

Kompendium i

Monovariable systemer og signaler

Trond Andresen

**Institutt for teknisk kybernetikk
NTNU
høst 1998**

(leses sammen med *Signals and Systems* av Oppenheim, Willsky, Young)

1	Residueregning, invers Laplace- og Fouriertransformasjon	1
1.1	Residueregning	1
1.2	Invers Laplacetransformasjon	2
1.2.1	Invers Laplacetransformasjon - et eksempel	5
1.3	Invers Fouriertransformasjon	6
1.4	Bemerkning om notasjon	7
2	Diskrete signaler og systemer	9
2.1	Innledning	9
2.2	Representasjon av diskrete systemer	9
2.2.1	Merknad om folding for kontinuerlige signaler	10
2.2.2	Merknad om frekvens for diskrete signaler	11
2.3	Den diskrete Fouriertransformasjon	11
2.3.1	Eksempel: Utregning av en diskret Fourier-transform	13
2.4	Z-transformasjonen	15
2.4.1	Bemerkning om notasjon	17
2.4.2	Invers Z-transformasjon - utledning	17
2.4.3	Eksempel: Utregning av en invers Z-transform	18
2.5	Z-transformen som transferfunksjon. Differensligning	19
3	Diskretisering av kontinuerlig system	21
3.1	Innledning	21
3.2	Diskretisering via tilstandsromform	21
3.3	Om poler i det diskrete og kontinuerlige system	23
3.3.1	Polenes plassering og diskret tidsrespons	23
3.3.2	Polenes plassering når et kontinuerlig system diskretiseres	25
4	Diskretisering og regulering av kontinuerlig system	27
4.1	Innledning	27
4.2	To diskretiseringsmetoder	27
4.2.1	Metode 1: Diskretisering via tilstandsromform	27
4.2.2	Metode 2: Diskretisering via impulsresponsen $h[k]$	29
4.3	Blokkdiagram for diskrete systemer	31
4.4	Balchens “Q-transformasjon”	32
4.5	ARMAX-modell	33
5	Regulering v.h.a. polplassering	35
5.1	Definisjoner, notasjon	35

5.2 Den hurtigste utgangssignal-dead-beat-regulator	36
5.2.1 To betydninger av symbolet z	38
5.2.2 Polene til det lukkede system	38
5.2.3 Eksempel 1: Dead-beat-regulator	39
5.2.4 Eksempel 2: En dead-beat-regulator som skaper problemer	40
5.2.5 Robusthet er lav med den hurtigste dead-beat-regulatoren	43
5.3 Robust utgangssignal-dead-beat-regulator	44
5.3.1 Følgeegenskaper for den robuste regulatoren	47
5.4 Om graden på F- og G-polynomet	49
5.5 En generell lineær regulator	50
5.6 Stabilitet og frekvensanalyse	51
5.6.1 Nyquists stabilitetskriterium for diskrete systemer	51
5.6.2 Eksempel: Stabilitet i kontinuerlig og diskret 1. ordens system	53
5.6.3 Bode(AFF) - diagram	56
5.6.4 Nyquists stabilitetskriterium for åpent ustabile systemer	56
5.6.5 Frekvensanalyse av systemer med dead-beat-regulator	57
5.7 Kildehenvisning	60
6 Stokastiske prosesser	61
6.1 Innledning	61
6.2 Matematisk grunnlag i sannsynlighetsregning	65
6.2.1 Tethetsfunksjon og fordelingsfunksjon	65
6.2.2 Betinget sannsynlighet og betinget tethetsfunksjon	66
6.2.3 Forventningsverdi, kovarians og korrelasjon	69
6.2.4 Et eksempel på korrelasjon	72
6.2.5 Statistisk uavhengighet og korrelasjon	73
6.2.6 Normalfordelingen	74
6.2.7 Simultant normalfordelte variable	76
6.3 Stokastiske prosesser beskrevet i tidsplanet	77
6.3.1 Beskrivelse ved middelverdi og korrelasjonsfunksjon	78
6.3.2 Signaler som er en sum av stokastiske og deterministiske komponenter	83
6.3.3 Bruk av autokorrelasjonsfunksjoner og krysskorrelasjonsfunksjoner	84
6.3.4 Autokorrelasjon direkte fra simultan tethetsfunksjon: et eksempel	85
6.3.5 Oppsummering av egenskaper til korrelasjonsfunksjoner	88
7 Stokastiske prosesser, effektspektra	89
7.1 Innledning	89
7.2 Sammenhengen mellom autoeffektspektrum og autokorrelasjon	90
7.3 Krysseffektspektrum og krysskorrelasjon	96
7.4 Hvit støy	99
7.4.1 Integralet av hvit støy	101

7.4.2	Rullebevegelsen for skip - bruk av støy i modellering	102
8	Stokastiske prosesser i lineære dynamiske systemer	105
8.1	Grunnleggende sammenhenger	105
8.2	Estimering av transferfunksjoner i systemer som er utsatt for støy	110
8.2.1	Pådragsstøy, prosess-støy	110
8.2.2	Målestøy	111
8.2.3	Støy i tilbakekoplede systemer	112
8.3	Diskrete stokastiske prosesser og diskrete systemer	114
8.3.1	Diskret prosess, korrelasjon, effektspektrum	114
8.3.2	Diskret hvit støy	116
8.3.3	Diskretisering når kontinuerlig hvit støy påvirker et kontinuerlig lineært system	116
8.4	Diskrete tilstandsrommodeller av monovariable stokastiske dynamiske systemer	119
9	Estimering av spektra, korrelasjons- og transferfunksjoner	121
9.1	Diskret Fourier-transform (“DFT”) av endelig måleserie	121
9.2	Estimering av spektra	123
9.2.1	Om forventningsfeilen til estimatet av effektspektret	126
9.2.2	Varians kontra forventningsfeil	129
9.2.3	Varians kontra forventningsfeil: Et eksempel	133
9.2.4	Krysseffektspektra	134
9.3	Estimering av korrelasjonsfunksjoner	135
9.4	Estimering av transferfunksjoner	137
10	Signalbehandling: Tilpassing, filtrering	141
10.1	Analog til digital omsetting	141
10.1.1	Tasting av signaler	141
10.1.2	Kvantisering av signaler	143
10.2	Digitale filtre	144
10.3	Overgang fra analogt til digitalt filter	145
10.3.1	Syntese av analogt Butterworth lavpassfilter	147
10.3.2	Fleire lav-ordens filtre eller et med høy orden?	148
10.3.3	Syntese av diskret (digitalt) lavpassfilter	149
10.3.4	Realisering av andre typer filtre fra lavpassfilter	151
10.4	"Wild-point"- utplukking	154
11	Regulering av stokastiske systemer - optimale regulatorer med gitt struktur	155
11.1	Godhetskriterier for regulatorer	155

11.2 Optimalisering av regulatorer med fast struktur	156
12 Regulering av stokastiske systemer - optimale regulatorer med optimal struktur	167
12.1 Regulatoralgoritme	167
12.2 Optimal m-skritts prediktor	170
12.3 Sammenheng MV- og dead-beat-regulering	172
12.3.1 Minimum-energi-regulator	172
12.4 Robust MV-regulator	173
12.5 Lagerstyringseksemplet	175

Kapittel 1: Residueregning, invers Laplace- og Fouriertransformasjon

1.1 Residueregning

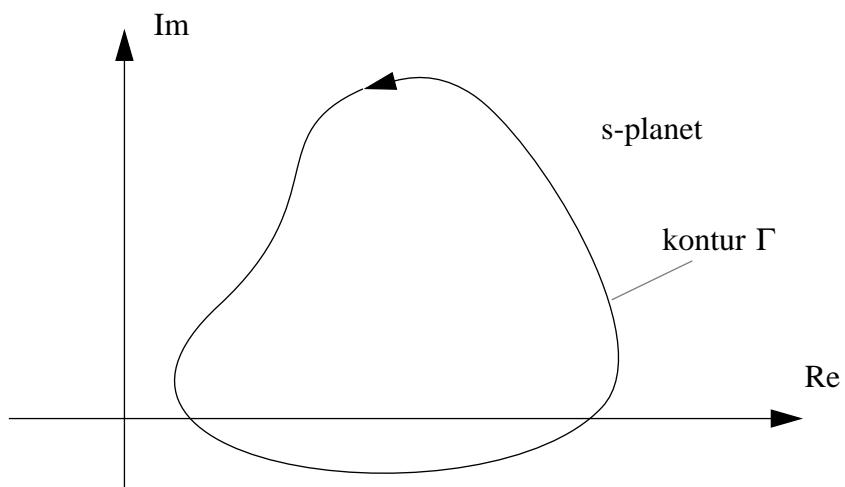
Det skal vises hvordan man kan finne den inverse Laplace- og Fouriertransform for funksjoner som er kontinuerlige i tidsplanet (tilsvarende for tidsdiskrete funksjoner i et seinere kapittel). Vi skal først konsentrere oss om Laplacetransformen, og etterpå se hvordan man kan betrakte Fourier-transformen som en Laplacetransform, og dermed bruke samme løsningsmetode for å finne den inverse Fouriertransformen.

(I et seinere kapittel skal vi se hvordan metoden som presenteres i dette kapittelet også kan brukes for å finne *invers z-transform* og *invers diskret Fouriertransform*, som vi bruker for å beskrive *diskrete* signaler.)

Den **inverse Laplacetransformasjon** $h(t)$ av en funksjon $H(s)$ er

$$h(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} H(s)e^{ts}ds \quad (1.1)$$

Merk at vi i dette kurset bruker små bokstaver for funksjoner i t -planet, og store i s -planet. σ er slik at den loddrette linjen $Re(s) = \sigma$ i det komplekse plan ligger innafor konvergensområdet til Laplacetransformen. Integralet (1.1) er vanskelig å løse direkte. Vi skal nå forklare hvordan man kan benytte **residueregning** for å løse (1.1). Denne metoden baserer seg på resultater fra kompleks funksjonsteori. Det dreier det seg om å finne integralet av en funksjon $F(s)$ i en kompleks variabel s , når integrasjonsbanen Γ er en lukket kontur *mot* urviseren i det komplekse plan som vist på figur 1.1.



Figur 1.1

Det kan da vises at hvis $F(s)$ er **analytisk** innafor konturen (“analytisk” $\Leftrightarrow F(s)$ og alle deriverte av $F(s)$, eksisterer overalt innafor konturen), så er konturintegralet

$$\oint_{\Gamma} F(s) ds = 0 \quad (1.2)$$

Hvis $F(s)$ derimot har en pol a av orden m innafor konturen Γ , slik at $F(s)$ kan skrives som

$$F(s) = \frac{G(s)}{(s-a)^m}, \text{ der } G(s) \text{ ikke har pol(er) i } s = a, \quad (1.3)$$

så sier den såkalte **residuesatsen** at

$$\frac{1}{2\pi j} \oint_{\Gamma} F(s) ds = \text{Res}[F(s)] = \frac{1}{(m-1)!} \left[\frac{d^{m-1}}{ds^{m-1}} [(s-a)^m F(s)] \right]_{s=a} \quad (1.4)$$

der $\text{Res}[F(s)]$ (uttales “residuet til $F(s)$ i a ”).

For $m = 1$ blir (1.4) atskillig enklere

$$\frac{1}{2\pi j} \oint_{\Gamma} F(s) ds = \text{Res}[F(s)] = [(s-a)F(s)]_{s=a} \quad (1.5)$$

Hvis det er flere poler a_i innafor konturen Γ har vi

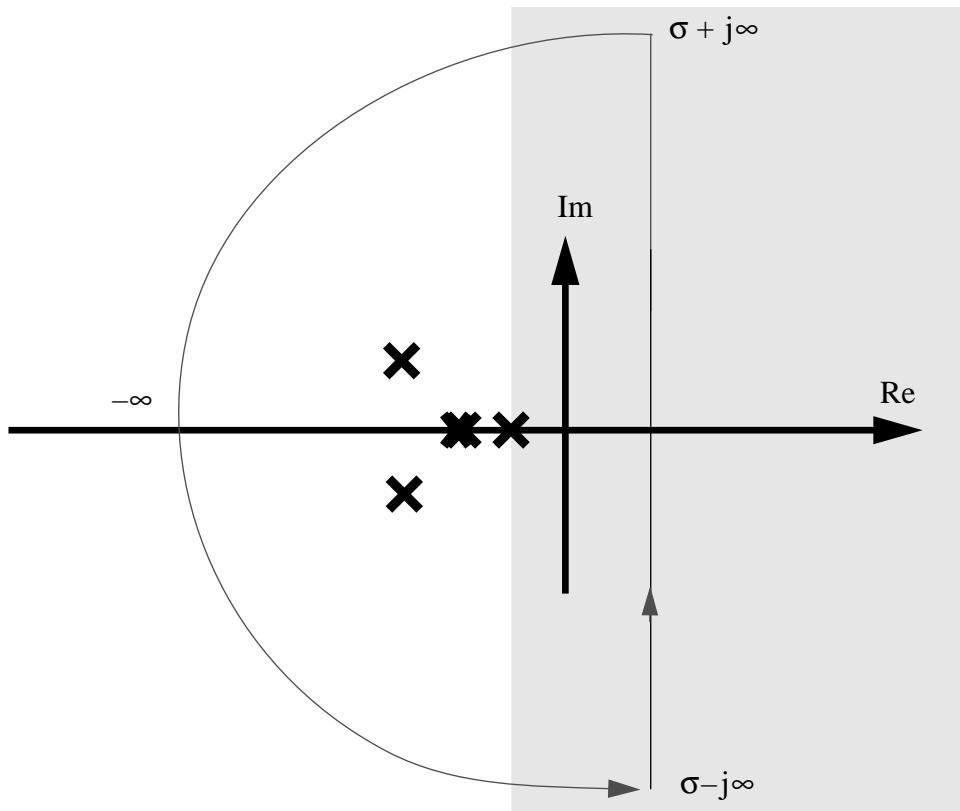
$$\frac{1}{2\pi j} \oint_{\Gamma} F(s) ds = \sum_i \text{Res}[F(s)] \quad (1.6)$$

I (1.6) anvender vi (1.5) eller (1.4) avhengig om polen a_i er enkel eller av høyere orden.

1.2 Invers Laplacetransformasjon

Hvordan kan så dette brukes for å finne en invers Laplacetransform? Venstresida i (1.6) ligner på (1.1), med $F(s) = H(s)e^{ts}$. Problemets er integrasjonsbanen til (1.1), som ikke er lukket. Vi må derfor bruke et knep for å lukke integrasjonsbanen. Betrakt figur 1.2:

1. I det generelle tilfelle skal høyresida i (1.4) være $\lim_{s \rightarrow a} \frac{1}{(m-1)!} \left(\frac{d^{m-1}}{ds^{m-1}} [(s-a)^m F(s)] \right)$, men for de funksjoner vi vil komme borti, trenger vi ikke foreta grenseovergang.



Figur 1.2

Gitt integranden $H(s)e^{ts}$ i (1.1). Et sett poler i denne integranden er antydet i figur 1.2. Et mulig konvergensområde er angitt ved skravering. De oppgitte poler og det angitte konvergensområdet forteller at Laplace-transformen svarer til et høyresidig tidsforløp (se notat 2). Polenes beliggenhet forteller at dette høyresidige tidsforløp er stabilt. Som det framgår av figur 1.2 har vi modifisert integrasjonsbanen. Vi har føyd til en uendelig stor halvsirkel inn i venstre halvplan, for å få til en lukket kontur.

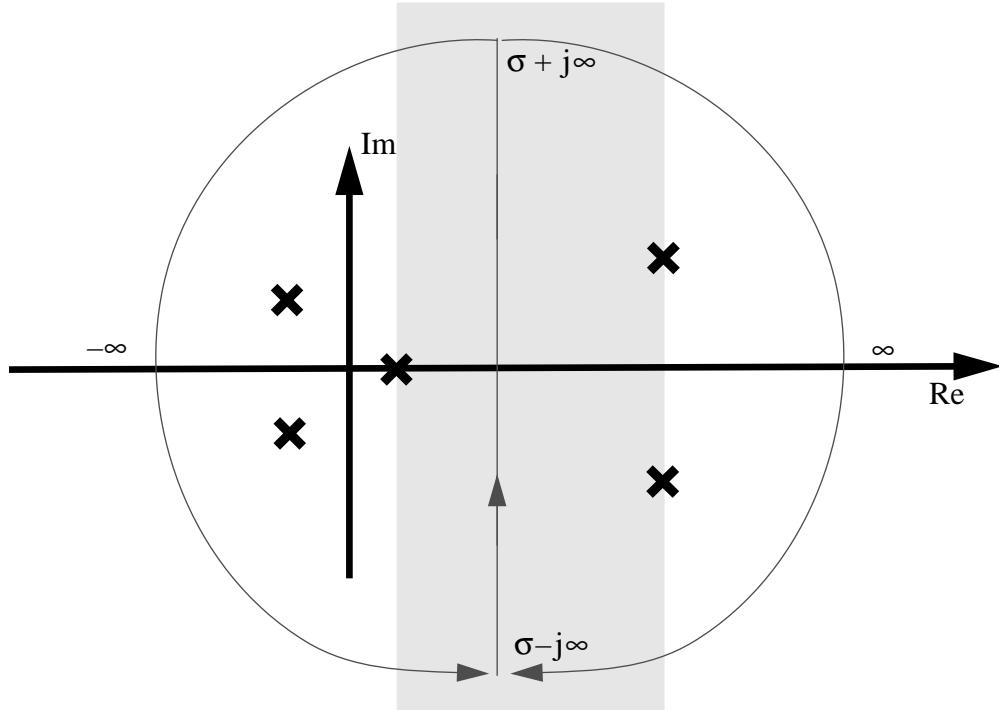
Dette kan vi tillate oss uten å endre verdien av integralet (1.1), så lenge $|H(s)e^{ts}| \rightarrow 0$ når $|s| \rightarrow \infty$.

Hvis integranden er null så lenge vi befinner oss på den uendelig store halvsirkelen, gir den intet bidrag til integralet. Hvis $H(s)$ er et rasjonalt uttrykk, så er dette oppfylt (dette er ikke helt triviert og må bevises, det gjøres ikke her) når graden av nevneren er minst en høyere enn graden av telleren. Vi kaller da $H(s)$ **strentg proper**. $H(s)$ vil alltid være strengt proper for de kontinuerlige signaler og systemer vi skal arbeide med. Vi kan altså konkludere med at den loddrette integrasjonsbanen i (1.1) uten videre kan lukkes med en uendelig stor halvsirkel, uten at dette berører verdien av integralet. **Dermed kan vi bruke residuesatsen ved invers Laplacetransformasjon.**

Hvis vi har dobbeltsidige eller venstresidige tidsforløp, blir det litt annerledes.

I figur 1.3 har vi poler i begge halvplan. Siden konvergensområdet i figur 1.3 er ei stripe, svarer dette til et dobbeltsidig tidsforløp (se læreboka¹). Den høyresidige delen av tidsforløpet er i dette eksemplet ustabilt, siden en av polene som svarer til det høyresidige tidsforløpet

1. Når det her og seinere henvises til "læreboka", menes *Signals and Systems* av Oppenheim, Willsky, Young.



Figur 1.3

ligger i høyre halvplan. (Dette medfører at den imaginære akse ligger utafor konvergensstripa; vårt tidsforløp er i dette tilfelle ikke Fouriertransformerbart). Vi samler residuene i to grupper, de som svarer til poler til venstre for den loddrette linja $s = \sigma$ (som må ligge et sted i konvergensområdet), og de som svarer til poler til høyre for den. Konturintegralet som skal omslutte poler til høyre for $s = \sigma$ må nå dannes ved hjelp av en uendelig stor halvsirkel inn i *høyre* halvplan. Denne konturen får da motsatt dreieretning av den forrige, *med* urviseren. Dette fører til den enkle endring i (1.4) og (1.5) at vi må bytte fortegn. Dermed er vi nå i stand til angi den endelige formel for invers Laplacetransformasjon ved hjelp av residueregning:

$$h(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} H(s)e^{ts} ds = \begin{cases} \mu_1(t) \sum_{i1} \operatorname{Res}[H(s)e^{ts}] & s = a_{i1} \\ -\mu_1(-t) \sum_{i2} \operatorname{Res}[H(s)e^{ts}] & s = a_{i2} \end{cases} \quad (1.7)$$

der a_{i1} svarer til poler til venstre for den loddrette linja $s = \sigma$, og a_{i2} poler til høyre for den.

1.2.1 Invers Laplacetransformasjon - et eksempel

Nå skal vi demonstrere metoden med et eksempel. Vi bruker symbolene $h(t)$ og $H(s)$, siden tidsfunksjonen svært ofte er en impulsrespons og den tilsvarende funksjonen av s er en transferfunksjon. Men dette gjelder sjølsagt for et vilkårlig Laplacetransform-par

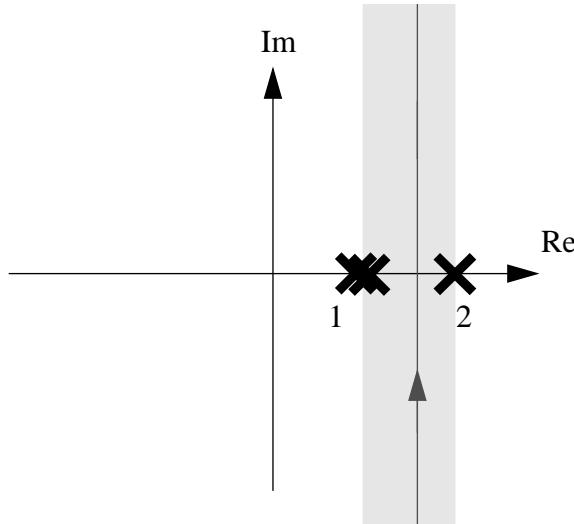
$$\xrightarrow{L} \\ x(t) \leftrightarrow X(s)$$

Gitt $H(s) = \frac{1}{(s-1)^2(1-0.5s)}$ (1.8)

Dette må vi først bringe på formen

$$H(s) = \frac{-2}{(s-1)^2(s-2)} \quad (1.9)$$

For at invertransformen skal bli entydig (se læreboka), må også konvergensområdet være oppgitt. Det er i dette tilfelle som vist i figur 1.4.



Figur 1.4

Vi ser av konvergensområdet og polenes plassering at vi har et dobbeltsidig tidsforløp $h(t)$, og at den høyresidige delen av det er ustabilt. Nå kan vi finne $h(t)$ ved hjelp av (1.7):

$$h(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{-2}{(s-1)^2(s-2)} e^{ts} ds = \begin{cases} \mu_1(t) \frac{1}{1!} \frac{d}{ds} \left[(s-1)^2 \left(\frac{-2}{(s-1)^2(s-2)} e^{ts} \right) \right]_{s=1} \\ -\mu_1(-t) \left[(s-2) \left(\frac{-2}{(s-1)^2(s-2)} e^{ts} \right) \right]_{s=2} \end{cases} \quad (1.10)$$

Det høyresidige tidsforløpet blir da

$$h_h(t) = \mu_1(t) \frac{d}{ds} \left[\frac{-2}{(s-2)} e^{ts} \right]_{s=1} \quad (1.11)$$

dvs.

$$h_h(t) = \mu_1(t) \left[\frac{-2}{(s-2)} t e^{ts} + \frac{2}{(s-2)^2} e^{ts} \right]_{s=1} = \mu_1(t)(2te^t + 2e^t) \quad (1.12)$$

Det venstresidige tidsforløpet blir:

$$h_v(t) = -\mu_1(-t) \left[\frac{-2}{(s-1)^2} e^{ts} \right]_{s=2} = \mu_1(-t) 2e^{2t} \quad (1.13)$$

Vi kan da skrive $h(t)$ som

$$h(t) = \mu_1(-t) 2e^{2t} + \mu_1(t)(2te^t + 2e^t) \quad (1.14)$$

Der $\mu_1(t)$ er enhetssprangfunksjonen. Så langt den inverse Laplacetransformasjon.

1.3 Invers Fouriertransformasjon

Hvis vi foretar et skifte av integrasjonsvariabel i (1.1) ved å forutsette $\sigma = 0$ og så sette $s = j\omega$, dvs. $ds = jd\omega$, blir (1.1) etter litt mellomregning:

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) e^{j\omega t} d\omega \quad (1.15)$$

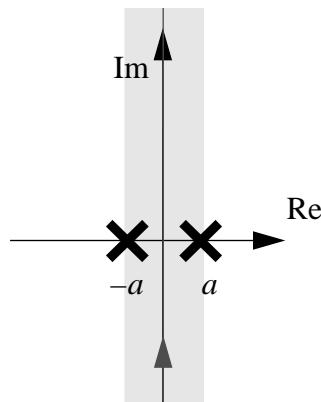
Vi gjenkjenner (1.15) som den inverse Fouriertransformasjonen. Dermed kan Fourier-transformasjonen betraktes som et spesialtilfelle av Laplacetransformasjonen, hvor den imaginære akse er integrasjonsbanen, dvs. konvergensområdet ikkebefatter den imaginære akse. La oss vise metoden med et eksempel. Vi tar utgangspunkt i Fouriertransformen i eksempel (4.8) i læreboka:

$$X(\omega) = \frac{2a}{a^2 + \omega^2} \quad (1.16)$$

Vi har $s = j\omega$, dvs. $\omega = s/j = -js$. Vi setter inn for ω og faktoriserer nevneren:

$$X(\omega) = \frac{2a}{a^2 + (-js)^2} = \frac{2a}{a^2 - s^2} = \frac{-2a}{(s+a)(s-a)} = H(s) \quad (1.17)$$

Polene og konvergensområdet er vist i figur 1.5:



Figur 1.5

Residueregning gir da (dette gjør du sjøl)

$$x(t) = \mu_1(-t)e^{at} + \mu_1(t)e^{-at} = e^{-a|t|} \quad (1.18)$$

1.4 Bemerkning om notasjon

I dette kurset vil vi av bekvemmelighetsgrunner ikke være helt stringente når det gjelder notasjon for uttrykk i frekvens- og s -plan. Et kontinuerlig signal med impulsrespons $h(t)$ har en Laplacetransform betegnet med $H(s)$, mens Fouriertransformen vil bli betegnet med $H(\omega)$. Dette er strengt tatt noe tvilsomt.. Ta transferfunksjonen

$$H(s) = \frac{1}{1 + Ts} \quad (1.19)$$

Fouriertransformen vet vi at vi får når vi setter inn $s = j\omega$. Vi får

$$H(j\omega) = \frac{1}{1 + Tj\omega} \quad (1.20)$$

som egentlig *ikke* burde skrives $H(\omega)$, fordi $H(\omega) = \frac{1}{1 + T\omega}$

Man skulle altså egentlig ha skrevet $H(s)$ og $H(j\omega)$, slik som man finner det i mye av litteraturen..

Vær derfor klar over at når vi for lettvinthets skyld skriver “ $H(\omega)$ ” mener vi egentlig $H(j\omega)$.

Kapittel 2: Diskrete signaler og systemer

2.1 Innledning

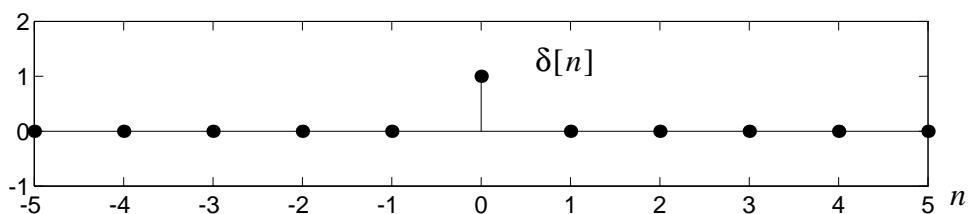
Vi vet at vi med utgangspunkt i Fourierrekka for periodiske kontinuerlige signaler, kan definere Fouriertransformasjonen for ikke-periodiske kontinuerlige signaler. Vi vet at hvis signalet er en impulsrespons for et LTI-system, så er Fouriertransformen til dette signalet frekvensresponsen for det samme systemet. Vi vet at vi kan betrakte Fouriertransformasjonen som et spesialtilfelle av Laplacetransformasjonen.

Vi kunne gått fram på akkurat samme måte for *diskrete* signaler og systemer: Fra diskret Fourierrekke til -transformasjon, til frekvensrespons, så til Z-transformasjon (som vi skal se er en parallel til Laplace-transformasjonen for kontinuerlige signaler). For å utnytte tida litt bedre, velger jeg imidlertid å gjøre dette på en litt annen måte. Dette medfører bl.a. at diskret Fourierrekke utsettes noe (men for de som ønsker en helt tilsvarende gjennomgang for diskrete signaler og systemer som vi gjorde for kontinuerlige, henvises til læreboka).

2.2 Representasjon av diskrete systemer

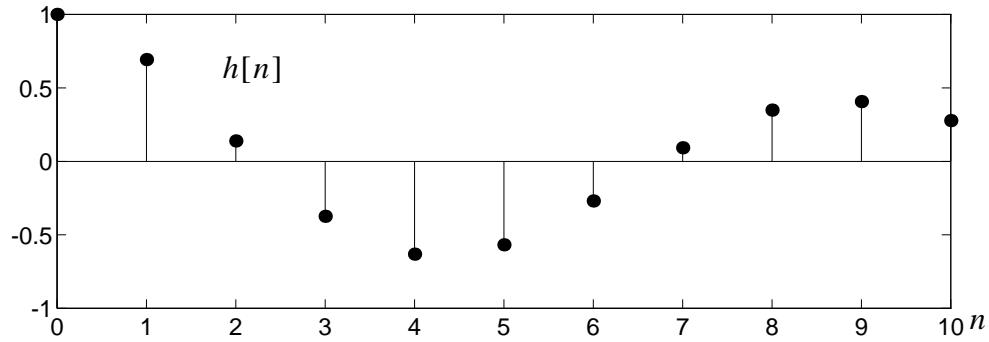
På samme måte som et kontinuerlig lineært system kan karakteriseres ved impulsresponsen $h(t)$, skal vi nå karakterisere et diskret lineært system ved den **diskrete impulsrespons** $h[n]$. Merk at vi bruker **hakeparentes** [] for å markere tidsdiskrete signaler. Først definerer vi den diskrete analogi til diracpulsen $\delta(t)$ for det kontinuerlig tilfelle, nemlig den **diskrete enhetspuls** $\delta[n]$:

$$\delta[n] = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \quad (2.1)$$



Figur 2.1

Når inngangssignalet $x[n] = \delta[n]$, blir utgangen en diskret impulsrespons $h[n]$ som f.eks. kan se ut som skissert i figur 2.2.



Figur 2.2

Anta nå at inngangsignalet er et vilkårlig tidsforløp $x[k]$.

Siden $h[n]$ er impulsresponsen til et diskret LTI-system, gjelder superposisjonsprinsippet. Det betyr at utgangen $y[n]$ kan betraktes som en sum av en rekke impulsresponser, hvor hver impuls som fører til en tilsvarende respons inntreffer ved tida k , og har amplituden $x[k]$ ved denne tida. Vi kan altså skrive $y[n]$ som

$$y[n] = x[0]h[n] + x[1]h[n-1] + \dots + x[n-1]h[1] + x[n]h[0] = \sum_{k=0}^n h[n-k]x[k] \quad (2.2)$$

I (2.2) er $x[k] = 0$ for $k < 0$. Det er også antatt at systemet er kausalt, fordi det i (2.2) ikke er noen bidrag til $y[n]$ fra $x[k]$ når $k > n$.

Den generelle versjon av (2.2) blir

$$y[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k] = h[n]^*x[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k] \quad (2.3)$$

(2.2) og (2.3) kaller vi diskret folding. Den er som antydet i (2.3), kommutativ, akkurat som i det kontinuerlige tilfellet. Den er også asosiativ og distributiv.

(Som vi vil oppdage etterhvert, gjelder det generelt at de egenskaper vi kjenner fra kontinuerlige signaler og systemer, har sin parallel i det diskrete tilfellet).

2.2.1 Merknad om folding for kontinuerlige signaler

Formelen (2.3) er analog med **foldingssintegralet**

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)d\tau = h(t)^*x(t) \quad (2.4)$$

for det kontinuerlige tilfellet.

Faktisk kan kontinuerlig folding gis en “fysisk”, intuitiv tolkning nettopp ved å betrakte (2.4) slik som vi betraktet (2.3): Kontinuerlig folding kan betraktes som en uendelig sum av små impulsresponser, hvor hver impuls (deltafunksjon) inntrer ved tida τ_i , og har arealet $x(\tau_i)\Delta\tau$ ved denne tida:

$$y(t) \approx \sum_{i=-\infty}^{\infty} h(t - \tau_i)x(\tau_i)\Delta\tau \quad , \text{her er } \Delta\tau = \tau_i - \tau_{i-1} \quad (2.5)$$

2.2.2 Merknad om frekvens for diskrete signaler

Pr. definisjon setter vi avstand i diskret tid mellom $x[n]$ og $x[n+1] = 1$ (**dimensjonsløs tid**). Anta at det diskrete signalet framkommer ved **tasting** (“sampling” på engelsk) av et kontinuerlig fysisk signal. M.a.o.: Det diskrete signalet representerer **punktpørøver** (“samples”) av det kontinuerlige, fysiske signal. Man har dermed en fysisk tidsavstand T mellom to “samples” (T kalles “**tastetid**”, “**samplingstid**”). Anta at et signal med fysisk frekvens ω tastes med tastetid T . Når vi beskriver det tastede, diskrete signalet med dimensjonsløs tidsavstand = 1 (vi skalerer det i tid) må det lede til en tilsvarende skalering i frekvens: ω svarer til en **dimensjonsløs frekvens** $\Omega = \omega T$.

Eksempel:

Er $T = 0.001s = 1ms$, og $\omega = 3000$ rad/s (dette svarer til $3000/2\pi = 477$ Hz), så blir den dimensjonsløse frekvens $\Omega = 3000 \cdot 0.001 = 3$.

Vi vil i dette kurset bruke stor Ω for å markere dimensjonsløs frekvens, som altså gjelder for diskrete signaler med tastetid 1 (dimensjonsløs tid). Se også læreboka side 30.

Vi skal nå finne et uttrykk for (dimensjonsløs) frekvensrespons, når impulsresponsen $h[n]$ er gitt. Denne frekvensresponsen er den *diskrete Fouriertransform* av tidsforløpet $h[n]$.

2.3 Den diskrete Fouriertransformasjon

Gitt et stabilt diskret system med impulsresponsen $h[n]$. Vi eksiterer det med det diskrete inngangssignalet $x[n] = \sin(\Omega[n])$, $-\infty \leq n \leq \infty$. Siden signalet har vært påsatt siden $n = -\infty$ og systemet er stabilt, er alle transienter dødd ut, og vi sitter igjen med den stasjonære responsen $y[n]$.

Vi kan skrive $\sin(\Omega[n]) = \text{Im}(e^{j\Omega[n]})$. Vi bruker diskret folding (2.3), høyre del, og får

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k] = \sum_{k=-\infty}^{\infty} h[k]\text{Im}(e^{j\Omega[n-k]}) \quad (2.6)$$

Dette kan omformes til

$$y[n] = \operatorname{Im} \left\{ \left(\sum_{k=-\infty}^{\infty} h[k] e^{-j\Omega[k]} \right) e^{j\Omega[n]} \right\} \quad (2.7)$$

Hvis vi nå definerer

$$H(\Omega) = \sum_{k=-\infty}^{\infty} h[k] e^{-j\Omega[k]} \quad (2.8)$$

kan vi skrive (2.7) som

$$\begin{aligned} y[n] &= \operatorname{Im}\{H(\Omega)e^{j\Omega[n]}\} = \operatorname{Im}(|H(\Omega)|e^{j\angle H(\Omega)}e^{j\Omega[n]}) \\ &= |H(\Omega)| \sin(\Omega[n] + \angle H(\Omega)) \end{aligned} \quad (2.9)$$

(2.9) viser at $y[n]$ svinger med en amplitud og en faseforskyvning gitt av hhv. $|H|$ og $\angle H$.

$H(\Omega)$ er derfor **frekvensresponsen** til systemet med impulsrespons $h[n]$. (Merk at frekvensen Ω er kontinuerlig, sjøl om tida n er diskret).

Dermed er **den diskrete Fouriertransformasjon**

$$H(\Omega) = \sum_{n=-\infty}^{\infty} h[n] e^{-j\Omega n}, \quad (2.10)$$

Vi symboliserer denne med $H(\Omega) = F(h[n])$, altså som i det kontinuerlige tilfellet.

Det gjenstår nå å finne inverstransformasjonen. Dette gjøres via følgende “knep”:

Vi merker oss at $e^{-j\Omega[n]}$ er periodisk i Ω med periode 2π . Da framgår det av (2.10) at også $H(\Omega)$ må bli periodisk i Ω med periode 2π . Dermed kan $H(\Omega)$ representeres ved en “Fourierrekke”. Men merk at periodisiteten nå er i frekvens- og ikke i tidsplanet. Matematisk sett har denne forskjellen ingen betydning. Vi betrakter formlene for Fourierrekka (læreboka ligning (4.34)-(4.35)):

$$x(t) = \sum_{k=-\infty}^{\infty} a[k] e^{jk\omega_0 t}, \text{ eller } x(t) = \sum_{k=-\infty}^{\infty} \bar{a}[k] e^{-jk\omega_0 t}, \text{ der } \bar{a}[k] = a[-k] \quad (2.11)$$

$$a[k] = \frac{1}{T_0} \int_{T_0} x(t) e^{-jk\omega_0 t} dt, \text{ eller } \bar{a}[k] = \frac{1}{T_0} \int_{T_0} x(t) e^{jk\omega_0 t} dt \quad (2.12)$$

Sammenligner vi (2.10) og høyre versjon av (2.11), samtidig som vi benytter at $H(\Omega)$ er periodisk i Ω med periode 2π , kan vi lage en tabell som viser hvilke størrelser som svarer til hverandre rent matematisk:

Fourierrekke	Diskret F.-transform
$x(t)$	$H(\Omega)$
$\bar{a}[k]$	$h[n]$
T_0	2π
ω_0	1

Tabellen og høyre del av (2.12) gir oss da den **inverse diskrete Fouriertransformasjonen**:

$$h[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\Omega) e^{j\Omega n} d\Omega \quad (2.13)$$

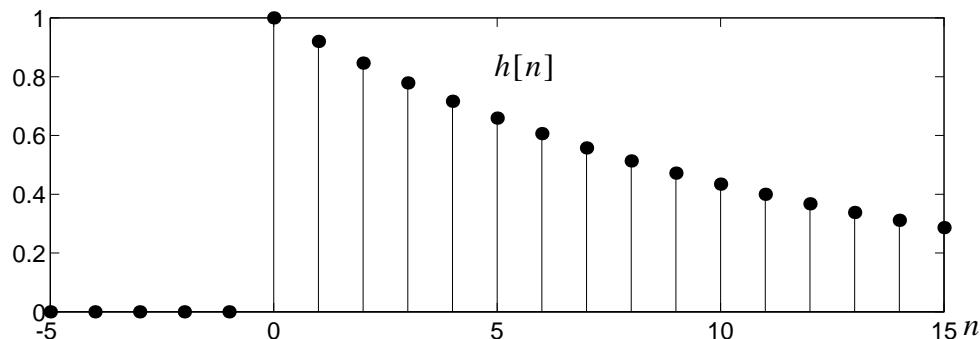
Dermed kan vi oppsummere de to ligningene for diskret Fouriertransformasjon. For sammenligningas skyld har jeg også tatt med de tilsvarende ligninger når tidsforløpet er kontinuerlig:

$$\begin{aligned} H(\Omega) &= \sum_{n=-\infty}^{\infty} h[n] e^{-j\Omega n} & H(\omega) &= \int_{-\infty}^{\infty} h(t) e^{-j\omega t} dt \\ h[n] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\Omega) e^{j\Omega n} d\Omega & h(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) e^{j\omega t} d\omega \end{aligned} \quad (2.14)$$

2.3.1 Eksempel: Utregning av en diskret Fourier-transform

Betrakt det stabile diskrete systemet med impulsrespons

$$h[n] = \mu_1[n] a^n, \quad 0 < a < 1 \quad (2.15)$$



Figur 2.3

$\mu_1[n]$ er den diskrete enhetssprangfunksjonen. Responsen er vist i figur 2.3.

Vi bruker (2.10) til å finne den diskrete Fouriertransformen til (2.15)

$$H(\Omega) = \sum_{n=0}^{\infty} a^n e^{-j\Omega n} = \sum_{n=0}^{\infty} \left(\frac{a}{e^{j\Omega}} \right)^n = \frac{1}{1 - \frac{a}{e^{j\Omega}}} = \frac{1}{1 - ae^{-j\Omega}}, \text{ eller } \frac{e^{j\Omega}}{e^{j\Omega} - a} \quad (2.16)$$

Her har vi brukt formelen for summen av en geometrisk rekke for å finne (2.16):

$$1 + \beta + \beta^2 + \dots = \frac{1}{1 - \beta}, \text{ når } |\beta| < 1 \quad (2.17)$$

Betingelsen $|\beta| < 1$ er oppfylt i (2.16), fordi vi har forutsatt $|a| < 1$ samtidig som $|e^{j\Omega}| = 1$. Hvis derimot $|a| > 1$, så hadde ikke (2.16) konvergert, og $H(\Omega)$ hadde ikke eksistert. Dette er som forventet, fordi $|a| > 1$ i (2.15) ville ha betydd ustabilt system, dvs. en eksponensielt økende tallfolge $h[n] = \mu_1[n]a^n$, og et slikt system er ustabilt og har dermed ingen frekvensrespons. Vi trenger forøvrig ikke begrense oss til systemers impulsrespons; vi kan sjølsgå betrakte $h[n]$ som et vilkårlig signal. Betingelsen for at vilkårlig signal er diskret Fouriertransformerbart er enten at det er **absolutt summérbart** eller at det har **endelig energi**,

$$\sum_{n=-\infty}^{\infty} |h[n]| < \infty \quad \text{eller} \quad \sum_{n=-\infty}^{\infty} |h[n]|^2 < \infty \quad (2.18)$$

Vi ser at (2.15) oppfyller (2.18), begge varianter. (2.18) er som forventet nokså lik betingelser for eksistens for Fouriertransformasjonen for kontinuerlige tidsforløp.

Vi får bekreftet perioden 2π for $H(\Omega)$ i (2.16), fordi Ω bare opptrer i ledd av typen $e^{j\Omega}$. Dette vil alltid være tilfelle for diskrete Fouriertransformer.

For flere eksempler på diskrete Fouriertransform-par, se læreboka s. 338 - 339.

Den diskrete Fouriertransformasjonen har de samme egenskaper som den kontinuerlige m.h.p. linearitet, folding (convolution), dualitet m.m. Se læreboka.

Når det gjelder hvordan regne ut den inverse diskrete Fouriertransform, dvs. (2.13), kan dette gjøres via invers z-transformasjon (beskrevet seinere i dette kapittelet).

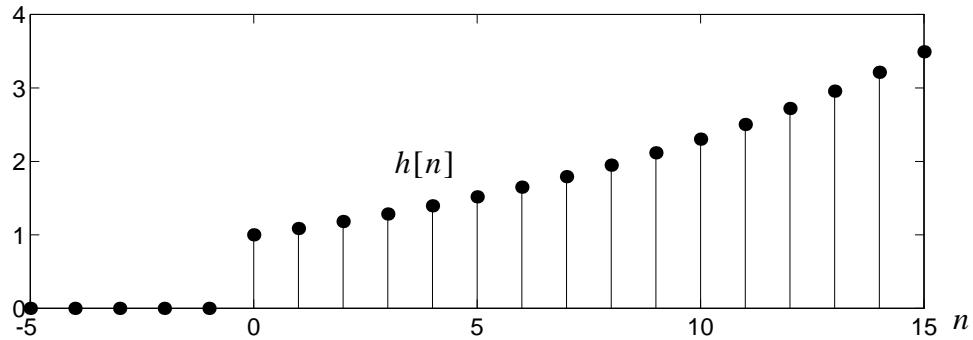
Vi skal nå gå over til z-transformasjonen for diskrete systemer, som spiller samme rolle for diskrete systemer som Laplacetransformasjonen spiller for kontinuerlige.

2.4 Z-transformasjonen

Vi husker at Laplacetransformasjonen kunne betraktes som en generalisering av den kontinuerlige Fouriertransformasjonen, ved å erstatte $j\omega$ med $s = \sigma + j\omega$, der σ var en konstant, valgt slik at Laplacetransformasjonen eksisterte, også for en rekke signaler som ikke kunne Fourier-transformeres.

Vi skal nå se at z-transformasjonen løser samme problem for diskrete signaler. Dette forklares enklest med et eksempel. Betrakt det ustabile systemet i figur 2.4 med impulsrespons

$$h[n] = \mu_1[n]a^n, \quad a > 1 \quad (2.19)$$



Figur 2.4

Dette er *ikke* diskret-Fourier-transformerbart, siden $a > 1$. Vi løser problemet slik: Vi velger en reell positiv konstant r slik at $r > a$, dvs. $a/r < 1$. Vi definerer en modifisert versjon av $h[n]$,

$$\hat{h}[n] = \mu_1[n]a^n\left(\frac{1}{r}\right)^n = \mu_1[n]\left(\frac{a}{r}\right)^n$$

Da vil Fouriertransformen til $\hat{h}[n]$ eksistere:

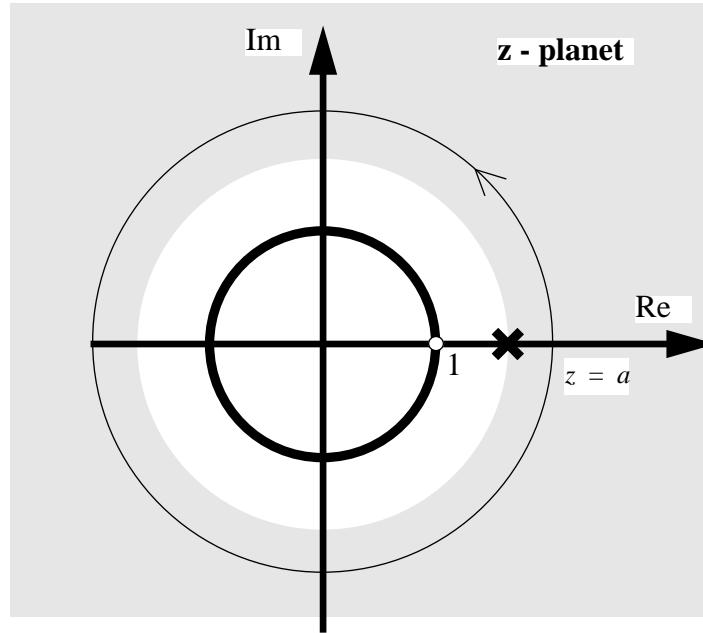
$$\hat{H}(\Omega) = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^n e^{-j\Omega n} = \frac{1}{1 - \frac{a}{r}e^{-j\Omega}} = \sum_{n=0}^{\infty} a^n (re^{j\Omega})^{-n} = \sum_{n=0}^{\infty} h[n](re^{j\Omega})^{-n} \quad (2.20)$$

Vi innfører nå en variabel $z = re^{j\Omega}$ (dette er analogien til $s = \sigma + j\omega$ i det kontinuerlige tilfellet). Da blir (2.20):

$$H(z) = \sum_{n=0}^{\infty} h[n]z^{-n} = \sum_{n=0}^{\infty} a^n z^{-n} = \frac{1}{1 - az^{-1}}, \text{ eller } \frac{z}{z - a} \quad (2.21)$$

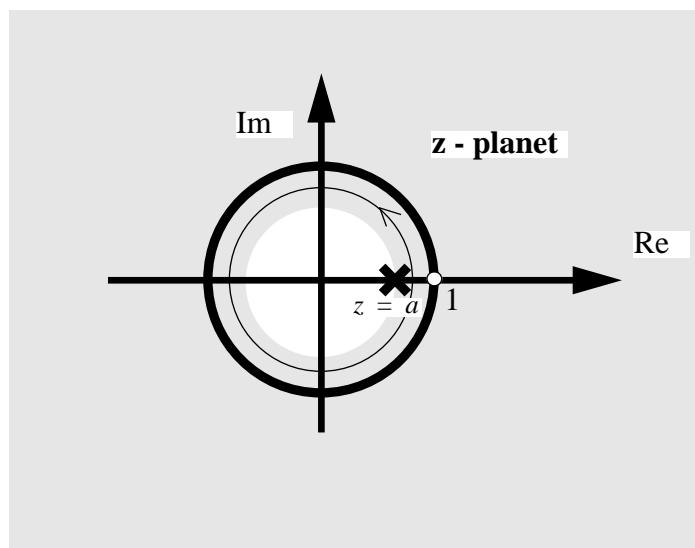
$H(z)$ kalles z-transformen til $h[n] = \mu_1(n)a^n$, og eksisterer da for $|az^{-1}| < 1 \Leftrightarrow |z| > a$.

Vi kan nå innføre begrepet *konvergensområde* i z-planet på samme måte som vi for Laplace-transformasjonen snakker om konvergensområde i s-planet. Konvergensområdet er vist skrevet på figur 2.5.



Figur 2.5

For signalet gitt i (2.15) blir det tilsvarende konvergensområdet som vist i figur 2.6:



Figur 2.6

Dette signalet er stabilt. Vi ser at sirkelen gitt av $|z| = 1$ nå er inne i konvergensområdet. Men $|z| = 1$ svarer til $|re^{j\Omega n}| = 1 = r$, dvs. $z = e^{j\Omega}$. Med andre ord: Hvis z befinner seg på sirkelen gitt av $|z| = 1$, så er z-transformen $H(z) = H(\Omega)$ = den diskrete Fouriertransform. Denne eksisterer i dette tilfelle fordi enhetssirkelen ligger inne i konvergensområdet.

Dette er en nøyaktig parallel til følgende forhold for kontinuerlige signaler: Hvis den imaginære akse ligger i konvergensområdet i s-planet, så eksisterer Fouriertransformen for det tilsvarende kontinuerlige signal, og man får den ved å substituere $s = j\omega$ i Laplacetransformen. Med andre ord spiller enhetssirkelen i z-planet samme rolle som den imaginære akse i s-planet.

Det henvises forøvrig til læreboka, som gir en rekke eksempler på konvergensområder for z-transformen til forskjellige typer diskrete signaler. For *tosidige* diskrete signaler blir konvergensområdet en sirkelformet stripe med sentrum i origo (igjen i pen analogi med det vi kjenner fra s-planet).

$$\text{Den generelle formelen for z-transformasjonen blir: } H(z) = \sum_{n=-\infty}^{\infty} h[n]z^{-n} \quad (2.22)$$

2.4.1 Bemerkning om notasjon

Vi har nevnt i avsnitt 1.4 for kontinuerlige systemer at læreboka i dette kurset ikke er helt konsekvent når det gjelder notasjon for uttrykk i frekvens- og s-plan.

For diskrete systemer/signaler med tidsforløp $h[n]$ og z-transform $H(z)$ gjelder tilsvarende: Frekvensresponsen burde skrives som $H(e^{j\Omega})$, ikke $H(\Omega)$. Av praktiske hensyn vil vi likevel - *med unntak av i utledninga i neste avsnitt* - bruke notasjonen $H(\Omega)$.

2.4.2 Invers Z-transformasjon - utledning

For å finne inverstransformasjonsformelen har vi bruk for $F(h[n]r^{-n})$, som blir

$$F(h[n]r^{-n}) = \sum_{n=-\infty}^{\infty} h[n](re^{\Omega})^{-n} = H(re^{j\Omega}) = H(z)|_{z=re^{j\Omega}} \quad (2.23)$$

Da er $h[n]r^{-n} = F^{-1}[H(re^{j\Omega})]$, dvs. $h[n] = r^n F^{-1}[H(re^{j\Omega})]$.

(2.13) gir da

$$h[n] = \frac{r^n}{2\pi} \int_{-\pi}^{\pi} H(re^{j\Omega})e^{j\Omega n} d\Omega \quad (2.24)$$

Vi substituerer $z = re^{j\Omega}$. Dette gir

$$dz = jre^{j\Omega}d\Omega = jzd\Omega \Rightarrow d\Omega = dz/jz. \text{ Videre blir } e^{j\Omega} = z/r.$$

Vi setter inn i (2.24), som da blir
$$h[n] = \frac{r^n}{2\pi} \int_{2\pi} H(z)z^n r^{-n} \frac{dz}{jz} \quad (2.25)$$

Dette gir den inverse z-transformasjon:

$$h[n] = \frac{1}{2\pi j} \oint_{|z|=r} H(z)z^{n-1} dz, \quad (2.26)$$

der r er slik at sirkelen med radius r er inne i konvergensområdet.

Integrasjonsintervallet for Ω i (2.13) er 2π . Dette svarer til at z beveger seg en omdreining langs en sirkel mot urviseren, og innafor konvergensområdet (se sirkler med tynn strek i figur 2.5 og 2.6). Integranden er et rasjonalt uttrykk med poler i nevner. Dermed kan vi bruke konturintegrasjon og residueregning (se kapittel 1) for å finne den inverse z-transform.

2.4.3 Eksempel: Utregning av en invers Z-transform

Vi tar for oss det stabile høyresidige signalet (2.15), som har z-transform

$$H(z) = \frac{z}{z-a} \quad (2.27)$$

Konvergensområdet og den ene polens plassering er som vist i figur 2.6. Enhetsirkelen omslutter den ene polen. Dermed kan vi bruke (2.26):

$$h[n] = \frac{1}{2\pi j} \oint_{|z|=r} \frac{z^n}{z-a} dz = \sum_i \text{Res}[f(z)] = \mu_1[n]z^n \Big|_{z=a} = \mu_1[n]a^n \quad (2.28)$$

som ventet. Merk at polen i $z = a$ gir stabilt tidsforløp fordi den ligger innafor enhetsirkelen. Igjen: Enhetsirkelen spiller samme rolle m.h.p. stabilitet for diskrete signaler og systemer som den imaginære akse spiller i s-planet (kontinuerlige systemer).

Hvis vi skulle invers-z-transformere et venstresidig signal, får vi et problem med invers z-transformasjon, fordi ingen integrasjonsbane som omslutter polen er i konvergensområdet. Et slikt signal er vist i eksempel 10.2 i læreboka, s. 632. Vi må bruke et “knep”:

Egenskap 10.5.4 (tabell 10.1 s.654 i læreboka) sier at

$$h[-n] \xrightarrow{Z} H\left(\frac{1}{z}\right), \text{ eller } h[-n] = \frac{1}{2\pi j} \oint_{|z|=r} H\left(\frac{1}{z}\right) z^{n-1} dz \quad (2.29)$$

Mens konvergensområdet for $H(z)$ for et venstresidig signal er på innsida av en sirkel med radius $|a_{\min}|$, der a_{\min} er den polen i $H(z)$ som har minst tallverdi, blir konvergensområdet for $H(1/z)$ nå på *utsida* av en sirkel med radius $1/|a_{\min}|$. Dermed kan vi bruke (2.29) og residuesatsen. Prosedyren blir som følger:

- Dann $H(1/z)$
- Omform $H(1/z)$ til et uttrykk $G(z)$, med polynomer i z i teller og nevner, slik at nevneren får faktorer av typen $z - \alpha_i$. (Du vil oppdage at $\alpha_i = 1/a_i$).
- Inverstransformering av $G(z)$ ved residuesatsen gir da $h[-n]$, men vær oppmerksom på følgende viktige poeng: Hvis graden i telleren til $G(z)$ er m lavere enn graden i nevneren, så vil $h[-n]$ bli på formen

$$h[-n] = \mu_1[n-m] \cdot \dots ,$$

og $h[n]$ vil tilsvarende få formen

$$h[n] = \mu_1[-n-m] \cdot \dots .$$

Med andre ord: Både venstre- og høyresidige signaler vil være = 0 i m tidsskritt målt fra og med origo, bakover for venstresidig signal, forover for høyresidig. Dette kan du om ønskelig utlede ved å bruke definisjonen (2.22) på z-transformasjon.

Når det gjelder dobbeltsidige signaler, kan de betraktes som summen av et venstre- og et høyresidig signal, og disse kan beregnes hver for seg som forklart.

2.5 Z-transformen som transferfunksjon. Differensligning

Vi har sett at hvis $h[n]$ er en impulsrespons, vil $H(z)$ med $z = e^{j\Omega}$, gi oss frekvensresponsen $H(\Omega)$ for det lineære diskrete systemet. Vi har videre linearitets- og foldingsegenskaper for z-transformasjonen (læreboka s. 649 og 652), slik at den med hensyn på å beregne utgangssignal hvis man kjenner inngangssignalet og impulsresponsen, kan brukes på samme måte som Laplacetransformasjonen i det kontinuerlige tilfellet. Derfor kaller vi også $H(z)$ for *transferfunksjonen* for diskrete systemer.

z-transformasjonen har også en annen uhyre nyttig egenskap, **tidsforskyvingsegenskapen**. Denne finner du oppsummert på side 650 i læreboka (prøv å utlede den, det er ganske enkelt):

$$h[n-m] \xleftrightarrow{Z} z^{-m}H(z) , \text{ eller} \quad h[n+m] \xleftrightarrow{Z} z^mH(z) \quad (2.30)$$

Denne egenskapen kan brukes på en uhyre enkel måte til å finne en rekursiv formel (som også kalles ei *differensligning*) for et utgangssignal $y[n]$ fra et system som eksiteres med et inngangssignal $x[n]$.

Dette vises enklest med et eksempel. Gitt transferfunksjonen

$$H(z) = \frac{z}{z-a} = \frac{1}{1-az^{-1}} = \frac{y(z)}{x(z)} \quad (2.31)$$

Dette betyr at

$$y(z)(1-az^{-1}) = x(z) \quad (2.32)$$

Vi kan nå invers-z-transformere på begge sider av (2.32). Leddene på venstresida kan invers-transformeres hver for seg, siden (invers) z-transformasjon er lineær. Når vi også benytter tidsforskyvingsegenskapen (2.30), får vi differensligninga:

$$y[n] - ay[n-1] = x[n] , \text{ eller } y[n] = ay[n-1] + x[n] \quad (2.33)$$

Ved hjelp av differensligninger - som jo er rekursive, dvs. vi beregner neste utgangsverdi som en funksjon av tidligere utganger og nåværende og tidligere inngangsverdier - kan vi umiddelbart beregne responsen når inngangssignalet er kjent.

I den forstand er det enklere for diskrete systemer, for i det kontinuerlige tilfellet må vi løse en differensiell ligning for å finne tidsresponsen.

Oppsummeringsvis kan vi nå merke oss disse parallellene mellom diskret og kontinuerlig beskrivelse:

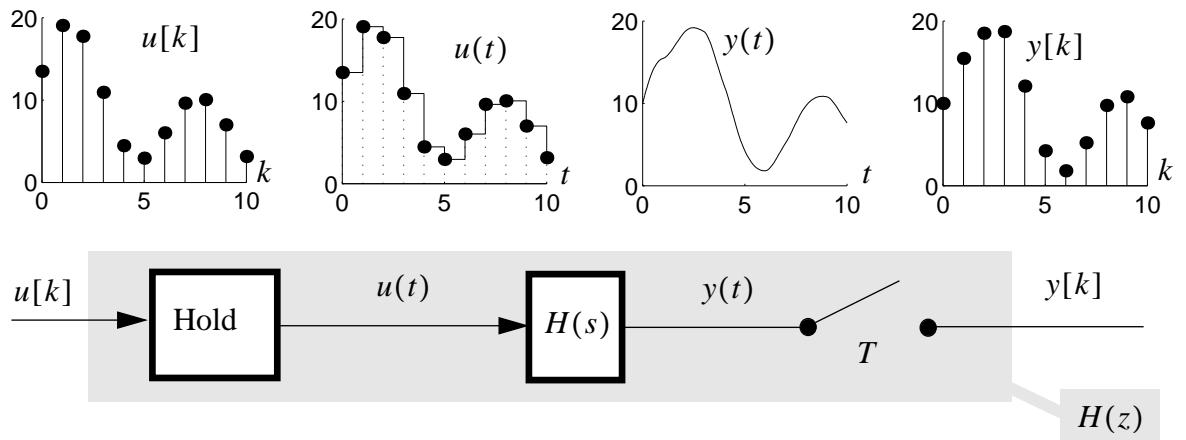
- Differensligning svarer til differensiell ligning.
- Vi finner differensligninga fra $H(z)$ ved å bruke tidsforskyving som følger av multiplikasjon med z .
- Vi finner differensiell ligninga fra $H(s)$ ved å bruke differensiering som følger av multiplikasjon med s .

Kapittel 3: Diskretisering av kontinuerlig system

3.1 Innledning

Det skal vises hvordan man kan finne en diskret ekvivalent for et kontinuerlig system med **holdeelement** foran og **taster** etter. Figur 3.1 viser strukturen. Holdeelementet omgjør det diskrete signalet $u[k]$ - som kommer fra en datamaskin - til et *kontinuerlig* "trappeformet" signal slik som vist. Dette signalet påvirker det kontinuerlige systemet, og gir en kontinuerlig respons $y(t)$, som tastes, og vi får et diskret utgangssignal $y[k]$.

Impulsresponsen $h(t)$ eller transferfunksjonen $H(s)$ for det kontinuerlige system er gitt, og man må velge en tastetid ("samplingtid") T . Da er en ekvivalent diskret transferfunksjon $H(z)$ entydig bestemt. Å finne $H(z)$ kalles "diskretisering". Den ekvivalente $H(z)$ er symbolisert med skravering i figur 3.1, og omfatter det kontinuerlige system med holdeelement foran og tasting etter.



Figur 3.1

Vi skal i dette kapittelet først utlede diskretisering ved hjelp av en tilstandsrom-basert metode. I kapittel 4 vil vi lære en annen metode som vil være raskere enn denne i de fleste tilfeller.

3.2 Diskretisering via tilstandsromform

Det monovariable kontinuerlige system bringes på tilstandsromformen

$$\begin{aligned}\dot{\underline{x}} &= A\underline{x} + Bu \\ y &= D\underline{x}\end{aligned}\tag{3.1}$$

Hvis utgangspunktet er transferfunksjonen $H(s)$, kan dette gjøres f.eks. ved hjelp av fasevariabel form (Balchen: "Reguleringsteknikk", s. 69). Vi kan da umiddelbart finne transisjonsmatrisa $\phi = e^{AT}$. Denne bruker vi i foldingsintegralet (Balchen: s. 61), mellom tidspunkt kT og $(k+1)T$:

$$\underline{x}((k+1)T) = e^{AT}\underline{x}(kT) + \int_{kT}^{(k+1)T} e^{A((k+1)T-\tau)}B u(\tau)d\tau \quad (3.2)$$

På grunn av "trappe-effekten" til holdeelementet i figur 3.1 har vi $u(t) = u[k]$ = konstant for $kT \leq t < (k+1)T$. Vi foretar også et skifte av integrasjonsvariabel, $\alpha = \tau - kT$. Dermed kan vi skrive (3.2) som

$$\underline{x}[k+1] = \phi\underline{x}[k] + \left(\int_0^T e^{A(T-\alpha)}B d\alpha \right) u[k] \quad (3.3)$$

Her bruker vi notasjonen $\underline{x}[k+1] = \underline{x}((k+1)T)$.

Integralet i parentes i (3.3) kan regnes ut slik at vi får¹:

$$\begin{aligned} \underline{x}[k+1] &= \phi\underline{x}[k] + \Delta u[k] \\ y[k] &= D\underline{x}[k], \text{ som før} \end{aligned} \quad (3.4)$$

der integralet blir

$$\Delta = A^{-1}(e^{AT} - I)B \quad (3.5)$$

(Jeg har for enkelhets skyld antatt A inverterbar, det er forøvrig ikke noe problem å beregne integralet i (3.3) uansett).

Nå kan vi z-transformere begge sider av (3.4) ved hjelp av tidsforskyvingsregelen (2.30):

$$z\underline{x}(z) = \phi\underline{x}(z) + \Delta u(z) \quad (3.6)$$

Merk at mens ϕ er en $n \times n$ -matrise av reelle tall, så er $\underline{x}(z)$ en z-transformert vektor, og $u(z)$ en z-transformert skalar. (3.4) kan omformes til

$$(zI - \phi)\underline{x}(z) = \Delta u(z) \Rightarrow \underline{x}(z) = (zI - \phi)^{-1}\Delta u(z) \quad (3.7)$$

Da blir $y(z) = D(zI - \phi)^{-1}\Delta u(z) = H(z)u(z) = D \frac{\text{adj}(zI - \phi)}{|zI - \phi|} \Delta u(z)$ (3.8)

1. (3.4) er den diskrete parallelle til den kontinuerlige tilstandsrommodellen (3.1).

Transferfunksjonen for det diskretiserte system er altså¹

$$H(z) = D \frac{\text{adj}(zI - \phi)}{|zI - \phi|} \Delta \quad (3.9)$$

(I kapittel 4 vil vi vise et eksempel på anvendelse av denne metoden.) $H(z)$ er et skalart (husk systemet er monovariabelt, sjekk gjerne dimensjonene på matrisene for å konstatere at telleren blir skalar. Nevneren er sjølsagt skalar fordi den er en determinant), rasjonalt uttrykk i z . $H(z)$ er **strent proper**, dvs. graden i telleren er minst *en* lavere enn i nevneren. Dette gjelder ikke generelt, men vil være tilfelle når $H(z)$ er en diskretisert versjon av et kontinuerlig strengt propert system, og systemet(3.1) er strengt propert. Hvorfor?

$$H(z) \text{ kan skrives som } H(z) = \frac{B(z)}{A(z)} = \frac{b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n}{z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n} \quad (3.10)$$

Ved å dividere med z^n over og under brøkstreken, får vi et ekvivalent uttrykk

$$H^*(z^{-1}) = \frac{B^*(z^{-1})}{A^*(z^{-1})} = \frac{b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}} \quad (3.11)$$

(3.11) er den vanligste måten å skrive transferfunksjonen på (vi vil i andre sammenhenger, når misforståelser ikke kan oppstå, droppe superskript “*”, og skrive bare $H(z^{-1})$).

