

Sammanfattning av EL1000 Reglerteknik, allmän kurs

Yashar Honarmandi
yasharh@kth.se

14 oktober 2019

Sammanfattning

Detta är en sammanfattning av EL1000 Reglerteknik, allmän kurs.

Innehåll

1	Grundläggande koncept	1
2	Prestanda och prestandamått	1
3	Negativ återkoppling	2
4	Frekvensanalys	6
5	Kompensering	10
6	Tillståndsrepresentationer	12
7	Alternativa regulatorstrukturer	16

1 Grundläggande koncept

Syftet med reglerteknik Reglerteknik handlar om att kontrollera olika storheter, ofta betecknad y , i ett system mot något värde, ofta betecknad r . I tillägg till systemets egna beteende påverkas storheten vi vill reglera typiskt av en yttre störning v . Vi kan reglera systemet genom att tillföra en påverkan, ofta betecknad u .

Strategi För att förstå systemet, tar vi först fram en modell som beskriver det. Ur denna modellen fås typiskt en differentialekvation. Denna löser vi med Laplacetransform över tid.

Överförningsfunktionen För linjära system fås en lösning i Laplacerummet på formen $Y(s) = G(s)U(s)$, där U är Laplacetransformen av u . Funktionen G är överförningsfunktionen. Notera att denna lösningsformen typiskt beror på att alla initialvärden är 0.

Poler Ett systems poler är rötterna till nämnarpolynomet (som typiskt finns) i överförningsfunktionen.

Stabilitet Ett system är stabilt om det tenderar mot ett visst läge efter lång tid. Systemets stabilitet är typiskt kopplad med dets poler. Detta kan man se i enkla fall, till exempel vanliga linjära ordinära differentialekvationer. Här är systemet stabilt om det inte finns några poler i högre halvplan, och avståndet längs med reella axeln anger hur snabbt lösningen tenderar mot det stabila läget.

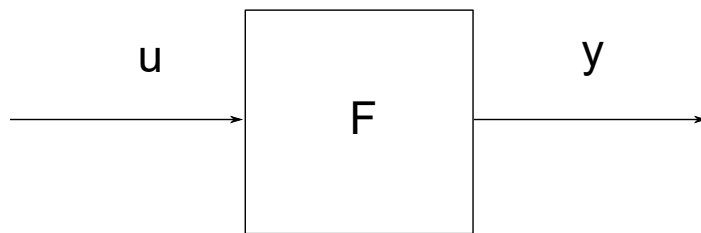
Nollställen Ett systems nollställen är rötterna till täljarpolynomet (som typiskt finns) hos överförningsfunktioner. Eftersom vi är intresserade av att styra y , är det viktigt hur vi ska välja u för att få det. Därmed är $\frac{1}{G}$ en viktig storhet, och nollställen kan därmed orsaka reglerproblem som är svårlösta.

Impulssvar Om lösningen för Y är på formen $Y = GU$, är lösningen för y på formen

$$y(t) = \int_0^t d\tau g(\tau)u(t - \tau).$$

g kallas för impulssvaret.

Blockschema Ett blockschema är ett systematiskt sätt att rita reglerade system på. För att förstå hur man läser dem, betrakta figur 1.



Figur 1: Illustration av ett enkelt blockschema.

Med denna figuren menas att $Y(s) = F(s)U(s)$.

Rotort En rotort är en plott av ett systems poler som funktion av någon parameter. Den är typiskt uppdelad i grenar, som är kurvor i planet som är parametriserade av parametervärdet. Polerna som motsvarar parametervärdet 0 är rotortens startpunkter, och polarna motsvarande parametervärdet ∞ är rotortens ändpunkter. Om rotorten närmar sig kurvor, är dessa rotortens asymptoter.

2 Prestanda och prestandamått

Stigtid Stigtiden definieras som $T_r = t_2 - t_1$, där vi typiskt har kriteriet $y(t_2) = 0.9$ och $y(t_1) = 0.1$, med y mätt i relativa enheter.

Insvängningstid Insvängningstiden definieras som $|y(t) - 1| < p$ när $t > T_s$, med y mätt i relativa enheter. p är typiskt lika med 0.05.

Översläng Överslänget definieras som $y_{\max} - 1$, med y mätt i relativa enheter.

Parametrar i svängningslika system Om du har ett system med ett andra ordningens polynom i överföringsfunktionens nämnare, skriv polynomet som $s^2 + 2\zeta\omega_0 s + \omega_0^2$. ω_0 är systemets resonansfrekvens och ζ dets dämpning. Det gäller för ett rent andra ordningens system att

$$T_r \propto \frac{1}{\omega_0}, \quad T_s \approx \frac{3}{\zeta\omega_0}, \quad M = e^{\frac{\pi\zeta}{\sqrt{1-\zeta^2}}}.$$

Stationärt fel Det stationära felet är felet $e = r - y$ som kvarstår efter lång tid.

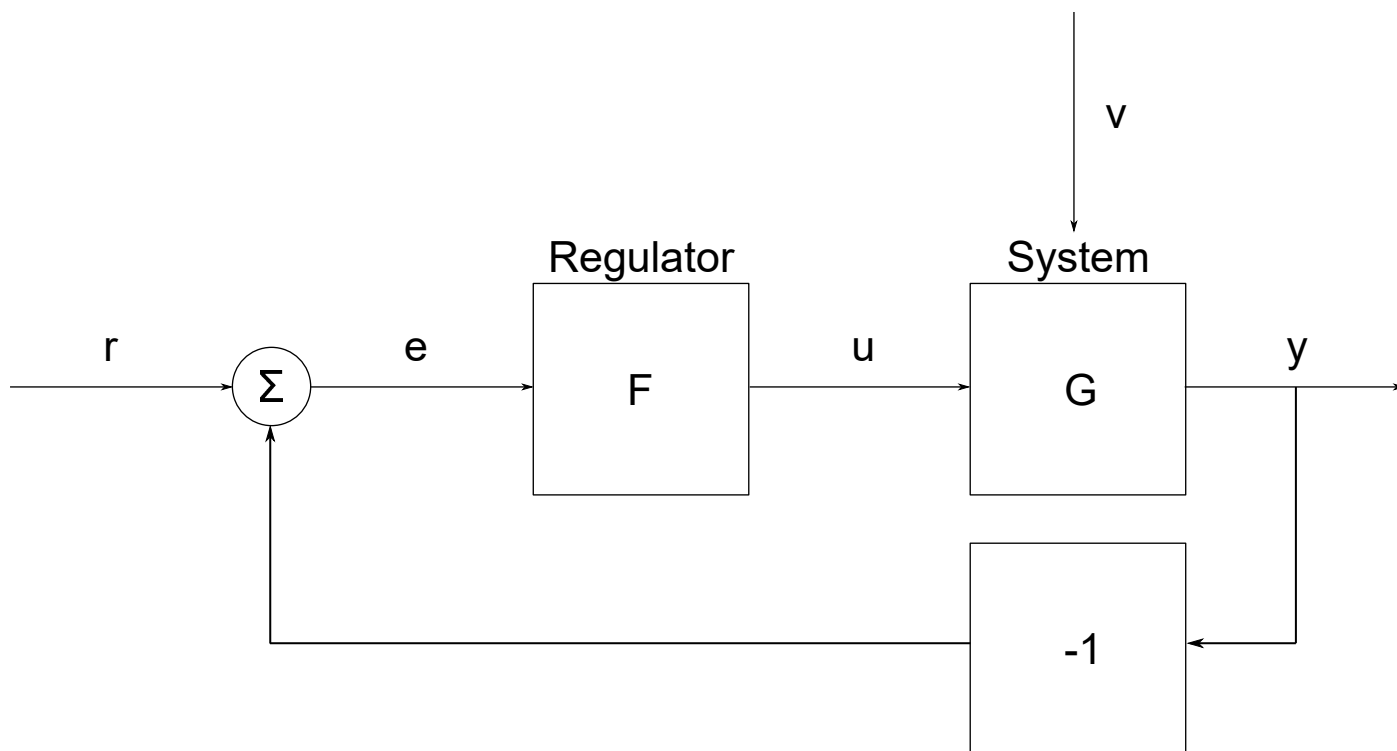
Felkoefficienter Det stationära felet beror både på systemets egenskaper och reglersignalen. Om reglersignalen är på formen $r_n = t^n\theta(t)$, där θ är Heavisidefunktionen, definieras felkoefficienterna som

