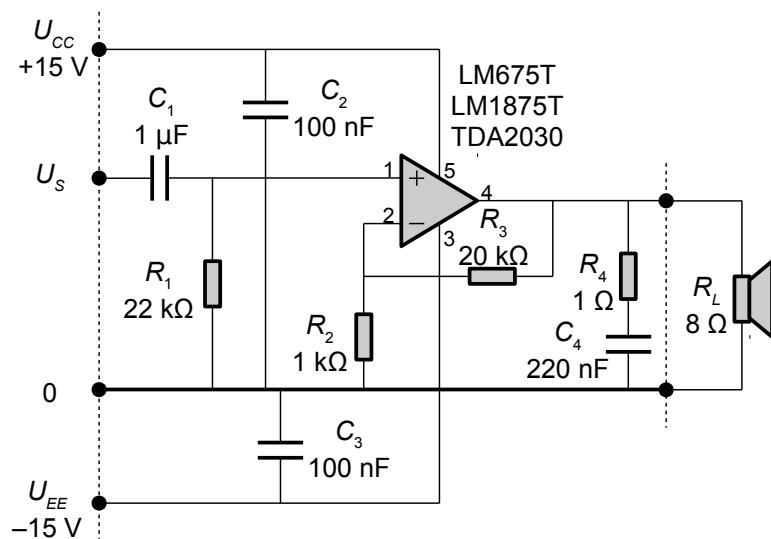


Analogteknik

T-005

Tore Skogberg

2015



<http://www.ida.dk/studerende>

VIDEN DER STYRKER
ida.dk

BLIV MEDLEM
på ida.dk

MAKE IT EASY

BLIV MEDLEM AF IDA OG FÅ MASSER AF FORDELE

I IDA bliver du en del af et stærkt fællesskab, der hjælper dig godt og grundigt igennem din studietid. Vi tilbyder bl.a. billige studieforsikringer, arrangementer på dit studiested, en attraktiv studiekonto, netværk og rådgivning. Herefter koster studiemedlemskabet 20 kr. pr. måned. Og du kan endda blive medlem i dag!

➤ **Læs mere på ida.dk/studerende**

KUN FOR INGENIØRSTUDERENDE OG STUD.SCIENT.ER

Forord

Velkommen til min bog om analog elektronik, der siden 2012 er taget i brug ved undervisningen på Ingenørhøjskolen i Århus for E3MSE mixed-signal elektronik og E4ASD analog systemdesign.

Bogen omhandler analog elektronik, men alt styres i dag af digital elektronik for billede og lyd er jo noget man streamer til sin PC fra internettet så hvor kommer den analoge elektronik ind?

Den musik du lytter til er optaget via en mikrofon der omsætter lydens svingninger til en dermed analog elektrisk svingning og den film du ser er optaget med et kamera der omsætter lyset til et analogt elektrisk signal. Disse signaler skal forstærkes op og omformes til digitale signaler og det kræver brug af analog elektronik. Ved transmission til din lyd- eller billedafspiller er de digitale bit båret som analoge signaler uanset om du modtager signalet via kabel eller parabol og med en DVD skal de digitale data aflæses som en analog modulation af en laserstråle. Dit TV gengiver billedet ved en analog justering af farve og lysintensitet for hver enkelt pixel og MP3 lyden skal forstærkes i analoge kredsløb for at drive din øreprop eller en højttaler.

Den høje præcision af dit armbåndsur og din PC skyldes en krystalstyret oscillator der kun kan bygges med analog elektronik. Effektforsyningen af din PC benytter en netadapter hvor analoge komponenter varetager energitransporten. Mikroprocessoren er en digital komponent, men det kræver viden om analog elektronik for at forstå hvorfor den fejler ved over clocking, hvorfor dens effektforbrug stiger med frekvensen og hvorfor spændingen i dens CPU kerne er omkring 1,5 V.

Jeg har valgt at indlede med to kapitler om det elektriske grundlag og den nødvendige matematik, men du kan roligt springe dem over hvis du er mere interesseret i operationsforstærkeren der introduceres i tredje kapitel ved nogle typiske kredsløb og hvor de vigtigste problemer uddybes. Fjerde kapitel viser hvordan man designet analoge filtre med operationsforstærkere og i femte kapitel forklares om hvordan halvlederne og operationsforstærkeren er bygget med eksempler på kredsløb. Sjette kapitel omhandler effektelektronik og bogen slutes med appendiks og indeks.

Der kræves en del viden om matematik for at kunne følge bogens kapitler, men jeg har forsøgt at holde den fri for en alt for avanceret tilgang. Hvis der alligevel forekommer matematik som du ikke forstår så kan du vælge én af to veje frem. Den ene er simpelthen at ignorere det svære og bare læse videre i teksten. Som en underviser i musik sagde: *Der er ikke noget der er så svært at det ikke kan springes over*. Det er en metode jeg selv anvender når jeg skal hurtigt igennem en tekst og ikke har overskud til detaljer eller jeg ikke har behov for at forstå i dybden. Den anden metode er at tage et ark papir og følge med i udledningerne. Den metode benytter jeg også, men kun når det er vigtigt at forstå detaljerne om hvordan man kommer frem til resultatet.

Hvis du skal sætte dig ind i et kapitel hvis stof er nyt for dig så er mit råd at du ser på billedeerne og læser billedteksten. Spring over tekstafsnit og formler. Når du derefter har skabt dig et overblik over kapitlet er det tid til at læse de dele af den ledsagende tekst som du har interesse i. Hvis du ikke er sat på denne jord for at læse formler så spring dem over for de formler du får brug for kan du nok klare at læse op på senere.

Bogens illustrationer har mange formler tilknyttet og kan sine steder virke som en designmanual, men det er dit ansvar at benytte oplysningerne korrekt. Jeg anbefaler derfor at du sætter dig ind i de afsnit der har interesse for dig og dit projekt inden du anvender formlerne. En kær lærer i folkeskolen sagde engang at *man må kun bruge en formel hvis man læser mindst fem linjers tekst inden og efter formlen*. Det er et godt råd som hermed er givet videre.

Forbehold for fejl. Jeg påtager mig intet ansvar for om et design gennemført efter mine anvisninger viser sig at virke eller ej og heller ikke for skader som dette design måtte påføre personer eller materiel.
--

Bogen indfrier en livsdrøm. Jeg begyndte at skrive den 2010 i min fritid uden at vide om den ville kunne udgives. De første tre kapitler var skrevet da der opstod behov for en lærebog til et fag på Ingeniørhøjskolen i Århus og bogen kom i brug fra foråret 2012. Det var planlagt at fortsætte med aktive filtre og halvledere og det passede godt med at benytte bogen til et nyt fag hvor fokus var på operationsforstærkerens opbygning og design af effekteletronik. Siden 2015 udkommer bogen som PDF for fri download dels for at udbrede bogen til et større publikum, men også for at kunne imødekomme piratkopiering og min indtægt kommer nu fra annoncer.

Min egen tilgang til elektronik startede i barndommen dog uden at være målrettet; jeg har blot altid været nysgerrig. Som dreng legede jeg med Märklin-tog og en bog om batterier og glødelamper startede min interesse for elektriske kredsløb. Jeg så en annonce for et Philips EE20 byggesæt i et Anders And blad og det gav mulighed for at bygge kredsløb med transistorer. Jeg lånte derefter det lokale bibliotek tør for bøger om elektronik og måtte efterfølgende ty til de engelske bøger. Det sprog kunne jeg ikke læse på det tidspunkt så for at stille min først læste jeg billeder og forsøgte den vej at forstå hvad der var meningen. Interessen var skabt og det førte til Josty Kit samlesæt der minder om de tilgængelige kits fra Brinck elektronik og Kontron og snart byggede jeg selv mine kredsløb fra grunden selv om jeg ikke vidste meget om elektronik. Senere kom interessen for højttalerens funktion og hvorfor nogle rum er bedre til lydgengivelse end andre og det førte via nogle omveje til ingeniørstudiet. Jeg byggede her en stereoforstærker, en analog musiksynthesizer og et måleapparat for højttalere. Dels fordi der var et behov, men også for at afprøve teorien.

Jeg blev akademiingeniør i elektronik fra Danmarks Ingeniørakademiet i 1983 og ti år senere tog jeg nogle fag i musikvidenskab fra Københavns Universitet. Omkring 1995 gik jeg i gang med at bygge et elektronisk orgel drevet af nysgerrighed. Siden 2006 er jeg civilingeniør i akustik fra Danmarks Tekniske Universitet. Jeg har herefter arbejdet som konsulent for den elektroakustiske side af et trådløst headset for hørehæmmede, magnetisk opladning af batterier, interface til tryktransducere og test af software. Mit professionelle arbejde siden 1983 omfatter udvikling af analog og digital elektronik, programmering af mikroprocessorer og brug af CAD software for diagramtegning, simulering og design af printkort. September 2010 blev jeg ansat som vikar ved Ingeniørhøjskolen, februar 2011 som adjunkt og fra april 2013 er jeg lektor ved ingeniørhøjskolen.

En varmt tak til mine kolleger Per Thorsen (pt@ase.au.dk) og Martin Kristensen (mk@inano.dk) for mange og dybe diskussioner om elektronik, optik, fysik og halvlederteori, dertil Lars G Johansen (lgj@ase.au.dk) for udbytterige diskussioner om lyd og akustik og Arne Justesen (aju@ase.au.dk) for gode diskussioner om DC-DC konvertere og magnetiske systemer. Endelig en meget stor tak til de studerende der har stillet spørgsmål til de første fem udgaver af min bog eller som har fundet fejl i bogen.

Det kan ikke udelukkes at der stadig er fejl i bogen og hvis du finder en til så må du gerne sende mig en oplysning på mail@torean.dk. Angiv bogens version, sidenummeret og din kommentar.

Bogen er skrevet med LibreOffice på operativsystemerne Ubuntu, Linux Mint og Windows 7. Jeg har testet algoritmerne med MATLAB og open-source programmet Octave. En del illustrationer har været en tur igennem GNOME og IrfanView billedredigeringerne.

Papirudgaven af bogen er trykt på Vester Kopi for Teknisk boglade ved Ingeniørhøjskolen i Århus.

Tore Skogberg.

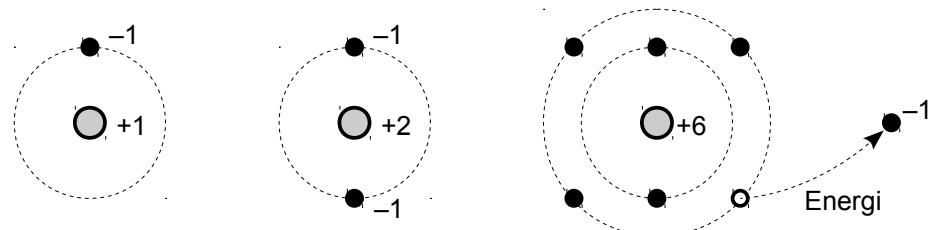
Indholdsfortegnelse

1 ELEKTRICITET.....	7
1.1 Ladning.....	8
1.2 Elektrisk felt.....	9
1.3 Modstand.....	10
1.4 Konventioner og love.....	15
1.5 Thevenin og Norton.....	19
1.6 Kondensator.....	20
1.7 Magnetisk felt.....	27
1.8 Selvinduktion.....	29
1.9 Magnetiserbar kerne.....	39
1.10 Resonans.....	42
1.11 Energi.....	43
1.12 Dæmpet svingning.....	44
1.13 Magnetiske kræfter.....	45
1.14 Definitioner.....	47
1.15 Sammenfatning.....	55
1.16 Opgaver.....	56
2 ANALYSE.....	59
2.1 Impedans.....	60
2.2 Komplekse tal.....	62
2.3 Differentialregning.....	66
2.4 Taylor polynomier.....	73
2.5 Eulers identitet.....	75
2.6 Fourier serier.....	77
2.7 Kompleks Fourier serie.....	79
2.8 Fourier transformation.....	81
2.9 Diskret Fourier transformation.....	85
2.10 Laplace transformation.....	89
2.11 Overføringsfunktion.....	96
2.12 Parasitiske komponenter.....	100
2.13 Analogier.....	101
2.14 Opgaver.....	104
3 OPERATIONSFORSTÆRKER.....	105
3.1 Lidt historie.....	106
3.2 En OpAmp.....	107
3.3 Spændingsforstærkning.....	108
3.4 Strømforstærkning.....	116
3.5 Filtrerende kredsløb.....	118
3.6 Tonegeneratorer.....	123
3.7 Støjgenerator.....	125
3.8 DC begrænsninger.....	127
3.9 Båndbredde.....	133
3.10 Ind- og udgangsimpedans.....	143
3.11 Støj.....	147
3.12 Sammenfatning.....	162
3.13 Effektforsyning.....	166
3.14 Opgaver.....	169
4 FILTRE.....	171
4.1 Filtre af første orden.....	172
4.2 Filtre af anden orden.....	180
4.3 Filtre af højere orden.....	187
5 HALVLEDER.....	193
5.1 Dioden.....	193
5.2 Lysdiode.....	212
5.3 Transistor ved DC.....	218
5.4 Transistor ved AC.....	233
5.5 Operationsforstærker.....	253
5.6 Specielle operationsforstærkere.....	268
5.7 Halvlederteori.....	271
5.8 Elektronrøret.....	282
5.9 Opgaver.....	284
6 EFFEKTREDSSLØB.....	287
6.1 Effektforstærker.....	287
6.2 Klassifikation af udgangstrin.....	288
6.3 Effektforsyning.....	310
6.4 Ensretter.....	310
6.5 Spændingsregulator.....	314
6.6 Shuntregulator.....	315
6.7 Serieregulator.....	321
6.8 DC-DC konverter.....	326
6.9 Opkonvertering.....	327
6.10 Nedkonvertering.....	330
6.11 Dimensionering.....	332
6.12 Konverterens tab.....	334
7 APPENDIKS.....	337
7.1 Konstanter.....	337
7.2 Variable.....	337
7.3 Forkortelser.....	338
7.4 Enhedssystemet.....	339
7.5 Omregning.....	341
7.6 Litteraturliste.....	342
8 INDEKS.....	343

1 ELEKTRICITET

Kapitlet tager sit udgangspunkt ved elektronens ladning og de elektrostatiske kræfter og det leder til en definition af elektrisk strøm og spænding. Ohms lov benyttes til at indføre den elektriske *modstand* som kan beregnes alene ud fra kendskab til materialet og de fysiske dimensioner af et emne. Det teoretiske fundament for kredsløbsanalyse præsenteres ved Kirchhoffs love. Energibetrægninger leder til *kapacitet* og *selvinduktion* og det vises hvordan man kan beregne disse værdier. Kapitlet afsluttes med magnetiske kredsløb og de kræfter der udnyttes i en elektrisk motor og i en højtalere.

Elektricitet skyldes elektronerne¹ i atomets yderste skal *valenselektronerne* som blandt andet er betydende for atomets kemiske egenskaber. Atomet er opbygget med elektronerne arrangeret i et antal skaller efter et sæt regler som den danske forsker Niels Bohr udledte i 1913. Brint (hydrogen) og helium har henholdsvis én og to elektroner i den første skal og alle andre atomer har mindst to skaller hvor elektronen traditionelt opfattes som kredsende omkring atomets kerne.



Figur 1 – Grundstofferne hydrogen, helium og kulstof samt en illustration at en elektron kan slås løs ved en udefrakommende energi som typisk kan være temperaturen. Kernens partikler er ikke vist, der er intet belæg for at elektronen tegnes som en partikel i en cirkelformet bane om kernen og størrelsesforholdet også er forvrænget.

Elektronerne har vedtægtsmæssigt en negativ ladning og atomets kerne har en positiv ladning der modsvarer antallet af elektroner så atomet er udadtil i elektrisk balance. I den klassiske opfattelse af atomet er elektronens bevægelse om kernen associeret til jorden og planeterne i deres kredsløb omkring solen. Sammenligningen skal dog ikke bringes for vidt for elektronerne i den yderste bane mærker kun en svag tiltrækning til den positive kerne på grund af frastødningen fra elektronerne nærmere ved kernen i modsætning til planeterne der alle tiltrækker hinanden på grund af gravitationen. Dertil kommer at kvantemekanikken behandler elektroner som energi fordelt over en skal omkring kernen og ikke som en roterende partikel.

Elektronerne i den yderste af skallerne kan slås væk fra atomet alene på grund af temperaturen idet varme er udtrykt ved elektronernes bevægelsesenergi og de vibrerer desto kraftigere når det er varmt. Vibrations amplitude er statistisk betinget og ikke bare en pæn oscillation omkring en ligevægt. Elektron efter elektron vil få tilstrækkelig med energi til at smitte væk fra atomet og de flyder så rundt indtil de igen bliver opfanget af et andet atom. Denne sky af frie elektroner kan påvirkes af kraften fra et elektrisk felt hvorved de skubbes væk fra den negative pol og trækkes hen mod den positive pol. Elektronen har en ladning og når en ladning flyttes så løber der en elektrisk strøm båret af de frie elektroner.

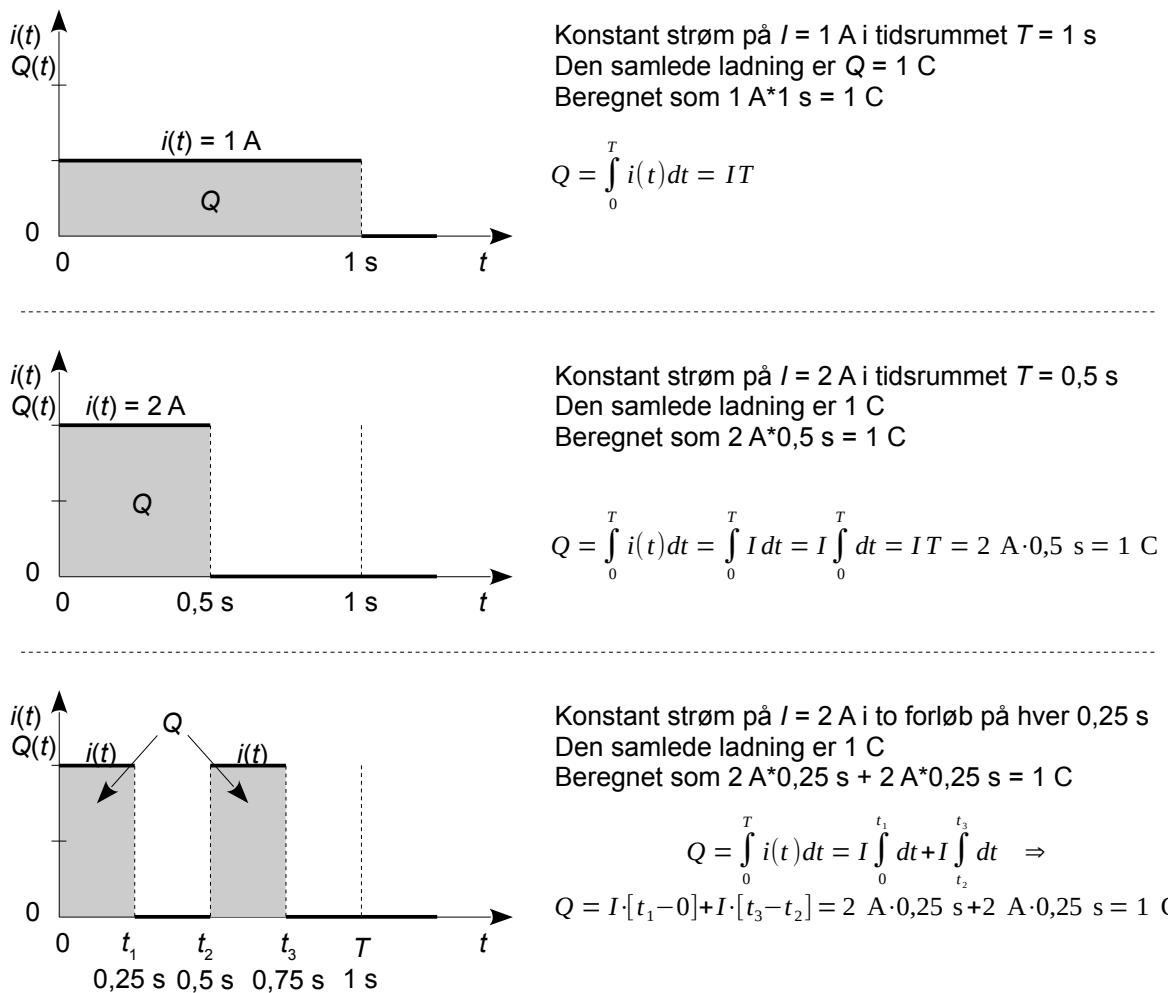
Både en ledet og en halvleders funktion er baseret på at elektroner kan slås fri af temperaturen og de kan derefter sættes i bevægelse af et elektrisk felt. Elektronerne kan opnå energi nok til at slå andre elektroner løs ved kollision med krystallets atomer som så bliver accelereret af det elektriske felt. Det udnyttes blandt andet i fotomultiplikatoren og visse halvledere. Elektronerne kan også slås fri af elektromagnetisk stråling, som fx lys, hvilket forklarer at en fotocelle kan producere elektrisk strøm ved belysning. Et stærkt elektrisk felt kan trække elektroner løs fra krystallet. Det kaldes for en tunneleffekt og udnyttes i zenerdioden samt et tunnelling mikroskop. Ved tilstrækkelig høj temperatur kan elektronen også opnå tilstrækkelig energi til at blive slynget helt ud af det opvarmede metal. Det blev udnyttet i 1900-tallets radiorør hvor katoden blev opvarmet til glødning så den afgav elektroner som derefter blev tiltrukket af den positive anode.

1 http://www.denstoredanske.dk/lt%2c_teknik_og_naturvidenskab/Fysik/Elektrontransport_i_faste_stoffer/elektrisk_felt

1.1 Ladning

Hver elektron har en fast værdi af den elektriske ladning, elementarladningen, der beskrives ved symbolet² q_0 . En samling af elementarladninger udgør en ladning (*charge*) der beskrives ved symbolet Q for en statisk ladning og $q(t)$ eller blot q for en tidsvarierende ladning. Ladning måles i coulomb (C) og den består af et helt antal elementarladninger idet ladningen ikke kan neddeles. Denne opdeling i partikler er ikke noget man ser ret meget til i praksis for selv en forholdsvis lille ladning på 1 C svarer til at der er samlet det ufatteligt store antal af $6,242 \cdot 10^{18}$ elektroner³.

En gruppe af elektroner i bevægelse udgør en elektrisk strøm (*electrical current*) der betegnes med symbolet $i(t)$ for en tidsvarierende strøm og med I for konstant strømstyrke. Elektrisk strømstyrke måles i ampere med enhedssymbolet A. En elektriske strømstyrke på 1 A betyder at der i løbet af ét sekund passere en ladning på 1 C igennem en flade der kan være grænsen mellem to dele af en transistor eller en tænk (virtuel) flade på tværs i en leder. Der har med andre ord passeret godt og vel 6 milliarder milliarder elektroner i det ene sekund. Det er helt den samme ladning som en strøm på 2 A står for i tidsrummet 0,5 s for det er det samme antal af elektroner. Den eneste forskel er at der er lidt mere fart over feltet.



Figur 2 – Den samlede ladning gennem en ledning som funktion af tiden.

Ladningen kan opbygges som summen af flere bidrag som vist underst i illustrationen. Det svarer til situationen hvor et bils batteri oplades af en pulserende strøm.

2 Andre muligheder er: q , q^- , e , e_0 og e^- .

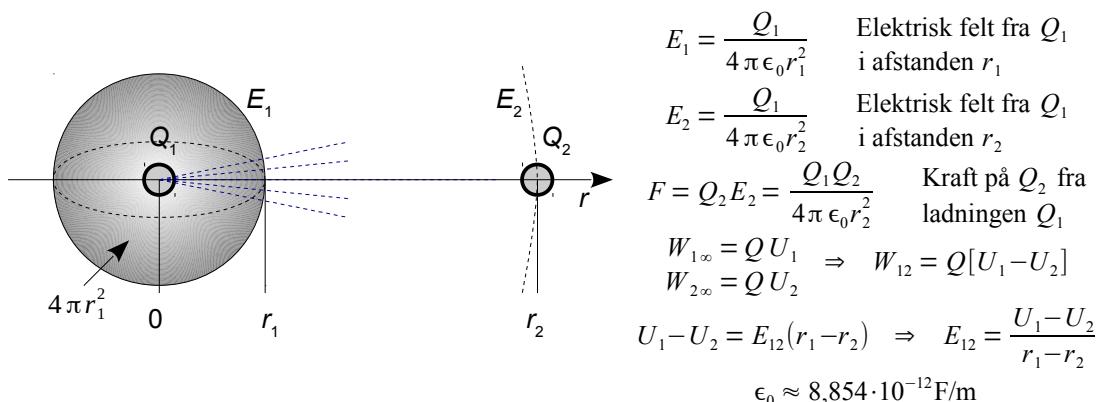
3 Se fx http://en.wikipedia.org/wiki/Electric_charge.

1.2 Elektrisk felt

En ladning kan bestå af elektroner, der vedtægtsmæssigt tillægges en negativ elektrisk værdi, eller ladningen kan bestå af atomer som har afgivet en elektron og derfor udgør en positiv ladning. Det betyder at ladningen Q kan have enten et negativt eller positivt fortegn.

Omkring en ladning er der et elektrisk felt med den samme styrke i alle retninger i samme afstand for ladningen har ikke en foretrukken udbredelsesretning i universet. I stor afstand kan en ladning opfattes som punktformet uanset ladningens form. Feltets styrke er proportional med ladningen så en fordobling af ladningens værdi fordobler feltstyrken og uden ladning er der intet felt. Det har stor teoretisk betydning da det elektriske felt derved er lineært og tillader addition af enkeltbidrag.

Feltets styrke beskrives ved en tæthed hvor man opfatter styrken som et antal feltlinjer der stråler ud fra ladningen i alle retninger. Feltets styrke kan derved beskrives ved det antal feltlinjer som krydser igennem et areal i en given afstand. Feltlinjerne udgår fra centrum af ladningen og i en given afstand er ladningen Q fordelt over en kugleskal med radius r der har arealet $4\pi r^2$. Ud over de to parametre for ladning og afstand behøves en konstant ϵ_0 , der tilpasser talværdi og enhed til SI-systemet. Den elektriske konstant ϵ_0 kaldes også for vacuumpermittiviteten og den gælder med god tilnærrelse i atmosfærisk luft. Feltet breder sig i alle retninger og det sker med lysets fart så man kan normalt antage at feltet blot følger med når værdien af en ladning ændres eller flyttes.



Figur 3 – Feltet omkring en ladning strækker sig uendeligt langt væk og er ligeligt fordelt i alle retninger for en enlig ladning. Feltstyrken beskrives ved antallet af feltlinjer per areal. To ladninger påvirker hinanden med en kraft der er funktion af de to ladninger og afstanden mellem dem.

Et elektrisk felt har den egenskab at det kan påvirke en anden elektriske ladning med en mekanisk kraft der kan være tiltrækkende eller frastødende. Kraften afhænger af fortegnet på de to ladninger og deres indbyrdes afstand. Det elektriske felt fra ladningen Q_1 vil påvirke en anden ladning Q_2 i afstanden r_2 med en kraft F der er givet ved feltstyrken E_2 i afstanden r_2 fra ladningen Q_1 og dertil styrken af ladningen Q_2 . Kraften er givet ved at samle udtrykkene og det leder til Coulombs lov⁴. Formlen er symmetrisk så man kan lige så vel tage udgangspunkt i feltstyrken fra ladningen Q_2 og betragte kraften på Q_1 . Kraften udgør en frastødning hvis fortegnet af produktet $Q_1 Q_2$ er positivt så kraften peger i r -aksens retning og det betyder at Q_1 og Q_2 har samme fortegn. Kraften udgør en tiltrækning hvis fortegnet af produktet er negativt og de to ladninger har forskellige fortegn.

Fra mekanikken har vi definitionen af arbejde W (work) som kraft gange vejlængde så der kan opbygges eller udvindes energi ved at flytte rundt på ladningerne i et elektrisk system. Ved at flytte en ladning fra afstanden r_1 til det uendeligt fjerne opbygges en potentiel energi på QU_1 og hvis ladningen derefter bringes tilbage til afstanden r_2 indvindes der en potentiel energi på QU_2 . Det svarer til en fjeder der spændes hvorefter der opbygges potentiel energi som senere kan indvindes igen.

4 Kilde: http://en.wikipedia.org/wiki/Coulomb's_law.

Den energi W_{12} der opbygges ved at føre ladningen fra afstanden r_1 til r_2 er uafhængig af om turen var et smut forbi det uendeligt fjerne eller om den gik direkte fra r_1 til r_2 og fortegnet angiver om der blev vundet energi (positivt fortegn) eller om der tabes energi (negativt fortegn). Den energi der potentielt er tilstede i en given ladning beskrives ved ladningen Q og et potentiale U der får en enhed som energi per ladning (joule per coulomb) som har fået enheden volt ($V = J/C$).

Energi er lig med kraft gange vejlængde, men kraften er ikke konstant under flytningen så det er ikke nok bare at gange den elektriske kraft med den strækning ladningen bevæges. Energien beregnes ved at addere små stykker af kraft gange vejlængde hvor hvert stykke er så lille at kraften kan regnes for konstant og den justeres så stykke for stykke. For at flytte ladningen fra afstanden r til afstanden $r + dr$ kræves der en energi på $dW = Fdr$. Energien beregnes ved at summere bidragene over den aktuelle strækning og ved de uendeligt små strækninger dr benyttes integration for at kunne gennemføre summeringen.

Den største kraftstigning opnås når en ladning flyttes fra afstanden r til det uendeligt fjerne hvor feltstyrken er nul på grund af afstandsrelationen. Som det ses af formagens symmetri er der ingen forskel på om det var ladningen Q_1 der blev flyttet i det elektriske felt fra Q_2 eller omvendt. Det arbejde der udføres er givet ved at integrere Fdr fra afstanden r til uendelig hvor r er afstanden mellem centrum af hver af de to ladninger. Energien repræsenterer en potentiel energi og svarer til den energi der oplagres i en fjeder når den spændes og som kan udnyttes på et senere tidspunkt hvorved fjedren afspændes.

$$W = \int_r^{\infty} F dr = \int_r^{\infty} \frac{Q_1 Q_2}{4 \pi \epsilon_0 r^2} dr = \frac{Q_1 Q_2}{4 \pi \epsilon_0} \int_r^{\infty} r^{-2} dr = \frac{Q_1 Q_2}{4 \pi \epsilon_0} [-r^{-1}]_r^{\infty} \Rightarrow W = \frac{Q_1 Q_2}{4 \pi \epsilon_0 r}$$

Hvis ladningen ikke føres langs forbindelseslinjen mellem de to ladninger, men langs en valgfri rute skal integralet inkludere $\cos(\theta)$ hvor θ er vinklen mellem feltets og rutens retninger. Det betyder at arbejdet er størst langs med feltet hvor $\theta = 0$ og dermed $\cos(\theta) = 1$ og nul på tværs af feltet hvor $\theta = 90^\circ$ og $\cos(\theta) = 0$.

For et sted med konstant feltstyrke kan spændingsforskellen beskrives ved afstanden a mellem de to potentialer⁵. Den elektriske feltstyrke E ses at få enheden volt per meter (V/m).

For luft er den øvre grænse på 3000 V ved en afstand på 1 mm ($E = 3 \text{ MV/m}$). I praksis benyttes den mere konservative værdi på 1 kV/mm ($E = 1 \text{ MV/m}$) så der er en sikkerhedsmargin. Ved en større feltstyrke end det tilladelige ioniseres luften så der dannes korona eller endda elektrisk udladning (gnister eller lysbue). Nogle lignende forhold gør sig gældende i halvede hvor grænsen på 30 MV/m for silicium let nås ved moderne integrationstæthed. Allerede ved 100 nm afstand mellem de spændingsførende dele vil 3 V give en feltstyrke på grænseværdien og derfor må arbejdsspændingen stadigt sænkes for at kunne fortsætte ønsket om forøgelse af tætheden i de integrerede kredse.

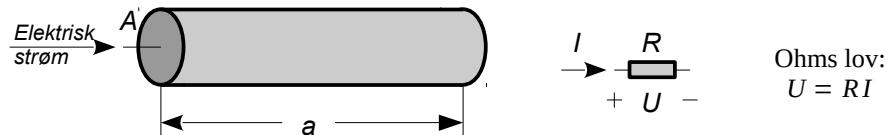
1.3 Modstand

Hvis der etableres en elektrisk forbindelse mellem to punkter med en spændingsforskell U så vil det elektriske felt påvirke elektronerne i den elektriske leder med en kraft så elektroner flyttes fra den negative pol til den positive. Der vil løbe en elektrisk strøm af gennem forbindelsen hvis styrke er givet ved antallet af elektroner der per tidsenhed passerer ind i lederen (og ud af lederen igen). Strømretningen blev fastlagt inden videnskaben havde fundet frem til at strømmen blev båret af negativt ladede elektroner og strømretningen blev defineret som positiv med udgangspunkt ved det positive punkts potentiale og frem til referencen ved det negative punkts potentiale. At elektronerne så i praksis løber den anden vej er lidt uheldigt, men det er til at leve med.

Strømmen af elektroner flyder ikke problemløst gennem en elektrisk leder for strømmen bæres af de frie elektroner og et materiale har kun en vis mængde af frie elektroner. En isolator har ingen frie elektroner og vil derfor ikke kunne lede en elektrisk strøm. Desuden rammer elektronerne ind i ledernes atomer, forureningselementer og fejl i krystallets gitterstruktur hvilket lægger en dæmper på bevægelsen som i gennemsnit over mange elektroner vil virke som en elektrisk modstand.

⁵ Det gælder eksempelvis for det indre af en kondensator, samt for en elektrostatisk højttaler og en kondensatormikrofon.

Hvis spændingen sættes op til det dobbelte vil der løbe dobbelt så mange elektroner indenfor et givet tidsrum og det ledte i 1827 Georg Simon Ohm til at formulere en lov der efterfølgende har vist sig at danne selve grundloven indenfor al elektricitet og elektronik. Loven blev introduceret på basis af empiri det vil sige ud fra eksperimenter i laboratoriet⁶.



Figur 4 – En leder fremstilles ved en lang stang med længden a og et areal i tværsnittet på A og beregning af modstandsværdien gælder for alle tværsnit (cirkulære, kvadratiske eller andet).

Spændingen U er den energikilde der bringer en strøm I af elektroner til at bevæge sig i en given retning og med modstanden R som proportionalitet mellem strøm og spænding.

Modstandsværdien skyldes sammenstødenes bremsende virkning på de frie elektroner. Det elektriske felt accelererer elektronerne op i fart, men ved kollision med de fastsiddende atomer afgives bevægelsesenergi i form af varme og hastigheden sænkes til en konstant gennemsnitlig værdi der er proportional med den eksterne spænding. Modstandsværdien R kan beregnes ud fra kendskab til emnets materiale udtrykt ved den specifikke modstand ρ (også kaldet resistivitet), som er angivet i tabellen herunder og dertil længden a og tværsnittets areal A . Modstandens variation med temperaturen angives ved temperaturkoefficienten α der oplyses ved en referencetemperatur θ_0 som er 0°C i tabellen⁷.

Materiale	ρ	α	Modstandsværdi ved $\theta_0 = 0^\circ\text{C}$:
Sølv	$14,7 \cdot 10^{-9} \Omega\text{m}$	$3,8 \cdot 10^{-3}/^\circ\text{C}$	$R = \rho \frac{a}{A}$
Kobber	$15,5 \cdot 10^{-9} \Omega\text{m}$	$3,9 \cdot 10^{-3}/^\circ\text{C}$	Modstandsværdi ved θ :
Aluminium	$26,0 \cdot 10^{-9} \Omega\text{m}$	$3,9 \cdot 10^{-3}/^\circ\text{C}$	$R = \rho \frac{a}{A} [1 + \alpha(\theta - \theta_0)]$
Jern	$90 \cdot 10^{-9} \Omega\text{m}$	$5,0 \cdot 10^{-3}/^\circ\text{C}$	
Platin	$97,7 \cdot 10^{-9} \Omega\text{m}$	$3,9 \cdot 10^{-3}/^\circ\text{C}$	
Konstantan	$490 \cdot 10^{-9} \Omega\text{m}$	$0,008 \cdot 10^{-3}/^\circ\text{C}$	
Havvand	$0,2 \Omega\text{m}$		$R_{\text{Serie}} = \rho \frac{a_1 + a_2}{A} = R_1 + R_2$
Silicium	$1600 \Omega\text{m}$	$-0,075/^\circ\text{C}$	$R_{\text{Parallel}} = \rho \frac{a}{A_1 + A_2} = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}$

En kobberleder med trådtykkelsen 1 mm har arealet $785 \cdot 10^{-9} \text{ m}^2$ så 10 m tråd har modstanden $0,2 \Omega$.

En kobberbane på et printkort har længden 100 mm, bredden 0,25 mm og tykkelsen 35 μm . Arealet af banens tværsnit bliver $8,75 \cdot 10^{-9} \text{ m}^2$ så modstanden er $0,18 \Omega$ ved 0°C . Ved $\theta = 25^\circ\text{C}$ bliver korrektionen på $1 + 3,9 \cdot 10^{-3} \cdot (25 - 0) = 1,1$ så kobberbanens modstand stiger til $0,20 \Omega$ ved 25°C .

Udtrykkene for sammenkobling af to modstande i serie eller parallel følger direkte af formlen for modstanden af en leder.

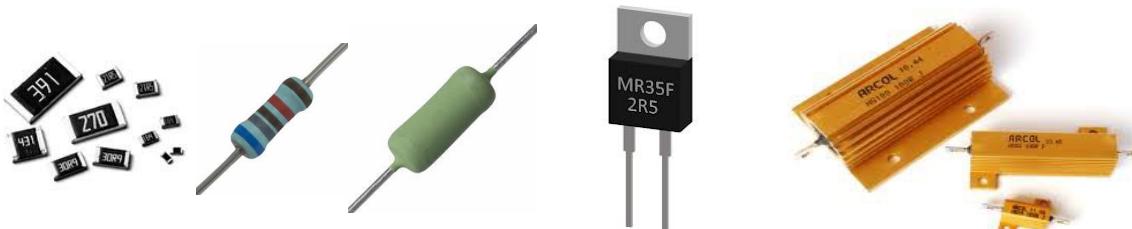
En serieforbindelse svarer til at to trådstykker med samme areal sidder i forlængelse så de to længder adderes. En parallelforbindelse svarer til at to trådstykker med samme længde sidder ved siden af hinanden så her adderes arealerne. Hvis arealet udtrykkes som $A_1 = pa/R_1$ og tilsvarende for A_2 følger udtrykket for en parallelforbindelse.

⁶ Egentlig burde man skrive Ohms lov som $I = GU$ hvor $G = 1/R$ er ledningsevnen af det materiale som lederen er fremstillet af. Begrundelsen er, at det er spændingen U , der gennem ledningsevnen G bringer strømmen I til at flyde.

⁷ http://en.wikipedia.org/wiki/Electrical_resistivity_and_conductivity.

1.3.1 Modstandstyper

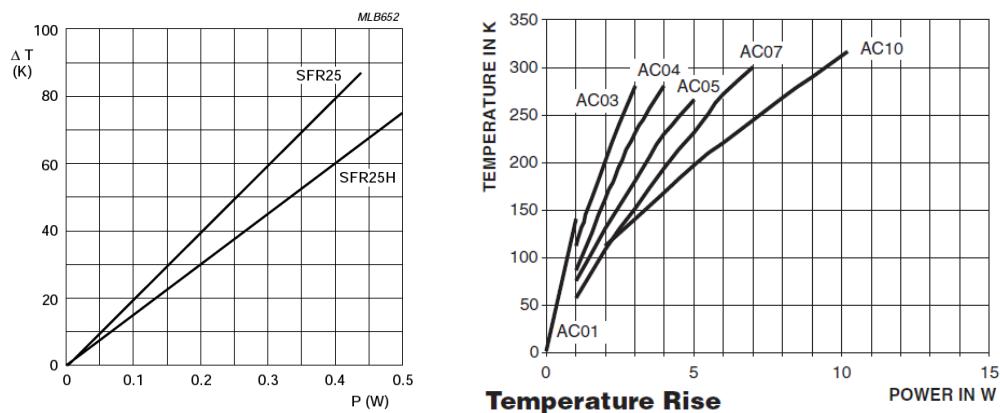
Normalt er en modstand en komponent man køber til en specifieret værdi og ofte bestilles de i ruller med nogle tusinde ens værdier for automatisk bestykning af printkortene. Det er sjældent at man beregner en modstandsværdi, dog med kobberbanen på et printkort som en undtagelse.



Figur 5 – SMD metalfilm indtil 0,5 W (Yaego). SFR25H metalfilm med trådender på 0,5 W (Vishay). AC01 trådviklet modstand på 1 W (Vishay). Modstand i TO-220 hus på 20 W med køleplade (MQEC). Trådviklede modstande i cement for 10 til 100 W med køleplade (Arcol).

Modstandsaterialet kan være en tynd metalfilm der dampes på en keramisk basis hvilket er det almindelige både ved overflademontage og for modstande med trådender. Modstandsaterialet kan også være en tynd tråd der er viklet omkring et keramisk materiale og forseglet med lak eller cement. For at undgå selvinduktion benyttes en bifilar vikling hvor strømmen føres i hver sin retning igennem to parallelle ledere for at reducere magnetfeltet (se side 31), men trods dette tiltag er trådviklede modstande ikke egnede til højfrekvens.

De mest populære modstande er SMD (*surface-mounted devices*) der måler nogle få millimeter på hver led og er beregnet for montage på overfladen af et printkort hvor de fastholdes af loddetinnet. De bør ikke belastes med mere end maksimalt 0,1 W da varmen primært afledes til printkortet der bliver skævvredet og den vil opvarme de nærmeste komponenter. For en større effektafsættelse benyttes modstande med trådender som kan afgive en del af effekten til omgivelserne gennem varmestråling og konvektion og ved at bruge afstandsstag opnås en afstand til printkortet som mindsker varmepåvirkningen. For endnu større effektafsættelse kan modstanden skrues fast til et køleprofil som så står for afgivelsen af den afsatte effekt.



Figur 6 – SFR25 af metalfilm er 2,5 mm i diameter og 6,5 mm lang og den kan klare indtil 500 mW (BC components). Den trådviklede modstand AC01 er 4,3 mm i diameter og 11 mm lang og kan klare indtil 1 W ved 40°C omgivelsestemperatur hvor temperaturen stiger til 180°C (Vishay).

En belastning på 100 mW ved en almindelig modstand med trådender (SFR25) betyder en stigning på 20°C over omgivelsestemperaturen så modstanden vil have en driftstemperatur på omkring 50°C hvilket kan være en udfordring ved krævende analog signalbehandling.

1.3.2 Effektafsættelse

Der afsættes energi i en modstand når der løber en strøm i den over et tidsrum og det varmer modstanden op. Energien skyldes at ladningen Q tvinges igennem spændingsfaldet U over tidsrummet T , men det er mere normalt at beregne den afsatte effekt P på et givet tidspunkt,

Ladningen kan udtrykkes ved den strøm der løber i et tidsrum og energi er effekt gange tid og herfra følger formlene for en effektafsættelse i en modstand.

$$\begin{aligned} W &= QU \\ Q &= IT \\ W &= PT \end{aligned} \Rightarrow P = \frac{W}{T} = \frac{ITU}{T} = IU \Rightarrow \begin{aligned} P &= I^2 R & \text{da} & U = IR \\ P &= \frac{U^2}{R} & \text{da} & I = \frac{U}{R} \end{aligned}$$

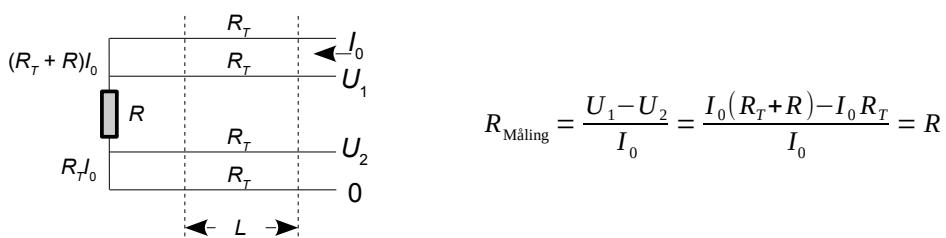
Når strømmen $I = 0,1$ A løber gennem en modstand på $50\ \Omega$ så er spændingsfaldet $U = 5$ V og der afsættes en effekt på $P = IU = (0,1\text{ A}) \cdot (5\text{ V}) = 0,5\text{ W}$.

Ofte glemmer man at tage hensyn til den afsatte energi i en komponent og det opdager man så ved at prototypen sender røgsignaler op. Modstande til overflademontage er begrænset til en afsat effekt på højest 0,1 W. For modstande med trådender (*leaded resistor*) er grænsen på 0,5 W. Et ofte overset problem er at modstandens overfladetemperatur er langt over 100°C ved den tilladte belastning og det kan generere andre komponenter i nærheden. Temperaturstigningen giver som et eksempel en mekanisk belastning af printkortet på grund af dets udvidelseskoefficient.

Det er ikke tilrådeligt at afsætte den maksimalt tilladelige effekt i henhold til databladet og det er der flere grunde til. Modstanden bliver seriøst varm når den belastes med sin maksimalt tilladelig effekt og det reducerer levetiden⁸. Indenfor visse produktkategorier kræves en såkaldt *derating* af arbejdsbetingelserne for at designet kan godkendes (fx militær elektronik), men det er generelt en sund praksis aldrig at gå til kanten af hvad en komponent specificeres til at kunne klare. Som en tommelfingerregel bør man holde effekten under 50 % af det databladet tillader.

1.3.3 Temperaturmåling med PT100

Modstanden af en metaltråd stiger med temperaturen og det kan anvendes til præcise termometre. En PT100 temperaturføler benytter ædelmetallet platin⁹ det har en modstandsværdi på $100\ \Omega$ ved 0°C og kan benyttes fra under -250°C til over 600°C .



Figur 7 – En 4-tråds måling af en lav modstandsværdi ved fx PT100 temperaturfølere.

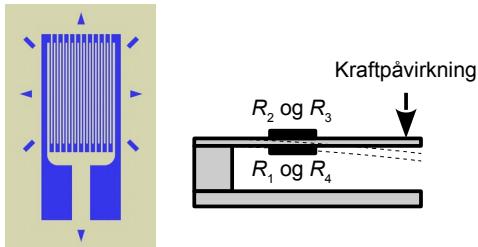
Modstandsværdien måles ved at sende en strøm igennem den og måle spændingsfaldet hvorefter Ohms lov benyttes til at beregne modstandsværdien. Tilledningens modstandsværdi R_T indgår i målingen og giver et problem ved lave modstandsværdier. Det kan løses ved at benytte en firetråds- eller Kelvinkobling hvor det yderste lederpar bærer strømmen I_0 mens midterste ledere blot benyttes til at monitorere spændingsdifferensen $U_1 - U_2$ uden at involvere en betydende strøm. Det tillader desuden en stor afstand L mellem modstand og måleapparatur uden at påvirke den opnåelige nøjagtighed. For at undgå indstråling af elektromagnetisk støj anbefales det at skærme lederene og skærmen forbindes til måleapparatets nul.

8 Søg efter Arrhenius på nettet.

http://en.wikipedia.org/wiki/Resistance_thermometer

1.3.4 Strain gauge

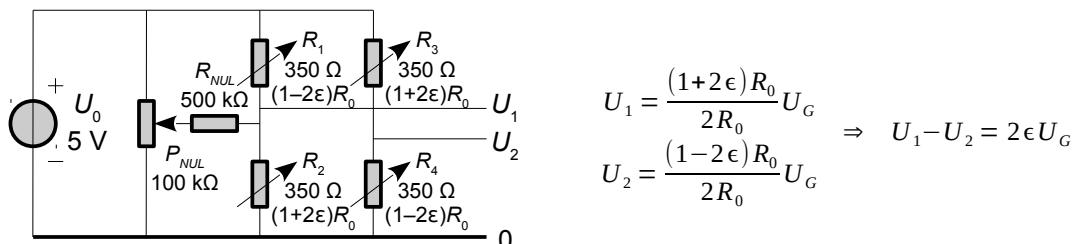
En metaltråd kan udnyttes til at måle mekanisk deformation af et emne i form af de såkaldte *strain gauge* modstande¹⁰ (udtales "stræin gaitsj"). Det kan oversættes til belastningsmåler, men der er ikke noget godt dansk navn.



$$\begin{aligned} \text{Modstand af tråd:} \\ R(\epsilon) = \rho \frac{a(1+\epsilon)}{A} \\ \text{Areal vs. volumen:} \quad A = \frac{V}{a(1+\epsilon)} \Rightarrow \text{Modstand:} \\ R(\epsilon) \approx (1+2\epsilon) R_0 \end{aligned}$$

Figur 8 – Opbygningen af en vægt med strain gauge modstande hvor længden a ændres med ϵ . Arealet reduceres A så voluminet V af tråden forbliver konstant (antallet af molekyler ændres ikke).

Modstandsværdien er funktion af trådens længde og areal og en strain gauge er ofte konstrueret ved at udfælde et tyndt metallag på en plastfolie som derefter kan limes fast til emnet. Værdien af modstanden ændres dobbelt så meget som længdeændringen så en relativ deformation med ϵ vil svare til en modstandsændring på 2ϵ i forhold til værdien R_0 uden deformation ($\epsilon = 0$) hvis ϵ er lille ($\epsilon < 0,01$). En større ændring kan give en blivende deformation af tråden, men det betyder også at modstandsændringen bliver lille når der er højst 10Ω ved en strain gauge på $1 \text{ k}\Omega$. Den lave værdi måles ved en firetrådkobling eller ved brug af en brokobling hvor to grene arrangeres så spændingen fra den ene gren stiger (herunder U_1 dannet fra R_1 og R_2) og den anden aftager (U_2 fra R_3 og R_4). Som en yderligere fordel vil de fire modstande korrigere for temperaturændring.



Figur 9 – Modstanden af en metaltråd stiger når den strækkes brøkdelen ϵ i længeretningen a og variationen forøges ved at arealet A mindskes idet trådens samlede volumen V antages konstant. Den største modstandsvariation opnås i den lodrette retning med den viste opbygning.

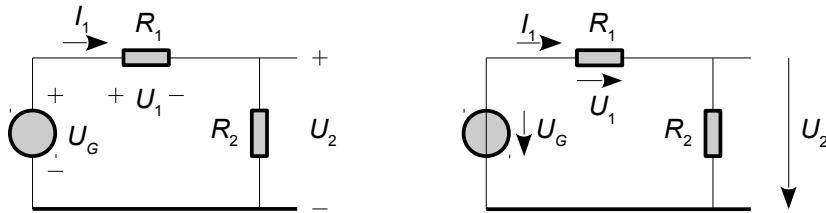
Det vil normalt være nødvendigt at kompensere for en ubalance i broen så der måles 0 V uden en ekstern påvirkning af opstillingen. Det kan opnås ved at indføre en svag strøm i den ene gren fra et potentiometer anbragt mellem 5 V og nul og med en passende modstand til knudepunktet. Hver gren har en Thevenin-modstand på 175Ω med de viste modstandsværdier så der skal bruges en strøm på $\pm 6 \mu\text{A}$ for at justere knudepunktet med $\pm 1 \text{ mV}$. Det kan opnås med et potentiometer på $100 \text{ k}\Omega$ og en seriemodstand på $500 \text{ k}\Omega$. En anden mulighed er at måle spændingen uden ekstern påvirkning med en mikroprocessor og efterfølgende trække værdien fra i software. En køkkenvægt bruger denne metode ved start samt når Tara-knappen aktiveres for at nulstille broen.

Udgangsspændingens område afhænger af den valgte effektforsyning så en nøjagtig måling vil kræve en konstant spænding der højst bør variere $\pm 5 \text{ mV}$ for at måle med $0,1 \%$ præcision. Med en effektforsyning på 5 V vil en mekanisk deformation på $\pm 1 \%$ svare til $\pm 200 \text{ mV}$ fra broen. For at holde dynamikområdet skal elektronikken kunne måle med $\pm 200 \mu\text{V}$ nøjagtighed.

¹⁰ http://en.wikipedia.org/wiki/Strain_gauge.

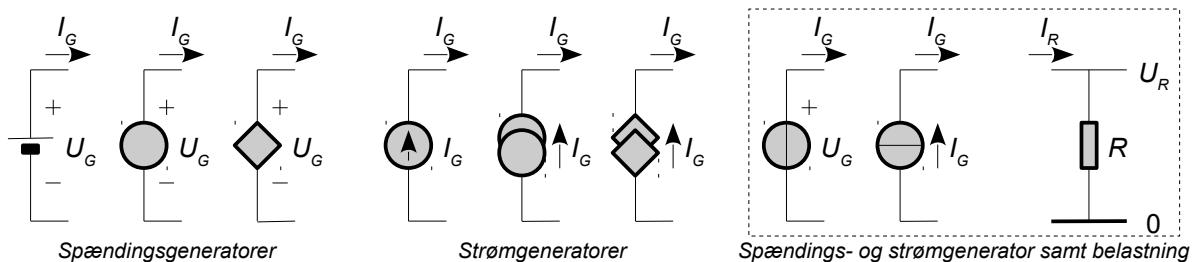
1.4 Konventioner og love

Ohms lov danner grundloven indenfor elektronikken og den medfører et behov for en definitionen af positiv spænding og positiv strøm. Strømretningen er positiv når den løber fra plus til minus. Det er imod elektronernes retning, men det har ingen praktisk betydning¹¹.



Figur 10 – Et elektrisk kredsløb med en spændingskilde U_G samt modstandene R_1 og R_2 . Positiv spænding noteres med plus/minus i den ældre notation og med pil i den internationale standard. Strøm noteres med en pil i begge standarder og med positiv retning imod elektronernes retning.

Den notation der benyttes for at angive positivt spændingsfald har der været en hel del forskellige bud på gennem tiden og herover vises de typiske. Til venstre vises den notation, som ofte bruges i europæisk og amerikansk litteratur. Spænding vises med plus og minus ved de refererede punkter, og strøm vises med en pil nær ved den leder der bærer strømmen. Til højre vises den konvention IEC vedtaget som international standard i 1970. Det er min foretrukne notation, men jeg har valgt at benytte en kompromisnotation i denne bog for ikke at skabe forvirring om at en pil betyder strøm. Der findes endnu en notation hvor spændingen angives med spændingspilene i modsat retning. Den blev jeg selv undervist i omkring 1980, men jeg er ude af stand til at indse dens fordele.



Figur 11 – Et udvalg af symboler for spændingsgeneratorer og strømgeneratorer samt den symbolik der vil blive benyttet i dette dokument.

Symbolerne for strøm- og spændingskilder er der mange bud på med et udvalg vist herover. Amerikanske bøger har batterisymbolet som DC spænding, cirklen som AC spænding og en firkant på spidsen som en styret generator. Det sidste vil sige at spændingen er funktion af en strøm eller spænding et andet sted i kredsløbet. Tilsvarende er der flere bud på strømkilden som vist ved de tre midterste symboler. Yderst til højre vises de symboler som vil blive benyttet i denne bog.

En kilde noteres med en cirkel. Ved en spændingskilde forbinder stregen de to poler og derved indikeres en lav indre modstand (0Ω). Strømkilden noteres ligeledes med en cirkel, men stregen indikerer en afbrydelse mellem de to poler og indikerer en høj indre modstand ($\infty \Omega$). Styrede kilder angives med uændret symbol og relationen vises i betegnelsen ved generatoren. En spænding der ikke er noteret med plus og minus regnes positiv overfor et fælles nul-potentiale her angivet som "0" hvor andre betegnelser er nul, stel, GND, ground og chassis. Strømretningen er negativ hvis strømmen løber fra minus til plus så afgiven effekt bliver negativ (fx afladning af et batteri) og afsat effekt bliver positiv (fx opvarmning af en modstand).

¹¹ Enkelte lærebøger har forsøgt at rette op på den fejl, men det har ikke slået an – det er svært at ændre på inddarbejdede vaner.

1.4.1 Kirchhoffs love

En god udvidelse af Ohms lov er Kirchhoffs love som jeg kun kan opfordre dig til at blive fortrolig med for det er et særdeles effektivt værktøj ved kredsløbsanalyse. En analyse skal lede til løsning af et problem, men det skal ikke føre til en opgave i sig selv for så er opmærksomheden flyttet fra det at løse et problem til det at anvende værktøjet. Med Kirchhoffs love skabes en fast procedure for at angribe et problem og med lidt øvelse vil analysen føre sikkert i mål.

