

FYS3230 Sensorer og måleteknikk

Kortfattet kompendium over noen emner

Ola Sveen

2007

Forord

Denne oversikten oppsummerer mange av emnene som foreleses i FYS3230 Sensorer og måleteknikk.

Noen emner vil mangle helt, noen vil være mangelfullt behandlet, mens noen få emner ikke vil bli berørt i forelesningene.

Innhold

1.	Målesystem.....	3
2.	Statisk karakteristik av målesystem	3
2.1.	Systemets egenskaper.....	3
3.	Statistiske egenskaper	5
3.1.	Statistiske variasjoner i utgangssignalet fra et enkelt element.....	5
3.2.	Kalibrering - standarder	6
4.	Nøyaktigheten til et målesystem i likevekt.....	6
4.1.	Systemfeil.....	6
4.2.	Reduksjon av systemfeil.....	7
4.2.1.	Kompensasjon	7
4.2.2.	Reduksjon av miljøeffekter.	7
5.	Systemer av 1. og 2. orden	8
5.1.	Systemer av 1. orden	8
5.2.	Systemer av 2. orden	10
5.3.	Sinus-respons til 1. og 2. ordens system, Bode-plot	12
5.3.1.	Systemer av 1. orden	12
5.3.2.	Systemer av 2. orden	13
6.	Tilbakekoblede systemer og reguleringssystemer	15
6.1.	Frekvensanalyse	18
6.1.1.	Stabilitetskrav.....	18
7.	Modell-tilpasning. Minste kvadraters metode.....	20
7.1.	Minste kvadraters metode.	20
7.2.	Det lineære tilfelle.....	20
7.3.	Eksempel	21
7.4.	Ikke-lineære problemer.	24
8.	Digital filtrering.....	25
8.1.	Ikke-rekursive filter. Utjevning.....	25
8.2.	Overføringsfunksjon for et digitalt filter.....	26

8.3.	Rekursive filter.....	27
8.3.1.	Eksempel: Digitalt lavpassfilter.....	27
8.4.	Beregning av responsen på grunnlag av filterets impulsrespons.....	28
9.	Analog til digital (A/D) og digital til analog (D/A) omsetter	32
9.1.	Sampling.....	32
9.2.	Kvantisering.....	33
9.3.	Koding.....	33
9.3.1.	Gray kode.....	33
9.4.	Polaritet.....	33
9.5.	Digital til analog (D/A) omsetter.....	33
9.5.1.	Feil i D/A omsettere.....	34
10.	A/D omsetter	34
10.1.	Rampe-type.....	34
10.2.	Suksessive approksimasjoner.....	35
10.3.	Integrerende omsettere.....	35
10.4.	Parallell omsetter (flash converter).....	36
10.5.	Spenning (strøm) til frekvens omsetter.....	37
10.6.	Sigma-delta type.....	37
11.	Datainnsamling.....	38
11.1.	LabVIEW.....	39
11.2.	Serie- og parallell I/O.....	39
11.2.1.	Serie-port.....	39
11.2.2.	Parallellport.....	40
12.	Frekvensanalyse	42
12.1.	Fourierrekker.....	42
12.2.	Effektspektrum.....	44
13.	Signal og støy i målesystem	45
13.1.	Statistisk beskrivelse av tilfeldige (random) signaler.....	45
13.1.1.	Middelverdi og standard avvik.....	45
13.1.2.	Sannsynlighets tetthetsfordeling.....	45
13.1.3.	Spektral effekt tetthetsfordeling.....	45
13.1.4.	Autokorrelasjon.....	46
13.2.	Virkning av støy på en målekrets.....	48
13.3.	Støykilder og koblingsmåter.....	49
13.3.1.	Interne støykilder.....	49
13.3.2.	Eksterne støykilder.....	50
13.3.3.	Koblingsmåter til eksterne støykilder.....	50
13.4.	Metoder for reduksjon av virkningen av støy.....	53
13.4.1.	Fysisk separasjon.....	53
13.4.2.	Elektromagnetisk skjerming.....	53
13.4.3.	Elektrostatisk skjerming.....	53
13.4.4.	Bruk av differanseforsterker.....	54
13.4.5.	Filtrering.....	55
13.4.6.	Modulasjon.....	55
13.4.7.	Midling.....	56
13.4.8.	Autokorrelasjon.....	56
14.	Fasefølsom detektor	57
15.	Sensorer.....	59
15.1.	Resistive sensorer.....	59
15.1.1.	Potensiometer for lineær forskyvning.....	59
15.1.2.	Motstandstermometere og termistorer.....	59

15.1.3.	Strekklapper (metallegering, halvledere)	60
15.1.4.	Fotomotstander	61
15.2.	Kapasitive sensorer	61
15.3.	Induktive sensorer	62
15.3.1.	Forskyvningssensorer	62
15.3.2.	Differensialtransformatorer	63
15.4.	Elektromagnetiske sensorer	64
15.4.1.	Hastighetssensorer	64
15.4.2.	SQUID	64
15.5.	Termoelementer	65
15.6.	Elastiske sensorer	66
15.7.	Piezoelektriske sensorer	67
15.8.	Piezoresistive sensorer	68
15.9.	Elektrokjemiske sensorer	68
16.	Behandling av sensorsignalet	71
16.1.	Brokoblinger	71
16.1.1.	Resistiv brokobling	71
16.1.2.	Reaktive broer	72
16.2.	Forsterkere	73
16.2.1.	Operasjonsforsterkere	73
16.2.2.	Instrumenteringsforsterker	74
17.	Data innsamling og telemetri	75
17.1.	Parallell dataoverføring	75
17.2.	Seriell dataoverføring	75
17.2.1.	Transmisjons-båndbredde	76
17.2.2.	Virkningen av støy på P.C.M. signal	76
17.3.	Deteksjon av feil og retting	77
17.3.1.	Paritetsbit	78
17.3.2.	Feilrettende koder	78
17.3.3.	Longitudinelle sjekk-summer	78
17.4.	Frekvensmodulering, F.S.K.	79

1. Målesystem

Et målesystems oppgave er å gi en numerisk verdi for en variabel.

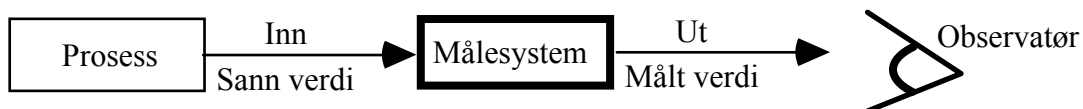


Fig. 1.1 Generalisert målesystem

Målesystemet kan gjerne deles opp i flere blokker.

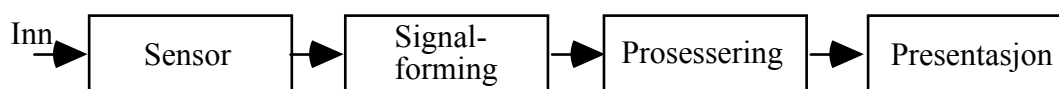


Fig. 1.2 Komponenter i et typisk målesystem

En sensor (føler, transducer) er en krets som vanligvis gir ut en spenning (strøm) som avhenger av en inngangsvariabel (f.eks. trykk, temperatur etc.). Blokkdiagram brukes ofte til å vise hovedtrekkene i et system. Noen vanlige symboler er:

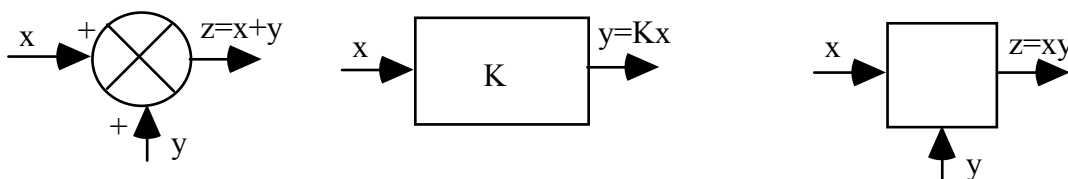


Fig. 1.3 Symboler i blokkdiagram

2. Statisk karakteristikk av målesystem

Utgangssignal som funksjon av inngangssignal når inngangssignalet er konstant eller varierer langsomt.

2.1. Systemets egenskaper

Systematiske egenskaper kan kvantifiseres eksakt matematisk eller grafisk i motsetning til statistiske egenskaper.

Måleområdet (range): Intervallet mellom nedre og øvre målegrense, I_{\min} til I_{\max} for inngangsverdier, og O_{\min} til O_{\max} for utgangsverdier.

Måleomfang (span): Den algebraiske forskjellen mellom øvre og nedre målegrense, dvs. $I_{\max} - I_{\min}$ og $O_{\max} - O_{\min}$.

Lineært element: Utgangssignalet er en lineær funksjon av inngangssignalet (ideell rett linje):

$$O_{ideel} = KI + a$$

$$K = \text{vinkelkoeffisient} = (O_{\max} - O_{\min}) / (I_{\max} - I_{\min})$$

$$a = \text{nullpunktsavvik} = O_{\min} - KI_{\min}$$

Ofte varierer ikke utgangssignalet helt lineært. Vi kan da skrive utgangssignalet som sum av en lineær funksjon og en korreksjon

$$O(I) = KI + a + N(I)$$

hvor $N(I)$ er forskjellen mellom det virkelige utgangssignalet og den ideelle rette linjen gjennom minimums og maksimums verdiene (I_{\min}, O_{\min}) og (I_{\max}, O_{\max}) .

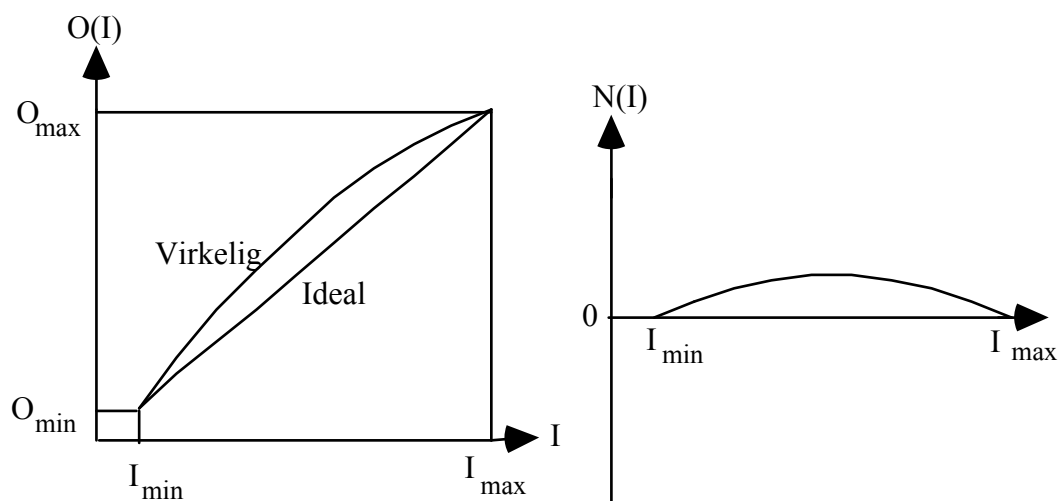


Fig. 2.1 Definisjon av ulinearitet

Ulineariteten spesifiseres ofte i % av full skala utslag (full-scale deflection, f.s.d.), dvs. i % av måleomfang.

Følsomheten (sensitivity) er definert som

$$dO/dI = K + dN/dI$$

Effekter som skyldes omgivelsene (**miljøeffekter, environmental effects**), f.eks. temperatur, fuktighet etc., vil ofte påvirke utgangssignalet. Hvis effekten forandrer følsomheten for et lineært element, K , kalles det en *modifiserende* effekt. Hvis effekten forandrer nullpunktsavviket, kalles det en *interfererende* effekt. For relativt små effekter kan vi regne med lineære korreksjonsledd. Det generelle uttrykket for utgangssignalet kan da skrives som

$$O = (K + K_M I_M)I + a + K_I I_I + N(I)$$

hvor I_M og I_I er avvik fra "standard" verdi (lik null ved standard betingelser), og K_M og K_I er miljø koblingskonstanter.

3. Statistiske egenskaper

Noen av feilkildene ved målinger kan være systematiske, dvs. de trekker i en bestemt retning. For eksempel vil bruk av en målestav som er litt for kort, gjøre at alle lengdemålinger får en systematisk feil. Andre feil kan være av tilfeldig natur og gjøre at gjentatte målinger ikke gir nøyaktig samme resultat. Dette skyldes tilfeldige effekter (statistiske variasjoner) i elementet og i inngangssignalene.

3.1. Statistiske variasjoner i utgangssignalet fra et enkelt element

Hvis vi for en bestemt verdi av inngangssignalet gjør en rekke målinger, vil resultatene fordele seg rundt en middelværdi. Fordelingen vil vanligvis ligne mer og mer på en Gaussfordeling (normalfordeling) når antall målinger øker. Normalfordelingen (sannsynlighetsfordelingen) er beskrevet av funksjonen

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

hvor μ er middelværdien og σ er standard avvik.

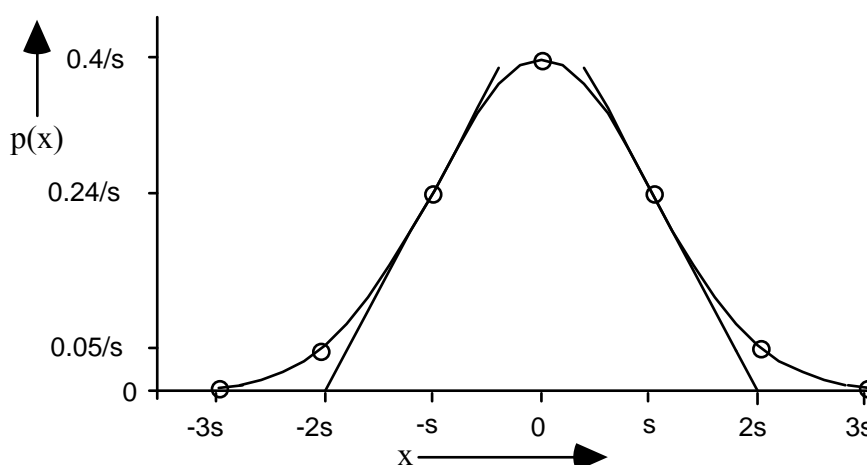


Fig. 3.1 Gauss sannsynlighetsfordeling for middelværdi $\mu=0$ og st.avvik s

Utgangssignalet O er en funksjon av inngangssignalene I , I_M og I_I . Hvis variasjonene i inngangssignalene er uavhengige av hverandre og normalfordelt med st.avvik σ_I , σ_{I_M} og σ_{I_I} , er st.avvik for utgangssignalet

$$\sigma_o = \sqrt{\left(\frac{\partial O}{\partial I} \sigma_I\right)^2 + \left(\frac{\partial O}{\partial I_M} \sigma_{I_M}\right)^2 + \left(\frac{\partial O}{\partial I_I} \sigma_{I_I}\right)^2}$$

dvs. roten av summen av kvadratene.

I det generelle tilfellet med mange "like" elementer, får vi også kvadratledd som skyldes variasjoner i parametrene K , a etc.

For et endelig antall målinger N som gir verdiene x_1 til x_N , kan vi beregne (estimerte) verdier for middelværdi

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

og st. avvik (root-mean-square of deviations)

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}$$

Hvis målingene er Gauss fordelt, vil 68% av verdiene ligge innen ett standard avvik fra middelværdien, og 95% innen to standard avvik.

3.2. Kalibrering - standarder

Kalibrering: Måle I og O og forstyrrende effekter med best mulige instrumenter. Langsamt varierende inngangssignaler.

Nøyaktighet(accuracy): Hvor nær målt verdi svarer til sann verdi.

Sann verdi (true value): Verdien vi ville ha målt ved å bruke standarder med høyeste presisjon.

Presisjon (precision): Mål for hvor nøyaktig resultatet er bestemt (hvor reproducerbart det er).

Sporbarhet (traceability) og målestige (traceability ladder): Spesifikasjon (sertifikat) for nøyaktighet i forhold til nasjonale/internasjonale standarder (justervesenet, prototyper).

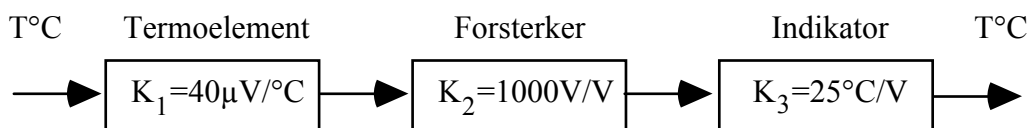
Systematiske feil: Feilaktige kalibreringer, måleinstrumenter etc. Systematiske og tilfeldige (statistiske) feil bør spesifiseres separat.

4. Nøyaktigheten til et målesystem i likevekt

4.1. Systemfeil

Komplett system: Inngangsverdi I , sann verdi.
 Utgangsverdi O , målt verdi.
 Nøyaktighet (systematisk feil) $E = O - I$
 Perfekt system (ingen systemfeil) hvis $E = 0$

Eksempel: Temperaturmåling og tre ideelle elementer $O_i = K_i I_i$.



$$O = K_1 K_2 K_3 I$$

$$E = O - I = 40 * 10^{-6} * 10^3 * 25 - 1 = 0$$

dvs. ingen system feil.

Vanligvis vil K 'ene variere og avvike fra ideelle verdier. Hvis variasjonene tilnærmes med Gauss fordelinger, kan vi beregne systematiske feil med standard avvik. Det er liten vits i å prøve å redusere statistiske feil så meget at de blir vesentlig mindre enn systematiske feil!

4.2. Reduksjon av systemfeil

4.2.1. Kompensasjon

Et system med et ulineært element kan forbedres ved å bruke en forsterker med ulineær karakteristikk slik at det tilsammen blir en tilnærmet lineær respons.

4.2.2. Reduksjon av miljøeffekter.

Isolasjon.

Velge elementer lite følsomme for miljøeffekter.

Bruke kompensasjon for å motvirke miljøeffekter, differensialkobling.

Negativ tilbakekobling.

Dette er en meget viktig metode. Utgangssignalet er tilnærmet uavhengig av forsterkningen i foroverretningen og avhenger bare av elementet i tilbakekoblingen.

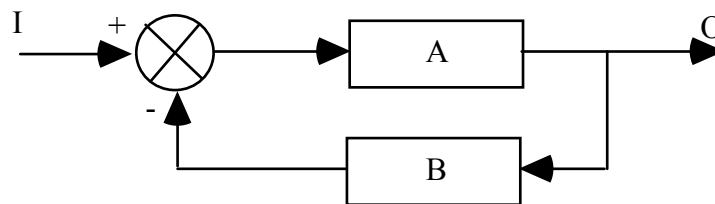


Fig. 4.1 Negativ tilbakekobling

Utgangssignalet er

$$O = (I - BO)A$$

Forsterkningen er

$$G = \frac{O}{I} = \frac{A}{1 + AB} \approx \frac{1}{B}$$

hvis sløyfeforsterkningen (loop gain) $AB \gg 1$.

Utgangssignalet er da bare avhengig av elementet i tilbakekoblings-sløyfen, og ulinearitet og modifierende effekter i foroverretningen har liten betydning.

Variasjonen i forsterkningen er gitt ved

$$\frac{dG}{G} = \frac{dA}{A} \frac{1}{1 + AB}$$

hvis B er konstant. Variasjonen i A er derfor redusert med en faktor $1 + AB \approx AB$, dvs. sløyfeforsterkningen. De dynamiske egenskapene til systemet blir også forbedret med negativ tilbakekobling (evnen til å følge høye frekvenser bedres).

5. Systemer av 1. og 2. orden

Denne gjennomgangen er grundigere enn den som foreleses i FYS3230. For en enda mer utfyllende gjennomgang henvises til FYS3220.

5.1. Systemer av 1. orden

Et system som kan beskrives med en 1. ordens lineær differensialligning, kalles et 1. ordens system. En tilnærmet masseløs kraft-sensor er et eksempel på et 1. ordens system. Systemet kan illustreres med følgende modell:

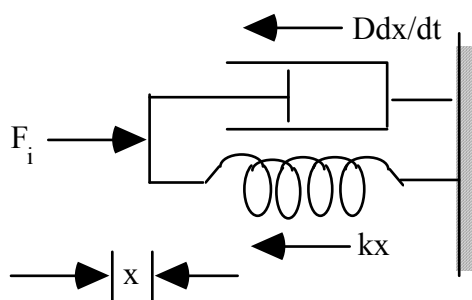


Fig. 5.1. Mekanisk 1. ordens system. Kraft-sensor uten masse.

Hvis sensoren påvirkes av en kraft F , vil fjæra gi en motkraft $F=kx$ hvor k er proporsjonalitetsfaktoren for fjæra og x er forskyvningen. Demperen vil gi en motkraft F_1 proporsjonal med hastigheten dx/dt , dvs. $F_1 = Ddx/dt$ hvor D er proporsjonalitetsfaktoren for demperen. Hvis vi som utgangspunkt antar $x=0$ og alle krefter lik null, og så plutselig påtrykker en kraft F_i , har vi at

$$F_i = kx + Ddx/dt = F + (D/k) dF/dt \quad \text{dvs.}$$

$$TdF/dt + F = F_i \quad \text{hvor } T = D/k \text{ er en "mekanisk" tidskonstant, eller}$$

$$Tdx/dt + x = F_i/k$$

Dette er en 1. ordens lineær differensialligning.

Et annet eksempel er en elektrisk krets med en motstand og en kondensator:

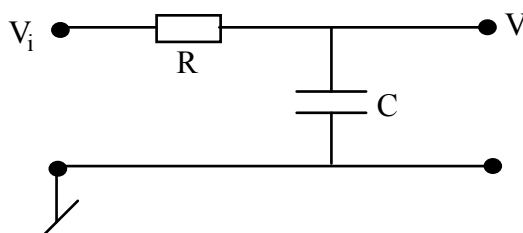


Fig. 5.2 Elektrisk 1. ordens system

Hvis inngangssignalet plutselig forandres fra null til verdien V_i , vil det gå en strøm i kretsen som er gitt ved

$$V_i - V = iR$$

(Vi antar ladningen q på kondensatoren er null opprinnelig.)

Med

$$i = dq/dt = C dV/dt$$

får vi

$$R C dV/dt + V = V_i$$

dvs.

$$T dV/dt + V = V_i$$

hvor $T=RC$ er "elektrisk" tidskonstant. Ligningen er helt analog til det mekaniske eksempelet.

Spenningen V_r over motstanden er

$$V_r = V_i - V$$

slik at vi kan skrive

$$T d(V_i - V_r)/dt = V_r$$

dvs.

$$dV_r/V_r = -(1/T) dt$$

hvis $V_i = \text{konstant}$. Ved integrasjon

$$\ln V_r = -t/T + \ln K$$

hvor K er en integrasjons-konstant. Dette gir

$$V_r = K e^{-t/T}$$

Da $V=0$ og $q=0$ for $t=0$, er $K = V_i$ dvs.

$$V_r = V_i e^{-t/T} \quad \text{og} \quad V = V_i (1 - e^{-t/T})$$

Spenningen over kondensatoren stiger derfor fra null med en tidskonstant T mot den asymptotiske verdien V_i . Spenningen over motstanden (og strømmen) synker på tilsvarende måte mot null.

Hvis inngangssignalet ikke er konstant, men varierer med tiden, kan vi bruke numeriske metoder for integrering av differensialligningen. En annen mulighet er å bruke Laplace transformasjon. Differensialligninger blir da omformet til algebraiske ligninger som kan være lettere å løse.

Den Laplace transformerte $F(s)$ av en tidsfunksjon $f(t)$ er definert ved

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

hvor $s = \sigma + j\omega$ (kompleks frekvens variabel).

Man kan vise at derivasjon i tidsplanet svarer til å multiplisere den Laplace transformerte med den komplekse variable s . Differensialligningen $TdV/dt + V = V_i$ for spenningen over kondensatoren blir derfor ved transformasjon omformet til

$$T s V(s) + V(s) = V_i(s)$$

dvs.

$$V(s)/V_i(s) = 1/(1 + sT)$$

Dette forholdet mellom de transformerte av utgangssignalet og inngangssignalet kalles systemets overføringsfunksjon eller systemfunksjon. Denne ligningen er altså systemfunksjonen for et 1. ordens system.

Hvis overføringsfunksjonen er kjent, finner vi den transformerte av utgangssignalet ved å multiplisere den transformerte av inngangssignalet med overføringsfunksjonen. Ved invers transformasjon finner vi så utgangssignalet som funksjon av tiden. Overføringsfunksjonen for elektriske kretser kan vi som regel finne ganske lett ved å bruke de transformerte impedansene

$$R \quad 1/sC \quad sL$$

(jfr. likheten med vanlige vekselstrømsimpedanser).

Skjemaet for den elektriske 1. ordens kretsen (lav-pass filter) med transformerte impedanser er vist i *fig. 5.3*.

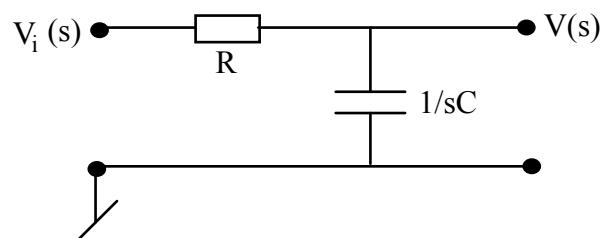


Fig. 5.3 Elektrisk 1. ordens system (lav pass filter)

Fra *fig. 5.3* får vi

$$V(s)/V_i(s) = \frac{1/sC}{R + 1/sC}$$

dvs.

$$V(s)/V_i(s) = 1/(1 + sT)$$

hvor $T = RC$ som før.

Hvis inngangssignalet er en sprang-funksjon (step-funksjon), er

$$V_i(s) = V_1/s$$

hvor V_1 er en konstant lik størrelsen på spranget.

Dette gir

$$V(s) = V_1/(s(1 + sT))$$

$$V(s) = V_1(1/s - 1/(s + 1/T))$$

Invers transformasjon (se tabell i Bentley) gir

$$V(t) = V_1(1 - e^{-t/T})$$

som før.

5.2. Systemer av 2. orden

Hvis kraft-sensoren har en masse m , kan vi bruke følgende modell:

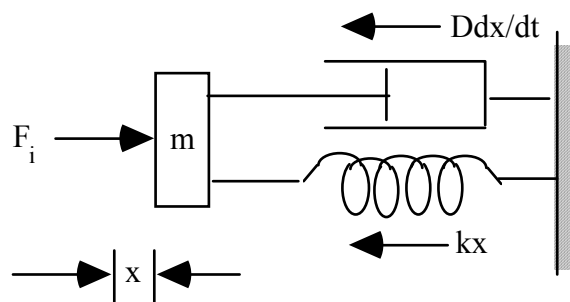


Fig. 5.4 Mekanisk 2. ordens system. Kraft-sensor med masse.