$$e_n = \lim_{t \rightarrow \infty} r_n - y.$$

3 Negativ återkoppling

Vad är negativ återkoppling? I denna kursen kommer vi huvudsakligen att studera hur man kontrollerar ett system vid att låta avvikelsen mot det önskade värdet kontrollera regleringen av storheten, alltså låta $U = FE$, där $e = r - y$.

Illustration i blockdiagram Ett enkelt negativt återkopplad system illustreras i figur 2.



Figur 2: Schematisk illustration av ett enkelt negativt återkopplad system.

Beskrivning av systemet Vi börjar beskrivningen av systemet med att inte betrakta störningar. I ena ändpunkten har vi

$$Y = GU = GFE.$$

Summationskomponenten till vänster ger oss

$$E = R - Y,$$

och därmed

$$Y = GFR - GFY.$$

Därmed kan vi skriva

$$Y = \frac{GF}{1 + GF}R.$$

Återkopplad överföringsfunktion För ett återkopplad system som kan skrivas som $Y = G_C R$ definieras G_C som den återkopplade överföringsfunktionen. För systemet ovan har vi alltså

$$G_C = \frac{GF}{1 + GF}R.$$

Samband mellan reglerfel och referens Alternativt kan vi lösa systemet ovan för att få

$$R - E = GFE, \quad E = \frac{1}{1 + GF}R.$$

Samband mellan referens och styrsignal Systemet ovan kan även lösas för att ge

$$U = FR - FY = FR - GFU, \quad U = \frac{F}{1 + GF}R.$$

Känslighetsfunktionen Känslighetsfunktionen S definieras som överföringsfunktionen från störning till utsignal.

Komplementära känslighetsfunktionen Komplementära känslighetsfunktionen definieras som $T = 1 - S$.

Slutna systems poler Vi ser att slutna system har poler där $1 + GF = 0$. Därmed bestäms systemets stabilitet av systemet och regulatorn.

P-reglering Principet i P-reglering är att välja en styrsignal som är proportionell mot storleken av felet, alltså

$$u = K(r - y) = Ke.$$

Det är här klart att för att få negativ återkoppling väljer vi $K > 0$.

Denna regleringsmetoden

- minskar inverkan av störning och modellfel för ett bra val av K .
- ökar snabbheten vid insvängning.
- stabiliserar instabila system.

Däremot kan regleringen gå fel om t.ex.

- systemet inte uppför sig som man tror.
- man har begränsningar i styrförmåga.
- man får instabilitet på grund av återkopplingen.

Det är även ett problem att om felet är stationärt, är även styrsignalen det, så även om du har ett nollskild fel klarar inte systemet nödvändigtvis anpassa sig.

PID-reglering PID står för proportionell integrerande deriverande. Denna sortens reglering löser många reglerproblem. Med PID-reglering väljer vi styrsignalen

$$u = K_P e + K_I \int_{t_0}^t d\tau e + K_D \frac{de}{dt}.$$

Alternativt kan vi skriva det som

$$u = K \left(e + \frac{1}{T_I} \int_{t_0}^t d\tau e + T_D \frac{de}{dt} \right).$$

De tre ingående termerna i styrsignalen är

- proportionell återkoppling, som betraktar det nuvarande felet.
- integrerande återkoppling, som betraktar hur felet har uppfört sig.
- deriverande återkoppling, som betraktar hur felet kommer att uppföra sig.

PI-reglering PI-reglering använder ej den deriverande återkopplingstermen. Vi ser härifrån att vid ett stationärt tillstånd är antingen $e = 0$, annars ökar eller minskar u på grund av integraltermen.

Vi vill nu betrakta systemets insvängning. Om det stationära \bar{u} krävs för att $e = 0$, har vi

$$\bar{u} = K \left(e + \frac{1}{T_I} \int_{t_0}^t d\tau e \right).$$

Vid att derivera detta fås

$$K \left(\frac{de}{dt} + \frac{1}{T_I} e \right) = 0,$$

med lösning proportionell mot $e^{-\frac{t}{T_I}}$.

Notera att om man har stort fel kan PI-reglering ge problem. Därför använder man det typiskt när felen är små.

PI-reglering i Laplacevärlden Vid att Laplacetransformera uttrycket för styrsignalen i en PI-regulator, nämligen

$$u = K \left(e + \frac{1}{T_I} \int_{t_0}^t d\tau e \right),$$

fås

$$U = K \left(E + \frac{1}{T_I s} E \right),$$

och därmed

$$F(s) = K \left(1 + \frac{1}{T_I s} \right).$$

Rotorter med negativt återkopplade system Betrakta ett system med överförningsfunktion G_O för det öppna systemet. Det slutna systemet kommer ha överförningsfunktion

$$G_C = \frac{G_O}{1 + G_O}.$$

G_O är ofta på formen

$$G_O = K \frac{Q}{P},$$

där K är en parameter. Då får vi

$$G_C = \frac{KQ}{P + KQ}.$$

Systemet har alltså poler som ges av

$$P + KQ = 0, \quad \frac{P}{Q} = -K.$$

Detta kriteriet anger systemets rotort.

Hur man ritar rotorter Vi antar nu att P och Q är polynom av grad n respektive m , där $n \geq m$. Då vet vi att kriteriet har n rötter, och rotorten har därför n grenar (annars skulle den inte kunna starta i n punkter).

Rötterna dyker antingen upp som reellvärda eller i par av komplexkonjugerade rötter, och därmed är rotorten symmetrisk med avseende på reella axeln. Vi vet även att rotorten har $n - m$ asymptoter, då det bara finns m ändpunkter.

Det gäller även att alla delar av reella axeln som har ett udda antal reella start- och ändpunkter till höger om sig tillhör rotorten. För att förstå detta, faktorisera polynomen på vänstersidan och para ihop de komplexkonjugerade paren så att alla koefficienter är reella. Om man startar långt till höger på reella axeln är kvoten positivt, och tillhör ej rotorten. Varje gång den passerar en start- eller ändpunkt byter kvoten tecken, och den kan då tillhöra rotorten. En bra matteövning kan vara att övertyga sig om att det även är sant för de komplexa nollställena.

