att finna en lämplig klassificering av signalerna. I praktiken begränsar man sig ofta till att undersöka systemets förmåga att reproducera vissa testsignaler, t. ex. stegfunktion, rampfunktion, parabel och sinusfunktion. För ett luftvärnsservo kan man t. ex. få viss uppfattning om instorheten genom att uppskatta det tänkta flygplanets största hastighet och undanmanöverförmåga.

I detta avsnitt skall vi ange några olika sätt att specificera ett system. Vi skall först redogöra för en uppsättning specifikationer, som utvecklades i samband med de heuristiska syntesmetoderna, nämligen specifikationer på stegsvar och frekvenssvar. Dessa specifikationer utvecklades för att kvantitativt uttrycka kompromissen mellan statisk noggrannhet och stabilitet för reproducerande servosystem. Specifikationerna är i allmänhet sådana att det finns många system som uppfyller dem. Vi skall sedan behandla specifikationer, som anger de önskade överföringsfunktionerna från referenssignal till utsignal och från störning till utsignal. Slutligen skall vi behandla specifikationer som formuleras med hjälp av en kriteriefunktion.

Specifikationer i tidsplanet

Överföringsfunktionen för ett linjärt tidsinvariant system är entydigt bestämd av ett insignal–utsignalpar under förutsättning att insignalen varierar tillräckligt oregelbundet (en enda sinussignal är ej tillräckligt). Det är alltså möjligt att specificera överföringsfunktionen för ett system genom att man anger systemets utsignal för en given insignal. För reproducerande servosystem är det vanligt att som insignaler välja stegfunktioner, rampfunktioner etc. Vi skall i anslutning till fig. 2.1 ange några vanliga specifikationer på stegsvaret, dvs. det slutna systemets svar på ett enhetssteg.

Följande storheter brukar anges för att specificera stegfunktionssvaret:

- *Översläng*, M, (Overshoot). Utsignalens maximalvärde minskat med 1. Överslängen anges vanligen i %. Dess tillåtna värde varierar med tillämpningen. Truxal har angivit följande riktvärden. Från 1 % i vissa mätinstrument, 5–10 % för video-förstärkare i television, upp till 25–30 % i vissa servomekanismer. Ibland anges även *maxtiden* T_M, dvs. den tid då utsignalen antar sitt största värde.
- *Stigtid*, T_r (Rise time). Inverterade värdet av maximum av utsignalens tidsderivata. Andra definitioner förekommer t. ex. den tid det tar för utsignalen att växa från 0,1 till 0,9.
- *Insvängningstid* eller lösningstid, T_s (Solution time, settling time). Det minsta värdet sådant att utsignalen för alla $t \geq T_s$ tar värden mellan $1-p$ och $1+p$. Vanligen är $p=0{,}05$.

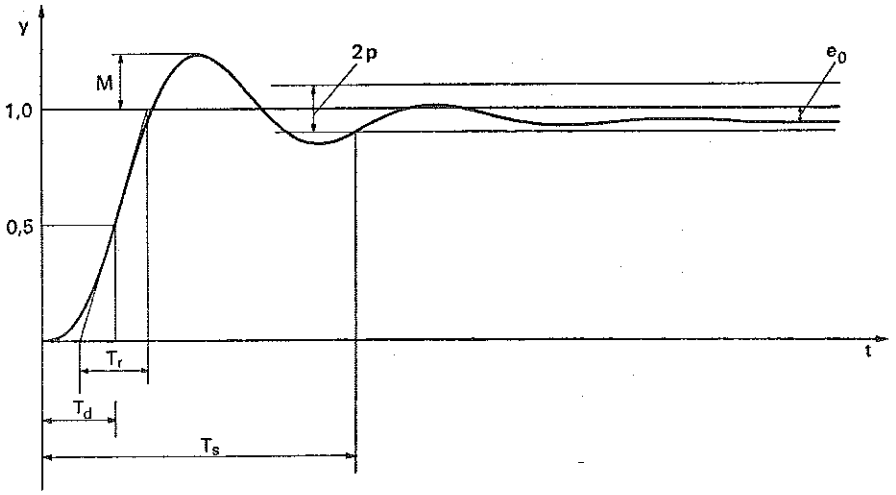


Fig. 2.1. Stegfunktionssvar för ett reproducerande servosystem med illustration av några vanliga specifikationer i tidsplanet.

- *Fördröjningstid*, T_d (Delay time). Den tid det tar för utsignalen att uppnå värdet 0,5.
- *Slutfel*, e_0 (Static error, Final value of error). Definieras som $e_0 = 1 - y(\infty)$. Slutfelet e_0 är identiskt med den nollte felkoefficienten. Jämför avsnitt 4:3.

För att karakterisera systemets förmåga att följa lågfrekventa insignaler brukar felkoefficienterna anges. Dessa kan direkt avläsas ur en registrering av systemets in- och utsignaler då insignalen är en potens av t (fig. 2.2 och 2.3).

Observera att även om samtliga nu nämnda storheter specificeras, så är systemets stegsvar ej entydigt bestämt.

Specifikationerna på stegsvaret hänför sig till den situation då systemet skall reproducera en given insignal. Om systemets syfte i stället är att eliminera störningar erhålls liknande specifikationer. I fig. 2.1 kan t. ex. kvantiteten $1 - y(t)$ tolkas som det fel, som erhålls vid en störning i form av ett enhetssteg. Felkoefficienten e_0 motsvaras då av styvhetskoefficienten c_0 med ombytt tecken.

Specifikationer i frekvensplanet

Specifikationerna kan även formuleras i frekvensplanet. Man låter då systemets omgivning karakteriseras av en klass av sinussignaler. I fig. 2.4 visas amplitudkurvan för överföringsfunktionen för ett reproducerande servosystem.

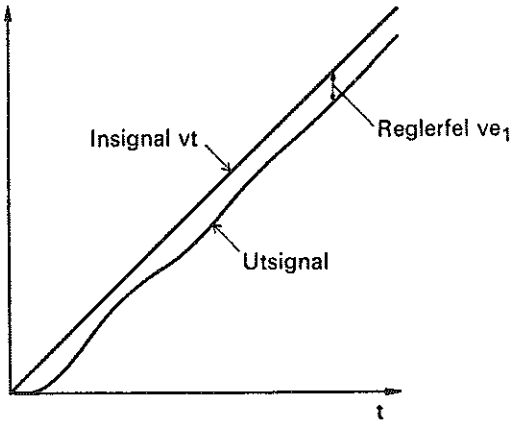


Fig. 2.2. Rampsvar för system med $e_0 = 0$, $e_1 \neq 0$.

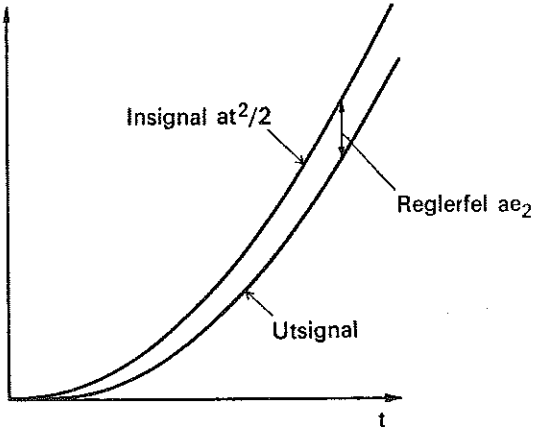


Fig. 2.3. Insignal och utsignal för system med $e_0 = e_1 = 0$, $e_2 \neq 0$, då insignalen är en parabel.

Följande storheter brukar anges för att specificera systemet:

- *Resonanstopp* M_p (M-peak, resonance ratio). Amplitudkurvans maximum. Anges ofta i dB. Vanliga värden på resonanstoppen är $M_p = 1{,}0$–$1{,}5$ dvs. 0–3,3 dB.
- *Resonansfrekvens* ω_p (Peak frequency). Frekvens för vilken amplitudkurvan har sitt maximum.
- *Bandbredd* ω_B (Bandwidth). Frekvens för vilken $|G(i\omega_B)| = 1/\sqrt{2}$.
- *Amplitudkurvans lutning* k (Slope). Vid höga frekvenser avtar $|G(i\omega)|$ som $|\omega|^{-k}$, vilket i ett logaritmiskt diagram motsvarar en rät linje med lutningen $-k$.

Liknande specifikationer kan ställas upp för överföringsfunktionen från störning till utsignal för system där man vill eliminera inverkan av störningar.

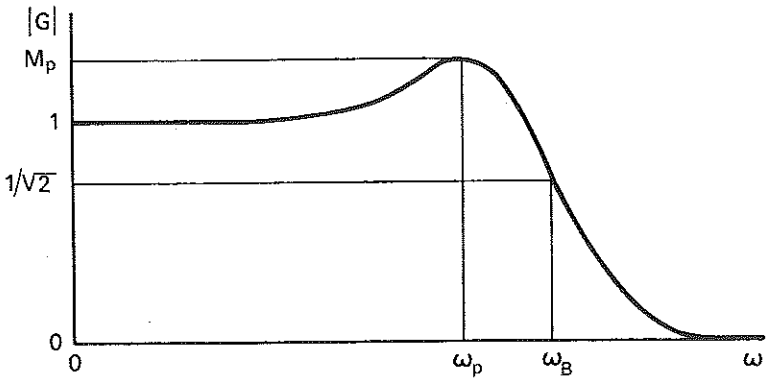


Fig. 2.4. Amplitudkurva för ett reproducerande servosystems överföringsfunktion, som har $G(0) = 1$, dvs. $e_0 = 0$. Observera linjära skalor!

Specifikationer på överföringsfunktionen

De specifikationer som hittills behandlats ställer vissa villkor på det slutna systemets överföringsfunktioner från insignal till utsignal och från störning till utsignal. Det finns emellertid många överföringsfunktioner, som satisfierar en given uppsättning specifikationer. För att undvika denna brist på entydighet, kan man därför direkt specificera de önskade överföringsfunktionerna, t. ex. genom att man fastlägger dess singulariteter.

Ett reproducerande servosystem har vanligen en överföringsfunktion enligt fig. 2.5. Överföringsfunktionen karakteriseras av:

- Ett par komplexkonjugerade poler $s = p_2$ och $s = p_3$
- En »dipol» $s = z_1$, $s = p_1$ på reella axeln nära origo. Observera att nollstället oftast ligger närmast origo (motiveras senare). Eventuellt kan flera dipoler finnas.
- Ett varierande antal poler och nollställen på negativa reella axeln eller dess närhet.

Systemets »transienta uppförande», dvs. stigtid, översläng, bandbredd, resonanstopp bestäms väsentligen av de komplexkonjugerade polerna p_2 och p_3. Dessa kallas *kontrollpoler* eller *dominerande poler* och betecknas

$$p_{2,3} = -\zeta\omega_k \pm i\omega_k\sqrt{1-\zeta^2}$$

där ζ kallas den *relativa dämpningen*. Systemets egenskaper vid mycket snabbt varierande insignaler bestäms väsentligen av skillnaden mellan totala antalet poler och nollställen.

Felkoefficienterna som karakteriserar systemets statiska egenskaper bestäms väsentligen av överföringsfunktionens egenskaper i närheten av origo. Se avsnitt 4:3.

Vi skall nu härleda ett samband mellan felkoefficienterna och det slutna

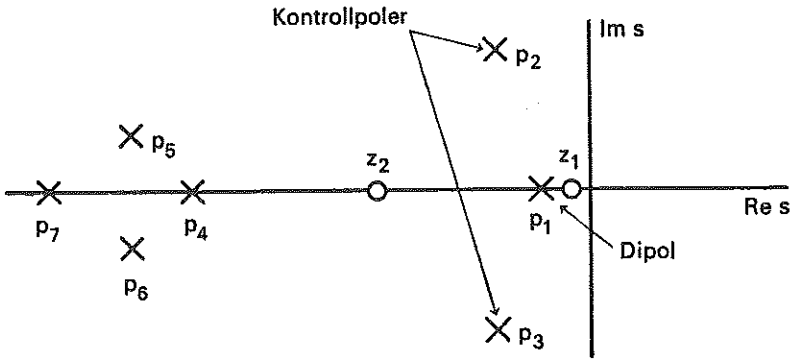


Fig. 2.5. Singularitetsdiagram för överföringsfunktionen $G(s)$ för ett typiskt följeservo. Vid låga frekvenser är förstärkningen 1, dvs. $G(0) = 1$.

systemets poler och nollställen. Låt det slutna systemets överföringsfunktion vara

$$G(s) = K\frac{(s-z_1)(s-z_2)\dots(s-z_m)}{(s-p_1)(s-p_2)\dots(s-p_n)} \tag{2.1}$$

Felkoefficienterna definierades i sats 4:3.1 som koefficienterna för s-potenser i Maclaurinutvecklingen av $1-G(s)$, dvs.

$$1-G(s) = e_0 + e_1 s + e_2 s^2 + \dots \tag{2.2}$$

Vi finner således

$$e_0 = 1 - G(0) \tag{2.3}$$

Derivering av (2.2) ger

$$e_1 = -\left[\frac{dG(s)}{ds}\right]_{s=0} = -G(0)\left[\frac{d\log G(s)}{ds}\right]_{s=0} = (1-e_0)\left[\sum_{i=1}^{m}\frac{1}{z_i} - \sum_{i=1}^{n}\frac{1}{p_i}\right] \tag{2.4}$$

För att beräkna e_2 deriveras (2.2) två gånger. Vi får

$$2e_2 = -\left[\frac{d^2G(s)}{ds^2}\right]_{s=0} = -\frac{1}{G(0)}\left[\frac{dG(s)}{ds}\right]^2_{s=0} - G(0)\left[\frac{d^2\log G(s)}{ds^2}\right]_{s=0}$$

dvs

$$2e_2 = \frac{-e_1^2}{1-e_0} + (1-e_0)\left[\sum_{i=1}^{m}\frac{1}{z_i^2} - \sum_{i=1}^{n}\frac{1}{p_i^2}\right] \tag{2.5}$$

Felkoefficienten e_1 är vanligen positiv. Det följer av ekvation (2.4) att det är möjligt att minska e_1 genom att placera ett nollställe nära origo på negativa reella axeln. Detta är motivet till nollstället z_1 i fig. 2.5. Motivet till att ha en pol p_1 omedelbart till vänster om nollstället z_1 i fig. 2.5 är att man ej vill att

nollstället z_1 skall påverka systemets stegsvar alltför mycket. Om polen p_1 och nollstället z_1 ligger tillräckligt nära varandra så har de ringa inflytande på stegsvaret. För att visa detta beräknar vi stegfunktionssvaret för överföringsfunktioner som endast skiljer sig med en dipol p_1, z_1.

Låt $G(s)$ vara en överföringsfunktion med stegsvaret $f(t)$. Vi skall nu uppskatta stegsvaret för överföringsfunktionen

$$\tilde{G}(s) = \frac{p_1}{z_1}\,\frac{s-z_1}{s-p_1}\,G(s)$$

där vi förutsätter att z_1 och p_1 är reella och $p_1 < 0$. Vi antar dessutom att överföringsfunktionen $G(s)$:s samtliga poler ligger i halvplanet $\operatorname{Re} s < p_1 - a$ där $a > 0$. Vi har följande identitet

$$\tilde{G}(s) = \frac{p_1}{z_1}\,G(s) + \frac{p_1}{z_1}\,\frac{p_1 - z_1}{s-p_1}\,G(s)$$

Låt $\tilde{f}(t)$ beteckna stegsvaret för $\tilde{G}(s)$. Det följer nu av faltningssatsen att

$$\tilde{f}(t) = \frac{p_1}{z_1}\,f(t) + \frac{p_1(p_1-z_1)}{z_1}\int_0^t e^{p_1(t-s)}\,f(s)\,ds \tag{2.6}$$

För små t, dvs $|p_1 t| < 1$, gäller approximativt

$$\tilde{f}(t) \approx \frac{p_1}{z_1}\,f(t) + \frac{p_1(p_1-z_1)}{z_1}\int_0^t f(s)\,ds$$

och för stora t gäller approximativt

$$\tilde{f}(t) \approx \frac{p_1}{z_1}\,f(t) + \left(1 - \frac{p_1}{z_1}\right)(1 - e^{p_1 t})$$

Det är viktigt att ha en känsla för hur systemets egenskaper såsom felkoefficienter, bandbredd m m. beror av polernas och nollställenas lägen. Detta kan fås genom att i detalj studera många exempel. Läsaren rekommenderas att själv göra detta. För enkla system räcker analys med penna och papper, men då det finns många poler och nollställen får räknetekniska hjälpmedel som analogimaskin och dator tillgripas. Några exempel ges nedan.

EXEMPEL 2.1. Första ordningens system

Betrakta ett slutet system med överföringsfunktionen

$$G(s) = \frac{a}{s+a} \tag{2.7}$$

Systemets Bodediagram framgår av fig. 2:9.4. Dess stegsvar är

$$c(t) = 1 - e^{-at}$$

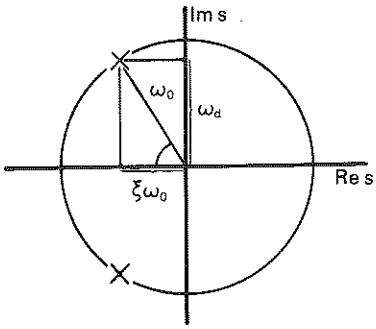


Fig. 2.6. Singularitetsdiagram för överföringsfunktionen

$$G(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}.$$

Enkla räkningar ger

Stigtid	$T_r = 1/a$
Fördröjningstid	$T_d = 0{,}69/a$
Insvängningstid (5 %)	$T_s = 3/a$
Översläng	$M = 0$
Felkoefficienter	$e_0 = 0, \quad e_1 = 1/a$
Bandbredd	$\omega_B = a$
Resonanstopp	$M_p = 1{,}0$
Resonansfrekvens	$\omega_p = 0$

Överföringsfunktionen (2.7) kan erhållas som ett enkelt slutet system enligt fig. 4:3.2 med kretsöverföringen

$$G_0(s) = a/s$$

Detta system har skärningsfrekvensen $\omega_c = a$ och fasmarginalen $\phi_m = 90°$.

Vi finner således att bandbredd och skärningsfrekvens är proportionell mot parametern a medan stigtid, fördröjningstid, insvängningstid och felkoefficienten e_1 är omvänt proportionell mot a.

EXEMPEL 2.2. Andra ordningens system

Betrakta ett slutet system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \tag{2.8}$$

Systemets singularitetsdiagram framgår av fig. 2.6. Orten för polerna då relativa dämpningen ζ varierar är sammansatt av en cirkel runt origo med radie ω_0 och reella axeln. Polerna är reella om $|\zeta| \geqslant 1$.

Systemets stegsvar är

$$c(t) = 1 - \left[\cosh \omega_d t + \frac{\zeta}{\sqrt{\zeta^2 - 1}} \sinh \omega_d t\right] e^{-\zeta\omega_0 t}, \quad |\zeta| > 1$$

$$c(t) = 1 - (1 + \omega_0 t)\, e^{-\omega_0 t}, \quad \zeta = 1$$

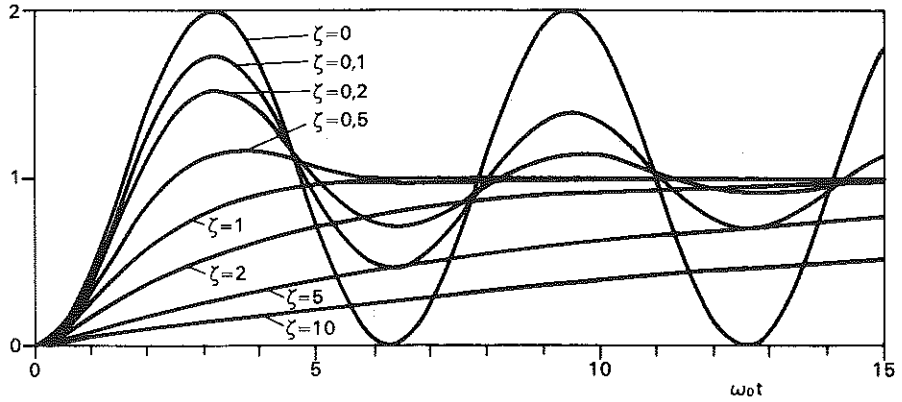


Fig. 2.7. Stegsvar för ett system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 + \omega_0^2}.$$

$$c(t) = 1 - \left[\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t\right] e^{-\zeta\omega_0 t}$$

$$= 1 - \frac{e^{-\zeta\omega_0 t}}{\sqrt{1-\zeta^2}}\sin(\omega_d t + \phi) \quad \zeta < 1$$

där $\omega_d = \omega_0\sqrt{|1-\zeta^2|}$ och vinkeln ϕ är definierad av $\cos\phi = \zeta$, $\sin\phi = \sqrt{1-\zeta^2}$. Jämför fig. 2.6. Stegsvaret visas i fig. 2.7. Enkla räkningar ger för $|\zeta| \leqslant 1$.

Stigtid	$T_r = \omega_0^{-1}\exp(\phi/\tan\phi)$
Insvängningstid (5 %)	$T_s \approx 3/(\omega_0\zeta) \quad 0 \leqslant \zeta < 0{,}9$
Maxtid	$T_M = \pi/\omega_d$
Översläng	$M = \exp(-\pi\zeta/\sqrt{1-\zeta^2})$
Felkoefficienter	$e_0 = 0, \quad e_1 = 2\zeta/\omega_0$
Bandbredd	$\omega_B = \omega_0\sqrt{1-2\zeta^2+\sqrt{(1-2\zeta^2)^2+1}}$
Resonanstopp	$M_p = \begin{cases} 1/(2\zeta\sqrt{1-\zeta^2}) & \zeta \leqslant 1/\sqrt{2} \\ 1 & \zeta > 1/\sqrt{2} \end{cases}$
Resonansfrekvens	$\omega_p = \begin{cases} \omega_0\sqrt{1-2\zeta^2} & \zeta \leqslant 1/\sqrt{2} \\ 0 & \zeta > 1/\sqrt{2} \end{cases}$

Överföringsfunktionen (2.8) kan erhållas som ett enkelt slutet system enligt fig. 4:3.2 med kretsöverföringen

$$G_0(s) = \frac{\omega_0^2}{s(s+2\zeta\omega_0)}$$

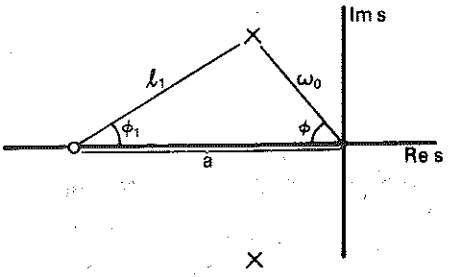


Fig. 2.8. Singularitetsdiagram för ett system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2}{a} \frac{s+a}{s^2+2\zeta\omega_0 s+\omega_0^2}.$$

Denna kretsöverföring har skärningsfrekvensen

$$\omega_c = \omega_0 \sqrt{\sqrt{4\zeta^4+1}-2\zeta^2}$$

och fasmarginalen

$$\phi_m = 90 - \arctan\,(\omega_c/(2\zeta\omega_0))$$

Sammanfattningsvis finner vi att bandbredd och skärningsfrekvens är proportionella mot parametern ω_0, medan stigtid, insvängningstid, maxtid, och felkoefficienten e_1 är omvänt proportionella mot ω_0. Vidare finner vi att överslängen, resonanstoppen, resonansfrekvensen, skärningsfrekvensen och bandbredden avtar med växande ζ, medan stigtiden och felkoefficienten växer med växande ζ. Insvängningstiden har lokala minima men är i huvudsak avtagande med växande ζ för $0 \leqslant \zeta < 0{,}9$.

Exempel 2.3. System med två poler och ett nollställe

Betrakta ett slutet system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2}{a} \cdot \frac{s+a}{s^2+2\zeta\omega_0 s+\omega_0^2} \tag{2.9}$$

I jämförelse med systemet i exempel 2.2 så har ett nollställe tillfogats. För $|\zeta| < 1$ kan systemets stegsvar skrivas på följande sätt

$$c(t) = 1 - \frac{l_1}{a\sqrt{1-\zeta^2}}\, e^{-\zeta\omega_0 t} \sin\,(\omega_d t + \phi + \phi_1)$$

där

$$\omega_d = \omega_0\sqrt{1-\zeta^2}, \quad \phi = \arctan\sqrt{1-\zeta^2}/\zeta,$$

$$\phi_1 = \arctan\,\omega_0\sqrt{1-\zeta^2}/(a-\omega_0\zeta) \quad \text{och} \quad l_1 = \sqrt{\omega_0^2(1-\zeta^2)+(a-\zeta\omega_0)^2}$$

storheterna ϕ, ϕ_1 och l_1 har enkel geometrisk tolkning, vilket framgår av fig. 2.8. Systemets stegsvar för $\zeta = 0.5$ och olika värden på $\alpha = a/\omega_0$ framgår av fig. 2.9. Systemet har bandbredden

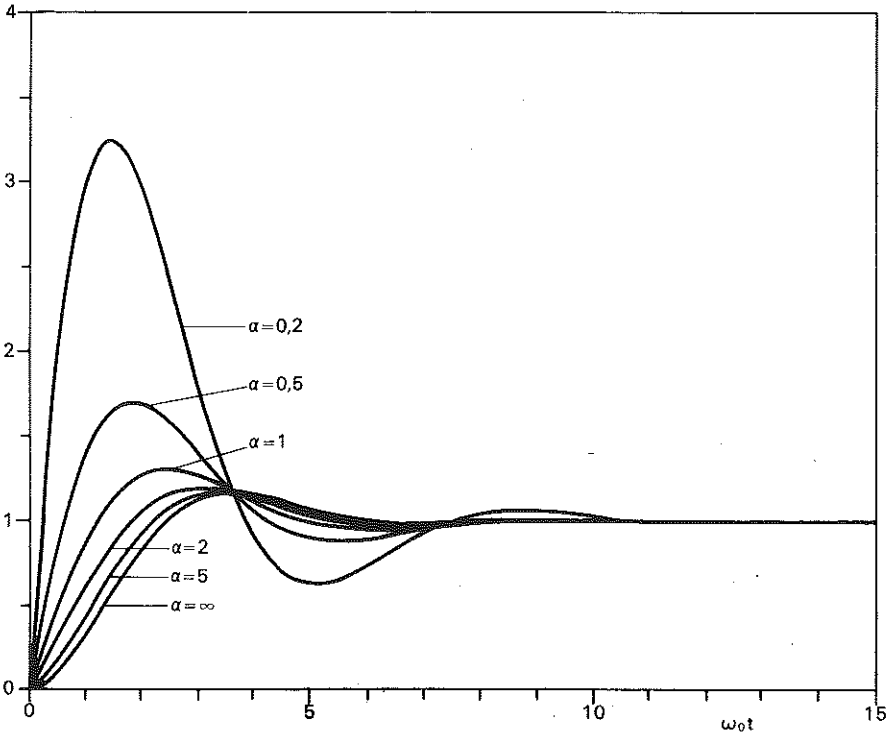


Fig. 2.9. Stegsvar för ett system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2}{a} \frac{s+a}{s^2+2\zeta\omega_0 s+\omega_0^2}$$

för $\zeta = 0{,}5$ och olika värden på $\alpha = a/\omega_0$.

$$\omega_B = \omega_0 \sqrt{1-2\zeta^2+(\omega_0/a)^2+\sqrt{(1-2\zeta^2+(\omega_0/a)^2)^2+1}}$$

Felkoefficienterna är

$$e_0 = 0, \quad e_1 = 2\zeta/\omega_0 - 1/a$$

Jämfört med systemet i exempel 2.2 så har stigtid och lösningstid minskat, bandbredden har ökat, överslängen och resonanstoppen har också ökat. Ändringarna är knappt märkbara för $a > 5\omega_0$ men förändringarna jämfört med systemet i exempel 2.2 ökar drastiskt med minskande α. Observera också att det är möjligt att få felkoefficienten e_1 noll.

EXEMPEL 2.4. System med tre poler

Betrakta ett slutet system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2 b}{(s+b)(s^2+2\zeta\omega_0 s+\omega_0^2)} \tag{2.10}$$

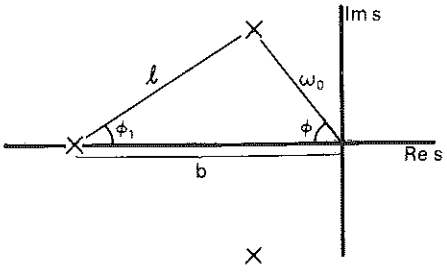


Fig. 2.10. Singularitetsdiagram för ett system med överföringsfunktionen

$$G(s) = \frac{\omega_0^2 b}{(s^2 + 2\zeta\omega_0 s + \omega_0^2)(s+b)}.$$

I jämförelse med systemet i exempel 2.2 så har en pol tillfogats. Genom att tillämpa formeln 2: 8.3 för inversa Laplacetransformen finner vi för $|\zeta| < 1$ att stegsvaret blir

$$c(t) = 1 - \frac{b}{l\sqrt{1-\zeta^2}} e^{-\zeta\omega_0 t} \sin(\omega_d t + \phi - \phi_1) - \frac{\omega_0^2}{l^2} e^{-bt}$$

där

$$\omega_d = \omega_0\sqrt{1-\zeta^2}, \quad \phi = \arctan \sqrt{1-\zeta^2}/\zeta,$$

$$\phi_1 = \arctan \omega_0\sqrt{1-\zeta^2}/(b - \zeta\omega_0), \quad l = \sqrt{(1-\zeta^2)\omega_0^2 + (b-\zeta\omega_0)^2}$$

Det framgår av fig. 2.10 att storheterna ϕ, ϕ_1 och l har enkel geometrisk tolkning. Stegsvaret för $\zeta = 0.5$ och olika värden på b/ω_0 framgår av fig. 2.11. Sammanfattningsvis finner vi i jämförelse med exempel 2.2 att ett tillägg av en pol på negativa reella axeln medför att stigtiden ökar och att överslängen minskar. Förändringen är knappt märkbar om $b > 5\omega_0$ men den ökar med avtagande b.

Bandbreddskompromissen

Att ett servosystem har bandbredden ω_B innebär att signaler, vars frekvensinnehåll ligger i intervallet $(0, \omega_B)$, kommer att reproduceras väl. Man skulle kunna tro att det alltid är önskvärt att ha system med största möjliga bandbredd. Detta medför att systemet blir snabbt. Felkoefficienterna blir också små. Vidare kommer systemets känslighet för parametervariationer också att bli liten. Det kan i vissa fall vara svårt att göra system med hög bandbredd. Tidsfördröjningar och icke-minimumfas egenskaper ger naturliga begränsningar. (Varför?) Dessutom finns det skäl som talar för att systemets bandbredd bör begränsas, nämligen

- Högfrekventa störningar i mätsignalerna
- Mekaniska resonanser
- Slitage av reglerventiler o. dyl.

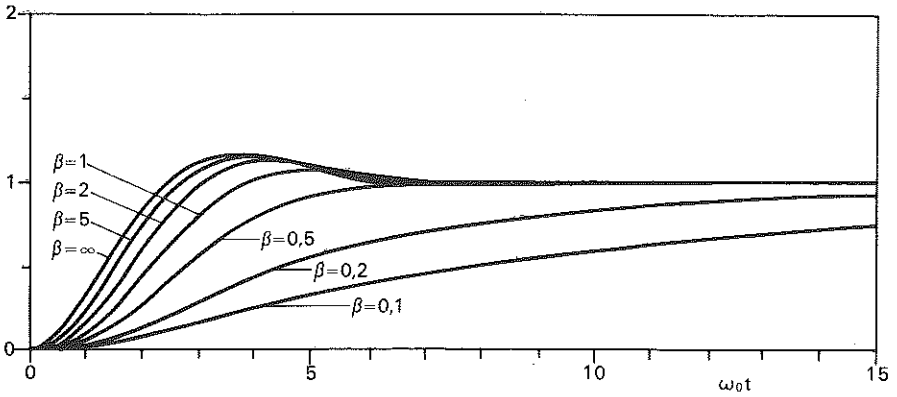


Fig. 2.11. Stegsvar för ett system med överföringsfunktionen

$$G(s) = \frac{\omega_0^3 b}{(s+b)(s^2 + 2\zeta\omega_0 s + \omega_0^2)}$$

för $\zeta = 0{,}5$ och olika värden på $\beta = b/\omega_0$.

En störning som uppkommer vid mätning av utsignalen inträder i reglersystemet på samma sätt som en insignal. Se t. ex. blockschemat i fig. 4:3.2. Mätfel vars frekvenser ligger under reglersystemets bandbredd kommer således att orsaka variationer i systemets utsignal. Då mätfelen ofta har höga frekvenser finns således skäl att begränsa bandbredden.

Högfrekventa mätfel kan uppkomma på många olika sätt. Om regleravvikelsen avkänns med en potentiometer kan störningar uppkomma då potentiometerarmen går från trådvarv till trådvarv. Vid tjockleksmätning genom mätning av strålningsabsorption erhålls störningar på grund av mätningens statistiska natur. Om mätgivaren är en elgon kan störningen bestå av överhörning från bärfrekvensen. Om mätvärdet erhålls från en digital dataöverföring eller från ett digitalt mätorgan, erhålls störningar på grund av kvantisering. Vid tryckmätning orsakar turbulens högfrekventa störningar.

Mekaniska resonanser motsvarar system med överföringsfunktioner av typen

$$G(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$$

med mycket små värden på dämpningskoefficienten ζ. Sådan resonans förekommer t. ex. på grund av torsionssvängningar i en axel eller böjsvängningar i en mekanisk konstruktion. För att undvika att excitera sådana mekaniska resonanser bör reglersystemets bandbredd ligga väl under resonansfrekvensen.

Slutligen slitagefrågan. En onödigt hög bandbredd på ett servosystem ökar slitaget och utmattningen av servokomponenterna t. ex. axlar och kuggväxlar.

Optimeringskriterier

De kriterier som vi hittills diskuterat utvecklades väsentligen för att specificera servomekanismer som skall reproducera vissa typer av signaler och eliminera inflytande av störningar. Specifikationerna är väsentligen ett försök att kvantifiera de kvalitativa synpunkter på återkopplade system, som behandlades i kap. 4. De väsentliga nackdelarna med dessa specifikationer är att de ej bestämmer ett system entydigt och att det kan vara svårt att upptäcka om specifikationerna är inbördes konsistenta. Vi skall nu behandla ett helt annat sätt att specificera systemets syfte. Bektrakta ett system som beskrivs av ekvationen

$$\frac{dx}{dt} = f(x, u, v, t) \tag{2.11}$$

där x är tillståndsvektorn, u styrvariabeln och v en störning. Bilda uttrycket

$$\int_0^T g(x(t), u(t))\, dt + G(x(T)) \tag{2.12}$$

där g är en skalär funktion av x och u, och där G är en skalär funktion av x. Uttrycket (2.12) kallas *förlustfunktion*, *vinstfunktion*, *kriterium* etc. och anger t. ex. den totala vinsten eller den totala kostnaden med att operera ett system. Vidare anges begränsningar på tillstånds- och styrvariablerna, t. ex.

$$\begin{aligned} &\alpha_i(t) \leq x_i(t) \leq \beta_i(t) \quad i = 1, 2, \ldots, n \\ &\gamma_i(t) \leq u_i(t) \leq \delta_i(t) \quad i = 1, 2, \ldots, r \end{aligned} \tag{2.13}$$

Systemets syfte anges nu som att minimera förlustfunktionen (2.12) med bivillkoret att begränsningarna som ges av ekvation (2.13) skall uppfyllas. Detta sätt att specificera ett systems syfte har flera fördelar:

- Metoden är generell. Den kan appliceras på olinjära, tidsvariabla och flervariabla system.
- Specifikationen leder ofta till en entydig styrlag.
- Det finns matematiska metoder med vars hjälp vi kan bestämma en styrlag som uppfyller syftet även i mycket komplicerade situationer.

Vi ger några exempel.

EXEMPEL 2.5. Satellituppskjutning

Sänd upp en satellit till given höjd med minsta bränsleförbrukning.

EXEMPEL 2.6. Stålugnar

Betrakta en stålframställningsprocess t. ex. martinugn eller en LD-process. Syftet här kan vara att på given tid uppnå givna koncentrationer av kol och legeringsämnen med minsta möjliga kostnad. Förlustfunktionen är i detta fall

totalkostnaden för råvaror, bränsle, syre, legeringsämnen m. m. Styrvariablerna kan vara tillsats av syrgas och legeringsämnen.

Metoder för att dimensionera servosystem var väl kända innan optimeringskriterier började användas för att specificera reglersystem. Man har därför försökt att i efterhand bestämma kriteriefunktioner som leder till de egenskaper för servosystem som tidigare befunnits bra. Ett sätt är att ansätta en överföringsfunktion för det slutna systemet på formen

$$G(s) = \frac{\omega_0^n}{s^n + p_1 s^{n-1}\omega_0 + \ldots + p_{n-1} s\omega_0^{n-1} + \omega_0^n} = \frac{\omega_0^n}{P_n(s)} \tag{2.14}$$

och sedan välja de värden på koefficienterna $p_1, \ldots, p_{n-1}$ som minimerar ett kriterium, t. ex.

$$I_1 = \int_0^\infty e^2(t)\,dt$$

eller

$$I_2 = \int_0^\infty t\,|e(t)|\,dt$$

där $e(t)$ är felet vid en stegstörning. Enkla räkningar visar t. ex. för $n=2$ att $p_1 = 1$, dvs. $\zeta = 0{,}5$ minimerar I_1 för ett andra ordningens system. Erfarenhetsmässigt är $\zeta = 0{,}7$ i många fall att föredra för ett system av andra ordningen. Dessutom har det visat sig att kriteriet I_1 ger flacka minima. Genom att undersöka ett stort antal olika kriterier har man visat att kriteriet I_2, som även kallas ITAE (Integral Time-multiplied Absolute Error), har många trevliga egenskaper. Exempel på optimala ITAE system ges nedan.