Vi har $A^*(z^{-1})y(z) = B^*(z^{-1})u(z)$. Ved å inverstransformere på begge sider, bruke tidsforskyvingsegenskapen til z , og løse m.h.p. $y[k]$, får vi et rekursivt uttrykk (en differensligning) for det diskretiserte systemets utgang som funksjon av tidligere utgangssignaler og inngangssignaler:

$$y[k] = -a_1 y[k-1] - a_2 y[k-2] - \dots - a_n y[k-n] + b_1 u[k-1] + b_2 u[k-2] + \dots + b_n u[k-n] \quad (3.12)$$

3.3 Om poler i det diskrete og kontinuerlige system

3.3.1 Polenes plassering og diskret tidsrespons

Vi kan diagonalisere ϕ (vi forutsetter for enkelhets skyld bare distinkte poler):

$$\phi = MPM^{-1} \quad (3.13)$$

1. (3.9) er den diskrete parallel til $H(s) = D \frac{\text{adj}(sI - A)}{|sI - A|} B$, og utledes som vi ser på identisk vis.

der M er en modalmatrise (=en matrise av egenvektorer), og P er diagonal. Vi har $P = \text{diag}(\rho_i)$, der ρ_i er egenverdiene til ϕ . Ved å bruke (3.13) i (3.4), får vi

$$\begin{aligned} M^{-1}\underline{x}[k+1] &= PM^{-1}\underline{x}[k] + M^{-1}\Delta u[k] \\ y[k] &= DMM^{-1}\underline{x}[k] \end{aligned} \quad (3.14)$$

Ved å innføre $q[k] = M^{-1}\underline{x}[k]$, $\tilde{\Delta} = M^{-1}\Delta$ og $\tilde{D} = DM$, kan (3.14) skrives som

$$\begin{aligned} q[k+1] &= Pq[k] + \tilde{\Delta}u[k] \\ y[k] &= \tilde{D}q[k] \end{aligned} \quad (3.15)$$

Impulsresponsen for systemet blir (vis det!)

$$h[k] = \mu_1[k-1]\tilde{D}P^{k-1}\tilde{\Delta}, \quad (3.16)$$

som er en lineærkombinasjon av ledd av typen $\mu_1[k-1]\rho_i^{k-1}$. (Dette er en parallel til at impulsresponsen for kontinuerlige systemer vil være en lineærkombinasjon av ledd av typen $\mu_1(t)e^{\lambda_i t}$)¹. Dette gjelder også for en generell respons $y[k]$: Ved hjelp av diskret folding (2.3) og (3.16), får vi at $y[k]$ blir

$$y[k] = \tilde{D}P^k\underline{q}[0] + \sum_{m=1}^k d_i\rho_i^{k-m}\tilde{\delta}_i u[m-1], \quad \text{her er er } \begin{bmatrix} \tilde{\delta}_1 \\ \vdots \\ \tilde{\delta}_n \end{bmatrix} = \tilde{\Delta} \quad (3.17)$$

og $\tilde{D} = [d_1 \dots d_n]$. På ikke-dekoplet (= ikke-diagonalisert) form gir (2.3) og (3.4) oss

$$y[k] = D\phi^k\underline{x}[0] + D \sum_{m=1}^k \phi^{k-m}\Delta u[m-1] \quad (3.18)$$

ρ_i i (3.17) kan være komplekse. Da vil ρ_i^k sammen med den konjugerte pol $(\rho_i^*)^k$ representere eksponensielt dempede ($|\rho_i| < 1$) eller økende ($|\rho_i| > 1$) sinusformede diskrete tidsforløp (læreboka, s. 21). $|\rho_i| < 1$ svarer til at ρ_i ligger innafor enhetssirkelen. Hvis alle ρ_i ligger innafor enhetssirkelen er systemet stabilt.

1. Hvis p av polene er sammenfallende, får vi ledd av typen $\mu_1[k-1]k^{p-1}\rho_i^{k-1}$,

hhv. $\mu_1(t)t^{p-1}e^{\lambda_i t}$

3.3.2 Polenes plassering når et kontinuerlig system diskretiseres

Sammenhengen mellom polenes plassering for det kontinuerlige system, og den tilsvarende plassering for det diskretiserte system, finner vi slik: Polene til det diskretiserte system er nullpunktene til nevneren i (3.7) eller (3.9):

$$\begin{aligned} |zI - \phi| &= |zI - e^{AT}| = |zIMM^{-1} - Me^{\Lambda T}M^{-1}| = |M||M^{-1}||zI - e^{\Lambda T}| \\ &= (z - e^{\lambda_1 T})(z - e^{\lambda_2 T})\dots(z - e^{\lambda_n T}) = (z - \rho_1)(z - \rho_2)\dots(z - \rho_n) \end{aligned} \quad (3.19)$$

mens polene i det kontinuerlige system er nullpunktene til nevneren i $H(s)$

$$\begin{aligned} |sI - A| &= |sIMM^{-1} - M\Lambda M^{-1}| = |M||(sI - \Lambda)||M^{-1}| \\ &= (s - \lambda_1)(s - \lambda_2)\dots(s - \lambda_n) \end{aligned} \quad (3.20)$$

Vi ser altså at polene ved diskretisering transformeres slik at poler i v.h.p. havner på innsiden av enhetssirkelen, og poler i h.h.p havner på utsiden av enhetssirkelen¹. Se eksempel i figur 3.2. Vi kan si det slik at v.h.p. avbildes inne i enhetssirkelen, h.h.p. utenfor. Linjer parallelle med den imaginære akse i s-planet svarer til konsentriske sirkler om origo i z-planet, linjer parallelle med den reelle akse svarer til stråler ut fra origo. Merk at linjene i rutenettet fortsatt krysser hverandre med 90° vinkel unntatt i det singulære punkt origo. Dette er en såkalt **konform avbildning**, hvor vinkler forblir de samme ved avbildninga.

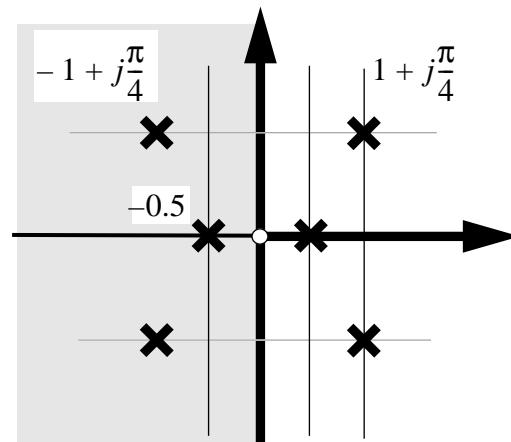
I eksemplet i figur 3.2 er $T = 1$. Hvis T økes, vil stabile poler bevege seg mot origo, ustabile vekk fra origo. Videre vil øvre stråle, hvis T økes helt til 4, svinge mot urviseren til vinkel π og nedre stråle *med* urviseren til $-\pi$.

Vi får vansker hvis T økes over dette: Egenfrekvensen $\pi/4$, gitt av imaginærdelen til polene for det kontinuerlige system, blir da større enn $\omega_s = \pi/T$. Vi gjenkjenner π/T som halve tastefrekvensen².

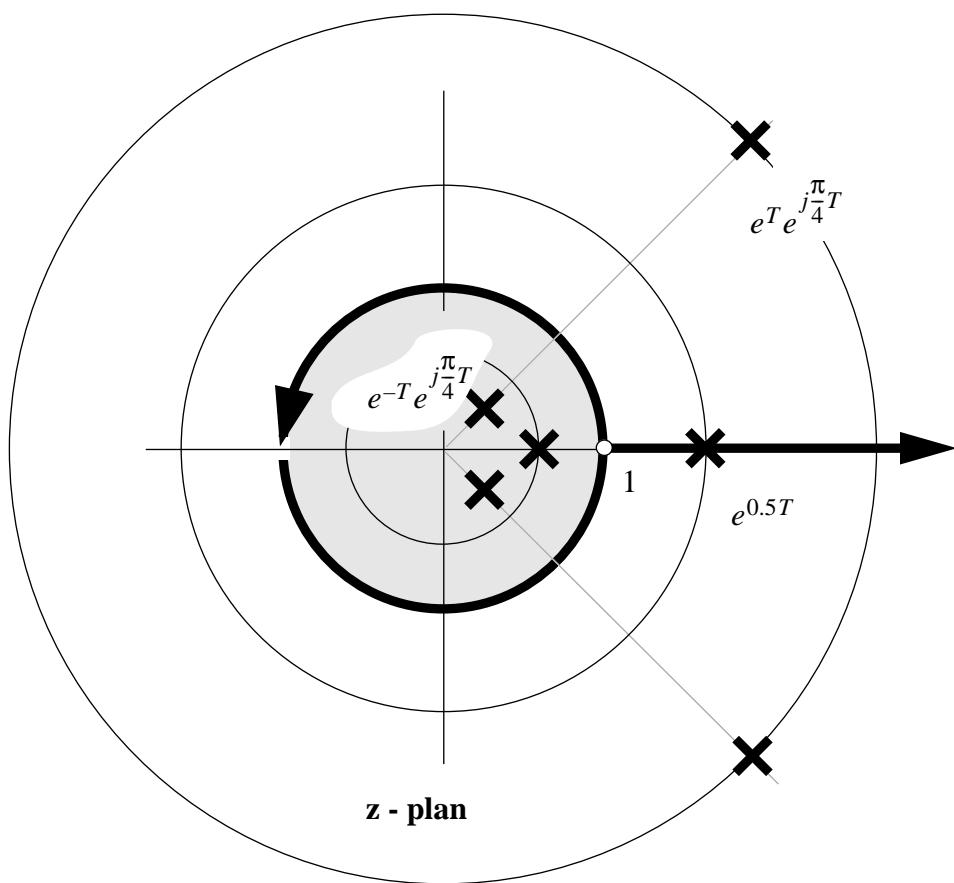
Samplingsteoremet sier da at vi ikke kan gjengi det gitte systemets svingemodi entydig.

1. Når det gjelder *nullpunkter* i telleren til transferfunksjonen $H(s)$ gjelder ikke noen entydig regel for hvordan de transformeres ved diskretisering; dette er generelt meget komplisert.

2. Tastefrekvensen $2\pi/T$ kalles også **Nyquistfrekvensen**



s - plan

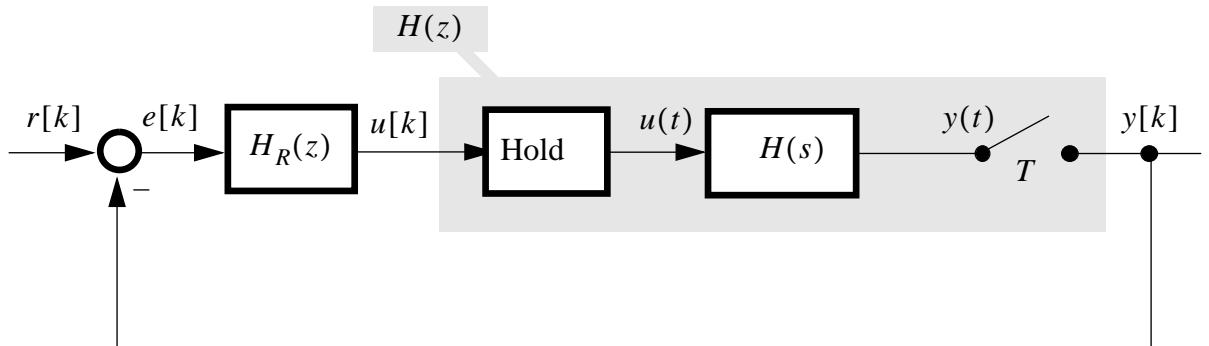


Figur 3.2

Kapittel 4: Diskretisering og regulering av kontinuerlig system

4.1 Innledning

Vi fortsetter der vi slapp i kapittel 3. I figur 4.1 under er vist reguleringsstrukturen: Et kontinuerlig system gitt ved impulsresponsen $h(t)$ eller transferfunksjonen $H(s)$ skal reguleres med en diskret regulator $H_R(z)$. Vi har i avsnitt 3.1 lært en metode for hvordan man kan finne en ekvivalent diskret transferfunksjon $H(z)$ for det kontinuerlige systemet med holdeelement foran og tasting etter. Dette kalles “diskretisering”. Den ekvivalente $H(z)$ er vist skrevet i figur 4.1.



Figur 4.1

Før vi går inn på tilbakekopling og regulering, skal vi vise et eksempel på diskretisering ved hjelp av den tilstandsrom-baserte metoden som ble presentert i 3.1. Så skal vi lære en annen metode som vil være raskere enn denne i de fleste tilfeller. Det samme eksemplet løses ved denne andre metoden. Deretter skal vi se på blokdiagrammer, og komme med noen innledende betraktninger om regulatorsyntese ved diskret regulering av kontinuerlige systemer.

4.2 To diskretiseringsmetoder

4.2.1 Metode 1: Diskretisering via tilstandsromform

Vi skal nå gi et eksempel på metoden beskrevet i avsnitt 3.1
Det monovariable kontinuerlige systemet er gitt ved

$$H(s) = \frac{1}{s(s+1)} \quad (4.1)$$

Ved å bruke fasevariabel form (Balchen: "Reguleringsteknikk", side 68-69 og side 100, eller Finn Haugen: "Anvendt Reguleringsteknikk", 1992-utgaven, side 162-164) får vi tilstandsrommodellen

$$\begin{aligned}\dot{\underline{x}} &= A\underline{x} + Bu, \quad \text{der } A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \end{bmatrix} \\ y &= D\underline{x}\end{aligned}\quad (4.2)$$

Vi søker nå den diskretiserte tilstandsrommodellen

$$\begin{aligned}\underline{x}[k+1] &= \phi(T)\underline{x}[k] + \Delta(T)u[k] \\ y[k] &= D\underline{x}[k]\end{aligned}\quad ,\quad (4.3)$$

dvs. vi trenger matrisene ϕ og Δ .

ϕ kan finnes på flere måter. Jeg velger her å bruke at

$$\phi(t) = L^{-1}\{(sI - A)^{-1}\}, \text{ eller skrevet på en annen måte: } \phi(t) \xrightarrow{L} (sI - A)^{-1} \quad (4.4)$$

Vi har for systemet (4.2) at

$$(sI - A)^{-1} = \frac{\text{adj}(sI - A)}{|sI - A|} = \frac{1}{s(s+1)} \begin{bmatrix} s+1 & 1 \\ 0 & s \end{bmatrix} = \begin{bmatrix} \frac{1}{s} & \frac{1}{s(s+1)} \\ 0 & \frac{1}{s+1} \end{bmatrix} \quad (4.5)$$

Fra (4.5) får vi (residueregning eller tabell over Laplace-transform-par):

$$L^{-1}\{(sI - A)^{-1}\} = \phi(t) = \begin{bmatrix} 1 & (1 - e^{-t}) \\ 0 & e^{-t} \end{bmatrix} \Rightarrow \phi(T) = \begin{bmatrix} 1 & (1 - e^{-T}) \\ 0 & e^{-T} \end{bmatrix} \quad (4.6)$$

når vi setter $t = T$, der $T =$ tastetida.

Vi må også finne et uttrykk for matrisa Δ i (4.3):

Δ finnes enklest ved å bruke ledet i parentes i formel (3.3) i kapittel 3:

$$\Delta(T) = \int_0^T e^{A(T-\tau)} B d\tau = \int_0^T e^{A\beta} B d\beta = \int_0^T \begin{bmatrix} 1 - e^{-\beta} \\ e^{-\beta} \end{bmatrix} d\beta = \begin{bmatrix} T - 1 + e^{-T} \\ 1 - e^{-T} \end{bmatrix} \quad (4.7)$$

Fra (4.6) finner vi

$$(zI - \phi)^{-1} = \begin{bmatrix} z-1 & e^{-T}-1 \\ 0 & z-e^{-T} \end{bmatrix}^{-1} = \frac{1}{(z-1)(z-e^{-T})} \begin{bmatrix} z-e^{-T} & 1-e^{-T} \\ 0 & z-1 \end{bmatrix} \quad (4.8)$$

(4.8) og (3.7) gir $H(z)$:

$$H(z) = D(zI - \phi)^{-1}\Delta = \begin{bmatrix} 1 & 0 \end{bmatrix} \frac{1}{(z-1)(z-e^{-T})} \begin{bmatrix} z-e^{-T} & 1-e^{-T} \\ 0 & z-1 \end{bmatrix} \begin{bmatrix} T-1+e^{-T} \\ 1-e^{-T} \end{bmatrix} \quad (4.9)$$

$$H(z) = \frac{(T-1+e^{-T})z + (1-e^{-T}(1+T))}{z^2 - (1+e^{-T})z + e^{-T}} = \frac{b_1 z + b_2}{z^2 + a_1 z + a_2} \quad (4.10)$$

Vi skriver (4.10) på formen

$$H^*(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \quad (4.11)$$

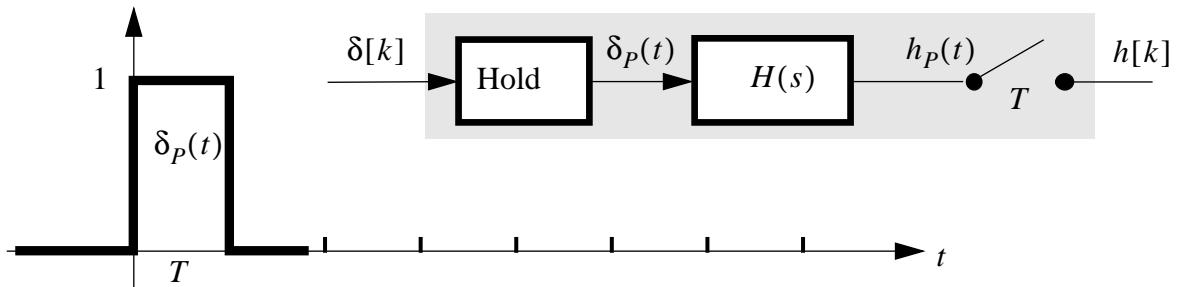
Dette gir, jfr. (3.10)-(3.11), den rekursive formelen

$$y[k] = (1+e^{-T})y[k-1] - e^{-T}y[k-2] + (T-1+e^{-T})u[k-1] + (1-e^{-T}(1+T))u[k-2] \quad (4.12)$$

4.2.2 Metode 2: Diskretisering via impulsresponsen $h[k]$

Denne er noe raskere enn metoden basert på tilstandsrombeskrivelse. Første trinn er å finne impulsresponsen til det diskretiserte system, $h[k]$. Den kan finnes slik:

Hvis vi sender en diskret enhetspuls $\delta[k]$ inn på holdelementet, vil dette igjen sende en firkantpuls $\delta_p(t)$ inn på det kontinuerlige systemet som vist på figur 4.2:



Figur 4.2

Denne pulsen kan oppfattes som summen av den kontinuerlige enhetssprangfunksjonen og en negativ, forsinket enhetssprangfunksjon:

$$\delta_P(t) = \mu_1(t) - \mu_1(t-T), \text{ dermed er } L\{\delta_P(t)\} = \frac{1}{s}(1 - e^{-Ts}) \quad (4.13)$$

Responsen til utgangssignalet blir da

$$h_P(t) = L^{-1}\left(\frac{1}{s}(1 - e^{-Ts})H(s)\right) \quad (4.14)$$

Da har vi

$$h[k] = h_P(kT) \quad (4.15)$$

Nå kunne vi ha funnet $H(z)$ som z-transformen til $h[k]$, gjerne via tabell.

Men følgende framgangsmåte er raskere:

Polene i $H(z) = B(z)/A(z)$ er allerede kjent. Vi finner nevnerpolynomet $A(z)$ i følge (3.19):

$$A(z) = (z - e^{\lambda_1 T})(z - e^{\lambda_2 T}) \dots (z - e^{\lambda_n T}) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n \quad (4.16)$$

Dermed gjenstår bare å finne koeffisientene i tellerpolynomet $B(z)$. Dette kan gjøres ved å sette opp et ligningssystem basert på inngangssignalet $\delta[k]$ (diskret enhetspuls), og den tilsvarende responsen $h[k]$, som vi allerede har funnet som (4.15). Fra (3.12) får vi

$$\begin{aligned} h[1] &= b_1 \delta[0] = b_1 \Rightarrow b_1 = h[1] \\ h[2] &= -a_1 h[1] + b_1 \delta[1] + b_2 \delta[0] \\ &= -a_1 h[1] + b_2 \Rightarrow b_2 = h[2] + a_1 h[1] \\ &\vdots \\ b_n &= h[n] + a_1 h[n-1] + \dots + a_{n-1} h[1] \end{aligned} \quad (4.17)$$

Transferfunksjonen $H(z) = B(z)/A(z)$ er nå funnet.

Prosedyren skal nå demonstreres på samme system som i eksemplet i 4.2.1. Gitt det kontinuerlige system

$$H(s) = \frac{1}{s(s+1)} \quad (4.18)$$

Først finner vi firkantpuls-responsen i følge (4.14), gjerne ved å bruke en Laplacetransformertabell:

$$h_P(t) = L^{-1}\left(\frac{1-e^{-Ts}}{s^2(s+1)}\right) = \mu_1(t)(t-1+e^{-t}) - \mu_1(t-T)(t-T-1+e^{-(t-T)}) \quad (4.19)$$

Dette gir

$$\begin{aligned} h[1] &= T-1+e^{-T} \\ h[2] &= 2T-1+e^{-2T}-(T-1+e^{-T}) = T+e^{-2T}-e^{-T} \end{aligned} \quad (4.20)$$

Nevnerpolynomet $A(z)$ blir i følge (4.16):

$$A(z) = (z-e^{-1T})(z-e^{0T}) = z^2-(1+e^{-T})z+e^{-T} \quad (4.21)$$

Vi bruker nå (4.21) og (4.17):

$$\begin{aligned} b_1 &= h[1] = T-1+e^{-T} \\ b_2 &= h[2]+a_1h[1] = T+e^{-2T}-e^{-T}-(1+e^{-T})(T-1+e^{-T}) = 1-e^{-T}(1+T) \end{aligned} \quad (4.22)$$

Dermed er også tellerpolynomet $B(z)$ funnet. Vi oppsummerer resultatet:

$$H(z) = \frac{(T-1+e^{-T})z+(1-e^{-T}(1+T))}{z^2-(1+e^{-T})z+e^{-T}} = \frac{b_1z+b_2}{z^2+a_1z+a_2} \quad (4.23)$$

som er identisk med (4.10), som forventet.

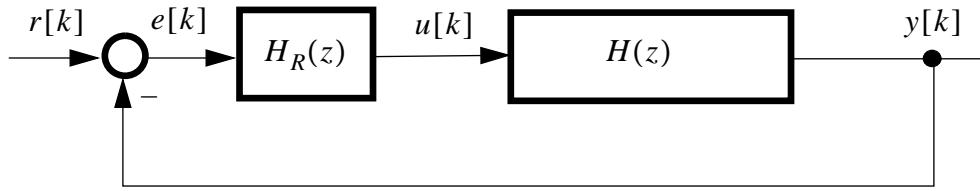
4.3 Blokkdiagram for diskrete systemer

Det ekvivalente diskrete system svarende til figur 4.1 er vist i figur 4.3. Begge blokker $H_R(z)$ og $H(z)$ er rasjonale uttrykk i z .

Dermed kan vi bruke de samme regler for reduksjon av blokkdiagram som vi kjenner fra Laplacetransformerte kontinuerlige systemer.

Vi har fra før definert notasjonen

$$H(z) = \frac{B(z)}{A(z)} \quad (4.24)$$



Figur 4.3

Nå definerer vi i tillegg

$$H_R(z) = \frac{G(z)}{D(z)} \quad (4.25)$$

Transferfunksjonen fra $r[k]$ til $y[k]$ blir

$$H_{ry}(z) = \frac{H_R(z)H(z)}{1 + H_R(z)H(z)} = \frac{B(z)G(z)}{A(z)D(z) + B(z)G(z)} , \text{ eller kompakt: } \frac{BG}{AD + BG} \quad (4.26)$$

Vi ser at $H_{ry}(z)$ blir et rasjonalt uttrykk. Vi kan dermed finne tidsresponsen fra $r[k]$ til $y[k]$ ved invers z-transformasjon. Og vi kan finne frekvensresponsen $H_{ry}(\Omega)$ ved å sette $z = e^{j\Omega}$.

4.4 Balchens “Q-transformasjon”

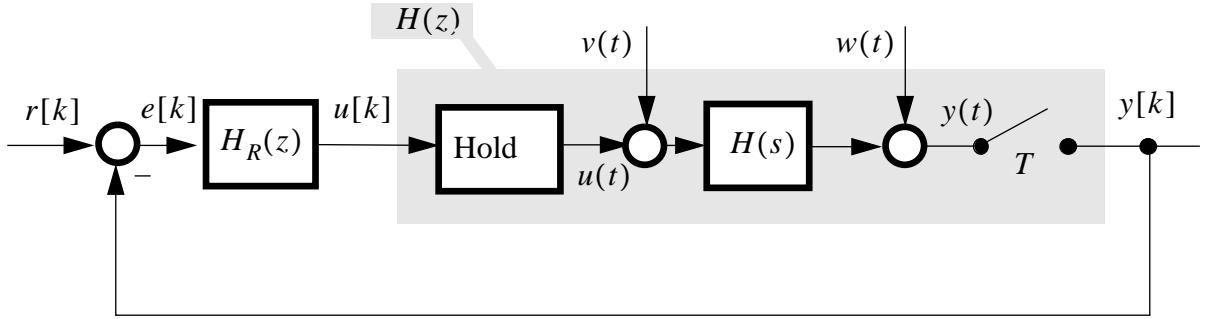
I Balchen: Reguleringsteknikk bind I, s. 397, er denne metoden beskrevet. Ved å substituere

$$z = \frac{1 + \frac{T}{2}q}{1 - \frac{T}{2}q} \quad (4.27)$$

i z-transformerte transferfunksjoner, vil de transformerte transferfunksjoner $H_R(q)$, $H(q)$, $H_{ry}(q)$ osv. forbli rasjonale. Poenget er at enhetssirkelen $|z| < 1$ blir avbildet i halvplanet $\text{Re}(q) < 0$. dette svarer til situasjonen i s-planet for kontinuerlige systemer. Dermed kan Bode-diagram, Nichols-diagram etc., som brukes for Laplacetransformerte kontinuerlige systemer, også anvendes på diskrete systemer som er q-transformert. Man kan bl.a. lage regulatorer v.h.a. seriekompenasjon i Bodediagram. Det henvises til Balchen: Regulerings-teknikk bind I. I dette kurset skal vi imidlertid syntetisere regulatorer ved hjelp av en annen metode en seriekompenasjon, såkalt *pollassering*, og vi skal bruke z-plan-beskrivelse.

4.5 ARMAX-modell

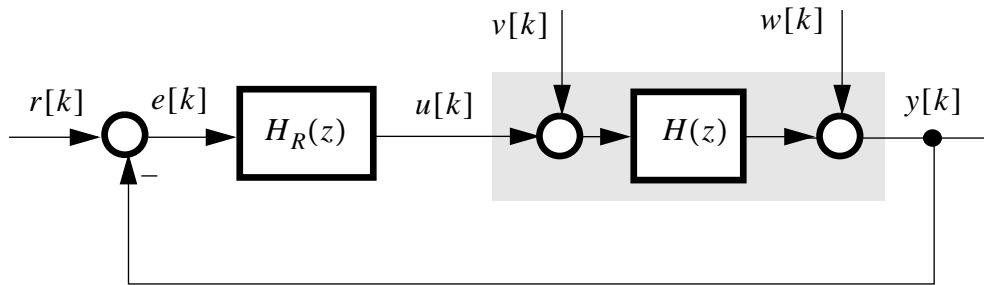
Betrakt reguleringsstrukturen i figur 4.4



Figur 4.4

De kontinuerlige forstyrrelsene $v(t)$ og $w(t)$ vil fra nå av bli kalt henholdsvis **prosess-støy** og **målestøy**. Vi ønsker nå å lage regulatorer som både får utgangen til best mulig å følge referansen $r[k]$, og samtidig motvirker effekten av $v(t)$ og $w(t)$.

For å kunne gjøre det må vi lage et ekvivalent diskretisert skjema, vist på figur 4.5:



Figur 4.5

Hvis vi nå studerer det diskrete systemet innenfor det skraverte området, kan vi definere en såkalt **ARMAX-modell**

$$A^*(z^{-1})y[k] = B^*(z^{-1})u[k] + C^*(z^{-1})\epsilon[k] \quad (4.28)$$

For diskretiserte kontinuerlige systemer blir innholdet i (4.28) i følge (3.11)

$$\begin{aligned} A^*(z^{-1}) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n} \\ B^*(z^{-1}) &= b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n} \end{aligned} \quad (4.29)$$

Hvis vi bare har prosess-støy og ikke målestøy, får vi

$$C^*(z^{-1}) = B^*(z^{-1}), \quad \varepsilon[k] = v[k] \quad (4.30)$$

og hvis vi bare har målestøy og ikke prosess-støy, får vi

$$C^*(z^{-1}) = A^*(z^{-1}), \text{ fordi} \quad \varepsilon[k] = w[k] \quad (4.31)$$

Hvis vi har begge deler, blir $C^*(z^{-1})$ mer komplisert. Dette kommer vi tilbake til seinere når vi skal behandle regulering av systemer som utsettes for støy som er *stokastiske* prosesser.

I kapittel 5 skal vi derimot forklare hvordan vi ved hjelp av polplasserings-metodikk kan lage regulatorer som motvirker støy som er *deterministisk*, dvs. impulser, sprangfunksjoner, etc. Slike regulatorer bygger på en modell av det diskretiserte system på formen (4.28).

Kapittel 5: Regulering v.h.a. polplassering

I avsnitt 4.5 definerte vi en såkalt ARMAX-modell. Vi skal i dette kapittelet utlede to regulatorer som kan anvendes på et system beskrevet av en ARMAX-modell. Regulatorene har det til felles at de bringer det lukkede systems utgang til ro etter et endelig antall tidsskritt. En slik egenskap ved en regulator gjør at den kalles en **dead-beat-regulator**, dvs. utgangen “legges død”. Dead-beat-regulatorer er en av flere klasser av regulatorer som syntetiseres ved hjelp av *polplassering*. Polplassering betyr at man spesifiserer hvor man vil at det lukkede systems poler skal ligge, og så velger man regulatoren slik at dette oppnås.

Vi skal se at hvis man velger regulatoren slik at alle poler i det lukkede diskrete system havner i origo i z-planet, så vil regulatoren være en dead-beat-regulator.

Polplasseringsmetoder er radikalt forskjellige fra den klassiske frekvensanalytiske metoden med seriekompenasjon i Bode-diagram. Begge er nyttige å kunne.

5.1 Definisjoner, notasjon

Gitt det åpne diskrete system (som kan være den diskretiserte representasjonen av et kontinuerlig system) ved ARMAX-modellen

$$A(z^{-1})y[k] = z^{-m}B(z^{-1})u[k] + C(z^{-1})\varepsilon[k] \quad (5.1)$$

der m er et positivt heltall og

$$\begin{aligned} A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_nz^{-n} \\ B(z^{-1}) &= b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_nz^{-n} \\ C(z^{-1}) &= 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_nz^{-n} \end{aligned} \quad (5.2)$$

Denne modellen er litt forskjellig fra (4.28)-(4.31) i kapittel 4: Polynomet $B(z^{-1})$ begynner nå med et konstantledd, mens vi i (4.28) forutsatte at første ledd inneholdt z^{-1} . Dette var en konsekvens av diskretiseringen av et kontinuerlig strengt propert system. Men dette er ikke noe problem: Ved å sette $m = 1$ i (5.1) så oppnår vi strukturen i (4.28) som et spesialtilfelle. Hensikten med å innføre m er å kunne ivareta effekten av en eventuell transportforsinkelse i pådraget eller målingen. Dette ivaretas med $m > 1$. Mer om det seinere.

Videre: De tre polynomene kan godt ha forskjellig grad, dvs. n_A, n_B, n_C , men vi har da uansett $n_B, n_C \leq n_A$. For enkelhets skyld setter vi stort sett $n_A, n_B, n_C = n$ i det følgende.

Merk også at vi nå bruker notasjonen $A(z^{-1})$ osv., ikke $A^*(z^{-1})$ osv. Vi kan sløyfe “*” når det går fram av sammenhengen at vi betrakter polynomer i z^{-1} , ikke polynomer i z .

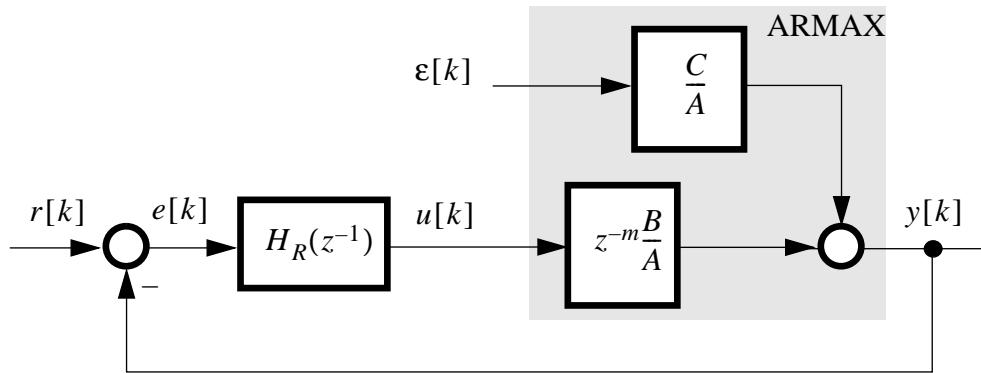
Vi vil også, for å spare plass, svært ofte skrive uttrykk av typen (5.1) som

$$Ay[k] = z^{-m}Bu[k] + C\varepsilon[k] \quad (5.3)$$

Det er da underforstått at store bokstaver betyr polynomer i z^{-1} .

Vi kan nå skissere en reguleringsstruktur som vist på figur 5.1:

Det åpne system, beskrevet ved hjelp av ARMAX-modellen, ligger innafor det skraverte området i figuren.

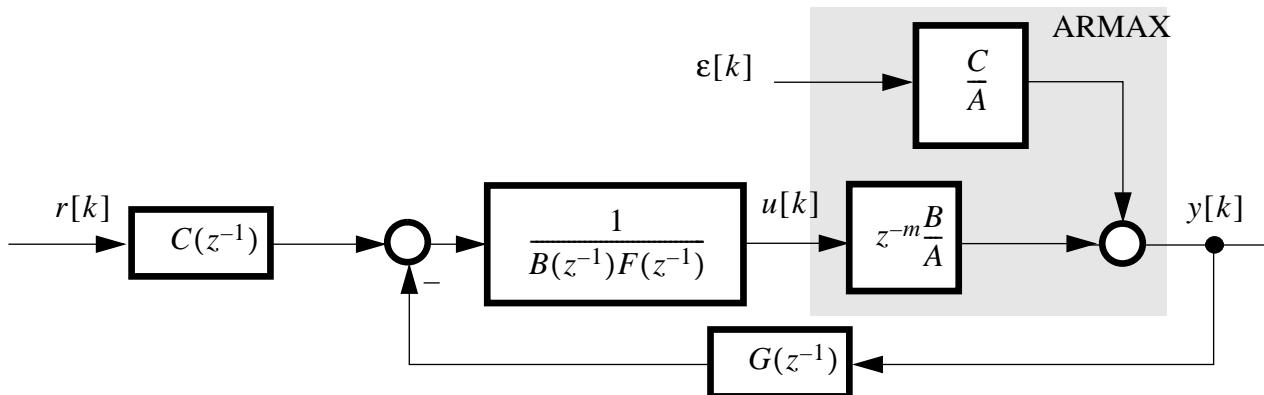


Figur 5.1

Problemet er nå å velge en regulator som plasserer alle poler i det lukkede system i origo. Den første regulatoren vi skal utlede, er den av alle tenkelige regulatorer som bringer系统的 utgang til referansen raskest, når referansen har endret seg, og/eller systemet blir utsatt for en forstyrrelse $\varepsilon[k]$. Først skal vi se på oppgaven med å regulere utgangen til ønsket verdi (dvs. $y[k]$ skal følge $r[k]$ best mulig).

5.2 Den hurtigste utgangssignal-dead-beat¹-regulator

For å realisere denne regulatoren foretar vi en tilsynelatende umotivert endring av regulator-delen av systemet. Se figur 5.2



Figur 5.2

1. Dead-beat = "legger død", dvs. regulatoren bringer systemet helt til ro

Vi ser at regulatoren “spres” på tre blokker i stedet for én, slik som i figur 5.1. Vi ser videre at polynomene C og B også inngår i regulatoren, samtidig som det er introdusert to nye og foreløpig ukjente polynomer G og F . La oss nå finne det lukkede systems transferfunksjon fra $r[k]$ til $y[k]$ ved reduksjon av blokkdiagram:

$$H_{ry}(z^{-1}) = \frac{\frac{C}{BF} z^{-m} \frac{B}{A}}{1 + \frac{G}{BF} z^{-m} \frac{B}{A}} = \frac{C z^{-m} B}{ABF + G z^{-m} B} = \frac{C z^{-m}}{AF + z^{-m} G} \quad (5.4)$$

når vi forkorter bort B -polynomet. Hvis vi nå kan velge de foreløpig ukjente polynomer G og F slik at

$$C(z^{-1}) \equiv A(z^{-1})F(z^{-1}) + z^{-m}G(z^{-1}) \quad (5.5)$$

så vil transferfunksjonen fra $r[k]$ til $y[k]$ bli

$$H_{ry} = \frac{C z^{-m}}{AF + z^{-m} G} = z^{-m}, \text{ dvs. } y[k] = z^{-m}r[k] = r[k-m] \quad (5.6)$$

Utgangen stiller seg med andre ord eksakt etter referansen, bare m tidsskritt forsinket.

Hurtigere er det ikke mulig å gjøre det, for husk at m representerer en transportforsinkelse mellom inngang og utgang i det uregulerte (fysiske, diskretiserte) system.

Symbolet \equiv i (5.5) betyr “identisk lik”, og illustrerer at venstre og høyre side skal være lik for alle verdier av argumentet z^{-1} . Dette betyr at når man multipliserer sammen polynomene på høyre side, skal koeffisientene i det resulterende polynom alle være lik de tilsvarende koeffisientene i $C(z^{-1})$.

Hvis vi studerer transferfunksjonen fra forstyrrelsen $\varepsilon[k]$ til $y[k]$, blir den

$$H_{\varepsilon y}(z^{-1}) = \frac{\frac{C}{A}}{1 + \frac{G}{BF} z^{-m} \frac{B}{A}} = \frac{CBF}{ABF + G z^{-m} B} = \frac{CF}{AF + z^{-m} G} \quad (5.7)$$

som forenkles til

$$H_{\varepsilon y}(z^{-1}) = \frac{CF}{AF + z^{-m} G} = F(z^{-1}) \quad (5.8)$$

hvis polynomene G og F er valgt slik at de tilfredsstiller (5.5).

Du kan nå sjøl kontrollere at hvis forstyrrelsen $\varepsilon[k]$ er en diskret enhetspuls, vil $y[k]$ stille seg inn på referansen etter et endelig antall tidsskritt. Mer om det nedenfor.

Vi finner koeffisientene i de ukjente polynomer G og F ved å bruke identiteten (5.5) til å sette opp en ligning for sammenhørende koeffisienter foran z^{-1} , z^{-2} , z^{-3} , osv.

Vi velger grad $n - 1$ på G -polynomet, og grad $m - 1$ på F -polynomet

$$\begin{aligned} G(z^{-1}) &= g_0 + g_1 z^{-1} + g_2 z^{-2} + \dots + g_{n-1} z^{-(n-1)} \\ F(z^{-1}) &= 1 + f_1 z^{-1} + f_2 z^{-2} + \dots + f_{m-1} z^{-(m-1)} \end{aligned} \quad (5.9)$$

Ved disse valg av grader, og under forutsetning av at $A(z^{-1})$ og $B(z^{-1})$ ikke har noen sammenfallende røtter, kan det vises (gjøres ikke i dette kurset) at da finnes det *ett og bare ett* sett med koeffisienter i polynomene G og F som tilfredsstiller identiteten (5.5). Med andre ord får vi en entydig løsning for regulatoren. Når vi har funnet polynomene, blir regulatoren (se figur 5.2)

$$u(z) = \frac{C}{BF} r(z) - \frac{G}{BF} y(z) \quad (5.10)$$

som ved invers z-transformasjon fører til den rekursive algoritmen for $u[k]$

$$BFu[k] = Cr[k] - Gy[k] \quad (5.11)$$

Vi kaller denne regulatoren for en utgangssignal-dead-beat-regulator

5.2.1 To betydninger av symbolet z

Størrelsen z brukes i litteraturen i *to* betydninger:

1. I forbindelse med z-transformen av en diskret tidsfunksjon, f.eks. en transferfunksjon $H(z^{-1}) \cdot z = e^{j\Omega}$ gir da frekvensresponsen.
I (5.10) har z denne tolkningen.
2. Som symbol for forskyvningsoperasjon i det diskrete *tidsplan*: $u[k-1] = z^{-1}u[k]$.
I (5.11) har z denne tolkningen.

Enkelte forfattere bruker symbolet q i stedet for z når det er forskyvningsoperasjon det gjelder. Vi kommer i det følgende til å bruke z i begge betydninger, og det går da fram av sammenhengen hvilken betydning som gjelder.

5.2.2 Polene til det lukkede system

Vi kan skrive (5.6) som

$$H_{ry} = z^{-m} = \frac{z^{n-m}}{z^n} \quad (5.12)$$

og (5.8) som

$$H_{\text{ey}} = F(z^{-1}) = \frac{z^n F(z^{-1})}{z^n} = \frac{z^n + f_1 z^{n-1} + f_2 z^{n-2} + \dots + f_{m-1} z^{n-m+1}}{z^n} \quad (5.13)$$

Vi ser at begge transferfunksjonene kan omformes slik at de får et nevnerpolynom med alle poler i origo, noe som er karakteristisk for dead-beat-regulerte systemer. Poler i origo er det lengste man kan komme fra enhetssirkelen, og dermed kan ikke systemet gjøres raskere.

5.2.3 Eksempel 1: Dead-beat-regulator

Gitt et system med

$$\begin{aligned} A(z^{-1}) &= 1 - 1.7z^{-1} + 0.7z^{-2} \\ B(z^{-1}) &= 1 + 0.5z^{-1} \quad , m = 2. Vi \text{ ser av } A(z^{-1}) \text{ at } n \text{ er } 2 \\ C(z^{-1}) &= 1 + 1.5z^{-1} + 0.9z^{-2} \end{aligned} \quad (5.14)$$

G -polynomet får grad $n - 1 = 1$, og F -polynomet får grad $m - 1 = 1$. Polynomidentiteten (5.5) blir da

$$1 + 1.5z^{-1} + 0.9z^{-2} \equiv (1 - 1.7z^{-1} + 0.7z^{-2})F(z^{-1}) + z^{-2}G(z^{-1}) \quad (5.15)$$

eller

$$1 + 1.5z^{-1} + 0.9z^{-2} \equiv (1 - 1.7z^{-1} + 0.7z^{-2})(1 + f_1 z^{-1}) + z^{-2}(g_0 + g_1 z^{-1}) \quad (5.16)$$

Ved å forlange like koeffisienter foran sammenhørende ledd på begge sider, får vi ligningssystemet

$$\begin{aligned} z^{-1}: 1.5 &= -1.7 + f_1 \\ z^{-2}: 0.9 &= 0.7 - 1.7f_1 + g_0 \\ z^{-3}: 0 &= 0.7f_1 + g_1 \end{aligned} \quad (5.17)$$

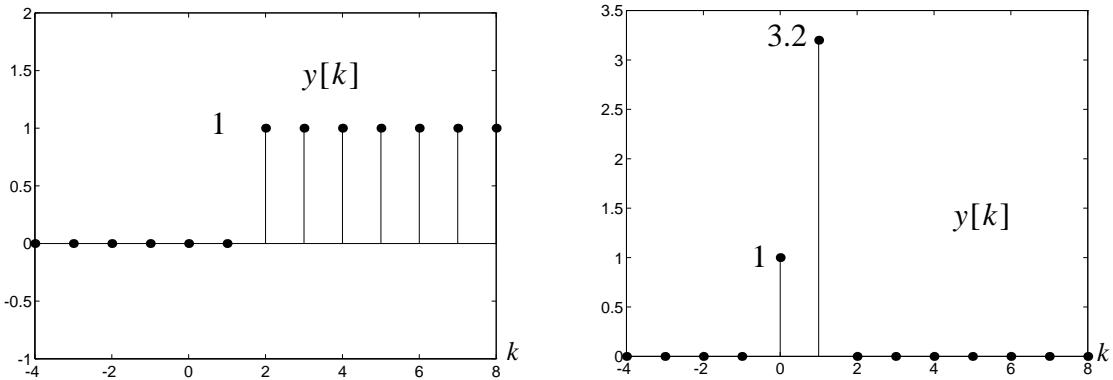
Dette gir $f_1 = 3.2$, $g_0 = 5.64$, $g_1 = -2.24$. Regulatoren blir da i følge (5.11):

$$(1 + 0.5z^{-1})(1 + 3.2z^{-1})u[k] = (1 + 1.5z^{-1} + 0.9z^{-2})r[k] - (5.64 - 2.24z^{-1})y[k] \quad (5.18)$$

som er det samme som det rekursive uttrykk

$$\begin{aligned} u[k] &= -3.7u[k-1] - 1.6u[k-2] + r[k] + 1.5r[k-1] + 0.9r[k-2] \\ &\quad - 5.64y[k] + 2.24y[k-1] \end{aligned} \quad (5.19)$$

Sprangresponsen for det lukkede system blir da i følge (5.6) som vist til venstre på figur 5.3. Responsen på en forstyrrelse $\varepsilon[k]$ som er en diskret enhetspuls blir i følge (5.8) som vist til høyre i figur 5.3 (her forutsettes for enkelhets skyld $r[k] \equiv 0$):



Figur 5.3

5.2.4 Eksempel 2: En dead-beat-regulator som skaper problemer

Anta at vi har følgende enkle diskret(isert)e system som skal reguleres:

$$y[k] = 0.7y[k-1] + u[k-1] + 0.99u[k-2] + \varepsilon[k] + 0.95\varepsilon[k-1] \quad (5.20)$$

Bruker (5.2) og får

$$\begin{aligned} A(z^{-1}) &= 1 - 0.7z^{-1} \\ B(z^{-1}) &= 1 + 0.99z^{-1}, \quad n = 1, \quad \text{og } m = 1 \\ C(z^{-1}) &= 1 + 0.95z^{-1} \end{aligned} \quad (5.21)$$

Polynomidentiteten (5.5) blir da

$$1 + 0.95z^{-1} \equiv (1 - 0.7z^{-1})F(z^{-1}) + z^{-1}G(z^{-1}) \quad (5.22)$$

eller $1 + 0.95z^{-1} \equiv (1 - 0.7z^{-1})(1) + z^{-1}(g_0) \quad (5.23)$

dette blir

$$1 + 0.95z^{-1} \equiv 1 + (-0.7 + g_0)z^{-1} \quad (5.24)$$

Dette gir én ligning for å finne én ukjent koeffisient:

$$0.95 = g_0 - 0.7 \Rightarrow g_0 = 1.65 \quad (5.25)$$

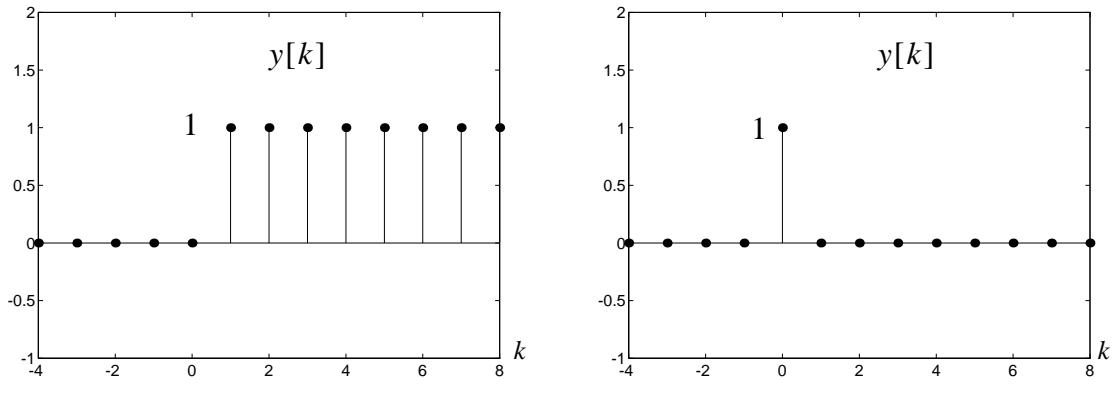
Regulatoren blir da i følge (5.11):

$$(1 + 0.99z^{-1})(1)u[k] = (1 + 0.95z^{-1})r[k] - (1.65)y[k] \quad (5.26)$$

d.v.s. den rekursive formel

$$u[k] = -0.99u[k-1] + r[k] + 0.95r[k-1] - 1.65y[k] \quad (5.27)$$

Sprangresponsen for systemet blir da i følge (5.6) som vist til venstre på figur 5.4. Responsen på en forstyrrelse $\varepsilon[k]$ som er en diskret enhetspuls blir i følge (5.8) som vist til høyre i figur 5.4:



Figur 5.4

Dette ser jo veldig bra ut. Men denne regulatoren er likevel ikke brukbar for det gitte system. Dette ser vi på følgende måte: Vi beregner nå det generelle uttrykk for transferfunksjonen fra referansen $r[k]$ til pådraget $u[k]$:

$$H_{ru}(z^{-1}) = \frac{\frac{C}{BF}}{1 + \frac{G}{BF}z^{-m}\frac{B}{A}} = \frac{AC}{ABF + Gz^{-m}B} = \frac{A}{B} \quad (5.28)$$

Tilsvarende får vi at transferfunksjonen fra forstyrrelsen $\varepsilon[k]$ til pådraget $u[k]$ blir

$$H_{\varepsilon u}(z^{-1}) = \frac{\frac{C}{A}\left(-\frac{G}{BF}\right)}{1 + \frac{G}{BF}z^{-m}\frac{B}{A}} = \frac{-CG}{ABF + Gz^{-m}B} = \frac{-G}{B} \quad (5.29)$$

Vi ser at polene til H_{ru} og $H_{\varepsilon u}$ blir lik nullpunktene til B -polynomet. Dermed vil $u[k]$ sine svingninger bli uakseptable hvis disse polene ligger nær innsida av enhetssirkelen - eller enda verre - på utsiden av den. I det siste tilfellet vil pådraget faktisk bli ustabilt.

Nullpunktet (det er bare ett) for B -polynomet i vårt eksempel gis av

$$1 + 0.99z^{-1} = 0 \Leftrightarrow z + 0.99 = 0, \Rightarrow z = -0.99 \quad (5.30)$$

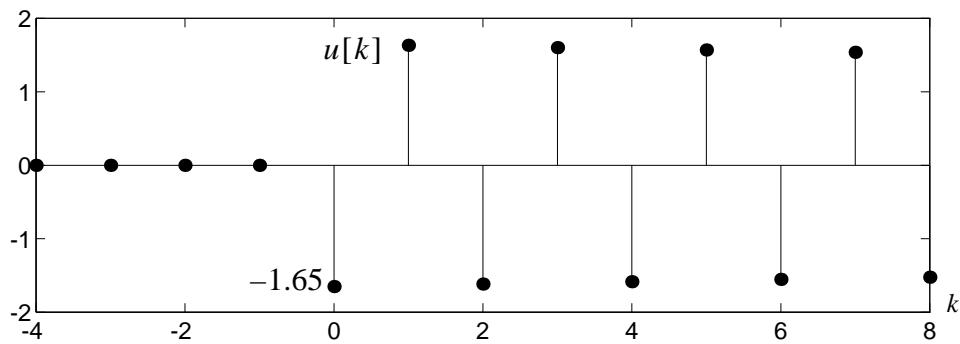
La oss nå betrakte hva som skjer med pådraget til systemet i vårt eksempel når systemet utsettes for en forstyrrelse $\varepsilon[k]$ som er en diskret enhetspuls. Utgangen $y[k]$ blir uhyre raskt regulert til null, som allerede kommentert, og skissert til høyre i figur 5.4. Pådraget kan beregnes etter den rekursive formel som vi finner på grunnlag av (5.29)

$$B(z^{-1})u[k] = -G(z^{-1})\varepsilon[k] \quad (5.31)$$

eller, når vi setter inn tall fra vårt eksempel og løser m.h.p. $u[k]$:

$$u[k] = -0.99u[k-1] - 1.65\varepsilon[k], \text{ der } \varepsilon[k] = \delta[k] \quad (5.32)$$

Dette gir et forløp som skissert i figur 5.5:



Figur 5.5

Vi ser at pådraget oscillerer med meget langsom innsvingningstid, bestemt av polen $z = -0.99$ i B -polynomet. Med andre ord: På tross av at utgangen som vist til høyre i figur 5.4 megt raskt kommer til ro, har man kraftige oscillasjoner “inne i systemet”.

Vi kunne ha forutsagt dette ved å betrakte (5.4):

$$H_{ry}(z^{-1}) = \frac{\frac{C}{BF}z^{-m}\frac{B}{A}}{1 + \frac{G}{BF}z^{-m}\frac{B}{A}} = \frac{Cz^{-m}B}{ABF + Gz^{-m}B} = \frac{Cz^{-m}}{AF + z^{-m}G}$$

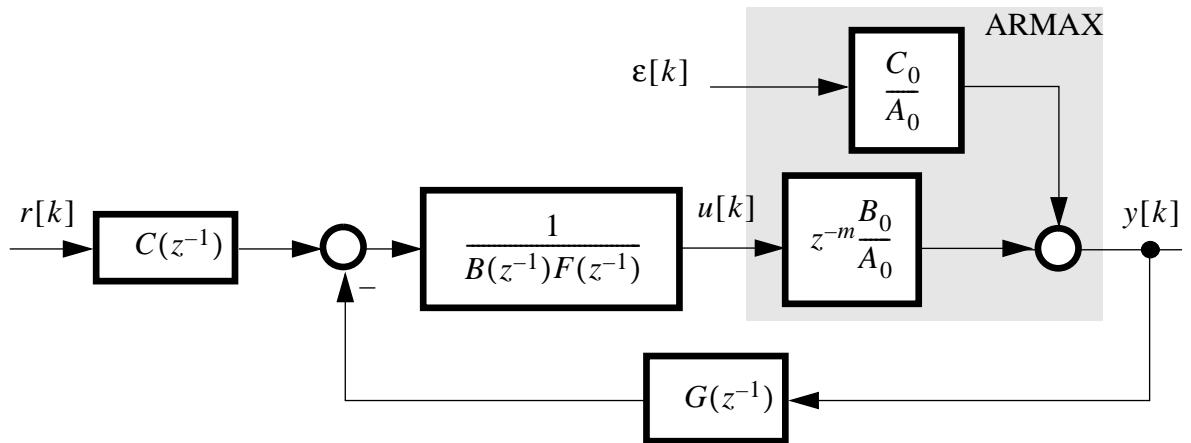
Uttrykket til høyre framkommer ved at vi *forkorter bort B-polynomet*. Men vi vet fra Reguleringsteknikk grunnkurs at når vi forkorter bort ledd i teller og nevner i transferfunksjoner, så fører det til at de tilsvarende modi blir ikke-observerbare og/eller ikke-styrbare. I dette tilfelle blir systemet ikke-observerbart; virkningen av polen i det lukkede system som skyldes B -polynomet, kan ikke spores på systemets utgang. Dette er bakgrunnen for betegnelsen *utgangssignal-dead-beat*.

Denne regulator kan derfor bare anvendes hvis alle nullpunkter i B -polynomet ligger godt innanfor enhetssirkelen.

I eksempel 1 hadde vi nullpunktet $z = -0.5$ i B -polynomet. Dette er så nær origo at vi kan akseptere å bruke utgangssignal-dead-beat-regulator.

5.2.5 Robusthet er lav med den hurtigste dead-beat-regulatoren

Begrepet robusthet i reguleringsteknikken forteller noe om et systems evne til å beholde de spesifiserte egenskaper på tross av at prosessen som skal reguleres er noe annerledes enn det vi antok da vi syntetiserte systemets regulator. Det viktigste robusthets-kriterium er om systemet forblir stabilt når parametre endres. Betrakt strukturen i figur 5.6:



Figur 5.6

Indeksene "0" i A_0 , B_0 og C_0 symboliserer nå at disse polynomene representer de virkelige, riktige fysiske parametre i prosessen som skal reguleres. Men vi beregner regulatoren på grunnlag av en antagelse om prosessen som er noe forskjellig, symbolisert ved polynomene A , B , C .

Nå kan vi regne ut transferfunksjonen fra forstyrrelsen $\varepsilon[k]$ til $y[k]$ (jfr. (5.7)), som blir

$$H_{\varepsilon y}(z^{-1}) = \frac{\frac{C_0}{A_0}}{1 + \frac{G}{BF} z^{-m} \frac{B_0}{A_0}} = \frac{C_0 BF}{A_0 BF + G z^{-m} B_0} \quad (5.33)$$

Vi ser at det ikke er mulig å forkorte, og vi ser at noen av polene i det lukkede system vil være nær nullpunktene i B - og B_0 -polynomet. Det første betyr at nå vil virkningen av disse polene synes på utgangen, dvs. systemet blir observerbart. Det siste betyr at systemet kan bli ustabilt hvis nullpunktene i B eller B_0 -polynomet ligger utefor enhetssirkelen, eller til og med om de måtte ligge nær den på innsida.

Dette skal nå illustreres med systemet fra eksempel 2, gitt ved (5.20):

$$y[k] = 0.7y[k-1] + u[k-1] + b_{01}u[k-1] + \varepsilon[k] + 0.95\varepsilon[k-1] \quad (5.34)$$

Endringen er nå at mens den fysiske, riktige parameter fortsatt er $b_{01} = 0.99$, er vår *antagelse* b_1 feil, hhv. 0.995 i et tilfelle og 0.95 i et annet. Regulatoren blir altså litt annerledes. For feil antagelse $b_1 = 0.995$ får vi nå regulatoralgoritmen

$$u[k] = -0.995u[k-1] + r[k] + 0.95r[k-1] - 1.65y[k] \quad , \quad (5.35)$$

og for feil antagelse $b_1 = 0.95$ får vi regulatoralgoritmen

$$u[k] = -0.95u[k-1] + r[k] + 0.95r[k-1] - 1.65y[k] \quad (5.36)$$

Figur 5.7 på neste side viser hva som skjer i disse to tilfellene, når forstyrrelsen $\varepsilon[k]$ er en diskret enhetspuls som før. Vi kan trekke den konklusjon at denne dead-beat-regulatoren er ubruklig for det aktuelle system, siden svært små feilantagelser gir elendig eller til og med ustabil oppførsel.

5.3 Robust utgangssignal-dead-beat-regulator

Vi skal nå presentere en annen dead-beat regulator som ikke er fullt så rask, men som til gjengjeld er mer robust. Vi lager en litt annen reguleringsstruktur enn den i figur 5.2:

Endringen er nå at B -polynomet faktoriseres i følge ligninga

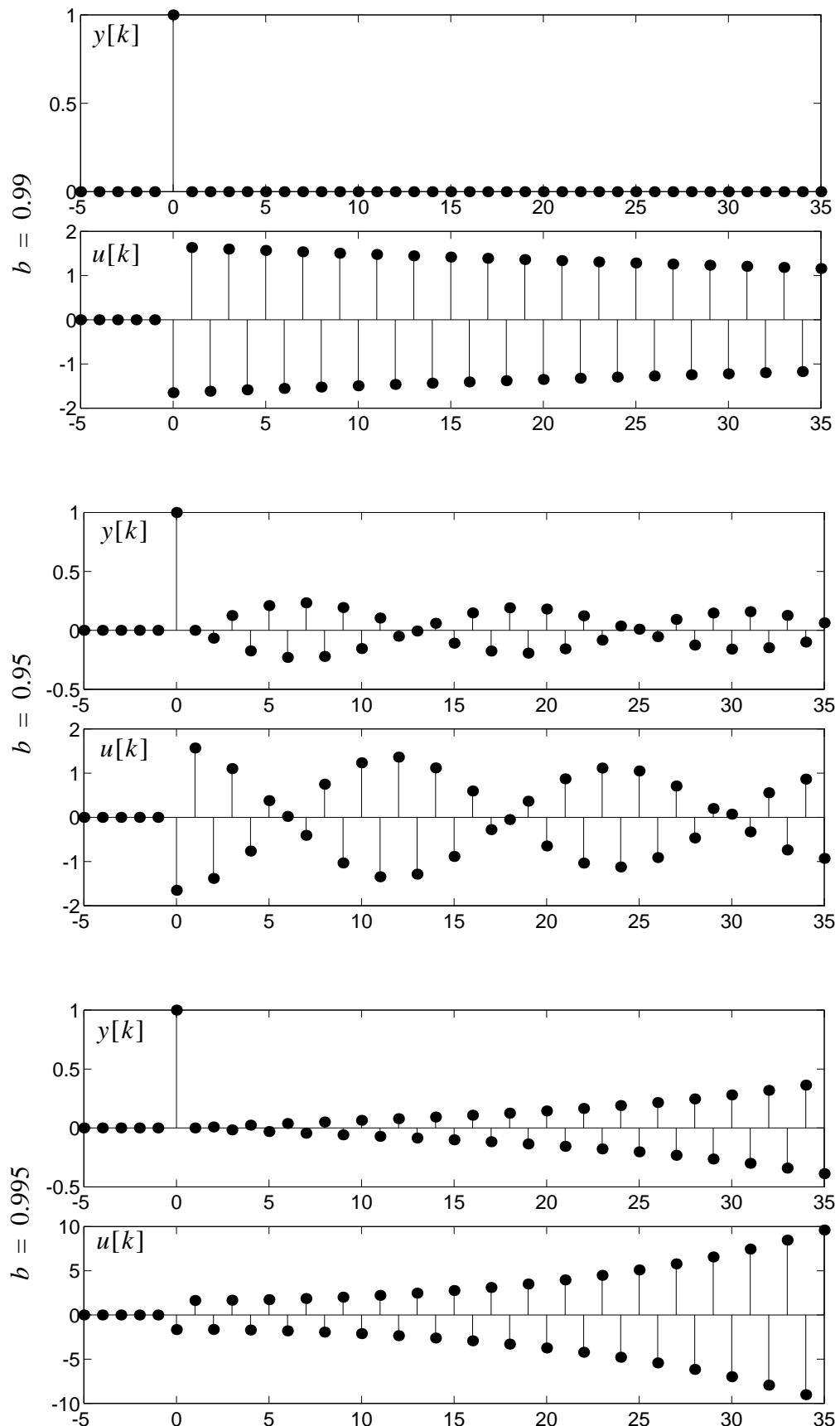
$$B(z^{-1}) = B_1(z^{-1})B_2(z^{-1}) \text{ , der } \text{grad}(B_1) = n_{B1} \text{ og } \text{grad}(B_2) = n_{B2}, n_{B1} + n_{B2} = n_B \quad (5.37)$$

Vi velger B_1 slik at alle nullpunktene i denne ligger godt innanfor enhetssirkelen. Hvis vi studerer transferfunksjonen fra forstyrrelsen $\varepsilon[k]$ til $y[k]$, blir den

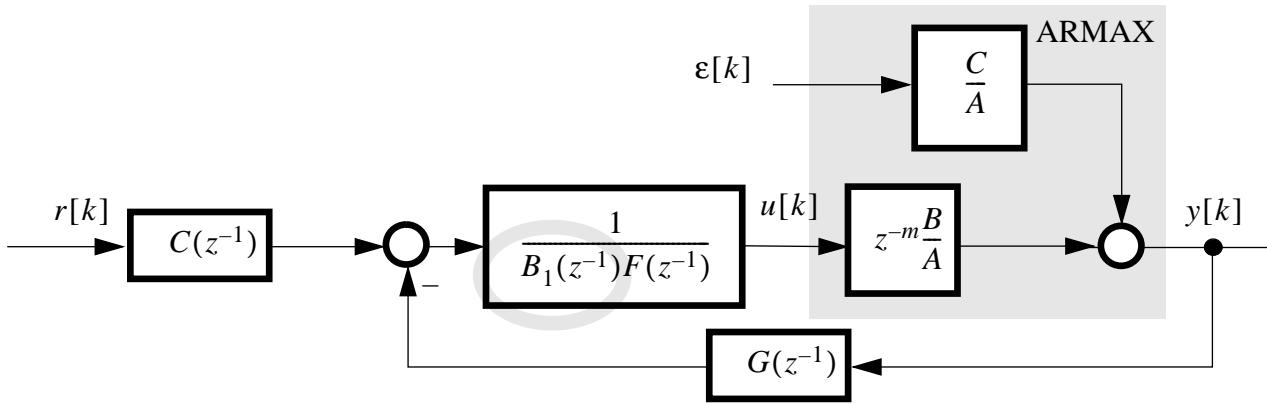
$$H_{\varepsilon y}(z^{-1}) = \frac{\frac{C}{A}}{1 + \frac{G}{B_1 F} z^{-m} \frac{B}{A}} = \frac{CB_1 F}{AB_1 F + G z^{-m} B_1 B_2} = \frac{CF}{AF + z^{-m} B_2 G} \quad (5.38)$$

Vi lager nå en litt mer komplisert (i forhold til (5.5)) polynomidentitet

$$C \equiv AF + z^{-m} B_2 G \quad (5.39)$$



Figur 5.7



Figur 5.8

Her er $\text{grad}(F) = m + n_2 - 1$, mens $\text{grad}(G) = n - 1$ som før. Med disse grader på F og G , vil (5.39) ha en entydig løsning for F og G (vises ikke her). (5.39) innsatt i (5.38) gir forkortning med C -polynomet som før, men nå bare med B_1 -polynomet i tillegg til C -polynomet.

Vi får da

$$H_{\varepsilon y}(z^{-1}) = F(z^{-1}) \quad (5.40)$$

Transferfunksjonen fra forstyrrelsen $\varepsilon[k]$ til pådraget $u[k]$ blir nå

$$H_{\varepsilon u}(z^{-1}) = \frac{\frac{C}{A} \left(-\frac{G}{B_1 F} \right)}{1 + \frac{G}{B_1 F} z^{-m} \frac{B}{A}} = \frac{-CG}{AB_1 F + G z^{-m} B_1 B_2} = \frac{-G}{B_1} \quad (5.41)$$

Vi ser at polene til $H_{\varepsilon u}$ blir lik nullpunktene *bare* til B_1 -polynomet. Dermed vil $u[k]$ sine svingninger bli akseptable, da vi jo nettopp valgte B_1 slik at disse polene ligger godt på innsida av enhetssirkelen.

Med denne regulatoren er det nå bare de “ufarlige” nullpunkter som blir ikke-observerbare poler i det lukkede system.

Du kan nå igjen kontrollere at hvis forstyrrelsen $\varepsilon[k]$ er en diskret enhetspuls, vil $y[k]$ stille seg inn på en null-referanse etter et endelig antall tidsskritt. Forskjellen blir at dette tar litt lengre tid, på grunn av F -polynomets høyere grad.

Også denne regulatoren skal nå prøves på systemet fra eksempel 2. Igjen tenker vi oss at vi har en viss uoverensstemmelse mellom den virkelige $b_{01} = 0.99$ og våre antagelser. Figur 5.9 viser hva som skjer hvis den feilaktige antagelse b_1 er henholdsvis 0.995 og 1.05, referansen er null, og forstyrrelsen $\varepsilon[k]$ er en diskret enhetspuls. Vi får bekreftet at denne regulatoren er langt mer robust.