Vi vill nu studera asymptoterna, och antar därför här att $n > m$. Vi skriver då

$$\frac{P}{Q} = \frac{s^n + a_1 s^{n-1} + \dots}{s^m + b_1 s^{m-1} + \dots} = s^{n-m} + (a_1 - b_1) s^{n-m-1} + \frac{P_1}{Q}.$$

Vi använder binominalsatsen för att skriva

$$\frac{P}{Q} = \left(s + \frac{a_1 - b_1}{n - m} \right)^{n-m} + \frac{P_2}{Q},$$

där utvecklingen av parentesen kommer återskapa de två potenstermerna ovan. P_2 har grad $n - 2$, så för s med stora belopp kommer parentesen dominera. Rotortkriteriet ger

$$\arg \left(s + \frac{a_1 - b_1}{n - m} \right) = \frac{\pi}{n - m} + \frac{2\pi}{n - m} k, \quad k = 0, \dots, m - n.$$

Detta är strålar från punkten $-\frac{a_1 - b_1}{n - m}$ som pekar i riktningarna $\frac{\pi}{n - m} + \frac{2\pi}{n - m} k$ för stora s . Eftersom asymptoterna strålar ut från denna punkten, ser vi att den är rotortens centroid. Binominalsatsen ger även att a_1 är summan av systemets poler och b_1 är summan av systemets nollställena.

Det finns en entydighetssats för rotortkriteriet, vilket innebär att om det finns en del av reella axeln mellan två poler som är med i rotorten, kommer rotorten behöva bryta ut från reella axeln någonstans imellan dessa två polerna. På samma sätt kommer rotorten bryta in i reella axeln i områden mellan två nollställena som är med i rotorten.

Om man vill, kan man räkna ut vinkeln i punkterna där rotorten bryter in och ut. Detta vet jag dock inte hur man gör. Man kan även räkna ut för vilka s rotorten bryter in och ut. För att göra detta, ställ upp karakteristiska ekvationen $P + KQ = 0$, derivera med avseende på s . Punkterna som uppfyller $\frac{dK}{ds} = 0$ som ger $K > 0$ är punkterna där rotorten bryter in och ut.

För att rita rotorten kan du följa dessa steg:

- Beräkna överförningsfunktionen för det slutna systemet. Skriv nämnaren som $P + KQ = 0$.
- Hitta startpunkterna, alltså rötterna till P .
- Hitta ändpunkterna, alltså rötterna till Q .

- Bestäm antal asymptoter, alltså gradtalet till P minus gradtalet till Q .
- Bestäm vilka delar av reella axeln som tillhör rotorten. Det är de delar som har ett udda antal start- och slutpunkter till höger om sig.
- Beräkna punkterna där rotorten korsar reella axeln. Såna ligger mellan två startpunkter. För att hitta dem kan man ansätta att karakteristiska ekvationen har en dubbelrot, och identifiera vilka kombinationer av s och K som ger detta, eller hitta stationära punkter för K som funktioner av s .
- Bestäm centroiden $-\frac{\sum p_i - \sum q_i}{n-m}$, där p_i är systemets poler och q_i dess nollställen.
- Bestäm asymptoternas riktningar relativt centroiden genom att betrakta argumentet till $\frac{P}{Q} = -K$ för stora s . Spoilers: De är $\frac{\pi}{n-m} + \frac{2\pi}{n-m}k$, $k = 0, \dots, m-n$.
- Bestäm korsningar med imaginära axeln genom att sätta $s = i\omega$. Kom ihåg att $K \geq 0$.
- Rita.

Nyquistkurvan Givet överförningsfunktionen

$$G_C = \frac{G_O}{1 + G_O}.$$

för ett slutet system, finns det några poler så att systemet är instabilt? En ide för att undersöka detta är att undersöka alla s i högre halvplan och se vilka värden på G_O man får. Detta gör vi genom att rita två halvcirklar med radier r respektive R i högre halvplan, förbinda dem med raka linjer i änderna och låta r gå mot 0 och R mot ∞ . G_O kommer då anta värden på en kurva som kallas för Nyquistkurvan.

Eftersom Nyquistkurvan i stort sett antingen skickas mot oändligheten eller mot origo, är den viktigaste delen av Nyquistkurvan som är avbildad från imaginära axeln. Man kan även observera att $G_O(\infty) = 0$ och $G_O(0) \approx \frac{K}{s^p}$, där p är antal poler i origo.

Nyquistkriteriet Betrakta ett öppet systems Nyquistkurva. Argumentvariationsprincipen från komplex analys ger oss att antalet poler i höger halvplan till ett återkopplat system är lika med antalet poler i höger halvplan hos G_O plus antalet varv som Nyquistkurvan omsluter punkten -1 . Detta kallas Nyquistkriteriet.

Speciellt, om G_O inte har poler i höger halvplan, är systemet stabilt om Nyquistkurvan inte omsluter -1 . Detta är varianten av Nyquistkriteriet som ofta används i denna kursen.

4 Frekvensanalys

Fundamental ide Eftersom periodiska funktioner kan skrivas som en summa av trigonometriska funktioner och funktioner som avtar tillräckligt snabbt kan skrivas som en integral över trigonometriska funktioner, vet vi att när vi studerar linjära system räcker det att studera systemets respons på en enda term, alltså en enda trigonometrisk funktion, och se hur den beror av frekvensen. Om vi tillför en signal $u = \sin \omega t$ till ett system med överförningsfunktion G får vi

$$\begin{aligned} y &= \int_0^\infty d\tau g(\tau) u(t - \tau) \\ &= \text{Im} \left(\int_0^\infty d\tau g(\tau) e^{i\omega(t-\tau)} \right) \\ &= \text{Im} \left(e^{i\omega t} \int_0^\infty d\tau g(\tau) e^{-i\omega\tau} \right) \\ &= \text{Im}(e^{i\omega t} G(i\omega)) \\ &= |G(i\omega)| \sin(\omega t + \arg G(i\omega)). \end{aligned}$$

Det kan även finnas transienta termer här, men om systemet är stabilt kommer dessa försvinna över tid. Vi ser alltså att systemets svar beror av $G(i\omega)$.

Nyquistdiagram Ett Nyquistdiagram är en uppritning av $G(i\omega)$ för $0 < \omega < \infty$. Vi kan se att detta i stort sett är den intressanta delen av Nyquistkurvan.

Bodediagram Ett Bodediagram är en diagram av $|G(i\omega)|$ och $\arg G(i\omega)$ som funktioner av ω .

Skärfrekvens Skärfrekvensen ω_c är frekvensen för vilken Nyquistkurvan skär enhetscirkeln. Det är en tummregel att denna är proportionell mot systemets bandbredd, som vi kommer diskutera sedan.

Fasmarginal Fasmarginalen ϕ_m är vinkeln mellan negativa x -axeln och punkten där Nyquistkurvan skär enhetscirkeln. Mer specifikt kan vi skriva $\phi_m = \pi + \arg G(i\omega_c)$.