Ved en maske, lække eller sløje følger man en tur rundt i kredsløbet langs med de serie forbundne komponenter indtil man ender ved udgangspunktet. Det kan sammenlignes med en gåtur der starter fra hoveddøren og slutter samme sted så en strækning op ad bakke må modsvares af en lige så lang strækning ned ad bakke for ellers vil turen ikke afsluttes i samme niveau som den begyndte. I det elektriske kredsløb svarer højdeforskel til spænding og turen rundt skal give resultatet nul. Det er Kirchhoffs spændingslov også kaldet maskelov (KVL, *Kirchhoffs voltage law*).

$$\text{Kirchhoffs maskelov: } \sum_{n=0}^N U_n = 0$$

En knude er et sted hvor flere komponenter mødes og den strøm der flyder til en knude fra én af grenene skal modsvares af summen af strømme der løber væk i de øvrige grene. Det kan lignes med et åløb der deler sig, og reglen undgår en ophobning i knudepunktet. Ved brug af Kirchhoffs knudelov skal man først beslutte om der skal regnes positivt hen til knuden for alle grene eller væk fra knuden for alle grene og derefter opstilles summen af strømme regnet med fortegn der skal være lig med nul. Det er Kirchhoffs strømlov også kaldet knudelov (KCL, *Kirchhoffs current law*).

$$\text{Kirchhoffs knudelov: } \sum_{n=0}^N I_n = 0$$

Opstillingen side 15 øverst danner en spændingsdeler og opgaven er at finde udgangsspændingen. Herunder vises brugen af Ohms lov for kredsløbet samt begge Kirchhoffs love. Kredsløbet er så enkelt at Ohms lov leder direkte til målet, men ved mere komplicerede kredsløb kan Kirchhoffs love være den eneste mulige vej frem.

Ohms lov udnytter at modstandene R_1 og R_2 er i serie og spændingen U_G driver strømmen I_1 gennem dem som givet ved Ohms lov. Spændingen over R_2 bestemmes også fra Ohms lov.

$$I_1 = \frac{U_G}{R_1 + R_2} \Rightarrow U_2 = R_2 I_1 = R_2 \frac{U_G}{R_1 + R_2} \Rightarrow U_2 = \frac{R_2}{R_1 + R_2} U_G$$

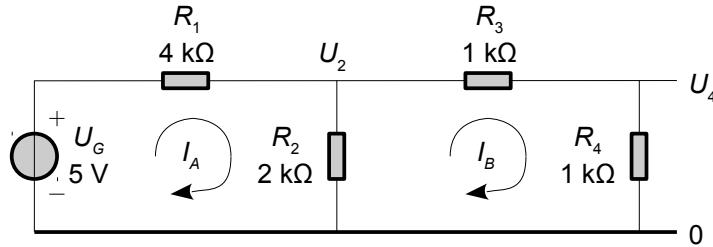
Kirchhoffs spændingslov bruges ved at vælge en positiv strømretning samt et udgangspunkt. Man går nu rundt langs med den valgte omløbsretning, startende fra det valgte udgangspunkt og indtil man er tilbage igen. Samtlige spændingsfald noteres med fortegn og for modstandene noteres spændingen efter Ohms lov. Ligningen giver strømmen ved samme udtryk som vist ovenover og udgangsspændingen findes til det samme udtryk.

$$-U_G + I_1 R_1 + I_1 R_2 = 0 \Rightarrow I_1 = \frac{U_G}{R_1 + R_2} \Rightarrow U_2 = I_1 R_2 = \frac{R_2}{R_1 + R_2} U_G$$

Kirchhoffs strømlov bruger et valg af positiv strømretning som her vises med strøm væk fra knuden. Bemærk at knudepunktets betegnelse (U_2) derved står forrest i begge led. Ligningen løses ved at samle led med fælles variabel (her U_2 og U_G) og at isolere det eftersøgte led.

$$\frac{U_2 - U_G}{R_1} + \frac{U_2 - 0}{R_2} = 0 \Rightarrow \left(\frac{1}{R_1} + \frac{1}{R_2} \right) U_2 = \frac{U_G}{R_1} \Rightarrow U_2 = \frac{U_G}{R_1 \left(\frac{1}{R_1} + \frac{1}{R_2} \right)} = \frac{U_G}{1 + \frac{R_1}{R_2}} = \frac{R_2}{R_1 + R_2} U_G$$

Uanset om man vælger maskeloven eller knudeloven så vil ligningerne for flere end tre masker eller knuder resultere i meget lange udtryk. Så er det bedre at opstille et ligningssystem i en matrix og benytte et matematisk værktøj for at løse det eller at anvende et simuleringssværktøj af SPICE typen selv om det kun giver resultatets spænding eller strøm og ikke et algebraisk udtryk.



Figur 12 – Kredsløb med to masker A og B hvor strømmene beregnes til $I_A = 1 \text{ mA}$ og $I_B = 0,5 \text{ mA}$.

Herover vises et eksempel på et kredsløb der kan repræsenterer ganske mange systemer og ikke kun et elektroniske kredsløb. Her skal strømmen I_B i modstanden R_4 bestemmes ved brug af Kirchhofs spændingslov (maskeloven, KVL).

Der opstilles to ligninger for hver af de to masker og strømmens retning er positiv med uret. For maske A løber strømmen fra maske B imod den valgte positive retning så strømmen i R_2 bliver derfor $I_A - I_B$, og noget tilsvarende gælder for strømmen i R_2 ved analysen for maske B.

$$\begin{aligned} \text{A: } & -U_G + R_1 I_A + R_2 (I_A - I_B) = 0 \Rightarrow \text{A: } -U_G + (R_1 + R_2) I_A - R_2 I_B = 0 \\ \text{B: } & R_2 (I_B - I_A) + R_3 I_B + R_4 I_B = 0 \quad \text{B: } (R_2 + R_3 + R_4) I_B - R_2 I_A = 0 \end{aligned}$$

For at finde strømmen I_B skal strømmen I_A fjernes fra ligningssystemet hvilket benytter ligningen for maske B da det er lettest at isolere I_A herfra. Derefter indsættes det fundne udtryk i ligningen for maske A der nu kun afhænger af I_B , og den ligning løses ved at dividere med R_2 .

$$\begin{aligned} \Rightarrow \text{Fra ligning B: } & I_A = \frac{R_2 + R_3 + R_4}{R_2} I_B \quad \text{Indsættes i ligning A: } R_2 I_B = -U_G + (R_1 + R_2) \frac{R_2 + R_3 + R_4}{R_2} I_B \\ \Rightarrow \text{Løses for } I_B: & I_B = \frac{-U_G}{R_2 - (R_1 + R_2) \frac{R_2 + R_3 + R_4}{R_2}} = \frac{-5 \text{ V}}{2 - (4+2) \frac{2+1+1}{2} \text{ k}\Omega} = \frac{-5 \text{ V}}{-10 \text{ k}\Omega} = 0,5 \text{ mA} \end{aligned}$$

Resultatet findes til $I_B = 0,5 \text{ mA}$ med de viste komponentværdier og det medfører $I_A = 1 \text{ mA}$ der giver de rigtige værdier af spændingerne ved kontrol.

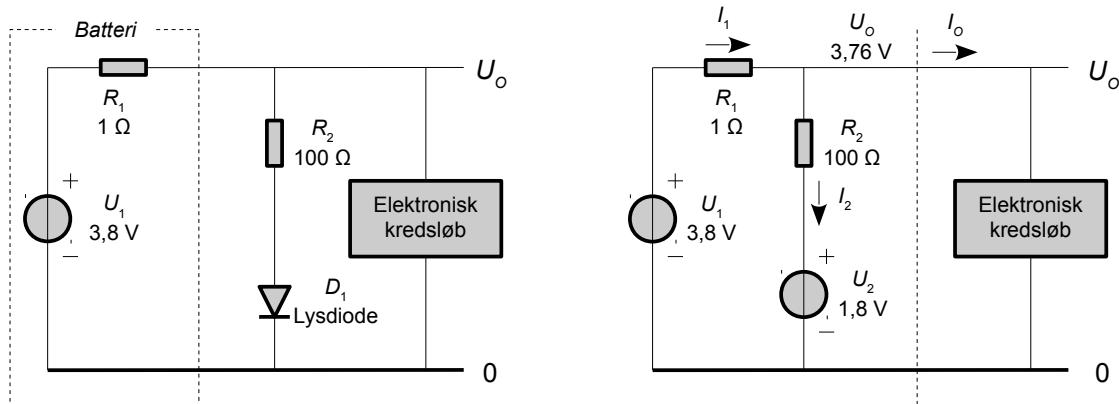
1.4.2 Superposition

En af de vigtigste regler indenfor analyse er superposition der oversat til almindelig dansk betyder at spænding og strøm kan opbygges af bidrag som ikke gensidigt påvirker hinanden. Reglen har sin bund i en antagelse om linearitet der er et bærende fundament indenfor naturvidenskab.

Linearitet betyder at hvis en påvirkning giver en vis respons så vil en fordobling af påvirkningen fordouble responsen og hvis påvirkningen reduceres til nul så vil responsen også blive nul. Det kan lyde ret trivelt for vi er så vant til antagelsen om linearitet at vi ikke skænker det en tanke, men der er ikke ret meget i den virkelige verden der er lineært. Et eksempel er addition af magnetiske felter, der gælder så godt som ubegrænset i vakuums, men ved et magnetiserbart materiale, som fx jern er lineariteten kompromitteret på grund af mætrin og remanens. Indenfor elektronik antages alle komponenter at opføre sig lineært. Man ignorerer derved at komponentens værdi er funktion af strøm og spænding, af temperaturen, af mekanisk påvirkning og muligvis af radioaktiv stråling samt at alt ændres over tid.

Superposition tillader at flere kilder beregnes en efter en og at deres bidrag adderes med hensyn til deres fortegn. For at analysere virkningen fra hver enkelt kilde skal alle andre kilder nulstilles hvilket betyder at en spændingskilde erstattes af en kortslutning og at en strømkilde erstattes af en afbrydelse.

Et eksempel på anvendelse af superposition vises for kredsløbet herunder hvor batteriet U_1 driver en lysdiode og et elektronisk kredsløb. Opgaven er at bestemme spændingen U_o til kredsløbet som følge af belastningen fra lysdioden på 20 mA og det elektroniske kredsløb på 100 mA.



Figur 13 – Et eksempel med et Li-ion batteri, der driver en lysdiode og et elektroniske kredsløb. Til højre vises en model af kredsløbet hvor dioden repræsenteres af en ideel spændingskilde.

Batteriet er Li-ion og beskrives ved spændingskilden på 3,8 V og en indre modstand på 1 Ω. Spændingen over lysdioden regnes for konstant på cirka 1,8 V når dioden lyser så i analysen repræsenteres den ved spændingskilden U_2 . Modstanden R_2 har til formål at give lysdioden en nogenlunde konstant strøm på 20 mA.

Analysen starter med den del af kredsløbet der ligger til venstre for det elektroniske kredsløb. Der er to spændingskilder (U_1 og U_2) som begge er forbundet til udgangen U_o gennem hver sin modstand (R_1 og R_2). Bidraget fra kilden U_1 beregnes ved at nulstille U_2 (kortslutning) hvorefter U_1 spændingsdeles af R_1 og R_2 med udgangen over R_2 . Bidraget fra kilden U_2 findes på samme måde ved at nulstille U_1 . Bidragene adderes og med talværdierne indsatt findes $U_T = 3,78$ V.

$$\begin{aligned} U_o(U_1)_{U_2=0} &= \frac{R_2}{R_1+R_2} U_1 & U_o = U_o(U_1)_{U_2=0} + U_o(U_2)_{U_1=0} \\ U_o(U_2)_{U_1=0} &= \frac{R_1}{R_1+R_2} U_2 & U_o = \frac{R_2}{R_1+R_2} U_1 + \frac{R_1}{R_1+R_2} U_2 \end{aligned}$$

I anden del af analysen skal udgangen leve strømmen I_o til det elektroniske kredsløb ved den stiplede linje. Ved superposition nulstilles U_1 og U_2 og belastningsstrømmen beskrives som en strømgenerator der trækker strøm ud af de parallelforbundne modstande R_1 og R_2 .

$$U_o(I_o)_{U_1=U_2=0} = -(R_1 \parallel R_2) I_o$$

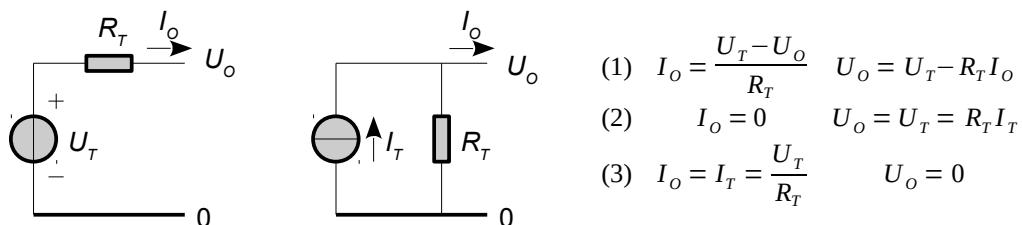
Dette udtryk kan adderes til det ovenstående resultat idet strømkilden I_o ikke indgik i analysen fra før og derfor kan regnes som en afbrudt strømkilde ved de to første beregningstrin.

$$\begin{aligned} U_o &= U_o(U_1)_{U_2=0} + U_o(U_2)_{U_1=0} + U_o(I_o)_{U_1=U_2=0} \\ U_o &= \frac{R_2}{R_1+R_2} U_1 + \frac{R_1}{R_1+R_2} U_2 - (R_1 \parallel R_2) I_o \\ &\text{Fast spændingsværdi} && \text{Belastning} \end{aligned}$$

Første led giver spændingen på udgangen når der ikke trækkes strøm fra opstillingen ($I_o = 0$) og herfra trækkes en spændingsværdi proportionel med belastningsstrømmen. Værdien heraf er i det analyserede kredsløb givet ved parallelforbindelsen af modstandene R_1 og R_2 , hvilket er en ret typisk situation. Det første led er tomgangsspændingen (*open-circuit voltage*) for det ubelastede kredsløb bestående af batteri, modstande og lysdiode. Det andet led er spændingen over den indre modstand som også kaldes for et Thevenin ækvivalent for kredsløbet.

1.5 Thevenin og Norton

Det gælder generelt for et lineært, elektrisk kredsløb, at det kan beskrives ved parametrene U_T for tomgangsspændingen og R_T for den indre modstand. Relationen skyldes oprindelig Hermann von Helmholtz (1853), men det betegnes et Thevenin ækvivalent efter Leon Charles Thevenin, der genopdagede relationen i 1883. Et Thevenin ækvivalent er således en spændingskilde U_T i serie med en modstand R_T , der ofte kaldes for den indre modstand. Det en populær model for sensorer, batterier, elektroniske reguleringer, samt for småsignal- og effektforstærkere.

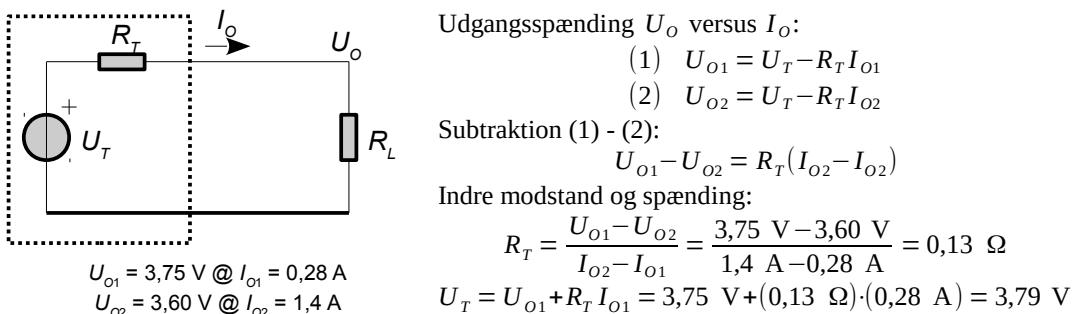


Figur 14 – To modeller for ethvert lineært elektrisk kredsløb: Et Thevenin ækvivalent til venstre og det tilsvarende Norton ækvivalent til højre. Ligning (1) giver udgangsstrømmen og -spændingen i det generelle tilfælde. Ligning (2) giver udgangsspændingen når udgangen er afbrudt og der ikke løber en strøm. Ligning (3) giver udgangsstrømmen når udgangen er kortsluttet.

I eksemplet side 18 var tomgangsspændingen $U_T = 3,78 \text{ V}$ (Thevenin spændingen) og den indre modstand var $R_T = 0,99 \Omega$ (Thevenin modstanden). Ved en strøm på $I_o = 100 \text{ mA}$ til det elektroniske kredsløb falder spændingen med $R_T I_o = 99 \text{ mV}$ til en resulterende værdi på $U_o = 3,68 \text{ V}$.

Det er ikke muligt at måle om strømmen skyldes en spændingskilde i serie med en modstand eller en strømkilde i parallel med den samme modstandsværdi. Det fik i 1926 Hans Ferdinand Mayer og Edward Lawry Norton til at foreslå en variant, der kaldes Mayer-Norton ækvivalentet, eller blot et Norton ækvivalent. Det er uden betydning for analysen om der benyttes et Thevenin eller Norton ækvivalent, så man står sig ved at vælge det, der forenkler analysearbejdet.

Kredsløbets indre modstand, R_T (Thevenin-modstanden) kan beregnes under forudsætning af linearitet. Herunder vises proceduren med et Li-ion batteri som taleksempel¹². Det er dog sjældent at et kredsløb er lineært, så man bør beregne værdien ved den aktuelle strømstyrke¹³.



Figur 15 – Den indre modstand R_T og for et Li-ion batteri kan bestemmes ved at foretage to målinger af udgangsspændingen U_o ved to forskellige belastningsstrømme I_o .

Først udnyttes ligningen for kredsløbets udgangsspænding U_o , der kan måles med et voltmeter, som funktion af udgangsstrømmen I_o , der kan beregnes fra Ohms lov, når man kender værdien af belastningsmodstanden R_L samt udgangsspændingen. Der opskrives to ligninger, en for hver af de to målinger. Derefter trækkes de to ligninger fra hinanden, hvilket eliminerer U_T og tillader beregning af den indre modstand R_T . Herefter kan U_T beregnes fra fx U_{o1} (eller måles).

12 Type 18650 på 2,8 Ah fra http://www.houseofbatteries.com/custom-batteries-custom-lithium-polymer-c-1_67-l-en.html.

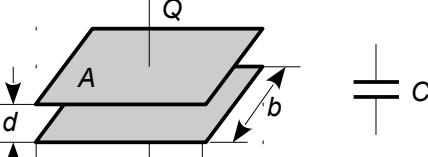
13 Beregningen giver $R_T = 0,2 \Omega$ ved en strømstyrke omkring 2 A.

1.6 Kondensator

Med en elektrisk spændingsforskæl mellem to ledere dannes der et elektrisk felt mellem dem. For at føre en ladning fra den ene ledere til den anden kræves tilførsel af energi, så der kan opbagres potentiel energi i et elektrostatisk felt. Det er muligt at trække energien ud igen, ved at danne en forbindelse fra den ene plade til den anden, så de to ledere opfører sig lidt som et batteri. Det kaldes for en kondensator (*capacitor*), og er ligesom modstanden en meget vigtig komponent.

For at optimere funktionen har en kondensator to ledere med stort overfladeareal A i kort afstand d . Med en ladning Q tilført den ene plade skubber det elektriske felt til ladningen på den anden plade, så den får lige så stor værdi, men med modsat fortegn. Der flyder strøm til den ene plade under opbygningen af det elektriske felt, mens en lige så stor strøm løber ud fra den anden plade, for at danne ladningen $-Q$. Der løber derfor en strøm i en kondensator under opladning.

Der er en proportionalitet mellem ladningen og spændingen, $Q = CU$, og proportionaliteten kaldes for kondensatorens kapacitet C , der er et mål for hvor meget spændingen stiger når der tilføres en elektrisk ladning. Enheden er coulomb per volt, som har fået symbolet farad ($F = C/V = As/V$).



$$C = \frac{\epsilon_r \epsilon_0 A}{d} \quad A = ab \quad d \ll a \quad d \ll b$$

$$\epsilon_0 = 8,854 \cdot 10^{-12} \text{ F/m}$$

$$C_{\text{Serie}} = \frac{\epsilon_r \epsilon_0 A}{d_1 + d_2} = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2}}$$

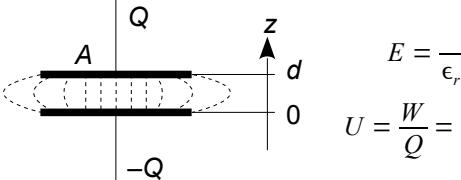
$$C_{\text{Parallel}} = \frac{\epsilon_r \epsilon_0 (A_1 + A_2)}{d} = C_1 + C_2$$

$\epsilon_r =$	1	Luft
$\epsilon_r =$	2,1	Teflon
$\epsilon_r =$	2...5	Plast
$\epsilon_r =$	3,9	SiO_2
$\epsilon_r =$	4,7	FR4
$\epsilon_r =$	5...10	Glas
$\epsilon_r =$	11,7	Si
$\epsilon_r =$	80	Vand

Figur 16 – Symbolet for en kondensator repræsenterer opbygningen med de to parallelle streger.

Et Europakort af glasfiber FR4 måler 160 mm gange 100 mm så arealet er $A = 16 \cdot 10^{-3} \text{ m}^2$ og ved tykkelsen $d = 1,6 \text{ mm}$ bliver kapaciteten $C \approx 400 \text{ pF}$ med kobberplaner på begge sider og $\epsilon_r = 4,7$.

Den elektriske feltstyrke E mellem pladerne beskrives ved ladningen per fladeelement Q/A . Med et andet materiale end vakuum imellem de to plader medtages egenskaberne af materialet, som den dimensionsløse materialekonstant ϵ_r , den relative dielektriske konstant eller permittivitet¹⁴.



$$E = \frac{Q}{\epsilon_r \epsilon_0 A}$$

$$U = E \int_0^d dz = Ed = \frac{Qd}{\epsilon_r \epsilon_0 A} = \frac{Q}{C}$$

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

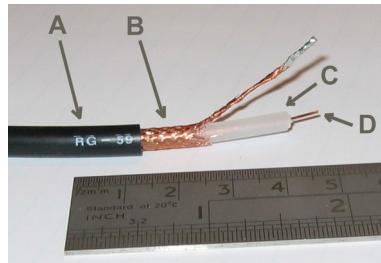
Arbejdet W ved at flytte en ladning fra den ene plade til den anden beregnes ved at integrere produktet af kraft og vejlængde. Kraften F er givet ved ladningen Q og feltstyrken E til $F = QE$. For strækningen dz er energien $W = Fdz = QE dz$. Feltstyrken antages konstant, hvorfed der ses bort fra spredningen af det elektriske felt ved kanten af pladerne, så integration fra den ene plade ($z = 0$) til den anden ($z = d$) giver afstanden d , og dermed bestemmes kapaciteten C .

En serieforbindelse af kondensatorer med samme areal, men forskellige afstande d_1 og d_2 svarer til at afstandene adderes. En parallellforbindelse af to kondensatorer med samme afstand d , men med forskellige arealer A_1 og A_2 svarer til at arealerne adderes. Det giver de to formler for serie- og parallellforbindelse af kondensatorer, der er "omvendt" i forhold til formlerne for to modstande.

14 http://en.wikipedia.org/wiki/Relative_permittivity
<http://www.rfcafe.com/references/electrical/dielectric-constants-strengths.htm>
<http://hyperphysics.phy-astr.gsu.edu/hbase/tables/diel.html>

1.6.1 Skærmkabel

Et populært kabel består af en elektrisk leder omgivet af en koncentrisk leder, kaldet en skærm (*shield*). Et skærmkabel bruges til at overføre svage signaler fra en mikrofon, samt højfrekvente signaler for TV og internet. Fordelen er at det elektriske felt er begrænset til det indre af kablet, med den ydre leder som fysisk afskærmning overfor et forstyrrende, eksternt elektrisk felt.



$$\frac{C}{a} = \frac{2\pi\epsilon_r\epsilon_0}{\ln(\frac{R_2}{R_1})}$$

$$\frac{R_2}{R_1} = 3,2 \Rightarrow \frac{C}{a} = 100 \text{ pF/m}$$

$$\epsilon_r = 2,1$$

Figur 17 – Skærmkablet RG-59 har den karakteristiske impedans $Z_0 = 75 \Omega$ ($R_2/R_1 \approx 6$), og RG-58 har $Z_0 = 50 \Omega$ ($R_2/R_1 \approx 3,2$). A er en isolerende kappe, B er skærmens, der danner den ydre leder, C er isolationsmateriale og D er den indre leder. I formlen er R_1 trådens ydre radius og R_2 er skærmens indre radius (fra Wikipedia).

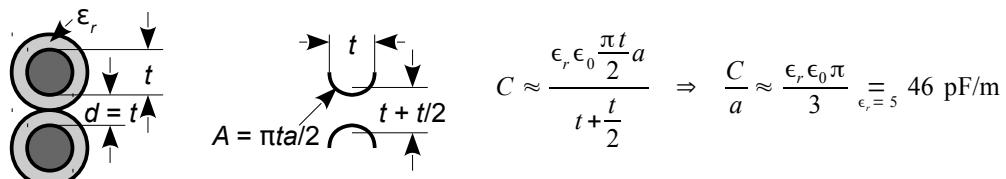
Den elektriske ladning er ens for skærmens og den indre leder, men arealet af de to leders overflader er forskellige. For at omgå det problem opdeles området mellem den indre og ydre leder i koncentriske rør med radius r og $r + dr$ for hvert af rørene. Kapaciteten C_r mellem to rør i afstanden r og $r + dr$ fra centrum beregnes med arealet A_r , der er givet ved omkredsen $2\pi r$, som ganges med længden af kablet a . Afstanden mellem rørene er $d = dr$. Summen af de mange kapaciteter beregnes efter formlen for serieforbundne kondensatorer, der kan udtrykkes mere hensigtsmæssigt ved at benytte det reciprokke af kapaciteterne.

$$C_r = \frac{\epsilon_r \epsilon_0 A}{d} = \frac{\epsilon_r \epsilon_0 2\pi r a}{dr} \Rightarrow \frac{1}{C} = \sum_{r=R_1}^{R_2} \frac{1}{C_r} = \int_{R_1}^{R_2} \frac{dr}{\epsilon_r \epsilon_0 2\pi r a} = \frac{1}{\epsilon_r \epsilon_0 2\pi a} \int_{R_1}^{R_2} \frac{dr}{r} = \frac{1}{\epsilon_r \epsilon_0 2\pi a} \ln\left(\frac{R_2}{R_1}\right)$$

1.6.2 To parallele ledere

Et kabel med to eller flere parallelle ledere benyttes flittigt til fx at overføre energi fra lysnettet, fra en effektforstærker til en højttaler og for transmission af digitale data som fx USB. For det typiske kabel med to ledere kan kapaciteten beregnes til cirka 50 pF per meter længde så 10 m kabel vil give en kapacitet på 500 pF. Det populære fladkabel (*ribbon cable*) består af mange parallelle ledere og benyttes ofte med hver anden leder til nul så de resterende ledere har en kapacitet på cirka 100 pF/m overfor nul.

Lederene tilnærmes ved to parallelle planer, hvis overfladeareal sættes til det halve af ledernes omkreds: $\pi t/2$ gange ledernes længde a . Afstanden mellem de to arealer sættes til $d = t$, plus et lille bidrag som følge af fladernes krumning (antaget til en fjerdedel af tykkelsen af lederen). Det antages at feltet hovedsageligt løber indenfor isolationen med den effektive værdi på $\epsilon_r = 5$. Ved at dividere med a opnås beregning af kapaciteten per længde af kablet.

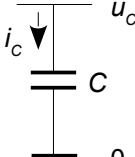


$$C \approx \frac{\epsilon_r \epsilon_0 \frac{\pi t}{2} a}{t + \frac{t}{2}} \Rightarrow \frac{C}{a} \approx \frac{\epsilon_r \epsilon_0 \pi}{3} \underset{\epsilon_r=5}{=} 46 \text{ pF/m}$$

Bidraget fra randfeltet er ignoreret så den beregnede værdi må forventes at være i underkanten. Til gengæld vil noget af magnetfeltet passere gennem luft og ikke gennem plastmaterialet og det vil reducere værdien. Den beregnede værdi er afrundet i teksten til $C/a = 50 \text{ pF/m}$.

1.6.3 Kondensatorligningen

Der er nu materiale nok til at udlede en overordentlig praktisk formel for kondensatorer. Ladningen Q er som tidligere vist givet ved strømmen I , der løber i tidsrummet T , som $Q = IT$. Men der er også en relation med kapaciteten som $Q = CU$. De to udtryk beskriver begge kondensatorens ladning og kan derfor samles til et fælles udtryk: kondensatorligningen.

	Kondensatorligning: $Q = C U = I T$	$\Delta Q = C \Delta U = I \Delta T$ $dq = C du = i dt$
i_c C 0	... på differensform: ... på differentialform:	
	Strømmen $i(t)$: $i = \frac{dq}{dt} = C \frac{du}{dt}$	
	Spændingen $u(t)$: $du = \frac{i dt}{C} \Rightarrow u = \frac{1}{C} \int i dt + U_0$	

Figur 18 – Ladningen på en kondensator er givet ved potentiølet og den tid en strøm har løbet.

Næst efter elektronikkens *grundlov*, som Ohms lov udgør, er det her *formlen for kondensatorer* for den er rigtig god ved opstillingen hvor en kondensator skal op- eller aflades over et tidsrum.

En meget stor kondensator (SuperCap) på $C = 10 \text{ F}$ er opladt til $U = 3 \text{ V}$ og indeholder derfor en ladning på $Q = CU = (10 \text{ F})(3 \text{ V}) = 30 \text{ C}$. Hvis kondensatoren aflades med en konstant strøm på $I = 10 \text{ mA}$ vil den være tømt efter $T = Q/I = (30 \text{ C})/(10 \text{ mA}) = 3000 \text{ s}$ så det varer lidt under en time.

Kondensatorens ladning ændres under op- eller afladning når en strøm I løber i tidsrummet T så det er ganske praktisk at beskrive kondensatoren ved det tidsinterval ΔT hvori strømmen I løber og ændringerne i ladning og spænding betegnes tilsvarende som henholdsvis ΔQ og ΔU . Der er ingen grundlæggende forskel fra den første ligning; det er mere et udtryk for konventioner og et ønske om at tydeliggøre at det kun er kapaciteten C og strømmen I , der regnes konstante.

En kondensator på $C = 1000 \mu\text{F}$ benyttes som energireservoir i en effektforsyning. Den skal kunne afgive en strøm på $I = 100 \text{ mA}$ over et tidsrum på $\Delta T = 10 \text{ ms}$. Spændingen over kondensatoren reduceres med $\Delta U = I \Delta T/C = (100 \text{ mA})(10 \text{ ms})/(1000 \mu\text{F}) = 1 \text{ V}$.

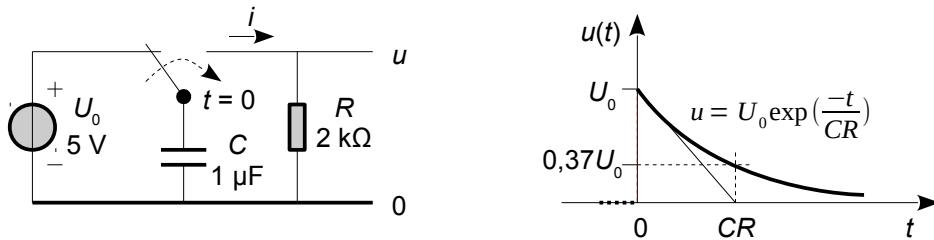
Ved analytisk brug erstattes differenserne med differentiale, så i stedet for ΔQ skrives der dq og tilsvarende for spænding og tid. Et lille bogstav repræsenterer en tidsvarierende størrelse, så ladningen er funktion af tiden $q = q(t)$, og differentialkvotienten dq/dt angiver den hastighed ændringen i q forløber med. Med den notation kan kondensatorligningen skrives med ladningen q , spændingen u og strømmen i som tidsafhængige, mens kapaciteten C er den eneste konstant. En omrøkering giver to relationer mellem henholdsvis ladning og strøm samt mellem spænding og strøm. De indeholder samme information, nemlig at den hastighed som ladning eller spænding ændres med er givet ved strømmen; eneste forskel er, at kondensatorens kapacitet C indgår i den ene ligning, men da værdien er konstant ændrer det ikke på konklusionen.

Spændingen på en kondensator kan bestemmes fra differentialet du ved integration. Derved findes en vigtig ligning for kondensatoren, der viser hvordan spændingen u over kondensatoren ændres som funktion af strømmen i , der løber igennem kondensatoren. Det kan benyttes ved kredsløbsanalyse, hvor kondensatoren behandles på samme måde som ved en modstand ($u = Ri$), blot benyttes udtrykket i illustrationen og det resulterer i en førstegrads differentialligning. Ligningen viser kun hvor meget spændingen bliver ændret under forløbet, og ikke giver den absolute værdi af spændingen, så det er almindeligt at indføre en startværdi U_0 , som bestemmes ved den aktuelle applikation. Hvis kondensatoren er tømt for ladning inden analysen startes er $U_0 = 0$.

Dette skal illustreres med nogle eksempler.

1.6.4 Afladning af kondensator

Herunder holdes kondensator C opladt på spændingen U_0 indtil tidspunktet $t = 0$ hvor omskifteren flyttes til modstanden R og en afladning påbegyndes. Det kredsløb har mange anvendelser og vil dukke op flere gange i de følgende kapitler. Det giver en eksponentiel afladning, hvor spændingen aftager fra startværdien U_0 til klokken $t = 0$, og derefter glider asymptotisk mod nul for $t > 0$.



Figur 19 – En kondensator holdes opladt fra en spændingskilde og når omskifteren flyttes over til modstanden vil kondensatoren aflades asymptotisk mod nul med en hastighed, der afhænger af produktet af kondensatorens og modstandens værdier.

Produktet CR kaldes for tidskonstanten og er en karakteristisk størrelse for kredsløbet, som her er på $CR = 2$ ms. Det betyder at i løbet af tidsrummet fra nul til $t = CR$ mindskes spændingen fra U_0 til $\exp(-1) U_0$ hvilket er $0,37 \cdot U_0$, efter to tidskonstanter er den mindsket til $0,135 \cdot U_0$ og så fremdeles. I teorien når spændingen aldrig nul, men den kommer nærmere og nærmere, og er under 1 % af det oprindelige niveau efter $t = 4,6 CR$, så man kan normalt regne en kondensator for helt afladet efter fem til syv gange tidskonstanten.

Spændingen over kondensatoren er givet af kondensatorligningen (se side 22), hvor der ikke er behov for integrationens konstant. Strømmen i er givet ved spændingen u over kondensatoren delt med modstanden R , og da strømmen flyder ud af kondensatoren er fortegnet negativt; der afgives energi.

$$u = \frac{1}{C} \int i dt \Rightarrow u = -\frac{1}{CR} \int u dt$$

Spændingen u optræder både på venstre og højre side og det er et klassisk problem indenfor matematikken. Løsningen til differentialligninger er ofte en eksponentialfunktion, i det mindste indenfor elektronikken, så løsningen findes ved at gætte på et funktionsudtryk baseret på en eksponentialfunktion, og så prøve om der er ramt rigtigt. Det kan lyde lidt løst, men er en ofte anvendt metode. Konstanterne A og B skal fastlægges, og hvis det kan lade sig gøre er det en verificering af gættet på løsningen, ellers må man prøve med et modifieret funktionsudtryk. Der findes en generel løsning til ligningen, men jeg foretrækker selv denne tilgang til opgaven.

$$u = A \exp(Bt) \Rightarrow \int u dt = \int A \exp(Bt) dt = \frac{A}{B} \exp(Bt)$$

De to udtryk indsættes i ligningen for u og det indses at ligningen kun gælder hvis B sammen med ledet $-CR$ giver et ét-tal, hvorved konstanten B er defineret.

$$A \exp(Bt) = -\frac{1}{CR} \frac{A}{B} \exp(Bt) \Rightarrow B = -1/CR \Rightarrow u = A \exp\left(\frac{-t}{CR}\right)$$

Til tiden $t = 0$ skal kondensatorens spænding være $u(0) = U_0$ hvilket giver konstanten A . Ved at løse ligningen for t kan man bestemme tiden for afladning til en given værdi. Som eksempel er der brug for tiden $t = 4,6 \cdot CR$ for afladning til 1 % af startværdien.

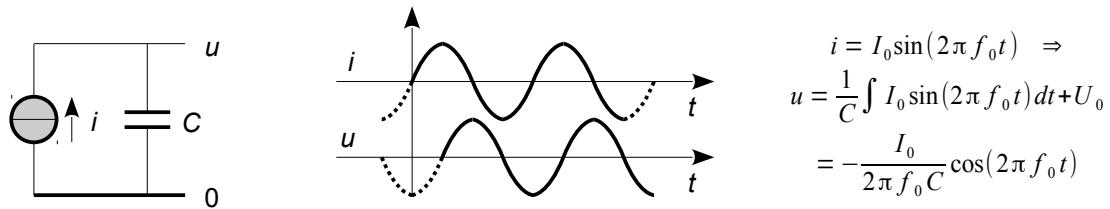
$$u(0) = U_0 \Rightarrow u = U_0 \exp\left(\frac{-t}{CR}\right) \Rightarrow t_{u(t)/U_0} = -CR \ln\left(\frac{u(t)}{U_0}\right)$$

1.6.5 Oscillerende strøm

Når en strøm føres til en kondensator så lades den op, og når strømmen er negativ aflades den, som vist i eksemplet ovenfor. Hvis strømmen er en svingning vil spændingen over kondensatoren også blive en svingning, men spændingen er ikke i fase med strømmen. Det er en sjælden gæst i moderne elektronik at integrere et analogt signal; dog foretager mange kredsløb tilnærmedesvist en integration, fx filterkredsløb, som beskrives i tredje og fjerde kapitel, og ved måling af signalet fra sensorer er integration ofte påkrævet^{15,16}.

Herunder analyseres hvad der sker ved integration af et harmonisk signal

En strømgenerator giver en harmonisk svingning (sinus) med amplituden I_0 og frekvensen f_0 . Det antages at situationen er "stationær" hvilket blot betyder at der ikke optræder andre forløb end det undersøgte. Integrationens konstant U_0 antages derfor til værdien nul ($U_0 = 0$). Resultatet er at kondensatorens spænding varierer med den samme frekvens som strømmen, men fasen er 90° bagud, og amplituden aftager med stigende værdi af oscillationens frekvens f_0 . Det er af stor betydning for det kommende kapitel, hvor det leder til definitionen af begrebet impedans.



Figur 20 – En strømkilde driver en kondensator med en sinus-svingning. Spændingen over kondensatoren får samme form, men fasen er 90° bagud og amplituden afhænger af frekvensen.

1.6.6 Genopladeligt batteri

Kondensatorligningen er også brugbar ved batterier, selv om ladningen her er kemisk bundet. Et batteris ladning måles almindeligvis i ampere timer (Ah), der dog grundlæggende er det samme som coulomb (C). Da der er 3600 sekunder i hver time er 1 Ah lig med 3600 As, som igen er lig med 3600 C.

Et batteri oplades med en konstant strøm på 0,1 A til ladningen 0,5 Ah er opnået, og opladningen behøver en tid på: $T = Q/I = (0,5 \text{ Ah})/(0,1 \text{ A}) = 5 \text{ h} = 18\,000 \text{ s}$.

Ikke alle batterityper tåler fuld afladning. Et Li-Ion batteri er fuldt opladt ved 4,2 V og afladning til en spænding under 3 V initierer en destruktiv proces. En del batterier indeholder derfor et kredsløb, der afbryder ved underafladning, da der ellers kan opstå en ekslosion. Et blybatteri (bilbatteri) er mere robust, og tåler fuld afladning, men giver tre gange mindre ladning per vægt¹⁷.

Der er en virkningsgrad knyttet til disse sekundærbatterier, så kun 80 ... 95 % af energien efter en opladning kan tappes fra batteriet. Dertil har batterier en selvafladning, der langsomt aflader det med cirka 10 % tab per måned for Li-Ion¹⁸ så efter et års tid er to-tredjedele af ladningen gået tabt. Der mistes en tilsvarende cirka 10 % af den oprindeligt, tilrædighed værende ladning for et typisk Li-ion batteri efter cirka 250 komplette op- og afladninger¹⁹.

¹⁵ Et piezoelektrisk accelerometer vil ved praktisk brug foretage en differentiation, så en integration er krævet for at regenerere signalet. Accelerometeret mäter accelerationen af en vibration, men ved analyse af slitage på maskinlede er man ofte interesseret i hastigheden af vibrationen, og det kræver endnu en integration idet $a = dv/dt$.

¹⁶ Ved måling af lydtrykket fra en højttaler ved lave frekvenser kan man placere en mikrofon inde i højttalerens kabinet. Lydtrykket er her proportionelt med membranens amplitude x , mens lydtrykket foran højttaleren er proportionelt med membranens acceleration. Her er der brug for to gange integration, idet $a = dv/dt = d^2x/dt^2$.

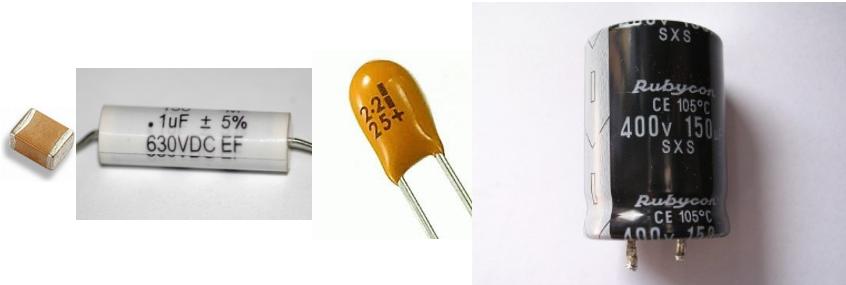
¹⁷ <http://da.wikipedia.org/wiki/Energi>.

¹⁸ http://en.wikipedia.org/wiki/Lithium-ion_battery.

¹⁹ http://batteryuniversity.com/learn/article/how_to_prolong_lithium_based_batteries.

1.6.7 Kondensatortyper

Enhver kondensator er opbygget med to plader adskilt af et dielektrikum, men der er mulighed for store variation, som den følgende oversigt over hovedtyperne viser. Forskellene ligger primært i det anvendte dielektriske materiale.



Figur 21 – Til venstre en keramisk kondensator for overflademontage (SMD) med længden 2 mm. Derefter en plastfilm kondensator på 20 mm længde. I midten en tantal elektrolyt, der ofte er under 10 mm høj og til højre en aluminium elektrolyt, der let kan være 50 mm i længde.

En keramisk kondensator er opbygget med flere lag af keramik som dielektrikum, vekslede med lag af metal for den elektriske forbindelse. Kapaciteten dækker fra 1 pF til 10 µF og komponenten er meget populær i industrien da den er velegnet til automatisk montering, tilbyder lave tab og generelt er associeret med god kvalitet for pengene²⁰.

Keramiske kondensatorer opdeles i klasser som angiver nøjagtigheden og temperaturens indflydelse. Klasse 1 (fx NP0) dækker op til cirka 100 nF med god nøjagtighed ($\pm 2\%$) og en temperaturkoefficient nær nul ($\pm 30 \text{ ppm}^{\circ}\text{C}$). Klasse 2 (fx X7R) dækker op til 10 µF, med et ringere temperaturområde og nøjagtighed end for klasse 1. Klasse 2 kondensatorerne kan lide af mikrofoni, der er en piezoelektrisk effekt hvor en mekanisk påvirkning giver en spænding over kondensatoren. De benyttes almindeligvis til opgaver som afkobling af analoge og digitale kredse og til ukritisk filtrering af signaler. Klasse 3 går op til 100 µF, men kapaciteten varierer med spændingen, den aftager over tid, tabene er høje og tolerancen er dårlig.

En filmkondensator består af metalbelagte plastfolier, der rulles sammen. De har større fysiske dimension og højere pris end keramiske kondensatorer. De findes i mange varianter, der dækker området fra 1 nF til 100 µF, og de har generelt god stabilitet og lave tab. Der er mange muligheder for plastmaterialet imellem pladerne, og de har alle specifikke anvendelsesområder²¹.

Polykarbonat har lav dielektrisk absorption (ingen "hukommelse"), så de benyttes til krævende formål. Temperaturkoefficienten er på $\pm 80 \text{ ppm}^{\circ}\text{C}$ og kan anvendes til 125°C . Polypropylen har meget lave tab, mulighed for at hele efter gennemslag ved overspænding, og de tåler høje spændinger og høj effekt. Polystyren har negativ temperaturkoefficient og kan anvendes op til 85°C .

En elektrolytkondensator benytter en oxid af aluminium eller tantal som dielektrikum mellem selve metallet og en elektrolyt. Kapaciteten rækker fra 1 µF til 10 mF, dog kan specielle typer (SuperCap) nå over 100 F. Den ene tilslutning er via metallet og den anden er via elektrolytten, som danner en slags diode overfor metallet, så den skal vendes korrekt. Temperaturkoefficienten er ringe²².

Aluminium elektrolytkondensatorer benyttes til afkobling og udglatning ved lave frekvenser, men undgås til signalbehandling på grund af en spændingsafhængig kapacitet, der giver forvrængning. De er populære i DC-DC konvertere, men højfrekvensegenskaberne er generelt ringe. Den ækvivalente seriemedstand er lav fra 100 µF og op. Lækstrømmen er høj, men aftager når komponenten har været i brug. Tolerancen er $+50/-20\%$ og kan nå $+100/-50\%$. De leveres i temperaturkategorier på 85°C og 105°C og den nedre driftstemperatur er cirka -5°C . Der tillades en svag negativ driftsspænding på nogle få volt.

Tantal elektrolytkondensatorer dækker området fra 1 µF til 100 µF og er fysisk mindre end ved tilsvarende kapacitetsværdier ved aluminium. Tolerancen er bedre end for aluminium elektrolytkondensatorer og den ækvivalente seriemedstand er lavere. Kondensatorerne tåler ikke ripplestrøm eller omvendt polarisering.

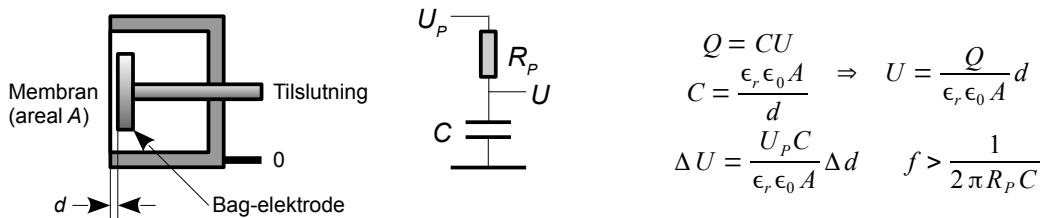
20 http://en.wikipedia.org/wiki/Ceramic_capacitor.

21 http://en.wikipedia.org/wiki/Film_capacitor.

22 http://en.wikipedia.org/wiki/Electrolytic_capacitor.

1.6.8 Kondensatormikrofon

En populær mikrofontype er baseret på kondensatorens princip. Den ene plade er en tynd metalfolie, som lydens tryksvingninger sætter i bevægelse, og den anden er en fast elektrode, der normalt forbindes til nul. Kapaciteten ændres ved trykvariationerne og det omsættes til spænding ved at give mikrofonen en fast ladning. Kapaciteten er på 10 til 50 pF, så mikrofonen kan ikke kobles til en forstærker gennem et kabel, for kablets kapacitet vil drastisk reducere den i forvejen beskedne amplitude af signalet. I stedet indbygges en buffer i mikrofonen.

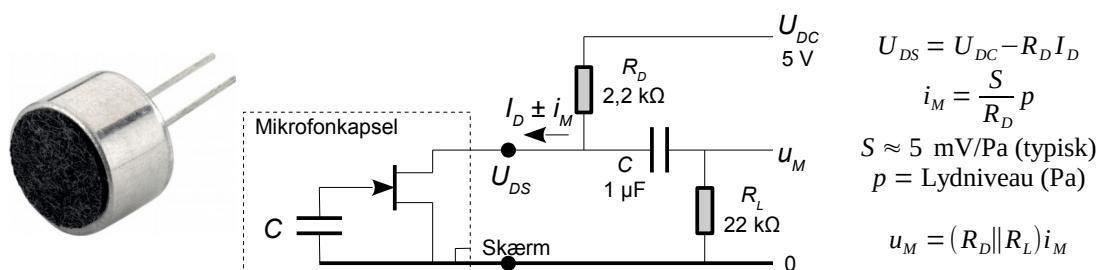


Figur 22 – En elektrostatisk mikrofon er opbygget med en tynd metalfolie som membran og en fast plade som bag-elektrode. Kondensatoren lades op fra en ekstern spænding gennem en stor modstand, og spændingen over kondensatoren ændres som funktion af kapacitetsændringerne.

Med $U_P = 200$ V, $C = 50$ pF, $\epsilon_r = 1$, $A = 10$ mm 2 og $\Delta d = 1$ μ m ($\Delta d \ll d$) bliver $\Delta U = 10$ mV.

Der er to metoder for at give mikrofonens kondensator en fast værdi af ladningen. Ved en ekstern polariseringsspænding²³ U_P oplades kondensatoren gennem modstanden R_P , og mikrofonens følsomhed bliver proportionel med spændingens værdi. Det filter, der dannes af modstand og kondensator, definerer den lave grænsefrekvens, og for at kunne gengive hele det hørbare område skal der bruges en høj værdi af modstanden (mindst 200 M Ω).

Den anden metode benytter et dielektrikum mellem membran og bag-elektrode, hvor ladningen er "indfrosset" i materialet. Det er den såkaldte elektretmikrofon, hvor en felteffekt transistor fungerer som buffer, og mikrofontypen produceres i store serier til lav pris²⁴. Egenstøjen er relativt kraftig, men der bygges elektretmikrofoner for lyd- og målebrug med langt bedre data²⁵.



Figur 23 – En elektretmikrofon er en med "fastfrosset" ladning og indbygget felteffekt transistor som buffer. Der kræves en ekstern effektforsyning på 5 ... 12 V typisk.

Den indbyggede transistor trækker en strøm på $I_D < 0,5$ mA, der leveres fra en effektforsyning U_{DC} gennem R_D . Signalets amplitude i_M moduleres ind som variationer omkring DC arbejdspunktet I_D for felteffekt-transistorens drain, og optræder derfor som en spændingsvariation u_M over R_D i parallel med belastningen R_L . Databladets "udgangsmodstand" på 2 k Ω giver følsomheden $S = 5$ mV/Pa som en typisk værdi. Signalets spændingsamplitude er $u_M = 5$ mV ved lydniveauet $p = 1$ Pa, der er relativt kraftigt²⁶ og svarer til en amplitude af signalet i transistorens drain på $i_M = 2,5$ μ A/Pa.

23 De typiske værdier er 48 V ved Phantom Powering og 200 V for professionelle målemikrofoner (fx GRAS og DPA).

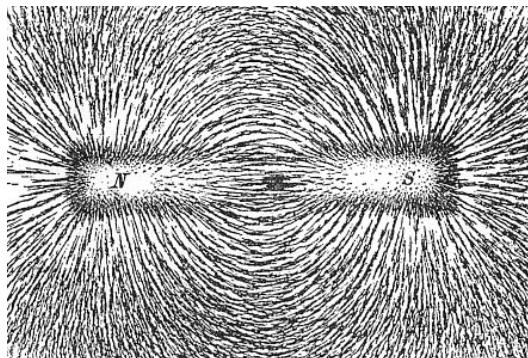
24 MCE-400 fra Monacor koster 20 kr. Følsomheden er 7,9 mV/Pa ± 3 dB ved 2,2 k Ω og den dækker 50 Hz til 16 kHz.

25 MCE-4000 fra Monacor. Følsomheden er 5,6 mV/Pa ± 1 dB og den dækker 30 Hz til 18 kHz. Phantom Powering.

26 Et RMS-niveau af lydtrykket på 1 Pa svarer til 94 dB SPL som er lydniveauet umiddelbart foran den talendes mund.

1.7 Magnetisk felt

Når elektronerne bevæger sig dannes der et magnetfelt. Det blev opdaget i 1820 af Ørsted, der så at en elektrisk strøm kunne påvirke en kompasnål. Hans offentliggørelse startede et arbejde med deltagelse af blandt andre Ampere, Faraday og Laplace samt Biot og Svart. Der kom snart mange anvendelser, fx telegrafen fra 1837. Maxwell samlede i 1861 de elektriske og magnetiske felter til en fælles beskrivelse som elektromagnetismen (*electromagnetism*)²⁷.

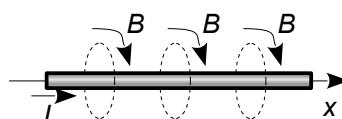


Figur 24 – Jernspåner i feltet fra en stangmagnet viser retningen af den magnetiske flux. Feltlinjerne går vedtægtsmæssigt fra nordpol til sydpol, og metalspånerne samler sig som små magneter, der klæber til hinanden med retning som små tangenter til linjerne (Wikipedia).

For en permanentmagnet løber feltlinjerne nok fra nordpol til sydpol, og de fortsætter inde i magneten fra sydpol til nordpol og danner derfor lukkede sløjfer.

I lighed med det elektriske felt betegnes et magnetfelts ved magnetfeltlinjer. Det skyldes formen af den måde jernspåner lægger sig på i et magnetfelt som det ses på billedet ovenfor. De magnetiske linjer er en lidt romantisk model, men den giver en kvantitativ forståelse blot linjerne ikke opfattes som den endegyldige sandhed. I kvantemekanikken er fotonen bærer elektromagnetismen så en ændring i feltet fra en elektrisk strøm i en ledning udbredes med lysets hastighed, men det vil jeg ikke tage hensyn til i denne fremstilling.

At magnetfeltet løber fra nordpol til sydpol er en vedtægt helt som den elektriske strøms retning. Man har simpelt hen besluttet at feltet højreskruer omkring strømmens retning (*højrehåndreglen*). Gruber man om lederen med tommelfingeren i strømmens retning vil fingrene angive magnetfeltets retning. Nok så vigtigt er at magnetfeltlinjerne danner lukkede sløjfer og at de aldrig starter eller ender ved en kilde som de elektriske feltlinjer gør. Der er ingen magnetiske monopoler²⁸.



Figur 25 – En elektrisk leder er omgivet af et magnetfelt, der er proportional med strømmens styrke og som skruer om lederen som en proptrækker (højreskruet). For en -+lang leder er den magnetiske induktion omvendt proportional med afstanden.

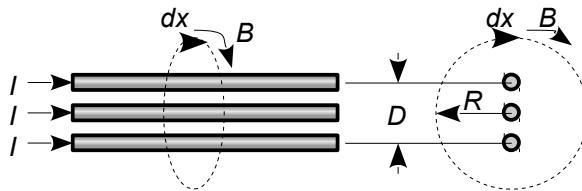
Et magnetfelts styrke betegnes ved den magnetiske induktion med symbolet B og enheden tesla ($T = \text{Wb/m}^2$). I modsætning til det elektrostatiske felt, der er et kildefelt hvor feltet starter ved en positiv ladning og løber til en negativ ladning, så består det magnetiske felt af feltlinjer i lukkede baner omkring den elektriske leder. Det kaldes for et hvirvelfelt.

²⁷ Brødrene Ramagnosi publicerede allerede i 1802 opdagelsen i italienske aviser. Senere genopdagede danskeren Ørsted i 1820 relationen mellem elektricitet og magnetisme og publiceret det gennem de videnskabelige kanaler. Fra 1861 til 1873 formulerede Maxwell elektromagnetismen, som knyttede elektricitet og magnetisme sammen som én kraft. Lysets hastighed blev samtidigt fastlagt ved to naturkonstanter; én for det elektriske system og én for det magnetiske.

²⁸ http://en.wikipedia.org/wiki/Magnetic_monopole.

1.7.1 Amperes lov

Det magnetiske induktion danner cirkulære baner omkring lederen med centrum på lederen. Den magnetiske induktions værdi i et punkt kan findes ved Amperes lov, der angiver den magnetiske induktion langs med en lukket bane omkring lederen. Definitionen herunder er generel og der vil blive benyttet en simpelere form til udledningerne senere i denne bog.



$$\begin{aligned} \text{Amperes lov: } & \mu_0 \int J dS = \oint \vec{B} \cdot d\vec{s} \\ \text{Magnetisk konstant: } & \mu = 4\pi 10^{-7} \text{ H/m} \\ \text{Kabelbundt: } & \mu_0 N I = \oint B \cos(\theta) dx \\ \text{Induktion langs med} & \mu_0 N I = 2\pi R B \\ \text{en cirkulær bane om} & \\ \text{lederne med radius } R: & \\ \text{Induktion: } & B = \frac{\mu_0 N I}{2\pi R}, \quad D \ll R \end{aligned}$$

Figur 26 – Amperes lov for den magnetiske induktion B i afstanden R fra et kabelbundt med N ledere, der hver bærer strømmen I .