I dette tilfelle må vi beregne nye parametre i F- og G-polynomet for hvert av de to tilfellene, siden B_2 inngår i polynomidentiteten (5.39).

For feil antagelse $b_1 = 0.995$ har vi

$$B_2(z^{-1}) = B(z^{-1}) = 1 + 0.995 z^{-1}, \quad (5.42)$$

og for feil antagelse $b_1 = 1.05$ har vi

$$B_2(z^{-1}) = B(z^{-1}) = 1 + 1.05 z^{-1} \quad (5.43)$$

Denne "robuste" regulatoren er også en utgangssignal-dead-beat-regulator. En annen type dead-beat-regulatorer er *tilstands-dead-beat-regulatorer*. Disse tillater ikke at noen modi i det lukkede system blir ikke-observerbare eller ikke-styrbare. Den robuste regulatoren presentert her er oppfyller ikke dette kravet, fordi vi fortsatt forkorter bort både B_1 -polynomet og C-polynomet.

5.3.1 Følgeegenskaper for den robuste regulatoren

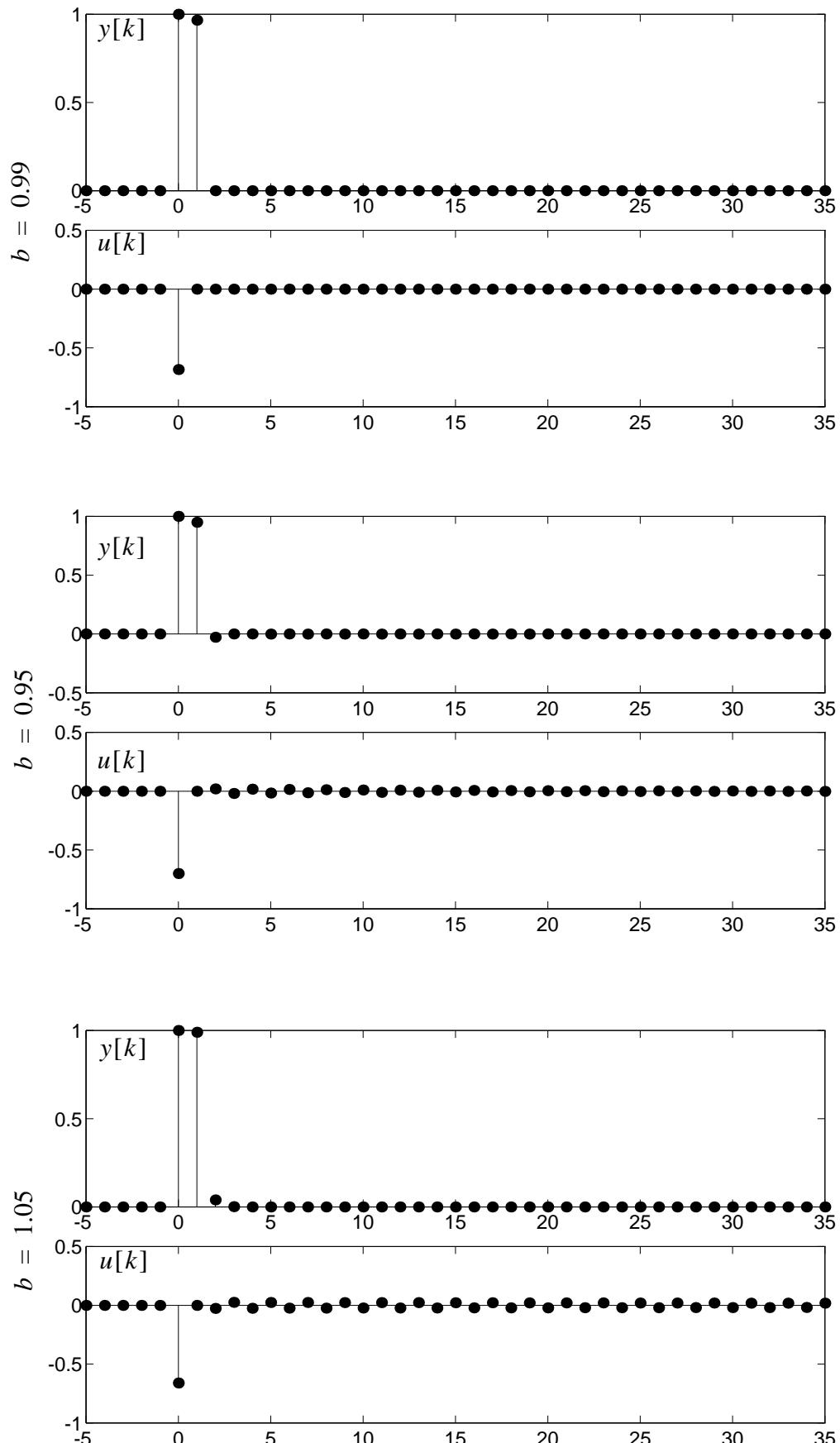
Det var ovenfor forutsatt at referansen var lik null. Fra nå av tenker vi oss at den kan variere. La oss regne ut det lukkede systems transferfunksjon fra $r[k]$ til $y[k]$:

$$H_{ry}(z^{-1}) = \frac{\frac{C}{B_1 F} z^{-m} B}{1 + \frac{G}{B_1 F} z^{-m} B} = \frac{C z^{-m} B}{AB_1 F + G z^{-m} B_1 B_2} = \frac{C z^{-m} B_2}{AF + z^{-m} B_2 G} = z^{-m} B_2 \quad (5.44)$$

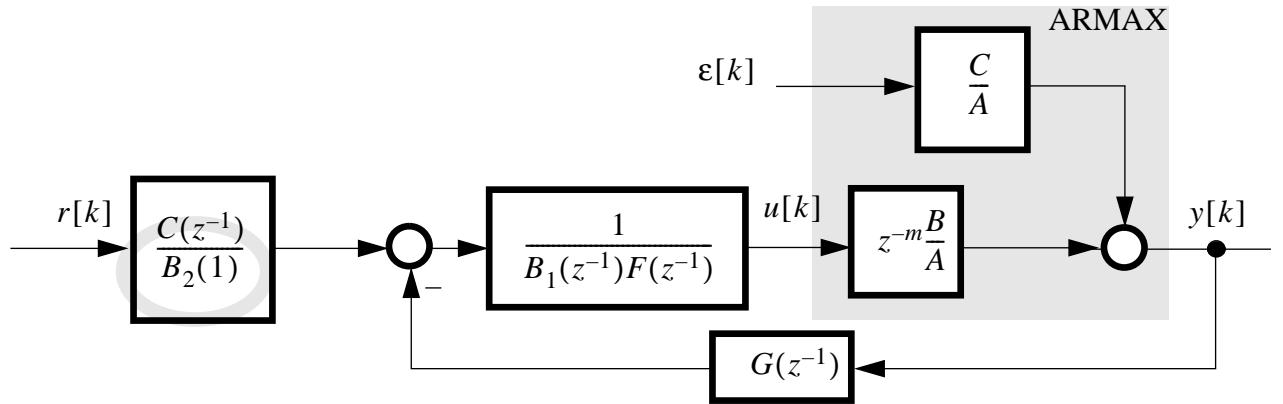
Vi ser at vi nå ikke får skikkelig følging av referansen, slik som vi fikk i (5.6) med den første regulatoren.

Men hvis $r[k] \equiv 0$ er dette som nevnt ikke noe problem, og hvis $r[k] = \mu_1[k]$ (et sprang) vil den robuste regulatoren gi perfekt regulering hvis vi legger inn en konstant forsterkningsfaktor $1/B_2(1)$, der $B_2(1)$ betyr $B_2(z^{-1})$ med $z = 1$. Se figur 5.10.

Hvorfor vil vi da få null stasjonært avvik etter et sprang i referansen?



Figur 5.9



Figur 5.10

5.4 Om graden på F- og G-polynomet

Graden på disse polynomene kan gis en intuitiv tolkning. Det er enklest å forstå dette hvis vi tar utgangspunkt i den første dead-beat-regulatoren (avsnitt 5.2). Vi har da grad $n - 1$ på G -polynomet, og grad $m - 1$ på F -polynomet:

$$\begin{aligned} G(z^{-1}) &= g_0 + g_1 z^{-1} + g_2 z^{-2} + \dots + g_{n-1} z^{-(n-1)} \\ F(z^{-1}) &= 1 + f_1 z^{-1} + f_2 z^{-2} + \dots + f_{m-1} z^{-(m-1)} \end{aligned} \quad (5.45)$$

Regulatorformelen er

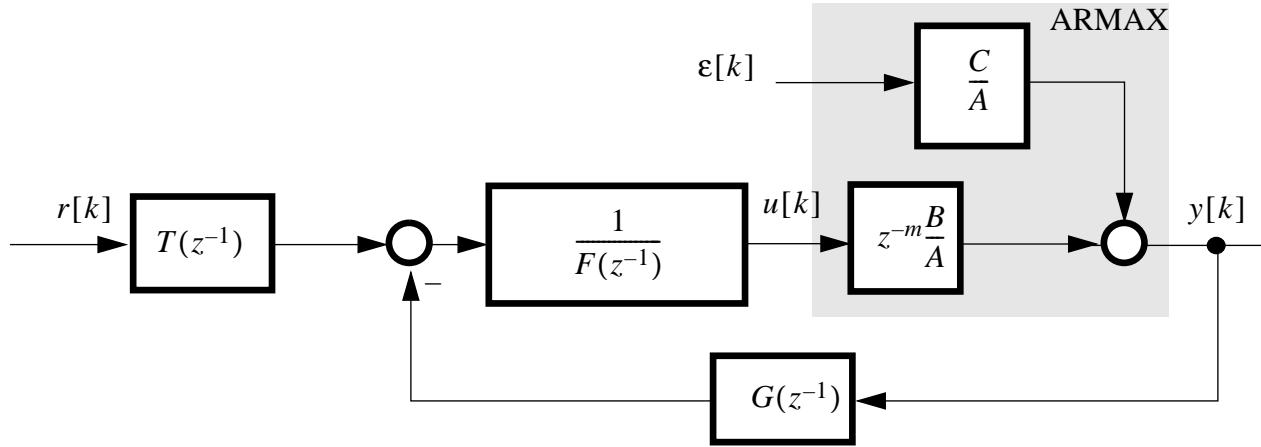
$$BFu[k] = Cr[k] - Gy[k] \quad (5.46)$$

Når $G(z)$ har grad $n-1$, så vil dette polynomet ivareta informasjon om det regulerte systems orden. Jo høyere orden på det åpne systemet (grad n på A -polynomet), jo flere gamle verdier av $y[k]$ inngår i regulatorformelen, for at valget av $u[k]$ skal tas under hensyn til den påvirkning gamle verdier av $y[k]$ har på systemet for diskret tid $t > k$.

Når $F(z)$ har grad $m - 1$, vil dette polynomet ivareta nødvendig informasjon om tidligere påsatte pådrag som ennå ikke har sluttet å virke på systemutgangen $y[k]$. Regulatorformelen må inneholde disse tidlige verdier av $u[k]$, siden de vil virke inn på systemet for $t > k$.

5.5 En generell lineær regulator

Strukturen i figur 5.11 er en generalisering av den i figur 5.8, med $B_1(z^{-1}) = 1$ dvs. $B_2(z^{-1}) = B(z^{-1})$:



Figur 5.11

Vi danner en polynomidentitet

$$P \equiv AF + z^{-m}BG, \quad \text{der } \text{grad}(A) = n_A \text{ og } \text{grad}(B) = n_B \quad (5.47)$$

Merk at nå er det forskjellige grader på A - og B -polynomet (men vi vil alltid ha $n_A \geq n_B$ siden systemet er forutsatt propert).

Hvis A og B ikke har felles faktorer som kan forkortes bort, kan polynomene F med $\text{grad}(F) = m + n_B - 1$, og G med $\text{grad}(G) = n_A - 1$, velges slik at $P(z^{-1})$ blir et *vilkårlig* polynom med maksimal grad = $m + n_A + n_B - 1$.

Det lukkede system blir

$$H_{ry}(z^{-1}) = \frac{\frac{T}{F} z^{-m} \frac{B}{A}}{1 + \frac{G}{F} z^{-m} \frac{B}{A}} = \frac{T z^{-m} B}{AF + G z^{-m} B} = \frac{z^{-m} TB}{P} \quad (5.48)$$

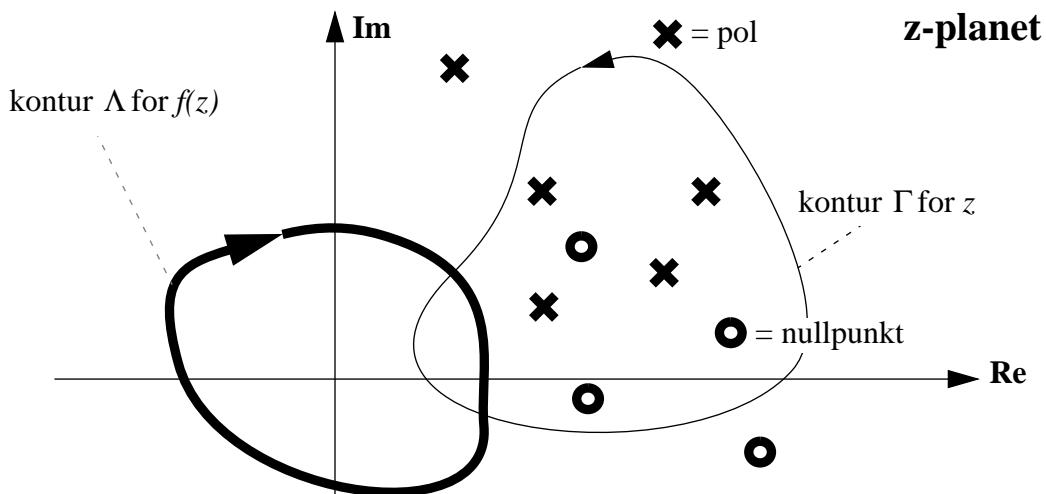
Vi ser altså at vi kan gi det lukkede systemet vilkårlig plassering av polene. Dead-beat er et spesialtilfelle av polpllassering, der alle poler plasseres i origo. Hvis man velger polene i P slik at de er noe mer vekk fra origo vil vi få langsommere innsvingning, og den vil bli asymptotisk slik som vi kjener den fra kontinuerlige systemer. Det er bare alle poler i origo som gir eksakt innsvingning etter et endelig antall tidsskritt, dvs. dead-beat. Men ved å velge polene noe vekk fra origo, vil man få en gevinst: Systemet blir mer robust, dvs. det tåler modellfeil i større grad uten å bli ustabilt. I motsetning til dead-beat-regulering som utnytter kunnskapen om det fysiske systemet maksimalt, med den ulempe at da blir følgene av feil modellering desto større.

5.6 Stabilitet og frekvensanalyse

Polplasseringsmetoder gir ingen informasjon om forsterkningsmargin og fasemargin, slik som frekvensanalytiske metoder gir. Men vi kan gjerne plotte transferfunksjoner for diskrete systemer i frekvensplanet. Hvis vi skal plotte en transferfunksjon H_0 , kan det f.eks. skje som $H_0(z)$ med $z = e^{j\Omega} = e^{j\omega T}$.

5.6.1 Nyquists stabilitetskriterium for diskrete systemer

Nyquists stabilitetskriterium (NS) kan brukes på samme måte for diskrete transferfunksjoner som for kontinuerlige (for kontinuerlige systemer henvises til Balchen: "Reguleringsteknikk", s. 227 og utover, eller Haugen: "regulering av dynamiske systemer", kap. 9.4). NS baserer seg på det såkalte **argumentvariasjonsprinsippet** fra kompleks funksjonsteori: Hvis en funksjon $f(z)$ har N nullpunkter og P poler innafor en gitt kontur slik som vist på fig 5.12., så vil konturen Λ for $f(z)$, som vi får når z gjennomløper Γ , omslutte origo $N - P$ ganger.



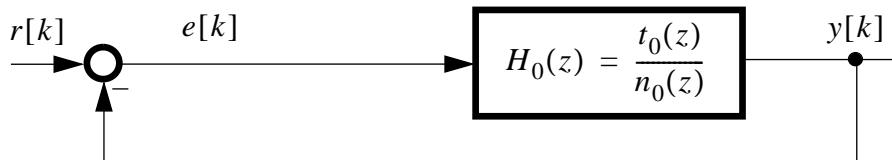
Figur 5.12

I figuren har vi én pol mer enn nullpunkter innafor Γ , ergo vil Λ omslutte origo én gang med urviseren (med urviseren, fordi poler er i *nevneren*, i motsetning til nullpunkter).

La oss bruke dette på diskrete systemer. For enkelhets skyld antar vi at vi har en situasjon med et uregulert (dvs. åpent) stabilt system gitt ved sløyfetransferfunksjonen $H_0(z)^{-1}$ (det som nå følger kan utmerket godt gjennomføres med transferfunksjonen på formen $H_0^*(z^{-1})$). Vi skal nå la funksjonen $R(z) = 1 + H_0(z)$ svare til $f(z)$, og konturen Γ for z vil være enhetssirkelen.

1. $H_0(z)$ svarer til **sløyfetransferfunksjonen** for kontinuerlige systemer, dvs. produktet av alle blokker rundt sløyfa, inkludert regulatortransferfunksjonen.

Vi antar at systemet reguleres med en enhetstilbakekopling, se figur 5.13



Figur 5.13

Det lukkede system blir

$$H(z) = \frac{H_0(z)}{1 + H_0(z)}, \text{ eller } H(z) = \frac{t_0(z)}{n_0(z) + t_0(z)}, \quad (5.49)$$

der $H_0(z) = t_0(z)/n_0(z)$. Vi forutsetter at $H_0(z)$ er *proper*, dvs.

$\text{grad}[t_0(z)] \leq \text{grad}[n_0(z)]$. Da blir i følge (5.49) antall poler til $H_0(z) =$ antall poler til $H(z)$. Polene til $H(z)$ finnes som *nullpunktene* til $1 + H_0(z)$. Det er disse som skal forbli innafor enhetssirkelen for å sikre stabilitet. *Polene* til $1 + H_0(z)$ er uforandret; de er de samme som polene til $H_0(z)$ og ligger som forutsatt (åpent stabilt system) innafor enhetssirkelen. Vi ønsker altså å forsikre oss om at ingen poler *innafor* enhetssirkelen forflytter seg utafor når vi kopler tilbake. Dette blir da ekvivalent med å kreve at uttrykket

$$R(z) = 1 + H_0(z) \quad (5.50)$$

har *like mange* nullpunkter som poler innafor enhetssirkelen, dvs. $N = P$.

Hvis vi lar Ω løpe fra $-\pi$ til π , svarer det til at $z = e^{j\Omega}$ følger enhetssirkelen én omdreining mot urviseren (konturen Γ). $H_0(z)$ gir da frekvensresponsen til det åpne system. Denne kan plottes i polardiagram (Nyquistdiagram). Hvis den polare stedkurve for $R(z)$ (konturen Λ) da omslutter origo, eller ekvivalent, konturen for $H_0(z)$ omslutter punktet -1, forteller det at antallet nullpunkter til $R(z)$ er blitt færre enn antall poler. $N - P$ nullpunkter til $R(z)$ (som jo er poler til $H(z)$) har flyttet seg utafor enhetssirkelen etter tilbakekopling, og $N - P$ kan avleses som antall ganger den polare stedkurve omslutter punktet -1. Derfor:

Den polare stedkurve for $H_0(z)$ må ikke omslutte punktet -1 hvis et åpent stabilt system gitt ved transferfunksjonen $H_0(z)$ skal forbli stabilt når det koples tilbake med enhetstilbakekoppling.

Dette er **Nyquists stabilitetskriterium** for åpent stabile diskrete systemer,

Som vi ser er det helt analogt med kriteriet for kontinuerlige systemer. Forskjellen er at konturen som gjennomløpes i z-planet er enhetssirkelen, mens den er den imaginære akse + en uendelig stor halvsirkel i s-planet for kontinuerlige systemer.

Dette stemmer bra med avsnitt 3.3.2, hvor det påpekes at venstre halvplan for kontinuerlige systemer svarer til innsida av enhetssirkelen for diskrete systemer.

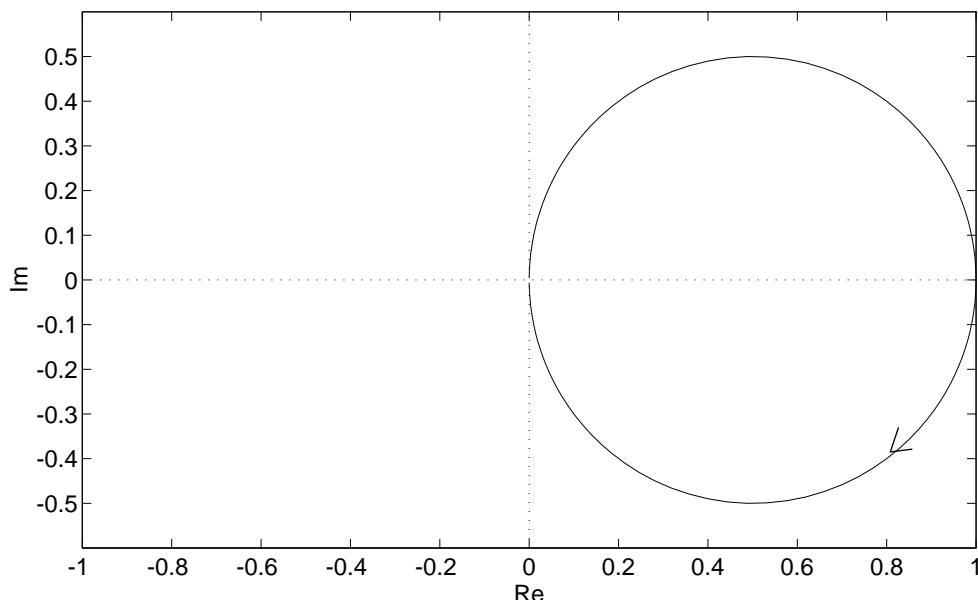
Vi skal se på et eksempel:

5.6.2 Eksempel: Stabilitet i kontinuerlig og diskret 1. ordens system

Nyquist-kurven (den polare stedkurve) for et 1.-ordens kontinuerlig system med proporsjonalregulering, gis av sløyfetransferfunksjonen

$$H_0(s) = \frac{K}{1 + T_1 s} , \quad (5.51)$$

der vi lar regulatorforsterkninga inngå i K . Kurven er vist på figur 5.14 for $K = 1$ og $T_1 = 1$. (Som en kuriositet: Kurven blir en sirkel. Prøv å vise det!)

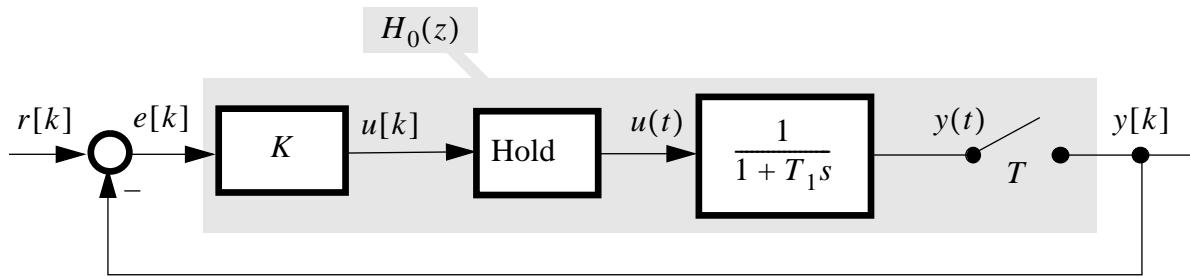


Figur 5.14

Systemet koples tilbake med enhetstilbakekopling slik som vist i figur 5.13. K er da regulatortransfertfunksjonen i serie med den egentlige prosessen, men inkorporert i $H_0(s)$. Som vi ser av figur 5.14 kan dette systemet aldri bli ustabilt uansett hvor mye vi øker K , noe vi vet fra reguleringsteknikken. Kurven vil bare blåses opp mot høyre, beholde sin form, men aldri omslutte -1. Sagt på en annen måte: Siden fasegangen aldri kan gå under -90° vil stedkurven aldri kunne krysse den negative reelle akse. Da måtte fasegangen i tilfelle ha gått under -180° .

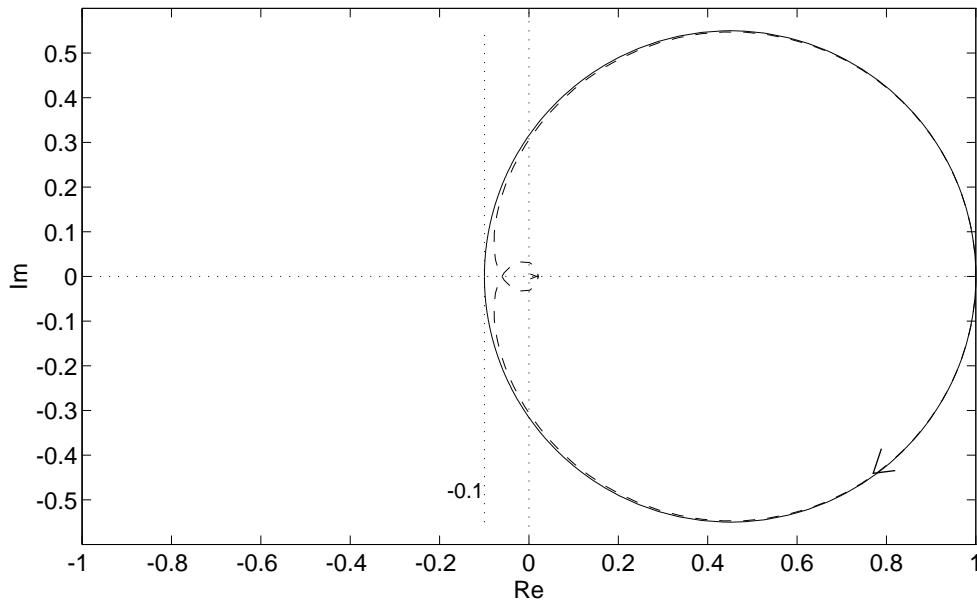
La oss nå i stedet se på et diskretisert 1.ordens system, styrt med (diskret) proporsjonalregulator. Strukturen blir som vist på figur 5.15: Vi har

$$H_0(z) = \frac{K(1 - e^{-T/T_1})}{z - e^{-T/T_1}} \quad (5.52)$$



Figur 5.15

Vi setter $T = 0.2$, med $T_1 = 1$ som før. Den heltrukne kurven i figur 5.16 er polardiagrammet til $H_0(z)$ (den stiplete kurven som i stor grad faller sammen med den heltrukne, kommer vi snart tilbake til):

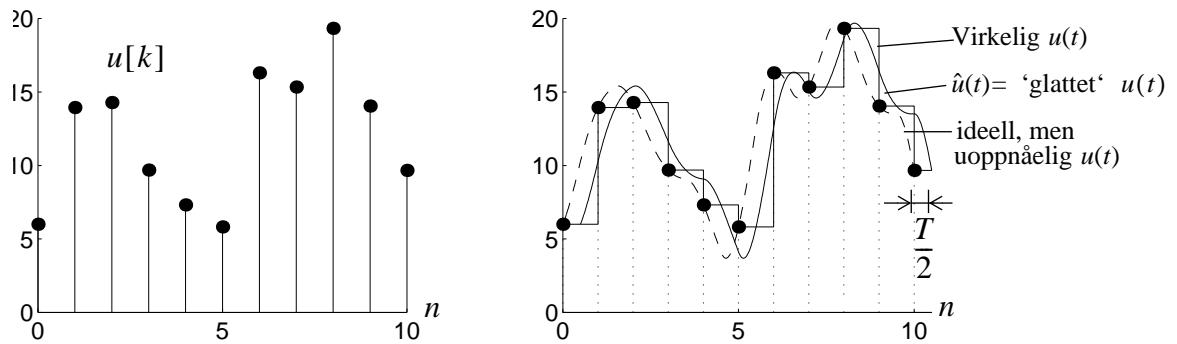


Figur 5.16

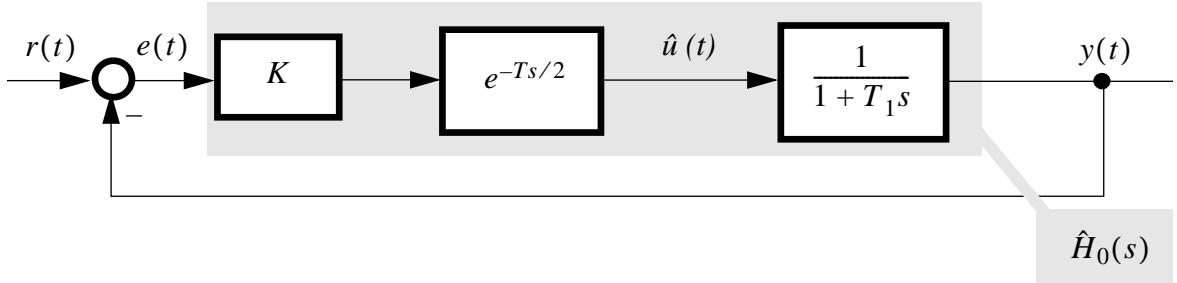
Som vi ser kan dette systemet gjøres ustabilt ved å øke K . Forsterkningsmarginen, avlest fra figuren, blir ca. $\Delta K \approx 1/0.10 \approx 10$. Systemet kommer altså på stabilitetsgrensa for $K = 10$.

Forklaringen på at systemet nå er mindre robust enn med kontinuerlig regulator, er først og fremst den faseforsinkelse som introduseres av holdeelementet, se figur 5.17:

Som vi ser svarer holdeelementets virkning grovt sett til et kontinuerlig system hvor det er lagt inn en tidsforsinkelse på $T/2$ i sløyfa, se figur 5.18:



Figur 5.17



Figur 5.18

Den faseforverring som dette innebærer, er den alvorligste negative bivirkning av diskret regulering. Vi kan innføre den kontinuerlige tilnærminga vist i figur 5.18 for å få et brukbart anslag over virkningen av holdelementet, ihvertfall hvis tastetida er rimelig kort, la oss si ca. 5 ganger mindre en minste tidskonstant i det kontinuerlige systemet. I figur 5.16 er polar kurven for det kontinuerlige systemet

$$\hat{H}_0(s) = \frac{Ke^{-Ts/2}}{1 + T_1 s} \quad (5.53)$$

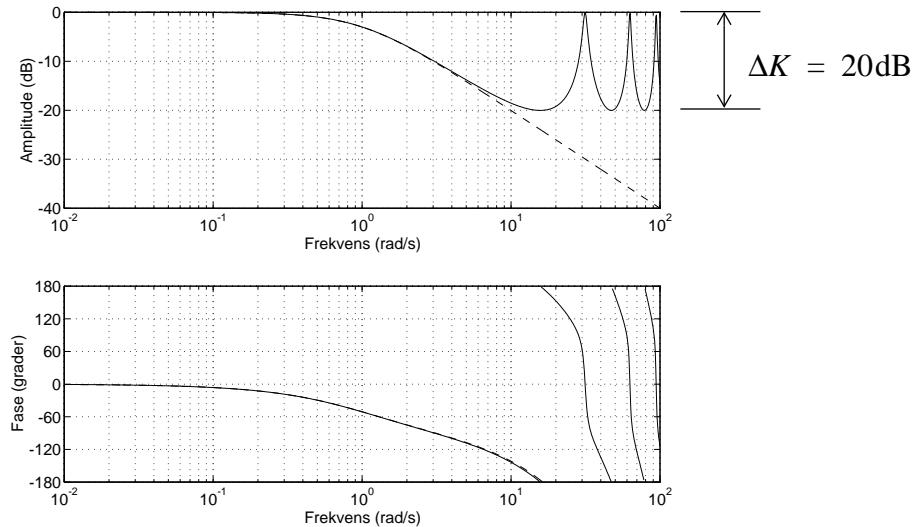
vist stiplet. Vi kjenner igjen den karakteristiske spiralformen som skyldes transportforsinkelsen. Ellers ser vi at det er svært bra overensstemmelse med frekvensresponsen for det diskrete systemet for lavere frekvenser. Vi ser også at forsterkningsmarginen blir av samme størrelsesorden, sjøl om den er litt bedre for det kontinuerlige systemet med transportforsinkelse.

5.6.3 Bode(AFF) - diagram

Som vi kjenner fra reguleringsteknikken for kontinuerlige systemer kan vi også bruke NS med utgangspunkt i et Bode-diagram. Kravet om at den polare stedkurve for H_0 ikke skal omslutte punktet -1 betyr i et Bode-diagram at amplitudeforløpet $|H_0|$ skal gå under 0 dB før fasevinkelen $\angle H_0$ går under -180° , dvs. kryssfrekvensen må tilfredsstille $\omega_c < \omega_{180}$.

Figur 5.19 viser Bode-diagrammet for den diskrete sløyfetransferfunksjonen $H_0(z)$ (heltrukken linje) og den kontinuerlige tilnærmelsen $\hat{H}_0(s)$ (stiplet linje):

$H_0(z)$ med
 $z = e^{j\Omega} = e^{j\omega T}$
og $\hat{H}_0(s)$ med
 $s = j\omega$



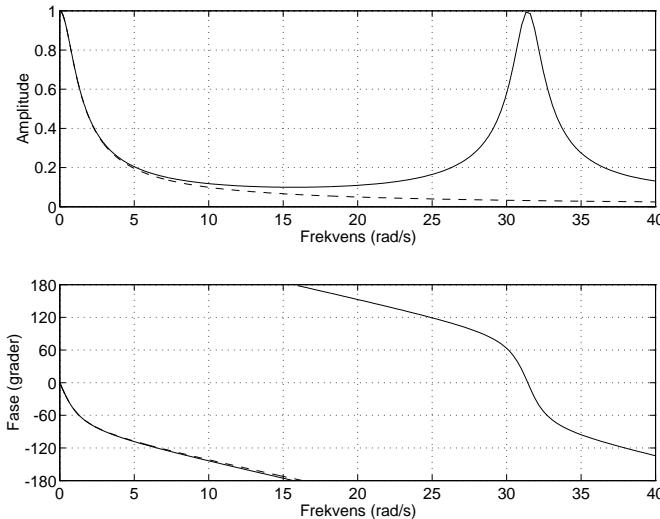
Figur 5.19

Periodisiteten i $H_0(z)$ kommer fram som bulker i amplitudeforløpet. Igjen ser vi den gode overensstemmelsen mellom $H_0(z)$ og $\hat{H}_0(s)$ ved lavere frekvenser. Frekvensgangen til $H_0(z)$ (heltrukken linje) og den kontinuerlige tilnærmelsen $\hat{H}_0(s)$ (stiplet linje) er også plottet i *lineær* skala langs begge akser i figur 5.20. Her blir ikke periodene sammentrykt mot høyre slik som med log- ω -skalaen i figur 5.19.

5.6.4 Nyquists stabilitetskriterium for åpent ustabile systemer

Vi forutsetter at $H_0(z)$ nå er åpent ustabilt, dvs. minst én pol ligger utefor enhetssirkelen. Polene til $1 + H_0(z)$ er uforandret; de er de samme som polene til $H_0(z)$. Nå blir kravet at den eller de poler som ligger utefor enhetssirkelen blir flyttet innafør den etter tilbakekoppling, for å sikre stabilitet. Dette blir da ekvivalent med å kreve at uttrykket

$$R(z) = 1 + H_0(z) \quad (5.54)$$



Figur 5.20

får så mange flere nullpunkter enn poler innafor enhetssirkelen, at det svarer til antall poler utafor enhetssirkelen i det ustabile åpne system.

Sagt med andre ord:

Den polare stedkurve for $H_0(z)$ må omslutte punktet -1 mot urviseren for at et åpent ustabilt system gitt ved transferfunksjonen $H_0(z)$ skal bli stabilt når det koples tilbake med enhetstilbakekopling. Omslutningen må skje like mange ganger som det er poler utafor enhetssirkelen i det åpne system $H_0(z)$.

Dette er Nyquists stabilitetskriterium for åpent ustabile diskrete systemer.

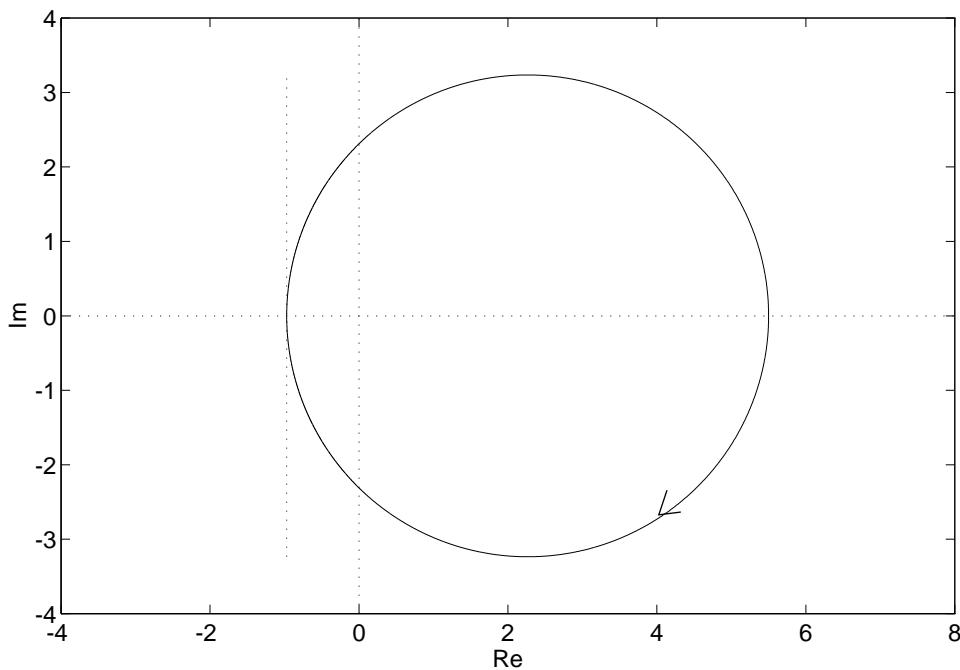
5.6.5 Frekvensanalyse av systemer med dead-beat-regulator

Vi tar utgangspunkt i et system med dead-beat-regulator type 1, se figur 5.2. Her har vi ikke enhetstilbakekopling, så hvis vi skal anvende NS, må vi sette

$$H_0(z^{-1}) = \frac{G}{BF} z^{-m} \frac{B}{A} \quad (5.55)$$

som vi kjenner igjen som sløyfetransferfunksjonen. (Her har vi brukt $H_0(z^{-1})$ i stedet for $H_0(z)$, men det gjør ingen forskjell når man skal se på polardiagram og frekvensrespons). Vi må nå først sjekke nullpunktene i A -, B - og F -polynomene, for å konstatere om uttrykket (5.55) er stabilt eller ustabilt. Deretter kan man plotte polarkurven og anvende en av de to versjonene av NS-kriteriet avhengig av om vi har åpen stabilitet eller ustabilitet.

La oss nå prøve dette på systemet i avsnitt 5.2.4, dvs. "eksempel 2". Vi så seinere at dette systemet var svært nær stabilitetsgrensa, sjøl ved perfekt overensstemmelse mellom modell og fysisk virkelighet. A -, B - og F -polynomene har alle nullpunktene innafor enhetssirkelen, så systemet er åpent stabilt, og derfor skal kurven ikke omslutte -1. Dett er NS-kriteriet versjon 1. Figur 5.21 viser polarkurven for H_0 ved perfekt overensstemmelse. Vi ser at forsterkningsmarginen er svært liten, som forventet.



Figur 5.21

Vi så seinere, i avsnitt 5.2.5, at systemet ble ustabil ved en liten uoverensstemmelse i en eneste parameter, nemlig koeffisienten foran z^{-1} -leddet i B -polynomet. Figur 5.22 viser polarkurven ved feil antagelse, med koeffisienten $b = 0.995$ i stedet for 0.99, som er den riktige.

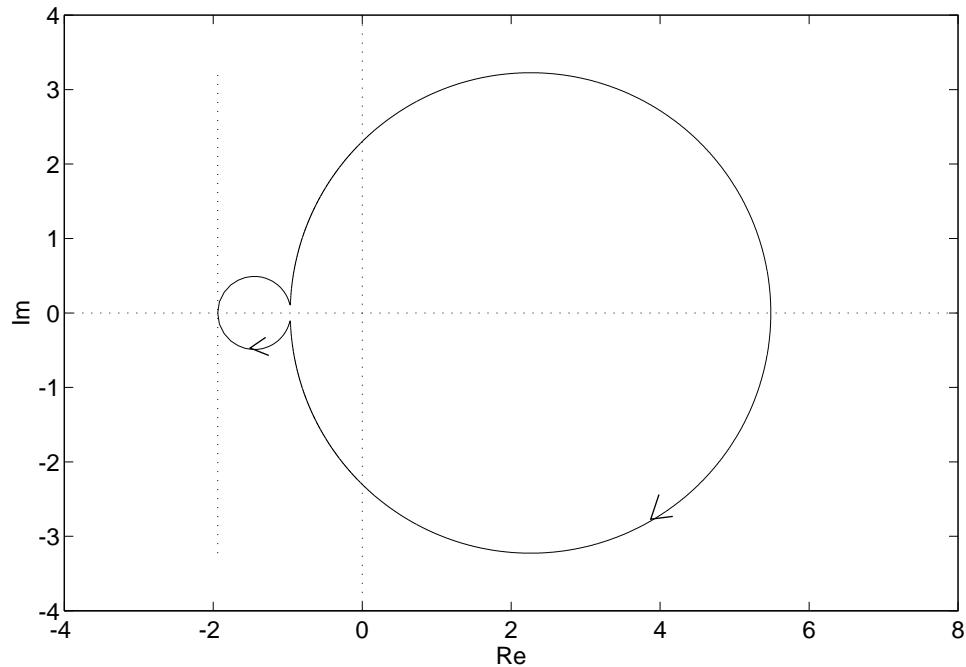
H_0 blir fortsatt åpent stabilt (det kan du sjekke hvis du vil), så vi må anvende NS versjon 1. Vi ser at det har oppstått en liten "utvekst" til venstre på polarkurven, som gjør at -1 nå ligger innafor kurven. Kurven omslutter -1 én gang, det betyr at én pol har flyttet seg utafor enhetssirkelen, det lukkede system er ustabilt.

Denne kurven svarer til de impulsresponser som vist i midten på figur 5.7. De er ustabile, og svinger seg eksponensielt ut på grunn av den ene polen vi har fått utafor enhetssirkelen.

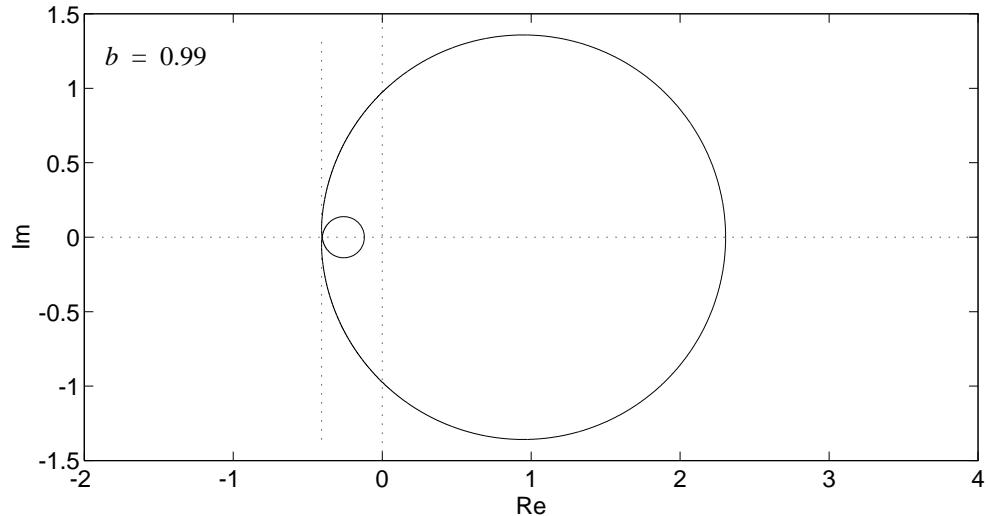
La oss nå istedet se på polarkurver med dead-beat-regulator type 2, altså den mer robuste men langsomme varianten. Da får vi, jfr. figur 5.8:

$$H_0(z^{-1}) = \frac{G}{B_1 F} z^{-m} \frac{B}{A} \quad (5.56)$$

Dette er også åpent stabilt, dvs. vi må bruke NS versjon 1. Polarkurven ved perfekt overensstemmelse, dvs. antatt $b = 0.99 =$ virkelig b , er vist på figur 5.23:



Figur 5.22



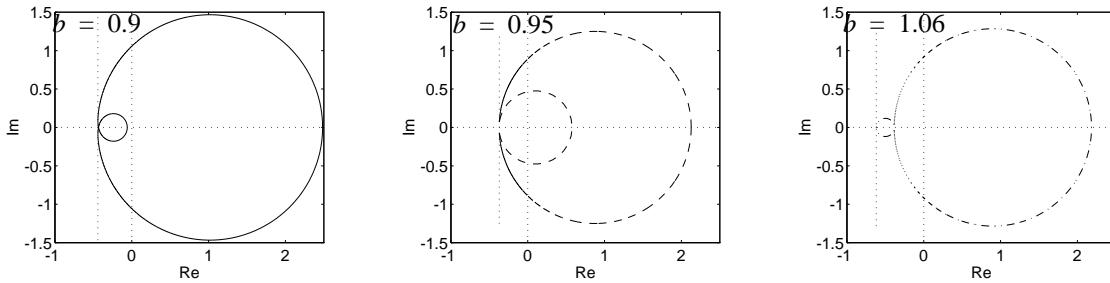
Figur 5.23

Vi ser at vi har god forsterkningsmargin. Robustheten til regulatoren kommer fram i figur 5.24, hvor vi har tre polarkurver for feilantagelsene $b = 0.95, 1.05$ og 1.1 . Systemet er åpent stabilt i alle tre tilfeller, og vi må anvende NS versjon 1.

For den generelle lineære regulator i avsnitt 5.5 får vi

$$H_0(z^{-1}) = \frac{G}{F} z^{-m} \frac{B}{A} \quad (5.57)$$

Sløyfeforsterkningen, og dermed forsterkningsmarginen, vil da kunne justeres ved å multiplisere alle koeffisienter i G -polynomet med en konstant.



Figur 5.24

Forsterknings- og fasemargin vil også kunne endres ved å spesifisere et annet polynom P , som dermed vil gi andre polynomer F og G via identiteten (5.43). Når man har valgt et annet P -polynom, må man først sjekke om systemet blir åpent stabilt eller ustabilt med det nye F -polynomet man da får. Så kan man velge den korresponderende versjon av Nyquist-kriteriet og sjekke stabilitetsmarginer.

Igjen: En tommelfinger-regel er at jo nærmere origo man insisterer på at polene i P skal legges (dvs. nærmere dead-beat-regulering), jo mindre stabilitetsmargin risikerer man å få.

Merk: Frekvensresponsanalyse basert på transferfunksjon fra inngang til utgang, forteller ingen ting om tilstander inne i systemet hvor de tilsvarende poler *forkortes bort* i transferfunksjonen fordi tilstanden ikke er observerbar, ikke styrbar, eller begge deler. Hvis polen svarende til en slik tilstand ligger utover enhetssirkelen vil man ha et system som er åpent ustabilt uten at det framkommer i polardiagrammet eller Bode-diagrammet til transferfunksjonen H_0 . Det samme gjelder sjølsagt for kontinuerlige systemer, og skulle være kjent fra grunnkurs i reguleringsteknikk.

5.7 Kildehenvisning

En meget bra bok om polplasseringssmetoder og diskrete systemer i sin alminnelighet er "Reglerteori- modrena analys- og syntesmetoder" av Lennart Ljung. Gitt ut på "Studentlitteratur", Lund i sverige.

Kapittel 6: Stokastiske prosesser

6.1 Innledning

I den klassiske reguleringsteknikken, som ble utviklet før den andre verdenskrigen, baserte man seg stort sett på deterministiske modeller, eller forestillinger, om de prosessene som man skulle regulere. Med begrepet **deterministisk modell** menes at kunnskapen om systemets nåværende tilstand og de framtidige inngangssignalene setter en i stand til å **prediktere** (forutsi) presis hva som vil skje i framtida. Dette er selvfølgelig aldri mulig i praksis, men likevel vil deterministiske modeller ofte gi en god nok beskrivelse av den prosessen eller det systemet man står overfor. Et eksempel på en god deterministisk modell er likningene for månens bevegelse. Denne modellen setter oss i stand til å forutsi månens posisjon og hastighet for meget lang tid fremover.

Imidlertid finnes det andre systemer som må modelleres som ikke-predikterbare, eller ikke-deterministiske. La oss se på et slikt system.

- En båt/plattform som beveger seg på havet, er et typisk ikke-deterministisk system. Fartøyet utsettes for bølger, strøm, vind og vær. Alle disse gir opphav til krefter som påvirker fartøyet. Kreftene er vanskelige eller umulige å modellere i deterministisk forstand. Vi vet alle at havbølger ikke kan beskrives som velformede tidsfunksjoner, men at disse har et mer eller mindre tilfeldig forløp. Hvis vi prøver å prediktere en bølgebevegelse inn i fremtiden, vil usikkerheten eller prediksjonsfeilen snart øke på grunn av bølgens tilfeldige karakter. Selv om bølgen har en tilfeldig karakter, er det imidlertid visse karaktertrekk ved bevegelsen. Vi innser umiddelbart forskjellen på knapp sjø og på store, lange havdønninger.
- Likeså vil vinden variere på en mer eller mindre usikker måte. Vi kunne f.eks. tenke oss å inkludere en modell av været for bedre å kunne forutsi vinden og dens innvirkning. Alle vet imidlertid at meteorologiske modeller, som de brukes i værvarsling, heller ikke er fri for usikkerhet (for å si det forsiktig).
- De langsomme havstrømmenes innvirkning kan forutsies med større sikkerhet da disse i stor utstrekning drives av månens og solas bevegelser (tidevann), men også her er usikkerheter.
- Alle disse usikre kreftene medvirker at fartøyets posisjon ikke kan predikteres nøyaktig. Usikkerheten blir dessuten større etterhvert som tiden går.

For å kunne regulere et slikt system er det ofte nødvendig å kunne ta hensyn til karakteren av de ulike forstyrrelsene og usikkerhetene. Slike “usikre systemer”, eller systemer som er utsatt for tilfeldige påvirkninger (**stokastiske prosesser**, **støy**, kjært barn har mange navn), vil vi i dette kurset kalte **stokastiske systemer**. Et “vanlig” lineært system blir altså “stokastisk” når det påvirkes av en eller flere stokastiske prosesser.

Det stokastiske eller ikke predikterbare elementet i en prosess kan ha flere årsaker:

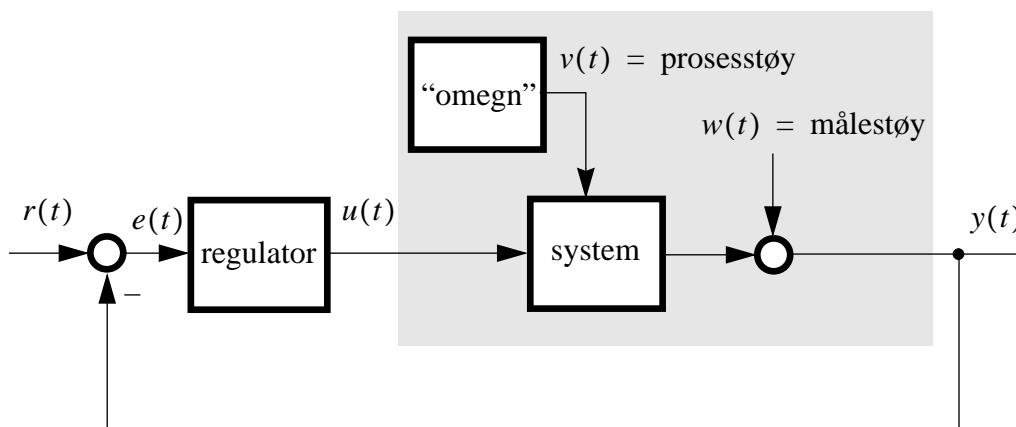
1. Inngangssignalene eller påvirkningene er usikre, slik som beskrevet ovenfor. Alle systemer er egentlig åpne: dvs. de er i interaksjon med andre systemer. Dette betyr at ingen prosess kan dekoples fra resten av universet hvis en beskrivelse skal være maksimalt nøyaktig. Avgrensning av det systemet en skal beskrive, gir en modell hvor påvirkninger ("støy") fra omgivelsene ("omgaven") eksiterer systemet på en stokastisk eller ikke predikterbar måte.
2. Våre målinger er unøyaktige eller ufullstendige. Dette gir oss en usikkerhet i tilstanden for systemet. Dette medfører prediksjonsfeil.
3. Vårt kjennskap til de fysiske lover er ufullstendig. Dette medfører prediksjonsfeil. (M.h.p. fartøyet i eksemplet over har vi f.eks. ufullstendig kunnskap om hvordan ror- og propell-krefter virker inn).
4. Selv med perfekt kunnskap om de fysiske lovene er vi nødt til å operere med enklere modeller enn det som er teoretisk korrekt. Et hvert system er i virkeligheten uendelig komplisert.

Sjøl om bare punkt 1. relaterer seg direkte til støy som fysisk påvirker et system, vil også 2. - 4. ofte kunne ivaretas ved å forutsette at systemets usikkerhet skyldes forstyrrelser som påvirker det. Denne delen av kurset skal omhandle analyse, beskrivelse og metoder for regulering i slike stokastiske systemer. Vi vil fortsatt begrense oss til å behandle systemer som er:

- monovariable, dvs.
én-inn, én-ut, eller på engelsk: SISO - Single Input, Single Output,
- lineære og tidsinvariante(LTI).

Både kontinuerlige og diskrete lineære systemer vil bli behandlet. Vi vil i første del av denne "stokastiske" delen av kurset utvikle den teorien som gjelder for *kontinuerlige* systemer og signaler. Seinere vil vi se at overgangen til diskrete systemer og signaler er triviell.

De aktuelle problemstillingene kan belyses med utgangspunkt i figur 6.1 (som viser situasjonen i det kontinuerlige tilfelle, men det blir tilsvarende for det diskrete):



Figur 6.1

Vi skal ta for oss følgende oppgaver:

1. Klassifisering av og matematisk beskrivelse av støyen som påvirker systemet.
2. Med utgangspunkt i dette, og en modell av systemet: Hva blir virkningen på systemets utgang på grunn av støy?
3. Hvordan velge en regulator som gir “best” resultat når støyens karakter og systemmodellen er kjent?

Nødvendigheten av å ta hensyn til støy og ikke-predikterbare forstyrrelser i reguleringssystemer ble man klar over relativt tidlig i fagets historie. Mange av de klassiske metodene innen reguleringsteknikken muliggjorde enkle og intuitivt baserte frekvensanalytiske metoder for å ta hensyn til dette, som vist i følgende sitat fra A.C. Halls bok “Frequency Response” utgitt på MacMillan i 1954:

I well remember an instance in which M.I.T. and Sperry were co-operating on a control of an air-borne radar, one of the first such systems to be developed. Two of us had worked all day in the Garden City Laboratories on Sunday, December 7, 1941, and consequently did not learn of the attack on Pearl Harbor until late in the evening. It had been a discouraging day for us because while we had designed a fine experimental system for test, we had missed completely the importance of noise with the result that the system's performance was characterized by large amounts of jitter and was entirely unsatisfactory. In attempting to find an answer to the problem we were led to make use of frequency-response techniques. Within three months we had a modified control system that was stable, had a satisfactory transient response, and an order of magnitude less jitter. For me this experience was responsible for establishing a high level of confidence in the frequency-response techniques.

Det kan altså slås fast at det ofte er nødvendig å innføre realistiske modeller for forstyrrelser og signaler av ikke predikterbar karakter. Et øyeblinkks refleksjon tilsier at det ikke er trivielt å modellere denne typen fenomener. Det er f.eks. nytteløst å modellere en stokastisk prosess ved bruk av analytiske funksjoner, for hvis en analytisk funksjon er kjent over et endelig lite intervall i tid, kan den forutsies for all framtid. Neste forsøk er da naturlig nok å modellere slike prosesser ved hjelp av stokastiske begreper. Fra tidlig litteratur innen dette feltet kan man se at dette heller ikke har vært lett. Et forsøk var å modellere en forstyrrelse, $x(t)$, ved

$$x(t) = \sum_{i=1}^n a_i(t) \xi_i \quad (6.1)$$

hvor $a_1(t), a_2(t), \dots, a_n(t)$ er kjente funksjoner og ξ_i er tilfeldige variable. Men da ser vi at vi kan beregne ξ_i etter å ha målt $x(t_1), x(t_2), \dots, x(t_n)$ fra likningene

$$\begin{aligned} x(t_1) &= a_1(t_1)\xi_1 + a_2(t_1)\xi_2 + \dots + a_n(t_1)\xi_n \\ x(t_2) &= a_1(t_2)\xi_1 + a_2(t_2)\xi_2 + \dots + a_n(t_2)\xi_n \\ &\vdots \\ x(t_n) &= a_1(t_n)\xi_1 + a_2(t_n)\xi_2 + \dots + a_n(t_n)\xi_n \end{aligned} \quad (6.2)$$

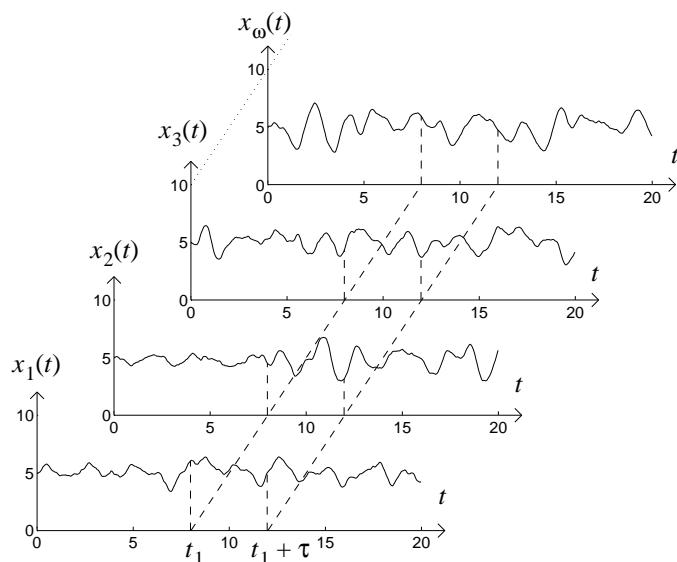
så sant disse er lineært uavhengige. For tida $t > t_n$ er dermed $x(t)$ fullstendig bestemt og det ikke predikterbare element faller bort.

Først etter at man var begynt å formulere modeller av stokastiske prosesser hvor prosessens variable til et hvert tidspunkt er tilfeldige variable i statistisk forstand, kom man inn i et fruktbart spor. Dette skjedde i perioder under og etter annen verdenskrig. De største bidragsytere til utviklingen på dette området var matematikere som Cramér, Khintchine, Kolmogorov og Wiener.

Definisjonen av en stokastisk prosess er gitt i det følgende. Denne definisjonen inkorporerer usikkerhetsaspektet med hensyn til fremtidige verdier.

*En stokastisk prosess er en samling (et **ensemble**) av funksjoner, $\{x_\omega(t)\}$ ¹, hvor t er et tidsargument som kan være diskret $t \in \dots, -1, 0, 1, 2, 3, \dots$ eller kontinuerlig ($t \in [-\infty, \infty]$). Hver enkelt av disse funksjonene $x_\omega(t)$ hvor t er variabel og ω er gitt, kalles en **utfallsfunksjon** eller en **realisasjon** av den stokastiske prosessen. $x_\omega(t)$ for en gitt t er en tilfeldig variabel.*

I praksis vil en utfallsfunksjon eller en realisasjon være en måleserie fra et gitt eksperiment. Vi kan f.eks. tenke oss at eksperimentet måler posisjonen til en ferge ved overfarten mellom Moss og Horten eller temperaturforløpet i en kjemisk reaktor over et døgn. Vi skal i det følgende ofte benytte betegnelsen **tidsserie** i stedet for realisasjon. Den stokastiske prosessen i seg selv er mengden av alle realisasjoner eller tidsserier som kan resultere fra den gitte eksperimentsituasjon. Disse realisasjonene nummereres så med parameteren ω (som i dette tilfelle er et positivt heltall, og ikke har noe med frekvens å gjøre!). Figur 6.2 indikerer et ensemble som altså utgjør en stokastisk prosess.



Figur 6.2 Ulike realisasjoner av en stokastisk prosess $\{x_\omega(t)\}$.

Merk at begrepet ensemble og stokastisk prosess er et teoretisk begrep, og at en stokastisk prosess ikke er det samme som en *realisasjon* av en stokastisk prosess. En stokastisk prosess kan imidlertid kun observeres gjennom sine realisasjoner.

1. { } symboliserer "ensemble" = "samling av" = "mengde av".

Etter å ha definert begrepet stokastisk prosess, kommer vi til problemet: Hvordan kan vi i større detalj karakterisere en slik prosess? Som angitt i definisjonen vil en stokastisk prosess *for et gitt tidspunkt* være en tilfeldig variabel. Beskrivelse og analyse av tilfeldige variable er utførlig behandlet innen sannsynlighetsregning og statistikk og danner derfor grunnlaget for beskrivelsen av stokastiske prosesser. Vi skal her betrakte dette grunnlaget som kjent og kun repetere noen viktige trekk av sannsynlighetsregningen:

6.2 Matematisk grunnlag i sannsynlighetsregning

Som nevnt innledningsvis vil det være naturlig å beskrive stokastiske prosesser ved hjelp av sannsynlighetsbegreper. For å få et grunnlag for den videre behandlingen skal vi derfor i dette kapittel repetere noen grunnleggende trekk av sannsynlighetsregningen.

6.2.1 Tetthetsfunksjon og fordelingsfunksjon

Hvis utfallet av et eller annet eksperiment kan ges en numerisk verdi på en skala (f.eks. en temperatur eller en posisjon), sies utfallet å være en realisasjon av en **tilfeldig variabel**. La oss betrakte det tilfellet at vi har en samling av måleserier fra en gitt prosess. La oss anta at disse måleseriene er tatt opp under like forsøksbetingelser så langt disse kan kontrolleres, men hvor tilfeldige forstyrrelser medvirker til en viss forskjell mellom de ulike måleseriene. Vi betrakter hver av serie som en realisasjon av en stokastisk prosess. De ulike måleseriene fra en slik forsøksrekke er vist i figur 6.3, symbolisert ved indekser $\omega = 1, 2, 3, \dots$:

Vi ser at når vi foretar et snitt på et vilkårlig sted t_1 langs tidsaksen, så vil realisasjonen bli en tilfeldig variabel, og vi gjenkjenner (**sannsynlighets**)**tetthetsfunksjonen** $f(x; t_1)$, som vi kan lese som “f av x, gitt t_1 ”. Ved dette tidspunktet vil sannsynligheten for¹ at størrelsen x vil passere mellom a og b som vist på figuren være

$$P[a < x(t_1) < b] = F(b; t_1) - F(a; t_1) = \int_a^b f(x; t_1) dx \quad (6.3)$$

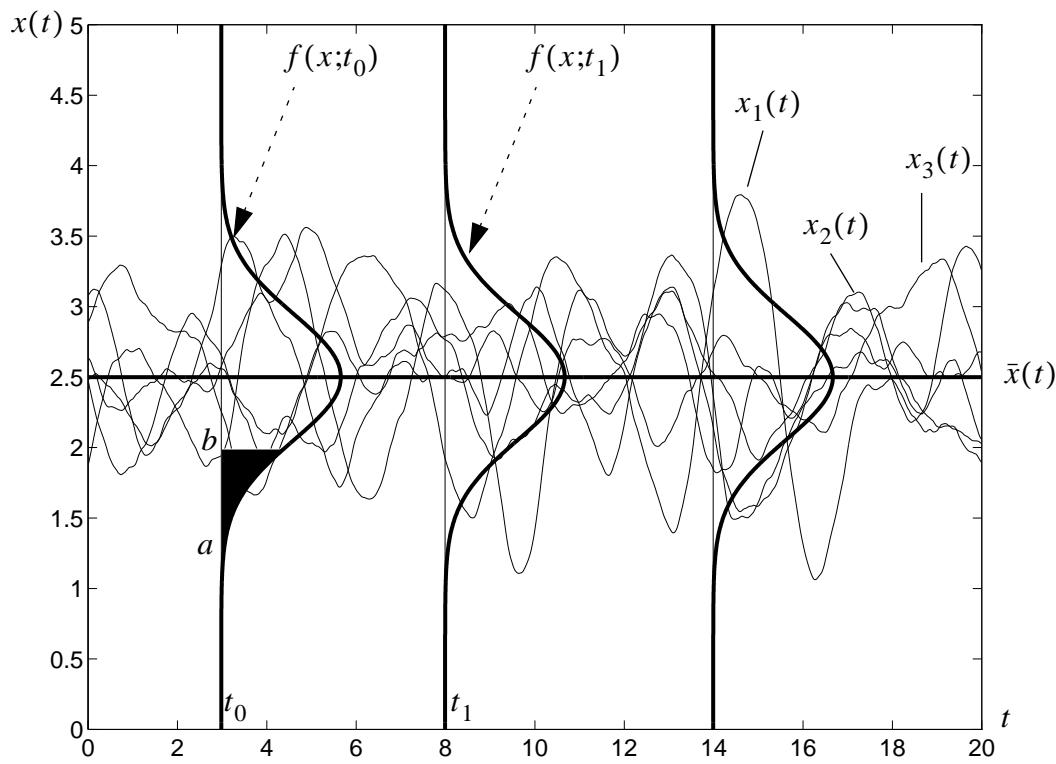
Her er $F(x; t_1)$ (**sannsynlighets**)**fordelingsfunksjonen** for x ved tidspunktet t_1 . Vi har

$$f(x; t_1) = \frac{\partial}{\partial x} F(x; t_1) \quad (6.4)$$

Sannsynligheten for at x vil inntrefte med en verdi lavere enn b ved tidspunktet t_1 blir

$$P[x(t_1) < b] = F(b; t_1) = \int_{-\infty}^b f(x; t_1) dx \quad (6.5)$$

1. Symbolet $P[]$ er stenografi for “sannsynligheten for”, mens $F()$ og $f()$ er funksjoner.



Figur 6.3

Vi har, som kjent fra sannsynlighetsregningen

$$F(\infty; t_1) = \int_{-\infty}^{\infty} f(x; t_1) dx = 1 \quad (6.6)$$

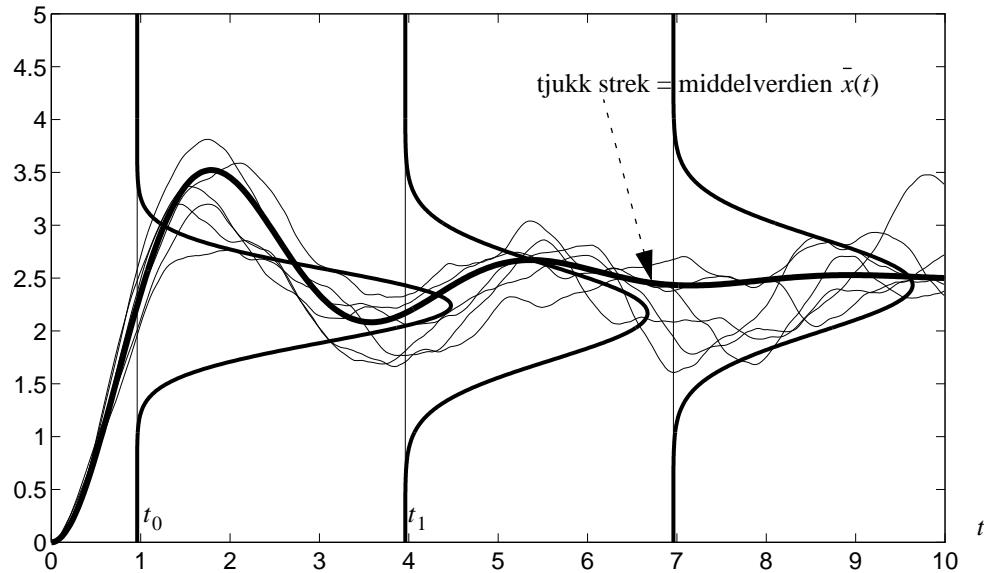
I figur 6.3 får vi inntrykk av at $f(x; t)$ er den *samme* for forskjellige verdier av t . Dette bekreftes også av den horisontale linjen $\bar{x}(t) = \text{konstant}$, som uttrykker middelverdien av $x(t)$. Dette er svært vanlig, og vi sier da at prosessen er **stasjonær**¹. Men slik trenger det ikke være: Figur 6.4 viser noen utfall av et stokastisk signal hvor tetthetsfunksjonen endrer seg med tida. Denne prosessen er ikke stasjonær.