Hvis alle krefter, posisjon og hastighet er null ved $t = 0$ og vi så plutselig påtrykker en kraft F_i , får vi

$$F_i - kx - Ddx/dt = md^2x/dt^2$$

dvs.

$$(m/k)(d^2x/dt^2) + (D/k)dx/dt + x = F_i/k$$

som er en 2. ordens lineær differensialligning. Systemer som kan beskrives med en slik ligning, kalles 2. ordens systemer. Svært mange systemer kan (iallfall tilnærmet) beskrives som 2. ordens systemer.

Et elektrisk analogt system får vi ved å kople en spole, en motstand og en kondensator i serie.

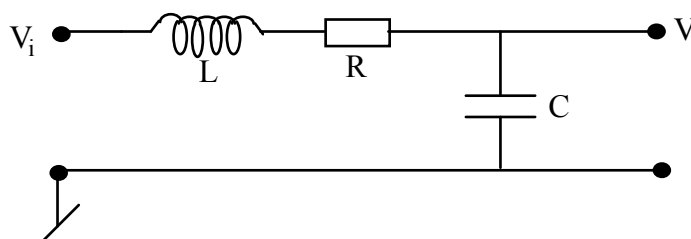


Fig. 5.5 Elektrisk 2. ordens system

For kretsen har vi at

$$V_i = Ldi/dt + Ri + q/C$$

dvs.

$$V_i = Ld^2q/dt^2 + Rdq/dt + q/C$$

eller

$$V_i = CLd^2V/dt^2 + RCdV/dt + V$$

Ved Laplace transformasjon får vi en 2. grads ligning i s :

$$V_i(s) = (CLs^2 + RCs + 1)V(s)$$

dvs. systemfunksjonen (overføringsfunksjonen) er altså

$$V(s)/V_i(s) = 1/(CLs^2 + RCs + 1)$$

Ved å bruke de Laplace-transformerte uttrykkene for impedanser får vi direkte fra fig. 5.5 at

$$\frac{V(s)}{V_i(s)} = \frac{1/sC}{sL + R + 1/sC} = \frac{1}{CLs^2 + RCs + 1}$$

som stemmer med det vi fikk ved transformasjon av differensialligningen. I det mekaniske tilfelle får vi en helt analog overføringsfunksjon.

Nevneren i overføringsfunksjonen skrives ofte på en standard form

$$(s/\omega_0)^2 + 2\zeta s/\omega_0 + 1$$

hvor vi har innført udempet egenfrekvens (resonansfrekvens)

$$\omega_0 = 1/\sqrt{CL} \quad \text{eller} \quad \sqrt{k/m}$$

og demningsforhold

$$\zeta = (R/2)\sqrt{C/L} \quad \text{eller} \quad (D/2)/\sqrt{km}$$

Polene til overføringsfunksjonen finnes av ligningen

$$(s/\omega_0)^2 + 2\zeta s/\omega_0 + 1 = 0$$

som gir

$$s_1/\omega_0 = -\zeta + \sqrt{\zeta^2 - 1}$$

og

$$s_2/\omega_0 = -\zeta - \sqrt{\zeta^2 - 1}$$

Polene ligger altså alltid i venstre halvplan. For $\zeta \geq 1$ har nevneren to reelle nullpunkter. For $0 \leq \zeta < 1$ har nevneren to kompleks konjugerte nullpunkter.

Hvis inngangssignalet til et 2. ordens system er en plutselig forandring (step-funksjon), avhenger responsen av verdien på ζ . Hvis $\zeta = 1$ (kritisk dempning), får vi ikke noe oversving. Hvis $\zeta < 1$, får vi mer eller mindre sterke oscillasjoner før systemet kommer til ro.

5.3. Sinus-respons til 1. og 2. ordens system, Bode-plot

Istedenfor å teste med en step-funksjon, kan vi bruke et sinus-signal hvor vi måler amplitude og fase til utgangssignalet som funksjon av frekvensen på inngangssignalet (Bode-plot). Vi er da interessert i den stasjonære AC-responsen, dvs. forholdene etter at sinus-signalet har stått på en stund slik at transienter har dødd ut. I overføringsfunksjonen skal da den komplekse frekvensen s erstattes med "vanlig" vekselstrøm-frekvens $j\omega$.

5.3.1. Systemer av 1. orden

For et 1. ordens system får vi når s erstattes med $j\omega$

$$A(j\omega) = V(j\omega)/V_i(j\omega) = 1/(1 + j\omega T)$$

som har modulverdi $|A(j\omega)| = 1/\sqrt{1+(\omega T)^2}$ og fase $\phi(\omega) = -\arctan(\omega T)$

For $\omega T \ll 1$:	$A = 1$	$ A = 0dB$	$\phi = 0$
For $\omega T = 1$:	$A = 1/(1+j)$	$ A = 1/\sqrt{2} = -3dB$	$\phi = -45\text{ deg}$
For $\omega T \gg 1$:	$A = 1/(j\omega T)$	$ A = 1/(\omega T)$	$\phi = -90\text{ deg}$

Ved grafisk fremstilling bør vi bruke log. skala for frekvens og dB-skala (dvs. log. skala) for modulverdi, og lineær skala for fase. Variasjonen i modulverdi er da godt beskrevet ved to asymptoter (rette linjer) som skjærer hverandre ved grensefrekvensen (knekk-frekvensen) $\omega_0 = 1/T$. Høyfrekvens-asymptoten synker med 20 dB/dekade.

Forløpet til fassen er godt beskrevet ved en rett linje fra et punkt på lavfrekvens-asymptoten ved $\omega = 0.1/T$ til et punkt på høyfrekvens-asymptoten ved $\omega = 10/T$ hvor T er tidskonstanten.

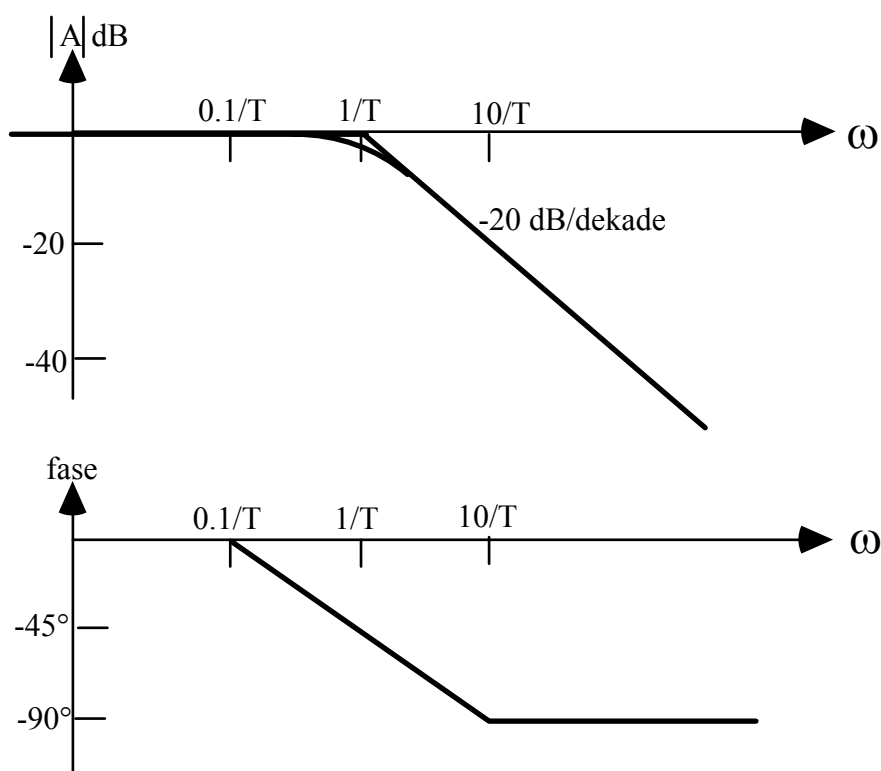


Fig. 5.6 Modul og faseforløp for 1. ordens system

For modulverdien er det maksimale avvik fra asymptotene -3 dB ved grensefrekvensen. For faseforløpet er maksimalt avvik ca. 6 grader. Fordelen ved å bruke en slik fremstilling er at ved seriekopling av flere kretser kan vi addere desibelverdiene og faseverdiene for å finne forløpene for den sammensatte kretsen.

5.3.2. Systemer av 2. orden

Responen til et mekanisk 2. ordens system med en sprang-funksjon som inngangssignal er et oscillatorisk innsving til den stasjonære verdien hvis dempingen er < 1 . Nå vil vi se på den stasjonære AC-responen med et sinus-signal som inngangssignal, og velger den elektriske 2. ordens kretsen som eksempel. Overføringsfunksjonen er

$$A(s) = V(s)/V_i(s) = 1/(CLs^2 + RCs + 1)$$

Frekvensresponsen finner vi ved å erstatte s med $j\omega$. Dette gir

$$A(j\omega) = 1/(1 - \omega^2 CL + j\omega RC)$$

Det asymptotiske forløp er vist i tabellen og fig 5.7. (Jfr. figurer i Bentley s. 56 og Kuo: Network Analysis s. 227 og 228.)

For $\omega \ll 1/\sqrt{CL}$ og $\omega \ll 1/RC$	For $\omega = 1/\sqrt{CL} = \omega_0$	For $\omega \gg 1/\sqrt{CL}$ og $\omega \gg 1/RC$
$A = 1$ $ A = 1$ $\phi = 0 \text{ deg}$	$A = -j\sqrt{L/C}/R$ $ A = \sqrt{L/C}/R$ $\phi = -90 \text{ deg}$	$A = -1/(\omega^2 CL)$ $ A = 1/(\omega^2 CL)$ $\phi = -180 \text{ deg}$

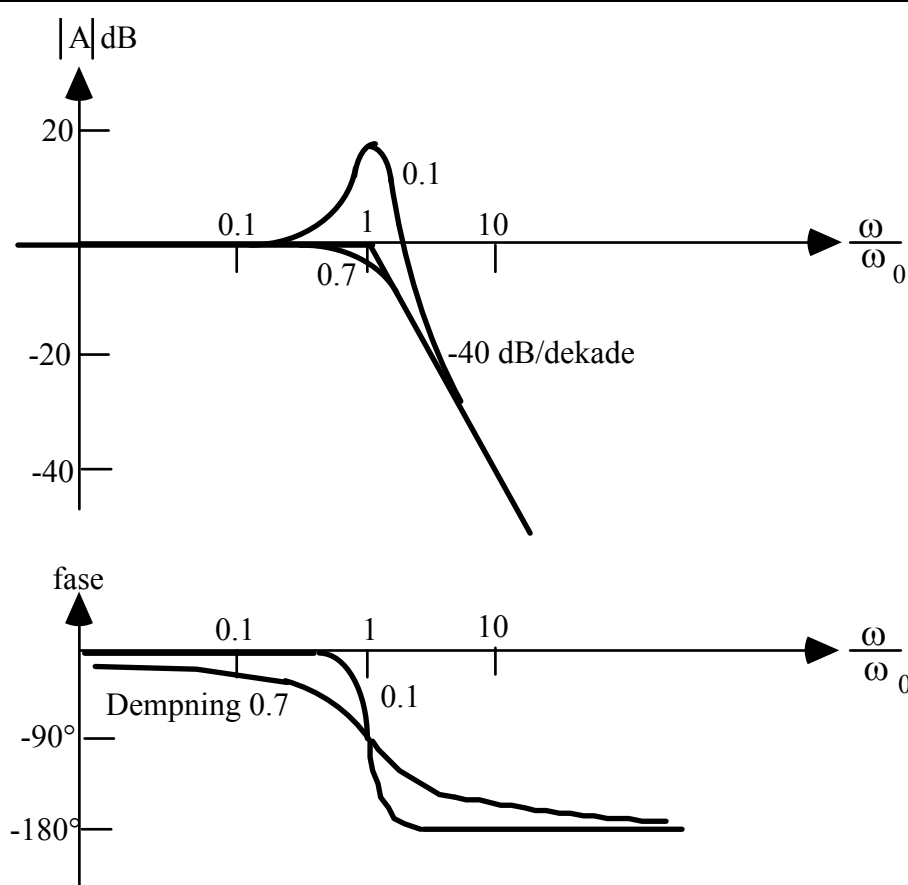


Fig. 5.7 Modul- og fase-forløp for dempning lik 0.1 og 0.7.

Forløpet for frekvenser omkring den udedempede egenfrekvensen $\omega_0 = 1/\sqrt{CL}$ er bestemt ved dempningen. Hvis $R\sqrt{C/L} \ll 1$, vil amplitudeforløpet få en høy topp nær egenfrekvensen, og fasen endres steilt fra 0 til -180 grader. Hvis dempningen er stor, $0.5R\sqrt{C/L} > 1$, vil overføringsfunksjonen ha to reelle poler (nullpunkter for nevneren), og amplitudeforløpet får to knekk-punkter. For kritisk dempning ($0.5R\sqrt{C/L} = 1$) er modulverdien -6 dB ved egenfrekvensen.

I forbindelse med resonanskretser brukes ofte Q -verdi (Q for *quality*) for å karakterisere hvor skarp resonanstoppen er. I en seriekrets med motstand R , spole L og kondensator C er Q -verdien lik reaktansen X til L eller C ved resonansfrekvensen dividert med seriemotstanden i kretsen. I seriemotstanden R inngår også eventuell ohmsk motstand i spolen. Resonansfrekvensen er

$$\omega_0 = 1/\sqrt{CL}$$

som gir

$$Q = X/R = \omega_0 L/R = 1/\omega_0 CR$$

Sammenhengen med dempningsparameteren blir

$$2\zeta = 1/Q$$

Hvis vi bruker spenningen over spolen i *fig 5.5* (i stedet for kondensatoren) som utgangssignal, får vi et 2. ordens høy-pass filter. Hvis vi bruker spenningen over motstanden som utgangssignal, får vi bånd-pass filter. Ved kombinasjoner av motstander, kondensatorer og operasjonsforsterkere kan vi lage aktive filtre av annen og høyere orden uten spoler, noe som kan være en stor fordel.

6. Tilbakekoblede systemer og reguleringsystemer

Vi vil nå se litt mer på stabilitet i et system med tilbakekobling. Et blokkskjema for et system med tilbakekobling er vist i *fig. 6.1*.

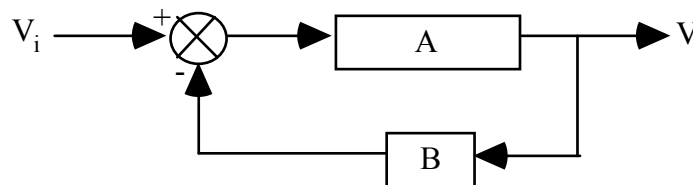


Fig. 6.1 Blokkskjema for et tilbakekoblet system

Fra figuren får vi

$$(V_i - VB)A = V$$

dvs. overføringsfunksjonen for systemet blir

$$\frac{V(s)}{V_i(s)} = \frac{A(s)}{1 + A(s)B(s)}$$

For $s = j\omega$ gir ligningen frekvenskarakteristikken for det tilbakekoblede systemet. Selv om komponentene i kretsen er valgt slik at vi har negativ tilbakekobling for de signalene som skal overføres, kan faseforskyvning i tilbakekoblingsløyfen føre til at tilbakekoblingen blir positiv for andre frekvenser, og dette kan føre til ustabilitet. Det avhenger av modul og fase til A og B .

Ved praktisk beregning kan det ofte være fordelaktig å flytte rundt på blokker. Dette kan vi gjøre hvis blokkene modifiseres slik at vi hele tiden får den samme

overføringsfunksjonen. F.eks. er skjemaet i *fig. 6.1* ekvivalent med blokkskjemaet i *fig. 6.2*.

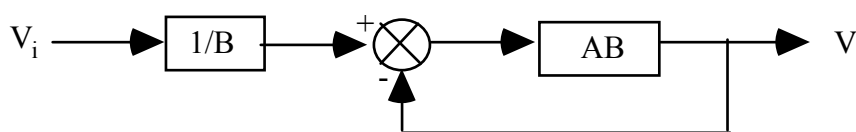


Fig. 6.2 Enhets-tilbakekobling

Denne typen tilbakekobling kalles enhets-tilbakekobling fordi hele utgangssignalet er tilbakekoblet. I det følgende vil vi holde oss til et slikt skjema fordi formlene da blir litt enklere (men det er altså fullt mulig å regne seg tilbake til det opprinnelige skjemaet).

En ikke-inverterende op.amp. kobling er et eksempel på det første blokkskjemaet, mens en op.amp. koblet som spenningsfølger svarer til en enhets-tilbakekobling.

Alle reguleringssystemer bruker en eller annen form for tilbakekobling for å oppnå ønsket resultat. Reguleringssystemer kan deles i to hovedtyper: På/av-regulering og kontinuerlig regulering. Et kjøleskap er et eksempel på et på/av system. Hvis temperaturen går over en viss verdi, kobler termostaten inn kjølekompressoren. Når temperaturen har sunket til en annen verdi, kobles så kompressoren ut igjen. Vi må ha litt hysteresis for å unngå altfor hyppige inn- og ut-koblinger.

Ved kontinuerlig regulering prøver vi hele tiden å påvirke systemet slik at det har den tilstanden vi ønsker. Hvis det blir utsatt for forstyrrelser, må det komme til ro igjen og ikke bli ustabil. I det følgende skal vi se litt på hvordan vi kan lage et kontinuerlig reguleringssystem og finne betingelser for stabilitet.

Som eksempel kan vi ta et vannbad som skal kunne varmes opp til en valgt temperatur, og holde seg på denne temperaturen selv om omgivelsestemperaturen varierer. Dette burde kunne løses ved å ha et potensiometer hvor vi stiller inn en referansespenning som styrer en spenningsforsterker som igjen styrer en effektforsterker koblet til et varmeelement i vannbadet. Temperaturen i vannet måles med en sensor som gir en spenning proporsjonal med temperaturen. Denne spenningen sammenlignes med referansespenningen fra potensiometeret, og avviket brukes som styrespenning for forsterkeren. Hvis vanntemperaturen varierer på grunn av varierende omgivelsestemperatur, korrigeres tilført effekt til varmeelementet. Et blokkskjema for systemet er vist i *fig. 6.3*. Fra figuren ser vi at utgangsverdien $y(t)$ sammenlignes med referanseverdien $r(t)$, og avviket (feil-signalet) $e(t) = r(t) - y(t)$ brukes som pådrag for regulatoren (*negativ tilbakekobling*). Vi kaller overføringsfunksjonen for regulatoren H_1 og for prosessen inkludert temperaturføleren H_3 . Overføringsfunksjonen for varierende omgivelsestemperatur (omregnet til varierende effekt) kalles H_2 .

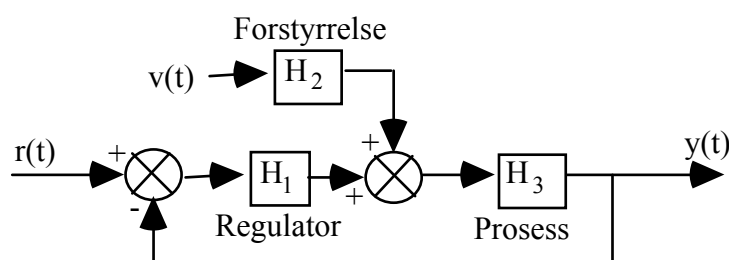


Fig. 6.3 Reguleringsystem

Dette gir følgende overføringsfunksjoner (Laplace-transformerte størrelser):

Når $v(t)$ er konstant:

$$\frac{Y(s)}{R(s)} = \frac{H_1 H_3}{1 + H_1 H_3}$$

$$\frac{E(s)}{R(s)} = \frac{1}{1 + H_1 H_3}$$

Når $r(t)$ er konstant:

$$\frac{Y(s)}{V(s)} = \frac{H_2 H_3}{1 + H_1 H_3}$$

I analysen er det hensiktsmessig å se på innvirkningen av $r(t)$ og $v(t)$ hver for seg, men prosessens utgang er gitt ved summen av overføringsfunksjonene.

Alle overføringsfunksjonene har

$$\frac{1}{1 + H_1 H_3}$$

som felles faktor. Produktsummen av overføringsfunksjonene i sløyfen:

$$A(s) = H_1 H_3$$

kalles *sløyfeforsterkningen (loop gain)* eller *sløyfe-transferfunksjonen*.

Vi ser først på det tilfellet at $v(t)$ er konstant:

$$\frac{E(s)}{R(s)} = \frac{1}{1 + A(s)} = \frac{1}{1 + H_1 H_3} = N(s)$$

kalles *avviksforholdet*.

$$\frac{Y(s)}{R(s)} = \frac{A(s)}{1 + A(s)} = N(s)A(s) = M(s)$$

kalles *følgeforholdet* og uttrykker hvor godt utgangen $y(t)$ følger referansen $r(t)$.

Hvis $|A(s)| \gg 1$ får vi $|M(s)| = 1$ og $Y(s) = R(s)$, dvs. *perfekt følgerregulering*.

Dette fører også til at avviksforholdet $|N|$ vil bli tilnærmet null.

For virkningen av forstyrrelsen $v(t)$ har vi (når $r(t)=0$) at med tilbakekobling er

$$Y(s) = \frac{V(s)H_2H_3}{1 + H_1H_3} = V(s)H_2H_3N(s)$$

Uten tilbakekobling ville virkningen være

$$V(s)H_2H_3$$

Forholdet mellom utgangssignalene med og uten tilbakekobling blir derfor $N(s)$, dvs. $N(s)$ uttrykker hvor godt forstyrrelser på utgangen undertrykkes med tilbakekobling. Hvis avviksforholdet

$$|N(s)| = 1$$

får vi ingen forbedring.

Hvis

$$|N(s)| = 0$$

får vi fullstendig undertrykkelse av forstyrrelsen.

6.1. Frekvensanalyse

Frekvensresponsen til det tilbakekoblede systemet kan vi studere ved å erstatte s med $j\omega$. Egenskapene til systemet er karakterisert ved

sløyfeforsterkningen $A(j\omega)$ og

avviksforholdet $N(j\omega) = 1/(1 + A(j\omega))$.

Vi ser at når $|A(j\omega)| \gg 1$ er $N(j\omega) = 1/A(j\omega)$

Når $|A(j\omega)| \ll 1$ er $N(j\omega) = 1$.

For modulverdiene i desibel (dB) har vi derfor at

$$|N(j\omega)|_{dB} = -|A(j\omega)|_{dB} \quad \text{når} \quad |A|_{dB} \gg 0$$

og $|N(j\omega)|_{dB} = 0 \text{ dB}$ når $|A|_{dB} \ll 0$

For fasevinkelen får vi

$$\text{fasen}(N(j\omega)) = -\text{fasen}(A(j\omega)) \quad \text{når} \quad |A|_{dB} \gg 0$$

og $\text{fasen}(N(j\omega)) = 0$ når $|A|_{dB} \ll 0$.

Vi ser at $|A|_{dB}$ og $|N|_{dB}$ kurvene er speilbilder om 0 dB linjen for $|A| > 0 \text{ dB}$.

Fasekurvene er speilbilder om 0 linjen. Dette er asymptotiske tilnærminger for $N(j\omega)$ når $A(j\omega)$ er kjent. Mere nøyaktige verdier for $N(j\omega)$ kan vi finne ved å bruke spesielle grafiske fremstillinger (Nichols-diagrammer).

6.1.1. Stabilitetskrav

For signalfrekvenser i det området kretsen er beregnet for, ønsker vi $|A| \gg 1$ og negativ tilbakekobling. For økende frekvenser vil vanligvis $|A|$ avta og fasen forandre seg. Hvis fasen forandrer seg så meget som 180 grader i forhold til det vi

startet med (dvs. negativ tilbakekobling), blir tilbakekoblingen positiv. For å unngå problemer (økende forsterkning, ustabilitet) må da samtidig $|A|$ (modulverdien av sløyfeforsterkningen) ha blitt mindre enn 1 (dvs. $0dB$). Et vanlig krav er at modulverdien skal ha avtatt til $-6dB$ ved den kritiske frekvensen hvor faseforandring blir 180 grader. Forsterkningen i det tilbakekoblede systemet blir da

$$|Y/R| = |A/(1 + A)| = 0.5/0.5 = 1$$

dvs. ingen økning i forsterkning.

En alternativ formulering er å si at sløyfeforsterkningen må ha avtatt til $0dB$ når faseforandringen har blitt 150 grader (30 grader fasemargin).

7. Modell-tilpasning. Minste kvadraters metode.

På grunnlag av et sett målinger ønsker vi ofte å foreta en sammenligning med en modell. Modellen kan være en teoretisk formel, eller det kan være at vi ønsker å beskrive målingene ved en enkel funksjonssammenheng (f.eks. et polynom). I alle tilfeller må vi velge et kriterium (*figure-of-merit function*) som måler hvor godt data og modell med et visst valg av parametre stemmer overens. Målingene vil typisk være beheftet med målefeil slik at de aldri eksakt vil stemme med modellen. Vi bør derfor på grunnlag av statistisk teori kunne avgjøre hvor store avvik vi kan godta, dvs. om modellen er brukbar eller må forkastes.

Et program som foretar en tilpasning av målte data til en modell bør derfor kunne gi :

- verdier for parametre
- estimat av usikkerhet i parametre
- statistisk mål for nøyaktigheten i tilpasningen (*goodness-of-fit*)

Den vanligste metoden for tilpasning er å bestemme parametrene slik at summen av det kvadratiske avvik mellom målte verdier og beregnede verdier blir minst mulig (minste kvadraters metode). Hvis målefeilene er uavhengige og normal-fordelte med konstant standard avvik, sier statistisk teori at dette gir det beste estimat av parametrene (*maximum likelihood estimatio*). Målefeilene er ofte ikke helt normal-fordelt, men metoden gir likevel en brukbar indikasjon på hvor sannsynlig det er at den valgte modellen passer med målingene.

7.1. Minste kvadraters metode.

Anta at vi har N målte verdier (x_i, y_i) og en modell med M justerbare parametre som beskriver en funksjonell sammenheng mellom x og y . Vi ønsker å bestemme parametrene a_1, a_2, \dots, a_M slik at vi får minimum for *chi-square*

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - y(x_i, a_1, \dots, a_M)}{\sigma_i} \right)^2$$

hvor σ_i er usikkerheten (standard avviket) for y_i .

Minimum finnes ved å sette de partiell-deriverte av χ^2 med hensyn på parametrene lik null. Dette gir M ligninger (kalles gjerne normal ligningene) som må løses for å finne verdiene på parametrene. Fra verdien av χ^2 ved minimum kan en regne ut sannsynligheten Q (gitt ved *incomplete gamma function*) for å få en større verdi på χ^2 enn den vi fant. Hvis Q er svært liten, er det derfor lite sannsynlig at modellen er riktig (eller usikkerhetene i de målte verdiene er vesentlig større enn antatt). En tommelfinger regel er å si at for en brukbar tilpasning er χ^2 omtrent lik antall frihetsgrader, dvs. lik $N - M$.

7.2. Det lineære tilfelle.

Modellen er her en lineær kombinasjon av funksjoner av x . I det enkleste tilfelle kan det være potensrekken $1, x, x^2, \dots, x^{M-1}$ slik at y er den lineære kombinasjonen

$$y(x) = a_1 + a_2x + a_3x^2 + \dots + a_Mx^{M-1}$$

dvs. et polynom av grad $M-1$.