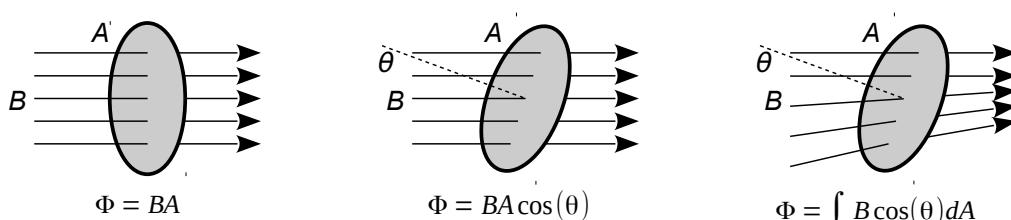
For tre ledere ($N = 3$) og en strømstyrke på $I = 1 \text{ A}$ i hver ledere bliver den magnetiske induktion $B = 6 \mu\text{T}$ i afstanden $R = 0,1 \text{ m}$. Til sammenligning er styrken af jordens magnetfelt nær overfladen 30 til 60 μT .

Første integral benytter strømtætheden J , der er strømstyrken per arealenhed (A/m^2), og dS der er en del af det areal strømmen løber igennem. Produktet $J dS$ angiver strømmen i en del af lederen, og ved integration over hele ledernes tværsnit giver det den strøm, der løber i hele lederen. Det er ganske uanset om strømtætheden J er konstant over arealet. Derfor kan første integral normalt skrives som $\mu_0 I$, hvor I er strømmen i lederen og μ_0 er den magnetiske konstant (Permeabilitet). Hvis løkken omfatter et bundt med N ledere, hver med strømmen I , så kan integralet løses til $\mu_0 NI$.

Andet integral angiver at der skal beregnes et skalært vektorprodukt af den magnetiske induktion (vektor B) langs med en lukket kurve omkring lederne. Kurven defineres ved linjeelementet dx , der også er en vektor, og som er givet ved tangenten til den valgte kurve. Hvis de to vektorer B og dx ikke peger i samme retning, men danner vinklen θ , bliver integralet B gange med cosinus til vinklen mellem vektor B og vektor dx . I stor afstand fra kabelbundtet beskrives løkken, som dx løber langs med, ved periferien af en cirkel med radius R og centrum i kabelbundtet. I det tilfælde har de to vektorer samme retning så vinklen mellem dem er nul og cosinus bliver én.

1.7.2 Magnetisk flux

Det magnetiske felts styrke beskrives ved den magnetiske flux Φ , der er antallet af magnetiske feltlinjer, som krydser gennem et areal. Den magnetiske flux Φ har enheden weber ($\text{Wb} = \text{Vs}$).

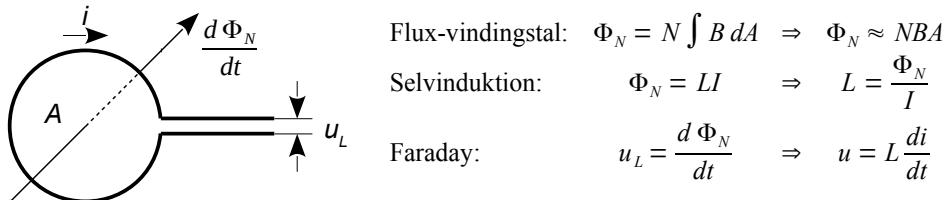


Figur 27 – Den magnetiske flux gennem et areal er givet ved den magnetiske induktion integreret over et areal under hensyn til variationen i den magnetiske induktion og vinklen med arealet.

Hvis den magnetiske induktion B er konstant over arealet A bliver den magnetiske flux lig med den magnetiske induktion gange med arealet. Hvis magnetfeltet ikke går vinkelret igennem arealet så skal der ganges med cosinus til feltets retning og normalen til arealet.

1.8 Selvinduktion

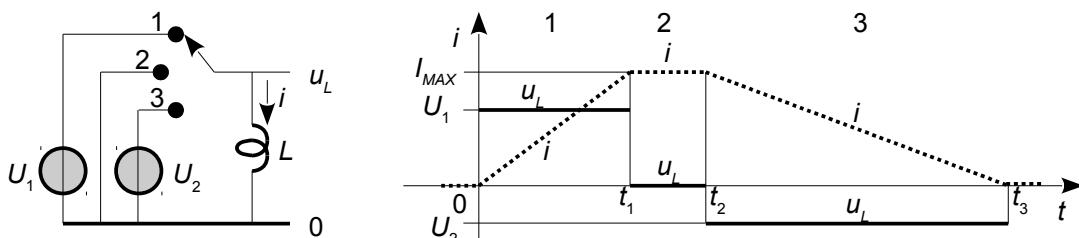
For brug ved beregning af selvinduktionen af en spole er det ikke nok at beregne den magnetiske flux gennem arealet af spolen. En spole har normalt flere vindinger og hver vinding genererer en magnetisk flux samt afgrænsler et areal som denne flux så passerer igennem. Med flere vindinger passerer den magnetiske flux derfor gennem mange indesluttede arealer. Det leder til en definition af flux-vindingstallet Φ_N hvor N er antallet af vindinger på en spole hver med arealet A .



Figur 28 – En variation i den magnetiske flux gennem et areal giver anledning til en spænding over den leder, der definerer den lukkede kurve.

For en lukket elektrisk sløjfe er der proportionalitet mellem den magnetiske flux igennem arealet A , som lederen omslutter, og strømmen i lederen I , og relationen kaldes selvinduktion (*self induction*), der har fået symbolet L for at hædre forskeren Lenz. Enheden henry ($H = Wb/A = Vs/A$) er valgt for at hædre endnu en af elektromagnetismens mange fædre. Ifølge Faradays induktionslov skal den magnetiske flux gennem et areal ændres for at inducere en elektrisk spænding over den leder der omkranser arealet. Hvis kredsen er kortsluttet genererer den inducerede spænding en strøm og det skaber et modsat rettet magnetfelt der søger at modvirke ændringen i feltet^{29,30}.

Hvis spændingen over spolen ændres vil strømmen variere som det illustreres herunder. Den hele streg er spændingen over spolen og den stippled streg er strømmen i spolen. En konstant værdi af spændingen over spolen giver en lineært varierende strøm i spolen og ved kortslutning af spolen vil strømmen være konstant hvis der kan ses bort fra tab.



Figur 29 – Strømmen i spolen er funktion af spændingen over den. Hvis spændingen er konstant vil strømmen ændres jævt (retlinet). En momentan afbrydelse af strømmen kræver derimod en uendelig stor spænding.

Spændingen over spolen u_L er givet af det ydre kredsløb og strømmen i spolen kan findes fra Faradays induktionslov. Integrationskonstanten I_0 angiver strømmen i spolen inden starten på et nyt forløb. Her er $I_0 = 0$ ved $t = 0$ samt $I_0 = I_{MAX}$ både ved $t = t_1$ og $t = t_2$.

$$di = \frac{u_L}{L} dt \Rightarrow i(t_1) = \int_0^{t_1} \frac{u_L}{L} dt + I_0 \quad u_L = U = \text{konstant} \quad i(t_1) = \int_0^{t_1} \frac{U}{L} dt + I_0 = \frac{U t_1}{L} + I_0$$

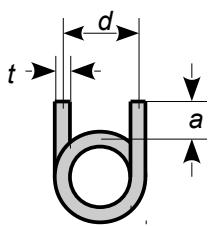
- (1) Med spolen forbundet til en spændingskilde U giver det en lineært stigende strøm i spolen.
- (2) En kortslutning af spole ($u_L = 0$) giver strømstyrken $i = I_0$ så strømmens styrke er uændret.
- (3) Med en negativ spænding over spolen ($u_L < 0$) reduceres strømmen lineært med tiden.

29 Se: http://en.wikipedia.org/wiki/Faradays_law.

30 Et populært fysikforsøg er at lade en magnet falde ned gennem et rør, hvilket tage en vis tid. Hvis røret erstattes af et elektrisk ledende rør vil tiden forlænges på grund af den inducerede strøm.

1.8.1 Beregning af selvinduktion

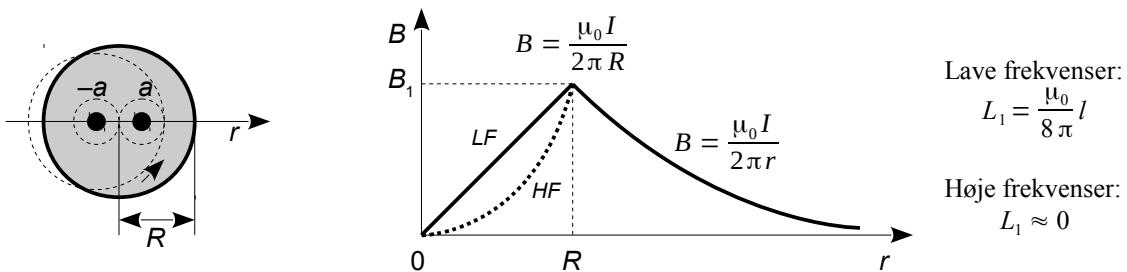
Selvinduktionen af en spole består af bidrag fra selvinduktionen af den tråd spolen vikles af (L_1), af den løkke tilslutningens tråd danner (L_2) og af spolen selv (L_x). Herunder er N antallet af vindinger, a er tilslutningens længde, d er spolens gennemsnitlige diameter og t er trådens tykkelse. Faktor K er relateret til den fysiske udformning af spolen og er i området fra 0,5 til 2 for almindelige spoler. De følgende afsnit viser beregningen af trådens og tilslutningens selvinduktioner og derefter vises hvordan beregningen af selvinduktionen L_x kan angribes for nogle enkle typer af spoler.



Samlet selvinduktion:	$L = L_1 + L_2 + L_x$
Trådens indre selvinduktion:	$L_1 = \frac{\mu_0}{8\pi} l, \quad l = 2a + \pi d$
Tilslutningens selvinduktion:	$L_2 \approx \frac{\mu_0 a}{\pi} \ln\left(\frac{2d}{t}\right)$
Spolens selvinduktion:	$L_x = K \mu_0 N^2 d$
Magnetisk konstant:	$\mu_0 = 4\pi \cdot 10^{-7} \text{ H/m}$

1.8.2 Enkeltleder

Der er et elektrisk felt inde i en ledning og det giver anledning til en selvinduktion på 50 nH for hver meter ledning. Det inkluderer både tilledningen og den leder spolen består af. Værdien gælder op til cirka 20 kHz for en leder med en trådtykkelse på 1 mm og aftager til nul ved cirka 1 MHz. Ved høje frekvenser løber strømmen stort set kun i overfladen på grund af strømfortrængning³¹.



Figur 30 – Et tværsnit igennem en leder med strømretning ud imod betragteren. Magnetfeltet er nul i centrum af den fysiske leder og vokser til maksimum overfladen. Ved høje frekvenser fortrænges strømmen fra ledernes kerne og løber i et smalt område nær ved overfladen af tråden.

Strømmen i lederen antages jævnt fordelt over hele lederen tværsnit, så den kan opfattes som bestående af mange parallelle "tråde" hver med en brøkdel af strømmen. På grund af symmetri vil en tråd og dens spejlbillede ligge i samme afstand fra centrum og felterne fra dette trådpair vil ophæve hinanden i centrum så ledernes centrum er altid feltfri. Feltet vokser fra nul til det viste udtryk ved overfladen (se side 28) og relationen er lineær ved lave frekvenser.

I positionen a langs radius ($0 < a < R$) er der et spejlbillede i punktet $-a$. For $a < R/2$ vil feltet i strækningen fra 0 til a gå gennem arealet mod uret, og feltet fra a til $2a$ vil gå gennem arealet med uret, så integralet fra 0 til $2a$ er nul, og kun strækningen fra $2a$ til R giver et positivt bidrag. For $a > R/2$ kan samme argumentation gennemføres fra $2a - R$ til R , mens strækningen fra 0 til $2a - R$ vil give et negativt bidrag af samme styrke som det positive bidrag for $a < R/2$. Det giver et integral fra $a = 0$ til R på nul for lederen ved a , så det er kun tråden ved $-a$, og dermed kun halvdelen af strømmen, der giver et bidrag til den magnetiske flux i lederen³². Den magnetiske induktion beregnes ved integration over et areal i ledningens længde l på $dA = l dr$.

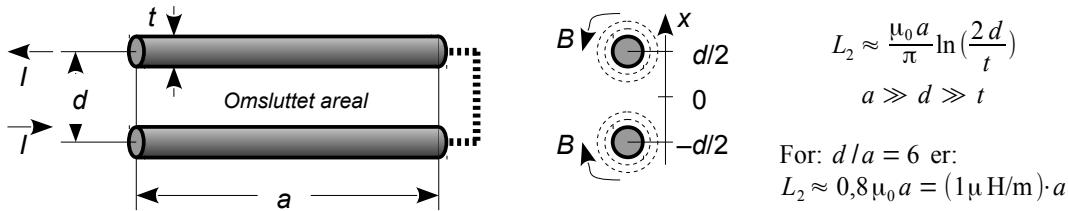
$$B(r) = \frac{\mu_0}{2\pi R} \frac{I}{2} \frac{r}{R} \quad \Rightarrow \quad \Phi_N = \int B dA = \frac{\mu_0 l I}{4\pi R^2} \int_0^R r dr = \frac{\mu_0 l I}{8\pi} \quad \Rightarrow \quad L_1 = \frac{\Phi_N}{I} = \frac{\mu_0 l}{8\pi}$$

31 Der findes en populær tommelfingerregel på 1 $\mu\text{H}/\text{m}$ for et lederykk, som skyldes en misforståelse, og det viser jeg senere.

32 Den eneste udledning jeg kender skyldes Rosa, der uden nogen argumentation skriver at faktor 1/2 optræder.

1.8.3 Dobbeltleder

En dobbeltleder repræsenterer stort set al elektrisk forbindelse med en signalvej og en returvej og mange elektriske forbindelser benytter her to ens ledere. Eksemplerne dækker den elektriske installation i huset, telefon- og datakommunikation samt ledningen til en højttaler. Herunder angiver a længden af lederen, d er afstanden mellem de to ledere og t er trådtynkkelsen.



Figur 31 – To lange ledere i parallel omspænder et areal mellem de to ledere og derved dannes en selvinduktion. Formlen kan anvendes for en almindelig dobbeltledning til fx en højttaler.

Et kabel med afstanden $d = 6$ mm og trådtynkkelsen $t = 1$ mm har selvinduktionen $L_2 = 1$ μH for $a = 1$ m. For en længde på 10 m er selvinduktionen $L_2 = 10$ μH og impedansen når 1Ω ved 16 kHz.

Den elektriske tilslutning benytter to ledere. Strømmen i lederne danner en magnetisk flux omkring hver leder og denne flux løber gennem det areal lederne danner og det giver en selvinduktion³³.

Lederen beskrives her som to parallelle tråde og den fjerne ende opfattes som en kortslutning, men den kan være forbundet til en belastning; fx en spole. Udenfor dobbeltlederen er feltet fra de to ledere modsatrettet og feltet aftager med kvadratet på afstanden fra ledertopparret mens feltet i området mellem de to ledere har samme retning. Hvis de to ledere er lange i forhold til afstanden mellem dem ($a \gg d$) kan formlen for B side 28 benyttes for feltet i arealet mellem de to ledere. Feltet består af to bidrag, der beregnes fra Amperes lov som: $B = \mu_0 I / 2\pi R$. Her skal indføres et koordinatsystem med $x = 0$ midt mellem de to ledere og retning på tværs af lederne. Afstanden fra den underste af de to ledere i figuren til punktet x er givet ved $R = b/2 + x$ og tilsvarende er afstanden fra den øverste leder er $R = b/2 - x$. Feltet i positionen x er givet ved summen af bidragene fra hver af lederne.

$$B_2 = \frac{\mu_0 I}{2\pi(d/2+x)} + \frac{\mu_0 I}{2\pi(d/2-x)} = \frac{\mu_0 I}{2\pi} \frac{d}{d^2/4-x^2}, \quad |x| < \frac{d}{2}-t$$

Trådens diameter t sætter en grænse for den magnetiske induktion så integrationen går ikke fra centrum til centrum, men fra overflade til overflade. Den magnetiske flux findes ved integration af et smalt bånd $dA = adx$ fra $x = -d/2 + t/2$ ved den nederste leder til $x = d/2 - t/2$ ved den øverste leder, hvor a er længden af båndet og dx er bredden af båndet. Der er kun én vinding i spolen ($N = 1$) og integrationen udføres med et tabelværk [RW-157].

$$\Phi_N = N \int B_2 dA = -\frac{\mu_0 ad}{2\pi} \int_{\frac{-d+t}{2}}^{\frac{d-t}{2}} \frac{dx}{x^2 - d^2/4} = -\frac{\mu_0 ad}{2\pi} \left[\frac{1}{2(d/2)} \ln \left| \frac{x-d/2}{x+d/2} \right| \right]_{\frac{-d+t}{2}}^{\frac{d-t}{2}}$$

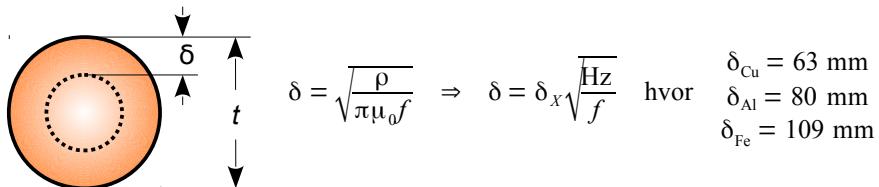
For at komme videre benyttes reglerne om subtraktion af logaritmer: $\ln(x) - \ln(y) = \ln(x/y)$, og opløftning til en eksponent: $\ln(x^2) = 2 \ln(x)$. Det antages at tråden er tynd i forhold til afstanden mellem dem ($t < d$) så nævneren simplificeres til d og en regel for logaritmer bruges for at fjerne det negative fortegn $-\ln(x) = \ln(1/x)$.

$$\Phi_N = -\frac{\mu_0 a}{\pi} \ln\left(\frac{t/2}{d-t/2}\right) \approx \frac{\mu_0 a}{2\pi} \ln\left(\frac{2d}{t}\right) \Rightarrow L_2 = \frac{\Phi_N}{I} \approx \frac{\mu_0 a}{\pi} \ln\left(\frac{2d}{t}\right)$$

³³ I visse kredse indenfor lydgengivelse er holdningen den at de to ledere i kablet til højttaleren skal holdes langt fra hinanden "for at mindske koblingen mellem lederne". Som det ses vil arealet, og dermed selvinduktionen, stige og kablet bliver mere betydnende som komponent. Det eneste der aftager ved at øge afstanden mellem lederne er kapaciteten mellem dem.

1.8.4 Strømfortrængning

Ved lave frekvenser løber strømmen jævnt fordelt over hele ledernes tværsnit, men ved stigende frekvens fortrænges strømmen fra centrum af lederen og koncentrerer til et bånd ved overfladen af lederen. Det kaldes for strømfortrængning. Det kan vises at 63 % af strømmen vil løbe indenfor et område kaldet indtrængningsdybden δ fra overfladen og ind i lederen.



Figur 32 – Feltfortrængning mindsker det effektive areal af lederen, som fører strømmen.

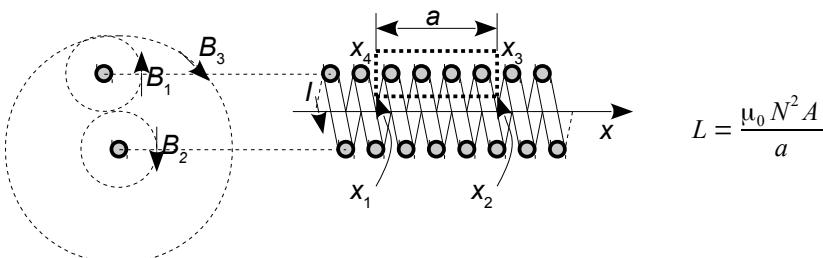
Materialekonstanten ρ kan aflæses på side 10, men for ledere af metal er indtrængningsdybden fra cirka 10 mm ved 50 Hz ned til 1 mm ved 20 kHz så den er uden større betydning ved hørbare frekvenser for typiske kabeldimensioner. Indtrængningsdybden er dog under 100 µm ved 1 MHz så en spole i en radio har strømmen løbende i trådens overflade og modstandsværdien stiger derfor.

Et højspændingskabel af aluminium har ved 50 Hz en indtrængningsdybde på 10 mm. Der benyttes en stålerne til at forøge styrken af kablet for der løber kun svag strøm i kernen. Ved en ydre diameter over 20 mm er det indre af lederen stort set fri for elektrisk strøm, så et større tværsnit kræver at lederen deles i to eller fire parallelle tråde for at udnytte materialet bedst.

Havvand har en indtrængningsdybde på 0,2 m ved 1 MHz, så radiokommunikation med undervandsbåde er vanskeligt for følget aftager med cirka 50 dB/m dybde. Ved 100 Hz er indtrængningsdybden 22 m, men der kan til gengæld kun sendes nogle få bit i sekundet. Det princip er i praktisk anvendelse når det er vigtigt for undervandsbåden at forblive usynlig.

1.8.5 Lang cylindrisk spole

Som introduktion til beregning af selvinduktionen vises et klassisk eksempel hvor lederen er viklet i skrueform. Inde i spolen vil magnetfeltet fra hver snoning af lederen give et magnetfelt i samme retning som nabovindingerne så felterne understøtter hinanden inde i spolen (B_1 og B_2). Udenfor spolen vil felterne være modsat rettet (B_1 og B_3) så feltet er svagt udenfor spolen og vil i stigende afstand fra spolen aftage med kvadratet på afstanden.



Figur 33 – En udsnit af en meget lang spole, hvor strømmen følger en lang led, som er viklet op som en skrue. Udenfor spolen modvirker magnetfelterne hinanden, men inde i spolen har de den samme retning og bidragene adderes. Selvinduktionens formel gælder kun for udsnittet af spolen, men er asymptotisk korrekt for meget lange spoler.

Formlen udledes herunder og er ofte citeret i lærebøger og formelsamlinger, det figurerer flittigt på internettet og det danner et godt udgangspunktet for beregning af den magnetiske induktion for en spole med jernkerne. Desværre gengives formlen tit som selvinduktionen af en spole selv om den kun gælder for et udsnit af en lang spole. Hvis spolens længde er lig med radius vil formlen give en værdi på det dobbelte af det korrekte og fejlen bliver endnu større ved helt korte spoler. Formlerne i de efterfølgende afsnit bør af samme årsag foretrækkes ved spoleberegninger.

Den magnetiske induktion i spolen bestemmes ved Amperes lov. Den lukkede kurve starter ved punktet x_1 inde i spolen, hvor feltet antages at være konstant med værdien B . Strækningen fra x_1 til x_2 har længden a så integralet giver aB . Derefter skiftes retning så kurven krydser ud gennem spolen fra x_2 til x_3 hvor strækningen vinkelret på feltet så vektorproduktet er nul, idet $\cos(\theta)$ er nul. Strækningen fra x_3 til x_4 er så langt væk fra spolen at feltet er nul. Der krydses tilbage til indersiden ved x_4 til x_1 , og igen er det vinkelret på feltet så vektorproduktet er nul. Der er kun et resterende bidrag fra det første stykke inde i spolen, så integralet langs med den lukkede kurve giver aB . Da kurven omslutter N vindinger er den omsluttede strøm NI .

$$\mu_0 NI = \int_0^a B dx = Ba \Rightarrow B = \frac{\mu_0 NI}{a}$$

Spolens areal i tværsnittet er A og den magnetiske flux gennem arealet beregnes ved BA under antagelse af, at den magnetiske induktion er konstant over tværsnittet. Spolen består af i alt N vindinger, der hver afgrænsler et areal på A , og den magnetiske flux løber igennem arealet for hver vinding, så den magnetiske flux krydser igennem N gange arealet af tværsnittet.

$$\Phi_N = N \int B dA = N B A$$

Arealets form har ikke større betydning, så et cirkulært tværsnit benytter $A = \pi R^2$ ved radius R , og for et rektangulært tværsnit er $A = bc$ ved sidemålene b og c . Selvinduktionen af spolen kan nu findes ved at dividere den magnetiske flux (flux-vindingstallet) med strømmen.

$$L = \frac{\Phi_N}{I} = \frac{N B A}{I} = N \frac{\mu_0 N}{a} A \Rightarrow L = \frac{\mu_0 N^2 A}{a}$$

1.8.6 Kort cylindrisk spole

Selvinduktionen for en cylindrisk spole med tråden liggende som en skrue i et enkelt lag beregnes meget ofte efter Wheelers formel, der nærmest har status som selve "formlen" for selvinduktion. Spolen beskrives ved længden a og arealet A af tværsnittet eller alternativt diameteren d regnet fra centrum til centrum af tråden. De N vindinger vikles ofte på en ikke-magnetiserbar kerne af plast eller spolen vikles så den kan bære sig selv. Jeg viser formler for et tværsnit med arealet A samt for et cirkulært tværsnit med diameteren d hvor arealet er $A = \pi d^2/4$.



Figur 34 – Wheelers formel for en spole viklet i ét lag med tynd tråd.

Formlen angives at være korrekt indenfor $\pm 0,35\%$ for spoler med en længde på mindst 40 % af diameteren. Det kræver dog at spolen er viklet med særdeles tynd tråd. Dertil kommer usikkerhed med hvordan spolens dimensioner skal opgøres når tråden har tykkelse og måske ligger i flere lag så i praksis er nøjagtigheden næppe bedre en $\pm 5\%$. Se også kommentarerne på side 39.

Formlen kan sandsynliggøres ved at antage at feltet i spolens akseretning aftager lineært til nul i afstanden R fra spolens ender, og at feltet er nul langs med spolens yderside. Integration langs med aksen giver bidraget Ba fra det indre i spolen og $BR/2$ for hver ende af spolen.

$$\mu_0 NI = \oint B dx = aB + RB \Rightarrow B = \frac{\mu_0 NI}{a+R} \Rightarrow L = \frac{\mu_0 N^2 A}{a+0,5 d}$$

Antages feltet til $(1 - \alpha)B$ inde i spolen og αB langs med spolens yderside vil faktor 0,5 erstattes med 0,45 for $\alpha = 0,05$ hvilket svarer til Wheelers formel. Med $\alpha = 0,08$ bliver værdien på 0,42 og formlen er korrekt indenfor $\pm 10\%$ for en længden ned til 10 % af diameteren.

1.8.7 Tyk cylindrisk spole

Hvis spolens tråd er viklet i lag kan en anden formel fra Wheeler benyttes. Nøjagtigheden kender jeg ikke og det vides heller ikke om der er nogle grænser for forholdet mellem dimensionerne.

$$L_{W2} = \frac{\pi}{4} \frac{8\mu_0 N^2 d^2}{9a + 10b + 3d}$$

$$d = \frac{d_y + d_i}{2} \quad b = \frac{d_y - d_i}{2}$$

$$d_y = D_y - t \quad d_i = D_i + t \quad a = a_y - t$$

Figur 35 – Wheelers formel for en spole viklet i flere lag med ydre diameter D_y , indre diameter D_i og ydre længde a_y . Trådtynkelsen t kan måles med en skydelærer eller blot ignoreres.

Spolens DC modstand er proportional med længden af tråd så et optimalt design vil maksimere selvinduktionen for den kortest tråd. Det resulterer i *Brooks coil*.

$$L_B = 0,676 \mu_0 N^2 d$$

$$d = 3b$$

$$b = a$$

Figur 36 – Brooks formel for en spole med lavest mulig DC modstand for en given selvinduktion.

1.8.8 Biot og Savarts lov

Amperes lov beskriver feltet omkring uendeligt lange ledere. Korte ledere er dog meget væsentlige indenfor elektronikken og det er muligt at analyser feltet omkring dem ved *Biot og Savarts lov*. Her angives den magnetiske induktion dB for et punkt i rummet hidrørende fra strømmen I i en kort tråd med dx for længden af en underdeling af tråden og med r for afstanden til punktet.

$$\vec{dB} = \frac{\mu_0 I}{4\pi r^2} \frac{\vec{dx} \times \vec{r}}{|\vec{r}|} \Rightarrow dB = \frac{\mu_0 I}{4\pi r^2} \sin(\theta) dx$$

$$|\vec{dx} \times \vec{r}| = dx r \sin(\theta)$$

$$|\vec{r}| = r$$

$$B = \frac{\mu_0 I}{4\pi} \int \frac{\sin(\theta)}{r^2} dx$$

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ H/m}$$

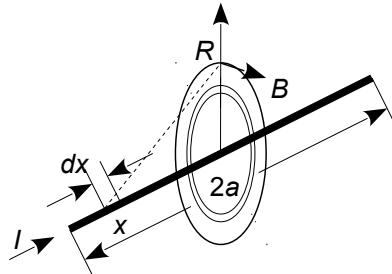
Figur 37 – Det magnetiske felt i et punkt kan beregnes ved at summere delbidrag fra lederen og det enkelte delbidrag beregnes efter Biot-Savarts lov.

Udtrykket hører ikke til de mest indbydende, men det kan forenkles. Længden af krydsproduktet er givet ved produktet af de to vektorers længder gange sinus til vinklen imellem dem. Længden af r kan derved divideres bort så der kun er dx og $\sin(\theta)$ tilbage og sinus til en vinkel kan som regel beskrives ved trekantrelationer. For at beregne den magnetiske induktion i et punkt i rummet adderes de enkelte bidrag dB ved integration af dx over den aktuelle længde af lederen hvilket i illustrationen herover er fra x_1 til x_2 .

Udtrykket skal benyttes for at finde selvinduktionen af en cirkulær spole med en enkelt vinding, men først vises det at den magnetiske induktion omkring en lang leder giver det samme resultat som ved brug af Amperes lov. Det viser at feltet tæt på en kort leder er meget tæt på værdien som er givet ved Amperes lov for en uendelig lang leder.

1.8.9 Feltet omkring en leder

Den magnetiske induktion B beregnes langs med cirkelperiferien i afstanden R fra en leder som vist i figuren herunder og beregningen gennemføres i eksemplet. Resultatet er det samme som Amperes lov for en uendelig lang leder og viser at man kan vælge den enkleste af metoderne.



$$\text{Ampere} \quad B_1 = \frac{\mu_0 I}{2\pi R}, \quad a \rightarrow \infty$$

$$\text{Biot-Savart} \quad B_1 = \frac{\mu_0 I}{2\pi R} \frac{a/R}{\sqrt{(a/R)^2 + 1}}$$

Figur 38 – Det magnetiske felt omkring lederen efter både Amperes lov og Biot-Savarts lov.

I afstanden R fra midten af en leder med længden $2R$ (det vil sige $a/R = 1$) er feltet på 71 % af værdien for en uendelig lang leder. I en afstand på 10 % af ledernes længde ($a/R = 5$) er feltet på 98 % af Amperes lov for en uendeligt lang leder.

Feltet skal beregnes langs en cirklen med radius R og centrum ved $x = 0$. Den elektriske leder strækker sig fra $x = -a$ ved den ende hvor strømmen føres ind til $x = a$ ved den anden ende, så lederen har længden $2a$. I det følgende antages det at nul ligger på ledernes midtpunkt for at simplificere udtrykkene. Der tages ikke hensyn til at strømmen skal føres frem til og væk fra ledernen for feltet herfra kan beregnes på samme måde. Linjestykket dx på ledernen er placeret ved x , og længden af vektor r beregnes af Pythagoras.

$$|\vec{r}| = \sqrt{R^2 + x^2}$$

Magnetfeltets vektor dB vil pege langs med tangenten til det punkt hvor R rører cirklen; det er ud af papirets plan øverst i tegningen. Sinus til vinklen mellem vektor r og x -aksen skrives som R delt med længden af vektor r ved at benytte lidt trigonometri, og længden skrives blot som r .

$$\sin(\theta) = \frac{R}{|\vec{r}|} = \frac{R}{r}$$

Ved indsættelse kan magnetfeltets induktion B beregnes langs med cirkelperiferien i afstanden R . En kvadratrod svarer til eksponenten $1/2$ så radius i tredje giver eksponenten $3/2$.

$$B = \frac{\mu_0 I}{4\pi} \int \frac{R/r}{r^2} dx = \frac{\mu_0 I}{4\pi} \int_{-a}^a \frac{R}{(x^2 + R^2)^{3/2}} dx$$

For at gennemføre integrationen normeres x med afstanden R så x ændres til $\alpha = x/R$ og derved er $x = R\alpha$ og differentialet bliver $dx = Rd\alpha$. Parentesen i ændres til $\alpha^2 + 1$ og grænserne ændres til $\alpha_1 = -a/R$ for den nedre og $\alpha_2 = a/R$ for den øvre. Nu er R^2 sat udenfor parentesen i nævneren og det giver R^3 når eksponenten på $3/2$ tages i ed. Med R^2 i tællereren svarer det til at konstanten nu indeholder R i nævneren. Integrationen klares med et tabelopslag [RW-160].

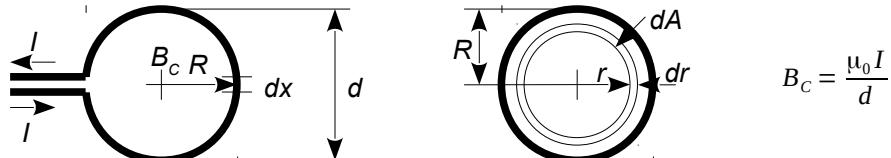
$$B = \frac{\mu_0 I}{4\pi} \int_{-a/R}^{a/R} \frac{R}{[R^2 \alpha^2 + R^2]^{3/2}} R d\alpha = \frac{\mu_0 I}{4\pi R} \int_{-a/R}^{a/R} \frac{1}{[\alpha^2 + 1]^{3/2}} d\alpha = \frac{\mu_0 I}{4\pi R} \left[\frac{\alpha}{\sqrt{\alpha^2 + 1}} \right]_{-a/R}^{a/R}$$

De to grænser indsættes for at give resultatet vist i illustrationen.

$$B = \frac{\mu_0 I}{4\pi R} \left[\frac{a/R}{\sqrt{(a/R)^2 + 1}} - \frac{-a/R}{\sqrt{(a/R)^2 + 1}} \right] = \frac{\mu_0 I}{2\pi R} \frac{a/R}{\sqrt{(a/R)^2 + 1}}$$

1.8.10 Feltet i centrum af en cirkel

Hvis lederen bukkes til en cirkulær spole med en enkelt vinding vil den magnetiske induktion fra de mange delelementer dx adderes i cirklens indre mens den magnetiske induktion i et punkt udenfor spolen vil udbalanceres på samme måde som ved spolen med to parallelle ledere på side 31. Den magnetiske induktion i centrum er π større end i afstanden R fra en lang leder (se side 28).



Figur 39 – Den magnetiske feltstyrke i centrum kan beregnes af Biot og Savarts lov.

Den magnetiske induktion i centrum er givet ved Biot og Savarts lov som en sum af bidragene fra dx på cirkelperiferien. I afstanden $r = R$ fra centrum løber dx én gang rundt i cirklen svarende til en længde på $x = 2\pi R$. Vinklen mellem dx og r er $\theta = \pi/2$ (90°) så $\sin(\theta) = 1$.

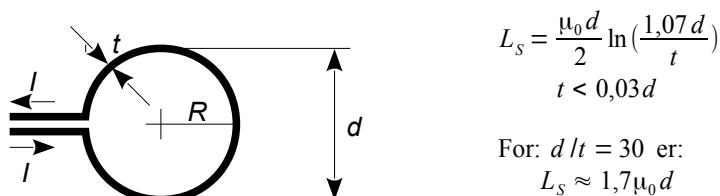
$$B_c = \frac{\mu_0 I}{4\pi} \int \frac{\sin(\theta)}{R^2} dx = \frac{\mu_0 I}{4\pi R^2} \int_0^{2\pi R} dx = \frac{\mu_0 I}{4\pi R} [x]_0^{2\pi R} = \frac{\mu_0 I}{2R} = \frac{\mu_0 I}{d}$$

Selvinduktionen kan estimeres hvis det antages at den magnetiske flux er konstant over arealet. Resultatet er dog lavere end den mere udbyggede beregning for en skiveformet spole.

$$L_c \approx \frac{B_c A}{I} = \frac{\mu_0 A}{2R} = \frac{\mu_0 \pi R^2}{2R} = \frac{\mu_0 \pi R}{2} = \frac{\pi \mu_0 d}{4} = 0,8 \mu_0 d$$

1.8.11 Skiveformet spole

En leder bukket til en cirkel med radius R danner en vigtig komponent i elektronikken for spoler har ofte et cirkulært tværsnit da det giver den bedste udnyttelse af tråden. Det vil sige at der opnås den største værdi af selvinduktionen ved den laveste værdi af trådens modstand. Den magnetiske flux fra vindingen løber igennem området i det indre af spolen og bidrager til selvinduktionen.



Figur 40 – En cirkulær spole har det største areal med mindst muligt trådlængde, og giver derfor den største selvinduktionen for en given modstand i tråden.

Den magnetiske induktion vokser fra centrum og på grund af symmetri kan den beskrives som funktion af radius. Værdien er B_c i centrum af spolen og B_1 i nærheden af lederen. Her benyttes en tilnærmelse hvor feltet beskrives ved summen af B_c og B_1 . Der er derved to bidrag i centrum så for at feltet i centrum skal blive korrekt sænkes niveauet fra B_c med faktor α så summen af de to led giver B_c ved centrum. Som vist før giver B_1 cirka 1/3 af B_c i centrum så B_c står for 2/3.

$$B_r = \alpha B_c + B_1 \quad \text{hvor} \quad B_c = \frac{\mu_0 I}{2R} \quad \text{og} \quad B_1 = \frac{\mu_0 I}{2\pi(R-r)} \quad \Rightarrow \quad \alpha = \frac{\pi-1}{\pi}$$

Den magnetiske induktion varierer med radius r . R er cirklens radius til centrum af tråden og t er trådens tykkelse. Den variable går derfor fra $r = 0$ ved centrum til $r = R - t/2$ ved overfladen af tråden. Den magnetiske flux gennem arealet findes ved at integrere over cirklens areal og på grund af rotationssymmetrien benyttes et areal dA formet som et cirkulært bånd i afstanden r fra centrum. Båndets længde er $2\pi r$ og dets tykkelse er dr så arealet bliver $dA = 2\pi r dr$. Der er kun

én vinding så $N = 1$ og den magnetiske flux beregnes.

$$\Phi_N = N \int_0^{R-\frac{t}{2}} B_r dA = \int_0^{R-\frac{t}{2}} \left(\frac{\pi-1}{\pi} \frac{\mu_0 I}{2R} + \frac{\mu_0 I}{2\pi(R-r)} \right) 2\pi r dr$$

De fælles konstanter kan sættes udenfor og integralet deles i to hvorved det første giver $r^2/2$. For at løse det andet anvendes et tabelværk [RW-153]. Substitution med $x = r/R$ giver $r = Rx$ og $dr = Rdx$ og den øvre bliver $1 - t/2R$.

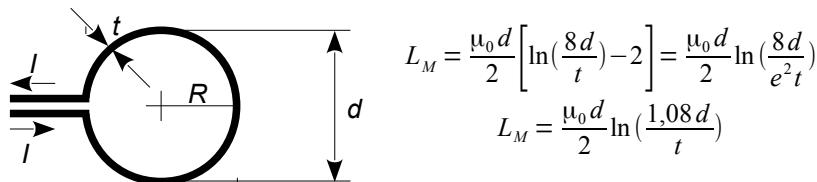
$$\Phi_N = \frac{\pi-1}{2R} \mu_0 I [r^2]_0^{R-\frac{t}{2}} - \mu_0 I R [x + \ln|x-1|]_0^{1-\frac{t}{2R}}$$

Grænserne indsættes og det antages at trådtykkelsen er lille i forhold til cirkelens radius. Det giver en fejl på 1 % for en trådtykkelse på 3 % af diameteren og på 5 % hvis trådtykkelsen er på 10 % af diameteren. Regnereglen $\ln(x) = -\ln(1/x)$ fjerner et minus og de konstante led samles til $0,071 = \ln(e^{0,071})$. Regnereglen for addition af logaritmer $\ln(x) + \ln(y) = \ln(xy)$ giver en faktor på $e^{0,071} = 1,073$ i tælleren. Udtrykket reduceres til det viste ved $d = 2R$.

$$\Phi_N = \mu_0 I R \left(\frac{\pi-1}{2} - 1 - \ln \left(\frac{t}{2R} \right) \right) = \mu_0 I R \left(0,071 + \ln \left(\frac{2R}{t} \right) \right) = \mu_0 I R \ln \left(\frac{2e^{0,071} R}{t} \right)$$

1.8.12 Maxwells formel

Litteraturen citerer ofte et udtryk kaldet Maxwells formel³⁴.



Figur 41 – Selvinduktionen af en spole med en enkelt vinding kaldes for Maxwells formel.

Formlen forudsætter så høje frekvenser at det er muligt at se bort fra den interne induktion. Tallet 2 kan skrives som $\ln(e^2) = 2$ og subtraktion af to logaritmer tillader at tallet trækkes ind i logaritmen som divisor e^2 . Divideret op i ottetallet giver det en faktor, der er meget tæt på resultatet ovenfor.

1.8.13 Kort spole

Formlen for L_s gælder kun for én vinding, men den kan udvides til at omfatte flere vindinger ved at inkludere kvadratet på vindingstallet. Det forudsætter dog at magnetfeltet fra hver enkelt viking når gennem alle vindingernes arealer. Det er dog ikke en realistisk mulighed alene på grund af trådens tykkelse. Det betyder at en del af feltet ikke når rundt om alle vindingerne så den estimerede værdi af selvinduktionen bliver for høj. Det kan der korrigeres for ved at ignorere faktoren på 1,07 og ved at erstatte trådtykkelsen med spolens dimension i længderetningen. Som vist i sammenfatningen har jeg god erfaring med formlen for korte spoler (se side 55).

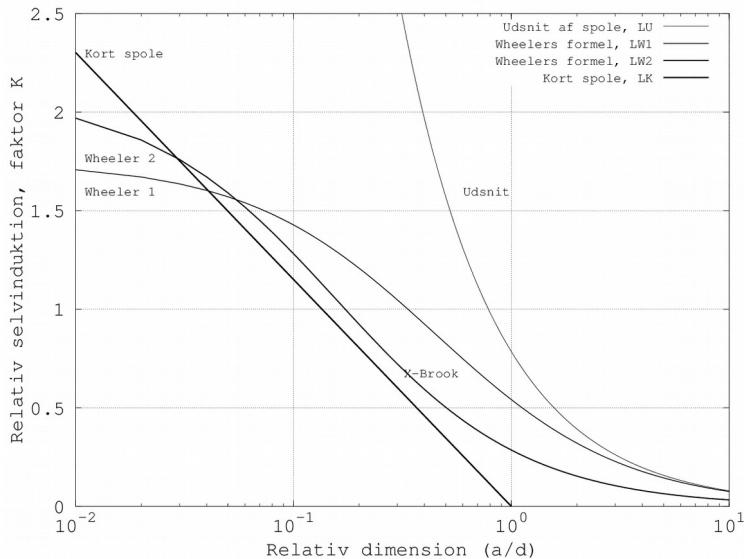


Figur 42 – Formlen for en skiveformet spole kan udvides til at gælde for korte spoler.

34 Se: <http://en.wikipedia.org/wiki/Inductance>.

1.8.14 Sammenligning af formlerne

Modellerne er gode i hvert sit område, men desværre overlapper områderne ikke så der vil ikke blive beregnet den samme selvinduktion for en spole ved brug af de forskellige formler. Wheelers formel L_{W1} er god for spoler med en længde på mindst 40 % af diameteren og formlen for en kort spole L_K bør vælges når længden er under 10 % af diameteren. Imellem de to kan L_{W2} benyttes. Brooks formel L_B er et bud på designet af en spole hvor DC modstanden er minimal.



Generel formel:

$$L_X = K \mu_0 N^2 d$$

Udsnit af spole:

$$L_U = \frac{\pi \mu_0 N^2 d^2}{4 a}$$

Wheelers formler:

$$L_{W1} = \frac{\pi \mu_0 N^2 d^2}{4 a + 0,45 d}$$

$$L_{W2} = \frac{\pi 8 \mu_0 N^2 d^2}{4 9 a + 10 b + 3 d}$$

Kort spole:

$$L_K = \frac{\mu_0 N^2 d}{2} \ln\left(\frac{d}{a}\right)$$

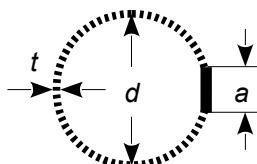
Brooks coil:

$$L_B = 0,676 \mu_0 N^2 d$$

Figur 43 – Sammenligning mellem formlerne, der er normaleret med $\mu_0 N^2 d$ og plottet som funktion af længden i forhold til diameteren. Brooks coil er vist som et kryds og passer med Wheelers formel L_{W2} for en tyk spole med $b = a$ (se side 34). I kurven er a spolens længde og d dens diameter.

1.8.15 Selvinduktion af ledning

Der er en populær tommelfingerregel om at en tråd har selvinduktionen 1 μH per meter. Jeg har set det demonstreret i praksis med et måleapparat, men selv om det tilsyneladende kan eftervises eksperimentelt så er resultatet forkert og det er der en enkel forklaring på. Enhver ledning indgår i et lukket kredsløb når der måles selvinduktion og den målte selvinduktion er derfor et resultat af den samlede løkke. Ved en måling indgår tilledningerne derfor som en del af målingen og hvis måleapparatet er kalibreret til at vise en selvinduktion på nul med kortsluttede prober vises kun ændringen fra kortsluttede prober til situationen hvor måleobjektet indgår.



$$\begin{aligned} \frac{L}{a} &= \frac{L_S}{\pi d} = \frac{\mu_0}{2\pi} \ln\left(\frac{1,07 d}{t}\right) = 0,6 \dots 1,4 \mu\text{H/m} \\ \Rightarrow \quad \frac{L}{a} &\approx 1 \mu\text{H/m} \end{aligned}$$

Figur 44 – Tommelfingerreglen om 1 $\mu\text{H/m}$ for en kort ledning er god til overslagsberegning, men den skyldes at ledningen nødvendigvis indgår i et system med en returvej og derfor er del af en spole med en enkelt vinding.

Med formlen for selvinduktionen af en cirkulær vinding kan selvinduktionen per meter beregnes ved at dividere med omkredsen af spolen. Konstanten $\mu_0/2\pi$ er 0,2 $\mu\text{H/m}$ og forholdet mellem diameter og trådt tykkelse vil ligge i området fra 20 til 1000 hvor logaritmen giver en værdi i det relativt begrænsede område fra 3,0 til 6,9. Selvinduktionen af en enkelt vinding er derfor fra 0,6 til 1,4 $\mu\text{H/m}$ med et gennemsnit på 1 $\mu\text{H/m} \pm 40\%$. Hvis måleopstillingen er af en størrelse, der kan ligge på et bord, vil den observerede variation fejlagtigt kunne afskrives som en måleusikkerhed.

1.8.16 Afslutning

Som en slags "facitliste" for beregning af selvinduktion benyttes et arbejde af Lorentz og Nagaoka hvis der ikke er mulighed for at måle selvinduktionen på en færdig spole. Problemet er dog at den matematiske abstraktion, som Lorentz benyttede, er en uendelig tynd folie formet som en cylinder, med strømmen kredsende perfekt cirkulært. Det er et såkaldt *current sheet*. En reel spole vikles derimod i skrueform og med en tråd, der har en endelig tykkelse, så der begås en tilnærmede alene ved at forsøge at sammenligne med Lorentz arbejde.

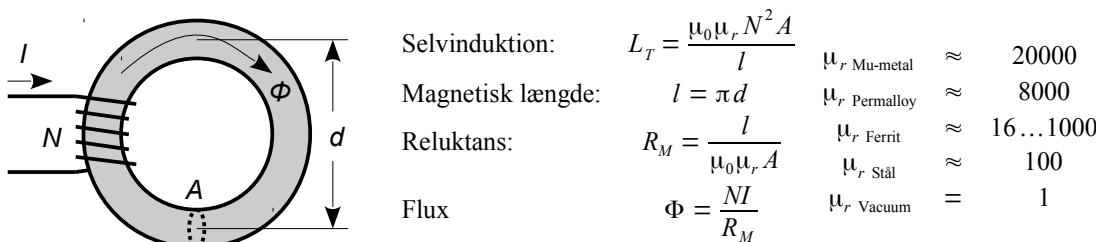
Materialet fra Lorentz og Nagaoka giver en for høj selvinduktion ved en kort spole idet et current sheet svarer til en tråd med tykkelsen nul. Her giver Maxwells formel et mere realistisk bud, men den tillader ikke flere vindinger så her er man på tynd is for den almindelige spole. Der er altså flere bud på en facitliste for spoler med realistiske dimensioner så en universel formel findes ikke.

Dertil kommer et ofte overset problem med at måle dimensionerne på spolen. For at få en idé om betydningen heraf så betragt en spole med 1 mm tråd og 25 vindinger. Længden er 25 mm, men center til center er den kun 24 mm (en mulig fejl på 5%). Trådens diameter er normalt opgivet for kobberet alene, altså uden isolation, men med blot 25 µm til lak bliver spolens 1,25 mm længere end antaget (en mulig fejl på 5%). Benyttes en skydelærer for måling af diametern er det svært ikke at deformere spolen under målingen og det er endnu en mulig fejl af samme størrelsesorden. Anvendes en form til vikling af spolen skal man tage hensyn til både trådtykkelse og isolation og så kan en håndviklet spole ikke vikles lige så tæt som det måske var planlagt.

Som en konklusion er det ret sikkert at målsætningen ikke nås og min erfaring er at der ikke kan forventes en nøjagtighed bedre end $\pm 10\%$ samt at der let opstår overraskelser.

1.9 Magnetiserbar kerne

Der er en grænse for hvor høj selvinduktion, der kan opnås med de såkaldte luftspoler, der er viklet på et ikke-magnetiserbart materiale, som ud over luft inkluderer plast og træ. Selvinduktionen kan øges radikalt ved at benytte en kerne af et magnetiserbart materiale, som jern og nikkel, hvori der kan være blandet andre stoffer. Ved vekselstrøm dannes der roterende strømmme i kernen, som modvirker det magnetiske felt, så kernen udformes normalt af små metalflager, der arrangeres så de bryder strømmen. Et populært materiale i elektronikken er ferrit, der er et keramisk materiale, som kan støbes i den ønskede form og ikke giver alt for store tab ved vekselstrøm.



Figur 45 – En ringkerne danner en lukket magnetisk sløjfe og det magnetiserbare materiale giver en forstærkning af den magnetiske induktion (relativ permeabilitet fra Wikipedia).

Den forstærkende virkning skyldes at atomerne i en magnetiserbar kerne har et "spin", der repræsenterer en roterende strøm. Det giver atomet et magnetfelt som peger i en eller anden tilfældig retning. Atomerne er arrangeret i domæner, hvilket blev opdaget af Weiss i 1906 og domænerne bærer hans navn. Indenfor et sådant domæne vil alle atomer pege i samme retning. Hvert af domænerne er 1 til 100 µm i udstrækning og med et stort antal domæner, vil nettoresultatet set udefra være fravær af magnetfelt. Hvis der påtrykkes et eksternt magnetfelt vil domænerne kunne skifte til samme retning som det eksterne felt, og derved vil de ikke længere opnæve hinandens magnetfelter, men derimod virke feltforstærkende. Det er en process, der virker indtil alle domæner pege i samme retning, og derefter kan feltet ikke forstærkes yderligere, hvilket kaldes for mætning (*saturation*). Det betyder at relationen mellem strøm og induktion kun er omkrents lineær indtil en øvre grænse, der for jern ligger ved cirka 2 T.

Det magnetiserbare materiale har en betydelig forstærkende effekt på den magnetiske induktion, og selvinduktionen kan forøges mere end 1000 gange blot ved at vikle en spole på en egnet kerne. Den feltforstærkende effekt beskrives ved en konstant μ_r for det anvendte materiale. Værdien er én for magnetisk neutrale materialer som luft, træ og plast, og større end én for de magnetiserbare materialer. Visse materialer som fx grafit modsætter sig magnetisering og har en værdi under én.

Selvinduktionen kan bestemmes på samme måde som vist fra side 30 og frem, med den tilføjelse at den magnetiske induktion er $\mu_r\mu_0B$ i kernen. Den mest effektive forøgelse af selvinduktionen opnås når magnetfeltet løber i en lukket sløjfe, som eksempelvis ved ringkernen (*toroid core*). Der findes mange andre udformninger og den aktuelle geometri af kernen er ikke specielt vigtig.

1.9.1 Magnetisk modstand

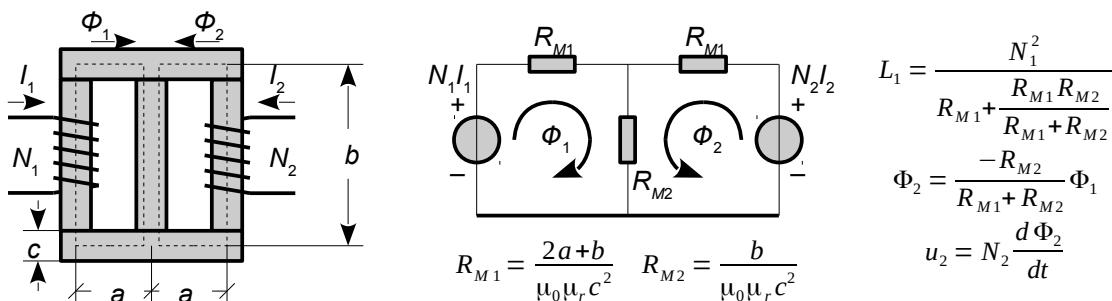
Den magnetiske modstand, R_M , som også kaldes reluktans, blev introduceret af Heaviside i 1888. Det er en metode til at beregne den magnetiske flux Φ i en specifik del af et magnetisk kredsløb. Han foreslog at produktet NI blev betragtet som en slags "motor", der driver den magnetiske flux Φ i de magnetiske kredsløb i analogi til Ohms lov. Det magnetiske "kredsløb" repræsenteres af den magnetiske modstand, der beregnes fra længden af det magnetiske materiale l , som feltlinjerne vil følge, den feltforstærkende virkning af materialet μ_r (den magnetiske permeabilitet) og arealet af materialets tværsnit A som den magnetiske flux passerer igennem.

$$NI = R_M \Phi \quad \text{hvor} \quad R_M = \frac{l}{\mu_0 \mu_r A}$$

Den magnetiske modstand R_M kan beregnes ved brug af formlen $\Phi_N = LI$ for selvinduktion, hvor Φ_N er flux-vindingstallet $\Phi_N = N\Phi$. Sidste beregningstrin benytter formlen for en lang spole.

$$R_M = \frac{NI}{\Phi} = \frac{N^2 I}{N\Phi} = \frac{N^2 I}{\Phi_N} = \frac{N^2 I}{LI} \Rightarrow R_M = \frac{N^2}{L} = \frac{N^2 I}{\mu_0 \mu_r N^2 A} = \frac{I}{\mu_0 \mu_r A}$$

Hvis det magnetiske system opbygges af flere sammenkoblede sektioner kan man beregne den magnetiske modstand for hver sektion og benytte kredsløbsanalysens værktøj til at bestemme den magnetiske flux i hver enkelt maske. Det vises herunder for et system med to spoler N_1 og N_2 på en fælles EI-kerne. Man kan så beregne selvinduktionen for L_1 samt fluxen Φ_2 og spændingen ved spole N_2 . Formlerne herunder forudsætter at $I_2 = 0$.



Figur 46 – Den populære EI kerne, der er dannet af tynde plader med form som et E og et I, der blæses sammen med skiftende retning for at mindske den effektive værdi af luftspalten.

Kredsløbet udgør en transformator, hvor midterbenet er indført for at vise hvordan en beregning på et magnetisk kredsløb kan gennemføres. Feltretningen bestemmes med højrehåndsreglen og det ses at begge spoler sender en flux i samme vej gennem midterbenet. Differentiationen $d\Phi_2/dt$ fra Faradays induktionslov tillader ikke overførelse af DC gennem transformatoren. En tonegenerator på spolen N_1 driver en AC strøm i spolen og det danner en magnetisk flux Φ_1 , der dæmpes til den magnetiske flux Φ_2 ved det viste udtryk med R_{M1} og R_{M2} .

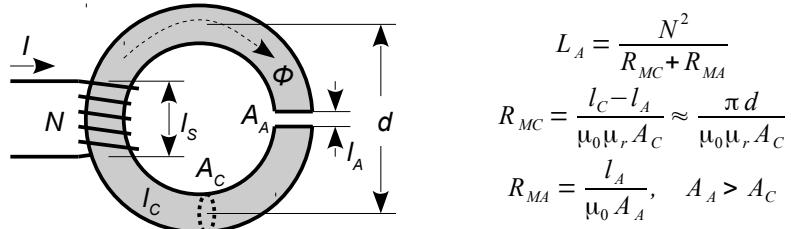
Udgangsspændingen fra spolen N_2 bliver en skaleret version af spændingen på indgangen.

$$\begin{aligned} u_1 &= U_1 \sin(\omega t) \Rightarrow u_2 = N_2 \frac{d\Phi_2}{dt} \Rightarrow u_2 = \frac{N_2}{N_1} \frac{R_{M2}}{R_{M1} + R_{M2}} U_1 \sin(\omega t) \\ i_1 &= \int \frac{u_1}{L_1} dt = -\frac{U_1}{\omega L_1} \cos(\omega t) \Rightarrow \Phi_2 = \frac{R_{M2}}{R_{M1} + R_{M2}} \frac{L_1 i_1}{N_1} \quad u_2 \underset{R_{M2} \rightarrow \infty}{=} \frac{N_2}{N_1} U_1 \sin(\omega t) \end{aligned}$$

Hvis midterbenet fjernes opstår relationen for en ideel transformator, hvor udgangsspændingen er givet ved indgangsspændingen gange med forholdet mellem antallet af vindinger.