Når vi har med fysiske, tidsvarierende målesignaler å gjøre bruker vi ofte begrepene **amplitudetetthetsfunksjon** og **amplitudefordelingsfunksjon**. Eller vi sier bare tetthetsfunksjon og fordelingsfunksjon.

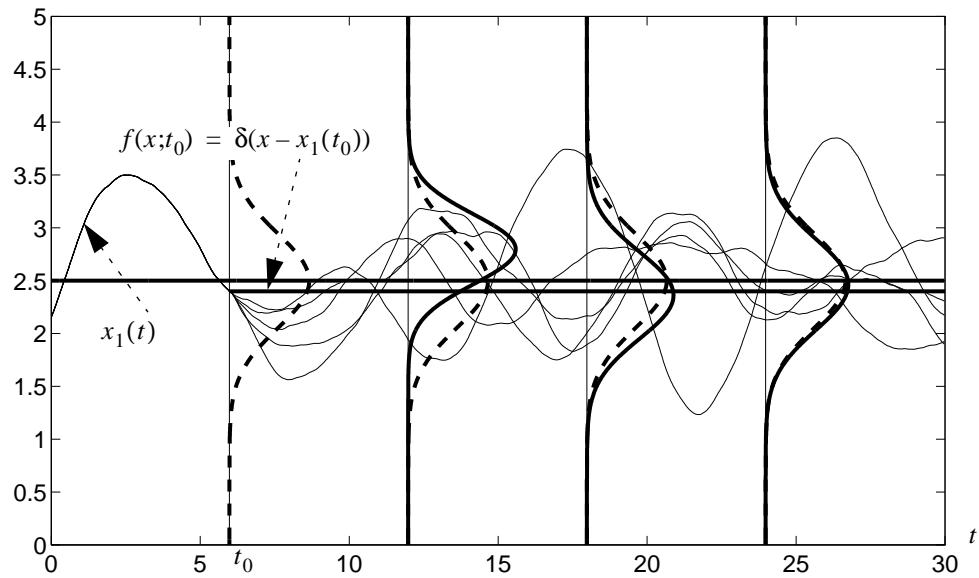
6.2.2 Betinget sannsynlighet og betinget tetthetsfunksjon

Dersom to tilfeldige variable betraktes simultant, har man definert den **simultane fordelingsfunksjonen** $F(x, y)$ og den tilhørende **simultane tetthetsfunksjonen** $f(x, y)$. x og y er her to tilfeldige variable. Vi vil også kunne være interessert i å bestemme tetthetsfunksjonen for x når vi vet at y allerede har funnet sted, og kjenner y sin verdi. Denne tetthetsfunksjonen kaller vi den **betingede tetthetsfunksjonen** og benevner den $f(x|y)$.

1. Grovt sett. Vi kommer tilbake til en mer presis definisjon på begrepet "stasjonær".



Figur 6.4



Figur 6.5

Figur 6.5 viser noen realisasjoner av samme prosess som i figur 6.3. Men nå har vi målt en realisasjon $x_1(t)$ fram til $t = t_0$, og da vet vi at den hadde verdien $x_1(t_0)$ ved $t = t_0$. Siden $x_1(t_0)$ nå er en deterministisk størrelse, blir "tetthetsfunksjonen" her en deltafunksjon¹. Den kunnskap vi har om $x(t_0)$ ved $t = t_0$ må påvirke de seinere tetthetsfunksjonene for $t > t_0$, dvs. vi får betingede tetthetsfunksjoner. Dette er vist på figur 6.5, som også antyder tetthetsfunksjonene fra figur 6.3, hvor man ikke hadde noen forhåndskunnskaper om bestemte verdier av x . Vi ser av figur 6.5 hvordan tetthetsfunksjonene likevel blir breiere med tida, dvs. usikkerheten øker, og etter en stund vil kunnskapen fra $t = t_0$ være uten verdi. Vi er tilbake i situasjonen i figur 6.3.

1. Faktisk kan et *deterministisk signal* $x(t)$ oppfattes som et spesialtilfelle av et stokastisk signal, hvor tetthetsfunksjonen er en deltafunksjon $f(x; t) = \delta(x - x(t))$ for alle verdier av t

Eksemplet over viste hvordan vi har behov for et begrepsapparat som håndterer simultan og betinget sannsynlighet når vi studerer *samme* stokastiske prosess ved forskjellige tidspunkter. En annen anvendelse er når vi studerer interaksjonen mellom to *forskjellige* signaler, $x(t)$ og $y(t)$. Vi kan f.eks. stille spørsmålet: På hvilken måte vil det at vi kjenner $x(t)$ ved $t = t_0$ influere på statistisk kunnskap om $y(t)$ ved $t = t_1$? Dette skal vi komme inn på seinere med de såkalte korrelasjonsfunksjoner, i denne omgang skal vi nøye oss med å repetere litt om simultan og betinget sannsynlighet.

Simultan sannsynlighet er definert gjennom den simultane tetthetsfunksjon:

Sannsynligheten for at størrelsen $x(t_0)$ vil inntreffe mellom a og b og at størrelsen $y(t_1)$ vil inntreffe mellom c og d vil være

$$P[a < x < b \text{ og } c < y < d] = \iint_{ca}^{db} f(x; t_0, y; t_1) dx dy \quad (6.7)$$

Simultan¹ sannsynlighet er også definert i det enda mer multidimensjonale (> 2) tilfelle. Det henvises til litteratur i sannsynlighetsregning.

Når vi ser på prosessens amplitude ved to gitte tidspunkt, eller to prosessers amplitude ved samme eller forskjellige tidspunkt, kan vi oppsummere følgende regler:

For to tilfeldige variable vil den betingede sannsynlighetstetthetsfunksjonen være gitt av

$$f(x|y) = \frac{f(x, y)}{f(y)} \quad (6.8)$$

Likeledes vil den betingede sannsynlighetsfordelingsfunksjonen være gitt av

$$F(x|y) = \frac{F(x, y)}{F(y)} \quad (6.9)$$

Vi ser altså at betingede sannsynligheter kan beregnes ut fra simultane sannsynlighetsmål og sannsynlighetsmål for y alene.

La oss vende tilbake til eksemplet i figur 6.5. Vi har målt realisasjonen av $x(t)$ opp til tiden t_0 . Hvis vi ønsker å uttale oss om sannsynligheten for $x(t_1), t_1 > t_0$ når vi kjenner verdien av $x(t_0)$, får vi med betingede sannsynligheter å gjøre, nemlig $f(x(t_1)|x(t_0))$. Fra (6.8) har vi

$$f(x(t_1)|x(t_0)) = \frac{f(x(t_1), x(t_0))}{f(x(t_0))} \quad (6.10)$$

1. Merk at begrepet "simultan" her ikke brukes i betydningen "på samme tidspunkt", men i betydningen "for flere variable".

Disse tetthetsfunksjonene er nettopp de som blir breiere og breiere etterhvert som t øker i figur 6.5.

Vi kan nå innføre begrepet **statistisk uavhengighet**:

To tilfeldige variable x og y er statistisk uavhengige hvis

$$f(x, y) = f(x)f(y) \quad (6.11)$$

Fra likning (6.8) ser vi at i dette tilfellet har vi

$$f(x|y) = \frac{f(x, y)}{f(y)} = \frac{f(x)f(y)}{f(y)} = f(x) \quad (6.12)$$

Relatert til eksemplet i figur 6.5, ville statistisk uavhengighet mellom $x(t_1), t_1 > t_0$ og $x(t_0)$ bety at tetthetsfunksjonen umiddelbart etter t_0 ville blitt lik den vi hadde i figur 6.3. Vi kan trekke den foreløpige konklusjon at dette ville betydd en hurtig varierende prosess, noe vi vil komme inn på seinere.

6.2.3 Forventningsverdi, kovarians og korrelasjon

Av grunner som vi skal komme tilbake til, kan amplitude-egenskapene til stokastiske prosesser ofte beskrives ut fra middelverdi og varians for et gitt tidspunkt. Disse midlere verdiene kan finnes enten fra rene teoretiske beregninger eller fra midlings-operasjoner utført på målte tidsserier.

Vi utelater nå tidsargumentet i $x(t)$, slik at vi bare skriver x . Det er da underforstått at vi betrakter prosessen på et bestemt sted på tidsaksen, og da blir x en tilfeldig variabel slik vi kjenner begrepet fra sannsynlighetsregningen. **Middelverdien** \bar{x} , eller **forventningsverdien**, defineres ved

$$\bar{x} = E(x) = \int_{-\infty}^{\infty} xf(x)dx \quad (6.13)$$

hvor x er den stokastiske prosessen ved et gitt tidspunkt (eller generelt en tilfeldig variabel), $p(x)$ er tetthetsfunksjonen for x og $E(\cdot)$ betegner **forventningsoperatoren**. Merk at $E(\cdot)$ bare er en stenografisk notasjon for integraloperasjonen (6.13). Vi kan skrive

$$E(\cdot) = \int_{-\infty}^{\infty} (\cdot) \cdot f(x)dx \quad (6.14)$$

Da blir forventningsverdien til en funksjon $g(x)$ av en tilfeldig variabel, gitt ved

$$E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x) dx \quad (6.15)$$

g og x kan gjerne være vektorer og utvidelsen til dette tilfellet er triviell, i det vi får

$$E(\underline{g}(\underline{x})) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \underline{g}(\underline{x}) f(\underline{x}) dx_1 dx_2 \dots dx_n , \quad (6.16)$$

hvor $f(\underline{x})$ nå er en simultan tetthetsfunksjon, $\dim(\underline{x})=n$. Merk at de to vektorene \underline{x} og $\underline{g}(\cdot)$ ikke trenger å ha samme dimensjon.

To viktige egenskaper ved forventningsoperasjonen vil bli hyppig benyttet:

1. Forventningsoperasjonen er lineær:

Gitt konstantene α og β , funksjonene $\underline{h}(x)$ og $\underline{g}(x)$, og den tilfeldige variable \underline{x} med (simultan) tetthetsfunksjon $f(\underline{x})$. Gitt en ny tilfeldig variabel \underline{z} , med $\underline{z} = \alpha \underline{h}(\underline{x}) + \beta \underline{g}(\underline{x})$. Hva blir $E(\underline{z})$?

Vi har

$$E(\underline{z}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [\alpha \underline{h}(\underline{x}) + \beta \underline{g}(\underline{x})] f(\underline{x}) dx_1 dx_2 \dots dx_n$$

Ved å bruke de kjente regler for integralet av en sum, og ved å sette konstantene utafor integralet, får vi til slutt

$$E(\underline{z}) = E[\alpha \underline{h}(\underline{x}) + \beta \underline{g}(\underline{x})] = \alpha E[\underline{h}(\underline{x})] + \beta E[\underline{g}(\underline{x})] \quad (6.17)$$

2. Integrasjon og forventningsoperasjon kan byttes om:

$$E\left[\int_{\beta_1}^{\beta_2} h(\tau) g(\underline{x}, \tau) d\tau\right] = \int_{\beta_1}^{\beta_2} h(\tau) E[g(\underline{x}, \tau)] d\tau \quad (6.18)$$

Beviset er trivielt: Forventningsoperasjonen representerer en eller flere integrasjoner. En ny integrasjon kan gjøres før eller etter de øvrige, fordi resultatet av et multippelt integral ikke avhenger av integrasjonsrekkefølgen.

Variansen til en tilfeldig skalar variabel x defineres som

$$\text{var}(x) = \sigma^2 = E\{(x - \bar{x})^2\} = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx \quad (6.19)$$

hvor $\bar{x} = E(x)$.

σ^2 er et mål for dispersjonen eller spredningen av x omkring sin midlere verdi \bar{x} .

Hvis vi i stedet har en vektor $\underline{x} = [x_1 \dots x_n]^T$ av tilfeldige variable, kan vi definere en multidimensjonal analogi til varians, nemlig **kovariansmatrisa** til \underline{x} :

$$X = E\{(\underline{x} - \bar{\underline{x}})(\underline{x} - \bar{\underline{x}})^T\} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\underline{x} - \bar{\underline{x}})(\underline{x} - \bar{\underline{x}})^T f(\underline{x}) dx_1 \dots dx_n \quad (6.20)$$

Kovariansmatrisa X er symmetrisk dvs. $X = X^T$ der elementet $x_{ij} = x_{ji}$ er gitt av

$$x_{ij} = E\{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\} \quad (6.21)$$

Diagonalleddene x_{ii} gir følgelig variansen til x_i .

Det er også vanlig å bruke betegnelsen $X = \text{cov}(\underline{x})$. Vi har

$$E\{(\underline{x} - \bar{\underline{x}})(\underline{x} - \bar{\underline{x}})^T\} = E\{\underline{x} \underline{x}^T - \bar{\underline{x}} \underline{x}^T - \underline{x} \bar{\underline{x}}^T + \bar{\underline{x}} \bar{\underline{x}}^T\} \quad (6.22)$$

Nå er $E(\bar{x}) = \bar{x}$, siden \bar{x} ikke er en tilfeldig variabel. Altså får vi at (6.22) blir

$$E(\underline{x} \underline{x}^T) - \bar{\underline{x}} \bar{\underline{x}}^T - \bar{\underline{x}} \bar{\underline{x}}^T + \bar{\underline{x}} \bar{\underline{x}}^T = E(\underline{x} \underline{x}^T) - \bar{\underline{x}} \bar{\underline{x}}^T \quad (6.23)$$

I det skalare tilfellet (en tilfeldig variabel) forenkles dette til

$$E[(x - \bar{x})^2] = \sigma^2 = \text{var}(x) = E(x^2) - [E(x)]^2 \quad (6.24)$$

Standardavviket σ_i til x_i er definert som

$$\sigma_i = \sqrt{x_{ii}} \quad (6.25)$$

Som tidligere nevnt indikerer x_{ii} (eller $\sigma_i = \sqrt{x_{ii}}$) spredningen omkring \bar{x}_i .

De ikke-diagonale leddene i X angir sammenhengen eller graden av samvariasjon mellom variablene x_i og x_j .

Den normaliserte kovariansen mellom x_i og x_j defineres som

$$\rho_{ij} = \frac{x_{ij}}{\sigma_i \sigma_j} \quad (6.26)$$

Den kalles **korrelasjonskoeffisienten** mellom x_i og x_j . Korrelasjonskoeffisienten vil alltid ligge i intervallet $[-1, 1]$.

6.2.4 Et eksempel på korrelasjon

La oss betrakte de to avleddede størrelsene

$$x_1 = z$$

$$x_2 = az + w$$

hvor z og w er stokastiske variable og forutsatt statistisk uavhengige. Videre forutsettes $E(w) = 0$, $E(z) = 0$, $E(z^2) = \sigma_z^2$, $E(w^2) = \sigma_w^2$.

Da blir kovariansmatrisa til $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$:

$$X = E \begin{bmatrix} z^2 & az^2 + zw \\ az^2 + zw & a^2z^2 + 2azw + w^2 \end{bmatrix} = \begin{bmatrix} E(z^2) & E(az^2 + zw) \\ E(az^2 + zw) & E(a^2z^2 + 2azw + w^2) \end{bmatrix} \quad (6.27)$$

$$X = \begin{bmatrix} \sigma_z^2 & a\sigma_z^2 \\ a\sigma_z^2 & a^2\sigma_z^2 + \sigma_w^2 \end{bmatrix} \quad (6.28)$$

Korrelasjonskoeffisienten (se (6.26) mellom x_1 og x_2 blir:

$$\rho_{12} = \frac{a\sigma_z^2}{(a^2\sigma_z^2 + \sigma_w^2)^{1/2}\sigma_z} = \frac{a\sigma_z}{\sqrt{a^2\sigma_z^2 + \sigma_w^2}} \quad (6.29)$$

Setter vi $a = 0$, får vi $\rho_{12} = 0$. Dette rimer med at da er $x_1 = z$ og $x_2 = w$, og x_1 og x_2 blir følgelig statistisk uavhengige.

Lar vi $a \rightarrow \infty$ eller $\sigma_w^2 \rightarrow 0$, får vi $\rho_{12} = 1$. Dette rimer med at da blir $x_2 = \text{konst.} \cdot x_1$, dvs. maksimal samvariasjon.

6.2.5 Statistisk uavhengighet og korrelasjon

Vi har tidligere definert statistisk uavhengighet mellom x_i og x_j . Da er den simultane tettettsfunksjon $f(x_i, x_j) = f(x_i)f(x_j)$, jfr. (6.11).

For $i \neq j$ får vi da

$$\begin{aligned} x_{ij} &= E\{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \bar{x}_i)(x_j - \bar{x}_j) f(x_i, x_j) dx_i dx_j \\ &= \int_{-\infty}^{\infty} (x_i - \bar{x}_i) f(x_i) dx_i \int_{-\infty}^{\infty} (x_j - \bar{x}_j) f(x_j) dx_j = 0 \end{aligned} \quad (6.30)$$

Vi ser av dette at **statistisk uavhengige variable er ukorrelerte**, dvs. at $x_{ij} = 0$ når $i \neq j$. Har vi derfor en vektor av statistisk uavhengige variable, vil kovariansmatrisa bli diagonal.

Nå er ikke ukorrelerte variable alltid statistisk uavhengige, som vi skal se av et eksempel:

Gitt to statistisk uavhengige tilfeldige variable x_1 og x_2 . Begge er **uniformt fordelte** i intervallet $[-1, 1]$, dvs. $f(x_1) = f(x_2) = \frac{1}{2}$, 0 ellers. Gitt en ny tilfeldig variabel $z = x_1 x_2$. Det er åpenbart at z og x_1 ikke er statistisk uavhengige siden x_1 inngår i z . Men det skal nå vises at x_1 og z er ukorrelerte.

Søker først middelverdien av z :

$$E(z) = \bar{z} = \int_{-1}^{1} \int_{-1}^{1} (x_1 x_2) f(x_1, x_2) dx_1 dx_2 \quad (6.31)$$

Siden vi har statistisk uavhengighet er $f(x_1, x_2) = f(x_1)f(x_2)$. Setter inn for $f(x_1)$ og $f(x_2)$ og får

$$E(z) = \int_{-1}^{1} x_1 \frac{1}{2} dx_1 \int_{-1}^{1} x_2 \frac{1}{2} dx_2 = \bar{x}_1 \bar{x}_2 = 0 \quad (6.32)$$

Vi har da

$$E[(z - \bar{z})(x_1 - \bar{x}_1)] = E[zx_1] = E[x_1^2 x_2] \quad (6.33)$$

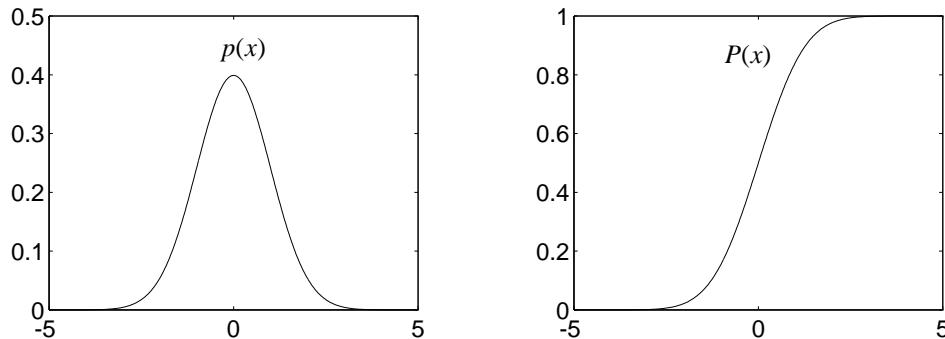
Nå er x_1^2 og x_2 statistisk uavhengige. Med samme resonnement som over, er da $E(x_1^2 x_2) = \bar{x}_1^2 \cdot \bar{x}_2 = \bar{x}_1^2 \cdot 0 = 0$. Altså er $E[(z - \bar{z})(x_1 - \bar{x}_1)] = 0$ dvs. z og x_1 er ukorrelerte på tross av at de ikke er statistisk uavhengige.

6.2.6 Normalfordelingen

Når tilfeldige variable karakteriseres ved sine tetthetsfunksjoner $f(x)$, er vi nødt til å skaffe til veie disse funksjonene. Dette kan skje ved teoretiske betrakninger, ved eksperimenter eller ved en kombinasjon. Dette er i mange tilfeller vanskelig og i mange tilfeller må man greie seg med å *anta* en tetthetsfunksjon.

Imidlertid kommer naturen oss i en viss grad til hjelp i mange tilfeller, idet svært mange stokastiske fenomener i naturen er tilnærmet **normalfordelte** (= **Gaussfordelte**). Den skalare (dvs. ikke-simultane) normalfordelingen, figur 6.6, er gitt av

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2}, \quad \sigma \text{ er standardavvik og } \bar{x} \text{ er middelverdien} \quad (6.34)$$



Figur 6.6 Normalfordeling for $\bar{x} = 0$ og $\sigma = 1.0$

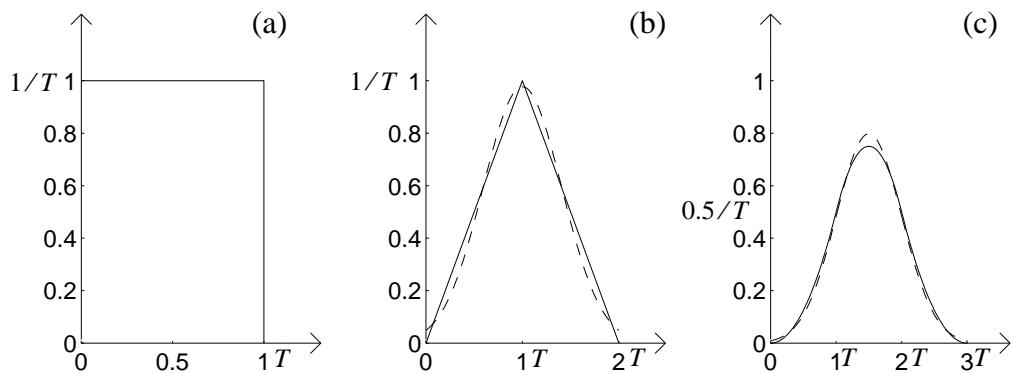
Normalfordelingen er ikke bare en matematisk abstraksjon. Den forekommer hyppig i naturen og fysikken. Dette kan forklares via det såkalte **sentralgrenseteoremet**. Det er et resultat fra statistikken som sier at **hvis en tilfeldig variabel er en sum av N statistisk uavhengige, tilfeldige variable, hver med en vilkårlig tetthetsfunksjon, så vil den resulterende variabel være normalfordelt når $N \rightarrow \infty$.**

Et annet resultat som er beslektet med sentralgrenseteoremet, er at **en veid sum av normalfordelte størrelser er også normalfordelt. Dette betyr at normalfordelte forstyrrelser som eksisterer et lineært dynamisk system gir normalfordelte responser.**

Sentralgrenseteoremet skal illustreres med to eksempler:

Eksempel 1:

La x_i være uniformt ("=rektagulaert") fordelte, tilfeldige og uavhengige variable på intervallet $[0, T]$ som vist i Figur 6.7a. Vi betrakter så $x = x_1 + x_2$. Tetthetsfunksjonen $f(x)$ er vist i Figur 6.7b som en heltrukket linje. Vi ser at den er en trekantformet funksjon. Den har $\bar{x} = T$ og $\text{var}(x) = \frac{1}{6}T^2$.

Figur 6.7 Tettethetsfunksjoner for x_1 , $(x_1 + x_2)$ og $(x_1 + x_2 + x_3)$

En normalfordeling med samme middelverdi og varians er gitt av

$$f_N(x) = \frac{1}{T} \sqrt{\frac{3}{\pi}} e^{-3(x-T)^2/T^2} \quad (6.35)$$

Denne normalfordelingen er vist i Figur 6.7b med stiplet linje. Vi ser at vi allerede i dette tilfellet har en viss likhet mellom $f(x)$ og $f_N(x)$.

Hvis nå $x = x_1 + x_2 + x_3$ får vi fordelingen $f(x)$ som er heltrukket i figur 6.7c med $\bar{x} = \frac{3T}{2}$ og $\text{var}(x) = (T^2/4)$. $f(x)$ består i dette tilfellet av tre parabolske deler. En normalfordeling $f_N(x)$ med samme middelverdi og varians blir nå

$$f_N(x) = \frac{1}{T} \sqrt{\frac{2}{\pi}} e^{-2(x-1.5T)^2/T^2} \quad (6.36)$$

$f_N(x)$ er gitt ved den stiplete linjen i figur 6.7c. Vi ser at forskjellen på $f(x)$ og $f_N(x)$ er ganske liten.

Eksempel 2:

Følgende prosess betraktes: En sekvens av tilfeldige variable midles gjennom et lavpassfilter. Inngangssignalet $x(t)$ antar verdiene +1 eller -1 og kan kun skifte verdi ved $t = 0, 1, 2, 3, \dots$. Sannsynligheten for å anta hver av verdiene er 0.5, slik at

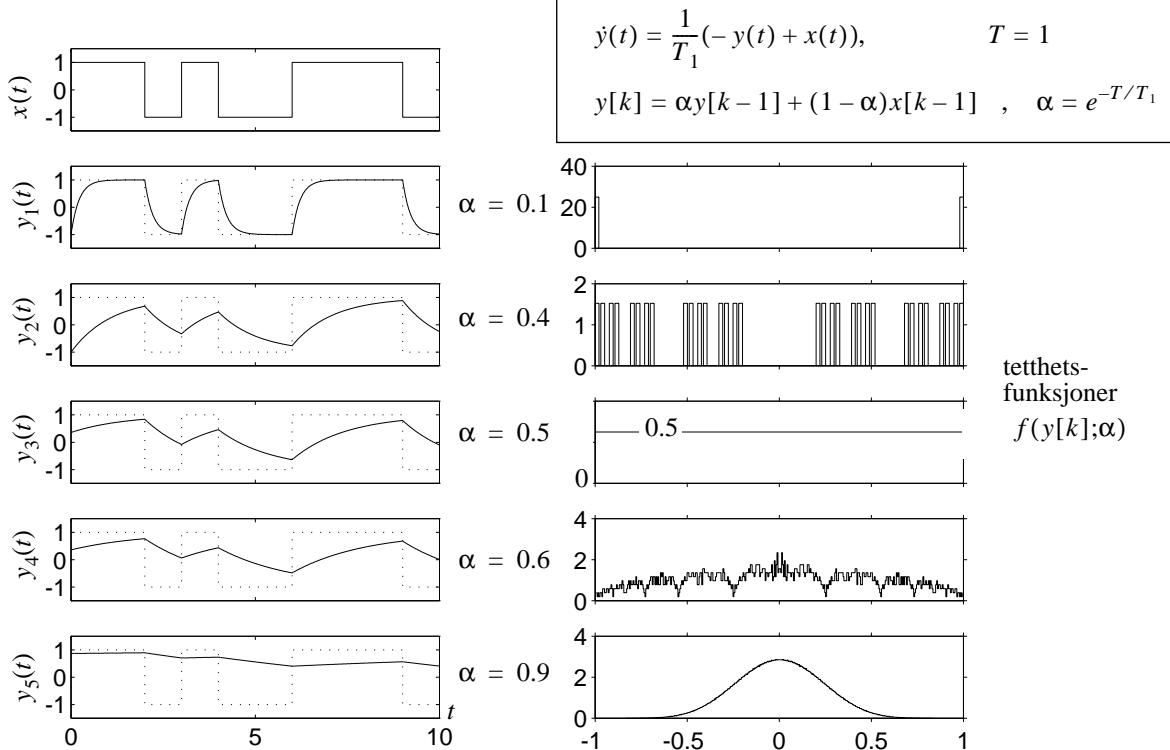
$$f(x) = \frac{1}{2}[\delta(-1) + \delta(1)] \quad (6.37)$$

Dette signalet sendes gjennom ulike lavpassfiltre (RC-nettverk) med forskjellig båndbredde. Figur 6.8 viser typiske forløp for inngangssignalet $x(t)$, for utgangssignalet $y(t)$ samt for tettethetsfunksjonene $f(y[k])$ for ulike båndbreddeverdier i lavpassfilteret. ($y[k] = y(t)$ for $t = 0, 1, 2, 3, \dots$). Vi ser at $f(y[k])$ går mot noe som likner en normalfordeling når båndbredden avtar. Dette kan forklares ut fra sentralgrenseteoremet, fordi filteret foretar en veiet summasjon av tidligere verdier av $x(t)$. Anta at filteret har impulsrespons $h(t)$:

Foldingsintegralet gir

$$y(t) = \int_{-\infty}^t h(t-\alpha)x(\alpha)d\alpha \quad (6.38)$$

$y(t)$ kan betraktes som en sum med uendelig mange ledd, og hvor hver inngangsverdi $x(\alpha)$ veies med faktoren $h(t-\alpha)d\alpha$. Gamle ledd veies kraftigere inn i summen jo lavere båndbredde (dvs. større α) filteret har, og sentralgrenseteoremetets betingelse om uendelig mange ledd nærmer seg derfor oppfyllelse¹.



Figur 6.8

6.2.7 Simulant normalfordelte variable

Til slutt i dette avsnittet skal vi si litt om **simulant normalfordelte variable**, dvs. vi har ikke lenger en skalar tilfeldig variabel x , men en vektor \underline{x} med normalfordelte komponenter med kovariansmatrise X . Vi har da

$$f(\underline{x}) = \frac{1}{[(2\pi)^n \det(X)]^{1/2}} \exp\left(-\frac{1}{2}(\underline{x} - \bar{\underline{x}})^T X^{-1} (\underline{x} - \bar{\underline{x}})\right), \text{ hvor } n = \dim(\underline{x}). \quad (6.39)$$

1. Kuriositet: For $\alpha < 0.5$ er tilstandsrommet for $y[k]$ ikke lenger kontinuerlig mellom -1 og 1; det finnes tilstander som aldri vil inntrefte. I tillegg blir tilstandsrommet *fraktalt*, fordi det for et vilkårlig lite intervall rundt en vilkårlig verdi av $y[k]$ *innafor* tilstandsrommet alltid finnes tilstander *utafor*. For $\alpha = 0.33$ vil tilstandsrommet for $y[k]$ være den spesielle fraktale mengde som kalles *Kantor-støv*.

Dersom elementene i \underline{x} er ukorrelerte, dvs. at X er diagonal, finner vi

$$f(\underline{x}) = \prod_{i=1}^n \frac{1}{(2\pi x_{ii})^{1/2}} \exp\left\{-\frac{1}{2}(x_i - \bar{x}_i)^2/x_{ii}\right\} = \prod_{i=1}^n f(x_i) \quad (6.40)$$

der

$$f(x_i) = \frac{1}{(2\pi x_{ii})^{1/2}} \exp\left\{-\frac{1}{2}(x_i - \bar{x}_i)^2/x_{ii}\right\} \quad (6.41)$$

Hvis vi sammenholder dette med definisjonen på statistisk uavhengighet, likn. (6.11), ser vi at -

for normalfordelte variable vil ukorrelerte variable også være statistisk uavhengige.

Dette er nyttig å vite fordi det som oftest er lettere å påvise ukorrelerthet enn statistisk uavhengighet. Korrelasjon kan beregnes på grunnlag av målte data, statistisk uavhengighet kan man bare *anta* ut fra innsikt i de prosesser som studeres. (Et eksempel på det siste er antagelsen om at neste utfall av å kaste mynt og kron er statistisk uavhengig av hva resultatet ble i forrige kast. Dette virker helt oppagt, men det er uansett en antagelse og kan ikke bevises).

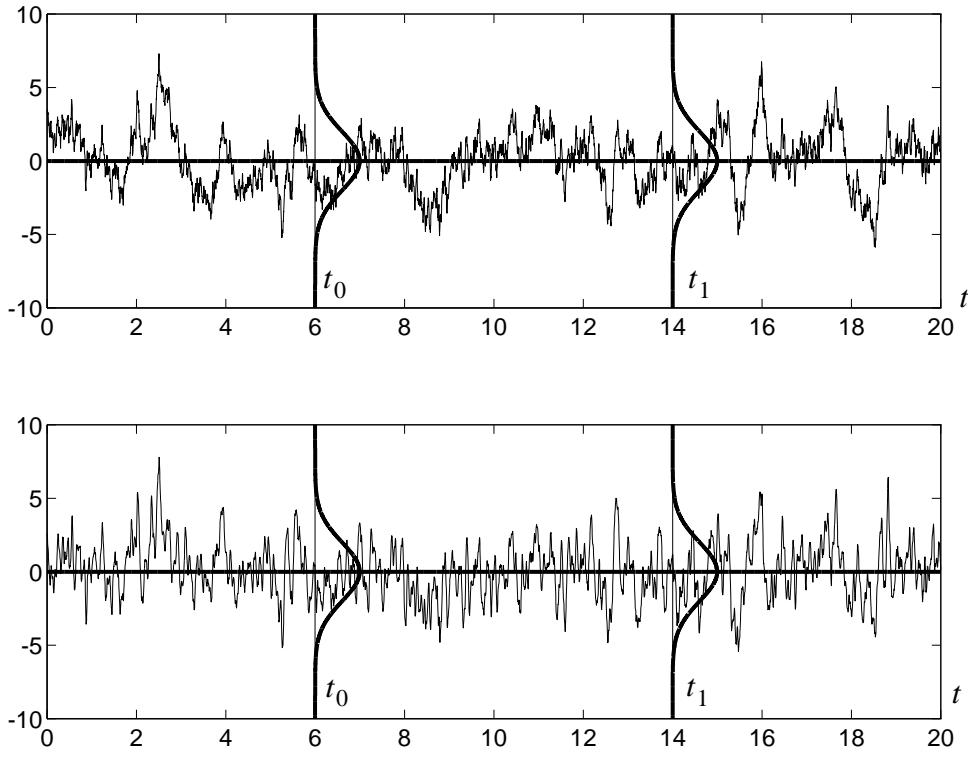
La oss aller sist understreke at i tillegg til å opptre ofte er normalfordelingen enkel å karakterisere, idet den er fullstendig gitt av middelverdi \bar{x} og kovariansen $X = \text{cov}(\underline{x})$.

Dette, sammen med de tidligere nevnte egenskapene gitt av sentralgrenseteoremet, setter normalfordelingen i en særstilling.

6.3 Stokastiske prosesser beskrevet i tidsplanet

Vi har hittil omtalt stokastiske prosesser i relativt intuitive vendinger. Vi har sagt at en stokastisk prosess er den totale mengde mulige tidsserier eller måleserier fra et gitt eksperiment eller fra et naturlig forløp av en fysisk prosess. Hver av disse mulige tidsseriene er en realisasjon av den stokastiske prosessen. Hvis vi betrakter den stokastiske prosessen ved et gitt tidspunkt, vil $x(t)$ være en tilfeldig (stokastisk) variabel. Denne tilfeldige variable kan beskrives eller modelleres ved hjelp av tetthetsfunksjoner, noe vi gikk inn på og repeterte i avsnitt 6.2. I dette tilfellet hvor en fordelingsfunksjon $P(x)$ eller en tetthetsfunksjon $p(x)$ refererer til en stokastisk prosess, benytter vi som nevnt ofte betegnelsen amplitudefordelingsfunksjon eller amplitudetetthetsfunksjon.

Vi kan intuitivt innse at tetthetsfunksjonene ved forskjellige tidspunkt ikke er tilstrekkelig for beskrivelse av en stokastisk prosess. Aleine gir $p(x)$ bare informasjon om hvordan verdien av en stokastisk prosess er fordelt i øyeblikket og ikke noe om hvordan tidsforløpet av en realisasjon ser ut. Figur 6.9 viser f.eks. to realisasjoner fra hver sin stokastiske prosess. Amplitudefordelingen er den samme i de to tilfellene, men vi ser at de to prosessene har ulike egenskaper med hensyn til variasjon med tida. Hvis vi ser nærmere på figur 6.9, ser vi at forskjellen har noe med variasjonshurtigheten å gjøre



Figur 6.9

Den øverste prosessen ser ut til å ha et mer predikterbart forløp enn den andre, som har et sterkere innslag av hurtige variasjoner.

For å skille mellom de to prosessene måtte vi i tillegg til tetthetsfunksjonene for gitte tidspunkter, også måtte formulere *simultane* tetthetsfunksjoner for

$x(t_1), x(t_2), x(t_3), x(t_4), \dots, x(t_n)$ for vilkårlige t_i og n :

$$f(x(t_1), x(t_2), x(t_3), x(t_4), \dots, x(t_n)) \quad (6.42)$$

Men dette blir altfor komplisert å arbeide med. Vi skal derfor nøye oss med å beskrive stokastiske prosesser ved hjelp av **korrelasjonsfunksjoner** i tidsplanet, og ved hjelp av **effektspektra** i frekvensplanet. Vi vil få se at disse to beskrivelsene er fullstendig ekvivalente, og at **de er Fouriertransform-par**. Vi vil få se at vi kan bruke de verktøy vi allerede har utviklet i forbindelse med kontinuerlige og diskrete systemer og signaler.

6.3.1 Beskrivelse ved middelverdi og korrelasjonsfunksjon

Korrelasjonsfunksjonene for $x(t)$ er definert ved $\phi_{xx}(t, \tau)$

$$\phi_{xx}(t, \tau) = E\{x(t)x(t + \tau)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \cdot f[x_1, x_2] dx_1 dx_2 \quad (6.43)$$

der vi har innført $x(t) = x_1$ og $x(t + \tau) = x_2$. Korrelasjonsfunksjonen måler som vi ser sammenhengen mellom den variable x ved tiden t og x ved tiden $t + \tau$. φ_{xx} sier derfor noe om variasjonsegenskapene til x langs tidsaksen. Hvis det er liten statistisk sammenheng mellom $x(t)$ og $x(t + \tau)$, vil $\varphi_{xx}(t, \tau)$ være liten, og omvendt hvis sammenhengen er stor, vil $\varphi_{xx}(t, \tau)$ være stor.

Denne korrelasjonsfunksjonen måler korrelasjonen av en tidsserie med seg selv langs tidsaksen og kalles derfor også **autokorrelasjonsfunksjonen**.

For å måle sammenhengen eller korrelasjonen mellom to ulike stokastiske prosesser, defineres **krysskorrelasjonsfunksjonen** $\varphi_{xy}(t, \tau)$

$$\varphi_{xy}(t, \tau) = E\{x(t)y(t + \tau)\} \quad (6.44)$$

hvor x og y er de to stokastiske prosessene.

Vi skal nå definere to begreper som videre vil kunne forenkle analysen og beskrivelsen av stokastiske prosesser i mange tilfeller.

STASJONÆR PROSESS

*En stokastisk prosess er **stasjonær** hvis den simultane tetthetsfunksjonen (6.42) er uavhengig av tidsparameteren. Dvs. at t_i kan erstattes med $t_i + T$ for vilkårlige T i (6.42).*

Spesielt vil da $\varphi_{xx}(t, \tau) = \varphi_{xx}(\tau)$ i (6.43).

ERGODISK PROSESS

En stasjonær stokastisk prosess er **ergodisk** hvis vi for en vilkårlig funksjon $g[x(t)]$ av en stokastisk prosess $x(t)$ har at

$$E\{g[x(t)]\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T g[x(t)] dt \quad (6.45)$$

En prosess er altså ergodisk hvis ensemble-middelverdien av $g(x)$ er lik tidsmiddelverdien av $g(x)$.

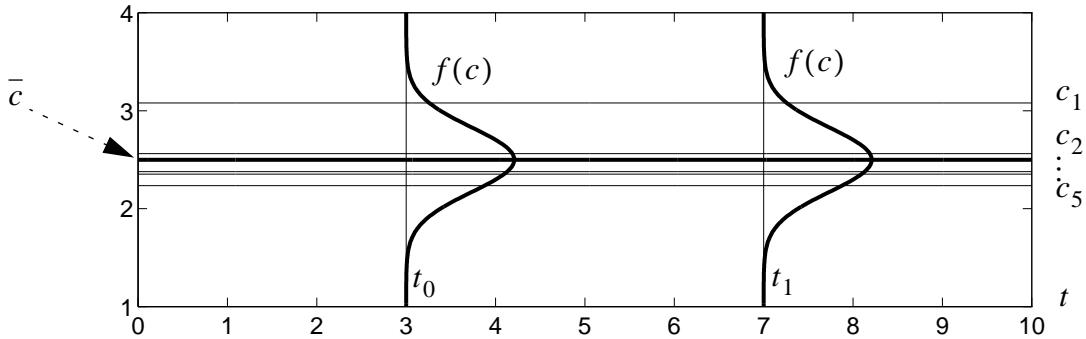
Hvis en prosess er ergodisk, kan vi finne ensemble-forventningsverdier ved i stedet å beregne tids-forventningsverdier. Dermed kan vi greie oss med en eneste lang tidsserie. Ergodisk prosess betyr at all statistisk informasjon om prosessen ligger i en enkelt realisasjon. Forutsetningen om ergodisitet er svært vanlig, og dessuten helt nødvendig for å beregne korrelasjonsfunksjoner fra målte data.

Et prosessreguleringssystem vil ofte ha tilnærmet ergodiske signaler når systemet har kommet inn i en stasjonær driftstilstand og transiente har dødd ut.

Imidlertid: En stasjonær prosess er ikke nødvendigvis ergodisk, som vi skal se i et enkelt eksempel: Prosessen er karakterisert ved at dens realisasjoner er konstante funksjoner,

$$x(t) = c, \quad (6.46)$$

hvor c er en tilfeldig variabel med sannsynlighetstetthetsfunksjon $f(c)$, dvs. c varierer fra realisasjon til realisasjon. Figur 6.10 viser noen realisasjoner, og middelverdien \bar{c} :



Figur 6.10

Denne prosessen er ikke ergodisk fordi ensemble-middelverdien er ulik tidsmiddelverdien for en bestemt realisasjon. Vi kan ikke regne ut ensemble-middelverdien \bar{c} ved å bruke tidsmiddelverdien.

Men prosessen er stasjonær, siden sannsynlighetstetthetsfunksjonen (6.41) i dette trivielle tilfellet er uavhengig av hvor på tidsaksen vi befinner oss.

Vi oppsummerer: *En ergodisk prosess er alltid stasjonær. For at en stasjonær prosess skal være ergodisk må i tillegg enhver realisasjon eller tidsserie være representativ for hele ensemblet.*

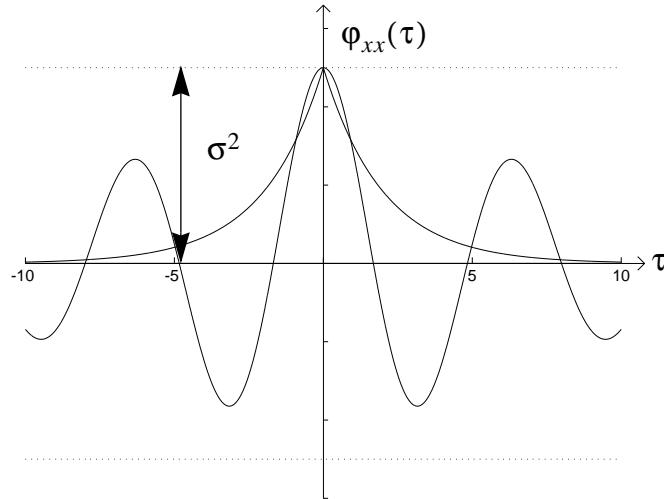
Ved ergodiske prosesser kan vi regne ut autokorrelasjonsfunksjonen v.h.a. en realisasjon av $x(t)$:

$$\Phi_{xx}(\tau) = E[x(t)x(t + \tau)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t + \tau) dt \quad (6.47)$$

og krysskorrelasjonsfunksjonen fra en realisasjon av prosessen $x(t)$ og en realisasjon av den andre prosessen $y(t)$.

$$\Phi_{xy}(\tau) = E[x(t)y(t + \tau)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)y(t + \tau) dt \quad (6.48)$$

Vi skal se på en del egenskaper til korrelasjonsfunksjonen. Vi forutsetter at de prosesser vi studerer er stasjonære, og inntil videre også at de har middelverdi = 0. Figur 6.11 viser to nokså vanlige forløp for autokorrelasjonsfunksjoner:

Figur 6.11 , prosessene har $\bar{x} = 0$

Autokorrelasjonsfunksjonen for en stasjonær prosess vil alltid være symmetrisk, dvs.

$$\varphi_{xx}(\tau) = \varphi_{xx}(-\tau) \quad (6.49)$$

For en ergodisk prosess sees dette enkelt ved å sette $t := t - \tau$ i (6.47).

En viktig egenskap utledes ved å betrakte

$$E[\{x(t + \tau) \pm x(t)\}^2] = 2[\varphi_{xx}(0) \pm \varphi_{xx}(\tau)] \geq 0 \quad (6.50)$$

som gir

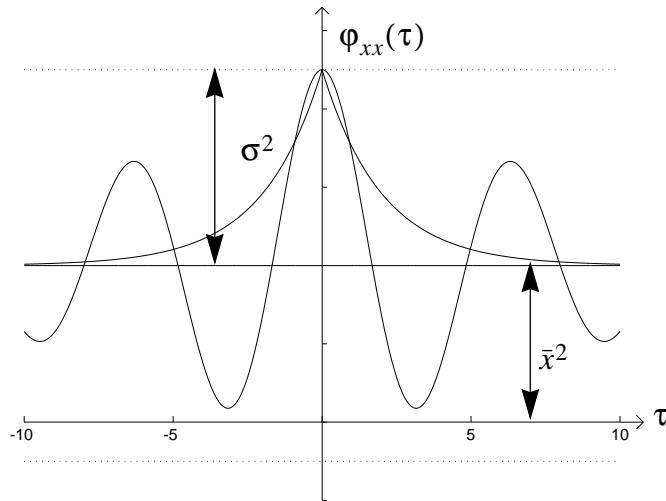
$$\varphi_{xx}(0) \geq |\varphi_{xx}(\tau)| \quad (6.51)$$

Autokorrelasjonsfunksjonens absoluttverdi er altså alltid mindre enn verdien for $\tau = 0$. For den oscillatoriske autokorrelasjonsfunksjonen i figur 6.11 er dette markert ved at funksjonen ikke kan gå mer negativ enn den nedre stiplede linjen, og mer positiv enn den øvre, som er bestemt av verdien til $\varphi_{xx}(0)$. Fysisk kan dette forklares ved at samvariasjonen mellom signalet $x(t)$ og $x(t + \tau)$ må avta med økende tallverdi av τ , fordi usikkerheten øker med tidsforskjellen.

Anta nå at en prosess har middelverdi $\bar{x} \neq 0$. Da gir (6.24)

$$\varphi_{xx}(0) = E[x(t + 0)x(t)] = E(x^2) = [E(x)]^2 + var(x) = \bar{x}^2 + \sigma^2 \quad (6.52)$$

Se figur 6.12. Hvis $x(t)$ er et fysisk signal, representerer $\varphi_{xx}(0)$ **effekten** i signalet. Dette kommer vi tilbake til.



Figur 6.12 , prosessene har $\bar{x} \neq 0$

For *krysskorrelasjonsfunksjonen* mellom de stokastiske prosessene x og y har vi

$$\varphi_{xy}(\tau) = \varphi_{yx}(-\tau) \quad (6.53)$$

Dette følger fra (6.48) ved å sette $t := t - \tau$. Dessuten vil $\varphi_{xy}(\tau) \equiv 0$ når x og y er ukorrelerte.

En egenskap ved krysskorrelasjonsfunksjonen som minner om (6.51) er

$$\varphi_{xy}^2(\tau) \leq \varphi_{xx}(0)\varphi_{yy}(0) \quad (6.54)$$

Vi har hittil kun betraktet stokastiske prosesser som er kontinuerlige i tid. Identiske forhold gjelder imidlertid for **diskrete stokastiske prosesser** $x[k]$ og $y[k]$. Likningene (6.47) og (6.48) blir da

$$\varphi_{xx}[n] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x[k]x[k+n] \quad (6.55)$$

$$\varphi_{xy}[n] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x[k]y[k+n] \quad (6.56)$$

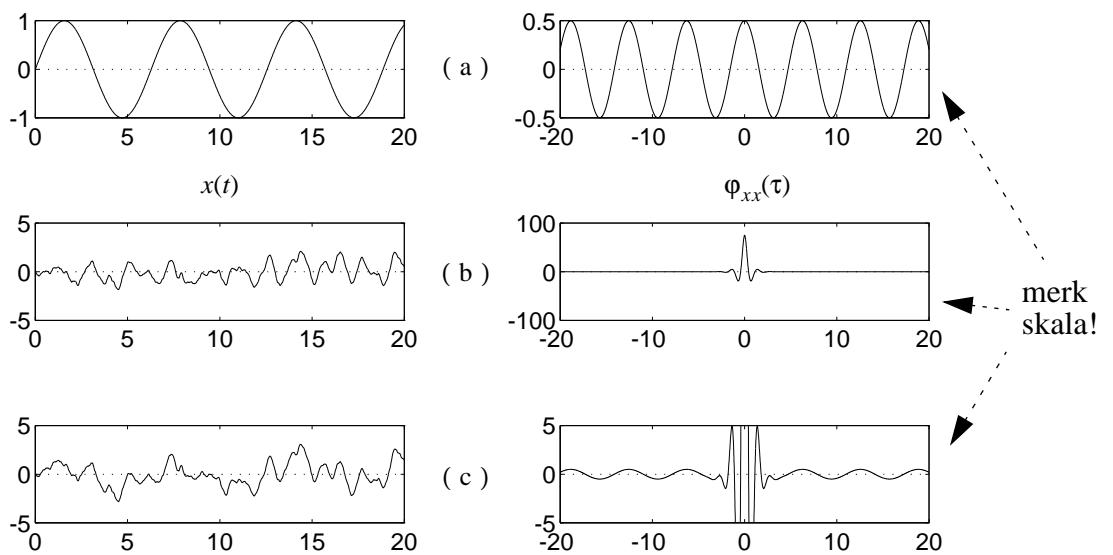
6.3.2 Signaler som er en sum av stokastiske og deterministiske komponenter

Vi vil først betrakte autokorrelasjonsfunksjonen for et deterministisk sinussignal. Det er da ikke riktig å snakke om autokorrelasjon i statistisk forstand, men vi vil anvende formelen (6.47) på dette signalet. (6.47) uttrykker nå ikke forventningsverdi, men det midlere produkt av signalet med seg selv τ tidsenheter senere. Gitt signalet $x(t) = a \sin(\omega t + \theta)$. (6.47) gir da

$$\begin{aligned}\varphi_{xx}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T a^2 \sin(\omega t + \theta) \sin(\omega t + \theta + \omega\tau) dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{1}{2} a^2 [(1 - \cos 2(\omega t + \theta)) \cos \omega\tau + \sin 2(\omega t + \theta) \sin \omega\tau] dt = \frac{1}{2} a^2 \cos \omega\tau\end{aligned}\quad (6.57)$$

Vi ser at $\varphi_{xx}(\tau)$ er uavhengig av den opprinnelige signalfasen, θ (hvorfor?). Generelt gjelder at en mengde signaler med samme form, men som er tidsforskjøvet i forhold til hverandre, kan gi samme autokorrelasjonsfunksjon, i det all faseinformasjon forsvinner. x og φ_{xx} er vist i figur 6.13 (a).

Anta nå at vi har et hurtig varierende stokastisk signal som vist i figur 6.13 (b). Autokorrelasjonsfunksjonen for dette signalet er vist i samme figuren, og vi ser at $\varphi_{xx}(\tau)$ går raskt mot null når τ øker. Dette er i motsetning til hva som er tilfelle i pkt. (a). Hurtigheten hvormed autokorrelasjonsfunksjonen går mot null er derfor et mål på predikterbarheten av den stokastiske prosessen. Hvis φ_{xx} ikke går mot null, må den stasjonære prosessen inneholde en deterministisk, periodisk komponent.



Figur 6.13 Nøen tidsforløp og tilhørende autokorrelasjonsfunksjoner

Hvis vi nå adderer støy av samme type som vist i (b) til sinussignalet vist i (a), vil autokorrelasjonsfunksjonen til summen være lik summen av autokorrelasjonsfunksjonen hvis signalene er ukorrelerte. En slik tidsserie og dens autokorrelasjon er vist i figur 6.13 (c).

Gjennom dette eksemplet ser vi hvordan begrepet autokorrelasjon via formel (6.47) **både** kan anvendes på en deterministisk, periodisk prosess og en stokastisk prosess. Vi kan detektere både periodiske deterministiske komponenter og stokastiske komponenter i et signal.

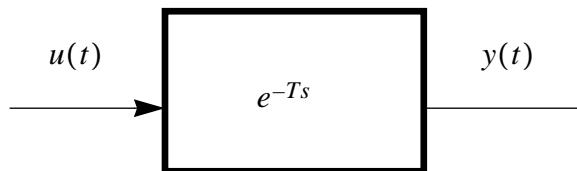
Og: Den stokastiske delen av prosessen representeres ikke lenger av noe så diffus som et “ensemble av realisasjoner”, men av en entydig autokorrelasjonsfunksjon.

6.3.3 Bruk av autokorrelasjonsfunksjoner og krysskorrelasjonsfunksjoner

Et av hovedformålene for måling av autokorrelasjoner i reguleringssystemer og i andre sammenhenger, er å måle graden av predikterbarhet, dvs. hvor stor innflytelsen av nåværende verdier er på fremtidige verdier. Som vi har sett i siste avsnitt vil autokorrelasjonsfunksjonen til et sinussignal aldri dø ut for store τ . Derfor vil måling av autokorrelasjonsfunksjonen også kunne klarlegge sinussignaler eller andre deterministiske signaler som er maskert i støysignaler. Andre og mer indirekte anvendelser skal vi komme inn på senere.

Krysskorrelasjonsfunksjoner benyttes for å fastslå en eventuell sammenheng mellom ulike signaler:

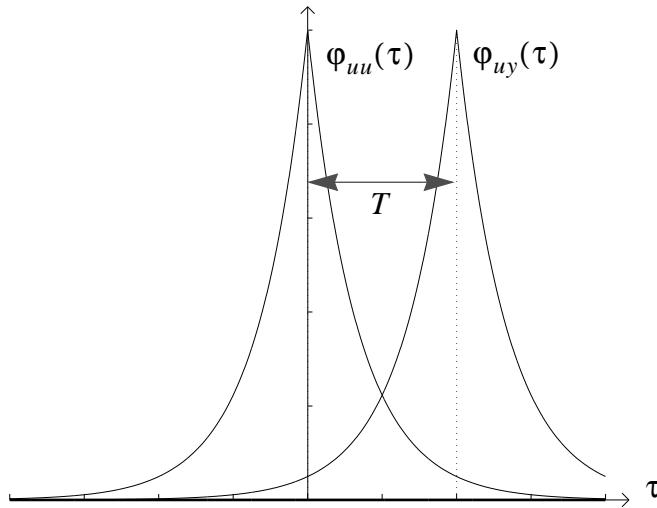
Krysskorrelasjonsberegninger kan bl.a. benyttes for **estimering (= anslag)** av tidsforsinkelse i en prosess. La oss anta gitt en prosess som en ren transportforsinkelse. $u(t)$ og $y(t)$ har null middelverdi.



Vi har da at $y(t) = u(t - T)$. Dermed er

$$\Phi_{uy}(\tau) = E[u(t)y(t + \tau)] = E[u(t)u(t + \tau - T)] = \Phi_{uu}(\tau - T) \quad (6.58)$$

Vi ser av figur 6.14 at $\Phi_{uy}(\tau)$ er lik $\Phi_{uu}(\tau)$ forskjøvet T sekunder langs τ -aksen. Dette kan vi benytte i praksis for måling av dominerende tidsforsinkelse idet krysskorrelasjonens topp gir denne.



Figur 6.14 ϕ_{uu} og ϕ_{uy} når $y(t) = u(t - T)$.

6.3.4 Autokorrelasjon direkte fra simultan tetthetsfunksjon: et eksempel

Ved å bruke forutsetningen om ergodisitet, har vi fått en enkle formler for å finne auto- og krysskorrelasjon, (6.47) og (6.48). Vi har tidligere påpekt at *egentlig* er det simultane tetthetsfunksjoner av typen (6.42) som gir den fullverdige beskrivelsen av en stokastisk prosess, men samtidig er det svært komplisert å bruke disse. Vi skal likevel gi et eksempel på at det går an å utlede autokorrelasjon direkte fra en simultan tetthetsfunksjon

La oss betrakte en situasjon hvor den stokastiske prosessen $x(t)$ er et *binært signal* med amplitude $+c$ eller $-c$. Antall fortegnskift i intervallet $[t, t + \tau]$ skjer tilfeldig og ved uavhengige tidspunkter med en midlere hyppighet λ ganger pr. sekund. Et slikt signal kalles ofte et **telegrafsignal**, se figur 6.15.

Det kan vises at sannsynlighetstettheten¹ for antall fortegnskift n i intervallet er **Poissonfordelt**

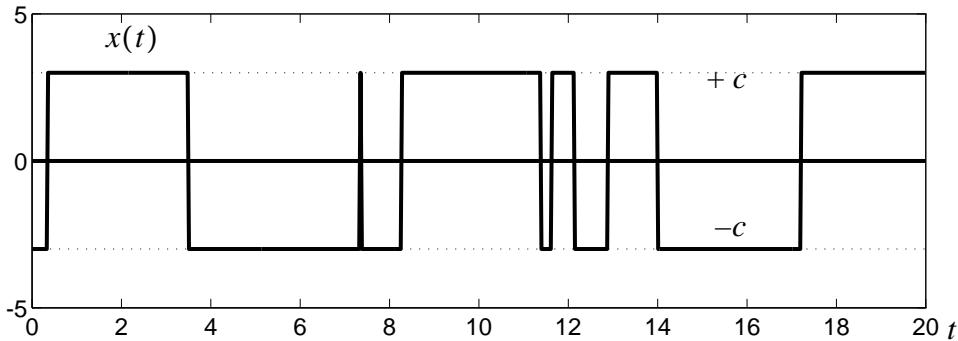
$$p[n] = \frac{\lambda^n |\tau|^n}{n!} e^{-\lambda|\tau|} \quad (6.59)$$

Vi skal finne autokorrelasjonsfunksjonen for $x(t)$. Vi har

$$\phi_{xx}(\tau) = E\{x(t)x(t + \tau)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \cdot f[x_1, x_2] dx_1 dx_2 \quad (6.60)$$

der vi har innført $x(t) = x_1$ og $x(t + \tau) = x_2$.

1. Vi skriver $p[n]$ med hakeparentes for markere at $p[n]$ er en *diskret* tetthetsfunksjon.



Figur 6.15

Vi har tidligere beregnet autokorrelasjonsfunksjon ved tidsmidling. Denne gangen skal vi altså gå veien om den simultane sannsynlighetstetthetsfunksjonen for $x(t)$ og $x(t + \tau)$. Vi må derfor finne $f(x(t), x(t + \tau)) = f(x_1, x_2)$.

Nå kan x_1 og x_2 bare anta to verdier, c eller $-c$, og likning (6.60) kan derfor skrives

$$\begin{aligned} \Phi_{xx}(\tau) &= ccP(x_1 = c, x_2 = c) + c(-c)P(x_1 = c, x_2 = -c) \\ &\quad + (-c)cP(x_1 = -c, x_2 = c) + (-c)(-c)P(x_1 = -c, x_2 = -c) \end{aligned} \quad (6.61)$$

Telegraftsignalets symmetri (statistisk sett) om tidsaksen gjør at vi kan sette

$$P(x_1 = c, x_2 = c) = P(x_1 = -c, x_2 = -c) \quad (6.62)$$

Disse to simultane sannsynligheter svarer til at antall tegnskift n i perioden τ er et like tall. Tilsvarende har vi

$$P(x_1 = -c, x_2 = c) = P(x_1 = c, x_2 = -c) \quad (6.63)$$

og da er n et ulike tall. Videre har vi $P(x_1 = c) = \frac{1}{2} = P(x_1 = -c)$.

Loven om betingede sannsynligheter (6.8) gir oss da

$$\begin{aligned} P(x_1 = c, x_2 = c) &= P(x_2 = c | x_1 = c) \cdot 0.5 \\ P(x_1 = -c, x_2 = c) &= P(x_2 = c | x_1 = -c) \cdot 0.5 \end{aligned} \quad (6.64)$$

Vi kan nå skrive (6.61) som

$$\Phi_{xx}(\tau) = 0.5 \cdot 2c^2[P(x_2 = c | x_1 = c) - P(x_2 = c | x_1 = -c)] \quad (6.65)$$

Det første ledet representerer sannsynligheten for et like antall tegnskift i intervallet τ , det andre et ulike antall.

(6.59) gir da

$$P(x_2 = c | x_1 = c) = e^{-\lambda|\tau|} \left(\frac{\lambda^0|t|^0}{0!} + \frac{\lambda^2|t|^2}{2!} + \dots \right) \quad (6.66)$$

$$-P(x_2 = c | x_1 = -c) = e^{-\lambda|\tau|} \left(-\frac{\lambda^1|t|^1}{1!} - \frac{\lambda^3|t|^3}{3!} - \dots \right) \quad (6.67)$$

Summerer vi de to rekrene over, kan vi skrive summen som

$$e^{-\lambda|\tau|} \left(\frac{(-\lambda)^0|t|^0}{0!} + \frac{(-\lambda)^1|t|^1}{1!} + \frac{(-\lambda)^2|t|^2}{2!} + \frac{(-\lambda)^3|t|^3}{3!} + \dots \right) = e^{-2\lambda|\tau|}, \quad (6.68)$$

og (6.65) blir derfor

$$\varphi_{xx}(\tau) = 0.5 \cdot 2c^2 [e^{-\lambda|\tau|} e^{-\lambda|\tau|}] = c^2 e^{-2\lambda|\tau|} \quad (6.69)$$

$\varphi_{xx}(\tau)$ har maksimum i $\tau = 0$ og faller eksponensielt for økende $|\tau|$, dvs. samme forløp som den ene av de to autokorrelasjonsfunksjonene som er skissert i figur 6.11.

Jo mindre λ , jo mindre avtar $\varphi_{xx}(\tau)$ for $|\tau| > 0$. Dette impliserer at telegraftsignalets tilstand i øyeblikket betyr mye for tilstanden en stund fremover, og det er jo nettopp det som er tilfelle når λ er liten, for da skifter signalet relativt sjeldent. Autokorrelasjonsfunksjoner av typen $ae^{-b|\tau|}$ er svært vanlige og det er viktig å merke seg at noen andre stokastiske prosesser enn telegraftsignalet har en slik autokorrelasjonsfunksjon. Dette gjelder f.eks. den enkleste varianten av såkalt lavpassfiltrert hvit støy, som vi kommer tilbake til seinere.

6.3.5 Oppsummering av egenskaper til korrelasjonsfunksjoner

$$\varphi_{xx}(\tau) = \varphi_{xx}(-\tau) \quad (6.70)$$

$$|\varphi_{xx}(\tau)| \leq \varphi_{xx}(0) \quad (6.71)$$

Gitt $z(t) = x(t) + y(t)$, da blir

$$\varphi_{zz}(\tau) = \varphi_{xx}(\tau) + \varphi_{yy}(\tau) + \varphi_{xy}(\tau) + \varphi_{yx}(-\tau) \quad (6.72)$$

eller $\varphi_{zz}(\tau) = \varphi_{xx}(\tau) + \varphi_{yy}(\tau)$ hvis $x(t)$ og $y(t)$ er ukorrelerte.

$$\varphi_{xy}(\tau) = \varphi_{yx}(-\tau) \quad (6.73)$$

$$\varphi_{xy}(\tau) = 0, \text{ når } x(t) \text{ og } y(t) \text{ er ukorrelerte.} \quad (6.74)$$

$$\varphi_{xy}^2(\tau) \leq \varphi_{xx}(0)\varphi_{yy}(0) \quad (6.75)$$

$$\varphi_{xy}(\tau) \leq \frac{1}{2}[\varphi_{xx}(0) + \varphi_{yy}(0)] \quad (6.76)$$

(6.76) følger av (6.75), i det man alltid har at $\frac{1}{2}(a + b) \geq \sqrt{ab}$ når a og $b > 0$
(den geometriske ulikhet).