Fasskärfrekvens Fasskärfrekvensen ω_p är frekvensen så att $\arg G_O(i\omega_p) = -\pi$.

Amplitudmarginal Amplitudmarginalen ges av $A_m = \frac{1}{|G_O(i\omega_p)|}$.

Alternativt förenklad Nyquist-kriterium Det förenklade Nyquistkriteriet kan nu formuleras som $\phi_m > 0$, $A_m > 1$.

Bandbredd Bandbredden är bredden på det frekvensintervallet där $|G(i\omega)| \geq \frac{1}{\sqrt{2}}$ i relativa enheter, vilket motsvarar förstärkning större än 3dB, och benämns ω_B . Bandbredden kan ge information om systemets tillväxt, då hög bandbredd typiskt betyder snabb tillväxt.

Resonansfrekvens Resonansfrekvensen ω_r är den frekvens som ger starkast respons i systemet.

Resonanstopp Resonanstoppen är $M_p = |G(i\omega_r)|$ i relativa enheter (?), och ger typiskt en indikation på hur mycket översläng man får. Man önskar typiskt att denna ska vara liten.

Dekomponering av Bodediagram Vi skriver överförningsfunktionen som

$$G(i\omega) = \frac{K \prod \left(1 + \frac{i\omega}{z_i}\right)}{s^p \prod \left(1 + \frac{i\omega}{p_i}\right)},$$

som har poler i $-z_i$ och nollställena i $-p_i$. En sådan överförningsfunktion ger

$$\log |G(i\omega)| = \log K + \sum \log \left|1 + \frac{i\omega}{z_i}\right| - \sum \log \left|1 + \frac{i\omega}{p_i}\right|, \quad \arg G(i\omega) = -p \frac{\pi}{2} + \sum \arctan \frac{\omega}{z_i} - \sum \arctan \frac{\omega}{p_i}.$$

Alltså är Bodeplottet en summa av bidrag från enstaka faktorer, eventuellt två komplexkonjugerade parade ihop. Vi studerar de därför var för sig.

Bodediagram för en potens Betrakta en faktor s^p i överförningsfunktionen. För den gäller att

$$\log |G(i\omega)| = p \log \omega, \quad \arg G(i\omega) = -p \frac{\pi}{2}.$$

Detta motsvarar en rät linje med lutning p i Bodediagrammet och en konstant fassänkning.

Bodediagram för en linjär faktor En enda faktor $1 + \frac{i\omega}{z_i}$, under antagandet att z_i är reell, har

$$\log |G(i\omega)| = \frac{1}{2} \log \left(1 + \frac{\omega^2}{z_i^2}\right) \quad \arg G(i\omega) = \arctan \frac{\omega}{z_i}.$$

Lågfrekvensasymptoten är 0, medan högfrekvensasymptoten är $\log \frac{\omega}{z_i}$, alltså en rät linje med lutning 1 i en log-log-plot. Dessa två asymptoterna skär varandra för $\omega = z_i$. I denna punkten ändrar bodeplotten brått sin lutning. Denna punkten kallas då bodeplottens brytpunkt.

Bodediagram för ett polynom av andra ordningen Betrakta två faktorer

$$\left(1 + \frac{i\omega}{z_i}\right) \left(1 + \frac{i\omega}{z_i^*}\right) = 1 - \frac{\omega^2}{|z_i|^2} + 2i\omega \left(\frac{1}{z_i} + \frac{1}{z_i^*}\right) = 1 - \frac{\omega^2}{|z_i|^2} + 2i\frac{\omega}{|z_i|^2}(z_i + z_i^*) = 1 - \frac{\omega^2}{|z_i|^2} + 2i\frac{\omega}{|z_i|} \cos \arg z_i.$$

Analogt med vad vi har gjort tidigare definierar vi $|z_i| = \omega_0$, $\zeta = \cos \arg z_i$. Vi har då

$$\log |G(i\omega)| = \frac{1}{2} \log \left(\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + 4\frac{\omega^2}{\omega_0^2} \zeta^2 \right), \quad \arg G(i\omega) = \arctan \frac{2\frac{\omega}{\omega_0} \zeta}{1 - \frac{\omega^2}{\omega_0^2}}.$$

Lågfrekvensasymptoten är 0. Högfrekvensasymptoten är $2 \log \frac{\omega}{\omega_0}$, vilket är en rak linje med lutning 2 i en log-log-plot. Nära $\omega = \omega_0$ försvinner den kvadratiske termen, och Bodediagrammet beter sig som $\log \frac{\omega}{\omega_0}$, alltså en rak linje med lutning 1 i en log-log-plot. Denna asymptoten skär alla andra precis när $\omega = \omega_0$, varför detta är en dubbel brytpunkt för Bodediagrammet.

Ritning av Bodeplottar För att rita amplituddelen av ett Bodediagram, gör följande:

- Faktoriserar täljare och nämnare och identifiera poler och nollställen.
- Identifiera exponenten p .
- Identifiera K .
- Starta Bodediagrammet i $\log K$ och fortsätt i en rak linje med lutning $-p$.
- När du kommer till en reell pol, sänk lutningen med 1.
- När du kommer till beloppet av ett par av komplexa poler, sänk lutningen med 2.
- Gör samma för nollställena, men höj lutningen i stället.

Bodes relation Låt G vara minimumsfas, dvs. ha alla sina nollställen och poler i vänstre halvplan, och $G(0) > 0$. Då gäller att om $|G(i\omega)|$ i ett visst frekvensområde avtar med 20 dB per dekad (en dekad är en ökning i frekvens med en faktor 10), är $\arg G(i\omega) \approx -90^\circ$, och om $|G(i\omega)|$ avtar med 40 dB per dekad, är $\arg G(i\omega) \approx -180^\circ$.

Snabbhet och svängighet Vi kan med tidigare resultat se att om $G_O(i\omega)$ är nära 1 blir $G_C(i\omega)$ stor, och om $G_O(i\omega)$ är liten blir även $G_C(i\omega)$ liten. Vi kan också se att ett ekvivalent kriterium för bandbredden är $|G_O(i\omega) - 1| \leq \sqrt{2}$ för $\omega \geq \omega_B$.

Resonanstopp och fasmarginal Vi har

$$M_p \geq |G(i\omega_c)| = \frac{1}{2 \sin(\frac{1}{2}\phi_m)}.$$

Speciellt ger liten fasmarginal stort översläng.