EXEMPEL 2.7. Optimala ITAE system med $e_0 = 0$

De optimala ITAE systemen av ordningstalen 2–6 med strukturen (2.14) ges av

$$P_2(s) = s^2 + 1{,}4\omega_0 s + \omega_0^2$$

$$P_3(s) = s^3 + 1{,}75\omega_0 s^2 + 2{,}15\omega_0^2 s + \omega_0^3$$

$$P_4(s) = s^4 + 2{,}1\omega_0 s^3 + 3{,}4\omega_0^2 s^2 + 2{,}7\omega_0^3 s + \omega_0^4$$

$$P_5(s) = s^5 + 2{,}8\omega_0 s^4 + 5{,}0\omega_0^2 s^3 + 5{,}5\omega_0^3 s^2 + 3{,}4\omega_0^4 s + \omega_0^5$$

$$P_6(s) = s^6 + 3{,}25\omega_0 s^5 + 6{,}6\omega_0^2 s^4 + 8{,}6\omega_0^3 s^3 + 7{,}45\omega_0^4 s^2 + 3{,}95\omega_0^5 s + \omega_0^6$$

EXEMPEL 2.8. Optimala ITAE system med $e_0 = e_1 = 0$

De system som angavs i exempel 2.7 har alla felkoefficienten $e_1 \neq 0$. Om det slutna systemets överföringsfunktion väljs som

$$G(s) = \frac{p_{n-1}\,\omega_0^{n-1} s + \omega_0^n}{s^n + p_1\,\omega_0 s^{n-1} + \ldots + p_{n-1}\,\omega_0^{n-1} s + \omega_0^n} = \frac{p_{n-1}\,\omega_0^{n-1} s + \omega_0^n}{P_n(s)}$$

erhålls ett slutet system med $e_1 = 0$. De optimala ITAE systemen ges då av

$$P_2(s) = s^2 + 3{,}2\omega_0 s + \omega_0^2$$

$$P_3(s) = s^3 + 1{,}75\omega_0 s^2 + 3{,}25\omega_0^2 s + \omega_0^3$$

$$P_4(s) = s^4 + 2{,}41\omega_0 s^3 + 4{,}93\omega_0^2 s^2 + 5{,}14\omega_0^3 s + \omega_0^4$$

$$P_5(s) = s^5 + 2{,}19\omega_0 s^4 + 6{,}5\omega_0^2 s^3 + 6{,}3\omega_0^3 s^2 + 5{,}24\omega_0^4 s + s^5$$

$$P_6(s) = s^6 + 6{,}12\omega_0 s^5 + 13{,}42\omega_0^2 s^4 + 17{,}16\omega_0^3 s^3 + 14{,}14\omega_0^4 s^2 + 6{,}76\omega_0^5 s + \omega_0^6$$

Känslighet

De dynamiska egenskaperna hos den process som skall regleras är sällan exakt kända på förhand. Processdynamiken kan också ändras av olika orsaker. Det är därför vanligt att i specifikationerna kräva att prestandakrav skall uppfyllas även om processens parametrar varierar inom föreskrivna gränser. Det följer av känslighetsanalysen i avsnitt 4.3 att sådana krav ofta kan översättas i krav på kretsförstärkningen G_0.

Samband mellan olika systemspecifikationer

Det är ett mycket svårt problem att korrekt specificera ett systems syfte. För konstruktören, som skall syntetisera ett reglersystem, är det bekvämt att ha specifikationerna i form av villkor på systemets överföringsfunktion eller i form av en förlustfunktion, eftersom dessa ligger till grund för systemanalysen. För beställaren eller användaren är det däremot mycket mer naturligt att ge specifikationer i form av krav på systemets svar på typiska insignaler då dessa direkt visar systemets egenskaper i specifika situationer. Sådana specifikationer är också lätta att kontrollera genom direkta mätningar på systemet. Beställaren kan även i många fall ha svårt att överblicka innebörden av överföringsfunktionens egenskaper eller innebörden av de olika termerna i en förlustfunktion som givits matematisk form. Det är därför mycket väsentligt att konstruktören har en överblick av sambanden mellan de olika sätten att specificera systemets egenskaper. Det händer mycket ofta att specifikationerna ändras under konstruktionsarbetets gång. Det kan t. ex. visa sig vara antingen mycket komplicerat eller mycket dyrt att göra ett system som uppfyller de primärt givna specifikationerna, medan en mindre ändring av specifikationerna kan leda till avsevärda förenklingar. Vid praktiska tillämpningar av optimeringskriterier har det ofta visat sig vara mycket svårt att vid den första analysen få med alla begränsningar och man får då successivt modifiera problemställningen.

Några samband mellan olika systemspecifikationer skall nu diskuteras. För det enkla systemet i exempel 2.1 kan alla specifikationer uttryckas med en enda parameter. I detta fall finns således mycket begränsade möjligheter att

välja specifikationerna oberoende av varandra. Genom att välja mer komplicerade överföringsfunktioner ökar dock möjligheterna att göra detta. Analysen av inverkan av en dipol, ekv. (2.6), visar t. ex. att felkoefficienterna kan väljas oberoende av stigtid och bandbredd. Det finns dock vissa samband mellan specifikationerna som är oberoende av överföringsfunktionens komplexitet. Stigtid, fördröjningstid och insvängningstid hänger t. ex. samman. Överslängen kan däremot väljas oberoende av dessa.

Det finns också ett samband mellan stigtid och bandbredd som säger att produkten av dessa storheter är approximativt konstant. För att inse detta betraktas ett stabilt system med överföringsfunktionen $G(s)$. Systemets stegsvar har då Laplacetransformen $G(s)/s$ och stegsvarets tidsderivata har alltså transformen $G(s)$. Genom att tillämpa formeln (2:8.3) för inverstransformering finner vi då

$$\frac{dc(t)}{dt} = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{st} G(s)\, ds$$

och

$$\max_{0\leqslant t<\infty} \frac{dc(t)}{dt} \leqslant \frac{1}{2\pi}\int_{-\infty}^{\infty} \max_{0\leqslant t<\infty} \left|e^{i\omega t} G(i\omega)\right| d\omega = \frac{1}{\pi}\int_{-0}^{\infty} \left|G(i\omega)\right| d\omega$$

Eftersom $\lim_{t\to\infty} y(t) = G(0)$ ger definitionen av T_r

$$T_r \cdot \int_0^{\infty} \left|G(i\omega)/G(0)\right| d\omega \geqslant \pi \cdot T_r \cdot \max_{0\leqslant t<\infty} \frac{dc}{dt} \Big/ G(0) = \pi \tag{2.15}$$

där T_r är stigtiden. Men för en överföringsfunktion med en frekvenskurva enligt fig. 2.4 gäller approximativt

$$\int_0^{\infty} \left|G(i\omega)/G(0)\right| d\omega \approx \omega_B \tag{2.16}$$

där ω_B är bandbredden. Vi finner således att produkten av stigtid och bandbredd är av storleksordningen 3. För att illustrera noggrannheten i de gjorda approximationerna ges nedan några exakta uttryck.

EXEMPEL 2.9.

För systemet i exempel 2.1 gäller exempelvis $\omega_B \cdot T_r = 1$. I detta fall är (2.16) en dålig approximation, ty frekvenskurvan avtar långsamt. Integralen i vänsterledet av (2.16) blir i själva verket

EXEMPEL 2.10.

För ett slutet system med överföringsfunktionen

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

gäller

$$\omega_B T_r = e\sqrt{\sqrt{2}-1} = 1{,}75$$

EXEMPEL 2.11.

För systemet i exempel 2.2 beror stigtids-bandbreddsprodukten av ζ. För $\zeta = 0{,}707$ gäller

$$\omega_B T_r = \exp(\pi/4) = 2{,}19$$

och för $\zeta = 0$ gäller

$$\omega_B T_r = 1{,}55$$

För servosystem är det vanligt att specificera snabbhet, dämpning, statiska fel och systemets förmåga att eliminera störningar. Systemets snabbhet kan anges med stigtid T_r, bandbredd ω_B eller med de dominerande polernas avstånd från origo. Systemets dämpning kan anges med översläng M, resonanstopp M_p eller de dominerande polernas relativa dämpning ζ. De statiska felen anges med felkoefficienterna och systemets förmåga att eliminera störningar med styvhetskoefficienter och med

$$M_d = \max_{\omega} |G_d(i\omega)|$$

där $G_d(s)$ är överföringsfunktionen från störning till utsignal.

3. Reglerprinciper och regulatorstrukturer

En *reglerprincip* anger hur en process skall regleras. Den beskriver i stora drag vilka processvariabler som skall hållas konstanta och vilka som skall varieras för att kompensera för olika störningar. Vi illustrerar med några exempel. Vid reglering av en kraftstation kan den effekt som stationen levererar till nätet regleras så att nätfrekvensen hålls konstant. För att reglera ett värmekraftverk kan ångtrycket hållas konstant vid varierande uteffekt (konstanttrycksreglering) eller alternativt kan ångtrycket variera med uteffekten (glidtrycksreglering). Reglering av en målsökande robot kan göras så att den hela tiden styr mot målet (hundkurvestyrning) eller så att en tänkt linje från roboten till målet har konstant orientering i rymden (syftbäringsprincipen). Reglersystemet för manövrering av en båt kan antingen utformas så att båtens vinkelhastighet är den primärt reglerade variabeln eller alternativt så att båtens girradie regleras (girradiestyrning). Reglerprinciper utformas på grundval av ingående studier av processerna. En bra reglerprincip medför ofta att regleringen blir enkel. Eftersom val av reglerprinciper baseras på en grundlig

förståelse av en speciell process kan man tyvärr ej ge några allmängiltiga regler för att finna reglerprinciper.

En *regulatorstruktur* anger mätstorheter och styrstorheter och deras inbördes koppling. Strukturen kan också ange regulatorns karaktär, *P*, *PI*, dynamiska system av ordning n etc. Däremot behöver regulatorns parametrar ej anges numeriskt. En regulatorstruktur anger således de grova dragen av ett reglersystem. Det har tidigare påpekats att syntesproblem sällan kan lösas rent analytiskt. I praktiken löses problemen ofta genom att ansätta en lämplig regulatorstruktur utgående från en reglerprincip. Det slutna systemets egenskaper undersöks sedan med analys och simulering och regulatorparametrarna och strukturen ändras successivt tills tillfredsställande prestanda erhålls. Som en förberedelse för en detaljerad genomgång av olika syntesmetoder skall några olika regulatorstrukturer diskuteras.

PID regulatorn

I samband med analysen i kapitel 4 visades att enbart återkoppling har många gynnsamma egenskaper. Många reglerproblem kan också lösas med mycket enkla regulatorer. Nästan alla reglerproblem i processindustrin löses t. ex. med *PI* eller *PID* regulatorer. Nedan ges exempel på problem som kan lösas med sådana enkla regulatorer.

EXEMPEL 3.1.

Betrakta en process som beskrivs av

$$m\frac{d^2y}{dt^2} + d\frac{dy}{dt} + ky = u \tag{3.1}$$

Processen kan t. ex. representera fjäderupphängd massa med viskös dämpning där styrvariabeln är kraften på massan. Vi skall nu visa hur systemets egenskaper kan modifieras med olika enkla regulatorer. Antag först att systemet är dåligt dämpat. Den deriverande regulatorn (*D* regulatorn)

$$u = y^r - k_1\frac{dy}{dt}$$

ger ett slutet system som beskrivs av

$$m\frac{d^2y}{dt^2} + (d + k_1)\frac{dy}{dt} + ky = y^r$$

En jämförelse med (3.1) visar att den deriverande återkopplingen ger en ökning av dämpningskoefficienten. Hur mycket dämpningen kan ökas beror på hur hög förstärkningen k_1 i den deriverande återkopplingen görs. I praktiken begränsas k_1 av störningarna i mätsignalen (mätbruset).

Antag vidare att systemet ej är tillräckligt snabbt, dvs. parametern k är för liten. Genom att införa den proportionella och deriverande regulatorn (PD regulatorn)

$$u = -k_1\frac{dy}{dt} - k_2(y - y^r)$$

erhålls det slutna systemet

$$m\frac{d^2y}{dt^2} + (d + k_1)\frac{dy}{dt} + (k + k_2)\,y = k_2 y^r$$

och koefficienten framför y-termen kan således väljas godtyckligt genom lämpligt val av parametern k_2. Observera att om parametrarna k_1 och k_2 kan väljas tillräckligt stora så blir det slutna systemets egenskaper oberoende av parametrarna d och k. Det slutna systemets egenskaper kommer däremot att variera med parametern m även om k_1 och k_2 är mycket stora. I praktiken begränsas parametrarnas storlek av den maximalt tillgängliga kraften (reglerauktoriteten).

Med en P eller PD regulator erhålls i stationärt tillstånd

$$y = k_2/(k + k_2)\,y^r$$

Systemet har således ett stationärt fel som är omvänt proportionellt mot parametern k_2. För att helt undvika ett stationärt fel kan en integrerande term införas och vi erhåller då följande regulator

$$u(t) = -k_1\frac{dy}{dt} - k_2(y - y^r) - k_3\int^t [y(\tau) - y^r(\tau)]\,d\tau \tag{3.2}$$

I regulatorn (3.2) är styrsignalen en summa av tre termer som är proportionella mot reglerfelet, reglerfelets tidsintegral och utsignalens tidsderivata. En sådan regulator kallas PID regulator. Observera att man ej deriverar reglerfelet, ty då skulle en stegändring i referensvärdet y^r ge upphov till mycket stora styrsignaler till processen. Parametrarna i en PID regulator brukar ofta anges på följande sätt:

$$u(t) = -K\left[T_D\frac{dy}{dt} + (y - y^r) + \frac{1}{T_I}\int^t [y(\tau) - y^r(\tau)]\,d\tau\right] \tag{3.3}$$

där parametern T_I kallas integraltid och T_D derivatatid. Ibland anges också $K = 100/P$ där P kallas proportionalband. Anledningen till detta är historisk. De första PID regulatorerna förverkligades med pneumatiska komponenter och det var enklast att göra parametrarna P, T_I och T_D justerbara. En elektrisk PID regulator kan förverkligas med återkopplade operationsförstärkare. I ett sådant fall är det lätt att fritt välja de justerbara parametrarna. Ett blockschema för regulatorn (3.3) visas i fig. 3.1.

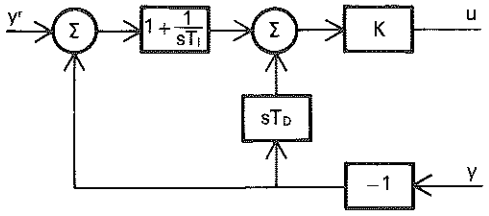


Fig. 3.1. Blockschema för en *PID* regulator. Regulatorn har två insignaler, mätsignalen y och referensvärdet y^r. Regulatorns utsignal u är processens styrsignal.

Det förekommer också att regulatorn parametriseras som

$$u(t) = -K\left(1+\frac{1}{T_1 p}\right)[(1+T_2 p)y - y^r]$$

där p betecknar differentialoperatorn $p = d/dt$. Det påstås ibland att det är lättare att justera parametrarna i denna regulator än i regulatorn (3.3).

Ekvationen (3.3) ger en idealiserad beskrivning av *PID* regulatorn. I verkligheten är regulatorns utsignal alltid begränsad. Regulatorns förstärkning vid höga och låga frekvenser är också begränsad.

I datorsystem för processreglering programmeras *PID* funktionen som en enkel differensapproximation av (3.3). Sådana program utgör en av byggstenarna i standardsystem för processreglering med dator, s. k. DDC program (Direct Digital Control Package).

En *PID* regulator är ofta utformad så att det går att koppla om mellan automatisk och manuell reglering. Vid omkoppling från manuell till automatisk reglering kan det inträffa att styrsignalen ändras språngvis även om $y = y^r$ på grund av att integralen i (3.3) har fel värde. Regulatorn brukar därför utformas så att integratorn automatiskt ställs in på rätt värde vid omkoppling. Detta kallas på engelska för »bumpless transfer», vilket skulle kunna översättas med stötfri omkoppling.

Det följer av (3.3) att integraltermen kan bli stor om reglerfelet har samma tecken under en längre tid. Förutom vid omkoppling till manuell reglering kan detta inträffa om processen stängs eller om processens styrsignal mättas eller begränsas. Detta kan orsaka stora överslängar och eventuellt också självsvängning. Fenomenet, som på engelska kallas »reset windup», kan elimineras genom att integratorn nollställs vid stora avvikelser.

Det är brukligt att *PID* regulatorer utformas så att derivatatermen går att koppla ifrån. I många fall är derivatatermen ej heller nödvändig och regulatorn kallas då en PI regulator. Analogt talar man också om P och PD regulatorer. Eftersom dessa typer är så vanliga, finns mycket empirisk kunskap om val av regulatorer och deras inställning, som tillhandahålls av regulatorleverantörer. Ett exempel ges i tabell 3.1. Många processer i processindustrin har monotont växande stegsvar. Sådana processer kan ofta regleras tillfredsställande med *PI* eller *PID* regulatorer. En vanlig tumregel säger att det ej är nödvändigt

Tabell 3.1. Exempel på anvisningar för regulatorval. I kolumnen med rubriken egenskaper ges först processens väsentliga karakteristika. Sedan ges rekommenderade regulatorinställningar och lämplig ventilkarakteristik.

	Process				
Egenskaper	Flöde och och vätsketryck	Gastryck	Vätskenivå	Temperatur och ångtryck	Koncentration
Linjär/olinjär	Kvadratisk	Linjär	Linjär	Olinjär	—
Ordningstal	>1	1	1	Högt	Variabelt
Mätbrus	Väsentligt	Litet	Väsentligt	Litet	—
Dödtid	Nej	Nej	Nej	Variabel	Variabel
Tidskonstant	1–10 s	Liten	1–10 s	min–tim	min–tim
Processförstärkning	0,5–5	Integrerande	Integrerande	1–2	10–1 000
Regulatorförstärkning	0,2–2	20–100	2–20	1–10	0,1–1
Integralverkan	Viktig	Behövs ej	Sällan	Ja	Viktig
Derivataverkan	Nej	Behövs ej	Nej	Viktig	Om möjligt
Ventilkarakteristik	Linjär	Logaritmisk	Linjär	Logaritmisk	Linjär

med en deriverande term om $T_d > 1.5T_r$, där T_d är fördröjningstiden och T_r stigtiden. Jämför fig. 2.1.

Det finns många sätt att ställa in PID regulatorer. Man kan t. ex. utgå från processens stegsvar som erhålls antingen genom processanalys eller genom direkt mätning. I stegsvaret dras en tangent i den punkt där stegsvaret har maximal lutning. Ur tangentens skärningar med koordinataxlarna bestäms sedan parametrarna a och b. Se fig. 3.2. Lämpliga regulatorinställningar framgår av tabell 3.2. En alternativ metod är att man först kopplar in en proportionell regulator. Regulatorns förstärkning ökas sedan tills det slutna systemet självsvänger med konstant amplitud. Det kritiska värdet på förstärkningen K_c och svängningens period T_0 observeras. Lämpliga regulatorinställningar kan sedan erhållas ur rekommendationer som givits av Ziegler och Nichols. Se tabell 3.3. Observera att denna metod endast kan användas då processen utan risk kan försättas i självsvängning.

Zieglers och Nichols rekommendationer är härledda för linjära system, vars dynamik kan beskrivas med en tidsfördröjning och två tidskonstanter. Inställningarna får därför betraktas som grova riktvärden, vilka kan utgöra utgångspunkter för trimning av regulatorerna. Om man har bättre kunskap om processens dynamik kan de metoder som behandlas i avsnitt 6.6 användas för att finna lämpliga regulatorinställningar. Det bör speciellt uppmärksammas att Zieglers och Nichols rekommendationer kan ge felaktiga inställningar för processer, vars dynamik approximativt kan beskrivas som första och andra ordningens system.

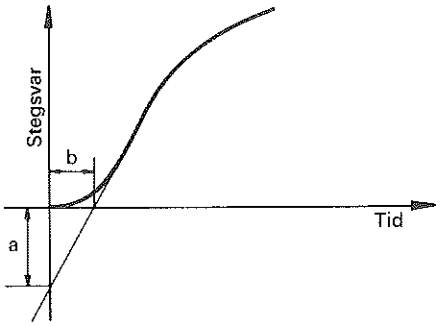


Fig. 3.2. Bestämning av parametrarna a och b ur processens svar på ett enhetssteg.

Kaskadreglering

Det finns reglersystem som har många mät- och styrstorheter. Det är attraktivt att bygga upp sådana system genom kombination av enkla regulatorer. Detta underlättar drifttagning, trimning och underhåll av system. Eftersom en *PID* regulator har två insignaler, mätsignalen och börvärdet, så kan komplicerade regulatorer byggas upp genom kombination av *PID* regulatorer genom att låta utsignalen från en *PID* regulator vara referensvärde till en annan *PID* regulator osv. Detta brukar kallas för *kaskadreglering*. Ett enkelt exempel på kaskadreglering visas i fig. 3.3. I detta fall regleras en ugn för uppvärmning av en produkt genom att reglera oljeflödet till brännarna på basis av mätning av oljeflöde och produkttemperatur. Den ena regulatorn som styr reglerventilen har oljeflödet som insignal. Börvärdet till denna regulator erhålls som utsignal från den andra regulatorn som har temperaturen som insignal.

Tabell 3.2. Rekommenderade regulatorinställningar enligt Ziegler och Nichols. Parametrarna a och b bestäms ur processens svar på ett enhetssteg enligt fig. 3.2.

Regulator	K	T_I	T_D
P	$1/a$		
PI	$0{,}9/a$	$3\,b$	
PID	$1{,}2/a$	$2\,b$	$b/2$

Tabell 3.3. Rekommenderade regulatorinställningar enligt Ziegler och Nichols.

Regulator	K	T_I	T_D
P	$0{,}5\,K_c$		
PI	$0{,}45\,K_c$	$T_0/1{,}2$	
PID	$0{,}6\,K_c$	$T_0/2$	$T_0/8$

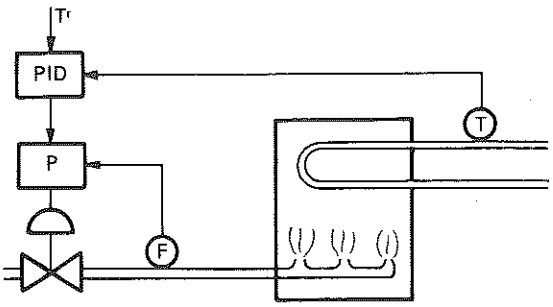


Fig. 3.3. Exempel på kaskadreglering.

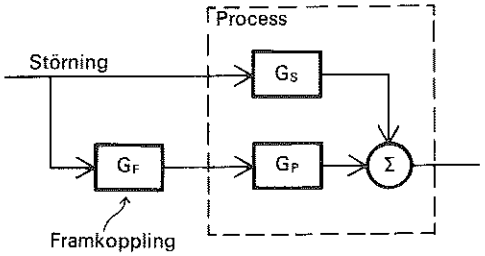


Fig. 3.4. Eliminering av en mätbar störning med framkoppling.

I detta fall är kaskadreglering mycket tilltalande, ty regulatorn i den inre kretsen eliminerar inverkan av den olinjära ventilkarakteristiken. Jämför Exempel 4: 3.2. Genom att flödet mäts undviks också den svårighet som orsakas av att flödet beror av såväl ventilöppning som oljetryck.

Framkoppling

Även om återkopplingsprincipen är den bärande idén inom reglertekniken så finns det också andra sätt att eliminera störningar. Med ett återkopplat system måste ju först ett reglerfel erhållas innan regulatorn kan vidtaga åtgärder för att eliminera störningen. I vissa fall är det möjligt att mäta störningen innan den givit upphov till ett reglerfel. Det är då naturligt att försöka motverka störningens inverkan innan den hunnit påverka utsignalen. Detta sätt att reglera kallas *framkoppling*, i engelsk litteratur *feedforward* och i tysk litteratur *Störgrössenaufschaltung*. Principen är enkel. Betrakta ett linjärt system. Låt G_S vara överföringsfunktionen från störning till utsignal och G_P överföringsfunktionen från styrsignal till utsignal. Se fig. 3.4. Om den mätta störningen kopplas till styrsignalen med ett system med överföringsfunktionen G_F där

$$G_S + G_P G_F = 0 \tag{3.4}$$

så kommer uppenbarligen störningens inverkan på utsignalen att elimineras.

Framkoppling är i många fall ett effektivt hjälpmedel. Dess stora fördel är

att snabbt varierande störningar kan elimineras därigenom att kompenseringsåtgärder vidtas innan verkan av störningen visat sig i avvikelser i utsignalen. Framkoppling kan utnyttjas för såväl linjära som olinjära system.

Framkopplingens användbarhet begränsas av att den endast kan användas för att eliminera mätbara störningar och ändringar i referensvärden. En annan nackdel är att framkoppling kräver en modell av processen. Uttrycken *statisk* respektive *dynamisk framkoppling* används ibland för att indikera om framkopplingen baserats på en statisk eller dynamisk modell av processen. Då framkopplingen baseras på att två signaler skall ta ut varandra blir framkopplingen också känslig för parametervariationer. Ett fel i modellen på 10 % medför således att störningen elimineras till 90 %. Återkopplingen behöver således endast ta hand om en tiondel av störningen. Felaktigheter i modellen återspeglas således direkt i reglerfel. Detta medför att framkoppling kräver bättre trimning än återkoppling. Då framkopplingen kräver en modell av processen kan framkopplingen också vara mer komplicerad än en enkel PI regulator. Detta har medfört att framkoppling ej är lika vanlig som PI reglering. Framkoppling har därför hittills mest använts för reglerproblem med höga prestandakrav. Typiska exempel är reglering av domnivån i en ångpanna, reglering av destillationskolonner och valsverksreglering. Det finns dock fall då en enkel proportionell framkoppling har gett goda resultat. Då kostnaderna för signalbehandling minskat kraftigt de senaste åren kan man vänta sig en ökad användning av framkoppling även för enklare reglerkretsar. Vissa standardregulatorer har en framkopplingsterm. Vid processreglering med datorer kan t. ex. framkoppling erhållas till priset av en mycket måttlig ökning i datorprogrammets storlek. Vid reglering av kontinuerliga processer, som består av flera seriekopplade processteg, består de väsentliga störningarna ofta av variationer i produktkvalitet och produktflöde från ett tidigare processteg. I sådana fall kan störningarna mätas och framkoppling kan tillämpas. Det förekommer också att börvärdesändringar i ett tidigare processteg används som framkopplingssignal.

En allmän regulatorstruktur som kombinerar framkoppling och återkoppling

I fig. 3.5 visas en generell regulatorstruktur som medger en kombination av framkoppling och återkoppling. Såsom framgår av fig. 3.5 kan framkoppling utnyttjas såväl för att eliminera inverkan av störningar som för att förbättra möjligheterna att följa insignaler.

Efter enkla räkningar finner vi att reglerfelets Laplacetransform blir

$$E(s) = -\frac{G_{P1}(1 - G_{P2}G_{F2})}{1 + G_P G_R} V(s) + \frac{1 - G_P G_{F1}}{1 + G_P G_R} Y^r(s) \tag{3.5}$$

där G_{F1} och G_{F2} är överföringsfunktionerna för framkopplingarna från referens-

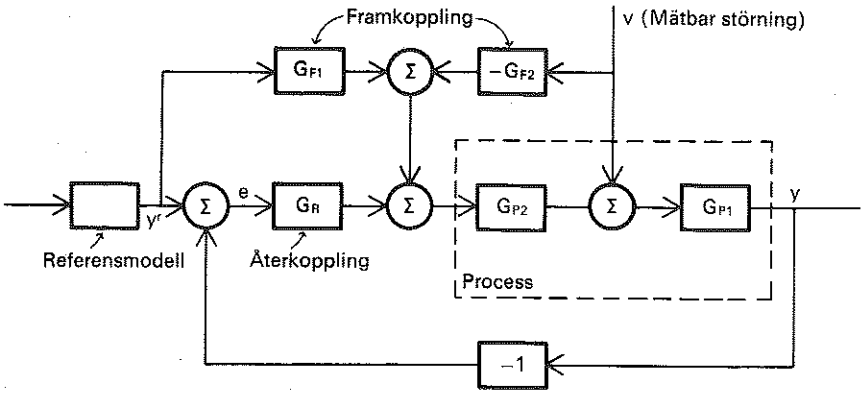


Fig. 3.5. En allmän regulatorstruktur som kombinerar framkoppling och återkoppling.

värde respektive störning, G_R är överföringsfunktionen för återkopplingen och

$$G_P = G_{P1} G_{P2} \tag{3.6}$$

är processens överföringsfunktion. Överföringsfunktionen G_{F1} som representerar framkoppling från referensvärdet kallas ibland additiv korrektion.

Överföringsfunktionen från referensvärde till reglerfel är således

$$G_{y'}(s) = \frac{1 - G_P G_{F1}}{1 + G_P G_R} \tag{3.7}$$

Även om det ej finns någon framkoppling, dvs. $G_{F1} = 0$, finner vi att överföringsfunktionen G_{yr} är liten om kretsöverföringen $G_0 = G_P G_R$ är stor.

Då kretsöverföringen $G_0 = G_R G_P$ vanligen är stor för frekvenser, som är lägre än systemets bandbredd eller dess skärningsfrekvens, är överföringsfunktionen G_{yr} liten för sådana frekvenser oavsett valet av framkoppling. Genom lämpligt val av återkoppling är det således möjligt att få G_{yr} liten i frekvensområdet $(0, \omega_B)$, där ω_B är systemets bandbredd, genom att se till att nämnaren i (3.7) är stor. Observera att genom att välja framkopplingen så att

$$G_P G_{F1} = 1 \tag{3.8}$$

blir $G_{yr} = 0$ för alla frekvenser. Villkoret (3.8) innebär att G_{F1} skall väljas som inversen av processens överföringsfunktion.

Det följer av ekvation (3.5) att överföringsfunktionen från störning till reglerfel är

$$G_v = -\frac{G_{P1}(1 - G_{P2} G_{F2})}{1 + G_P G_R} \tag{3.9}$$

Vi finner analogt med den tidigare diskussionen att G_v är noll om framkopplingen väljs så att

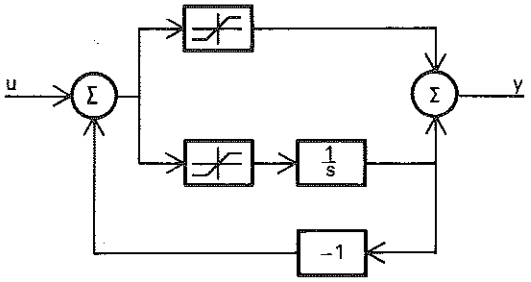


Fig. 3.6. Olinjär referensmodell. Utsignalen följer insignalen exakt vid små signalamplituder. Vid stora insignalamplituder är dock utsignalens ändringshastighet begränsad.

$$G_{P2}G_{F2} = 1 \tag{3.10}$$

Detta innebär att framkopplingen väljs så att den signal som passerar blocken med överföringsfunktionerna G_{P2} och G_{F2} exakt eliminerar störningen v. Jämför fig. 3.4. Nämnaren i överföringsfunktionen G_v kan liksom tidigare göras stor genom lämpligt val av återkoppling.

Framkoppling och återkoppling har således delvis komplementära egenskaper. Med återkoppling är det således möjligt att eliminera störningar som har frekvenser lägre än systemets bandbredd. Framkoppling tillåter elimination av störningar inom ett större frekvensband. Återkoppling är robust i den meningen att den är relativt okänslig för variationer i processmodellen medan framkoppling kräver en processmodell och dessutom är känslig för parametervariationer. Det kan också noteras att återkoppling kan leda till instabilitet medan framkoppling ej ger några stabilitetsproblem. För att få ett bra reglersystem är det fördelaktigt att kombinera framkoppling och återkoppling.

Det bör slutligen observeras att regulatorn i fig. 3.5 också innehåller ett block som kallas referensmodell. Dess dynamik anger det önskade sambandet mellan referensvärde och utsignal. Ett annat skäl till att använda en referensmodell är att det finns fall då styrsignalens extremvärden är sådana att det tar relativt lång tid för utsignalen att variera över hela sitt variationsområde. För att undvika att systemet mättas kan referensmodellen t. ex. utformas enligt fig. 3.6, vilket medför att ändringshastigheten i y^r är begränsad oberoende av kommandosignalen.

Framkoppling användes första gången omkring 1925 i samband med reglering av domnivån i en ångpanna. Detta exempel behandlas nedan.

EXEMPEL 3.1. Trepunktskoppling

En förenklad bild av en ångpanna visas i fig. 3.7. Vattnet i stigtuberna värms av brännarna. Vid uppvärmningen kokar vattnet och ånga bildas. Blandningen av vatten och ånga har lägre densitet än vatten. Därigenom uppstår självcirkulation i dom, falltuber, stigtuber. Ångan skiljs från ång-vattenblandningen i domen. Med hjälp av turbinventilen bestäms ångmängden till turbinen. Vattentillförseln till domen regleras med matarvattenventilen. Det är viktigt att vattennivån i domen hålls konstant. Vid för låg vattennivå blir kylningen

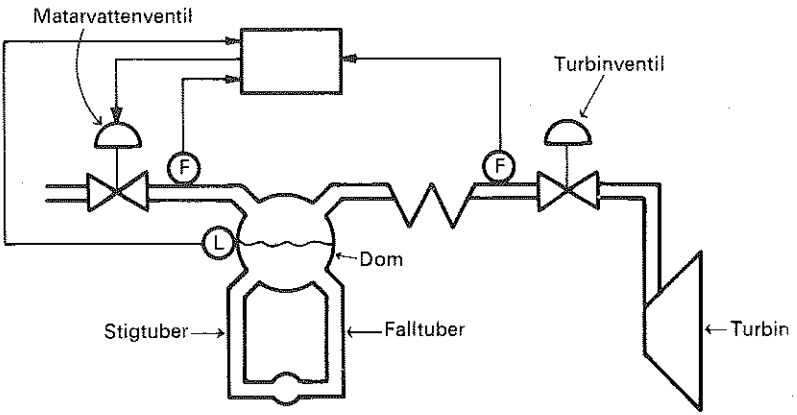


Fig. 3.7. Förenklad bild av en ångpanna med trepunktsreglering av domnivån.

av stigtuberna sämre och det finns risk att de bränns sönder. Vid för hög vattennivå finns risk att vatten kan sugas in i turbinen, vilket kan resultera i haveri. För att se till att nivån hålls inom rimliga gränser finns dels ett reglersystem som eftersträvar att hålla nivån konstant och dels ett säkerhetssystem som stänger av brännarna om nivån trots regleringen kommer utanför vissa gränser.

Problemet att reglera vattennivån i domen är svårt, vilket kan inses av följande fysikaliska resonemang. Anta att systemet är i jämvikt med konstant nivå och att matarvattenventilen hålls konstant. Om ånguttaget minskar genom att turbinventilen stryps så kommer trycket i domen att stiga. Därmed stiger också trycket i stigtuberna. Förångningstemperaturen stiger och mängden ångbubblor i stigtuberna minskar. Domnivån avtar därför till att börja med. Eftersom ånguttaget minskat är matarvattenflödet större än ånguttaget och domnivån kommer därför så småningom åter att stiga. Vi finner således att vid en minskning i ånguttaget kommer domnivån först att avta och sedan raskt öka igen. Fenomenet har i engelsk litteratur givits den talande benämningen »shrink and swell phenomenon». En formell analys visar att det dynamiska system som beskriver processen ej är ett minimum-fas-system. Effekten blir ännu större om ångpannan försetts med en regulator. Då domnivån sänks i början så kommer matarvattenventilen att öka vattentillförseln på grund av regleringen. Efter tillräckligt lång tid efter störningen kommer regleringen dock att se till att nivån intar rätt värde. Det reglerade systemet kommer således att få en kraftig översläng. Överslängen ökar med ökande förstärkning hos regulatorn. Vi finner således att reglersystemets reaktion till en början är helt motsatt den önskade. Om ånguttaget minskar så borde rimligtvis också matarvattenflödet minskas. Detta kan åstadkommas genom att mäta differensen mellan ångflöde och matarvattenflöde och framkoppla denna signal till domnivåregulatorn. Regulatorns blockschema framgår av fig. 3.7. Användning av framkoppling i detta fall betraktades på sin tid (1925) som en väsentlig reglerteknisk innovation. Den speciella kombinationen av fram- och

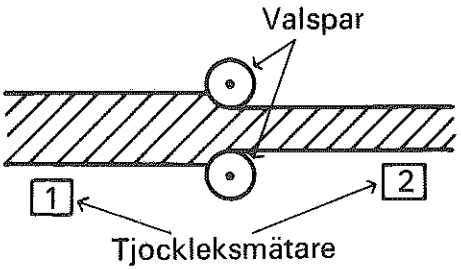


Fig. 3.8. Schematisk bild av bandvalsverk med två tjockleksmätare.

återkoppling som visas i fig. 3.7 kallas trepunktskoppling och är än idag standardmetoden för reglering av domnivån.

Kombinationen av fram- och återkoppling belyses med ytterligare två exempel.