1.9.2 Luftspalte

Hvis et stykke af kernen skæres bort efterlades en luftspalte med længden l_A . Den magnetiske modstand i luftspalten er meget højere end i materialet ($\mu_r = 1$), og blot 1 mm luft vil svare til den magnetiske modstand af en længde på omkring en meter af kernens materiale ($\mu_r \approx 1000$). Den magnetiske modstand består af en serieforbindelse med R_{MA} fra luftspalten og R_{MC} fra kernen. Selvinduktionen er givet som vist herunder, hvor index C står for core og A for air.



Figur 47 – Et kort luftgab indskyder en forholdsvis stor magnetisk modstand i kredsløbet.

Selvinduktionen er bestemt af kernen når R_{MC} er dominerende, og af luftspalten når R_{MA} er størst, og for et magnetiske materiale med $\mu_r > 1000$ svarer det til en luftspalte på 0,1 % af længden af kernen hvis forholdet A_A/A_C antages at ligge tæt på én. Det er dog kun tilnærmelsesvist tilfældet for feltet spredes i luftspalten, så luftspalten bliver meget hurtigt betydelige.

Med en diameter på $d = 30$ mm er vejlængden i kernen $l_C = \pi d = 94$ mm så med en relativ magnetisk konstant på $\mu_r = 1000$ vil en luftspalte på blot 94 µm kunne halvere værdien af selvinduktionen, og hvis der tages hensyn til spredningen af feltet vil fordoblingen indtræde før luftspalten er nået op på 94 µm.

Det er ikke enkelt at bestemme det gennemsnitlige areal A_A for feltet i luftspalten, for arealet er en funktion af længden af luftspalten og derfor af spolens mekaniske design.

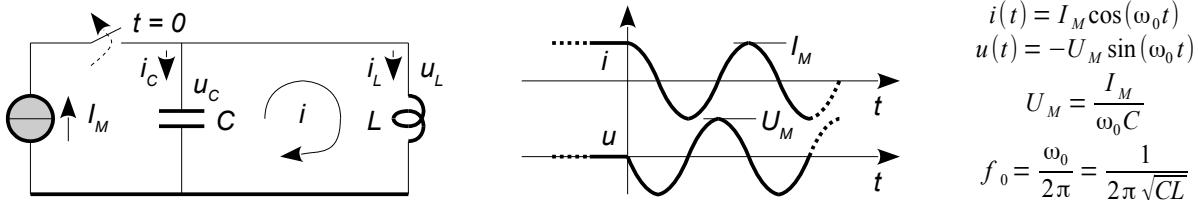


Figur 48 – Eksempler på spoler med kerne af magnetiserbart materiale. Til venstre en ringkerne, der er skåret op for at udnytte luftspaltens induktion. I midten en kommersiel spole med en kerne, der ikke er lukket. Til højre et "samlesæt" bestående af spoleform af plast og to kernedele.

For kommersielt tilgængelige kerner og spoleforme er luftspalten specifiseret ved en faktor, der angiver hvor meget selvinduktionen øges ved at bruge produktet. Det gælder for de lukkede spoleformer, vist til højre i illustrationen, hvor luftspalten er en integreret del af designet.

1.10 Resonans

De to reaktive komponenter, kondensatoren og spolen, er været flittigt benyttet i elektronikken, da de tilsammen har den vigtige egenskab at kunne danne en svingningskreds. Den er fundamental for forståelsen af radiomodtagere, oscillatorer og filtrering.



Figur 49 – En sammenkobling af en kondensator og en spole danner en svingningskreds.

Et passivt kredsløb foretager sig intet uden tilførsel af energi, der her er en ekstern strømkilde, som opretholder en DC strøm i spolen. For en ideel spole er DC modstanden nul, så spændingsfaldet i kredsløbet er nul. Tid tiden $t = 0$ afbrydes strømkilden, hvorefter kondensator og spole er overladt til sig selv. Strømmen fra spolen oplader kondensatoren og spolens magnetfelt aftager, idet der overføres energi fra spolen til kondensatoren. Når kondensatorens spænding er maksimal vil strømmen i spolen være nul. Derefter starter den modsatte proces hvor kondensatoren afgiver energi tilbage spolen, og genopbygges en strøm i kredsløbet.

Det kaldes for resonans. Spændingen vil oscillere 90° forskudt i forhold til strømmen, og den ene er maksimal når den anden er nul. I et mekanisk system svarer det til at bevægelsesenergi veksler med potentiel energi, som fx pendulet i et standerur.

Kredsløbet undersøges med Kirchhoffs maskelov for situationen efter at kontakten er åbnet. Der benyttes positiv retning med uret og startbetingelsen er $i(0) = I_M$. Det er ønsket her, at finde både strømmen $i(t)$ og spændingen $u(t)$ som funktion af tiden. Kondensatorens strøm har et negativt fortegn for strømmen på grund af den valgte positive retning (se side 22).

$$-u_C + u_L = 0 \Rightarrow -\frac{1}{C} \int (-i dt) + L \frac{di}{dt} = 0 \Rightarrow \frac{1}{C} \int i dt + L \frac{di}{dt} = 0 \quad \text{Differentiation} \Rightarrow i = -LC \frac{d^2 i}{dt^2}$$

Ved differentiation fjernes integrationen og relationen bliver en anden-ordens differentialligning. Løsningen (strømmen i) skal give sig selv med modsat fortegn ved to gange differentiation, og det er tilfældet for sinus og cosinus. Strømmens tidsmæssige forløb forventes at være baseret på en cosinus da strømmen er maksimal til $t = 0$. Prøves med sinus fås et tilsvarende resultat, men fasen skal justeres så spolens strøm bliver maksimal til tiden $t = 0$.

$$i = I_M \cos(\omega_0 t) \Rightarrow \begin{aligned} \frac{di}{dt} &= \frac{d[I_M \cos(\omega_0 t)]}{dt} = -\omega_0 I_M \sin(\omega_0 t) \\ \frac{d^2 i}{dt^2} &= \frac{d[-\omega_0 I_M \sin(\omega_0 t)]}{dt} = -\omega_0^2 I_M \cos(\omega_0 t) \end{aligned}$$

Med udtrykkene indsattes det at $I_M \cos(\omega_0 t)$ kan divideres bort, hvilket viser at antagelsen om løsningen var korrekt, for ellers ville resultatet ende med en uløselig konflikt.

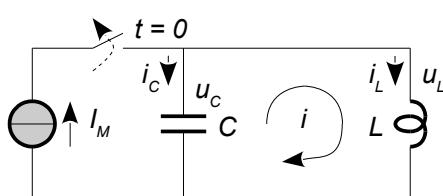
$$I_M \cos(\omega_0 t) = LC \omega_0^2 I_M \cos(\omega_0 t) \Rightarrow 1 = LC \omega_0^2 \Rightarrow \omega_0 = \frac{1}{\sqrt{LC}}$$

Strømamplituden er kendt (I_M) og spændingen over opstillingen (u_C) kan findes ved indsættelse af strømmens forløb i definitionen for kondensatorens spænding.

$$u_C = \frac{1}{C} \int i_C dt = \frac{1}{C} \int (-i) dt = -\frac{1}{C} \int I_M \cos(\omega_0 t) dt = -\frac{I_M}{\omega_0 C} \sin(\omega_0 t)$$

1.11 Energi

Kondensatoren opmagasinerer energi givet ved spændingen over den, og spolen opmagasinerer energi når strømmen løber i den. Det er den samme energimængde, der flyttes frem og tilbage mellem de to komponenter i kredsløbet, så den totale energi i kredsløbet er til ethvert tidspunkt givet ved summen af energierne i de to komponenter.



Energi i en kondensator:

$$w_C(t) = \frac{1}{2}Cu^2(t) \Rightarrow W_{C MAX} = \frac{1}{2}CU_M^2$$

Energi i en spole:

$$w_L(t) = \frac{1}{2}Li^2(t) \Rightarrow W_{L MAX} = \frac{1}{2}LI_M^2$$

Energi i kredsløbet:

$$W_{TOT} = w_C(t) + w_L(t)$$

$$W_{TOT} = W_{C MAX} = W_{L MAX}$$

Figur 50 – Energien svinger frem og tilbage mellem kondensatoren og spolen i kredsløbet, og den samlede energi er givet ved strømmens eller spændingens maksimum.

Den oplagrede energi i en reaktiv komponent er af interesse ved sikkerhedsberegninger, fx indenfor eksplosionsfarlige områder. Hvis et kredsløb eksempelvis har en kondensator på $1 \mu\text{F}$, som er opladet til 5 V , så er der opmagasineret en beskeden energi på $12,5 \mu\text{J}$, som næppe kan anrette megen skade, men hvis kondensatoren oplades til 500 V vil energien på 125 mJ kunne danne en tilstrækkelig kraftig gnist til at starte en forbrænding. Ved EMC test af elektronisk udstyr skal ind- og udgange kunne tåle en afladning af en kondensator på omrent denne værdi gennem en modstand på 42Ω , der vil medføre en startstrøm på små 12 A . Almindelig integreret elektronik kan ikke klare at optage en energimængde ret meget over 1 mJ , så producenten må indbygge et beskyttelseskredsløb bestående af modstande, for at begrænse strømmen, og dertil zenerdioder (som fx TranZorb-serien), varistorer og gasudladningsrør, for at optage energien.

For en spole er problemet mere konkret, for energien opbygges ved at der løber en strøm i kredsløbet. Afbrydes for strømvejen vil spændingen over spolen stige indtil der dannes en ny strømvej, hvilket som regel er en gnist gennem luften eller et sammenbrud i en anden komponent. Den princip benyttes i benzindrevne biler til at starte forbrændingen i motoren, og ved design af DC-DC konvertere er det et ret konkret problem, da man netop omsætter energien ved at skifte mellem opbygning og nedbrydning af energien i en spole. I en H-bro for styring af motorer og i effektforstærkere indbygges altid dioder for at danne en strømvej fra pludselige skift i signalet til den induktive belastning, som en motor og en højttaler udgør.

Kondensatorens energi opbygges af en strøm af elementarladninger, som drives igennem et stadigt stigende spændingsfald u , hvilket kræver energien dW for opladning fra et givet niveau til det næste, hvor der tilføres ladningen dq . Opladning fra $u = 0$ til slutværdien $u = U_M$ giver den samlede energi ved integration af dW over spændingsændringen, og det resulterer i at den potentielle energi, der er lagret i kondensatoren bliver som vist herunder.

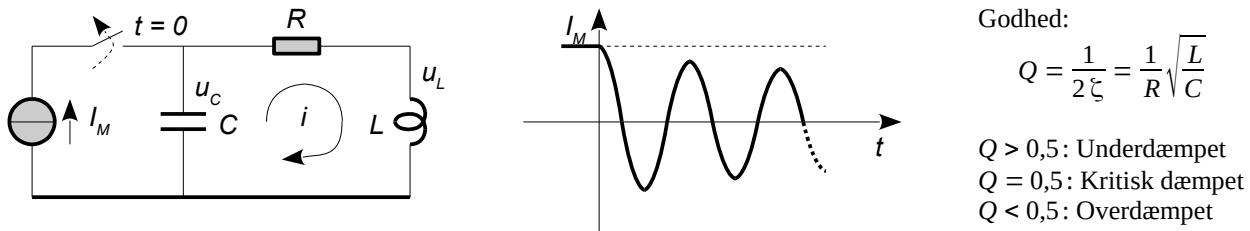
$$\frac{dW}{dq} = u \frac{dq}{dt} \Rightarrow dW = C u du \Rightarrow W_C = \int dW = \int_0^{U_M} C u du = C \left[\frac{u^2}{2} \right]_0^{U_M} \Rightarrow W_C = \frac{1}{2} C U_M^2$$

Spolens energi opbygges ved strømmen, der igen kan skrives ved ladningen dq , der føres igennem spændingsfaldet u , men nu med ladningen dq givet som strømmen i over tiden dt . Spændingen over spolen udtrykkes ved definitionsligningen og integration fra stilstand ($i = 0$) til maksimal strøm I_M giver selvinduktionens oplagrede kinetiske energi.

$$\begin{aligned} dW &= u dq & u &= L \frac{di}{dt} & W_L &= \int dW = \int_0^{I_M} L i di = L \left[\frac{i^2}{2} \right]_0^{I_M} & \Rightarrow W_L &= \frac{1}{2} L I_M^2 \\ \frac{dq}{dt} &= i dt & \frac{dW}{dt} &= L i di & & & & \\ dW &= ui dt & dW &= L i di & & & & \end{aligned}$$

1.12 Dæmpet svingning

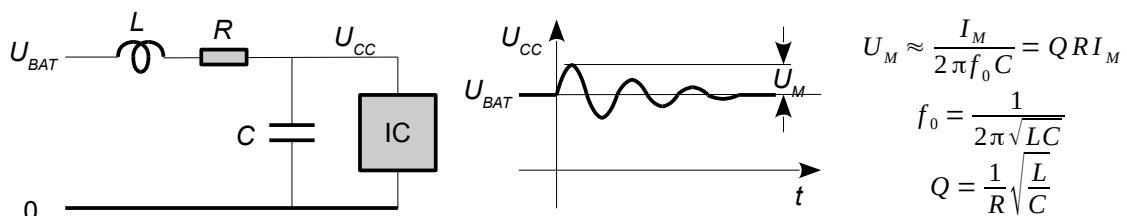
Energiudvekslingen mellem kondensatoren og spolen er ikke tabsfri, for ingen komponent er uden tab, så svingningen vil klinge ud i takt med at energien omsættes til varme. Det er især spolen, der afgiver energi til omgivelserne, dels på grund af modstandsværdien i den tråd, som spolen er viklet med, men også på grund af afgivet energi gennem magnetfeltet, der kan danne hvirvelstrømme i metaldele i nærheden af spolen. Kondensatoren har en lav tabsmodstand, men den er ikke nul. Tabene vises herunder ved modstanden R , som repræsenterer de samlede tab uanset årsagen. Effekttabet er proportionel med strømmens amplitude, så så strømmens amplituden vil gradvist blive reduceret mod nul med en fast brøkdel for hver periode, og det kan vises at lede til en eksponentielt aftagende amplitude.



Figur 51 – Amplituden af svingning vil aftage eksponentielt mod nul på grund af tabene.

Et mål for dæmpningen beskrives ved godheden Q (*quality factor*), der for kredsløb med en høj godhed omrentlig angiver antallet af perioder af svingningen inden amplituden er nær nul³⁵.

Kredsløbet herunder repræsenterer en hyppigt mødt situation hvor en elektronisk komponent (IC) har en kondensator som afkobling og bliver effektforsyнет via en kobberbane på et printkort. Banen repræsenterer en selvinduktion, så der dannes en svingningskreds, der dæmpes af modstanden i kobberbanen. Hvis effektforsyningens indre modstand regnes lig med nul, er kredsløbet identisk med det der lige er blevet analyseret. Et pludseligt strømtræk (eller et pludseligt ophør) vil starte en dæmpet svingning og amplituden kan blive generende for funktionen da den er overlejret U_{BAT} og derfor giver en variation i komponentens driftsbetingelse.



Figur 52 – En analog eller digital komponent er effektforsyнет gennem en kobberbane på et print.

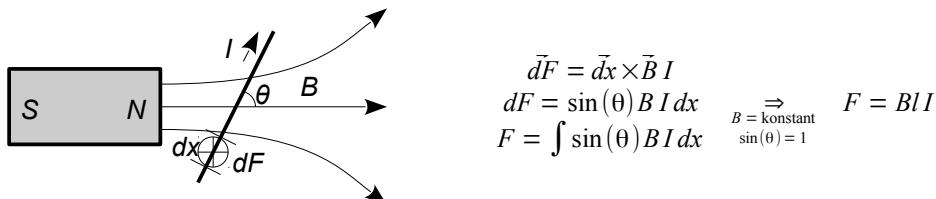
En kort leder har en selvinduktion af størrelsesordenen 1 nH/mm (se side 38) så en bane på 100 mm vil kunne nå en selvinduktion på 100 nH , og med $0,5 \text{ mm}$ bredde og $35 \mu\text{m}$ tykkelse af kobberbanen er DC modstanden på $R = 0,1 \Omega$ (se side 10). Med en kondensator på $C = 100 \text{ nF}$ bliver resonansen $f_0 = 1,6 \text{ MHz}$. En pludselig strømændring på $I_M = 0,1 \text{ A}$ giver en startværdi af amplituden på $U_M \approx \pm 0,1 \text{ V}$, hvor fortegnet er positivt ved en strømreduktion. Godheden bliver $Q = 10$ så svingning en klinger ud over cirka $6 \mu\text{s}$.

Kondensatoren C er en nødvendighed for at få stabilt fungerende kredsløb. Digitale kredse har fx kortvarige strømtræk når de skifter niveau, og det strømtræk kan ikke leveres af effektforsyningen på grund af den induktive egenskab af tilledningen. Der må derfor være en kondensator nær ved elektronikken for at leverer strømmen, og så har man skabt en svingningskreds. Analoge kredse har brug for en lav impedans i effektforsyningen for at være stabile, men der opstår kun sjældent stærke strømtræk.

35 http://en.wikipedia.org/wiki/Q_factor.

1.13 Magnetiske kræfter

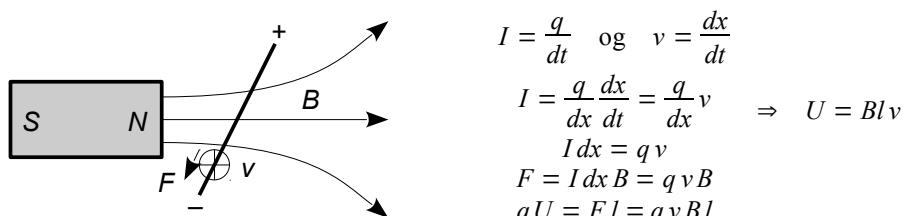
En elektrisk ladning bliver som nævnt på side 9 påvirket med en mekanisk kraft i et elektrisk felt, og hvis ladningen er i bevægelse vil den også blive påvirket af en kraft fra et ydre magnetfelt, og det anvendes i motorer, drejespoleinstrumenter og højtalere. Kraften beskrives af Lorentz³⁶ som kraftbidraget dF fra vektoren dx på lederen, ved magnetfeltets induktion B og strømmen I i lederen. Hvis feltretningen er vinkelret på den elektriske leder bliver kraften maksimal da sinus bliver én. Hvis B er konstant, så findes en vigtig ligning, hvor *kraftfaktoren* Bl gange med strømmen vil søge at skubbe lederen sidelæns i magnetfeltet med kraften F .



Figur 53 – En elektrisk leder i et magnetfelt vil blive påvirket af en mekanisk kraft, hvis der løber en strøm og kraftens retning er ind mod papiret med den viste orientering af strøm og magnetfelt.

De to vektorer dx og B udspænder et plan (i tegningen er det parallelt med papirets plan) og krydsproduktet angiver at kraften står vinkelret på planet³⁷ (den peger væk fra iagttageren). Den skalære værdi af kraften findes ved produktet af dx , B og sinus til vinklen imellem de to. Den resulterende kraft beregnes ved at integrere langs med den lederen, med længden l . Der skal tages hensyn til variationen i magnetfeltet og vinklen mellem leder og magnetfelt.

En elektrisk ladning q i bevægelse repræsenterer en elektrisk strøm I , givet ved at ladningen q , der passerer et punkt i tidsrummet dt . Relationen er også gyldig for ladningen i en leder, der bevæger sig, idet ladningen blot følger med lederen. Hastigheden v er givet af den strækning dx , der gennemløbes i tidsrummet dt . Hvis de to relationer kombineres fås et udtryk for at spændingen over en lederrørstrekning er givet ved kraftfaktoren Bl og hastigheden [HJ-87]. Resultatet udnyttes i den elektriske dynamo, der omsætter en rotation af en leder i et magnetfelt til en elektrisk strøm.



Figur 54 – En leder i bevægelse vinkelret på magnetfeltet, der på tegningen er en bevægelse ind mod papiret, vil give en spændingsforskell over lederen.

Produktet $I dx$ angiver en strøm gennem en strækning, og efter Laplaces lov giver det en kraftpåvirkning hvis der er et magnetfelt B omkring lederen. Ved hjælp af ligningen ovenfor kan det udtrykkes ved ladningen q og dens hastighed v ved at gange med B . Kraften påvirker de frie ladningsbærere i lederen ved at dreje deres bevægelse væk fra den retning vektorerne for hastigheden peger, så ladningen dirigeres imod den ene ende af lederen, og dermed opstår der en spændingsforskell over de to ender af lederen. I tegningen er kraften rettet langs med papiret (pilen), så de negative elektroner ophobes i den nederste ende. Spændingsforskellen er givet ved det arbejde, som det elektriske felt udfører ved at en ladning q føres igennem en spændingsforskell U . Arbejdet er givet ved kraft gange vejlængde, og må være numerisk lig med det arbejde, som det magnetiske felt udfører ved at kraften F virker over længden l af lederen.

36 Kilder: <http://en.wikipedia.org/wiki/Magnetism> og HJ-86.

37 Kraftens retning kan bestemmes ved de første tre fingre på højre hånd; tommel-, pege- og langfinger, der holdes så de udspænder et tre-retvinklet koordinatsystem. Tommelfingeren peger i ledernes retning (dx), pegefinger angiver magnetfeltets retning (B) og langfinger viser den retning kraften påvirker lederen (dF).

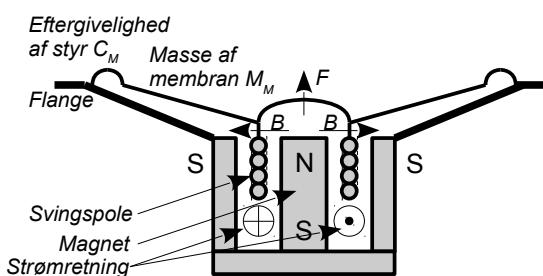
1.13.1 Den elektro-dynamiske højtalere

De to relationer for magnetiske kredsløb gør det muligt at forstå et kombineret elektrisk-mekanisk system som den elektro-dynamiske højtalere, der vises i snit herunder. Afsnittet er af introducerende karakter, og for en uddybende analyse henvises til litteraturen^{38,39,40,41}.

Kort fortalt er højttaleren et mekanisk system, hvor en masse accelereres af en elektrisk kraft, som følge af en elektrisk strøm, og accelerationen omsættes til hørbare lydbølger.

En membran (*diaphragm*) er fleksibelt ophængt i nogle styr så den kan bevæges i én retning og dermed flytte luften. Membranen er kegleformet for at opnå stor styrke uden at massen forøges. En spole (svingspole) er monteret på membranen og befinner sig i et magnetfelt, så en strøm i spolen vil påvirke den med en mekanisk kraft F , der er proportional med den elektriske strøm i spolen. Massen af svingspolen og den medsvingende luft inkludere almindeligvis i massen M_M . Der indgår flere komponenter i en højttaler, men det er tilstrækkeligt at koncentrere analysen om mag-

netsystemets kraftfaktor Bl (se side 45) og massen af det bevægelige system M_M .



$$\begin{aligned} F &= Bl i = M_M \frac{dv}{dt} \Rightarrow v = \frac{Bl}{M_M} \int i dt \\ i &= I_0 \sin(\omega t) \Rightarrow v = \frac{Bl I_0}{\omega M_M} \cos(\omega t) \\ \omega &= 2\pi f \\ u &= Blv = \frac{(Bl)^2 I_0}{\omega M_M} \cos(\omega t) \\ f > f_0 & \quad f_0 = \frac{1}{2\pi\sqrt{M_M C_M}} \end{aligned}$$

Figur 55 – Den elektro-dynamiske højtalere genererer en kraft på membranen som accelererer dens masse op. Kraften genereres af strømmen i svingspolen og magnetfeltets induktion i det smalle luftgab. Derved bevæges svingspolen i magnetfeltet og der induceres en spænding over spolen.

Membranens hastighed er omvendt proportional med frekvensen, hvilket kompenseres af at koblingen fra membran til den omgivende luft er proportional med frekvensen, så lydtrykket foran højttaleren er konstant i det frekvensområde hvor formlen er gyldig.

Magnetfeltets styrke og længden af tråden på svingspolen udgør kraftfaktoren Bl , der altid opgives i databladet for en højttaler og en typisk værdi er $Bl = 5 \text{ N/A}$. For en mindre bashedhøjttaler er massen af membran og svingspole på $M_M = 0,02 \text{ kg}$. Hvis der drives en strøm i svingspolen med en amplitude på $I_0 = 1 \text{ A}$ ved en harmonisk svingning med frekvensen $f = 100 \text{ Hz}$ beregnes kraftens amplitude til $F = 5 \text{ N}$, og det giver en amplitude af membranens svingning på $v = 0,4 \text{ m/s}$. Ved en højere frekvens aftager hastigheden.

Når membranen svinger vil der induceres en elektrisk spænding over svingspolen og med de givne værdier er den inducerede spændingsamplitude $u = 2 \text{ V}$.

Ved lave frekvenser er det membranens styr, der bestemmer bevægelsen som en balance mellem kraften fra svingspolen og deformationen af den fjeder som styret udgør. En typisk specifikation af fjedrens eftergivelighed er $C_M = 1 \text{ mm/N}$, så med eksemplets værdier, hvor kraften oscillerer med en amplitude på 5 N, vil membranens udsving være $\pm 5 \text{ mm}$ uanset frekvensen.

Grænsen mellem en bevægelse givet ved accelerationen af massen M_M og en bevægelse givet ved eftergiveligheden C_M udgør højttalerens resonansfrekvens f_0 , der med eksemplets værdier er på 36 Hz.

38 Tore Skogberg "Elektroakustik", 2014, der benyttes ved fagene ATLYAK og ATAUTE på sjette semester.

39 Knud Rasmussen "Højttalere", Danmarks Tekniske Universitet, Note 2109, 1996.

40 W. Marshall Leach "Introduction to Electroacoustics & Audio Amplifier Design", Third Edition, Kendall/Hunt, 2003.

41 Joseph D'Appolito "Testing Loudspeakers", Audio Amateur Press, 1998.

1.14 Definitioner

Elektronikken gør brug af nogle definitioner, som skal gennemgås i det følgende. Efter decibel, der benyttes flittigt ved elektriske, mekaniske og akustiske systemer, beskrives beregning af et signals middelværdi og effektive værdi, og det vises hvordan beregningen foretages i en computer.

1.14.1 Decibel

Elektronik og akustik gør udstrakt brug af signalers amplitude relateret til en reference, som kan være indgangssignalet, udgangssignalet eller en underforstået værdi. Relateret betyder her, at der dannes et forhold imellem signalet og referencen, og da det meget ofte skal bruges i forbindelse med menneskers opfattelse af tingenes tilstand er det naturligt at benytte den logaritmiske værdi af forholdet. Det skyldes at menneskets sanser reagerer på den relative ændring i en fysisk variabel og ikke ændringen selv, så en ændring på 10 % føles nogenlunde lige stor uanset om det er en ændring ud fra en lille påvirkning eller ud fra en stor påvirkning.

Definitionen af decibel har sit ophav fra telefonindustrien, hvor en relation udtrykt i enheden *neper* benyttede den naturlige logaritme til forholdet mellem to effektudtryk. I elektronikken og akustikken benyttes derimod 10-tals logaritmen, men relationen er stadig defineret ud fra forholdet mellem to effekter, dog ofte udtrykt ved kvadratet på spænding, strøm, etc. Enheden bliver bel (B), men af historiske årsager foretrækkes underdelingen decibel, med enhedssymbolet dB, som betyder 10 gange logaritmen til forholdet imellem to effekter, eller 20 gange forholdet mellem to spændinger (eller mellem to strømme).

Definition	Kommentar
$L = 10 \cdot \log_{10} \left(\frac{P}{P_0} \right) \text{ dB}$	Generel definition af "decibel" som ti gange tals-logaritmen til forholdet mellem effekter.
$L = 10 \cdot \log_{10} \left(\frac{P}{P_0} \right) \text{ dBm}, \quad P_0 = 1 \text{ mW}$	Indenfor radiokommunikation benyttes "dBm" som absolut målestørrelse af udstrålet effekt med 0 dBm defineret som 1 mW.
$L = 20 \cdot \log_{10} \left(\frac{U}{U_0} \right) \text{ dB}$	Afledning af "decibel" som tyve gange logaritmen til forholdet mellem spændinger, strømme, hastigheder, etc.
$L = 20 \cdot \log_{10} \left(\frac{U}{U_0} \right) \text{ dBu}, \quad U_0 = 0,775 \text{ V}$	I lydverdenen betegner "dBu" den spænding, der giver 1 mW i 600 Ω. Modstanden bruges ikke længere, men spændingen er bibeholdt.
$L = 20 \cdot \log_{10} \left(\frac{p}{p_0} \right) \text{ dB SPL}, \quad p_0 = 20 \cdot 10^{-6} \text{ Pa}$	Indenfor akustikken er lydtrykket 20 μPa det svageste niveau, der kan opfattes ved 1 kHz.

To signaler med ens amplitude, $P = P_0$, giver forholdet $P/P_0 = 1$ så logaritmen bliver nul og den logaritmiske niveauforskæl bliver 0 dB. En effektfordobling ($P_1 = 2P_0$) svarer til 3 dB, som ved spændingsudtryk bliver til en ændring på $U_1 = \sqrt{2} \cdot U_0 = 1,41 \cdot U_0$. En fordobling af spændingen svarer til 6 dB, der igen svarer til fire gange forøgelse af effekten.

Omregning fra brug af spænding, frem for definitionens effekt, benytter relationen til spænding gennem $P = U^2/R$ (og ved strøm $P = I^2R$). Ved samme modstandsværdi for signal og reference bliver udtrykket lig med 20 gange forholdet mellem spændingerne, fordi $\log(x^2) = 2 \cdot \log(x)$.

$$10 \cdot \log_{10} \left(\frac{P_1}{P_0} \right) = 10 \cdot \log_{10} \left(\frac{U_1^2/R}{U_0^2/R} \right) = 10 \cdot \log_{10} \left(\left[\frac{U_1}{U_0} \right]^2 \right) = 20 \cdot \log_{10} \left(\frac{U_1}{U_0} \right)$$

Tabellen er ikke komplet og der er desværre en del uenighed om hvad det enkelte symbol står for. Som eksempel benyttes dBm både indenfor radiokommunikation for en udsendt effekt på 1 mW, mens den ved lydformål står det for den spænding, der afsætter 1 mW i en belastning på 600Ω . Det svarer til 0,775 V, som efterfølgende er blevet en standard i lydverdenen, uanset om det er brugt 600Ω eller ej. En del referencer kalder den sidste definition for dBu. Som hovedregel bør man angive hvad man mener, og ikke regne med at læseren kan gætte det.

Et typisk referenceniveau i den professionelle lydverden er 4 dBm. Med udgangspunkt i $U_0 = 0,775 \text{ V}$ ses det at svare til en amplitude af signalet på $U_0 \cdot 10^{4/20} = 1,23 \text{ V}$.

Selv om decibel er defineret fra et forhold mellem to størrelser, så benyttes enheden også som en absolut angivelse af spænding, effekt eller lydtryk. Et eksempel er dBV, der ikke står i tabellen, og som benytter referencen 1 V. Ved digital lydtransmission benyttes et 0 dB niveau, der ligger nogle få dB under overstyringsgrænsen, som igen kan relateres til det spændingsniveau den aktuelle hardware kører på. Her er grænsen sat af overstyring i det digitale system⁴².

I akustikken angives et lydtryk ud fra referencelydtrykket 20 μPa , der svarer nogenlunde til den mindste vibrationsamplitude i luften som den menneskelige hørelse kan opfatte ved 1 kHz. Det resulterende lydniveau kaldes dB SPL, hvor akronymet SPL betyder *Sound Pressure Level* og underforstår den valgte reference.

Hvis vi ignorerer problemer med linearitet så er den øvre grænse for amplituden af lydtryk givet ved det atmosfæriske lufttryk på 100 kPa, hvor lydtrykket varierer mellem 0 og 200 kPa. Det svarer til lydniveauet 194 dB SPL som den absolut øverste grænse – måske med undtagelse af lydtrykket fra en ekspllosion.

Et akustisk signal med amplituden $p = 1 \text{ Pa}$, det vil sige et lydsignal, der svinger $\pm 1 \text{ Pa}$ omkring det atmosfæriske lufttryk, har lydniveauet 94 dB, hvilket undertiden skrives 94 dB re. 20 μPa . Det er det lydniveau, der kan måles direkte udenfor læberne når der tales højt, men uden at råbe.

Akustikken benytter i vid udstrækning en vægtning af frekvensområdet inden der måles, og det mest kendte eksempel er dBA, der både angiver en logaritmisk relation baseret på et lydniveau på 20 μPa og at der anvendes A-vejning. Det sidste er en relativt enkel filtrering af signalet, hvor frekvenser i området 1 til 5 kHz prioriteres og signalandele udenfor dette interval bliver dæmpet. Filtreringen angives undertiden som dB(A). Ved introduktion af vejning var idéen at en A-vejning skulle svare til den menneskelige hørelse ved lavt lydniveau, B-vejningen ved middel lydniveau og C-vejningen ved højt lydniveau, men de to sidste bruges ikke mere. Til gengæld er der kommet en D-vejning til, som skal eftergøre den psykoakustiske genevirkning som flystøj har på mennesker. For at gøre forvirringen komplet benyttes A-vejning i dag for vurdering af skadenvirkningen af et højt lydniveau; det er baseret på solid erfaringsmateriale og det faktum at A-vejningen er vidt udbredt, men det gør det ikke lettere at forklare sammenhængen.

Et af problemerne med decibel er at målemetoden angives i enheden, for standardiseringsmaterialet bag SI siger direkte at enheden ikke må afspejle målemetoden, så der burde kun være én definition af decibel, men enheden er slidt til ved praktisk brug, så det er altså ikke tilfældet⁴³.

Den modsatte relation, fra decibel til spænding findes ved at dele med 20 dB ved strøm, spænding og lydtryk eller 10 dB ved effekter, tage eksponentalfunktionen base 10 og gange med referencen. Som et eksempel en omregning fra lydtrykniveau L i decibel til lydtrykkets amplitude p i pascal.

$$p = p_0 \cdot \exp_{10} \left(\frac{L}{20 \text{ dB}} \right)$$

⁴² Her svinger signalet mellem laveste og højeste talværdi og amplituden kan ikke øges.

⁴³ Der er en parallel til bit og byte i computerverdenen hvor SI kun angiver bit som informationsenhed. Reklameverdenen har så skabt byte, der ikke har nogen officiel definition, men oftest tolkes som 8 bit, og dertil kilobyte, der snart betyder 1.000 byte og 1.024 byte, og der er ikke enighed om hvorvidt præfix er k eller K eller sågar Ki. Ved datatransmission kan en hastighed fx opgives til "100 Mb", der kan være alt fra 100 Mbit/s og op til 105 Mbytes/s (da 1.024×1.024 er 1.048.576). Der er sågar problemer med præfix M.

1.14.2 Middelværdi (AVE)

Middelværdien af et signal angiver den gennemsnitlige værdi, som signalet svinger omkring. Et signal beskrevet ved talserien 1, 2, 3, ... 9, 10 har en middelværdi på 5,5. En harmonisk svingning som sinus svinger symmetrisk omkring nul, så den har middelværdien nul.

$$\text{Middel} = \frac{1+2+\dots+10}{10} = 5,5$$

I den matematiske formulering skal signalet integreres over tiden T , som kan være en periode af signalet, eller et helt antal perioder, eller blot "længe nok" til at usikkerheden jævnnes ud og der nås et tilfredsstillende estimat. Herunder vises det for en spænding u , men det gælder for alle variable og ikke kun ved elektronik. På engelske kaldes middelværdi for average.

$$\text{Middelværdi: } U_{\text{AVE}} = \frac{1}{T} \int_0^T u dt$$

En firkantsvingning, der skifter mellem U_0 i den første halvdel af perioden og nul i resten af perioden, vil have middelværdien $U_0/2$.

$$u(t) = \begin{cases} U_0 & \text{for } 0 \leq t < T/2 \\ 0 & \text{for } T/2 \leq t < T \end{cases} \Rightarrow U_{\text{AVE}} = \frac{1}{T} \int_0^{T/2} U_0 dt = \frac{U_0}{2}$$

En del måleinstrumenter viser en AC spænding som om der var foretaget en RMS måling, men de mäter en middelværdi og ganger resultatet med 1,111 så visningen bliver korrekt ved sinus.

Det målte signal helbølge-ensrettes, hvilket matematisk svarer til den numeriske værdi, og herefter findes middelværdien. Ved beregning af middelværdi for en helbølge-ensrettet sinus er den numeriske værdi af sinus for et argument fra 0 til π præcis det samme som fra π til 2π , så integralet kan beregnes som to gange integralet over alene den første af perioderne. Med grænserne indsat bliver resultatet $U_{\text{AVE}} = 0,637 \cdot U_0$. Hvis det skal vises på displayet som 0,707 $\cdot U_0$ skal den målte værdi ganges med 1,111.

$$U_{\text{AVE}} = \frac{1}{T} \int_0^T U_0 |\sin(\omega t)| dt = \frac{2U_0}{T} \int_0^{T/2} \sin\left(\frac{2\pi}{T}t\right) dt = \frac{2U_0}{T} \frac{T}{2\pi} \left[-\cos\left(\frac{2\pi}{T}t\right)\right]_0^{T/2} = \frac{2}{\pi} U_0$$

1.14.3 Effektivværdi (RMS)

Med en strøm i en modstand afsættes effekt og den varmes op. Det er relativt let at beregne den effekt, der over tid akkumuleres til energi, når det er en jævnspænding eller en jævnstrøm, men det bliver lidt mere besværligt når signalet oscillerer, har karakter af støj eller kun eksisterer som korte pulser. Derfor har man indført definitionen af den effektive værdi af strømmen (eller spændingen), som netop afsætter den samme effekt som en DC strøm ville gøre. På engelsk aldes det for RMS, stående for root mean square, som beskriver metoden i beregningen af effektivværdien.

$$\text{Effektivværdi: } U_{\text{RMS}} = \sqrt{\frac{1}{T} \int_0^T u^2 dt}$$

Hvis signalet ikke er periodisk er det almindeligt at bevare definitionen og så blot kræve brug af en tilstrækkelig lang til integrationen for at resultatet giver fornuft. Det er ofte muligt at angive hvor lang tid det ikke-periodiske signal varer, selv om det indebærer en del antagelse. Ved støj, hvor signalet aldrig repeterer sig selv, og heller ikke klinger ud, integrerer man over en tid, der er lang sammenlignet med tidskonstanten af den nedre grænsefrekvens⁴⁴ (se side 52).

⁴⁴ Indenfor akustikken måles et støjfyldt signal ved at integrere over to gange 5 sekunder, hvor den første periode får beregningen til at falde på plads og den anden periode benyttes til selve målingen.

Effekten afsat i et givet øjeblik kaldes for p og er givet ved produktet af spænding og strøm, hvor alle variable er funktioner af tiden, og det antages herunder at signalerne er i fase. I lighed med tidligere kan effekten udtrykkes ved kvadratet på spænding eller strøm, så effekten er positiv eller nul, men aldrig negativ.

$$p = ui = \frac{u^2}{R} = i^2 R \quad u \text{ og } i \text{ er i fase}$$

For en periodisk funktion beregnes middelværdien af den effekt, der afsættes i perioden T som energien over tidsrummet delt med tiden. Energien beregnes ved at summere små bidder at energi $dW = p dt$ over perioden T . Hvis vi benytter spændingen som eksempel bliver integralet givet ved af kvadratet på spændingen delt med modstanden.

$$P = \frac{1}{T} \int_0^T p dt = \frac{1}{T} \int_0^T \frac{u^2}{R} dt$$

Den beregnede middeleffekt relateres nu til en gennemsnitsværdi U_{RMS} , der ville give samme resultat som den gennemsnitligt afsatte effekt. Ved at tage kvadratroden på begge sider bestemmes den effektive værdi af spændingen.

$$P = \frac{U_{RMS}^2}{R} = \frac{1}{T} \int_0^T \frac{u^2}{R} dt \Rightarrow U_{RMS}^2 = \frac{1}{T} \int_0^T u^2 dt \Rightarrow U_{RMS} = \sqrt{\frac{1}{T} \int_0^T u^2 dt}$$

Herunder vises eksempler på beregningen med signaler af typen sinus og firkant.

En sinus svingning med en amplitude på U_0 er et typisk eksempel på beregning af effektivværdi. Kvadratet på en sinus kan ved tabelopslag skrives med cosinus til den dobbelt vinkel.

$$u = U_0 \sin(\omega t) \quad \omega = \frac{2\pi}{T} \Rightarrow U_{RMS} = \sqrt{\frac{1}{T} \int_0^T U_0 \sin^2(\omega t) dt} = \sqrt{\frac{1}{T} \int_0^T U_0 \frac{1-\cos(2\omega t)}{2} dt}$$

Integralet af cosinus giver nul over et helt antal perioder så er det kun den konstante faktor, der giver et bidrag. Med grænser indsæt findes det velkendte udtryk, hvor den effektive værdi af en spænding (eller en strøm) er givet ved amplituden delt med kvadratroden af to.

$$U_{RMS} = \sqrt{\frac{1}{T} U_0^2 \left[\frac{1}{2} t \right]_0^T} = U_0 \sqrt{\frac{1}{2}} = \frac{U_0}{\sqrt{2}} \Rightarrow U_{RMS} \approx 0,707 \cdot U_0$$

En firkantsvingning med amplituden én i halvdelen af intervallet og minus én i resten kvadreres til værdien én i hele perioden. Integralet fra nul til T giver $U_0 T$, der skal ganges med $1/T$ og efterlader et resultat på U_0 . Det er et intuitivt korrekt resultat, for det er lige meget om strømmen løber den ene vej igennem modstanden eller den anden vej.

$$\begin{aligned} u(t) &= U_0 & \text{for } 0 \leq t < T/2 \\ u(t) &= -U_0 & \text{for } T/2 \leq t < T \end{aligned} \Rightarrow U_{RMS} = \sqrt{\frac{1}{T} \int_0^T U_0^2 dt} = U_0$$

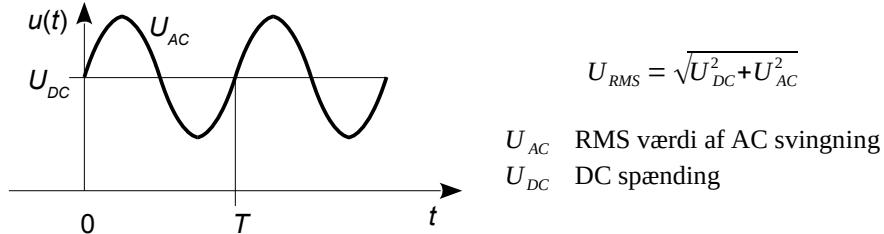
En firkantsvingning, der skifter mellem U_0 i den første halvdel af perioden og er nul i resten vil have en effektiv værdi på $0,707 \cdot U_0$.

$$\begin{aligned} u(t) &= U_0 & \text{for } 0 \leq t < T/2 \\ u(t) &= 0 & \text{for } T/2 \leq t < T \end{aligned} \Rightarrow U_{RMS} = \sqrt{\frac{1}{T} \int_0^{T/2} U_0^2 dt} = \frac{U_0}{\sqrt{2}}$$

For afklaring af hvor lang tid der skal integreres over i en reel situation, hvor man ikke på forhånd kender signalets frekvens, henvises til diskussionen side 52.

1.14.4 Signal med DC og AC komponenter

Meget ofte vil et signal bestå af en DC værdi overlejret med et AC signal hvor man kender RMS værdien af AC signalet og skal beregne RMS værdien for det sammensatte signal. Beregningen er ikke intuitiv og foretages ved at kvadrere RMS værdien af både DC signalet og AC signalet før man lægger de to sammen og derefter tager kvadratroden. Det kaldes for "RMS addition".



Figur 56 – En svingning kan bestå af en AC spænding der er overlejret på et DC niveau. Her beskrives AC spændingen ved en RMS værdi U_{AC} og DC niveauet ved U_{DC} .

Signalet beskrives ved et DC niveau U_{DC} , der er overlejret med et AC signal, hvor amplituden beskrives ud fra RMS værdien U_{AC} . Ved at bruge Fourier serierne fra næste kapitel kan AC signalet repræsentere enhver periodisk svingning, så den følgende udledning er generel.

$$u(t) = U_{DC} + \sqrt{2} U_{AC} \sin(\omega t) \quad \text{hvor } U_{AC} \text{ er RMS værdien}$$

Udtrykket for $u(t)$ indsættes i formlen for beregning af RMS værdien.

$$U_{RMS} = \sqrt{\frac{1}{T} \int_0^T u^2(t) dt} = \sqrt{\frac{1}{T} \int_0^T [U_{DC} + (\sqrt{2} U_{AC} \sin(\omega t))^2] dt}$$

Parentesen ganges ud og integrationen deles op.

$$U_{RMS} = \sqrt{\frac{1}{T} \int_0^T U_{DC}^2 dt + \frac{1}{T} \int_0^T (\sqrt{2} U_{AC})^2 \sin^2(\omega t) dt + \frac{1}{T} \int_0^T 2 U_{DC} \sqrt{2} U_{AC} \sin(\omega t) dt}$$

Første led er uproblematisk fordi spændingen U_{DC} ikke varierer over tid, så integralet bliver U_{DC}^2 gange med tiden for perioden og ved multiplikation med $1/T$ er der U_{DC}^2 tilbage. Andet led kan reduceres til $2 \cdot U_{AC}^2$ gange integralet af kvadratet på en sinus, der tidligere er fundet til $1/2$ så det led giver U_{AC}^2 . Tredje led bliver nul fordi sinus integrerer over en hel periode er nul.

$$U_{RMS} = \sqrt{U_{DC}^2 + U_{AC}^2}$$

1.14.5 Misforståelsen RMS watt

Ved beregning af udgangseffekten fra en effektforstærker benyttes trømmens effektive værdi I_{RMS} ganget med spændingens effektive U_{RMS} , og det ligger lige for at kalde produktet for "RMS watt". Det lyder logisk, men det er det ikke, selv om det benyttes bredt i brochuremateriale. Det korrekte er angive middelværdien af effekten.

Strøm og spænding skrives som effektivværdier og udtrykkene samles under en fælles kvadratrod hvor produktet $i u$ repræsenteres som effekt p . Det resterende udtryk er den ovenstående definition af middelværdien blot nu udtrykt som middelværdien af effekten.

$$P_{AVE} = I_{RMS} U_{RMS} = \sqrt{\frac{1}{T} \int_0^T i^2 dt} \sqrt{\frac{1}{T} \int_0^T u^2 dt} = \frac{1}{T} \sqrt{\left(\int_0^T p dt \right)^2} = \frac{1}{T} \int_0^T p dt$$

Det er her forudsat at strøm og spænding er i fase; ellers skal der inkluderes $\cos(\theta)$, hvor θ er fasen mellem strøm og spænding. Det betyder at den afgivne effekt er nul ved $\theta = 90^\circ$.

1.14.6 Numerisk beregning

Ved dataopsamlingsudstyr (*data acquisition system*) beregnes signalets middelværdi og effektive værdi ud fra data målt på diskrete tidspunkter. Signalet opfattes som om det var kontinuerligt med en øjebliksværdi, der kun varierer lidt fra et måletidspunkt til det næste. Signalet $u(t)$ er kun kendt ved N diskrete værdier over tidsforløbet T .

$$U_{AVE} = \frac{1}{N} \sum_{n=0}^{N-1} u\left(\frac{nT}{N}\right) \quad \text{og} \quad U_{RMS} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} u\left(\frac{nT}{N}\right)^2}$$

Signalet opsamles (*samples*) ved frekvensen $f_s = N/T$, og for at signalet kan opfattes som konstant mellem tidspunkterne skal det være båndbegrænset, hvilket i praksis sikres ved at indskyde et filter i signalvejen. I den teoretiske litteratur nævnes Nyquist eller Shannon for det krav, at der ikke må forefindes frekvenskomponenter over halvdelen af f_s .

Indgangssignalet antages i det følgende at være en spænding $u(t)$, der måles over N lige store intervaller fra $t = 0$ til $t = T$, og spændingen repræsenteret ved måleværdien til tidspunkterne nT/N hvor n er heltalig. Det antages at $u(t)$ varierer så langsomt indenfor tidsrummet T/N at det er tilladeligt at regne værdien for konstant.

$$t = \frac{nT}{N} \quad \text{hvor } n = 0, 1, 2, \dots, N-1 \Rightarrow u(t) = u\left(\frac{nT}{N}\right)$$

Dette udtryk indsættes i definitionen af middelværdi.

$$U_{AVE} = \frac{1}{T} \int_0^T u dt = \frac{1}{T} \left[\int_0^{T/N} u(0) dt + \int_{T/N}^{2T/N} u\left(\frac{T}{N}\right) dt + \dots + \int_{(N-1)T/N}^T u\left(\frac{(N-1)T}{N}\right) dt \right]$$

Integralet fra $t = 0$ til $t = T/N$ benytter den konstante værdi $u(0)$ som skal integreres over et tidsrum på T/N . En konstant værdi deltager ikke i integrationen så opgaven reduceres til at finde integralet af dt og det giver tiden t , som med øvre og nedre grænse indsæt giver T/N . Følgelig er integralet af det første interval $u(0)$ gange med T/N , og helt tilsvarende gælder for de andre intervaller, så det resulterende udtryk bliver en sum af de målte signalværdier delt med antallet.

$$U_{AVE} = \frac{1}{T} \left[u(0) \frac{T}{N} + u\left(\frac{T}{N}\right) \frac{T}{N} + \dots + u\left(\frac{(N-1)T}{N}\right) \frac{T}{N} \right] \Rightarrow U_{AVE} = \frac{1}{N} \sum_{n=0}^{N-1} u\left(\frac{nT}{N}\right)$$

Helt tilsvarende beregnes den effektive værdi som følger.

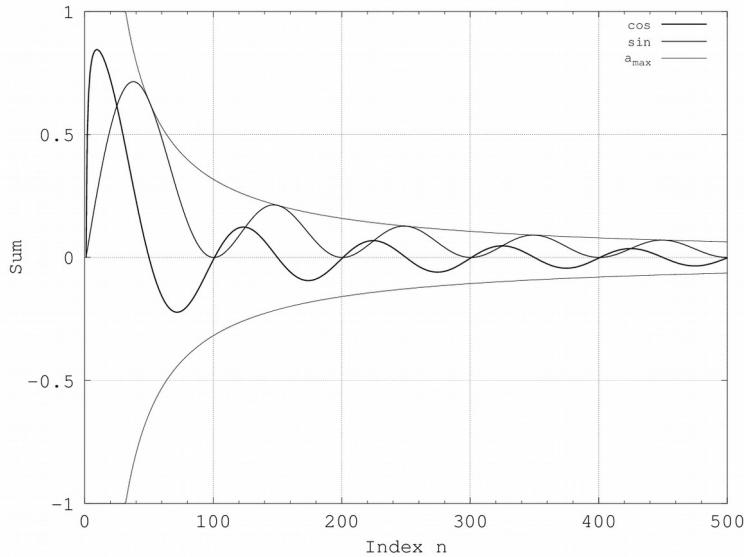
$$U_{RMS} = \sqrt{\frac{1}{T} \sum_{n=0}^{N-1} \int_{nT/N}^{(n+1)T/N} u\left(\frac{nT}{N}\right)^2 dt} \Rightarrow U_{RMS} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} u\left(\frac{nT}{N}\right)^2}$$

1.14.7 Integrationstid

Eksemplerne fra de tidligere afsnit benyttede integration over en enkelt periode af signalet, men ved praktisk brug af dataopsamlingsudstyr skal man ofte måle på signaler hvis udseende i langt de fleste tilfælde først kendt efter endt dataopsamling. Der kan fx være mange frekvenser involveret, så det ikke er givet at der kan måles over et helt antal perioder. Dertil kommer at støjsignaler, som støj fra modstande og forstærkere, slet ikke er periodiske. Det betyder at man ved måling på et signal må antage en værdi af integrationstiden, der forventes at være tilstrækkelig til den aktuelle opgave.

For at holde fejlen indenfor $\pm 1\%$ skal det samlede antal data N være mindst 32 gange perioden N_P af den laveste frekvens, der skal behandles korrekt, så der skal generelt opsamles data over lang tid for at give en valid måling.

Elektricitet



Figur 57 – En beregning af middelværdien af en cosinus (tyk streg) og en sinus (medium streg) som funktion af den tid der integreres over, med maksimum afvigelse ved tynde streger. Perioden af signalet er $N_P = 100$ samples, og der vises de første fem perioder af signalerne.

Ved brug af et kortere tidsrum end signalets periode ($N < N_P$) kan der ikke integreres over en hel periode og middelværdien ligger langt fra den korrekte værdi på nul. Ved brug af længere tidsrum ($N > N_P$) svinger resultatet for cosinus omkring det korrekte resultat på nul og krydser gennem nul for hver halve periode. Resultatet for sinus svinger mellem nul og en positiv værdi, og begge kurver holder sig indenfor et maksimum, som aftager med N . Da man ikke på forhånd ved om et signal er opbygget med sinus eller cosinus står man sig ved at indordne sig efter det værst tænkelige, og det er vist som indhyldningskurven for sinus (og minus sinus).

Indhyldningskurven for sinus er givet ved a_{max} , der med god tilnærmelse også gælder for en cosinus. Ved at specificere et krav til amplitudens værdi kan det nødvendige antal samples N beregnes. I udtrykket er N_P antallet af data i det periodiske signal og N er det samlede antal data.

$$a_{max} = \frac{N_P}{\pi N} \Rightarrow N = \frac{N_P}{\pi a_{max}} \quad \begin{array}{ll} N & \text{Nødvendigt antal samples} \\ N_P & \text{Antal samples i én periode} \\ a_{max} & \text{Krævet fejlgrænse} \end{array}$$

For et krav om en nøjagtighed på 1 % skal der benyttes godt og vel 30 gange så mange data som selve det periodiske signal opsamles over.

$$a_{max} < 0,01 \Rightarrow N > \frac{100 N_P}{\pi} \approx 32 N_P$$

1.14.8 Algoritme for middelværdi og RMS

For beregning af middelværdi og RMS ved en måleserie, er det hensigtsmæssigt at dele opgaven op i to trin. I første trin beregnes to talværdier undervejs i måleserien, og efter måleserien kan de to tal kombineres til beregning af middelværdien og den effektive værdi (RMS). Undervejs opsamles en simpel sum af måledata i variablen X_{AVG} samt en sum af kvadratet på måledata i variablen X_{SQR} . Ved afslutning af måleserien beregnes de to parametre ved følgende enkle formler.

$$X_{AVG} = \frac{X_{AVG}}{N} \quad \text{og} \quad X_{RMS} = \sqrt{X_{SQR} - X_{AVG} \cdot X_{AVG}}$$

Algoritmen kan udledes fra de tidligere fundne formler for addition af to RMS værdier.

Ved at kvadrere formlen for beregning af RMS for et sammensat signal, derefter omrokere og til sidst tage kvadratroden sider ændres formlen til beregning af RMS værdien af et AC signal.

$$U_{RMS} = \sqrt{U_{DC}^2 + U_{AC}^2} \Rightarrow U_{AC} = \sqrt{U_{RMS}^2 - U_{DC}^2} \Rightarrow U_{AC} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} u\left(\frac{nT}{N}\right)^2 - \left(\frac{1}{N} \sum_{n=0}^{N-1} u\left(\frac{nT}{N}\right)\right)^2}$$

Herunder vises en mulig opbygning af algoritmen i C. Metoderne StartConvert() og StopConvert() vil henholdsvis starte og afslutte en A/D konvertering og GetResult() aflæser rådata fra A/D konverteren. Metodernes navne må naturligvis afpasses til den aktuelle implementering.