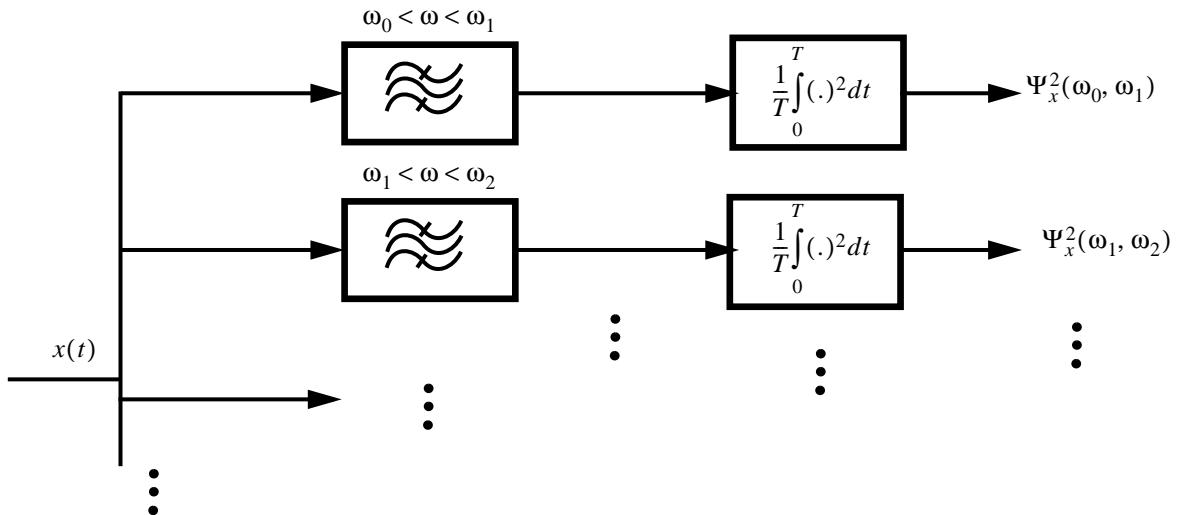
Kapittel 7: Stokastiske prosesser, effektspektra

7.1 Innledning

Vi har sett at en stasjonær stokastisk prosess kan beskrives ved hjelp av sannsynlighetstethetsfunksjoner eller ved hjelp av middelverdi og autokorrelasjon (som er entydig gitt av tethetsfunksjonene).

Vi har også sett at noen stokastiske prosesser varierer hurtigere enn andre, dvs. disse har et forholdsvis stort innhold av høyfrekvente variasjoner.

Vi skal nå utvikle et verktøy for systematisk analyse av innholdet av ulike frekvenskomponenter i et stasjonært stokastisk signal. En intuitiv metode ville være å filtrere en tidsserie gjennom en rekke båndpassfiltre som ikke overlapper hverandre og som hver dekker et lite frekvensområde. Utgangen i kvadrat fra disse filtrene kan så midles over tid. Se figur 7.1.



Figur 7.1 Undersøkelse av effektspektrum ved båndpassfiltrering

I figuren er $\omega_1 = \omega_0 + \Delta\omega$, $\omega_2 = \omega_1 + \Delta\omega$, osv. I lignings form kan vi skrive

$$\Psi_x^2(\omega, \omega + \Delta\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x^2(t, \omega, \omega + \Delta\omega) dt \quad (7.1)$$

hvor $x(t, \omega, \omega + \Delta\omega)$ er utgangen fra et båndpassfilter med passbånd fra ω til $\omega + \Delta\omega$. $\Psi_x^2(\omega, \omega + \Delta\omega)$ er effekten av signalet $x(t, \omega, \omega + \Delta\omega)$. Når $\Delta\omega \rightarrow 0$ og $T \rightarrow \infty$, definerer vi

$$G_x(\omega) = \lim_{\Delta\omega \rightarrow 0} \frac{\Psi_x^2(\omega, \omega + \Delta\omega)}{\Delta\omega} \quad (7.2)$$

Vi forstår intuitivt at hvis **effekt-tetthets-spektret** $G_x(\omega)$ har størst verdi ved lave frekvenser, er den stokastiske prosessen langsomt varierende. Hvis $G_x(\omega)$ er stor ved høye frekvenser har $x(t)$ hurtig varierende komponenter. Hvis $x(t)$ består av kun et deterministisk sinussignal, vil $G_x(\omega)$ ha en uendelig høy effekttetthet når ω er lik sinussignalets frekvens. Merk her at vi snakker om effekt-*tetthet*. Den *totale* effekten i signalet blir i følge (7.2)

$$P = \int_{\omega=0}^{\infty} G_x(\omega) d\omega \quad (7.3)$$

For enkelthets skyld dropper vi ofte “-tetthet-” og sier bare **effektspektrum**. Effektspektra er svært viktige ved analyse av dynamiske systemer fra målinger. Siden vi her studerer bare *en* prosess, $x(t)$, snakker vi om **autoeffektspektrum** Autoeffektspektret, sammen med **kryss-effektspektrret** som vi senere skal definere, gir oss muligheter til å beregne transferfunksjoner og dynamiske modeller for en prosess (som vi vil få se i et seinere kapittel). Vi skal nå utlede at *et effektspektrum er Fouriertransformen tilden tilsvarende korrelasjonsfunksjon*. Dermed er den informasjon som gis av et effektspektrum for en stokastisk prosess den samme som man finner via korrelasjonsfunksjonen.

7.2 Sammenhengen mellom autoeffektspektrum og autokorrelasjon

Vi betrakter den stokastiske prosessen $\{x(t)\}$. Hvis vi nå ser på den Fourier-transformerte til *en* realisasjon $x(t)$ i intervallet $[-T, T]$, benevnt $X_T(\omega)$, og definert som

$$X_T(\omega) = \int_{-T}^{T} x(t) e^{-j\omega t} dt \quad (7.4)$$

bør vi kunne vente at $X_T(\omega)$ sier noe om innholdet av ulike frekvenskomponenter i $\{x(t)\}$. Siden $X_T(\omega)$ er beregnet på basis av *en* realisasjon av prosessen $\{x(t)\}$, vil også $X_T(\omega)$ måtte betraktes som et utfall av en ny stokastisk prosess, avledet av (7.4).

Hvis vi nå definerer

$$\phi_T(\omega) = \frac{1}{2T} X_T(\omega) X_T(-\omega) \quad (7.5)$$

blir $\phi_T(\omega)$ også en ytterligere avledet realisasjon av en stokastisk prosess. Vil så denne si oss noe om effekten av de ulike frekvenskomponentene i realisasjonen $x(t)$ for $-T \leq t \leq T$?

Vi ønsker egentlig å se på ensemble-middelet til $\phi_T(\omega)$, altså $E[\phi_T(\omega)]$. Men vi skal først vise at tolkningen av $\phi_T(\omega)$ som et effektspektrum er rimelig.

Midlere effekt (totalt over alle frekvenser) for $x(t)$ er gitt som

$$\begin{aligned} \frac{1}{2T} \int_{-T}^T x^2(t) dt &= \frac{1}{2T} \int_{-T}^T \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} X_T(\omega) e^{j\omega t} d\omega \right] x(t) dt = \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2T} X_T(\omega) \left[\int_{-T}^T x(t) e^{j\omega t} dt \right] d\omega &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2T} X_T(\omega) X_T(-\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_T(\omega) d\omega \end{aligned} \quad (7.6)$$

hvor vi har benyttet (7.5) til slutt. Vi ser her at totaleffekten (i tidsintervallet $[-T, T]$) er lik integralet av $\frac{1}{2\pi} \phi_T(\omega)$ over alle frekvenser.

Merk: (7.6) er forøvrig identisk med den såkalte Parsevals sats, se læreboka s. 211.

Med andre ord: Det er naturlig å tolke $\frac{1}{\pi} \phi_T(\omega) \Delta\omega$ som den delen av effekten til $x(t)$ som ligger i frekvensområdet $\omega \pm \frac{\Delta\omega}{2}$ (eller om vi måtte ønske: ω til $\omega + \Delta\omega$). La oss nå definere autoeffektspektret til $x(t)$ som ensemble-middelverdien

$$\phi(\omega) = \lim_{T \rightarrow \infty} E[\phi_T(\omega)] \quad (7.7)$$

For å finne et uttrykk for (7.7) må vi nå først gjennomføre forventningsoperasjonen, og deretter foreta grenseovergangen. Vi ser først på $\phi_T(\omega)$. Vi har fra (7.4) og (7.5)

$$\begin{aligned} \phi_T(\omega) &= \frac{1}{2T} X_T(\omega) X_T(-\omega) = \frac{1}{2T} \int_{-T}^T x(t) e^{-j\omega t} dt \int_{-T}^T x(\tau) e^{j\omega \tau} d\tau \\ &= \frac{1}{2T} \int_{-T}^T \left[\int_{-T}^T x(t) x(\tau) e^{-j\omega(t-\tau)} dt \right] d\tau \end{aligned} \quad (7.8)$$

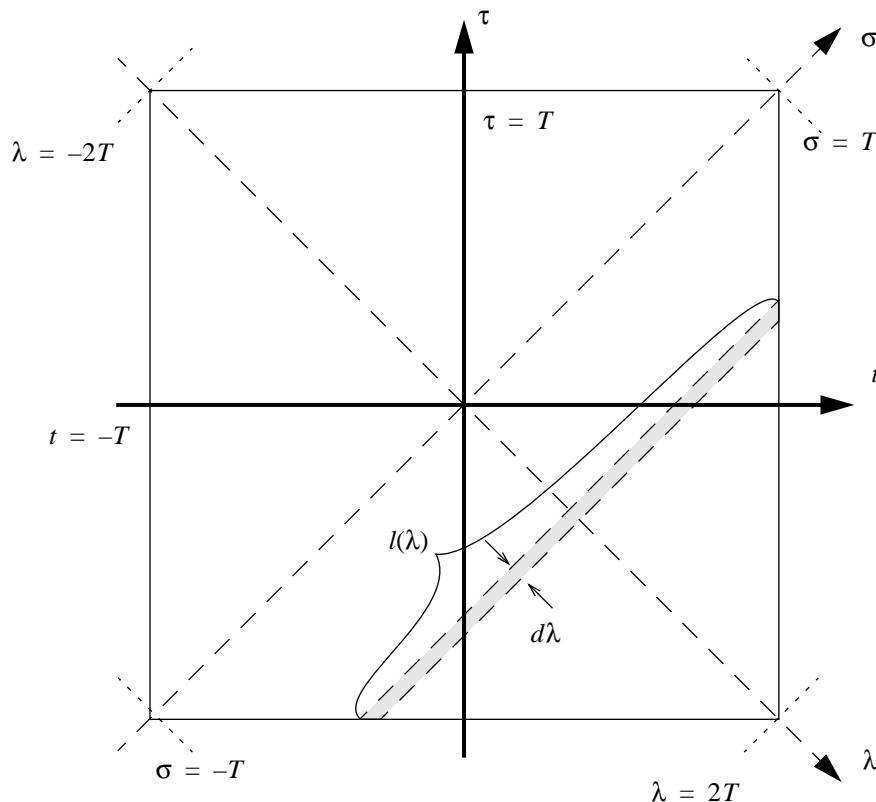
Vi får så, når vi flytter forventningsoperasjonen inn i integranden (jfr. (6.18)):

$$\phi(\omega) = \lim_{T \rightarrow \infty} E[\phi_T(\omega)] = \lim_{T \rightarrow \infty} \left(\frac{1}{2T} \int_{-T}^T \int_{-T}^T \phi_{xx}(t-\tau) e^{-j\omega(t-\tau)} dt d\tau \right) \quad (7.9)$$

Integralet ovenfor er et dobbeltintegral over området vist i figur 7.2.

For å løse dobbeltintegralet innfører vi en koordinat-transformasjon

$$\lambda = t - \tau \quad \text{og} \quad \sigma = \frac{1}{2}(t + \tau), \quad \Leftrightarrow t = \sigma + \lambda/2 \quad \text{og} \quad \tau = \sigma - \lambda/2 \quad (7.10)$$



Figur 7.2

Denne transformasjonen innebærer at vi innfører et nytt σ, λ -koordinatsystem som er dreiet 45° i forhold til t, τ -systemet.

$\lambda = t - \tau$ gir $\varphi_{xx}(\lambda)e^{-j\omega\lambda}$ i integralet over, og dette er motivet for transformasjonen. I slike smale striper med lengde $l(\lambda)$ som antydet i figur 7.2, er da $\varphi_{xx}(\lambda)e^{-j\omega\lambda} = \text{konst}$.

Følgende regel gjelder for variabelskifte i et dobbeltintegral

$$\int \int_A f(t, \tau) dt d\tau = \int \int_A f(t[\sigma, \lambda], \tau[\sigma, \lambda]) \begin{vmatrix} \frac{\partial t}{\partial \sigma} & \frac{\partial t}{\partial \lambda} \\ \frac{\partial \tau}{\partial \sigma} & \frac{\partial \tau}{\partial \lambda} \end{vmatrix} d\sigma d\lambda \quad (7.11)$$

Tallverdien av Jacobi-matrisens determinant

I vårt tilfelle blir Jacobi-matrisens determinant

$$\begin{vmatrix} 1 & 0.5 \\ 1 & -0.5 \end{vmatrix} = -1, \text{ og } |-1| = 1 \quad (7.12)$$

Integralet blir da

$$\begin{aligned}
 I &= \int_{\lambda = -2T}^{\lambda = 2T} \int_{\sigma_n(\lambda)}^{\sigma_o(\lambda)} \varphi_{xx}(\lambda) e^{-j\omega\lambda} \cdot 1 \cdot d\sigma d\lambda = \int_{\lambda = -2T}^{\lambda = 2T} \int_{\sigma_n(\lambda)}^{\sigma_o(\lambda)} \varphi_{xx}(\lambda) e^{-j\omega\lambda} d\sigma d\lambda \\
 &= \int_{\lambda = -2T}^{\lambda = 2T} \varphi_{xx}(\lambda) e^{-j\omega\lambda} [\sigma_o(\lambda) - \sigma_n(\lambda)] d\lambda = \int_{\lambda = -2T}^{\lambda = 2T} \varphi_{xx}(\lambda) e^{-j\omega\lambda} l(\lambda) d\lambda \\
 &= \int_{\lambda = -2T}^{\lambda = 2T} \varphi_{xx}(\lambda) e^{-j\omega\lambda} (2T - |\lambda|) d\lambda
 \end{aligned} \tag{7.13}$$

σ_n og σ_o er lineære funksjoner av λ som avgrenser det ruter-formede arealet i figur 7.2.

At $[\sigma_o(\lambda) - \sigma_n(\lambda)] = l(\lambda) = 2T - |\lambda|$, sees av figur 7.2.

(7.9) blir nå

$$\begin{aligned}
 \phi(\omega) &= \lim_{T \rightarrow \infty} E[\phi_T(\omega)] = \lim_{T \rightarrow \infty} \frac{1}{2T} I = \lim_{T \rightarrow \infty} \left(\frac{1}{2T} \int_{\lambda = -2T}^{\lambda = 2T} \varphi_{xx}(\lambda) e^{-j\omega\lambda} (2T - |\lambda|) d\lambda \right) \\
 &= \int_{\lambda = -\infty}^{\lambda = \infty} \varphi_{xx}(\lambda) e^{-j\omega\lambda} d\lambda - \lim_{T \rightarrow \infty} \left(\int_{\lambda = -2T}^{\lambda = 2T} \frac{|\lambda|}{2T} \varphi_{xx}(\lambda) e^{-j\omega\lambda} d\lambda \right)
 \end{aligned} \tag{7.14}$$

Hvis nå

$$\int_{\lambda = -\infty}^{\lambda = \infty} |\lambda| \varphi_{xx}(\lambda) e^{-j\omega\lambda} d\lambda < \infty, \tag{7.15}$$

så vil det siste integralet i (7.14) gå mot null når $T \rightarrow \infty$. Dette vil være oppfylt for alle stokastiske prosesser som ikke inneholder reine periodiske komponenter, og som har null middelverdi. Vi får dermed:

$$\lim_{T \rightarrow \infty} E[\phi_T(\omega)] = \phi_{xx}(\omega) = \int_{-\infty}^{\infty} \varphi_{xx}(\tau) e^{-j\omega\tau} d\tau \tag{7.16}$$

Det betyr at $\varphi_{xx}(\tau)$ og $\phi_{xx}(\omega)$ er Fouriertransform-par.

Vi har innført indeks "xx" i autoeffektspektret for å skille det fra krysseffektspektret, som vi kommer til seinere i dette kapitlet.

Begge disse representasjonene - tidsplanbeskrivelsen gitt av $\phi_{xx}(\tau)$ og frekvensplanbeskrivelsen gitt av $\phi_{xx}(\omega)$ - inneholder akkurat like mye informasjon om prosessen.

(De er ekvivalente på samme måte som et lineært system kan representeres på to likeverdige måter: Ved frekvensresponsen $H(\omega)$ eller ved impulsresponsen $h(t)$).

Hvis betingelsen (7.15) ikke er oppfylt, dvs. prosessen inneholder periodiske komponenter, og/eller har middelverdi forskjellig fra null (som forøvrig kan tolkes som en periodisk komponent med frekvens = null), vil autokorrelasjonsfunksjonen inneholde cos-funksjoner, som vist i kapittel 6, avsnitt 6.3.2. Det kan vises at hvis man tillater deltafunksjoner i $\phi_{xx}(\omega)$, så vil $\phi_{xx}(\tau)$ og $\phi_{xx}(\omega)$ også i dette grensetilfellet være Fouriertransform-par. Figur 7.3 gir eksempler på noen par av autokorrelasjonsfunksjoner og autoeffektspektra.

La oss nå kommentere $\phi_{xx}(\omega)$ og noen av dens egenskaper:

1. $\phi_{xx}(\omega)$ er reell og symmetrisk. Dvs. at $\phi_{xx}(\omega) = \phi_{xx}(-\omega)$. Det ser vi av

$$\phi_{xx}(\omega) = \int_{-\infty}^{\infty} \phi_{xx}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} \phi_{xx}(\tau) \cos(\omega\tau) d\tau - j \int_{-\infty}^{\infty} \phi_{xx}(\tau) \sin(\omega\tau) d\tau \quad (7.17)$$

Leddet med sinus faller bort, da $\phi_{xx}(\tau)$ er symmetrisk om $\tau = 0$.

2. $\phi_{xx}(\omega) \geq 0$. Dette følger av (7.5) og utledningen videre.

Eller av følgende enkle betrakting: Effekten kan ikke være negativ i noe frekvensområde!

3. Totaleffekten i signalet er gitt av $E(x^2) = \phi_{xx}(0)$; jfr. (6.52).

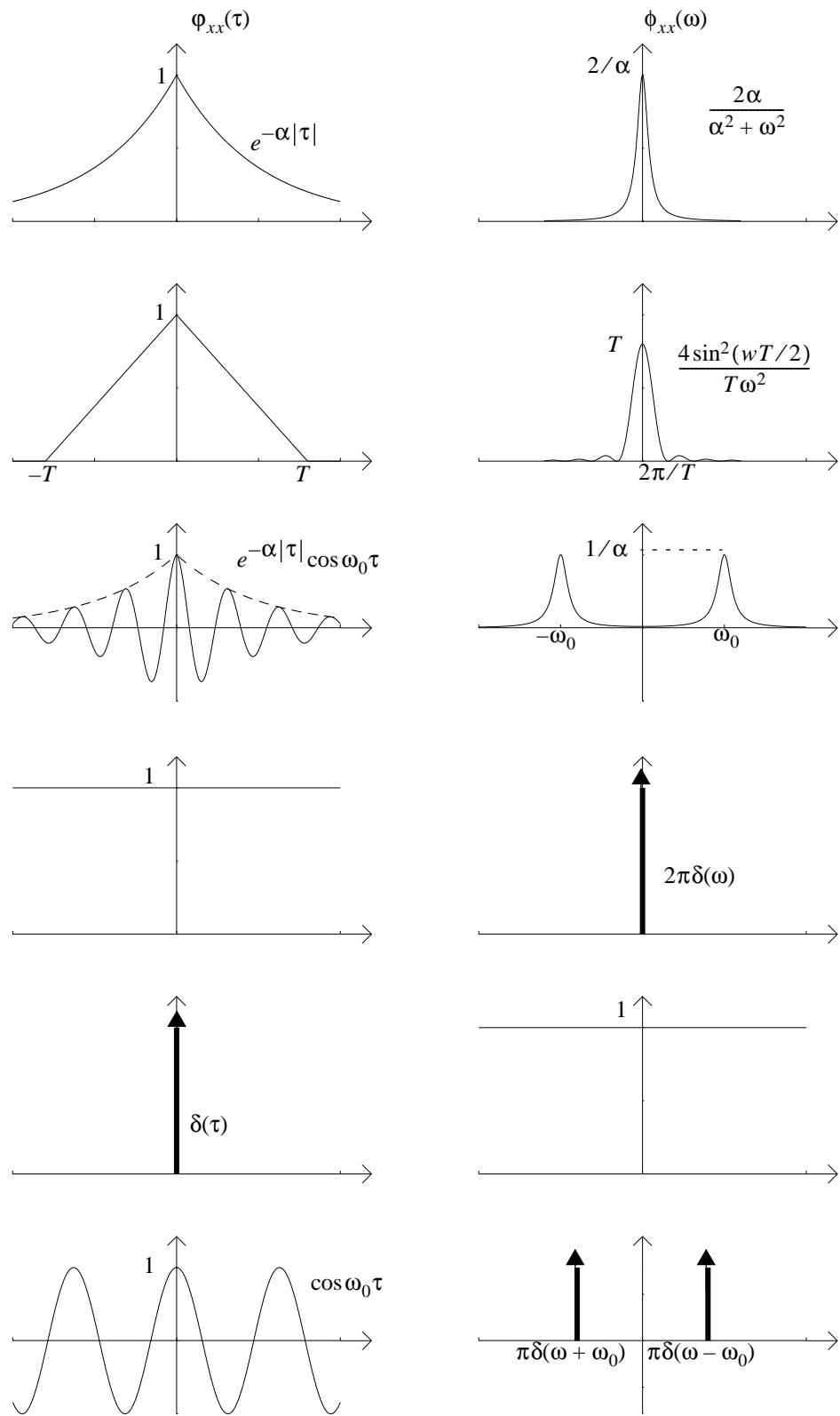
Men vi har (invers Fouriertransformasjon) at

$$P = \phi_{xx}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{xx}(\omega) e^{j\omega 0} d\omega = \frac{1}{\pi} \int_0^{\infty} \phi_{xx}(\omega) d\omega \quad (7.18)$$

Forenklinga med nedre integrasjonsgrense = 0 får vi fordi $\phi_{xx}(\omega)$ er en like funksjon. $\phi_{xx}(\omega)$ angir hvordan signaleffekten i $x(t)$ er spredt i de ulike frekvensområdene, og totaleffekten blir flateinnholdet under denne kurven multiplisert med $1/\pi$.

4. Effekten i et frekvensintervall $\omega \in [\omega_1, \omega_2]$ vil være gitt av $P_{\omega_1 - \omega_2}$

$$P_{\omega_1 - \omega_2} = \frac{1}{\pi} \int_{\omega_1}^{\omega_2} \phi_{xx}(\omega) d\omega \quad (7.19)$$



Figur 7.3 Noen autoeffektspektrum/autokorrelasjons-par.

5. Sammenhengen mellom $G_x(\omega)$ definert i (7.2) og $\phi_{xx}(\omega)$ blir

$$G_x(\omega) = \frac{1}{\pi} \phi_{xx}(\omega) \quad (7.20)$$

6. Dersom $x(t)$ inneholder periodiske komponenter med frekvensene ω_i , vil $\phi_{xx}(\omega)$ ha impulser (linjer) ved frekvensene ω_i . Det har den fysiske tolkningen av effekttetheten er uendelig for disse diskrete frekvensene.

7.3 Krysseffektspektrum og krysskorrelasjon

Vi har sett at effektspektret til en prosess kan finnes ved å Fourier-transformere autokorrelasjonsfunksjonen til den stokastiske prosessen. Begrepet **krysseffektspektrum** for to stokastiske prosesser $\{x(t)\}$ og $\{y(t)\}$ fremkommer helt analogt som den Fourier-transformerte til krysskorrelasjonsfunksjonen. Siden krysskorrelasjonsfunksjonen i motsetning til autokorrelasjonsfunksjonen generelt ikke er en like funksjon i τ (ikke symmetrisk i τ), er krysseffektspektret generelt en kompleks funksjon av ω . Vi har

$$\begin{aligned} \phi_{xy}(\omega) &= \int_{-\infty}^{\infty} \phi_{xy}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} \phi_{xy}(\tau) \cos(\omega\tau) d\tau - j \int_{-\infty}^{\infty} \phi_{xy}(\tau) \sin(\omega\tau) d\tau \\ &= C_{xy}(\omega) - jQ_{xy}(\omega) \end{aligned} \quad (7.21)$$

$C_{xy}(\omega)$ kalles **cospektret** til $x(t)$ og $y(t)$, mens $Q_{xy}(\omega)$ kalles **kvadraturspektret**.

Hvis vi ønsker å tolke $C_{xy}(\omega)$ og $Q_{xy}(\omega)$ ut fra frekvensbetrakninger, kan cospektret tolkes som det midlere produktet av $x(t)$ og $y(t)$ i et smalt frekvensbånd. Dvs.

$$\frac{1}{\pi} C_{xy}(\omega) = \lim_{\Delta\omega \rightarrow 0} \left[\frac{1}{\Delta\omega} \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t, \omega, \omega + \Delta\omega) y(t, \omega, \omega + \Delta\omega) dt \right) \right] \quad (7.22)$$

hvor $x(t, \omega, \omega + \Delta\omega)$ er utgangen fra et båndpassfilter med passbånd fra ω til $\omega + \Delta\omega$.

(Vi har $1/\pi$ i (7.22) - ikke $1/(2\pi)$ - fordi $C_{xy}(\omega)$ er definert slik at det har verdi både for positive og negative frekvenser). Tilsvarende får vi

$$\frac{1}{\pi} Q_{xy}(\omega) = \lim_{\Delta\omega \rightarrow 0} \left[\frac{1}{\Delta\omega} \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t, \omega, \omega + \Delta\omega) y^\circ(t, \omega, \omega + \Delta\omega) dt \right) \right] \quad (7.23)$$

hvor y° er lik y med 90° faseforsinkelse.

En annen måte å uttrykke et krysseffektspektrum på er

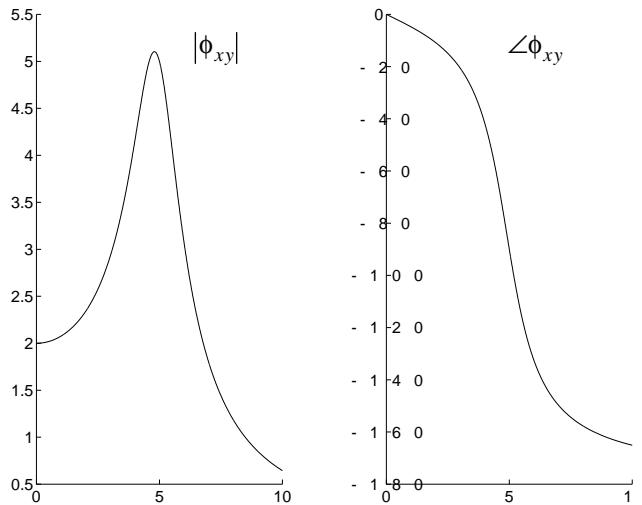
$$\phi_{xy}(\omega) = |\phi_{xy}(\omega)| e^{-j\angle\phi_{xy}(\omega)} \quad (7.24)$$

hvor

$$|\phi_{xy}(\omega)| = \sqrt{C_{xy}^2(\omega) + Q_{xy}^2(\omega)} \quad (7.25)$$

$$\angle\phi_{xy}(\omega) = \text{atan}\left(\frac{Q_{xy}(\omega)}{C_{xy}(\omega)}\right) \quad (7.26)$$

I dette tilfellet kan $\angle\phi_{xy}(\omega)$ tolkes som faseforsinkelsen for den delen av y som er korrelert med x i forhold til x , og for den gitte frekvensen. Et eksempel på et forløp av $|\phi_{xy}(\omega)|$ og $\angle\phi_{xy}(\omega)$ er vist i figur 7.4. Forløpene minner sterkt om frekvensresponsen til en transferfunksjon, noe vi vil komme tilbake til i kapittel 8 og 9.



Figur 7.4

Vi skal se litt på ulike egenskaper for $\phi_{xy}(\omega)$. Det kan vises at

$$|\phi_{xy}(\omega)|^2 \leq \phi_{xx}(\omega)\phi_{yy}(\omega) \quad (7.27)$$

Med basis i (7.27) defineres **koherensfunksjonen**

$$\gamma_{xy}^2(\omega) = \frac{|\phi_{xy}(\omega)|^2}{\phi_{xx}(\omega)\phi_{yy}(\omega)} \quad (7.28)$$

Det framgår av (7.27) og (7.28) at vi alltid har $\gamma_{xy}^2(j\omega) \leq 1$. Hvis γ_{xy}^2 er null ved en gitt frekvens, sies x og y å være **inkoherente** ved denne frekvensen. Med andre ord er x og y ukorrelerte ved denne frekvensen. Hvis x og y er statistisk uavhengige, er disse inkohrente for alle frekvenser. Hvis $\gamma_{xy}^2 = 1$, sies x og y å være **komplett koherente**. Dette gjelder f.eks. for det tilfellet hvor $y(s) = H(s)x(s)$, og $H(s)$ er lineær.

To andre viktige egenskaper ved $\phi_{xy}(\omega)$ er gitt av

$$\phi_{xy}(\omega) = \phi_{yx}(-\omega) \quad (7.29)$$

og

$$\phi_{xy}(\omega) = \phi_{yx}^*(\omega) \quad (* \text{ betyr her "kompleks konjugert"}) \quad (7.30)$$

Dette følger fra definisjonen (7.21) og sammenhengen $\phi_{xy}(\tau) = \phi_{yx}(-\tau)$.

Vi skal nå se på et eksempel på krysseffektspektrum

La oss betrakte telegraftsignalet fra kapittel 6, avsnitt 6.3.4. Dette hadde autokorrelasjonsfunksjonen (6.69)

$$\phi_{xx}(\tau) = c^2 e^{-2\lambda|\tau|} \quad (7.31)$$

Autoeffektspektret er gitt av den Fourier-transformerte til $\phi_{xx}(\tau)$ som kan finnes ved hjelp av figur 7.3, øverst:

$$\phi_{xx}(\omega) = \frac{4c^2\lambda}{4\lambda^2 + \omega^2} \quad (7.32)$$

Som vi ser er ϕ_{xx} reell og symmetrisk om $\omega = 0$. Anta så at vi danner $y(t) = x(t - T)$, dvs. utgangssignalet er en tidsforsinket versjon av inngangssignalet. Vi ønsker å finne $\phi_{xy}(\omega)$. Fra (6.58) har vi

$$\phi_{xy}(\tau) = \phi_{xx}(\tau - T) \quad (7.33)$$

Fra definisjonen av $\phi_{xy}(\omega)$ og (7.33) får vi

$$\phi_{xy}(\omega) = \int_{-\infty}^{\infty} \phi_{xx}(\tau - T) e^{-j\omega\tau} d\tau \quad (7.34)$$

Setter $\sigma = \tau - T$ og får

$$\phi_{xy}(\omega) = \int_{-\infty}^{\infty} \phi_{xx}(\sigma) e^{-j\omega\sigma} d\sigma \cdot e^{-j\omega T} = \frac{4c^2\lambda}{4\lambda^2 + \omega^2} e^{-j\omega T} \quad (7.35)$$

Vi ser at $\phi_{xy}(\omega)$ ikke lenger er rent reell. Vi har

$$\phi_{xy}(\omega) = \phi_{xx}(\omega) \cos(\omega T) + j\phi_{xx}(\omega) \sin(\omega T) \quad (7.36)$$

slik at $C_{xy}(\omega) = \phi_{xx}(\omega) \cos \omega T$ og $Q_{xy}(\omega) = \phi_{xx}(\omega) \sin \omega T$. $\angle\phi_{xy}(\omega)$ er gitt av $\angle\phi_{xy}(\omega) = \omega T$.

7.4 Hvit støy

Av alle typer stokastiske prosesser er det én som spiller en dominerende rolle, nemlig det vi kaller **hvit støy**. Denne stokastiske prosessen er egentlig bare et matematisk begrep, fordi den aldri helt ut kan realiseres i praksis. Blant annet har hvit støy uendelig stor effekt. Likevel vil hvit støy ofte benyttes for modellering av prosesser som *tilnærmet* kan betraktes som hvitstøyprosesser.

En hvitstøy-prosess er karakterisert ved at verdien av den variable $x(t_1)$ ved tidspunktet t_1 er ukorrelert med verdien $x(t_2)$ ved *alle andre* tidspunkter, $t_2 \neq t_1$. Av dette forstår vi at hvitstøy-prosessen varierer meget hurtig. I motsatt fall ville vi fått korrelasjon mellom $x(t_1)$ og $x(t_2)$. Autokorrelasjonsfunksjonen for hvit støy er null for $t_1 - t_2 \neq 0$.

For å behandle egenskapene til hvitstøy-prosessen, skal vi først betrakte en tilnærmet hvitstøy-prosess og foreta en grenseovergang til det eksakte tilfelle.

La $\{x_\Delta(t)\}$ være den kontinuerlige prosess definert ved at

$$x_\Delta(t) = x[k], \quad k\Delta \leq t \leq (k+1)\Delta, \quad (7.37)$$

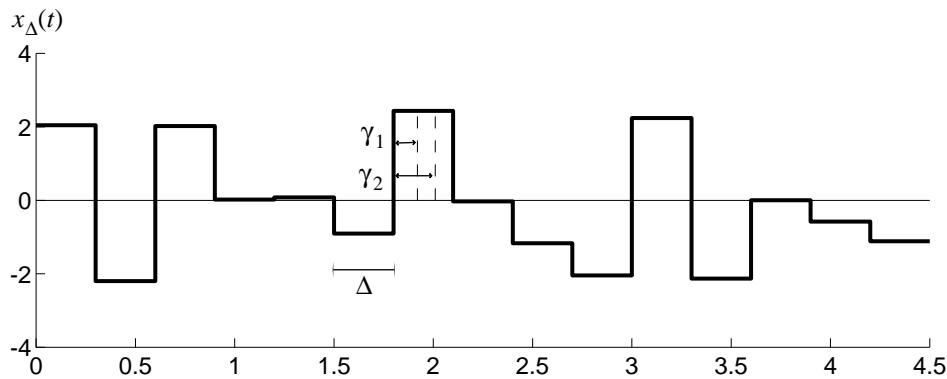
og hvor

$$E(x[k]) = 0$$

$$E(x[k]x[l]) = \begin{cases} 0, & k \neq l \\ Q_0, & k = l \end{cases} \quad (7.38)$$

$$Q_0 = Q/\Delta \quad (7.39)$$

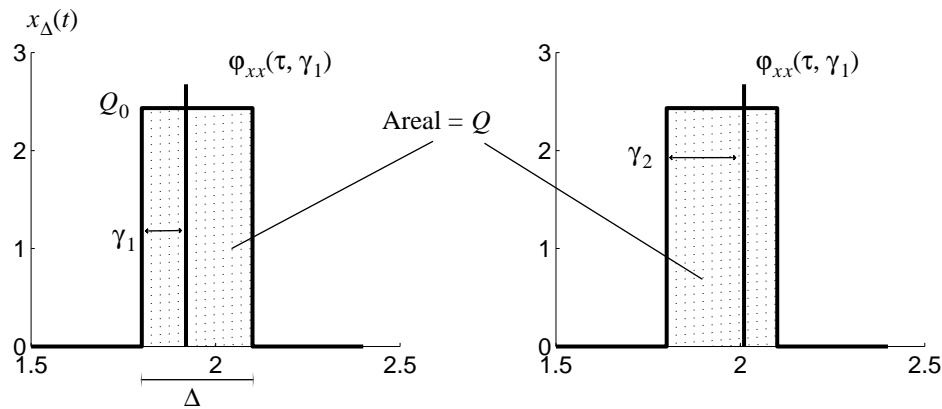
Q er en konstant og Δ er et tidsinkrement. $x_\Delta(t)$ er konstant over tidsinkrementet Δ . En realisasjon av $x_\Delta(t)$ er vist i figur 7.5.



Figur 7.5

Vi ønsker å la intervallet Δ gå mot null og observere effekten av dette. Av (7.39) ser vi umiddelbart at slik Q_0 er valgt, vil variansen gå mot uendelig når $\Delta \rightarrow 0$. Dette valget av Q_0 vil bli rettferdiggjort under. Vi ønsker å finne autokorrelasjonsfunksjonen til prosessen. Nå er φ_{xx} avhengig av hvor i intervallene av lengde Δ den er beregnet. Med andre ord er denne prosessen ikke stasjonær.

Figur 7.6 illustrerer dette.



Figur 7.6

φ_{xx} blir en funksjon av tidsforskyvningen τ og parameteren γ , som er definert i figur 7.5 og 7.6. Med utgangspunkt i (7.38) og (7.39) har vi da umiddelbart at

$$\varphi_{xx}(\tau, \gamma) = \begin{cases} Q/\Delta, & -\gamma \leq \tau < \Delta - \gamma \\ 0, & \text{ellers} \end{cases} \quad (7.40)$$

Når vi lar $\Delta \rightarrow 0$, vil begge autokorrelasjonsfunksjoner, uavhengig av γ , bli smalere og høyere og i grensen gå mot en Dirac-puls (deltafunksjon), og dermed også bli identiske. Prosessen er altså blitt stasjonær, dens statistiske egenskaper er de samme over alt langs tidsaksen. Arealet under $\phi_{xx}(\tau)$ er Q . Fordi dette arealet er konstant når grenseovergangen gjennomføres.

Denne prosessen er en kontinuerlig hvit støy.

Vi kan derfor oppsummere: En kontinuerlig hvitstøy-prosess har en autokorrelasjonsfunksjon

$$\phi_{xx}(\tau) = Q \delta(\tau) \quad (7.41)$$

Q er en konstant og $\delta(t)$ er en deltafunksjon. Hvis $Q = 1$, kalles hvitstøyprosessen i tillegg **rudimentær**.

Ved å Fouriertransformere (7.41) får vi at effektspektret til en kontinuerlig hvitstøy-prosess blir

$$\phi_{xx}(\omega) = \int_{-\infty}^{\infty} \phi_{xx}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} Q \delta(\tau) e^{-j\omega\tau} d\tau = Q \int_{-\infty}^{\infty} \delta(\tau) d\tau = Q \quad (7.42)$$

Herav ser vi berettigelsen av begrepet "hvit" støy: Hvit støy inneholder alle frekvenskomponenter med samme intensitet. Akkurat som hvitt lys kan brukes til å generere alle farger i spektret, kan vi modellere eller generere andre støyprosesser med vilkårlig spektrum ved å sende hvit støy inn på et lineært filter. Dette er hovedårsaken til at hvit støy er et så viktig begrep. (7.42) bekrefter også at hvit støy har uendelig effekt, da flateinnholdet under $\phi_{xx}(\omega) = Q = \text{konst.}$, blir ∞ .

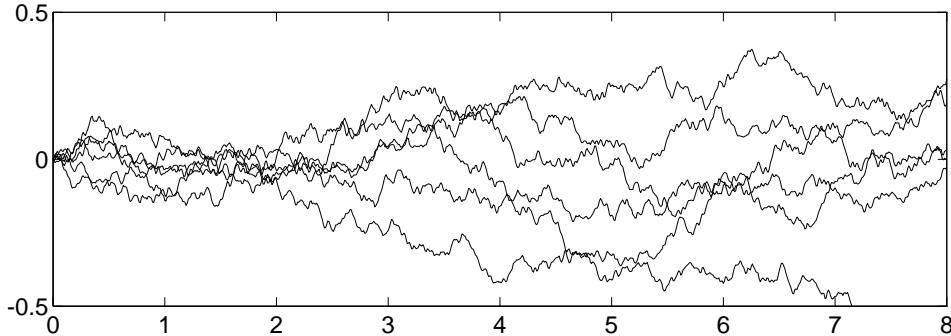
7.4.1 Integralet av hvit støy

Vi skal se på en stokastisk prosess $y(t)$ som er lik integralet av hvit støy. Den hvite støyen er $v(t)$, gitt ved $E[v(t)] = 0$ og $\phi_{vv}(\tau) = Q\delta(\tau)$. Vi søker middelverdi og varians for utgangen $y(t)$:

$$E[y(t)] = E\left[\int_0^t v(\alpha) d\alpha\right] = \int_0^t E[v(\alpha)] d\alpha = 0 \quad (7.43)$$

$$\begin{aligned} E[y^2(t)] &= E\left[\int_0^t v(\alpha) d\alpha \int_0^t v(\beta) d\beta\right] = E\left[\int_0^t \int_0^t v(\alpha) v(\beta) d\beta d\alpha\right] \\ &= \int_0^t \left[\int_0^t Q \delta(\beta - \alpha) d\beta \right] d\alpha = Q \int_0^t [1] d\alpha = Qt \end{aligned} \quad (7.44)$$

Denne spesielle prosessen kalles en Wiener-prosess eller en Brownsk prosess. Som vi ser, øker variansen til $y(t)$ lineært med tiden. $\{y(t)\}$ er derfor ikke en stasjonær prosess i motsetning til en hvitstøy-prosess. Noen ulike realisasjoner av $y(t)$ er vist i figur 7.7.



Figur 7.7 Noen realisasjoner av Wiener-prosessen $\{y(t)\}$.

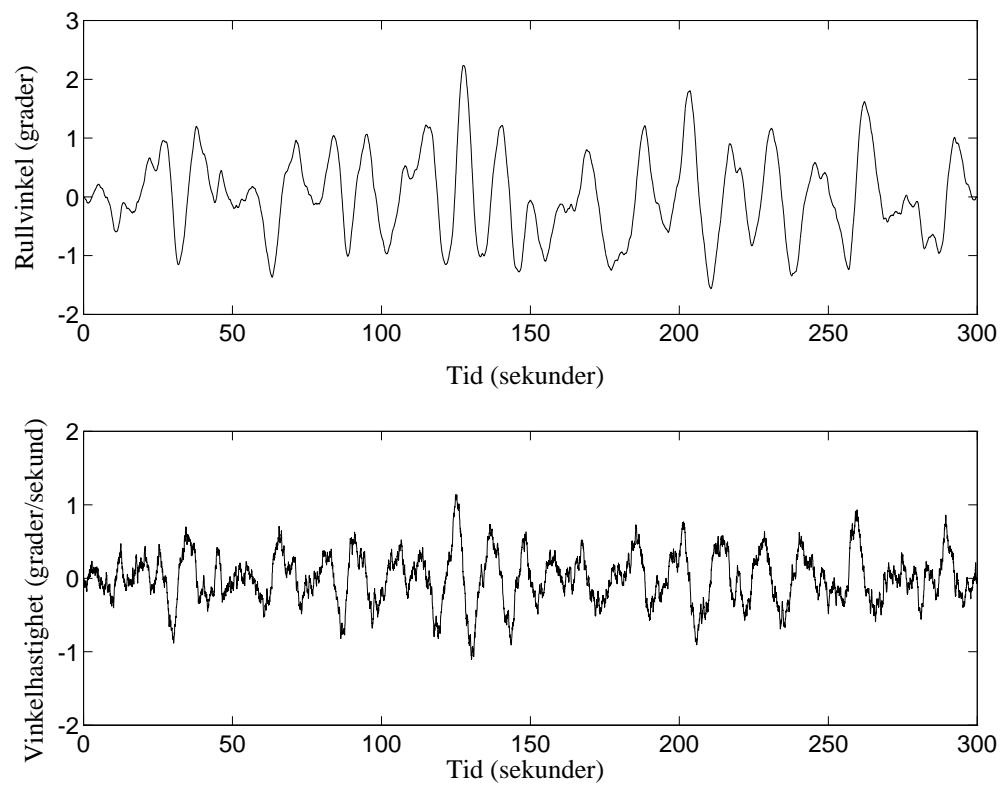
7.4.2 Rullebevegelsen for skip - bruk av støy i modellering

Som et eksempel på en stasjonær stokastisk prosess kan en tenke seg rullebevegelsen for et sjøgående fartøy under stasjonære værbetingelser. Man kan tenke seg en slik bevegelse modellert ved en dempet harmonisk oscillator eksitert av støy, i dette tilfelle bølger. Dette er selvsagt ingen eksakt beskrivelse, men gir for mange forhold en god nok beskrivelse. Modellen får følgende form

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -cx_2 - \omega_0^2 x_1 + \nu\end{aligned}\tag{7.45}$$

hvor x_1 er rullvinkel, x_2 rullhastighet, c er dempningskoeffisient, ω_0 er egenfrekvensen for rullebevegelsen og ν er stasjonær støy (ikke hvit, spektret er ikke flatt, men avtar ved lave og høye frekvenser). En simulert realisasjon av prosessens variable er vist i figur 7.8 for $c = 0.15$ og $\omega_0 = 0.5$. Disse prosessene kan betraktes som stokastisk stasjonære fordi de statistiske parametrene for ν , og dermed for x_1 og x_2 , er tilnærmet konstante i tid, sett i forhold til den relativt hurtige dynamikken til fartøyet. Hvis imidlertid værforholdene endrer seg raskt, vil de statistiske parametrene for ν endre seg og prosessene må betraktes som ikke-stasjonære.

Dette eksemplet kan ses å som en innledning til kapittel 8 hvor vi skal ta for oss hvordan stokastiske prosesser påvirker dynamiske systemer og hvordan sammenhengen er mellom ulike variable i dynamiske systemer som er eksitert av stokastiske prosesser.



Figur 7.8 Simulerte realisasjoner av rull og rullhastighet.

Kapittel 8: Stokastiske prosesser i lineære dynamiske systemer

Vi skal i dette kapittelet analysere forholdene i et lineært, monovariabelt dynamisk system som påvirkes av stasjonære stokastiske prosesser.

8.1 Grunnleggende sammenhenger

Vi skal først betrakte det forholdet at vi har et veldefinert og målbart pådrag, $u(t)$, som eneste eksitasjon til det lineære dynamiske systemet med frekvensresponsen $H(\omega)$ svarende til impulsresponsen $h(t)$.

Vi antar at $u(t)$ er en realisasjon av den stokastiske prosessen $\{u(t)\}$. Utgangen fra systemet er $y(t)$. Vi har foldingsintegralet

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)d\tau , \quad \text{alternativt} \quad y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau \quad (8.1)$$

Øvre integrasjonsgrense i venstre uttrykk kan alltid settes lik ∞ , også for kausale systemer (som blir et spesialtilfelle). Dette kan gjøres fordi $h(t-\tau)$ er null for $t-\tau < 0$ for kausale systemer. Et tilsvarende resonnement gjelder for høyre uttrykk.

Autokorrelasjonsfunksjonen $\varphi_{uu}(\tau)$ til $u(t)$ er kjent. Vi skal nå beregne autokorrelasjonsfunksjonen til $y(t)$, $\varphi_{yy}(\tau) = E[y(t)y(t+\tau)] = E[y(t)y(t-\tau)]$.

Vi bruker høyre del av (8.1) og finner først

$$y(t)y(t-\tau) = \int_{-\infty}^{\infty} h(\alpha)u(t-\alpha)d\alpha \int_{-\infty}^{\infty} h(\beta)u(t-\tau-\beta)d\beta \quad (8.2)$$

Dette kan gjøres om til et dobbeltintegral:

$$y(t)y(t-\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha)h(\beta)u(t-\alpha)u(t-\tau-\beta)d\alpha d\beta \quad (8.3)$$

Hvis vi nå tar forventningsverdi på begge sider, får vi

$$\varphi_{yy}(\tau) = E[y(t)y(t-\tau)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha)h(\beta)\varphi_{uu}(\tau-\alpha+\beta)d\alpha d\beta \quad (8.4)$$

Vi ser hermed at autokorrelasjonsfunksjonen til y kan beregnes fra (8.4) når autokorrelasjonsfunksjonen til u og impulsresponsen fra u til y er kjent.

Nå skal vi vise at det er langt letttere å gjennomføre denne beregninga i frekvensplanet, dvs. ved hjelp av autoeffektspektret til u og transferfunksjonen $H(\omega)$. Vi Fouriertransformerer (8.4), og får:

$$\phi_{yy}(\omega) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha)h(\beta)\varphi_{uu}(\tau - \alpha + \beta) d\alpha d\beta \right) e^{-j\omega\tau} d\tau \quad (8.5)$$

(8.5) kan skrives

$$\begin{aligned} \phi_{yy}(\omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha)h(\beta)\varphi_{uu}(\tau - \alpha + \beta) e^{-j\omega\tau} e^{j\omega\alpha} e^{-j\omega\alpha} e^{j\omega\beta} e^{-j\omega\beta} d\alpha d\beta d\tau \\ &= \int_{-\infty}^{\infty} h(\alpha) e^{-j\omega\alpha} \int_{-\infty}^{\infty} h(\beta) e^{j\omega\beta} \left(\int_{-\infty}^{\infty} \varphi_{uu}(\tau - \alpha + \beta) e^{-j\omega(\tau - \alpha + \beta)} d\tau \right) d\beta d\alpha \end{aligned} \quad (8.6)$$

som jo er det samme som

$$\phi_{yy}(\omega) = H(\omega)H(-\omega)\phi_{uu}(\omega) \quad (8.7)$$

Dette kan alternativt skrives

$$\phi_{yy}(\omega) = |H(\omega)|^2\phi_{uu}(\omega) \quad (8.8)$$

Vi kan tilsvarende finne følgende relasjon mellom $\varphi_{uu}(\tau)$ og krysskorrelasjonsfunksjonen $\varphi_{uy}(\tau)$:

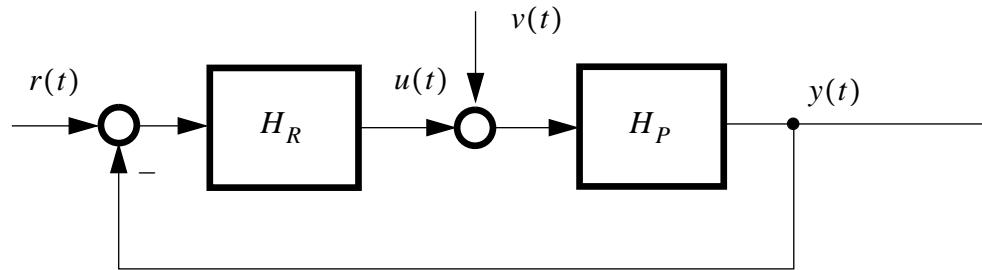
$$\varphi_{uy}(\tau) = \int_{-\infty}^{\infty} h(\alpha)\varphi_{uu}(\tau - \alpha) d\alpha = h(\tau)^*\varphi_{uu}(\tau) \quad (8.9)$$

som i frekvensplanet blir

$$\phi_{uy}(\omega) = H(\omega)\phi_{uu}(\omega) \quad (8.10)$$

Et interessant spesialtilfelle av (8.10) er når $u(t)$ er rudimentær hvitstøy; $\varphi_{uu}(\omega) = 1$. Da blir $\phi_{uy}(\omega) = H(\omega)$, dvs. krysseffektspekteret blir lik frekvensresponsen. Dette er tilfellet med krysseffektspekret vist i figur 7.4. I dette tilfellet blir sjølsagt også $\varphi_{uy}(\tau) = h(t)$.

Nå til sammenhengene (8.7), (8.8) er svært viktige og nyttige relasjoner. Hvis vi f.eks. betrakter reguleringssystemet i figur 8.1, kan vi kvantitativt beregne systemets støyreduserende egenskaper.



Figur 8.1 Reguleringssystem med forstyrrelse $v(t)$.

Vi antar at $v(t)$ er et ergodisk signal med null middelverdi. Da er effekten til $y(t)$ med reguleringen utkoplet gitt av variansen

$$\sigma_y^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y^2(t) dt = \phi_{yy}(0) = \frac{1}{\pi} \int_0^\infty \phi_{yy}(\omega) d\omega = \frac{1}{\pi} \int_0^\infty |H_p(\omega)|^2 \phi_{vv}(\omega) d\omega \quad (8.11)$$

Med regulatoren innkoplet, får vi

$$\sigma_{yR}^2 = \frac{1}{\pi} \int_0^\infty |H_p(\omega)N(\omega)|^2 \phi_{vv}(\omega) d\omega \quad (8.12)$$

der

$$N(\omega) = \frac{1}{1 + H_R(\omega)H_p(\omega)} \quad (8.13)$$

er **reguleringsgraden** til systemet; dette begrepet er kjent fra grunnkurs i reguleringsteknikk.

Forbedringen, eller den **effektive reguleringsgraden** kan defineres som

$$N_{\text{eff}} = \sigma_{yR}/\sigma_y \quad (8.14)$$

Vi ser at for å oppnå en effektiv motvirkning av støyen må de energirike delene av spektret $\phi_{vv}(\omega)$ falle utenfor den delen av spektret hvor $N(\omega)$ har verdier større eller lik 1. H_R må velges slik at det lukkede system får større båndbredde enn høyeste frekvens hvor $\phi_{vv}(\omega)$ er signifikant større enn 0. Vi skal komme tilbake til dette senere hvor vi skal se på metoder for bl.a. å minimalisere σ_{yR}^2 .

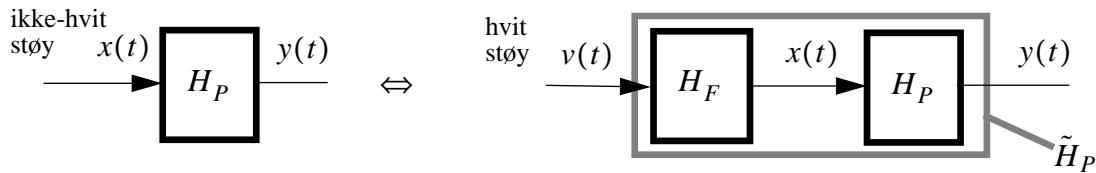
Fra likning (8.7) ser vi at vi faktisk kan generere en stokastisk prosess med et vilkårlig spektrum ved å la hvit støy passere et dynamisk system med en passende transferfunksjon. Vi lar $v(t)$ være rudimentær hvit støy med autokorrelasjonsfunksjonen

$$\phi_{vv}(\tau) = \delta(\tau) \text{ og effektspekret } \phi_{vv}(\omega) = 1 \quad (8.15)$$

Et dynamisk system med transferfunksjon $H_F(\omega)$ som påtrykkes denne støyen får et utgangssignal med effektspektrum gitt av

$$\phi_{xx}(\omega) = H_F(\omega)H_F(-\omega)\phi_{vv}(\omega) = H_F(\omega)H_F(-\omega) \quad (8.16)$$

med andre ord kan en stokastisk prosess $\{x(t)\}$ med autoeffektspektrum $\phi_{xx}(\omega)$ tolkes som hvit støy som er filtrert gjennom en transferfunksjon $H_F(\omega)$. Dette er den såkalte **spektralfaktoriseringssatsen** som sier at et vilkårlig spektrum kan faktoriseres i to kompleks konjugerte deler. Det betyr igjen at man for system som er eksitert av en vilkårlig støyprosess kan trekke transferfunksjonen for dette tenkte filteret inn i systemmodellen og la det nye systemet man da får, være eksitert av hvit støy. Se figur 8.2.



Figur 8.2

Spektralfaktoriseringssatsen gir oss en enorm fordel: Uansett hva slags spektrum den virkelige stokastiske prosess $\{y(t)\}$ har, kan dette omformes til en ekvivalent modell hvor $\{y(t)\}$ er et system av noe høyere orden (i figur 8.2: \tilde{H}_P) er eksitert av hvit støy.

Vi trenger m.a.o. bare å utvikle reguleringsalgoritmer for systemer som er eksitert av hvit støy.

Vi skal nå se på en hyppig forekommende stokastisk prosess som kan betraktes som generert av hvit støy. Hvis vi antar $H(\omega) = K / (1 + j\omega T)$, altså et 1.ordens lavpassfilter med tidskonstant T, og sender rudimentær hvit støy inn på dette, får vi

$$\phi_{yy}(\omega) = \frac{K}{1 + j\omega T} \cdot \frac{K}{1 - j\omega T} \cdot 1 = \frac{K^2}{1 + \left(\frac{\omega}{\omega_0}\right)^2} \quad (8.17)$$

hvor $\omega_0 = 1/T$ er filterets båndbredde. Vi kaller gjerne $y(t)$ for **lavpassfiltrert hvit støy**.

Totaleffekten i $y(t)$ blir

$$E[y^2(t)] = \Phi_{yy}(0) = \frac{1}{\pi} \int_0^\infty \phi_{yy}(\omega) d\omega = \frac{1}{\pi} \int_0^\infty \frac{K^2}{1 + \left(\frac{\omega}{\omega_0}\right)^2} d\omega = \frac{K^2}{\pi} \omega_0 \operatorname{atan} \frac{\omega}{\omega_0} \Big|_0^\infty = \frac{K^2 \omega_0}{2} \quad (8.18)$$

Vi ser at utgangseffekten er proporsjonal med båndbredden ω_0 . Når $\omega_0 \rightarrow \infty$, blir $y(t)$ hvit og (8.18) bekrefter det vi visste - at effekten i hvit støy er uendelig.

Autokorrelasjonsfunksjonen finnes ved invers Fourier-transformasjon. Her kan du omgjøre $H(\omega)$ til en transferfunksjon $H(s)$ og bruke residueregning (se kapittel 1). Du må da ta hensyn til at $\Phi_{yy}(\tau)$ er et tosidig signal. Siden $\Phi_{yy}(\tau)$ er en like funksjon, kan du likevel nøye deg med å finne $\Phi_{yy}(\tau)$ for $\tau > 0$. Vi får

$$\Phi_{yy}(\tau) = \frac{K^2 \omega_0}{2} e^{-\omega_0 |\tau|} \quad (8.19)$$

Når $\omega_0 \rightarrow \infty$, går $\Phi_{yy}(\tau)$ mot $K^2 \delta(\tau)$ ¹, fordi arealet under (8.19) er invariant = K^2 .

Vi merker oss forøvrig at (8.19) har prinsipielt samme forløp som autokorrelasjonen til telegraftsignalet, jfr. avsnitt 6.3.4 i kapittel 6. M.h.p. middelverdi, varians, korrelasjon og effektspektrum, vil derfor et lineært system som eksiteres enten av et telegraftsignal eller lavpassfiltrert hvit støy, oppføre seg liketan.

De svært viktige spektrale relasjonene som vi har kommet frem til har to hovedanvendelser:

1. For det første design og analyse av reguleringsystemer hvor transferfunksjonene, eller systemet vi skal regulere, er kjent. Vi kan dermed beregne varians og spektral fordeling av ulike signaler og benytte dette til å finne egnede regulatorstrukturer og parametre i disse.
2. Vi kan dessuten benytte disse relasjonene til å finne et systems transferfunksjon (modell av systemet), i det vi fra likning (8.10) ser at hvis vi beregner anslag for Φ_{uy} og Φ_{uu} ut fra måleserier av $u(t)$ og $y(t)$, kan et anslag for (et **estimat** av) H beregnes fra formelen

$$H(\omega) = \frac{\Phi_{uy}(\omega)}{\Phi_{uu}(\omega)} \quad (8.20)$$

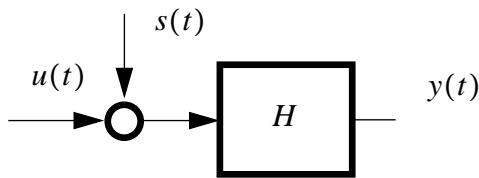
Vi skal i senere kapitler gå nærmere inn på **regulatordesign-problemet** i punkt 1. Vi skal i det følgende se på problemet i punkt 2, å finne et systems systemtransferfunksjon mer eller mindre direkte fra likning (8.20). Vi kan kalle dette et **systemidentifikasjonsproblem** eller et **estimeringsproblem**.

1. Denne eksponensielt formede symmetriske funksjonen gir faktisk en alternativ måte å definere deltafunksjonen på, i stedet for å definere den som grensetilfellet av en smal og høy firkantpuls.

8.2 Estimering av transferfunksjoner i systemer som er utsatt for støy

8.2.1 Pådragsstøy, prosess-støy

Anta gitt systemet i figur 8.3. Vi ser at vi i tillegg til det kjente inngangssignalet $u(t)$ også har et ikke målbart støysignal $s(t)$ som eksiterer prosessen. Vi kan for eksempel tenke oss en kjemisk prosess som mates med et pådrag i form av finknust materiale av ett eller annet slag. $u(t)$ representerer da det ønskede pådraget som vi setter ut til mateorganet, mens $s(t)$ representerer tilfeldige variasjoner i matemekanismen.



Figur 8.3 System med ønsket pådrag $u(t)$ pluss støy $s(t)$.

Det egentlige pådraget er da $x(t) = u(t) + s(t)$. Foldingsintegralet gir oss

$$y(t) = \int_{-\infty}^{\infty} h(\alpha)[s(t - \alpha) + u(t - \alpha)]d\alpha \quad (8.21)$$

hvor $h(\alpha)$ er impulsresponsen. Videre

$$\begin{aligned} \varphi_{uy}(\tau) &= E[u(t)y(t + \tau)] = E\left(\int_{-\infty}^{\infty} u(t)h(\alpha)[s(t + \tau - \alpha) + u(t + \tau - \alpha)]d\alpha\right) \\ &= \int_{-\infty}^{\infty} h(\alpha)[\varphi_{us}(\tau - \alpha) + \varphi_{uu}(\tau - \alpha)]d\alpha = h(\tau)^*[\varphi_{us}(\tau) + \varphi_{uu}(\tau)] \end{aligned} \quad (8.22)$$

I frekvensplanet blir dette

$$\phi_{uy}(\omega) = H(\omega)[\phi_{uu}(\omega) + \phi_{us}(\omega)] \quad (8.23)$$

Vi får derfor ved bruk av formelen (8.20)

$$\hat{H}(\omega) = H(\omega)\left[1 + \frac{\phi_{us}(\omega)}{\phi_{uu}(\omega)}\right] \quad (8.24)$$

hvor nå $\hat{H}(\omega)$ er et estimat av (et anslag for) $H(\omega)$.

Vi ser at estimatet blir korrekt hvis $s(t)$ og $u(t)$ er ukorrelerte.

Den tidligere definerte koherensfunksjonen blir (det er underforstått at forskjellige størrelser alle er funksjoner av ω , men dette utelates for enkelhets skyld):

$$\gamma_{uy}^2 = \frac{|\phi_{uy}|^2}{\phi_{uu}\phi_{yy}} = \frac{\phi_{uy}\phi_{uy}^*}{\phi_{uu}\phi_{yy}} = \frac{|H|^2(\phi_{uu} + \phi_{us})(\phi_{uu} + \phi_{su})}{\phi_{uu}\phi_{yy}} \quad (8.25)$$

(hvor vi har benyttet (8.23) og at $\phi_{su} = \phi_{us}^*$). Uttrykket kan igjen omformes til

$$\gamma_{uy}^2 = \frac{|H|^2\phi_{uu}(\phi_{uu} + \phi_{su} + \phi_{us} + \phi_{ss} - \phi_{ss}) + |H|^2\phi_{us}\phi_{su}}{\phi_{uu}\phi_{yy}} \quad (8.26)$$

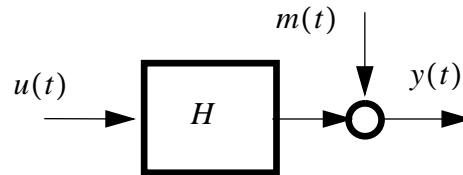
Nå er $\phi_{yy} = |H|^2\phi_{xx}$. Siden $\phi_{xx} = \phi_{ss} + \phi_{uu} + \phi_{su} + \phi_{us}$, blir
 $\phi_{yy} = |H|^2(\phi_{ss} + \phi_{uu} + \phi_{su} + \phi_{us})$. Vi benytter dette, og får

$$\gamma_{uy}^2 = 1 - \frac{|H|^2(\phi_{uu}\phi_{ss} - \phi_{us}\phi_{su})}{\phi_{uu}\phi_{yy}} \quad (8.27)$$

Vi så av (8.24) at estimatet $\hat{H}(\omega)$ kun er korrekt hvis $s(t)$ er ukorrelert med $u(t)$.

Da ser vi at koherensfunksjonen blir lik 1 minus signal/støy-forholdet mellom virkningen av s på utgangen y . Det kan enkelt vises (gjør det!) at hvis s er fullstendig gitt av u ved en transferfunksjon $G_{us}(\omega)$, så vil $\gamma_{uy}^2 = 1$. Dette illustrerer det faktum at koherensfunksjonen gir et mål for i hvor stor grad utgangssignalet er lineært avhengig av innsignalet.

8.2.2 Målestøy



Figur 8.4 Prosess med målestøy

La oss nå i stedet anta at vi har målestøy på utgangen som vist i figur 8.4.

Vi har at

$$y(t) = \int_{-\infty}^{\infty} h(\alpha)u(t-\alpha)d\alpha + m(t) \quad (8.28)$$

Dermed blir krysseffektspektret mellom u og y

$$\phi_{uy}(\omega) = H(\omega)\phi_{uu}(\omega) + \phi_{um}(\omega) \quad (8.29)$$

Estimatet $\hat{H}(\omega)$ av $H(\omega)$ etter likning (8.17) blir da

$$\hat{H}(\omega) = H(\omega) + \frac{\phi_{um}(\omega)}{\phi_{uu}(\omega)} \quad (8.30)$$

Koherensfunksjonen blir

$$\gamma_{xy}^2 = 1 - \frac{(\phi_{uu}\phi_{mm} - \phi_{um}\phi_{mu})}{\phi_{uu}\phi_{yy}} \quad (8.31)$$

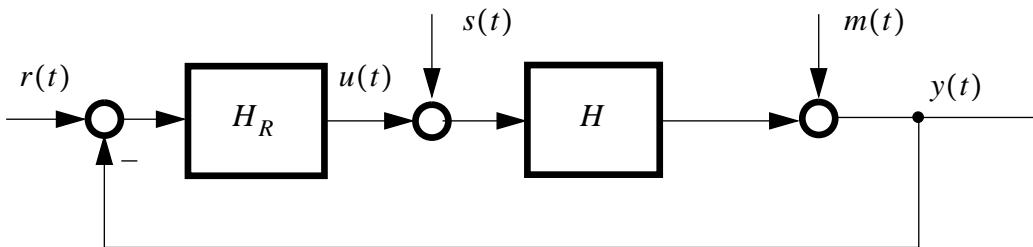
Vi ser igjen at $\hat{H}(\omega)$ gir et korrekt estimat av $H(\omega)$ hvis støyen er ukorrelert med pådraget.

Koherensfunksjonens betydning kan generelt sammenfattes slik: Kall systemets inngang x og utgangen y . Hvis $\gamma_{xy}^2 < 1$, gjelder en eller flere av tre mulige situasjoner:

- a. Det er støy på målingene.
- b. Det er støy i inngangssignalet $x(t)$ i tillegg til pådraget (det styrte signal).
- c. Systemet som gir $y(t)$ som funksjon av $x(t)$ er ikke lineært.

8.2.3 Støy i tilbakekoplede systemer

Anta at vi har et reguleringsssystem som vist i figur 8.5.



Figur 8.5 Reguleringsssystem med forstyrrelser.