EXEMPEL 3.2. Bandvalsning

Betrakta ett valspar i ett bandvalsverk (fig. 3.8). Jämför exempel 2:8.3. Det är önskvärt att tjockleksvariationerna hos bandet som lämnar valsparet är så små som möjligt. Variationerna i tjocklek orsakas bland annat av variationer i hårdhet och tjocklek hos bandet vid inträde i valsparet. För att eliminera dessa variationer styrs valsparets ansättning från tjockleksmätaren 2, som mäter det utgående bandets tjocklek. Reglersystemet består i detta fall av en återkoppling från tjockleksmätaren 2 till valsparets ansättning. På grund av den tid det tar för bandet att passera från valsparet till tjockleksmätaren 2, kan återkopplingen ej eliminera snabba variationer i det ingående bandets tjocklek. Om en tjockleksmätare placeras före valsparet kan även snabba variationer elimineras, ty bandets tjocklek före inträdet i valsparet är då känd. Vi skall nu göra en enkel analys i anslutning till ovanstående resonemang. Låt bandets tjocklek och hårdhet vid passage av tjockleksmätaren 1 vara $v_1(t)$ respektive $v_2(t)$. Låt styrvariabeln $u(t)$ vara valsparets ansättning, och låt systemets utsignal $y(t)$ vara bandets tjocklek vid passage av tjockleksmätaren 2. Om vi endast studerar små avvikelser, kan systemet beskrivas med följande ekvation

$$Y(s) = G_1(s)\,U(s) + G_2(s)\,V_1(s) + G_3(s)\,V_2(s) \tag{3.11}$$

där G_1 är överföringsfunktionen från valsparets ansättning till utgående bandets tjocklek, G_2 är överföringsfunktionen från ingående bandets tjocklek till utgående bandets tjocklek och G_3 är överföringsfunktionen från ingående bandets hårdhet till utgående bandets tjocklek. Det följer nu av ekvation (3.11) att styrlagen

$$U(s) = -\frac{G_2(s)}{G_1(s)}\,V_1(s) \tag{3.12}$$

ger

$$Y(s) = G_3(s)\, V_2(s)$$

Inverkan på utsignalen av variationer i det ingående bandets tjocklek har således helt eliminerats med hjälp av styrlagen (3.12). Vi kan då fråga om tjockleksmätaren 2 kan slopas helt. För att belysa denna fråga, skall vi närmare studera hur styrlagen (3.12) fungerar. Anta att det band som passerar tjockleksmätaren 1 plötsligt blir tjockare. Styrlagen (3.12) innebär då att valsparets ansättning skall ändras så att inflytandet av tjockleksökningen i det ingående bandet helt elimineras. Detta kräver uppenbarligen en ingående kunskap om hur en ändring i valsparets ansättning påverkar det utgående bandets tjocklek såväl statiskt som dynamiskt, med andra ord en exakt kunskap om överföringsfunktionen $G_1(s)$. Då överföringsfunktionen $G_1(s)$ bestäms av ett mycket komplicerat fysikaliskt förlopp, kan det vara mycket svårt att bestämma $G_1(s)$ noggrant. Dessutom kommer parametrarna i överföringsfunktionen $G_1(s)$ att variera med valsrullarnas slitage m. m. Om vi däremot gör ett återkopplat system, dvs. om valsparets ansättning styrs från tjockleksmätaren 2, finner vi att det reglersystem som erhålls kan göras mycket okänsligt för variationer i överföringsfunktionen $G_1(s)$. Jämför känslighetsanalysen i avsnitt 4:3. Vi kan t. ex. på vanligt sätt garantera att felet efter en stegstörning alltid är noll genom att införa en integration i återkopplingskretsen. För att i ett fall som detta erhålla en bra reglering, använder man därför båda tjockleksmätarna. Styrlagen blir då av formen

$$U(s) = H_1(s)\, Y(s) + H_2(s)\, V_1(s)$$

där $H_1(s)$ representerar återkopplingstermen (feedback) och $H_2(s)$ framkopplingstermen (feedforward). I praktiken brukar man välja överföringsfunktionerna $H_1(s)$ och $H_2(s)$ så att de högfrekventa störningarna elimineras genom inverkan av framkopplingen och de lågfrekventa störningarna elimineras av återkopplingen. För höga frekvenser gäller således

$$H_2(s) \sim -\frac{G_2(s)}{G_1(s)}$$

$$H_1(s) \sim 0$$

Övergången från det frekvensområde där framkopplingstermen dominerar till det där återkopplingstermen dominerar bestäms med ledning av störningarnas karaktär och osäkerheten i överföringsfunktionerna $G_1(s)$ och $G_2(s)$.

EXEMPEL 3.3. Temperaturreglering

I fig. 3.9 visas en förenklad bild av ett temperaturregleringssystem. En stabil oscillator kan vara ett praktiskt tillämpningsexempel. För att stabilisera en kristalloscillator önskar man hålla oscillatorkristallens temperatur konstant.

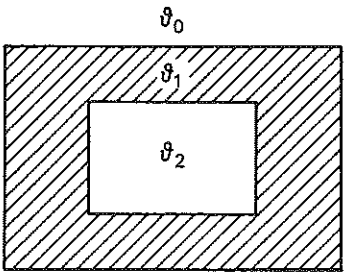


Fig. 3.9. Schematisk bild av temperaturstabiliserad ugn.

Man placerar då kristallen i ett kopparblock vars temperatur kan styras t. ex. med ett peltiereffektelement.

Låt yttertemperaturen vara ϑ_0, kopparblockets temperatur ϑ_1, kristallens temperatur ϑ_2, och låt styrvariabeln u vara det av peltiereffektelementet tillförda värmet. Om vi antar små avvikelser erhålls då följande samband

$$\theta_2(s) = G_1(s)\,U(s) + G_2(s)\theta_0(s)$$

Om yttertemperaturen ϑ_0 kan mätas exakt finner vi att styrlagen

$$U(s) = -\frac{G_2(s)}{G_1(s)}\theta_0(s)$$

helt eliminerar inverkan av variationer i yttertemperaturen. Jämför med ekvation (3.4). Även i detta fall kan man rikta invändningar mot denna styrlag av skäl som är helt analoga med dem som framfördes mot styrlagen (3.12) i exempel 3.2. Man använder därför i stället styrlagen

$$U(s) = H_1(s)\theta_2(s) + H_2(s)\theta_0(s)$$

som är en kombination av återkoppling (H_1) och framkoppling (H_2).

Vid reglering av temperaturen i bostadshus är det vanligt att man eliminerar inverkan av variationer i utomhustemperaturen såsom visats i exemplet men enbart med en framkoppling från en termometer placerad på husets norrsida.

Otto Smith-regulatorn

Ett system som innehåller tidsfördröjningar är icke-minimum-fas och kan vara svårt att reglera. Det finns en naturlig begränsning genom att det tar en viss tid innan en förändring av insignalen ger sig tillkänna i utsignalen. Då en tidsfördröjning har överföringsfunktionen $\exp(-sT)$, finner vi också att tidsfördröjningen orsakar en negativ fasförskjutning vars storlek växer linjärt med frekvensen. Återkoppling av ett system med tidsfördröjningar leder därför lätt till instabilitet. Den amerikanske forskaren Otto Smith har hittat på ett finurligt sätt att reglera ett system med en tidsfördröjning. Den regulator-

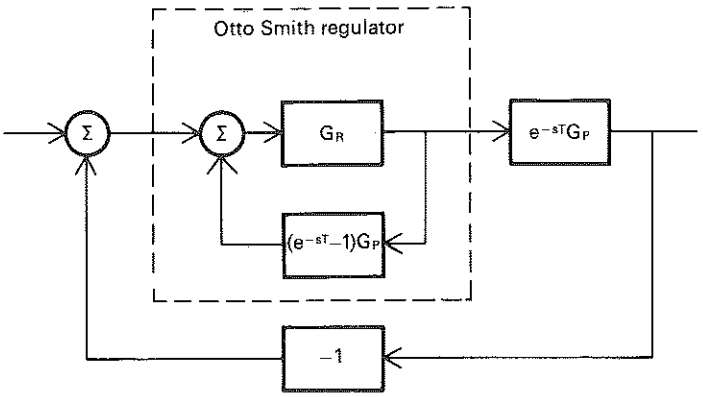


Fig. 3.10. Blockschema för en process med en Otto Smith-regulator.

struktur som Smith föreslagit brukar kallas *dödtidskompensering* eller *Otto Smith-regulator*.

För att beskriva idén betraktar vi först ett system som erhållits genom enkel återkoppling av en process med överföringsfunktionen G_P med en regulator med överföringsfunktionen G_R. Anta att regulatorn valts så att ett slutet system med önskade prestanda erhållits. Det slutna systemets överföringsfunktion är

$$G(s) = \frac{G_R(s)\,G_P(s)}{1 + G_R(s)\,G_P(s)} \tag{3.13}$$

Antag nu att processen ändras genom att dess insignal fördröjs T sekunder. Det slutna systemets överföringsfunktion blir då

$$G(s) = \frac{G_R(s)\,G_P(s)\,e^{-sT}}{1 + G_R(s)\,G_P(s)\,e^{-sT}}$$

Om tidsfördröjningen T är tillräckligt stor blir det slutna systemet instabilt. (Det räcker med att välja T så att $\omega_c T \geqslant \phi_m$ där ω_c är skärningsfrekvensen och ϕ_m fasmarginalen för $G_R G_P$.) För att få en stabil regulator med en tidsfördröjning har Otto Smith föreslagit den regulator som visas i fig. 3.10. Om blockschemat i fig. 3.10 ritas om såsom visas i fig. 3.11 finner vi att signalerna $-y$ och y_c tar ut varandra och det slutna systemet får överföringsfunktionen

$$G(s) = \frac{G_R(s)\,G_P(s)\,e^{-sT}}{1 + G_R(s)\,G_P(s)} \tag{3.14}$$

En jämförelse med ekv. (3.13) visar alltså att sånär som på faktorn $\exp(-sT)$ i täljaren så har det kompenserade systemet samma överföringsfunktion som i det fall då processen ej hade någon tidsfördröjning.

För att dimensionera en Otto Smith-regulator för en process med en tidsfördröjning så dimensioneras först en regulator som ger bra prestanda för en

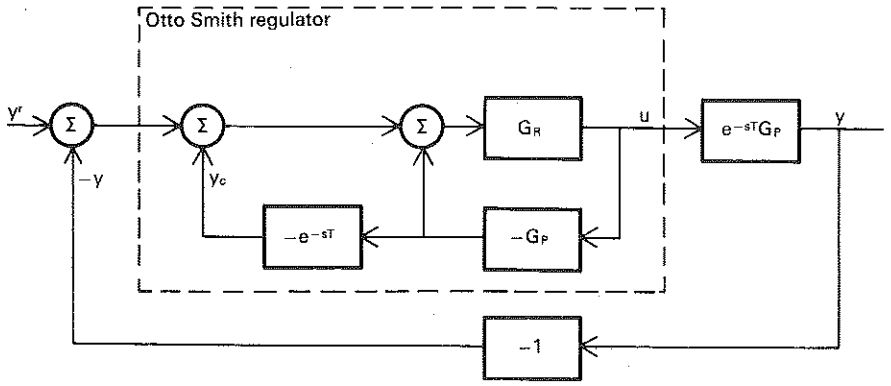


Fig. 3.11. Modifiering av blockschemat i fig. 3.10.

process utan tidsfördröjning. Därefter kompletteras regulatorn G_R såsom visas i fig. 3.10. Observera att regulatorn innehåller en modell av processdynamiken. Termen e^{-sT} i regulatorn betyder också att regulatorn måste innehålla komponenter som gör att styrsignalens värde fördröjs i T tidsenheter. Detta kan vara komplicerat att göra för kontinuerliga system men det är lätt att göra om en dator används för regleringen.

Statisk optimering och extremalsökande regulatorer

Vi har hittills uteslutande diskuterat regulatorer som strävar att hålla processvariablernas värden konstanta eller servosystem där utsignalen om möjligt skall följa insignalens variationer. Vid processreglering finns dessutom ett annat problem, nämligen att bestämma lämpliga referensvärden för reglering av de olika processvariablerna. Om regleringen fungerar någorlunda väl kan analysen göras med utgångspunkt från processvariablernas tidsmedelvärden. Det gäller således att välja processens driftpunkt så att t. ex. verkningsgrad eller driftsekonomi maximeras med hänsynstagande till begränsningar. Detta kan formuleras som ett rent statiskt optimeringsproblem, nämligen att maximera en funktion med bivillkor. Problem av detta slag kallas därför *statisk optimering*. Det finns många olika metoder att lösa sådana problem. Värdet med att införa en bättre reglering erhålls indirekt genom att en bättre reglering tillåter att processvariablernas värden kan väljas närmare begränsningarna.

Statisk optimering förutsätter att matematiska modeller för processen och dess prestationsfunktion är tillgängliga. Det finns fall då prestationsfunktionen kan bestämmas direkt genom mätningar och det är då möjligt att göra en regulator som ställer in processvariablerna så att prestationsfunktionen maximeras utan att någon matematisk modell för processen är känd. En sådan regulator kallas *extremalsökande*. Ett exempel på ett problem där en sådan regulator kan användas ges nedan.

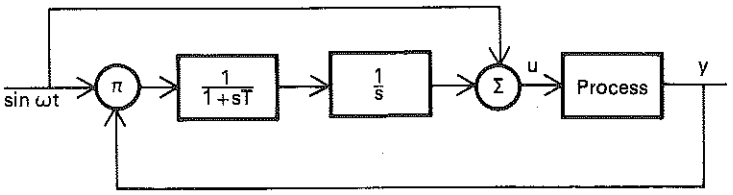


Fig. 3.12. En extremalsökande regulator enligt Draper och Li.

EXEMPEL 3.4.

Betrakta ett system som består av en gaseldad varmvattenberedare. Systemet skall leverera vatten med konstant temperatur. Vattenflödet kan variera inom vida gränser beroende på behovet. Då vattenmängden varierar måste naturligtvis också gasflödet till brännaren varieras så att temperaturen hela tiden är konstant. Detta kan ordnas t. ex. med en *PI* regulator som styr gasflödet och som har vattentemperaturen som mätsignal. Brännaren måste ha tillförsel av såväl luft som gas. Då förbränningsverkningsgraden beror på förhållandet mellan luft och gasflödet är det också nödvändigt att reglera luftflödet. Med hjälp av en s. k. *kvotregulator* kan luftflödet styras så att kvoten mellan luftflöde och gasflöde är konstant. Detta garanterar emellertid ej att verkningsgraden alltid är optimal. Det är dock möjligt att få ett approximativt uttryck på verkningsgraden genom att multiplicera vattenflödet med vattnets temperaturstegring och dividera med gasflödet. Man vill sedan ha en regulator som ställer in luftflödet så att förbränningsverkningsgraden är så stor som möjligt.

Det finns många sätt att göra en extremalsökande regulator. En metod visas i blockschemat i fig. 3.12. I denna regulator störs insignalen hela tiden perio-

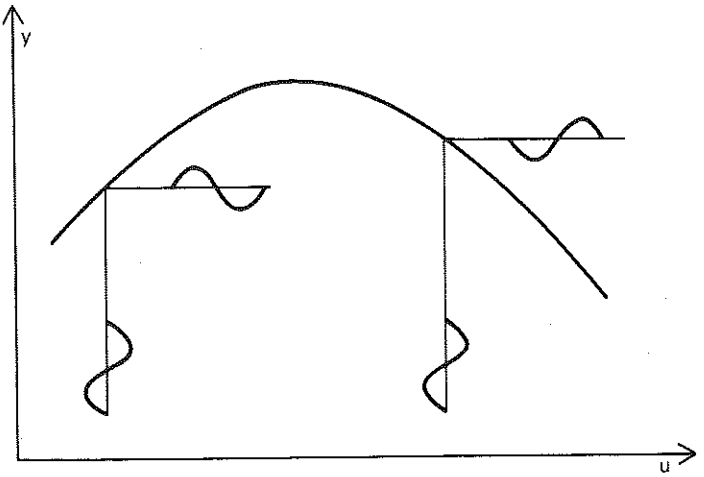


Fig. 3.13. Illustration av den extremalsökande regulatorns funktion.

diskt. Regulatorns funktion illustreras i fig. 3.13. Om insignalens värde är mindre än det optimala får processens utsignal en periodisk komponent som är i fas med insignalen. En multiplikation med insignalen ger då en positiv signal som efter lågpassfiltrering driver integratorn, vilken ökar insignalen. Om insignalen ligger över det optimala värdet blir utsignalen i motfas mot insignalen. Produkten blir negativ och insignalen avtar. Regulatorn i fig. 3.12 föreslogs av Draper och Li 1951. Den provades praktiskt på optimering av tändning och bränsle/luftkvot för en flygplansmotor.

Extremalsökande regulatorer används också vid numeriskt styrda verktygsmaskiner. Det har under årens lopp gjorts många försök att bygga olika typer av extremalsökande regulatorer. En mycket intressant möjlighet är att bygga upp regulatorn kring en liten dator.

Dynamisk optimering

Statisk optimering kan användas för att välja arbetspunkter (börvärden) för reglering av processer där målsättningen är att hålla processvariablernas värden nära konstanta värden. Typexempel är reglering av kontinuerliga processer med få kvalitetsomställningar. Däremot passar ej formuleringen som ett statiskt optimeringsproblem om man önskar göra en regulator som åstadkommer effektiva kvalitetsomställningar. I ett sådant fall kan man t. ex. vara intresserad av en så snabb omställning som möjligt eller en omställning som minimerar energi eller råmaterialförluster. Detta leder till s. k. *dynamisk optimering*. Om systemet beskrivs av ekvationen (2.11) och prestanda kan anges med kriteriefunktionen (2.12) med bivillkor av typen (2.13) så kan man såsom nämnts i avsnitt 2 ange nödvändiga villkor för extremum med hjälp av variationskalkyl, Pontryagins maximumprincip eller dynamisk programmering. Med hjälp av olika iterativa numeriska beräkningar kan den optimala styrsignalen och motsvarande nominella tidsförlopp för tillståndsvariablerna, s. k. *trajektorier* eller *banor*, bestämmas. Observera dock att man på detta sätt erhåller en programstyrning, dvs. ett öppet system. Om man försöker att implementera ett sådant system så kommer de verkliga banorna på grund av störningar att avvika från de nominella banorna. För att korrigera för dessa störningar, som ej elimineras av den öppna styrningen, vore det önskvärt att införa återkoppling genom att bestämma de optimala regleringreppen i en omgivning till de nominella banorna. Detta kan göras genom att approximera reglerproblemet i en omgivning till den nominella banan. Genom att linjärisera den olinjära ekvationen (2.11) kring den nominella banan erhålles ett linjärt tidsvariabelt system

$$\frac{d}{dt}\delta x = A(t)\,\delta x + B(t)\,\delta u \tag{3.15}$$

som approximativt beskriver systemet i en omgivning till den nominella banan. En approximation av kriteriet (2.13) ger på samma sätt approximationen

$$J = J_0 + \delta x^T(T)\, Q_0\, \delta x(T) + \int_0^T [\delta x^T(t)\, Q_1(t)\, \delta x(t) + 2\, \delta x^T(t)\, Q_{12}(t)\, \delta u(t) + \delta u^T(t)\, Q_2(t)\, \delta u(t)]\, dt \qquad (3.16)$$

Matrisera Q_0, Q_1, Q_{12} och Q_2 definieras av

$$Q_0 = g_{xx}, \quad Q_1 = g_{xx}, \quad Q_{12} = g_{xu}, \quad Q_2 = g_{uu}$$

där g_{xx} betecknar matrisen av partiella derivator av andra ordningen av g och högerleden utvärderas längs de nominella banorna. Vidare betecknar δx och δu avvikelser från nominella tillstånd x^0 och nominella styrsignaler u^0. De linjära termerna i approximationen av kriteriet försvinner, eftersom den nominella banan var optimal.

Problemet att finna en styrsignal för systemet (3.15) som minimerar kriteriet (3.16) brukar kallas det *»linjär-kvadratiska problemet»*, eftersom systemet är linjärt och kriteriet kvadratiskt. Detta problem är mycket lättare att lösa än det ursprungliga olinjära optimeringsproblemet. Lösningen kan skrivas som

$$\delta u(t) = -L(t)\delta x(t)$$

dvs.

$$u(t) = u^0(t) - L(t)[x(t) - x^0(t)] \qquad (3.17)$$

där matrisen L kan bestämmas.

De korrektioner i styrsignalen som skall göras blir således proportionella mot tillståndsvariablernas avvikelser från de nominella värdena. Regulatorn (3.17) är alltså en proportionell regulator. Den proportionella förstärkningen är i regel tidsvariabel. Styrlagen (3.17) brukar kallas en *tillståndsåterkoppling*, eftersom styrsignalen ges som en återkoppling från alla tillståndsvariabler. Om alla tillståndsvariabler ej kan mätas direkt så kan de rekonstrueras. Detta skall behandlas mer detaljerat i det följande.

I fig. 3.14 visas ett blockschema för den regulator som erhålls vid tillämpning av dynamisk optimering. Dynamisk optimering har tillämpats vid styrning av rymdraketer och containerkranar. Problemformuleringen passar också väl för att styra satsvisa processer (batchprocesser) liksom för att åstadkomma effektiva omställningar av kontinuerliga processer.

Varierande processdynamik — parameterstyrning och adaptiva regulatorer

I många fall kan processdynamiken variera med processens driftsförhållanden. En reglerventil har t. ex. i allmänhet en olinjär karakteristik. Förstärkningen

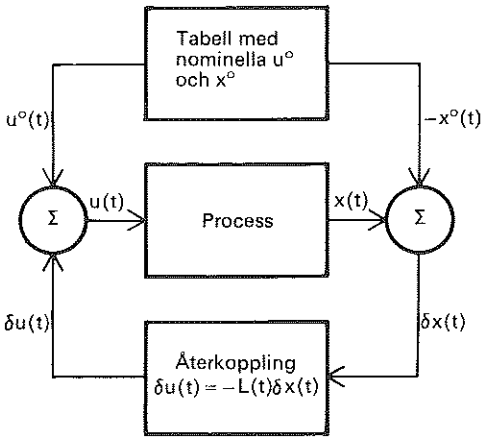


Fig. 3.14. Blockschema för regulator baserad på dynamisk optimering.

för en reglerventil kommer således att variera med ventilöppningen. Jämför exempel 4:3.2. Vid processreglering av kontinuerliga processer är tidsfördröjningar och tidskonstanter ofta omvänt proportionella mot flödet genom processen och därmed även mot produktionen. Dynamiken för ett värmekraftverk varierar med effektnivån. Dynamiken kan också ändras på grund av avlagringar i ångpannans tuber. Dynamiken för ett flygplan varierar kraftigt med flygplanets höjd och hastighet. Fartygs dynamik varierar med fart och lastförhållanden.

Om processdynamiken varierar måste speciella åtgärder tillgripas för att få en bra reglering. Följande åtgärder är tänkbara:

- lokal återkoppling
- olinjär kompensering
- programmering av regulatorns parametrar
- adaptiv reglering

I samband med känslighetsanalysen i avsnitt 4:3 visades att ett återkopplat system är relativt okänsligt för parametervariationer om kretsförstärkningen kan göras tillräckligt stor. I vissa fall kan verkan av variationerna i processdynamiken minskas genom att lägga en regulator runt den process vars dynamik varierar s. k. *lokal återkoppling*. Exempel 4:3.2 är ett typfall på användning av lokal återkoppling.

Variationerna i processdynamiken kan också kompenseras genom att regulatorn förses med olinjära element som kompenserar för variationerna. Olinjära ventilkarakteristika används t. ex. för att kompensera för olinjäriteter i processen. Detta kan göras om processmodellerna ej är alltför komplicerade. Metoden är allmänt förekommande vid reglering av temperatur och fukthalt i byggnader (komfortreglering) och i enkla processreglersystem.

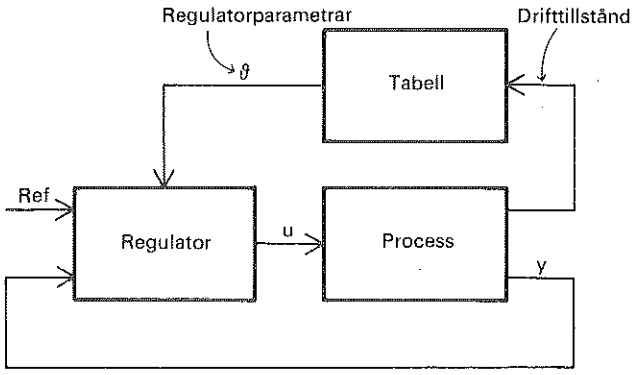


Fig. 3.15. Blockschema för regulator med parameterstyrning. Regulatorns parametrar ϑ bestäms som funktioner av drifttillståndet genom tabelluppslagning och interpolation.

I vissa fall kan förändringarna i processdynamiken vara entydigt bestämda av mätbara storheter. För många processer är t. ex. produktionsnivån en sådan storhet. För reglering av flygplan kan processdynamiken väl beskrivas som funktion av höjd och hastighet. I sådana fall kan man dimensionera regulatorn för vissa drifttillstånd och göra en tabell av motsvarande regulatorparametrar. De regulatorparametrar som används i en given driftsituation bestäms sedan genom interpolation i tabellen. I stället för att utnyttja en tabell och interpolation kan man ibland hitta funktionsuttryck som ger regulatorparametrarna som funktion av drifttillståndet. Blockschema för en regulator av denna typ framgår av fig. 3.15. Drifttillståndet kan uppenbarligen också betraktas som en utsignal. Orsaken till uppdelningen är att drifttillstånden antas variera avsevärt långsammare än utsignalerna. Detta sätt att lösa problemet kallas på engelska *gainscheduling*, vilket skulle kunna översättas med *parameterstyrning*. Regulatorn kan också tolkas som en variant av olinjär kompensering. Dimensioneringen kan vara arbetskrävande. Det är nödvändigt att först kartlägga hur processdynamiken varierar med driftsförhållandena. Regulatorparametrarna måste sedan bestämmas för ett stort antal driftfall. Man måste sedan undersöka att systemet blir stabilt inom hela driftområdet. Då det slutna systemet är olinjärt är detta mycket besvärligt, eftersom man endast i begränsad omfattning kan använda teori. I praktiken övertygar man sig om att systemet fungerar genom omfattande simuleringar. Det är också ett problem att bestämma hur många driftfall som skall finnas i tabellen och hur interpolationen skall utföras.

Det finns flera system där parameterstyrning används. I datorstyrda processreglersystem kan parametrarna i PID regulatorer göras beroende av produktionen. Autopiloter för överljudsflygplan är paradexemplet på parameterstyrning. I detta fall görs regulatorparametrarna beroende av höjd och hastighet. Det finns autopiloter med tabeller för ett 30-tal flygfall.

Det finns fall då det ej är möjligt att på förhand kartlägga hur process-

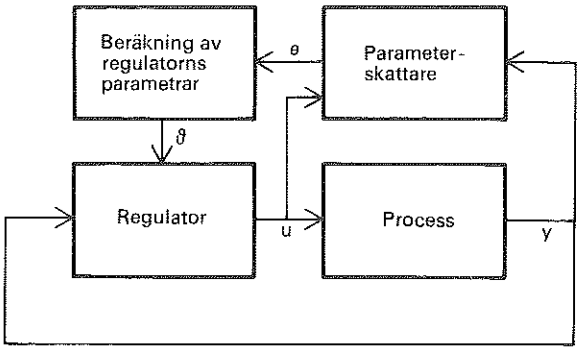


Fig. 3.16. Blockschema för adaptiv regulator baserad på skattning av processens parametrar.

dynamiken beror av drifttillståndet. I en sådan situation kan *adaptiv reglering* tillgripas. Ett adaptivt system visas i fig. 3.16. I detta system skattas processparametrarna med hjälp av en parameterskattare. Parameterskattaren bestämmer processens parametrar θ ur observationer av processens styr- och mätsignaler. Med utgångspunkt från de skattade parametrarna kan sedan regulatorparametrarna bestämmas genom tabelluppslagning eller genom en funktionsberäkning. Eftersom parameterskattaren måste ha tillgång till processens in- och utsignaler över ett visst intervall, erhålls processparametrarna med viss tidsfördröjning. Det system som visas i fig. 3.16 kan därför ha svårigheter om processens parametrar varierar mycket snabbt. Med parameterstyrning finns det däremot möjlighet att följa snabbare parametervariationer. Parameterstyrning kan därför kombineras med adaption på samma sätt som framkoppling och återkoppling kombinerades för vanliga reglersystem.

Teorin för adaptiva system är ett aktuellt forskningsområde inom reglerteknikен. Det finns i litteraturen förslag på många olika typer av regulatorer. De har också provats i autopiloter, för motorstyrning och för industriell processreglering.

4. Syntes av framkoppling och inversa system

Det visades i föregående avsnitt att framkopplingsproblemet för linjära system reduceras till följande: Givet ett linjärt system med överföringsfunktionen G, bestäm ett linjärt system vars överföringsfunktion H har egenskapen

$$G(s)H(s) = I \tag{4.1}$$

Jämför med ekvationerna (3.8) and (3.10) och blockschemat i fig. 4.1. Matematiskt innebär (4.1) att (höger)inversen till G skall bestämmas. Syntes av framkoppling är således ekvivalent med att bestämma en (höger)invers till ett dynamiskt system.

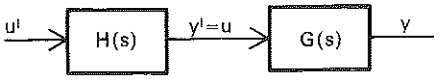


Fig. 4.1. Kaskadkoppling av ett system och dess högerinvers.

Inverser av enkla linjära system

Betrakta ett linjärt tidsinvariant system med en insignal och en utsignal. Låt G vara systemets överföringsfunktion. Det följer av ekvation (4.1) att det inversa systemet har överföringsfunktionen

$$H(s) = \frac{1}{G(s)} \tag{4.2}$$

Vi skall nedan ge några exempel.

EXEMPEL 4.1. Första ordningens system
Antag att

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

då blir inversen

$$H(s) = s+a$$

Observera att inversen innehåller en direktterm och en ren derivation.

EXEMPEL 4.2. Rationell funktion
Låt

$$G(s) = \frac{B(s)}{A(s)}$$

där $A(s)$ och $B(s)$ är polynom. Inversen blir då

$$H(s) = \frac{A(s)}{B(s)}$$

Det inversa systemets ordningstal är således lika med täljarens gradtal. Det inversa systemets poler (nollställen) är lika med det ursprungliga systemets nollställen (poler).

Vi kan i detta exempel lägga märke till att om överföringsfunktionen G har ett nollställe i högra halvplanet så blir det inversa systemet instabilt. Jämför avsnitt 2:9. Detta belyser att icke minimum-fas-system är svåra att reglera.

EXEMPEL 4.3. Tidsfördröjning

Antag att systemet har en ren tidsfördröjning. Jämför exempel 2:8.3. Överföringsfunktionen är då

$$G(s) = e^{-sT}$$

och det inversa systemet för överföringsfunktionen

$$H(s) = e^{sT}$$

Observera att denna överföringsfunktion ej är fysikaliskt realiserbar då den svarar mot ett system med insignal–utsignal-relationen

$$y(t) = u(t+T)$$

Utsignalens värde vid tiden t är alltså lika med insignalens värde T tidsenheter i framtiden. Ett sådant system kallas icke-kausalt och är ej fysikaliskt realiserbart.

Vi finner således att det är lätt att beräkna det inversa systemets överföringsfunktion men vi observerar i exempel 4.3 en potentiell svårighet då i detta fall det inversa systemets överföringsfunktion ej är realiserbar. Vi skall senare återvända till denna svårighet.

Inverser av flervariabla system

För flervariabla linjära tidsinvarianta system med samma antal in- och utsignaler så är G en kvadratisk matris, vars element är rationella funktioner. Jämför avsnitt 2: 8. Det inversa systemets överföringsfunktion H kan då bestämmas entydigt om $\det G(s)$ är en rationell funktion som ej är identiskt noll. Vi illustrerar med några exempel.

EXEMPEL 4.4. Rosenbrocks system

Betrakta ett system med överföringsfunktionen

$$G(s) = \begin{bmatrix} \dfrac{1}{s+1} & \dfrac{2}{s+3} \\ \dfrac{1}{s+1} & \dfrac{1}{s+1} \end{bmatrix}$$

Vi finner att

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

Kramers regel ger följande uttryck på det inversa systemets överföringsfunktion

$$H(s)=\frac{1}{1-s}\begin{bmatrix}(s+1)(s+3) & -2(s+1)^2\\ -(s+1)(s+3) & (s+1)(s+3)\end{bmatrix}$$

Rosenbrocks system är svårt att reglera trots att överföringsfunktionen ser mycket enkel ut. En indikation på detta är att det inversa systemet är instabilt, vilket bl. a. medför att kompensering med framkoppling leder till svårigheter. Begreppet minimum-fas definierades i avsnitt 2:9 endast för system med en insignal och en utsignal. En alternativ definition är att ett system är minimum-fas om systemet och dess invers är stabila. Utgående från denna definition är systemet i exempel 4.4 icke-minimum-fas. Det finns också system som i motsats till Rosenbrocks system ser ut att vara svåra att reglera men som i själva verket är lätta att reglera. Ett exempel ges nedan:

EXEMPEL 4.5.

Betrakta ett system med överföringsfunktionen

$$G(s)=\frac{1}{(s+1)^2}\begin{bmatrix}s-1 & 5s+1\\ -1 & s-1\end{bmatrix}$$

Vi finner att det inversa systemet har överföringsfunktionen

$$H(s)=\frac{s+1}{s+2}\begin{bmatrix}s-1 & -5s-1\\ 1 & s-1\end{bmatrix}$$

Observera att det inversa systemet är stabilt trots att diagonalelementen i G representerar icke-minimum-fas system.

Det finns system som ej kan inverteras. För överföringsfunktionen

$$G(s)=\begin{bmatrix}\dfrac{1}{s+1} & \dfrac{2}{s+1}\\[2ex] \dfrac{2}{s+1} & \dfrac{4}{s+1}\end{bmatrix}$$

gäller att det $G(s)=0$ och systemet kan således ej inverteras.

Vi kan alltså återigen konstatera att det är lätt att beräkna inverser för flervariabla system som har samma antal in- och utsignaler. Om antalet utsignaler är större än antalet insignaler finns ingen (höger)invers och om antalet utsignaler är mindre än antalet insignaler så finns eventuellt flera (höger)inverser.

EXEMPEL 4.6.

Betrakta ett system med överföringsfunktionen

$$G(s)=\begin{bmatrix}\dfrac{s-1}{s+2} & \dfrac{s-2}{s+2}\end{bmatrix}$$

Systemet har bl. a. följande inverser

$$H_1(s)=\begin{bmatrix}\dfrac{s+2}{s-1}\\ 0\end{bmatrix},\quad H_2(s)=\begin{bmatrix}0\\ \dfrac{s+2}{s-2}\end{bmatrix},\quad H_3(s)=\begin{bmatrix}4\\ -3\end{bmatrix}$$

Observera att inverserna med överföringsfunktionerna H_1 och H_2 båda är instabila medan överföringsfunktionen H_3 är stabil.

Om man försöker att reglera systemet i exempel 4.6 genom att endast utnyttja en av styrvariablerna uppträder svårigheter då överföringsfunktionen har ett nollställe i högra halvplanet. Svårigheten kan undvikas genom att använda båda styrvariablerna. Exemplet visar att det kan vara en fördel att ha många styrvariabler.

Inverser av system givna på tillståndsform

I kapitel 2 fann vi att dynamiska system kunde beskrivas på två principiellt olika sätt, med externa modeller (insignal–utsignal beskrivningar) eller med interna modeller (tillståndsekvationer). Vi skall nu visa hur man beräknar inversen av ett system som är givet på tillståndsform. Vi börjar med att behandla två exempel.

Exempel 4.7.
Betrakta följande observerbara och styrbara system

$$\frac{dx}{dt}=Ax+Bu \tag{4.3}$$

$$y=Cx+Du$$

Anta att antalet insignaler och utsignaler är lika och att matrisen D är inverterbar. Låt u^I beteckna det inversa systemets insignal och y^I dess utsignal. Se fig. 4.1. Problemet att bestämma det inversa systemet kan formuleras på följande sätt. Givet en signal v. Bestäm ett system så att om dess insignal väljs som $u^I=v$, så skall utsignalen y^I vara sådan att då den väljs som insignal till processen, dvs. $u=y^I$, blir processens utsignal lika med den givna signalen, dvs. $y=v$. Det följer av ekvation (4.3) att

$$y^I=u=D^{-1}(-Cx+v)$$

Insättning i ekvation (4.3) ger

$$\frac{dx}{dt}=[A-BD^{-1}C]x+BD^{-1}v$$

Det inversa systemet till (4.3) ges således av

$$\frac{dz}{dt} = [A - BD^{-1}C]z + BD^{-1}u^I \tag{4.4}$$

$$y^I = -D^{-1}Cz + D^{-1}u^I$$

där inversens tillstånd kallats z. Observera att resonemanget som gav ekv. (4.4) grundades på antagandet att inversens (4.4) tillstånd z är lika med systemets (4.3) tillstånd x. Ekv. (4.4) är således en invers till (4.3) endast om $z = x$. För att se vad som händer om $z \neq x$ antas att inversen och systemet kaskadkopplas såsom visas i fig 4.1. En enkel räkning ger

$$\frac{d}{dt}(z - x) = A(z - x)$$

Om matrisen A har alla egenvärden i vänstra halvplanet så kommer således differensen mellan inversens och processens tillstånd att gå mot noll exponentiellt.

EXEMPEL 4.8.

Betrakta systemet

$$\frac{dx}{dt} = Ax + Bu \tag{4.5}$$

$$y = Cx$$

Anta att antalet insignaler och utsignaler är lika samt att matrisen CB är inverterbar. För att bestämma inversen föreskrivs processens utsignal y. Derivation med avseende på tiden ger

$$\frac{dy}{dt} = C\frac{dx}{dt} = CAx + CBu$$

Eftersom matrisen CB är inverterbar ges processens insignal u och inversa systemets utsignal av

$$y^I = u = (CB)^{-1}\left[-CAx + \frac{dy}{dt}\right]$$

Insättning i (4.5) ger

$$\frac{dx}{dt} = [I - B(CB)^{-1}C]Ax + B(CB)^{-1}\frac{dy}{dt}$$

$$y^I = u = -(CB)^{-1}CAx + (CB)^{-1}\frac{dy}{dt}$$

vilket är en representation av det inversa systemet. Införs tillståndsvariabeln

$$z = x - B(CB)^{-1}y$$

så kan det inversa systemet skrivas som

$$\frac{dz}{dt} = [I - B(CB)^{-1}C]\,A[z + B(CB)^{-1}u^I]$$

$$y^I = -(CB)^{-1}CAz - (CB)^{-1}CAB(CB)^{-1}u^I + (CB)^{-1}\frac{du^I}{dt} \tag{4.6}$$

Observera att i exempel 4.7 så har det inversa systemet samma ordningstal som det ursprungliga systemet. Det följer t. ex. av exempel 4.2 att det inversa systemet mycket väl kan ha annat ordningstal än det ursprungliga systemet. Inversen (4.6) är av ordningen n. Det följer dock av (4.6) att

$$\frac{d}{dt}(Cz) = 0$$

oavsett hur insignalen u^I väljs. Detta innebär att systemet (4.6) har p egenvärden i origo som motsvarar icke styrbara moder. Talet p betecknar antalet utsignaler för systemet (4.3). Det finns således ett system av ordningen $n-p$ som har samma insignal–utsignal relation som (4.6).

Den metod som användes för att härleda inversen i exempel 4.8 kan lätt utvidgas till det fall då CB är singulär. Principen är den att varje komponent av utsignalen deriveras. De derivator av tillståndsvektorn som erhålls substitueras sedan med hjälp av systemekvationerna. Man deriverar ända tills styrvariablerna uppträder i högerledet. Därefter löses styrvariablerna ut och uttrycket på dessa substitueras tillbaka i tillståndsekvationerna. Villkoret för inverterbarhet är att ekvationerna kan lösas. Detta är i princip lätt att genomföra. Resultatet blir att inversen för ett system med lika många in- och utsignaler får formen

$$\frac{dz}{dt} = A_1 z + B_1 u^I$$

$$y^I = C_1 z + D_0 u^I + D_1\frac{du^I}{dt} + D_2\frac{d^2u^I}{dt^2} + \ldots + D_k\frac{d^ku^I}{dt^k} \tag{4.7}$$

Uttrycken som ger dimensionen av z samt matriserna, A_1, B_1, C_1, D_0, ..., D_k som funktioner av A, B, C och D är komplicerade. Det går dock bra att ge algoritmer som genererar dessa storheter. I specifika fall kan den procedur som givits ovan också tillämpas.