Algoritmen skal forberedes før den kan bruges. Tællervariablen n skal løbe fra 0 til N – 1 og N skal sættes til det antal punkter, der skal benyttes i algoritmen (her er det 4096). Data læses fra A/D konverteren til variabel Vadc og algoritmen benytter tre variable for beregning af middelværdi Vavg for average, Vsqr for kvadrateret på måleværdien og Vrms for den beregnede RMS værdi. Derefter startes A/D konverteren og mikroprocessoren afventer data.

```
int      n = 0;
int      N = 4096;
float    Vadc;
float    Vavg = 0;
float    Vsqr = 0;
float    Vrms = 0;
StartConvert();
```

A/D konverteren giver interrupt, eller tilkendegiver på anden måde at data er rede, og afhængig af implementeringen kan det være krævet at stoppe A/D konverteren. Derefter læses data ind til variabel Vadc og benyttes til at danne en sum for senere beregning af middelværdien og en sum af kvadrateret for senere beregning af RMS værdien. Derefter tælles variabel n én op og hvis den er mindre end N startes A/D konverteren for måling af den næste dataværdi. Hvis n er lig med N foretages selve beregningen af middelværdien til Vavg ved division med antallet af data og tilsvarende beregnes middelværdien Vsqr af kvadraterne på data. Herefter kan RMS værdien beregnes som kvadratroden af Vsqr minus kvadrateret på Vavg. På dette tidspunkt er data klar til brug, og algoritmen kan startes forfra for en ny beregning.

```
StopConverter();
Vadc = GetResult();
Vavg = Vavg + Vadc;
Vsqr = Vsqr + Vadc*Vadc;
n = n + 1;
if (n == N)
{
    Vavg = Vavg / N;
    Vsqr = Vsqr / N;
    Vrms = sqrt(Vsqr - Vavg*Vavg);
    n = 0;
}
StartConvert();
```

1.15 Sammenfatning

Regler for spænding U , modstand R , strøm I , effekt P , energi W og tid T (side 10). Kondensator C , spole L , ladning q , spænding u og strøm i (22, 29). Energi W (43). Selvinduktionen består af tråden selv L_1 (30) plus tilledningen L_2 (31) plus spolens bidrag fra en af formlerne: L_w for Wheeler (33) eller L_s for skivespole (36). Flere formler på side 34. Kraften F på en ledet ved induktionen B , længden l og strømmen i . Spændingen U over en ledet, der bevæges med hastigheden v i induktionen B (side 45). Bevægelse, felt og ledet er vinkelrette på hinanden i sammenfatningen.

$$\begin{aligned} U &= RI \quad u_c = \frac{1}{C} \int i dt \quad u_L = L \frac{di}{dt} \quad dq = C du = i dt \\ P &= WT \quad W_c = \frac{1}{2} CU^2 \quad W_L = \frac{1}{2} LI^2 \quad P = IU = I^2 R = \frac{U^2}{R} \\ F &= Bli \quad U = Blv \end{aligned}$$

Modstand:

$$R = \rho \frac{a}{A} \quad \rho_{Cu} = 15,5 \cdot 10^{-9} \Omega m$$

$$\rho_{Fe} = 47 \cdot 10^{-9} \Omega m$$

Kapacitet mellem to plader:

$$C = \frac{\epsilon_0 \epsilon_r A}{d} \quad \epsilon_0 = 8,85 \text{ pF/m}$$

Kapacitet mellem to ledere:

$$C \approx \frac{\pi}{2} \epsilon_0 \epsilon_r l \quad \epsilon_r = 5 \text{ for plast}$$

Selvinduktion:

$$L = L_1 + L_2 + L_x \quad \mu_0 = 4\pi \cdot 10^{-7} \text{ H/m}$$

Intern induktion af en ledet:

$$L_1 = \frac{\mu_0 l}{8\pi} = (50 \text{ nH/m}) \cdot l \quad f < 20 \text{ kHz}$$

Selvinduktion af dobbeltleder:

$$L_2 \approx \frac{\mu_0 a}{\pi} \ln\left(\frac{2b}{t}\right) \quad a \gg b \gg t$$

Selvinduktion af cylinderspole:

$$L_w = \frac{\pi}{4} \frac{\mu_0 N^2 d^2}{a + 0,45d} \quad a > 0,4d$$

Selvinduktion af en kort spole:

$$L_k = \frac{\mu_0 N^2 d}{2} \ln\left(\frac{d}{a}\right) \quad d \gg a > t$$

Ydre mål:

$$D = d + t \quad \text{Diameter}$$

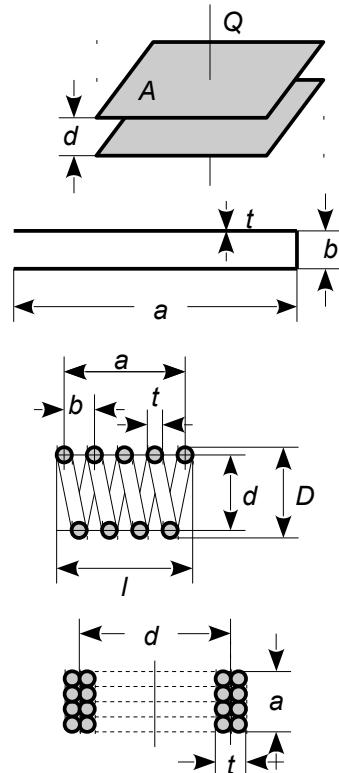
$$l = a + t \quad \text{Længde}$$



Cylinderspole. Spolen har $N = 25$ vindinger, diameteren $d = 32,5$ mm, tråden har diameteren $t = 0,25$ mm og spolens ydre længde er 12,5 mm. Wheelers formel giver: $L_{w1} = 24,0 \mu\text{H}$. Tilledningens længde er 0,2 m og trådene ligger 5 mm fra hinanden (løst snoet). Heraf $L_2 = 0,3 \mu\text{H}$. Trådlængden er $N\pi td = 2,6$ m hvortil kommer i alt 0,4 m fra tilledningen til i alt $l = 3,0$ m tråd. Heraf $L_1 = 0,15 \mu\text{H}$. Den estimerede selvinduktion bliver $L = 24,45 \mu\text{H}$. Producenten leverede to spoler med de målte værdier 24,2 μH og 24,5 μH så det estimerede ligger indenfor den grænse som de to måleværdier indikerer.

Skivespole. Spolen har $N = 4$ vindinger, diameteren $d = 180$ mm (centrum til centrum), tråden har tykkelsen $t = 0,5$ mm og isolations tykkelse er 0,5 mm så beviklingens længde er $a = 5$ mm. Kort spole giver: $L_k = 6,48 \mu\text{H}$. Tilledning har længden $a = 1$ m og de to tråde ligger $b = 5$ mm fra hinanden (løst viklet). Heraf $L_2 = 1,20 \mu\text{H}$. Trådlængden er $N\pi td = 2,3$ m hvortil kommer to gange 1 m fra tilledningen til i alt $l = 4,3$ m tråd. Heraf $L_1 = 0,22 \mu\text{H}$, så den estimerede værdi af selvinduktionen bliver $L = 7,90 \mu\text{H}$. Spolen er blevet målt til præcis den samme værdi på en målebro ved 10 kHz.

Tommelfingerregler for modstandsværdien af en kobberleder er $0,02 \Omega/\text{m}$ ved 1 mm tråddiameter (side 10), kapaciteten mellem to parallelle, isolerede ledere er 70 pF/m (side 20), og selvinduktion af en kort ledning er $1 \mu\text{H}/\text{m}$ (side 38).



1.16 Opgaver

1 – Benyt formlerne på side 8 til at beregne den opnåede ladning Q_1 for en strøm, der varierer som en sinus over en halv periode. Det vil sige at $i(t) = I_p \sin(\omega t)$, hvor $\omega = 2\pi/T$ og T er periodetiden. **Svar:** $Q_1 = I_p T / \pi$.

2 – Beregn ladningen for $I_p = 1 \text{ A}$ og $T = 20 \text{ ms}$ ved brug af formlen fra det forrige spørgsmål. **Svar:** $Q_1 = 6,4 \text{ mC}$.

3 – Brug en af formlerne side 9 til at beregne den kraft som ladningen Q_1 fra forrige spørgsmål vil påvirke en elektron med i afstanden $r_2 = 1 \mu\text{m}$. **Svar:** $F = 9,2 \text{ N}$.

4 – Brug formlerne side 11 til at beregne modstanden af en printkorts kobberbane og den afsatte effekt ved en strøm på $I = 2 \text{ A}$. Banen har tykkelsen $35 \mu\text{m}$, bredden $0,25 \text{ mm}$ og længden $0,1 \text{ m}$. **Svar:** $R = 0,18 \Omega$, $P = 0,71 \text{ W}$.

5 – Selv om kobberbanen ikke er udført på samme måde som ved en trådviklet modstand, så kan opvarmningen af printet i nærheden af kobberbanen groft estimeres ved at benytte kurven for modstanden AC01 (se side 12). Hvad bliver den tilnærmede værdi af temperaturen i nærheden af banen? **Svar:** $\theta = 125^\circ\text{C}$ ved en omgivelsestemperatur på 25°C .

6 – Temperaturmåling med en PT100 føler (se side 13) skal foretages ved en strøm på $I_0 = 1 \text{ mA}$. Beregn modstanden af PT100 føleren ved $\theta = 100^\circ\text{C}$ og spændingsændringen ΔU over føleren ved formlerne side 11. **Svar:** $R_{100} = 129,3 \Omega$ og spændingsændringen er $\Delta U = 29,3 \text{ mV}$.

7 – En strain gauge benyttes i en brokobling som vist på side 14. Påvirkningen af emnet ændrer strain gaugens længde med 1 %. Beregn ændringen i broens spænding. **Svar:** $U_1 - U_2 = 0,1 \text{ V}$.

8 – For kredsløbet side 17 opstilles ligningerne ved brug af Kirchhoffs strømlov (KCL), og det vises at resultatet bliver det samme som i eksemplet.

9 – For kredsløbet side 18 opstilles ligningerne ved brug af enten Kirchhoffs strømlov (KCL) eller spændingslov, og det vises at resultatet bliver det samme som i eksemplet.

10 – For batteriet side 19 er $U_T = 3,75 \text{ V}$ og $R_T = 0,13 \Omega$. Beregn værdierne for I_T og R_T ved et Norton ækvivalent af batteriet. **Svar:** $I_T = 28,8 \text{ A}$ og $R_T = 0,13 \Omega$.

11 – En kondensator på $C = 100 \mu\text{F}$ er opladt til $U_0 = 5 \text{ V}$, og oplades yderligere med $I = 1 \text{ A}$ over $\Delta T = 10 \mu\text{s}$. Beregn spændingen på kondensatoren efter endt opladning (se side 22). **Svar:** $U_0 = 5,1 \text{ V}$.

12 – Kondensatoren aflades nu gennem en modstand på $R = 5,1 \text{ k}\Omega$. Beregn kredsløbets tidskonstant og hvor lang tid der går inden spændingen er aftaget til 51 mV . **Svar:** $RC = 0,51 \text{ s}$ og $t_{0,01} = 2,34 \text{ s}$ (se side 23).

13 – Elektret-kondensatormikrofonen fra side 26 ændres til at arbejde fra en effektforsyning på 15 V . Beregn den maksimalt mulige værdi af R_D og den opnåede følsomhed, uden R_L . **Svar:** $R_D = 30 \text{ k}\Omega$ og $S = 75 \text{ mV/Pa}$.

14 – Udledningen af formlen for en kort spole på side 33 skal gennemføres som beskrevet i forklaringen. Sæt a til de angivne værdier (0,05 og 0,08).

15 – For Wheelers formel side 34 sættes parametrene a , b og d lig med forudsætningen for Brooks coil på samme side. Vis at formen giver meget nær det samme som er vist for Brooks coil.

16 – Selvinduktionen af en ledning angives på side 38 til tomelfingerværdien $1 \mu\text{H}$ per meter. Beregn værdien for en dobbeltleder på 10 m og sammenlign med beregningen på side 31 for dine egne værdier af parametrene d og t .

Svar: Tommelfingerreglen giver $20 \mu\text{H}$ for to ledere og med $d = 10 \text{ mm}$ og $t = 1 \text{ mm}$ findes $12 \mu\text{H}$.

17 – For en luftspole med 10 mm diameter og 10 vindinger med trådtykkelsen $0,1 \text{ mm}$, beregnes den magnetiske induktion i centrum ved strømstyrken 10 mA (se side 28). Beregn derefter den magnetiske induktion for samme spole på en toroid med en diameter på 30 mm og en relativ permeabilitet på 1000 . Brug formlerne på side 39.

Svar: $B = 4 \mu\text{T}$ uden kerne og $B = 13 \text{ mT}$ med toroidkerne.

Tip: Find den gennemsnitlige længde $l = 94 \text{ mm}$, arealet af toroidens tværsnit $A = 78,5 \cdot 10^{-6} \text{ m}^2$, reluktansen af feltets vej $R_M = 955 \text{ H}^{-1}$, den magnetiske flux $\Phi = 1,05 \cdot 10^{-6} \text{ A/H}$ og den magnetiske induktion findes fra side 28 ved (igen) at dele med arealet af toroidens tværsnit.

18 – Beregn toroiden selvinduktion. **Svar:** $L = 105 \mu\text{H}$.

19 – Toroiden på forsynes med en luftspalte på 1 mm og selvinduktionen beregns (se side 41). **Svar:** $L = 99 \text{ nH}$.

20 – LC kredsen på side 42 skal svinge på 1 MHz med en præcision på 1 %. Beregn den højeste værdi af spolens selvinduktion og foreslå en værdi af spole og kondensator når spredningskapaciteten (usikkerheden) er på 10 pF .

Svar: Spolen danner resonans på 1 MHz med 10 pF og en spole på $2,5 \text{ mH}$, så for at kunne benytte en 100 gange større værdi af kondensatoren (over 1 nF) bliver kravet til spolen $L < 25 \mu\text{H}$. Ved valget $L = 10 \mu\text{H}$ findes $C = 2,5 \text{ nF}$.

21 – Spolens DC modstand skal vælges så der kan opnås en godhed på 50 . Bestem den største værdi af spolens DC modstand (se side 44). Kobbertrådens har $0,1 \text{ mm}$ diameter og den længste mulige tråd skal bestemmes (se side 11). **Svar:** $R < 1,3 \Omega$, $a < 0,65 \text{ m}$.

22 – Tråden vikles til en spole med diameteren 5 mm . Beregn det maksimalt antal vindinger, længden af en spole i to lag og spolens selvinduktion (se fx side 37). Beregn den nye værdi af kondensatoren for at fastholde frekvensen på 1 MHz . **Svar:** $N = 42$ vindinger, som rundes ned til $N = 40$. Længden bliver $a = 2 \text{ mm}$, selvinduktionen bliver $L = 4,6 \mu\text{H}$, og der behøves $C = 5,5 \text{ nF}$.

23 – Højtaleren på side 46 har en kraftfaktor på $Bl = 5 \text{ Tm}$, massen af det bevægelige system er $M_M = 5 \text{ g}$ og styrenes eftergivelighed er $C_M = 1 \text{ mm/N}$. Beregn den frie resonans. Højtaleren drives ved 100 Hz med 1 A som strømamplitude. Beregn kraften F og hastigheden af membranen v . **Svar:** $f_0 = 71 \text{ Hz}$, $F = 5 \text{ N}$, $v = 1,6 \text{ m/s}$.

Note 1.

Elektricitet

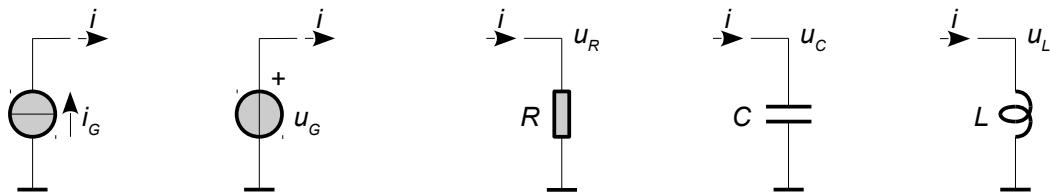
Note 2.

2 ANALYSE

Ohms og Kirchhoffs love udvikles i dette kapitel så et elektronisk kredsløbs reaktion på et AC signal kan beskrives. De passive komponenter repræsenteres ved de komplekse impedanser så vi har brug for de komplekse tal og for at kunne beskrive reaktionen på andet end simple, oscillerende signaler har vi brug for Fourier serier og Laplace transformationen som gør det muligt at arbejde med signaler der abrupt skifter værdi. Til sidst vises det at operationsforstærkeren og større analoge systemer med tilbagekoppling kan betragtes som ikke-tilbagekoppledte systemer uden at der mistes information. Endelig vises det at det matematiske værktøj kan benyttes på meget andet end lige de elektriske systemer idet der er et nært slægtsskab med mekaniske, akustiske og termiske systemer, som alle kan repræsenteres og analyseres ved elektronikkens værktøj fx gennem SPICE simulering.

Første kapitel definerede de funktioner der knytter forbindelse mellem strøm i og spænding U for de tre passive komponenter modstanden R , kondensatoren C og selvinduktionen L der ofte blot kaldes for en spole. Komponenterne er passive fordi de ikke selv genererer noget signal⁴⁵.

Nedenfor vises de tre symboler for de passive komponenter med den valgte positive notation. For at drive energi ind i kredsløbene er der to typer af kilder én for strøm og én for spænding. De kan valgfrit benyttes som repræsentant for et batteri (konstant værdi), en tonegenerator (amplitude og frekvens) eller en styret kilde der er en funktion af noget andet som angives ved det ledsagende symbol. For eksempel kan strømgeneratoren afhænge af spændingen et andet sted i kredsløbet.



Figur 58 – En oversigt over de to kilder: strømgenerator i_G og spændingsgenerator u_G samt de tre passive komponenter: modstand R , kondensator C og spole L . Spændingskilden indikerer en lav indre modstand ved det "kortsluttede" symbol mens strømkilden indikerer høj indre modstand ved "afbrydelsen". Modstanden vises som en kasse der også kan repræsentere kompleks impedans. Kondensatoren viser de to adskilte plader og spolen viser en opviklet tråd.

Definitionen af spændingen u over hver af de tre komponenter som funktion af strømmen i gennem dem er vist herunder. Et lille bogstav betegner en tidsvarierende kvantitet så i det formelle sprog repræsenterer betegnelserne derfor $u = u(t)$ og tilsvarende $i = i(t)$.

$$u_R = R i \quad \text{og} \quad u_C = \frac{1}{C} \int i dt \quad \text{og} \quad u_L = L \frac{di}{dt}$$

For modstanden er der streng proportionalitet mellem spænding og strøm mens kondensatoren har relation til en ophobning af ladning som er strøm integreret over tid. Spændingen over spolen er relateret til ændringer i strømmen. Ligningerne har grundlæggende betydning for elektronikken og betyder at et kredsløb kan bringes til at fungere forskelligt alt afhængigt af den hastighed et signal ændrer sig med. Som eksempel kan spændingen over kondensatoren kun ændres gradvist for det er nødvendigt at opsamle ladning så virkningen af hurtige strømvariationer kan udjævnes. For spolen er det modsatte tilfældet for den reagerer på hurtige variationer, men ignorerer til gengæld langsomme eller statiske signaler for ved en konstant strømstyrke er $di/dt = 0$.

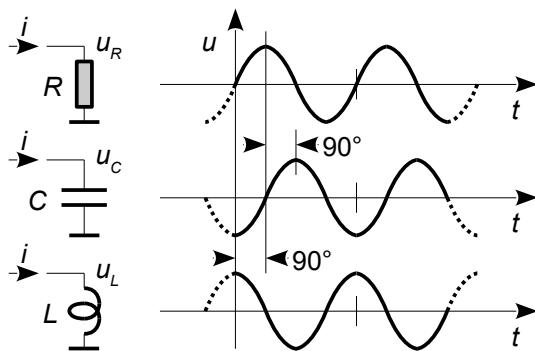
Indledningsvis vil de tre passive komponenter blive opfattet som ideelle. Det vil sige at funktionen kan beskrives matematisk eksakt med de viste ligninger. De virkelige komponenter er dog plaget af en række fejl og mangler og det vender jeg tilbage til på side 100.

⁴⁵ Der ses bort fra termisk genereret spænding og haglstøj, samt Seebeck-effekten og piezo-elektriske egenskaber. Det antages tillige at både kondensator og spole er beskyttet imod ydre elektriske og magnetiske felter.

2.1 Impedans

Det blev illustreret i første kapitel at hvis man driver en kondensator med en oscillérende strøm vil der over komponenten dannes en oscillérende spænding med samme frekvens, men 90° bagud i forhold til strømmen og spændingens amplitude aftager med stigende frekvens. Spolen har nogle lignende egenskaber dog vil spændingen over den stige med frekvensen og fasen er 90° fremme i forhold til strømmen.

Det illustreres herunder hvor de tre komponenter modtager en oscillérende strøm $i(t)$. Den stippled del af kurverne angiver at signalerne længe har oscilléret som vist og at de fortsætter med at gøre det i lang tid. Det kaldes for den stationære tilstand (*stationary condition*) selv om både strøm og spænding hele tiden ændres for det overordnede billede ændres ikke. Den stationære tilstand er vigtig i elektronikken helt uagtet at der måske sker store ændringer indenfor en periode af signalet som det eksempelvis er tilfældet i en DC-DC konverter.



Figur 59 – En strømkilde oscillerer som en sinus og driver en modstand R , en kondensator C og en spole L . Spændingen over komponenterne har forskellig fase.

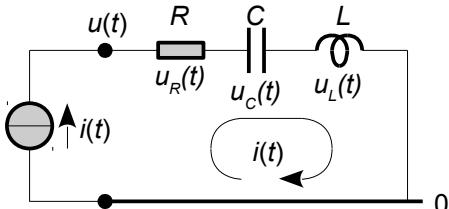
Parameteren $\omega = 2\pi f$ har dimensionen per sekund (s^{-1}) og angiver den hastighed hvormed fasen ωt øges; det er fasehastigheden der beskrives ved radianer per sekund (rad/s) og i en stor del af litteraturen lidt misvisende kaldes for vinkelfrekvensen (*angular frequency*).

Modstanden giver en spænding u_R der er i samme takt som strømmen (samme fase) og værdien er givet ved RI_M så spændingen er proportional med strømmen, men er ikke afhængig af frekvensen.

Kondensatoren giver en spænding u_C der svinger med samme frekvens som strømmen, men nu er funktionen en cosinus og med negativt fortegn. Det betyder at signalet er 90° bagefter strømmen og nok så vigtigt er faktor $1/\omega C$ en funktion af frekvensen og aftager ved stigende frekvens.

Spolen giver en spænding u_L der svinger med samme frekvens som strømmen, men faktor ωL stiger med frekvensen og signalet er 90° forude for strømmen.

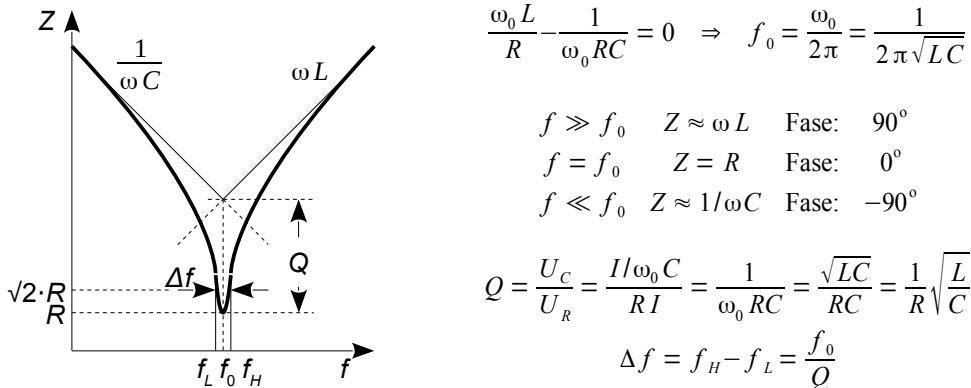
Herunder vises et kredsløb hvor de tre komponenter er i serie og drives af strømkilden $i(t)$. Det er opgaven at beregne spændingen $u(t)$ over serieforbindelsen og det klares ved at addere de tre bidrag med hensyn til den valgte positive retning. Vi ender med et led i fase med det drivende signal (sinus) og et led hvor amplituden er frekvensafhængig og fasen er $\pm 90^\circ$ (cosinus).



$$\begin{aligned} \text{Spænding: } & u(t) = u_R(t) + u_C(t) + u_L(t) \\ \text{Drivende signal: } & i(t) = I \sin(\omega t), \quad \omega = 2\pi f \Rightarrow \\ & u(t) = RI \sin(\omega t) - \frac{1}{\omega C} I \cos(\omega t) + \omega L I \cos(\omega t) \\ & u(t) = RI \sin(\omega t) + \left(\omega L - \frac{1}{\omega C} \right) I \cos(\omega t) \end{aligned}$$

Figur 60 – Et eksempel på et elektrisk kredsløb med modstand R , kondensator C og spole L . Til højre vises overføringsfunktionen for spændingen versus strømmens amplitude og frekvens.

Parentesen har dimensionen Ω og er derfor en slags modstand selv om den både afhænger af frekvensen og modifierer fasen af det oscillérende signal. Den kaldes for kredsløbets impedans Z (*impedance*) og i næste afsnit skrives det mere elegant ved brug af komplekse tal. Det kan dog her konstateres at leddene i parentesen bliver nul ved en karakteristisk frekvens f_0 der kaldes for kredsløbets resonansfrekvens (*resonance frequency*) hvor amplituden af spændingen bliver lig med RI og er i fase med den drivende strøm. Impedansen Z har samme funktion som modstanden i Ohms lov med den forskel at impedansen er funktion af frekvensen.



Figur 61 – Resonans for et RCL seriekredsløb (tyk linje). De skrå linjer repræsenterer strømmen i spolen (lave frekvenser) og i kondensatoren (høje frekvenser). Ved den karakteristiske frekvens f_0 er impedansen af kondensatoren og spolen ens, men de har modsat fortegn og ophæver hinanden så kredsløbets funktion er ved den frekvens bestemt alene af modstanden R .

Hvis kredsløbet alene bestod af kondensatoren ville amplituden af spændingen ved frekvensen f_0 være lig med $U = I/\omega_0 C$, men med spolen inde i kredsløbet bliver spændingen $U_R = IR$ ved den samme frekvens. Denne evne til at kondensator og spole kompenserer for hinanden definerer en meget vigtig egenskab ved kredsløbet der kaldes for dets godhed (*quality factor*). Godheden Q bestemmer det frekvensområde Δf hvor impedansens værdi ikke afviger alt for meget fra værdien af modstanden alene. Som ved alt andet indenfor elektronikken defineres grænseværdien som en ændring på $\sqrt{2}$ af den aktuelle parameter og det svarer til 3 dB.

Frekvensbåndets grænsefrekvenser defineres ved en forøgelse af impedansens værdi Z_0 med kvadratorden af 2. For at finde grænsefrekvenserne kvadreres for at fjerne kvadratorden, der divideres med R^2 og det hele sættes på fælles brøkstreg. Kvadratorden tages på begge sider hvilket giver en fortugnsusikkerhed og nævneren ganges over til en andengrads ligning i ω_1 som løses for at finde de to grænsefrekvenser.

$$Z(f_1) = \sqrt{R^2 + \left(\omega_1 L - \frac{1}{\omega_1 C}\right)^2} = \sqrt{2} R \Rightarrow 1 + \left(\frac{\omega_1^2 LC - 1}{\omega_1 CR}\right)^2 = 2 \Rightarrow \omega_1^2 LC \pm \omega_1 CR - 1 = 0$$

Løsning af andengrads ligningen benytter det klassiske værkøj og der deles med CR i både tæller og nævner for at introducere Q under kvadratordenen. Det udnyttes at ω_1 skal være positiv og L skrives som $(L/C)^{0.5}$ gange $(LC)^{0.5}$ hvorfed både Q og ω_0 indføres i udtrykket.

$$\omega_1 = \frac{\pm CR \pm \sqrt{(CR)^2 + 4LC}}{2LC} \Rightarrow \omega_1 = \frac{\pm 1 \pm \sqrt{1+4Q^2}}{2L} \xrightarrow{\omega_1 > 0} \left(\frac{\omega_H}{\omega_L}\right) = \omega_0 \frac{\sqrt{1+4Q^2} \pm 1}{2Q}$$

Båndbredden hvor impedansen er steget til maksimalt $\sqrt{2} R$ er givet ved differensen mellem de to grænsefrekvenser. Resultatet gælder også for en parallel-resonans.

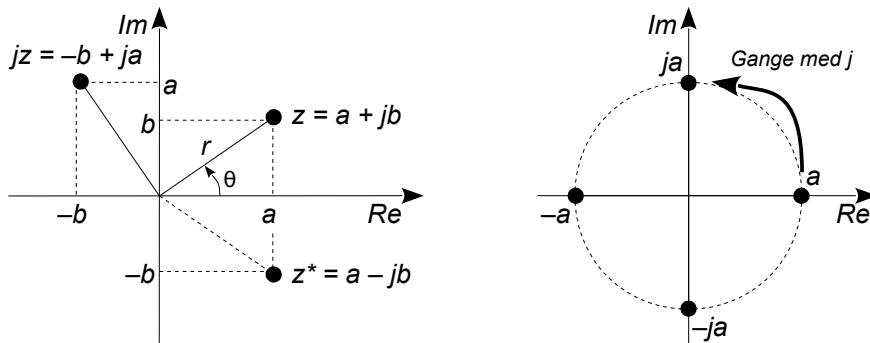
$$\Delta f = \frac{\omega_H - \omega_L}{2\pi} = \frac{\omega_0}{2\pi} \frac{(\sqrt{1+4Q^2} + 1) - (\sqrt{1+4Q^2} - 1)}{2Q} = \frac{f_0}{Q}$$

2.2 Komplekse tal

For at komme frem til et brugbart udtryk for impedansen er der brug for en effektiv notation der kan give os de $\pm 90^\circ$ fasedrejning og her kommer de komplekse tal ind i billedet⁴⁶.

$$\begin{aligned} z &= a + jb & z &= r \exp(j\theta) & r &= |z| = \text{mod}(z) = \sqrt{a^2 + b^2} & y &= r \sin(\theta) \\ \text{Re}(z) &= a & z &= r e^{j\theta} & \theta &= \arg(z) = \arctan\left(\frac{b}{a}\right) & x &= r \cos(\theta) \\ \text{Im}(z) &= b \end{aligned}$$

De komplekse tal er en udvidelse af de reelle tal som vi er vant med. Et komplekst tal z er en sum af to reelle tal a og b hvor det ene af tallene er ganget med den imaginære enhed, kaldet i eller j , eksempelvis det komplekse tal $z = 3 + 2j$. Det komplekse tal danner et punkt i en talplan med den reelle værdi langs den horizontale akse (Re-aksen) og den imaginære værdi langs den vertikale akse (Im-aksen). Her er 3 den reelle del af det komplekse tal z og 2 er den imaginære del. Det er lidt i stil med vektorer så man kan opfatte den imaginære enhed som en enhedsvektor langs den imaginære akse. Forskellen ses når man regner på tallene for de komplekse tal benytter kun de sædvanlige fire regnearter og er derfor fri for vektorregningens skalar- og krydsprodukter. Et komplekst tal kan dog også her beskrives ved dets længde (modulus, r) og vinkel (argument, θ).



Figur 62 – Til venstre den komplekse talplan med et komplekst tal $a + jb$ (første kvadrant, øverst til højre), og virkningen af at gange det med den imaginære enhed (andet kvadrant, øverst til venstre) samt z^* der er den komplekst konjugerede af z (fjerde kvadrant, nederst til højre) hvor fortegnet til b er skiftet. Til højre et eksempel hvor det reelle tal a roteres med den imaginære enhed.

Den imaginære enhed fungerer som en drejning på 90° i positivt omløb (mod uret). Et almindeligt reelt tal, for eksempel $a = 3$, bliver til et imaginært tal ved multiplikation med j hvor $3j$ skal opfattes som at tallet er drejet 90° i et koordinatsystem så det ved multiplikationen er drejet til at ligge ud af den imaginære akse. Multiplieres videre med den imaginære enhed bliver tallet $3j$ til $3j^2$ og da det nu er blevet roteret til positionen -3 på den reelle akse er den imaginære enhed defineret.

$$j^2 = -1$$

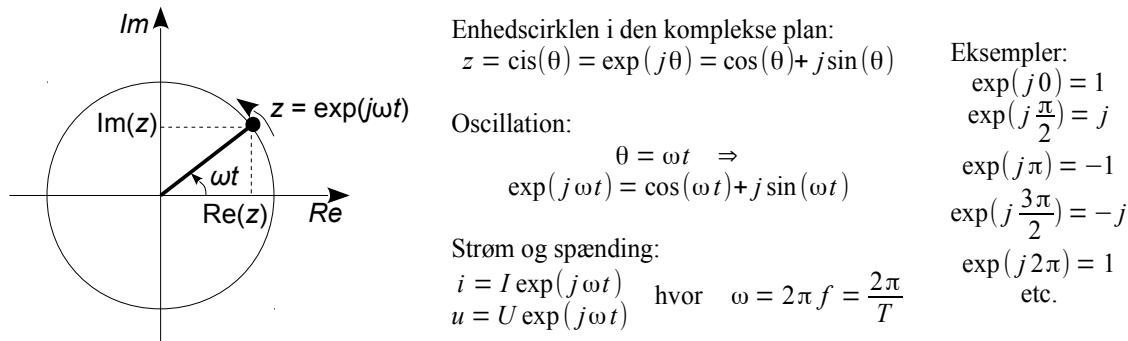
Definitionen skal bare accepteres for det kan ganske enkelt ikke forstås. Det er præcis derfor det hedder den imaginære enhed. Multipliceres $3j^2$ igen med j kommer vi frem til $3j^3 = -3j$ der nu peger nedad den imaginære akse og ved endnu en multiplikation med j er vi tilbage ved udgangspunktet på den reelle akse $3j^4 = 3$, så det kan konkluderes at $j^4 = 1$ i god tråd med $j^4 = j^2 j^2 = (-1)(-1) = 1$.

Fortsat multiplikation med den imaginære enhed får et tal til at rotere i spring af 90° så der ses et potentiale for at kunne beskrive noget der roterer som et hjul i en maskine. Svingninger er meget vigtige indenfor naturvidenskab og det er derfor de komplekse tal er blevet så populære.

⁴⁶ Cardano opdagede i 1545 at tredjegrads ligninger altid gav tre løsninger hvis man kunne acceptere kvadratroden af negative tal og i 1572 indførte Bombelli en notation for denne imaginære enhed. Moivre opdagede i 1730 at trigonometriske relationer blev lettere at løse ved brug af komplekse tal. Euler fandt i 1748 frem til den eksponentielle formel (se side 75) og i 1799 noterede Wessel at det komplekse tal er et punkt i en talplan.

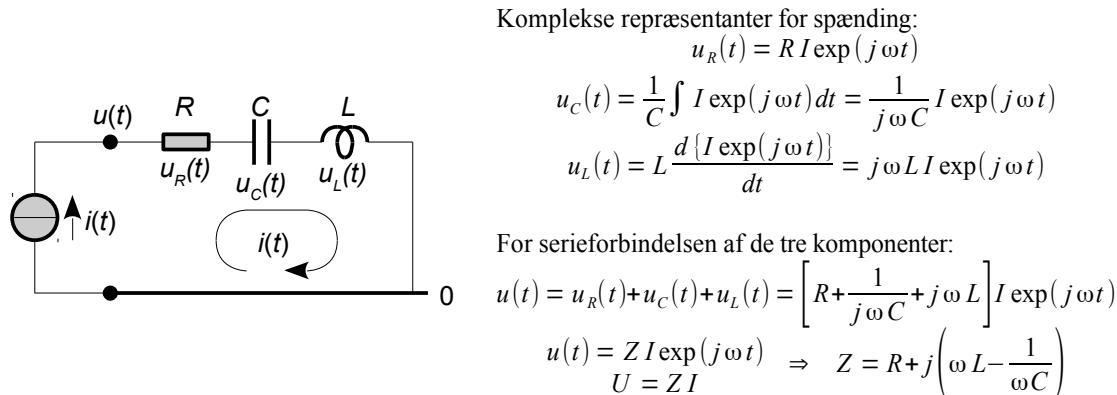
Punkterne på periferien af enhedscirklen er givet ved $x = \cos(\theta)$ og $y = \sin(\theta)$ så i den komplekse talplan skal cosinus pege ud af den reelle akse og sinus op af den imaginære akse. Det skrives ved funktionen $\text{cis}(\theta)$, der står for cosinus plus den imaginære enhed ganget med sinus. 1748 viste matematikeren Euler at eksponentialfunktionen $\exp(x) = e^x$ giver netop denne relation hvis den får et imaginært argument $x = j\theta$ og det uddybes på side 75.

En oscillation i den komplekse talplan kan nu beskrives ved $\exp(j\theta)$ for realdelen er en cosinus og imaginærdelen en sinus. Eksponentialfunktionen vil rotere en hel omgang for θ gående fra 0 til 2π . Det er normalt at angive vinklen som $\theta = \omega t$ hvor ω er vinkelhastigheden (radianer per sekund) for en frekvens f af oscillationen med t som tiden⁴⁷.



Figur 63 – Den komplekse eksponentialfunktion beskriver cirkelperiferien af enhedscirklen i den komplekse talplan for et imaginært argument.

Svingningen indføres i udtrykkene for modstand, kondensator og spole hvor det bliver udnyttet at eksponentialfunktionen giver sig selv ved differentiation og integration på nær et konstant led $j\omega$ der optræder som henholdsvis faktor ved differentiation og divisor ved integration (se side 70).

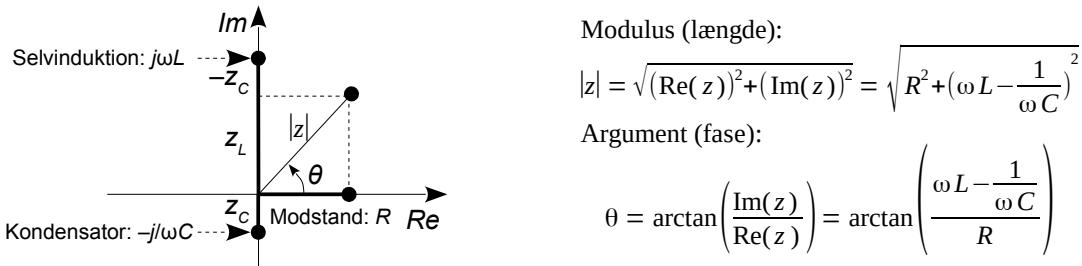


Figur 64 – Kredsløbet med en serieforbindelse af de tre komponenter R, C og L analyseres med en repræsentation ved de komplekse tal.

Forskellen fra det tidligere resultat med strømmen defineret som $i = I \sin(\omega t)$ er at alle led nu har den oscillérende funktion $i = I \exp(j\omega t)$. Faseskiftet på -90° for kondensatoren og $+90^\circ$ for spolen beskrives ved den imaginære enhed. For kondensatorens $1/j\omega C$ er minusset blot gemt lidt af vejen ved at den imaginære enhed står i nævneren i faktoren til den oscillérende strøm. Det kan ses ved at gange med den imaginære enhed i både tæller og nævner. Det ændrer udtrykket til $-j/\omega C$ i tråd med definitionen af kvadratet på den imaginære enhed. Det unikke ved den komplekse impedans er at oscillationen ikke er af central betydning det er derimod svingningens amplitude og fase.

⁴⁷ At strøm og spænding ved dette kunstgreb bliver komplekse ser mange lærebøger som et stort problem og derefter bruger de ret så mange kræfter på at løse det selvskabte problem. Så er det bedre at man blot accepterer at matematikken er en abstraktion. Strøm og spænding kan aldrig være komplekse, det er der intet fysisk der kan være, men matematikken bliver elegant. Hvis et resultat skal være reel kan man benytte modulus til z eller realdelen $\text{Re}(z)$ som lærebøgerne da også kommer frem til.

Summen af de tre komponenters impedans kan vises som en vektor der angiver den resulterende længde (*modulus*) og fasevinkel (*argument*). Det illustreres herunder for en frekvens over f_0 hvor spolens impedans dominerer over både modstand og kondensator. Den samlede længde er større end for spolen og modstanden alene og fasevinklen θ er positiv da spolens impedans er numerisk større end kondensatorens impedans.



Figur 65 – En afbildning af de tre komponenters impedans i den komplekse talplan sammen med det resulterende komplekse tal og de tilhørende formler.

Modstanden repræsenteres af den konstante og reelle værdi $Z_R = R$ og som følge heraf afbides den ud af den reelle akse.

Kondensatoren repræsenteres af en frekvensafhængig og imaginær værdi $Z_C = 1/j\omega C = -j/\omega C$ der giver en negativ imaginær værdi så den afbides ud af den imaginære akse og i negativ retning svarende til en fasevinkel på -90° . Impedansen aftager med stigende frekvens så en kondensator kan ved meget lave frekvenser (DC) repræsentere en afbrydelse.

Spolen repræsenteres af en frekvensafhængig og positiv imaginær værdi $Z_L = j\omega L$ så den afbides i den positive retning af den imaginære akse svarende til en fasevinkel på 90° . Impedansen stiger med stigende frekvens så en spole kan ved meget lave frekvenser (DC) repræsentere en kortslutning.

Med Z for impedansen af modstand, kondensator og spole kan Ohms lov udtrykkes generelt for en kompleks impedans hvis værdi er frekvensafhængig. Det er sædvanlig at benytte store bogstaver indenfor elektronikken for de komplekse størrelser der oftest kun angiver signalamplitude og fase hvor frekvensen underforstået, men der ses mange måder at notere komplekse tal på⁴⁸.

$U = Z I$ $u(t) = U \exp(j\omega t)$ $i(t) = I \exp(j\omega t)$ $\omega = 2\pi f$	$\text{Modstand: } U_R = R I \quad \Rightarrow \quad Z_R = \frac{U_R}{I} = R$	$\text{Kondensator: } U_C = \frac{1}{j\omega C} I \quad \Rightarrow \quad Z_C = \frac{U_C}{I} = \frac{1}{j\omega C} = -j \frac{1}{\omega C}$	$\text{Spole: } U_L = j\omega I \quad \Rightarrow \quad Z_L = \frac{U_L}{I} = j\omega L$
---	---	--	--

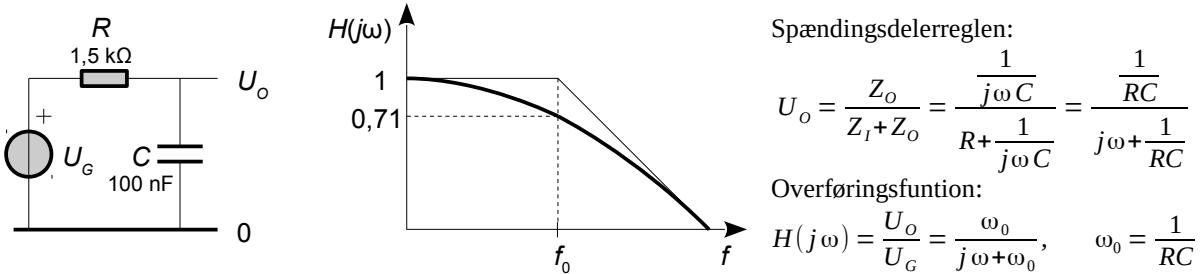
Den tidsafhængige relation gennem $\exp(j\omega t)$ er fjernet fra ligningerne hvilket er tilladeligt da vi kun beskæftiger os med den stationære tilstand for kredsløbet altså at svingningen altid har været der og at hver periode er identisk med den foregående. Denne definition af impedans sætter os i stand til at analysere et hvilket som helst analogt kredsløb med konstante værdier for modstand, spole og kondensator. Det betyder ikke at et potentiometer eller en drejkondensator ikke kan benyttes i et kredsløb for værdien forudsættes blot konstant når frekvensen varieres.

Med dette materiale er vi godt rustet til at analysere et elektrisk kredsløb.

48 Almindelige notationer: $z, \hat{z}, Z, r/\theta$.

2.2.1 Eksempler på analyse

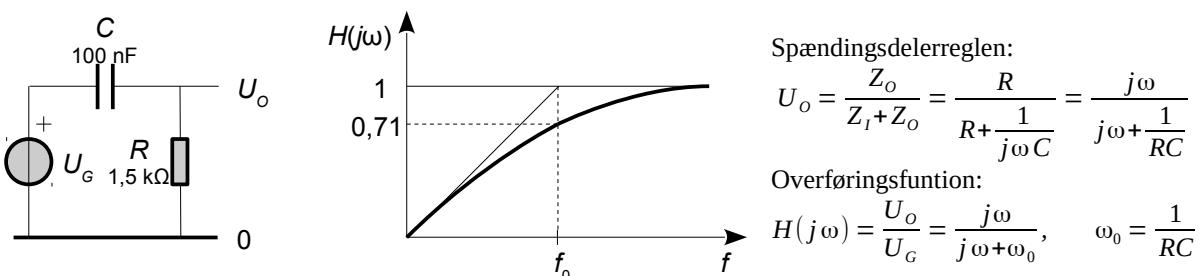
Mange kredsløb kan analyseres direkte ved spændingsdelerreglen og som et eksempel vises en analyse af filtret herunder. Her er Z_o belastningen af udgangen og Z_i er impedansen i indgangen. Der er kun en enkelt kondensator i udgangen som har impedansen $1/j\omega C$ og impedansen i indgangen er modstanden R . Ved at gange med $j\omega$ samt dividere med R i tæller og nævner vil brøkens værdi forblive uændret. Udtrykket ses nu at bestå af $j\omega$ og et led $1/RC$ der skal kaldes ω_0 og som udgør en vinkelhastighed. Dimensionen er radianer per sekund.



Figur 66 – Et lavpasfilter tillader lavfrekvente signaler at passere mens de højfrekvente dæmpes.

Ved lave frekvenser er $j\omega$ lille i forhold til ω_0 så overføringsfunktionen $H(j\omega)$ er næsten lig med ω_0 delt med ω_0 og det vil sige at $H(j\omega) \approx 1$. Det betyder at en svingning med en lav frekvens stort ses ikke dæmpes og vil passere filtret uhindret. Ved høje frekvenser er $j\omega$ stor i forhold til ω_0 så overføringsfunktionen $H(j\omega)$ næsten identisk med $\omega_0/j\omega$ så en svingning med en høj frekvens vil blive dæmpet. For en specifik frekvens vil modulus af $j\omega$ være identisk med ω_0 og her vil overføringsfunktionen blive $1/(j + 1)$ der er lig med 0,71 vinkel -45° .

Hvis modstanden og kondensatoren ombyttes skabes et højpasfilter som blandt andet benyttes til at blokere for DC. Spændingsdelerreglen danner næsten samme overføringsfunktion, dog er ω_0 i tælleren erstattet med $j\omega$ og det har stor betydning. Ved lave frekvenser blokerer kondensatoren for signaler og i analysen vil en DC svare til $j\omega = 0$ som da også viser at overføringsfunktionen her er nul så "signalet" vil ikke kunne passere filtret. Ved høje frekvenser er $j\omega$ større end ω_0 og det betyder at overføringsfunktionen bliver $j\omega$ delt med $j\omega$ som giver et et-tal. Filtret vil derfor spærre for DC, dæmpe lavfrekvente signaler og lade højfrekvente signaler passere uden nævneværdig dæmpning.



Figur 67 – Et højpasfilter tillader højfrekvente signaler at passere mens de lavfrekvente dæmpes.

En væsentlig observation er at nævnerens polynomium $j\omega + \omega_0$ er fælles for de to filtre og at kun tælleren er forskellig. Nævneren viser hvordan forløbet er over frekvens og tælleren afgør om det er et lavpasfilter eller et højpasfilter. I begge tilfælde er filtret karakteriseret ved tidskonstanten RC der giver en grænsefrekvens f_0 på lidt over 1 kHz.

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi RC} \quad \text{Med viste komponentværdier: } f_0 = \frac{1}{2\pi(1,5 \text{ k}\Omega)(100 \text{ nF})} \approx 1,06 \text{ kHz}$$

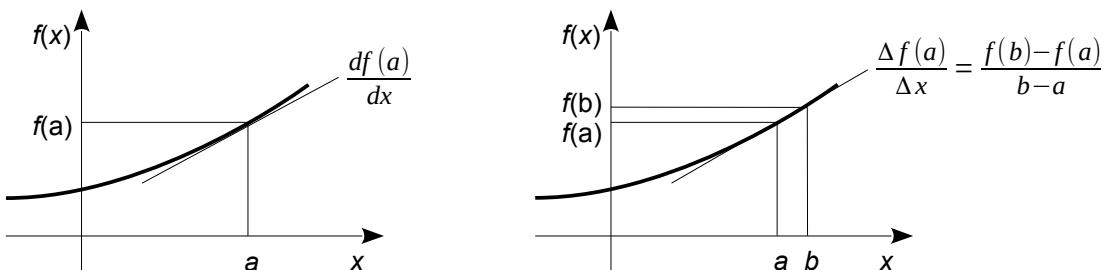
2.3 Differentialregning

Meget videnskabeligt arbejde beskæftiger sig med at opstille og løse differentialligninger og det er ikke kun ved analyse af elektriske kredsløb for ligningerne er også vigtige indenfor de mekaniske, akustiske, termiske og kemiske områder. Indenfor elektronikken kan vi ofte forudsætte en relation til eksponentialfunktionen eller de trigonometriske funktioner sinus og cosinus hvilket skyldes at vi oftest arbejder med systemer der enten glider på plads efter en påvirkning eller oscillerer omkring en tilstand. Vi kan så koncentrere os om hvor hurtigt bevægelsen ændres og ikke bekymre os om udseendet af selve svingningen.

Matematikken bag differentialligningerne blev udviklet af Leibnitz i 1684 og uafhængigt heraf af Newton i 1666 der dog først publicerede resultaterne fra 1687. Matematikken er ikke begrænset til svingninger for den kan også beskrive hvordan en opløsning diffunderer gennem en væske eller varmeenergi flyder i et materiale og begge beskrives med helt den samme matematik som for strømmen i en kondensator der oplades gennem en modstand.

2.3.1 Differentiation

Det matematiske værktøj "differentiation af en funktion" giver oplysning om hvordan funktionen kan forventes at udvikle sig i den nærmeste omegn for et punkt. Som en illustration kan man sige at "hvis det går op ad bakke så viser en positiv værdi af differentiationen at det bliver værre endnu for bakken krummer opad". Differentiation kan også referere til et punkt i planet (x,y) hvilket netop svarer til et geografisk kort med højdeangivelser og man kan så studere om det går op eller ned ad bakke i den ene eller anden retning. Funktionen kan også beskrive et punkt i rummet (x,y,z) og differentiationen kan eksempelvis fortælle hvilken vej en trykbølge vil følge i en gas eller hvordan de mekaniske kræfter fordeler sig i en bærende konstruktion.



Figur 68 – Ved differentiation af en funktion bestemmes hældningen i et punkt som et bud på hvordan funktionen varierer omkring punktet.

Hele idéen med differentiation er at finde hældningen af funktionen $f(x)$ i et punkt der herover kaldes for a . Hældningen i punktet kan tilnærmet beregnes som hældningen af en ret linje der skærer $f(x)$ i to punkter hvor det ene er det eftersøgte punkt a og det andet punkt b er i nærheden. Hvis de to punkter på aksen ligger langt fra hinanden vil resultatet blive en grov tilnærrelse og måske endda helt forkert hvis funktionen pludseligt drejer nedad i stedet for at fortsætte med at stige. Hvis de to punkter ligger tæt sammen vil estimatet ramme næsten helt rigtigt. Det kan så generaliseres til at differentialkvotienten af funktionen $f(x)$ i punktet x kan skrives som følger:

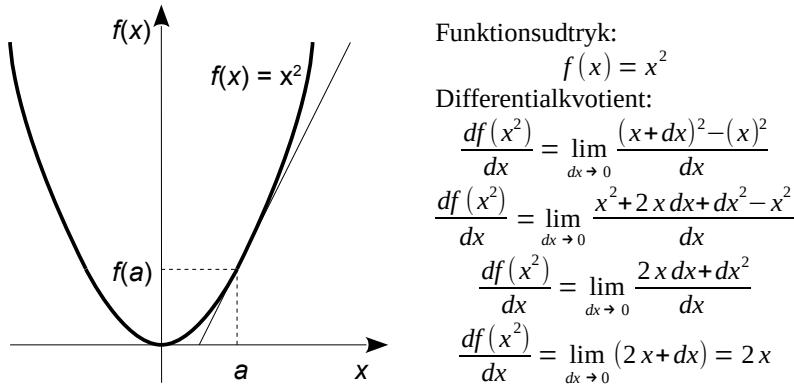
$$f'(x) = \frac{df(x)}{dx} = \lim_{dx \rightarrow 0} \frac{f(x+dx)-f(x)}{dx}$$

Hvis dx gøres mindre og mindre vil beregningen gradvist ende med den korrekte værdi hvilket skrives med **lim**, der står for *limes*, og som betyder "det, det vil ende med". Det betyder at vi skal finde værdien af "nul delt med nul". Det kan lyde som det højere skoleridt indenfor matematik, men det danner en milepæl indenfor den analytiske matematik. Kort sagt skiftede videnskaben fra at kunne beskrive noget som det var til at beskrive den måde det vil udvikle sig fremover.

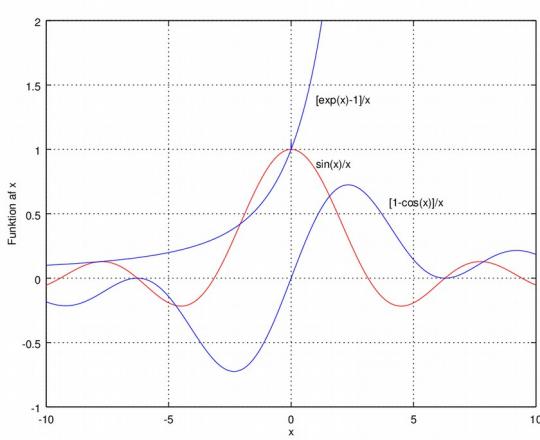
2.3.2 Eksempler på differentiation

Herunder vises hvordan differentiationen gennemføres for $f(x) = x^2$ og derefter for $f(x) = \exp(x)$ og $f(x) = \sin(x)$. Parablen repræsenterer polynomier der er funktioner som $f(x) = a_0 + a_1x + a_2x^2 + \dots$ og som kan benyttes til at opbygge komplicerede funktioner (fx Bessel funktionerne).

Polynomier skal her repræsenteres ved parablen x^2 og formålet er at beregne dens hældning (differentialkvotient) i punktet a der her generaliseres til x . Det leder til det velkendte resultat at hældningen bliver $2x$ og det kan ret enkelt udvides til at differentialkvotienten af x^n bliver nx^{n-1} .



Eksponentielle og trigonometriske funktioner er lidt tricky at bestemme for $\sin(dx)/dx$ giver nul delt med nul når dx går imod nul og det samme gør $\exp(dx) - 1$ og $1 - \cos(dx)$ delt med dx . Det kan løses mere elegant end det jeg gør herunder hvor jeg blot plotter leddene og derefter konkluderer at $\sin(dx)/dx$ og $[\exp(dx) - 1]/dx$ begge går én for dx gående imod nul mens leddet $[1 - \cos(dx)]/dx$ går mod nul for dx gående imod nul. Eksponentialfunktionen $\exp(x + dx)$ kan skrives som $\exp(x)\exp(dx)$ og $\sin(x + dx)$ kan skrives som vist ved en trigonometrisk tabel [RW-127]. På helt samme vis kan det vises at $\cos(x)$ differentieres til $-\sin(x)$.



Differentiation af $\exp(x)$:

$$\frac{d \exp(x)}{dx} = \frac{\exp(x+dx) - \exp(x)}{dx}$$

$$= \exp(x) \frac{\exp(dx) - 1}{dx} \xrightarrow{dx \rightarrow 0} \exp(x)$$

Differentiation af $\sin(x)$:

$$\frac{d \sin(x)}{dx} = \frac{\sin(x+dx) - \sin(x)}{dx}$$

$$= \frac{\sin(x)\cos(dx) + \cos(x)\sin(dx) - \sin(x)}{dx}$$

$$= -\sin(x) \frac{1 - \cos(dx)}{dx} + \cos(x) \frac{\sin(dx)}{dx} \xrightarrow{dx \rightarrow 0} \cos(x)$$

2.3.3 Regneregler ved differentiation

En konstant faktor til funktionen $f(x)$ indgår ikke i selve differentiationen og bibringes blot som en fast faktor. Funktioner med en konstant faktor til den variable, af typen $f(ax)$, kan differentieres ved substitution hvor dx erstattes af dax og funktionen kan nu differentieres med hensyn til ax . For at tillade denne substitution skal der kompenseres ved at gange hele udtrykket med a .