Mer generelt kan vi ha

$$y(x) = \sum_{k=1}^M a_k X_k(x)$$

hvor $X_m(x)$ er funksjoner av x . Ett eksempel kan være Chebyshev polynomer som anvendes mye i data analyse. Man kan også bruke ikke-lineære funksjoner i x , men modellens avhengighet av parametrene skal være lineær.

Noen tilsynelatende ikke-lineære problemer kan gjøres lineære. F.eks. ligningen for en termistor

$$R = a e^{b/T}$$

med parametre a og b kan skrives som

$$\ln R = c + b/T$$

med $c = \ln(a)$. Denne relasjonen er lineær i parametrene c og b .

Det er laget mange program for tilpasning med minste kvadraters metode. Problemet er at ligningssettet som skal løses, ofte er tilnærmet singulært slik at det kan være nødvendig å bruke spesielle metoder for å få fornuftige resultater. En bør også merke seg at en tilpasning av data direkte til en potensrekke i x kan skape problemer. Derfor anbefales bruk av ortogonale polynomer (f.eks. Chebyshev polynomer).

7.3. Eksempel

Vi ønsker å tilpasse en rett linje til fire målte verdier y_i (avhengig variabel) svarende til fire forskjellige verdier x_i (fri variabel). Standard avvik for de målte verdiene er σ_i . Den rette linjen skrives som

$$y = a_1 + a_2 x$$

og oppgaven er å

- 1) bestemme parametrene a_1 og a_2
- 2) bestemme usikkerheten (st. avvik) for a_1 og a_2
- 3) bedømme om tilpasningen er akseptabel, dvs. avgjøre om det er sannsynlig at y varierer lineært med x .

Data-verdiene er

x_i :	0.0	1.0	2.0	3.0
y_i :	1.4	1.5	3.7	4.1
σ_i :	0.5	0.5	0.5	0.5

Vi skal altså finne de verdiene av a_1 og a_2 som gjør

$$\chi^2 = \sum_{i=1}^4 (y_i - a_1 - a_2 x_i)^2 / \sigma_i^2 = \sum_{i=1}^4 w_i (y_i - a_1 - a_2 x_i)^2$$

minst mulig.

Ved minimum må de deriverte av χ^2 m.h.p. a_1 og a_2 være lik null. Dette gir to ligninger

$$\sum_i w_i (y_i - a_1 - a_2 x_i) = 0$$

$$\sum_i w_i (y_i - a_1 - a_2 x_i) x_i = 0$$

til bestemmelse av de to ukjente. Herav får vi

$$\left(\sum_i w_i \right) a_1 + \left(\sum_i w_i x_i \right) a_2 = \sum_i w_i y_i$$

$$\left(\sum_i w_i x_i \right) a_1 + \left(\sum_i w_i x_i^2 \right) a_2 = \sum_i w_i x_i y_i$$

Ligningene kan skrives på matrisform

$$\begin{Bmatrix} \sum w_i & \sum w_i x_i \\ \sum w_i x_i & \sum w_i x_i^2 \end{Bmatrix} \begin{Bmatrix} a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} \sum w_i y_i \\ \sum w_i x_i y_i \end{Bmatrix}$$

Alle summene går over indeks i fra 1 til 4. Kaller vi de tre matrisene B , A og C , kan vi skrive matriseligningen

$$BA = C$$

som har løsningen

$$A = B^{-1}C$$

hvor B^{-1} er den inverterte av B -matrisen. Setter vi inn tallene, får vi

$$\begin{Bmatrix} 16 & 24 \\ 24 & 56 \end{Bmatrix} \begin{Bmatrix} a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} 42.8 \\ 84.8 \end{Bmatrix}$$

Herav får vi

$$\begin{Bmatrix} a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} 56/320 & -24/320 \\ -24/320 & 16/320 \end{Bmatrix} \begin{Bmatrix} 42.8 \\ 84.8 \end{Bmatrix}$$

som gir

$$a_1 = 1.13 \quad \text{og} \quad a_2 = 1.03$$

Den inverterte matrisen B^{-1} kalles korrelasjonsmatrisen. Det kan vises at standard avvik for parametrene a_1 og a_2 er gitt ved kvadratroten av diagonalelementene i denne matrisen. Vi får derfor standard avvikene $\sigma(a_1) = 0.42$ og $\sigma(a_2) = 0.22$.

Når parametrene er bestemt, kan vi regne ut chi-square og får verdien

$$\chi^2 = 3.13$$

Tabeller over chi-square viser at sannsynligheten for å få denne verdien eller en større verdi for chi-square, er 0.21. Det betyr at det er sannsynlig at data indikerer en lineær sammenheng mellom y og x . En stor verdi for chi-square (eller en svært liten verdi) ville bety at det var usannsynlig at det var noen slik sammenheng.

For å få et inntrykk av hvor god tilpasningen er, bør målte og beregnede verdier fremstilles grafisk. Ofte får en best inntrykk av systematiske avvik ved å plote forskjellen mellom målt verdi og beregnet verdi (dvs. "residuals"), jfr. figurene nedenfor.

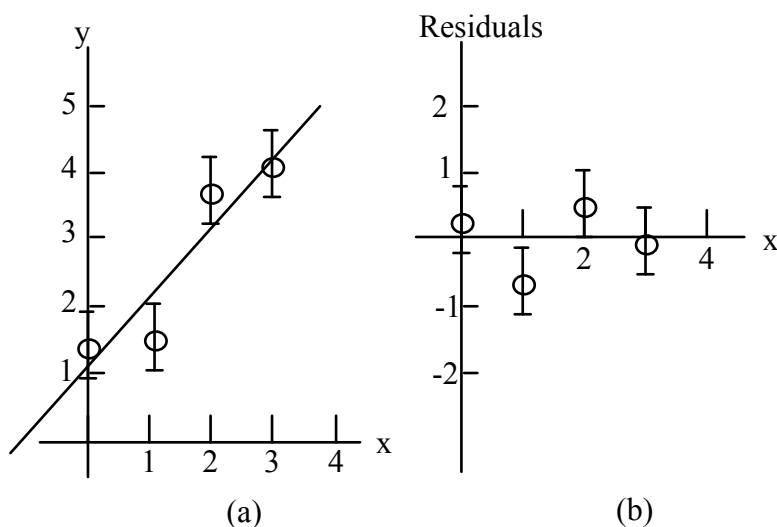


Fig. 7.1 (a) Tilpasning av rett linje til målte verdier med standard avvik.
(b) Forskjell mellom målte og beregnede verdier.

I korrelasjonsmatrisen angir elementene utenom diagonalen korrelasjonen mellom parametrene. Når disse elementene er forskjellige fra null, vil det si at en forandring i en parameter delvis kan kompenseres med en forandring i en annen parameter. Hvis vi i dette eksempelet f.eks. øker verdien for a_2 med ett st. avvik, øker chi-square til 5.9, dvs. en fordobling. Hvis vi nå varierer a_1 , finner vi at chi-square har minimums verdi lik 4.1 for a_1 lik 0.80. Tilsvarende finner vi at hvis verdien for a_1 økes med ett st. avvik, får vi minimum lik 4.1 for chi-square hvis a_2 forandres til 1.23.

I tabellen er det vist en del verdier for parametrene a_1 og a_2 som gir verdien 4.1 for chi-square. Plotter vi sammenhørende verdier av a_1 og a_2 , får vi en ellipse ("korrelasjons-ellipse"). Innenfor ellipsen er chi-square mindre enn 4.1, og den minste verdien er altså 3.13 svarende til $a_1=1.13$ og $a_2=1.03$.

a_1	a_2	chi-square
0.72	1.20	4.1
0.80	1.25	4.1
0.80	1.10	4.1
1.00	1.21	4.1
1.00	0.96	4.1
1.26	0.85	4.1
1.35	1.05	4.1
1.46	0.81	4.1
1.52	0.90	4.1

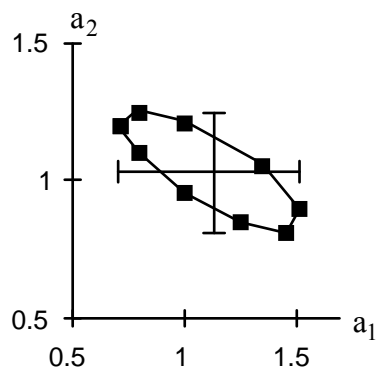


Fig. 7.2 Tabell over sammenhørende verdier av parametrene som gir chi-square lik 4.1 og den tilsvarende korrelasjons-ellipsen.

Hvis ikke standard avvik for de målte verdiene er kjent, settes vektene $w_i = 1/\sigma_i^2 = 1$. I dette tilfelle er det vanlig å estimere usikkerheten i parametrene slik: Kvadratrotten av diagonalelementene i korrelasjonsmatrisen multipliseres med en faktor lik

$$(\chi^2/(N - \nu))^{1/2}$$

hvor ν er antall parametre som skal tilpasses. I vårt eksempel får vi:

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 10.7 \\ 21.2 \end{pmatrix}$$

Herav får vi

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 14/20 & -6/20 \\ -6/20 & 4/20 \end{pmatrix} \begin{pmatrix} 10.7 \\ 21.2 \end{pmatrix}$$

som gir $a_1 = 1.13$ og $a_2 = 1.03$ som tidligere.

For chi-square får vi nå 0.783. Dette gir de estimerte st. avvik for parametrene

$$\sigma(a_1) = \sqrt{14/20} \sqrt{0.783/2} = 0.52 \quad \text{og} \quad \sigma(a_2) = \sqrt{4/20} \sqrt{0.783/2} = 0.28$$

som stemmer bra med de tidligere verdiene.

7.4. Ikke-lineære problemer.

Hvis modellen ikke avhenger lineært av parametrene som f.eks. ved en Gauss-kurve

$$y = a e^{-\left(\frac{x-b}{c}\right)^2}$$

med parametrene a , b og c må en bruke spesielle metoder. Det består i å anta rimelige start verdier for parametrene og så finne bedre verdier ved iterasjon.

8. Digital filtrering.

Med et signal forstår vi normalt en målbar fysikalsk størrelse som forandrer seg som funksjon av tiden. Hvis signalet er definert ved alle relevante tidspunkt, kaller vi det et analogt signal. Hvis vi tenker oss at vi måler signalet bare ved visse tidspunkt ("sampler" det) med en analog til digital omvandler, får vi en sekvens med tall som vi kaller et digitalt signal. Vanligvis vil vi måle med konstante samplingsintervall (tidsintervall) svarende til en konstant samplingsfrekvens.

Digital signalbehandling består i å analysere og manipulere det digitale signalet, for eksempel

- produsere en ny sekvens av tall der frekvensinnholdet er modifisert i forhold til det opprinnelige signalet (digital filtrering).
- produsere en sekvens av tall som gir uttrykk for frekvensinnholdet i det gitte signalet (diskret fouriertransformasjon).

Hvis tidsintervallet mellom hvert sample er 1 sek, er samplingsfrekvensen $f_s = 1$ Hz. Den høyeste frekvensen som gjengis korrekt, er gitt ved Nyquist-frekvensen $f_s / 2 = 0.5$ Hz. Høyere frekvenser vil (hvis de ikke er fjernet av et lavpass filter foran AD-konverteren), bli speilet om denne frekvensen ned i lavfrekvensområdet. For eksempel vil en signalfrekvens på 1.25×0.5 Hz gi et "alias-signal" med frekvens 0.75×0.5 Hz. Videre har vi at hvis vi tar N (f.eks. 100) sampler av et analogt signal, og samplingsfrekvensen er f_s (f.eks. 1 Hz), vil vi få med akkurat en hel periode av et sinus signal med frekvens $f_s / N = 0.01$ Hz.

8.1. Ikke-rekursive filter. Utjevning.

Et vanlig resultat av en måleserie er at det digitale signalet (sekvensen av samplede verdier) inneholder støy som vi ønsker å redusere, i tillegg til det signalet vi er interessert i. Hvis signalet vi ønsker å fremheve er lavfrekvent, kan vi prøve en form for lavpass filter.

En måte er å bruke en form for utjevning eller glatting. Vi lager da et nytt digitalt signal (sekvens av verdier) hvor hver verdi er en eller annen form for midling over noen naboeverdier i den opprinnelige sekvensen. Sekvensen inneholder $N = 512$ verdier (sampler). Hvis vi for eksempel midler over 5 naboeverdier, får vi

$$y(n) = (u(n-2) + u(n-1) + u(n) + u(n+1) + u(n+2)) / 5$$

for $n = 2$ til 509 hvor $u(n)$ er en verdi i den opprinnelige sekvensen og $y(n)$ en verdi i den nye. For de to første og de to siste verdiene må vi modifisere formelen, men for enkelhets skyld setter vi disse verdiene lik de gamle, dvs.

$$y(0) = u(0), \quad y(1) = u(1), \quad y(510) = u(510), \quad y(511) = u(511).$$

Dette er et eksempel på et ikke-rekursivt filter som generelt kan skrives som

$$y(n) = \sum_{k=-K}^K c(k) u(n-k)$$

hvor K angir hvor langt ut til hver side vi går i filtreringen, og $c(k)$ er filterkoeffisientene. I vårt tilfelle er $K = 2$ og alle koeffisientene $1/5$. Vi kan se på filteringen som en prosess hvor vi ser på den opprinnelige sekvensen gjennom et

vindu (filteret). Etter å ha beregnet den nye verdien, forskyver vi vinduet et step og beregner neste filtrerte verdi, osv.

I stedet for å ta middelverdien av noen punkter (som svarer til tilpasning med en rett linje), brukes det ofte mer kompliserte polynomer. Vi kunne for eksempel bruke minste kvadraters metode og finne en tilpasning til et tredjegrads polynom gjennom fem punkter. I stedet for koeffisientene $\{1, 1, 1, 1, 1\} / 5$ (midling over 5 punkter) får vi da $\{-3, 12, 17, 12, -3\} / 35$. I litteraturen er det angitt mange forskjellige koeffisient sett for glatting over flere punkter.

8.2. Overføringsfunksjon for et digitalt filter.

For analoge signaler kan vi regne ut en overføringsfunksjon (systemfunksjon) som multiplisert med inngangssignalet gir utgangssignalet. Når vi er interessert i frekvensavhengigheten (stasjonære signaler), tegner vi gjerne opp et Bode-plot som viser forsterkning som funksjon av frekvens.

For digitale filter kan vi også finne en overføringsfunksjon som fremstilt grafisk som funksjon av frekvensen, svarer helt til et Bode-plot. Som eksempel kan vi bestemme overførings-funksjonen for et filter som en 5 punkts midling slik:

Vi starter med å se på hva som hender hvis inngangssignalet er en enkel frekvens

$$u(n) = e^{j\omega n}.$$

Dette gir

$$\begin{aligned} y(n) &= \{e^{j\omega(n-2)} + e^{j\omega(n-1)} + e^{j\omega n} + e^{j\omega(n+1)} + e^{j\omega(n+2)}\} / 5 \\ &= e^{j\omega n} \{e^{-j2\omega} + e^{-j\omega} + 1 + e^{j\omega} + e^{j2\omega}\} / 5 \\ &= e^{j\omega n} \{2\cos(2\omega) + 2\cos(\omega) + 1\} / 5 \end{aligned}$$

Inngangsfrekvensen blir altså multiplisert med en faktor som avhenger av frekvensen og filter koeffisientene. Et plot av denne funksjonen viser hvordan hver enkelt frekvens blir forsterket (forminsket) på grunn av filteret. (Vi plotter da gjerne som funksjon av "vanlig" frekvens i stedet for sirkelfrekvens.)

I stedet for å utlede et analytisk uttrykk for overføringsfunksjonen kan vi beregne noen utgangsverdier når vi sender inn et signal med bare en frekvens. Forholdet mellom en utgangsverdi og en inngangsverdi med samme nummer i sekvensen, er verdien for overførings-funksjonen for denne frekvensen. Ved å gjøre dette for en del frekvenser kan vi så tegne opp kurven for funksjonen. Som eksempel kan vi tenke oss at vi sampler et kosinus signal med samplefrekvens 1 Hz, dvs. 1 sek mellom hvert sample. Hvis signal frekvensen er 0.25 Hz, får vi 4 sampler pr. periode. Bruker vi for enkelthets skyld bare 3 punkts midling, får vi

<i>Inn</i>	<i>Ut</i>	<i>Forsterkning</i>
$u(0)=1$		
$u(1)=0$	$y(1)=0$	
$u(2)=-1$	$y(2)=-0.33$	0.33
$u(3)=0$	$y(3)=0$	
$u(4)=1$	$y(4)=0.33$	0.33
	<i>osv.</i>	

Med signalfrekvens 0.5 Hz får vi to sampler per periode og resultatene

<i>Inn</i>	<i>Ut</i>	<i>Forsterkning</i>
$u(0)=1$		
$u(1)=-1$	$y(1)=0.33$	-0.33
$u(2)=1$	$y(2)=-0.33$	-0.33
$u(3)=-1$	$y(3)=0.33$	-0.33
	<i>osv.</i>	

Ved å undersøke flere frekvenser kan en se at funksjonen har ett nullpunkt mellom 0.25 Hz og 0.50 Hz. Nullpunktet for 3 punkts midling ligger ved ca. 0.33 Hz. Hvis frekvensen går mot 0, går funksjonen mot 1.

8.3. Rekursive filter.

Et filter hvor de nye verdiene beregnes etter en formel

$$y(n) = \sum_{k=0}^K c(k) u(n-k) + \sum_{k=0}^K d(k) y(n-k)$$

kalles et rekursivt filter fordi vi ved beregningen av $y(n)$ også tar med tidligere beregnede verdier ($y(n-1)$, $y(n-2)$ osv.). Filteret har derfor en slags hukommelse, dvs. utgangsverdiene påvirkes av tidligere beregnede verdier. Vi skal se at det digitale filteret som svarer til et 1. ordens analogt lavpass filter, er et slikt rekursivt filter.

8.3.1. Eksempel: Digitalt lavpassfilter.

For å finne formelen for et digitalt lavpass filter starter vi med et differensial ligningen for et 1. ordens analogt lavpass filter:

$$v_2(t) + RC \frac{dv_2(t)}{dt} = v_1(t)$$

med $v_1(t)$ inngangssignal, $v_2(t)$ utgangssignal og RC tidskonstanten T . På samme måte som i lab.oppg. 1 bruker vi tilnærmelsen

$$\frac{dv_2}{dt} = (v_2(n) - v_2(n-1))/S$$

dvs. den deriverte settes lik forskjellen mellom to nabo sampler dividert med sample intervallet (tidsintervallet). Dette gir differens ligningen

$$v_2(n) = b v_2(n-1) + (1-b) v_1(n) \quad \text{med} \quad b = \frac{T/S}{1 + T/S}$$

hvor $T = RC$ er filterets tidskonstant og S er sample intervall (tidsintervall) Dette er et rekursivt filter, dvs. utgangsverdien avhenger både av inngangsverdi og utgangsverdi, og er altså den digitale ekvivalenten til et 1. ordens analogt lavpass filter. Forholdet mellom tidskonstanten T og sample intervallet S bestemmer hvor sterkt det digitale filteret virker.

For det analoge filteret vil en liten tidskonstant T svarer til en høy verdi for grensefrekvensen. Tilsvarende ser vi for det digitale filteret at en liten verdi for T/S gir en liten verdi for b og liten virkning av det digitale filteret. For b gjelder at

$0 < b < 1$. Med $b = 0$ får vi ingen filtrering, den nye sekvensen er identisk med den opprinnelige.

8.4. Beregning av responsen på grunnlag av filterets impulsrespons.

Ved å bruke Laplace transformasjon omformes differensial ligninger til algebraiske ligninger som kan være lettere å løse.

Den Laplace transformerte $F(s)$ av en tidsfunksjon $f(t)$ er definert ved

$$F(s) = \int_0^{\infty} f(t) e^{-st} dt \quad \text{hvor} \quad s = \sigma + j\omega \quad (\text{kompleks frekvens variabel}).$$

Man kan vise at derivasjon i tidsplanet svarer til å multiplisere den Laplace transformerte med den komplekse variable. Den Laplace transformerte av differensial ligningen for lavpass filteret blir derfor

$$v_2(s) + sTv_2(s) = v_1(s) \quad \text{med} \quad T = RC, \text{ dvs.}$$

$$v_2(s) = \frac{v_1(s)}{1 + sT}$$

Vi antar ingen ladning på kondensator ved $t = 0$. Størrelsen som inngangs signalet $v_1(s)$ skal multipliseres med for å gi utgangssignalet $v_2(s)$, kalles systemets overføringsfunksjon. Denne kan vi ofte lett sette opp ved å bruke de transformerte impedansene R , $1/sC$ og sL . (Obs! Merk likheten med de vanlig komplekse impedansene!)

For lavpass filteret får vi f.eks.

$$v_2(s) = \frac{v_1(s) / \frac{1}{sC}}{R + \frac{1}{sC}} \quad \text{dvs.} \quad v_2(s) = \frac{v_1(s)}{1 + sT} \quad \text{med} \quad T = RC$$

som er det samme som det vi fikk fra differensialligningen. Når $v_1(t)$ er gitt, er det altså "bare" å transformere til $v_1(s)$ og multiplisere med overføringsfunksjonen for å finne $v_2(s)$.

Ofte vil pol-plasseringen i s -planet (nullpunktene til nevneren i produktet) gi nok informasjon om systemet, eller vi kan transformere tilbake til tidsplanet ved hjelp av delbrøk oppspalting, residue regning eller bruk av tabeller.

Eksempler på inngangs signal:

a) Enhets sprangfunksjon (unit step function)

$$v_1(t) = 0 \quad \text{for} \quad t < 0$$

$$v_1(t) = 1 \quad \text{for} \quad t = 0 \quad \text{og} \quad t > 0 \quad \text{gir} \quad v_1(s) = 1/s.$$

b) Enhets impulsfunksjon (delta funksjon) kan tenkes som en uendelig smal, uendelig høy puls (areal = 1) (jfr. kraften som virker ved elastisk støt). Den svarer til den deriverte av enhets sprangfunksjon og den transformerte blir lik 1.

Enhets impulsfunksjon (unit impulse)

$$\text{delta}(t) = \delta(t) \quad \text{transformert} : 1.$$

En slik puls inn på et system gir derfor en respons som er lik overføringsfunksjonen for systemet.

Punkt b) indikerer at det kan være lurt å starte med å studere responsen til et digitalt filter når inngangssignalet er en

digital enhetsimpuls (unit-sample)

definert som sekvensen

$$\begin{aligned} \delta(n) &= 1 \quad \text{for } n = 0 \\ &= 0 \quad \text{for } n \neq 0 \end{aligned}$$

Vi vil bruke dette som inngangssignal i differensligningen (fra lavpass filteret)

$$v_2(n) = av_1(n) + bv_2(n-1), \quad \text{hvor } a = \frac{1}{1 + T/S} \quad \text{og } b = \frac{T/S}{1 + T/S}$$

Med

$$\begin{aligned} v_1(n) &= 1 \quad \text{for } n = 0 \\ &= 0 \quad \text{for } n \neq 0 \end{aligned}$$

får vi sekvensen

$$\begin{aligned} v_2(0) &= a \\ v_2(1) &= ab \\ v_2(2) &= abb \\ v_2(3) &= abbb \\ &\text{osv} \end{aligned}$$

Da b er positiv og mindre enn 1, er dette en sekvens av verdier som avtar eksponentielt mot 0. Dess mindre b er, dess mindre er tidskonstanten i forhold til sample intervallet og dess fortere avtar verdiene mot 0.

Dette passer med det vi fant i lab. oppg. 1 når vi sendte en kort puls (en som avtok eksponentielt med kort tidskonstant i forhold til filterets tidskonstant) inn på et analogt lavpassfilter. Utgangssignalet steg med en kort tidskonstant svarende til pulsens *decay* konstant, og avtok med en lang tidskonstant svarende til filterets tidskonstant. Her har vi gått til grensetilfellet uendelig kort puls og ser bare filterets tidskonstant.

Dette viser hvordan et 1. ordens lavpassfilter reagerer på en enhets impuls. Reaksjonen overfor enhets impulsen kalles filterets impuls respons og betegnes i det følgende med

$$\{h(n)\} \quad (\text{obs.: sekvens av verdier}).$$

Vi skal så vise at hvis vi kjenner systemets impuls respons, kan vi regne ut hvordan det reagerer på et vilkårlig inngangssignal.

Et generelt digitalt signal (dvs. en sekvens av verdier) kan skrives som en sum av enhets impulser med forskjellig vekt og forsinkelse. Eksempel:

$$x(n) = a_0\delta(n) + a_1\delta(n-1) + a_2\delta(n-2) + a_3\delta(n-3) \dots$$

En vilkårlig sekvens kan skrives som

$$x(n) = \sum_{k=0}^{\infty} x(k) \delta(n-k) \dots$$

Hvis vi har et system (f.eks et filter) som er lineært og skift-invariant, gjelder følgende:

Hvis: $\{\delta(n)\}$ gir responsen $\{h(n)\}$
 vil en forsinket enhetsimpuls $\{\delta(n-k)\}$ gi responsen $\{h(n-k)\}$

Tilsvarende gjelder for en vilkårlig sekvens:

Hvis $x(n)$ gir responsen $y(n)$
 vil $x(n-k)$ gi responsen $y(n-k)$.

Hvis transformasjonen av $x(n)$ til $y(n)$ skrives symbolsk som $y(n) = T[x(n)]$

får vi derfor

$$y(n) = T\left[\sum_{k=0}^n x(k) \delta(n-k)\right] = \sum_{k=0}^n x(k) T[\delta(n-k)],$$

dvs.

$$y(n) = \sum_{k=0}^n x(k) h(n-k).$$

Et lineært skift-invariant system er derfor fullstendig karakterisert ved $h(n)$, responsen til en enhets impuls. Responsen er gitt ved *foldning (convolution)* av inngangssekvensen $x(n)$ med systemets impuls respons (sekvensen) $h(n)$. Ligningen ovenfor kalles *convolution sum*. Ordet *foldning* kommer av at vi ved utregningen kan tenke oss de to sekvensene skrevet på to linjer, den ene under den andre. Vi tar så og folder den ene om origo. Sekvensen som er foldet forskyves så mot høyre ett trinn av gangen. I hver posisjon skal vi så summere produktene av de to samplene som står rett overfor hverandre for å beregne $y(n)$.

Eksempel 1: Beregne $y(2)$

$$\begin{array}{cccccc} & x(0) & x(1) & x(2) & x(3) & \dots \\ \dots & h(3) & h(2) & h(1) & h(0) & \dots \end{array}$$

$$y(2) = x(0)h(2) + x(1)h(1) + x(2)h(0).$$

Foldingssummen kan også skrives som $y(n) = \sum_{k=0}^n h(k) x(n-k)$.

Det blir jo det samme hvilken sekvens som foldes.