Inverser av olinjära system

Inversa system definierades med hjälp av ekvation (4.1) för linjära system. Denna definition kan uppenbarligen ej användas för olinjära system då dessa

ej har någon överföringsfunktion. Däremot kan resonemanget som användes i exempel 4.7 genomföras även för olinjära system. Vi illustrerar detta med ett exempel.

EXEMPEL 4.9. Pendeln

Betrakta den olinjära modellen för pendeln i exempel 2:2.2, dvs.

$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_2}{dt} = -\sin x_1 + u \cos x_1$$

$$y = x_1$$

Om vi följer det recept som angavs för linjära system skall utsignalen deriveras. Efter två derivationer erhålls

$$\frac{d^2y}{dt^2} = -\sin x_1 + u \cos x_1$$

Lösning av u ger

$$u = \frac{1}{\cos x_1}\left[\sin x_1 + \frac{d^2y}{dt^2}\right] = \tan y + \frac{1}{\cos y}\frac{d^2y}{dt^2}$$

vilket är det inversa systemets insignal–utsignal-relation.

Lägg märke till att inversen ej existerar för $\cos x_1 = 0$. Om $\cos x_1 = 0$ så har insignalen ej heller någon inverkan på systemet.

Förkortning av poler och nollställen

Vi skall nu undersöka vad som händer då ett system kaskadkopplas med sin invers. Jämför blockschemat i fig. 4.1. Om systemet har en invers gäller enligt ekvation (4.1) att det sammankopplade systemet har enhetsmatrisen som överföringsfunktion. Vid ett ytligt betraktande skulle man då kunna förledas att tro att insignal–utsignal-relationen

$$y(t) = u^I(t) \tag{4.8}$$

skulle ge en uttömmande beskrivning av det kaskadkopplade systemet. Detta är ej sant. Ty om det ursprungliga systemet är ett dynamiskt system så måste det sammansatta systemet ha ett ordningstal skilt från noll. Insignal–utsignal-relationen (4.8) är ett nollte ordningens system och det följer då av Kalmans uppdelningssats (sats 3:6.1) att det sammansatta systemet ej är observerbart och styrbart. Vi skall närmare undersöka detta. Speciellt skall vi visa att de icke observerbara och icke styrbara delsystemen under vissa omständigheter kan ge upphov till svårigheter. Vi börjar med att analysera ett enkelt exempel.

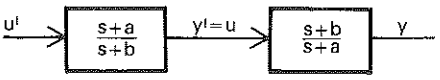


Fig. 4.2. Kaskadkoppling av system med överföringsfunktionerna $(s+b)/(s+a)$ och $(s+a)/(s+b)$.

EXEMPEL 4.10.

Systemet med överföringsfunktionen

$$G(s)=\frac{s+b}{s+a} \tag{4.9}$$

har inversen

$$H(s)=\frac{s+a}{s+b} \tag{4.10}$$

Om systemet och dess invers kaskadkopplas på det sätt som visas i fig. 4.2 erhålls ett system med överföringsfunktionen

$$G(s)\,H(s)=\frac{s+b}{s+a}\,\frac{s+a}{s+b}=1$$

Eftersom G och H är rationella funktioner är det uppenbarligen matematiskt helt korrekt att förkorta faktorerna $s+a$ och $s+b$. Då både systemet och dess invers är system av första ordningen så är det sammankopplade systemet uppenbarligen ett andra ordningens system. Eftersom överföringsfunktionen GH motsvarar ett nollte ordningens system, så följer av Kalmans uppdelningssats att det sammansatta systemet har icke styrbara och tysta delsystem. Vi skall nu analysera vad som händer. Överföringsfunktionen (4.9), som beskriver systemet, har realisationen

$$\frac{dx_1}{dt}=-ax_1+(b-a)\,u$$

$$y=x_1+u$$

och överföringsfunktionen (4.10) har realisationen

$$\frac{dx_2}{dt}=-bx_2+(a-b)\,u^I$$

$$y^I=x_2+u^I$$

Det kaskadkopplade systemet i fig. 4.1 har då följande realisation

$$\frac{d}{dt}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}-a & b-a\\0 & -b\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}+\begin{bmatrix}b-a\\a-b\end{bmatrix}u^I$$

$$y=x_1+x_2+u^I$$

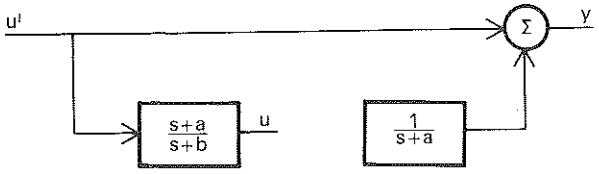


Fig. 4.3. Kalmans uppdelning av systemet i fig. 4.2.

vilket erhålls genom att insätta $u = y^I$ i ovanstående realisationer. Inför nya tillståndsvariabler genom transformationen

$$z_1 = x_1 + x_2$$

$$z_2 = x_2$$

Det transformerade systemet ges av

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ a-b \end{bmatrix} u^I$$

$$y = z_1 + u^I$$

Systemet är på diagonalform. Om $a \neq b$ kan sats 3:4.5 tillämpas. Det följer då att moden z_1 är observerbar men ej styrbar och att moden z_2 är styrbar men ej observerbar. Att moden $z_1 = x_1 + x_2$ ej är styrbar innebär att inversens och processens tillstånd alltid har summan noll om $z_1(0) = 0$. Ett blockschema som illustrerar Kalmans uppdelning av systemet visas i fig. 4.3. Det är nu klart hur kombination av systemet och dess invers fungerar. Det sammansatta systemet är av andra ordningen. Systemet kan uppdelas i tre delsystem, ett statiskt system med överföringsfunktionen $G(s) = 1$, ett styrbart men tyst delsystem med egenvärdet $\lambda = -b$ och ett observerbart men ej styrbart delsystem med egenvärdet $\lambda = -a$. Om den observerbara moden z_1 exciteras så kommer variationer i x direkt att märkas i utsignalen. Däremot kan moden ej exciteras av insignalen. Den styrbara moden $s = -b$ kan exciteras av insignalen. Eftersom moden ej är observerbar har den inget inflytande på utsignalen. Däremot påverkas processens insignal u.

Observera att z_1 är summan av processens tillstånd x_1 och inversa systemets tillstånd x_2. Det finns ingen garanti för att initialvärdet av z_1 är noll. Detta betyder att utsignalen y kommer att avvika från insignalen u^I. Om processen är stabil, dvs. a positiv, så kommer z_1 dock att konvergera mot noll. Processens utsignal y kommer då också att närma sig u^I. Jämför fig. 4.4.A.

Det är också illustrativt att se vad som händer om processen är instabil, dvs. a negativ. I detta fall kommer tillståndsvariabeln z_1 att växa utan gräns om den exciteras. Eftersom z_1 ej är styrbar, så kan den ej påverkas av insignalen u^I. Processens utsignal y kommer således att växa exponentiellt. Jämför fig. 4.4.B.

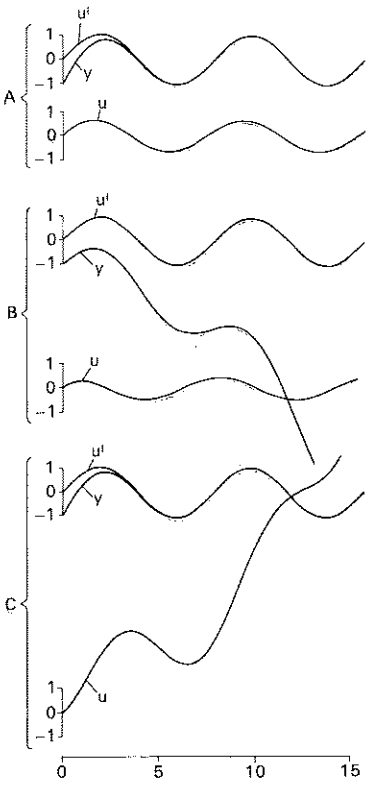


Fig. 4.4. Simulering av systemet i fig. 4.2. I fall A är processen och inversen stabila ($a = 1, b = 2$). Efter en transient kommer y och u^I att överensstämma. I fall B ($a = -0{,}2$, $b = 2$) är processen instabil. Processens utsignal y kommer då att växa utan gräns. I fall C ($a = 1, b = -0{,}2$) är processen stabil men inversen instabil. Processens insignal u växer då som $\exp - bt$. Nollstället $s = -b$ blockerar emellertid denna komponent av insignalen och y överensstämmer med u^I efter en transient.

Om processen ej är ett minimum-fas-system, dvs. om nollstället $s = -b$ ligger i högra halvplanet, så är det inversa systemet instabilt. Vid en excitation, t. ex. förändring i insignalen u^I, kommer då z_2 att växa exponentiellt, vilket medför att processens insignal u också kommer att växa exponentiellt. Detta kommer dock ej att märkas i processens utsignal y då moden z_2 ej är observerbar. Jämför fig. 4.4.C.

Sammanfattningsvis finner vi således att det ej går bra att kaskadkoppla systemet (4.9) och dess invers (4.10) om något av talen a aller b är negativt, dvs. om systemet eller dess invers är instabila.

Det är lätt att generalisera analysen i exemplet till system med en godtycklig rationell överföringsfunktion. Betrakta en process med överföringsfunktionen

$$G(s) = \frac{B(s)}{A(s)} \tag{4.11}$$

där $A(s)$ och $B(s)$ är polynom. Detta system har inversen

$$H(s) = \frac{A(s)}{B(s)} \tag{4.12}$$

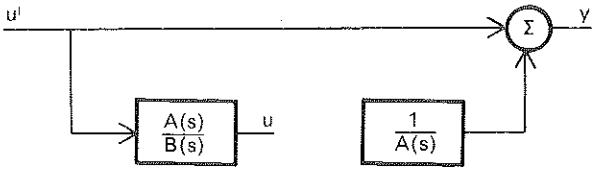


Fig. 4.5. Kalmans uppdelning av det system som erhålls då systemet med överföringsfunktionen $B(s)/A(s)$ kaskadkopplas med sin invers.

I fig. 4.5 visas Kalmans uppdelning av det system som erhålls om systemet och dess invers kaskadkopplas. Vi finner således att det kaskadkopplade systemet har icke styrbara moder som svarar mot polerna till $G(s)$ och tysta moder svarande mot nollställena till $G(s)$. En kaskadkoppling av systemet (4.11) och dess invers (4.12) kan ej förverkligas om systemet (4.11) har poler eller nollställen i högra halvplanet. Om (4.11) har en pol i högra halvplanet så kommer processens utsignal att växa exponentiellt därest den exciteras. Om (4.11) har ett nollställe i högra halvplanet så blir inversen instabil och processens insignal växer i stället exponentiellt. Då denna insignal blockeras av ett motsvarande nollställe i processen, kommer processens utsignal dock att vara begränsad.

Det skall visas i avsnitt 5 att ett styrbart system alltid kan stabiliseras med återkoppling. En instabil process kan då först stabiliseras med återkoppling och det stabiliserade systemet kan sedan framkopplas. Om framkoppling och återkoppling kombineras på detta sätt finner vi att nollställen i högra halvplanet utgör den väsentliga begränsningen.

Exemplet visar att man bör vara försiktig med förkortning av faktorer vid formell manipulation med överföringsfunktioner. Vi belyser detta med ytterligare ett exempel:

EXEMPEL 4.11. Otto Smith-regulatorn

Vid bestämning av överföringsfunktionen (3.14) för ett system med en Otto Smith-regulator i avsnitt 3 förkortades faktorn $G_P(s)\,e^{-sT}$. En noggrannare analys visar att det slutna systemet har den Kalmanuppdelning som visas i fig. 4.6. Det följer av figuren att kompenseringen ej kommer att fungera om processen är instabil då de observerbara men icke styrbara moderna ger bidrag till utsignalen som växer utan gräns. En process som är instabil måste därför först stabiliseras innan dödtidskompensering kan tillämpas.

Analysen visar också fördelen med att kunna behandla system såväl från externa (insignal–utsignal modeller, Laplacetransformer) som interna (tillståndsmodeller) synpunkter. Det är uppenbarligen lätt att beräkna inversen av ett system vars överföringsfunktion är given. Ett formellt manipulerande med överföringsfunktionen kan dock leda till felaktiga slutsatser om man ej kommer ihåg att överföringsfunktionen endast representerar den del av

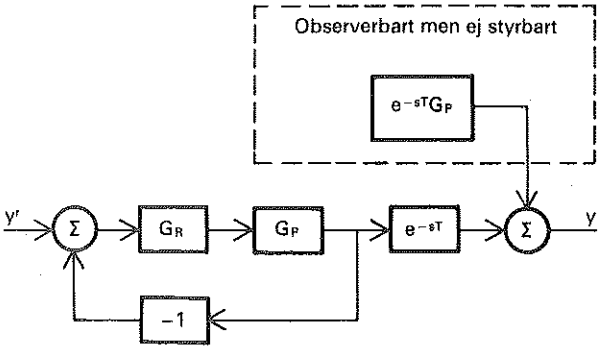


Fig. 4.6. Kalmans uppdelning av systemet i fig. 3.10.

ett linjärt system som är observerbar och styrbar. En god regel vid formellt räknande med Laplacetransformer är att man bör tänka sig för vad som händer då poler och nollställen förkortas.

Approximativa inverser

Vi har givit flera exempel på system vars inverser ej kan förverkligas. I exempel 4.1 innehöll inversen en ren derivation. Systemet i exempel 4.3 innehöll en ren tidsfördröjning, vilket medförde att inversen ej var ett kausalt system. Vi har också sett att inverser till system som ej är minimum-fas ej kan förverkligas. Vid syntes av framkoppling för sådana system får approximativa inverser tillgripas. Några exempel på approximationer ges nedan.

Vid invertering av system med rationella överföringsfunktioner där nämnaren har högre gradtal än täljaren erhålls en invers som innehåller derivatorer. En derivation kan ej förverkligas exakt, ty man skulle då få ett system som har mycket hög förstärkning vid höga frekvenser. Man förverkligar därför en approximativ invers. Detta kan t. ex. göras med approximationen

$$s \approx \frac{s}{1+sT} \tag{4.13}$$

Detta ger en approximation som är någorlunda bra för signaler vars frekvens ω_0 är sådan att $\omega_0 T \ll 1$. Ett sätt att förverkliga denna överföringsfunktion visas i fig. 4.7. För detta system gäller att $T = 1 - a$. Det är således möjligt att få mycket små värden på tidskonstanten T. I praktiken är T begränsad underifrån av störningarna i den signal som skall deriveras. Det bör också observeras att i många fall används framkoppling i kombination med en referensmodell såsom framgår av fig. 3.5. I ett sådant fall kan derivation av signaler ofta undvikas genom att kombinera systeminversen med referensmodellen.

Det har visats att ett system som ej är minimum-fas har en instabil invers. Svårigheterna med att förverkliga en sådan invers har också diskuterats. I ett

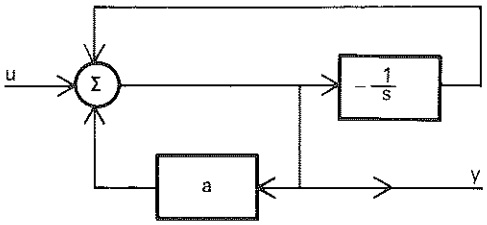


Fig. 4.7. Blockschema för förverkligande av en approximativ derivation. Systemet har överföringsfunktionen

$G(s) = s/[1 + (1-a)s]$.

Observera att det krävs att $a < 1$ för stabilitet.

sådant fall är det naturligt att undersöka om det är möjligt att finna en approximativ invers som är stabil. Ett svar ges av nedanstående resultat.

SATS 4.1.

Antag att en rationell överföringsfunktion G har alla poler i vänstra halvplanet och inga nollställen på imaginära axeln. Låt

$$G(s) = G^+(s)\,G^-(s) \tag{4.14}$$

vara en faktorering av G sådan att $G^+(s)$ innehåller alla poler och har alla nollställen i vänstra halvplanet och $G^-(s)$ alla nollställen i högra halvplanet. En approximativ stabil invers som minimerar tidsintegralen av kvadraten på felet $e = u^I - y$ vid en insignal i form av ett steg ges av

$$H^+(s) = \frac{1}{G^+(s)\,G^-(-s)} \tag{4.15}$$

Vidare gäller

$$\min \int_0^\infty e^2(t)\,dt = -2a^2 \left[\frac{d}{ds} \log G^-(s)\right]_{s=0} \tag{4.16}$$

där a är insignalens amplitud.

Bevis:

Låt e vara felet om insignalen är ett enhetssteg. Laplacetransformen av e ges av

$$E(s) = [1 - G(s)\,H(s)]\frac{1}{s}$$

där $H(s)$ är den approximativa inversens överföringsfunktion. Enligt Parsevals sats gäller att tidsintegralen av felets kvadrat kan skrivas som

$$J[H] = \int_0^\infty e^2(t)\,dt = -\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} [1 - G(s)\,H(s)]\,[1 - G(-s)\,H(-s)]\,s^{-2}ds \tag{4.17}$$

För att integralen skall vara ändlig måste det krävas att

$G(0)H(0) = 1$

För att minimera J med avseende på stabila H utnyttjas den i satsen givna faktoreringen av G och integranden kan då skrivas om genom att utnyttja identiteten

$$[1-G(s)H(s)][1-G(-s)H(-s)]$$
$$= [1-G^+(s)G^-(-s)H(s)][1-G^+(-s)G^-(s)H(-s)$$
$$+[G^-(-s)-G^-(s)]G^+(s)H(s)+[G^-(s)-G^-(-s)]G^+(-s)H(-s)$$

Integralen (4.17) skall tolkas som

$$\int_{-i\infty}^{i\infty}\dots ds = \lim_{R\to\infty}\lim_{r\to 0}\left\{\int_{-iR}^{-ir}\dots ds+\int_{ir}^{iR}\dots ds\right\}$$

Det följer av residuesatsen att

$$\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}[G^-(-s)-G^-(s)]G^+(s)H(s)s^{-2}ds$$
$$+\frac{1}{2\pi i}\int_{\gamma^+}[G^-(-s)-G^-(s)]G^+(s)H(s)s^{-2}ds=0$$

där γ^+ är en infinitesimal halvcirkel till höger om origo, ty integranden har inga poler i högra halvplanet och integralen längs en halvcirkel med imaginära axeln som diameter försvinner. Vidare följer av residuesatsen att

$$\frac{1}{2\pi i}\int_{\gamma^+}[G^-(-s)-G^-(s)]G^+(s)H(s)s^{-2}ds$$
$$+\frac{1}{2\pi i}\int_{\gamma^-}[G^-(s)-G^-(-s)]G^+(-s)\mathrm{H}(-s)s^{-2}ds$$
$$=\frac{1}{2\pi i}\int_{\gamma^+\gamma^-}[G^-(s)-G^-(-s)]G^+(-s)H(-s)s^{-2}ds$$
$$=-dG^-/ds|_{s=0}G^+(0)H(0)=-dG^-/ds|_{s=0}/G^-(0)$$

Vi finner således

$$J[H]=-\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}[1-G^+(s)G^-(-s)H(s)][1-G^+(-s)G^-(s)H(-s)]s^{-2}ds$$
$$-2\,dG^-/ds|_{s=0}/G^-(0)\geqslant -2\,dG^-/ds|_{s=0}/G^-(0)$$

där likhet erhålls om överföringsfunktionen H väljs som (4.15). Resultatet erhålls nu om vi observerar att felet är proportionellt mot stegets storlek.

Vi illustrerar satsen med ett exempel.

EXEMPEL 4.12.

Låt överföringsfunktionen G vara

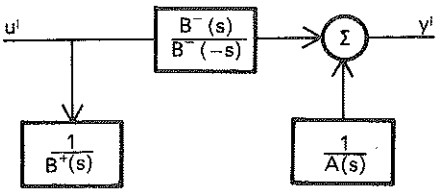


Fig. 4.8. Kalmans uppdelning av det system som erhålls då en process med överföringsfunktionen $B^+(s)B^-(s)/A(s)$ kaskadkopplas med den approximativa inversen $A(s)/[B^+(s)B^-(-s)]$.

$$G(s)=\frac{\prod(s+\beta_i)\prod(s-\gamma_i)}{\prod(s+\alpha_i)}=\frac{B^+(s)B^-(s)}{A(s)}$$

där α_i, β_i och γ_i alla är positiva. Den approximativa inversen ges då av

$$H^+(s)=\frac{\prod(s+\alpha_i)}{\prod(s+\beta_i)\prod(-s-\gamma_i)}=\frac{A(s)}{B^+(s)B^-(-s)}$$

Vidare är

$$\log G^-(s)=\Sigma\log(s-\gamma_i)$$

dvs.

$$\frac{d}{ds}\log G^-(s)=\sum\frac{1}{s-\gamma_i}$$

Om e är felet vid ett enhetssteg finner vi således

$$\min\int_0^\infty e^2(t)\,dt=2\sum 1/\gamma_i$$

där minimum tas med hänsyn till alla stabila inverser. Vi observerar att felet ökar med minskande γ_i. Det är således till nackdel att ha nollställen i högra halvplanet med liten realdel. I fig. 4.8 visas Kalmans uppdelning av det system som erhålls om en process med överföringsfunktionen G framkopplas med den approximativa inversen. Observera att det kompenserade systemet har överföringsfunktionen

$$G_k(s)=\frac{B^-(s)}{B^-(-s)}=\frac{\prod(s-\gamma_i)}{\prod(-s-\gamma_i)}$$

Moderna som svarar mot processens nollställen i vänstra halvplanet är tysta och moderna som svarar mot processens poler är icke styrbara. Jämför också med fig. 4.3. I fig. 2:9.8 visas stegsvaret i det fall då processen endast har en pol $s=1$ i högra halvplanet.

Den approximativa invers som ges av sats 4.1 beror uppenbarligen av den valda insignalen (ett steg). Det är möjligt att härleda andra approximativa inverser för andra typer av insignaler. Det är också möjligt att på liknande

sätt bestämma approximativa kausala inverser för system som innehåller tidsfördröjningar.

Sammanfattning

Det konstaterades inledningsvis i detta avsnitt att syntes av framkoppling är ekvivalent med att bestämma högerinverser till ett system. Vi har visat hur systeminverser beräknas för olika system. Svårigheter som uppträder då inverser innehåller rena derivationer, då de är instabila eller icke kausala har berörts. Möjligheter att beräkna approximativa inverser i sådana fall har också diskuterats. Det bör avslutningsvis betonas att den ur praktisk synpunkt kanske viktigaste frågan vid syntes av framkoppling består i att avgöra vilken modell av det fysikaliska systemet som skall ligga till grund för dimensioneringen. Sedan modellen väl fastlagts är ju beräkningen av inversen relativt enkel. De modeller som används kan variera avsevärt från enkla statiska modeller baserade på balansekvationer (massa, energi och rörelsemängd) till komplicerade modeller som väl beskriver processens uppförande vid snabba variationer i insignalen. För enkla servon används ibland en proportionell framkoppling från referensvärdet som bestämts enbart med ledning av den statiska förstärkningen. Vid industriell processreglering används ofta framkoppling som enbart baserats på statiska mass- och värmebalanser. Det är ofta till god vägledning att fråga sig inom vilka frekvensområden man önskar att framkopplingen skall vara verksam. Det är också bra att komma ihåg att framkoppling nästan alltid används i kombination med återkoppling.

Inversen av ett linjärt system är ett dynamiskt system eventuellt kombinerat med derivatorer. För att förverkliga en framkoppling skall således en apparat byggas som realiserar ett dynamiskt system. Det visades i exempel 4:3.3 att integratorer och summatorer lätt kan realiseras med hjälp av återkopplade operationsförstärkare. Med hjälp av operationsförstärkare och potentiometrar kan man således få komponenter som motsvarar de block som behandlades vid den grafiska beskrivningen av dynamiska system i avsnitt 2:10. De blockschema som visades i detta avsnitt kan alltså direkt tolkas som ritningar för apparater som realiserar dynamiska system. En rationell överföringsfunktion kan t. ex. konstrueras antingen på det sätt som anges av fig. 2:10.10 eller som anges av fig. 2:10.11. Vid digitala realiseringar kan man i stället programmera differensapproximationer av motsvarande differentialekvationer.

5. Syntes av system med given överföringsfunktion

I detta avsnitt skall syntes av en återkoppling som ger ett system med föreskriven överföringsfunktion diskuteras. Vi skall således utgå ifrån att specifikationerna på det slutna systemet givits genom att ange det slutna systemets

överföringsfunktion såsom diskuterats i avsnitt 2, t. ex. som ett system som minimerar ITAE. De resultat som erhålls är också viktiga ur en annan synpunkt. De visar klart vad som kan åstadkommas med hjälp av återkoppling och vad som begränsar de resultat som kan uppnås. Vi skall först visa att om alla tillståndsvariabler kan mätas så kan det slutna systemets poler placeras godtyckligt under förutsättning att systemet är styrbart. Däremot påverkas systemets nollställen ej av återkopplingen. Om alla tillståndsvariabler ej mäts så kan de beräknas ur processens in- och utsignaler. Vi skall ange några sätt att konkret utföra denna beräkning. Den regulator som erhålles kan då tolkas som bestående av två delsystem: dels ett system där tillståndsvariablerna beräknas, dels en linjär regulator som består av en tillståndsåterkoppling från de rekonstruerade tillståndsvariablerna. Den erhållna regulatorstrukturen härleds först utgående från interna modeller (tillståndsekvationer). En alternativ härledning utgående från externa modeller skall också anges. Detta leder till alternativa metoder för att bestämma regulatorn. Resultaten illustrerar också att det är fruktbart att se ett problem från två synpunkter. Möjligheterna att påverka det slutna systemets nollställen med återkoppling är begränsade. För att ändra nollställena måste därför framkoppling tillgripas. Detta behandlas kortfattat. Avsnittet avslutas med en diskussion av de faktorer som begränsar ett systems prestanda.

Polplacering genom tillståndsåterkoppling

Betrakta en process som kan beskrivas som ett enkelt linjärt tidsinvariant system. Anta att systemet är observerbart och styrbart samt att $D=0$, vilket innebär att utsignalen ej påverkas direkt av insignalen (»direktterm saknas»). Systemet kan då beskrivas med ekvationerna

$$\frac{dx}{dt} = Ax + Bu \tag{5.1}$$

$$y = Cx$$

Det följer av ekvation 3:8.2 att systemet har överföringsfunktionen

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + \ldots + a_n} = \frac{B(s)}{A(s)} \tag{5.2}$$

där n är systemets ordningstal.

Anta att specifikationerna kräver att överföringsfunktionen från referenssignal y^r till utsingal y skall vara

$$G_r(s) = \frac{q_1 s^{n-1} + q_2 s^{n-2} + \ldots + q_n}{s^n + p_1 s^{n-1} + \ldots + p_n} = \frac{Q(s)}{P(s)} \tag{5.3}$$

Specifikationerna kan uppenbarligen satisfieras med hjälp av framkoppling

genom att seriekoppla det givna systemet med ett system med överföringsfunktionen

$$G_k(s) = \frac{Q(s)\,A(s)}{P(s)\,B(s)}$$

Det följer emellertid av analysen i avsnitt 4 att en sådan lösning ej är brukbar om systemet (5.1) är instabilt, om det påverkas av störningar eller om det är icke-minimum-fas. Dessutom leder framkopplingen till ett system som är känsligt för parametervariationer. Vi skall därför undersöka vad som kan åstadkommas med hjälp av återkoppling.

För att genomföra analysen införs först ett koordinatsystem i tillståndsrummet som ger enkla räkningar. Eftersom systemet (5.1) är styrbart finns enligt sats 2:6.2 en transformation

$$z = Sx \tag{5.4}$$

så att det transformerade systemet erhålls på styrbar kanonisk form, dvs.

$$\frac{dz}{dt} = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u$$

$$y = [b_1\, b_2 \dots b_n]z \tag{5.5}$$

Inför nu en återkoppling definierad av följande styrlag

$$u = y^r - l_1 z_1 - l_2 z_2 - \dots - l_n z_n = y^r - \bar{L}z \tag{5.6}$$

där y^r är referenssignalen (det återkopplade systemets insignal). Den ekvation, som beskriver det återkopplade systemet, erhålls genom insättning av (5.6) i (5.5). Vi får

$$\frac{dz}{dt} = \begin{bmatrix} -l_1 - a_1 & -l_2 - a_2 \dots -l_{n-1} - a_{n-1} & -l_n - a_n \\ 1 & 0 \quad \dots \quad 0 & 0 \\ \vdots & & \\ 0 & 0 \quad \dots \quad 1 & 0 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} y^r$$

$$y = [b_1\, b_2 \dots b_n]z$$

Välj nu $l_i = p_i - a_i$, $i = 1, \dots, n$. Det återkopplade systemet får då följande överföringsfunktion

$$\frac{Y(s)}{U_r(s)} = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \dots + b_n}{s^n + p_1 s^{n-1} + \dots + p_n} = \frac{B(s)}{P(s)}$$

där koefficienterna i nämnarpolynomet $P(s)$ kan väljas fritt genom val av l_i.

Sammanfattningsvis finner vi således

SATS 5.1. Polplaceringssatsen

Betrakta ett linjärt tidsinvariant system (5.1) av n:te ordningen med en insignal. Anta att systemet är styrbart. Systemets poler kan placeras godtyckligt genom en linjär återkoppling från alla tillståndsvariabler.

$$u = y^r - Lx \tag{5.7}$$

Systemets nollställen påverkas ej av återkopplingen.

Det följer av ekvationerna (5.4) och (5.6) att den tillståndsåterkoppling som ger ett slutet system med den karakteristiska ekvationen $P(s)$ ges av

$$L = [p_1 - a_1 \; p_2 - a_2 \; \dots \; p_n - a_n]S \tag{5.8}$$

där matrisen S ges av ekvationen 2:6.23. Det finns datorprogram för beräkning av vektorn L. För system av låg ordning är det ofta lättare att direkt ansätta en återkoppling av formen (5.7) och sedan bestämma elementen i vektorn L så att det slutna systemet får föreskrivna poler.

Sats 5.1 kan generaliseras till system med flera insignaler. Observera också att det finns andra formuleringar av syntesproblemet som också leder till en tillståndsåterkoppling av formen (5.7). Jämför t. ex. med ekvation (3.17) som beskriver den styrlag som erhålls genom att tillämpa s. k. linjärkvadratisk teori.

Observera att styrlagen (5.7) ej innehåller dynamiska element. Den kan tolkas som en proportionell återkoppling från alla tillståndsvariabler och kallas därför tillståndsåterkoppling. Lägg också märke till att sats 5.1 nästan är självklar om vi observerar att ett styrbart system kan representeras med blockschemat i fig. 2:10.11 eller signalflödesdiagrammet i fig. 2:10.16.

Sats 5.1 innebär således att ett system med överföringsfunktionen (5.3) kan erhållas genom återkoppling av systemet (5.1) om polynomen $P(s)$ och $A(s)$ har samma gradtal, om polynomen $Q(s)$ och $B(s)$ är lika och alla tillståndsvariabler mätbara. Vi illustrerar resultatet med ett exempel:

EXEMPEL 5.1. Servomotor med takometeråterkoppling

En servomotor med drivförstärkare kan beskrivas med följande överföringsfunktion

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b}{s(s+a)}$$

där insignalen u är drivförstärkarens insignal och utsignalen y är motorns vridningsvinkel. Hur skall systemet återkopplas för att det slutna systemet skall få överföringsfunktionen

$$G_r(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2} \tag{5.9}$$

Vi observerar att täljarna till överföringsfunktionerna $G(s)$ och $G_r(s)$ endast skiljer sig på en konstant faktor. Denna kan justeras med hjälp av en förstärkare med lämpligt vald förstärkning. Vi kan således utan att förlora i allmängiltighet direkt anta att $b=\omega^2$. För att undersöka hur systemet skall återkopplas för att ge den önskade överföringsfunktionen, skriver vi systemet på följande kanoniska form genom att införa tillståndsvariablerna x_1 och x_2 definierade av

$$X_1(s)=\frac{\omega^2}{s+a}\,U(s)$$

$$X_2(s)=\frac{1}{s}\,X_1(s)$$

Systemet kan då beskrivas med ekvationerna

$$\frac{dx}{dt}=\begin{bmatrix}-a & 0\\ 1 & 0\end{bmatrix}x+\begin{bmatrix}\omega^2\\ 0\end{bmatrix}u$$

$$y=[0 \quad 1]\,x$$

Vi finner att styrlagen

$$u=y^r+(a\omega^{-2}-2\zeta\omega^{-1})x_1-x_2$$

ger det önskade resultatet. Ty med denna återkoppling erhålls följande ekvation för det återkopplade systemet

$$\frac{dx}{dt}=\begin{bmatrix}-2\zeta\omega & -\omega^2\\ 1 & 0\end{bmatrix}x+\begin{bmatrix}\omega^2\\ 0\end{bmatrix}y^r$$

$$y=[0 \quad 1]\,x$$

Överföringsfunktionen från insignal y^r till utsignal y ges av (5.9).

Tillståndsvariablen x_1 som motsvarar motoraxelns vinkelhastighet kan mätas med en takometer och tillståndsvariabeln x_2, som motsvarar axelns vridningsvinkel, kan mätas med en potentiometer.

Observera att den kanoniska form som användes i exempel 5.1 skiljer sig från den styrbara kanoniska formen. I ekvation (5.5) är koefficienten framför u ett, medan motsvarande koefficient är $b=\omega^2$ i exempel 5.1. Detta medför att tillståndsvariablerna x_1 och x_2 motsvarar motoraxelns vinkelhastighet och vinkelläge oavsett hur drivförstärkaren ställts in. I det allmänna fallet svarar detta mot att styrlagen (5.6) parametriseras som

$$u=b[y^r-\lambda_1 x_1-\ldots-\lambda_{n-1}x_{n-1}-x_n]$$

Detta har fördelen att reglerfelet y^r-x_n förekommer explicit i styrlagen.

6:5

Rekonstruktion med hjälp av derivering

Sats 5.1 förutsätter att alla tillståndsvariabler kan mätas direkt. Vi skall nu förutsätta att endast utsignalen y kan mätas direkt. Om systemet (5.1) är observerbart, följer av definition 3:3.2 att samtliga tillståndsvariabler kan rekonstrueras ur mätning av utsignalen. Vi skall nu studera hur rekonstruktionen skall genomföras i detalj. Successiva deriveringar av observationsekvationen ger:

$$y = Cx$$

$$\frac{dy}{dt} = C\frac{dx}{dt} = CAx + CBu$$

$$\frac{d^2y}{dt^2} = CA\frac{dx}{dt} + CB\frac{du}{dt} = CA^2x + CABu + CB\frac{du}{dt}$$

$$\vdots$$

$$\frac{d^{n-1}y}{dt^{n-1}} = CA^{n-1}x + CA^{n-2}Bu + CA^{n-3}B\frac{du}{dt} + \ldots + CB\frac{d^{n-2}u}{dt^{n-2}}$$

Omskrivning av dessa ekvationer ger

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x = \begin{bmatrix} y \\ \dfrac{dy}{dt} - CBu \\ \vdots \\ \dfrac{d^{n-1}y}{dt^{n-1}} - CA^{n-2}Bu - CA^{n-3}B\dfrac{du}{dt} - \ldots - CB\dfrac{d^{n-2}u}{dt^{n-2}} \end{bmatrix} \tag{5.10}$$

Om systemet $S(A, B, C, 0)$ är observerbart så är matrisen i vänsterledet av ekvation (5.10) inverterbar (sats 3:5.1). Tillståndsvektorn x kan då alltid beräknas ur insignalerna och utsignalerna och deras derivator med hjälp av ekvation (5.10). Vi har således angivit en konstruktiv metod för rekonstruktion av tillståndsvariabler som ej kan mätas direkt. Observera att konstruktionen kräver $n-1$ derivationer av utsignalen. En jämförelse med ekvation (2:7.5) visar att den högsta derivata av styrvariabeln som förekommer motsvarar gradtalet för överföringsfunktionens täljare minus ett. Vi illustrerar med ett exempel.

EXEMPEL 5.2. Rekonstruktion med derivatorer

Betrakta motorn i exempel 5.1. Med de tillståndsvariabler som infördes i exemplet gäller

$$x_1 = dy/dt$$

$$x_2 = y$$

I stället för att mäta vinkelhastigheten med en takometer så kan motsvarande tillståndsvariabel x_1 rekonstrueras genom att derivera utsignalen.

Rekonstruktion med dynamiska system

Rekonstruktionen (5.10) kan ej alltid realiseras praktiskt då den innebär att såväl styrsignal som mätsignal deriveras ett antal gånger. Det kan ofta vara svårt att derivera en signal. Vi skall därför ge ett alternativ till rekonstruktionen (5.10) som ej kräver att mätsignalen deriveras. Betrakta vektorn $\hat{x}$ definierad av ekvationen

$$\frac{d\hat{x}}{dt} = A\hat{x} + Bu + K[y - C\hat{x}] \tag{5.11}$$

Ekvationen (5.11) kan tolkas som en model av systemet (5.1) som återkopplats med hjälp av utsignalerna y från systemet (5.1). Vi skall visa att det är möjligt att välja matrisen K så att tillståndsvektorn $\hat{x}$ i (5.11) är en god approximation av tillståndsvektorn x i (5.1).

Det följer av ekvationerna (5.1) och (5.11) att differensen

$$\tilde{x} = x - \hat{x} \tag{5.12}$$

satisfierar följande ekvation

$$\frac{d\tilde{x}}{dt} = [A - KC]\,\tilde{x} \tag{5.13}$$

Om $\tilde{x}(t_0) = 0$ så gäller uppenbarligen $\tilde{x}(t) = 0$ för att alla $t \geqslant t_0$. Ekvation (5.11) har således den anmärkningsvärda egenskapen att dess lösning $\hat{x}$ överensstämmer med lösningen till ekvation (5.1) för någon tidpunkt t_0 så överensstämmer $\hat{x}(t)$ och $x(t)$ för alla tider $t \geqslant t_0$. Det förefaller således rimligt att ekvation (5.11) skulle kunna användas för att rekonstruera tillståndsvektorn för systemet (5.2). Låt oss nu undersöka vad som händer om $\hat{x}(0) \neq x(0)$. Om samtliga egenvärden till matrisen $A - KC$ har negativ realdel, så är ekvation (5.13) stabil och samtliga lösningar kommer att konvergera mot noll. Sats 2:11.1 ger följande uppskattning

$$\|\tilde{x}(t)\| < ke^{-\alpha t} \text{ där } -\alpha > \max \operatorname{Re} \lambda[A - KC] \tag{5.14}$$

Om det är möjligt att välja K så att (5.13) blir asymptotiskt stabil så kommer således $\hat{x}(t)$ att närma sig $x(t)$ exponentiellt.