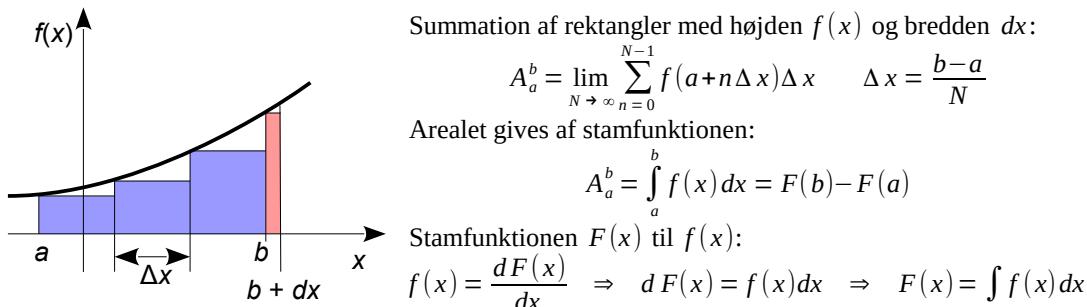
$$\frac{d}{dx} \{af(x)\} = a \frac{d}{dx} \{f(x)\}$$

$$\frac{d}{dx} f(ax) = a \frac{d}{dax} f(ax) \quad \xrightarrow{\text{Eksempel}} \quad \frac{d}{dt} f(t) = A \sin(\omega t) + B \cos(\omega t)$$

$$\frac{d}{dt} f(t) = \omega A \cos(\omega t) - \omega B \sin(\omega t)$$

2.3.4 Integration

Integration er opstået som bestemmelse af arealet under en funktion⁴⁹ og indenfor et afgrænset område lidt som at tælle antallet af fliser på badeværelsets væg. For at bestemme arealet under funktionen $f(x)$ fra positionen a til b opdeles stykket i rektangler med højden givet ved funktionens værdi $f(x)$ i det pågældende punkt gange med bredden Δx . Ved at addere rektanglerne fra a til b findes et estimat af arealet med den vandrette akse som nedre grænse. Estimatet er ikke eksakt for funktionen tilnærmes med rektangler, men hvis Δx er lille vil antallet af rektangler blive meget stort og tilnærmlsen til det eksakte resultat vil være god. Påstanden er at for Δx gående mod nul vil arealet gå mod den eksakte værdi.



Figur 69 – Ved integration bestemmes arealet af et område defineret vertikalt ved funktionen $f(t)$ og den vandrette akse samt horisontalt ved et startpunkt a og et slutpunkt b .

Det elegante som Newton og Leibnitz fandt frem til er at der findes en funktion $F(x)$ som direkte giver arealet A under funktionen $f(x)$ fra a til b . Funktionen kaldes for *stamfunktionen* til $f(x)$ og den betegnes med $F(x)$. Stamfunktionen er givet ved den omvendte proces af differentiation hvilket i dag kaldes for infinitesimalregningens hovedsætning (*the fundamental relation of calculus*)⁵⁰. Nogle eksempler på de to regneregler vises på side 70.

Funktionen $F(x)$ kan bestemmes ved at forøge integralet fra a til b med et bidrag fra b til $b+dx$ hvor arealet af det nye areal er $f(b)dx$. Tilvæksten i arealet kan skrives som differensen mellem det nye areal og det tidligere areal.

$$\begin{aligned} A_a^{b+dx} &= A_a^b + f(b)dx \\ f(b)dx &= A_a^{b+dx} - A_a^b \\ f(b)dx &= F(b+dx) - F(a) - [F(b) - F(a)] \\ f(b)dx &= F(b+dx) - F(b) \end{aligned}$$

Ved at dele med dx findes selve definitionen af den afledte til funktionen $F(x)$ i punktet b :

$$f(b) = \lim_{dx \rightarrow 0} \frac{F(b+dx) - F(b)}{dx} = \frac{dF(b)}{dx}$$

Punktet b kan ligge hvor som helst så $F(b)$ kan generaliseres til $F(x)$ og differentialet $dF(x)$, altså ændringen i $F(x)$ ved en lille ændring i x , bestemmes ved at gange udtrykket ovenfor med dx på begge sider af lighedstegetnet. Den omvendte process af differentiation er integration og den kan defineres som den proces der danner funktionen $F(x)$ fra differentialet $dF(x)$.

$$dF(x) = f(x)dx \Rightarrow F(x) = \int dF(x) = \int f(x)dx$$

Da integration og differentiation er hinandens ækvivalente operationer kan relationen benyttes begge veje. Man kan derfor finde stamfunktionen $F(x)$ fra $f(x)$ ved at integrere $f(x)$ eller man kan finde $f(x)$ fra $F(x)$ ved at differentiere $F(x)$.

49 Noget tyder på at en romer arbejdede med integration for bestemmelse af voluminet af emner allerede omkring Kristi fødsel.

50 http://en.wikipedia.org/wiki/Fundamental_theorem_of_calculus.

2.3.5 Eksempler på integration

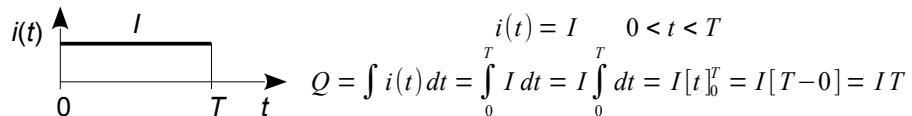
Herunder vises hvordan integration kan benyttes for beregninger indenfor elektronikken som fx hvor meget ladning der er tilført en kondensator eller et batteri. Integralet herunder beskriver den ladning Q der er involveret ved at strømmen $i(t)$ har været aktiv i tidsrummet fra $t = 0$ til $t = T$.

$$Q = \int_0^T i(t) dt$$

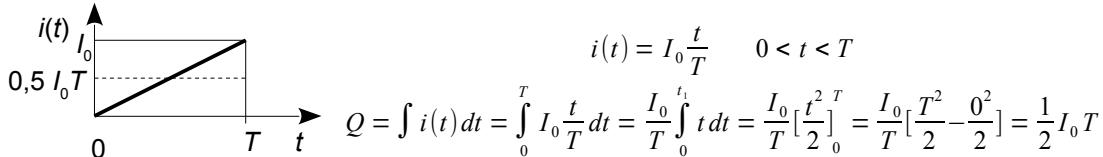
Produktet strøm gange tid har dimensionen ampere gange sekund og det er ladning (coulomb).

Hvis strømmen $i(t)$ er konstant kan man erstatte den med symbolot I for en konstant strøm. Da en konstant ikke deltager i integrationen kan den stilles ud foran integrationen som en fast faktor og integralet er nu reduceret til at bestemme hvor lang tid opladningen har stået på. Den konstante faktor kan ligeledes stilles ud foran i de efterfølgende eksempler hvor strømmen beskrives ved en maksimalværdi gange med et matematisk udtryk som en lineær ændring, en eksponentiel ændring og en halvbølge sinus. Alle udtryk er overordentligt relevante indenfor elektronikken.

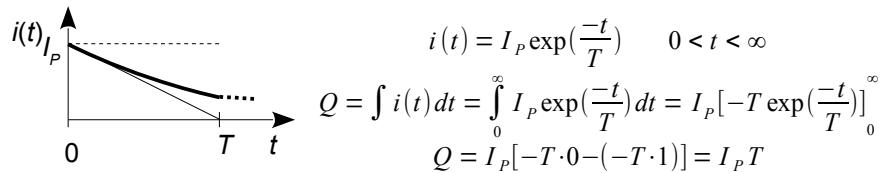
En konstant strømstyrke svarer til opladningen af et bilbatteri fra en ladestation. Strømmen antages konstant og forløbet dækker et tidsrum på T . Resultatet er at ladningen er givet ved strømmens DC værdi ganget med tidsrummets længde.



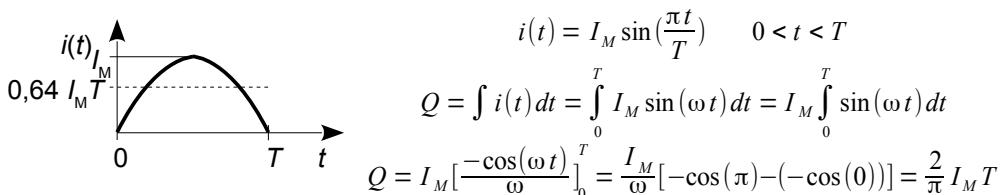
En lineært voksende strømstyrke mødes ved en DC-DC konverter (se sjette kapitel) og den resulterer i en ladning på halvdelen af strømmen I_0 gange med tidsrummet T .



En eksponentielt aftagende strømstyrke svarer til afladning af en kondensator gennem en modstand. Strømmen aftager fra startværdien I_P mod nul med tidskonstanten T . Strømmen vil i principippet aldrig nå nul så integrationen forløber over "al resterende tid i universet". Resultatet er en ladning på startstrømmen I_P gange med tidskonstanten T . Integrationen kan standses efter 5 tidskonstanter (integration til 5 T) hvor resultatet er 0,7 % fra det korrekte.



En halvbølge sinus giver den ladning som overføres i en DC effektforsyning hvor lysnettets svingning ensrettes af en diode for opladning af en kondensator. Resultatet er at der overføres en ladning på 64 % af topværdien gange med tidsrummet for en halvperiode.



2.3.6 Tabel over differentiation og integration

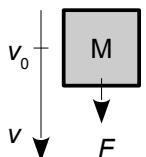
Herunder gengives nogle vigtige resultater fra differential- og integralteorien. Funktioner af den variable x kaldes $f = f(x)$, $g = g(x)$ og $h = h(x)$ og konstanter kaldes a , b og c . Heltal kaldes for n . Funktioner af andre variable vil blive betegnet som $f(y)$ og $g(z)$ og er konstante ved differentiation og integration med hensyn til x . En sammensat funktion skrives som $g(h)$ hvor det underforstås at den variable for funktionen h er x så det kunne også skrives $g(h(x))$. Der optræder en konstant c ved ubestemt integration som må fastlægges ud fra krav til løsningen. Ved bestemt integration fra a til b er resultatet dog korrekt så her er der ingen integrationskonstant.

Differentialkvotient	Funktion	Integrale
$f' = f'(x) = \frac{df(x)}{dx}$	$f = f(x)$	$F = F(x) = \int f(x) dx + c$ $F(b) - F(a) = \int_a^b f(x) dx$
$f' = ag' + bh'$	$f = ag + bh$	$F = aG + bH$
$f' = a[g'h + gh']$	$f = agh$	$F = a[Gh - \int Gh' dx]$
$f' = a \frac{g'h' - gh'}{g^2}$	$f = a \frac{g}{h}$	
$f' = ag'(h)h'$	$f = ag(h)$	
$f' = a$	$f = ax$	$F = \frac{ax^2}{2}$
$f' = anx^{n-1}$	$f = ax^n$	$F = \frac{a}{n+1}x^{n+1} \quad n \neq -1$
$f' = a \frac{n}{x^{n+1}}$	$f = \frac{a}{x^n} = ax^{-n}$	$F = \ln x \quad n = 1$ $F = \frac{a}{(n-1)x^{n-1}} \quad n \neq 1$
$f' = \frac{a}{2\sqrt{x}}$	$f = a\sqrt{x}$	
$f' = ab\exp(bx)$	$f = a\exp(bx)$	$F = \frac{a}{b}\exp(bx)$
$f' = \frac{a}{x}$	$f = a\ln(bx)$	
$f' = ab\cos(bx)$	$f = a\sin(bx)$	$f = -\frac{a}{b}\cos(bx) + c$
$f' = -ab\sin(bx)$	$f = a\cos(bx)$	$F = \frac{a}{b}\sin(bx) + c$
$f' = ab[1 + \tan^2(bx)] = \frac{ab}{\cos^2(bx)}$	$f = a\tan(bx)$	$F = -\frac{a}{b}\ln \cos(bx) $
$f' = ab\cosh(bx)$	$f = a\sinh(bx)$	$F = \frac{a}{b}\cosh(bx)$
$f' = ab\sinh(bx)$	$f = a\cosh(bx)$	$F = \frac{a}{b}\sinh(bx)$
$f' = ab[1 - \tanh^2(bx)] = \frac{ab}{\cosh^2(x)}$	$f = a\tanh(bx)$	$F = \frac{a}{b}\ln(\cosh(bx))$

2.3.7 Eksempler på differentialligninger

Eksempler på differentialligninger skal gives. Den første beskriver Newtons anden lov for en faldende masse og viser en løsning med et polynomium. De andre eksempler viser afladning af en kondensator og opladning af en spole og viser en løsning med en eksponentialfunktion.

Newton's anden lov beskriver accelerationen af en masse M der påvirkes med en kraft F som her er givet ved jordens tiltrækning af loddet. Hvis loddet slippes vil kraften accelerere loddet og hastigheden v beregnes til tidspunktet t efter at loddet er sluppet. Samme differentialligning kan benyttes til at finde positionen x som funktion af tiden efter at loddet er sluppet ved at udnytte at hastigheden er defineret som differentialkvotienten af positionen.



Newton's anden lov: $F = M \frac{dv}{dt}$ hvor kraften er $F = Mg$ og $g = 9,82 \text{ m/s}^2$

Bestemmelse af hastighed:

$$g = \frac{dv}{dt} \Rightarrow dv = g dt \Rightarrow v = \int dv = \int g dt = g \int dt = v_0 + gt$$

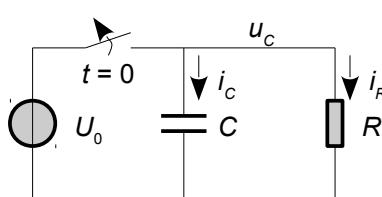
Bestemmelse af position:

$$v = \frac{dx}{dt} \Rightarrow dx = v dt \Rightarrow x = \int dx = \int v dt = \int (v_0 + gt) dt = x_0 + v_0 t + \frac{g}{2} t^2$$

Ved kombination af ligningerne for Newtons anden lov og kraften ses at massen kan forkortes væk og det efterlader en differentialligning med differentialkvotienten af accelerationen dv/dt og tyngdeaccelerationen g . Differentialet til hastigheden dv kan isoleres ved at gange med dt og hastigheden v bestemmes ved at integrere på begge sider. Da tyngdeaccelerationen anses for konstant kan den flyttes ud foran integrationen. Herefter skal vi integrere differentialet af tiden dt og det giver tiden t . Løsningen vil indeholde en konstant v_0 fordi løsningen altid kan forskydes med et konstant led og stadig være en legal løsning (konstanten forsvinder ved differentiation). Det betyder at en differentialligning *altid* stiller et krav om en ekstra oplysning for at fastlægge konstanten og det er startbetingelsen (*initial condition* eller *boundary condition*). En mulighed er at loddet er i ro lige inden det slippes hvorved hastigheden er nul til tidspunktet $t = 0$ så $v_0 = 0$.

Positionen x til tidspunktet t kan beregnes fra definitionsligningen $v = dx/dt$. Ved at isolere dx og integrere findes et anden-ordens polynomium hvor x_0 er integrationsvariablen der angiver den position loddet befandt sig i lige inden det blev sluppet og det kan fx være højden over jorden.

Afladning af en kondensator. En kondensator C holdes opladt på spændingen U_0 gennem en kontakt og efter tidspunktet $t = 0$ aflades den gennem en modstand R .



$$u_C = \frac{1}{C} \int i_C dt \quad u_C = u_R = Ri_R \Rightarrow i_R = \frac{u_C}{R} \quad i_C = -i_R$$

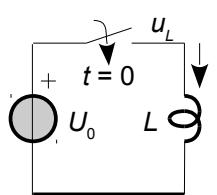
$$\Rightarrow u_C = \frac{1}{C} \int -\frac{u_C}{R} dt \quad \text{Differentiation} \quad \frac{du_C}{dt} = -\frac{1}{C} \frac{u_C}{R}$$

$$u_C = U_1 \exp(at) + U_2 \quad \frac{du_C}{dt} = a U_1 \exp(at) \quad a = -\frac{1}{CR}$$

Spændingen over kondensatoren u_C er givet ved ligningen for kondensatorens ladning, der er integralet af strøm over tid, og af strømmen i modstanden i_R . De to strømretninger for positiv er modsat rettede så strømmen i kondensatoren er negativ svarende til afladning. Ligningerne kombineres og en måde at løse ligningen på er ved at differentiere på begge sider for at fjerne integrationen og det efterlader en første-ordens differentialligning.

Der findes en eksakt løsning til denne type differentialligning, men oftest er man henvist til et såkaldt kvalificeret gæt som gennem differentiation og indsættelse viser sig at passe eller at lede til en uløselig konflikt. Her gættes på en eksponentialfunktion med U_1 , U_2 og a som nogle konstanter. Efter indsættelse optræder $U_1 \exp(at)$ på begge sider af lighedstegnet og kan divideres bort og det definerer konstanten a som tidskonstanten. Efter uendelig lang tid er kondensatoren er tømt så $U_2 = 0$ og startbetingelsen for $t = 0$ er at spændingen skal være lig med batteriet så $U_1 = U_0$.

Opladning af en ideel spole. Spændingen over en spole u_L afhænger af hvordan strømmen i_L i spolen varierer. For en ideel spole er der ikke et spændingsfald hvis strømmen i den er konstant og det er uanset værdien af strømmen (en reel spole behandles i næste afsnit). Hvis spolen til tidspunktet $t = 0$ tilsluttes en spændingskilde U_0 så tvinges spændingen over spolen til $u_L = U_0$ og det betyder at strømmen i spolen må ændres for at få ligningen til at passe.



Definitions ligning for spole:

$$u_L(t) \stackrel{\text{def}}{=} L \frac{di_L(t)}{dt}$$

Strømmen i spolen:

$$di_L(t) = \frac{u_L(t)}{L} dt, \quad t \geq 0$$

Kontakt ON for $0 \leq t \leq t_{ON}$:

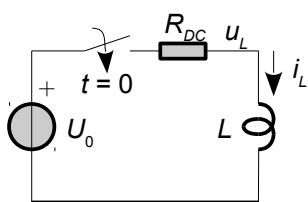
$$u_L(t) = U_0$$

Løses ved integration:

$$i_L(t) = \int_0^{t_{ON}} di_L = \int_0^{t_{ON}} \frac{U_0}{L} dt = \frac{U_0}{L} t_{ON}$$

Strømmen vokser lineært med tiden og for spændingen $U_0 = 5$ V med en spole på $L = 100 \mu\text{H}$ vokser strømmen til $i_L(t_{ON}) = 0,5$ A hvis kontakten er ON over tidsrummet $t_{ON} = 10 \mu\text{s}$. Det er en typisk situation for en DC-DC konverter hvor kontakten derefter går OFF og energien i spolen overføres til udgangen ved et kredsløb der ikke er vist.

Opladning af en reel spole. En spole har ud over selvinduktionen L en DC modstand R_{DC} fra den tråd den er viklet af så der er en grænse for hvad strømmen kan vokse til da spændingen over modstanden ikke kan overstige spændingskilden. Det er ikke muligt at løse ligningen så enkelt som vist ovenfor så i stedet opstilles en differentialligning og der gættes på en løsning⁵¹. Første ordens differentialligninger har eksponentiale funktionen som løsning og ved indsættelse af det viste generelle udtryk for i_L i differentialligningen bestemmes konstanterne a , b og c .



Kirchhoff på kredsløbet:

$$U_0 = R_{DC} i_L(t) + u_L(t)$$

Differentialligning:

$$U_0 = R_{DC} i_L(t) + L \frac{di_L(t)}{dt}, \quad t \geq 0$$

Prototype på en løsning:

$$i_L(t) = a + b \exp(ct)$$

Strømmen starter fra nul:

$$t = 0 \Rightarrow a + b = 0 \Rightarrow b = -a$$

Revision af løsning:

$$i_L(t) = a[1 - \exp(ct)]$$

Indsættelse:

$$U_0 = R_{DC} a [1 - \exp(ct)] - L [a c \exp(ct)]$$

Strømmen starter fra nul:

$$t = 0 \Rightarrow U_0 = -L a c \Rightarrow c = -\frac{U_0}{L a}$$

Strømmen er begrænset:

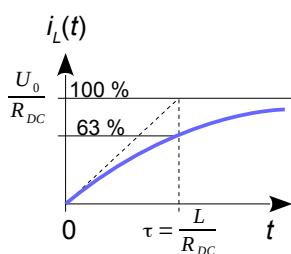
$$t \rightarrow \infty \Rightarrow U_0 = R_{DC} a \Rightarrow a = \frac{U_0}{R_{DC}}$$

Krav: $c < 0$

$$c = -\frac{R_{DC}}{L}$$

Heraf sidste konstant:

$$i_L(t) = \frac{U_0}{R_{DC}} \left[1 - \exp\left(-\frac{R_{DC}}{L} t\right) \right]$$



Løsning af differentialligninger kræver oplysninger om hvordan forløbet starter og slutter for selve differentialligningen beskriver kun det generelle om forløbet. Den nødvendige information kaldes for randværdier (*initial values*) eller grænsebetegnelser (*boundary conditions*). Her vides (eller antages) at strømmen er nul når processen startes og den information udnyttes hele to gange i eksemplet for at fastlægge udtryk for konstanterne b og c som funktion af konstanten a . Endelig ved vi at "træerne vokser ikke ind i himlen" så der er en slutværdi for strømmen når tiden vokser over alle grænser og den fastlægger værdien af konstanten a .

Strømmen i spolen vokser fra nul i det øjeblik kontakten går ON og vokser eksponentielt op mod slutværdien givet af spændingskilden og modstands værdien.

51 Det kan virke noget løst, men differentialligninger løses ofte på denne måde. Undtagelserne er så simple ligninger at der findes en eksakt løsning, hvilket også er tilfældet her, men jeg vil hellere vise en generel *metode* for at løse differentialligninger. Idéen er at hvis det er muligt så er udtrykket løsningen og ellers må man søge efter et andet udtryk.

2.4 Taylor polynomier

En funktion $f(x)$ kan udtrykkes ved et polynomium: $c_0 + c_1x + c_2x^2 + \dots$ hvor $c_0, c_1, c_2 \dots$ bestemmes så polynomiet udvikler sig på samme måde som funktionen i et område nær ved nul. Det kræver selvfølgelig at funktionen er defineret omkring nul så for logaritmfunktionerne kan man udvikle polynomiet omkring et punkt a indenfor funktionens definitionsområde hvor $a > 0$. Det betegnes et Taylor polynomium for funktionen omkring punktet a .

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots \quad \text{hvor} \quad c_n = \frac{1}{n!} \frac{d^n f(a)}{dx^n}$$

Konstanten c_n bestemmes ved fortsat differentiation af funktionen $f(x)$ og $d^2f(a)/dx^2$ betyder at funktionen $f(x)$ er differentieret to gange og derefter sættes $x = a$. Udtrykket $n! = 1 \cdot 2 \cdot 3 \cdots n$ kaldes for fakultet n .

Ud over det interessante i at det er muligt at opskrive enhver funktion som et polynomium så er det en overordentlig praktisk metode når komplicerede ligninger skal løses og det er tilstrækkeligt at få en tilnærmelse til resultatet. Det er ofte nok at benytte de to første led af serien der både giver en funktionsværdi c_0 og en hældningskoefficient c_1 så funktionens variation er kendt i den nærmeste omgivelse til $x = a$. Et kompliceret problem kan ofte reduceres til en overkommelig opgave.

Taylor-serien er provokerende enkel at eftervise. Som første trin opskrives Taylor polynomiet og ved at indsætte $x = a$ udgår alle led på nær c_0 som nu er defineret ud fra funktionsværdien.

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots \Rightarrow c_0 = f(a)$$

Som andet trin differentieres Taylor polynomiet hvilket fjerner c_0 og ved at indsætte $x = a$ udgår alle led på nær c_1 som nu kan defineres ud fra funktionsværdien af den første afledte.

$$\frac{df(x)}{dx} = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \dots \Rightarrow c_1 = \frac{df(a)}{dx}$$

Som tredje trin differentieres Taylor polynomiet hvilket fjerner c_1 og ved at indsætte $x = a$ udgår alle led på nær c_2 som nu kan defineres ud fra funktionsværdien af den anden afledte. De øvrige koefficienter bestemmes ved fortsat differentiation efter samme opskrift.

$$\frac{d^2f(x)}{dx^2} = 2c_2 + 3 \cdot 2c_3(x-a) + \dots \Rightarrow c_2 = \frac{1}{2} \frac{d^2f(a)}{dx^2}$$

Som eksempel på opbygningen af et Taylor polynomium vises udledningen for den trigonometriske funktion sinus. Som det ses vil sinus ved små argumenter kunne simplificeres til $\sin(x) \approx x$ hvor x er en vinkelbue (radianer). Dermed ses det at $\sin(x)/x$ fra side 67 vil blive én for $x = 0$.

Differentialkvotienten af sinus giver cosinus og en fortsat differentiation giver $-\sin(x)$ og $-\cos(x)$ hvorefter vi er tilbage ved sin(x) og forløbet starter forfra. Serien udvikles her omkring $a = 0$.

$$\begin{aligned} c_0 &= f(a) = \sin(0) = 0 \\ c_1 &= \frac{df(a)}{dx} = \cos(0) = 1 \\ c_2 &= \frac{1}{2} \frac{d^2f(a)}{dx^2} = -\sin(0) = 0 \\ c_3 &= \frac{1}{2 \cdot 3} \frac{d^3f(a)}{dx^3} = -\cos(0) = -1 \quad \Rightarrow \quad f(x) = \sin(x) = x - \frac{x^3}{2 \cdot 3} + \frac{x^5}{2 \cdot 3 \cdot 4 \cdot 5} + \dots \\ c_4 &= \frac{1}{2 \cdot 3 \cdot 4} \frac{d^4f(a)}{dx^4} = \sin(0) = 0 \\ c_5 &= \frac{1}{2 \cdot 3 \cdot 4 \cdot 5} \frac{d^5f(a)}{dx^5} = \cos(0) = 1 \end{aligned}$$

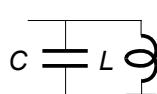
2.4.1 Tabel over Taylor polynomier

Tabel over approksimationer til almindelige funktioner hvor variabel x kan være ethvert reelt eller komplekst tal indenfor de angivne grænser. Bemærk at $-1 < x < 1$ ikke inkluderer grænserne for ved $x = -1$ er eksempelvis $\ln(0)$ ikke defineret da værdien går imod minus uendelig og for $x = 1$ vil rækken ikke konvergere. Polynomierne kan kun tilnærmes med de første led hvis x er nær nul.

Taylor polynomium	Forudsætning
$e^x = \exp(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$	$-\infty < x < \infty$
$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \dots$	$-1 < x < 1$
$\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} \mp \dots$	$-\infty < x < \infty$
$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} \mp \dots$	$-\infty < x < \infty$
$\tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \dots$	$\frac{-\pi}{2} < x < \frac{\pi}{2}$
$\sinh(x) = x + \frac{x^3}{6} + \frac{x^5}{120} + \dots$	$-\infty < x < \infty$
$\cosh(x) = 1 + \frac{x^2}{2} + \frac{x^4}{24} + \dots$	$-\infty < x < \infty$
$\tanh(x) = x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} \pm \dots$	$\frac{-\pi}{2} < x < \frac{\pi}{2}$
$\frac{1}{1+x} = 1 - x + x^2 - x^3 \pm \dots$	$-1 < x < 1$
$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$	$-1 < x < 1$
$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{x^5}{128} \pm \dots$	$-1 < x < 1$
$\frac{1}{\sqrt{1+x}} = 1 - \frac{x}{2} + \frac{3x^2}{8} - \frac{5x^3}{16} + \frac{35x^5}{128} \mp \dots$	$-1 < x < 1$

Et eksempel på brug af Taylor-serien vises hvor meget frekvensen f_0 af et parallelt LC led ændres ved en ændring i selvinduktionen på ΔL eller kapaciteten på ΔC . Det kaldes en sensitivetsanalyse. På grund af kvadratroden vil $+2\%$ ændring i selvinduktionen kun betyde at frekvensen ændres med -1% og tilsvarende gælder for kondensatoren. I beregningen herunder udnyttes det at ændringerne antages så små at produktet $\Delta L \Delta C$ kan ignoreres.

Resonansfrekvens ved variation i L med ΔL og i C med ΔC :



$$f_{0\Delta} = \frac{1}{2\pi\sqrt{(L+\Delta L)(C+\Delta C)}} = \frac{1}{2\pi\sqrt{LC\left(1+\frac{\Delta L}{L}\right)\left(1+\frac{\Delta C}{C}\right)}}$$

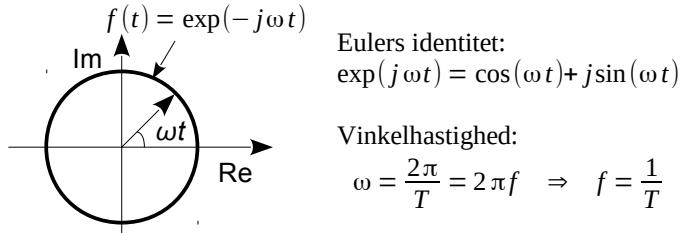
$$f_0 = \frac{1}{2\pi\sqrt{LC}}$$

$$f_{0\Delta} \underset{\Delta C \ll C}{\approx} \frac{1}{2\pi\sqrt{LC\left(1+\left[\frac{\Delta L}{L} + \frac{\Delta C}{C}\right]\right)}}$$

$$\underset{x = \frac{\Delta L}{L} + \frac{\Delta C}{C}}{\approx} \frac{1}{2\pi\sqrt{LC}} \left(1 - \frac{1}{2}\left[\frac{\Delta L}{L} + \frac{\Delta C}{C}\right]\right)$$

2.5 Eulers identitet

En af de vigtigste relationer indenfor de komplekse tals matematik er den identitet som Euler fandt i midten af 1700-tallet. Det er den følgende relation der har betydning indenfor elektronik og andre områder hvor der er behov for at beskrive noget der roterer eller svinger. Relationen fra cosinus og sinus til den komplekse eksponentialfunktion gør det meget lettere at løse de differentialligninger der resulterer af en kredsløbsanalyse som eksemplet på side 63 viser.



Figur 70 – Den komplekse eksponentialfunktion beskriver en cirkel i den komplekse talplan.

Eulers identitet beskriver punkterne langs med en enhedscirkel i den komplekse talplan og de gennemløbes én gang for ωt fra 0 til 2π hvorefter bevægelsen gentages fra 2π til 4π , etc. Parameteren ω er den tilbagelagte vinkelbue i løbet af sekund og kaldes for vinkelhastigheden. Dimensionen er radianer per sekund (rad/s), det er $2\pi \text{ s}^{-1}$, men megen litteratur kalder desværre enheden for s^{-1} og det kan let forveksles med hertz som også har dimensionen s^{-1} . Det indfører en risiko for at miste 2π i en beregning. Vinkelhastigheden ω defineres af periodetiden T for en hel svingning eller den frekvens f som bevægelsen gentages med.

At sinus og cosinus indgår i udtrykket for den komplekse eksponentialfunktion betyder at de trigonometriske funktioner kan udtrykkes alene ved den komplekse eksponentialfunktion. Det har den betydning at sammensatte udtryk kan simplificeres væsentligt.

$$\begin{aligned}\cos(\theta) &= \frac{\exp(j\theta) + \exp(-j\theta)}{2} & \cos(j\theta) &= \frac{\exp(-\theta) + \exp(\theta)}{2} = \cosh(\theta) \\ \sin(\theta) &= \frac{\exp(j\theta) - \exp(-j\theta)}{2j} \quad \text{og} & \sin(j\theta) &= \frac{\exp(-\theta) - \exp(\theta)}{2j} = j \sinh(\theta)\end{aligned}$$

Eulers identitet kan også benyttes til at definere de trigonometriske funktioner med et komplekst argument. De komplekse trigonometriske funktioner ses at være identiske med de hyperboliske funktioner cosh og sinh.

Eulers identitet skal vises for cosinus. Først opskrives Eulers identitet for både positivt og negativt argument. Det udnyttes at cosinus er en lige funktion $\cos(-\theta) = \cos(\theta)$ mens sinus er en ulige funktion $\sin(-\theta) = -\sin(\theta)$ og ved at addere de to udtryk opnås udtrykket for cosinus.

$$\begin{aligned}\exp(j\theta) &= \cos(\theta) + j \sin(\theta) \\ \exp(-j\theta) &= \cos(\theta) - j \sin(\theta)\end{aligned} \stackrel{\text{sum}}{\Rightarrow} \exp(j\theta) + \exp(-j\theta) = 2 \cos(\theta) \Rightarrow \cos(\theta) = \frac{\exp(j\theta) + \exp(-j\theta)}{2}$$

Som et eksempel på anvendelserne vises at kvadratet på sinus kan skrives ved cosinus til det dobbelte argument. Sinus erstattes med Eulers identitet og derefter kvadreres udtrykket.

$$\sin^2(x) = \left(\frac{\exp(jx) - \exp(-jx)}{2j} \right)^2 = \frac{\exp(j2x) + \exp(-j2x) - 2\exp(0)}{-4} = \frac{1}{2}(1 - \cos(2x))$$

Ved at gange på begge sider af Eulers identitet med eksponentialfunktion $\exp(\sigma)$ hvor argumentet σ er et reelt tal kan relationen udvides til at gælde for ethvert komplekst tal $\sigma + j\omega$ og ikke kun for de imaginære tal. Ved en negativ værdi af σ hvilket er det normale indenfor elektriske kredsløb vil det beskrive en svingning der lidt efter lidt dør ud med en hastighed bestemt af parameteren σ .

$$\exp(\sigma t + j\omega t) = \exp(\sigma t)[\cos(\omega t) + j \sin(\omega t)]$$

Dette udtryk beskriver hvordan et elektrisk, mekanisk eller akustisk system vil svinge efter det først er startet. Det ses at et fysisk realisabelt system vil have $\sigma + j\omega$ placeret i venstre halvplan for at sikre en hendøende bevægelse, at $\sigma = 0$ vil svare til en tonegenerator og at $\sigma > 0$ give en stadigt stigende amplitude. Desuden ses at $\omega = 0$ give en almindelig eksponentiel relation.

Herunder udledes de viste formler. Eulers identitet vil blive benyttet i de efterfølgende kapitler mens der ikke vil forekomme referencer til de trigonometriske funktioner med komplekst argument i denne bog.

De trigonometriske funktioner kan udledes ved henholdsvis addition og subtraktion af Eulers formler med både positivt og negativt argument. Ved at addere de to ligninger elimineres sinus og tilsvarende elimineres cosinus ved subtraktion. Udtrykkene fremkommer ved at løse de to udtryk for henholdsvis cosinus og sinus.

$$\begin{aligned}\exp(j\theta) &= \cos(\theta) + j\sin(\theta) \\ \exp(-j\theta) &= \cos(\theta) - j\sin(\theta)\end{aligned} \Rightarrow \begin{array}{lll}\text{Sum} & \exp(j\theta) + \exp(-j\theta) &= 2\cos(\theta) \\ \text{Differens} & \exp(j\theta) - \exp(-j\theta) &= 2j\sin(\theta)\end{array}$$

De komplekse trigonometriske funktioner kan udledes på samme måde blot erstattes θ med $j\theta$ som derefter indsættes i Eulers identitet.

$$\begin{aligned}\exp(j(j\theta)) &= \cos(j\theta) + j\sin(j\theta) \\ \exp(-j(j\theta)) &= \cos(j\theta) - j\sin(j\theta)\end{aligned} \Rightarrow \begin{array}{lll}\text{Sum} & \exp(-\theta) + \exp(\theta) &= 2\cos(j\theta) \\ \text{Differens} & \exp(-\theta) - \exp(\theta) &= 2j\sin(\theta)\end{array}$$

Eulers identitet kan eftervises ved at opstille det følgende udtryk for funktionen $f(\theta)$ der herefter differentieres. For at forberede denne proces defineres to funktioner $g(\theta)$ og $h(\theta)$ der står for henholdsvis de trigonometriske udtryk og den komplekse eksponentialfunktion.

$$f(\theta) = [\cos(\theta) + j\sin(\theta)]\exp(-j\theta) \Rightarrow f(\theta) = \frac{\cos(\theta) + j\sin(\theta)}{\exp(j\theta)} = \frac{g(\theta)}{h(\theta)}$$

For at differentiere udtrykket har vi brug for en regel for differentiation af en brøk af funktioner.

$$f(\theta) = \frac{g(\theta)}{h(\theta)} \Rightarrow \frac{df(\theta)}{dt} = \frac{\frac{dg(\theta)}{d\theta}h(\theta) - g(\theta)\frac{dh(\theta)}{d\theta}}{[h(\theta)]^2}$$

Ved at differentiere den sammensatte funktionen ses det at resultatet bliver nul.

$$\begin{aligned}\frac{df(\theta)}{dt} &= \frac{[-\sin(\theta) + j\cos(\theta)]\exp(j\theta) - j[\cos(\theta) + j\sin(\theta)]\exp(j\theta)}{\exp(j2\theta)} \\ \frac{df(\theta)}{dt} &= \frac{[-\sin(\theta) + j\cos(\theta) - j\cos(\theta) - \sin(\theta)]\exp(j\theta)}{\exp(j2\theta)} = 0\end{aligned}$$

Når differentialkvotienten er nul må funktionen $f(\theta)$ være konstant og den konstante værdi kan findes ved at bestemme funktionsværdien for en værdi hvor det er let at foretage beregningen. Her benyttes $\theta = 0$ der giver $f(0) = 1$. Den værdi gælder uanset værdien af θ idet funktionen er konstant og dermed uafhængig af θ . Det giver Eulers identitet ved at gange på begge sider af ligningen med den komplekse eksponentialfunktion.

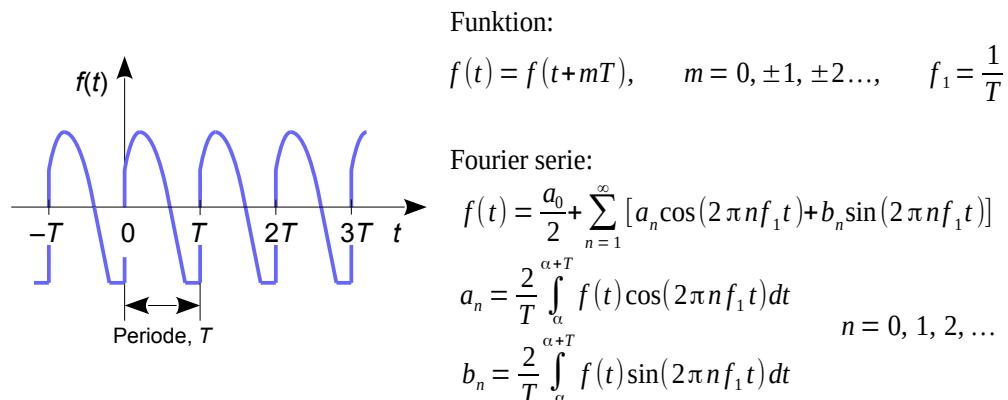
$$f(\theta) = [\cos(\theta) + j\sin(\theta)]\exp(-j\theta) = 1 \Rightarrow \exp(j\theta) = \cos(\theta) + j\sin(\theta)$$

2.6 Fourier serier

1807 publicerede Fourier et banebrydende værk om hvordan varmeenergi breder sig igennem en plade fra en varmekilde. Samtiden kendte til løsninger hvor varmekildens styrke varierede som en harmonisk svingning (sinus eller cosinus) og man eftersøgte en løsning hvor varmekilden kunne beskrives generelt. Det var Fouriers idé at opbygge løsningen for en vilkårlig kurveform ved brug af harmoniske svingninger for derved at basere løsningen på det allerede kendte⁵².

Fouriers arbejde prægede udviklingen indenfor matematikken også selv om han mødte en del modstand for den påne sinus var ifølge Fourier i stand til at eftergøre diskontinuerte svingninger. Det er funktioner som springer abrupt mellem værdierne og det havde samtidens matematikere lidt svært ved at acceptere. Fouriers arbejde har fundamental betydning for den moderne matematik og finder anvendelse ved komprimering af data for HDTV (*high-definition television*) og tillader en kredsløbsanalyse at beskrive stepresponsen baseret på analyse med en harmonisk svingning.

En Fourier serie efterligner en svingning der fx kan være et musikinstrument svingende streng. For at kunne definere en Fourier serie kræves det at signalet er periodisk så det skal gentage sig selv i det uendelige med en fast periodetid der i det følgende vil blive kaldt for T . Funktionen $f(t)$ har derved en repetitionsfrekvens f_1 der kaldes for den første harmoniske. Fourier serien $f(t)$ defineres som en sum af harmoniske svingninger med frekvenserne $f_1, 2f_1, 3f_1, \dots$ og med individuelle vægte ved koefficienterne a_n og b_n . Serien går imod $f(t)$ når antallet af svingninger n går imod uendelig.



Figur 71 – Svingningen skal være periodisk og alle perioder skal være identiske.

Det er relativt let at vise at koefficienterne bestemmes som vist ovenfor hvor α kan vælges fri. Det er kun krævet at der integreres over et helt antal perioder. Alle integraler kan derfor foretages hvor det er lettest at slippe igennem fx fra 0 til T eller fra $-T/2$ til $T/2$.

Bemærk at for $n = 0$ er $\cos(2\pi n f_1 t) = 1$ hvorev a_0 bliver 2 gange middelværdien af $f(t)$ så $a_0/2$ angiver middelværdien af funktionen (DC værdien). For $n = 0$ er $\sin(2\pi n f_1 t) = 0$ så b_0 er altid nul.

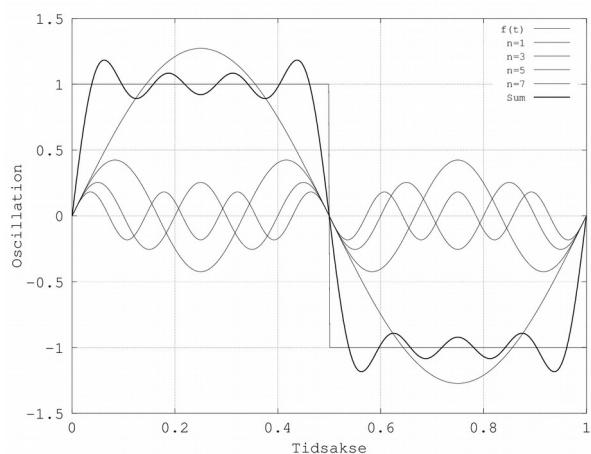
Integralerne udfører en test af om eksempelvis svingningen $\cos(2\pi f_1 t)$ er indeholdt i funktionen $f(t)$. Hvis det er tilfældet vil integralet indeholde kvadratet på $\cos(2\pi f_1 t)$ som giver et resultat for a_1 . Hvis $\cos(2\pi f_1 t)$ ikke er indeholdt vil integrationen give $a_1 = 0$ da sinus og cosinus giver nul ved integration over en periode. Tilsvarende tests udføres for koefficienten b_1 og for de højere harmoniske a_2, b_2 , etc. Se også side 80.

En vigtig konklusion på Fouriers arbejde er at de analyser vi laver på elektriske kredsløb kan bruges på alle former for komplicerede signaler blot signalerne er periodiske. Det er uanset om der er tale om svingninger i en maskine eller et akustisk signal. Analysen kan derfor vise om der er specifikke frekvenser hvor eksempelvis en maskine vil svinge så kraftigt at der er risiko for slitage. Vi kan derfor karakterisere et elektronisk system ved brug af en harmonisk svingning og derefter konkludere hvordan det vil fungere ved eksempelvis en firkantsvingning.

52 Det er ikke kun sinus og cosinus der har denne egenskab. Andre muligheder er Bessel, Chebychev og Walsh funktioner.

2.6.1 Firkantsvingning

En vigtig form for svingning er den symmetriske firkant der svinger mellem 1 og -1 med perioden T svarende til frekvensen $f_1 = 1/T$. Det er en typisk brugt påvirkning af et system. Eksemplet viser at svingningen ikke indeholder lige harmoniske hvilket gælder generelt for ulige funktioner⁵³.



$$f(t) = \begin{cases} 1 & 0 < t < \frac{T}{2} \\ -1 & \frac{T}{2} < t < T \end{cases} \quad f(t+mT) = f(t) \quad m \text{ heltal}$$

$$f(t) = \sum_{n \text{ ulige}} \frac{4}{n\pi} \sin(n2\pi\frac{t}{T})$$

Harmoniske:

$n = 1$	$\frac{4}{\pi} = 1,2732$	+2,1 dB
$n = 3$	$\frac{4}{3\pi} = 0,4244$	-7,4 dB
$n = 5$	$\frac{4}{5\pi} = 0,2546$	-11,9 dB

Figur 72 – Eksempel på opbygning af en firkantsvingning med harmoniske svingninger.

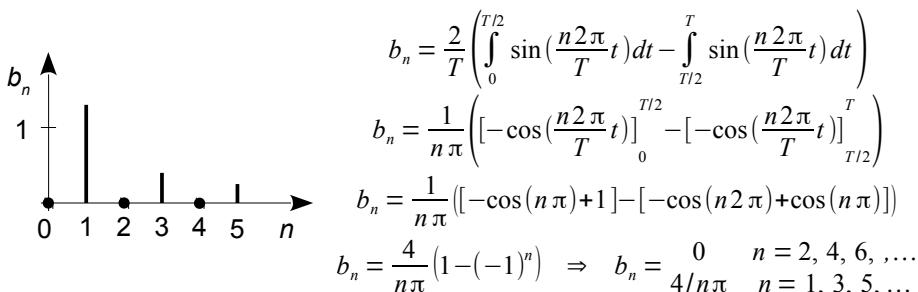
Den første harmoniske har en amplitude på $4/\pi = 1,27$ gange firkantsvingningens amplitude hvilket betyder at lavpasfiltrering af en firkant vil give en 27 % højere amplitude på udgangen af filtret.

Svingningen i området før og efter niveauskiftene er typisk for Fourier serierne og er et resultat af at serien afsluttes abrupt (Gibbs fænomen) så n ikke vokser mod uendelig. Fourier serierne viser at ethvert periodisk signal kan repræsenteres ved en sum af svingninger af typen cosinus og sinus. De enkelte leds frekvens er arrangeret efter den harmoniske række $(1, 2, 3, \dots)$ og der er ikke mulighed for svingninger i "mellemrummene". I den ældre litteratur tales om en grundtone (første harmoniske) og dertil en række overtoner, men der opstår let forvirring om nummereringen så det er sikrest at benytte betegnelsen harmoniske.

Integralet af $f(t)\cos(2\pi nf_1t)$ deles op ved $T/2$ så $f(t)$ er konstant indenfor intervallet og værdien kan sættes udenfor som et fortegn. Beregningen af a_n udnytter at integralet af cosinus over et multiplum af halve perioder er nul så alle koefficienterne til cosinus er nul.

$$a_n = \frac{2}{T} \left(\int_0^{T/2} \cos\left(\frac{n2\pi}{T}t\right) dt - \int_{T/2}^T \cos\left(\frac{n2\pi}{T}t\right) dt \right) = 0 \Rightarrow a_n = 0, \quad n = 0, 1, 2, \dots$$

Signalets symmetri medfører at alle lige harmoniske er nul som det skal vises for beregning af koefficienterne b_n . Det udnyttes at $\cos(n\pi)$ svinger med ± 1 for lige og ulige værdi af index n og at $\cos(2n\pi)$ er lig med én for alle værdier af n .



53 En ulige funktion opfylder betingelsen $f(x) = -f(-x)$ som fx sinus og det kaldes også for punktsymmetri hvor punktet her er $(0, 0)$. En lige funktion opfylder betingelsen $f(x) = f(-x)$ som fx cosinus og det kaldes også for spejlsymmetri hvor spejlet her er ordinataksen.

2.7 Kompleks Fourier serie

En hensigtsmæssig omformning af Fourier serien benytter de komplekse repræsentanter for sinus og cosinus ved brug af Eulers identitet fra side 75 der indsættes i formlerne for a_n og b_n og koefficienterne samles til en kompleks amplitude defineret som $c_n = (a_n - jb_n)/2$. Efter lidt arbejde kan den kompleks Fourier serie skrives som vist herunder (bemærk at der er to måder at skrive den på) og udtrykkene til højre beregner amplituden (modulus) og fasen af den n 'te harmoniske.

$$f(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_1 t) \quad f_1 = \frac{1}{T}$$

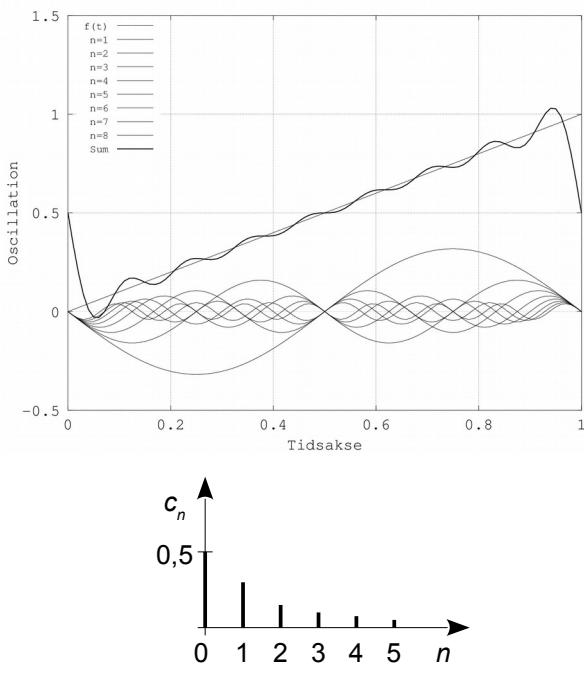
$$f(t) = c_0 + \sum_{n=1}^{\infty} (c_n \exp(j2\pi n f_1 t) + c_{-n} \exp(-j2\pi n f_1 t)) \quad |c_n| = \sqrt{(\operatorname{Re}(c_n))^2 + (\operatorname{Im}(c_n))^2}$$

$$c_n = \frac{1}{T} \int_a^{a+T} f(t) \exp(-j2\pi n f_1 t) dt \quad \theta_n = \arctan\left(\frac{\operatorname{Im}(c_n)}{\operatorname{Re}(c_n)}\right)$$

Der er principielt set intet nyt i den kompleks Fourier serie der dog i visse situationer kan være lettere at løse analytisk. Herunder skal den illustreres med et eksempel og derefter vises at den med en let omskrivning fører til Fourier transformationen.

2.7.1 Savtaksvingning

I lighed med firkantsvingningen er savtaksvingningen et eksempel på et repeteret signal med stor praktisk værdi. En savtak beskrives her ved en lineær stigning fra 0 ved $t = 0$ til 1 ved $t = T$ og den repeteres herefter. Ved indsættelse af funktionen t/T i udtrykket for c_n findes følgende resultat.



Figur 73 – Eksempel på opbygning af en savtaksvingning med harmoniske svingninger.

Udgangspunktet er en reel funktion så Fourier serien for $f(t)$ består kun af reelle led og resultatet bliver en sum af harmoniske funktioner uden imaginære dele. Springet fra slutværdien på én til startværdien på nul går gennem middelværdien på 0,5 ved $t = 0$ og $t = T$ og forløbet bliver skarpere (mere abrupt skift fra 1 til 0) ved et højere antal af harmoniske.

Savtak:

$$f(t) = \frac{t}{T} \quad 0 \leq t < T$$

Periode T :

$$f(t+mT) = f(t) \quad m \text{ heltal}$$

Harmoniske amplitude:

$$c_n = \frac{1}{T} \int_0^T \frac{t}{T} \exp(-j2\pi n f_1 t) dt, \quad f_1 = \frac{1}{T}$$

For $n = 0$ findes DC værdien:

$$c_0 = \frac{1}{T} \int_0^T \frac{t}{T} dt = \frac{1}{T^2} \left[\frac{t^2}{2} \right]_0^T = \frac{1}{2}$$

For $n > 0$ findes ved tabelopslag [RW-175]:

$$c_n = \frac{1}{T^2} \left[\frac{-j2\pi n f_1 t - 1}{(-j2\pi n f_1)^2} \exp(-j2\pi n f_1 t) \right]_0^T$$

$$c_n = \frac{j}{2\pi n}, \quad n = 1, 2, 3, \dots$$

Fourier serien bliver:

$$f(t) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{1}{\pi n} \sin(2\pi n f_1 t)$$

2.7.2 Beregning af koefficienterne

Det kan virke mystisk at koefficienterne kan beregnes blot ved at integrere funktionen $f(t)$ ganget med en harmonisk svingning så det uddybes i dette afsnit. Den følgende matematik viser egentlig blot at en funktion $f(t)$ kan "testes" for om den indeholder en svingning som $\cos(2\pi f_1 t)$ ved at gange funktionen med svingningen og integrere over en hel periode. Hvis svingningen optræder som del af funktionen vil $\cos(2\pi f_1 t)$ blive kvadreret og er dermed positiv eller nul hvorved integrationen vil give et bidrag, men ellers vil resultatet blive nul.

For at gennemføre eksemplet er der brug for at vise at en sinus eller cosinus integreret over en hel periode giver nul. Forklaringen er at svingningen er akkurat lige så meget positiv indenfor perioden som den er negativ. Resultatet skal benyttes lidt senere i eksemplet.

$$\int_0^T \sin(2\pi f_1 t) dt = \left[\frac{-\cos(2\pi f_1 t)}{2\pi f_1} \right]_0^T \Big|_{f_1 = \frac{1}{T}} = \frac{-\cos(2\pi) - (-\cos(0))}{2\pi f_1} = 0$$

$$\int_0^T \cos(2\pi f_1 t) dt = \left[\frac{\sin(2\pi f_1 t)}{2\pi f_1} \right]_0^T \Big|_{f_1 = \frac{1}{T}} = \frac{\sin(2\pi) - \sin(0)}{2\pi f_1} = 0$$

Det er lettest at vise beregningen for den komplekse Fourier serie, men da cosinus og sinus indgår i den komplekse eksponentialfunktion er eksemplet herunder generelt og inkluderer derfor Fourier serien fra side 77.

$$\exp(j 2\pi f_1 t) = \cos(2\pi f_1 t) + j \sin(2\pi f_1 t)$$

Den komplekse Fourier serie konstruerer en funktion $f(t)$ ved addition af harmoniske svingninger på grundfrekvensen f_1 (første harmoniske) og de harmoniske ved $2f_1, 3f_1$, etc. Koefficienterne c_n angiver vægtningen (betydningen) af hver enkelt harmoniske.

$$f(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j 2\pi n f_1 t), \quad n = 1, 2, 3, \dots$$

Begge sider af lighedstegnet ganges med $\exp(-j 2\pi m f_1 t)$ hvor m er et helt tal ligesom n er det. De to eksponentialfunktioner samles til en enkelt ved at addere argumenterne:

$$f(t) \exp(-j 2\pi m f_1 t) = \sum_{n=-\infty}^{\infty} c_n \exp(j 2\pi n f_1 t) \exp(-j 2\pi m f_1 t)$$

$$f(t) \exp(-j 2\pi m f_1 t) = \sum_{n=-\infty}^{\infty} c_n \exp(j 2\pi(n-m)f_1 t)$$

Der integreres på begge sider over en hel periode. Det udnyttes at integrationen er en lineær operation og derfor kan foretages isoleret for hvert enkelt led i en sum. Det udnyttes også at koefficienten c_n er uafhængig af tiden og kan sættes ud foran integrationen. For $m = n$ vil differensen $n - m$ blive nul og eksponentialfunktionen bliver lig med én hvilket simplificerer integrationen til tiden T . For alle andre m og n vil integrationen benytte et helt antal perioder af den komplekse eksponentialfunktions cosinus og sinus og integralerne bliver derfor nul. Det eneste bidrag kommer ved $m = n$ og ved at løse ligningen for c_n fremkommer relationen.

$$\int_0^T f(t) \exp(-j 2\pi m f_1 t) dt = \int_0^T \left(\sum_{n=-\infty}^{\infty} c_n \exp(j 2\pi(n-m)f_1 t) \right) dt$$

$$\int_0^T f(t) \exp(-j 2\pi m f_1 t) dt = \sum_{n=-\infty}^{\infty} \left(c_n \int_0^T \exp(j 2\pi(n-m)f_1 t) dt \right)$$

$$m = n \Rightarrow \int_0^T f(t) \exp(-j 2\pi n f_1 t) dt = c_n \int_0^T dt = c_n T \Rightarrow c_n = \frac{1}{T} \int_0^T f(t) \exp(-j 2\pi n f_1 t) dt$$

$$m \neq n \Rightarrow \int_0^T f(t) \exp(-j 2\pi n f_1 t) dt = 0$$

2.8 Fourier transformation

Bestemmelsen af koefficienten c_n for et periodisk signal kan udvides til at omfatte signaler der ikke repeteres. Det kan sandsynliggøres ved at integrere fra $-T/2$ til $T/2$ og studere hvad der sker hvis perioden T går imod uendelig. Repetitionsfrekvensen f_1 vil gå imod nul og ligeledes vil afstanden mellem de harmoniske så resultatet bliver et kontinuerligt spektrum. Den komplekse koefficient c_n bliver til en kontinuerlig funktion af frekvensen som kaldes for $F(f)$. Faktoren $1/T$ foran integralet vil gå imod nul, men der summeres nu over et uendeligt stort tidsområde.

$$\text{Transformation af } f(t) \text{ til frekvensdomænet: } F(f) = \int_{-\infty}^{\infty} f(t) \exp(-j2\pi ft) dt$$

$$\text{Transformation af } F(f) \text{ til tidsdomænet: } f(t) = \int_{-\infty}^{\infty} F(f) \exp(j2\pi ft) df$$

Transformationen omformer en tidsafhængig funktion $f(t)$ til en frekvensafhængig funktion $F(f)$ der indeholder den samme information. Det kan udnyttes til at løse et kompliceret problem ved først at transformere funktionen $f(t)$ til frekvensdomænets $F(f)$ hvor problemet løses og så transformere det fundne tilbage til tidsdomænet. Alternativt kan en datamængde reduceres ved at transformere den til frekvensdomænet hvor frekvensområdet begrænses og signalet derefter transformeres tilbage til tidsdomænet som en ny $f(t)$ der fylder mindre, men mangler de højfrekvente data.