Eksempel 2:

Inngangssekvens 1 1 1 0 0 0 0 0 inn på filteret med $T/S = 4$, dvs. $b=0.8$

$v_1(n)$	$h(n)$	respons
1	0.2	0.2
1	0.16	0.2+0.16=0.36
1	0.128	0.2+0.16+0.128=0.488
0	0.1024	0.16+0.128+0.1024=0.3904
0	0.0819	0.3127
0	0.0655	0.2502
0	0.0524	0.1998
0	0.0419	0.1598
0	0.0336	0.1279

Dette svarer til et analog lavpass filter påtrykt en firkantpuls.

Hvis systemet ikke svarer til et vanlig filter, men har en impuls respons som f.eks. er gitt ved sekvensen 1 1 1 0 0 0 0 0, vil inngangssignalet ovenfor gi sekvensen 1 2 3 2 1 0 0 0, dvs. en trekantet form.

For lange sekvenser blir det mye regnearbeid. Da det er en nær sammenheng mellom Fourier transformasjon og "folding", kan ofte et program for Fourier transformasjoner også brukes til "folding". Mange dataprogrammer for Fourier transformasjoner forutsetter at antall sampler (målte verdier, N) er en potens av 2, spesielt effektive algoritmer som bruker $N \log_2(N)$ operasjoner er da anvendbare uten at ekstra mellomlagringsplass er nødvendig ("sommerfugl"-algoritmer). Hvis N ikke er en potens av 2, trengs N^2 operasjoner (kan reduseres til $3N \log_2(N)$ ved å anvende "neste" 2er potens).

Hvis det ikke er periodiske forløp (funksjoner), må en være forsiktig når en tolker resultatene. For å unngå randeffekter kan det være nødvendig å forlenge sekvensene med nuller.

"Folding" er et viktig begrep i alle slags målinger. Alle målinger påvirkes av måleapparaturen, og de målte verdiene er de opprinnelige verdiene "foldet" med responsfunksjonen til måleapparaturen. En optisk spektrograf eller en magnetisk spektrograf for ladede partikler vil for eksempel gi en utsmøring, en fordeling som skyldes spalte åpninger osv. Tilsvarende gjelder for en detektor for ladede partikler. Selv om partiklene har en veldefinert energi, vil vi måle en Gaussformet fordeling. I prinsippet kan vi da gjøre to ting. Vi kan enten starte med en eller annen modell ligning for inngangsverdiene med justerbare parametre, folde den med responsfunksjonen og tilpasse parametrene slik at vi får best mulig overensstemmelse med målingene. Den andre muligheten er å starte med målingene og prøve å fjerne innflytelsen av apparaturen ved den omvendte prosessen av folding (*unfolding*, *deconvolution*). Hva som er den beste fremgangsmåten, vil variere.

Program som er laget for folding, kan som regel benyttes for den inverse prosessen også.

9. Analog til digital (A/D) og digital til analog (D/A) omsetter

Analoge signaler er kontinuerlige i tid og verdi, mens digitale signaler består av en sekvens av tall som angir verdien til den variable i diskrete (gitte) tidspunkt. Analoge signaler må digitaliseres, dvs. kvantiseres i tid og verdi for å kunne leses inn og behandles i en datamaskin. Denne prosessen kalles sampling (også kalt tasting, prøvetaking). Sampleintervall er tiden mellom hvert sample, og samplefrekvens er antall sampler pr. tidsenhet (dvs. det inverse av sampleintervall).

9.1. Sampling

Kvantiseringen av et analogt signal kan i prinsippet gjøres med en bryter og en kondensator. Ved regelmessige tidsintervall (sampleintervall) gir bryteren kontakt og kondensatoren lades til signalets øyeblikksverdi. Kondensatoren holder denne verdien til neste måling (sample and hold). Kretsen kan typisk være slik som vist på figuren. Inngangstrinnet skal gi stor strøm for hurtig ladning av kondensatoren, og utgangstrinnet skal gi liten lekkasje av ladningen i hold-perioden.

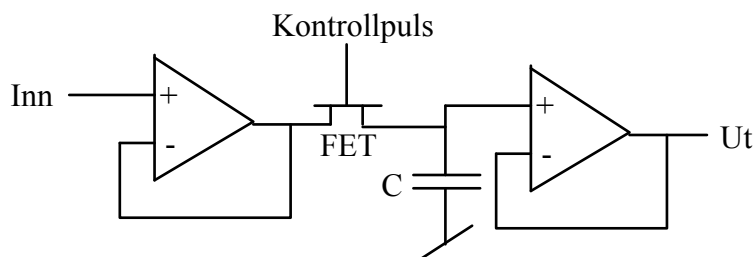


Fig. 9.1 Typisk sample og hold krets

Nyquist sample teorem: For å få korrekt resultat må samplefrekvensen f_{sample} være minst dobbelt så stor som høyeste frekvens f_{max} i det analoge signalet, dvs.

$$f_{sample} > 2f_{max}$$

For et sinusformet signal svarer det til minst to sampler på en periode av signalet. Hvis ikke sample frekvensen er høy nok, får vi *aliasing* ved rekonstruksjon, dvs. et signal med en lavere frekvens enn den opprinnelige.

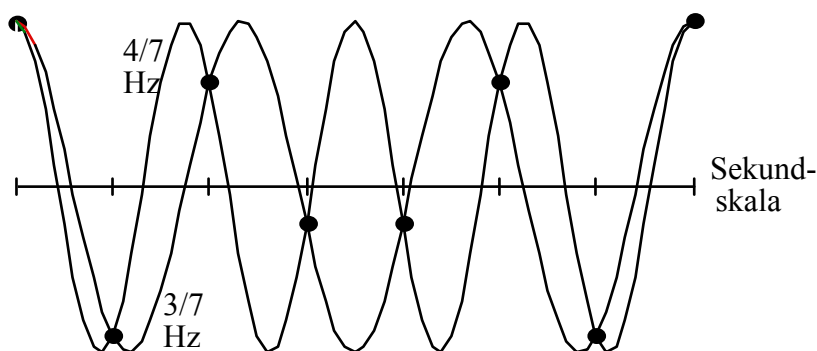


Fig. 9.2 viser et eksempel på aliasing ved for lav samplefrekvens.

$f_{sample} = 1$ Hz. Signalfrekvensene må derfor være mindre enn 0.5 Hz for å unngå aliasing. Signalet har en frekvens på 4/7 Hz, og det er inntegnet en alias-frekvens på 3/7 Hz som passer perfekt til samplene. (symmetri om $3.5/7 = 0.5$ Hz = $f_{sample}/2$).

9.2. Kvantisering

En n-bits A/D omsetter har 2^n forskjellige kvantiseringsverdier 0, 1, 2, ..., $2^n - 1$. (Vi antar at minste verdi er 0.) Dette gir

$$\text{maksimal kvantiseringsfeil} = \frac{V_{max}}{2^n - 1}$$

hvor V_{max} er maksimal inngangsspenning som kan konverteres.

9.3. Koding

Den kvantiserte verdien blir gjort om til et parallelt digitalt signal svarende til en binær kodet verdi av desimaltallet.

I forbindelse med digitale voltmeterer brukes ofte *binary coded decimal* (BCD-) kode. Hvert desimaltall (0,...,9) er da representert med en gruppe på fire binære siffer.

Ofte kan det være hensiktsmessig å bruke *oktal* kode, dvs. bruke 8 som grunntall, eller *hexadesimal* kode med 16 som grunntall.

9.3.1. Gray kode

Dette er en binær kode som er slik at bare *ett bit* skifter av gangen. Denne koden brukes i mekaniske lineære eller vinkel enkodere.

9.4. Polaritet

Forskjellige metoder brukes til å skille mellom positive og negative analoge signaler:

- Fortegn og tallverdi: Bruker ett ekstra bit for fortegn, 0 for positive og 1 for negative tall.
- Binært nullpkt. forskjøvet: Området -5 til +5 volt skiftes til 0 til 10 volt.
- Toers komplement: Det negative tallet fåes ved å komplementere det tilsvarende positive tallet og addere 1. Alle positive tall har mest signifikant bit (MSB) lik 0, og alle negative har MSB lik 1. F.eks. -3 i fire bits system blir:
 $1100 + 0001 = 1101$. Med toers komplement kan vi bruke vanlig binær addisjon for både positive og negative tall.

9.5. Digital til analog (D/A) omsetter

D/A omsetteren er en krets som gjør om et digitalt tall til et analogt signal proporsjonalt med tallet. Det analoge signalet er en viss brøkdel av en referanseverdi. Hvis referanseverdien kan variere avhengig av et analogt signal, får vi et utgangssignal som er produktet av et analogt signal og et digitalt tall (multipliserende D/A). Oppløsningen til en D/A omsetter er den minste verdien vi kan endre utgangssignalet med, og tilsvarer verdien av det minst signifikante bit (LSB).

To varianter av en 4-bits D/A omsetter er vist i *fig. 9.3* og *fig. 9.4*. Den første kretsen krever mange forskjellige motstandsverdier på inngangen, mens den siste bare bruker to verdier. Begge kretsene bruker den inverterende inngangen på en operasjonsforsterker som summeringspunkt.

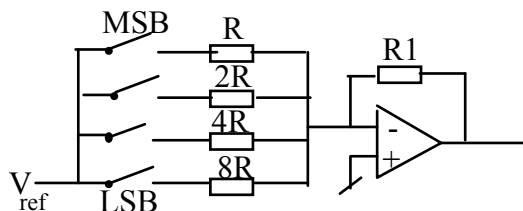


Fig. 9.3 D/A omsetter med binært motstands nettverk.

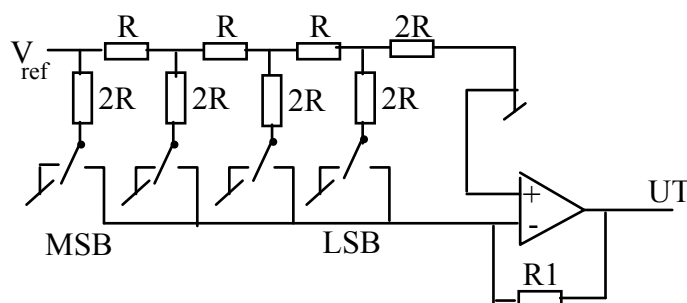


Fig. 9.4 D/A omsetter med R-2R stige nettverk.

9.5.1. Feil i D/A omsettere

For D/A omsettere spesifiseres ofte:

Maksimal differensiell feil som betyr største avvik i trinnhøyde fra den ideelle verdien som er lik 1 LSB.

Monotonitet : Hvis den differensielle feilen er større enn 1 LSB vil ikke utgangsverdien alltid øke når inngangsverdien øker. Dette kan være uheldig.

10. A/D omsetter

A/D omsettere kan deles i noen hovedtyper etter prinsippet for omsettingen:

- 1) Rampe-type (counter-ramp).
- 2) Suksessive approksimasjoner. Mest vanlig kommersielt.
- 3) Integrasjonstype. Langsom, DVMT.
- 4) Parallell type (flash). Hurtigst.
- 5) Indirekte omsetter (spenning (strøm) til frekvens/teller.) Langsom, integrerende.
- 6) Sigma-delta type

10.1. Rampe-type

Prinsippet for denne typen er vist nedenfor. Etter nullstilling av telleren og sampling av inngangsspenningen er utgangen på komparatoren høy og porten åpen for klokkepulser. Pulsene telles i telleren og spenningen fra D/A omsetteren øker inntil den blir større (maks. 1 LSB) enn inngangsspenningen. Da stenges porten og utlesning kan finne sted.

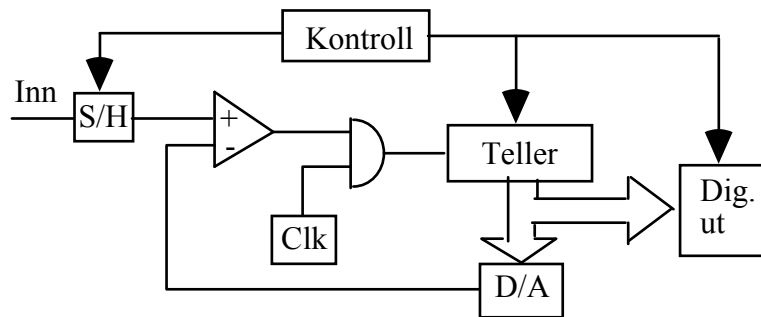


Fig. 9.5 Rampe-type omsetter

En variant av denne omsetteren er en *tracking* -omsetter, eller *servo* -omsetter. Telleren er da en opp/ned teller. Telleren teller opp eller ned slik at spenningen fra D/A omsetteren hele tiden prøver å følge inngangsspenningen. Hvis inngangsspenningen forandres i store sprang, kan D/A spenningen ofte ikke følge med raskt nok.

En annen variant er *Wilkinson* -typen som brukes mye i forbindelse med pulshøyde analyse. En kondensator lades til topp-spenningen for pulsen (peak-detector) og lades så ut lineært med en konstant strøm. Tiden for utladning måles ved å telle klokkepulser. Fordelen med denne typen er at den differensielle lineariteten er meget god.

10.2. Suksessive approksimasjoner

Prinsippet for denne typen er vist på *fig. 9.6*. Virkemåten svarer til bruk av en skålvekt med et sett med binære lodd. Først testes mest signifikante bit, deretter neste osv. inntil alle bitene er testet. Uansett hvor i området inngangssignalet ligger, er omsetningstiden lik antall bit multiplisert med tiden den bruker til å teste ett bit pluss et lite tillegg for utlesning, nulling etc. Slike omsettere kan derfor være meget raske.

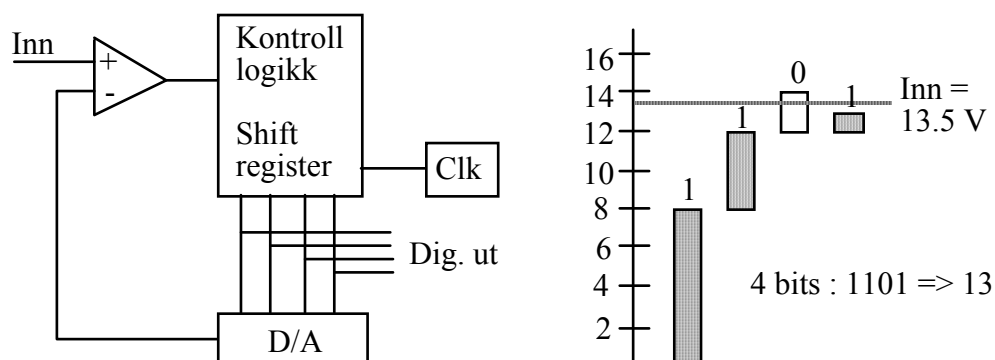


Fig. 9.6 Prinsippet for en 4 bits suksessiv approksimasjons-omsetter.

10.3. Integrerende omsettere

Denne omsetteren kjennetegnes av bra nøyaktighet og god støyundertrykkelse, men er relativt langsom. Prinsippet for omsetteren er vist i *fig. 9.7*.

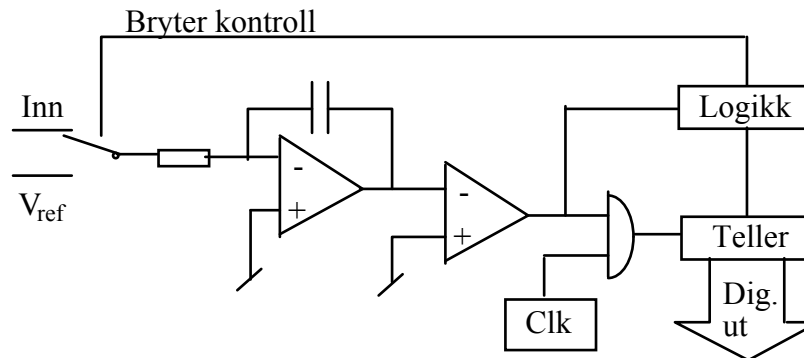


Fig. 9.7 Prinsippet for integrerende omsetting.

Omsettingen er delt i to perioder (dual slope converter). Integratoren er først koblet til inngangssignalet i en konstant tid T_0 og kondensatoren lades til en spenning som er proporsjonal med inngangssignalet. I neste periode kobles inngangen til en referansespenning med motsatt fortegn, og kondensatoren utlades med en konstant strøm. Utgangsspenningen på integratoren avtar derfor med en konstant hastighet. Tiden T det tar til den er null, er proporsjonal med inngangssignalet. I praksis kan det f.eks. være gjort slik at telleren først settes til et negativt tall $-N_0$. I den første perioden teller telleverket pulser inntil det når null. Da kobles integratoren til referanseskilden og telleverket teller pulser inntil utgangsspenningen på integratoren er blitt null. Det binære tallet i telleren er den digitale representasjon av inngangsspenningen.

Omsetters nøyaktighet er vesentlig bestemt ved referansespenningen, og integratorens linearitet og nullpunktsfeil. Fordi signalet integreres, blir tilfeldig støy undertrykket. Hvis støyen er periodisk, blir den helt eliminert hvis integrasjonen går over en hel, eller et multiplum av en hel periode. For å dempe 50Hz støy er det derfor vanlig å bruke 20ms (eller et multiplum av 20ms) som integrasjonstid.

Feil i integratoren og komparatoren kan reduseres ved å bruke en ekstra periode til korreksjon (quad-slope principle).

10.4. Parallell omsetter (flash converter)

I denne omsetteren kobles inngangssignalet til N komparatorer i parallell som vist på *fig. 9.8*. Komparatorene sammenligner inngangsspenningen med en brøkdeler av referansespenningen. Komparatorene er koblet til logikk som f.eks. kan gi resultatet som vanlig binær kode. Omsettingen er meget rask, men f.eks. en 8-bits omsetter krever 255 komparatorer.

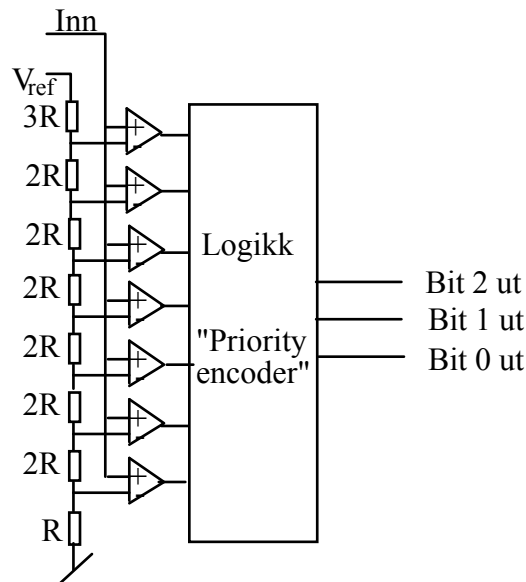


Fig. 9.8 Parallell omsetter (3 bits)

10.5. Spenning (strøm) til frekvens omsetter

Det finnes mange integrerte kretser som gjør om spenning eller strøm til pulser med en frekvens som varierer lineært med inngangssignalet. Frekvensen kan måles ved å telle pulser i et visst tidsintervall, og gir da et digitalt mål for det analoge signalet. Dette er ofte en enkel, men meget fin metode fordi signalet i form av pulser er lite følsomt for støy og derfor kan overføres over store avstander. Det kan også lett kobles til en datamaskin med optokobler for å unngå jordingsproblem. Ulempen er at det er en relativt langsom metode med tidsintervall i millisekund-området.

10.6. Sigma-delta type

Denne ADC-typen er mye anvendt til digitalisering i forbrukerelektronikk. Den passer godt som del av et digitalt system, siden de analoge delene av ADCen er redusert til det minimale. Analog elektronikk er her en ett-bits omformer, som gir en sekvens av enkeltbits med frekvens langt høyere enn sample raten. Problemer med Nyquist frekvens og 'aliasing' blir derved mindre enn med andre metoder.

Digitale kretser (gjerne programmerbare) behandler deretter bitsekvensen, og kan lage både lite nøyaktige samples og svært nøyaktige samples ved å anvende korte eller lange bitsekvenser til å generere verdien for den enkelte sample. Forholdet mellom enkeltbit frekvens og sample frekvens kalles 'oversampling'.

11. Datainnsamling

11.1. LabVIEW

Programmeringsystemet LabVIEW med sitt grafiske programmeringsspråk G er et programmeringsverktøy som er spesielt vel egnet for data innsamling, bearbeiding og presentasjon. Laboratoriedelen som inngår i FYS3230 anvender ferdige LabVIEW programmer som illustrerer digital måleteknikk. Både LabVIEW og programmene ligger klare til bruk på laboratoriemaskinene.

Hensikten med bruk av LabVIEW i dette kurset er ikke å lage LabVIEW-spesialister av kursdeltagerne, men å la dem få et lite innblikk i hvordan moderne grafiske brukergrensesnitt kan anvendes. Og husk : det er ikke finurlighetene i LabVIEW som er pensum i måleteknikkurset, LabVIEW er bare et verktøy som illustrerer endel måletekniske metoder som utgjør deler av pensum.

11.2. Serie- og parallell I/O

Datamaskinen kan kommunisere med omverden ved hjelp av serie- eller parallellporter. På mange personlige datamaskiner er det minst en serie-port (RS-232) og en parallellport av typen *Centronics* . Denne parallellporten er en 'begrenset' parallellport. En generell parallell port krever et eget kort som settes inn på en ledig plass i maskinen.

11.2.1. Serie-port

Ved serie-transmisjon blir ett og ett bit overført. Det krever bare en signallinje og jord. Overføringen blir imidlertid langsom da det tar minimum åtte transmisjoner pr byte. Overføringen skjer som regel asynkront, dvs. det er ingen overføring av klokkepulser, men sender og mottager har blitt enige om overføringshastighet (baud rate). Videre må antall bit pr ord med start bit og stopp bit være bestemt, og om det skal være paritetsbit. Hvis det f.eks. er bestemt *like paritet* , vil det si at hvis det opprinnelig er et ulike antall bit, blir det føyd til et paritetsbit slik at det totale antall bit alltid er et like tall. Mottageren tester på antall bit og kan reagere hvis det ikke stemmer.

Ved asynkron overføring har altså sender og mottager hver sin klokke. For å synkronisere disse brukes start og stop bit. *Fig. 11.1* viser overføringen av et 7-bit ord.

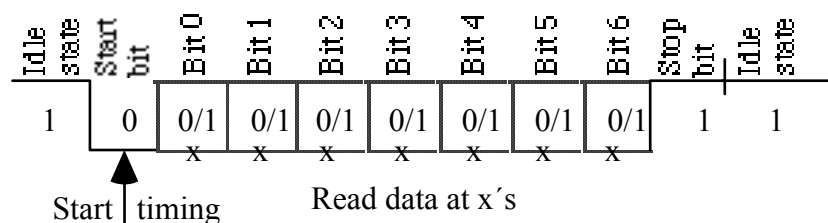


Fig. 11.1 Bit mønster ved serieoverføring av et 7 bit ord uten paritet og med et stop bit.

Datalinjen er opprinnelig i hviletilstand. Her er det vist som logisk høy.

1. Senderen starter med overgang til logisk lav for en klokke periode, start bit.
2. Senderen sender 7 bit i rekkefølge, LSB først. Hver bit varer en klokke periode.

3. Mottageren registrerer overgangen høy til lav for start bit. For å være sikker venter den en halv klokke periode og sjekker om linjen fortsatt er lav. Hvis den er det, venter den en klokke periode og starter lesing av 7 bit.
4. Hvis det er overføring med stopbit, sjekker mottageren at linjen er høy på siste bit, og signaliserer feil hvis det ikke er tilfelle.

Fordi mottageren leser nivåene midt i bitene, behøver ikke de to klokkene ha nøyaktig samme frekvens, avvik opp til 3% kan tolereres. For å kunne lese midt i bitene, er master klokkeperioden 16, 32 eller 64 ganger kortere en bitperioden. En bit vil derfor være 16, 32 eller 64 master perioder, og for å komme til midten av et bit, telles halvparten så mange perioder.

RS-232 serie overføring bruker ikke TTL-nivåer, men -3 til -20 volt som logisk sann, og +3 til +20 volt som logisk falsk (negativ logikk). Overføringen kan være *simplex* (bare en vei) eller *full duplex* (begge veier på hver sine linjer). Serieporten på en PC er typisk en DB-25 hun-kontakt, jfr. tabellen nedenfor. Det er ofte problemer med sammenkobling av to enheter via RS-232. Det var opprinnelig en standard for å knytte data terminal utstyr (data terminal equipment, DTE) til data kommunikasjons utstyr (data communication equipment, DCE). Når en DTE kobles til en DCE skal det bare være å forbinde korresponderende pinner på kontaktene. Hvis to like enheter skal forbindes, må derimot ledningene fra pinne 2 og 3 krysses fordi en DTE sender på pinne 2 og mottar på pinne 3. Det kan også være problemer med å finne ut hvordan *handshaking* -signalene skal forbindes.

Pin nr.	Retning	Navn	
1		gnd	
2	to DCE	TxD	transmitted data
3	to DTE	RxD	received data
4	to DCE	RTS	request-to-send
5	to DTE	CTS	clear-to-send
6	to DTE	DSR	data set ready
7		gnd	logic ground
8	to DTE	DCD	carrier detected
20	to DCE	DTR	data terminal ready

De viktigste signalene for RS-232 overføring.

Det kan defineres opp til fire serieporter (COM1, COM2....).

11.2.2. Parallellport

Mange PC-er har en parallellport av typen *Centronics* . Denne er primært beregnet for tilkobling av printer, og er typisk en DB-25 hun-kontakt som vist i tabellen nedenfor . Kontakten har 8 *datalinjer* ut. I tillegg er det en del *handshakinglinjer* hvorav fem kan brukes som datalinjer inn.

	Pin nr.	Bit
Data	2	0
Ut	3	1
	4	2

	5	3	
	6	4	
	7	5	
	8	6	
	9	7	
Status		0	
Inn		1	
		2	
	15	3	-ERROR(aktiv lav)
	13	4	SLCT
	12	5	PE
	10	6	ACK (aktiv lav)
	11	7	BUSY
Kontroll	1	0	STROBE
Pin 18 til 25 jord			

Centronics (printer) signaler

12. Frekvensanalyse

12.1. Fourierrekker

Dette emne gjennomgås på en langt mer stringent måte i kurset FYS3220. Det er tatt med her som en introduksjon til kommende kapitler.

Ved Fourier analyse (frekvensanalyse) kan vi bestemme hvilke frekvenser som et periodisk signal inneholder, og amplitudene til disse frekvenskomponentene. Ut fra dette kan vi ofte finne ut hvordan f.eks. signalet skal behandles for å bedre forholdet mellom ønsket signal og forstyrrende støy.