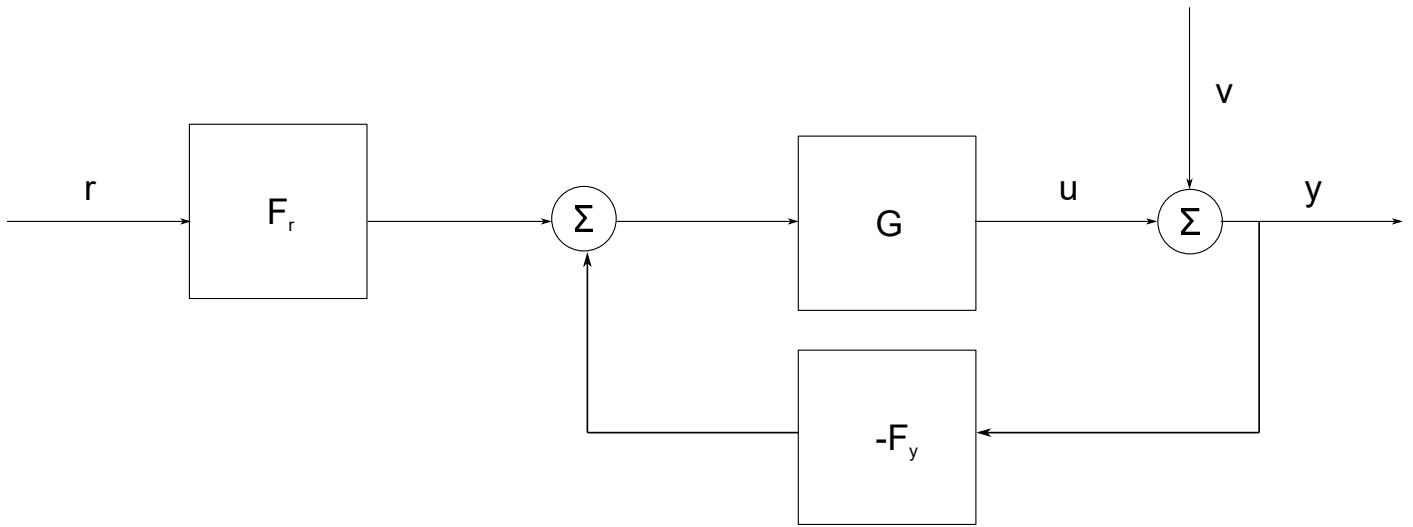
Frekvensanalys av känslighetsfunktionen Betrakta ett system som i figur 3. Om $F_r = F_y = F$, skulle systemet vara ekvivalent med det välkända negativt återkopplade systemet. Vi är nu intresserade av att studera systemets känslighet för störningen.

Vi har allmänt

$$Y = V + U = V + G(RF_r - YF_y), \quad Y = \frac{GF_r}{1 + GF_y} R + \frac{1}{1 + GF_y} V,$$

och känslighetsfunktionen ges då av

$$S = \frac{1}{1 + GF_y}.$$



Figur 3: Illustration av tänkt reglersystem.

Vi har

$$|S(i\omega)| = \left| \frac{1}{1 + G(i\omega)F_y(i\omega)} \right| \leq M_s,$$

vilket kan skrivas som

$$|G(i\omega)F_y(i\omega) - (-1)| \leq \frac{1}{M_s}.$$

Alltså måste Nyquistkurvan för GF_y vara utanför en cirkel med mittpunkt i -1 och radie $\frac{1}{M_s}$.

Robusthet Medan en modell kan ha en överförningsfunktion G , kan ett reellt system ha en överförningsfunktion $G' = G(1 + \Delta G)$, där ΔG är det relativa felet i överförningsfunktionen. Om man återkopplar systemet får man då

$$\begin{aligned} Y' &= \frac{G'F}{1 + G'F}R \\ &= \frac{G'F(1 + GF)}{(1 + GF)(1 + G'F)}R \\ &= \frac{GF(1 + \Delta G)(1 + GF)}{(1 + GF)(1 + G'F)}R \\ &= \frac{GF(1 + \Delta G + GF + \Delta GGF)}{(1 + GF)(1 + G'F)}R \\ &= \frac{GF\Delta G + FG(1 + GF + \Delta GGF)}{(1 + GF)(1 + G'F)}R \\ &= \frac{GF\Delta G + FG(1 + G'F)}{(1 + GF)(1 + G'F)}R \\ &= \left(\frac{GF\Delta G}{(1 + GF)(1 + G'F)} + \frac{FG}{1 + GF} \right) R. \end{aligned}$$

Om återkopplingen är $U = FE$ fås

$$\begin{aligned} Y' &= \frac{FG}{1 + GF}R + \frac{GF\Delta G}{(1 + GF)(1 + G'F)}R \\ &= (1 + S'\Delta G)Y, \end{aligned}$$

där S' är känslighetsfunktionen för det verkliga systemet. Det relativa felet i Y blir alltså $S'\Delta G$ blir det relativa felet i Y . Detta betyder att om systemets känslighetsfunktion är liten, får modellfel liten inverkan.

Robusthetskriteriet Antag att G ger ett system som är stabilt. Är då ett motsvarande verkligt system stabilt? För att svara på det, kan vi använda Nyquistkriteriet. Kravet är att $F_y G'$ ej får omsluta -1 , vilket vi anser som uppfyllt om avståndet mellan $F_y G'$ och $F_y G$ är mindre än avståndet mellan $F_y G$ och -1 . Detta kan formuleras som

$$|F_y(i\omega)G'(i\omega) - F_y(i\omega)G(i\omega)| < |F_y(i\omega)G(i\omega) + 1|$$

och slutligen robusthetskriteriet

$$\left| \frac{F_y(i\omega)G'(i\omega)}{1 + F_y(i\omega)G(i\omega)} \right| < \frac{1}{|\Delta G(i\omega)|}.$$

Vi definierar nu den komplementära känslighetsfunktionen

$$T = \frac{GF_y}{1 + GF_y} = 1 - S.$$

Normalt är ΔG ej känd, men vi har en övre skattning g av denna. Det verkliga systemet är då stabilt om

$$|T(i\omega)| < \frac{1}{g(\omega)},$$

vilket är ett tillräckligt, men ofta inte nödvändigt, villkor.

Nackdelar med återkoppling Återkoppling kan ge följande problem:

- Hög krets förstärkning ökar risken för instabilitet, speciellt vid modellfel.
- Om GF_y är stor, krävs stora styrsignaler. Det samma gäller i frekvensband där $|G|$ är liten.
- Mätbrus suger, typ.

Bodes sats Om G' har relativt gradtal större än 2, gäller att

$$\int_0^\infty d\omega \log |S(i\omega)| = \pi \sum_{k=1}^m p_k,$$

där p_k är polerna till G' i högre halvplan.

5 Kompensering

Ideen Vi ser att det enklaste sättet att konstruera en bra regulator på är att ändra konstruktionen av det öppna systemet. Vi bestämmer alltså regulatorn F utifrån krav på

- snabbhet, alltså skärffrekvens.
- dämpning, alltså fasmarginal.
- stationärt fel, alltså krav på $|G_O(0)|$.

Denna delen kommer att diskutera hur man kan designa en kompensator som får systemet att uppfylla kraven på dessa.