Antag att systemet (5.1) endast har en mätsignal. I analogi med lösningen till polplaceringsproblemet kan vektorn K bestämmas på följande sätt. Eftersom systemet (5.1) antagits vara observerbart så finns enligt avsnitt 2:6 en transformation

$$z = Tx$$

som transformerar (5.1) på observerbar kanonisk form, dvs.

$$\frac{dz}{dt} = \begin{bmatrix} -a_1 & 1 & 0 \dots 0 \\ -a_2 & 0 & 1 \dots 0 \\ \vdots & & \\ -a_{n-1} & 0 & 0 \dots 1 \\ -a_n & 0 & 0 \dots 0 \end{bmatrix} z + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} u = \tilde{A}z + \tilde{B}u$$

$$y = [1 \ 0 \ 0 \dots 0]z = \tilde{C}z$$

Transformationsmatrisen T ges av ekvation (2:6.15). För rekonstruktionen

$$\frac{d\hat{z}}{dt} = \tilde{A}\hat{z} + \tilde{B}u + \tilde{K}[y - \tilde{C}\hat{z}]$$

där

$$\tilde{K} = \begin{bmatrix} d_1 & -a_1 \\ d_2 & -a_2 \\ \vdots & \\ d_{n-1} & -a_{n-1} \\ d_n & -a_n \end{bmatrix} \tag{5.15}$$

gäller då

$$\det [sI - \tilde{A} + \tilde{K}\tilde{C}] = s^n + d_1 s^{n-1} + \dots + d_n$$

En transformation till de ursprungliga koordinaterna ger vidare

$$K = T^{-1}\tilde{K} \tag{5.16}$$

Eftersom den karakteristiska ekvationen är invariant vid byte av koordinater i tillståndsrummet så är vektorn K sådan att

$$\det [sI - A + KC] = s^n + d_1 s^{n-1} + \dots + d_n \tag{5.17}$$

Observera att analysen är helt analog med den som genomfördes vid beviset av sats 5.1.

Vektorn K kan således väljas så att matrisen $A - KC$ får godtyckliga egenvärden, dvs. vi kan genom val av vektorn K godtyckligt förfoga över parametern α i ekvation (5.14). Om vektorn K väljs så att talet α är stort konvergerar $\hat{x}(t)$ mycket snabbt mot $x(t)$ oavsett initialtillståndet $\hat{x}(0)$. Med ett stort värde på parametern α blir emellertid beräkningen som definieras av ekvation (5.11) känslig för fel i mätningen av utsignalen y. I praktiken får man därför kompromissa mellan konvergenshastighet och störningskänslighet. För att bestämma den bästa K-vektorn måste man ha detaljerad kunskap om störningarnas karaktär. Saknas sådan information kan man emellertid ofta uppnå mycket bra resultat genom att bestämma vektorn K med hjälp av simulering.

Genom att beskriva processtörningar och mätfel med statistisk teori går det att visa att det finns en i viss mening bästa uppskattning av tillståndsvektorn som ges av ett dynamiskt system av formen (5.11) med ett speciellt värde på K. Denna teori utvecklades av Kalman och vi skall därför kalla (5.11) för *Kalmanfilter*.

För att bestämma vektorn K så att polynomet (5.17) får föreskrivna rötter kan ekvationerna (5.15) och (5.16) utnyttjas. Detta kräver dock att transformationsmatrisen T beräknas med hjälp av ekvation (2:6.15). En alternativ metod är att ansätta ett K och direkt utnyttja de villkor som krävs för att $\det[sI - A + KC]$ skall bli ett föreskrivet polynom. Vi ger ett exempel.

EXEMPEL 5.3. Rekonstruktion med Kalmanfilter

Betrakta motorn i exempel 5.1 och anta att endast motorns vridningsvinkel mäts. I det speciella fallet kan rekonstruktionen (5.11) skrivas som

$$\frac{d\hat{x}}{dt} = \begin{bmatrix} -a & 0 \\ 1 & 0 \end{bmatrix} \hat{x} + \begin{bmatrix} b \\ 0 \end{bmatrix} u + \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} [y - \hat{x}_2]$$

Det gäller att

$$\det[sI - A + KC] = s^2 + s(a + k_2) + k_1 + ak_2$$

För att rekonstruktionsfelet skall gå mot noll måste parametrarna väljas så att $k_2 > -a$ och $k_1 > -ak_2$. Det sätt på vilket rekonstruktionsfelet går mot noll bestäms av rötterna till det karakteristiska polynomet. Anta t. ex. att det krävs att det karakteristiska polynomet har en dubbelrot i $s = -10$. Vi finner då $k_1 = 100 - 20a + a^2$ och $k_2 = 20 - a$. Rekonstruktionsfelet är sammansatt av termer proportionella mot $\exp(-10t)$ och $t \cdot \exp(-10t)$.

Den uppskattade tillståndsvektorns Laplacetransform ges av

$$\hat{X}_1(s) = \frac{b(s + k_2)\, U(s) + k_1 s\, Y(s)}{s^2 + s(a + k_2) + k_1 + ak_2}$$

$$\hat{X}_2(s) = \frac{bU(s) + (k_2 s + k_1 + ak_2) Y(s)}{s^2 + s(a + k_2) + k_1 + ak_2}$$

Ekvationerna (5.10) och (5.11) är således två olika alternativ till att rekonstruera tillståndsvektorn ur mätning av in- och utsignaler. Observera att både (5.10) och (5.11) kan tolkas som dynamiska system. Systemet (5.10) är av nollte ordningen och innehåller endast derivatorer. Systemet (5.11) är ett dynamiskt system av n:te ordningen. Den rekonstruktion som ges av ekvation (5.11) är flexibel därigenom att vi kan förfoga över vektorn K inom vida gränser.

Eftersom rekonstruktionen (5.10) motsvarar ett system med enbart derivatorer och (5.11) ett system av n:te ordningen, så är det naturligt att fråga om det ej finns några andra alternativ som ger rekonstruktioner av ordningstal

mellan noll och n. Det finns i litteraturen flera sådana alternativ. De brukar kallas *observerare* och de är alla av formen

$$\frac{d\hat{z}}{dt} = A_1\hat{z} + B_1 u + B_2 y$$

$$\hat{x} = C_1\hat{z} + D_1(p)u + D_2(p)y \tag{5.18}$$

där $D_1(p)$ och $D_2(p)$ är polynom i derivationsoperatorn p.

Låt n vara antalet tillståndsstorheter och p antalet mätstorheter. Eftersom mätstorheterna är kända, behöver vi endast rekonstruera $n-p$ storheter för att hela tillståndsvektorn skall vara känd. Den observerare som rekonstruerar tillståndsvektorn på detta sätt kallas Luenbergers observerare. Den kan härledas på följande sätt: Välj en stabil $(n-p)\times(n-p)$ matrix A_1 och en $(n-p)\times p$ matrix $\tilde{K}$. Bestäm sedan $(n-p)\times n$ matrisen T så att

$$TA - A_1 T = \tilde{K}C \tag{5.19}$$

Följande ekvation

$$\frac{d\hat{z}}{dt} = A_1\hat{z} + TBu + \tilde{K}y \tag{5.20}$$

är nu en observerare för $z = Tx$, ty det följer av (5.1) och (5.20) att

$$\frac{d}{dt}(\hat{z} - Tx) = A_1\hat{z} + TBu + \tilde{K}y - TAx - TBu = A_1(\hat{z} - Tx)$$

Eftersom A_1 är stabil, kommer $\hat{z}$ att konvergera mot Tx. För att få en uppskattning av tillståndsvektorn löses sedan $\hat{x}$ ur ekvationssystemet

$$\begin{bmatrix} T \\ C \end{bmatrix}\hat{x} = \begin{bmatrix} \hat{z} \\ y \end{bmatrix} \tag{5.21}$$

Det kan visas att matrisen T kan lösas ur ekvation (5.19) om A och A_1 ej har gemensamma egenvärden. Vidare kan ekvationssystemet (5.21) lösas om systemet $\mathcal{S}(A_1, \tilde{K}, 0, 0)$ är styrbart och om $\mathcal{S}(A, 0, C, 0)$ är observerbart.

Följande exempel visar hur Luenbergers observerare kan bestämmas.

Exempel 5.4. Rekonstruktion med Luenbergers observerare

Betrakta servomotorn i exempel 5.1. Det gäller att $n=2$ och $p=1$. Välj $A_1 = -\alpha$ och $\tilde{K} = k$. Ekvation (5.19) lyder då

$$[t_1 \;\; t_2]\begin{bmatrix} -a & 0 \\ 1 & 0 \end{bmatrix} + \alpha[t_1 \;\; t_2] = k[0 \;\; 1]$$

Denna ekvation kan lösas med avseende på t_1 och t_2 om $\alpha \neq a$. Enkla räkningar ger

$$t_1 = k/[\alpha(a-\alpha)], \quad t_2 = k/\alpha$$

Rekonstruktionen (5.20) lyder

$$\frac{d\hat{z}}{dt} = -\alpha\hat{z} + t_1 bu + ky$$

Vidare ges $\hat{x}$ av (5.21), dvs.

$$\hat{x}_1 = \hat{z}/t_1 - t_2 y/t_1$$

$$\hat{x}_2 = y$$

Enkla räkningar ger

$$\hat{X}_1(s) = \frac{b}{s+\alpha} U(s) + \frac{(\alpha - a)\, s}{s+\alpha} Y(s)$$

vilket är Laplacetransformen för Luenbergers observerare för tillståndsvariabeln x_1. Observera att koefficienten framför Y försvinner då $\alpha = a$. Rekonstruktionen kan utföras i detta fall även om ekvationssystemet för t_1 och t_2 ej har någon lösning.

Vi har således funnit att om systemet är observerbart så kan tillståndsvariabler som ej mäts direkt uppskattas på flera olika sätt med hjälp av derivation, Kalmanfilter eller Luenbergers observerare. Vi fann t. ex. i exempel 5.2, 5.3 och 5.4 att för servomotorn i exempel 5.1 så kan vridningsvinkelns tidsderivata uppskattas med

$$\hat{X}_1(s) = s\, Y(s) \qquad \text{(derivator)}$$

$$\hat{X}_1(s) = \frac{bU(s) + (\alpha - a)\, s\, Y(s)}{s+\alpha} \qquad \text{(Luenberger)}$$

$$\hat{X}_1(s) = \frac{b(s+k_2)\, U(s) + k_1 s\, Y(s)}{s^2 + s(a+k_2) + k_1 + ak_2} \qquad \text{(Kalman)}$$

De olika metoderna att uppskatta rotationshastigheten skiljer sig dels genom komplexitet och dels genom att observerarna reagerar olika på mätbrus och processtörningar. Derivation är enklast men den är också olämplig om det finns högfrekvent mätbrus. Kalmanfiltret är mest komplicerat men överföringsfunktionen från mätsignal y till uppskattning $\hat{x}_1$ dämpar högfrekvent mätbrus.

Vi kan också lägga märke till att om T är en $n \times n$ matris som satisfierar (5.19) så ger (5.20) en rekonstruktion av Tx. Detta kan utnyttjas för att ge följande alternativmetod för att bestämma matrisen K i rekonstruktionen (5.11).

- Välj en stabil $n \times n$ matris A_1 så att A och A_1 ej har gemensamma egenvärden
- Välj en matrix $\tilde{K}$ sådan att $\mathrm{S}(A_1, \tilde{K}, 0, 0)$ är styrbar

- Lös T ur det linjära ekvationssystemet (5.19)
- Matrisen K i ekv. (5.11) ges sedan av $K = T^{-1}\tilde{K}$

Analys av en regulator baserad på tillståndsåterkoppling och rekonstruktion

Det har visats att det slutna systemets poler kan ges godtyckliga värden genom en tillståndsåterkoppling. Vidare har vi visat hur tillstånd som ej mäts kan rekonstrueras. Det system som erhålls genom kombination av rekonstruktion och tillståndsåterkoppling från de rekonstruerade tillståndsvariablerna skall nu analyseras. Ett blockschema för systemet visas i fig. 5.1. Systemet beskrivs av

$$\frac{dx}{dt} = Ax + Bu$$

$$y = Cx \tag{5.22}$$

Enligt ekvation (5.11) ges rekonstruktionen av

$$\frac{d\hat{x}}{dt} = A\hat{x} + Bu + K[y - C\hat{x}] \tag{5.23}$$

Vidare antas att styrlagen har formen

$$u = u^r - L(\hat{x} - x^r) = -L\hat{x} + u^r + Lx^r \tag{5.24}$$

där u^r och x^r är styrvariabelns respektive tillståndsvariablernas referensvärden. Observera att då u^r och x^r endast förekommer i (5.24) i kombinationen $y^r = u^r + Lx^r$, så kan det synas onödigt att införa båda referensvärdena. Det finns dock en fördel med att skriva regulatorn på formen (5.24), ty om modellens överföringsfunktion överensstämmer med processens överföringsfunktion så genereras inga korrektionssignaler av återkopplingen, oavsett vilket värde matrisen L har. Man får således på detta sätt en effektiv separation av framkoppling och återkoppling. Det finns många sätt att generera u^r och x^r. En möjlighet är att använda optimeringsteori som angavs i avsnitt 3. Jämför fig. 3.13. En annan möjlighet är att generera x^r från u^r med hjälp av en modell som ger det önskade sambandet. En tredje möjlighet är att utnyttja framkoppling som beskrivits i detalj i avsnitten 3 och 4. Signalen u^r genereras då av det inversa systemet och signalen x^r kan t. ex. erhållas ur det inversa systemets tillstånd. Jämför diskussionen i anslutning till exempel 4.7. Om en approximativ invers används så kan x^r genereras ur en modell med u^r som insignal. Detta visas i fig. 5.1.

Det slutna systemet ges således av ekvationerna (5.22), (5.23) och (5.24). För att analysera systemet införs tillståndsvariablerna x och $\tilde{x} = x - \hat{x}$. Enkla räkningar visar att systemet då beskrivs av

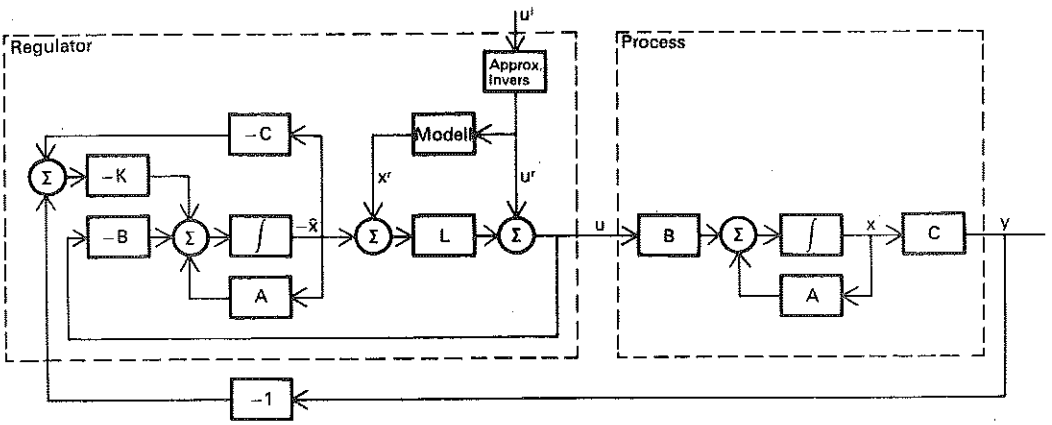


Fig. 5.1. Blockschema för en regulator baserad på tillståndsåterkoppling och rekonstruktion av icke mätbara tillståndsvariabler med ett Kalmanfilter.

$$\frac{d}{dt}\begin{pmatrix} x \\ \tilde{x} \end{pmatrix} = \begin{pmatrix} A-BL & BL \\ 0 & A-KC \end{pmatrix}\begin{pmatrix} x \\ \tilde{x} \end{pmatrix} + \begin{pmatrix} B \\ 0 \end{pmatrix} u^r + \begin{pmatrix} BL \\ 0 \end{pmatrix} x^r$$

$$y = Cx$$

$$u = -Lx + L\tilde{x} + u^r + Lx^r \tag{5.25}$$

Vi finner således att det slutna systemet är av ordningen $2n$. Systemets poler är egenvärdena till matriserna $A-BL$ och $A-KC$. Det följer vidare av (5.25) att tillståndet $\tilde{x}$ ej är styrbart vare sig från u^r eller x^r. Jämför med analysen i avsnitt 3:6. De moder som genereras vid rekonstruktionen av tillståndsvariablerna är således ej styrbara.

Enligt Kalmans uppdelningssats (sats 3:6.1) kan det återkopplade systemet uppdelas på det sätt som visas i fig. 5.2. Insignal–utsignal-relationen från u^r och x^r till y ges av

$$Y(s) = C[sI - A + BL]^{-1}B[U^r(s) + LX^r(s)]$$

Överföringsfunktionen är således entydigt bestämd av A, B, C och L. Den är helt oberoende av hur matrisen K väljs. Man kan visa att analoga förhållanden gäller oavsett hur tillståndsvariablerna rekonstrueras. De moder som införs i observeraren blir alltid icke styrbara.

Den regulatorstruktur som visas i fig. 5.1 är mycket flexibel. Den kan också användas vid reglering av flervariabla system. För att dimensionera regulatorn krävs kunskap om matriserna A, B och C, som kan erhållas ur en matematisk modell av processen. Vidare måste matriserna K och L bestämmas. För ett styrbart system med en insignal ger sats 5.1 ett L sådant att det slutna systemet får givna poler. Det finns också utvidgningar av sats 5.1 till flervariabla system. Val av matrisen K påverkar ej det slutna systemets över-

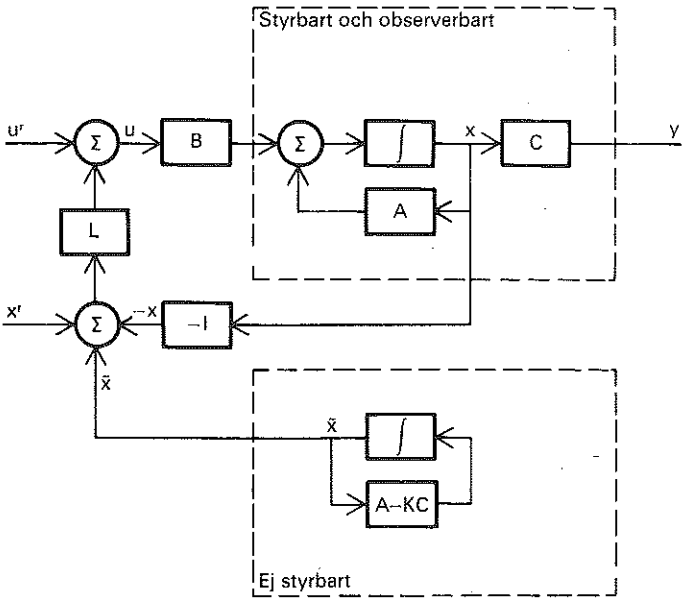


Fig. 5.2. Kalmans uppdelning av systemet i fig. 5.1. Det finns två delsystem: ett är styrbart och ej observerbart, det andra är observerbart men ej styrbart.

föringsfunktion. Däremot bestämmer K hur snabbt de rekonstruerade tillståndsvariablerna närmar sig de verkliga tillståndsvariablerna. Observera också att det finns flera andra syntesmetoder som leder till en regulator med den struktur som visas i fig. 5.1, t. ex. den linjär-kvadratiska metoden som kortfattat behandlades i avsnitt 3.

Regulatorn i fig. 5.1 innehåller en modell av den process som skall regleras. Dynamiken i regulatorn uppträder genom att tillståndsvariablerna rekonstrueras med hjälp av modellen. Man bör också lägga märke till att modellen kan utvidgas så att den beskriver systemets omgivning. Vi belyser detta med ett exempel.

EXEMPEL 5.5. *PI* regulatorn

Betrakta ett system av första ordningen som beskrivs av

$$\frac{dx}{dt} = -ax + bu$$

Modellen ovan har en allvarlig brist. I stationärt tillstånd gäller

$$x^0 = (b/a)u^0$$

Om $u^0 = 0$ gäller således $x^0 = 0$. För att skriva upp modellen är det således nödvändigt att givare och ställdon är perfekt kalibrerade. Även om parametrarna

a och b kan bestämmas med god noggrannhet kan det vara mycket svårt att uppnå en perfekt kalibrering av styrsignalen. En mer realistisk modell är därför

$$\frac{dx}{dt} = ax + b(u + v)$$

där v är en konstant som representerar styrvariabelns okända nollpunkt. Eftersom v är konstant gäller

$$\frac{dv}{dt} = 0$$

och systemet och dess omgivning kan således representeras av följande dynamiska system

$$\frac{d}{dt}\begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ v \end{pmatrix} + \begin{pmatrix} b \\ 0 \end{pmatrix} u$$

$$y = x$$

som har två tillståndsvariabler. Observera att endast tillståndsvariablen x är styrbar.

Styrlagen

$$u = -l(x - x^r) - \hat{v}$$

medför att en av det slutna systemets poler kan ges ett föreskrivet värde. Vidare elimineras inverkan av störningen v om $\hat{v}$ är en god skattning av v. Storheten kan skattas t. ex. med ett Kalmanfilter (5.11), dvs.

$$\frac{d\hat{v}}{dt} = K(x - \hat{x})$$

En enklare approximativ skattning erhålls genom att ersätta $\hat{x}$ med x^r, dvs.

$$\frac{d\hat{v}}{dt} \approx K(x - x^r)$$

Detta leder till styrlagen

$$u = -l(x - x^r) - K\int^t [x(\tau) - x^r(\tau)]d\tau$$

vilken är en PI-regulator. Vi finner således i detta fall att I-termen i PI-regulatorn kan tolkas som en approximativ rekonstruktion av ett konstant men okänt kalibreringsfel i styrvariabeln. Observera att v också skulle kunna tolkas som en konstant störning som påverkar processen.

Genom att utvidga modellen så att den också beskriver systemets omgivning

går det att göra en regulator som är anpassad till störningar av speciell karaktär. Störningar i form av steg, ramper och n:te ordningens polynom kan beskrivas med

$$\frac{dx}{dt} = \begin{pmatrix} 0 & 1 & 0 \dots 0 \\ 0 & 0 & 1 \dots 0 \\ \vdots & & \\ 0 & 0 & 0 \dots 1 \\ 0 & 0 & 0 \dots 0 \end{pmatrix} x$$

och störningar i form av en periodisk signal med frekvensen ω beskrivs med

$$\frac{dx}{dt} = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix} x$$

På samma sätt kan man behandla störningar som är lösningar till ordinära differentialekvationer med konstanta koefficienter.

Polplacering genom utsignal–återkoppling

Den analys som genomförts var uteslutande baserad på en intern beskrivning, tillståndsmodellen (5.1), av det system som skall regleras. Vi skall nu visa att samma resultat också kan erhållas direkt utgående från en extern (insignal–utsignal) beskrivning av systemet. Resultatet är av intresse, ty dels får vi andra metoder att bestämma regulatorn och dels får vi andra tolkningar till de tidigare resultaten.

Vi skall utgå från att det system som skall regleras har överföringsfunktionen

$$G(s) = \frac{B(s)}{A(s)}$$

där polynomet $A(s)$ har graden n. Det kommer att vara betydelsefullt att veta gradtalet för polynomet $B(s)$ och vi antar därför att

$$B(s) = b_0 s^m + b_1 s^{m-1} + \dots + b_m$$

där $b_0 \neq 0$ och $m < n$. Vidare införs beteckningen grad $B(s)$ för att beteckna polynomets gradtal.

Vi skall nu visa att man direkt kan bestämma överföringsfunktionen för en regulator som ger föreskrivna poler för överföringsfunktionen från referensvärdet y^r till utsignalen y. Resultatet skall erhållas genom att först översätta det tidigare resultatet till insignal–utsignal-relationer.

Den regulator som består av ett dynamiskt system för rekonstruktion av tillståndsvariablerna och en återkoppling från de rekonstruerade tillståndsvariablerna beskrivs av ekvationerna (5.22), (5.23) och (5.24). Enkla räkningar visar att regulatorn också kan beskrivas med ekvationerna

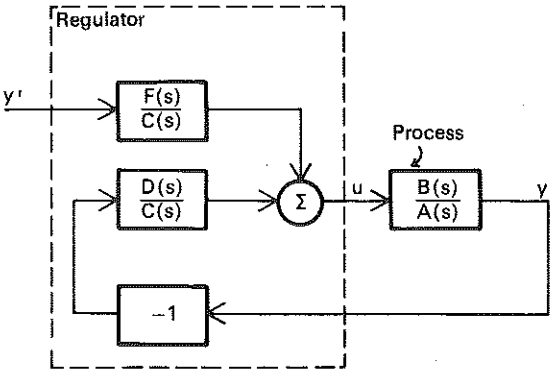


Fig. 5.3. Blockschema för den regulator som ges av ekvation (5.30).

$$\frac{d\hat{x}}{dt} = [A - BL - KC]\hat{x} + B[u^r + Lx^r] + Ky$$

$$u = -L\hat{x} + u^r + Lx^r \tag{5.26}$$

Regulatorn kan således beskrivas som ett dynamiskt system av n:te ordningen med insignalerna $y^r = u^r + Lx^r$ och y. Laplacetransformation av (5.26) visar att regulatorn har insignal–utsignal-relationen

$$C(s)\,U(s) = -D(s)\,Y(s) + F(s)[U^r(s) + LX^r(s)]$$

där

$$C(s) = \det[sI - A + BL + KC] \tag{5.27}$$

och polynomen $D(s)$ och $F(s)$ definieras av

$$C^{-1}(s)\,D(s) = L[sI - A + BL + KC]^{-1}K \tag{5.28}$$

$$C^{-1}(s)\,F(s) = I - L[sI - A + BL + KC]^{-1}B \tag{5.29}$$

Då u^r och x^r endast förekommer i kombinationen $y^r = u^r + Lx^r$ i ekv. (5.26) kan regulatorn (5.26) skrivas

$$C(s)\,U(s) = -D(s)\,Y(s) + F(s)\,Y^r(s) \tag{5.30}$$

Observera att regulatorn kan tolkas som en kombination av fram- och återkoppling. Regulatorns blockschema framgår av fig. 5.3. Blockschemat visar endast insignal–utsignal-relationerna och är därför mycket enklare än blockschemat i fig. 5.1 som ger en intern beskrivning av process och regulator. Regulatorns insignal–utsignal egenskaper är uppenbarligen entydigt bestämda av tre polynom $C(s)$, $D(s)$ och $F(s)$. Dessa polynom skall nu bestämmas så att det slutna systemet får den önskade överföringsfunktionen.

Det följer av fig 5.3 att överföringsfunktionen från y^r till y är

$$G_1(s) = \frac{B(s)\,F(s)}{A(s)\,C(s) + B(s)\,D(s)}$$

Enligt specifikationerna skall denna överföringsfunktion vara lika med $G_r(s)$ som ges av (5.3). Det krävs således att

$$G_1(s) = \frac{B(s)F(s)}{A(s)C(s) + B(s)D(s)} = \frac{Q(s)}{P(s)} = G_r(s) \tag{5.31}$$

Syntesproblemet kan alltså reduceras till det rent algebraiska problemet att finna polynom $C(s)$, $D(s)$ och $F(s)$ som uppfyller ovanstående ekvation.

Det slutna systemets nollställen är de faktorer i $B(s)F(s)$ som ej är faktorer i $A(s)C(s) + B(s)D(s)$. Nollställen kan alltså tillfogas genom lämpligt val av polynomet $F(s)$. Vidare kan nollställen elimineras genom förkortning. Det följer emellertid av analysen av framkoppling i avsnitt 4 att nollställen i högra halvplanet ej kan förkortas. Jämför med exempel 4.10. För att lösa syntesproblemet måste vi således kräva att om $B(s)$ har nollställen i högra halvplanet så har $Q(s)$ motsvarande nollställen. Inför

$$B(s) = B_1(s)Q_2(s) \tag{5.32}$$

$$Q(s) = Q_1(s)Q_2(s) \tag{5.33}$$

där $Q_2(s)$ är den största gemensamma faktorn till polynomen $B(s)$ och $Q(s)$. För att (5.31) skall gälla måste $Q_1(s)$ vara en faktor i $F(s)$. Inför vidare

$$F(s) = Q_1(s)F_1(s) \tag{5.34}$$

Det måste gälla att $B_1(s)F_1(s)$ är en faktor i $A(s)C(s) + B(s)D(s)$. Det följer av konstruktionen att $B_1(s)$ ej har nollställen i högra halvplanet. Det krävs vidare att $F_1(s)$ ej har nollställen i högra halvplanet. Polynomen $C(s)$ och $D(s)$ skall således bestämmas så att ekvationen

$$A(s)C(s) + B(s)D(s) = P(s)B_1(s)F_1(s) \tag{5.35}$$

är uppfylld.

Sammanfattningsvis blir syntesproceduren följande:

- Problemet kan endast lösas om polynomen $B(s)$ och $Q(s)$ har samma nollställen i högra halvplanet.
- Bestäm största gemensamma faktorn $Q_2(s)$ till polynomen $B(s)$ och $Q(s)$.
- Välj ett godtyckligt polynom $F_1(s)$ som har alla nollställen i vänstra halvplanet.
- Lös ekvationen (5.35) med avseende på $C(s)$ och $D(s)$.
- Regulatorn ges sedan av (5.30) där polynomet $F(s)$ ges av (5.34).

Om proceduren kan genomföras resulterar den uppenbarligen i ett slutet system med den önskade överföringsfunktionen (5.3).

Om polynomen $A(s)$, $B(s)$ och $P(s)B_1(s)F_1(s)$ är givna så har ekvation (5.35)

i allmänhet många lösningar. En möjlighet att välja ut lämpliga lösningar ges av

SATS 5.2.
Låt $A(s)$, $B(s)$ och $H(s)$ vara givna polynom. Ekvationen

$$A(s)C(s)+B(s)D(s)=H(s) \tag{5.36}$$

har en entydig lösning $C(s)$, $D(s)$ om polynomen $A(s)$ och $B(s)$ saknar gemensamma faktorer och ett av villkoren

(i) grad $D(s)<$ grad $A(s)$

eller

(ii) grad $C(s)<$ grad $B(s)$

är uppfyllda.

Bevis:
Det existerar alltid en lösning, ty om $A(s)$ och $B(s)$ saknar gemensamma faktorer så har ekvationen

$$A(s)X(s)+B(s)Y(s)=1$$

en lösning sådan att grad $X(s)<$grad $B(s)$ och grad $Y(s)<$grad $A(s)$. Det följer då att ekvation (5.36) har lösningen

$$C(s)=H(s)X(s)+R(s)B(s)$$
$$D(s)=H(s)Y(s)-R(s)A(s)$$

där $R(s)$ är ett godtyckligt polynom. Genom att välja $R(s)$ som kvoten av $H(s)Y(s)$ vid division med $A(s)$, och $D(s)$ som resten, finner vi att det är möjligt att finna en lösning till (5.36) som uppfyller (i). På samma sätt kan vi bestämma ett polynom $R(s)$ så att (ii) gäller.

De lösningar till (5.36) som uppfyller (i) eller (ii) är entydiga därför att motsvarande homogena ekvation

$$A(s)C(s)+B(s)D(s)=0 \tag{5.37}$$

har endast den triviala lösningen $C(s)=D(s)=0$. För att visa detta skrivs (5.37) som

$$A(s)C(s)=-B(s)D(s)$$

Dela upp båda leden i denna ekvation i faktorer. Faktorerna till polynomet $A(s)$ måste då återfinnas i högerledet. I fall (i) har $D(s)$ lägre gradtal än $A(s)$. Polynomen $A(s)$ och $B(s)$ har då åtminstone en gemensam faktor. På

samma sätt finner vi i fall (ii) att faktorerna till $B(s)$ måste finnas i vänsterledet. Eftersom $C(s)$ är av lägre grad än $B(s)$, måste $A(s)$ och $B(s)$ ha åtminstone en gemensam faktor. Vi får en motsägelse och satsen är bevisad.

Villkoret att polynomen $A(s)$ och $B(s)$ saknar gemensamma faktorer är ekvivalent med att systemet (5.1) är styrbart och observerbart.

Observera att syntesproceduren innehåller ett moment av godtycke i och med att polynomet $F_1(s)$, som förkortas bort i (5.31), kan väljas fritt. Detta motsvaras i den tidigare analysen av att det fanns många val av observerare, som alla ger samma överföringsfunktion från y^r till y därigenom att de moder som införs genom observeraren ej är styrbara. En jämförelse med analysen av en regulator bestående av en observerare och tillståndsåterkoppling visar att polynomet $F_1(s)$ motsvarar observerarens karakteristiska ekvation. Vid syntes med tillståndsåterkoppling påverkas ej processens nollställen. I detta fall gäller således $Q(s) = B(s)$, vilket medför $B_1(s) = Q_1(s) = 1$ och $F_1(s) = F(s)$. Ekvation (5.35) reduceras då till

$$A(s)\,C(s) + B(s)\,D(s) = P(s)\,F(s) \tag{5.35'}$$

där grad $P(s) =$ grad $A(s)$. De olika metoderna för rekonstruktion av tillståndsvariabler motsvarar olika val av ordningstal för polynomet $F(s)$. Vi finner

- Om tillståndsvariablerna rekonstrueras genom derivation så följer det av (5.10) att processens insignal deriveras grad $B(s) - 1$ gånger (grad $B(s) > 1$). Eftersom processens insignal erhålls genom en återkoppling från de rekonstruerade tillståndsvariablerna blir regulatorn således ett dynamiskt system av ordningen grad $B(s) - 1$ om grad $B(s) > 1$. Detta innebär att regulatorn som beskrivs av (5.30) är sådan att grad $C(s) =$ grad $B(s) - 1$, vilket motsvarar fall (ii) i sats 5.2. Det följer vidare av (5.10) att processens utsignal deriveras grad $A(s) = 1$ gånger, dvs grad $D(s) =$ grad $A(s) - 1$. Ekvation (5.35') medför vidare att grad $F(s) =$ grad $C(s)$. Sammanfattningsvis finner vi således att rekonstruktion med derivation motsvarar

  $$\text{grad } C(s) = \text{grad } F(s) = \max\,[0, \text{grad } B(s) - 1]$$
  $$\text{grad } D(s) = \text{grad } A(s) - 1 \tag{5.38}$$

- Vid rekonstruktion med Luenbergers observerare blir regulatorn ett dynamiskt system av ordningstalet grad $A(s) - 1$. Polynomet $C(s)$ i regulatorn (5.30) skall då väljas så att grad $C(s) =$ grad $A(s) - 1$. Då den regulator som erhålls med Luenbergers observerare ej innehåller några derivatorer, måste vi dessutom välja grad $D(s) \leqslant$ grad $A(s) - 1$, vilket motsvarar fall (i) i sats 5.2. Sammanfattningsvis finner vi att rekonstruktion med Luenbergers observerare motsvarar

  $$\text{grad } C(s) = \text{grad } F(s) = \text{grad } A(s) - 1$$
  $$\text{grad } D(s) < \text{grad } A(s) \tag{5.39}$$

- Vid rekonstruktion med Kalmanfilter följer det av (5.26) att regulatorn blir ett dynamiskt system av ordningen grad $A(s)$. Polynomet $C(s)$ skall då också ha samma grad. Jämför ekvation (5.27). Det följer vidare av (5.28) att polynomet $D(s)$ är av lägre grad, vilket motsvarar fall (i) i sats 5.2. Sammanfattningsvis finner vi för rekonstruktion med Kalmanfilter

$$\text{grad } C(s) = \text{grad } F(s) = \text{grad } A(s)$$

$$\text{grad } D(s) < \text{grad } A(s) \tag{5.40}$$

Observera emellertid att det finns många andra sätt att välja polynomen. Vi illustrerar syntesmetoden med några exempel.

EXEMPEL 5.6. Rekonstruktion med derivatorer

Betrakta servomotorn i exempel 5.1. Anta liksom tidigare att specifikationerna kräver att överföringsfunktionen från y^r till y skall vara

$$G_r(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2} = \frac{\omega^2}{P(s)}$$

Genom att välja förstärkningen $b = \omega^2$ erhålls $B(s) = Q(s)$, vilket medför att $Q(s) = 1$ och $F_1(s) = F(s)$. Det följer av (5.38) att polynomen $C(s)$ och $F(s)$ kan väljas av graden noll och polynomet $D(s)$ av graden ett. Ansätt

$$C(s) = 1$$

$$D(s) = d_0 s + d_1$$

$$F(s) = 1$$

Ekvationen (5.35) ger

$$A(s)C(s) + B(s)D(s) = P(s)$$

Identifiering av koefficienterna för s och s^0 ger

$$a + d_0\omega^2 = 2\zeta\omega$$

$$d_1\omega^2 = \omega^2$$

dvs.

$$d_0 = 2\zeta/\omega - a/\omega^2$$

$$d_1 = 1$$

Regulatorn kan således beskrivas med insignal–utsignal-relationen

$$u(t) = -[2\zeta/\omega - a/\omega^2]\frac{dy}{dt} - y + y^r$$

En jämförelse med exempel 5.1 och 5.2 visar att regulatorn är densamma som

erhålls med tillståndsåterkoppling och rekonstruktion genom derivation av processens utsignal.

EXEMPEL 5.7. Rekonstruktion med Luenbergers observerare

Betrakta motorn i exempel 5.1. Välj förstärkningen så att $b = \omega^2$. Eftersom grad $A(s) = 2$, följer det av (5.39) att polynomen $C(s)$ och $F(s)$ också kan väljas som förstagradspolynom. Ansätt

$$C(s) = s + c_1$$

$$D(s) = d_0 s + d_1$$

$$F(s) = s + f_1$$

där f_1 är ett godtyckligt positivt tal. Ekvationen (5.35) ger

$$A(s)C(s) + B(s)D(s) = F(s)P(s)$$

Identifiering av koefficienterna för potenserna s^2, s och s^0 ger följande ekvationer

$$a + c_1 = f_1 + 2\zeta\omega$$

$$ac_1 + d_0\omega^2 = \omega^2 + 2\zeta\omega f_1$$

$$\omega^2 d_1 = f_1\omega^2$$

vilka har lösningen

$$c_1 = f_1 + 2\zeta\omega - a$$

$$d_0 = 1 + (a/\omega)^2 - 2\zeta(a/\omega) + [2\zeta - a/\omega]f_1/\omega$$

$$d_1 = f_1$$

En jämförelse med rekonstruktion av tillståndsvariabler enligt exempel 5.4 och en tillståndsåterkoppling enligt exempel 5.1 från de rekonstruerade tillståndsvariablerna visar att samma regulator erhållits.