Vi søker prosessens transferfunksjon $H(\omega)$, som er ukjent. Referansen er $r \equiv 0$, men systemet er eksitert av støykildene $s(t)$ og $m(t)$, og vi har på basis av måleserier av $u(t)$ og $y(t)$ beregnet estimatorer av $\phi_{uu}(\omega)$ og $\phi_{uy}(\omega)$.

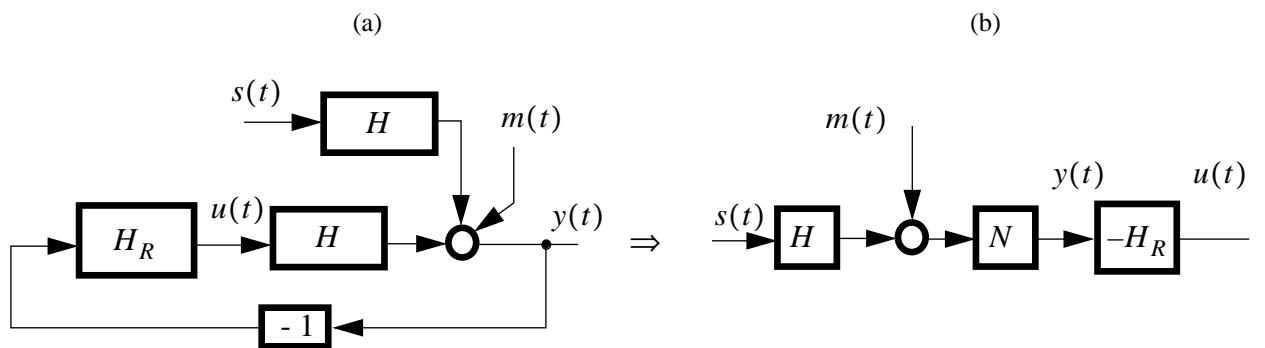
Vi søker prosessens transferfunksjon som (jfr. (8.20)):

$$H = \frac{\phi_{uy}}{\phi_{uu}} \quad (8.32)$$

Vi skal imidlertid vise at dette ikke går for det gitte systemet. Vi har reguleringsgraden

$$N = \frac{1}{1 + H_R H} \quad (8.33)$$

og kan da omforme systemet som vist i figur 8.6.



Figur 8.6

Ved hjelp av figur 8.6 b og likn. (8.10) får vi $\phi_{uy} = \phi_{yu}^* = -H_R^* \phi_{yy}$.

Videre har vi at $\phi_{uu} = |-H_R|^2 \phi_{yy} = H_R H_R^* \phi_{yy}$ (se (8.7) og (8.8)). Da blir

$$\frac{\phi_{uy}}{\phi_{uu}} = -\frac{1}{H_R} \quad (8.34)$$

Vi har fått et helt uakseptabelt resultat, estimatet av H er lik den negative inverse regulator-funksjonen. Dette skyldes **kausaliteten** (årsak/virkning) i systemet: Som det framgår av figur 8.6 b, er ϕ_{uy} eller riktigere, ϕ_{yu} , avhengig av H_R og ikke H . Intuitivt kan vi si at støykildene eksiterer u via y , og ikke y via u .

Generelt må man alltid være årvåken når man forsøker å estimere en transferfunksjon i en lukket sløyfe. En måte å unngå vanskelighetene på er å eksitere systemet ved å variere referansen slik at systemet hovedsakelig eksiteres fra referanseendringene og ikke fra m og s : Da H_R er kjent, kan H finnes, siden H inngår i

$$H_{ry} = \frac{H_R H}{1 + H_R H} \quad (8.35)$$

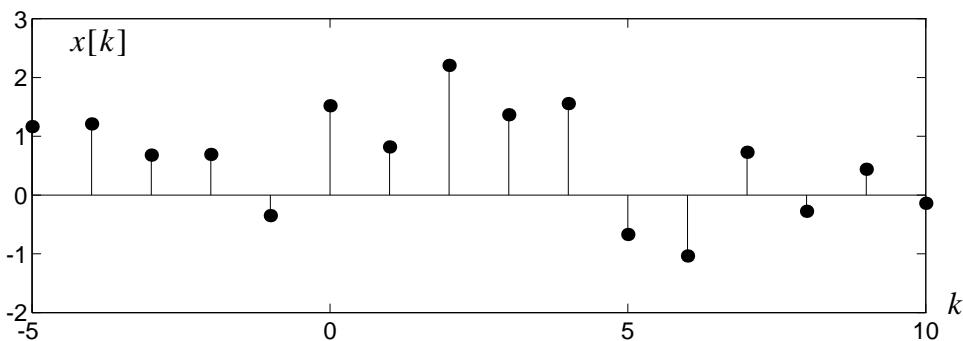
Vi må da først ha funnet et estimat av H_{ry} , via sammenhengen $\phi_{ry} = H_{ry} \phi_{rr}$.

8.3 Diskrete stokastiske prosesser og diskrete systemer

En diskret stokastisk prosess kan være fremkommet ved sampling av en kontinuerlig stokastisk prosess. Dette er ofte i praksis hvor en regulator eller et styresystem kan være digitalt realisert, mens prosessen i seg selv er kontinuerlig. Alle de grunnleggende statistiske begrepene og definisjonene av ulike egenskaper ved stokastiske prosesser gjelder også for diskrete prosesser. Dette gjelder begreper som middelverdi, korrelasjonsfunksjoner og ergodisitet. På samme måte som vi kan bruke Fourier- og Laplacetransformasjonen når vi studerer korrelasjon/effektspektrum for kontinuerlige prosesser, så kan vi anvende diskret Fourier- og z-transformasjon på diskrete stokastiske prosesser og deres korrelasjonsfunksjoner/effektspektra. Når det gjelder effektspektra får vi periodisitet, noe vi burde kunne forvente fra det vi vet om teorien for diskrete signaler og systemer.

8.3.1 Diskret prosess, korrelasjon, effektspektrum

Figur 8.7 viser en realisasjon av en diskret stokastisk prosess $\{x[k]\}$:



Figur 8.7

For diskrete stokastiske prosesser $\{x[k]\}$ og $\{y[k]\}$ har vi

$$\varphi_{xx}[n] = E(x[k]x[k+n]) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x[k]x[k+n] \quad (8.36)$$

$$\varphi_{xy}[n] = E(x[k]y[k+n]) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x[k]y[k+n] \quad (8.37)$$

Tidsmidlingen til høyre i (8.36) og (8.37) gjelder bare når prosessene er ergodiske.

Når det gjelder overganger fra korrelasjonsfunksjoner til effektspektra, gjelder diskret Fourier- og z-transformasjon, og vi kan umiddelbart sette opp sammenhengene:

Autokorrelasjon:

$$\phi_{xx}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{xx}(\Omega) e^{j\Omega n} d\Omega \quad (8.38)$$

Effektspektrum:

$$\phi_{xx}(\Omega) = \sum_{n=-\infty}^{\infty} \phi_{xx}[n] e^{-j\Omega n} \quad (8.39)$$

Krysskorrelasjon:

$$\phi_{xy}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{xy}(\Omega) e^{j\Omega n} d\Omega \quad (8.40)$$

Krysseffektspektrum:

$$\phi_{xy}(\Omega) = \sum_{n=-\infty}^{\infty} \phi_{xy}[n] e^{-j\Omega n} \quad (8.41)$$

Vi har videre de svært viktige ligningene for forholdene mellom inngang og utgang for et system som utsettes for et stokastisk signal

$$\phi_{xy}(\Omega) = H(\Omega)\phi_{xx}(\Omega) \quad (8.42)$$

$$\phi_{yy}(\Omega) = H(\Omega)H(-\Omega)\phi_{xx}(\Omega) = H(\Omega)H^*(\Omega)\phi_{xx}(\Omega) = |H(\Omega)|^2\phi_{xx}(\Omega) \quad (8.43)$$

De svarer til ligningene (8.10) og (8.7)/(8.8) for kontinuerlige stokastiske signaler.

Vi har videre

$$\phi_{xy}[n] = \sum_{k=-\infty}^{\infty} \phi_{xx}[n-k]h[k] = \phi_{xx}[n]^*h[n] \quad (8.44)$$

$$\phi_{yy}[n] = \sum_{k=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} h[k]h[i]\phi_{xx}[n+k-i] \quad (8.45)$$

som svarer til henholdsvis (8.9) og (8.4). Integrasjon er erstattet av summasjon.

8.3.2 Diskret hvit støy

Vi definerer en **diskret** hvit støy-prosess ved

$$\phi_{vv}[n] = \begin{cases} Q_D, & n = 0 \\ 0, & n \neq 0 \end{cases} \quad (8.46)$$

Dette innebærer at $E(v[n]) = 0$ og at **effekten** $E(v^2[n]) = \phi_{vv}[0] = Q_D$. Effekten blir altså ikke lenger uendelig, slik som ved kontinuerlig hvit støy.

Vi kan nå finne effektspektret til diskret hvit støy ved å diskret-Fourier-transformere (8.46), dvs. vi bruker (8.39).

$$\phi_{vv}(\Omega) = \sum_{n=-\infty}^{\infty} \phi_{vv}[n] e^{-j\Omega n} = \phi_{vv}[0] e^{-j\Omega_0} = Q_D \quad (8.47)$$

Som forventet får diskret hvit støy et konstant autoeffektspektrum.

Vi kan nå kontrollere at effekten i et diskret hvit støy-signal blir det vi skulle forvente (jfr. (8.38)):

$$\phi_{vv}[0] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{vv}(\Omega) e^{j\Omega_0} d\Omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q_D d\Omega = Q_D \quad (8.48)$$

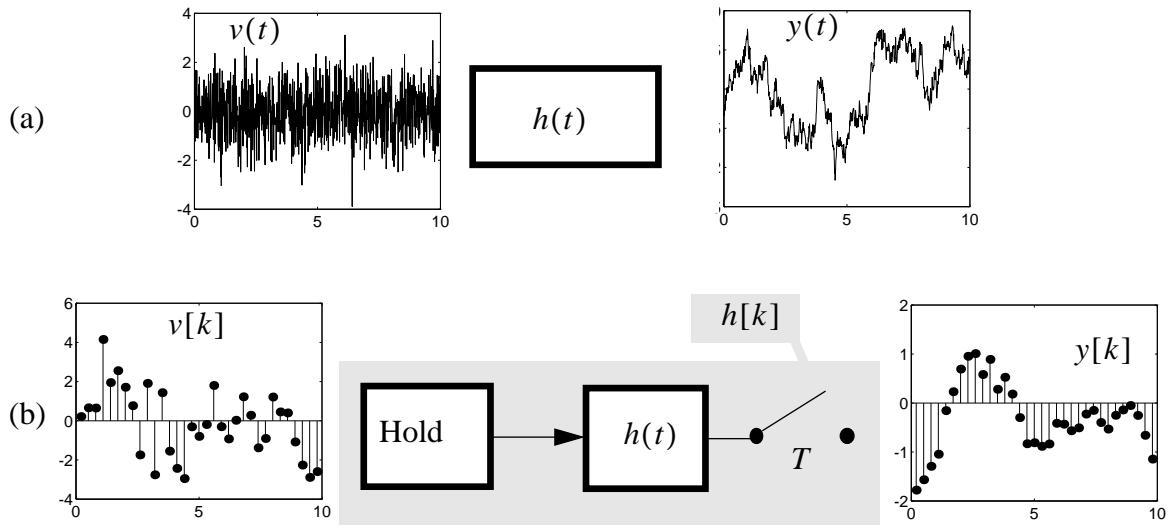
8.3.3 Diskretisering når kontinuerlig hvit støy påvirker et kontinuerlig lineært system

Vi skal se på hvordan vi må velge varians for en diskret hvitstøy-prosess, når det kreves at denne prosessen skal ha samme virkning på et diskretisert kontinuerlig system som det en gitt kontinuerlig hvitstøy-prosess har på det samme kontinuerlige system. Problemstillingen er skissert i figur 8.8:

Gitt en kontinuerlig hvitstøy-prosess $v(t)$, med $E[v(t)] = 0$ og $\phi_{vv}(\tau) = Q\delta(\tau)$. Denne prosessen påvirker systemet slik som vist i figur 8.8 a. Det kontinuerlig systemet er definert ved impulsresponsen $h(t)$. Systemet diskretiseres og er da gitt ved sin diskrete impulsrespons $h[k]$. Vi ser umiddelbart at $E(y[k]) = E(y(t)) = 0$.

Vi ønsker nå å bestemme Q_D slik at $v[k]$ definert i (8.46) får samme virkning på det diskretiserte systemet gitt ved $h[k]$, som $v(t)$ har på det kontinuerlige systemet gitt av $h(t)$.

Dette betyr å sørge for at variansen på utgangen blir den samme, da er oppgaven løst.



Figur 8.8

For å løse problemet må vi først se på sammenhengen mellom $h(t)$ og $h[k]$: Et monovariabelt lineært kontinuerlig system har altså impulsresponsen $h(t)$. Ved tastetidspunktene får vi utgangsverdiene $y(t) = h(kT)$, der T er tasteintervallet og k angir diskret tid $0, 1, 2, \dots$. Merk at vi her snakker om den *kontinuerlige* impulsreponsen, men dens verdier ved tastetidspunktene.

Hva blir så sammenhengen mellom den *diskrete* impulsresponsen $h[k]$ og den kontinuerlige $h(t)$ ved tastetidspunktene $t = 0, T, 2T, \dots, kT, \dots$? Vi skal nøyne oss med en tilnærmet betrakting, vi betrakter situasjonen for "små" T , dvs. $T \ll 1/|\lambda_i|$, der λ_i representerer det lineære systemets raskeste modus. Vi har med andre ord "hurtig" tasting.

Hvis vi nå sender inn en diskret puls med amplituden $1/T$ på inngangen til systemet $h[k]$ i figur 8.8 (b), vil holdelementet holde denne amplituden i T tidsenheter. Med andre ord vil holdelementet utsette det kontinuerlige systemet for et inngangssignal som er tilnærmet lik en Dirac-puls $\delta(t)$.

Vi får dermed for små T :

$$\frac{1}{T}h[k] \approx h(kT) \Leftrightarrow h[k] \approx Th(kT) \quad (8.49)$$

Merk her at $h[k]$ ikke er lik $h(kT)$! Dette henger sammen med *definisjonen* på en diskret enhetspuls og diskret impulsrespons: Den diskrete enhetspulsen har amplituden 1, uansett tasteintervallets lengde. Med andre ord vil $h[k] \rightarrow 0 \quad \forall k \text{ når } T \rightarrow 0$.

Vi skal nøye oss med å finne Q_D under forutsetning av at T er “liten”, slik som nevnt over. Vi har

$$\begin{aligned} E(y^2[k]) &= E\left(\sum_{j=-\infty}^{\infty} h[j]v[k-j] \sum_{n=-\infty}^{\infty} h[n]v[k-n]\right) \\ &= \sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h[j]h[n]E(v[k-n]v[k-j]) = \sum_{j=-\infty}^{\infty} h[j] \sum_{n=-\infty}^{\infty} h[n]Q_D\delta[j-n] \\ &= Q_D \sum_{j=-\infty}^{\infty} h[j] \sum_{n=-\infty}^{\infty} h[n]\delta[j-n] = Q_D \sum_{j=-\infty}^{\infty} h^2[j], \text{ evt. } Q_D \sum_{j=0}^{\infty} h^2[j] \end{aligned} \quad (8.50)$$

I siste tilfellet er systemet kausalt, nedre summasjonsgrense er $j=0$ fordi $h[j] = 0$ for $j < 0$. Men dette er et spesialtilfelle, (8.50) gjelder også for ikke-kausale systemer.

Før et (kausalt) kontinuerlig system får vi, via foldingsintegralet og lignende mellomregninger som ikke gjennomføres her

$$E(y^2(t)) = Q \int_0^{\infty} h^2(t)dt \quad (8.51)$$

Resultatene (8.50) og (8.51) er viktige i seg sjøl, så merk deg dem!

For små T kan vi sette $h[j] \approx Th(jT)$, jfr. (8.49). Vi setter inn for $h[j]$ i siste del av (8.50) og setter (8.50) = (8.51). Dette skal gi oss den ukjente Q_D som funksjon av Q og T .

$$Q_D \sum_{j=0}^{\infty} h^2[j] \approx Q_D \sum_{j=0}^{\infty} T^2 h^2(jT) = Q \int_0^{\infty} h^2(t)dt \quad (8.52)$$

Når T er liten, har vi dessuten

$$\int_0^{\infty} h^2(t)dt \approx \sum_{j=0}^{\infty} h^2(jT)T \quad (8.53)$$

(8.53) innsatt i (8.52) gir

$$Q \approx Q_D T \Rightarrow Q_D \approx \frac{Q}{T} \quad (8.54)$$

Altså: Jo kortere diskretiseringsintervall, jo mer må vi øke variansen på den diskrete hvite støyen for å få samme varians på systemutgangen som for det kontinuerlige tilfellet.

Hvis T er stor i forhold til systemets tidskonstanter, vil tilnærmlsen at Q_D bare avhenge av T , slik som i (8.54), ikke gjelde lenger. Da må vi ha med at Q_D også avhenger av systemets parametre. Det blir da mer komplisert å finne den Q_D som gir $E(y^2[k]) = E[y^2(t)]$.

8.4 Diskrete tilstandsrommodeller av monovariable stokastiske dynamiske systemer

Vi har hittil i dette kapittelet betraktet diskrete impulsresponser og diskrete transferfunksjoner som modeller for et diskret lineært én-inn-én-ut-system. Vi skal se på modelleringssproblemet på tilstandsromform. Fra kapittel 3, ligning (3.4) vet vi at

$$\begin{aligned} \underline{x}[k+1] &= \phi \underline{x}[k] + \Delta u[k] \\ y[k] &= D \underline{x}[k] \end{aligned} \quad (8.55)$$

representerer et monovariabelt system med inngang $u[k]$ og utgang $y[k]$. La oss nå utvide (8.55) til et stokastisk system

$$\begin{aligned} \underline{x}[k+1] &= \phi \underline{x}[k] + \Delta_u u[k] + \Delta_v v[k] \\ y[k] &= D \underline{x}[k] + w[k] \end{aligned} \quad (8.56)$$

hvor $v[k]$ og $w[k]$ er hvitstøy-prosesser, hhv. "prosess-støy" med $\phi_{vv}(\Omega) = V$ og "målestøy" med $\phi_{ww}(\Omega) = W$.

Ved z-transformasjon kan vi skrive (jfr. (3.8) i kapittel 3)

$$y(z) = D(zI - \phi)^{-1} \Delta_u u(z) + D(zI - \phi)^{-1} \Delta_v v(z) + w(z) \quad (8.57)$$

der vi med utgangspunkt i (3.8) -(3.11) i kapittel 3 har

$$D(zI - \phi)^{-1} \Delta_u = \frac{B(z)}{A(z)} = \frac{B^*(z^{-1})}{A^*(z^{-1})} \quad (8.58)$$

Høyre ledd framkommer ved å dele på z^n i teller og nevner¹. $A^*(z^{-1})$ gjenkjenner vi som A^* -polynomet i ARMAX-modellen. Vi kan skrive

$$y(z) = H^*(z^{-1})u(z) + H_v^*(z^{-1})v(z) + w(z) \quad (8.59)$$

$A^*(z^{-1})$ er også nevneren i $H_v^*(z^{-1})$, jfr. (8.57). Hvis vi nå antar $u[k] \equiv 0$, for å kunne studere virkningen av de to støyprosessene på utgangen, får vi

$$y(z) = H_v^*(z^{-1})v(z) + w(z) = \frac{P^*(z^{-1})v(z) + A^*(z^{-1})w(z)}{A^*(z^{-1})} \quad (8.60)$$

der $P^*(z^{-1})$ er et polynom som vi ikke trenger å regne ut her.

1. Vi bruker i dette underavsnitt *-symbolet for å markere funksjoner av z^{-1} . Som nevnt tidligere, trenger vi vanligvis ikke gjøre dette, når alle funksjoner er av z^{-1} , slik at dette er underforstått.

Antar vi for enkelhets skyld at $v[k]$ og $w[k]$ er innbyrdes ukorrelerte (denne antagelsen er ikke nødvendig, men jeg gjør den for å forenkle framstillinga her), så får vi ved hjelp av (8.43) og (8.60)

$$\phi_{yy}(\Omega) = |H_v(\Omega)|^2 \phi_{vv}(\Omega) + \phi_{ww}(\Omega) = |H_v(\Omega)|^2 V + W \quad (8.61)$$

(8.61) er et rasjonalt uttrykk¹ i $e^{-j\Omega} = z^{-1}$. Dette uttrykket vil alltid kunne faktoriseres (jfr. (8.16) som gjelder for kontinuerlige signaler, men tilsvarende gjelder altså i det diskrete tilfelle sjøl om det ikke vises her).

Faktorisering betyr at vi kan finne et ekvivalent uttrykk $H_\epsilon(\Omega)$ slik at

$$\phi_{yy}(\Omega) = |H_v(\Omega)|^2 V + W = |H_\epsilon(\Omega)|^2 = \left| \frac{C(\Omega)}{A(\Omega)} \right|^2 = \left| \frac{C^*(z^{-1})}{A^*(z^{-1})} \right|^2 \quad (8.62)$$

Høyre ledd framkommer når vi setter $e^{-j\Omega} = z^{-1}$.

Betydningen av (8.62) er at vi får det samme effektspektrum på utgangen hvis $v[k]$ og $w[k]$ erstattes av en ny hvit-støy-kilde $\epsilon[k]$ med effektspektrum $\phi_{\epsilon\epsilon}(\Omega) = 1$, som påvirker gjennom en modell

$$y[k] = \frac{C^*(z^{-1})}{A^*(z^{-1})} \epsilon[k] \quad (8.63)$$

Med andre ord: Polynomet $C^*(z^{-1})$ vil være gitt av autoeffektspektra V , W og koeffisientene i $P^*(z^{-1})$ og $A^*(z^{-1})$, jfr. siste ledd til høyre i (8.60) og (8.62).

Ved å innføre $C^*(z^{-1})$ får vi, når vi nå igjen lar pådraget være forskjellig fra 0, og benytter superposisjonsprinsippet som alltid gjelder for lineære systemer:

$$A^*(z^{-1})y[k] = B^*(z^{-1})u[k] + C^*(z^{-1})\epsilon[k] \quad (8.64)$$

Dermed ser vi at (8.64) er ekvivalent med (8.56) på den måten at inngangs/utgangs-egenskapene er identiske og slik at autoeffektspektret til $y[k]$ er det samme i de to tilfellene.

Vi gjenkjenner (8.64) som ARMAX-modellen (kapittel 4, formel (4.28)).

Det at vi nå har tatt vare på både prosesstøyens og målestøyens virkning i en tilstandsrom-modell, ved hjelp av en enklere ARMAX-modell med bare en ekvivalent støykilde, er mye av forklaringen på ARMAX-modellen sentrale rolle.

Effektspektret på systemets utgang når $u[k] \equiv 0$ er altså gitt av (8.62).

(8.62) og (8.63) forteller at et vilkårlig effektspektrum kan dannes ved å la hvit støy passere et lineært filter.

1. Egentlig har vi som tidligere påpekt uttrykk av typen $\phi_{yy}(e^{-j\Omega})$, men skriver dem for enkelhets skyld som $\phi_{yy}(\Omega)$ etc.

Kapittel 9: Estimering av spektra, korrelasjons- og transferfunksjoner

Vi har tidligere vist hvordan transferfunksjoner kan estimeres (dvs.: anslås) som et forhold mellom estimatorer for krysseffektspektret og autoeffektspektret:

$$\hat{H}(\Omega) = \frac{\hat{\phi}_{xy}(\Omega)}{\hat{\phi}_{xx}(\Omega)} \quad (9.1)$$

hvor x er inngang og y er utgang. Denne prosedyren krever kjennskap til effektspektrene, og vi skal se hvordan disse kan estimeres ved hjelp av **endelig diskret Fourier-transformasjon**, forkortet **“DFT”**, som er den numeriske, datamaskin-realiserbare versjon av den diskrete Fouriertransformasjonen. Vi må derfor først definere denne:

9.1 Diskret Fourier-transform (“DFT”) av endelig måleserie

Fourier-transformasjon er et meget effektivt hjelpemiddel for estimering av korrelasjonsfunksjoner, spektra og transferfunksjoner. I tillegg kan det ofte være fordelaktig å filtrere signaler ved å utnytte det faktum at konvolusjon (folding) i tidsplanet motsvarer multiplikasjon i frekvensplanet. Bruk av Fouriertransformasjonsteknikk har fått en særlig sterkt økning etter at Cooley og Tukey i 1965 publiserte en avhandling over en hurtig algoritme for beregning av en DFT. Denne algoritmen kalles **Fast Fourier Transform - FFT**. FFT skal ikke behandles her.

Vi skal nå definere den diskrete Fouriertransformasjonen av en endelig N -punkters måleserie. Det gjør vi med utgangspunkt i den ideelle diskrete Fouriertransformasjonen (2.14), som er basert på at man har en uendelig lang tidsserie tilgjengelig:

$$X(\Omega) = \sum_{k=-\infty}^{\infty} x[k] e^{-j\Omega k} \quad (9.2)$$

$$\text{Inverstransformasjonen er } x[k] = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\Omega) e^{j\Omega k} d\Omega = \frac{1}{2\pi} \int_0^{2\pi} X(\Omega) e^{j\Omega k} d\Omega \quad (9.3)$$

(Vi kan velge integrasjonsgrensene hvor vi vil p.g.a. integrandens periodisitet, bare intervallets lengde er 2π). Vi approksimerer integralet med en sum over N ekvidistante punkter i frekvensplanet: $d\Omega$ erstattes med $\Delta\Omega = \frac{2\pi}{N}$. Vi definerer $\Omega_n = \frac{2\pi}{N}n$ og $X_n = X(\Omega_n)$.

Merk at nå får vi også en summasjonsvariabel, n , i frekvensplanet, i tillegg til k i tidsplanet.

I dette avsnittet vil vi ofte bruke subskript, f.eks . X_n og x_k i stedet for $X[n]$ og $x[k]$. Begge varianter finnes i litteraturen, men betydningen er den samme.

(9.3) erstattes nå med tilnærmelsen

$$x[k] = \frac{1}{2\pi} \sum_{n=0}^{N-1} X_n e^{j\Omega_n k} \Delta\Omega = \frac{1}{2\pi} \sum_{n=0}^{N-1} X_n e^{j\Omega_n k} \left(\frac{2\pi}{N}\right) = \frac{1}{N} \sum_{n=0}^{N-1} X_n e^{j\frac{2\pi kn}{N}} \quad (9.4)$$

Betrakter vi så den ideelle diskrete Fourier-transformasjonen (9.2) vil den for en endelig måleserie med N punkter måtte erstattes med

$$X_n = \sum_{k=0}^{N-1} x[k] e^{-j\frac{2\pi kn}{N}} \quad (9.5)$$

(9.4) og (9.5) er altså utledet fra (9.3) og (9.2), og basert på N punkter i henholdsvis frekvens- og tidsplan.

Det gjenstår nå å vise at (9.4) og (9.5) er transformasjonspar: Vi setter inn (9.4) for $x[k]$ i (9.5). Dette gir

$$\sum_{k=0}^{N-1} \left(\frac{1}{N} \sum_{m=0}^{N-1} X_m e^{\frac{2\pi jm}{N} k} \right) e^{-\frac{2\pi jn}{N} k} = \frac{1}{N} \sum_{m=0}^{N-1} X_m \sum_{k=0}^{N-1} e^{(2\pi j/N)(m-n)k} \quad (9.6)$$

Nå er

$$\sum_{k=0}^{N-1} e^{(2\pi j/N)(m-n)k} = \begin{cases} N & \text{for } m = n \\ 0 & \text{ellers} \end{cases} \quad (9.7)$$

Tilfellet $m = n$ er opplagt. Når $m - n = p \neq 0$, kan summen i (9.7) betraktes som en sum av N enhetsvektorer fra origo. Endepunktene til vektorene utgjør et regulært polygon innskrevet i enhetssirkelen. Dermed må vektorsummen alltid bli null når $p \neq 0$. Vi får derfor at (9.6) forenkles til

$$\frac{1}{N} X_n N = X_n \quad (9.8)$$

Vi har nå vist at (9.4) og (9.5) er transformasjonspar når tidsserien har N punkter. Vi har nå formler som kan benyttes når vi i praksis vil Fourier-transformere en diskret tidsserie. Denne praktiske diskrete Fourier-transformasjonen vil vi som nevnt referere til med forkortelsen DFT, i motsetning til den ideelle diskrete Fourier-transformasjonen.

Merknad 1:

$\Omega_n = 2\pi n / N$ svarer til en dimensjonsløs samplingstid $T = 1$. Med fysisk tastetid T , vil den dimensjonsløse frekvensen Ω_n svare til den fysiske frekvensen

$$\omega_n = \frac{\Omega_n}{T} = \frac{2\pi n}{NT} \quad (9.9)$$

Merknad 2:

Studerer vi (9.4), ser vi at $x[k]$ blir periodisk med periode N . *DFT* medfører altså at en endelig diskret tidsserie *matematisk sett* blir å betrakte som **en** periode i en periodisk diskret tidsserie. *Derfor er egentlig DFT det samme som den diskrete Fourierrekke, på engelsk discrete Fourier series, DFS.* Men for våre formål er det mest interessant å bruke (9.4) og (9.5) på en tidsserie over N punkter, og å se bort fra at dette matematisk sett er å betrakte som **en** periode av en periodisk tidsserie. Denne periodisiteten skaper ingen komplikasjoner hvis man er klar over forholdet. I avsnitt (9.3) behandles dette nærmere.

9.2 Estimering av spektra

Med redskapet DFT definert kan vi nå gå løs på oppgaven med estimering av spektra. Man kunne tenke seg alternative metoder for estimering av effektspektret for en ergodisk prosess. Man kan for eksempel estimere autokorrelasjonsfunksjonen ved tidsmidling og deretter Fourier-transformere den estimerte autokorrelasjonsfunksjonen. Ved bruk av FFT (en ekstremt hurtig måte å beregne en DFT på) viser det seg imidlertid at direkte Fouriertransformasjon av en måleserie og bruk av en diskretisert versjon av ligning (7.5), kapittel 7, er å foretrekke. Ligning (7.5) var:

$$\phi_T(\omega) = \frac{1}{2T} X_T(\omega)X_T(-\omega) \quad (9.10)$$

Denne ligninga uttrykte et anslag over et effektspektrum basert på *en* realisasjon, som i tillegg var avkortet i begge ender. Siden vår tidsserie ikke er kontinuerlig, men diskret, og siden den er N punkter lang og forutsatt $= 0$ for $k < 0$ og $k > N - 1$, er det naturlig å prøve en diskretisert versjon av (9.10)

$$\phi_N(\Omega) = \frac{1}{N} X_N(\Omega)X_N(-\Omega) \quad (9.11)$$

hvor

$$X_N(\Omega) = X(\Omega_n) = X_n = \sum_{k=0}^{N-1} x[k] e^{-j \frac{2\pi n}{N} k} \quad (9.12)$$

(Merk at vi i dette avsnittet bytter ut notasjonen “ X_n ” for DFT ved frekvensen Ω_n , med “ $X_N(\Omega)$ ”. Dette er fordi vi i det følgende har behov for å angi lengden N på tidsserien som ligger til grunn for DFT'en. Vi viser heller ikke indeks i Ω_n , da den kan ansees underforstått. Symbolet “n” vil i dette avsnitt i stedet bli anvendt for å markere diskret tid, sammen med “k”).

Spørsmålet er nå om $\phi_N(\Omega)$ er et “godt” estimat av $\phi(\Omega)$, som er det virkelige autoeffekt-

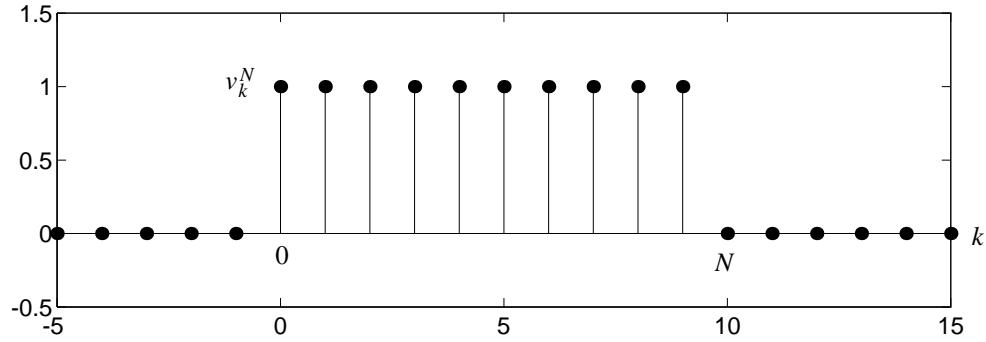
spektret til $x[k]$ (vi kommer snart til hva som er våre kriterium for “godt estimat”). Vi skal nå finne ut hvilken skade det gjør at realisasjonen $x[k]$ er forkortet i begge ender, i stedet for å være uendelig lang, slik som i det ideelle, men uoppnåelige tilfelle. Vi har

$$\begin{aligned}\phi_N(\Omega) &= \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{+j\Omega k} \sum_{l=0}^{N-1} x_l e^{-j\Omega l} \\ &= \frac{1}{N} \sum_{k=-\infty}^{\infty} v_k^N x_k e^{+j\Omega k} \sum_{n=-\infty}^{\infty} v_{n+k}^N x_{n+k} e^{-j\Omega(n+k)}\end{aligned}\quad (9.13)$$

hvor n er valgt slik at $l = n + k$, og v_k^N er en funksjon gitt av :

$$v_k^N = \begin{cases} 1; & 0 \leq k \leq N-1 \\ 0 & \text{ellers} \end{cases}, \quad (9.14)$$

(se figur 9.1).



Figur 9.1 Funksjonen v_k^N .

Likning (9.13) kan omformes til

$$\phi_N(\Omega) = \frac{1}{N} \sum_{k=-\infty}^{\infty} \left[\sum_{n=-\infty}^{\infty} v_k^N v_{k+n}^N x_k x_{k+n} e^{-j\Omega n} \right] \quad (9.15)$$

Vi tar forventningsverdien til $\phi_N(\omega)$ og får, når vi flytter forventningsoperasjonen innafor summasjonene slik at den bare omfatter de størrelser som er stokastiske,

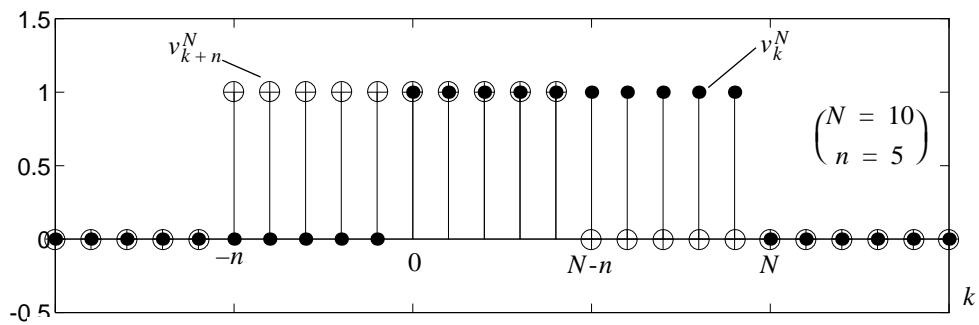
$$E[\phi_N(\Omega)] = \frac{1}{N} \sum_{n=-\infty}^{\infty} \left[\sum_{k=-\infty}^{\infty} v_k^N v_{k+n}^N \right] \phi[n] e^{-j\Omega n} \quad (9.16)$$

v_k^N og v_{k+n}^N er vist i figur 9.2. Fra figuren ser vi at summering over k gir $(N - |n|)$ bidrag $\neq 0$, slik at

$$E[\phi_N(\Omega)] = \sum_{-\infty}^{\infty} w_n^N \phi[n] e^{-j\Omega n} \quad (9.17)$$

hvor

$$w_n^N = \frac{N - |n|}{N} = 1 - \frac{|n|}{N} \quad \text{for } |n| < N \text{ og } w_n = 0 \text{ ellers.} \quad (9.18)$$



Figur 9.2 Fremstilling av v_k^N og v_{k+n}^N .

$E[\phi_N(\Omega)]$ er den diskrete Fourier-transformen av produktet mellom en triangulær funksjon og autokorrelasjonsfunksjonen. Vi ser at $E[\phi_N(\Omega)]$, gitt av (9.17), på grunn av w_n^N ikke blir **forventningsrett** fordi den er forskjellig fra den diskrete Fouriertransform av den sanne autokorrelasjon (som jo er det samme spektrum). Matematisk kan vi si det slik

$$E[\phi_N(\Omega)] \neq \sum_{-\infty}^{\infty} \phi[n] e^{-j\Omega n} = \phi(\Omega) \quad (9.19)$$

$\phi_N(\Omega)$ er et **estimat** av $\phi(\Omega)$ og kalles et **periodogram**. Funksjonen w_n^N kalles en **vindusfunksjon**, fordi vi kan tenke oss at den opprinnelige autokorrelasjonsfunksjonen $\phi[n]$ omformes ved å multipliseres med “vinduet” w_n^N .

Tilsvarende kan vi si at den uendelig lange tidsserien x_n endres til en endelig serie x_0, \dots, x_{N-1} ved å passere vinduet v_n^N .

Merknad:

Uttrykket $Nw_n = \sum_{k=-\infty}^{\infty} v_k^N v_{k+n}^N$, jfr. (9.16), kan skrives som

$$Nw_n = \sum_{k=-\infty}^{\infty} v_k^N v_{k-n}^N = \sum_{k=-\infty}^{\infty} v_k^N \bar{v}_{n-k}^N = v_n^N \bar{v}_n^N \quad (9.20)$$

der $\bar{v}_k^N = v_{-k}^N$, dvs. $\bar{v}_k^N = 1$ for $-(N-1) \leq k \leq 0$, $\bar{v}_k^N = 0$ ellers.

Dette skal vi bruke nedenfor.

9.2.1 Om forventningsfeilen til estimatet av effektspektret

Fra (9.19) ser vi at estimeringsprosedyren gir en forventningsfeil i $\phi_N(\Omega)$, estimatet av $\phi(\Omega)$. La oss se litt nærmere på dette. Fra (9.17) har vi at $E[\phi_N(\Omega)]$ og $(w_n^N \phi[n])$ er Fourier-transformpar. Vi vet videre at multiplikasjon i tidsplanet tilsvarer folding i frekvensplanet.¹ Dermed får vi

$$E[\phi_N(\Omega)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} W^N(\Omega - \lambda) \phi(\lambda) d\lambda \quad (9.21)$$

hvor $W^N(\Omega)$ er den diskret Fourier-transformerte til w_n^N , som blir

$$W^N(\Omega) = \frac{1}{N} \left[\frac{\sin(N\Omega/2)}{\sin(\Omega/2)} \right]^2 \quad (9.22)$$

Det gjenstår å utlede (9.22). Fra (9.20) har vi:²

$$F(v_n^N * \bar{v}_n^N) = NW^N(\Omega) = F(v_n^N)F(\bar{v}_n^N) \quad (9.23)$$

“Folding i tidsplanet svarer til multiplikasjon i frekvensplanet”. Vi må nå først finne $F(v_n^N)$:

$$F(v_n^N) = \sum_{n=0}^{N-1} e^{-j\Omega n} = \frac{1 - e^{-j\Omega N}}{1 - e^{-j\Omega}} = \frac{e^{j\Omega N/2} - e^{-j\Omega N/2}}{e^{j\Omega/2} - e^{-j\Omega/2}} \left(\frac{e^{-j\Omega N/2}}{e^{-j\Omega/2}} \right) = \frac{\sin(N\Omega/2)}{\sin(\Omega/2)} e^{-j\frac{(N-1)\Omega}{2}} \quad (9.24)$$

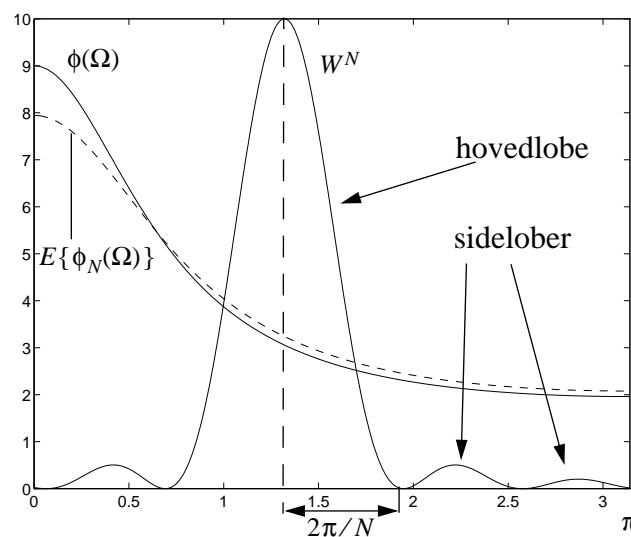
1. Se avsnitt 5.7 side 333 i “Signals and Systems”.

2. Vi bruker her symbolene “ F ” og “ F^{-1} ”, sjøl om dette vanligvis gjelder for kontinuerlig Fourier-transform. “ F ” betyr “DFT” når det framgår av sammenhengen.

Siden $\bar{v}_n^N = v_{-n}^N$, vil $F(\bar{v}_n^N) = [F(v_n^N)]^*$. Nå er $F(v_n^N)[F(v_n^N)]^* = |F(v_n^N)|^2$, og vi får

$$W^N(\Omega) = \frac{1}{N}|F(v_n^N)|^2 = \frac{1}{N} \left[\frac{\sin(N\Omega/2)}{\sin(\Omega/2)} \right]^2 \quad (9.25)$$

W^N er vist i figur 9.3. Av(9.21) følger at topper i $\phi(\Omega)$ blir midlet ut (smurt utover) på grunn av foldingen. Graden av "utsmøring" avhenger av (halve) bredden på **hovedloben** i $W^N(\Omega)$ som er gitt av $\Delta\Omega = 2\pi/N^1$, og av de tilsvarende bredder på **sidelobene**. Av dette kan vi konkludere at forventningsfeilen i $\phi_N(\Omega)$ avtar når N øker.



Figur 9.3 Virkelig spektrum $\phi(\Omega)$ og vindusfunksjon i frekvensplanet.

Fra figur 9.3 ser vi at sidelobene til W^N er relativt store og derfor bidrar betydelig til forventningsfeilen i $\phi_N(\Omega)$. Dette forholdet kan forbedres ved å velge en annen veiing av måleserien enn den rektangulære avkortingen definert ved v_n^N i (9.14). Dette vil endre W^N , fordi den er et resultat av v_n^N foldet med seg selv og Fourier-transformert, jfr. (9.25).

Hvis man velger v_n^N liten i hver ende av tidsserien og stor på midten, vil W^N få mye mindre sidelober, og forventningsfeilen blir redusert. En populær type vindusfunksjoner er

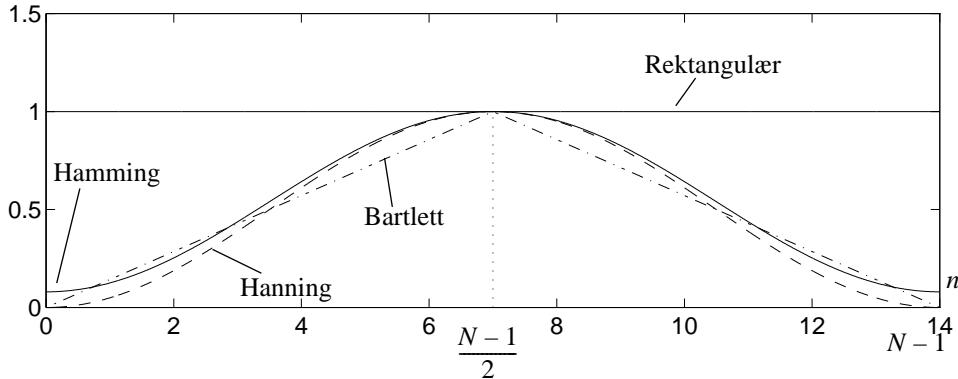
$$v_n^N = \frac{1}{\sqrt{\frac{3}{2}\alpha^2 - \alpha + \frac{1}{2}}} \left[\alpha - (1 - \alpha) \cos\left(\frac{2\pi n}{N-1}\right) \right] \quad (9.26)$$

For parameteren $\alpha = 0.50$ kalles vinduet et Hanning-vindu, og for $\alpha = 0.54$ et Hamming-vindu. Disse vinduene, sammen med noen andre, er vist i figur 9.4. Høyden på alle vinduer er i denne figuren normalisert = 1.

1. som vi også gjenkjenner som frekvensoppløsningen, jfr. (9.4)

Amplituden for $W^N = \frac{1}{N} |F(v_n^N)|^2$ for noen av disse vinduene er vist i figur 9.5, både i

lineær og logaritmisk (dB-) skala. Vi ser at Hamming-vinduet gir en W^N med svært små sidelober: Første sidelobe ligger 80 dB under hovedloben. Prisen man må betale er at hovedloben til W^N blir bredere når man benytter et Hamming-vindu i stedet for rektangulær avkortning.



Figur 9.4

(9.26) inneholder en multiplikativ faktor som er > 1 . Dette er nødvendig, for hvis vi bare reduserer amplituden til tidsserien på endene uten å kompensere ved å skalere opp midten, vil forventet midlere effekt i signalet bli for lav (slik er det i figur 9.4). I frekvensplanet innebærer dette at arealet under periodogrammet ikke blir forventningsrett. Den multiplikative faktoren kompenserer for dette. Den beregnes ved å forlange at

$$E\left[\frac{1}{N} \sum_{k=0}^{N-1} (v_k^N x[k])^2\right] = E\left[\frac{1}{N} \sum_{k=0}^{N-1} x[k]^2\right] \quad (9.27)$$

Her er $x[k]$ en realisasjon av signalet. Flytter vi forventningsoperasjonen inn i summen, faller $E(x[k]^2)$ bort på begge sider, og vi får ganske enkelt

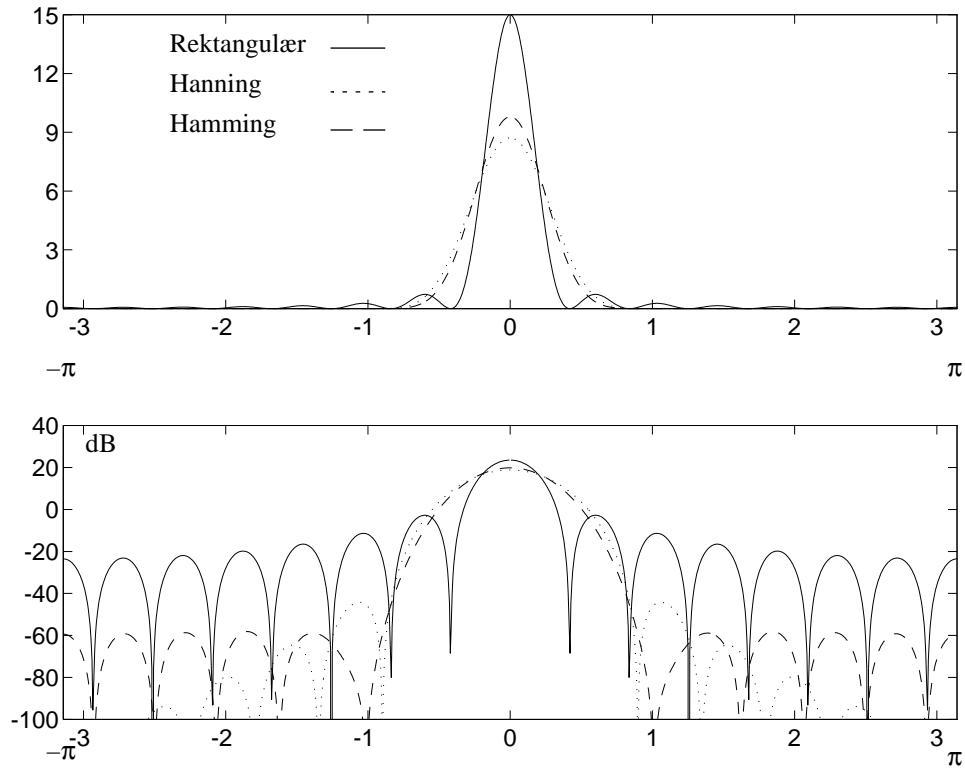
$$\frac{1}{N} \sum_{k=0}^{N-1} (v_k^N)^2 = 1 \quad (9.28)$$

Dette kan brukes til å beregne den rette multiplikative faktoren.

Med rett multiplikativ faktor vil $W^N(\Omega) \rightarrow 2\pi\delta(\Omega)$ når $N \rightarrow \infty$, uansett type vindusfunkjon. Da vil (9.21) ganske enkelt bli:

$$E[\phi_\infty(\Omega)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} 2\pi\delta(\Omega - \lambda)\phi(\lambda)d\lambda = \phi(\Omega) \quad (9.29)$$

Hvis vi beregner periodogrammet på grunnlag av en uendelig lang tidsserie, blir det forventningsrett, som allerede påpekt tidligere.



Figur 9.5

9.2.2 Varians kontra forventningsfeil

Vi har vist at periodogrammet ikke kan bli forventningsrett, men at feilen kan gjøres liten ved å øke N .

Vi kan nå stille oss spørsmål om *variansen* til periodogrammet. Vil den minke når N øker?

Dette er dessverre ikke tilfellet. Det kan vises (gjøres ikke i dette kurset) at variansen er tilnærmet gitt av

$$\text{var}[\phi_N(\Omega)] \approx \phi^2(\Omega) \quad (9.30)$$

D.v.s. standardavviket er av samme størrelsesorden som $\phi(\Omega)$ sjøl.

For å oppnå et estimat med akseptabel varians, midler man vanligvis flere periodogrammer basert på kortere måleserier. Hvis vi deler opp den opprinnelige måleserien i m mindre måleserier av lengde M slik at $MM = N$, får vi følgende estimat av $\phi(\Omega)$:

$$\phi_M^m(\Omega) = \frac{1}{m} \left[\frac{1}{M} \sum_{i=1}^m X_M^i(\Omega) X_M^i(-\Omega) \right] \quad (9.31)$$

hvor $X_M^i(\Omega)$ er den diskret Fourier-transformerte til den i' te måleserien definert ved

$$X_M^i(\Omega) = \sum_{k=M(i-1)}^{Mi-1} x_k e^{-j\Omega k} \quad (9.32)$$

Variansen til $\phi_N^m(\Omega)$ vil nå være gitt av

$$\text{var}[\phi_N^m(\Omega)] = \frac{1}{m} [\phi(\Omega)]^2 \quad (9.33)$$

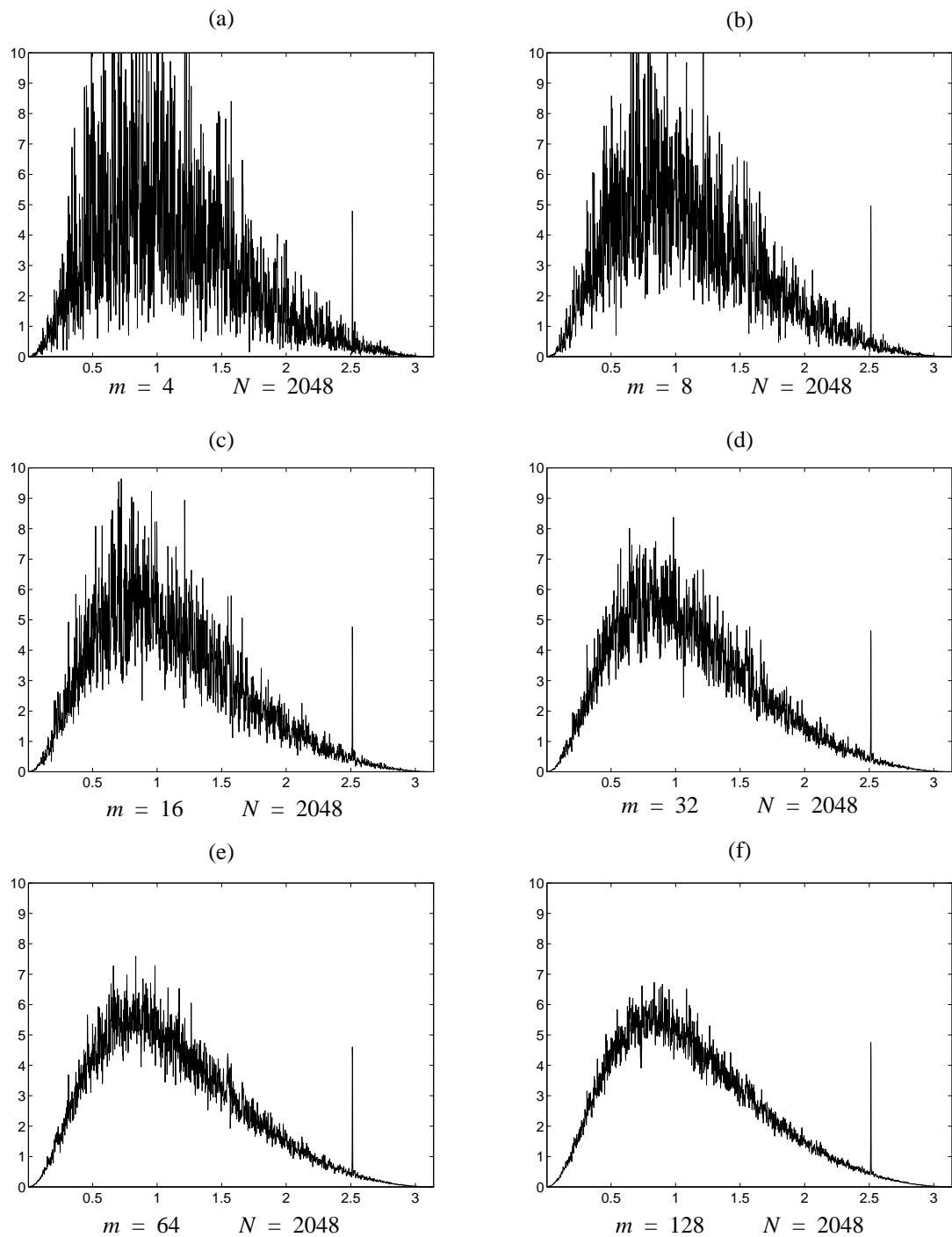
Dette fordi ϕ_M^m fremkommer som en sum av m tilnærmet uavhengige tilfeldige variable, hver med varians $\approx [\phi(\Omega)]^2$.

Fenomenet med minkende varians når flere spektra midles er illustrert i figur 9.6: Den aktuelle støysekvens er båndpassfiltrert hvit støy med en sinus addert på høyfrekvensiden. I 9.6 (a) er vist ett periodogram basert på 2048 målepunkter. Vi har $m = 1$. Vi ser at standardavviket er av samme størrelsesorden som ϕ . I 9.6 (e) derimot er $m = 256$, dvs. 256 periodogrammer á 2048 punkter er midlet. Variansen er da liten. Vi merker oss at sinusfunksjonen gir den samme "spikeren" hele veien, fordi dette er et deterministisk signal og har derfor ingen usikkerhet.

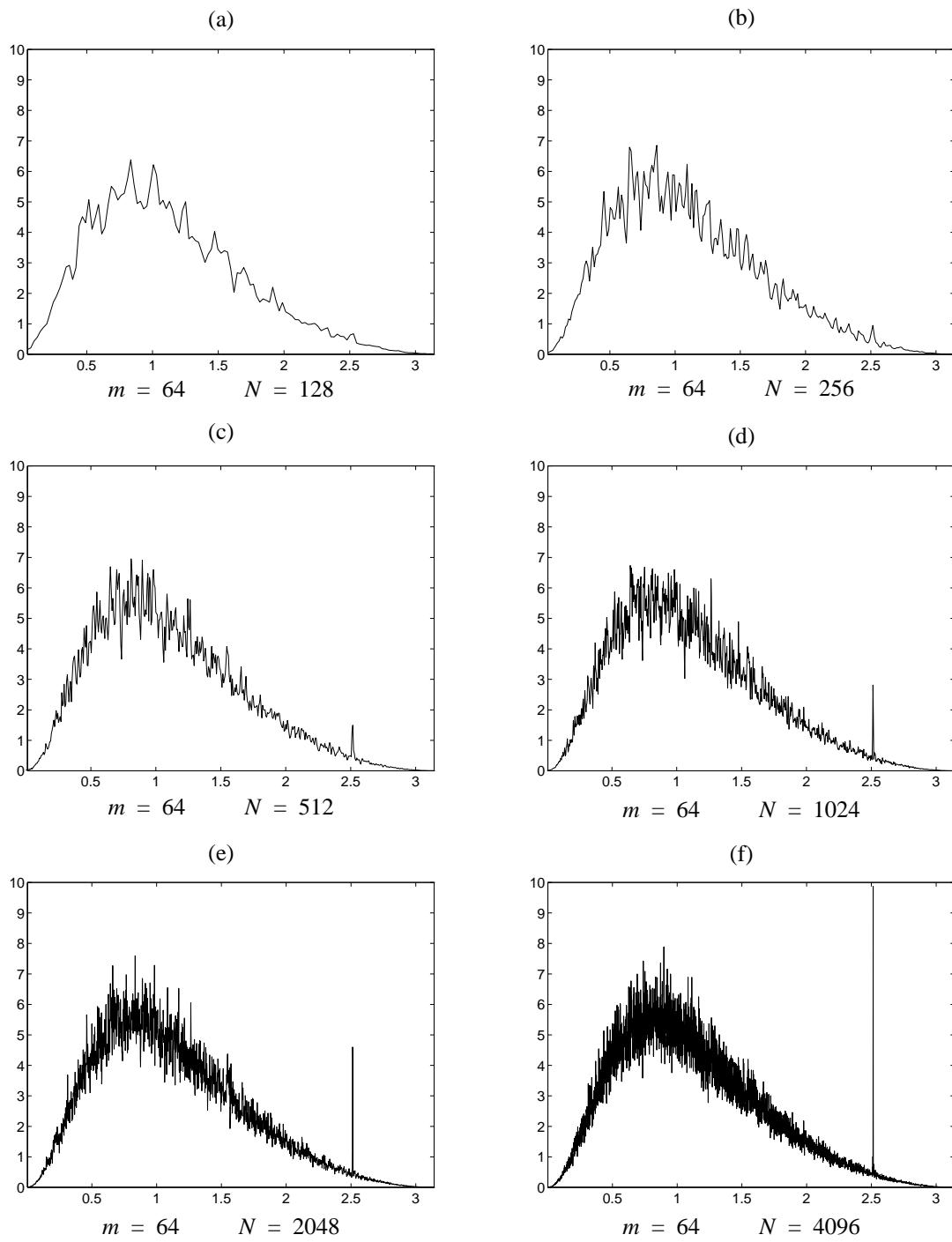
For å illustrere fenomenet med forventningsfeil ("utsmøring") er prosessen fra figur 9.6 analysert med forskjellige lengder på måleseriene, 128 til 4096 punkter, mens antall spektra som midles er holdt konstant, $m = 64$. Resultatet er vist i figur 9.7. Vi ser at for $N = 128$ blir utsmøringseffekten så kraftig at sinussignalet nesten forsvinner fra spektret. Tilsynelatende er variansen mindre for liten N . Dette skyldes at estimatet blir mindre "hårete" når frekvensopplosningen reduseres, men variansen er den samme fra 9.7 (a) til (f) siden $m = \text{konst.} = 64$.

Oppsummering:

For å redusere variansen i estimatet, må vi dele opp tidsserien i flere mindre serier og midle periodogrammene for hver av disse. Dermed vil utsmøringen og den tidligere definerte forventningsfeil øke. Vi må derfor inngå et kompromiss mellom varians i estimatene og forventningsfeil i estimatene. Men for en gitt varians (antall midlede periodogrammer = m) kan forventningsfeilen gjøres mindre ved å anvende en annen avkorting av tidsserien enn den rektangulære.



Figur 9.6



Figur 9.7

9.2.3 Varians kontra forventningsfeil: Et eksempel

Det er generert data ved hjelp av modellen

$$x[i+1] = -0.9x[i] + n[i+1] \quad (9.34)$$

hvor $n[i]$ er rudimentær hvit støy. Vi ønsker å beregne effektspektret $\phi(\Omega)$ til $x[i]$.

$$\text{Vi har } \phi(\Omega) = H(\Omega)H(-\Omega) \cdot 1 \quad (9.35)$$

$$H(\Omega) \text{ kan finnes via z-transformen, } H(z^{-1}) = \frac{1}{1 + 0.9z^{-1}} \quad (9.36)$$

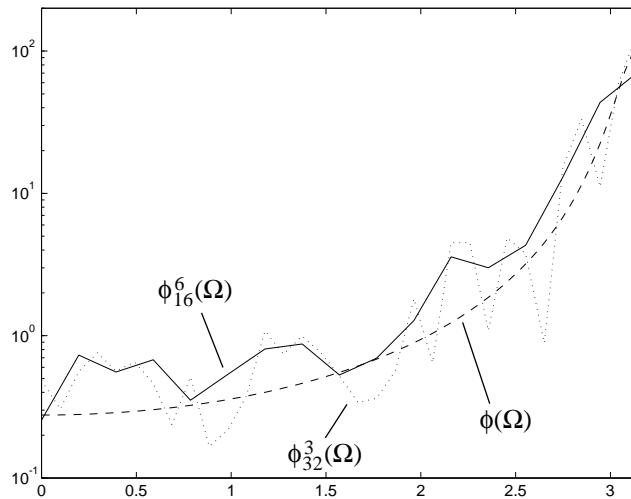
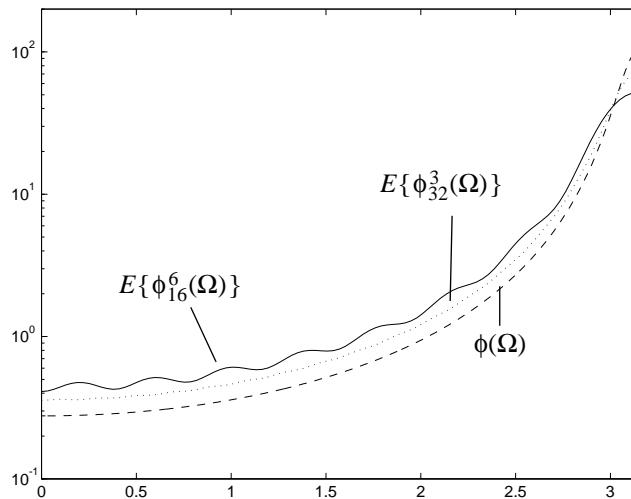
$$\text{Innsetting av } z^{-1} = e^{-j\Omega} \text{ gir } H(\Omega) = \frac{1}{1 + 0.9e^{-j\Omega}} \quad (9.37)$$

$$\phi(\Omega) = H(\Omega)H(-\Omega) = \frac{1}{1 + 0.9e^{+j\Omega} + 0.9e^{-j\Omega} + 0.81} \quad (9.38)$$

$$\phi(\Omega) = \frac{0.555}{1.005 + \cos \Omega} \quad (9.39)$$

Systemet (9.34) er simulert på datamaskin, og på basis av tidsserier av utgangen er periodogrammer beregnet. Den virkelige verdien (= forventningsverdien) for $\phi(\Omega)$ kan beregnes fra (9.35), og er vist i figur 9.8 sammen med to midlere periodogrammer. $\phi_{16}^6(\Omega)$ er fremkommet ved midling av seks periodogrammer á 16 samples. $\phi_{32}^3(\Omega)$ er framkommet ved midling av tre periodogrammer á 32 samples. I begge tilfeller er benyttet 96 samples totalt. Vi ser at $\phi_{16}^6(\Omega)$ har minst varians, men størst forventningsfeil.

Forventningsverdiene for $\phi_{16}^6(\Omega)$ og $\phi_{32}^3(\Omega)$ kan beregnes i dette tilfellet hvor vi kjenner modellen (9.34)-(9.35). Disse er vist i figur 9.9.

Figur 9.8 $\phi(\Omega)$, $\phi_{16}^6(\Omega)$ og $\phi_{32}^3(\Omega)$ Figur 9.9 $E\{\phi_{16}^6(\Omega)\}$ og $E\{\phi_{32}^3(\Omega)\}$

9.2.4 Krysseffektspektra

Estimering av krysseffektspektra foregår helt analogt med den ovenfor beskrevne metoden for estimering av autoeffektspektra. Krysseffektspektret estimeres via formelen

$$\phi_N(\Omega) = \frac{1}{N} X_N(-\Omega) Y_N(\Omega) \quad (9.40)$$

hvor $X_N(\Omega)$ er DFT'en til $x[n], n = 0, 1, \dots, N-1$ og $Y_N(\Omega)$ er DFT'en til $y[n], n = 0, 1, \dots, N-1$. De samme forholdene gjør seg gjeldende her som ved estimering av autoeffektspektret. Forventningsfeilen kan reduseres ved å modifisere dataene med ikke-rettangulære datavindu før man Fouriertransformerer, og variansen kan reduseres ved å bruke flere periodogrammer basert på kortere dataserier.

9.3 Estimering av korrelasjonsfunksjoner

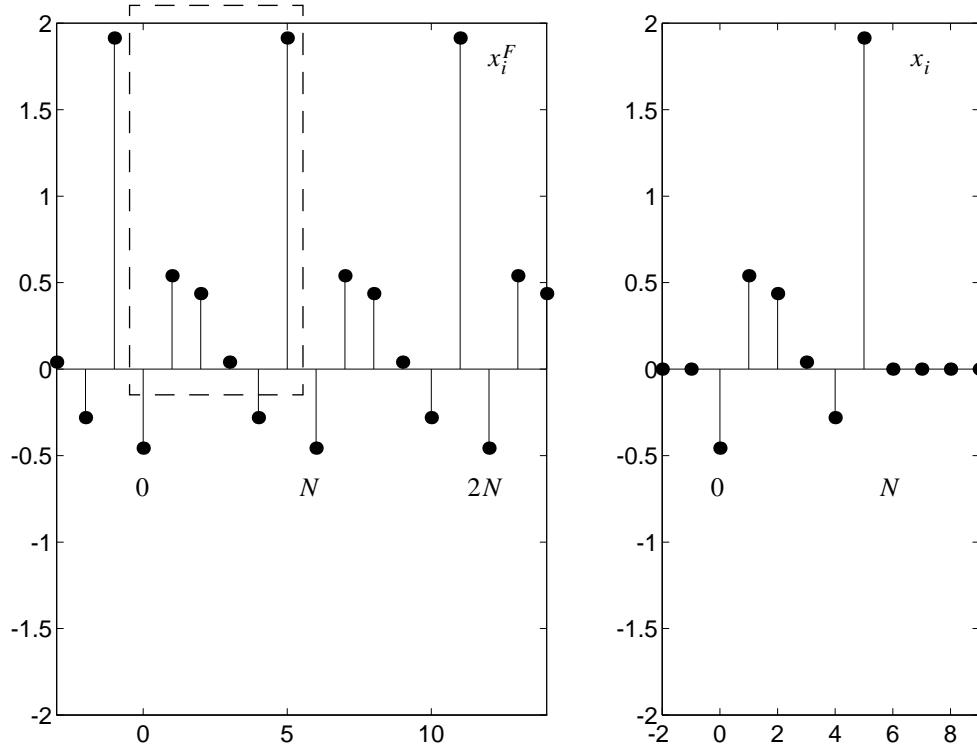
Korrelasjonsfunksjoner beregnes raskt og effektivt ved først å estimere det tilhørende spektrum for deretter å invers Fourier-transformere dette. Her kan man støte på et problem hvis man ikke er påpasselig. Den invers diskret Fourier-transformerte er gitt av¹

$$\varphi[n] = F^{-1} \left[\frac{1}{N} X_N(\Omega_p) X_N(-\Omega_p) \right] \text{ hvor } \Omega_p = \frac{2\pi p}{N} \quad (9.41)$$

Nå har vi at den inverse til $X_N(\Omega_p)$ er gitt av

$$x^F[i] = \frac{1}{N} \sum_{p=0}^{N-1} X_N(\Omega_p) e^{\frac{j2\pi pi}{N}} \quad (9.42)$$

Tidsserien $x^F[i]$ definert fra (9.42) vil nå være forskjellig fra den opprinnelige tidsserien $x[i]$, idet $x^F[i]$ vil være periodisk med periode N, mens $x[i] = 0$ når $i < 0$ eller $i \geq N$. Dette er eksemplifisert i figur 9.10.