Fasavancerande länk För att höja fasan kan man använda en deriverande länk, alltså en regulator med överföringsfunktion

$$F = K(\tau_D s + 1).$$

Typiskt kan man inte låta deriveringen verka fullt ut, så överföringsfunktionen blir i stället på formen

$$F_{\text{lead}} = K \frac{\tau_D s + 1}{\beta \tau_D s + 1}.$$

Vi har

$$\arg F_{\text{lead}}(i\omega) = \arctan \frac{(1 - \beta)\tau_D\omega}{1 + \beta\tau_D^2\omega^2},$$

och får därmed att den maximala fasförskjutningen är

$$\phi_{\text{max}} = \arctan \frac{1 - \beta}{2\sqrt{\beta}}$$

för frekvensen

$$\omega = \frac{1}{\sqrt{\beta}\tau_D}.$$

När vi designar kompensatorn väljer vi typiskt att den maximala fasförskjutningen skall ske för $\omega = \omega_c$. När vi vet den önskade fashöjningen ϕ , väljs β som

$$\beta = 2 \tan^2(\phi) + 1 - 2\sqrt{\tan^2(\phi) + \tan^4(\phi)}.$$

Notera att hur mycket fasen behöver höjas beror på önskad fasmarginal, systemets egna fasförskjutning och ett bidrag från en eventuell fasfördröjande länk som kommer diskuteras sen. Notera även att man typiskt inte använder fasavancerande länkar för fashöjningar större än 60° .

Vidare väljer vi

$$\tau_D = \frac{1}{\sqrt{\beta}\omega_c},$$

så att maximal fashöjning sker vid skärfrekvensen. Notera att den fashöjande länken även kan behöva kompensera för fasfördröjning i den andra länken.

Lågfrekvensförstärkning Lågfrekvensförstärkning kan ta bort stationärt fel. För att lågfrekvensförstärka kan man använda en integrerande länk, alltså en regulator med överförningsfunktion

$$F = \frac{\tau_I s + 1}{\tau_I s}.$$

Denna har dock oändligt hög förstärkning för låga frekvenser, vilket i praktiken kan vara svårt att realisera. Då kan man i stället använda en fasretarderande länk, alltså en regulator med överförningsfunktion

$$F_{\text{lag}} = \frac{\tau_I s + 1}{\beta\tau_I s + \gamma}.$$

Vi har

$$\arg F_{\text{lag}}(i\omega) = -\arctan \frac{(1 - \gamma)\tau_I\omega_c}{\gamma + \tau_I^2\omega_c^2}.$$

γ väljs baserad på önskad lågfrekvensförstärkning. Vi ser nu att fasfördröjningen beror på produkten $\tau_I\omega_c$. Som tumregel sätts denna till att vara 10, vilket ger en fasfördröjning på 6° som den fashöjande länken även kommer behöva kompensera för.

Placering av skärfrekvens För att placera skärfrekvensen lägger man helt enkelt på ett K så att systemets förstärkning vid skärfrekvensen är 1.

Lead-länkens förstärkning ges av

$$\begin{aligned}
|F_{\text{lead}}(i\omega)| &= K \left| \frac{i\tau_D\omega + 1}{i\beta\tau_D\omega + 1} \right| \\
&= \frac{K}{1 + \beta^2\tau_D^2\omega^2} |1 + \beta\tau_D^2\omega^2 + i(1 - \beta)\tau_D\omega| \\
&= \frac{K}{1 + \beta^2\tau_D^2\omega^2} \sqrt{(1 + \beta\tau_D^2\omega^2)^2 + (1 - \beta)^2\tau_D^2\omega^2} \\
&= \frac{K}{1 + \beta^2\tau_D^2\omega^2} \sqrt{1 + 2\beta\tau_D^2\omega^2 + \beta^2\tau_D^4\omega^4 + (1 - 2\beta + \beta^2)\tau_D^2\omega^2} \\
&= \frac{K}{1 + \beta^2\tau_D^2\omega^2} \sqrt{1 + \beta^2\tau_D^2\omega^2 + \beta^2\tau_D^4\omega^4 + \beta^2\tau_D^2\omega^2} \\
&= \frac{K}{1 + \beta^2\tau_D^2\omega^2} \sqrt{(1 + \beta^2\tau_D^2\omega^2)(1 + \tau_D^2\omega^2)} \\
&= K \sqrt{\frac{1 + \tau_D^2\omega^2}{1 + \beta^2\tau_D^2\omega^2}}.
\end{aligned}$$

Speciellt, vid skärfrekvensen, fås

$$\begin{aligned}
|F_{\text{lead}}(i\omega_c)| &= K \sqrt{\frac{1 + \frac{1}{\beta}}{1 + \beta}} \\
&= K \sqrt{\frac{\beta + 1}{\beta + \beta^2}} \\
&= \frac{K}{\sqrt{\beta}}.
\end{aligned}$$

Lag-länkens förstärkning ges av

$$\left| \frac{i\tau_I\omega + 1}{i\tau_I\omega + \gamma} \right| = \sqrt{\frac{\tau_I^2\omega^2 + 1}{\tau_I^2\omega^2 + \gamma^2}}.$$

Med valet av $\tau_I^2\omega^2$ som allmänt stort, blir detta ungefär 1. Skärfrekvensen kan därmed placeras genom att välja

$$K = \frac{\sqrt{\beta}}{|G(i\omega_c)|}.$$

6 Tillståndsrepresentationer

Ideen Den fundamentala ideen vi vill åt nu är att representera ett systems tillstånd på ett annat sätt än bara utsignalens tidsutveckling. Vi vill även inkludera andra storheter i systemet.

Kort om notation y kommer fortfarande att representera utsignalen och u insignalen. x kommer vara systemets tillstånd, karakteriserad av ett visst antal storheter. Däremot skulle en eller flera av dessa vara vektorer. Just för att detta kan variera från system till system och betraktning till betraktning, väljer jag att vara ambivalent med notationen.

Linjarisering Verkliga system är ofta olinjära, men vi ska försöka behandla dem som linjära ändå.

Betrakta ett system

$$\dot{x} = f(x, u), \quad \dot{y} = g(x, u).$$

Anta konstant styrsignal u_0 , och antag att systemet då tenderar mot ett stationärt tillstånd \mathbf{x}_0 . Denna punkten uppfyller då

$$f(x_0, u_0) = \mathbf{0}, \quad h(x_0, u_0) = y_0.$$

När vi linjäriserar, betraktar vi små variationer $\Delta x, \Delta u, \Delta y$ kring denna punkten. Vi får

$$\frac{d\Delta x}{dt} = f(x_0 + \Delta x, u_0 + \Delta u) - f(x_0, u_0) = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial u} \Delta u = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial u} \Delta u.$$

Observera att $\frac{\partial f}{\partial x}$ allmänt är en matris.

På samma sätt fås

$$\frac{d\Delta y}{dt} = h(x_0 + \Delta x, u_0 + \Delta u) - y_0 = h(x_0, u_0) - y_0 + \vec{\nabla}_x h \Delta x + \frac{\partial h}{\partial u} \Delta u = \vec{\nabla}_x h \Delta x + \frac{\partial h}{\partial u} \Delta u.$$

Det totala systemet

$$\begin{aligned} \frac{d\Delta x}{dt} &= \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial u} \Delta u, \\ \frac{d\Delta y}{dt} &= \vec{\nabla}_x h \Delta x + \frac{\partial h}{\partial u} \Delta u \end{aligned}$$

är alltså linjärt för små ändringar.