Et periodisk signal med periode T er et signal som har egenskapen

$$f(t) = f(t + T) = f(t + 2T) \dots$$

Et slikt signal kan skrives som en sum av cosinus- og sinus-ledd

$$f(t) = a_0/2 + \sum_{n=1}^{\infty} [a_n \cos(n\omega_1 t) + b_n \sin(n\omega_1 t)],$$

hvor $a_0 = 2/T \int_{-T/2}^{T/2} f(t) dt$ dvs. to ganger den midlere verdi av $f(t)$ over perioden T og

$$a_n = 2/T \int_{-T/2}^{T/2} f(t) \cos(n\omega_1 t) dt \quad \text{og} \dots b_n = 2/T \int_{-T/2}^{T/2} f(t) \sin(n\omega_1 t) dt.$$

Her er $\omega_1 = 2\pi/T$ grunnfrekvensen og $2\omega_1, 3\omega_1, 4\omega_1, \dots$ osv. er de høyere harmoniske frekvensene. Koeffisientene a_n og b_n er amplitudene svarende til frekvensen $n\omega_1$. (Jfr. Bentley side 59).

Ved hjelp av Eulers formler for cosinus og sinus kan vi skrive fourierrekken på kompleks form (Jfr. Kuo: Network analysis and synthesis side 55)

Hvis vi definerer

$$\beta_n = (a_n - jb_n)/2, \quad \beta_{-n} = (a_n + jb_n)/2, \quad \beta_0 = a_0/2,$$

får vi

$$f(t) = \sum_{n=-\infty}^{\infty} \beta_n e^{jn\omega_1 t}$$

hvor

$$\beta_n = 1/T \int_0^T f(t) e^{-jn\omega_1 t} dt$$

Her er β_n i alminnelighet kompleks. Amplitude-spektret er gitt ved modulverdiene, og fase-spektret ved

$$\Psi_n = \arctan\left(\frac{\text{Im}(\beta_n)}{\text{Re}(\beta_n)}\right)$$

For å analysere et signal på en digital regnemaskin må vi regne med diskrete verdier for $f(t)$. Vi sampler signalet med en frekvens N/T (dvs. bestemmer signalverdien i N ekvidistante punkter over en periode T) og får en sekvens av verdier

$$y(i) = f\left(\frac{T}{N}i\right) \quad i = 0, 1, 2, \dots, N-1$$

Det samlede signalet kan bare representere frekvenser opp til Nyquist-frekvensen, som er lik halve samplings-frekvensen. Fourier rekken for diskrete signaler (samlede signaler) blir derfor en endelig sum av ledd (jfr. Bentley side 91) :

$$y(i) = f\left(\frac{T}{N}i\right) = a_0/2 + \sum_{n=1}^{N/2-1} \left[a_n \cos\left(\frac{2\pi n}{N}i\right) + b_n \sin\left(\frac{2\pi n}{N}i\right) \right]$$

med fourierkoeffisientene (amplitudene) $a_0 = 2/N \sum_{i=0}^{N-1} f\left(\frac{T}{N}i\right)$,

$$a_n = 2/N \sum_{i=0}^{N-1} f\left(\frac{T}{N}i\right) \cos\left(\frac{2\pi n}{N}i\right) , \quad b_n = 2/N \sum_{i=0}^{N-1} f\left(\frac{T}{N}i\right) \sin\left(\frac{2\pi n}{N}i\right)$$

hvor $n = 1, 2, \dots, N/2 - 1$ (frekvens komponenter)

$i = 0, 1, 2, \dots, N - 1$ (sample nummer).

Den komplekse formen for Fourier rekken for diskrete signaler blir

$$y(i) = f\left(\frac{T}{N}i\right) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X(n) e^{j2\pi n i / N}$$

hvor samplene har nummer $i = 0, 1, \dots, N - 1$.

De komplekse fourierkoeffisientene $X(n)$ (amplitudene) er gitt ved

$$X_n = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} f\left(\frac{T}{N}i\right) e^{-j2\pi n i / N}$$

Koeffisientene er symmetriske omkring $n = N/2$ som svarer til Nyquist-frekvensen f_c . (Vi antar at N er et like tall.)

Altså :

$n = 0$ svarer til frekvens 0 (DC-verdi)
 $n = 1$ svarer til grunnfrekvens = $1/T$ hvor T er perioden for signalet som er samlet
 $n = N/2$ svarer til f_c .

og $N/2 + 1 < n < N$ svarer til de negative frekvensene, dvs. $-f_c < f < 0$.

Faktoren $1/\sqrt{N}$ i formlene er en skaleringsfaktor. Noen forfattere sløyfer den i den første formelen og får da $1/N$ i den siste formelen.

Sammenhengen mellom de komplekse koeffisientene $X(n)$ og koeffisientene a_n og b_n er gitt ved

$$X(n) = \frac{\sqrt{N}}{2} (a_n - j b_n)$$

Det finnes mange algoritmer for å beregne fourierkoeffisienter. Hvis antall sampler (N) er to opphøyet i et helt tall, kan spesielle algoritmer anvendes som bærer fellesnavnet FFT (Fast Fourier Transform). Disse er spesielt raske, og trenger utføring av $N \log_2(N)$ operasjoner. Hvis antallet sampler ikke er en potens av 2, trengs N^2 operasjoner, dette kan reduseres til $3N \log_2(N)$ ved å anvende "neste" 2er potens.

12.2. Effektspektrum

Hvis signalet representerer et spenningsignal, er øyeblikksverdien for effekten proporsjonal med kvadratet av amplitudene. Målt over en motstand på 1Ω er den midlere verdien for effekten svarende til n -te komponent

$$w_n = \frac{1}{T} \int_0^T [a_n \cos(n\omega_1 t)^2 dt] + \frac{1}{T} \int_0^T [b_n \sin(n\omega_1 t)^2 dt] = (a_n^2 + b_n^2)/2$$

Den totale effekten for de første n komponentene kalles kummulativ effekt

$$W_n = w_0 + w_1 + \dots + w_n$$

For grensetilfellet $T \rightarrow \infty$ og $\omega_1 \rightarrow 0$ blir W_n en kontinuerlig funksjon av ω . Dette er den kumulative effekt-funksjon definert som

$$W(\omega) = \lim_{\omega_1 \rightarrow 0} W_n$$

Den deriverte av denne funksjonen er spektral (effekt) tetthet

$$\phi(\omega) = \frac{dW}{d\omega}$$

med enhet *watt s/rad* eller *watt/Hz*.

Intern støy i elektriske kretser har ofte en tilnærmet uniform spektral tetthet, dvs. alle frekvenser er like sannsynlige. Dette kalles *hvit støy* og er altså kjennetegnet ved

$$\phi(\omega) = A$$

hvor A er en konstant. I mange tilfelle er støyen uniformt fordelt opp til en grensefrekvens, og lik null for høyere frekvenser. Dette kalles *båndbegrenset hvit støy*.

13. Signal og støy i målesystem

13.1. Statistisk beskrivelse av tilfeldige (random) signaler

Vi skiller mellom to typer signaler:

- 1) Deterministiske signaler: Verdien til signalet ved et senere tidspunkt kan beregnes.
- 2) Tilfeldig (random) signaler: Verdien ved et senere tidspunkt kan ikke beregnes, men ut fra observasjon over en viss tid kan vi angi statistiske verdier for hvordan signalet vil være. De statistiske størrelsene vi kan bruke er
 - 1) Middelerdi og standard avvik.
 - 2) Sannsynlighets-tetthetsfordeling.
 - 3) Spektral effekt tetthetsfordeling (power spectral density).
 - 4) Autokorrelasjonsfunksjon.

13.1.1. Middelerdi og standard avvik

Middelerdien for et kontinuerlig signal er
$$\bar{y} = \frac{1}{T} \int_0^T y(t) dt$$

og for et samplet signal
$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

hvor T er tidsintervallet vi midler over, og N er antall sampler.

Standard avvik σ er gitt ved
$$\sigma^2 = \frac{1}{T} \int_0^T (y(t) - \bar{y})^2 dt$$

og for et samplet signal
$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 .$$

Hvis spesielt $\bar{y} = 0$, så er standard avvik σ lik r.m.s. (root-mean-square) verdien.

13.1.2. Sannsynlighets tetthetsfordeling

Dette er et mål for sannsynligheten for at signalet vil ha en verdi mellom to bestemte grenser. Hvis de N samlede verdiene fordeles etter stigende verdi i m intervaller, er sannsynligheten P_j for å finne en verdi i intervallet j lik $P_j = n_j/N$ hvor n_j er antall sampler i intervallet j . For grensetilfellet uendelig mange sampler og uendelig små intervaller, får vi en kontinuerlig fordeling. Sannsynlighets tetthetsfordelingen integrert mellom to grenser gir sannsynligheten for å finne en verdi mellom disse grensene. Amplitudefordelingen for tilfeldig støy er som regel godt beskrevet ved en Gauss-fordeling.

13.1.3. Spektral effekt tetthetsfordeling

Hvis vi registrerer et tilfeldig (random) signal i flere observasjonsperioder, vil formen på signalet ikke være den samme i hver observasjonsperiode. Beregner vi midlere effekt for signalet i hver periode, vil imidlertid denne være nokså konstant, dvs. dette er en størrelse som kan brukes til å kvantifisere et tilfeldig signal. Vi kan approksimere et tilfeldig signal med et periodisk signal som svarer til det vi registrerte i den første observasjonsperioden, hvis vi bruker en lang observasjonstid og bare bruker resultatet til å beregne frekvensfordelingen av midlere effekt.

13.1.4. Autokorrelasjon

Autokorrelasjonsfunksjonen for en kontinuerlig funksjon i intervallet 0 til T er definert som

$$R_{yy}(\beta) = \lim_{T \rightarrow \infty} \int_0^T y(t) y(t - \beta) dt$$

hvor β er tidsforsinkelsen. Funksjonen har maksimal verdi for $\beta = 0$.

For et sinus-signal $y(t) = b \sin(\omega_1 t + \phi)$, får vi etter litt regning at

$$R_{yy}(\beta) = \frac{b^2}{2} \cos \omega_1 \beta$$

Det gjelder generelt at autokorrelasjonsfunksjonen for et periodisk signal har samme periode som det opprinnelige signalet. Faseinformasjon blir imidlertid borte.

Hvis vi sampler et signal og sample-intervall er ΔT , kan vi beregne N korrelasjonskoeffisienter

$$R_{yy}(m\Delta T) = \frac{1}{N} \sum_{i=1}^N y_i y_{i-m}$$

hvor y_i er sampleverdien ved tiden $i\Delta T$, og y_{i-m} er sampleverdien m sample-intervaller tidligere. Ved beregning av summen skal vi alltid ha N ledd. F.eks. har vi for $m=2$ at

$$\begin{aligned} R_{yy}(2) &= (y_1 y_{-1} + y_2 y_0 + \dots + y_N y_{N-2}) / N \\ &= (y_1 y_{N-1} + y_2 y_N + \dots + y_N y_{N-2}) / N \end{aligned}$$

dvs. vi betrakter samplesekvensen som en hel periode som gjentar seg.

Det kan vises at fourieranalyse av autokorrelasjonsfunksjonen gir effekt tetthetsspektret $\phi(\omega)$. Tilsvarende får vi autokorrelasjonsfunksjonen ved syntese av effekt spektret. Dette kan sannsynliggjøres ved å se på et periodisk signal som er en sum av tre harmoniske

$$y(t) = b_1 \sin(\omega_1 t) + b_2 \sin(2\omega_1 t) + b_3 \sin(3\omega_1 t)$$

Autokorrelasjonsfunksjonen blir

$$R_{yy}(\beta) = \frac{b_1^2}{2} \cos(\omega_1 \beta) + \frac{b_2^2}{2} \cos(2\omega_1 \beta) + \frac{b_3^2}{2} \cos(3\omega_1 \beta)$$

Effektsspektret til $y(t)$ blir tre linjer med høyde $b_1^2/2$, $b_2^2/2$ og $b_3^2/2$ ved frekvensene ω_1 , $2\omega_1$ og $3\omega_1$. Dette er det samme som vi får ved fourieranalyse av $R_{yy}(\beta)$.

For tilfeldige (random) signaler, hvit støy, har vi at

$$\begin{aligned} R_{yy}(\beta) &= \int_0^\infty \phi(\omega) \cos(\omega \beta) d\omega \\ \phi(\omega) &= \frac{2}{\pi} \int_0^\infty R_{yy}(\beta) \cos(\omega \beta) d\beta \end{aligned}$$

For båndbegrenset hvit støy, $\phi(\omega)$ konstant opp til ω_c får vi derfor

$$R_{yy}(\beta) = \int_0^{\infty} A \cos(\omega\beta) d\omega = A \omega_c \frac{\sin(\omega_c \beta)}{\omega_c \beta}$$

Korrelasjonsfunksjonen $R_{yy}(\beta)$ er derfor lik $A\omega_c$ for $\beta = 0$ og bredden av den sentrale delen er $2\pi/\omega_c$, jfr. fig. 13.1. En høy verdi for ω_c gir derfor en smal korrelasjonsfunksjon, og omvendt.

Videre har vi en viktig sammenheng ved tidsforsinkelse $\beta = 0$:

$$R_{yy}(\beta) = \lim_{T \rightarrow \infty} \int_0^T y^2 dt = \int_0^{\infty} \phi(\omega) d\omega$$

som viser at

$$R_{yy}(\beta) = y_{RMS}^2 = W_{TOT}$$

hvor W_{TOT} er den totale effekten som signalet produserer i en motstand på 1Ω .

En viktig anvendelse av autokorrelasjon er å bruke det til å undersøke om det er noen periodisk variasjon i et signal eller en samplesekvens. Formen av signalet kan ikke vises, men perioden vil komme tydeligere frem hvis signalet er blandet med mye støy. Hvis signalet f.eks. er et sinus-signal blandet med båndbegrenset hvit støy, vil korrelasjonsfunksjonen bli

$$R_{yy}(\beta) = \frac{b^2}{2} \cos(\omega_1 \beta) + A \frac{\sin(\omega_c \beta)}{\beta}$$

Det siste leddet som skyldes støyen, vil gå mot null, mens det første leddet vil vedvare slik at perioden for signalet kommer frem. For et sinus-formet signal ser vi at også amplituden kan bestemmes.

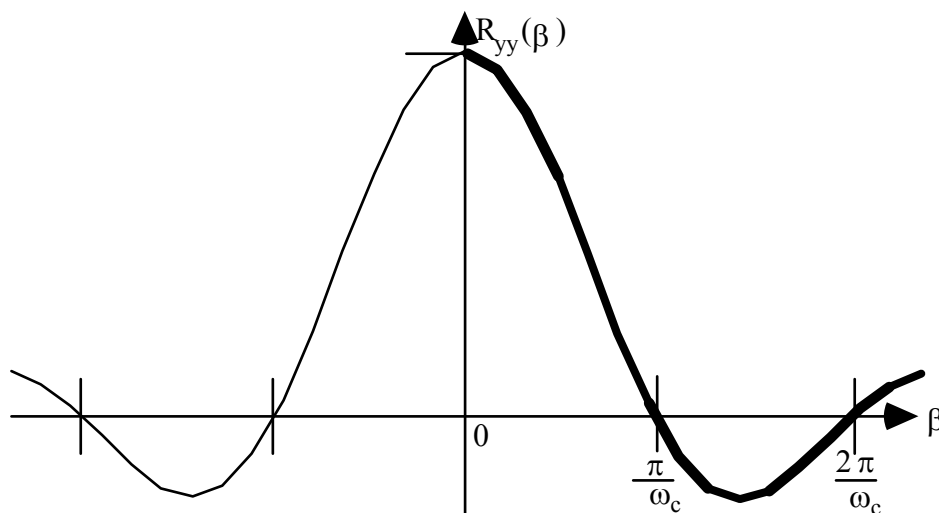


Fig. 13.1 Korrelasjonsfunksjon for båndbegrenset hvit støy.

13.2. Virkning av støy på en målekrets

En Thevenin spenningskilde E_{Th} med indre motstand Z_{Th} er koblet i serie med en støykilde V_{SM} til en lastmotstand Z_L som vist i *Fig. 13.2*. Overføringsledningene har en motstand $R_c/2$. Strømmen gjennom lastmotstanden er

$$i = \frac{E_{Th} + V_{SM}}{Z_{Th} + R_c + Z_L}$$

og spenningen over lasten blir

$$V_L = \frac{Z_L}{Z_{Th} + R_c + Z_L} (E_{Th} + V_{SM}).$$

Normalt velges $Z_L \gg R_c + Z_{Th}$, som gir

$$V_L \approx E_{Th} + V_{SM}$$

Hele støyspenningen kommer derfor over lasten og kan ikke skilles fra signalet.

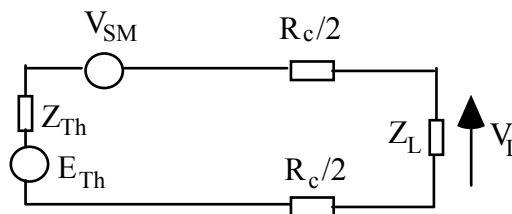


Fig. 13.2 Spenningsoverføring, signal og støy i serie

Signal til støy forholdet defineres som $S/N = 20 \log \frac{E_{Th}}{V_{SM}}$

For et strømoverføringssystem, *fig. 13.3*, får vi

$$i_N (Z_N \parallel (R_c + Z_L)) = i (R_c + Z_L)$$

dvs.

$$i = \frac{Z_N}{Z_N + R_c + Z_L} i_N$$

for strømmen gjennom lasten på grunn av signalkilden når $V_{SM} = 0$. Støykilden gir en strøm

$$i_{SM} = \frac{V_{SM}}{Z_N + R_c + Z_L}$$

Dette gir for spenningen over lasten (superposisjonsprinsippet)

$$V_L = i Z_L + i_{SM} Z_L = i_N Z_L \frac{Z_N}{Z_N + R_c + Z_L} + \frac{V_{SM}}{Z_N + R_c + Z_L} Z_L$$

Normalt vil $Z_N \gg R_c + Z_L$ slik at

$$V_L \approx i_N Z_L + \frac{Z_L}{Z_N} V_{SM} \approx i_N Z_L.$$

Et system med strømoverføring er derfor meget mindre følsomt for seriestøy enn et system med spenningsoverføring.

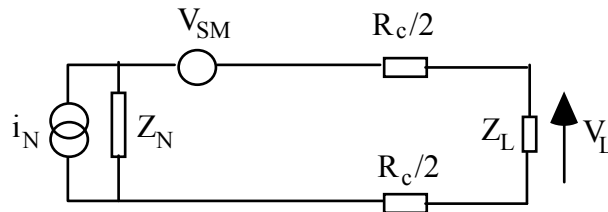


Fig. 13.3 Strømoverføring, støyspenning i serie

Hvis støykilden er koblet i liketakt (common mode) slik som i Fig. 13.4 vil begge sider av lasten bli påvirket like meget. Ideelt kan derfor signalet renses for støy ved å ta differansen mellom de to signalene (differanseforsterker).

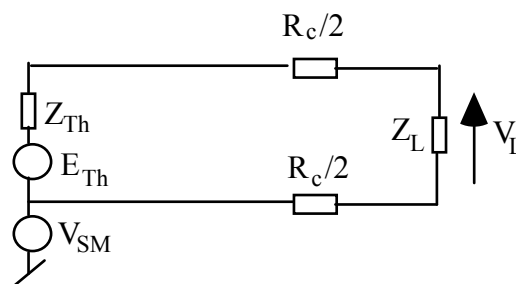


Fig. 13.4 Spenningsoverføring, liketaktkobling av støykilde (common mode interference)

13.3. Støykilder og koblingsmåter

13.3.1. Interne støykilder

Tilfeldige (random) temperaturinduserte bevegelser av elektroner og ladningsbærere i motstander og halvledere gir opphav til tilfeldige spenninger som kalles termisk eller Johnson støy. Den spektrale tetthetsfordeling for denne støyen er uniform (hvit støy) og proporsjonal med den absolutte temperatur θ . Den spektrale tetthetsfordelingen er gitt ved

$$\phi = 4Rk\theta \text{ watt / Hz}$$

hvor R er motstanden og $k = 1.4 \times 10^{-23} \text{ J/K}$ er Boltzmanns konstant. Den totale effekten i frekvensområdet f_1 til f_2 er

$$W = \int_{f_1}^{f_2} 4Rk\theta = 4Rk\theta(f_2 - f_1) \text{ watt}$$

eller fra avsnittet om autokorrelasjon

$$V_{RMS} = \sqrt{4Rk\theta(f_2 - f_1)} \text{ volt}$$

En lignende type støy er "shot noise" som forekommer i transistorer. Den skyldes tilfeldige fluktasjoner i diffusjonshastigheten til ladningsbærere gjennom en halvlederovergang.

13.3.2. Eksterne støykilder

Den mest vanlig eksterne støykilde er 50 Hz støy fra lysnettet. Støypulser fra inn- og utkobling av større belastninger på nettet (motorer osv.) kan også gi forstyrrelser. Lysstoff- rør (100Hz), elektrisk sveising og RF-sendere kan også være plagsomme.

13.3.3. Koblingsmåter til eksterne støykilder

Støy kan bli indusert i en leder ved induktiv kobling til magnetisk felt og ved kapasitiv kobling til elektrisk felt.

Induktiv kobling

Hvis målekretsen er i nærheten av ledere som fører store vekselstrømmer, (til f.eks. motorer) kan det induseres støy på grunn av det magnetiske feltet ved induktiv kobling. Den induserte støyen vil svare til en spenningskilde i serie med signalkilden og størrelsen er gitt ved

$$V_{SM} = M \frac{di}{dt}$$

hvor M er gjensidig induktanskoefisient som avhenger av geometrien.

Kapasitiv kobling

Støy fra f.eks. lysnett kan induseres i en målekrets ved kapasitiv (elektrostatisk) kobling. Signalledningene kan ha en viss kapasitans i forhold til lysnett-ledningene, og hvis det ikke er perfekt balanse blir det indusert støy. I *fig. 13.5* er det vist et eksempel med kobling til lysnett.

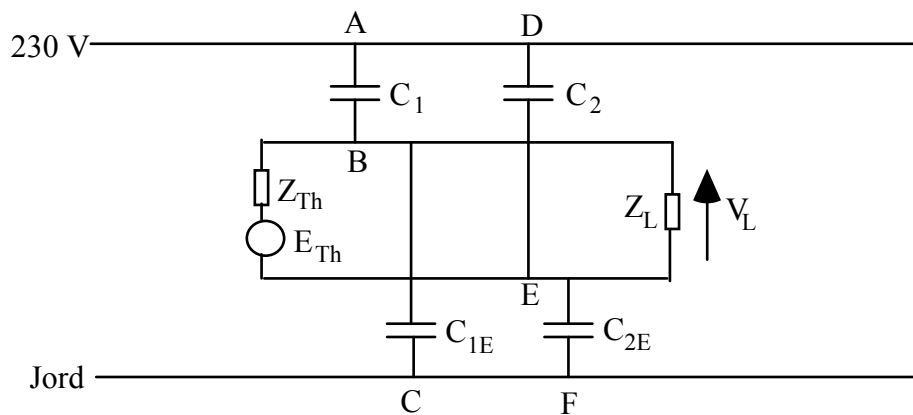


Fig. 13.5 Kapasitiv (elektrostatisk) kobling til lysnett.

Når vi ser bort fra signalspenningen E_{Th} , er spenningsene i punktene B og E

$$V_B = 230 \frac{1/(j\omega C_{1E})}{1/(j\omega C_{1E}) + 1/(j\omega C_1)} = 230 \frac{C_1}{C_1 + C_{1E}}$$

$$V_E = 230 \frac{1/(j\omega C_{2E})}{1/(j\omega C_{2E}) + 1/(j\omega C_2)} = 230 \frac{C_2}{C_2 + C_{2E}}$$

Dette gir en liketaktsspenning (common mode) $V_{CM} = V_E$ og en seriespenning

$$V_{SM} = V_B - V_E = 230 \left\{ \frac{C_1}{C_1 + C_{1E}} - \frac{C_2}{C_2 + C_{2E}} \right\}$$

Det vil derfor bli indusert en støyspenning i serie med signalspenningen hvis det ikke er perfekt balanse mellom koblingskapasitansene.

Jording og jordingsløyfer

Utstyr skal kobles til fysisk jord for å gjøre det berøringssikkert. Denne type jording kalles sikkerhetsjord. Hvis nettspenningen ved en feil blir koblet til utstyrets chassis, vil nettspenningen bli kortsluttet til jord og sikringer skal gå.

Signalmessig jording av systemer gjøres for å redusere støyproblemer. Ved feilaktig jording kan imidlertid forholdene bli betydelig verre. Signaljord skal virke som et felles potensial som hele systemet refereres til. Det kan være aktuelt å ha tre forskjellige jordkretser:

- **Signaljord** Forbinder enheter med lave signalnivåer og stor følsomhet for jordstøy.
- **Kraftjord** Jording for motorer o.l. store strømbrukere.
- **Chassisjord** Sikkerhetsjord for berøringssikkerhet.

For å redusere problemer med at jord ikke ligger på samme potensial overalt, bør signaljord fra forskjellige enheter kobles sammen til ett felles jordpunkt, slik at de i det minste går til et felles potensial. Det er viktig å ha oversikt over jordingssystemet for å unngå jordingsløyfer som kan gi problemer. *Fig. 13.6* viser to enheter som er jordet i et felles punkt, men i tillegg er de forbundet med skjermen i en kabel. Sløyfen mellom enhetene og jord får et stort areal, og det kan induseres merkbare strømmer i jordledningen, vanligvis 50 Hz.

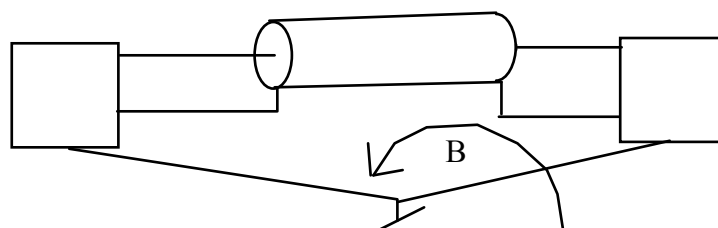


Fig. 13.6 Eksempel på jordsløyfe

Selv om jordforbindelsen i den ene enden av skjermen sløyfes, vil det være en viss kapazitiv impedans til jord, slik at det kan bli en sløyfe.