Figur 9.10

1. Egentlig skulle vi skrive $\Phi_{xx, N}[n]$ her, det er altså underforstått at vi nå studerer estimatet av en autokorrelasjonsfunksjon basert på en tidsserie på N punkter. Det samme resonnement vil forøvrig gjelde for en krysskorrelasjonsfunksjon.

Husk at DFT matematisk forutsetter at signalet er periodisk i tidsplanet, slik at DFT egentlig er det samme som den diskrete Fourierrekke, DFS, slik som nevnt på side 123. Nå vet vi at produkter i frekvensplanet tilsvarer folding i tidsplanet, slik at (9.41) motsvarer folding av de invers Fourier-transformerte til $X_N(\Omega_p)$ i tidsplanet. Vi har

$$\begin{aligned} F^{-1}\left[\frac{1}{N}X_N(\Omega_p)X_N(-\Omega_p)\right] &= \frac{1}{N}x^F[i]^*x^F[-i] = \frac{1}{N}\sum_{k=0}^{N-1}x^F[k]x^F[-(i-k)] \\ &= \frac{1}{N}\sum_{k=0}^{N-1}x^F[k]x^F[k-i] = \phi^F[i] \end{aligned} \quad (9.43)$$

Da $x^F[i]$ er den inverse Fourier-transformen til $X_N(\Omega_p)$, vil denne foldingen skje mellom $x^F[i]$ og seg selv. Periodisiteten i $x^F[i]$ vil derfor medføre feil. Med andre ord: Det vi egentlig ønsker å finne er

$$\phi[i] = \frac{1}{N}\sum_{k=0}^{N-1}x[k]x[k+i] \quad (9.44)$$

mens det vi får ved invers-Fourier-transformasjonen gitt i (9.43) er dessverre

$$\phi^F[i] = \frac{1}{N}\sum_{k=0}^{N-1}x^F[k]x^F[k+i] \quad (9.45)$$

Vi kan imidlertid omgå dette problemet ved å utvide signalet $x[i]$ med N nuller i tillegg til de opprinnelige datapunktene. Vi får dermed en dobbelt så lang tidsserie, hvorav den siste halvparten er nuller. Dermed unngår vi overlappingsfenomenene som vi ellers ville fått, siden $\phi^F[i] = \phi[i]$ for $0 \leq i \leq N-1$. Slik utfylling med nuller (eng: "*Padding with zeroes*") er derfor vanlig å foreta før DFT-operasjonen utføres. Men dette er ikke nødvendig hvis man bare skal studere spektra og transferfunksjoner, dvs. når man ikke skal inverstransformere etterpå for å finne korrelasjonsfunksjoner.

Utfylling med nuller i tidsplanet gir tilsvarende flere punkter i frekvensplanet. Tilsynelatende gir dette bedre frekvensoppløsning. Men utsmøringen er fortsatt like framtredende (jfr. figur 9.8). Derfor kan man si det slik at utfylling med nuller i tidsplanet bare medfører en interpolasjon i frekvensplanet, og ikke en reell forbedring av frekvensoppløsningen.

Når vi skal beregne korrelasjonsfunksjoner er det enda en ting vi bør ta hensyn til for å få best mulige estimerater. Anta at N nuller er lagt til, slik at vi har $2N$ punkter. Vi kaller den tilsvarende Fouriertransform for X_{2N} . En *forventningsrett* variant av (9.44) blir

$$\phi[i] = \frac{1}{2N-|i|}\sum_{k=0}^{2N-1}x[k]x[k+i] \quad (9.46)$$

Vi ser at i tillegg til at vi nå har dobbelt så mange punkter (som forutsatt), så er det et ledd $2N - |i|$ i nevneren, ikke $2N$ som vi kunne forvente. Grunnen er nevneren må reduseres proporsjonalt med graden av forskyvning mellom $x[k]$ og $x[k + i]$ for å kompensere for at tilsvarende ledd i summen blir = 0. Bare da blir $E(\phi[i]) =$ den virkelige korrelasjonsfunksjon, altså forventningsrett estimat.

Merknad:

(9.45) kan skrives med følgende symbolbruk:

$$\phi^F[i] = \frac{1}{N} \sum_{\langle N \rangle} x^F[k] x^F[k-i] = \frac{1}{N} x^F[i] \otimes x^F[-i] \quad (9.47)$$

Her symboliserer $\langle N \rangle$ at intervallet kan starte og slutte hvor det vil, bare det er N punkter langt. Denne friheten til å velge startpunkt for intervallet skyldes periodisiteten i signalet $x^F[k]$. Denne typen folding hvor intervallet er endelig kalles ofte syklistisk folding. Symbolet som brukes for syklistisk folding er ofte \otimes slik som vist i (9.47), men ellers kan man godt bruke $*$, som vi kjenner fra før.

9.4 Estimering av transferfunksjoner

Vi har tidligere sett på estimering av transferfunksjoner under forutsetning av ideell måling av signalspektrene. Vi skal nå ut fra den foregående diskusjonen omkring estimering av spektra, se hvordan feil i estimatorer slår ut i feil i transferfunksjonsestimatene. Vi antar at vi har estimert inngangsspektret til $\hat{\phi}_{xx}(\Omega_n)$ og krysseffektspektret $\hat{\phi}_{xy}(\Omega_n)$. Vi antar videre at m ulike periodogrammer er midlet for beregning av spektrene. Transferfunksjonsestimatene er da gitt av

$$\hat{H}(\Omega_n) = \frac{\hat{\phi}_{xy}(\Omega_n)}{\hat{\phi}_{xx}(\Omega_n)} \quad (9.48)$$

Det kan utledes formler (utledning gjøres ikke her) som gir spredningen i estimatet av \hat{H} når \hat{H} baseres på (9.48). I disse formlene brukes begrepet konfidensintervall: Man søker en størrelse $\hat{r}^2(\Omega_n)$ slik at

$$|\hat{H}(\Omega_n) - H(\Omega_n)|^2 \leq \hat{r}^2(\Omega_n) \quad (9.49)$$

og sannsynligheten for at (9.49) er oppfylt er $1 - \alpha$. M.a.o. må konfidensintervallet \hat{r}^2 for en gitt frekvens være en funksjon av α .

\hat{r}^2 må avta monoton med α ; eller ekvivalent, øke med $1 - \alpha$. Jo større sannsynlighet for at estimatet \hat{H} skal befinne seg mindre enn avstanden \hat{r} fra den virkelige transferfunksjonen H , jo større må \hat{r} være. For \hat{r} gitt ved (9.49) gjelder

$$\hat{r}^2(\Omega_n) = \frac{2}{m-2} [1 - \hat{\gamma}_{xy}^2(\Omega_n)] \frac{\hat{\phi}_{yy}(\Omega_n)}{\hat{\phi}_{xx}(\Omega_n)} F_{2, m; \alpha} \quad (9.50)$$

Her er $F_{2, m; \alpha}$ abscisseverdien til den såkalte F-fordelingen (se tabeller i lærebøker i statistikk). Verdien $F_{2, m; \alpha}$ framkommer ved at arealet under tetthetsfunksjonen skal være henholdsvis $1 - \alpha$ og α på hver side av $F_{2, m; \alpha}$. Indeksene 2 og m kalles frihetsgrader og er parametre i F-fordelingen sammen med α som er definert over. m er slik som tidligere antall spektra som midles. Jo større sikkerhet (konfidens) vi forlanger, jo mindre blir α , og $1 - \alpha$ går mot 1. $F_{2, m; \alpha}$ vil øke, og dermed øker \hat{r}^2 i følge (9.50). Av (9.50) ser vi også at \hat{r}^2 avtar når m øker, dvs. sikrere estimat $|\hat{H}|$ ved flere midlede spektra. Estimatet $|\hat{H}|$ blir også bedre hvis $x(t)$ og $y(t)$ er relativt koherente, dvs. $\gamma_{xy}^2 \rightarrow 1$.

M.a.o. kan vi få et meget godt estimat av $|\hat{H}|$ sjøl om $\hat{\phi}_{xx}$ og $\hat{\phi}_{yy}$ har stor varians, hvis bare x og y er koherente.

Formel (9.50) sier noe om feilen i tallverdien til $|\hat{H}|$. Når det gjelder fasfeilen til $|\hat{H}|$, kan man utlede et tilsvarende konfidensintervall med lengde $2\Delta\theta(\Omega_n)$, der

$$\Delta\theta(j\omega) = \arcsin\left(\frac{\hat{r}(\Omega_n)}{|\hat{H}(\Omega_n)|}\right) \quad (9.51)$$

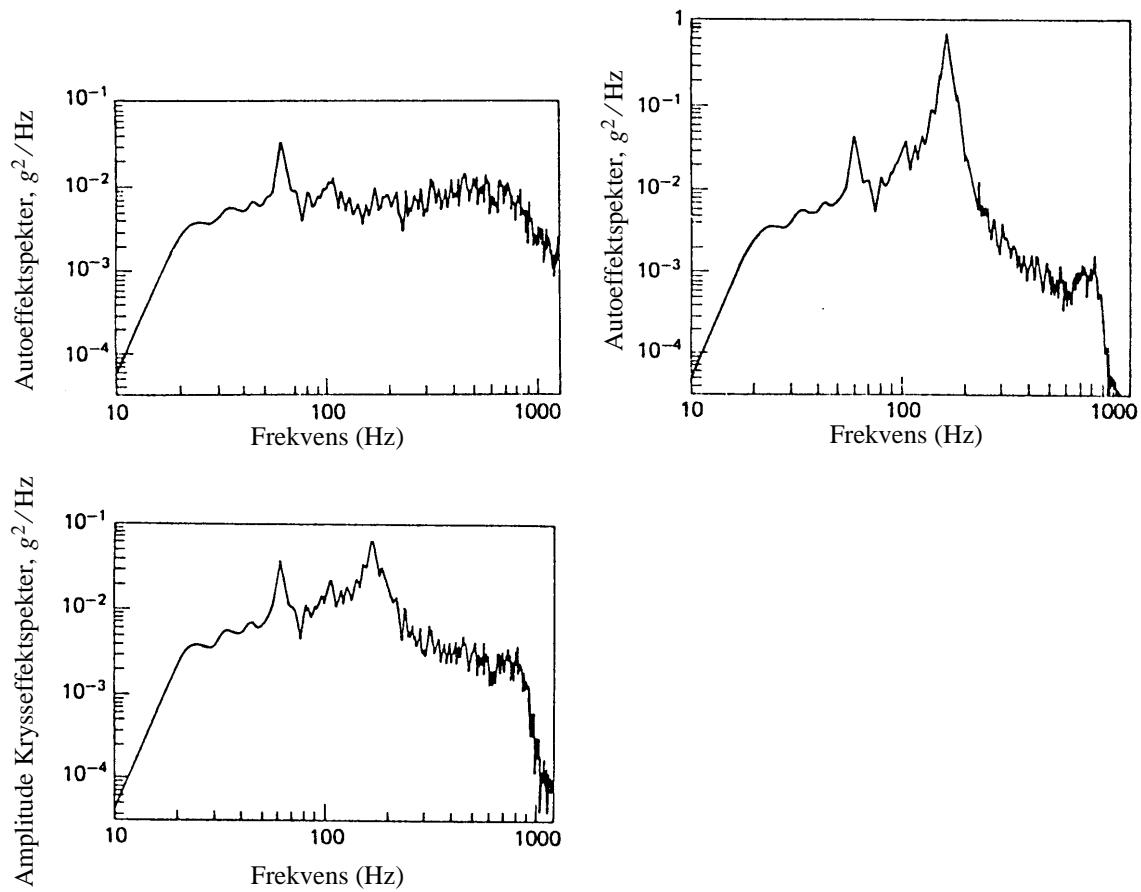
Man må altså først velge α , så regne ut \hat{r} ved hjelp av (9.50), og deretter bruke \hat{r} i (9.51) for å få et konfidensintervall for fasfeilen. Da vil virkelig fasegang $\theta(\Omega_n)$ være slik at

$$\hat{\theta}(\Omega_n) - \Delta\hat{\theta}(\Omega_n) \leq \theta(\Omega_n) \leq \hat{\theta}(\Omega_n) + \Delta\hat{\theta}(\Omega_n) \quad (9.52)$$

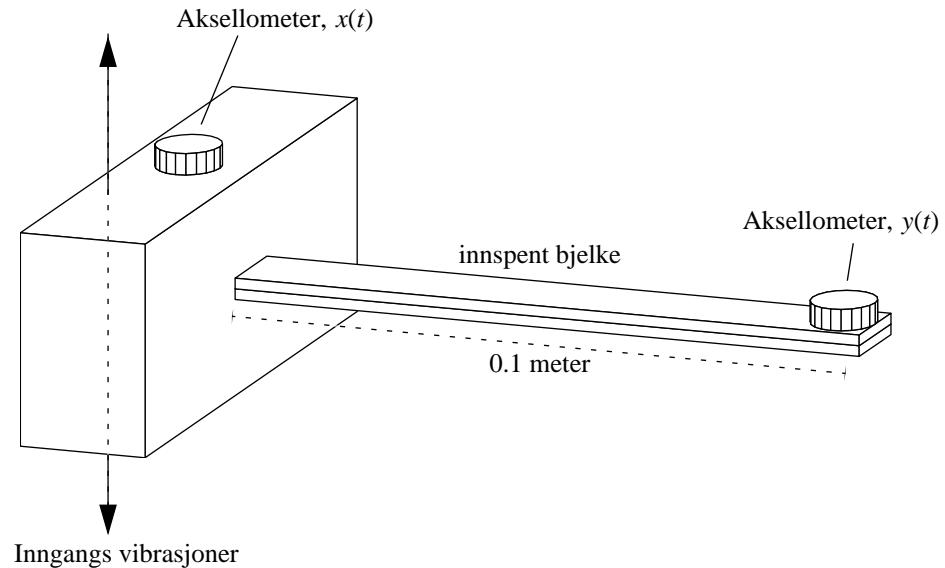
med sannsynlighet $1 - \alpha$.

Eksempel:

Vi skal vise et eksempel på estimering av transferfunksjoner for en innspent bjelke som vist i figur 9.12. Bjelken eksiteres ved vertikale akselerasjoner på innspenningspunktet og akselasjonen her og ytterst på bjelken måles med akselerometre. Inngangssignalets spektrum er vist i 9.11 (a), utgangssignalets spektrum i 9.11 (b) og krysseffektspektret i 9.11 (c). Estimatet av koherensfunksjonen og den estimerte transferfunksjonen er vist i figurene 9.13 og 9.14. Analysen er utført ved bruk av FFT og midling av 29 periodogrammer.

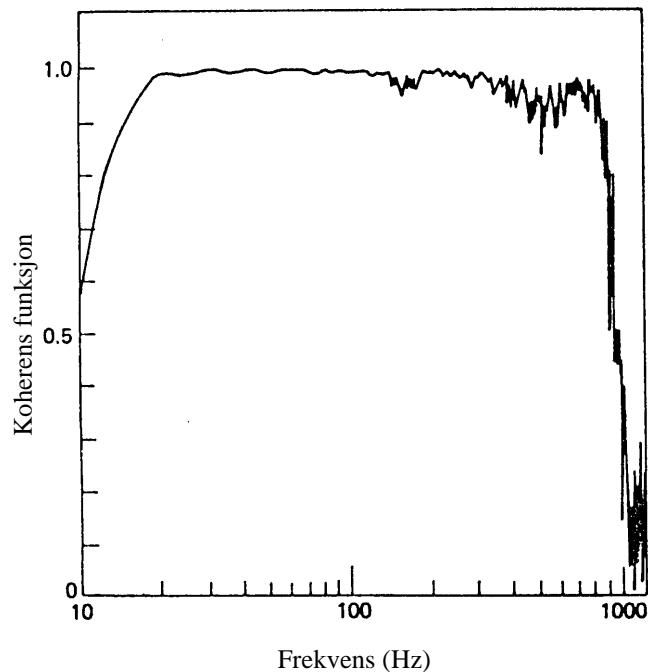


Figur 9.11 Spektra fra forsøk med den innspente bjelken.

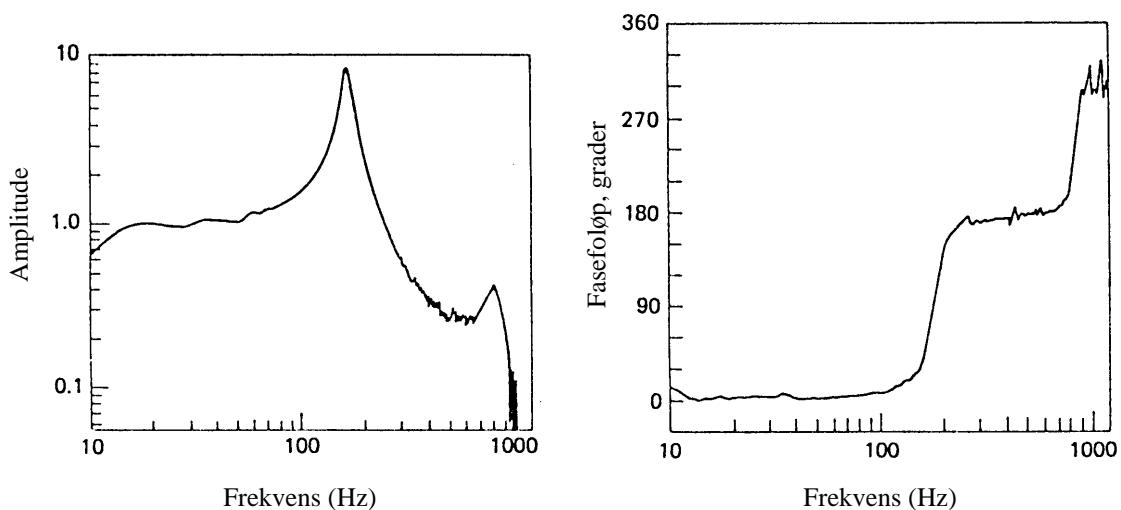


Figur 9.12 Innspent bjelke for estimering av vibrasjons-transferfunksjon.

Vi ser at $\hat{\gamma}_{xy}^2$ er mindre enn 1 ved $f < 20 \text{ Hz}$ og $f > 800 \text{ Hz}$. Det betyr at estimatene av \hat{H} er dårligere i disse områdene. Merk også toppen i eksitasjonen ved 60 Hz. Denne skyldes nettfrekvensen som utilsiktet har blitt introdusert i eksitasjonen. $m = 29$ tilsier en usikkerhet i estimatene $\hat{\phi}_{xy}$ og $\hat{\phi}_{xx}$ på ca. 20%. Usikkerheten i \hat{H} ser ut til å være mindre. Dette stemmer med (9.50) og (9.51).



Figur 9.13 Koherensfunksjon (estimat: $\hat{\gamma}_{xy}^2 = |\hat{\phi}_{xy}|^2 / \hat{\phi}_{xx} \hat{\phi}_{yy}$).



Figur 9.14 Estimert transferfunksjon, $|\hat{H}|$ og $-\angle H$.

Kapittel 10: Signalbehandling: Tilpassing, filtrering

I forbindelse med reguleringsstekniske anvendelser har man ofte behov for å tilpasse signaler og måleserier til det utstyret man benytter eller til den metoden som anvendes. Hvis man f.eks. benytter en digital regneenhet må analoge signaler omformes til diskrete signaler gjennom **prefiltrering** og **tasting ("sampling")**. Og diskrete signaler må omformes til analoge signaler for utsetting av pådrag til prosessen.

Hvis man ønsker å analysere tidsserier fra en prosess, må man i mange tilfeller foreta enkle typer forbehandling av disse signalene for å kunne benytte disse i ulike metoder for modelltilpasning eller estimering. Disse enklere formene for signalbehandling kan være ulike former for **filtrering**, fjerning av **middeverdier**, fjerning av **oppdagte gale målinger (wild points)** og konvertering av to tidsserier med **ulik tastetid** til to tidsserier med **felles tastetid**. Vi skal i dette kapittelet ta for oss noen slike problemer og løsningsmetoder.

10.1 Analog til digital omsettning

10.1.1 Tasting av signaler

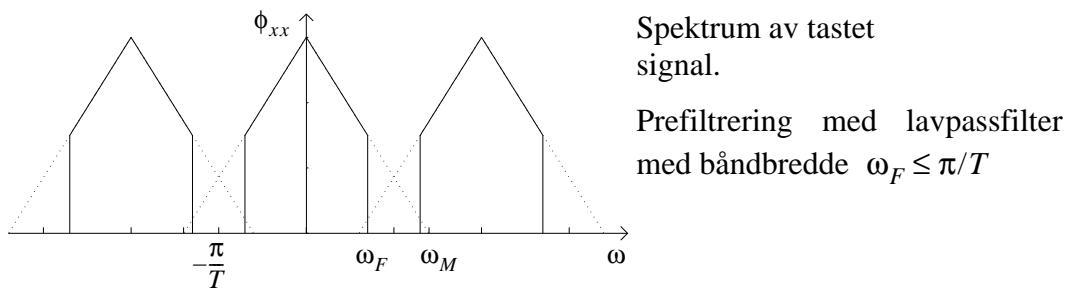
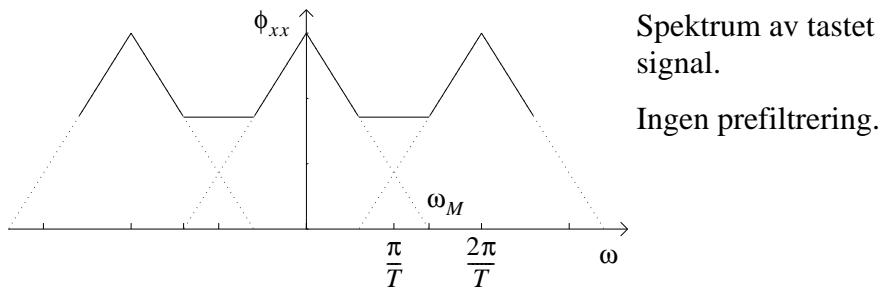
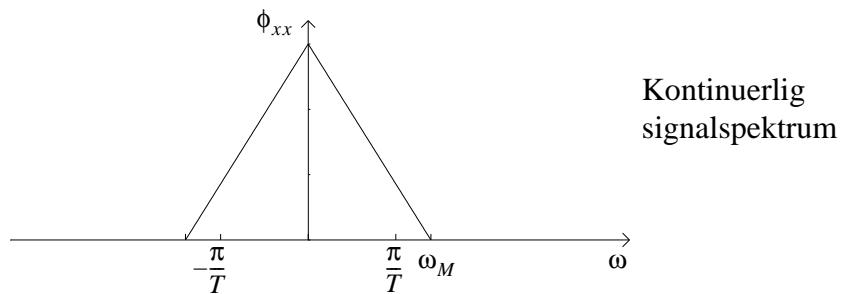
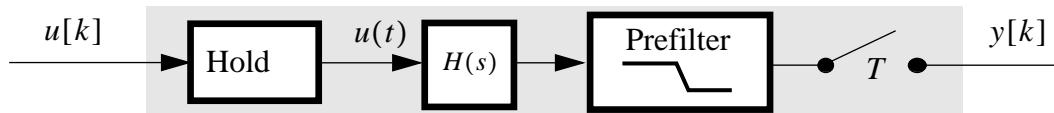
Som vist i avsnitt 8.1 i "Signals and Systems" vil tasting av et signal føre til modifikasjon av spektret: Det tastede signalets spektrum (Man sier ofte "spektrum" når man egentlig mener "Fouriertransform"). Dette går bra så lenge sammenhengen gjør at det ikke kan misforstås) består av en sum av gjentagelser av det opprinnelige spektret forskjøvet langs frekvensaksen (se figur 8.4 i S & S). Hvis disse forskjøvede spektra overlapper, fører dette til sammenblanding av frekvenser i det en hurtig frekvens (frekvenser i signalet over Nyquist-frekvensen) vil tolkes som et lavfrekvent signal. For å unngå nedfolding benytter man ved tasting av signaler som regel et analogt **prefilter** eller et **antinedfoldingsfilter** før (derav betegnelsen *prefilter*) tasting av signalet (se figur 10.1).

Merknad:

Med "analogt" mener man her "kontinuerlig", dvs. filteret er et elektronisk filter som behandler et kontinuerlig varierende signal. Dette i motsetning til digitale eller diskrete filtere som vi kommer til seinere i dette kapittelet.

Prefilteret må være et lavpassfilter. Hvis man ønsker å taste et signal med tastefrekvensen $2\pi/T$ bør signalet prefiltreres med et lavpassfilter med båndbredde $\omega_F \leq \pi/T$, slik at alle frekvenser i det kontinuerlige signal over π/T fjernes.

Situasjonen med og uten filter er vist i figur 10.1.



Figur 10.1 Prefiltrering før tasting

Vi kan oppsummere:

Ved tasting av et kontinuerlig signal bør vi

1. *taste raskere enn det dobbelte av den høyeste interessante frekvens i signalet ω_0 . Dvs.: $\pi/T > \omega_0$.*
2. *Introduksjon et pretastingsfilter med lavpasskarakter og båndbredde ω_F slik at $\omega_0 < \omega_F < \pi/T$.*

I praksis velger man f.eks. $\frac{\pi}{T} \approx 10 \cdot \omega_0$ og samtidig $\omega_F \approx \frac{1}{3} \frac{\pi}{T}$ hvor ω_0 er høyeste interessante frekvens i signalet og ω_F er båndbredden i lavpassfilteret. I reguleringssammenheng kan ω_0 ofte settes litt i overkant av systemets båndbredde.

10.1.2 Kvantisering av signaler

Når vi skal behandle et analogt (=kontinuerlig) signal i datamaskin må vi taste signalet. Tasting av signalet innebærer en digitalisering av signalet *langs tidsaksen*. Men vi må også digitalisere signalet i *amplituderetning*, dvs. signalets amplitude må uttrykkes som et binært tall med et visst antall bits. Denne siste formen for digitalisering kaller vi **kvantisering**.

Et eksempel er CD-spilleren. Tastingen skjer 44100 ganger i sekundet. Ved hvert tastetidspunkt omgjøres signalet til et binært 16 bits tall (vi har egentlig to signaler - stereo - men la oss bare betrakte den ene kanalen i denne sammenhengen). Vi sier da at kvantiseringen skjer via en "16-bits-A/D-omsetter" (A/D = "analog til digital"). Vi har tidligere studert virkningen av tasting. Vi skal nå se på effekten av kvantisering av signalet og den **kvantiseringfeilen** som oppstår. La oss anta at vi har en n-bits A/D-omsetter med maksimal utgang svarende til E fysikalske enheter (f.eks. 1 Volt).

Feilen i det digitaliserte signalet antas da å være stokastisk og uniformt (rektagulært) fordelt i området $\left[-\frac{1}{2} \frac{E}{2^n}, \frac{1}{2} \frac{E}{2^n}\right]$.

Sannsynlighetstetthetsfunksjonen for feilen e kan da skrives:

$$p(e) = \begin{cases} 2^n/E & \text{for } e \in \left[-\frac{1}{2} \frac{E}{2^n}, \frac{1}{2} \frac{E}{2^n}\right] \\ 0 & \text{ellers} \end{cases} \quad (10.1)$$

Vi har åpenbart at $E(e) = 0$. Variansen blir

$$E(e^2) = \int_{-\frac{1}{2} \frac{E}{2^n}}^{\frac{1}{2} \frac{E}{2^n}} e^2 \frac{2^n}{E} de = \left(\frac{1}{3} e^3 \cdot \frac{2^n}{E} \right) \Big|_{-\frac{1}{2} \frac{E}{2^n}}^{\frac{1}{2} \frac{E}{2^n}} = \frac{1}{12} \cdot \left(\frac{E}{2^n}\right)^2 \quad (10.2)$$

Kvantiseringstøy kan derfor tilnærmet anses som hvit støy med null middelverdi og varians (= effekt i dette tilfelle siden vi har $E(e) = 0$) gitt ved

$$\sigma_e^2 = \frac{1}{12} \left(\frac{E}{2^n}\right)^2 \quad (10.3)$$

10.2 Digitale filtre

Digitale filtre benyttes (som analoge filtre) til mange formål. For reguleringstekniske formål kan vi f.eks. tenke oss lavpassfiltrering av et signal, fjerning av et spesielt frekvensområde (f.eks. 50 Hz støy) osv. Design av dynamiske tilbakekoplinger og regulatorer kan også betraktes som digital filter-design.

Digitale filtre kan prinsipielt deles i to hovedkategorier: IIR-filtre og FIR-filtre. **IIR** betyr **Infinite Impulse Response**. Disse filtrene vil i diskret versjon ha formen

$$\begin{aligned} y[k] = & -a_1y[k-1] - a_2y[k-2] - \dots - a_ny[k-n] \\ & + b_0u[k] + b_1u[k-1] + \dots + b_nu[k-n] \end{aligned} \quad (10.4)$$

Betegnelsen IIR skyldes at impulsresponsen til et slikt filter aldri dør helt ut. Fordelen med et IIR-filter er at man ikke trenger å lagre mange gamle verdier av $y[k]$ og $u[k]$, siden formelen er rekursiv. Dette i motsetning til FIR-filteret behandlet nedenfor.

Merknad:

(10.4) er en ARMAX-modell uten støy-ledd (den kalles da ofte en ARX-modell). Vi ser til forskjell fra (5.1) og (5.2) i kapittel 5 at ledet $u[k]$ opptrer på høyre side i (10.4), mens det nyeste leddet som kan forekomme i (5.1) er $u[k-1]$. Forskjellen skyldes at (10.4) beskriver en diskret algoritme, mens (5.1) representerer et diskretisert, strengt propert kontinuerlig system med holdelement foran.

Betegnelsen **FIR står for Finite Impulse Response**. I motsetning til IIR-filtre som også realiseres analogt, er FIR-filtre noe som finnes utelukkende i diskret versjon. FIR-filtre kalles også transversale filtre. Et FIR-filter har formen

$$y[k] = b_0u[k] + b_1u[k-1] + \dots + b_nu[k-n] \quad (10.5)$$

Vi ser umiddelbart at et FIR-filter aldri kan bli ustabilt. Alle poler ligger i origo. Vi ser at impulsresponsen rett og slett blir lik koeffisientene i B -polynomet, og at $h[k]$ dermed blir = 0 for $k > n$. Derav betegnelsen *Finite Impulse Response*. Impulsresponsen til et system med dead-beat-regulering (jfr. kapittel 5) er slik. Dermed vil et slikt lukket system - i matematisk forstand - være et FIR-filter. Men vi reserverer helst betegnelsen "FIR" for diskrete filtere. FIR-filteret kan lages med vilkårlig (men endelig) impulsrespons rett og slett ved å velge passende verdier på koeffisientene, samtidig som man gjør antallet koeffisienter stort nok. Impulsresponsen kan også gjøres ikke-kausal og symmetrisk om $k = 0$, og dette betyr at filterets fase blir $\angle H(\Omega) = 0 \forall \Omega$; dvs. intet fase-eterslep for noen frekvens. Dette er en fordel man kan realisere når filteret ikke trenger å være kausalt, f.eks. ved etterbehandling av måleserier som ikke trenger å foregå i sann tid, eller ved sanntidsanwendelser hvor det er akseptabelt med en transportforsinkelse i databehandlingen. Hvis man forskyver impulsresponsen til høyre slik at den starter for $k = 0$, og filteret blir kausalt, men man samtidig beholder den symmetriske formen på impulsresponsen, blir fasengangen ikke lenger $\angle H = 0 \forall \Omega$, men $\angle H = \text{konst.} \cdot \Omega$. Amplituderesponsen forblir den samme. Slik lineær fasengang - som man alltid får med symmetri i impulsresponsen - er spesielt gunstig ved

signalbehandling av Hi-Fi-signaler, fordi det betyr at alle frekvenser tidsforsinkes like mye, og vi unngår innbyrdes faseforskyvning mellom de enkelte frekvenskomponenter i signalet, som fører til det som kalles faseforvrengning.

Ulempen med FIR-filtre er at et FIR-filter med like bratte flanker i overgangsområdet i frekvensplanet som et IIR-filter, må lagre svært mange flere inngangsverdier $u[k]$ og ha tilsvarende mange flere vektkoeffisienter b_k enn IIR-filteret. FIR-filteret har også “bulker” i pass- og stoppbåndet som kan skape problemer.

Vi skal i det følgende nøye oss med å gjennomgå en nokså utbredt metode for å konstruere et IIR-lavpassfilter. IIR-lavpassfiltre finnes i mange varianter (Tsjebyscheff-, elliptiske-, m.fl.), men vi skal ta utgangspunkt i det såkalte **Butterworth-filteret**. Dette er det lavpassfilter som har maksimalt flat respons i passbåndet for en gitt orden på filtertransferfunksjonen. Det kommer vi til i avsnitt 10.3.1. men først til et annet tema:

10.3 Overgang fra analogt til digitalt filter

Her skal vi vise hvordan vi kan lage et diskret filter ved først å spesifisere et analogt filter som tilfredsstiller de gitte krav, og så danne et digitalt filter på basis av det analoge (metoden kan altså brukes til bare å lage et analogt filter, f.eks. et prefilter, hvis det er målet).

Metoden for å syntetisere et digitalt filter ved hjelp av et analogt, går som følger:

Vi tar utgangspunkt i basisoperasjonen i en analog transferfunksjon, nemlig integrasjonen. Vi betrakter likningen

$$\dot{x} = u \quad (10.6)$$

hvor x er integralet av u . Denne likningen kan approksimeres på følgende måte, idet vi definerer $x(kT) = x[k]$ og $u(kT) = u[k]$. T er tasteintervallet:

$$\frac{x[k+1] - x[k]}{T} = \frac{1}{2}(u[k+1] + u[k]) \quad (10.7)$$

I z-planet blir dette:

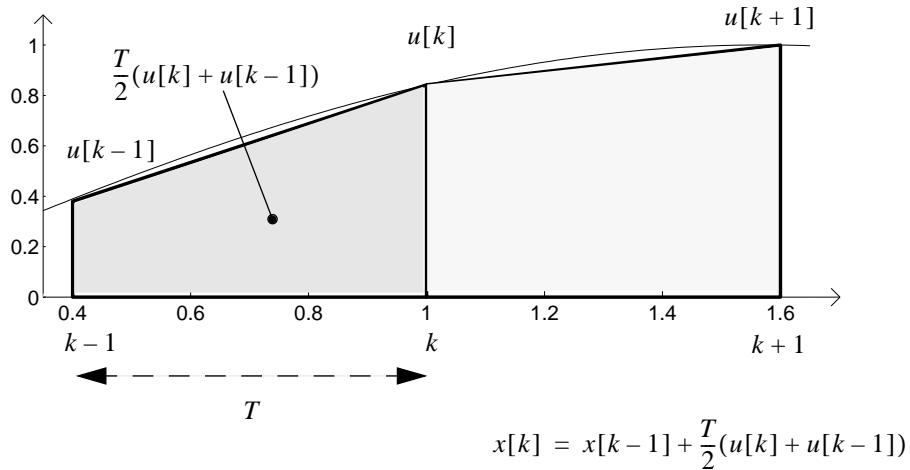
$$x(z) = \frac{T}{2} \frac{z+1}{z-1} u(z) \quad (10.8)$$

Vi ser at integraloperasjonen i analog form, $\frac{1}{s}$, svarer til $(T/2)(z+1)/(z-1)$ med den approksimasjonen som er antatt i (10.8). Vi har altså at

$$\frac{1}{s} \text{ svarer til } \frac{T}{2} \frac{z+1}{z-1}, \quad (10.9)$$

$$\text{eller } s \text{ svarer til } \frac{2}{T} \frac{z-1}{z+1} \quad (10.10)$$

Approksimasjonen (10.9) svarer til en såkalt **trapesintegrasjon**, se fig. 10.2:



Figur 10.2

Vi kan derfor tenke oss å konstruere diskrete filtre ved å substituere $\frac{2}{T} \frac{z-1}{z+1}$ for s (jfr. (10.10)) i den analoge transferfunksjonen som vi ønsker å realisere diskret.

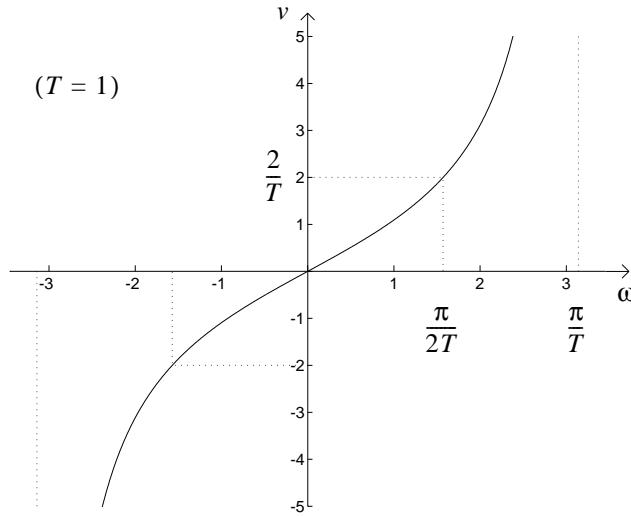
Vi har tidligere sett at $z = e^{j\Omega} = e^{j\omega T}$. Innsetting av dette i (10.10) gir

$$s = \frac{e^{j\omega T} - 1}{e^{j\omega T} + 1} \frac{2}{T} = j \frac{2}{T} \tan\left(\frac{\omega T}{2}\right) \quad (10.11)$$

eller

$$\nu = \text{Im}(s) = \frac{2}{T} \tan\left(\frac{\omega T}{2}\right) \quad (10.12)$$

Vi har her innført en "frekvens" ν som bare er en hjelpestørrelse, og ikke har noen direkte fysisk tolkning. Den variable ν i det analoge filteret vil svare til den fysiske frekvens ω gitt av (10.12) i det digitale filteret. (10.12) er vist i figur 10.3. Vi ser at når $\frac{\omega T}{2} \rightarrow \frac{\pi}{2}$ eller ekvivalent, når $\omega \rightarrow \frac{\pi}{T}$ (som også er Nyquistfrekvensen), vil $\nu \rightarrow \infty$.



Figur 10.3 Relasjonen (10.12).

Med andre ord: $v \in (-\infty, \infty)$ avbildes i frekvensområdet $\omega \in \left(-\frac{\pi}{T}, \frac{\pi}{T}\right)$: Det er - rimeligvis - en absolutt øvre grense (avhengig av tastefrekvensen) for det diskrete filterets evne til å approksimere et kontinuerlig filter. Men for små verdier av T har vi at

$$\frac{dv}{d\omega} = \frac{2}{1 + \cos(\omega T)} \approx 1 \quad (10.13)$$

Dvs, for lave frekvenser vil det være tilnærmet likhet mellom v i det kontinuerlige filteret og ω i det diskrete filteret.

Merk igjen at v er en hjelpestørrelse. Hvis vi ønsker et digitalt filter med en spesifisert frekvensrespons i ω , kan vi først konstruere et analogt filter med en passende respons i v og deretter benytte (10.10).

Det henvises til eksemplet i avsnitt 10.3.3.

10.3.1 Syntese av analogt Butterworth lavpassfilter

Butterworth-filteret er et av flere alternative analoge lavpassfiltre som allerede nevnt. Det særegne med BW-filteret er at frekvensresponsen er maksimalt flat i passbåndet, og at frekvensresponsen er monotont fallende (ingen "bulker" eller oscillasjoner). Et analogt BW-filter av orden N har en transferfunksjon $H(s)$ slik at

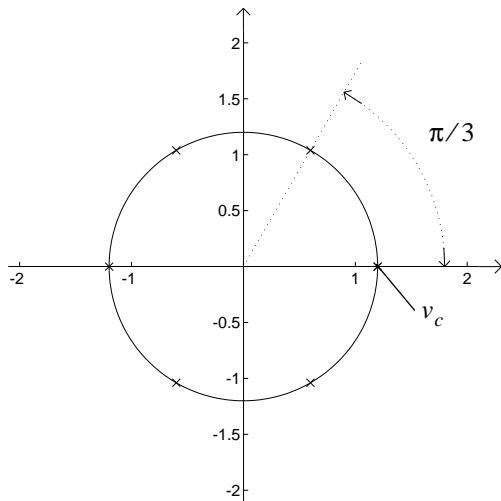
$$H(s) \cdot H(-s) = \frac{1}{1 + \left(\frac{s}{jv_c}\right)^{2N}} \quad (10.14)$$

hvor v_c er filterets båndbredde.

Ved å sette nevneren = 0 finner vi at (10.14) har poler i

$$s_p = (-1)^{\frac{1}{2N}} j v_c \quad (10.15)$$

Filteret har derfor N poler plassert på venstre halvsirkel om origo med radius v_c i det komplekse plan (polene til $H(-s)$ ligger på høyre halvsirkel). Polene i $H(s)$ og $H(-s)$ er symmetrisk plassert i forhold til hverandre om den imaginære aksen. Se eksempel i figur 10.4.



Figur 10.4 s-plan-poler for 3. ordens Butterworth-filter.

Vi har dermed en oppskrift for å lage et *analogt* lavpassfilter, som f.eks. kan brukes som prefilter før tasting, for eksempel i et reguleringssystem.

10.3.2 Flere lav-ordens filtre eller et med høy orden?

Vi skal i det følgende på hvordan vi kan lage et diskret lavpassfilter ved å gå via et analogt "hjelpefilter".

Men først et viktig punkt ved praktisk digital filter-design: Filterets følsomhet for avrundingsfeil (husk at filteret er ikke annet enn en algoritme i en datamaskin med begrenset regnøyaktighet), som igjen bestemmer dets numeriske stabilitet. Nå er det slik at lavere ordens filtre er mer numerisk stabile enn direkte implementerte høyere ordens filtre. Videre vil avrundingsfeil også skape mer støy i det høyere ordens filteret, dvs. det blir mindre støy ved flere lavere ordens filtere i serie. Det er derfor hensiktsmessig å koble f.eks. tre 2. ordens filtere i serie istedet for et direkte implementert 6. ordens filter.

10.3.3 Syntese av diskret (digitalt) lavpassfilter

For å konstruere det *digitale* lavpassfilteret svarende til $H(s)$ går vi fram som følger: Vi velger først tastetid T , båndbredden $\omega_c = \Omega_c/T$ og orden N for det digitale filteret. Deretter finner vi båndbredden for et tilsvarende analogt filter, gitt av

$$\nu_c = \frac{2}{T} \tan\left[\frac{\omega_c T}{2}\right] \quad (10.16)$$

Polene i det analoge filteret er gitt av (10.15). Vi kan dermed danne $H(s)$ som et produkt av flere andre ordens transferfunksjoner i s svarende til kompleks konjugerte poler i $H(s)$, og eventuelt én 1. ordens transferfunksjon hvis N er odde. Det digitale filteret realiseres så ved å substituere $s = (2/T)(1 - z^{-1})/(1 + z^{-1})$ fra (10.10).

Vi skal se på et eksempel:

Vi ønsker å konstruere et 6. ordens diskret Butterworth-filter slik at vi har en dempning på 15 dB ($=10^{-0.75}$) ved frekvensen $\omega_k = 0.3 \pi$. Tastetida $T = 1$.

Fra (10.14) ser vi at da må vi ha

$$H(s) \cdot H(-s) = 10^{-0.75} 10^{-0.75} = \frac{1}{1 + \left(\frac{\frac{2}{T} \tan\left(\frac{\omega_k T}{2}\right)}{\nu_c}\right)^{12}} \quad (10.17)$$

Ligning (10.17) gir

$$\frac{\frac{2}{T} \tan\left(\frac{\omega_k T}{2}\right)}{\nu_c} = [10^{1.5} - 1]^{1/12} \quad (10.18)$$

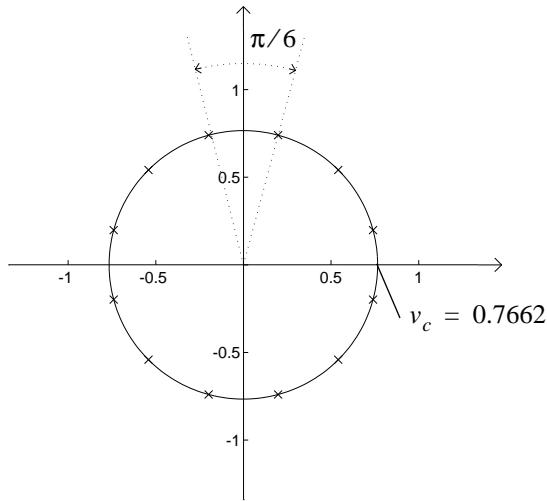
Med $T = 1$ og $\omega_k = 0.3 \pi$ får vi ved utregning

$$\nu_c = 0.7662 \quad (10.19)$$

Den resulterende analoge polplasseringen er vist i figur 10.5.

De 6 polene i venstre halvplan gir den analoge filtertransferfunksjonen

$$H(s) = (0.5871^3)H_1(s)H_2(s)H_3(s) = 0.20238 H_1(s)H_2(s)H_3(s) \quad (10.20)$$



Figur 10.5 Polplassering i s-plan for 6. ordens Butterworthfilter.

hvor

$$\begin{aligned} H_1(s) &= \frac{1}{s^2 + 0.396s + 0.5871}, \quad H_2(s) = \frac{1}{s^2 + 1.083s + 0.5871} \\ H_3(s) &= \frac{1}{s^2 + 1.4802s + 0.5871} \end{aligned} \quad (10.21)$$

Ved nå å sette inn $s = 2 \cdot \frac{1 - z^{-1}}{1 + z^{-1}}$ i (10.21) får vi

$$H_1(z^{-1}) = \frac{1}{\left(2 \cdot \frac{1 - z^{-1}}{1 + z^{-1}}\right)^2 + 0.396 \cdot 2 \cdot \frac{1 - z^{-1}}{1 + z^{-1}} + 0.5871}, \quad H_2(z^{-1}) = \dots, \quad H_3(z^{-1}) = \dots \quad (10.22)$$

$$H_1(z^{-1}) = \frac{(1 + z^{-1})^2}{(2 \cdot 1 - \dots)^2 + 0.396 \cdot 2(1^2 - \dots) + 0.5871 \cdot (1 + \dots)^2}, \quad \text{etc.} \quad (10.23)$$

$$H_1(z^{-1}) = \frac{1}{5.379} \cdot \frac{(1 + z^{-1})^2}{(1 - 1.2686z^{-1} + 0.7051z^{-2})}, \quad \text{etc.} \quad (10.24)$$

Dette gir til slutt

$$H(z^{-1}) = 0.0007378 \cdot H_1(z^{-1}) \cdot H_2(z^{-1}) \cdot H_3(z^{-1}) \quad (10.25)$$

der

$$\begin{aligned} H_1(z^{-1}) &= \frac{(1+z^{-1})^2}{1-1.2686z^{-1}+0.7051z^{-2}}, & H_2(z^{-1}) &= \frac{(1+z^{-1})^2}{1-1.0106z^{-1}+0.3583z^{-2}} \\ H_3(z^{-1}) &= \frac{(1+z^{-1})^2}{1-0.9044z^{-1}+0.2155z^{-2}} \end{aligned} \quad (10.26)$$

Konstantleddet foran i (10.25) kan om ønskelig kontrolleres slik: For $\omega = 0$, så skal

$H(z^{-1}) = H(e^{-j\Omega}) = H(e^{-j\omega T}) = H(e^0) = H(1)$ være = 1. Dette fordi filteret slipper gjennom lave frekvenser med forsterkning = 1.

For eksempelvis $y(z^{-1}) = H_1(z^{-1})x(z^{-1})$ rekursivt i tidsplan får vi da:

$$y[k] = 1.2686 \cdot y[k-1] - 0.7051 \cdot y[k-2] + u[k] + 2u[k-1] + u[k-2] \quad (10.27)$$

Den rekursive realisasjonen av den totale transferfunksjonen $H(z^{-1})$ foregår ved å kople $H_1(z)$, $H_2(z)$ og $H_3(z)$ i serie (utgangssignalet fra en algoritme blir inngangssignalet for den neste).

Frekvensresponsen for $H(z^{-1})$ er vist i figur 10.6 på neste side.

10.3.4 Realisering av andre typer filtre fra lavpassfilter

Med utgangspunkt i det kontinuerlige Butterworth-filteret, som jo er et *lavpassfilter* kan man konstruere andre typer kontinuerlige filter (**høypass-**, **båndstopp-**, **båndpass-**). Et slikt kontinuerlig filter kan så omformes til et digitalt filter ved at vi substituerer

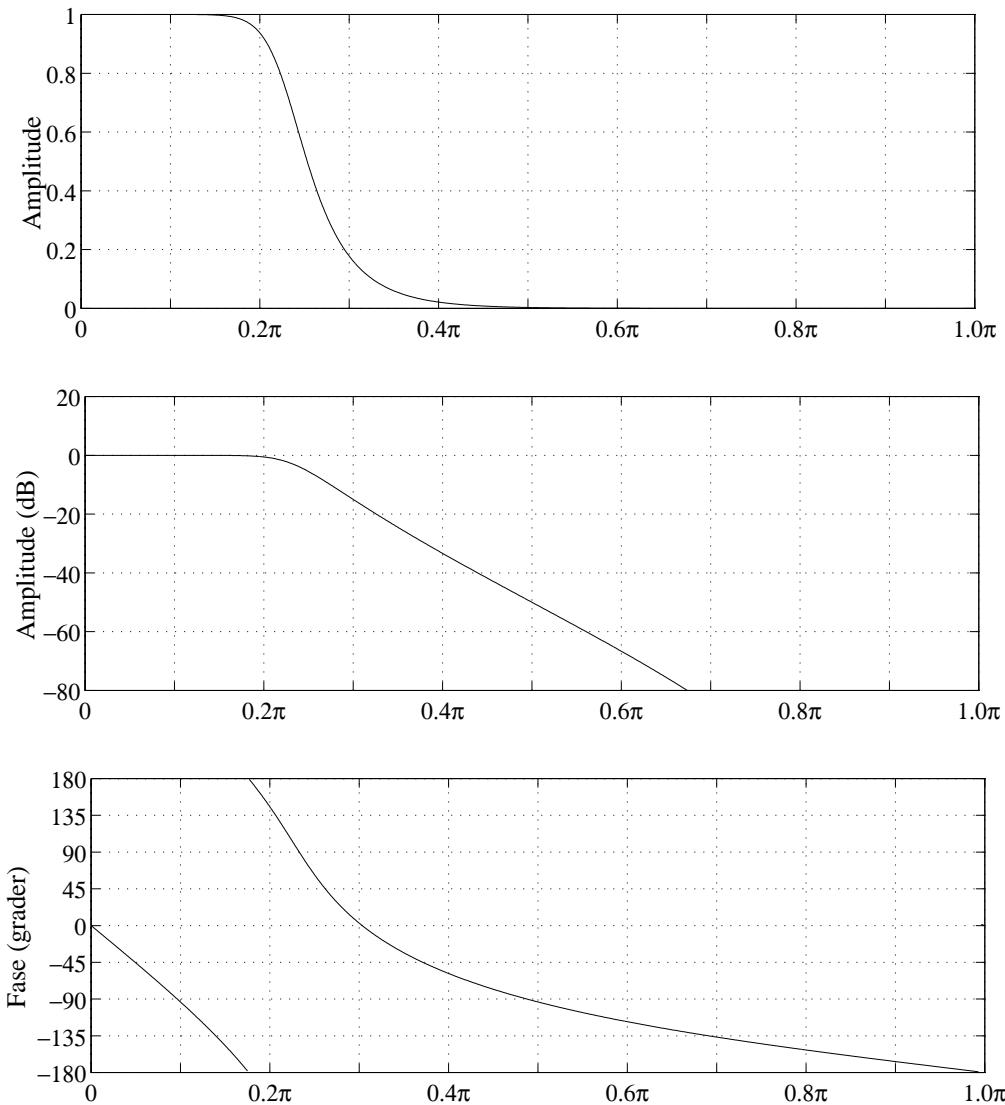
$$s = (2/T)(1-z^{-1})/(1+z^{-1}) \text{ som vist tidligere.}$$

Anta at du skal lage et kontinuerlig *høypassfilter* som gir en spesifisert forsterkning A_k ved en spesifisert frekvens v_k . Gå fram slik: Du lager først et kontinuerlig *lavpassfilter* etter prosedyren beskrevet i avsnitt 10.3.1. La oss kalle filteret du får $H_L(s)$. Erstatt så alle s i $H_L(s)$ med v_k^2/s . Det nye filteret vil da bli et *høypassfilter* med egenskapen

$$|H_H(s)|_{s=jv_k} = |H_L(s)|_{s=jv_k} = A_k \quad (10.28)$$

slik som ønsket. Hvis du skal lage et digitalt filter kan du så substituere for s .

Et *båndpassfilter* der v_l er nedre grensefrekvens og v_u er øvre grensefrekvens, har ideelt amplitudeforløp som vist i figur 10.7 (a). Forløpet til et *båndstoppfilter* er vist i figur 10.7 (b).



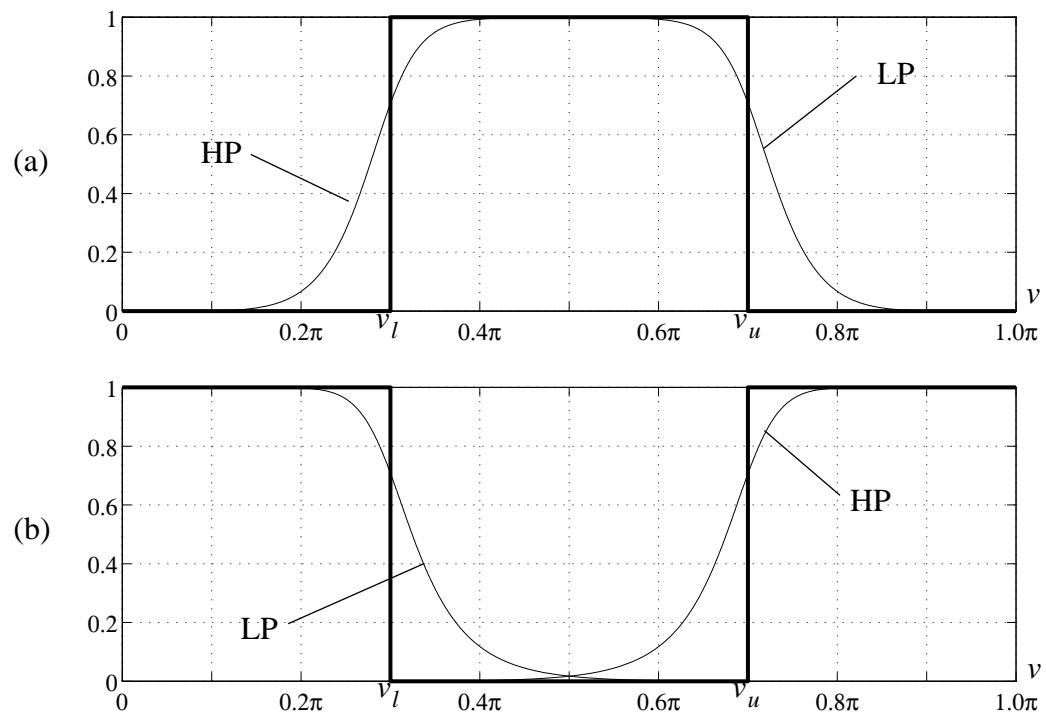
Figur 10.6 Frekvensrespons av 6. ordens Butterworthfilter med 15 dB dempning ved $\omega = 0.3\pi$.

Ved å kople et høypass- og et lavpassfilter i *serie* får vi et båndpassfilter.

Ved å kople et høypass- og et lavpassfilter i *parallel* får vi et båndstoppfilter.

Amplitudeforløpene for disse lavpass- og høypassfiltrene er vist med tynne linjer i figur 10.7 (a) og (b).

Dermed har vi prosedyrer for å lage alle fire filtertyper i analog eller digital utgave.



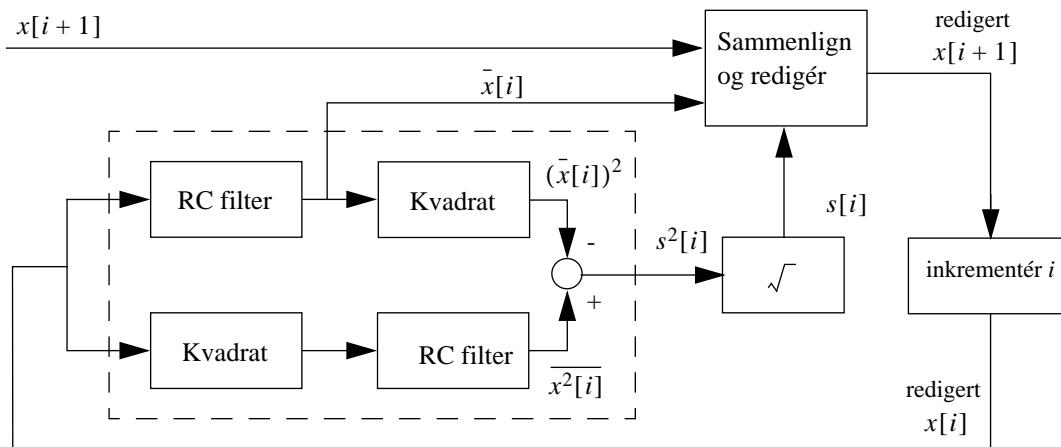
Figur 10.7

10.4 "Wild-point"- utplukking

Helt åpenbare feil kan forekomme i måledata tatt opp fra en prosess. Man ønsker å kunne automatisere prosessen med å lokalisere slike gale punkter ("wild-points"). Disse kan framkomme p.g.a. feil i datatransmisjonen, A/D-omsetter-feil, eller formatteringsfeil. Et eller flere slike gale punkter vil fort kunne innføre betydelig feil ved for eksempel beregning av effektspektra. Vi skal se på en metode for automatisk deteksjon av slike feil. Metoden baserer seg på antagelser om relativt glatte data og at feilmålingene avviker betydelig fra denne oppførselen.

Vi kaller den målte variable for x . Vi benytter to lavpassfiltre for løpende estimering av $E(x^2)$ og $E(x)$ ($= \bar{x}$).

En blokkskjema-fremstilling av algoritmen er vist i figur 10.8.



Figur 10.8 Identifisering av gale punkter.

Lavpassfiltrene gir et kontinuerlig og glidende estimat av $E(x^2)$ og \bar{x} , og dermed er estimatet $s^2[i]$ for variansen gitt ved

$$s^2[i] = E(x^2[i]) - (\bar{x}[i])^2 , \quad (10.29)$$

som svarer til (6.24) i kapittel 6.

Verdien $x[i+1]$ aksepteres så hvis f.eks.

$$\bar{x}[i] - ks[i] < x[i+1] < \bar{x}[i] + ks[i] \quad (10.30)$$

Her velges k av brukeren, f.eks. $k = 3$ til $k = 9$ med $k = 6$ som et typisk valg. Hvis man finner gale punkter basert på (10.30) kan disse løpende erstattes med en interpolert verdi.

Kapittel 11: Regulering av stokastiske systemer

- optimale regulatorer med gitt struktur

I dette kapittelet skal vi se hvordan vi med basis i det vi introduserte i kapittel 8, avsnitt 8.1, kan konstruere optimale regulatorer og reguleringssystemer for en prosess.

Vi skal forutsette stasjonære, ergodiske signaler. Vi skal utvikle teorien både for kontinuerlige systemer og signaler, og for det diskrete tilfelle.

11.1 Godhetskriterier for regulatorer

I et stokastisk system er det naturlig å assosiere kriterier for reguleringsgodhet med statistiske parametere i systemet. Hvis vi for eksempel ønsker å regulere en prosessutgang til $y(t) = y_R(t)$, der $y_R(t)$ er en referanse (ofte = 0), kan det være naturlig å minimalisere

$$J = E[\{y(t)-y_R(t)\}^2] \quad (11.1)$$

som gir midlere kvadratisk avvik mellom $y(t)$ og referanseverdien. Vi forutsetter at regulatoren gir null stasjonært avvik når vi ignorerer støyen som påvirker systemet. Hvis denne støyen har null middelverdi, vil $E(y(t)) = y_R$, slik at J representerer variansen til $y(t)$.

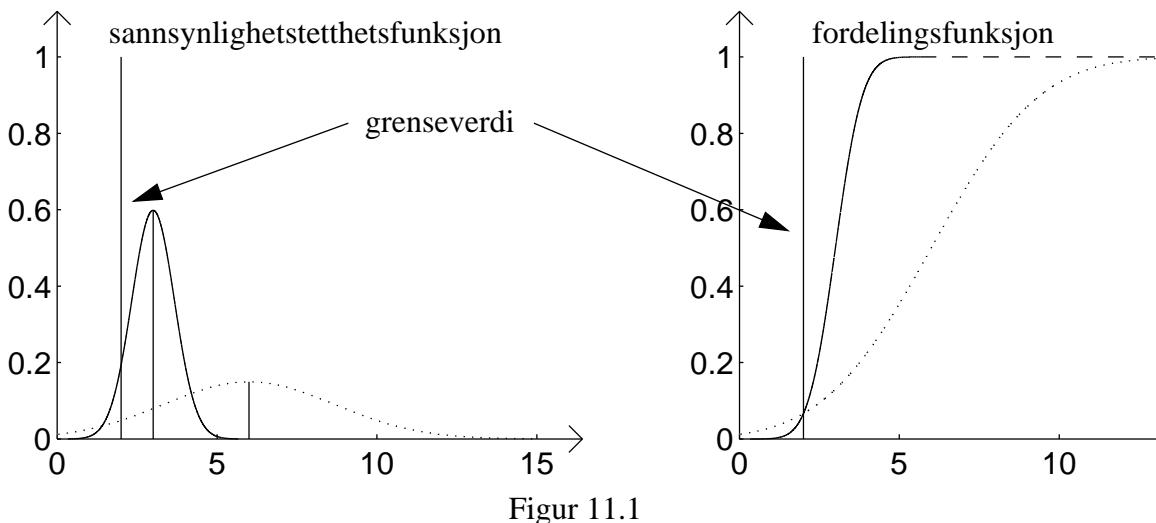
Det er imidlertid slik at en regulator hvor parametrene er valgt slik at (11.1) minimaliseres, fører til mer bruk av pådrag enn ønskelig for visse prosesser. Derfor velger man ofte et alternativt kriterium for reguleringsgodhet, nemlig

$$J = E[\{y(t)-y_R(t)\}^2] + \mu E[\{u(t)-u_R(t)\}^2] \quad (11.2)$$

hvor μ er en valgt konstant > 0 , og hvor $u_R = E[u(t)]$.

Bruk av slike **minimum-variанс**-kriterier kan være aktuelt av flere årsaker:

1. I en produksjonsprosess hvor man for eksempel produserer papir, er variansen av ulike variable et opplagt kvalitetsmål.
2. I andre prosesser spiller kanskje ikke spredningen i kvalitet og/eller dimensjoner i seg selv så stor rolle. Man vil imidlertid ligge så nær en kritisk grense som mulig. Dette ønsker man ved lagerstyring: Varianser i lagernivået er uten betydning. Men lavest mulig midlere lagerbeholdning er viktig, da denne bestemmer lagerkostnadene. Samtidig ønsker man ikke at lageret skal gå tomt når man trenger en vare. Man ønsker altså et kompromiss mellom midlere lavt lagernivå og sannsynligheten for at lageret går tomt. Problemet løses ved en regulering som minsker variansen i nivået. Man kan dermed ligge nærmere den kritiske grensa. Dette er skissert i figur 11.1.



Figur 11.1

11.2 Optimalisering av regulatorer med fast struktur

Vi skal anta i det følgende at regulatoren for et bestemt system har en gitt struktur (dvs. orden på regulatoren og type regulator er valgt, f.eks. P eller PI-regulator), og at vi ønsker å tilpasse parametrerne i denne regulatoren slik at et kriterium for reguleringens godhet av typen (11.1) eller (11.2) minimaliseres. Vi antar videre at vi har beregnet eller estimert spektrene til prosessens støysignaler, utgangssignaler, osv. Hvis vi nå for enkelhets skyld antar at $y_R = u_R = 0$ i (11.2), har vi fra (8.11) i kapittel 8 at J kan uttrykkes som

$$J = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\phi_{yy}(\omega, \underline{\alpha}) + \mu \phi_{uu}(\omega, \underline{\alpha})] d\omega \quad (11.3)$$

hvor $\underline{\alpha}$ indikerer at effektspektrene ϕ_{yy} og ϕ_{uu} avhenger av en regulatorparametervektor $\underline{\alpha}$ (for en PI-regulator vil $\underline{\alpha}$ være $\alpha_1 = K_p$ og $\alpha_2 = T_i$). I det diskrete tilfellet får vi tilsvarende (se (8.38) i kapittel 8 med $n = 0$)

$$J = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\phi_{yy}(\Omega, \underline{\alpha}) + \mu \phi_{uu}(\Omega, \underline{\alpha})] d\Omega \quad (11.4)$$

ϕ_{yy} og ϕ_{uu} er rasjonale funksjoner i $j\omega$, henholdsvis $e^{j\Omega}$, i det kontinuerlige respektive det diskrete tilfellet. Da kan vi i det kontinuerlige tilfellet skrive

$$\phi_{yy}(\omega) = \frac{P(j\omega)P(-j\omega)}{Q(j\omega)Q(-j\omega)} = \frac{P(s)P(-s)}{Q(s)Q(-s)} \quad (11.5)$$

og i det diskrete tilfellet $\phi_{yy}(\Omega) = \frac{B(e^{j\Omega})B(e^{-j\Omega})}{A(e^{j\Omega})A(e^{-j\Omega})} = \frac{B(z)B(z^{-1})}{A(z)A(z^{-1})}$ (11.6)

Tilsvarende spektral faktorisering kan foretas for ϕ_{uu} . Det er underforstått at ϕ_{yy} og ϕ_{uu} fortsatt er funksjoner av regulatorparameter-vektoren $\underline{\alpha}$, selv om vi ikke eksplisitt skriver dette.