Representation av linjära system Tidsutvecklingen av ett system tillstånd kan i linjära fall skrivas som

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

där u beskriver en styrsignal. Utsignalen är typiskt på formen

$$y = Cx + Du.$$

Linjära system i Laplacedomänen Genom att Laplacetransformera ekvationen som beskriver ett systems tillstånd fås

$$sX = AX + BU,$$

givet att systemets starttillstånd är 0. Detta kan skrivas som

$$X = (sI - A)^{-1}BU.$$

Insatt i uttrycket för Y fås

$$Y = CX + DU = (C(sI - A)^{-1}B + D)U,$$

och vi identifierar överföringsfunktionen som

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

Man kan visa att

$$A^{-1} = \frac{1}{|A|} A^\dagger,$$

vilket ger

$$G = \frac{1}{|sI - A|} C(sI - A)^\dagger B + D$$

Poler i tillståndsrepresentation Om inte $|sI - A|$ och $C(sI - A)^\dagger$ delar faktorer, blir då systemets poler egenvärdena till A .

Stabilitet Baserad på detta kan vi använda kunskapen om polernas inverkan på systemets stabilitet för att dra följande slutsats: Ett linjärt system är stabilt om och endast om alla egenvärden till A har negativa realdelar.

Lösning av system i representation Betrakta ett system på formen

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \quad y = Cx.$$

Vi använder att

$$\frac{d}{dt}e^{-At} = -Ae^{-At},$$

vilket ger

$$\begin{aligned} e^{-At}\dot{x} &= e^{-At}Ax + e^{-At}Bu, \\ \frac{d}{dt}(e^{-At}x) &= e^{-At}Bu, \\ x &= x_0e^{At} + \int_0^t d\tau e^{-A(t-\tau)}Bu. \end{aligned}$$

Styrbarhet Ett tillstånd x är styrbart om det finns en insignal som kan styra det motsvarande systemet från 0 till x på ändlig tid. Ett system är styrbart om alla tillståndsvektorer är styrbara.

Test av styrbarhet Vi noterar först att Cayley-Hamiltons sats ger

$$A^n + \sum_{i=1}^n a_i A^{n-i} = 0,$$

där a_i är koefficienterna i A 's karakteristiska polynom. Därmed kan alla potenser av A av högre ordning skrivas som en linjärkombination av potenser av A upp till och med $n-1$, vilket betyder att

$$e^{At} = \sum_{i=0}^{n-1} f_i(t)A^i$$

för några funktioner f_i .

Betrakta nu ett system och dets representation. Om systemets starttillstånd är 0, medför detta

$$x = \left(\sum_{i=0}^{n-1} \gamma_i(t)A^i \right) B$$

där

$$\gamma_i = \int_0^t d\tau f_i(\tau)u(\tau).$$

Om x är styrbart, måste det enligt resonemanget ovan vara en linjärkombination av de olika $A^i B$. Detta betyder att x måste ligga i bildrummet till styrbarhetsmatrisen

$$\mathcal{S} = [B, AB, \dots, A^{n-1}B].$$

Systemet är styrbart om $\det(\mathcal{S}) \neq 0$.

Observerbarhet Ett tillstånd x är icke observerbart om utsignalen y är identiskt noll då initialvärdet är x och insignalen identisk noll. Ett system är observerbart om inga tillståndsvektorer är icke-observerbara.

Test av observerbarhet Vi vill nu testa om ett tillstånd x_0 är observerbart. Om vi har styrsignal $u = 0$, gäller det att

$$x = e^{At}x_0, y = Ce^{At}x_0, \frac{d^n y}{dt^n} = CA^n e^{At}x_0.$$

Speciellt är $y = 0$ för alla t om

$$y(0) = Cx_0 = 0, \frac{dy}{dt} = CAx_0 = 0, \dots, \frac{d^{n-1}y}{dt^{n-1}} = CA^{n-1}x_0 = 0,$$

eftersom om detta stämmer, ger Cayley-Hamiltons sats att även högre ordningens derivator av y kommer vara 0. Därmed ligger de icke-observerbara tillstånden i nollrummet till observerbarhetsmatrisen

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}.$$

Systemet är observerbart om $\det(\mathcal{O}) \neq 0$.

Minimalhet En tillståndsbeskrivning är en minimal realisation av ett insignal-utsignalsamband om och endast om den är både styrbar och observerbar.

Proportionell återkoppling Antag att vi återkopplar systemet med $u = l_0 r - Lx$. Om utsignalen ej beror av u kan det återkopplade systemet skrivas som

$$\dot{x} = (A - BL)x + Bl_0 r, \quad y = Cx.$$

Detta ger att systemets poler ges av egenvärdena till $A - BL$. Det finns n såna, och L har n parametrar. Om systemet är styrbart kan dets poler därmed placeras godtyckligt vid lämpligt val av L .

Vidare studerar vi ett statistiskt scenario, som uppfyller

$$\begin{aligned} x &= (BL - A)^{-1} Bl_0 r, \\ y &= C(BL - A)^{-1} Bl_0 r + D l_0 r - DL(BL - A)^{-1} Bl_0 r \\ &= ((C - DL)(BL - A)^{-1} B + D) l_0 r, \end{aligned}$$

som är proportionellt mot l_0 . Genom att välja

$$l_0 = ((C - DL)(BL - A)^{-1} B + D)^{-1}$$

fås att $y = r$ när systemet är statiskt. Detta kräver dock att man känner de involverade matriserna och att inga störningar påverkar systemet. Därför inför vi I-reglering.

I-reglering När vi I-reglerar, inför vi extra tillstånd

$$x_{n+1} = \int_0^t d\tau e.$$

Detta ger

$$\dot{x}_{n+1} = r - y = r - Cx.$$

Då kan vi utvidga modellen till

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{x}_{n+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix} \begin{bmatrix} x \\ x_{n+1} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r = A\mathbf{x} + Bu + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r.$$

Strategin är nu att återkoppla det nya systemet med återkoppling på formen

$$u = -Lx - l_{n+1}x_{n+1} = -L\mathbf{x}.$$

Då kan L väljas så att $A - BL$ får önskade egenvärden. Stationärt har vi

$$\dot{x} = 0, \quad \dot{x}_{n+1} = 0.$$

Skattning av tillstånd Betrakta ett system där man endast kan mäta utsignal och insignal. I såna fall kan man inte göra tillstånd återkoppling direkt. Däremot kan man försöka göra detta genom att skatta systemets tillstånd.

Mer precist, antag att vi har en modell

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad x(0) = x_0$$

som vi simulerar enligt

$$\dot{\hat{x}} = A\hat{x} + Bu, \quad \hat{y} = C\hat{x}, \quad \hat{x}(0) = \hat{x}_0.$$

Simuleringsfelet är

$$y - \hat{y} = y - C\hat{x}.$$

Vi försöker återkoppla det simulerade systemet med hjälp av simuleringsfelet. Detta beskrivs av

$$\dot{\hat{x}} = A\hat{x} + Bu + K(y - C\hat{x}) = (A - KC)\hat{x} + Bu + Ky.$$

Systemet ovan kallas för en observatör för det ursprungliga systemet.