Transformationen ændrer dimensionen idet der ganges med tid. Hvis funktionen $f(t)$ er en spændingskilde med dimensionen volt vil den Fourier-transformerede have dimension af volt gange sekund.

En del litteratur erstatter frekvensvariablen df med vinkelhastigheden $d\omega = 2\pi df$ der medfører division med 2π ved den inverse Fourier-transformation. En del matematiske værker "fordeler" så denne konstant ligeligt mellem transformationens to dele ved at gange med $1/(2\pi)$ begge steder. Det er en misforståelse, men den er desværre meget udbredt og mange tabelværker har denne konstant inkluderet.

2.8.1 Rektangulær puls

Som eksempel på Fourier transformationen vises herunder resultatet for en rektangulær puls $f(t)$ der er nul for al negativ tid, der til tidspunktet $t = 0$ springer til værdien én hvor den forbliver til $t = T$ for så at springe tilbage til nul. Det er en ofte benyttet puls i videnskabelig teori og ved praktiske målinger på elektroniske, mekaniske og akustiske systemer og desuden danner den fundamentet for at forstå digital signalbehandling. For at Fourier transformere pulsen udnyttes det at den er nul for $t < 0$ og igen nul for $t > T$ hvilket klares ved at justere på integrationens grænser.

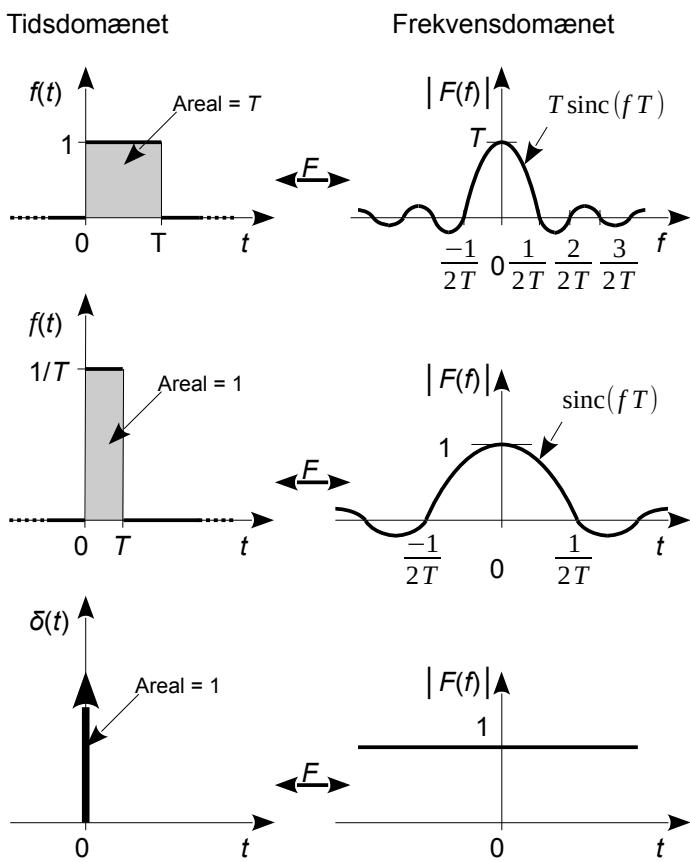
Resultatet kan tolkes med kurverne side 67, men det er mere sigende at beregne modulus af $H(f)$ som giver en sinc-funktion. Den mødes i de situationer hvor noget skifter abrupt som eksempelvis dataopsamling indenfor et givet tidsrum hvor et tidsvindue åbnes og lukkes igen. Den beskriver frekvensdomænets reaktion på de abrupte skift mellem tænd og sluk samt varigheden. Funktionen er defineret som $\text{sinc}(x) = \sin(\pi x)/\pi x$ og det kan vises at den er én for $x = 0$ og ellers periodisk.

Hvis den rektangulære puls gøres smal vil sinc-funktionen bredes ud så den dækker et større frekvensområde og ved en meget smal puls vil sinc-funktionen blive nærmest konstant over et stort frekvensområde. Da energiindholdet i pulsen er funktion af varigheden af pulsen vil energien dog aftage. Lydmæssigt bliver den smalle rektangulære puls mere "lys" i sin klang da den dækker et stort frekvensområde mens den samtidigt bliver svagere da energien aftager.

En puls med varighed $T = 1$ ms giver et støjsignal med båndbredden $1/2T = 500$ Hz regnet fra 0 Hz til første nulpunkt. Højere frekvenser forekommer, men amplituden er -15 dB eller derunder. Med $T = 20$ μ s vil frekvensbåndet dække det hørbare område med omtrent konstant amplitude ($-2,4$ dB ved 20 kHz) mens amplituden er reduceret med små 35 dB i forhold til pulsen på 1 ms.

Fourier transformationen viser at en hurtig ændring svarer til en stor båndbredde som vi allerede så ved Fourier serierne hvor de fine strukturer i firkant- og savtaksvingningernes skift mellem lav og høj beskrives ved mange høje harmoniske der til gengæld har lav amplitude.

Analyse

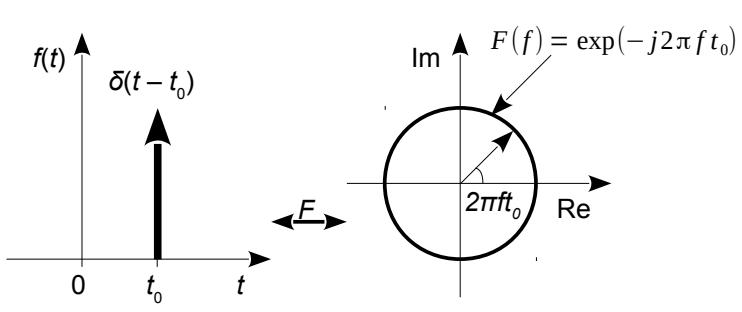


Figur 74 – En rektangulær puls har mest signalenergi ved lave frekvenser mens energien ved de høje frekvenser aftager i styrke og er nul ved de specifikke frekvenser $f = n/2T$ hvor n er et helt tal.

For at opnå konstant energi uanset varigheden af den rektangulære puls kan amplituden sættes til $1/T$ så arealet forbliver én uanset pulsens varighed. Hvis T går imod nul vil amplituden vokse mod uendelig mens arealet forbliver én. Det kaldes en Dirac puls der er væsentlig indenfor teoretisk fysik og ved digital signalbehandling.

2.8.2 Dirac puls

En Dirac puls $\delta(t)$ fyrer når argumentet er nul og det noteres ved en lodret pil ved det pågældende tidspunkt der i illustrationen herunder er $t = t_0$. Integralet af Dirac pulsen er én uanset hvornår den fyrer og hvis $\delta(t - t_0)$ ganges på en funktion $f(t)$ vil integrationen aflevere funktionsværdien til t_0 . Ved en Fourier transformation er resultatet derfor en kompleks oscillation i frekvensdomænet.



Definition af Dirac puls:

$$\int_{-\infty}^{\infty} \delta(t) dt = 1$$

$$\int_{-\infty}^{\infty} \delta(t-t_0) f(t) dt = f(t_0)$$

Fourier transformation af Dirac puls:

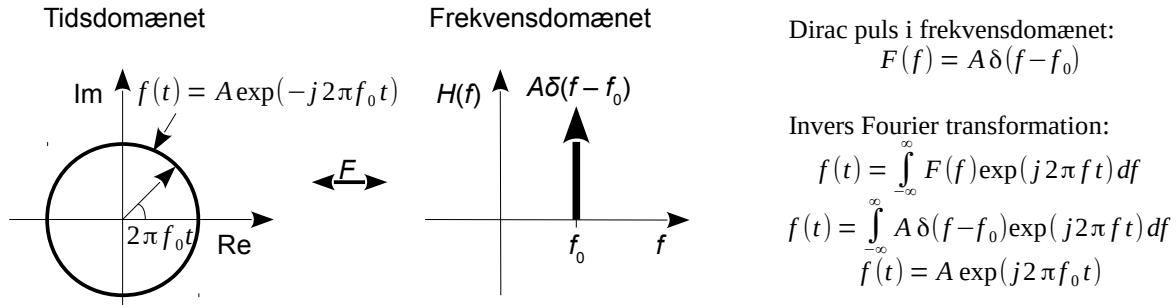
$$f(t) = \delta(t-t_0) \xrightarrow{F} F(f) = \int_{-\infty}^{\infty} \delta(t-t_0) \exp(-j2\pi f t) dt$$

$$F(f) = \exp(-j2\pi f t_0)$$

Figur 75 – Dirac pulsen fyrer når argumentet er nul. Fourier transformationen giver en kompleks eksponentialefunktion hvor amplituden er én mens fasen roterer med en hastighed givet ved t_0 .

2.8.3 Kompleks harmonisk svingning

En kompleks eksponentialfunktion i tidsdomænet bliver Fourier transformert til en Dirac puls i frekvensdomænet. Det betyder at Fourier transformationen omformer en tidsmæssig svingning til en Dirac puls på svingningens frekvens. Amplituden af svingningen er bibeholdt som faktor A.



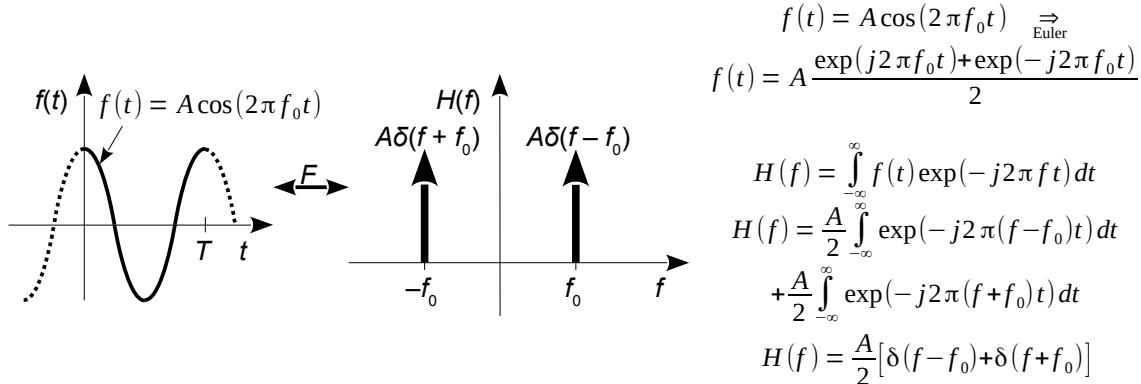
Figur 76 – En harmonisk oscillation transformeres til en Dirac puls i frekvensdomænet.

Det er lettest at udlede den modsatte vej. Man planter en Dirac puls i frekvensdomænet og bruger den inverse Fourier transformation til tidsdomænet for at se hvad der sker.

Vi så tidligere at en kompliceret svingning kunne opbygges ved en serie af harmoniske svingninger som eksemplificeret side 78 og 79. I frekvensdomænet optræder disse svingninger som en serie af Dirac pulser placeret ved de harmoniske frekvenser og med individuel vægtning.

2.8.4 Reel harmonisk svingning

De oscillerende funktioner sinus og cosinus kan Fourier transformeres ved brug af Eulers identitet. Da de hver beskrives ved to eksponentialfunktioner bliver der to Dirac pulser i frekvensdomænet. Med frekvensen f_0 vil svingningen optræde i frekvensdomænet med Dirac pulser ved f_0 og $-f_0$ som vist herunder for en cosinus. Forskellen mellem sinus og cosinus ligger alene i fasen da sinus vil introducere den imaginære enhed i Eulers identitet samt et fortegnsskifte for den ene Dirac puls.



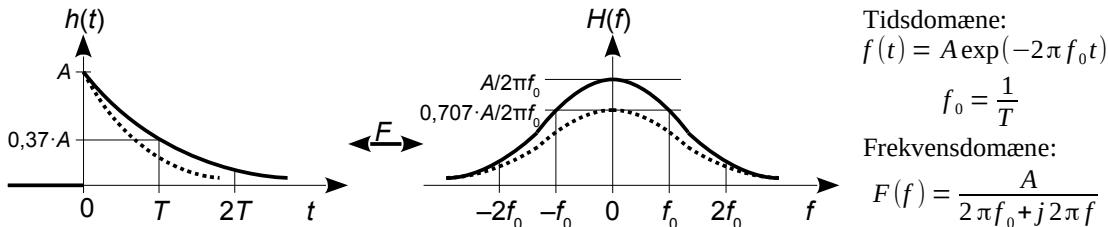
Figur 77 – En cosinus transformeres til to Dirac pulser i frekvensdomænet.

2.8.5 Negativ frekvens

Fourier transformationen introducerer en negativ frekvens idet definitionsområdet er fra minus uendelig til plus uendelig for både tid og frekvens. Man kan opfatte negativ tid som fortid og positiv tid som fremtid, men negativ frekvens har ikke en intuitiv tolkning. Det er lidt som at acceptere at en fysisk variabel kan beskrives med komplekse tal. En fysisk variabel kan aldrig være kompleks, men der opnås en stor fordel ved at acceptere abstraktionen. Det gælder det samme med negativ frekvens som derfor må opfattes som endnu et lag i den matematiske abstraktion fra det fysiske.

2.8.6 Eksponentiel puls

Som det sidste eksempel på Fourier transformationen vises resultatet af en eksponentiel puls der starter ved $t = 0$ hvor dens funktionsværdi springer op på én og som derefter klinger ud. Det kan fx repræsentere udgangssignalet fra en AC koblet forstærker hvor indgangen pludseligt tilsluttes en DC spænding. Udgangen vil så springe op på en maksimalværdi og derefter glide tilbage til udgangspunktet med tidskonstanten T .



Figur 78 – En eksponentiel puls med varigheden $2T$ vil i frekvensdomænet svare til et støjniveau der aftager over frekvensen $f_0 = 1/2\pi T$ med $-6 \text{ dB}/\text{oktav}$.

Frekvensspektret er tosidigt hvilket vil sige at Fourier-transformationen benytter både positive og negative frekvenser. Det har primært akademisk interesse så for en ingeniørsmæssig anvendelse kan man nøjes med at betragte den positive del af frekvensområdet idet de to sider er spejlbilleder af hinanden. Der er dog en forskel i fortegnet af fasen for de to frekvensområder.

I frekvensdomænet optræder signalet som et klik med en båndbredde givet ved den hastighed hvormed udgangssignalet glider på plads. Grænsefrekvensen defineres som -3 dB der nås når frekvensen er lig med $f = f_0 = 1/T$ idet nævneren af polynomiet bliver $1 + j$ gange $2\pi f_0$ som er lig med kvadratroden af to gange $2\pi f_0$.

Ved høje frekvenser ($f > f_0$ og kort tidskonstant) er niveauet af den Fourier transformerede omvendt proportional med frekvensen. Det ses ved at udtrykket kan simplificeres til $F_{HF}(f) = A/2\pi f$ der er aftagende ved stigende frekvens. Det betyder at niveauet af de højfrekvente signaler ved ikke er funktion af grænsefrekvensen så længe den er lavere end det betragtede frekvensområde.

Ved lave frekvenser ($f < f_0$ og lang tidskonstant) er den Fourier transformeredes niveau konstant. Det ses ved at udtrykket kan simplificeres til $F_{DC}(f) = A/2\pi f_0$ der ikke er funktion af frekvensen, men kun af den eksponentielle puls tidskonstant. Det betyder at DC niveauet $F_{DC}(f)$ vokser for aftagende grænsefrekvens og går imod uendelig hvis signalet aldrig dør ud, men forbliver på niveauet A .

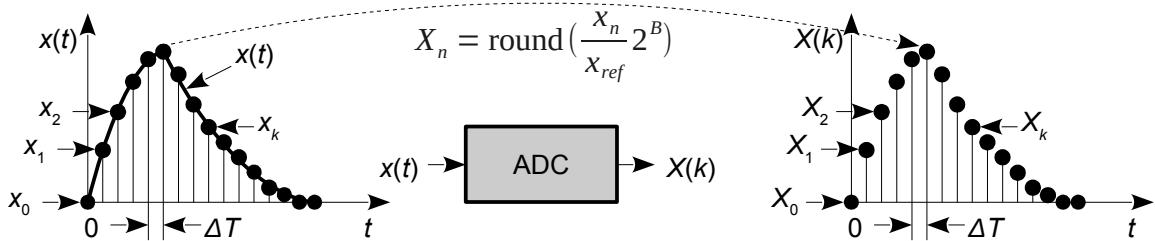
Beregningen af den Fourier-transformerede vises herunder.

Funktionen defineres og indsættes Fourier-transformationen. Da funktionen er nul for negativ tid kan integrationen startes ved $t = 0$. Produktet af to eksponentialfunktioner kan skrives som en enkelt eksponentialfunktion hvor de to argumenter er adderet. Integrationen gennemføres og det udnyttes at $\exp(-\infty)$ er lig med nul.

$$\begin{aligned} f(t) &= \begin{cases} A \exp(-2\pi f_0 t) & t \geq 0 \\ 0 & t < 0 \end{cases} \quad f_0 = \frac{1}{T} \\ F(f) &= A \int_0^\infty \exp(-2\pi f_0 t) \exp(-j 2\pi f t) dt \\ F(f) &= A \int_0^\infty \exp(-[2\pi f_0 + j 2\pi f] t) dt \\ F(f) &= -\frac{A}{2\pi f_0 + j 2\pi f} [\exp(-[2\pi f_0 + j 2\pi f] t)]_0^\infty \\ F(f) &= \frac{A}{2\pi f_0 + j 2\pi f} \end{aligned}$$

2.9 Diskret Fourier transformation

Bestemmelse af koefficienterne for Fourier serien er et værktøj som finder anvendelse til analyse af opsamlede data fra mikrofon, accelerometer og andre målesystemer. Analysen udtrykkes som amplitude versus frekvens og tillader dertil måling af harmonisk forvrængning og støjniveau. Det er sjældent at signalerne er repetitive, men det viser sig at Fourier serierne er robuste og kan bruges for ikke-periodiske signaler blot de indeholder meget gentagen information og i det følgende vises hvordan det kan implementeres i en computer.



Figur 79 – Et tidskontinuert indgangssignal bliver omsat til en tilsvarende numerisk værdi på diskrete tidspunkter. Her er signalet kun positivt, men det er normalt at definere et nulpunkt i midten af konverterens område så der kan håndteres signaler som svinger omkring nul.

Når indgangssignalet ikke er en matematisk funktion, men kommer fra den fysiske virkelighed, så skal det omsættes fra et analogt signal til digitale talværdier (*sampling*). Signalet føres gennem en analog til digital konverter (ADC) hvor det omsættes til en følge af numeriske værdier der er en gengivelse af signalet ved et fast tidsinterval mellem talværdierne. Illustrationen ovenfor benytter indgangssignalet $x(t)$ der repræsenteres af punkterne x_0, x_1, x_2, \dots til faste tidspunkter og de bliver konverteret til en følge af talværdier X_0, X_1, X_2, \dots . Hvad der sker imellem tidspunkterne er tabt og en efterfølgende analyse kan blot antage at signalet er konstant mellem værdierne. Det er derfor vigtigt at indgangssignalet ikke varierer så hurtigt at det samplede signal ikke er en repræsentativ gengivelse af indgangssignalet. Hvis det sker vil der optræde fejlagtig information i spektret og det kaldes for *aliasing* som ikke har en tilsvarende dansk betegnelse.

Konverteringen er i figuren vist som en normering af indgangssignalet med referencespændingen x_{ref} til et interval fra 0 til 1, eller alternativt fra -1 til +1, og der omformes til en binær repræsentation med heltal i et område bestemt ved det antal bit B som konverteren har. Analog konvertering vil ofte benytte 16 bit mens en mikroprocessor ofte tilbyder 8 bit med 24 bit for den mere avancerede signalbehandling. Med $B = 16$ dækker konverteringen tal fra 0 til 65535 for positive signaler der svarer til området fra -32768 til 32767 for bipolare signaler. Illustrationens "round" funktion skal blot skitsere det faktum at en ADC kun kan aflevere binære heltal.

Omsætningen sker med hastigheden f_s der kaldes for *sampling frequency* og som kan ligge fra brøkdele af en hertz for langsomt varierende signaler som fx luftens temperatur til over 1 GHz for digital signalbehandling af radiosignaler. Høreapparater og telekommunikation benytter 8 kHz eller højere, krævende analoge systemer benytter 44 kHz for at opnå CD kvalitet og ved oscilloskoper er frekvensen over 250 MHz for at kunne analysere signaler indtil 100 MHz.

Når et signal opsamles (*sampling*) ændres det fra at være tidskontinuert til tidsdiskret hvor signalet udelukkende kendes til fastsatte tidspunkter. I det følgende vil den diskrete tid blive repræsenteret ved heltallet k som et index for tidspunktet i forløbet. Det interval der opsamles data over starter til tidspunktet $t = 0$ hvor index er $k = 0$ og sidste data er $k = K - 1$ hvor K er det totale antal af data. Data opsamles med tidsintervallet ΔT og det betyder at et signal beskrevet ved funktionen $h(t)$ kun kendes ved de diskrete tidspunkter $t = k\Delta T$.

$$h(t) = h(k\Delta T), \quad k = 0, 1, 2, \dots, K-1; \quad \Delta T = \frac{1}{f_s}$$

Den fundamentele frekvens f_1 er givet ved det krav at netop én periode af det oscillérende signal skal vare tiden T så grundfrekvensen (første harmoniske) bliver $f_1 = 1/T$ og de harmoniske er givet ved $2f_1, 3f_1$ etc. Signalets DC niveau, der er en konstant forskydning væk fra nul, beskrives ved en frekvens på nul. Den øvre grænse er givet ved den hastighed f_s hvormed data blev opsamlet, men det oprindelige signal kan ikke rekonstrueres hvis der er mindre end to værdier fra henholdsvis cosinus og sinus for ellers er der ikke både en positiv og en negativ komponent. Signalet må ikke indeholde signalkomponenter over halvdelen af samlingsfrekvensen (Nyquists og Shannons teorem). For ikke at havne i nulgennemgangen for svingninger skal den højeste frekvens Nf_1 være under halvdelen af samplingfrekvensen. En tommelfingerregel er at sample ved en frekvens på cirka fire gange den højeste frekvens der skal indgå i rekonstruktionen af signalet ($N \approx K/4$).

Højeste frekvenskomponent:	$f_{max} = Nf_1 < \frac{f_s}{2}$	f_1 Mindste frekvensopløsning
	$N < \frac{K}{2}$	f_s Opsamlingsfrekvens
Frekvens for dataopsamling:	$f_s = \frac{K}{T} = Kf_1$	T Tid for en periode af signalet
		K Totale antal data
		N Højeste (diskrete) frekvens

Hvis indgangssignalet indeholder frekvenskomponenter af betydende amplitude over den halve værdi af f_s så vil disse frekvenskomponenter foldes ned i frekvensbåndet fra nul til $f_s/2$ og kan ikke efterfølgende fjernes (*aliasing*) så det er vigtigt at der ikke er betydende rester af de højfrekvente signalkomponenter over $f_s/2$.

I praksis kan det være nødvendigt at inkludere et stejlt lavpasfilter i signalopsamlingen for at kunne fjerne frekvenskomponenter over halvdelen af samplingfrekvensen. Ved at designe efter $N < K/4$ kan der derved reserveres plads til at dette filter kan arbejde. Hvor meget signaler over den halve samplingfrekvens skal dæmpes afhænger af den krævede kvalitet af analysen. Ved 8 bit vil et niveau på 1 % svare til LSB for de frekvenser der havner over den halve samplingfrekvens. Det kræver en reduktion til -40 dB i løbet af én oktav fra $f_s/4$ til $f_s/2$ så filtret skal mindst være et sjette ordens Chebychev (se fjerde kapitel).

Koefficienterne til Fourier serien er givet ved de tidlige udtryk for a_n og b_n der herunder kun vises for a_n og for beregning af koefficienten b_n skal cosinus blot erstattes med sinus. Tidsvariablen t skal erstattes med den diskrete tid $t = k\Delta T/K$ og repetitionsfrekvensen med $f_1 = 1/T$. Integrationen fra 0 til T opdeles i små skridt fra 0 til $1/\Delta T$, fra $1/\Delta T$ til $2/\Delta T$ og så videre indtil $(K-1)/\Delta T$ til T . Integralets funktioner antages konstante og integralet af dt over intervallets længde giver blot $\Delta T = T/K$ så det resulterende udtryk bliver en sum med funktionsværdierne til de faste tidspunkter.

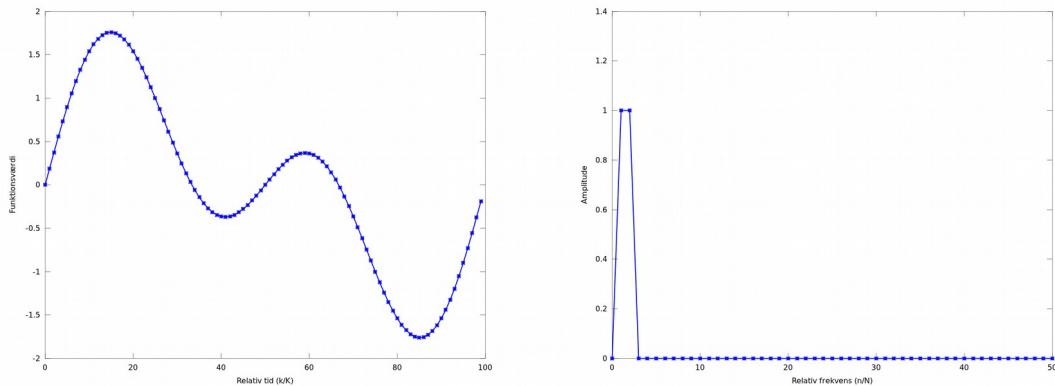
$$a_n = \frac{2}{T} \sum_{k=0}^{K-1} \int_{kT}^{(k+1)T} h(k\Delta T) \cos\left(2\pi \frac{n}{T} k \Delta T\right) dt \Rightarrow a_n = \frac{2}{K} \sum_{k=0}^{K-1} h(k\Delta T) \cos\left(2\pi \frac{n}{K} k\right)$$

Et systems respons udtrykt som amplitude og fase som funktion af frekvensen beregnes som vist herunder hvor det bedste resultat for fasen findes hvis arctan kan bringes til at dække over alle fire kvadranten i den komplekse plan. Den funktion hedder Atan2 i mange programmeringssprog.

$$|c_n| = \sqrt{a_n^2 + b_n^2} \quad \text{og} \quad \theta_n = \arctan\left(\frac{b_n}{a_n}\right)$$

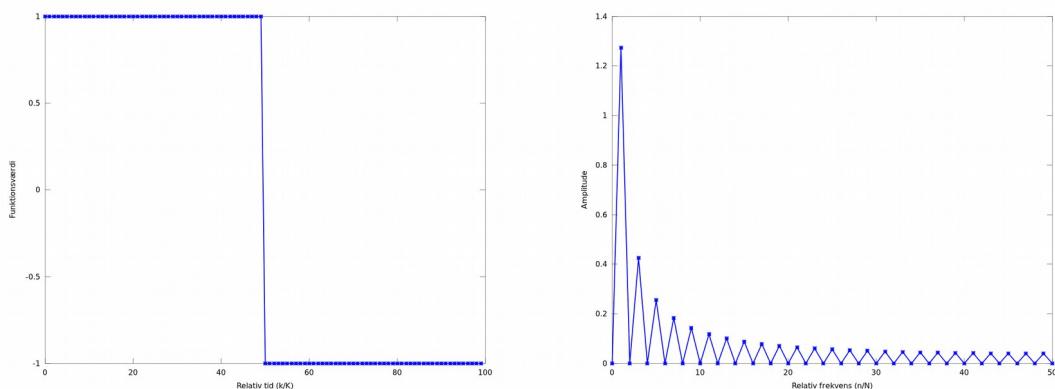
Algoritmen skal illustreres ved to signaler hvor den første består af to harmoniske svingninger på henholdsvis repetitionsfrekvensen f_1 og to gange denne frekvens med samme amplitude for hver svingning så det beregnede spektrum har derfor markering ved de relative frekvenser 1 og 2 samt nul for alle andre frekvenser. Disse frekvenspunkter kaldes undertiden for *bins*. Kurven til højre kaldes for signalets spektrum og viser to frekvenskomponenter. Det andet signal er en symmetrisk firksntsvingning hvor der kun er ulige harmoniske. Plottene kæder punkterne sammen med linjer, men det var mere korrekt at benytte lodrette linjer ved de mulige frekvenser i Fourier serien.

Analyse



Figur 80 – To svingninger med samme amplitude og frekvenserne f_1 og $2f_1$ giver kun uddata for de tilsvarende frekvenspunkter 1 og 2.

Den symmetriske firkantsvingning er analyseret tidligere i kapitlet og har amplituden af første harmoniske ved $4/\pi = 1,27$ og der er kun koefficienter for ulige frekvenser som det ses herunder. Niveauet ved 0 Hz repræsenterer DC niveauet i svingningen og er her lig med nul.

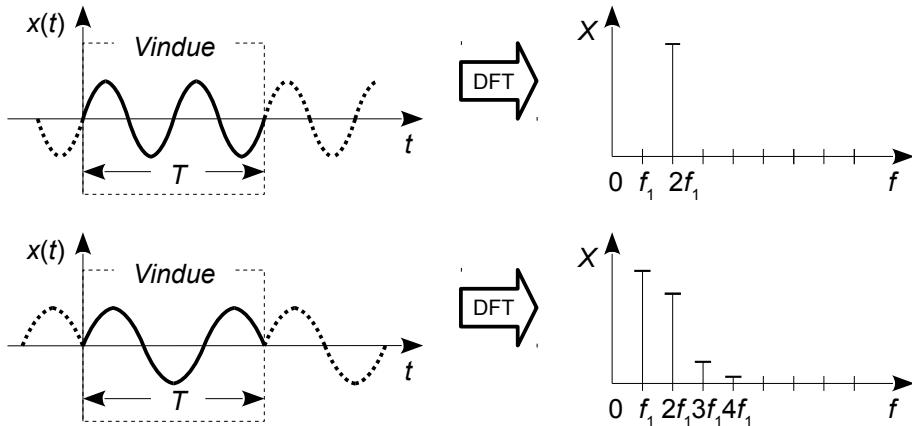


Figur 81 – En symmetrisk firkantsvingning har kun ulige harmoniske.

Den diskrete Fourier transformation (DFT) er populær for beregning af det harmoniske indhold i et reelt fysisk signal der ikke nødvendigvis er periodisk. Fourier serien forudsætter dog et periodisk signal så det opsamlede signal vil brutalt blive gjort periodisk. En dataopsamling starter og stopper på nogle givne tidspunkter og uden hensyn til det der analyseres. Det kan opfattes som om man kigger på et forløb gennem et vindue. Det der ligger til venstre for vinduets ramme kan ikke ses og tilsvarende for det der ligger til højre for vinduets ramme. Signalet $f(t)$ vil derfor begynde med en værdi der sandsynligvis er forskellig fra nul og tilsvarende gælder for afslutningen. Det er dertil meget lidt sandsynligt at start- og slutværdierne er ens. Signalet er ikke periodisk med T , men det forudsætter Fourier transformationen så resultatet bliver et signal med kraftigt niveau i de høje frekvenskomponenter som et forsøg på at eftergøre det opsamlede kurveforløb.

I figuren herunder består signalet øverst af to hele svingninger indenfor tidsvinduet og da signalet dermed er periodisk med T bliver resultatet gengivet korrekt med en markering ved den anden harmoniske. I nederste billede er frekvensen af signalet sænket til 1,5 gange repetitionsfrekvensen og algoritmen vil finde frem til amplituder af de harmoniske til repetitionsfrekvensen på $f_1 = 1/T$ i et forsøg på at kunne gengive dette forløb. Som illustrationen viser er det ikke en god tilnærmelse da der tilføres adskillige højere harmoniske til det oprindeligt harmoniske signal.

Analyse



Figur 82 – Signalet opfattes af den diskrete Fourier transformation som periodisk med $1/T$ og det vil derfor beskrives som harmoniske af denne frekvens uanset det oprindelige signals frekvens.

For at mildne skaderne kan man gange inddata med en funktion der har nulpunkter ved start og slut af tidsvinduet for tvinges signalet til at starte og slutte med værdien nul. Det beskrives normalt som brug af vinduesfunktioner (*window functions*) hvor en enkel er at gange data med en trigonometrisk funktion.

Hvis signalet ønskes analyseret med større opløsning kan man addere nulle hvilket på engelsk kaldes for *zero padding*. Idéen er at gøre perioden T længere så repetitionsfrekvensen bliver lavere, men det tilfører ikke ny information om det oprindelige signal.

For beregning i en mikroprocessor kan algoritmen implementeres med to for-løkker hvor den ydre dækker frekvensen og den indre dækker tiden (data). Herunder vises en MATLAB kode der er testet med freeware-programmet Octave. Frekvensen er givet ved en vektor for n og den refereres efterfølgende i for-løkkens for funktionsværdierne. De to løkker kan derved reduceret til en enkelt for-løkke fordi koeficienterne $a(n)$ og $b(n)$ er repræsenteret ved vektorer. Når der er adderet én til index skyldes det at MATLAB kræver et index der starter fra 1.

```

clear all
K=100; % Antal målepunkter.
N=K/2; % Antal frekvenspunkter.
t=0:K-1; % Tid.
% TESTFUNKTION.
%x(t+1)=sin(2*pi*t/K)+sin(4*pi*t/K); % 2 gange sinus.
%x(t+1)=t/K; % Rampe med DC.
x(1:K/2)=1; x(K/2+1:K)=-1; % Firkant uden DC.
% SPEKTRUM.
n=0:N; a(n+1)=0; b(n+1)=0; % Forbered frekvens.
for k=0:K-1
    a=a+x(k+1)*cos(2*pi*n*k/K);
    b=b+x(k+1)*sin(2*pi*n*k/K);
end
c=(2/K)*sqrt(a.^2+b.^2); % Amplitude.
c(1)=c(1)/2; % Korrigér DC.
figure(1), plot(t,x,'-')
xlabel("Relativ tid (k/K)", ylabel("Funktionsværdi")
print -dpng printfil1.png
figure(2), plot(n,c,'-')
xlabel("Relativ frekvens (n/N)", ylabel("Amplitude")
print -dpng printfil2.png

```

En DFT har frekvensopløsningen f_1 der også er den laveste analyserbare frekvens. For at gøre værktøjet anvendeligt bruges der relativt mange data ($K > 1000$) hvorved frekvensområdet er fra DC til $Kf_1/2$ med en opløsning på f_1 . Ved højere frekvenser end den halve sampling-frekvens gentages billedet blot nu spejlvendt og uden at tilføje mere information.

2.10 Laplace transformation

I elektricitetens barndom var et af de store problemer at forstå hvorfor telegrafstationens signaler blev forvansket undervejs til modtagerstationen. Forklaringen blev givet af Heaviside der havde opbygget sin egen matematik baseret på Fourier transformationen og ved hjælp af den fandt han frem til en løsning på problemet. Da isen var brudt blev Laplace transformationen genopdaget som det teoretiske fundament. Det er en matematik som Heaviside ikke kendte til.

Laplace transformationen tilhører en gruppe af matematiske operationer der er udviklet til løsning af differentialligninger og som finder anvendelse indenfor matematik og fysik. Som det vil blive klart i løbet af denne bog så er analyse af elektroniske netværk avanceret brug af matematiske metoder for opstilling og løsning af differentialligninger så den matematiske fortid er ikke bare en historisk detalje, men derimod kernen i hvad vi skal beskæftige os med fremadrettet.

$$F(s) = \int_0^{\infty} f(t) \exp(-st) dt \quad \text{hvor} \quad f(t) = 0, \quad t < 0$$

Hvis du sammenligner med Fourier transformationen fra side 81 ses at integralet starter fra $t = 0$ og derfor ikke omfatter al den tid universet har eksisteret. Det betyder at funktionen $f(t)$ er nul inden starttidspunktet så der må være et eller andet som sætter gang i kredsløbet. Det er en af de helt store styrker ved Laplace transformationen at den giver mulighed for at studere hvad der sker i et kredsløb efter tilslutning af en energikilde eller omskiftning af en kontakt. Dertil kommer at Laplace transformationen også kan vise hvad der sker i kredsløbet ved brug af en sinus-generator og det er helt i tråd med $j\omega$ metoden. Det er med andre ord et særdeles effektivt værktøj.

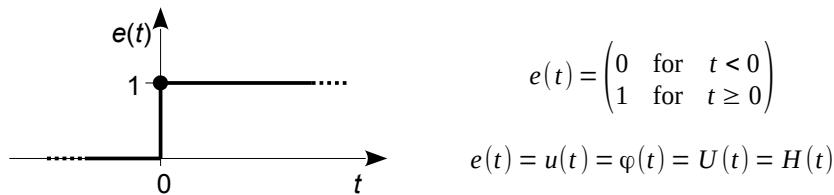
Laplace transformationen omsætter en tidsafhængig funktion $f(t)$ til en funktion $F(s)$ hvor s er en kompleks funktion i et plan givet af dæmpningen σ og frekvensen gennem ω . At s indeholder $j\omega$ giver et fingerpeg om at hele værktøjet fra de forrige afsnit om impedans kan benyttes og det er klart tilfældet. Det vil dog være forkert blot at beskrive s som en frekvensvariabel for matematisk set er det en *operator* og analysen er ikke begrænset til noget oscillerende.

$$s = \sigma + j\omega, \quad \omega = 2\pi f$$

I den matematiske formulering transformeres kredsløbet fra tidsdomænet til s -domænet. Det er normalt ledsaget af definitionen af en lang række formler der kan slå selv en matematisk stærk person til jorden så jeg vil følge en anden vej. Først analyseres følgerne af indkobling af en spændingskilde på et kredsløb og dernæst et kredsløb hvor en kondensator aflades.

2.10.1 Enhedsstep

Det første eksempel på anvendelse af Laplace transformationen benytter det såkaldte enhedsstep (*unit-step function*) der også kaldes for stepfunktion, springfunktion eller Heavisides funktion. Den har én af to diskrete værdier. For et negativt argument er værdien nul og for nul eller positiv værdi af argumentet er funktionen én. Kært barn har mange navne og jeg vil her benytte symbolet $e(t)$.

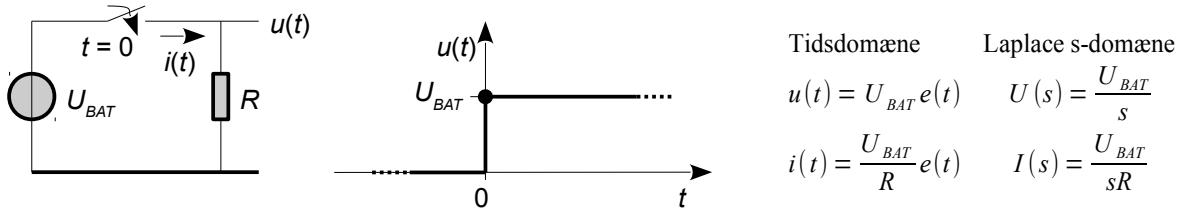


Figur 83 – Enhedsstrinfunktionen repræsenterer en kontakt, der sluttes til et givet tidspunkt.

Funktionen kan bruges til at beskrive en kilde der indkobles til tidspunktet $t = 0$ og derefter kan man så studere hvad der sker i kredsløbet. Det skal illustreres ved et par eksempler.

2.10.2 Indkobling af kilde

I eksemplet herunder benyttes en kontakt for at indkoble spændingskilden U_{BAT} til en modstand R til tidspunktet $t = 0$. Det er åbenlyst at strømmen $i(t)$ er nul inden kontakten aktiveres og at strømmen momentant springer op på værdien U_{BAT}/R for $t > 0$.



Figur 84 – Momentan indkobling af et batteri på et kredsløb.

Funktionen $i(t)$ er nul inden integrationen starter så alt hvad der behøves er at erstatte $f(t)$ i Laplace transformationen med U_{BAT}/R og integrere over tid. Argumentet er komplekst $(-st)$, men det følger de sædvanlige regneregler for integration.

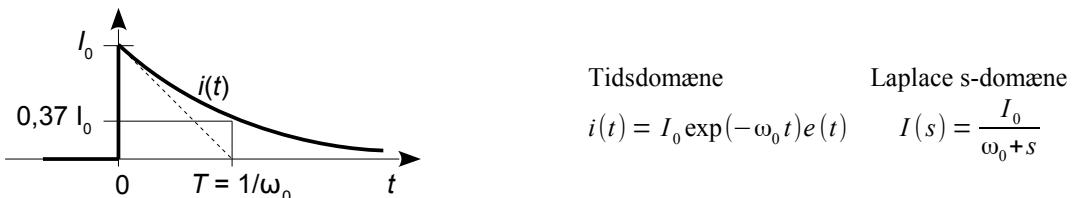
$$I(s) = \int_0^\infty \frac{U_{BAT}}{R} \exp(-st) dt = -\frac{U_{BAT}}{R} \frac{1}{s} [\exp(-st)]_0^\infty = -\frac{U_{BAT}}{R} \frac{1}{s} [0 - 1] = \frac{U_{BAT}}{sR}$$

Udtrykket kan virke mystisk, men Laplace transformationen leder til nogle lidt pudsige udtryk. Lidt fornuft er der dog i det for s repræsenterer en frekvens gennem variablen $\omega = 2\pi f$ og værdien af $I(s)$ vokser over alle grænser for s gående imod nul der svarer til "lang tid". Det betyder at der afsættes mere og mere energi i modstanden desto lavere frekvens der benyttes i en analyse, men når en effektforsyning er forbundet til et system så vokser den afsatte energi jo med den tid der er gået og en lav værdi af s svarer netop til at der er gået lang tid. En høj værdi af s svarer til noget der ikke tager ret lang tid og her er energiafsættelsen da også lav.

Bemærk at udgangssignalet $I(s)$ ikke er en strøm for enheden er ampere gange sekund (As). Det betyder at Laplace transformationen påvirker enheden af det man arbejder med så en strøm vil ikke bare transformeres til strøm, men til en *abstraktion* fra strøm og tilsvarende gælder for en spænding og hvad man ellers arbejder med.

2.10.3 Eksponentielt aftagende funktion

En kondensator der aflades gennem en modstand giver en eksponentielt aftagende strøm og det tilhørende udtryk er et af de vigtigste i elektronikken da det dukker op alle steder hvor der er en begrænsning i et frekvensområde. Sammenlign med resultatet af Fourier transformationen af den samme funktion på side 84.



Figur 85 – En eksponentielt aftagende funktion repræsenterer flertallet af fysiske situationer hvor et eller andet aftager som fx strømmen i en kondensator der lades op gennem en modstand.

Laplace transformationen af $i(t)$ beregnes ved at indsætte udtrykket i definitionen og det udnyttes at enhedsstepfunktionen er én i hele integrationens område fra nul og op.

$$I(s) = \int_0^\infty I_0 \exp(-\omega_0 t) \exp(-st) dt = I_0 \int_0^\infty \exp(-(\omega_0 + s)t) dt = \frac{I_0}{\omega_0 + s}$$

2.10.4 Definitioner

Et par definitioner er nødvendige for de efterfølgende analyser og de præsenteres uden bevis.

Den tidsafledte af den variable $f(t)$ transformeres til s gange det Laplace transformerede af $f(t)$ der som udgangspunkt er $sF(s)$. I udtrykket herunder er $f(0)$ funktionsværdien til det tidspunkt hvor analysen starter ($t = 0$). Det er startværdien (*initial value*). Ved brug til kredsløbsanalyse sættes startværdien ofte til nul med mindre der er behov for at benytte en specifik startværdi.

$$\frac{df(t)}{dt} \Leftrightarrow sF(s) - f(0)$$

Spændingen over en spole er givet ved dens selvinduktionen L gange med den afledte af strømmen di/dt . Spændingen $u_L(t)$ transformeres til $U_L(s)$ og strømmen til $I(s)$. Selvinduktionen L er konstant så den deltager ikke i transformationen, men forbliver en fast faktor. Sidste led med $i(0)$ er den værdi strømmen i spolen har til starttidspunktet $i(0)$. Denne ligning er fundamentet i en DC-DC konverter hvor $i(0)$ er middelværdien af DC strømmen i spolen.

$$u_L = L \frac{di}{dt} \Leftrightarrow U_L = sLI - Li(0)$$

Integralet af den variable $f(t)$ transformeres til $1/s$ gange med det Laplace transformerede af $f(t)$ plus $1/s$ gange værdien til tidspunktet $t = 0$.

$$\int_0^t f(\tau) d\tau \Leftrightarrow \frac{F(s)}{s} + \frac{f(0)}{s}$$

Spændingen over en kondensator er givet ved integralet af strømmen som den blev opladt med. Spændingen transformeres til $U(s)$ og strømmen til $I(s)$. Kapaciteten C er konstant og deltager ikke i processen. Sidste led med $u(0)$ er spændingen på kondensatoren til tidspunktet $t = 0$ og delt med s bliver spændingen indkoblet til tidspunktet $t = 0$.

$$u_C = \frac{1}{C} \int idt + u(0) \Leftrightarrow U_C = \frac{1}{sC} I + \frac{u(0)}{s}$$

Impedansen af modstand, kondensator og spole defineres som følger. Hvis realdelen af s sættes til nul ($\sigma = 0$) bliver $s = j\omega$ og de resulterende udtryk er nu identiske med resultatet fra $j\omega$ metoden fra de indledende afsnit.

$$Z_R = R \quad Z_C = \frac{1}{sC} \quad Z_L = sL$$

Det er muligt at beregne hvilken funktionsværdi tidssvaret $f(t)$ vil have i tidsdomænet ud fra et givet udtryk $F(s)$ i s -domænet. Ved startværdidisætningen (*initial value theorem*) findes funktionsværdien $f(0)$ i startøjeblikket til $t = 0$ ved at gange s på $F(s)$ og lade s gå imod uendelig. På tilsvarende vis benyttes slutværdidisætningen (*final value theorem*) til at finde funktionsværdien $f(\infty)$ når der er gået så lang tid at kredsløbet er faldet til ro blot sætter vi her $s = 0$. Forudsætningen er at polerne i $F(s)$ ligger i venstre halvplan så reglen gælder ikke for oscillerende funktioner.

$$\text{Startværdidisætningen: } \lim_{t \rightarrow 0} \{f(t)\} = \lim_{s \rightarrow \infty} \{sF(s)\}$$

$$\text{Slutværdidisætningen: } \lim_{t \rightarrow \infty} \{f(t)\} = \lim_{s \rightarrow 0} \{sF(s)\}$$

Der er mange andre definitioner indenfor Laplace transformationen, men det angivne vil vise sig at være tilstrækkeligt langt hen i en kredsløbsanalyse.

2.10.5 Tabel over Laplace transformationen

A og B samt a og b er konstanter der kan være reelle eller komplekse. Funktionerne $f(t)$, $g(t)$ og $h(t)$ er fysisk realisable. $F(s)$, $G(s)$ og $H(s)$ er de tilsvarende Laplace transformerede udtryk og $e(t)$ er et enhedsstep der er nul for $t < 0$ og én for $t \geq 0$.

	Funktion i tidsdomænet	Funktion i s-domænet	Kommentar
1	$h(t) = f(t) \cdot e(t)$	$H(s) = \int_0^{\infty} f(t) \exp(-st) dt$	Definition af Laplace transformationen.
2	$h(t) = [A f(t) + B g(t)] \cdot e(t)$	$H(s) = A F(s) + B G(s)$	Transformation af hvert led for sig.
3	$h(t) = A \frac{df(t)}{dt} \cdot e(t)$	$H(s) = s A F(s) - A f(0)$	Differentiation med startværdi $f(0)$.
4	$h(t) = A \int f(t) dt \cdot e(t)$	$H(s) = A \frac{f(s)}{s} + A \frac{f(0)}{s}$	Integration med startværdi $f(0)$.
5	$h(t) = f(t-T) \cdot e(t)$	$H(s) = F(s) \exp(-sT)$	Tidsforskydning.
6	$h(t) = f(t) \exp(-at) \cdot e(t)$	$H(s) = F(s+a)$	Dæmpet funktion.
7	$h(t) = \delta(t)$	$H(s) = 1$	Diracpuls til $t = 0$.
8	$h(t) = A \cdot e(t)$	$H(s) = A \frac{1}{s}$	Enhedsstep til $t = 0$ med værdien A.
9	$h(t) = At \cdot e(t)$	$H(s) = A \frac{1}{s^2}$	Indkobling af rampe til $t = 0$.
10	$h(t) = A \exp(-at) \cdot e(t)$	$H(s) = A \frac{1}{s+a}$	Eksponentielt aftagende fx RC-led med $a = 1/CR$.
11	$h(t) = A[1 - \exp(-at)] \cdot e(t)$	$H(s) = A \frac{a}{s(s+a)}$	Eksponentielt stigende fx RC-led med $a = 1/CR$.
12	$h(t) = At \exp(-at) \cdot e(t)$	$H(s) = A \frac{1}{(s+a)^2}$	Anden-ordens lavpas med to ens poler.
13	$h(t) = \frac{A}{a-b} [a \exp(at) - b \exp(bt)] \cdot e(t)$	$H(s) = A \frac{s}{(s+a)(s+b)}$	Anden-ordens båndpas med to forskellige poler.
14	$h(t) = \frac{A}{b-a} [\exp(-at) - \exp(-bt)] \cdot e(t)$	$H(s) = A \frac{1}{(s+a)(s+b)}$	Anden-ordens lavpas med to poler.
15	$h(t) = A \left[\frac{1}{ba} + \frac{\exp(-at)}{a(a-b)} - \frac{\exp(-bt)}{b(b-a)} \right] \cdot e(t)$	$H(s) = A \frac{1}{s(s+a)(s+b)}$	Integralet af et anden-ordens lavpas med to poler.
16	$h(t) = A \sin(\omega_0 t) \cdot e(t)$	$H(s) = A \frac{\omega_0}{s^2 + \omega_0^2}$	Sinus, der starter til tidspunktet $t = 0$.
17	$h(t) = A \cos(\omega_0 t) \cdot e(t)$	$H(s) = A \frac{s}{s^2 + \omega_0^2}$	Cosinus, der starter til tidspunktet $t = 0$.
18	$h(t) = A \sin(\omega_d t) \exp(-at) \cdot e(t)$	$H(s) = A \frac{\omega_d}{(s+a)^2 + \omega_d^2}$	Dæmpet sinus $0 < d < 1$.

19	$h(t) = A \sin(\sqrt{1-d^2} \omega_0 t) \exp(-d \omega_0 t) \cdot e(t)$	$H(s) = A \frac{\sqrt{1-d^2} \omega_0}{s^2 + 2d\omega_0 s + \omega_0^2}$	Dæmpet sinus $0 < d < 1$.
10	$h(t) = A \cos(\omega_d t) \exp(-at) \cdot e(t)$	$H(s) = A \frac{s+a}{(s+a)^2 + \omega_d^2}$	Dæmpet cosinus $0 < d < 1$.
21	$h(t) = A \cos(\sqrt{1-d^2} \omega_0 t) \exp(-d \omega_0 t) \cdot e(t)$	$H(s) = A \frac{s+\sqrt{1-d^2} \omega_0}{s^2 + 2d\omega_0 s + \omega_0^2}$	Dæmpet cosinus $0 < d < 1$.

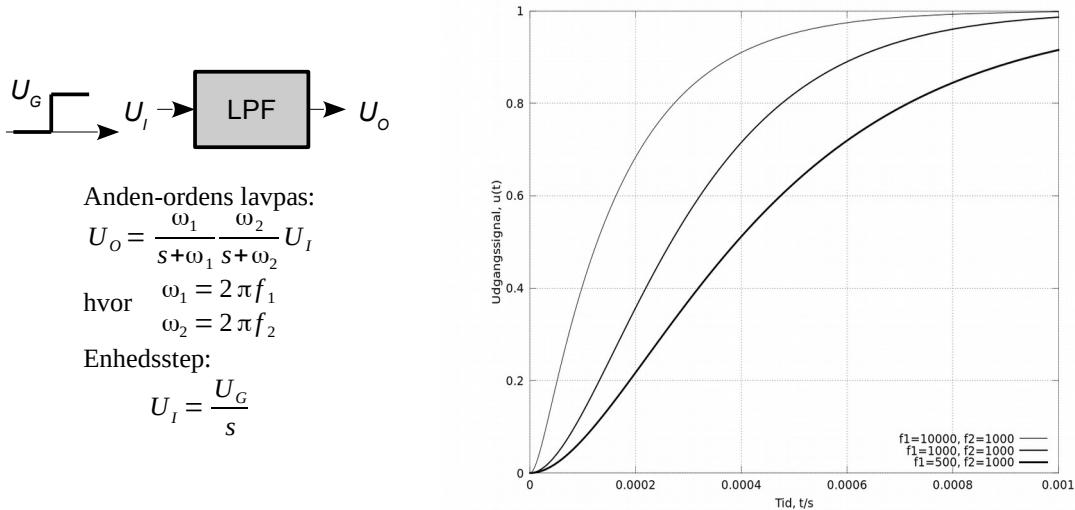
Se også referencerne RW og NA samt internettet.

2.10.6 Transformation tilbage til tidsdomænet

Det kan være nødvendigt at transformere tilbage til tidsdomænet hvis man vil vide hvor hurtigt et elektronisk kredsløb falder på plads efter en pludselig påvirkning. Der er ganske vist defineret en invers Laplace transformation, men det er almindeligvis lettest at benytte tabellen.

2.10.7 Steprespons

Et lavpasfilter tillader DC at passere, men spørgsmålet er hvor hurtigt spændingen på udgangen vil vokse mod slutværdien efter et pludseligt skift på filtrets indgang. Eksemplet herunder illustrerer hvordan det beregnes for et anden-ordens filter med reelle poler (fx to RC-led) og signalet ses at falde på plads efter tiden $T \approx 1/f_1$ hvor f_1 er den laveste grænsefrekvens for de to filter.



Figur 86 – Indsvingningsforløbet for et anden-ordens lavpasfilter med to reelle poler. Øverste kurve har poler ved 10 kHz og 1 kHz og ligger tæt på billedet af et første-ordens filter ved 1 kHz.

Indgangssignalet er $U_I = U_G/s$ og ved indsættelse ses at det er muligt at bringe udtrykket på en form som tabellen benytter.

$$U_O = \frac{\omega_1}{s + \omega_1} \frac{\omega_2}{s + \omega_2} \frac{U_G}{s} = \omega_1 \omega_2 \frac{1}{s(s + \omega_1)(s + \omega_2)} U_G$$

I tidsdomænet bliver resultatet som følger hvor funktionen består af et ettal og desuden to eksponentialfunktioner så funktionen angiver noget der roligt glider på plads. Funktionen starter fra nul til starttidspunktet ($t = 0$) da de to eksponentialfunktioner hver er én.

$$u_O(t) = [1 + \frac{\omega_2}{\omega_1 - \omega_2} \exp(-\omega_1 t) + \frac{\omega_1}{\omega_2 - \omega_1} \exp(-\omega_2 t)] \cdot U_G e(t)$$

Analyse

At tidssvaret fra filtret starter fra nul kan ses ved at bruge startværdidisætningen fra side 91.

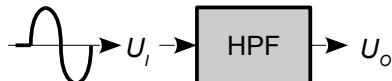
$$u_o(0) = \lim_{s \rightarrow \infty} \left\{ s \frac{\omega_1}{s+\omega_1} \frac{\omega_2}{s+\omega_2} \frac{U_G}{s} \right\} = \lim_{s \rightarrow \infty} \left\{ \frac{\omega_1}{s+\omega_1} \frac{\omega_2}{s+\omega_2} U_G \right\} = 0$$

Tilsvarende kan det ses fra slutværditeoremet at funktionen efter tilstrækkelig lang tid vil være nået op på sluttværdien U_G .

$$u_o(0) = \lim_{s \rightarrow 0} \left\{ s \frac{\omega_1}{s+\omega_1} \frac{\omega_2}{s+\omega_2} \frac{U_G}{s} \right\} = \lim_{s \rightarrow 0} \left\{ \frac{\omega_1}{s+\omega_1} \frac{\omega_2}{s+\omega_2} U_G \right\} = U_G$$

2.10.8 Indsvingning

En anden vigtig situation er en pludselig tilslutning af et oscillérende signal på et højpasfilter hvor det er vigtigt at vide hvor lang tid der går inden udgangen er på plads og svinger omkring nul. Det vil sige hvor lang tid der går før chokket fra den pludselige ændring er døet ud. Resultatet er i tråd med det tidligere fundne hvor svingningen falder på plads indenfor 0,2 % efter $T \approx 1/f_0$ hvor f_0 er filtrets grænsefrekvens (-3 dB). Analysen benytter et første-ordens filter og ved et filter af højere orden med komplekst konjugerede poler vil det tage længere tid at falde på plads.