EXEMPEL 5.8. Rekonstruktion med Kalmanfilter

Betrakta servomotorn i exempel 5.1. Då polynomet $A(s)$ är av andra graden, följer det av (5.40) att polynomen $C(s)$, $D(s)$ och $F(s)$ kan väljas som

$$C(s) = s^2 + c_1 s + c_2$$

$$D(s) = d_0 s + d_1$$

$$F(s) = s^2 + f_1 s + f_2$$

där f_1 och f_2 är godtyckliga positiva tal. Identifikation av koefficienterna för potenserna s^3, s^2, s^1 och s^0 i ekvation (5.35) ger följande ekvationssystem

$$a + c_1 = f_1 + 2\zeta\omega$$

$$ac_1 + c_2 = \omega^2 + f_2 + 2\zeta\omega f_1$$
$$ac_2 + \omega^2 d_0 = \omega^2 f_1 + 2\zeta\omega f_2$$
$$\omega^2 d_1 = \omega^2 f_2$$

Detta ekvationssystem kan lätt lösas och vi finner

$$c_1 = f_1 + 2\zeta\omega - a$$
$$c_2 = \omega^2 + a^2 - 2\zeta a\omega + (2\zeta\omega - a) f_1 + f_2$$
$$d_0 = f_1 + 2\zeta f_2/\omega - ac_2/\omega^2$$
$$d_1 = f_2$$

Regulatorn är densamma som erhålls genom rekonstruktion med Kalmanfilter och återkoppling från de rekonstruerade tillståndsvariablerna.

De regulatorer som bestämts i exempel 5.6, 5.7 och 5.8 ger alla samma överföringsfunktion från y^r till y. Regulatorerna har följande insignal–utsignal-relationer

$$U(s) = -[d_0 s + d_1]\, Y(s) + Y^r(s) \qquad \text{(Exempel 5.6)}$$
$$[s + c_1]\, U(s) = -[d_0 s + d_1]\, Y(s) + [s + f_1]\, Y^r(s) \qquad \text{(Exempel 5.7)}$$
$$[s^2 + c_1 s + c_2]\, U(s) = -[d_0 s + d_1]\, Y(s) + [s^2 + f_1 s + f_2]\, Y^r(s) \qquad \text{(Exempel 5.8)}$$

Snabba variationer i processens mätsignal förekommer i vissa fall. Dessa variationer kan orsakas av mätbrus eller av snabbt varierande störningar. Eftersom det alltid tar viss tid innan en förändring i processens styrvariabel märks i utsignalen, så kan snabba variationer ej elimineras med reglering. Det är då viktigt att se till att regulatorn är sådan att de snabba variationerna i processens utsignal ej genererar variationer i processens styrvariabel. De olika regulatorerna som givits ovan är mycket olika i detta avseende.

Eftersom regulatorn i exempel 5.6 förstärker höga frekvenser kraftigt, är den mycket olämplig om det finns snabba variationer i processens utsignal. Regulatorn i exempel 5.7 är något bättre i detta avseende, eftersom den har konstant förstärkning vid höga frekvenser. För regulatorn i exempel 5.8 gäller $U(s) \approx -d_0/s\ Y(s)$ för $|s| \gg 1$, vilket medför att snabba variationer i mätsignalen y orsakar små variationer i processens styrvariabel. Det är också möjligt att göra regulatorer som ger den föreskrivna överföringsfunktionen från y^r till y men som är ännu effektivare att eliminera verkan av snabba variationer i mätsignalen. Det följer nämligen av sats 5.2 att om gradtalet för polynomet $D(s)$ väljs mindre än grad $A(s)$, så kan gradtalet för polynomet $C(s)$ väljas godtyckligt stort. Detta utnyttjas t. ex. vid dimensionering av autopiloter för flygplan som är mycket böjliga. Då ett sådant flygplan utsätts för turbulens uppstår elastiska svängningar av höga frekvenser i vingar och flygkropp. Svängningarna registreras av gyroskop och accelerometrar. Om över-

föringsfunktionen från givarna till roderservon är sådana att svängningarna fortplantas till rodren så kan svängningarna förstärkas. Genom att välja F-polynomets gradtal tillräckligt högt så kan man se till att detta ej inträffar.

Godtycket vid val av polynomen $C(s)$, $D(s)$ och $F(s)$ kan också utnyttjas för att eliminera processtörningar. Vi nöjer oss med att belysa detta med ett exempel.

EXEMPEL 5.9

Betrakta motorn i exempel 5.1. Anta att ett störmoment verkar på motoraxeln. Störmomentet är praktiskt taget konstant och det antas att mätfelen är försumbara. Det krävs, liksom tidigare, att överföringsfunktionen från y^r till y skall vara den som angavs i exempel 5.6. Det följer av analysen i kapitel 4 att en integrator måste införas i systemet för att eliminera störningen. Detta är formellt likvärdigt med att multiplicera polynomet $A(s)$ med s. Det följer av sats 5.2 att en möjlig lösning ges av ett polynom $C(s)$ av första graden och ett polynom $D(s)$ av andra graden. Polynomet $F(s)$ kan väljas av första graden. Ansätt således

$$C(s) = s$$

$$D(s) = d_0 s^2 + d_1 s + d_2$$

$$F(s) = s + f_1$$

Ekvation (5.35) ger

$$(s^2 + as)s + \omega^2(d_0 s^2 + d_1 s + d_2) = (s + f_1)(s^2 + 2\zeta\omega s + \omega^2)$$

Identifikation av koefficienterna för olika potenser av s ger följande ekvationssystem:

$$a + \omega^2 d_0 = f_1 + 2\zeta\omega$$

$$\omega^2 d_1 = 2\zeta\omega f_1 + \omega^2$$

$$\omega^2 d_2 = f_1\omega^2$$

Lösningen är

$$d_0 = [f_1 + 2\zeta\omega - a]\omega^{-2}$$

$$d_1 = 1 + 2\zeta\omega^{-1} f_1$$

$$d_2 = f_1$$

där f_1 kan väljas fritt. Regulatorn blir således

$$U(s) = [1 + f_1/s]\, Y^r(s) - [d_0 s + d_1 + d_2/s]\, Y(s)$$

dvs. en PID regulator. Observera emellertid att då $d_1 \neq 1$ införs referensvärdet på något annat sätt än i en vanlig PID regulator. Jämför med exempel 5.6.

Om det dessutom finns högfrekvent mätbrus så kan i stället lösningar svarande mot de i exempel 5.7 och 5.8 bestämmas.

De beräkningar som gjordes i exemplen kan generaliseras. Ansätt polynomen $C(s)$ och $D(s)$ på formen

$$C(s) = c_0 s^k + c_1 s^{k-1} + \ldots + c_k$$
$$D(s) = d_0 s^l + d_1 s^{l-1} + \ldots + d_l$$

För att bestämma parametrarna i dessa polynom identifieras koefficienterna för potenser av s i ekvationen (5.35). Vi får då ett linjärt ekvationssystem med grad $C(s)$ + grad $D(s)$ + 2 obekanta. Antalet ekvationer är max [grad $A(s)$ + grad $C(s)$, grad $B(s)$ + grad $D(s)$] + 1.

Genom att kräva att antalet ekvationer skall vara lika med antalet obekanta erhålls följande kompatibilitetsvillkor

$$\begin{aligned} &\text{grad } C(s) \geqslant \text{grad } B(s) - 1 \\ &\text{grad } D(s) \leqslant \text{grad } A(s) - 1 \\ &\text{grad } H(s) \leqslant \text{grad } A(s) + \text{grad } C(s) \end{aligned} \tag{5.41}$$

eller

$$\begin{aligned} &\text{grad } C(s) \leqslant \text{grad } B(s) - 1 \\ &\text{grad } D(s) \geqslant \text{grad } A(s) - 1 \\ &\text{grad } H(s) \leqslant \text{grad } B(s) + \text{grad } D(s) \end{aligned} \tag{5.42}$$

vilka motsvarar fallen (i) och (ii) i sats 5.2.

Följande ekvationssystem erhålls då:

$$[c_0 c_1 \ldots c_k d_0 d_1 \ldots d_l]\, R(A, B) = [h_0 h_1 \ldots h_{k+l+1}] \tag{5.43}$$

I fallet $n + k = m + l$ ges matrisen $R(A, B)$ av

$$R(A,B) = \underbrace{\begin{bmatrix} a_0 & a_1 & \ldots & a_n & & \\ & a_0 & a_1 & \ldots & a_n & \\ & & \ddots & & & \\ & & & a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_m & & \\ & b_0 & b_1 & \ldots & b_m & \\ & & \ddots & & & \\ & & & b_0 & b_1 & \ldots & b_m \end{bmatrix}}_{\max[m+k+2,\, m+l+2]} \begin{matrix} \} k+1 \\ \\ \\ \\ \} l+1 \\ \\ \\ \end{matrix} \tag{5.44}$$

Element som ej markerats är noll. Det följer av villkoren (5.41) och (5.42) att matrisen är kvadratisk. I det allmänna fallet har matrisen $R(A, B)$ en liknande

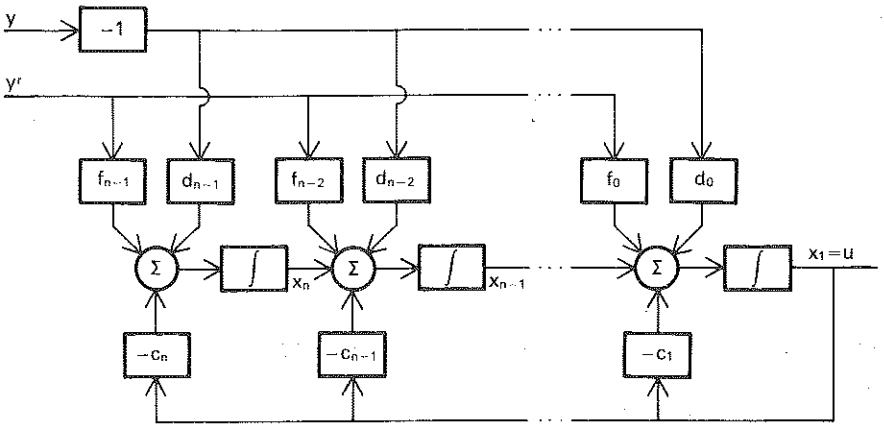


Fig. 5.4. Blockschema för regulatorn (5.45). Regulatorns insignaler är referensvärdet y^r och processens mätsignal y. Regulatorns utsignal är processens styrvariabel u.

uppbyggnad. Blocken som innehåller koefficienterna a_i och b_i kan emellertid vara förskjutna i förhållande till varandra. Om $n+k=m+l+i$ så uppträder t. ex. den första koefficienten b_0 i kolonn $i-1$. Matrisen $R(A, B)$ kallas *resultanten* för polynomen $A(s)$ och $B(s)$. Det följer av sats 5.2 att ett nödvändigt villkor för att $R(A, B)$ är inverterbar är att polynomen $A(s)$ och $B(s)$ saknar gemensamma faktorer. Observera att lösningen mycket väl kan vara sådan att c_0 eller d_0 är noll, vilket innebär att grad $C(s) <$ grad $B(s)-1$ i (5.42) respektive grad $D(s) <$ grad $A(s)-1$ i (5.41).

Den regulator som beskrivs av ekvation (5.30) kan lätt förverkligas med hjälp av återkopplade integratorer och approximativa inverser. Den regulator som motsvarar Kalmans rekonstruktion har $k=n$, $l=n-1$ och kan realiseras som

$$\frac{dx}{dt} = \begin{pmatrix} -c_1 & 1 & 0 \dots 0 \\ -c_2 & 0 & 1 \dots 0 \\ \vdots & & \\ -c_{n-1} & 0 & 0 \dots 1 \\ -c_n & 0 & 0 \dots 0 \end{pmatrix} x - \begin{pmatrix} d_0 \\ d_1 \\ \vdots \\ d_{n-2} \\ d_{n-1} \end{pmatrix} y + \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{pmatrix} y^r$$

$$u = [1 \ 0 \ 0 \dots 0]x \tag{5.45}$$

Ett blockschema för realiseringen visas i fig. 5.4.

Jämförelse mellan metoderna för polplacering

Två olika metoder som ger det återkopplade systemet föreskrivna poler har nu angivits. Utgående från interna representationer visades att det önskade resultatet kunde erhållas genom tillståndsåterkoppling i kombination med

rekonstruktion av tillståndsvariabler. Vi har också visat att den önskade återkopplingen kan bestämmas genom direkt manipulation av överföringsfunktioner utgående från en extern beskrivning av systemet. Beräkningstekniskt är metoderna likvärdiga. Båda kräver lösning av algebraiska ekvationer. Om man räknar för hand är det något lättare att direkt manipulera med överföringsfunktionerna. Om en representation av systemet väljs så att tillståndsvariablerna har goda fysikaliska tolkningar kan det vara fördelaktigt att arbeta med de interna beskrivningarna. Bland regulatorns interna variabler kan då de rekonstruerade variablerna återfinnas. Dessa kan ofta vara till nytta för andra funktioner i ett system. Ett typexempel är t. ex. autopiloter för flygplan och båtar där tillståndsvariablerna kan väljas så att de motsvarar farkostens hastighets- och vinkelhastighetskomponenter. Dessa variabler kan utnyttjas bl. a. för presentations- och navigationsfunktioner. Det finns dock andra fall då det ej är möjligt att finna tillståndsvariabler med naturlig fysikalisk tolkning. I ett sådant fall är det både bekvämt och estetiskt tilltalande att ha en syntesmetod som direkt utnyttjar insignal–utsignal-beskrivningar. Det är också intressant att notera hur störningar behandlas i de båda metoderna. Utgående från de interna beskrivningarna är det naturligt att införa störningar explicit genom att ta med dem i modellen för systemets omgivning på det sätt som gjordes i exempel 5.1. Detta resulterar i att man försöker att rekonstruera störningarna i observeraren. Om man i stället utgår från de externa beskrivningarna får man i stället ta hänsyn till störningarna genom att postulera att polynomen $C(s)$ och $D(s)$ skall ha vissa egenskaper på det sätt som gjordes i exempel 5.9.

Det är också av intresse att jämföra strukturen på de regulatorer som erhålls med de olika metoderna. Se t. ex. fig. 5.1, fig. 5.3 och fig. 5.4. Under förutsättningen grad $P(s) = \text{grad } A(s)$ och $B(s) = Q(s)$ krävs i båda fall samma antal integratorer motsvarande observeraren respektive polynomet $C(s)$. Regulatorstrukturen i fig. 5.1 är dock mer komplicerad än den i fig. 5.4. I gengäld är regulatorn i fig. 5.1 mer flexibel därigenom att de tre funktionerna, följning av insignaler, eliminering av störningar och polplacering, är separerade i följande mening. För systemet i fig. 5.1 bestäms överföringsfunktionens poler enbart av vektorn L. De påverkas ej av observeraren (vektorn K) eller av systeminversen. Förmågan att eliminera störningar påverkas av observeraren och förmågan att följa insignaler av systeminversen. För systemet i fig. 5.4, däremot, påverkas samtliga funktioner av variationer i parametrarna c_i och d_i. Genom att de olika funktionerna är separerade i regulatorn i fig. 5.1 är denna lättare att trimma.

Vad begränsar ett systems prestanda

Vi har visat att linjära tidsinvarianta system som är styrbara och observerbara kan återkopplas så att det slutna systemets poler antar föreskrivna värden. Nollställen som ligger i vänstra halvplanet kan elimineras och ett godtyckligt

antal nollställen kan tillfogas genom framkoppling. Nollställen i högra halvplanet, dvs. det faktum att systemet är icke-minimum-fas, är således en begränsning. Vi belyser denna begränsning med ett exempel.

EXEMPEL 5.10. Inverkan av nollställe i högra halvplanet

Betrakta ett styrbart system av första ordningen vars överföringsfunktion har ett nollställe $s=a$ i högra halvplanet. Då systemet är styrbart kan det alltid återkopplas så att dess pol antar ett givet värde, säg $s=-b$. Låt oss vidare anta att det slutna systemet skall kunna reproducera en stegsignal utan bestående fel. Det slutna systemets överföringsfunktion blir då

$$G(s)=\frac{b(a-s)}{a(s+b)}$$

där b kan ges ett godtyckligt värde. För att få ett snabbt system är det uppenbarligen bra att välja b stor. Detta leder emellertid till att systemet får andra ej önskade egenskaper. Betrakta t. ex. systemets stegsvar

$$y(t)=\mathcal{L}^{-1}\left[\frac{b(a-s)}{as(s+b)}\right]=1-\frac{a+b}{a}\,e^{-bt}$$

Det gäller således att

$$y(0)=-b/a,\quad y(\infty)=1$$

Systemets utsignal antar således i första ögonblicket ett negativt värde $y(0)=-b/a$ och växer sedan exponentiellt mot slutvärdet $y=1$ med tidskonstanten b^{-1}. Ett stort värde på b medför visserligen att systemet blir snabbt men den negativa överslängen blir samtidigt stor. Om vi vill hålla den negativa överslängen på ett rimligt värde måste kvoten b/a begränsas. Kräver vi att utsignalens storlek i första ögonblicket skall vara mindre än slutvärdet får vi t. ex. villkoret $b/a<1$. Vi finner således att det faktum att systemets överföringsfunktion har ett nollställe $s=a$ i högra halvplanet begränsar det återkopplade systemets snabbhet. Det återkopplade systemets stigtid och bandbredd blir av storleksordningen a^{-1}.

Observera att vattenturbinen som behandlas i exempel 2:9.2 ej är ett minimum-fas-system. Jämför överföringsfunktionen för en vattenturbin som ges av ekvation (2:9.8). Vid reglering av vattenturbiner kan man således ej få ett system vars stigtid väsentligen understiger anloppsrörets tidskonstant T.

Förhållandena är likartade i det allmänna fallet. I fig. 5.5 visas t. ex. stegsvaret för ett andra ordningens system som ej är minimum-fas.

Om ett system ej är minimum-fas kan man försöka att införa fler mätgivare och fler styrvariabler, flytta givare eller modifiera processen så att den blir ett minimum-fas-system. Jämför exempel 3.1 och 4.6. Bl. a. av detta skäl är det

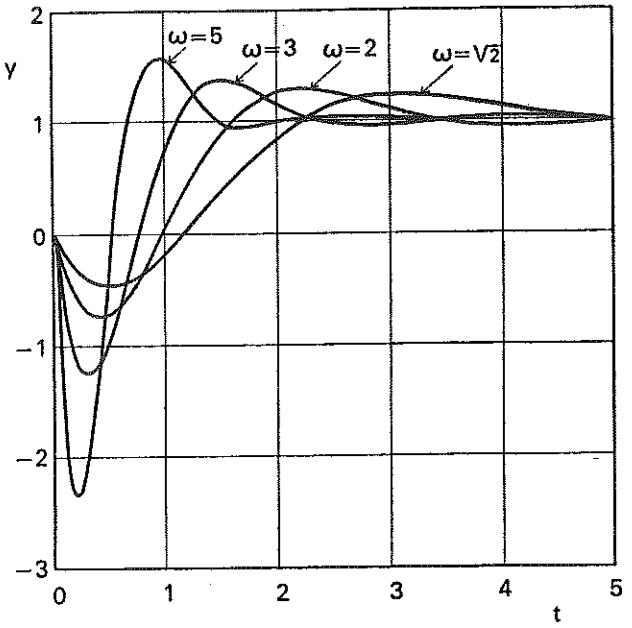


Fig. 5.5. Stegsvar för icke minimum-fas-system av andra ordningen. Systemet har överföringsfunktionen

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

önskvärt att om möjligt göra reglertekniska analyser redan på projektstadiet då det är lättare att göra ändringar i processen.

I realiteten finns också andra viktiga begränsningar, vilka beror av mätfel, störningar och olinjäriteter i processen. Den analys som genomförts baseras på antagandet att processen kan beskrivas med en linjär modell. I de flesta fall är den linjära modellen giltig endast för vissa signalnivåer och inom vissa frekvensområden. Det är således mycket viktigt att kontrollera att de signaler som uppträder i det återkopplade systemet är sådana att linjäritetsantagandena är uppfyllda. Det är t. ex. alltid värdefullt att undersöka den insignal till processen som genereras av regulatorn i olika situationer.

Det gäller som regel att det fordras höga förstärkningar för att erhålla system med stor snabbhet. För systemet i exempel 5.1 fann vi t. ex. att förstärkningen K var proportionell mot ω^2, dvs. kvadraten på det slutna systemets egenfrekvens. En fördubbling av kraven på bandbredd leder således till att förstärkningen måste kvadreras. Höga förstärkningar medför att mycket obetydliga reglerfel ger stora styrsignaler. Detta kan leda till slitage av ställdon och ej önskvärda materialpåkänningar.

6. Kompensering

Föregående avsnitt behandlade syntesmetoder som kunde användas för att erhålla ett system där det slutna systemets poler antog föreskrivna värden. Det kan i vissa fall vara onödigt restriktivt att föreskriva läget på det slutna systemets samtliga poler. Man kan då också nöja sig med en enklare regulator som dimensioneras med s. k. kompensering. Kompensering utgjorde stommen i den klassiska teorin för servomekanismer. Metoder för kompensering utvecklades ur försök att modifiera existerande system som ej uppfyller specifikationerna. Sådana problem uppträder naturligt om man försöker att återkoppla ett system med en proportionell regulator, såsom visades i avsnitt 4:1, men ej erhåller ett system med tillfredsställande egenskaper. Då ett system kan avvika från specifikationer på en mångfald olika sätt är det ej praktiskt att tabellera alla tänkbara fall och i tur och ordning diskutera dessa. I stället skall vi välja ut några vanligt förekommande fall och diskutera hur kompensering skall genomföras i dessa situationer. Följande typfall skall studeras:

- Systemets snabbhet är tillfredsställande men systemet har stora stationära fel vid följning av rampsignaler och liten styvhet för yttre störningar. (Stegsvaret är således bra. Specifikationerna på insvängningstid, stigtid och översläng, bandbredd och resonanstopp är uppfyllda men felkoefficienterna e_i och styvhetskoefficienterna c_i är för stora.)
- Systemet är instabilt eller systemet är stabilt men stegsvaret har för stor översläng. Alternativt är resonanstoppen för hög.
- Systemet är stabilt men stigtiden och felkoefficienterna är för stora.

Vi skall nu undersöka hur systemet skall modifieras för att dessa brister skall avhjälpas. Kompenseringen består vanligen i att systemet tillförs länkar, vars förstärkning är frekvensberoende. Kompenseringslänkarna kan vara passiva filter, återkopplade operationsförstärkare, fjädrar, viskösa dämpcylindrar, hävstänger etc.

Kompenseringslänkarna kan införas på många ställen i systemet, t. ex. efter givaren som ger reglerfelet (kaskad- eller seriekompensering), i återkopplingskretsen (parallellkompensering) eller genom lokal återkoppling av enstaka element. Se fig. 6.1, 6.2 och 6.3.

I detta avsnitt skall vi uteslutande behandla servoproblemet, dvs. det fall då de okompenserade systemen är enkla återkopplade system där utsignalen skall reproducera insignalen. Analysen inskränks till kaskadkompensering. Med hänvisning till fig. 6.1 skall vi således bestämma överföringsfunktionen G_k så att det slutna systemet får givna egenskaper. Observera att dimensioneringsmetoderna uteslutande utnyttjar återkoppling. Det kan därför ofta vara fördelaktigt att komplettera lösningen med framkoppling såsom diskuterats i avsnitten 3 och 4. Två kompenseringsmetoder: Nyqvist–Bode–Nichols och rotortmetoden skall behandlas. Implementering av kompenseringsnät diskuteras också kortfattat.

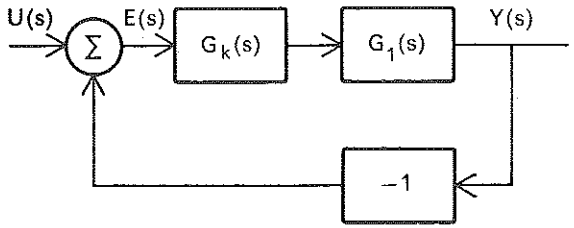


Fig. 6.1. Serie- eller kaskadkompensering. Kompenseringslänken $G_k(s)$ inkopplas efter felsignalgivaren, dvs. i serie med de givna länkarna $G_1(s)$.

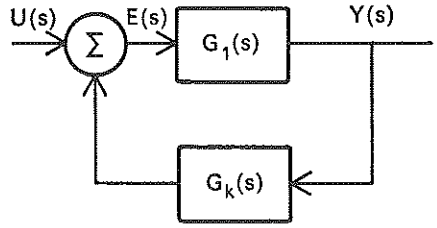


Fig. 6.2. Parallellkompensering eller kompensering i återkopplingskretsen. Kompenseringslänken $G_k(s)$ inkopplas i återkopplingskretsen.

Nyquist–Bode–Nichols kompenseringsmetod

Metoden förutsätter att specifikationerna ges som:

- Fasmarginal φ_m och skärningsfrekvens ω_c alternativt amplitudmarginal A_m och frekvens då faskurvan skär linjen $\arg G_0(i\omega) = -180^0$.
- Den första icke försvinnande felkoefficienten e_k.
- De givna länkarnas överföringsfunktion antingen i form av ett algebraiskt uttryck såsom

$$G_0(s) = \frac{k_n}{s^n} \cdot \frac{1 + b_1 s + \ldots + b_m s^m}{1 + a_1 s + \ldots + a_k s^k}$$

eller som belopp- och argumentkurvor i ett Bodediagram.

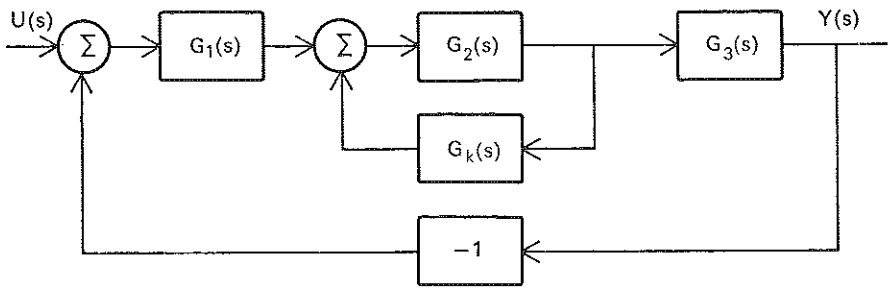


Fig. 6.3. Kompensering genom lokal återkoppling. En parallellkompensering av en del av systemet tillgrips ofta för att eliminera inverkan av parametervariationer i överföringsfunktionen $G_2(s)$.

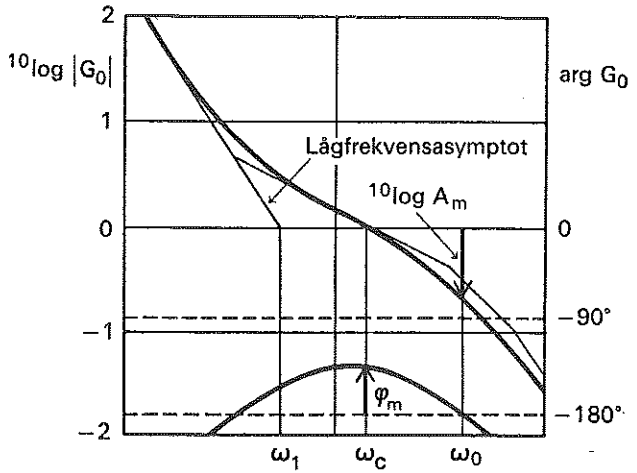


Fig. 6.4. Bodediagram för det öppna systemets överföringsfunktion.

Metoden baseras på grafisk konstruktion med hjälp av Nyquist–Bode–Nicholsdiagrammen, som behandlats i avsnitt 2:9 och 5:8. För att använda metoden är det väsentligt att man i Bodediagrammet för det öppna systemets överföringsfunktion kan utläsa systemspecifikationerna, ω_c, φ_m, A_m och e_k (fig. 6.4).

Den första icke försvinnande felkoefficienten ges av lågfrekvensasymptotens lutning $-n$ och lågfrekvensasymptotens skärning med axeln $|G_0(i\omega)| = 1$. Det gäller

$$e_i = 0 \quad i = 0, 1, \ldots, n-1 \quad n \geqslant 1$$

$$e_n = \omega_1^{-n}$$

Vidare kan förstärknings- och fasmarginaler direkt utläsas ur Bodediagrammet. Jämför fig. 6.4. Det slutna systemets bandbredd och resonanstopp M_p kan approximativt uppskattas ur förstärkningsmarginal och fasmarginal. För att uppskatta dessa kvantiteter noggrannare kan Nicholsdiagrammet användas. Det är vanligt att man försöker att ge specifikationerna direkt i form av krav på kretsöverföringen $G_k G_1$, t. ex. i den form som visas i fig. 6.5. Kraven vid låga frekvenser kan t. ex. bero på specifikationer på felkoefficienter och styvhetskoefficienter. Kraven vid höga frekvenser kan t. ex. vara betingade av specifikationer som kräver att snabba variationer i processens mätsignaler ej skall generera styrsignaler till processen. Figuren antyder de svårigheter som kan uppstå. Det gäller att få kretsöverföringen att avta tillräckligt brant. För att få ett system med tillräcklig dämpning så måste amplitudkurvans lutning dock vara nära -1 i närheten av skärningsfrekvensen.

Vi kan nu inse hur olika kaskadkompenseringar påverkar det slutna systemets egenskaper.

- En höjning av förstärkningen vid låga frekvenser medför att felkoefficienterna minskar och eventuellt försvinner.

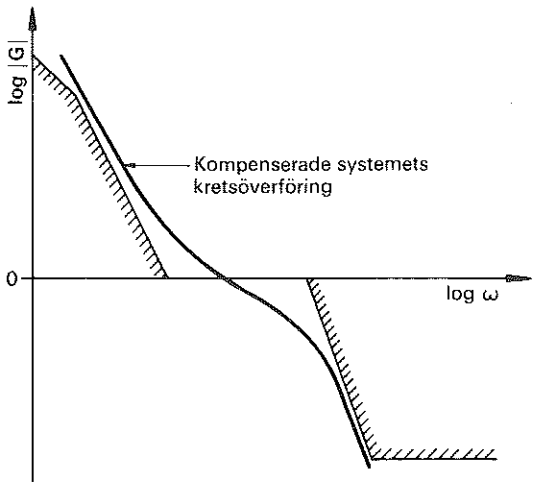


Fig. 6.5. Illustrerar typiska krav på kretsöverföringen $-G_1 G_k$. Jämför fig. 6.4.

- En höjning av förstärkningen med ett minimum-fas-nät i närheten av ω_c medför att bandbredden ökar (stigtiden minskar).
- En minskning av fasförskjutningen för frekvenser i närheten av ω_c ökar fasmarginalen och förbättrar därigenom stabiliteten.

Med ledning av detta kan vi nu avgöra hur kaskadkompenseringen skall väljas. Man kan t. ex. följa nedanstående procedur:

- Justera det öppna systemets förstärkning så att skärningsfrekvensen överensstämmer med specifikationen.
- Om fasmarginalen ej är tillfredsställande införs ett fasavancerande nät, som ökar fasmarginalen. Justera om så behövs kretsförstärkningen så att skärningsfrekvensen överensstämmer med specifikationen.
- Bestäm med ledning av specifikationerna på felkoefficienterna lågfrekvensasymptotens lutning och dess (eventuellt extrapolerade) skärning med axeln $|G_0(i\omega)| = 1$. Ange ett nät som ändrar det givna systemets förstärkning vid låga frekvenser så att den överensstämmer med den önskade lågfrekvensasymptoten. Placera den högfrekventa brytpunkten för detta nät så att fasmarginalen ej sänks alltför mycket.

Metodens största fördel är att den är enkel att tillämpa även i de fall då de givna länkarnas överföringsfunktion är mycket komplicerad. Det går t. ex. utmärkt att behandla system med tidsfördröjningar eller system med överföringsfunktioner av den typ som ges av ekvation (2:8.12). Metodens främsta nackdel är att det kan vara svårt att uppskatta det slutna systemets dämpning ur fasmarginal och amplitudmarginal. Man får därför ofta komplettera med att undersöka systemets stegsvar med hjälp av analogimaskin eller dator. Vi skall nu ge några exempel som illustrerar metoden.

EXEMPEL 6.1. Minskning av felkoefficienterna med bibehållande av systemets transienta egenskaper

Betrakta ett system där de givna länkarna har överföringsfunktionen

$$G_1(s)=\frac{K}{s(s+1)(s+2)} \tag{6.1}$$

Det återkopplade systemet har analyserats i tidigare avsnitt. Systemets stegsvar för olika värden på förstärkningsparametern K framgår av fig. 5:5.3. I fig. 5:8.4 visas Bodediagrammet för kretsöverföringen med $K=1$ och i fig. 5:8.8 visas Nicholsdiagrammet för systemet. Den tidigare analysen har visat att $K=1$ är ett rimligt värde på förstärkningsparametern. Detta ger en skärningsfrekvens $\omega_c=0{,}5$ rad·s^{-1} och en fasmarginal $\varphi_m=54°$. Jämför fig. 5:8.4 och fig. 6.6. Med $K=1$ erhålls följande värden på felkoefficienterna

$e_0=0$

$e_1=2/K=2$ s

Systemet kan således reproducera en stegfunktion utan stationärt fel. Däremot blir det stationära följfelet $2v$ vid insignalen $u_r=vt$. Anta nu att detta är 10 gånger för stort, dvs. att specifikationen är

$e_0=0$

$e_1=0{,}2$ s

men att systemets egenskaper i övrigt är acceptabla. Vi skall ange en kaskadkompensering sådan att specifikationen kan uppfyllas.

Om felkoefficienten e_1 skall minskas från 2,0 till 0,2, måste det öppna systemets förstärkning vid låga frekvenser ökas med en faktor 10. Lågfrekvensasymptoten i Bodediagrammet skall således höjas med $^{10}\log 10$. Se fig. 6.6.

Den önskade höjningen av lågfrekvensasymptoten kan åstadkommas genom kaskadkompensering med ett system vars överföringsfunktion är

$$G_k(s)=\frac{s+a}{s+a/M} \tag{6.2}$$

Denna överföringsfunktion har förstärkningen M vid låga frekvenser ($\omega \ll a/M$) och förstärkningen 1 vid höga frekvenser ($\omega \gg a$). För att minska e_1 med en faktor 10 bör vi således välja $M=10$. Genom att välja a tillräckligt litet kommer värdena av det öppna systemets överföringsfunktion i närheten av skärningsfrekvensen endast att påverkas obetydligt. Det gäller

$$|G_k(i\omega)|=\sqrt{\frac{1+(a/\omega)^2}{1+(a/(\omega M))^2}} \tag{6.3}$$

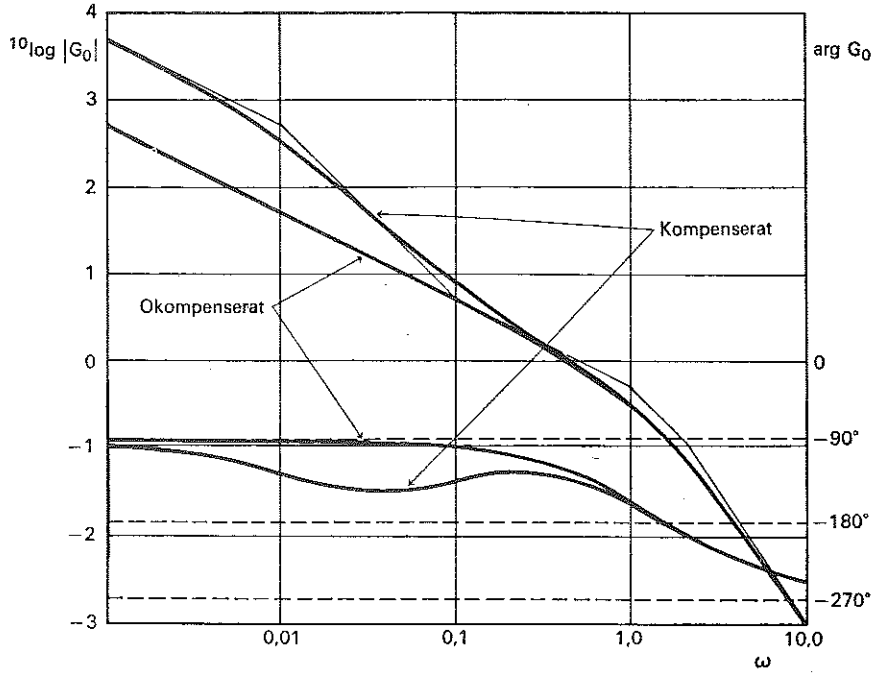


Fig. 6.6. Bodediagram för enkla återkopplade system med kretsöverföringarna

$$G_0(s) = \frac{1}{s(s+1)(s+2)} \quad \text{(okompenserade systemet)}$$

och

$$G_0(s) = \frac{s+0{,}1}{s(s+0{,}01)(s+1)(s+2)} \quad \text{(kompenserade systemet)}$$

$$\arg G_{lc}(i\omega) = \arctan(\omega/a) - \arctan(\omega M/a) = -\arctan\frac{\omega(M-1)}{a(1+\omega^2 M/a^2)} \tag{6.4}$$

Genom att parametern a väljs tillräckligt liten, kan kompenseringsnätets inverkan på kretsöverföringens värden i närheten av skärningsfrekvensen göras godtyckligt liten. I fig. 6.7 visas det slutna systemets rampfunktionssvar med $M = 10$ och några olika värden på a. Observera att det stationära följfelet för en enhetsramp alltid är 0,2 för alla värden på parametern a. Fig. 6.7 har erhållits genom simulering på analogimaskin. Med ökande värden på a minskar följfelet snabbare i början men systemet blir mer oscillativt.

För att ytterligare belysa inverkan av parametern a visar vi i fig. 6.8 det kompenserade systemets stegsvar för några olika a-värden.

Det framgår av fig. 6.8 att med $a = 0{,}1$ erhålls en översläng på 35 % jämfört med 15 % för det okompenserade systemet. Vi väljer tämligen godtyckligt $a = 0{,}1$.