Et eksempel med flere jordingspunkter er illustrert i *fig. 13.7*

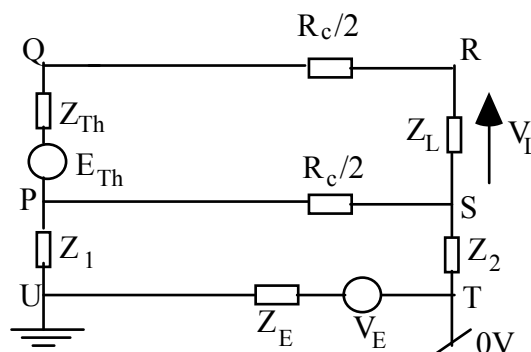


Fig. 13.7 Krets med to jordingspunkter

I praksis vil det ofte være slik at jordingspunkter som ligger et stykke fra hverandre, ikke har nøyaktig samme potensial. I *fig. 13.7* er dette illustrert med en spenningskilde V_E i serie med en motstand Z_E mellom to jordingspunkter. Hvis signalkilden og mottakeren er helt isolert fra jord, er det ingen problemer. Ofte vil det imidlertid være en viss lekkasjemotstand til jord, Z_1 og Z_2 på *fig. 13.7*, og det kan bli induert støy. Hvis $Z_L \gg R_c + Z_{Th}$, er strømmen i kretsen PQRS neglisjerbar og $V_L \approx E_{Th}$. I kretsen UPST går det en strøm

$$i_E = \frac{V_E}{Z_E + Z_1 + R_c/2 + Z_2}$$

og potensialene i P,Q,R og S blir

$$\begin{aligned} V_P &= V_E - i_E(Z_E + Z_1) \\ V_Q = V_R &= V_P + E_{Th} \\ V_S &= i_E Z_2 \end{aligned}$$

Det er derfor en liketakts (common mode) spenning

$$V_{CM} = V_S = V_E \frac{Z_2}{Z_E + Z_1 + R_c/2 + Z_2}$$

Den serieinduserte spenningen finner vi ved å se på spenningen over lasten Z_L

$$\begin{aligned} V_L = V_R - V_S &= V_P + E_{Th} - i_E Z_2 \\ &= V_E - i_E(Z_E + Z_1 + Z_2) + E_{Th} \\ &= E_{Th} + i_E R_c/2 \end{aligned}$$

dvs. det er en serieindusert støyspenning

$$V_{SM} = V_E \frac{R_c/2}{Z_E + Z_1 + R_c/2 + Z_2}$$

Ideelt bør både Z_1 og Z_2 være så store som mulig. Ofte er det vanskelig å unngå at Z_1 blir liten (sensor i forbindelse med jord), men da må Z_2 være stor, dvs. bare ett jordpunkt.

Eksempel : Med $Z_1 = 10 \Omega$, $Z_E = 1 \Omega$, $R_c/2 = 10 \Omega$, $V_E = 1V$ og $Z_2 = 10 M\Omega$

blir

$$V_{SM} = \frac{10}{1 + 10 + 10 + 10^6} \approx 10 \mu V$$

Hvis også $Z_2 \approx 0$, blir $V_{SM} = 0.48V$.

13.4. Metoder for reduksjon av virkningen av støy

13.4.1. Fysisk separasjon

Gjensidig induktans og koblingskapasitans er invers proporsjonale med avstanden. Avstand til støykilder må derfor være størst mulig.

13.4.2. Elektromagnetisk skjerming

En strømførende leder omgir seg med et magnetfelt som er proporsjonalt med strømstyrken. Magnetfeltet kan indusere strøm i en leder hvis den danner en lukket sløyfe. Strømmen som induseres er proporsjonal med sløyfens areal og magnetfeltets styrke loddrett på dette arealet. Den enkleste måten å redusere virkningen av denne støykilden er å tvinne sammen ledningsparene (twisted pair). Arealet blir da minst mulig samtidig med at de induserte spenninger vil parvis oppheve hverandre.

13.4.3. Elektrostatisk skjerming

Den ideelle løsningen er å ha hele målekretsen i en metallboks (skjerm) uten forbindelse til boksen, og boksen jordet i ett punkt. Dette er som regel ikke mulig. Ofte vil det være lekkasjemotstander mellom målekrets og skjerm, og mellom skjerm og jord som vist på *fig. 13.8*.

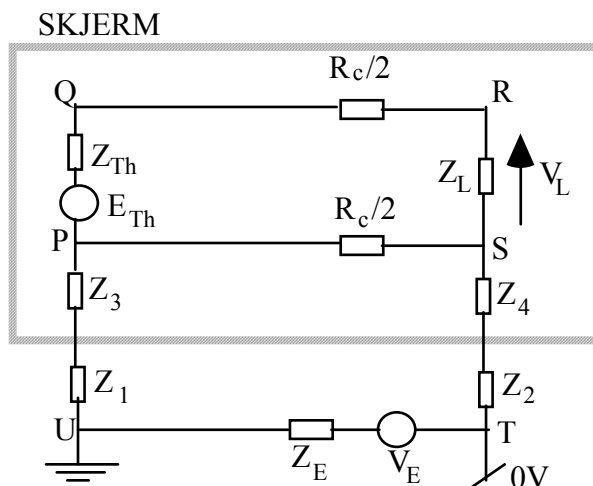


Fig. 13.8 Målekrets i en skjermet boks

På grunn av V_E får vi en liketakts støyspenning

$$V_{CM} = V_E \frac{Z_2 + Z_4}{Z_E + Z_1 + Z_3 + R_c/2 + Z_2 + Z_4}$$

og en serie støyspenning

$$V_{SM} = V_E \frac{R_c/2}{Z_E + Z_1 + Z_3 + R_c/2 + Z_2 + Z_4}$$

Dårlige kombinasjoner :

- Hvis Z_1 og Z_3 og Z_2 og Z_4 er små, har vi en jordsløyfe.
- Hvis Z_1 og Z_2 er små, har vi en jordsløyfe.
- Hvis Z_3 og Z_4 er små, har vi en intern sløyfe.
- Hvis Z_1 og Z_2 er store, har vi ingen lav impedanse vei for kapazitiv koblet strøm.

Brukbare kombinasjoner:

- Enten Z_1 eller Z_2 må være liten, da har vi ett jordpunkt.
- Z_3 og/eller Z_4 må være store, da har vi ingen intern sløyfe.

13.4.4. Bruk av differanseforsterker

En ideell differanseforsterker har et utgangssignal som er (fig. 13.9a)

$$v = \frac{R_1}{R_2}(v_2 - v_1)$$

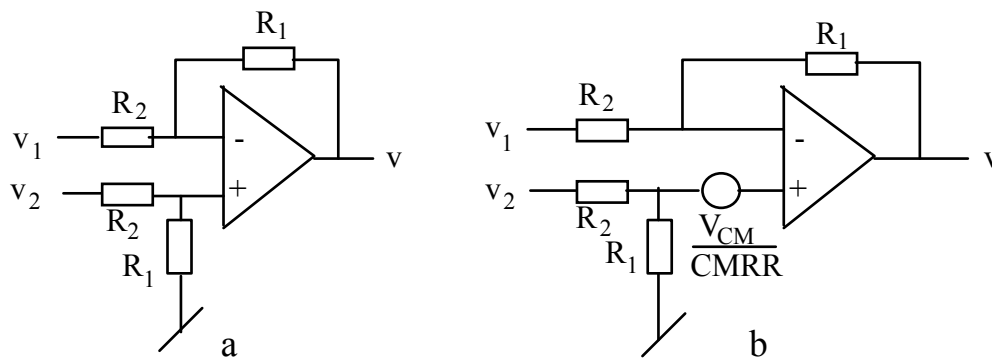


Fig. 13.9 Differanseforsterker

En virkelig operasjonsforsterker har en differanseforsterkning A_{OL} og en fellessignal (common mode) forsterkning A_{CM} slik at utgangssignalet uten tilbakekobling (open loop) er

$$v_{OL} = A_{OL}(v_+ - v_-) + A_{CM}v_{CM}$$

A_{CM} spesifiseres ofte ved *common mode rejection ratio*

$$CMRR = A_{OL} / A_{CM}$$

hvor $CMRR \approx 10^4 - 10^5$. Utgangssignalet for en differanseforsterker blir derfor (jfr. fig. 13.9b)

$$v = \frac{R_1}{R_2}(v_2 - v_1) + \frac{R_1 + R_2}{R_2} \frac{V_{CM}}{CMRR}$$

Hvis f.eks. $E_{Th} = 1mV$, $V_{CM} = 1V$, $CMRR = 100dB$, $R_2 = 1k\Omega$, $R_1 = 1M\Omega$ blir utgangssignalet

$$v = 1000 \times 10^{-3} + 1000 \times 10^{-5} = 1.0V + 0.01V$$

dvs. fellessignalet (liketaktssignalet) er 1%.

13.4.5. Filtrering

Hvis signal og støy har forskjellig frekvensområde, kan filtrering forbedre signal/støy forholdet. Det kan være aktuelt med lavpass, båndpass eller høypass avhengig av frekvensområdet som skal fjernes. Hvis det spesielt er 50Hz støy, kan det brukes stoppfilter for denne frekvensen. Etter at signalet er digitalisert kan det være aktuelt med digital filtrering.

13.4.6. Modulasjon

En effektiv metode for å bedre signal/støy forhold er ofte å la en sensor som gir et DC-signal eller bare lave frekvenser, amplitudemodulere et høyfrekvenssignal, f.eks. 5kHz . Dette overføres til en mottaker hvor det kan være et båndfilter som dekker det frekvensområdet omkring 5kHz som inneholder signalfrekvensene, men lite av støyfrekvenser, før signalet blir demodulert. Hvis det umodulerte høyfrekvens signalet er

$$V_c \cos \omega_c t$$

og dette moduleres med et signal

$$V_m \cos \omega_m t$$

blir det modulerte signalet

$$v_{AM}(t) = V_c (1 + m \cos \omega_m t) \cos \omega_c t$$

hvor $m = V_m / V_c < 1$ kalles modulasjonsindeks. Med

$$\cos a \cos b = (\cos(a + b) + \cos(a - b)) / 2$$

får vi det amplitudemodulerte signalet

$$v_{AM}(t) = V_c \cos \omega_c t + \frac{m}{2} V_c \cos(\omega_c + \omega_m) t + \frac{m}{2} V_c \cos(\omega_c - \omega_m) t$$

dvs. signalet består av tre frekvenser, og amplitudene til signalene med sum- og differanse-frekvensene avhenger av modulasjonsindeks. Modulasjonssignalet kan vi få tilbake ved likeretting og lavpass filtrering.

For DC-signaler (eller signaler med meget lave frekvenser) kan det brukes en fasefølsom likeretter, som med passende valg av komponenter, virker som et meget smalt båndfilter. (Jfr. lab.oppg. 6.)

Frekvensmodulasjon gir også en overføring av signaler som er lite følsom for støy. For å frekvensmodulere en bærebølge kan vi bruke en spenningsstyrt oscillator VCO . Frekvensen f til oscillatoren er bestemt ved størrelsen på inngangsspenningen v til oscillatoren slik at

$$f = f_c + kv$$

hvor f_c er frekvensen for $v = 0$ og k er følsomheten Hz/V for oscillatoren.

Utgangssignalet fra VCO blir

$$V_1 \sin 2\pi(f_c + kv)t$$

hvor V_1 er amplituden til VCO -signalet. Hvis modulasjonssignalet er et enkelt sinussignal

$$v = V_m \sin 2\pi f_m t$$

får vi det frekvensmodulerte signalet

$$v_{FM}(t) = V_1 \sin 2\pi(f_c + D \sin 2\pi f_m t)t$$

hvor $D = kV_m$ er maksimalt avvik i frekvensen fra f_c . Motakeren er gjerne en faselåst sløyfe som følger overføringsfrekvensen, og gir en utgangsspenning som varierer på samme måte som signalet som modulerte sendefrekvensen. (Jfr. lab. oppg. 8.)

13.4.7. Midling

Signalmidling kan være meget effektivt for å få frem et signal som er blandet med mye støy. Signalet må repetere seg og samples med nøyaktig samme fase hver gang, dvs. sampling startes hver gang på samme tidspunkt i forhold til f.eks. starten på signalet. Korresponderende sampler fra hver periode midles. Hvis det er p måleserier (perioder) og N sampler i hver, blir midlere verdi av sample i :

$$y_i^{AV} = (y_{i1} + y_{i2} + \dots + y_{ip})/p$$

Korresponderende verdier av signalkomponenten fra hver periode er tilnærmet like så den midlere verdi har en tilsvarende størrelse. Korresponderende verdier av støykomponenten er helt forskjellige, noen positive og noen negative. Disse vil derfor reduseres ved midling.

Hvis vi f.eks. har p signaler y_1 til y_p som er Gauss fordelte med standard avvik σ_1 til σ_p , får vi

$$y_{AV} = (y_1 + y_2 + \dots + y_p)/p$$

med standard avvik

$$\sigma_{AV} = \frac{1}{p} \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_p^2}$$

Hvis alle st.avvik er like store, får vi

$$\sigma_{AV} = \sigma/\sqrt{p}$$

dvs. r.m.s. verdien er redusert med faktoren \sqrt{p} .

13.4.8. Autokorrelasjon

Autokorrelasjon kan brukes til å undersøke om det er noen periodisk variasjon (signal) i en datastrøm (datafil). Selve formen på signalet kan ikke bestemmes. Bare amplitude og periode kan bestemmes fra korrelasjonsfunksjonen.

Autokorrelasjonsfunksjonen signal+støy er summen av korrelasjonsfunksjonen for signal og støy hver for seg, dvs.

$$R_{yy}^{S+N}(\beta) = R_{yy}^S(\beta) + R_{yy}^N(\beta)$$

For et sinusformet signal blandet med båndbegrenset støy får vi derfor

$$R_{yy}^{S+N} = \frac{b^2}{2} \cos \omega_1 \beta + A \frac{\sin \omega_1 \beta}{\beta}$$

Det siste leddet som representerer støyen, vil dø ut, og for store verdier av β kan derfor amplituden b og perioden $2\pi/\omega_1$ for signalet bestemmes.

14. Fasefølsom detektor

Når et målesignal er overlappet mye støy, kan en fasefølsom detektor være et viktig hjelpemiddel til å få frem signalet. Prinsippet for en slik detektor (likeretter) er vist på *fig. 14.1*. Den gjennomgås i lab.oppgave 3.

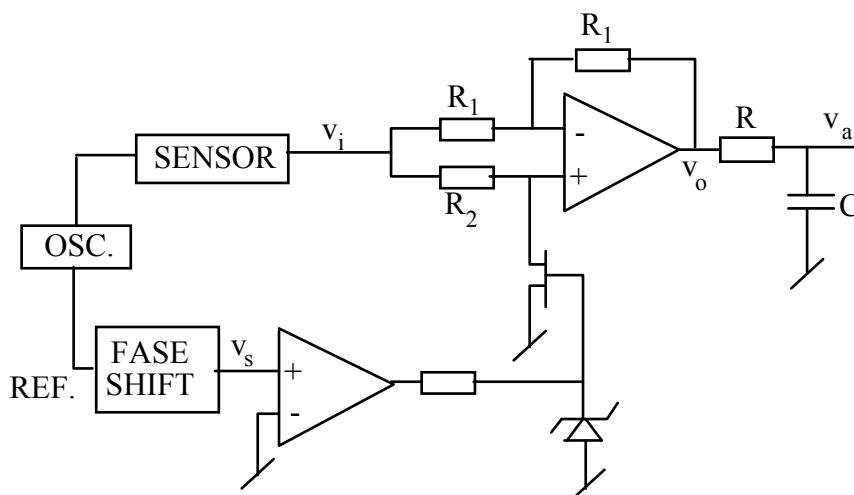


Fig. 14.1 Prinsippet for fasefølsom detektor

Signalet fra en oscillator sendes til et måleobjekt (sensor) som påvirker amplituden. Signalet forsterkes om nødvendig i en smalbandet forsterker før det kommer til detektoren. Oscillatoren er også koblet til en faseskifter slik at vi kan justere faseforskjellen mellom målesignalet og referansesignalet. På utgangen av operasjonsforsterkeren vil vi kunne se signaler svarende til kurvene på *fig. 14.2* avhengig av faseforskjellen mellom de to signalene. Etter lavpassfilteret får vi et likerettet signal

$$v_a = \frac{2}{T} \int_0^{T/2} V_m \sin(\omega t - \phi) dt = \frac{2}{\pi} \cos \phi$$

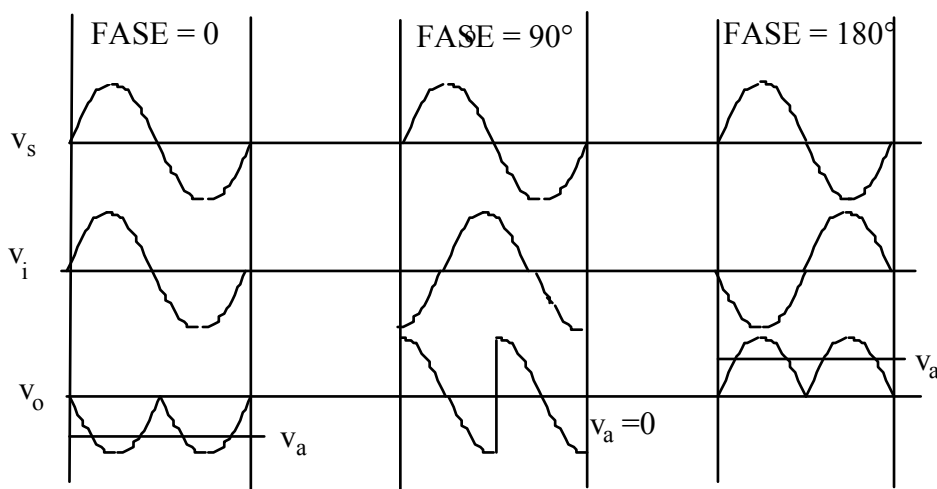


Fig. 14.2 Kurveformer for fasefølsom detektor

Detektoren vil fungere som et båndfilter med båndbredde bestemt av lavpassfilteret (R og C):

$$2\Delta f = \frac{1}{\pi RC}$$

Ved passende valg kan filteret gjøres meget smalt.

Betingelsen for å kunne bruke denne deteksjonsmetoden er at målesignalet varierer med en konstant frekvens, og at denne frekvensen er tilgjengelig som et referansesignal. Hvis signalet egentlig er et DC-signal, kan vi f.eks. ha en bryter på inngangen til signalforsterkeren som kobler forsterkeren vekselvis til jord og til signalet. Det samme signalet som brukes til å styre denne bryteren, brukes så som referansesignal.

En vanlig likeretter vil likerette signal og støy uansett frekvens. En fasefølsom likeretter vil likerette bare den delen av støyen som faller innenfor båndbredden bestemt av lavpass filteret.

Liknende prinsipp for undertrykkelse av støy er brukt i forsterkere som på engelsk kalles "lock-in" forsterkere.

15. Sensorer

Et målesystem består generelt av flere elementer: Sensor (føler, transduser), signalforming, prosessering og presentasjon. Sensoren er det første elementet i kjeden, og inngangsverdien til dette elementet er den sanne verdien til det som skal måles. Sensoren mottar energi fra det som skal måles og omsetter det til en hensiktsmessig form for målesystemet. For elektroniske målesystem vil det si at utgangssignalet kan f.eks. være spenning, strøm, motstand, frekvens, pulsbredde, eller kapasitans.

Det er seks former for energi som kan omsettes til elektrisk signal med en sensor: Mekanisk, termisk, ioniserende stråling, elektromagnetisk, magnetisk og kjemisk. I tabellen er det gitt eksempler på noen representative sensorer som omdanner disse energiformene til elektriske signaler.

Energiform	Representativ sensor
Mekanisk	Propell med pulsgiver
Termisk	Termoelementer
Ioniserende stråling	Ionisasjonskammer
Elektromagnetisk	Antenne
Magnetisk	Hall-effekt sensor
Kjemisk	pH elektrode

Sensorene kan klassifiseres etter anvendelsesområde, eller etter det fysiske prinsippet de baserer seg på. I det følgende vil vi diskutere sensorer ut fra den siste inndelingsmåten.

15.1. Resistive sensorer

Motstandsvariasjon er et av de mest vanlige prinsipp for sensorer. Motstanden i en resistiv sensor kan varieres ved en kontakt som forskyves langs en motstandstråd, ved temperaturforandring, mekanisk påvirkning og belysning av fotosensitivt materiale.

15.1.1. Potensiometer for lineær forskyvning

Sensoren kan i enkleste form være en glidekontakt langs en motstandstråd, eller et multiturns potensiometer koblet med f.eks. tannhjul og tannstand til måleobjektet. Linearitet og repeterbarhet kan gjøres meget god.

15.1.2. Motstandstermometere og termistorer

Motstanden i de fleste metaller varierer tilnærmet lineært med temperaturen i området -100 til $+800$ °C. Generelt kan motstanden R_T som funksjon av temperaturen T (i °C) skrives som en potensrekke

$$R_T = R_0 (1 + \alpha T + \beta T^2 + \gamma T^3 + \dots)$$

hvor R_0 er motstanden ved 0 °C og α , β , γ er temperaturkoeffisienter. Platina er det metallet som generelt har de beste egenskapene (kjemisk stabilt, høyt

smeltepunkt, mest lineært) og som foretrekkes for presisjonsmålinger. Motstandstermometre kan også lages av halvledermaterialer og kalles da termistor. De har en motstand som vanligvis avtar meget sterkt med temperaturen (negativ temp. koeffisient), men kan også lages med positiv temperaturkoeffisient. Motstanden R_θ ved temperaturen θ °K er

$$R_\theta = K e^{\beta/\theta}$$

hvor K og β er konstanter. Formelen kan omskrives til

$$R_\theta = R_1 e^{\beta\left(\frac{1}{\theta} - \frac{1}{\theta_1}\right)}$$

hvor R_1 er motstanden ved en referansetemperatur θ_1 , f.eks. 298 °K.

Motstandsvariasjon med temperaturen kan også brukes til å måle vakuum i området 0.1 til 10 *torr* (*mmHg*). Et filament hvor det går litt strøm vil ha større varmetap i dårlig vakuum enn i godt vakuum. Temperaturen og motstanden vil derfor øke når trykket synker, dvs. det blir bedre vakuum. (Pirani-manometer.)

15.1.3. Strekkklapper (metallegering, halvledere)

Motstanden i en tynn metalltråd vil øke hvis den strekkes fordi tverrsnittet minsker og lengden øker. Den spesifikke motstanden, særlig i halvledere, varierer også litt (piezoresistans). Strekkklapper er gjerne et folie med en tynn motstandstråd lagt i et mønster som gjør den maksimalt følsom for strekk i en retning og lite følsom for påvirkning i andre retninger. Motstandsendringen er gitt som

$$\frac{dR}{R} = G \frac{dl}{l}$$

hvor proporsjonalitetsfaktoren G er omkring 2.0 for metaller og 100 til 200 for halvledere. Halvledermotstander er mer følsomme enn metallmotstander, men også mer ulineære og temperaturavhengige.

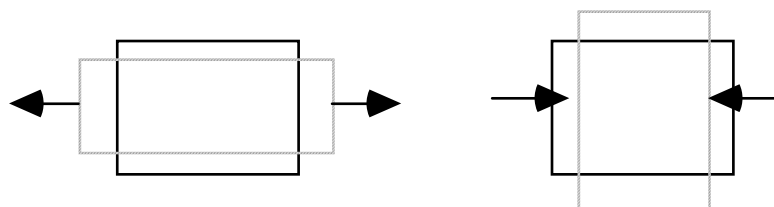


Fig. 15.1 Lineær strekk og sammentrykning

Trykk (stress) defineres som forholdet mellom kraft F og flate A (tverrsnitt). En positiv verdi skal svare til at legemet deformeres slik at det blir lengre (tensile stress), negativ verdi svarer til at det blir sammentrykket (compressive stress). Deformasjonen (strain)

$$e_L = \frac{\Delta l}{l} > 0$$

for lineært strekk. Elastisitetsmodulen er forholdet mellom påvirkning (stress) og deformasjon (strain). For lineært strekk eller sammentrykning kalles den Youngs modul. For skjærkrefter kalles det tilsvarende forholdet skjærmodul. Lineært strekk fører til transversal kompresjon. Poissons forhold

$$\nu = -\frac{e_T}{e_L}$$

er forholdet mellom transversal og lineær elastisitetsmodul. Verdien varierer fra 0.25 til 0.4.

Motstanden i en metalltråd med lengde l , tverrsnitt A og resistivitet ρ er

$$R = \frac{\rho l}{A}$$

Ved logaritmisk derivasjon får vi

$$\frac{\Delta R}{R} = \frac{\Delta l}{l} - \frac{\Delta A}{A} + \frac{\Delta \rho}{\rho}$$

Med $A = wt$ hvor w og t er bredde og tykkelse, får vi

$$\frac{\Delta A}{A} = \frac{\Delta w}{w} + \frac{\Delta t}{t} = 2e_T = -2\nu e_L$$

som gir

$$\frac{\Delta R}{R} = e_L + 2\nu e_L + \frac{\Delta \rho}{\rho} = (1 + 2\nu)e_L + \frac{\Delta \rho}{\rho}$$

Strekklappens følsomhet er definert som

$$G = \frac{\Delta R/R}{\Delta l/l} = \frac{\Delta R/R}{e_L}$$

som gir

$$G = 1 + 2\nu + \frac{1}{e_L} \frac{\Delta \rho}{\rho}$$

For metaller er $\nu \approx 0.3$ og det siste leddet (piezoresistivt ledd) ≈ 0.4 som gir

$$\frac{\Delta R}{R} \approx 2.0e_L = 2.0 \frac{\Delta l}{l}$$

Strekklapper laget av halvledere (p eller n dopet Si) har meget større følsomhet ($G \approx 100$ til 200), men stor temperaturkoeffisient. Ved å bruke flere elementer i en brokobling, kan temperaturfølsomheten reduseres vesentlig.

15.1.4. Fotomotstander

Materialer som cadmiumsulfid (CdS) og cadmiumselen (CdSe) har en motstand som avhenger av lysstyrken som treffer overflaten. Motstanden er stor når det er mørkt og avtar med økende lysstyrke.

15.2. Kapasitive sensorer

En kondensator med to parallelle metallplater har en kapasitans

$$C = \frac{\epsilon_0 \epsilon A}{d}$$

hvor d er avstanden mellom platene, A er arealet av platene som overlapper hverandre, ϵ_0 er permitiviteten i vakuum og ϵ er den relative permitivitet til isolatoren mellom platene. Kapasitansen kan derfor varieres ved å variere avstanden mellom platene. Dette gir en ulineær variasjon.

Hvis platene forskyves i forhold til hverandre, vil A og dermed kapasitansen C variere, dette blir en lineær funksjon av forskyvningen.

Hvis det mellom en del av platene er en isolator med permittivitet ε_1 mens det mellom resten av platene er en permittivitet ε_2 , vil kapasitansen variere lineært med forskyvningen av isolatoren.

Hvis det mellom to plater med avstand $2d$ settes inn en tredje plate, som kan forskyves mot eller fra en av ytterplatene, får vi en differensiell forskyvningssensor. Settes de to elementene i en brokobling, kan vi få en lineær sammenheng mellom kapasitansen og forskyvning i forhold til midtstillingen.