Merk at B, A, P, Q nå er definert slik at de er polynomer i s eller z (eller z^{-1} , resultatet blir det samme uansett, som påpekt tidligere). Q har nullpunkter i venstre halvplan, mens A har nullpunkter innafor enhetssirkelen (dette p.g.a. forutsetningen om stasjonære signaler, effekt-spektrene ville ikke ha eksistert hvis dette var tilfelle).

Vi skal gå inn på hvordan J beregnes, og vi betrakter først det kontinuerlige tilfellet. Ved innføring av den nye variable $s = j\omega$, og ved å benytte (11.5) får vi, i det μ antas å være null (dette er kun for å lette framstillingen): $d\omega = ds/j$. Når ω går fra $-\infty$ til $+\infty$ går s fra $-j\omega$ til $+j\omega$, og da blir (11.3)

$$J = \Phi_{yy}(0) = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \frac{P(s)P(-s)}{Q(s)Q(-s)} ds \quad (11.7)$$

$$\begin{aligned} P(s) &= p_0 + p_1 s + p_2 s^2 + \dots + p_{n-1} s^{n-1} \\ \text{hvor } Q(s) &= q_0 + q_1 s + q_2 s^2 + \dots + q_n s^n \end{aligned} \quad (11.8)$$

Et bestemt sett regulatorparametre $\underline{\alpha}$ fører til et tilsvarende sett koeffisienter p_i og q_i , og dermed til en tilsvarende integrand i (11.7). Vi kan nå benytte residueregning slik som beskrevet i kapittel 1 for å finne J for dette settet av regulatorparametre $\underline{\alpha}$.

En systematisk oversikt over J som funksjon av p_i og q_i er vist i tabell form i tabell 11.1 for systemets orden $n \leq 4$.

$J_1 = \frac{p_0^2}{2q_0 q_1}$
$J_2 = \frac{p_1^2 q_0 + p_0^2 q_2}{2q_0 q_1 q_2}$
$J_3 = \frac{p_2^2 q_0 q_1 + (p_1^2 - 2p_0 p_2) q_0 q_3 + p_0^2 q_2 q_3}{2q_0 q_3 (-q_0 q_3 + q_1 q_2)}$
$J_4 = \frac{p_3^2 (-q_0^2 q_3 + q_0 q_1 q_2) + (p_2^2 - 2p_1 p_3) q_0 q_1 q_4}{2q_0 q_4 (-q_0 q_3^2 - q_1^2 q_4 + q_1 q_2 q_3)} \\ + \frac{(p_1^2 - 2p_0 p_2) q_0 q_3 q_4 + p_0^2 (-q_1 q_4^2 + q_2 q_3 q_4)}{2q_0 q_4 (-q_0 q_3^2 - q_1^2 q_4 + q_1 q_2 q_3)}$

Tabell 11.1 J_n for $n = 1, 2, 3, 4$

En FORTRAN subrutine for beregning av J i det kontinuerlige tilfellet er vist i det følgende: $B(s)$ i programmet tilsvarer $P(s)$, og $A(s)$ tilsvarer $Q(s)$.

```

SUBROUTINE COLOSS (A, B, N, IERR, V, IN)
C
C PROGRAM FOR EVALUATING THE INTEGRAL OF THE RATIONAL
C FUNCTION
C      1/(2*PI*I)*B(S)*B(-S)/(A(S)*A(-S))
C ALONG THE IMAGINARY AXIS
C
C A—VECTOR WITH THE COEFFICIENTS OF THE POLYNOMIAL
C      A(1)*S**N + A(2)*S**(N-1) + ...+ A(N + 1)
C IT IS ASSUMED THAT A(I) IS POSITIVE
C B—VECTOR WITH THE COEFFICIENTS OF THE POLYNOMIAL
C      B(1)*S**(N-1) + B(2)*S**(N-2) + ...+ B(N)
C
C      THE VECTORS A AND B ARE DESTROYED
C
C N—ORDER OF THE POLYNOMIALS A AND B
C IERR—WHEN RETURNING IERR=1 IF ALL ZEROS OF A ARE IN LEFT
C      HALF PLANE IERR=0 IF THE POLYNOMIAL A DOES NOT HAVE
C      ALL ZEROS IN LEFT HALF PLANE OR IF A(1) IS NOT POSITIVE
C V—THE RETURNED LOSS
C IN—DIMENSION OF A AND B IN MAIN PROGRAM
C
C SUBROUTINE REQUIRED
C      NONE
C
C      DIMENSION A(IN), B(IN)
C
C IERR = 1
C V = 0.0
C IF (A(1)) 70, 70, 10
10 DO 20 K = 1, N
    IF (A(K + 1)) 70, 70, 30
30 ALFA = A(K)/A(K + 1)
    BETA = B(K)/A(K + 1)
    V = V + BETA**2/ALFA
    K1 = K + 2
    IF (K1=N) 50, 50, 20
50 DO 60 I = K1, N, 2
    A(I) = A(I) - ALFA*A(I + 1)
60 B(I) = B(I) - BETA*A(I + 1)
20 CONTINUE
    V = V/2.0
    RETURN
70 IERR = 0
    RETURN
END

```

Betegnelsen "COLOSS" henspiller på at man beregner J for et "COntinuous" system, og at kriteriet J i engelsk terminologi ofte kalles "LOSS function" (tapsfunksjon). Algoritmen er rekursiv og beregner J_n ved først å beregne objektfunksjoner av orden 1, 2, ..., $n - 1$.

(Det forlanges ikke at du skal forstå virkemåten til algoritmen, poenget her er bare å illustrere at det finnes et rekursivt alternativ til residueregning, siden residueregning blir uhyre komplisert for store n).

For det *diskrete* tilfellet gjelder (11.4). Vi antar igjen at $\mu = 0$, innfører $z = e^{j\Omega}$ og får $dz = je^{j\Omega}d\Omega \Rightarrow d\Omega = \frac{1}{jz} dz$; og dette gir (jfr. (2.26) i kapittel 2, med diskret tid $n = 0$):

$$J = \Phi_{yy}[0] = \frac{1}{2\pi j} \oint_{|z|=1} \frac{B(z)B(z^{-1})}{A(z)A(z^{-1})} \frac{1}{z} dz \quad (11.9)$$

hvor integrasjonen foregår langs enhetssirkelen definert ved $|z| = 1$.

Vi har at

$$A(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n$$

$$B(z) = b_0 z^n + b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n$$
(11.10)

Man kan også her benytte residue-regning for å finne J i (11.9).

En rekursiv subroutine for beregning av J er gitt i det følgende.

```

SUBROUTINE SALOSS (A, B, N, IERR, V, IN)
C
C PROGRAM FOR EVALUATING THE INTEGRAL OF THE RATIONAL
C FUNCTION
C   1/(2*PI*I)*B(Z)*B(I/Z)/(A(Z)*A(I/Z))**(1/Z)
C AROUND THE UNIT CIRCLE
C
C A—VECTOR WITH THE COEFFICIENTS OF THE POLYNOMIAL
C   A(1)*Z**N + A(2)*Z***(N-1) + ... + A(N+1)
C B—VECTOR WITH THE COEFFICIENTS OF THE POLYNOMIAL
C   B(1)*Z**N + B(2)*Z***(N-1) + ... + B(N+1)
C
C THE VECTORS A AND B ARE DESTROYED
C
C N—ORDER OF THE POLYNOMIALS A AND B (MAX 10)
C IERR—WHEN RETURNING IERR=1 IF A HAS ALL ZEROS INSIDE UNIT
C CIRCLE IERR=0 IF THE POLYNOMIAL A HAS ANY ROOT OUTSIDE
C OR ON THE UNIT CIRCLE OR IF A(1) IS NOT POSITIVE
C V—THE RETURNED LOSS
C IN—DIMENSION OF A AND B IN MAIN PROGRAM
C
C SUBROUTINE REQUIRED
C   NONE
C
C DIMENSION A(IN), B(IN), AS(11)
C
C
      A0 = A(1)
      IERR=1
      V = 0.0
      DO 10 K = 1, N
      L=N+1-K
      L1 = L+1
      ALFA = A(L1)/A(1)
      BETA = B(L1)/A(1)
      V = V + BETA*B(L1)
      DO 20 I= 1, L
      M=L+2-I
      AS(I) = A(I)-ALFA*A(M)
 20   B(I) = B(I)-BETA*A(M)
      IF (AS(1)) 50, 50, 30
 30   DO 40 I = 1, L
 40   A(I) = AS(I)
 10   CONTINUE
      V = V + B(1)**2/A(1)
      V = V/A0
      RETURN
 50   IERR = 0
      RETURN
      END

```

“SALOSS” henspiller her på “SAmpled LOSS”.

Til seinere bruk skal vi gi analytisk beregnede uttrykk for J når $n = 1$ og $n = 2$:

$$J_1 = \frac{a_0(b_0^2 + b_1^2) - 2b_0b_1a_1}{a_0(a_0 - a_1)(a_0 + a_1)}$$
(11.11)

$$J_2 = \frac{a_0(a_0 + a_2)(b_0^2 + b_1^2 + b_2^2) - 2a_0a_1b_1(b_0 + b_2) - 2b_0b_2(a_0a_2 - a_1^2 + a_2^2)}{a_0(a_0 - a_2)(a_0 - a_1 + a_2)(a_0 + a_1 + a_2)}$$
(11.12)

Vi skal i det følgende se på to eksempler.

Eksempel 1: Å styre et lager optimalt (diskret regulering):

(Hentet fra L. Ljung: "Reglerteori - moderna analys- og syntesmetoder", Lund 1981).

Vi antar at lagerbeholdningen kan uttrykkes som et reelt tall, dvs. enten som volum eller vekt, eller som et antall enheter som er så stort at lageret kan sies å variere kontinuerlig. La $\bar{y} + y[k]$ være lagerbeholdningen ved slutten av dag k . La $\bar{u} + u[k]$ være en bestilling i løpet av dag k for komplettering av lageret, og la $\bar{z} + z[k]$ være lageruttaket i løpet av dag k . Merk at de tre variable alle er modellert som en konstant middelverdi + avvik fra denne middelverdien. $y[k]$, $u[k]$ og $z[k]$ kan gjerne bli negative uten at dette er ufysisk, siden de representerer variasjoner om midlere verdier.

Oppførselen til kundene betraktes som en stokastisk prosess. Etter lengre tids logging av denne, har man oppdaget at uttak fra lageret er korrelert i tid, dvs. stort uttak en dag medfører vanligvis stort uttak også neste dag. En modell på formen

$$\bar{z} + z[k] = \bar{z} + (v[k] + av[k-1]) \quad (11.13)$$

viser seg å gi en god beskrivelse av de faktiske tidsserier. Her er $v[k]$ hvit støy med null middelverdi og varians 1.

I det man regner to dager fra utsendt bestilling på nye varer for å komplettere lageret, til mottak av varene, blir modellen:

$$\bar{y} + y[k] = \bar{y} + y[k-1] + \bar{z} + z[k] + \bar{u} + u[k-2] \quad (11.14)$$

her er \bar{y} og \bar{u} konstante verdier for henholdsvis midlere lagernivå og midlere bestilt varemengde pr. dag. \bar{y} kan trekkes fra på begge sider, og \bar{u} vil være lik $-\bar{z}$, slik at disse kansellerer hverandre. Da blir den dynamiske delen av modellen, når vi setter inn for $z[k]$ med $a = 0.9$:

$$y[k] = y[k-1] + u[k-2] + v[k] + 0.9v[k-1] \quad (11.15)$$

Vi vil nå forsøke å styre lageret med en proporsjonalregulator

$$u[k] = -Ky[k] \quad (11.16)$$

Dette gir $y[k] = \frac{1 + 0.9z^{-1}}{1 - z^{-1} + Kz^{-2}}v[k] = \frac{B(z^{-1})}{A(z^{-1})}v[k]$ (11.17)

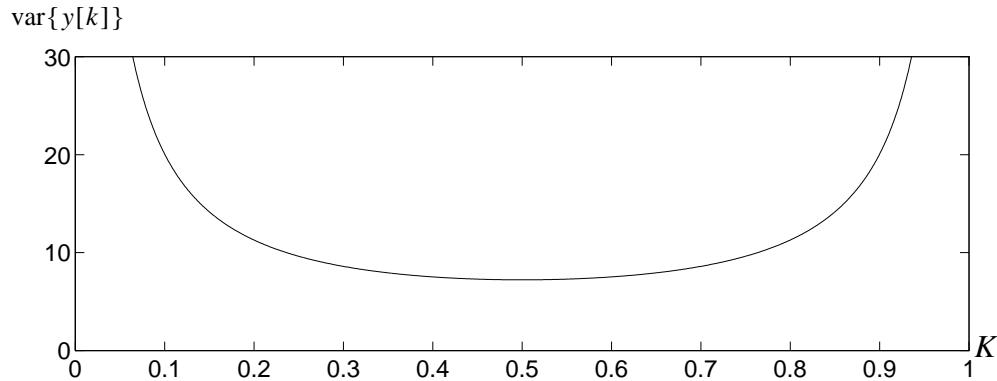
Siden $\phi_{vv}(\Omega) = 1$, får vi

$$\phi_{yy}(z) = \frac{B(z)B(z^{-1})}{A(z)A(z^{-1})} \quad (11.18)$$

Her er $a_0 = 1$, $a_1 = -1$, $a_2 = K$, $b_0 = 1$, $b_1 = 0.9$ og $b_2 = 0$. Ved bruk av (11.12) får vi

$$J_2 = \text{var}(y[k]) = \frac{(1+K)1.81 + 1.8}{(1-K)(K+2)K} \quad (11.19)$$

Denne funksjonen av K er vist i figur 11.2.



Figur 11.2 $\text{var}(y[k])$ som funksjon av proporsjonalforsterkningen, K .

J_2 får minimum $J_{2,MIN} = 7.23$ for $K = 0.5$. Vi ser at $J_2 \rightarrow \infty$ når $K \rightarrow 1$. Dette skyldes at systemet da blir ustabilt. Når $K \rightarrow 0$, vil også $J_2 \rightarrow \infty$, men her av andre årsaker, nemlig at integralet av en stokastisk prosess vil ha lineært økende varians med tida (ligning (11.15) inneholder en diskret integrator, noe vi også ser av (11.17) hvis vi setter $K = 0$).

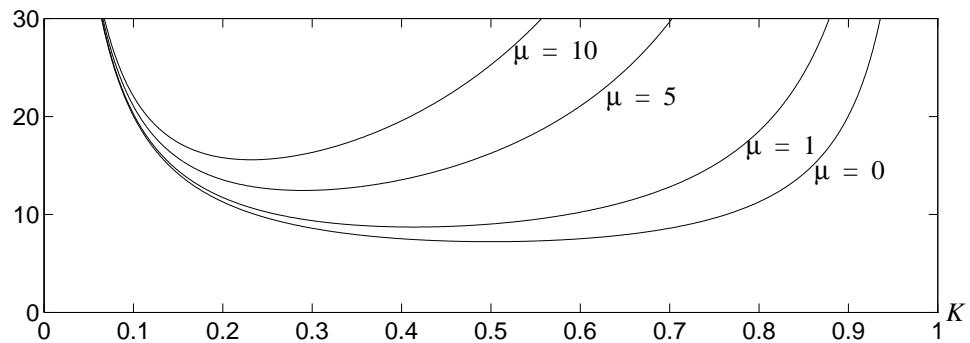
Anta så at store avvik i bestillingen fra gang til gang er uønsket, dvs. avvik ($u[k] \neq 0$) fra normalbestillingen bør unngås hvis mulig. Vi benytter da $J_2 = E(y^2[k]) + \mu E(u^2[k])$ som det nye kriteriet hvor nå $\mu > 0$. Da $u[k] = -Ky[k]$, blir

$$J_2 = (1 + \mu K^2) \text{var}(y[k]) = \frac{(1 + \mu K^2)(3.61 + 1.81K)}{(1 - K)(K + 2)K} \quad (11.20)$$

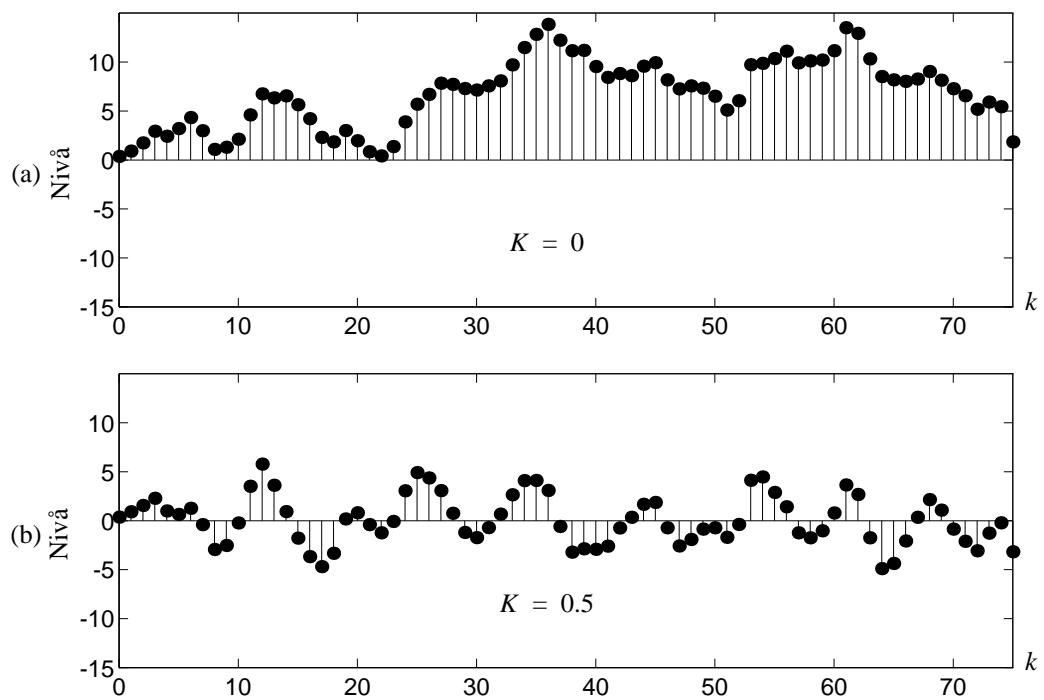
J_2 er nå skissert som funksjon av K og for ulike μ i figur 11.3.

Vi ser at den optimale K minker når μ øker. For $\mu = 1$ vil for eksempel $K_{MIN} = 0.41$ og $J_{2,MIN} = 8.72$. Når $\mu = 0$, blir $J_{2,MIN} = 7.23$ for $K_{MIN} = 0.5$ som før.

I figur 11.4 er dessuten vist lagernivåsvingninger for det uregulerte tilfellet (a), og når $K = 0.5$ (b). I tilfellet (a) ser vi at variansen øker med tiden. Prosessen er da ikke stasjonær.



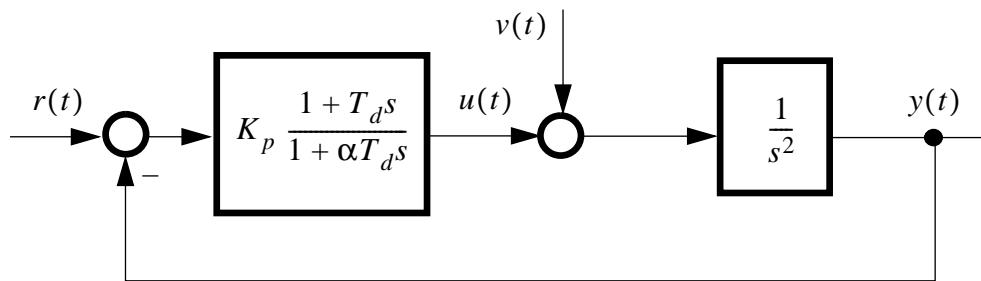
Figur 11.3 $J = \text{var}(y[k]) + \mu \text{ var}(u[k])$ for ulike μ og K .



Figur 11.4 Svingninger i lagerbeholdning uten og med regulering.

Eksempel 2: Dobbeltintegrator (kontinuerlig regulering):

Vi skal nå se på syntese av kontinuerlig regulator for et kontinuerlig system. Systemet er kjent fra Balchen: "Reguleringsteknikk", s. 286 - 288, eller Finn Haugen: "Anvendt reguleringsteknikk", s. 428 - 432. Vi har en dobbelt integrator. Vi vet at en rein proporsjonalregulator er ubrukelig for et slikt system, for da vil systemet - uansett forsterkning K_p - ligge og oscillere på stabilitetsgrensa med egenfrekvens $\omega_0 = \sqrt{K_p}$ (= kryssfrekvensen). Vi må løfte fasen rundt kryssfrekvensen, og fra Balchen eller Haugen vet vi at en akseptabel løsning i dette tilfellet er en PD-regulator, som i praksis alltid må være en *begrenset* PD-regulator. Det regulerte system er vist i figur 11.5.



Figur 11.5

Her er $\alpha < 1$, typisk verdi = 0.1.

Systemet antas utsatt for hvit støy, med autoeffektspektrum $\phi_{vv}(\omega) = 1$.

Det lukkede systems transferfunksjon fra støyen til utgangen blir

$$H_{vy}(s) = \frac{1 + \alpha T_d s}{\alpha T_d s^3 + s^2 + K_p T_d s + K_p} \quad (11.21)$$

Effektspektret $\phi_{yy}(\omega)$ på utgangen blir nå

$$\phi_{yy}(\omega) = H_{vy}(j\omega)H_{vy}(-j\omega)\phi_{vv}(\omega) \quad (11.22)$$

Vi skal nå benytte (11.7) og tabell 11.1. $\phi_{yy}(\omega)$ kan faktoriseres som

$$\phi_{yy} = H_{vy}(s)H_{vy}(-s)\phi_{vv} = \frac{P(s)P(-s)}{Q(s)Q(-s)} \cdot 1 \quad (11.23)$$

dvs. vi har

$$\begin{aligned} P(s) &= p_0 + p_1 s + p_2 s^2 = 1 + \alpha T_d s \\ Q(s) &= q_0 + q_1 s + q_2 s^2 + q_3 s^3 = K_p + K_p T_d s + s^2 + \alpha T_d s^3 \end{aligned} \quad (11.24)$$

Hvis vi bruker tabell 11.1, 3. linje, får vi etter forenkling og forkortning

$$J_3 = \frac{1 + K_p(\alpha T_d)^2}{2K_p^2 T_d(1 - \alpha)} \quad (11.25)$$

Vi ser umiddelbart at vi kan redusere $\text{var}[y(t)]$ så mye vi vil ved å øke K_p , så lenge $T_d > 0$. Nå vil vi måtte velge en endelig K_p av praktiske årsaker (blant annet vil pådraget bli svært kraftig hvis K_p er stor. Dette kunne man tatt hensyn til ved å trekke pådraget inn i optimalkriteriet slik som vi gjorde i siste del av lagerstyringseksemplet. Men for å få et enklere regneeksempel gjør vi ikke det her).

Vi antar i det følgende at K_p er valgt ut fra en betrakting om kompromiss mellom hurtighet og pådragsbruk. Vi ønsker nå å finne den optimale T_d når K_p er gitt.

Vi søker T_d slik at $\frac{dJ_3}{dT_d} = 0$. Dette vil gi den T_d som gir minimum av $\text{var}[y(t)]$ når K_p er valgt. Fra (11.25) får vi

$$\begin{aligned} \frac{\partial J_3}{\partial T_d} &= \frac{\partial}{\partial T_d} \left(\frac{1 + K_p(\alpha T_d)^2}{2K_p^2 T_d(1 - \alpha)} \right) = \\ &= \frac{2K_p^2 T_d(1 - \alpha)K_p(2\alpha^2 T_d) - (1 + K_p(\alpha T_d)^2)(2K_p^2(1 - \alpha))}{[2K_p^2 T_d(1 - \alpha)]^2} \end{aligned} \quad (11.26)$$

Dette gir $\frac{dJ_3}{dT_d} = 0$ når

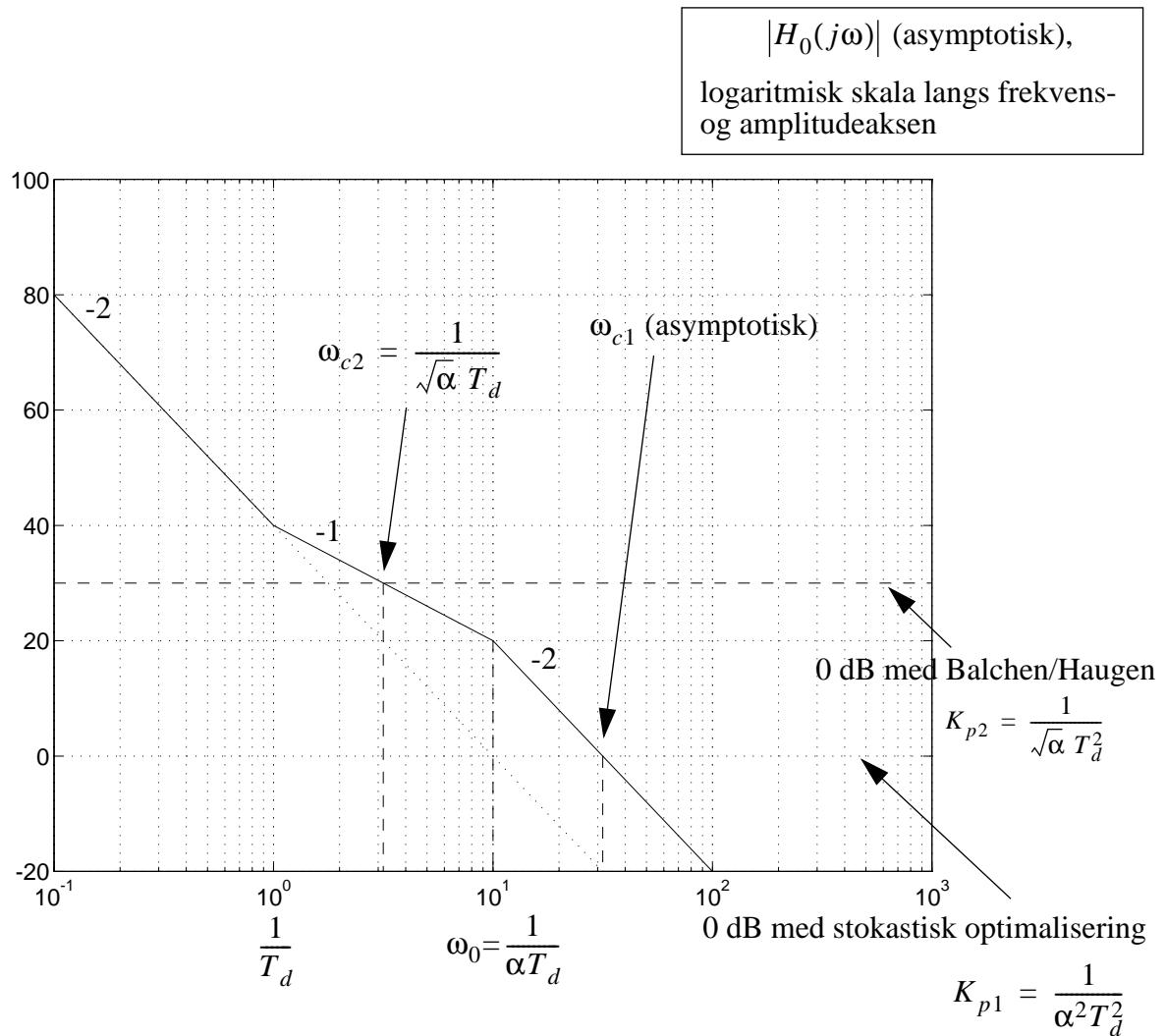
$$T_d K_p (2\alpha^2 T_d) - (1 + K_p(\alpha T_d)^2) = 0, \text{ dvs. når } T_d = \frac{1}{\alpha \sqrt{K_p}} \quad (11.27)$$

(Hvis støyen ikke hadde vært hvit, ville den optimale T_d blitt en annen. Beregningene hadde da blitt mer kompliserte. Vi måtte da først ha faktorisert støyspektrum og brukt metoden i kapittel 8. Den er forklart i tilknytning til figur 8.2)

Hvis vi løser (11.27) m.h.p. K_p får vi:

$$K_{p1} = \frac{1}{\alpha^2 T_d^2} \quad (11.28)$$

og 0 dB-linja finnes ved å forlange $\omega_0 = \sqrt{K_{p1}}$, se Bode-diagrammet i figur 11.6.



Figur 11.6 Amplitudeforløp av sløyfetransferfunksjon

Vi finner så den asymptotiske kryssfrekvensen $\omega_c1 = \frac{1}{\sqrt{\alpha\alpha T_d}}$ som vist på figur 11.6. Vi ser at den blir større enn den ω_c2 som vi får med den seriekompensasjons-baserte metoden brukt i Balchen eller Haugen, forklart nedenfor.

Seriekompensasjon i følge Balchen/Haugen går som følger:

ω_c2 velges som det geometriske middel mellom de to knekkfrekvensene i den begrensede PD-regulatoren. Dette fører nemlig til at ω_c2 gir et maksimum av $\angle H_0(j\omega)$, dvs. maksimal fasemargin. Geometrisk middel betyr at ω_c2 blir

$$\omega_c2 = \sqrt{\frac{1}{\alpha T_d} \cdot \frac{1}{T_d}} = \frac{1}{\sqrt{\alpha} T_d} \quad (11.29)$$

Vi kan så bruke ligninga

$$|H_0(j\omega_{c2})| = K_{p2} \frac{1}{\omega_{c2}^2} \sqrt{\frac{1 + (T_d \omega_{c2})^2}{1 + (\alpha T_d \omega_{c2})^2}} = 1, \quad (11.30)$$

til å finne K_{p2} , som - etter innsetting for ω_{c2} og en del mellomregning - blir

$$K_{p2} = \frac{1}{\sqrt{\alpha} T_d^2} \quad (11.31)$$

Det viktige å oppsummere her, er at forsterkinga og kryssfrekvensen blir markant høyere ved stokastisk optimalisering enn ved Balchen/Haugen-metoden, som er basert på kravet om maksimal fasemargin. Forskjellen i forsterkning er stor. (11.28) og (11.31) gir

$$K_{p1} = \frac{1}{\alpha \sqrt{\alpha}} K_{p2} \quad (11.32)$$

For $\alpha = 0.1$ gir (11.32) en forskjell på så mye som 30 dB.

Den stokastiske optimaliseringen tar utelukkende hensyn til kravet om å minimalisere effekten på utgangen, uten å skjele til fasemargin. Vi må dermed forvente mindre fasemargin med denne metoden (den blir bare ca. 15° for $\alpha = 0.1$, mens Balchen/Haugen-metoden gir 55°).

Dette er risikabelt, fordi det alltid vil være et visst avvik mellom den modell vi bruker, og det virkelige fysiske system. Derfor ønsker vi en viss stabilitetsmargin. I en praktisk regulerings-situasjon ville vi derfor redusere forsterkingen noe ut fra Bode-diagram og en stabilitetsmargin-betraktnng, hvis vi først hadde beregnet et parametersett K_p og T_d via en slik stokastisk optimaliseringssprosedyre.

Kapittel 12: Regulering av stokastiske systemer - optimale regulatorer med optimal struktur

I forrige avsnitt diskuterte vi minimalisering av varians¹ når regulatoren har en fast, gitt struktur, men med ukjente parametere. Vi skal i dette kapittelet betrakte minimalisering av varians når også regulatorstrukturen kan velges optimalt.

Vi antar som før at systemet er monovariabelt, lineært og tidsinvariant, og at forstyrrelsen som eksiterer systemet er en stasjonær, stokastisk prosess med rasjonalt effektspektrum. Prosessmodell og støyspektrum forutsettes kjent. Systemer og signaler antas diskrete i tid.

Vi skal etterhvert se at den såkalte dead-beat-regulator som vi har utledet tidligere, faktisk også vil være akkurat den regulator som vil oppfylle kravet til minimum effekt, og at det ikke finnes noen annen diskret regulator som gir mindre effekt på utgangen. Derfor vil jeg hyppig henvise til kapittel 5 i det følgende.

12.1 Regulatoralgoritme

Vi antar at prosessen kan beskrives av en ARMAX-modell (ligning (5.1) i kapittel 5)

$$A(z^{-1})y[k] = z^{-m}Bz^{-1}u[k] + C(z^{-1})\varepsilon[k] \quad (12.1)$$

Men nå er $\varepsilon[k]$ ikke lenger en deterministisk forstyrrelse, f.eks. en diskret enhetspuls. $\varepsilon[k]$ er nå diskret hvit støy med autoeffektspektrum = konst. = $\phi_{\varepsilon\varepsilon}$. Som i kapittel 11 ønsker vi å minimalisere $E(y^2[k])$ ved tilbakekopling (vi skal ikke gå inn på veiing av pådraget i optimalkriteriet i dette kapittelet). Vi skal nå omforme (12.1) for å kunne vise at dead-beat-regulatoren (5.11) i kapittel 5 gir minimum varians når $\varepsilon[k]$ er hvit støy.

I det følgende dropper vi for enkelhets skyld argumentet z^{-1} i polynomene A, B, C , det er underforstått.

Ved å inkrementere tida m skritt framover, og ved å dele på A på begge sider, gir (12.1)

$$y[k+m] = \frac{B}{A}u[k] + \frac{C}{A}\varepsilon[k+m] \quad (12.2)$$

Ved å benytte polynomidentiteten (5.5) fra kapittel 5

$$C \equiv AF + z^{-m}G \quad (12.3)$$

kan vi skrive (12.2) som

$$y[k+m] = F\varepsilon[k+m] + \frac{B}{A}u[k] + \frac{G}{A}\varepsilon[k] \quad (12.4)$$

1. Egentlig: effekt, men effekten er lik variansen når utgangens middelverdi = 0)

Hvis vi løser (12.1) m.h.p. $\varepsilon[k]$, får vi

$$\varepsilon[k] = \frac{A}{C}y[k] - \frac{B}{C}z^{-m}u[k] \quad (12.5)$$

Dette innsatt i (12.4) gir

$$y[k+m] = F\varepsilon[k+m] + \frac{B}{A}u[k] + \frac{G}{A}\left(\frac{A}{C}y[k] - \frac{B}{C}z^{-m}u[k]\right) \quad (12.6)$$

eller $y[k+m] = F\varepsilon[k+m] + \left[\frac{B}{A} - \frac{GB}{AC}z^{-m}\right]u[k] + \frac{G}{C}y[k] \quad (12.7)$

Uttrykket i hakeparentesen kan skrives som

$$\frac{B}{C}\left[\frac{C}{A} - \frac{z^{-m}G}{A}\right] \quad (12.8)$$

Ved å benytte polynomidentiteten (12.3) får vi

$$\frac{B}{C}\left[\frac{C}{A} - \frac{z^{-m}G}{A}\right] = \frac{B}{C}\left[\frac{AF}{A}\right] = \frac{BF}{C} \quad (12.9)$$

og (12.7) blir dermed

$$y[k+m] = F\varepsilon[k+m] + \frac{BF}{C}u[k] + \frac{G}{C}y[k] \quad (12.10)$$

(12.10) er ekvivalent med ARMAX-modellen (12.1), men (12.10) gir oss ny innsikt. Det første ledet i (12.10) er

$$\begin{aligned} F\varepsilon[k+m] &= (1 + f_1z^{-1} + f_2z^{-2} + \dots + f_{m-1}z^{-(m-1)})\varepsilon[k+m] \\ &= \varepsilon[k+m] + f_1\varepsilon[k+m-1] + f_2\varepsilon[k+m-2] + \dots + f_{m-1}\varepsilon[k+1] \end{aligned} \quad (12.11)$$

Siden $\varepsilon[k]$ er hvit støy med middelverdi = 0, får vi da at den betingede forventningsverdien til $y[k+m]$ når $y[i]$ og $u[i]$ er kjent for $i \leq k$ blir

$$\begin{aligned} E(y[k+m|k]) &= \bar{y}[k+m|k] = \\ (1 + f_1 + \dots + f_{m-1})E(\varepsilon[k]) + \frac{BF}{C}u[k] + \frac{G}{C}y[k] &= \frac{BF}{C}u[k] + \frac{G}{C}y[k] \end{aligned} \quad (12.12)$$

og effekten som vi skal minimalisere blir

$$\begin{aligned}
 E(y^2[k]) &= E(y^2[k+m]) = E\left[\left(\frac{BF}{C}u[k] + \frac{G}{C}y[k]\right) + F\varepsilon[k+m]\right]^2 = \\
 &E\left[\left(\frac{BF}{C}u[k] + \frac{G}{C}y[k]\right)^2\right] + 2E(\bar{y}[k+m|k]F\varepsilon[k+m]) + E(F\varepsilon[k+m])^2
 \end{aligned} \tag{12.13}$$

Nå er $\varepsilon[k+1], \varepsilon[k+2], \dots, \varepsilon[k+m]$ ukorrelerte med $y[i]$ og $u[i]$ for $i \leq k$. Vi bruker $y[i]$ for $i \leq k$ og $u[i]$ for $i < k$, for å regne ut nytt pådrag $u[k]$ ved diskret tid k . Dermed kan vi ikke påvirke det siste ledet i (12.13). Det inneholder bare framtidige utfall av $\varepsilon[k]$. Det midterste ledet (kryss-ledet) blir = 0, fordi $\bar{y}[k+m|k]$ er ukorrelert med framtidige utfall av $\varepsilon[k]$. Dermed er det bare det første ledet, som er identisk med $\bar{y}[k+m|k]$, vi kan påvirke ved vårt valg av pådrag $u[k]$. Minimumsverdien av dette ledet er sjølsagt 0 når $\bar{y}[k+m|k] = 0$.

(12.12) gir

$$\frac{BF}{C}u[k] + \frac{G}{C}y[k] = 0 \Leftrightarrow u[k] = -\frac{G}{BF}y[k] \tag{12.14}$$

(12.14) er det pådrag som gir minimum effekt i $y[k+m]$. Denne regulatoren kalles i litteraturen for **“minimum-varians-regulatoren”** for ARMAX-modellen (12.1).

Vi oppsummerer:

Anta en prosess beskrevet av

$$A(z^{-1})y[k] = z^{-m}B(z^{-1})u[k] + C(z^{-1})\varepsilon[k] \tag{12.15}$$

hvor $\varepsilon[k]$ er diskret hvit støy med autoeffektspektrum = konst. = $\phi_{\varepsilon\varepsilon}$.

Da er minimum-varians-regulatoren gitt av

$$u[k] = -\frac{G}{BF}y[k] \tag{12.16}$$

hvor polynomene F og G er av orden $m - 1$ og $n - 1$, og definert ved identiteten

$$C \equiv AF + z^{-m}G \tag{12.17}$$

Vi ser at (12.14) er identisk med dead-beat-regulatoren for det deterministiske tilfelle. Mer om det seinere.

Merknad 1:

Siden $E(y[k]) = 0$ vil effekten i signalet $E(y^2[k]) = \text{var}(y[k])$. Effekten på utgangen med MV-regulering blir (jfr. (12.13) med ledet til venstre = 0):

$$\begin{aligned}\text{var}(y[k]) &= E(F\varepsilon[k+m])^2 \\ &= E(\varepsilon[k+m] + f_1\varepsilon[k+m-1] + \dots + f_{m-1}\varepsilon[k+1])^2 \\ &= (1 + f_1^2 + \dots + f_{m-1}^2)E(\varepsilon^2[k]) = (1 + f_1^2 + \dots + f_{m-1}^2)\phi_{\varepsilon\varepsilon}\end{aligned}\quad (12.18)$$

Hvis vi anvender en annen regulator, men slik at det lukkede system fortsatt er stabilt, vil effekten (= variansen) bli større, og gitt av (12.13), nå med ledet til venstre $\neq 0$.

Siste linje i (12.18) følger av at alle ledd av typen $E(\varepsilon[i]\varepsilon[j]) = 0$ når $j \neq i$, og at $E(\varepsilon[i]\varepsilon[i]) = \phi_{\varepsilon\varepsilon}$ pr. forutsetning.

Forøvrig har vi et resultat i (8.50) som stemmer med (12.18)

$$E(y^2[k]) = Q_D \sum_{j=0}^{\infty} h^2[j] = \phi_{\varepsilon\varepsilon} \sum_{j=0}^{m-1} f^2[j] \quad (12.19)$$

Merknad 2:

Reguleringsavviket, definert som forskjellen mellom virkelig utgang og det vi predikerte at den ville bli da vi var m tidsskritt tilbake i tid, er gitt ved

$$\Delta y[k+m] = y[k+m] - \bar{y}[k+m|k] = F\varepsilon[k+m] \quad (12.20)$$

som ekvivalent kan skrives

$$\Delta y[k] = y[k] - \bar{y}[k|k-m] = F\varepsilon[k] \quad (12.21)$$

Merknad 3:

Fra (12.21) har vi at reguleringsfeilen er gitt av $\Delta y[k] = F\varepsilon[k]$. Da $\varepsilon[k]$ er hvit støy og F har orden $m-1$, betyr dette at $E(\Delta y[k]\Delta y[k+p]) = \phi_{\Delta y\Delta y}[p] = 0$ når $p \geq m$. Dette betyr at en løpende, numerisk beregnet autokorrelasjon kan benyttes for å etterprøve om hvorvidt en minimum-varians-regulator fungerer riktig.

12.2 Optimal m-skrifts prediktor

La oss nå tenke oss at vi har et system som ikke kan påvirkes, dvs. utgangen skyldes den hvite støyen og ikke noe annet. Vi får en **ARMA**-modell

$$A(z^{-1})y[k] = C(z^{-1})\varepsilon[k] \quad (12.22)$$

Nå står vi ikke lenger overfor et reguleringsproblem, siden vi ikke har noe pådrag. Men det kan være interessant å prøve å prediktere (forutsi) verdien av utgangen $y[k+m]$ når målinger fram til tidspunktet k er kjent. Vi kan finne denne **m-skritts-prediktoren** ved hjelp av utledningene i avsnitt 12.1, bare ved å sette $u[k] \equiv 0$.

$$(12.10) \text{ gir da} \quad y[k+m] = F\varepsilon[k+m] + \frac{G}{C}y[k] \quad (12.23)$$

Tar vi forventningsverdi på begge sider får vi m-skritts-prediktoren, jfr. (12.12)

$$\begin{aligned} E(y[k+m|k]) &= \bar{y}[k+m|k] = (1 + f_1 + f_2 + \dots + f_{m-1})E(\varepsilon[k]) + \frac{G}{C}y[k] \\ &\Rightarrow \bar{y}[k+m|k] = \frac{G}{C}y[k] \end{aligned} \quad (12.24)$$

som kan skrives på rekursiv form

$$\begin{aligned} \bar{y}[k+m|k] &= -c_1\bar{y}[k+m-1|k-1] - \dots - c_n\bar{y}[k+m-n|k-n] \\ &\quad + g_0y[k] + g_1y[k-1] + \dots + g_{n-1}y[k-n+1] \end{aligned} \quad (12.25)$$

Hvis vi nå vender tilbake til problemstillingen i avsnitt 12.1, kan vi umiddelbart slå fast at når vi har minimum-varians-regulering, så blir m-skritts-prediktoren gitt av (12.12):

$$\bar{y}[k+m|k] = \frac{BF}{C}u[k] + \frac{G}{C}y[k] \quad (12.26)$$

Vi ser derfor at minimum-varians-regulatoren inneholder en m-skritts prediktor og at pådraget velges slik at den predikerte utgangen blir lik null. Den egentlige verdien av $y[k+m]$ blir lik prediksjonsfeilen $F\varepsilon[k+m]$.

Hvis vi ikke ønsker $y[k]$ regulert til null, men til en referanse $y^R[k]$, velges pådraget slik at den predikerte målingen blir lik $y^R[k]$. Dvs. fra (12.26):

$$y^R[k+m] = \frac{BF}{C}u[k] + \frac{G}{C}y[k], \quad (12.27)$$

$$\text{eller} \quad u[k] = \frac{C}{BF}y^R[k+m] - \frac{G}{BF}y[k] \quad (12.28)$$

Vi ser av dette at det stokastiske reguleringsproblemet er separert i et prediksjonsproblem og et reguleringsproblem hvor man velger pådraget på basis av predikert verdi av $y[k]$. Man regulerer systemet ved hjelp av en dead-beat-regulator (jfr. (5.10) og (5.11) i kapittel 5) som om det var deterministisk (alle $\varepsilon[k] = 0$), men oppnår likevel minimum effekt av (det virkelige) stokastiske system. Dette viktige resultatet kalles **separasjonsteoremet**.

12.3 Sammenheng MV- og dead-beat-regulering

Vi har altså funnet ut at dead-beat-regulatoren og MV-regulatoren er den samme. Denne likheten blir lettere å forstå ved følgende betraktnign. Anta at vår ARMAX-modell eksiteres av en diskret enhetspuls $\varepsilon[k]$. Ligning (5.8) i kapittel 5 gir at impulsresponsen med dead-beat-regulator blir

$$h_{\varepsilon y}[k] = f_k \quad (12.29)$$

der f_k er koeffisientene i F -polynomet.

(Merknad: Impulsresponsen blir "FIR", Finite Impulse Response, se kapittel 10).

Ser vi på det stokastiske tilfelle hvor $\varepsilon[k]$ er hvit støy, kan vi nå betrakte denne støyen som en serie med diskrete impulser som inntreffer etter hverandre i tid, og som har forskjellig amplitude. Dermed kan vi bruke superposisjonsprinsippet og betrakte den stokastiske responsen som summen av responser av typen (12.29) på hver av disse pulsene. Dermed blir det rimelig at en dead-beat-regulator også vil gi minimum effekt på utgangen når $\varepsilon[k]$ er hvit støy.

12.3.1 Minimum-energi-regulator

Uttrykket (12.19) kan nå gis to tolkninger. Den første er som *effekten* til $y[k]$ når $\varepsilon[k]$ er hvit støy

$$E(y^2[k]) = \phi_{\varepsilon\varepsilon} \sum_{j=0}^{m-1} f^2[j] \quad (12.30)$$

Den andre er som *energien* til $y[k]$ når $\varepsilon[k]$ er en diskret enhetspuls med amplitude A :

$$Q = A^2 \sum_{j=0}^{m-1} f^2[j] \quad (12.31)$$

Dead-beat-regulatoren er dermed også en **minimum-energi-regulator** for systemet!

12.4 Robust MV-regulator

Avsnitt 5.3 i kapittel 5 beskriver en litt annerledes dead-beat-regulator, hvor systemet kommer noe seinere til ro, fordi man forlanger at nullpunkter i B -polynomet nær eller utover enhetssirkelen ikke skal kanselleres i det lukkede system. Denne regulatoren kan anvendes når $\varepsilon[k]$ er hvit støy, med akkurat samme resonnement som for dead-beat-regulatoren i avsnitt 12.1. Forskjellen blir at variansen blir noe større, fordi F -polynomet får noe høyere grad, $\text{grad}(F) = m + n_2 - 1$ som vist i kapittel 5, avsnitt 5.3.

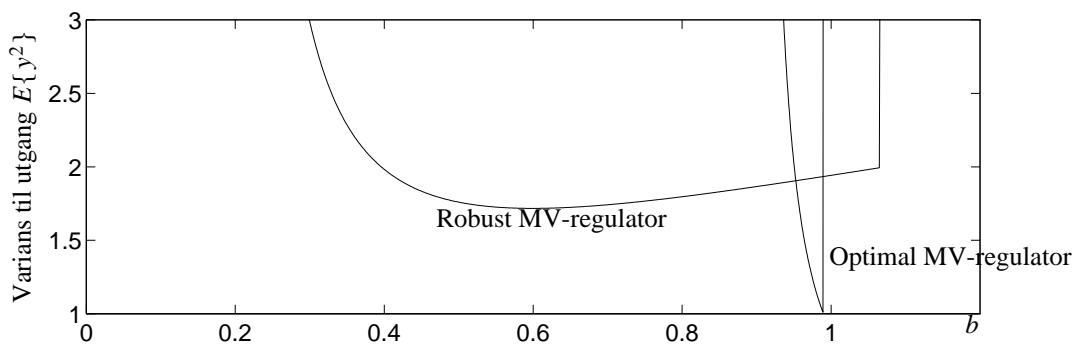
Effekten på utgangen blir nå - jfr. (12.30):

$$E(y^2[k]) = \phi_{\varepsilon\varepsilon} \sum_{j=0}^{m+n_2-1} f^2[j] \quad (12.32)$$

Merk øvre summasjonsgrense!

Hvis vi nå vender tilbake til eksempel 2 i avsnitt 5.2.4, så er responsen for dette systemet når $\varepsilon[k]$ er en diskret enhetspuls med amplitude 1 (dvs. impulsresponsen) vist i figur 5.7 (dead-beat, én riktig og to gale parameterverdier) og i figur 5.9 (robust dead-beat, én riktig og to gale parameterverdier).

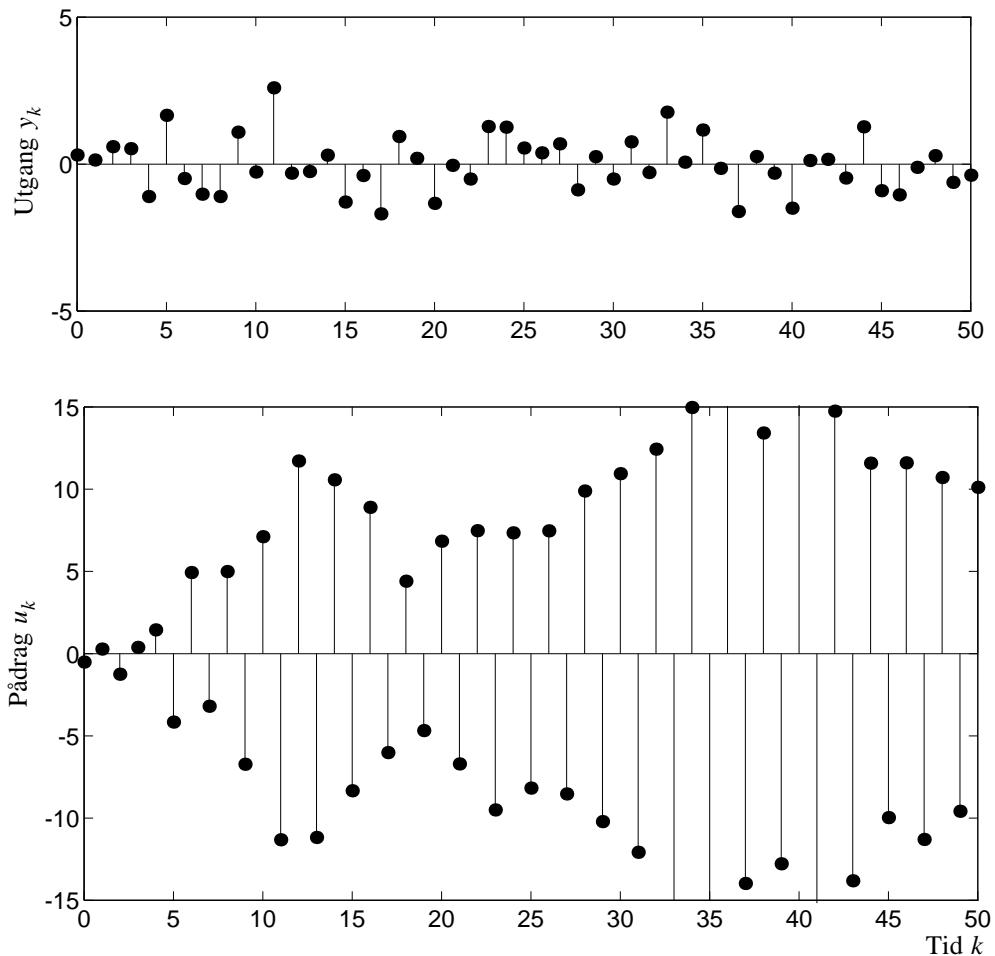
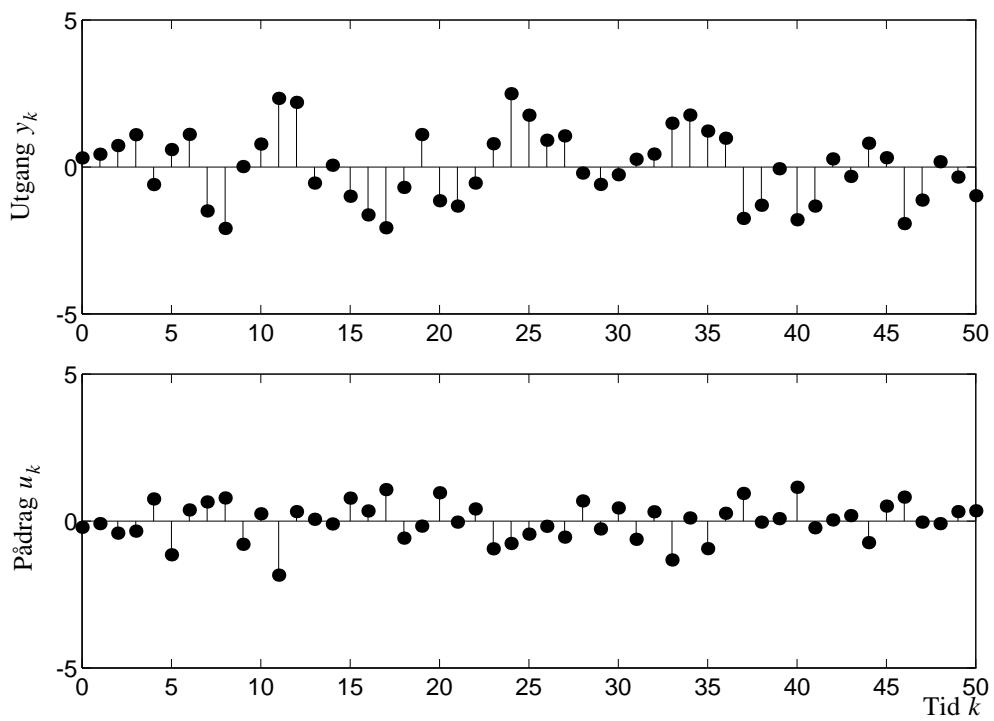
Tilsvarende resultat for systemet med to forskjellige regulatorer, og med varierende parameterverdier omkring den riktige, er vist i figur 12.1. $\varepsilon[k]$ er nå hvit støy med $\phi_{\varepsilon\varepsilon} = 1$.



Figur 12.1 Variansen på utgangen som funksjon av parameteren b

Sammenhold denne figuren med figurene 5.7 og 5.9!

Figurene 12.2 og 12.3 neste side viser pådrag og utgang med henholdsvis MV-regulator og robust MV-regulator når parameteren b har korrekt verdi = 0.99.

Figur 12.2 Med MV-regulator, $b = 0.99$ Figur 12.3 Med robust MV-regulator, $b = 0.99$

12.5 Lagerstyringseksemplet

Vi skal prøve å regulere lageret fra kapittel 11 med en bedre regulator enn den proporsjonalregulatoren vi brukte der. Vi har modellen (11.15)

$$y[k] = y[k-1] + u[k-2] + v[k] + 0.9v[k-1] \quad (12.33)$$

vi skal nå regulere lageret med en MV-regulator, og siden B-polynomet nå har nullpunkt i origo (det bare består av én koeffisient = 1) kan vi bruke den hurtigste av de to MV-regulatorene (altså den “ikke-robuste”). Vi har

$$\begin{aligned} A(z^{-1}) &= 1 - z^{-1} \\ B(z^{-1}) &= 1 \quad , \text{og } m = 2 \\ C(z^{-1}) &= 1 + 0.9z^{-1} \end{aligned} \quad (12.34)$$

Identiteten blir da

$$1 + 0.9z^{-1} \equiv (1 - z^{-1})(1 + f_1 z^{-1}) + z^{-2}(g_0) \quad (12.35)$$

$$\begin{aligned} z^{-1}: \quad 0.9 &= -1 + f_1 \\ z^{-2}: \quad 0 &= -f_1 + g_0 \end{aligned} \quad (12.36)$$

Dette gir $f_1 = 1.9$ og $g_0 = 1.9$. Regulatoren blir da i følge (12.16):

$$(1)(1 + 1.9z^{-1})u[k] = -(1.9)y[k] \quad (12.37)$$

som er det samme som det rekursive uttrykk

$$u[k] = -1.9u[k-1] - 1.9y[k] \quad (12.38)$$

Variansen på utgangen blir i følge (12.18) eller (12.19)

$$E(y^2[k]) = \phi_{vv} \sum_{j=0}^1 f^2[j] = 1 \cdot (1 + 1.9^2) = 4.61 = J_{MIN} \quad (12.39)$$

som er atskillig bedre enn det vi fikk med proporsjonalregulator, som var $J_{MIN} = 7.23$.

Regulatoren (12.38) kan intuitivt forklares slik: Vi har med ledet med $u[k-1]$ fordi dette pådraget ennå ikke har slått ut på utgangen $y[k]$. Den tidligste y som påvirkes blir $y[k+1]$. Dermed må vi ha med ledet $u[k-1]$ for at beslutningen $u[k]$ skal bli tatt på grunnlag av all relevant informasjon, dvs. for at beslutningen skal bli mest mulig riktig.

Og da kan vi også tillate oss en kraftigere forsterkning (= 1.9) enn det vi fant ut var optimalt med proporsjonalregulering (= 0.5).

A	amplitudefordelingsfunksjon	66
	amplitudetethetsfunksjon	66
	analytisk funksjon	2
	antinedfoldingsfilter	141
	ARMAX-modell	33
	autoeffektspektrum	90, 91
	diskret system	115
	egenskaper	94
	lineært dynamisk system	106
	og autokorrelasjon	93
	autokorrelasjon	
	diskret	114
	diskret system	115
	effekt	81
	lineært dynamisk system	105
	og autoeffektspektrum	90, 93
	autokorrelasjonsfunksjon	79
	beregning fra simultan tetthetsfunksjon	85
	deterministisk signal	83
	diskret prosess	82
	egenskaper	81
	ergodisk prosess	80
B	Balchens Q-transformasjon	32
	båndpassfilter	
	konstruksjon fra lavpassfilter	151
	båndstoppfilter	
	konstruksjon fra lavpassfilter	151
	blokkdiagram diskret system	31
	Brownsk prosess	101
	Butterworth-filter	147
C	COLOSS	158
	cospektrum	96
D	dead-beat-regulator	35
	eksempel 1	39
	eksempel 2	40
	følgeegenskaper	37, 41
	graden på F- og G polynomet	49
	polenes plassering	38
	utgangssignal-d-b-regulator	36
	deterministisk modell	61
	DFT	121, 123
	differensligning	19
	digitalt filter	144
	FIR-filter	144
	IIR-filter	144
	syntese	145
	dimensjonsløs tid	11
	diskret	
	enhetspuls	9
	folding	10
	frekvens, diskrete signaler	11
	frekvensrespons	12
	impulsrespons	9
	system	9
	ARMAX-modell	33
	tilstandsrommodell	
	stokastisk dynamisk system	119
	transferfunksjon	19
	diskret enhetspuls	117
	diskret filter	
	orden	148
	syntese	
	eksempel	149
	diskret Fourierrekke	123
	diskret hvit støy	116
	autoeffektspektrum	116
	effekt	116
	fra kontinuerlig hvit støy	116
	kontinuerlig hvit støy	118
	diskret impulsrespons	117
	diskret prediktor	170
	diskret regulator	
	dead-beat-regulator	35
	generell lineær regulator	50
	polplassering	35
	seriekompensasjon	32
	stabilitet og frekvensanalyse	51
	utgangssignal-d-b-regulator	36
	diskret signal	9
	diskret system	
	blokkdiagram	31
	diskretisering	
	holdeelement	21
	impulsrespons metode	

eksempel	30
impulsrespons-metode	29
konform avbildning	25
nullpunktene plassering	25
polenes plassering	23, 25
tilstandsrom metode	21
tilstandsrom-metode	
eksempel	27
E	
effekt	
autokorrelasjonsfunksjon	81
varians	107
effektiv reguleringsgrad	107
effektspektrum	89, 90
diskret	115
effekttetthetsspektrum	90
enhetspuls,diskret	9
ergodisk prosess	79
estimat	109
estimat av effektspektrum	
forventning	126
utfylling med nuller	136
varians kontra forventningsfeil	130
eksempel	133
estimat av korrelasjonsfunksjon	135
utfylling med nuller	136
estimat av krysseffektspektrum	134
estimering av spektrum	123
forventningsverdi	125
varians	129
estimering av transferfunksjon	137
eksempel	138
målestøy	112
prosess-støy	110
tilbakekoblet system	113
estimeringsproblem	109
F	
Fast Fourier Transform	121
FFT	121
Finite Impulse Response Filter	144
FIR-filter	144
lineær fasegang	144
folding	10
syklisk	136
foldingsintegral	105
foldingsintegralet	10, 76
fordelingsfunksjon	65
simultan	66
forventningsfeil	
utsmøring	130
forventningsoperasjon	
egenskaper	70
forventningsrett	125
forventningsverdi	69
Fourier-transform	
diskret	
endelig måleserie	121
Fourier-transformasjon	
endelig diskret	121
Fouriertransformasjon	
diskret	11
invers	13
kontinuerlig	
invers	6
frekvens diskrete signaler	11
frekvensrespons	12
G	
Gaussfordeling	74
H	
hakeparentes	9
Hamming-vindu	127
Hanning-vindu	127
holdeelement	21
høypassfilter	
konstruksjon fra lavpassfilter	151
hvit støy	99
diskret	116
diskret hvit støy	116, 118
eksempel	101
kontinuerlig	
egenskaper	101
lavpassfiltrert	108
telegrafsignal	109
rudimentær	101, 108

I	kvantiseringssstøy	143
	IIR-filter	144
	impulsrespons	9
	Infinite Impulse Response Filter	144
J		
	Jacobi-matrise	92
K		
	kausalitet	113
	koherensfunksjon	97
	betydning	112
	komplett koherens	98
	målestøy	112
	prosess-støy	111
	konfidensintervall	137
	konform avbilding	25
	korrelasjon	
	eksempel	72
	statistisk uavhengighet	73, 77
	korrelasjonsfunksjon	78
	oppsummering egenskaper	88
	korrelasjonskoeffisient	72
	kovariansmatrise	71
	kryssefektspektrum	
	egenskaper	98
	krysseffektspektrum	
	diskret	115
	diskret system	115
	eksempel	98
	krysskorrelasjon	96
	lineært dynamisk system	106
	krysskorrelasjon	
	diskret	114
	diskret system	115
	kryssefektspektrum	96
	lineært dynamisk system	106
	krysskorrelasjonsfunksjon	79
	bruk	84
	diskret prosess	82
	egenskaper	82
	ergodisk prosess	80
	kvadraturspektrum	96
	kvantisering	143
	kvantiseringssfeil	143
L	Laplacetransformasjon	
	invers	1, 2
	lavpassfiltrert hvit støy	108
	lineær fasegang	144
	Ljung, Lennart	60
M	målestøy	111
	estimering av transferfunksjon	112
	koherensfunksjon	112
	middelverdi	69
	minimum-energi-regulator	172
	minimum-varians-kriterium	155
	eksempel - dobbeltintegrator	163
	eksempel - styring av lager	160
	minimum-varians-regulator	169
	dead-beat-regulering	172
	eksempel - styring av lager	175
	eksempel robusthet	173
	optimal regulator	167
	robust regulator	173
N	normalfordeling	74
	simulant normalfordelte variable	76
	Nyquistfrekvensen	25
	Nyquist-kurve	
	1. ordens lineært system	53
	dead-beat-regulator	58
	diskret 1. ordens system	
	kontinuerlig tilnærming	55
	diskretisert 1. ordens system	53
	robust-dead-beat-regulator	58, 59
	Nyquists stabilitetskriterium	
	åpent stabilt system	52
	åpent ustabilt system	57
	argumentvariasjonsprinsippet	51
	Bode-diagram	56
	diskrete systemer	51
O	optimal prediktor	170
P	pådragsstøy	110

periodogram	125
Poisson-fordeling	85
pollassering	35
prediksjon	61
prediksjonsproblem	171
prediktor	170
prefilter	141
båndbredde	142
proper	23
prosess-støy	110
estimering av transferfunksjon	110
koherensfunksjon	111
punktprøver	11
 Q	
Q-transformasjon	32
 R	
regulator	
godhetskriterium	155
minimum-varians-kriterium	155
regulatorordesign-problem	109
reguleringsgrad	107
effektiv	107
reguleringsproblem	171
residueregning	1
residuesatsen	2
robust-dead-beat-regulator	
følgeegenskaper	47
rudimentær hvit støy	101
 S	
SALOSS	159
sampling	11, 141
sannsynlighetsfordelingsfunksjon	65
betinget	68
sannsynlighetstethetsfunksjon	65
betinget	68
sentralgrenseteorem	74
sentralgrenseteoremet	
eksempel 1	74
eksempel 2	75
separasjonsteoremet	171
seriekompensasjon	32
Signalbehandling	141
simultan fordelingsfunksjon	66
simultan tethetsfunksjon	66
spektralfaktoriseringssats	108
anvendelser	109
standardavvik	71
stasjonær prosess	79
ergodisk	79
stasjonær, ikke ergodisk prosess	
eksempel	80
statistisk uavhengighet	69
korrelasjon	73, 77
stokastisk	
fordelingsfunksjon	65
kontinuerlig system	62
prosess	64
realisasjon	64
system	61
tethetsfunksjon	65
tidsserie	64
utfallsfunksjon	64
stokastisk prosess	
stasjonær	66
stokastiske prosess	
diskret	114
støy	
estimering i tilbakekoplede systemer	112
hvit	99
målestøy	33, 111
pådrag	110
prosess-støy	33
støy i dynamiske systemer	102
strengt proper	3, 23
syklisk folding	136
systemidentifikasjons-problem	109
 T	
taster	21
tasting	11, 141
frekvens	142
telegraftsignal	85
lavpassfiltrert hvit støy	109
tethetsfunksjon	65
betinget	66
simultan	66, 68, 78

tidsforskyvingsegenskapen	19
tilbakekoblet system	
estimering av transferfunksjon	113
tilfeldig variabel	65
transferfunksjon	
strentg proper	3, 23
trapesintegrasjon	146
U	
utfylling med nuller	136
utsmøring	130
V	
varians	71
effekt	107
vindusfunksjon	124, 125
Hamming-vindu	127
Hanning-vindu	127
skalering for konstant effekt	128
W	
Wiener-prosess	101
wild points	141
wild-point- fjerning	154
Z	
z-transformasjon	15
tidsforskyvingsegenskapen	19
to betydninger av "z"	38