Kompenseringsnätets överföringsfunktion blir då

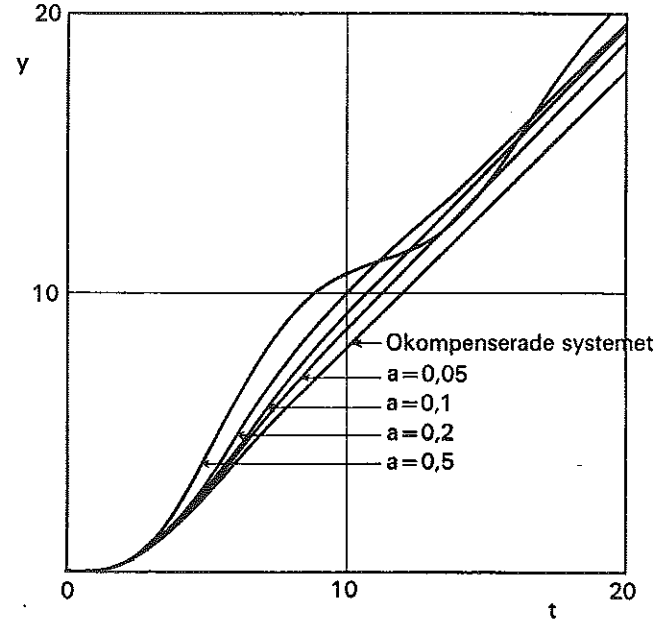


Fig. 6.7. Rampfunktionssvar för ett enkelt återkopplat system med kretsöverföringen

$$G_0(s) = \frac{s+a}{s(s+0{,}1a)\,(s+1)\,(s+2)}$$

Med ökande värden på parametern a minskar följfelet snabbare i början, men systemet blir även mer oscillativt.

$$G_k(s) = \frac{s+0{,}1}{s+0{,}01} \tag{6.5}$$

och det kompenserade systemet får kretsöverföringen

$$G_1(s)\,G_k(s) = \frac{(s+0{,}1)}{s(s+0{,}01)(s+1)(s+2)} = \frac{5(1+10s)}{s(1+100s)(1+s)(1+0{,}5s)} \tag{6.6}$$

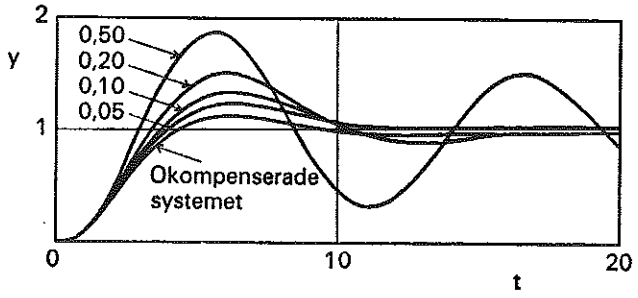


Fig. 6.8. Stegfunktionssvaret för det kompenserade systemet för olika värden på parametern a.

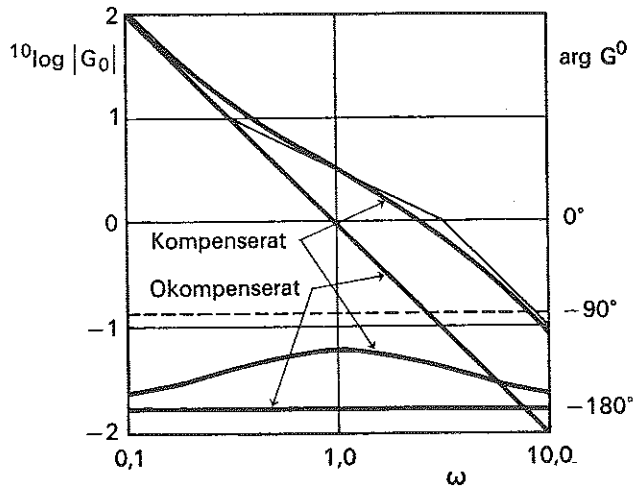


Fig. 6.9. Bodediagram för enkla återkopplade system med kretsöverföringarna $G_0(s) = 1/s^2$ (okompenserade systemet)

och

$$G_0(s) = \frac{10(s+0{,}32)}{s^2(s+3{,}2)} \quad \text{(kompenserade systemet)}$$

Bodediagrammet för denna överföringsfunktion framgår av fig. 6.6. Observera att bortsett från valet av parametern a hade kompenseringsnätets överföringsfunktion kunnat bestämmas enbart med hjälp av grafiska konstruktioner i Bodediagrammet. En ofta använd tumregel för bestämning av a är att välja $a = 0{,}1\ \omega_c$ där ω_c är det okompenserade systemets skärningsfrekvens. För ett kompenseringsnät av typen (6.2) med $M > 10$ innebär denna tumregel att skärningsfrekvensen ej ändras märkbart och att fasmarginalen minskar med ca 5,7°. Jämför ekvationerna (6.3) och (6.4).

Det kompenserade systemet får då ungefär samma bandbredd och samma resonanstopp som det okompenserade systemet. De »transienta egenskaperna» kommer således att bibehållas. Tillämpas tumregeln på vårt exempel erhålls $a = 0{,}05$ i stället för $a = 0{,}1$. Detta ger ett system med mindre översläng och långsammare insvängning av rampsvaret. Jämför fig. 6.7 och 6.8.

EXEMPEL 6.2. Stabilisering av instabilt system

Betrakta ett system där de givna länkarna har överföringsfunktionen

$$G_1(s) = \frac{K}{s^2} \tag{6.7}$$

Anta att specifikationerna kräver skärningsfrekvensen 1 rad·s^{-1} och fasmarginalen 55°. Vi följer den procedur som presenterades inledningsvis och bestämmer förstärkningsparametern K så att den önskade skärningsfrekven-

sen erhålls. Vi finner då det nominella värdet $K=1$. Jämför Bodediagrammet i fig. 6.9. Fasmarginalen är noll och det slutna systemet är således instabilt.

För att erhålla ett system med den önskade fasmarginalen 55°, måste vi alltså införa ett kaskadkompenseringsnät som minskar fasförskjutningen i närheten av skärningsfrekvensen. Detta kan uppnås med ett system med överföringsfunktionen

$$G_k(s)=\frac{a+s}{a+s/N}=N\ \frac{s+a}{s+aN} \tag{6.8}$$

Det gäller att

$$\arg G_k(i\omega)=\arctan\frac{\omega(1-1/N)}{a(1+\omega^2/(a^2N))} \tag{6.9}$$

Uttrycket (6.9) är alltid positivt och har maximum för $\omega=a\sqrt{N}$

$$\max_{\omega}[\arg G_k(i\omega)]=\arctan\frac{1}{2}\,(\sqrt{N}-1/\sqrt{N})$$

För att erhålla en fasmarginal på 55° väljer vi N så att

$$\frac{1}{2}\,(\sqrt{N}-1/\sqrt{N})=\tan 55^\circ$$

Detta ger $N=10$. Vidare väljer vi a så att den maximala fasförskjutningen uppnås för $\omega=1$. Detta ger $a=1/\sqrt{10}=0{,}32$. Kompenseringsnätets överföringsfunktion har beloppet

$$|G_k(i\omega)|=N\sqrt{\frac{\omega^2+a^2}{\omega^2+a^2N^2}}$$

För den önskade skärningsfrekvensen $\omega=\omega_c=1$ erhålls då

$$|G_k(i)|=\sqrt{N}=\sqrt{10}$$

Vidare gäller

$$|G_0(i\omega)|=|G_k(i\omega)|\,|G_1(i\omega)|=K|G_k(i\omega)|/\omega^2$$

Med det nominella värdet på kretsförstärkningen $K=1$ blir skärningsfrekvensen ω_c för hög (fig. 6.9). För att uppfylla specifikationen på skärningsfrekvensen modifieras K. Vi finner

$$1=|G_0(i\omega_c)|=K\sqrt{10}$$

dvs.

$$K=1/\sqrt{10}=0{,}32$$

Vi finner således att kompenseringsnätet skall ha överföringsfunktionen

$$G_k(s)=10\,\frac{s+1/\sqrt{10}}{s+\sqrt{10}}=10\,\frac{s+0{,}32}{s+3{,}2} \tag{6.10}$$

Det kompenserade systemets kretsöverföring blir

$$G_0(s) = G_k(s)\,G_1(s) = 10\,\frac{s+0{,}32}{s+3{,}2}\,\frac{0{,}32}{s^2}\,\frac{3{,}2}{s^2(s+3{,}2)} \tag{6.11}$$

Sammanfattning

De resonemang som genomförts i exemplen kan tillämpas på många andra system. Vi sammanfattar här egenskaperna hos fasretarderande och fasavancerande kompensering liksom tumregler för val av kompenseringsnätens parametrar.

Felkoefficienter och styvhetskoefficienter kan påverkas med fasretarderande kompensering. Kompenseringsnätet har överföringsfunktionen

$$G_k(s) = \frac{s+a}{s+a/M}$$

Med denna kompensering minskas den första felkoefficient, som är skild från noll, med en faktor M. Den fasretarderande kompenseringen ger en minskning av fasmarginalen som ges av ekv. (6.4). Talet a/ω_c, där ω_c är skärningsfrekvensen, bör vara litet för att dämpningen ej skall minska alltför mycket. Ett rimligt värde är $a \approx 0{,}1\ \omega_c$. Observera att kompenseringsnätet övergår i en PI-regulator då $M \to \infty$. Om man vill påverka flera felkoefficienter så får flera fasretarderande nät kaskadkopplas.

För att öka snabbheten eller bandbredden för ett system ökas först förstärkningen så att den önskade skärningsfrekvensen erhålls. Systemet blir då i allmänhet instabilt eller dåligt dämpat. Dämpningen kan sedan ökas med hjälp av en fasretarderande kompensering. Kompenseringsnätet har överföringsfunktionen

$$G_k(s) = N\,\frac{s+b}{s+bN}$$

Med denna kompensering erhålls en ökning av fasmarginalen med högst arctan $\frac{1}{2}(\sqrt{N} - 1/\sqrt{N})$ vid frekvensen $b\sqrt{N}$. Talet N väljs så att önskad ökning i fasmarginalen erhålls och parametern b väljs så att $b\sqrt{N} \approx \omega_c$. Systemets förstärkning och talet b får sedan justeras så att systemets stegsvar blir bra. Observera att kompenseringsnätet övergår i en PD-regulator då $N \to \infty$. Om den önskade ökningen i fasmarginalen ej kan uppnås så får flera fasavancerande nät kaskadkopplas. Det är svårare att välja parametrarna i ett fasavancerande än i ett fasretarderande nät och man får ofta göra justeringar för att få önskat resultat.

Vid kompensering med frekvenskurvor kan man ha stor hjälp av interaktiva datorprogram som kan rita Bodediagram och stegsvar på en bildskärm.

Kompensering med rotortmetoden

Rotorten infördes i avsnitt 5:5 som geometriska orten för det slutna systemets rötter då någon parameter i systemet varierar. Rotortkurvan ger en mycket åskådlig bild av hur det slutna systemets poler varierar då systemets parametrar ändras. Detta kan utnyttjas för att bestämma en lämplig kompensering av ett givet system. Metoden förutsätter att systemets specifikationer är givna på följande sätt.

- Systemets snabbhet är specificerad med stigtid, bandbredd eller alternativt med kontrollpolernas avstånd från origo och deras dämpning.
- Systemets statiska fel är specificerat med felkoefficienterna.
- De givna delarnas överföringsfunktion är given i form av en rationell funktion

$$G_1(s) = \frac{K}{s^n} \frac{(s+z_1)\dots(s+z_m)}{(s+p_1)\dots(s+p_l)}$$

Metoden utnyttjar det faktum att man med rotorten får en direkt representation av det slutna systemets överföringsfunktion.

Betrakta ett enkelt återkopplat system. Låt kretsöverföringen vara

$$G_0(s) = KQ_0(s)/P_0(s) \tag{6.12}$$

Det slutna systemet har överföringsfunktionen

$$G(s) = \frac{KQ_0(s)}{P_0(s) + KQ_0(s)} \tag{6.13}$$

Det slutna systemets nollställen överensstämmer således med det öppna systemets nollställen och det slutna systemets poler erhålls direkt ur rotorten. Då systemets rotort är given, är således det slutna systemets poler och nollställen explicit givna. Med hjälp av rotorten har vi alltså goda möjligheter att överblicka specifikationer givna som villkor på det slutna systemets överföringsfunktion. Man utnyttjar härvid de samband som härleddes i avsnitt 2, med vars hjälp felkoefficienterna kan beräknas direkt ur det slutna systemets poler och nollställen. Se ekvationerna (2.4) och (2.5). Systemets snabbhet bestäms väsentligen av kontrollpolernas avstånd från origo, ω_k, och deras dämpning ζ. Snabbheten påverkas även av nollställen, vars avstånd från origo är mindre än ω_k. För att göra noggrannare utsagor om snabbheten, brukar man utnyttja tabeller som visar bandbredd och procentuell översläng för överföringsfunktioner av låg ordning eller simuleringar av systemet. Vid kaskadkompensering med rotortmetoden används följande procedur.

- Låt e_k beteckna den första icke försvinnande felkoefficienten. Kompenseringslänkens överföringsfunktion $G_k(s)$ skall då ha $k-n$ poler i origo.
- Skissa rotortkurvan med avseende på förstärkningsparametern K för ett system med kretsöverföringen

$$\frac{K}{s^k}\frac{(s+z_1)\dots(s+z_m)}{(s+p_1)\dots(s+p_l)} \tag{6.14}$$

- Inför extra nollställen på reella axeln så att det slutna systemet blir stabilt och så att specifikationen på icke försvinnande felkoefficienten uppfylls. Inför eventuellt även för varje nollställe en motsvarande pol på stort avstånd från origo.
- Skissa rotortkurvan med avseende på förstärkningsparametern och kontrollera om specifikationerna på snabbhet kan uppfyllas för något värde på förstärkningsparametern. Om detta ej går, modifieras lägena hos de poler och nollställen som införts i föregående steg.

Rotortmetodens stora förtjänst är att den ger det slutna systemets poler explicit. Metodens nackdel är att en noggrann bestämning av rotorten kan vara tidsödande speciellt för system av hög ordning. Ganska grova skisser av rotortkurvorna brukar emellertid ofta ge en god insyn i hur systemets egenskaper påverkas genom kompenseringen. I stället för att rita rotorten exakt, kan man beräkna systemets stegsvar och rampsvar t.ex. med hjälp av analogimaskin. Rotorten kan också lätt ritas med hjälp av dator. Vi skall nu illustrera rotortmetoden med några exempel.

EXEMPEL 6.3. Stabilisering av instabilt system

Betrakta ett system där de givna länkarna har överföringsfunktionen

$$G_1(s) = \frac{K}{s^2} \tag{6.15}$$

Bestäm ett kompenseringsnät så att det kompenserade systemets stegsvar svängt in sig till 5 % på 10 sek (lösningstiden) och så att överslängen är mindre än 25 %.

Om systemet återkopplas på vanligt sätt får det slutna systemet följande karakteristiska ekvation

$$s^2 + K = 0$$

Det slutna systemet är således instabilt för samtliga värden på kretsförstärkningen.

För att få ett stabilt system måste vi införa en kompensering sådan att det kompenserade systemets poler i ligger i vänstra halvplanet. Detta kan uppnås genom att ett nollställe införs på reella axeln (inses t. ex. med hjälp av potentialanalogin). Kompenseringslänken skall således ha överföringsfunktionen

$$G_k(s) = (s+a)/a \tag{6.16}$$

Det kompenserade systemet har kretsöverföringen

$$G_1(s)\,G_k(s) = \frac{K(s+a)}{as^2} \tag{6.17}$$

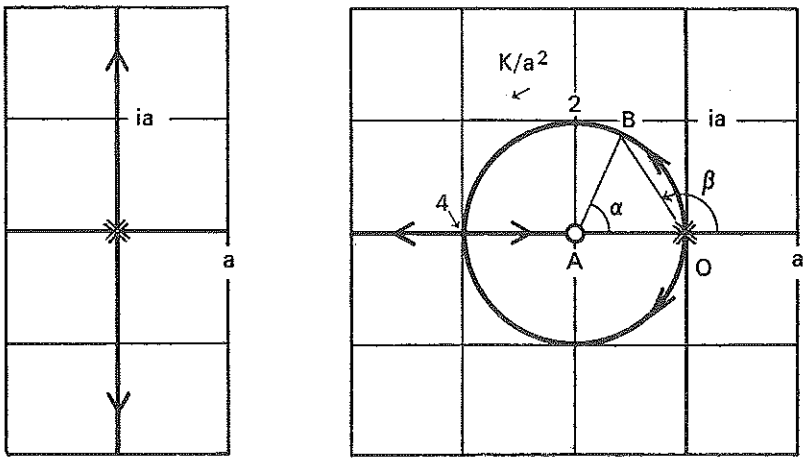


Fig. 6.10. Rotorterna för karakteristiska ekvationerna för enkla återkopplade system med kretsöverföringarna

$G_0(s) = \frac{K}{s^2}$ (okompenserade systemet)

och

$G_0(s) = \frac{K(s+a)}{a\,s^2}$ (kompenserade systemet)

Den karakteristiska ekvationen för det slutna systemet är

$$as^2 + K(s+a) = 0 \tag{6.18}$$

I fig. 6.10 visas rotorterna för kompenserade och okompenserade systemen. Det kompenserade systemets rotort utgörs till en del av en cirkel. För att visa detta införs beteckningar enligt fig. 6.10. Argumentvillkoret ger $\alpha - 2\beta = -\pi$. Vidare är $\wedge ABO = \beta - \alpha = \pi - \beta = \wedge AOB$. Sträckorna OA och AB är då lika.

För att fastlägga värdena på a och K förfar vi på följande (godtyckliga) sätt. Vi bestämmer först förstärkningsparametern K så att kontrollpolerna får en relativ dämpning $\zeta = 0{,}707$. Vi får då $p_1 = -a + ia$. Förstärkningsparametern K bestäms ur amplitudvillkoret, se ekvation (5:5.10). Om kontrollpolen p_1 skall ligga på rotorten erhålls ur (6.18)

$$K = a|p_1|^2/|p_1 + a| = 2a^2$$

Vi kontrollerar värdet genom att bestämma det kompenserade systemets stegsvar (fig. 6.11).

Vi finner att $K = 2a^2$ ger en översläng på 20 % vilket är acceptabelt. Då den karakteristiska ekvationens rötter har realdelen $-a$ uppskattas lösningstiden T_s ur ekvationen

$$\exp(-aT_s) = 0{,}05$$

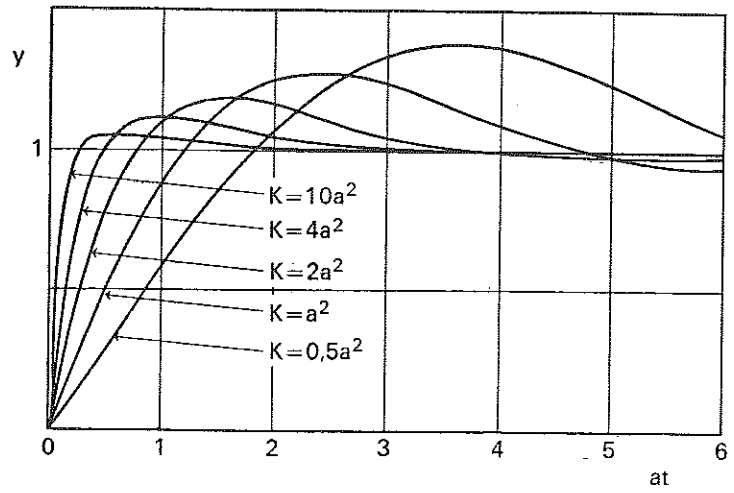


Fig. 6.11. Stegfunktionssvaret för enkelt återkopplat system med kretsöverföringen

$$G_0(s) = \frac{K(s+a)}{as^2}$$

Med $T_s = 10$ sek erhålls $a = 0{,}33$ rad/sek. En kontroll med stegsvaret i fig. 6.11 verifierar uppskattningen.

Kompenseringsnätet (6.16) motsvarar en proportionell och deriverande (PD) kompensering. Införandet av extra nollställen i det öppna systemets överföringsfunktion motsvarar således deriverande kompensering. Överföringsfunktionen (6.16) innebär att höga frekvenser förstärks mycket kraftigt. Om vi vill ha en kompenseringslänk, som är lättare att realisera praktiskt, kan vi välja

$$G_k(s) = \frac{a+s}{a+s/N} = \frac{N(s+a)}{s+aN} \tag{6.19}$$

Om N väljs stort verkar det troligt att systemen med kompenseringsnäten (6.16) och (6.19) skall ha samma egenskaper. I fig. 6.13 visas stegsvaren för $N = 10$. Observera att systemen med kompenseringsnäten (6.16) respektive (6.19) skiljer sig vid höga värden på förstärkningsparametern. Detta framgår t. ex. av rotortkurvan i fig. 6.12 och stegfunktionssvaret i fig. 6.13. Vi finner i fig. 6.13 att kompenseringen (6.19) ger ett acceptabelt system om $K = 2a^2$. Överslängen är 25 % och med $a = 0{,}33$ uppfylls specifikationen på lösningstiden.

EXEMPEL 6.4. Minskning av felkoefficienterna med bibehållande av systemets transienta egenskaper

Betrakta samma system som i exempel 6.1. Anta liksom i detta exempel att stegfunktionssvaret är tillfredsställande men kräv i detta fall att det stationära felet vid följning av en linjärt växande insignal skall vara noll. Det krävs

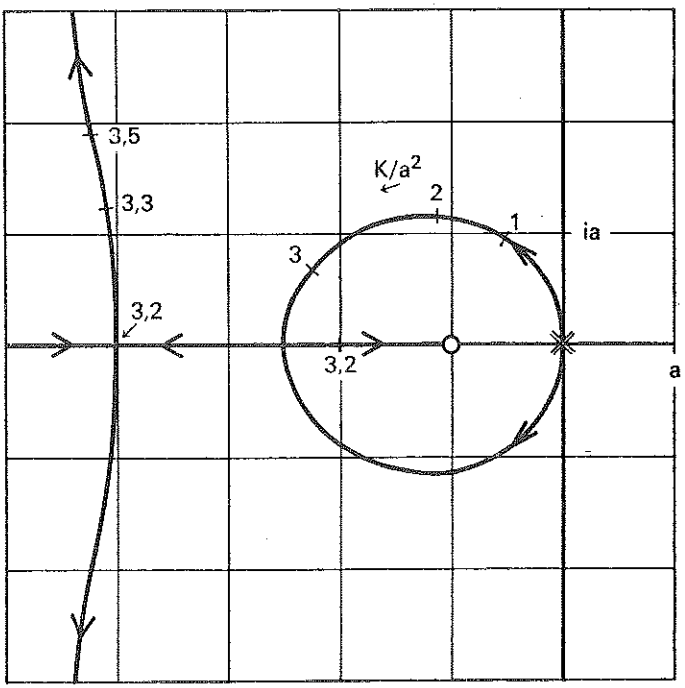


Fig. 6.12. Rotort för karakteristiska ekvationen för enkelt slutet system med kretsöverföringen

$$G_0(s) = \frac{10K(s+a)}{s^2(s+10a)}$$

Jämför fig. 6.10

således att de två första felkoefficienterna försvinner, dvs. $e_0 = e_1 = 0$. Rotortkurvan för det okompenserade systemet framgår av fig. 5:5.2. För att klara kraven på felkoefficienterna måste det öppna systemet förses med ytterligare en integration. Som ett första försök till kompensering erhålls således

$$G_k'(s) = 1/s$$

Det kompenserade systemets kretsöverföring är då

$$G_k'(s)\, G_1(s) = \frac{K}{s^2(s+1)(s+2)}$$

Rotortkurvan för systemet framgår av fig. 6.14.

Som synes är systemet instabilt för alla värden på förstärkningsparametern. För att få ett stabilt system, måste vi tvinga in rotorten i vänstra halvplanet. För att göra detta inför vi ett nollställe på reella axeln, dvs. vi försöker med kompenseringen

$$G_k(s) = (s+a)/s \qquad (6.20)$$

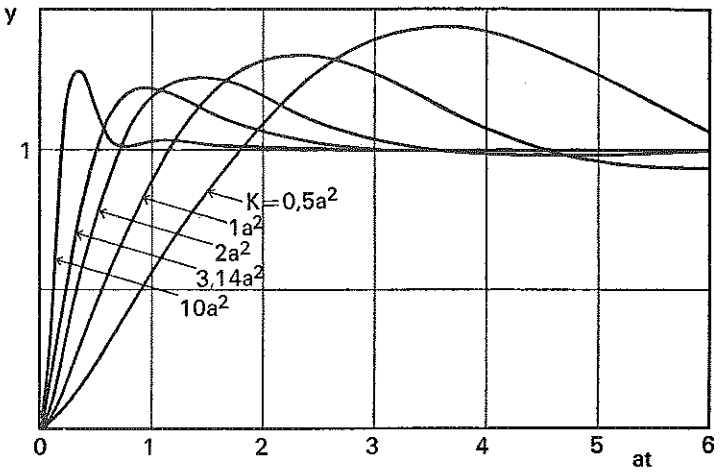


Fig. 6.13. Stegfunktionssvar för enkelt återkopplat system med kretsöverföringen. Jämför fig. 6.11.

$$G_0(s) = \frac{10K(s+a)}{s^2(s+10a)}$$

Det kompenserade systemets kretsöverföring är då

$$G_1(s)\,G_k(s) = \frac{K(s+a)}{s^2(s+1)(s+2)} \tag{6.21}$$

Vi skall först undersöka hur a skall väljas för att systemet skall bli stabilt. För detta användes Hurwitz' teorem (sats 5:6.2). Systemets karakteristiska ekvation lyder

$$s^4+3s^3+2s^2+Ks+aK=0 \tag{6.22}$$

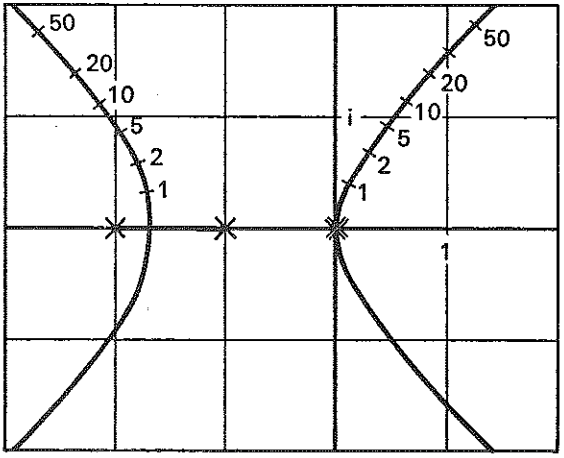


Fig. 6.14. Rotort för karakteristiska ekvationen för enkelt återkopplat system med kretsöverföringen

$$G_0(s) = \frac{K}{s^2(s+1)(s+2)}$$

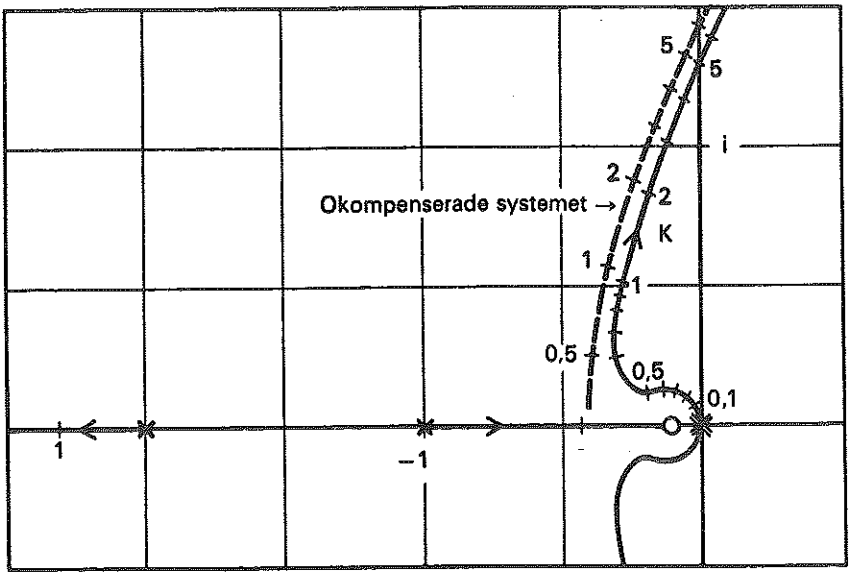


Fig. 6.15. Rotort för karakteristiska ekvationen för enkelt återkopplat system med kretsöverföringen

$$G_0(s) = \frac{K(s+0{,}1)}{s^2(s+1)(s+2)}$$

Huvudgrenen av det okompenserade systemets rotort visas streckad i figuren.

Hurwitzkriteriet (sats 5:6.2) ger följande villkor $3>0$, $6-K>0$, $(6-K)K-9aK>0$ och $aK>0$. Efter elimination av överflödiga villkor erhålls $0<a<(6-K)/9$ och $0<K<6$. Vi finner således att parametern a måste vara mindre än 2/3 om systemet skall vara stabilt. Vi väljer godtyckligt $a=0{,}1$ och finner då av ovanstående olikheter att systemet är stabilt om $K<5{,}1$.

Vi skall nu bestämma rotortens utseende då $a=0{,}1$. Detta resultat bör jämföras med rotorten för det okompenserade systemet (fig. 5:5.2). Vi finner då följande (siffrorna inom parentes ger motsvarande resultat för okompenserade systemet.) Startpunkter: 0, 0, -1, -2; (0, -1, -2). Asymptoternas skärningspunkt $-0{,}967$ ($-1{,}000$). Rotortens skärning med imaginära axeln 1,305 (1,414). I fig. 6.15 visas rotorten för det kompenserade systemet. Huvudgrenen av det okompenserade systemets rotort visas även i figuren.

Om vi bestämmer förstärkningsparametern K så att kontrollpolerna har relativa dämpningen $\zeta=0{,}7$ erhålls $K=0{,}66$ och $K=0{,}65$ för kompenserade resp. okompenserade systemen. Det kompenserade systemets överföringsfunktion är då

$$G(s) = \frac{0{,}66(s+0{,}10)}{(s^2+0{,}61s+0{,}19)(s+0{,}16)(s+2{,}23)}$$

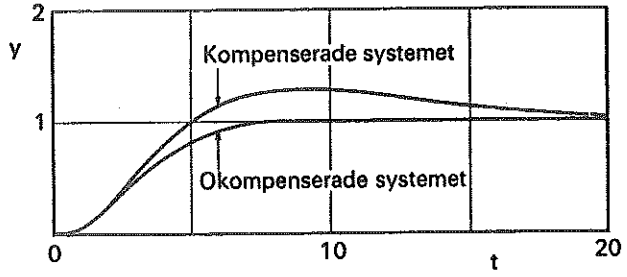


Fig. 6.16. Stegfunktionssvar för kompenserade och okompenserade systemen med $K = 0{,}65$. Jämför fig. 5:5.3.

Med $K = 0{,}65$ har det okompenserade systemet överföringsfunktionen

$$G(s) = \frac{0{,}65}{(s^2 + 0{,}76s + 0{,}29)(s + 2{,}24)}$$

Det kompenserade systemet avviker således från det okompenserade i två avseenden. »Dipolen» $(s+0{,}10)/(s+0{,}16)$ har tillkommit, de övriga polernas lägen har ändrats något. Enligt analysen i avsnitt 2 kommer det kompenserade systemets stegsvar att avvika något från det okompenserade. Jämför ekvation (2.6). Med $K = 0{,}65$ har emellertid det okompenserade systemet en mycket liten översläng så att det kompenserade systemet borde vara acceptabelt. Vi kontrollerar med simulering och erhåller resultatet i fig. 6.16. Det framgår av fig. 6.16 att det kompenserade systemet är väl »trögt» och vi försöker därför att höja kretsförstärkningen något. Resultatet framgår av fig. 6.17. Vi väljer tämligen godtyckligt $K = 1{,}0$.

Exempel 6.5. Ökning av systemets snabbhet

Betrakta ett enkelt återkopplat system med kretsöverföringen

$$G_0(s) = \frac{K}{s(s+1)(s+2)} \tag{6.23}$$

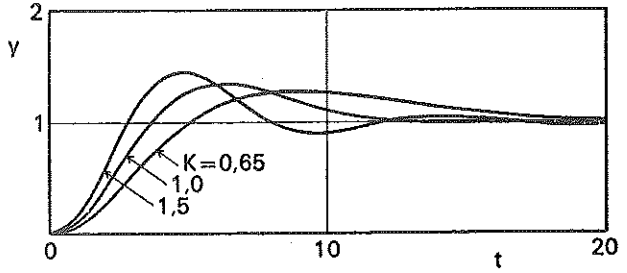


Fig. 6.17. Stegfunktionssvaret för det kompenserade systemet vid olika värden på förstärkningsparametern.

Vi har tidigare i avsnitt 5:8 funnit att $K = 1{,}0$ är ett lämpligt värde på förstärkningsparametern K. För detta parameterval är systemets bandbredd ca 0,8 rad $\cdot$ s^{-1} (fig. 5:8.8). Anta att specifikationerna kräver en bandbredd som är 10 gånger större. Alternativt 10 gånger mindre stigtid. Detta innebär att kontrollpolerna måste förskjutas en faktor 10 i radiell riktning. Detta kan uppnås genom en kompenseringslänk med överföringsfunktionen

$$G_k(s) = 100\,\frac{(s+1)(s+2)}{(s+10)(s+20)} \tag{6.24}$$

Det kompenserade systemet har således kretsöverföringen

$$G_1(s)\,G_k(s) = \frac{100K}{s(s+10)(s+20)}$$

Det är ingen fara med att förkorta bort faktorerna $s+1$ och $s+2$, eftersom de motsvarar nollställen i vänstra halvplanet. Jämför med diskussionen i avsnitt 4. Rotorten för det kompenserade systemet är alltså identisk med rotorten för det okompenserade systemet om skalorna ändras med en faktor 10. Vi finner att $K = 10$ är ett lämpligt värde på förstärkningsparametern. Kompenseringsnätet med överföringsfunktionen (6.24) förstärker höga frekvenser med en faktor 100, vilket innebär att vi kan få svårigheter med mättning av förstärkare. Om kompenseringen skall lyckas är det även nödvändigt att ekvationen (6.23) beskriver systemet väl upp till frekvenser på ca 100 rad/s.

Praktisk realisering av kompenseringsnät

Vi har således angivit två metoder för att bestämma överföringsfunktionen för en kaskadkompensering, som ger ett slutet system med givna specifikationer. Vi skall nu kortfattat diskutera hur man skall konstruera system med givna överföringsfunktioner. Kompenseringslänkarna kan praktiskt utformas på många olika sätt. Det speciella utförandet beror framför allt på i vilken form signalen föreligger. För vanliga servomekanismer kan signalerna vara läge, vridningsvinkel, hastighet, kraft, tryck, flöde, likström, likspänning eller modulerad spänning beroende på om systemet är mekaniskt, pneumatiskt, hydrauliskt eller elektriskt. Problemet är således att konstruera fysikaliska system med givna överföringsfunktioner, som är anpassade att motta och avge signaler i de önskade signalformerna. I många fall är signalerna elektriska. Kompenseringsnäten är då antingen passiva nät, t.ex. RC-kretsar, eller aktiva nät bestående av återkopplade operationsförstärkare, eventuellt små analogimaskiner. Även i andra fall omvandlas signalerna först till elektriska signaler med hjälp av mätomvandlare (transor). Kompenseringsnäten realiseras sedan som elektriska nät och signalerna omvandlas åter till önskad signalform. I nya system för processreglering är det mycket vanligt att alla signaler först omvandlas till elektriska standardsignaler. Dessa konverteras sedan till

digital form med hjälp av analog-digital omvandlare. Kompenseringsnäten realiseras som numeriska algoritmer i en dator och utsignalerna från kompenseringsnäten konverteras till elektriska signaler med digital-analog omvandlare. Eventuellt sker sedan ytterligare omvandling till andra signalformer.

Motiven till dessa skenbart mycket komplicerade system med mängder av signalomvandlingar är att det är lätt att praktiskt realisera även mycket komplicerade överföringsfunktioner med elektriska nät eller numeriska algoritmer i en processdator. Det är även lätt att med dessa utföringsformer ändra kompenseringsnätens egenskaper. Jämför t. ex. flexibiliteten hos en analogimaskin eller en dator med ett dynamiskt system, som är realiserat i form av ett hydrauliskt system med dämpkolvar, fjädrar och massor. Det starkaste motivet är dock ekonomiskt. Det har praktiskt visat sig billigare med de skenbart mer komplicerade systemen. Detta är speciellt markant med den nuvarande kostnadsutvecklingen inom mikroelektroniken.

I det följande skall vi visa hur kompenseringen skall realiseras med elektriska nät. Vi observerar först att om vi tillåter aktiva nät, t. ex. i form av operationsförstärkare, så har vi redan löst problemet. Med hjälp av operationsförstärkare kan vi erhålla integratorer och summatorer. Jämför exempel 5:3.3. Vi kan då direkt realisera vilken som helst överföringsfunktion som representerar ett linjärt tidsinvariant system. Överföringsfunktionen

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + \ldots + a_n}$$

kan t. ex. erhållas som det system som visas i fig. 2:10.10 eller fig. 2:10.11. Vi skall ge några exempel på hur enkla kompenseringsnät kan förverkligas med hjälp av en operationsförstärkare med återkopplingsnät.

EXEMPEL 6.6. Fasretarderande kompensering eller PI-kompensering

I exempel 6.1 skulle kompenseringen höja kretsförstärkningen vid låga frekvenser för att minska felkoefficienterna. Vi fann i exemplet att överföringsfunktionen

$$G_k(s) = \frac{s+a}{s+a/M} = M\frac{s+a}{Ms+a} \tag{6.2}$$

hade denna egenskap. För överföringsfunktionen (6.2) gäller att $\arg G_k(i\omega) \leqslant 0$ för varje ω. Kompenseringen har således alltid negativ fasförskjutning. Den kallas därför *fasretarderande* (*lag compensation*). Vi observerar vidare att då $M \to \infty$ övergår $G_k(s)$ i

$$\lim_{M\to\infty} G_k(s) = 1 + a/s$$

vilket är överföringsfunktionen för en PI-regulator. Den fasretarderande kompenseringen kallas därför även *integrerande* kompensering.