15.3. Induktive sensorer

15.3.1. Forskyvningssensorer

Sammenhengen mellom magnetisk flukstetthet B (tesla) og magnetisk feltstyrke H (A/m) er gitt ved

$$B = \mu H$$

hvor μ er permeabiliteten. For en torus hvor det går en strøm i gjennom n vindinger, er flukstettheten

$$B = \frac{\phi}{A} = \mu \frac{ni}{l}$$

dvs.

$$\phi = \frac{ni}{\frac{l}{\mu A}}$$

Her er A tverrsnittet og l lengden av senterlinjen i torusen. I analogi med motstanden i en leder definerer vi

$$R = \frac{l}{\mu A}$$

som magnetisk motstand, reluktans. Hvis det gjennom en vinding (sløyfe) går en magnetisk fluks ϕ på grunn av en strøm i , er selvinduktansen $L = \phi/i$. For n vindinger får vi derfor

$$L = \frac{n\phi}{i} = \frac{n^2}{\frac{i}{\mu A}} = \frac{n^2}{R}$$

Hvis reluktansen i en magnetiske krets forandres på grunn av f.eks. et varierende luftgap, vil selvinduktansen variere, og dette kan måles. I en magnetisk krets med et luftgap er reluktansen

$$R = R_0 + kd$$

hvor R_0 er reluktansen for jernet, d er luftgapet og k er en proporsjonalitetskonstant. Induktansen kan derfor uttrykkes som

$$L = \frac{L_0}{1 + \alpha d}$$

Induktansen varierer derfor ulineært med luftgapet. Hvis vi bruker en differensiell kobling, kan vi få en lineær variasjon. Et eksempel på en differensiell reluktanssensor som kan måle forskyvning, er vist på *Fig. 15.2*.

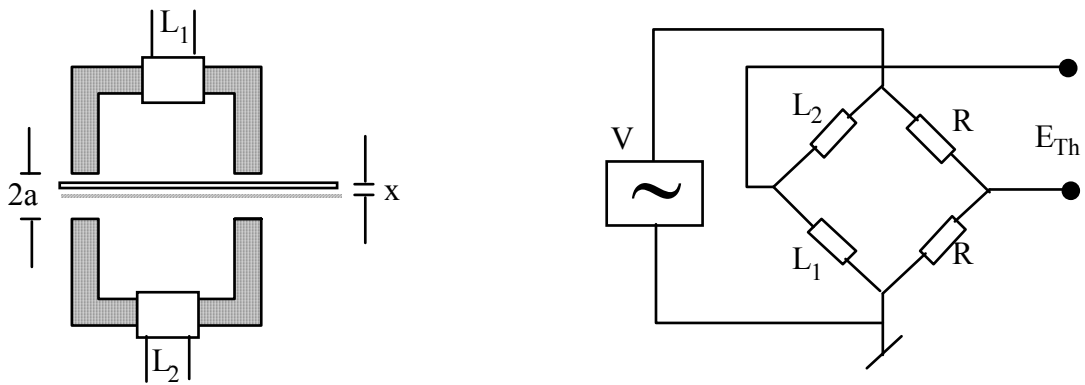


Fig. 15.2 Differensiell reluktanssensor og brokobling.

Induktansene kan skrives som

$$L_1 = \frac{L_0}{1 + \alpha(a - x)} \quad \text{og} \quad L_2 = \frac{L_0}{1 + \alpha(a + x)}$$

hvor x er avviket fra midtstillingen (jfr. fig. 15.2). Signalet fra brokoblingen blir

$$E_{Th} = V \left(\frac{j\omega L_1}{j\omega L_1 + j\omega L_2} - \frac{1}{2} \right) = \frac{V}{2} \frac{\alpha x}{1 + \alpha a}$$

som er en lineær funksjon av avstanden x .

15.3.2. Differensialtransformatorer

Differensialtransformatoren (linear variable differential transformer - LVDT) består av en primær og to like sekundærviklinger viklet på en spoleform. Signalene fra de to sekundærspolene adderes i motfase. Inne i spolen er det en jernkjerne som kan forskyves. Når den står i midtstillingen, er spenningen i de to sekundærviklingene like store, men motsatt rettet, og summen blir lik null. Trekkes kjernen vekk fra midten, blir spenningen i den ene spolen større enn i den andre. Differansen øker tilnærmet proporsjonalt med forskyvningen. Ved passasje av midtpunktet snur sekundærspenningens fase 180 grader i forhold til primærspenningen. Ved å bruke en fasefølsom likeretter kan vi se hvor kjernen er i forhold til midtstillingen. Denne sensoren kan lages for måleområder opp til en del desimeter med god linearitet (1%). Signalet varierer kontinuerlig med posisjonen og oppløsningen er bare begrenset av støy. Hvor raske posisjonsendringer som kan måles, avhenger av komponentene i den fasefølsomme likeretteren.

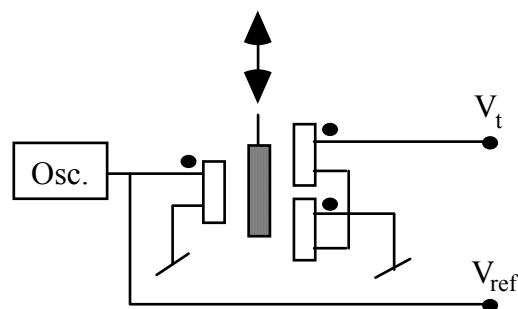


Fig. 15.3 Lineært varierende differensialtransformator.

15.4. Elektromagnetiske sensorer

15.4.1. Hastighetssensorer

Disse sensorene kan brukes til å måle lineær hastighet og rotasjonshastighet. De baserer seg på Faradays lov som sier at hvis fluksen N i en spole varierer, induseres det en spenning

$$E = -\frac{dN}{dt}$$

En sensor som utnytter dette for å måle rotasjonshastighet, kan lages som vist på *fig. 15.4*.

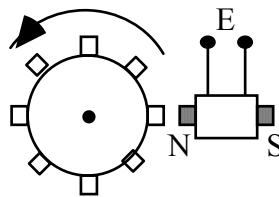


Fig.15.4 Variabel reluktans omdreiningsmåler

Et ferromagnetisk tannhjul roterer i nærheten av en spole som er viklet på en permanent magnet. Reluktansen for den magnetiske kretsen varierer med luftgapet mellom tannhjulet og magneten. Fluksen er omvendt proporsjonal med reluktansen og varierer tilnærmet som

$$N \approx a + b \cos(m\omega t)$$

hvor a og b er konstanter, m er antall tenner på tannhjulet og ω er vinkelhastigheten. Den induserte spenningen blir derfor

$$E = bm\omega \sin(m\omega t)$$

Vinkelhastigheten ω kan derfor bestemmes fra amplituden $bm\omega$, eller fra frekvensen $m\omega$, det siste er som regel best.

15.4.2. SQUID

"**S**ingle **Q**uantum **I**nterference **D**etectors" hører til blandt de mest sensitive av alle detektortyper, og måler magnetfelt-endringer. SQUID sensorer baserer seg på 3 fysiske prinsipper :

- Superledning
- Josephson tunnel effekt
- Kvantisering av magnetisk flux.

For å oppnå superledning vil kjøling trengs i en grad avhengig av materialet sensoren er laget av. En såkalt DC-SQUID kan være utformet som to Josephson barrierer som det skjer "tunnelling" gjennom når det påtrykkes en spenning. "Tunnellingen" utgjør en strøm som lager et magnetfelt. Hvis strømmen holdes konstant ved å variere spenningen, vil nødvendig spenning variere med magnetfeltet som vist på *fig. 15.5* :

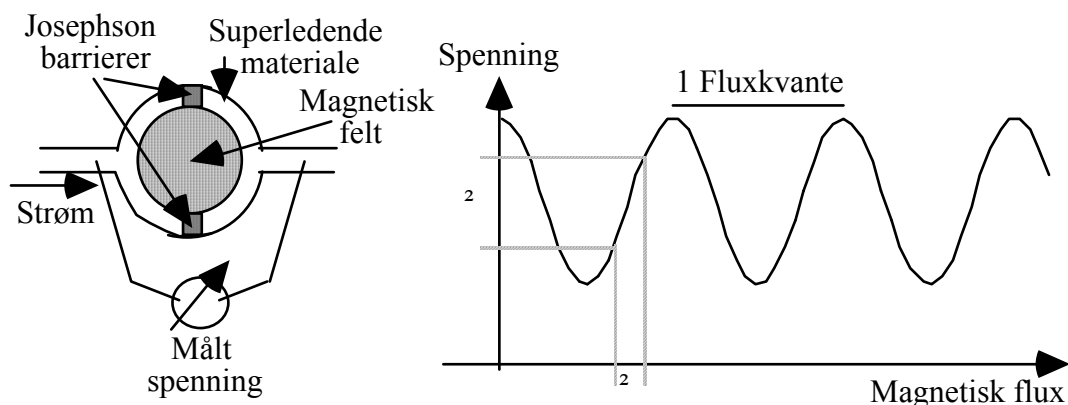


Fig. 15.5 Prinsippskisse og oppførsel for en DC-SQUID

Avstandene mellom spenningstoppene langs x-aksen tilsvarer flux-kvanter, fluxoider, som har en størrelse på ca. $2 \times 10^{-5} \text{ Tm}^2$. Jordmagnetismen er ca. 36×10^6 fluxoider/ mm^2 .

Ved å endre feltstyrken slik at man beveger seg langs den bratte delen av kurven, kan man måle ørsmå flux-endringer, tilsvarende den energimengde som trengs for å løfte ett elektron 0.1 mm i jordas tyngdefelt. En bevegelsesdetektor som kan måle forskyvninger på 10^{-18} meter er også konstruert (en typisk atomkjerne har diameter på 10^{-15} meter).

En SQUID brukes ofte sammen med en magnetisk flux transformator som kan ha større dimensjoner enn SQUIDen selv. Slike transformatorer vil ha en utforming avhengig av applikasjonen, noen av dem er spesialkonstruert for å fange opp flux-ndringer p.g.a. elektriske strømmer i biologisk materiale. Her kan det måles Magneto-Encephalo-Gram (p.g.a. hjernens elektriske aktivitet), Magneto-Cardio-Gram (p.g.a. hjertets elektriske aktivitet) på fostre etc.

15.5. Termoelementer

Seebeck oppdaget i 1821 at hvis to forskjellige metaller A og B sammenføres, oppstår det et kontaktpotensial som avhenger av metallene og temperaturen T °C i kontaktpunktet. Spenningen (termospenningen) kan skrives som

$$E_T^{AB} = a_1 T + a_2 T^2 + a_3 T^3 + \dots$$

hvor konstantene a_1 , a_2 osv. avhenger av metallene. For å måle kontaktpotensialet må vi koble til et høyohmig voltmeter, men dette medfører minst to nye kontaktpotensial. Forutsatt at koblingsledningene til voltmeteret er av samme metall C , og tilkoblingspunktene til instrumentet har samme temperatur (jfr. figuren), er den målte spenningen lik forskjellen i kontaktpotensial, dvs.

$$E_{T_1, T_2}^{AB} = a_1 (T_1 - T_2) + a_2 (T_1^2 - T_2^2) + \dots$$

De to koblingsmåtene på Fig. 15.6 er ekvivalente, og spenningen avhenger kun av de to temperaturene T_1 og T_2 i sammenføyningene av metallene A og B . Det spiller ingen rolle hvordan temperaturen varierer langs med ledningene.

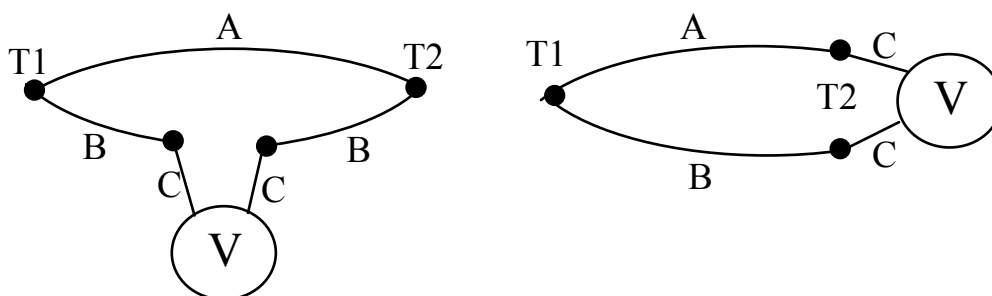


Fig. 15.6 Termopar med voltmeter.

For å bestemme temperaturen $T1$ må referansetemperaturen $T2$ være kjent, vanligvis brukes et isbad, eller en automatisk referanse spenning. Vanlige termopar er jern/konstantan (legering av kobber og nikkel) og kobber/konstantan, den første angir den med mest positiv termospennning. Termospenningen $E_{100,0}$ for jern/konstantan er $5.2mV$ og for kobber/konstantan $4.3mV$. Ved bruk av termopar er det noen lover det kan være nyttig å være klar over. F.eks. kan man bestemme termospenningen for kobber/jern hvis termospenningen for kobber/konstantan og konstantan/jern er kjent (4. lov):

$$E_{T1,T2}^{AB} = E_{T1,T2}^{AC} + E_{T1,T2}^{CB}$$

Den 5. lov om bruk av mellomliggende temperatur $T3$ er også ofte nyttig:

$$E_{T1,T2}^{AB} = E_{T1,T3}^{AB} + E_{T3,T2}^{AB}$$

Hvis $T2 = 0\text{ }^{\circ}\text{C}$, har vi

$$E_{T1,0}^{AB} = E_{T1,T3}^{AB} + E_{T3,0}^{AB}$$

Her kan $T3$ være romtemperatur og $E_{T3,0}^{AB}$ en automatisk kompensasjonsskilde som adderer en spenning svarende til termospenningen mellom romtemperatur og isbad. Hvis romtemperaturen varierer, skal kompensasjonsspenningen variere tilsvarende. Spenningskilden kan f.eks. lages med et motstandstermometer eller en halvleder-temperaturmåler.

15.6. Elastiske sensorer

En kraft som virker på en fjær, gir en forlengelse eller sammentrykning som er tilnærmet proporsjonal med kraften. Forandringen kan gjøres om til et elektrisk signal med f.eks. en resistiv sensor. Dette er prinsippet for elastiske sensorer. De kan også brukes til å måle

$$\text{akselerasjon} = \text{kraft/masse}$$

$$\text{moment} = \text{arm} * \text{kraft}$$

$$\text{trykk} = \text{kraft/areal}$$

$$\text{vinkelakselerasjon} = \text{moment/treghetsmoment}$$

Prinsippet for en sensor som kan måle akselerasjon i en retning, er vist på fig. 15.7. En masse m (seismisk masse) som kan bevege seg på et par skinner, er festet med en fjær og en demper til veggen i akselerometeret.

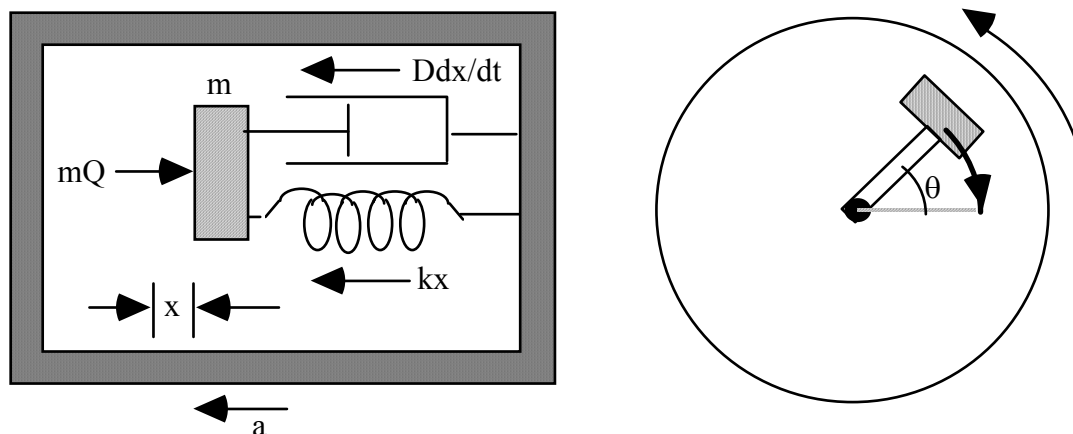


Fig. 15.7 Sensorer for lineærakselerasjon og vinkelakselerasjon.

Hvis akselerometeret plutselig får en akselerasjon a , vil massen m utsettes for en kraft

$$m d^2 x / dt^2 = ma - kx - D dx / dt$$

Dette er analogt med den kraften en person i en bil føler når bilen akselererer. Differensialligningen blir

$$m d^2 x / dt^2 + D dx / dt + kx = ma$$

Massen m får et dempet innsving til en ny posisjon x_1 hvor fjærkraften balanserer ma . Akselerometeret er et 2. ordens system med en (udempet) egenfrekvens og demping.

En tilsvarende sensor for å måle vinkelakselerasjon består i prinsippet av en masse festet på en arm som er kan dreie seg om sentrum i en skive. Hvis skiven får en vinkelakselerasjon, utsettes massen for et dreimoment i motsatt retning som er lik produktet av treghetsmomentet I og vinkelakselerasjonen ϕ . Armen er festet med en spiralfjær, og det er demping av bevegelsen på samme måte som i det lineære akselerometeret. Den tilsvarende differensialligningen blir

$$I d^2 \theta / dt^2 + D d\theta / dt + k\theta = I\phi$$

hvor θ er vinkelen mellom en linje på skiven og den bevegelige armen.

15.7. Piezoelektriske sensorer

Hvis en kraft virker på en krystall, forskyves atomene i krystallgitteret litt. Forskyvningen x er tilnærmet proporsjonal med kraften F . I likevekt er

$$x = \frac{1}{k} F$$

hvor k er stivheten i krystallen. Den er stor, typisk $2 \times 10^9 \text{ N/m}$. Den dynamiske sammenhengen mellom x og F svarer til et 2. ordens system med høy egenfrekvens (10 til 100 kHz) og meget liten demping (typisk 0.01).

Deformasjon av krystallgitteret gir en netto ladning q som er proporsjonal med x

$$q = Kx = \frac{K}{k} F$$

som gir

$$i = \frac{dq}{dt} = \frac{K}{k} \frac{dF}{dt} = K \frac{dx}{dt}$$

dvs. kraften gir direkte et elektrisk utgangssignal. Det er også en invers effekt hvor spenningsvariasjoner får krystallen til å vibrere. Dette brukes i ultralydsendere.

For å måle strømmen eller ladningsvariasjonene, legges det metallelektroder på krystallen slik at den blir en kondensator med kapasitans C_N . Sensoren kan da ekvivaleres med en strømkilde i_N i parallell med kapasitansen C_N . Hvis en slik signalkilde kobles til en spenningsforsterker, blir størrelsen på signalet avhengig av kapasitansen C_C til forbindelseskabelen. Dette er uheldig og kan unngås hvis signalkilden istedet kobles til en ladningsfølsom forsterker som vist på *fig. 15.8*.

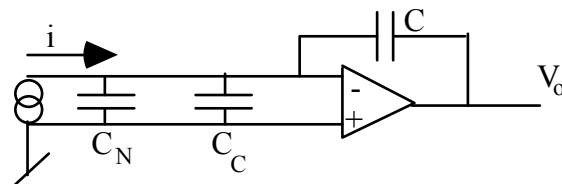


Fig. 15.8 Ekvivalentkrets for en piezoelektrisk sensor med ladningsforsterker

For en ideel operasjonsforsterker får vi

$$i = \frac{dq}{dt} = -C \frac{dV_0}{dt}$$

som gir

$$V_0 = -\frac{q}{C} = -\frac{Kx}{C}$$

Utgangssignalet fra ladningsforsterkeren er altså uavhengig av kabel-kapasitansen. Selv de beste operasjonsforsterkere har alltid en hvilestrøm slik at i praksis må tilbakekoblings-kondensatoren shuntes med en motstand, men denne kan velges meget stor (10 til 100 M Ω) og forandrer responsen lite.

Piezoelektriske sensorer kan f.eks. brukes sammen med seismiske masser for å måle akselerasjon og vibrasjon.

15.8. Piezoresistive sensorer

Dette er halvleder sensorer, Si dopet med n eller p , som gir stor piezoresistiv effekt dvs. motstandsvariasjon ved kraftpåvirkning. De brukes som strekkklapper til å måle deformasjon og krefter, men har stor temperaturfølsomhet.

15.9. Elektrokjemiske sensorer

For å måle konsentrasjon av et bestemt ion i en oppløsning, brukes såkalte ione-selektive elektroder. Når en elektrode settes ned i en oppløsning, får vi en reaksjon mellom ioner i oppløsningen og overflaten på elektroden. Det etableres et

likevektspotensial som avhenger av ionekonsentrasjonen. Hvis det er en selektiv elektrode, får vi et potensial som vesentlig avhenger av konsentrasjonen av et bestemt ion i oppløsningen. For å kunne måle potensialet må vi ha en elektrode til, en referanseelektrode. Denne skal gi et konstant potensial uavhengig av ionekonsentrasjon. Systemet svarer til en spenningskilde i serie med en meget høy motstand, 100 til 1000 M Ω .

Som eksempel på en ioneselektiv elektrode kan nevnes *pH* elektroder for måling av surhetsgrad, dvs. konsentrasjon av hydrogenioner i en løsning. I en sterk syre er det 1 gram H^+ ioner per liter, mens i en sterkt alkalisk løsning er det 10^{-14} gram per liter. I rent vann er det 10^{-7} gram H^+ ioner per liter. På grunn av den store variasjonen, brukes det en logaritmisk skala, og surhetsgraden *pH* defineres som

$$pH = -\log[H^+]$$

Surhetsgraden målt i *pH* er derfor 0 for en sterk syre, 7 for rent vann og 14 for en sterk alkalisk løsning. En såkalt kombinasjonselektrode for måling av *pH* er vist på *fig. 15.9*. Referanseelektroden er bygget inn i et ytre skall hvor sølv-sølvklorid elektroden er i kontakt med en kaliumklorid løsning. Ionisk kontakt med væsken som skal undersøkes skjer gjennom en kjeramisk plugg. Denne virker som en ikke-selektiv membran som gir elektrisk kontakt. Senter-elektroden er *pH* elektroden som også består av en sølv-sølvklorid elektrode i f.eks. saltsyre. Glassmembranen i bunnen slipper bare hydrogenioner igjennom slik at det etablerte potensialet avhenger bare av konsentrasjonen av hydrogenioner.

Følsomheten for *pH* elektroden er 59mV/pH ved 25 °C.

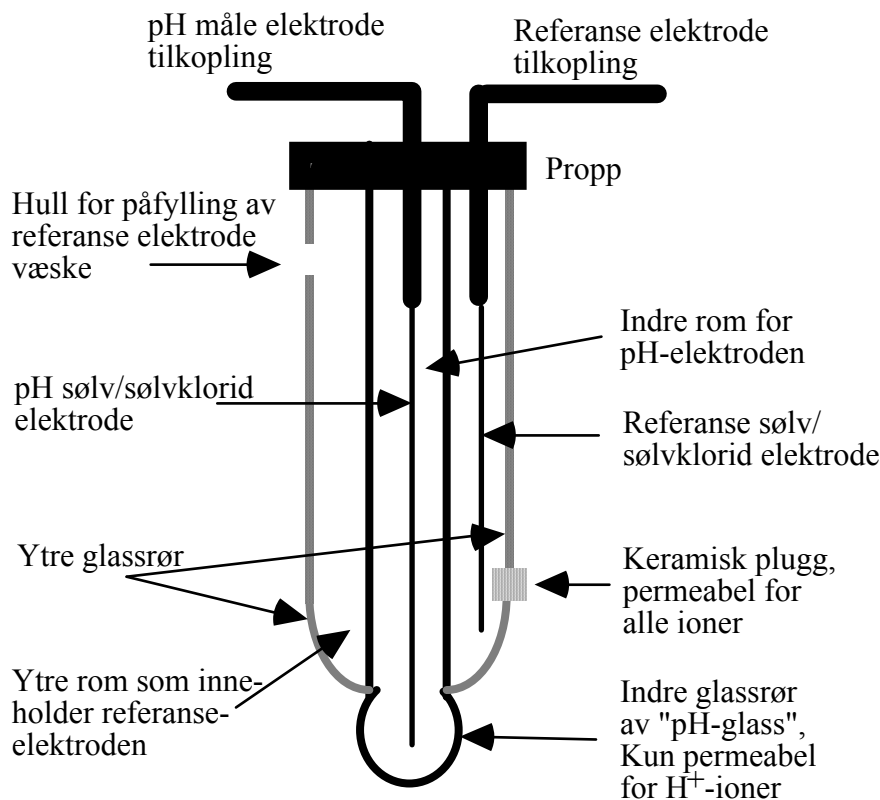


Fig. 15.9 Kombinasjonselektrode for *pH* måling

16. Behandling av sensorsignalet

Signalet fra en sensor må gis en passende form for videre prosessering. Dette vil si at det skal være et spenningsignal eller strømsignal.

16.1. Brokoblinger

Brokoblinger brukes for å få et spenningsignal fra resistive, kapasitive og induktive sensorer. En vanlig brokobling med en spenningskilde V_S er vist i *fig. 16.1*.

Utgangssignalet E_{Th} er

$$E_{Th} = V_S \left\{ \frac{Z_1}{Z_1 + Z_4} - \frac{Z_2}{Z_2 + Z_3} \right\}$$

Den ekvivalente utgangsimpedansen E_{Th} er den vi ser mellom utgangsklemmene når spenningskilden V_S erstattes med sin indre impedans. Utgangsimpedansen blir derfor

$$Z_{Th} = Z_1 \parallel Z_4 + Z_2 \parallel Z_3 = \frac{Z_1 Z_4}{Z_1 + Z_4} + \frac{Z_2 Z_3}{Z_2 + Z_3}$$

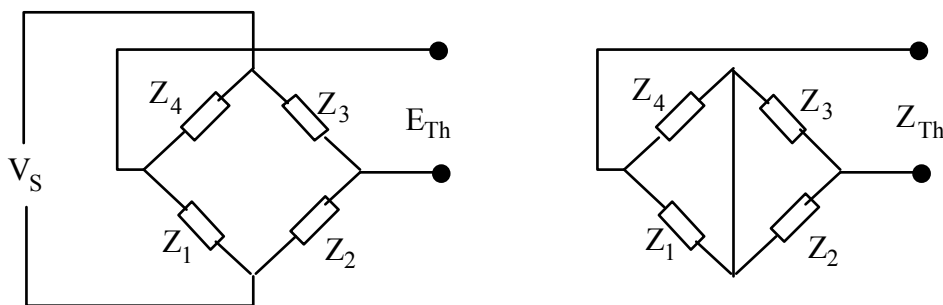


Fig. 16.1 Brokobling og ekvivalent utgangsimpedans

16.1.1. Resistiv brokobling

I en resistiv bro er alle impedansene rent ohmske motstander. Som eksempel kan vi se på en resistiv bro hvor en av motstandene er en resistiv sensor R_I , og R_2 , R_3 og R_4 er konstante motstander. Dette gir

$$E_{Th} = V_S \left\{ \frac{1}{1 + R_4/R_I} - \frac{1}{1 + R_3/R_2} \right\}$$

Hvis motstanden til sensoren varierer fra R_{Imin} til R_{Imax} , vil utgangssignalet variere fra

$$E_{Min} = V_S \left\{ \frac{1}{1 + R_4/R_{Imin}} - \frac{1}{1 + R_3/R_2} \right\} \quad \text{til} \quad E_{Max} = V_S \left\{ \frac{1}{1 + R_4/R_{Imax}} - \frac{1}{1 + R_3/R_2} \right\}$$

Ofte ønsker vi at $E_{Min} = 0$, dvs. balansert bro for R_{Imin} . Da må $\frac{R_4}{R_{Imin}} = \frac{R_3}{R_2}$.