Skattningsfelet \tilde{x} beskrivs av

$$\dot{\tilde{x}} = \dot{x} - \dot{\hat{x}} = Ax + Bu - A\hat{x} - Bu - K(Cx - C\hat{x}) = (A - KC)\tilde{x}.$$

Detta har lösning

$$\tilde{x} = e^{(A-KC)t}\tilde{x}(0).$$

Skattningsfelet tenderar alltså mot 0 om egenvärdena till $A - KC$ är negativa, och hur snabbt det tenderar mot 0 beror av egenvärdenas belopp. Det finns en kompromiss mellan hur snabbt skattningsfelet avtar och hur känsligt systemet är för störningar.

För att studera huruvida skattningsfelet kan fås att avta godtyckligt snabbt, studerar vi egenvärdena till $A - KC$. Detta är samma som för $A^T - C^T K^T$. Att få detta system att avta godtyckligt snabbt är en fråga om att kunna placera polerna för matrisen ovan godtyckligt. Detta är exakt samma problem som vi har studerat med proportionell återkoppling, och vi vet då att systemet kan fås att avta godtyckligt snabbt om systemet baserad på A^T och C^T är styrbart.

Kalmanfilter Betrakta ett system vars utsignal störs av ett mätfel v och vars tillstånd störs av ett fel w . Skattningsfelet beskrivs av

$$\dot{\tilde{x}} = (A - KC)\tilde{x} + e - Kv.$$

Stora K ger som vi ser höga mätfel, men det ger även snabba system. Optimeringen där görs av ett Kalmanfilter. Vi väljer då systemet så att egenvärdena till $A - BL$ har lägre belopp än egenvärdena till $A - KC$.

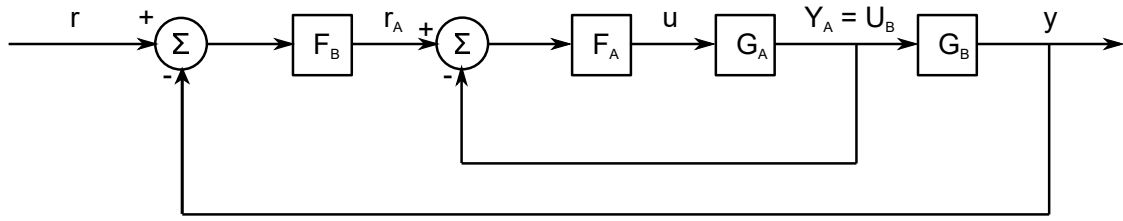
Vid återkoppling av såna system är överföringsfunktionen för det slutna systemet

$$G = C(sI - (A - BL))^{-1}Bl_0,$$

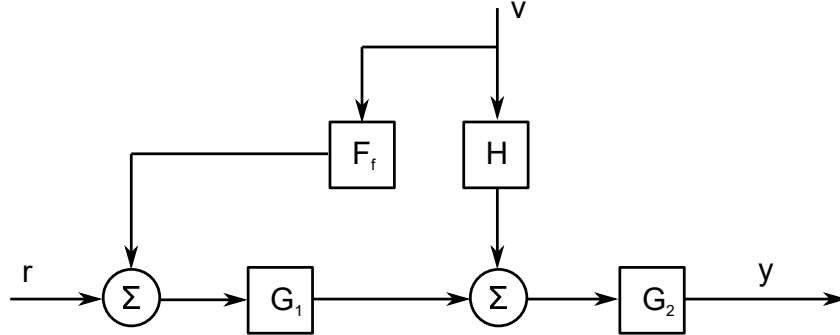
alltså likadant som tidigare.

7 Alternativa regulatorstrukturer

Kaskadreglering Ideen bakom kaskadreglering är att försöka reglera ett system med hjälp av en mellanliggande signal. Detta beskrivs av blockschemat i figur 4. Kretsen i mitten reglerar mellansignalen u mot ett visst värde U_B . Detta är värdet som regulatorn F_B vill ha. Om detta skall fungera, måste den inre regleringen vara snabbare än den yttre.



Figur 4: Blockschemat av kaskadreglering.



Figur 5: Blockschemat av framkoppling.

Framkoppling Idéen bak framkoppling är att försöka reglera bort en mätbar störsignal. Detta visas i blockschemat i figur 5. Vi önskar nu välja F_f så att störsignalen ej har någon inverkan.

Vi har

$$Y = G_2(HV + G_1(R + F_f H)) = G_2((H + G_1 F_f)V + G_1 R).$$

Därmed kommer V -termen försvinna om vi väljer

$$F_f = -\frac{H}{G_1}.$$

En fördel med detta är att man kan reglera sitt system innan störningen får en inverkan. Det kan dock uppstå problem med att man får rena deriveringstermer, men dessa kan lösas genom att approximera F_f nära $s = 0$.

Otto Smith-regulatorn Syftet med Otto Smith-regulatorn är att reglera system med tidsfördröjning. Färdöröjningen från en sån ökar med ω , vilket gör det svårt att reglera.

För att reglera detta, bestäm först ett F så att ditt äntliga system (till exempel det slutna reglersystemet) får bra egenskaper om man ignorerar tidsfördröjningen. Om reglering av det icke-fördröjda öppna systemet G ger en regulator F , inför nu regulatorn

$$F' = \frac{F}{1 + (1 - e^{-sT})FG},$$

där T är systemets fördröjning. Detta kommer ge

$$G_c = \frac{FG}{1 + FG} e^{-sT}.$$

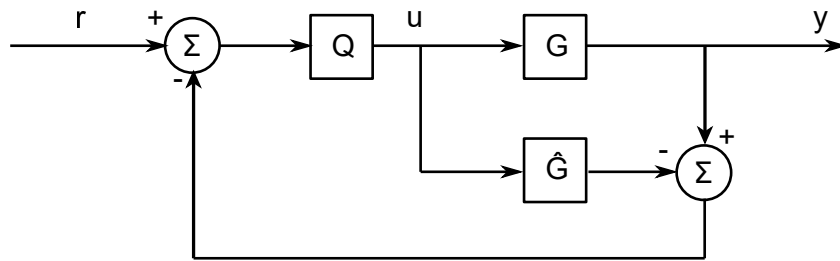
Internal Model Control Idéen bak IMC illustreras i figur 6. Vi har här introducerat \hat{G} som vår modell av systemet.

Flödesschemat ger

$$U = \frac{Q}{1 - Q\hat{G}}(R - Y).$$

Om vi jämför detta med standardvarianten $U = F(R - Y)$, ser vi att

$$F = \frac{Q}{1 - Q\hat{G}}.$$



Figur 6: Blockschema av IMC.

Detta betyder att alla problem vi har studerat kan omformuleras som ett IMC-problem.

Varför inför vi detta? Först av allt kan vi på något sätt hitta alla regulatorer som stabiliserar systemet genom att titta på stabila Q , vad nu det betyder. Vidare är det enkelt att studera det slutna systemet, ty om vi har en perfekt modell, försvinner återkopplingen, och $Y = GQR$. Detta är linjärt i Q , medan fall vi har studerat tidigare inte har varit linjära i F .

Betrakta specialfallet där vi vil ha $Y = HR$. Då väljer vi $Q = \frac{H}{G}$. Detta ger

$$F = \frac{1}{G} \frac{H}{1 - H},$$

som är resultatet vi känner från tidigare. På liknande sätt kan andra reglerproblem formuleras med IMC.