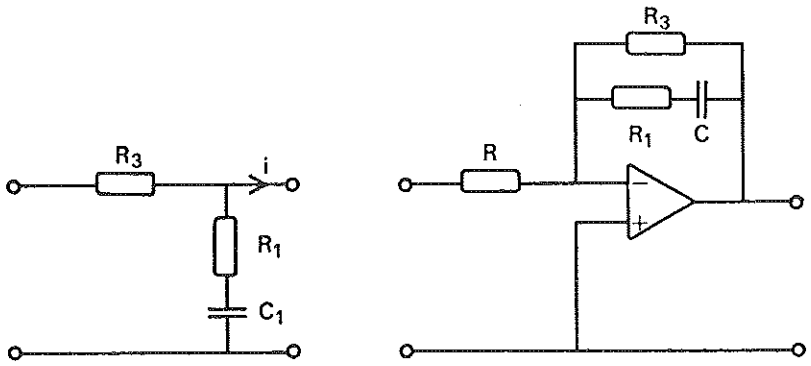


Fig. 6.18. Elektriska nät för fasretarderande kompensering. Om det passiva nätet är obelastat ($i = 0$) har det överföringsfunktionen

$$G(s) = \frac{s+a}{sM+a}$$

Om operationsförstärkaren har mycket hög förstärkning har det aktiva nätet överföringsfunktionen

$$G(s) = -KM\frac{s+a}{sM+a}$$

där $a = 1/(R_1C_1)$, $M = 1 + R_3/R_1$ och $K = R_1R_3/[R(R_1+R_3)]$

I fig. 6.18 visas två fasretarderande RC nät. Då ett passivt nät ger förstärkningen ett vid låga frekvenser måste systemets kretsöverföring ökas med en faktor N för att överföringsfunktionen (6.2) skall erhållas. Kompenseringsnätens Bodediagram framgår av fig. 6.19. Det aktiva nätet i fig. 6.18 ger en PI-regulator om motståndet R_3 avlägsnas ($R_3 \to \infty$).

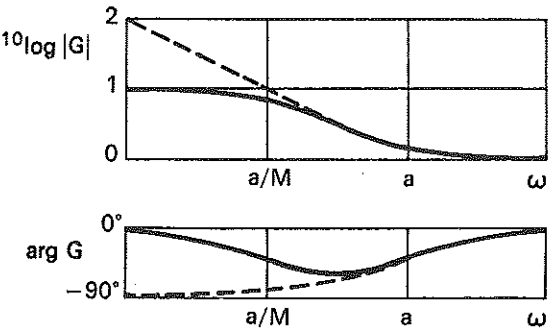


Fig. 6.19. Bodediagram för fasretarderande kompenseringsnät med överföringsfunktionen

$$G(s) = M\frac{s+a}{sM+a}$$

De streckade kurvorna visar Bodediagrammet för motsvarande PI-regulator med överföringsfunktionen $G(s) = 1 + a/s$, som erhålls då $M \to \infty$. I figuren är $M = 10$.

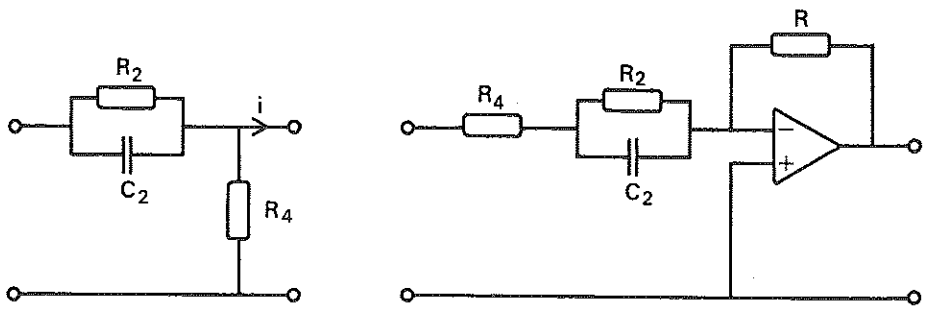


Fig. 6.20. Elektriska nät för fasavancerande kompensering. Om det passiva nätet är obelastat ($i = 0$) har det överföringsfunktionen

$$G(s) = \frac{s+b}{s+bN}$$

Om operationsförstärkaren har mycket hög förstärkning ($A = \infty$) har det aktiva nätet överföringsfunktionen

$$G(s) = -KN\,\frac{s+b}{s+bN}$$

där $b = 1/(R_2C_2)$, $N = 1 + R_2/R_4$ och $K = R/(R_2 + R_4)$

EXEMPEL 6.7. Fasavancerande kompensering eller PD-kompensering

För att stabilisera det system som behandlades i exempel 6.2, infördes ett kompenseringsnät med överföringsfunktionen

$$G_k(s) = \frac{b+s}{b+s/N} = N\,\frac{s+b}{s+bN} \tag{6.8}$$

För denna överföringsfunktion gäller $\arg G(i\omega) \geq 0$, för varje ω. Kompenseringen kallas därför *fasavancerande* (*lead compensation*). Vi observerar vidare att då $N \to \infty$ övergår $G_k(s)$ i

$$\lim_{N\to\infty} G_k(s) = 1 + s/b$$

vilket är överföringsfunktionen för en PD-(proportionell och deriverande)-regulator. Den fasavancerande kompenseringen kallas därför även *deriverande* kompensering.

I fig. 6.20 visas två fasavancerande elektriska nät. Då det passiva nätet har förstärkningen ett vid höga frekvenser, måste detta kombineras med en förstärkning på N gånger för att erhålla överföringsfunktionen (6.8). Kompenseringsnätets Bodediagram framgår av fig. 6.21. Formellt ger det aktiva nätet i fig. 6.20 en PD-regulator om motståndet R_4 kortsluts. I praktiken kan detta ej göras, ty operationsförstärkaren blir instabil eller mycket dåligt dämpad med $R_4 = 0$.

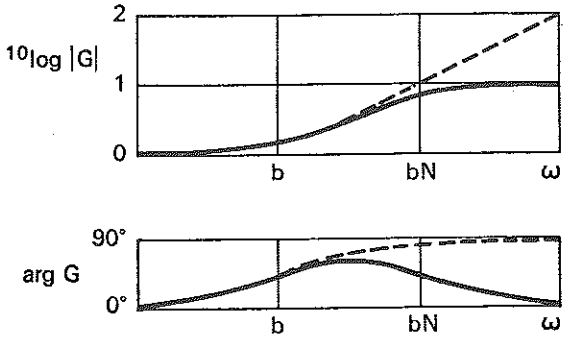


Fig. 6.21. Bodediagram för fasavancerande kompenseringsnät med överföringsfunktionen

$$G(s) = N\frac{s+b}{s+bN}$$

De streckade kurvorna visar Bodediagrammet för motsvarande PD-regulator med $G(s) = 1 + s/b$ som erhålls då $N \to \infty$. I figuren är $N = 10$.

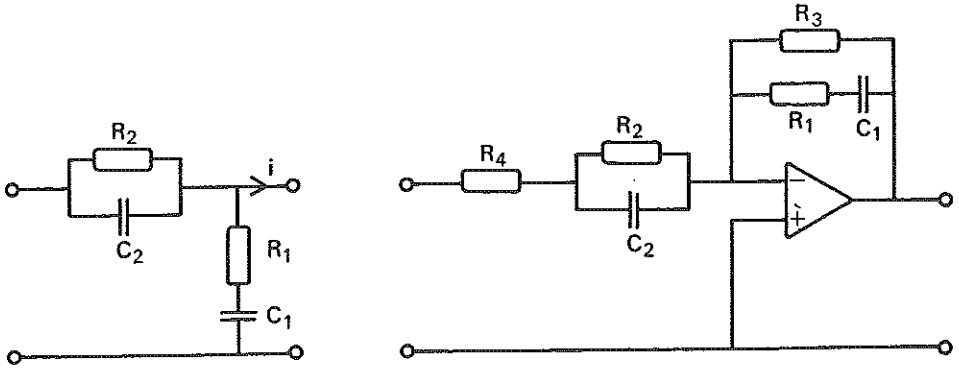


Fig. 6.22. Elektriska nät för fasretarderande och fasavancerande kompensering.

Om det passiva nätet är obelastat har det överföringsfunktionen

$$G(s) = L\frac{(s+a)(s+b)}{(sL+a)(s+bL)}$$

där L är den positiva roten till ekvation $R_1C_1L + R_2C_2/L = R_1C_1 + R_2C_2 + R_2C_1$.

Om operationsförstärkaren har mycket hög förstärkning har det aktiva nätet överföringsfunktionen

$$G(s) = -KMN\frac{(s+a)(s+b)}{(sM+a)(s+bN)}$$

Det gäller att $a = 1/(R_1C_1)$, $b = 1/(R_2C_2)$, $M = 1 + R_3/R_1$, $N = 1 + R_2/R_4$,
$K = R_1R_3/[(R_1+R_3)(R_2+R_4)]$

EXEMPEL 6.8. Fasretarderande och fasavancerande kompensering

Genom att kombinera kompenseringsnäten i exempel 6.6 och exempel 6.7 erhålls ett kompenseringsnät med följande överföringsfunktion

$$G_k(s) = \frac{s+a}{s+a/M}\,\frac{b+s}{b+s/N} = MN\,\frac{s+a}{a+sM}\,\frac{s+b}{s+bN} \tag{6.25}$$

En kaskadkompensering med denna överföringsfunktion har således potentiella möjligheter att såväl minska felkoefficienterna som att förbättra systemets stabilitet. Denna typ av kompensering förekommer mycket ofta. Den kallas i den amerikanska litteraturen *lead-lag-kompensering*. Då M och $N \to \infty$ övergår överföringsfunktionen (6.25) i

$$\lim_{M,\,N\to\infty} G_k(s) = (1 + a/s)(1 + s/b) = 1 + a/b + a/s + s/b$$

vilket är överföringsfunktionen för en PID-regulator. I fig. 6.22 visas elektriska nät med överföringsfunktionen (6.25). Det aktiva nätet i fig. 6.20 ger en PID-regulator om motståndet R_3 avlägsnas

7. Referenser

De flesta av de läroböcker som nämnts i tidigare avsnitt behandlar syntesproblemet. Förtjänstfulla diskussioner av specifikationer förekommer i

Truxal, J. G., *Control Systems Synthesis*, McGraw-Hill, New York 1955

Elgerd, O., *Control Systems Theory*, McGraw-Hill, New York 1967.

Truxal behandlar bl. a. sambanden mellan olika specifikationer.

Diagram och tabeller över stegsvar för vanliga konfigurationer av poler och nollställen finns t. ex. i

Truxal, J. G. (utg.), *Control Engineers Handbook*, McGraw-Hill, New York 1958.

Olika regulatorstrukturer behandlas i

Oppelt, W., *Kleines Handbuch Technischer Regelvorgänge*, Verlag Chemie, Weinlein 1964

Buckley, P. S., *Techniques of Process Control*, Wiley, New York 1964.

Framkoppling används ofta i processindustrin. Flera exempel ges i

Shinskey, F. G., *Process Control Systems*, McGraw-Hill, New York 1967

Luyben, W. L., *Process Modeling, Simulation, and Control for Chemical Engineers*, McGraw-Hill, New York 1973.

Dessa böcker innehåller också synpunkter på val av regulatorer liksom regler för inställning av regulatorer.

En uttömmande behandling av Otto Smith-regulatorn ges i
Smith, O. J. M., *Feedback Control Systems*, McGraw-Hill, New York 1958.
Denna lärobok innehåller också många värdefulla synpunkter på syntesproblemet.

Statisk optimering behandlas i
Lee, T. H., Adams, G. E. & Gaines, W. M., *Computer Process Control: Modeling and Optimization*, Wiley, New York 1968.

Dynamisk optimering diskuteras i
James, H. M., Nichols, N. B. & Phillips, R. S., *Theory of Servomechanisms*, McGraw-Hill, New York 1947
Newton, G. C., Gould, L. A. & Kaiser, J. F., *Analytical Design of Linear Feedback Controls*, Wiley, New York 1957.

Moderna framställningar har givits av
Bellman, R., *Adaptive Control Processes—A Guided Tour*, Princeton University, New Jersey 1961
Pontryagin, L. S. m. fl., *The Mathematical Theory of Optimal Processes*, Wiley, New York 1962
Bellman, R., *Introduction to the Mathematical Theory of Control Processes*, Academic Press, New York 1967
Bryson, A. E. & Ho, Y. C., *Applied Optimal Control*, Ginn and Company, Waltham Mass. 1969
Lee, E. B. & Marcus, L., *Foundations of Optimal Control Theory*, Wiley, New York 1967.

Den s. k. linjär-kvadratiska teorin diskuteras uttömmande i
Anderson, B. D. O. & Moore, J. B., *Linear Optimal Control*, Prentice Hall, New Jersey 1971
Kwakernaak, H. & Sivan, R., *Linear Optimal Control Systems*, Wiley, New York 1972.
Det finns ingen bra samlad framställning om adaptiva reglersystem. För att lära sig mer om sådana system är man därför hänvisad till tidskriftslitteraturen.

Syntes av framkoppling behandlas i doktorsavhandlingen
Bengtsson, G., *A Theory for Control of Linear Multivariable Systems*, Report 7341, Department of Automatic Control, Lund Institute of Technology, November 1973.

Syntes av reglersystem med hjälp av tillståndsåterkoppling är ett jämförelsevis nytt tillskott till reglerteorin. Mycket material finns i tidskriftslitteraturen men det finns få sammanställningar i samlad form. En behandling utifrån statistiska synpunkter ges i
Åström, K. J., *Introduction to Stochastic Control Theory*, Academic Press, New York 1970.

Kompenseringsmetoderna diskuteras uttömmande i läroböckerna i servomekanismteori. Nyquist–Bode–Nichols metoder behandlas i

Chesnut, H. & Mayer, R. W., *Servomechanisms and Regulating Devices*, Wiley, New York 1959

som även innehåller många tabeller, som underlättar de grafiska konstruktionerna.

Många genomräknade exempel finns i

D'Azzo, J. J. & Houpis, C. H., *Feedback Control System Analysis and Synthesis*, McGraw-Hill, New York 1960

Horowitz, I. M., *Synthesis of Feedback Systems*, Academic Press, New York 1963

Shinners, S. M., *Modern Control System Theory and Application*, Addison–Wesley, Reading Mass. 1972

Takahashi, Y., Rabins, M. J. & Auslander, D. M., *Control and Dynamic Systems*, Addison-Wesley, Reading Mass. 1970.

Utvidgningar till syntes av flervariabla system ges i

Rosenbrock, H. H., *State-Space and Multivariable Theory*, Nelson, London 1970

Wolowich, W. A., *Linear Multivariable Systems*, Springer, New York 1974.

I Rosenbrocks bok presenteras också ett nytt sätt att beskriva dynamiska system med en s. k. systemmatris.

Rosenbrock har även utvidgat kompenseringsmetoderna till flervariabla system. I detta fall krävs dock att man har tillgång till en dator. Se

Rosenbrock, H. H., *Computer-Aided Control System Design*, Academic Press, New York 1974.

Register

Accelerationskonstant 141
Accelerometer 22
Adams, G. E. 298
Adaptiv reglering 227
Adaptiv regulator 226
Administrativ process 11
Ahlfors, L. V. 187
AIAA 30
AIChE 30
Amplitudkurva 69
Amplitudmarginal A_m 180, 275
Analog-digitalomvandlare 293
Analogimaskin 21, 148, 292
Anderson, B. D. O. 298
Approximativa inverser 241
Approximativ derivation 241
Arbib, M. A. 93
Argumentkurva 69
Argumentvariationsprincipen 158, 171
Argumentvillkoret 162
ASME 30
Asymptoter
  Bodediagram 70
  rotort 163
Asymptoternas skärningspunkt
  Bodediagram 70
  rotort 164
Asymptotisk stabilitet 153, 157
Athans, M. 27
Atomreaktorer 19
Automatateori 25, 88, 93
Automatisk inriktning av väderkvarn 14
Automatiska landningssystem 15, 20
Autopiloter 15, 18, 19, 20, 28, 228, 271

Bandbredd 194, 198, 199, 207
  bestämning av bandbredd med hjälp av Nicholsdiagram 186
  syntes av system med specificerad bandbredd 284
Bandbreddskompromissen 202
Bandvalsverk 63, 219
Barnes, J. L. 91
Beckenbach, E. F. 92
Begränsning av ett systems prestanda 271
Bellman, R. 24, 26, 27, 91, 92, 182, 298
Beloppkurva 69
Beloppvillkoret 162
Bengtsson, G. 298
Bernoullis teorem 75
Biologisk process 11, 30, 31, 37
Birkhoff, G. D. 31, 93
Björk, Å. 118
Black, H. S. 21, 120, 144, 149, 173
Blackburn, J. F. 28
Blakelock, J. H. 28
Blockschema 77
Bode, H. W. 20, 31, 32, 93
Bodediagram 69
  användning för kompensering 275
  asymptoter 70, 71
  avläsning av amplitudmarginal och fasmarginal med hjälp av Bodediagram 181
  brytfrekvens 70
  brytpunkt eller asymptoternas skärningspunkt 70
  uppritning 69
Bodes relationer 73, 74
Boltyanskii, V. G. 27
Bombsikten 15
Brinck, I. 172, 187
Brown, G. S. 27
Brown, J. 29, 30
Bryson, A. E. 298
Brytfrekvens 70
Brytpunkt 72
Buckley, P. S. 28, 297
Buncker-Ramo 24
Byte av koordinater i tillståndsrummet 46
  till diagonalform 46
  till observerbar kanonisk form 47
  till styrbar kanonisk form 51

Campbell, D. P. 27, 29
Cayley–Hamiltons sats 48, 56, 99, 118
Cementugnar 15
Centrifugalregulator 16, 122, 131
Chestnut, H. 21, 27, 91, 149, 187, 299
Chipart 167, 171
Clynes, M. 30
Coddington, E. A. 92
Companion matris 48, 52

Cooke, K. L. 91
Cybernetik 26, 29

Dahlquist, G. 118
Dator 24, 37, 191, 283
Datorstyrning 23, 24, 25, 29, 211
D'Azzo, J. J. 27, 299
DDC-program 211
Decaulne, P. 27
Decibel 69
Decilog 69
Dekad 69
Desoer, C. A. 91
Destillationskolonner 15
Diagonalform 46, 47
Differensekvationer 37
Differential-differensekvationer 36, 91
Differentialekvationer 34
Differentialoperator 56
  approximation av 241, 242
Differential-integralekvationer 37
Digital-analog omvandlare 293
Dipol 196
Dirac's deltafunktion 54, 92
Direkta metoder för stabilitetsanalys 157
Distribution 59, 92
Doetsch, G. 92
Dominerande poler 195
Domnivåreglering 217
Dorf, R. C. 28
Drifttagning 188, 189
Dualitet 110
Dubbelintegratorn, se endimensionell partikel
Dynamisk framkoppling 215
Dynamisk optimering 225
Dynamisk programmering 24
Dynamiska system 31, 87
Dödtidskompensering 222

Eckman, D. P. 29
Egenvärden 46, 84, 153
Ekonomisk process 11, 29
Elektriska system 91
Elektroniska förstärkare 20, 144, 173
Elgerd, O. 28, 187, 297
Elgoner 21
Endimensionell partikel 34
  tillståndsekvation 39
  lösning av systemekvation 44
  insignal–utsignal-relation 56
  överföringsfunktion 62
Energialstring 14
Energidistribution 14
Enkelt system 37
Erdelyi, A. 92
Evans, W. R. 158, 186
Explosionsmotor 137
Exponentialmatrisen 44
Extern beskrivning 53
Extremalsökande regulator 223, 224

Falb, P. L. 27
Farkoststyrning 15
Fasavancerande kompensering (*PD*-kompensering) 283, 295
Fasavancerande nät 295
Faskurva 69
Faslåsning 15
Fasmarginal φ_m 180, 200, 275
Fasretarderande och fasavancerande kompensering (*PID*-kompensering) 283, 297
Fasretarderande kompensering (*PI*-kompensering) 283, 293
Felkoefficienter 120, 139, 196, 198, 199, 201, 275
Felkonstanter 141
  accelerationskonstant 141
  hastighetskonstant 141
  statisk konstant 141
Feller, W. 92
φ-cirklar 183
Flervariabelt system 38
Forrester, J. W. 29
Framkoppling 214, 216, 229
  statisk 215
  dynamisk 215
Frekvensanalys 67, 188
Frekvenskurva 66, 68, 175
Fröberg, C. E. 118
Fukthaltsreglering i pappersmaskin 13
Funktionalekvation 45, 89
Fördröjningstid T_d 193, 198
Förgrening 79
Förkortning av poler och nollställen 96, 113, 108, 236–241
Förlustfunktion 24, 204
Förstärkningsmarginal A_m 180
Förverkligande 188, 189

Gaines, W. M. 298
Gamkrelidze, R. V. 27
Gann, D. 29
Gantmacher, F. R. 92, 168, 187

Gardner, M. F. 91
Gatukorsning 11, 13
Gilberg, E. G. 118
Gille, J. C. 27
Goel, N. S. 30
Gould, L. A. 27, 298
Grafisk representation
  dynamiska system 77
  överföringsfunktionen 65
Gyroscopic stabiliser 19

Halldiagram 183, 184, 187
Hammond, P. H. 149
Harrer, J. M. 29
Harriott, P. 28
Hastighetskonstant 141
Hermite 167
Ho, Y. C. 298
Holt, C. C. 29
Horowitz, I. M. 299
Houpis, C. H. 27, 299
Hurwitz 18, 167, 170
Hurwitzdeterminanter 170
Hurwitzkriteriet 169
Hydrauliska system 28, 91
Högfrekvensasymptot 70

Icke minimum-fas-system 273
  approximativ invers 242
  definition 74
  exempel 74, 77
  svårighet att invertera 238
  svårighet att reglera 272
Identifiering 188
IEEE 29
IFAC 29
Impulssvar 55
Inbäddning 160
Indirekta metoder för stabilitetsanalys 157
Initialtillstånd 34, 53, 60, 155
Initialvärdesproblem 58
Insignal 11, 33, 87
Insignal–utsignal-relation 53
Insignal–utsignal-stabilitet 155
Instorhet 33
Instrumentation Laboratory 21
Instrumenteringsteknik 21
Insvängningstid T_s 192, 198, 199
Integrator 148
Integrerande eller I-reglering 17, 210
Interna beskrivningar 53
Inversa Laplace-transformen 59
Inversa system 229
  enkla linjära system 230
  flervariabla system 231
  olinjära system 235
Inverterad pendel 41
ISA 29
ITAE 205, 206, 246

JACC 29
Jacobs, J. 30
James, H. M. 21, 26, 149, 298

Kalman, R. E. 118
Kalmanfilter 253
Kalmans uppdelningssats 108, 258
Kanonisk form
  diagonal 46
  observerbar 47
  styrbar 51
Kaskad- eller seriekompensering 274
Kaskadreglering 213
Kausalitetsprincipen 33, 45, 55
Kedjebråk 169
Kemisk process 11
Kemiska reaktorer 15
Koenig, H. E. 91
Kompensering 274
  deriverande eller fasavancerande 295
  fasretarderande 293, 294
  genom lokal återkoppling 274, 275
  kaskad- eller seriekompensering 274, 275
  med rotortmetoden 284
  Nyquist–Bode–Nichols metod 275
  parallellkompensering 274, 275
  PI-kompensering 293
  tumregler 283
Kompenseringsmetoder 21, 274
Kompenseringsnät 293, 294, 295, 296
Kompromissen mellan stabilitet och förmåga
  att eliminera störningar 16, 19
Kontrollpoler 195
Konvergensabscissan 58
Kramers regel 61
Kretsförstärkning 139
Kretsöverföring 138
Kriteriefunktion 24, 204, 226
Kriterier
  observerbarhet 102
  optimering 24, 204, 205
  stabilitet 85, 157
  styrbarhet 98
Kvadratisk kriteriefunktion 226

Kvotregulator 224
Kwakernaak, H. 298
Känslighet 146, 206
  för parametervariationer 143
  samband med stabilitetsanalys 178
Kärnkraftteknik 19

Lag kompensering, se fasretarderande kompensering
Lagerhållning 13
Lagrange 18, 31
Lagranges metod med variation av konstanten 43
Laplacetransform 58, 92
LaSalle, J. 186
Lead-kompensering, se fasavancerande kompensering
Lead-lag kompensering, se fasavancerande och fasretarderande kompensering
Lee, E. B. 298
Lee, T. H. 298
Lefschetz, S. 92, 186
Leondes, C. T. 91
Levinsson, N. 92
Liénard 167, 171
Lighthill 92
Likformig asymptotisk stabilitet 153
Likformig stabilitet 153
Linjära system 38, 88
Linjära tidsinvarianta system 38, 89
Linjärisering 39, 225
Linjär-kvadratiska problemet 226
Logaritmiska frekvenskurvor 69
Lokal återkoppling 144, 227, 274
Luenbergers observerare 254
Luyben W. L. 297
Lyapunov 18, 152
Lyapunovs andra metod 158
Lågfrekvensasymptot 70
Långdistanstelefoni 144
Lösning 151
Lösningstid 192

MacCall 20
MacMillan 20
Magnus, W. 92
Maitra, S. C. 30
Marcus, L. 298
Matematisk modell 91
Matematisk pendel 35, 153
Matrisnorm 42, 155
Maximumprincipen 24
Maxtid T_m 199
Maxwell, J. C. 32, 119, 120, 128, 149, 167
Mayer, R. W. 21, 27, 91, 149, 187, 299
M-cirklar 183
Mekaniska resonanser 203
Melsa, J. M. 28
Mikusinski 92
Milhorn, H. T. 25, 29
Militära vapensystem 15
Milsum, J. H. 29, 30
Minimal realisation 112
Minimum-fas-system 73, 74
Minskning av felkoefficienter 278, 287
Mishchenko, E. F. 27
MIT Radiation Laboratory 21
Mod 47, 84
  styrbar 102
  observerbar 105
Modellbygge 188
Moderator 128
Montroll, E. W. 30
Moore, J. B. 298
M_p -värde 181
Multiplikation med konstant 79
Murphy, A. I. 91
Målsökning 15
Mätsignaler 11, 33
Mätvariabel 33
Mätvärdesomvandlare 23

Narendra, K. S. 118
Nemytskii, V. V. 93
Neper 69
Neurophysiology 26
Newton, G. C. 27, 298
Nichols, N. B. 21, 26, 149
Nicholsdiagram 184, 187
Nivåreglering 137
Nollösning 152
Nollställen 65
Numeriskt styrda verktygsmaskiner 15
Nyquist, H. 20, 173, 187
Nyquist–Bode–Nichols kompenseringsmetod 275
Nyquistdiagram 68
Nyquistkriteriet 20, 158, 173

Oberhettinger, F. 92
Observerare 254
Observerbar kanonisk form 47
Observerbarhet
  definition 97
  kriterier 102
  moder 105

system 97
tillstånd 97
Observerbarhetsmatris 103
Ogota, K. 28
Oldenburger, R. 173
Olinjär kompensering 227
Omgivning 11
Operationsanalys 25
Operationsförstärkare 23, 147, 292
Oppelt, W. 28, 297
Optimering
statisk 223
dynamisk 225
Optimeringskriterier 204, 226
Optimeringsteori 24, 27, 204, 223, 225
Originalfunktion 58
Oscillator 35
Ostörd lösning 151
Otto Smith-regulatorn 221, 222, 298

Pappersmaskiner 11
Parallellkompensering 274, 275
Parallellresonanskretsen 95, 100, 103, 109
Parameterfri representation 111
Parameterstyrning 226, 228
Parametervariationer 20
Partiella differentialekvationer 36, 91
Pelegrin, M. J. 27
Pendeln 35
insignal–utsignal-relation 57
inverterade systemet 236
linjärisering 40
lösning av systemekvationen 44
överföringsfunktion 62
Persson, A. 172, 187
Phillips, R. S. 21, 26, 149, 298
Physiological control systems 25
P-regulator, proportionell regulator 15, 17, 119, 126, 136, 210
PD-regulator 210, 283, 295
PI-kompensering 293
PI-regulator 23, 119, 127, 136, 209, 211, 258, 283, 293
PID-regulator 209, 210, 211, 212, 297
Poincaré 31
Poler 65
Polplacering
insignal–utsignal-återkoppling 260
jämförelse 270, 271
tillståndsåterkoppling 246, 248
tillståndsåterkoppling och rekonstruktion 256
Polära diagrammet 68
Pontryagin, L. S. 24, 27, 298
Porter, W. A. 91
Potentiometerskrivare 23
Praktisk realisering av kompenseringsnät 292
Praktisk stabilitet 178
Prestandabegränsning 271
Processdator 25
Processdynamik 23
Processindustri 24
Processreglering 23
Processreglersystem 25
Produktionsplanering 25
Profos, P. 28
Programstyrning 12
Proportionell regulator, se *P*-regulator
Proportionell och integrerande regulator, se *PI*-regulator
Proportionell och integrerande återkoppling, se *PI*-regulator
Proportionell och integrerande reglering, se *PI*-regulator
Proportionell reglering, se *P*-regulator
Psykologisk process 11

Raketstyrning 15
Rampsvar 194
Realisation
av dynamiskt system 112
av överföringsfunktion 112
minimal 112
observerbar 113
styrbar 115
Reethof, G. 28
Reglerventil 144, 212
Reglerprinciper 208
Regulator
P-regulator 15, 17, 119, 126, 236, 210
PD-regulator 210, 283, 295
PI-regulator 23, 119, 127, 136, 209, 211, 258, 283, 293
PID-regulator 209, 210, 211, 212, 297
Regulatorinställningar 212
Regulatorstrukturer 208, 209
Regulatorval 212
Referensmodell 216, 217, 257
Rekonstruerbart tillstånd 98
Rekonstruktion
av tillståndsvariabler 250
med derivatorer 250, 265
med dynamiska system 251
med Kalmanfilter 253, 266

med Luenbergers observerare 254, 266
med matematisk modell 251
Relativ dämpning 195
Representation av dynamiska system 110
Resonansfrekvens ω_p 194, 198, 199
Resonanstopp M_p 182, 194, 198, 199
Resultant 270
Richardson, H. H. 91
Riggs, D. S. 30
Rosenbrock, H. H. 299
Rosenbrocks system 231
Rotort 158
antal grenar 162
användning för analys av reglersystem 159
användning för syntes av reglersystem 284
definition 159
funktionsteoretisk tolkning 166
multipelrötter 164
start- och ändpunkter 162
symmetri 162
uppritning 162
Rouchés sats 172
Routh 18, 167
Routh-Hurwitzkriteriet 166
Rouths algoritm 167
Rouths tablå 167
Rymdteknologi 15

Samband mellan olika systemspecifikationer 206
Samplade system 37
Satellituppskjutning 12, 15, 204
Savant, C. J. 28
Savas, E. S. 29
Schultz, D. G. 28
Schwartz 92
Servomechanism laboratory 21
Servomekanismer 11, 15, 21
Servomotor 248
polplacering genom tillståndsåterkoppling 248
polplacering genom utsignal-återkoppling 265–267
rekonstruktion genom derivation 250
rekonstruktion med Kalmanfilter 253
rekonstruktion med Luenbergers observerare 254
Shearer, J. L. 91
Shinner, S. M. 28, 299
Shinskey, F. G. 29, 297
Signal 33
Signalflödesdiagram 77, 82
Simulering 188
Singularitetsdiagram 65
Sivan, R. 298
Skärningsfrekvens ω_c 181, 275, 276
Sluten styrning 12
Slutet system 139
frekvenskurva 183
överföringsfunktion 139, 183
Slutfel e_0 193
Smith, O. J. M. 298
Snabbhet 208
Specifikationer 191
i frekvensplanet 193
i tidsplanet 192
med optimeringskriterier 204
på överföringsfunktionen 195
samband mellan olika specifikationer 206
Sperry, E. A. 19
Spiegel, M. R. 92
Spirule 66
Stabiliserat spänningsaggregat 15
Stabilisering av instabilt system 274, 281, 285
Stabilitet 150
asymptotisk 153
definitioner 83, 151, 155
elementär teori 83
historik 17, 150, 166, 173
insignal–utsignal-stabilitet 155
kompromiss mellan stabilitet och statisk noggrannhet 16
kriterier 85, 158, 166, 171
likformig 153
likformig asymptotisk 153, 157
stabil lösning 152
stabila system 83, 154
Stark, L. 30
Stationär lösning 39
Statisk framkoppling 215
Statisk konstant 141
Statisk noggrannhet 139
Statisk optimering 223
Statiska fel 139
Statiskt system 31
Stegsvar 192
andra ordningens system 198, 199
första ordningens system 197
system med två poler och ett nollställe 200
system med tre poler 203
Stepanov, V. V. 93
Stigtid T_r 192
andra ordningens system 199
första ordningens system 198

Stigtid–bandbreddsprodukten 207, 208
Stodola, A. 18, 149, 167
Stokastiska processer 24
Stokastiska system 37, 298
Styrbar kanonisk form 51
Styrbara underrummet 99, 100, 101, 105
Styrbarhet 94
  definition 97
  kriterier 98
  moder 102
  system 97
  tillstånd 97
Styrbarhetsmatris 99
Styrbart system 97
Styrbart tillstånd 97
Styrd differentialekvation 34
Styrlag 119, 128, 190
Styrsignal 11
Styrvariabel 11, 33
Styvhetskoefficienter 120, 141, 142, 193, 208
Stålugn 204
Störd lösning 151
Störningar i begynnelsevillkor 151
Störningskänslighet 267
Summation 78
Superpositionsprincipen 139, 141
Syfte 11, 191
Syntes av reglersystem 188
  framkoppling 229
  kompensering Nyquist–Bode–Nichols 275
  kompensering rotort 284
  med given överföringsfunktion 245
System 11
  av n:te ordningen 38
  dynamiskt 31, 87
  enkelt 37
  flervariabelt 38
  linjärt 38, 88
  oändligt dimensionellt 36, 88
  samplade 37
  statiskt 31
  stationärt 38
  stokastiskt 37
  tidsdiskret 37, 88
  tidsinvariant 38
  tidskontinuerligt 88
  ändligt dimensionellt 36, 88

Tachometergenerator 21
Tachometeråterkoppling 248
Takahashi, Y. 299
Temperaturreglering 137, 220
Teorin för servomekanismer 21
Tidsdiskret system 37, 88
Tidsfördröjning 63, 77, 221, 231
Tidsinvariant system 88
Tillstånd 33, 87
  tyst 97
  rekonstruerbart 98
  uppnåeligt 97
  styrbart 97
Tillståndsekvationer 110
Tillståndsrum 46, 87
Tillståndsvariabel 33
Tillståndsvektor 33, 87
Tillståndsåterkoppling 226, 246
Tillståndsåterkoppling och rekonstruktion 256
Titchmarsh, E. C. 187
Topologiskt rum 87
Trafikreglering 13
Trajektorie 151
Transformation av systemekvationer
  till diagonalform 46
  till observerbar kanonisk form 47
  till styrbar kanonisk form 51
Transformteori 91
Transientanalys 188
Transienta uppförande 195
Transportfördröjning 63, 77, 221, 231
Trepunktskoppling 217
Trimning 188, 189
Truxal, J. G. 27, 91, 149, 187, 297
Tsien, H. S. 21, 27
Tumregler för kompensering 283
Turbiner 75, 137
Turbinreglering 149
Typnummer 139
Tysta underrummet 104, 106

Ugnsreglering 213
Underrum
  »icke-styrbart» 106
  »observerbart» 107
  styrbart 99, 100, 101, 105
  tyst 104, 106
Uppnåeligt tillstånd 97
Uppritning av Bodediagram 69
Uppritning av rotorten 162
Utsignal 11, 33, 38, 87
Utstorhet 33

Variationer i systemparametrarna 119
Variationskalkyl 24

Varierande processdynamik 226
Varvtalsreglering av ångpanna 12
Vattenklockor 14
Vattenturbin 75
Vektornorm 155
Viktfunktion 53, 54, 55, 60, 111
Vinstfunktion 24, 204
Vyshnegradskii, J. 17, 119, 120, 131, 149, 167
Värmeledning 64
Värmeväxlare 15

Watts centrifugalregulator 15
Widder, D. V. 92
Wiener, N. 24, 26, 29
Wolowich, W. A. 299
Wonham, W. M. 118
Wright 18

Zadeh, L. A. 91
Ziegler–Nichols metod för regulatorinställning 212

Ångkraftverk 11
Ångmaskin 14, 120
Ångmaskinreglering 15, 120
  med *P*-regulator 124
  med *PI*-regulator 125
Ångpanna 77, 218
Åström, K. J. 28, 298
Återkopplade förstärkare 15, 20, 144
Återkopplade system 124, 137, 173
Återkoppling 12, 20, 119, 216

Ökning av systemets snabbhet 291
Översläng M 192, 198, 199
Öppen styrning 12
Överföringsfunktion 58, 111
  bestämning av 188
  definition 60
  grafisk representation 65
  realisation av 112
  samband med viktfunktion 60
  slutna systemet 139
  öppna systemet 138, 139

sid	rad	står	skall stå
69	-2	log G(iω)	log(|G(iω)|)
102	-13	exempel 3:3.1	2
107	-2	$(\ldots A_{41}')$	$(\ldots A_{44}')$
108	-14	$G(s)=C_2'(sI-A_{11}')^{-1}B_1'+D$	$G(s)=C_1'(sI-A_{11}')^{-1}B_1'+D$
183	-8	2tg(φ/2)	2tg(φ)
183	-6	2sin (φ/2)	2sin (φ)
192	12-	minskat med 1.	... minskat med G(0).
	11-	i %.	i % av slutvärdet.
			(OBS!) Definitionen av M i fig 2.1 är följaktligen fel.
194	6-	$= 1/\sqrt{2}$	$=\lvert G(0)\rvert/\sqrt{2}$
199	Figur text	$\frac{\omega_o^2}{s^2+2\zeta\omega_o + \omega_o^2}$	$\frac{\omega_o^2}{s^2+2\zeta\omega_o s+\omega_o^2}$
207	-4	...i själva verket	..i själva verket ej konvergent
243	5		fattas en] längst till höger
243	8	$\int_{iR}$	$\int_{ir}^{iR}$
243	12	-	tillägg: γ^- är en infinitesimal halvcirkel till vänster om origo
243	-9	$= -dG^-/ds\|_{s=0}G^+(0)H(0) =$	$= -2dG^-/ds\|_{s=0}G^+(0)H(0) =$
		$= -dG^-/ds\|_{s=0}/G^-(0)$	$= -2dG^-/ds\|_{s=0}/G^-(0)$
244	4-	pol s=1	nollställe s=1
247	-3	U_r	Y_r
258	Figur text	...och ej observerbart	...och observerbart...

258 Fig 5.2 felaktig: den steckade ramen "styrbart och observerbart skall även omsluta blocket B.

269	-3	m+k+2, m+l+2	n+k+1, m+l+1
276	-15	uppskattas ur förstärkningsmarginal och fasmarginal	uppskattas ur skärfrekvens, förstärkningsmarginal och fasmarginal
283	-14	..hjälp av en fasretarderande	...hjälp av en fasavancerande

Tillägg: Formelsamling sid 9 formel 22: För att formeln skall stämma måste τ väljas så att $0<\tau<\infty$.