Lineariteten og følsomheten til broen avhenger av valget av forholdet R_3/R_2 og variasjonsområdet for sensormotstanden R_l . Hvis variasjonsområdet er lite (10 - 20 %), er broen ganske lineær selv om vi velger $R_3/R_2 = 1.0$ som gir størst følsomhet. Hvis variasjonsområdet er betydelig større, må motstandsforholdet velges lik 10 til 100 for at utgangssignalet skal variere lineært med sensormotstanden. Fig. 16.2 viser hvordan utgangssignalet varierer med sensormotstanden for forskjellige valg av R_3/R_2 . Utgangssignalet er regnet ut for en balansert bro, dvs.

$$\frac{E_{Th}}{V_S} = \frac{1}{1 + (R_3/R_2)(R_{lmin}/R_l)} - \frac{1}{1 + R_3/R_2}$$

eller

$$v = \frac{1}{1 + r/x} - \frac{1}{1 + r} = \frac{x}{x + r} - \frac{1}{1 + r}$$

hvor

$$v = E_{Th}/V_S \quad r = R_3/R_2 \quad x = R_l/R_{lmin}$$

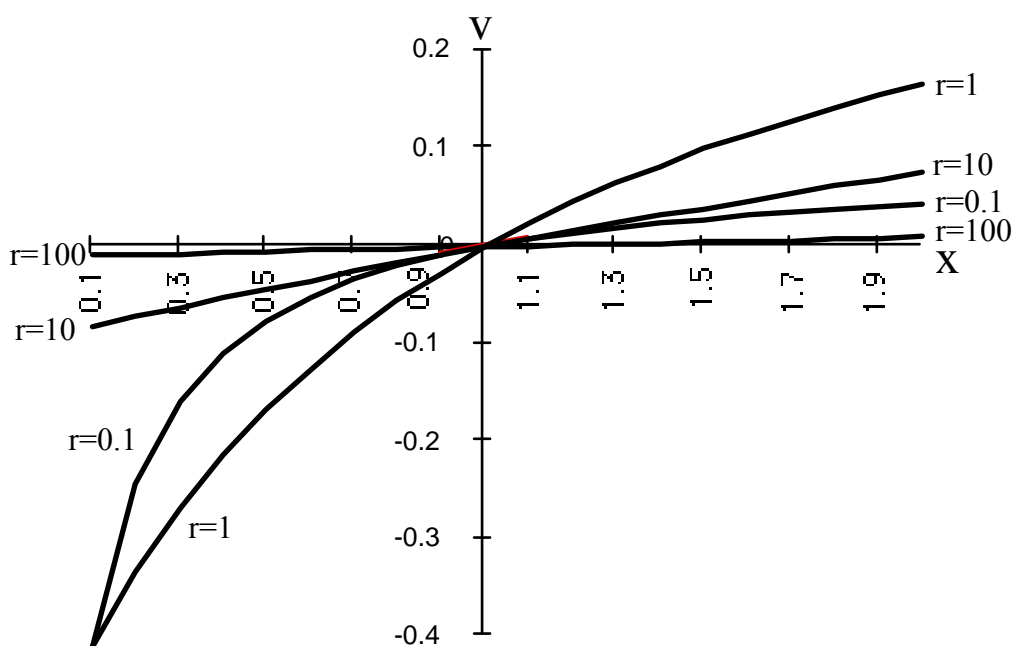


Fig. 16.2 Brosignal $v = E_{Th}/V_S$ som funksjon av $x = R_l/R_{lmin}$ for forskjellige verdier av $r = R_3/R_2$

Fra figuren ser vi at v alltid er null for $x=1$ som svarer til at broen er balansert for $R_l = R_{lmin}$. Vi ser at følsomheten er størst for $r=1$, men hvis R_l varierer meget, må vi velge $r=10$ til 100 for å få god linearitet.

16.1.2. Reaktive broer

En reaktiv bro har vekselstrøms spenningskilde. Som regel er to av elementene i broen reaktive (kapasitanser eller induktanser) og to er resistive. Broer for en kapasitiv og en induktiv differensiell forskyvningssensor er vist i fig. 16.3.

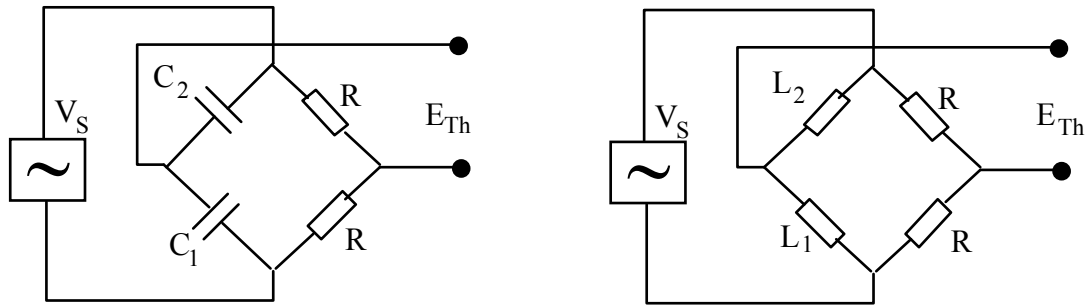


Fig. 16.3 Bro for kapasitiv og induktiv differensiell forskyvningssensor

Begge broene gir et signal som er uavhengig av frekvensen, og som varierer lineært med forskyvningen.

16.2. Forsterkere

Signaler fra mange sensorer som f.eks. termopar og strekkklapper må forsterkes. Det mest vanlige er nå å bygge forsterkere ved hjelp av integrerte operasjonsforsterkere.

16.2.1. Operasjonsforsterkere

I tillegg til stoffet presentert her, må man kjenne til vanlige opamp koplinger og deres anvendelser.

For en ideel operasjonsforsterker gjelder

$$i_- = i_+ = 0 \quad \text{og} \quad v_- = v_+ .$$

Den første ligningen svarer til at inngangsimpedansen er uendelig stor. Den andre ligningen svarer til at forsterkningen (open loop gain) er uendelig stor. Ved hjelp av disse ligningene kan de vanlige formlene for forsterkning med tilbakekobling utledes.

For ikke-ideelle operasjonsforsterkere gjelder begge ligningene med god approksimasjon for likespenning og lave frekvenser. Men en virkelig op.amp. uten tilbakekobling har typisk en forsterkning $A_{OL} = 100 \text{ dB}$ opp til en grensefrekvens lik 10 Hz , og for høyere frekvenser avtar forsterkningen med 20 dB/dekade . En slik op.amp. vil derfor med tilbakekobling ha en båndbredde lik 100 kHz når forsterkningen er 10 ganger. Produktet av forsterkningen og båndbredden, begrepet "Gain-bandwidth product" er altså 1 MHz . Dette gjelder for "små" utgangssignaler ($< 1 \text{ V}$). For større signaler kan "slew rate" begrense båndbredde ytterligere. For å få tilstrekkelig båndbredde i en forsterker må det derfor ofte brukes flere trinn med liten forsterkning i hvert trinn. Operasjonsforsterkere har også litt forsterkning A_{CM} for liketakt (common mode) signaler (felles-signaler), som regel spesifisert ved liketaktsdempningen (common mode rejection ratio, CMRR)

$$CMRR = \frac{A_{OL}}{A_{CM}}$$

CMRR er typisk 80 til 90 dB . Dette gir

$$\begin{aligned} v_{ut} &= A_{OL}(v_+ - v_-) + A_{CM}v_{CM} \\ &= A_{OL} \left[(v_+ - v_-) + \frac{v_{CM}}{CMRR} \right] \end{aligned}$$

I brokoblinger med sensorer som gir svært små signaler, kan liketaktsignalet være et problem. Virkningen av hvilestrøm i noen operasjonsforsterkere kan også være betydelig. I en inverterende forsterker kan den kompenseres ved å ha en motstand mellom pluss-inngang og jord. Motstanden skal svare til parallellkoblingen av de motstandene som hvilestrømmen til minus-inngangen trekkes igjennom.

16.2.2. Instrumenteringsforsterker

En instrumenteringsforsterker brukes gjerne i forbindelse med brokoblinger med resistive sensorer. Den vanligste typen er i prinsippet en sammenkobling av tre operasjonsforsterkere som vist på figuren. Den har to helt symmetriske innganger med meget høy inngangsmotstand, typisk mange hundre $M\Omega$ og fellessignaldempning på 100 dB eller mer. Forsterkningen kan varieres fra 1 til 1000 ved å variere på motstanden R_g på figuren.

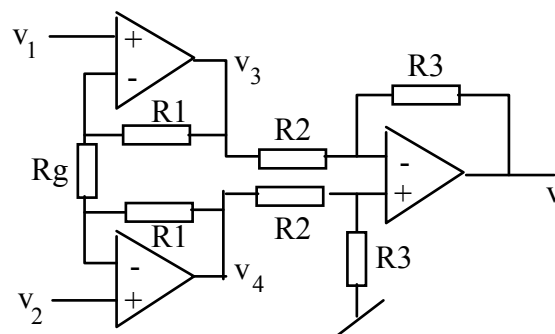


Fig. 16.4 Typisk instrumenteringsforsterker.

Forsterkningen kan regnes ut slik: For første trinn er

$$\frac{v_3 - v_1}{R1} = \frac{v_1 - v_2}{Rg} = \frac{v_2 - v_4}{R1}$$

Herav følger

$$v_3 = v_1 + (v_1 - v_2)R1/Rg \quad \text{og} \quad v_4 = v_2 - (v_1 - v_2)R1/Rg$$

Dette gir differanseforsterkningen $A1$ i 1. trinn

$$A1 = \frac{v_3 - v_4}{v_1 - v_2} = 1 + 2 R1/Rg$$

Forsterkningen $A2$ i 2. trinn er $A2 = -R3/R2$.

Total forsterkning blir $A = A1 \times A2 = -\left(1 + \frac{2 R1}{Rg}\right) \frac{R3}{R2}$

Når v_1 og v_2 beveges i takt, vil v_1 , v_2 , v_3 og v_4 måtte bli like, og derved

$$v_3 + v_4 = v_1 + v_2$$

Liketaktsforsterkningen (common mode forsterkningen) blir da $A_{CM} = \frac{v_3 + v_4}{v_1 + v_2} = 1$

for 1. trinn og *null* etter 2. trinn. Liketaktssignalet er altså tilstede internt.

På samme måte som for vanlige operasjonsforsterkere, må det alltid være en likestrømsforbindelse mellom inngangene og jord.

17. Data innsamling og telemetri

Ved innsamling av data fra mange målestasjoner er det ofte mest rasjonelt å bruke tidsmultipleksing. Ved hjelp av en multiplexer blir da hver målestasjon etter tur koblet til en analog til digital omvandler. Denne sender så data videre på parallell eller serie form for videre behandling. Tidsdelingen mellom de forskjellige målestasjonene kan f.eks. foregå slik at alle sensorene samples med samme samplefrekvens $f_s = 1/\Delta T$. Sampleintervallet for hver sensor er altså ΔT . Med m samplede signaler blir antall sampler pr. sekund etter multiplexeren

$$f_s^m = m f_s$$

En annen måte å gjøre det på er å sample de sensorene hvor signalene varierer raskest, oftere enn de andre sensorene. LabVIEW-systemet sampler alle sensorene så raskt som mulig, slik at verdiene skal være tilnærmet samtidige verdier. Så vil systemet vente f.eks. 10 sek. hvis samplefrekvensen er valgt lik 0.1Hz.

De digitale data fra analog til digital omvandleren kan enten direkte gjøres om til en form så de kan vises på f.eks. et flytende krystall display, eller de kan sendes som parallelle eller serielle data til en datamaskin.

17.1. Parallell dataoverføring

Overføring av digitale data på parallell form krever mange linjer og egner seg derfor best for korte avstander. Det er forskjellige standarder for antall databit (f.eks. 8, 16 eller 32) og antall styrebit. En samling av ledninger for styre og datasignaler kalles en *buss* når flere enheter har tilgang til den.

Eksempler på parallell-standarder er *CAMAC* og *GPIB* (eller HP-IB, IEEE488). Den første benytter 24 linjer og ble utviklet for bruk i store forskningslaboratorier (CERN, atomenergi osv.). Et litt enklere system for sammenkobling av måleinstrumenter ble utviklet av Hewlett-Packard som HP-IB systemet. Det er nå standardisert (GPIB) og brukt av mange. Dette er basert på 8 datalinjer for overføring av ASCII-kode (byte seriell), 3 for styring av overføringen (handshaking) og 5 linjer for kontroll, dvs. 16 linjer tilsammen. Det er beregnet for sammenkobling av opp til 15 enheter og avstander under 20 meter. Hver enhet er programmert med en unik adresse. En av enhetene er sjef (*master*). Den skal passe på at bare en enhet sender ad gangen. De andre enhetene kan være rene mottakere (*listener*), sendere (*talker*) eller begge deler. Overføringshastigheten kan være 250 til 500 kbyte/sek.

17.2. Seriell dataoverføring

For dataoverføring over større avstander brukes seriell overføring. Data blir sendt som et bit av gangen, logisk 1 og 0 svarer til to spenningsnivåer. Omgjøringen fra parallell til seriell og fra seriell til parallell form kan gjøres med et shift-register som vist på *figur 17.1*. Det minst signifikante bit sendes først. Serie overføring kalles også **pulskode-modulasjon P.C.M.**

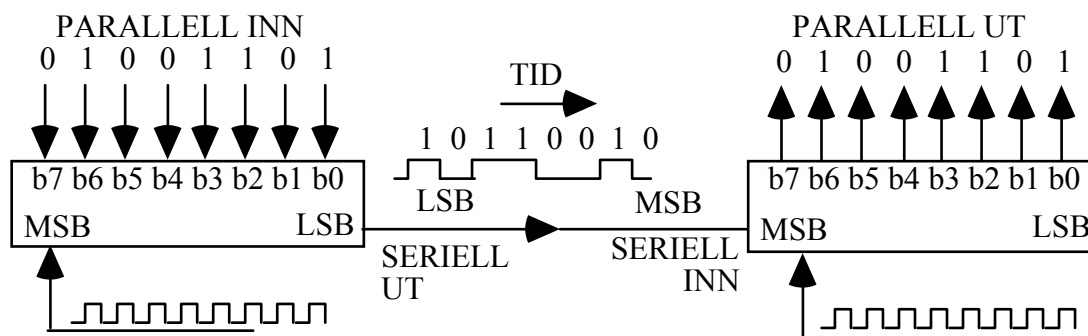


Fig. 17.1 Parallell til serie og serie til parallell omsetting.

Ved asynkron overføring begynner hver byte (eller "ramme") med data med et start bit og avsluttes med ett eller flere stop bit. Mottakeren kan derfor skjønne når data starter og slutter. En mye brukt standard for seriell overføring er RS-232 som er omtalt i avsnitt 11. På steder med mye elektrisk støy kan det være nødvendig å bruke strømsløyfe eller lysledere for overføring av data.

17.2.1. Transmisjons-båndbredde

Den båndbredden som trenges ved seriell overføring kan estimeres slik: Hvis et signal samples f_s ganger pr. sekund, og hvert sample kodes til n bit, blir *bithastigheten* (*bit rate*) lik nf_s . Hvis m signaler multiplexes, blir *bithastighet*

$$R = nm f_s$$

For ett signal skal n bit sendes i løpet av tiden mellom hver sampling, dvs. tiden $\Delta T = 1/f_s$. Bredden av hver bit svarer til denne tiden dividert på antall bit, dvs. $\Delta T/n$. Det bit-mønsteret som har de høyeste frekvenskomponentene, er det hvor 0 og 1 veksler hele tiden. For 8 bit vil det være 01010101 eller 10101010. Grunnfrekvensen i dette signalet er $4/\Delta T$ Hz. Hvis m signaler multiplexes, blir tiden for hvert signal m ganger kortere. Kravet til båndbredde for P.C.M. overføringen for m signaler som kodes til n bit, blir derfor at den iallfall må være så stor som

$$m n f_s / 2 = R / 2$$

hvor f_s er sample frekvens for hvert signal. Hvis f.eks. hvert av 16 signaler samples en gang i sekundet, multiplexes og kodes i 12 bit, må overføringslinjen minst ha en båndbredde på $16 \times 12/2 = 96$ Hz. I forbindelse med seriell overføring brukes ofte enheten *baud* som angir antall transisjoner pr. sekund. Hvis hver bit i meldingen fører til en endring i det overførte signalet, blir hastigheten i *baud* lik bithastigheten.

17.2.2. Virkningen av støy på P.C.M. signal

Forbindelseslinjen mellom P.C.M. senderen og mottakeren kan påvirkes av støy. Hvis det er *hvit støy*, vil effekt-tetthets spektret ha en konstant verdi A watt/Hz, uavhengig av frekvensen. En del av støyen kan filtreres bort ved å ha et lavpass filter med grensefrekvens $R/2$ på inngangen til P.C.M. mottakeren. Den støy-effekten som blir igjen, blir da

$$W_N = \frac{AR}{2} \text{ watt}$$

Hvis vi antar middelveiden for støyamplituden er null, blir standard avvik for støyen etter filteret

$$\sigma = \sqrt{W_N} = \sqrt{\frac{AR}{2}} \text{ volt}$$

P.C.M. signalet fra senderen har en verdi V_p volt for logisk 1 og 0 volt for logisk 0. Etter filteret i mottakeren er det en komparator med terskelverdi $V_p/2$. Hvis signalverdien overskrider denne terskelen, tolkes signalet som logisk 1, hvis ikke, tolkes signalet som logisk 0. Sannsynligheten for feiltolkning kan beregnes fra sannsynlighet-tetthets fordelingen $p(y)$ for støyen. Sannsynligheten for å tolke 0 som 1 er lik arealet under tetthetsfordelingen for $y > V_p/2$ når middelverdien for fordelingen er 0.

Sannsynligheten for å tolke 1 som null er arealet under fordelingen for $y < V_p/2$ når middelverdien for fordelingen er V_p , se *fig. 17.2*.

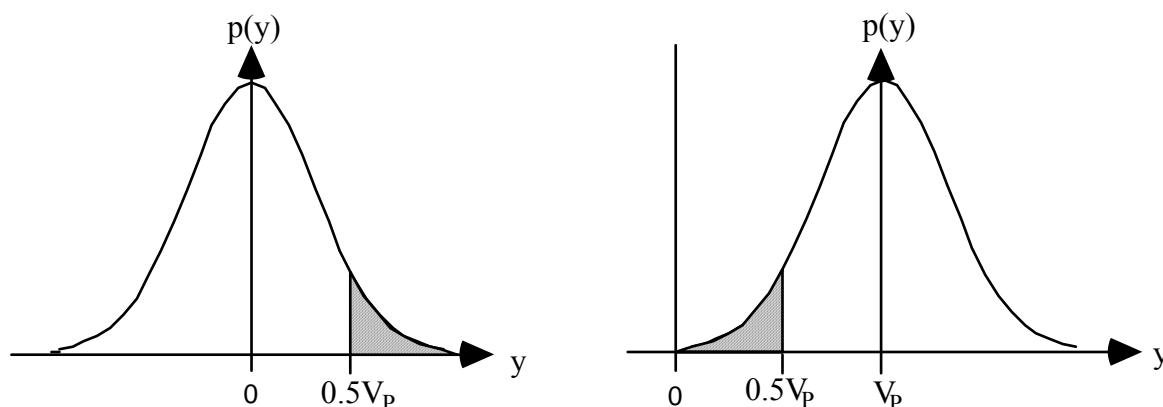


Fig. 17.2 Virkningen av støy på tolkning av signal.

Hvis støyen er normal fordelt (Gauss fordeling), er begge sannsynligheter like store og kan beregnes. Noen sannsynligheter for feil er angitt i tabellen :

V_p/σ	Sannsynlighet for feil
4	0.1
8	10^{-4}
10	10^{-6}
12	10^{-8}

Virkningen av støy på P.C.M. signal.

17.3. Deteksjon av feil og retting

Ved overføring av tekst behøver ikke enkelte bitfeil ha så stor betydning. Ved overføring av målinger og andre numeriske verdier kan feil (f.eks. i mest signifikante bit) være svært alvorlig. Det blir derfor ofte lagt stor vekt på å kunne oppdage feil. Dette kan gjøres ved å overføre kontroll bit i tillegg til informasjonsbitene.

17.3.1. Paritetsbit

Det enkleste feilsystemet er å bruke ett ekstra bit, *paritetsbit*, i tillegg. Hvis det skal være like paritet, velges ekstra bit slik at det komplette kodeordet får et like antall 1'ere. Hvis det skal være odde paritet, skal antallet være odde. Dette systemet vil reagere når det er en, tre, fem, ... feil, og kan lett implementeres.

17.3.2. Feilrettende koder

Mere avansert feildetekterende systemer bruker flere bit til å sjekke om det oppstår feil. Det kan anvendes algoritmer som retter feil i enkelte bit, mens alle samtidige feil i to bit oppdages. Et eksempel på dette er 3 kontroll bit (C0, C1 og C2) som anvendes til å feilsjekke 4 data bit (I0, I1, I2 og I3).

Ved avsendelse av data sendes alle 7 bit av gårde. Kontrollbitene er generert fra databitene ved følgende algoritme (H betyr "eksklusive or", ulike antall input 1'ere gir output 1, like antall 1'ere gir 0) :

$$C0 = I0 \oplus I1 \oplus I3$$

$$C1 = I0 \oplus I2 \oplus I3$$

$$C2 = I1 \oplus I2 \oplus I3$$

Når de 7 bit kommer frem til mottageren, brukes de 4 databit til å generere 3 nye sjekk-bit (G0, G1 og G2) ved å bruke samme algoritme som over. Deretter sammenliknes C0 med G0, C1 med G1 og C2 med G2. Hvis alle tre sammenlikninger gir likhet, er det ikke oppdaget noen feil (det *kan* være feil i to eller flere av de 7 bit). Hvis ikke alle tre sammenlikninger gir likhet, betyr kombinasjonene av mulige ulikheter følgende :

SAMMENLIGNINGER			RESULTAT
C0=G0	C1=G1	C2=G2	Ingen feil
C0=G0	C1=G1	C2≠G2	Feil i C2
C0=G0	C1≠G1	C2=G2	Feil i C1
C0=G0	C1≠G1	C2≠G2	Feil i I2
C0≠G0	C1=G1	C2=G2	Feil i C0
C0≠G0	C1=G1	C2≠G2	Feil i I1
C0≠G0	C1≠G1	C2=G2	Feil i I0
c0≠G0	C1≠G1	C2≠G2	Feil i I3

Når det oppdages en feil, kan den rettes automatisk når slike koder anvendes. Det er da nødvendig å vedlikeholde en *logg* over automatisk rettede feil, ellers vil de aldri oppdages slik at de kan repareres.

17.3.3. Longitudinelle sjekk-summer

Ved synkron serieoverføring av data overføres store blokker av data. Hver slik blokk har en såkalt longitudinell sjekk-sum bakerst, som ved mottak sammenliknes med en sjekk-sum generert på grunnlag av de mottatte data i blokken. En vanlig algoritme av dette slaget kalles "Cyclic Redundancy Check-sum" (CRC).

Ved å anvende mange sjekk-bit kan også slike longitudinelle metoder gjøres feil-korrigerende.

17.4. Frekvensmodulering, F.S.K.

P.C.M. signaler inneholder frekvenser fra likespenning til $R/2 \text{ Hz}$. Mange transmisjonslinjer, telefonlinjer f.eks., kan ikke overføre frekvenser under en viss grense. P.C.M. signalet må derfor ofte moduleres på en bærefrekvens som passer for linjen (frequency shift keying, F.S.K.). Frekvensmodulasjon brukes fordi det har bedre signal/støy forhold enn amplitudemodulasjon. På telefonlinjer kan data sendes som to forskjellige frekvenser innen det hørbare området.

F.S.K. senderen består av en spenningsstyrt oscillator (VCO). Mottakeren er en faselåst sløyfe som består av en spenningsstyrt oscillator, multiplikator og lavpassfilter i en lukket sløyfe. (Jfr. avsnitt 13) Multiplikatoren gir et feilsignal proporsjonalt med faseforskjell mellom inngangssignalet og signalet fra VCO, og tilbakekoblingen gjør at frekvensen til VCO låses til frekvensen til inngangssignalet.

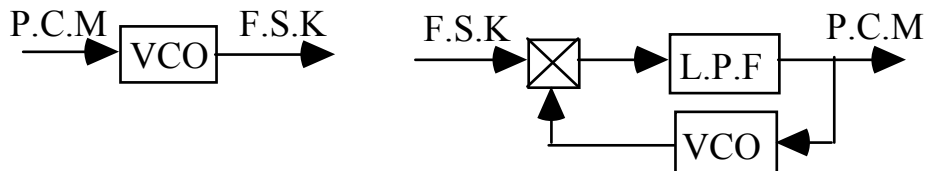


Fig. 17.3 F.S.K. sender og mottaker.

Utgangssignalet fra VCO er

$$V_1 \sin(2\pi(f_c + kv)t)$$

hvor V_1 er amplituden til VCO-signalet, f_c er frekvensen for inngangsspenningen $v=0$ og k er følsomheten (Hz/V) for oscillatoren. Hvis modulasjonssignalet er et enkelt sinussignal $v = V_m \sin 2\pi f_m t$ blir det frekvensmodulerte signalet

$$v_{FM}(t) = V_1 \sin 2\pi(f_c + D \sin 2\pi f_m t)t$$

hvor $D = kV_m$ er maksimalt avvik i frekvensen fra f_c . Forholdet D/f_m kalles *modulasjonsindeks*.

Hvis modulasjonssignalet ikke bare inneholder en enkelt frekvens f_m , men alle frekvenser mellom 0 og f_M , består det frekvensmodulerte signalet av alle frekvenser i et bånd omkring bærefrekvensen f_c . Bredden av frekvensbåndet avhenger av forholdet D/f_M . Frekvensspektret for PCM signaler går fra 0 til $R/2$, dvs. $f_M = R/2$ hvor R er bit-hastigheten. Dette krever en

$$\begin{array}{llll} \text{båndbredde:} & f_c - (D + R/2) & \text{til} & f_c + (D + R/2) & \text{hvis} & 2D/R \ll 1 \\ \text{og båndbredde:} & f_c - (D + R) & \text{til} & f_c + (D + R) & \text{hvis} & 2D/R \approx 1. \end{array}$$

For et F.S.K. signal med $f_c = 1080 \text{ Hz}$, $D = 120 \text{ Hz}$ og $R = 200 \text{ baud}$, blir $2D/R = 1.2$. og nødvendig båndbredde tilnærmet 760 Hz til 1400 Hz.