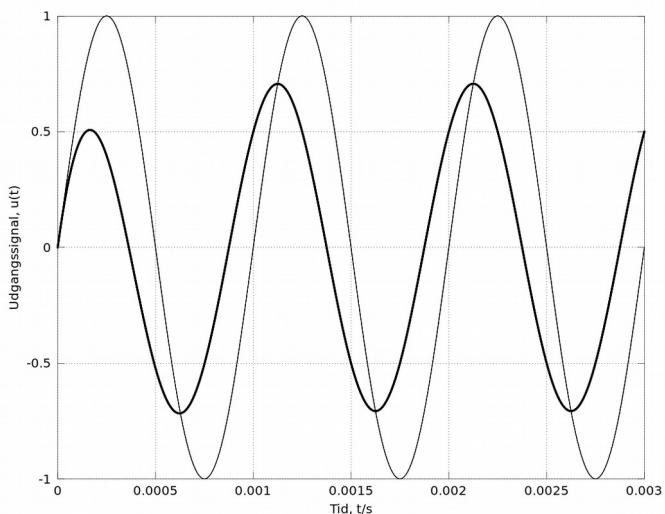
Første ordens højpas:

$$U_O = \frac{s}{\omega_0 + s} U_I$$

Tilkoblet sinus:

$$u(t) = \begin{pmatrix} U_M \sin(\omega_{osc} t) \cdot e(t) \\ \omega_{osc} = 2\pi f_{osc} \end{pmatrix}$$

$$U(s) = U_M \frac{\omega_{osc}}{s^2 + \omega_{osc}^2}$$



Figur 87 – Indsvingning (tyk streg) ved påtrykning af en sinus (tynd streg) på et højpasfilter. Svingningen har frekvensen 1 kHz der også er højpasfiltrets -3 dB grænsefrekvens.

Analysen er forholdsvis kompliceret og tabellen kan ikke bruges direkte så fremgangsmåden er at bringe udtrykket på en form hvor tabellen bliver anvendelig.

En harmonisk svingning (her en sinus) indkobles pludseligt på et første-ordens højpasfilter. Resultatet vil bestå af en transient fra den pludselige tilkobling og dertil et stationært udtryk. Først Laplace transformere indgangssignalet med amplituden U_M og frekvensen f_{osc} og udtrykket ganges på overføringsfunktionen. Resultatet står ikke i tabellen så det er nødvendigt at splitte overføringsfunktionen op og det tager udgangspunkt i nævnerens to polynomier. Det giver en enkelt konstant A for første-ordens ledet og et polynomium med to konstanter $Bs + C$ for andet-ordens ledet. Opgaven er nu at finde de tre konstanter A , B og C .

$$U(s) = U_M \frac{s \omega_{osc}}{(s + \omega_0)(s^2 + \omega_{osc}^2)} = U_M \left[\frac{A}{s + \omega_0} + \frac{Bs + C}{s^2 + \omega_{osc}^2} \right] \text{ og } A, B, C = \text{konstanter}$$

For at komme videre stilles de to udtryk i den firkantede parentes på samme nævner og ved en

sammenligning med det oprindelige udtryk findes tre ligninger for bestemmelse af konstanterne.

$$s\omega_{osc} = (A+B)s^2 + (B\omega_0 + C)s + A\omega_{osc}^2 + C\omega_0 \Rightarrow \begin{cases} A+B=0 \\ B\omega_0 + C = \omega_{osc} \\ A\omega_{osc}^2 + C\omega_0 = 0 \end{cases}$$

Ligningssystemet kan løses ved at substituere $-B$ ind i ligning tre i stedet for A og C kan nu fjernes ved at kombinere ligning to og tre. Herved kan B bestemmes og dermed A og C .

$$\begin{cases} A = -B \\ C = -B\omega_0 + \omega_{osc} \\ C\omega_0 = B\omega_{osc}^2 \end{cases} \Rightarrow B = \frac{\omega_{osc}\omega_0}{\omega_{osc}^2 + \omega_0^2} \Rightarrow C = \frac{\omega_{osc}^3}{\omega_{osc}^2 + \omega_0^2}$$

Det oprindelige udtryk kan nu skrives ved tre led der kan transformeres hver for sig.

$$U(s) = U_M \left[\frac{A}{s+\omega_0} + \frac{Bs}{s^2 + \omega_{osc}^2} + \frac{C}{\omega_{osc}} \frac{\omega_{osc}}{s^2 + \omega_{osc}^2} \right]$$

Tidssvaret bliver en sum af en aftagende eksponentialfunktion med koeficienten A , som svarer til transienten fra den pludselige tilkobling af den harmoniske svingning, og dertil to oscillerende led med den samme frekvens (cosinus og sinus med B og C), svarende til en dæmpning og et fasedrej i forhold til den sinus der var påtrykt indgangen.

$$u(t) = U_M \left[A \exp(-\omega_0 t) + B \cos(\omega_{osc} t) + \frac{C}{\omega_{osc}} \sin(\omega_{osc} t) \right] e(t)$$

Ved den samme frekvens for oscillator og højpasfilter ($f_{osc} = f_0$) bliver $A = -0,5$ og $B = 0,5$ så transienten starter ved $-0,5$ og klinger ud med en tidskonstant på $\tau = 1/\omega_0$ og er effektivt set væk efter cirka 1 ms (se forrige eksempel). For $f_{osc} = 1$ kHz findes $C/\omega_{osc} = 0,5$ så de to led med cosinus og sinus kan kombineres til $0,71 \cdot \sin(\omega_{osc} t + 45^\circ)$.

2.10.9 Frekvensdomænet

Det er almindeligt at repræsentere overføringsfunktionen som en funktion af frekvensen og ikke som en funktion af den komplekse variable fra Laplace transformationen. Det betyder at der må være en relation mellem $s = \sigma + j\omega$ og frekvensen f og det er der gennem ω . Laplace variablen s er et punkt i den komplekse plan hvorimod frekvensen er et punkt langs med en akse så i s -planen er frekvensen et punkt langs med den imaginære akse.

$$s = \sigma + j\omega = \sigma + j2\pi f \Rightarrow f = \frac{\omega}{2\pi} = \left| \frac{s}{2\pi} \right|_{\sigma=0}$$

Det er denne relation der knytter forbindelsen til $j\omega$ -metoden fra den indledende del af kapitlet. Differentialligningerne blev her løst for harmoniske svingninger med de komplekse eksponentialefunktioner. Ved Laplace transformationen er løsningsmængden udvidet til enhver funktion.

For et lavpasfilter skal amplituden opskrives som funktion af frekvensen.

$$H(s) = \frac{\omega_0}{s + \omega_0} \Rightarrow H(f) = \frac{\omega_0}{j2\pi f + \omega_0} = \frac{f_0}{jf + f_0}$$

Amplituden $M(f)$ er absolutværdien af funktionen $H(f)$ og M står for *magnitude*. Den kan omregnes til følgende handy relation der gælder for ethvert lavpasfilter af første-orden.

$$M(f) = |H(f)| = \left| \frac{\omega_0}{j2\pi f + \omega_0} \right| = \left| \frac{f_0}{jf + f_0} \right| = \frac{f_0}{\sqrt{f^2 + f_0^2}} = \frac{1}{\sqrt{1 + (f/f_0)^2}}$$

2.11 Overføringsfunktion

Laplace transformationen vil i denne bog primært blive anvendt til udledning af det der kaldes for et kredsløbs overføringsfunktion (*transfer function*). Den beskriver udgangssignalet U_o som følge af et indgangssignal U_I ved at gange indgangssignalet med overføringsfunktionen H . Signalet på indgangen er ofte en harmonisk svingning, men overføringsfunktionen kan benyttes med ethvert periodisk signal, der jo kan beskrives ved en Fourier serie, såvel som med ikke-periodiske signaler, der kan opbygges med kontinuerte funktioner, enhedsstep og Dirac delta pulser.

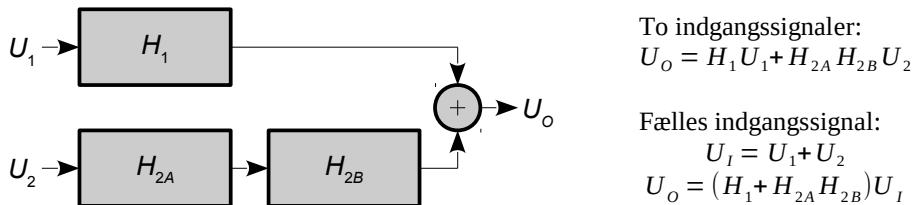
$$\begin{array}{c}
 U(s) \rightarrow \boxed{H(s)} \rightarrow U_o(s) \\
 \begin{aligned}
 U_o &= H U_I \\
 U_o &= |U_o| \exp(j\theta_o) \exp(j\omega t) \\
 U_I &= |U_I| \exp(j\theta_I) \exp(j\omega t) \\
 H &= \frac{U_o}{U_I} = \left| \frac{U_o}{U_I} \right| \exp(j(\theta_o - \theta_I))
 \end{aligned}
 \end{array}$$

Figur 88 – Et systems overføringsfunktion beskriver hvordan udgangen reagerer på et signal påført indgangen. Alt er komplekse funktioner af s.

Overføringsfunktionen kan beskrives ved de to signalers komplekse amplituder og er funktion af frekvensen. Selve oscillationen beskrives ved $\exp(j\omega)$ og kan forkortes væk som vist ovenfor. Det er muligt at måle overføringsfunktionen i laboratoriet ved amplituden og fasen af U_I og U_o og det er nærliggende at definere tonegeneratorens signal som referencen med U_I som 0 dB og $\theta_I = 0^\circ$. Andre muligheder er brug af et enhedsstep eller en Dirac puls. Endelig kan overføringsfunktionen findes analytisk ved brug af kredsløbsligninger baseret på Ohms og Kirchhoffs love.

2.11.1 Sammensatte kredsløb

Overføringsfunktionen kan umiddelbart udvides til at beskrive større kredsløb med flere ind- og udgange som det eksemplificeres herunder for et kredsløb med tre blokke og to indgangssignaler U_1 og U_2 der føres til udgangen U_o langs med hver sin rute.



Figur 89 – Kredsløbet kan bestå af flere delkredse. Ved serieforbundne delkredsløb er den samlede funktion givet ved produktet af overføringsfunktionerne og ved parallellforbundne delkredsløb er funktionen givet ved addition. Der kan være ét eller flere signaler ind til systemer og udgangen kan tilsvarende deles op over flere individuelle udgange.

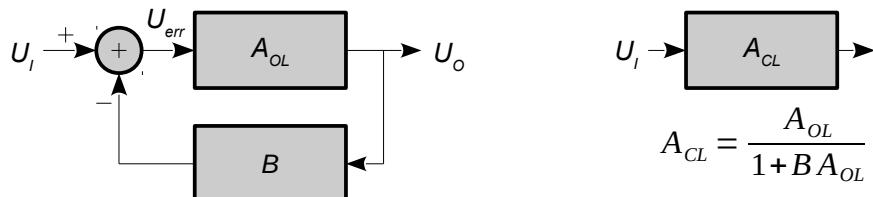
I eksemplet ledes det ene signal U_1 via overføringsfunktionen H_1 til udgangen U_o hvor det adderes med det andet signal U_2 via de serieforbundne overføringsfunktioner H_{2A} og H_{2B} . Hvis signalet på indgangen U_I er fælles for de to veje ($U_1 = U_2 = U_I$) vil den resulterende overføringsfunktion blive summen af de to funktionsudtryk⁵⁴.

Denne fremgangsmåde er meget anvendelig ved både analog og digital signalbehandling idet et signals vej gennem et system ofte kan deles op i flere veje som efter manipulation med amplitude og fase igen kombineres til et fælles udgangssignal. Et eksempel inkluderer lydeffekterne chorus og phasing hvor indgangssignalet blandes med en tidsforsinket version af sig selv. Det centrale er at tidsforsinkelsen varierer og det giver et varierende system af konstruktive og destruktive interferenser som øret opfatter som den karakteristiske lyd.

54 Det kan eksempelvis repræsentere det resulterende signal i lyttepositionen fra en højttaler med et direkte signal H_1 og et reflekteret signal H_{2A} hvor dæmpning og fasedrejning fra refleksionen i gulvet er givet ved H_{2B} .

2.11.2 Tilbagekobling

Funktionen af elektroniske kredsløb baseres i vid udstrækning på tilbagekobling (*feed back*) som benyttes for at sikre at kredsløbet opfører sig som forventet og her er operationsforstærkeren et godt eksempel. Princippet er at benytte et overskud af forstærkning (A_{OL}) til at reducere fejlene ved kredsløbet. Det tilbagekoblede signal kan være hele udgangssignalet ($B = 1$) som svarer til en spændingsfølger eller en brøkdel ($0 < B < 1$) der normalt realiseres ved en spændingsdeler, men tilbagekoblingen kan også indeholde komplekse komponenter såsom filtre og tidsforsinkelse.



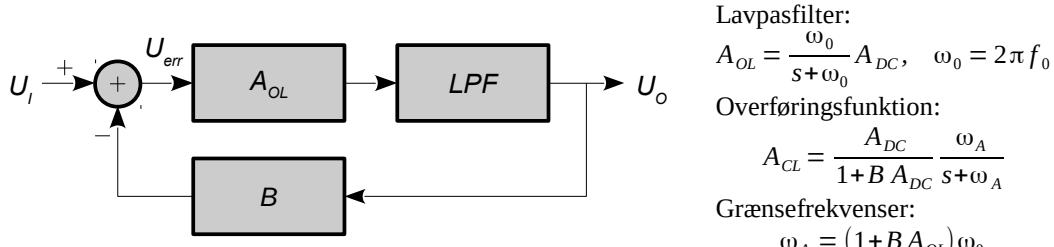
Figur 90 – Tilbagekobling af et kredsløb med overføringsfunktionen A_{OL} benytter brøkdelen B af udgangssignalet som fratrækkes indgangssignalet. Resultatet er en ny overføringsfunktion A_{CL} .

Der dannes en differens U_{err} mellem indgangssignalet U_I og brøkdelen B af udgangssignalet U_O og differensen forstærkes op med åben-sløjfe forstærkningen A_{OL} . Leddet BA_{OL} i nævneren er tilbagekoblingsfaktoren (*feed-back factor*) som normalt har en meget høj værdi. Hvis A_{OL} er så høj at den kan betragtes som uendelig vil forstærkningen blive $A_{CL} = 1/B$.

$$\begin{aligned} U_O &= A_{OL} U_{err} \\ U_{err} &= U_I - BU_O \end{aligned} \Rightarrow U_O = A_{CL} U_I \quad \text{hvor} \quad A_{CL} = \frac{A_{OL}}{1 + BA_{OL}} \xrightarrow{A_{OL} \rightarrow \infty} \frac{1}{B}$$

2.11.3 Båndbredde

Ved stigende frekvens vil forstærkningen nødvendigvis aftage for ethvert kredsløb har en endelig båndbredde og en model er at beskrive det ved en DC forstærkning A_{DC} og et lavpasfilter med en grænsefrekvens på f_0 (-3 dB) hvorover forstærkningen falder med -6 dB/oktav.



Med DC forstærkningen $A_{DC} = 10^5$ og tilbagekoblingen $B = 0,01$ bliver tilbagekoblingen på $BA_{DC} = 10^3$ så grænsefrekvensen i kredsløbet på fx $f_0 = 1$ kHz flyttes op til 1 MHz mens forstærkningen på 99,9 ved lave frekvenser kun er en promille fra de forventede $1/B = 100$.

Tilbagekoblingen giver en forøgelse af båndbredden på $1 + BA_{DC}$ mod en tilsvarende reduktion af forstærkningen til $1/B$. Det væsentlige er her at åben-sløjfe forstærkningen A_{OL} er reduceret til en ikke-dominerende parameter så funktionen af kredsløbet kan i det væsentlige styres af B gennem eksterne komponenter der typisk er et par modstande hvor tolerancen kan specificeres.

Overføringsfunktionen A_{OL} kan repræsentere en operationsforstærkers DC forstærkning A_{DC} med den dominante pol som et første-ordens lavpasfilter. Det kan også være et elektro-mekanisk system hvor forstærkningen skyldes et reguleringssystem med motorsens træghed repræsenteret gennem et lavpas-filter. Filtrert kan have flere poler hvorefter tilbagekoblingen udfordrer stabiliteten af systemet så det er ikke givet at man altid kan tvinge et system til at makke ret blot ved en tilbagekobling. Det emne diskuteres herunder.

2.11.4 Stabilitet

En operationsforstærker kan ofte beskrives ved en enkelt pol, men ved tilbagekobling af et større system er der ofte flere poler i overføringsfunktionen. Stabiliteten af det tilbagekoblede system er en funktion af hvor kraftig tilbagekoblingen er og kan lide under en stor værdi af BA_{DC} . Det kan udtrykkes ved en dæmpningsfaktor d der bør være nær ved én for at opnå stabilitet.

Overføringsfunktionen for det ikke-tilbagekoblede system A_{OL} beskrives ved to poler ω_1 og ω_2 samt forstærkningen A_{DC} ved lave frekvenser. Overføringsfunktionen for dette system bliver A_{CL} ved tilbagekobling med brøkdelen B af udgangssignalet.

$$\begin{aligned} A_{OL} &= \frac{\omega_1 \omega_2}{(s+\omega_1)(s+\omega_2)} A_{DC} \\ A_{CL} &= \frac{A_{OL}}{1+BA_{OL}} = \frac{\frac{\omega_1 \omega_2}{(s+\omega_1)(s+\omega_2)} A_{DC}}{1+B \frac{\omega_1 \omega_2}{(s+\omega_1)(s+\omega_2)} A_{DC}} = \frac{\omega_1 \omega_2 A_{DC}}{(s+\omega_1)(s+\omega_2) + B \omega_1 \omega_2 A_{DC}} \\ A_{CL} &= \frac{\omega_1 \omega_2 A_{DC}}{s^2 + (\omega_1 + \omega_2)s + (1+BA_{DC})\omega_1 \omega_2} = \frac{A_{DC}}{1+BA_{DC}} \frac{(1+\beta A_{DC})\omega_1 \omega_2}{s^2 + (\omega_1 + \omega_2)s + (1+BA_{DC})\omega_1 \omega_2} \end{aligned}$$

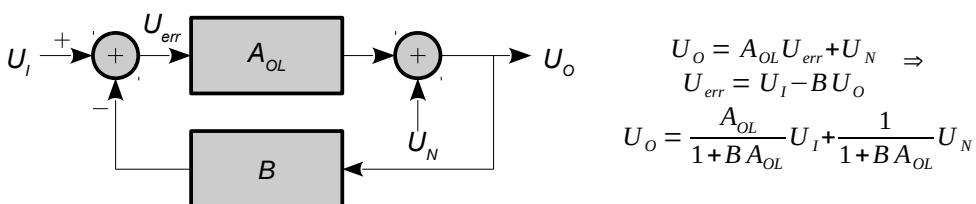
Den resulterende anden-ordens overføringsfunktion har opnået en større båndbredde ved multiplikationen af $\omega_1 \omega_2$ med $1+BA_{DC}$ og systemets stabilitet aflæses af dæmpningsfaktoren hvor det kan udnyttes at dæmpningsfaktoren er lavest ved to ens poler. Systemet er svagt underdæmpet ved $d = 0,7$ (Butterworth) som opnås for $1+BA_{DC} = 2$.

$$\begin{aligned} (\omega_1 + \omega_2)s &= 2d\omega_0 s \Rightarrow d = \frac{\omega_1 + \omega_2}{2\omega_0} = \frac{\omega_1 + \omega_2}{2\sqrt{(1+BA_{DC})\omega_1 \omega_2}} \quad \omega_2 = \omega_1 \Rightarrow d = \frac{1}{\sqrt{1+BA_{DC}}} \\ \omega_0^2 &= (1+BA_{DC})\omega_1 \omega_2 \end{aligned}$$

2.11.5 Reduktion af uønskede signaler

En vigtig egenskab ved et tilbagekoblet kredsløb er dets evne til at reducere problemer med eksterne signaler kobler ind i kredsløbet. Et typisk eksempel er brum fra effektforsyningen i en effektforstærker, men kredsløbet kan også selv være årsag til uønskede signaler som kan dæmpes med en tilbagekobling. Eksemplerne er en variation i åben-sløjfe forstærkningen fra en ændring i temperaturen eller en ikke-linearitet i et internt kredsløb som ved effektforstærkere.

En model af kredsløbet vises herunder hvor et uønsket signal U_N indkobles i signalvejen ved en addition til udgangssignalet. Det er uden betydning om U_N er et eksternt signal eller er genereret internt i A_{OL} .

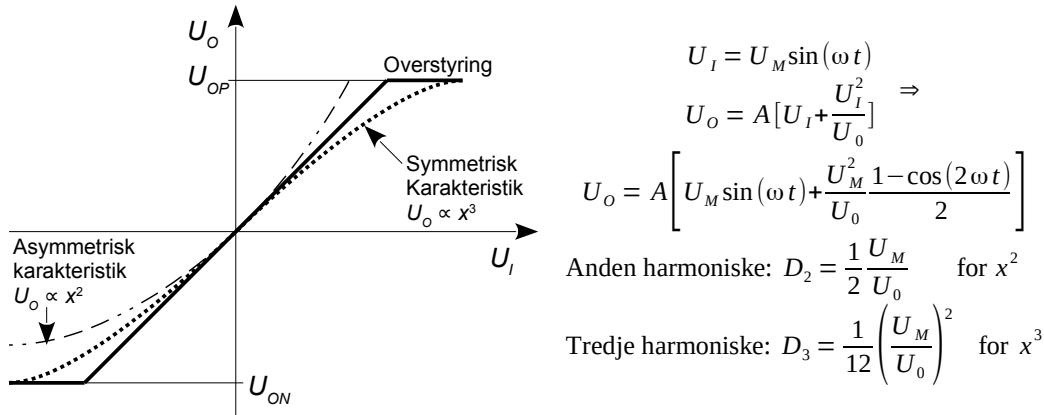


Figur 91 – Tilbagekobling af et kredsløb kan reducere uønskede signaler der kan repræsentere et eksternt støjsignal eller kredsløbets egen forvrængning.

Den matematiske model udvides med tilføjelse af støjsignalet U_N og efter reduktion af ligningerne ses at forstærkningen af indgangssignalet er det samme som før samt at det uønskede signal er reduceret med $1+BA_{OL}$ der kan være en meget høj talværdi. Det kan repræsentere stort set alt hvad der kan være galt i en forstærker og det er populært at antage at problemet derfor reduceres med tilbagekoblingsfaktoren. Der er dog grænser for hvad der er muligt at korrigere for og det uddybes i det følgende afsnit.

2.11.6 Forvrængning

Udgangssignalet fra en forstærker er ikke blot en forstærket kopi af indgangssignalet for selv den bedste forstærker vil "pynte" på resultatet og det kaldes forvrængning. Den mest almindelige type forvrængning er en krum karakteristik der kan beskrives ved at åben-sløjfe forstærkningen er en funktion af udgangens spændingsværdi. Det gælder for alle operations- og effektforstærkere så en sinus på indgangen bliver til en sum af harmoniske svingninger på udgangen hvilket kaldes for harmonisk forvrængning. Det kan beskrives som vist i afsnittet om Fourier serier.



Figur 92 – En forstærkers udgangssignal er sjældent en lineær funktion af indgangssignalet for oftest er relationen afhængig af spændingsniveauet hvilket giver en krum karakteristik (stiplet). Ved et kraftigt signal når udgangssignalet til et maksimum der ikke kan overskrides og som giver en momentan begrænsning (vist som U_{OP} i positiv retning og U_{ON} i negativ retning). Udledningen viser kun processen for anden-harmonisk forvrængning.

De simpleste former for forvrængning er en kvadratisk eller kubisk relation (x^2 eller x^3) der er en direkte følge af den måde signaler behandles elektronisk internt i forstærkeren og som vil blive mere udførligt beskrevet senere i bogen. Ved en kvadratisk relation består udgangssignalet af en forstærket udgave af dels indgangssignalet (AU_I) og dels en uønsket komponent der svinger på den dobbelte frekvens. Ved kubisk relation svinger det uønskede signal på tre gange frekvensen af indgangssignalet⁵⁵.

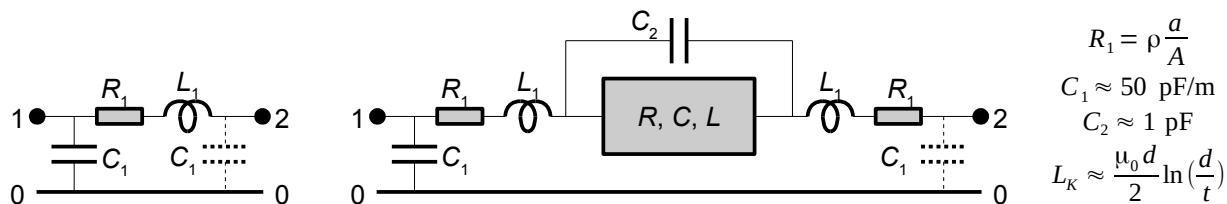
Som vist i forrige afsnit kan forvrængningen opfattes som et uønsket signal der kan reduceres ved tilbagekobling. Det udnyttes i enhver opstilling med operationsforstærkere og effektforstærkere. Hvis forstærkningen reduceres med faktor 10 vil den harmoniske forvrængning reduceres med et tilsvarende beløb så det gælder om at opnå en stor værdi af åben-sløjfe forstærkningen. Hvor langt ned forvrængningen skal bringes afhænger af applikationen, men som et udgangspunkt kan det menneskelige øre ikke opfatte en harmonisk forvrængning under 0,25 %.

Alle forstærkere har en øvre og nedre grænse for det mulige spændingssving på udgangen og et forsøg på at forstærke et signal for meget vil uundgåeligt begrænse signalets top og bund til det givne niveau. Denne forvrængning kan ikke reduceres ved tilbagekobling for forstærkningen falder til nul ved forsøg på at tvinge udgangen udenfor grænserne. Et tilsvarende problem optræder ved en effektforstærkers cross-over forvrængning hvor forstærkningen falder imod nul ved overgangen mellem svage positive og negative udgangsspændinger på grund af udgangstrinnet. Det gælder primært for de såkaldte klasse B forstærkere. Da $A_{OL} = 0$ kan forvrængningen ikke reduceres ved tilbagekobling.

⁵⁵ H F Olson viste omkring 1950 at lyden fra en almindelig højttaler med en 40 Hz tone på indgangen primært bestod af 120 Hz. Det oprindelige signal på 40 Hz kunne højttaleren kun vanskeligt gengive så det blev dæmpet. Den tredje-harmoniske forvrængning på 120 Hz blev dermed det kraftigste signal på udgangen.

2.12 Parasitiske komponenter

De passive komponenter er indtil nu betragtet som om de kunne realisere definitionsligningerne. Ingen komponent er dog ideel og det kan være nødvendigt at tage hensyn til afigelserne. Her vises modeller af både en ledere og en passiv komponent der kan være modstand, kondensator eller selvinduktion. Modellerne er relativt omfattende, men det er ikke nødvendigt at benytte alle dele ved en analyse. Man medtage det der er af betydning for den aktuelle applikation.



Figur 93 – Modeller af en lederbane og en passiv komponent sammen med tommelfingerregler for beregning af de parasitiske komponenters værdier.

Den enklest mulige komponent er en stump ledning eller en lederbane på et printkort som er vist til venstre i illustrationen. Den består af modstandsværdien R_1 fra ledernes materiale ρ , længden a i strømretningen og arealet A af dens tværsnit (se første kapitel). Hvis lederen er af kobber eller aluminium kan man ofte se bort fra modstandsværdien ($R_1 \approx 0$).

En ledere har en selvinduktion fra det magnetiske felt. En tommelfingerreglen for selvinduktionen er på $1 \mu\text{H}/\text{m}$, men lederen danner altid en løkke med det omkringliggende kredsløb så en bedre start kan være formlen L_K for en spole med én vinding, diameteren d og trådtynkelsen t .

En afkoblingskondensator på 100 nF benyttes ofte med integrerede kredse. Hvis komponenter og ledere danner en løkke med $d \approx 10 \text{ mm}$ og $t \approx 1 \text{ mm}$ findes $L_1 \approx 15 \text{ nH}$. Derved opnås serieresonans ved 4 MHz . Det betyder at afkoblingen af et analogt eller digitalt kredsløb har en lav impedans op til denne frekvens, men ved højere frekvenser er afkoblingen induktiv med stigende impedans. En selvinduktion på 15 nH vil nå op på $Z_L \approx 1 \Omega$ allerede ved 10 MHz og afkoblingen er illusorisk væsentligt over serieresonansen.

Tråden kobler med andre ledere eller med et stelplan og danner en kapacitet til omgivelserne. Som vist i første kapitel er værdien mellem to parallelle ledere på cirka 50 pF/m og lidt mere må antages for en lederbane på et printkort over et stelplan. En bane på 1 cm længde vil svare til 1 pF .

Når en passiv komponent anvendes i et kredsløb vil den være forbundet til de andre komponenter gennem ledere eller baner på et printkort og det påvirker funktionen. For en modstand er ledernes modstandsværdi ikke af betydning, men den kobler til omgivelserne gennem 1 pF selv ved meget korte lederebaner. En modstand på $1 \text{ M}\Omega$ holder derfor kun værdien under 100 kHz .

En kondensator har en seriemarkstand fra sin konstruktion. Det er den ækvivalente seriemarkstand *ESR* (equivalent series resistance) som opgives i databladet. For keramiske kondensatorer er den under $0,1 \Omega$ og vil kun sjældent genere, men for aluminium elektrolytkondensatorer kan den nå op over 10Ω . Værdien er højest for kondensatorer i området $0,1 \mu\text{F}$ til $10 \mu\text{F}$. Hvis en lav værdi er vigtig bør man ikke benytte aluminium elektrolytkondensatorer med værdier under $100 \mu\text{F}$ uden at studere databladet først. Dertil kommer kondensatorens lækstrøm der kan komme over $1 \mu\text{A}$ for en aluminium elektrolytkondensator. Keramiske kondensatorer kan regnes for nær ved ideelle i de fleste henseender, men de når ikke langt op i kapacitetsværdi.

En spole vil have en betydnende modstandsværdi fra den tråd den er viklet med og den beregnes af formlen for R_1 . Dertil en parallelkapacitet fra beviklingen der oftest er $C_2 < 100 \text{ pF}$ så en spole på $100 \mu\text{H}$ kan gå i resonans med sig selv på cirka 2 MHz . Spolen har en parallelmodstand til C_2 fra tab i en ferritkerne eller fra magnetisk kobling til omgivelserne.

2.13 Analogier

Det beskrevne materiale i denne bog kan anvendes til meget andet end elektroniske kredsløb for der er helt grundlæggende tale om et værktøj til løsning af differentialligninger og de kan bruges til at beskrive kemiske processer, mekaniske bevægelse, akustiske systemer, varmetransport, etc. Herunder vises hvordan man kan benytte relationerne mellem strøm- og spænding for analyse af en række systemer der er centrale for denne bog.

System	Kilder		Komponenter		
Elektrisk Impedans: $V/A = \Omega$	Spænding	Strøm	Kapacitet	Modstand	Selvinduktion
	V	A	$F = As/V$	$\Omega = V/A$	$H = Vs/A$
	$u(t)$	$i(t)$	$u = \frac{1}{C} \int i dt$	$u = Ri$	$u = L \frac{di}{dt}$
Mekanisk Impedans: $N/ms^{-1} = kg/s$	Kraft	Hastighed	Fjeder	Modstand	Masse
	N	m/s	m/N	Ns/m	kg
	$F(t)$	$v(t)$	$F = \frac{1}{C} \int v dt$	$F = Rv$	$F = M \frac{dv}{dt}$
Akustisk Impedans: $Nm^{-2}/m^3s^{-1} = kg/m^4s$	Tryk	Vol.hast.	Fjeder	Modstand	Masse
	$Pa = N/m^2$	m^3/s	m^5/N	Ns/m^5	kg/m^4
	$p(t)$	$q(t)$	$p = \frac{1}{C} \int q dt$	$p = Rq$	$p = M \frac{dq}{dt}$
Termisk Impedans: K/W	Temperatur	Effekt	Kapacitet	Modstand	(Ingen)
	K	W	J/K	K/W	$(ingen)$
	$T(t)$	$P(t)$	$T = \frac{1}{C} \int P dt$	$T = RP$	$(ingen)$

Det elektriske system benytter spænding u og strøm i . Spændingen over hver af komponenterne beskrives som vist for spolen L , modstanden R og kondensatoren C . Laplace transformationen benyttes til at erstatte differentiation og integration så impedansen er sL for selvinduktionen og $1/sC$ for kondensatoren.

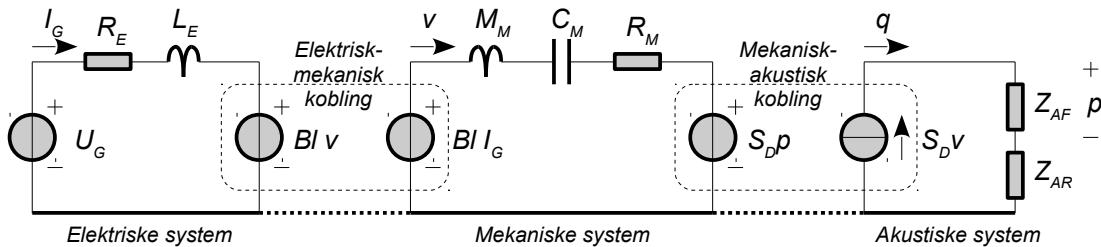
Det mekaniske system benytter kraft F og hastighed v for massen M , modstanden R og fjedren F . De tre komponenter beskrives ved de samme ligninger som for det elektriske systems komponenter så en masse kan repræsenteres ved en spole og fjedrens eftergivelighed C (*compliance*), der er reciprok stivhed, kan repræsenteres af en kondensator. Newtons anden lov $F = Mdv/dt$ giver relationen mellem kraft og masse og positionen beregnes fra definitionen af hastighed $v = dx/dt$. Fjederkraften er givet ved dens udvidelse x som integralet af hastigheden til $x = v/\omega$. Tab vil ofte antages at være proportionel med hastigheden, men det udelukker tab fra friktion.

Det akustiske system benytter tryk p (kraft per areal) og volumehastighed $q = S_D v$ (et rumfang luft der flyttes med en vis hastighed v af et stempel med arealet S_D). Når volumehastigheden er kendt beregnes trykket ved kilden som $p = Z_A q$, hvor Z_A er en akustiske impedans der almindeligvis er funktion af frekvensen. Lydtrykket aftager med afstandens kvadrat når lyden kan udbrede sig frit. Z_A kan overføres til det mekaniske system som et mekanisk serie-element ved at gange med kvadratet på højtalerns areal S_D^2 idet der er direkte proportionalitet mellem v og q .

Det termiske system beskrives ved temperatur T og effekt P . Den termiske inertি findes, men den mødes kun i ekstremt sjældne tilfælde såsom et system på størrelse med Jordkloden så i praksis antages det at et termisk system kun består af en termisk modstand og en termisk kapacitet.

2.13.1 Elektro-dynamisk højttaler

I første kapitel blev den elektro-dynamiske højttaler introduceret og det blev vist at membranens hastighed var omvendt proportionel med frekvensen. I dette afsnit udbygges modellen og det vises at den gælder indenfor et frekvensområde der er begrænset til cirka to dekader. Som model skal benyttes en højttaler monteret i en meget stor plade (uendelig baffel). Bl er kraftfaktoren som forbinder de elektriske og mekaniske systemer og der er tilsvarende en kobling til det akustiske system ved membranens areal $S_D = \pi a^2$ hvor a er radius af membranen.



Figur 94 – En elektro-dynamisk højttaler kan beskrives ved tre systemer der kobles sammen gennem styrede kilder. Modellen er velegnet for SPICE analyse.

Svingspolens DC modstand er R_E og dens selvinduktion er L_E . Højttalerens kraftfaktor er Bl . Membranens hastighed er v , dens masse er M_M , ophængets eftergiveleghed er C_M og de mekaniske tab er R_M . På den akustiske side er volumehastigheden q givet ved membranens areal S_D samt hastigheden v og højttaleren er belastet af strålingsimpedansen Z_{AF} ved højttalerens forside og Z_{AR} ved højttalerens bagside.

Lydtrykket p måles tæt på højttalerens membran. Ved en afstand større end radius af membranen vil lydtrykket aftage med kvadratet på afstanden hvis der ikke er reflekterende flader i nærheden.

2.13.2 Elektrisk system

På den elektriske side driver en generator højttaleren med spændingsamplituden U_G . Højttalerens svingspole har en DC modstand R_E på typisk tre fjerdedele af den nominelle impedans, så en 8Ω højttaler har $R_E \approx 6\Omega$. Selvinduktionen af svingspolen er $L_E \approx 1\text{ mH}$ for en basenhøjttaler og $100\mu\text{H}$ for en diskant. Når svingspolen bevæges i magnetfeltet induceres der en spænding i spolen, som er givet ved kraftfaktoren Bl gange med hastigheden v . Strømmen i svingspolen I_G er funktion af generatorens amplitude og den elektriske impedans af svingspolen, dog reduceret med Bl gange amplituden af membranens hastighed.

$$-U_G + R_E I_G + sL_E I_G + Blv = 0 \Rightarrow I_G = \frac{U_G - Blv}{R_E + sL_E} \quad \text{og} \quad f_s = \frac{R_E}{2\pi L_E}$$

For at lette bestemmelsen af membranens hastighed v ignoreres svingspolens selvinduktion, som med eksempelvis $R_E = 6\Omega$ og $L_E = 1\text{ mH}$ giver øvre grænse for modellen på $f_s < 1\text{ kHz}$.

2.13.3 Mekanisk system

I det mekaniske system omsættes strømmen I_G til bevægelse gennem acceleration af membranen med kraftfaktoren Bl gange strømmen I_G , hvor massen M_M inkluderer membran, svingspole og styr. Membranens hastighed er dertil påvirket af den akustiske side, idet trykket p omsættes til en kraft på membranen, og det har stor betydning ved montage i et kabinet.

$$-Bl I_G + sM_M v + \frac{1}{sC_M} v + R_M v + S_D p = 0 \Rightarrow v = \frac{Bl I_G - S_D p}{sM_M + 1/sC_M + R_M}$$

Det antages at trykvariationerne er små ($p = 0$), hvilket kun er tilladeligt for en umonteret højttaler, mens den indespærrede luft vil reducere eftergivelegheden C_M for en højttaler i et kabinet.

Membranens hastighed skrives nu som en anden-ordens funktion, der er bestemt af højttalerens parametre. Membranens hastighed ses at vokse med frekvensen (sC_M) for lave frekvenser ($s \approx 0$), mens hastigheden aftager over en karakteristisk frekvens hvor nævneren er domineret af ledet $s^2 M_M C_M$ og højttaleren er herved massestyret. Grænsen mellem de to områder er højttalerens resonansfrekvens f_0 hvor membranens hastighed er i fase med spændingen. Dæmpningen R_M er ofte lille sammenlignet med det elektriske systems $(Bl)^2/R_M$, der derfor er bestemmende hvordan højttaleren opfører sig omkring f_0 , og det er detaljeret i første kapitel.

$$\text{Generelt} \quad v = \frac{sC_M}{s^2 M_M C_M + sC_M [R_M + (Bl)^2/R_E] + 1} Bl \frac{U_G}{R_E} \quad \text{hvor} \quad L_S = 0 \quad \text{og} \quad p = 0$$

$$f_0 < f < f_s \quad v \approx \frac{Bl}{s M_M} \frac{U_G}{R_E} \quad f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi \sqrt{M_M C_M}}$$

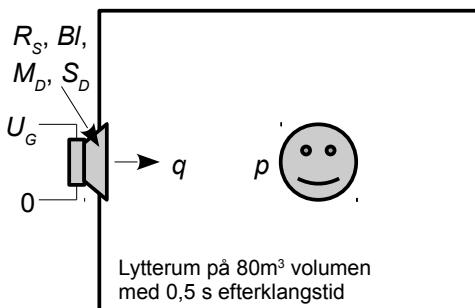
Med højttalerens værdier $M_M = 0,02 \text{ kg}$ og $C_M = 1,4 \text{ mm/N}$ findes resonansen til $f_0 = 30 \text{ Hz}$.

2.13.4 Akustisk system

På den akustiske side omsættes membranens hastighed til lyd ved at fungerer som et stempel, der flytter en luftmængde med arealet S_D og hastigheden er v , så volumehastigheden bliver $q = S_D v$. Den lydeffekt, der udstråles fra det akustiske system er givet ved strålingsimpedansens realdel r_A som $P_A = r_A q^2$ i stil med elektrisk effekt ($P = RI^2$). Det forudsættes at dimensionerne er små i forhold til lydens bølgelængde, hvilket udtrykkes ved faktor ka , der beskriver forholdet mellem højttalerens omkreds ($2\pi a$) og bølgelængden ($\lambda = c/f$). Højttaleren kan regnes for akustisk lille når $ka < 1$ og det er ved 550 Hz for en 200 mm højttaler (8 tommer).

$$ka = \frac{2\pi a}{\lambda} = \frac{2\pi a f}{c} < 1 \Rightarrow f < \frac{c}{2\pi a}$$

Under denne forudsætning er strålingsimpedansens givet ved udtrykket herunder [KR-3] og det er muligt at beregne det nødvendigt membranudsving x_{MAX} for en ønsket akustisk effekt og frekvens.



$$P_A = r_A q_{RMS}^2 \Rightarrow P_A = \frac{\rho_0 (\omega S_D v_{RMS})^2}{4\pi c}$$

$$r_A = \frac{\rho_0 c k^2}{4\pi}$$

$$v_{RMS} = \frac{1}{\omega S_D} \sqrt{\frac{4\pi P_A c}{\rho_0}}$$

$$x_{MAX} = \sqrt{2} \int v_{RMS} dt = \frac{\sqrt{2}}{\omega^2 S_D} \sqrt{\frac{4\pi P_A c}{\rho_0}}$$

Figur 95 – Lydtrykket i et beboelsesrum kan beregnes når den udstrålede akustiske effekt er kendt, og derved kan kravet til membranens amplitude beregnes.

Med en udstrålet akustisk effekt på 15 mW, er hastigheden af membranen $v_{RMS} = 0,93 \text{ m/s}$ ved $f = 40 \text{ Hz}$ for en højttaler med radius 100 mm ($S_D = 31,4 \cdot 10^{-3} \text{ m}^2$), og amplituden af membranens bevægelse bliver på $x_{MAX} = 5 \text{ mm}$, der svarer til 10 mm spids-til-spids.

Lydtrykket er funktion af den udstrålede effekt og en akustisk effekt på 1 W som middel over tid er rigeligt til hjemmebiografen selv ved høje lydtryk⁵⁶. Højttaleren begrænser lydtrykket ved lave frekvenser for amplituden af svingningen er omvendt proportionelt med kvadratet på frekvensen, og det er svært at bygge en højttaler, der kan klare mere end $\pm 10 \text{ mm}$; og selv om man kunne så ville lydgengivelsen lide under forvrængning på grund af Dopplereffekten.

56 Den akustiske effekt kan omsættes til lydtryk i et almindelig beboelsesrum idet et lydtryk på 1 Pa (94 dB SPL) opnås ved en effekt på $P_A = 15 \text{ mW}$ når rummet er på 80 m^3 med en efterklangstid på 0,5 s [KR-4]. For en højttaler med 1 % virkningsgrad svarer det til en tilført elektrisk effekt på 1,5 W som middelværdi over et passende langt stykke tid.

2.14 Opgaver

1 – Benyt formlerne på side 61 til at beregne frekvensen f_0 ved resonans for en serieforbindelse af en kondensator på $C = 1 \mu\text{F}$, en selvinduktionsspole med $L = 1 \text{ mH}$ og en seriemedstand på $R = 10 \Omega$. Beregn godheden Q og båndbredden Δf . **Svar:** $f_0 = 5 \text{ kHz}$, $Q = 3,2$ og $\Delta f = 1,6 \text{ kHz}$.

2 – Tegn et koordinatsystem for den komplekse talplan som vist på side 62. Marker tallet $z_1 = 3 + j4$. Beregn de komplekse tal $z_2 = jz_1$, $z_3 = jz_2$ og $z_4 = jz_3$ og plot dem i talplanen. **Svar:** $z_2 = -4 + j3$, $z_3 = -3 - j4$, $z_4 = 4 - j3$.

3 – Brug formlerne fra side 62 til at beregne modulus r og argument θ for de komplekse tal z_1 , z_2 , z_3 og z_4 fra forrige opgave. **Svar:** Modulus er $r_1 = r_2 = r_3 = r_4 = 5$, og argumenterne er $\theta_1 = 53^\circ$, $\theta_2 = 143^\circ$, $\theta_3 = -127^\circ$, $\theta_4 = -37^\circ$. Det er lettets at indse fortegnene grafisk.

4 – Beregn summen $z_1 + z_2$. **Svar:** $z_1 + z_2 = -1 + j7$.

5 – Beregn differensen $z_1 - z_2$. **Svar:** $z_1 - z_2 = 7 + j$.

6 – Opskriv den komplekst konjugerede z_3^* for z_3 fra de to tidligere opgaver og beregn produktet af disse to tal. **Svar:** $z_3^* = -3 + j4$ og $z_3 z_3^* = 25$. Bemærk at produktet er et reelt, positivt tal.

7 – Beregn produktet $z_1 z_2$. Komplekse tal ganges sammen som almindelige reelle tal, dog udnyttes det at $j^2 = -1$. **Svar:** $z_1 z_2 = -24 - j7$.

8 – Beregn kvotienten z_1/z_2 . For at bringe udtrykket på formen $a + jb$ forlænges brøken ved at multiplicere tæller og nævner med det komplekst konjugerede til nævneren. **Svar:** $z_1/z_2 = -j$

9 – Brug eksemplet på side 67 til at vise at $f(x) = x^3$ giver $f'(x) = 3x^2$ ved differentiation.

10 – Brug eksemplet på side 67 til at vise at $f(x) = \cos(x)$ giver $f'(x) = -\sin(x)$ ved differentiation.

11 – Vis at differentialkvotienten af $f(x) = \exp(\omega t)$ bliver til $f'(x) = \omega \exp(\omega t)$ hvor ω er en konstant.

12 – Brug eksemplet side 69 for halvbølge sinus til at vise at integralet over cosinus fra 0 til π er nul.

13 – Brug tabellen side 70 til at vise at for $f(x) = 1$ er differentialkvotienten til $f'(x) = 1$ og integralet $F(x) = x + c$ hvor c er en ukendt konstant. Bestem konstanten c når det kræves at funktionen $F(x)$ skal gå gennem origo (0, 0). **Svar:** $c = 0$.

14 – Brug tabellen side 70 til at vise at $f(t) = \exp(\omega t)$, hvor ω er en konstant, integreres til $F(t) = \exp(\omega t)/\omega$.

15 – Relationen for spændingen over en spole u_L som funktion af strømmen i_L igennem den er vist nedenfor. Isolér ændringen i strømmen di_L og brug tabellen side 70 for at vise at udtrykket for strømmen i_L som funktion af tiden er som angivet herunder hvor integrationskonstanten

er I_0 (begyndelsesværdien af i_L til $t = 0$). Beregn værdien af strømmen til tidsrummet $t_{ON} = 1 \mu\text{s}$ for $L = 100 \mu\text{H}$ efter tilslutning af spændingskilden $U = 5 \text{ V}$. Spolens strøm er nul ved start af beregningen. Hvor er strømstyrken vokset til ved $t_{ON} = 2t_{ON}$? **Svar:** $i_L = 100 \text{ mA}$.

$$u_L = L \frac{di_L}{dt} \Rightarrow i_L = \frac{U}{L} t_{ON} + I_0 = 50 \text{ mA}$$

16 – Fakultetsregning introduceres side 73 og beregnes ved at gange tallene fra 1 til n . Beregn værdien for tallene $n = 1 \dots 5$. **Svar:** $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$ og $5! = 120$.

17 – Taylor serierne beregnes ved fortsat differentiation af et polynomium. Beregn alle afledte af $f(x) = a + bx + cx^2$.

Svar: $f(x) = b + 2cx$, $f'(x) = 2c$, $f''(x) = 0$ og alle højere forbliver på nul: $f^{(n)}(x) = 0$ for $n > 3$.

18 – Brug tabellen side 74 til at vise at $[1 - \cos(x)]/x$ bliver nul for $x = 0$.

19 – Brug Eulers identitet side 75 til at udlede udtrykket for sinus.

20 – En sammenkobling af spolen L og kondensatoren C danner resonans ved frekvensen f_1 som er givet herunder. Brug tabellen side 74 til at vise at en beskeden forøgelse af spolens værdi fra L til $L + \Delta L$ på 1 % sænker frekvensen fra f_1 til f_2 med kun halvdelen af ændringen (0,5 %). Tjek resultatet mod formlen nederst på tabellens side.

$$f_1 = \frac{1}{2\pi\sqrt{LC}} \Rightarrow f_2 = \frac{1}{2\pi\sqrt{LC}} \left(1 - \frac{\Delta L}{2L}\right)$$

21 – Brug Eulers identitet side 75 til at eftervises den følgende formel for produktet af to sinus-funktioner.

$$\sin(x)\sin(y) = \frac{\cos(x+y) - \cos(x-y)}{2}$$

22 – Koefficienterne for en Fourier serie skal beregnes efter formlerne side 77 for en firkantet puls der er én fra nul til τ og nul for resten af perioden på T . Amplituden af harmoniske c_n beregnes som vist og opgaven er at vise udtrykkene for koefficienterne a_n og b_n samt c_n og at finde værdien af de første seks harmoniske for $\tau/T = 1/3$.

$$f(t) = \begin{cases} 1 & 0 \leq t \leq \tau \\ 0 & \tau < t < T \end{cases} \quad \text{periode: } T$$

$$a_n = \frac{\sin(2\pi n\tau/T)}{\pi n} \quad c_n = \sqrt{a_n^2 + b_n^2} = \left| \frac{2}{\pi n} \sin\left(\frac{\pi n\tau}{T}\right) \right|$$

$$b_n = \frac{1 - \cos(2\pi n\tau/T)}{\pi n}$$

Svar: $c_0 = 0,667$ (DC værdi 0,333), $c_1 = 0,551$, $c_2 = 0,276$, $c_3 = 0$, $c_4 = 0,138$, $c_5 = 0,110$, $c_6 = 0$.

23 – Vis at modulus af $H(f)$ side 82 giver sinc-funktionen. Du får brug for definitionen af modulus side 62 og et par regneregler for trigonometriske funktioner.

3 OPERATIONSFORSTÆRKER

Dette kapitel omhandler en universelt anvendelig komponent, der kan programmeres til opgaven ved hjælp af nogle få eksterne komponenter. Den kan bruges til forstærkning af spændinger i μV området fra en termokabler eller strømme i nA området ved lysfølsomme dioder, den kan anvendes for tilpasning af en transducers signal ved integration eller differentiation, den kan undertrykke uønsket brum og støj fra en transmissionskanal, den kan varetage tonekontrol i et lydanlæg eller dele frekvensområdet mellem flere kanaler i et højttaleranlæg og den kan stå for ikke-lineær signalbehandling som ensretning, logaritmering og begrænsning. Der findes desuden varianter for brug i en radios mellemfrekvens, for videosignaler eller for at drive en DC motor eller en højttaler. I dette kapitel introduceres denne fleksible komponent.

Det er elektronikkens arbejdshest, der skal introduceres. Basalt set er den bare en sort boks, der udfører en opgave som den bliver programmeret til af det kredsløb den sættes ind i; den udfører en given operation og derfor kaldes det en operationsforstærker (*OpAmp, operational amplifier*). Der er selvfølgelig nogle krav som den anvendte operationsforstærker skal opfylde for at kunne løfte en opgave, men når kravene først er defineret bør enhver kompatibel forstærker kunne fungere i den aktuelle opstilling, og denne fleksibilitet er årsagen til operationsforstærkerens store succes.



Figur 96 – De elektroniske kredsløb samles på en fælles siliciumchip og indstøbes i plast som afskærmning mod lys. Yderst til venstre den populære LM324 fra 1972, som indeholder fire ens trin og på et tidspunkt var den mest solgte integrerede operationsforstærker. Billedet i midten viser DIL pakkerne (Dual In-Line) med de elektriske tilslutninger i 2,54 mm raster (1/10 tomme), derunder SOIC (Small Outline Integrated Circuit) med 1,27 mm raster og sidst en SIL pakke (Single In-Line), der er ofte brugt til små effektorstærkere. Til højre vises en SOT pakke (Small Outline Transistor), hvor målene er cirka 1 mm gange 2 mm for plasthuset, og der kan være fra tre til seks tilslutninger.

Søger man i producenternes hjemmesider⁵⁷ findes et helt overvældende udbud af valgmuligheder, der dels viser hvor voldsom udbredelsen er blevet siden operationsforstærkerens introduktion i 1966, og dels viser at der er rigtig mange parametre at tage stilling til når et design skal optimeres. Som en erfaren udvikler er det et overflødighedshorn af valgmuligheder, men for den nyttilkomne er den barokke overflod nok mere skræmmende end tilløkkende. For at komme ind i det fine selskab skal du jo starte et eller andet sted, og mit råd er at grib i posen og derefter at stole på at den fundne komponent kan løfte opgaven. Når du senere kommer til den erkendelse at en af dens parametre skal optimeres; så er du inde i varmen.

Der er mange opfattelser af hvorvidt en operationsforstærker er egnet til en given opgave, eller om det var bedre med en diskret tilgang, og indenfor lydgengivelse er der en udbredt holdning imod at bruge komponenten. Det skyldes dels dårlige erfaringer med de tidlige operationsforstærkere, der ikke alle var lige velegnede til krævende analog signalbehandling, og dels konservativisme efter den velkendte devise: hvis det er håndbygget så er det nok også godt. Faktum er dog at der i dag er integrerede operationsforstærkere for signalbehandling i selv kostbart topudstyr.

Mere reelt er problemet med at finde egnede operationsforstærkere for forstærkning af signaler i radiofrekvensområdet, og ved et spændingsniveau over 30 V eller en strømstyrke over 3 A.

⁵⁷ Analog Devices, Linear Devices, Maxim Semiconductors, Motorola, National Instruments, Philips (NXP).

3.1 Lidt historie

Opgaven med at designe kredsløb er siden 1900-tallets start drejet fra at overveje detaljer om den måde de skal opbygges på, og over til at vælge en egnet komponent fra producentens datablad. Det er radikal nytænkning, og paradigmeskiftet sker omkring 1970, selv om mange udviklere blot fortsætter med individuelt at designe kredsløb til den aktuelle opgave; for det kan være svært at omstille sit sind fra at optimere hver del af en opstilling, til at stole blindt på de data, der kan hentes i dokumentationen fra leverandøren. Lidt efter lidt bliver designerne dog vant med databladene, og i stedet for at udtænke kreative løsninger bruges overskuddet så på anden form for kreativitet, som for eksempel mere komplekse opstillinger eller en bedre kvalitet af produktet.

Signalforstærkning starter i 1902 med opfindelsen af radiorøret⁵⁸, der 20 år senere er kommersielt interessant med verdensomspændende radiotransmissioner og opbygning af et marked hos den voksende mellemklasse. Julius Edgar Lilienfeld patenterer 1925 en halvlederkomponent med den samme funktion som radiorøret, men uden en sårbar glaskolbe og glødetråd, men komponenten slår ikke an, muligvis på grund af de stærke kommersielle interesser i radiorøret. Alan Blumlein patenterer 1936 en balanceret kobling med to radiorør, som danner forløberen for den moderne operationsforstærker og som udnyttes i den analoge computer. Den bipolære transistor patenteres 1947 af John Bardeen og Walter Brattain ved Bell lab, og muliggør fra cirka 1960 introduktionen af den integrerede kreds med digitale funktioner, og kort efter operationsforstærkeren, eksempelvis den velkendte uA741 fra 1968. Lilienfelds transistor dukker op som en MOSFET og den danner forudsætningen for at introducere mikroprocessoren, og herefter eskalerer udviklingen.

Batteriforsyning er interessant og LM324 fra 1972 kan arbejde fra 5 V ved 1,5 mA. Et designtrick tillod indgangen at arbejde ved spændinger under nul og udgangen tæt på nul, hvilket foregreb forstærkere med rail-to-rail på både indgang og udgang, som fx den moderne ADA4051, der kan køre fra 1,8 V med blot 13 μ A i forbrug. Højtydende operationsforstærkere som NE5534 fra 1980 og AD797 fra 1992 vandt indpas i den hurtigt voksende og meget krævende lydverden på grund af lav støj og forvrængning. Ligeledes blev designet i stigende grad baseret på felteffekt transistorer på grund af den meget lave indgangsstrøm, fx TL071 fra 1978 med JFET i indgangen, der blev meget populær på grund af et fornuftigt kompromis mellem kvalitet og pris, og lidt senere TLC271, der udelukkende består af MOSFET.

De beskrevne komponenter har til fælles at de kun kan leve cirka \pm 30 mA til belastningen. Det er nok til at drive en hovedtelefon, men langt fra tilstrækkeligt for en højttaler eller en lille DC motor. Det område kan dækkes af de mere specielle forstærkere som LM675T og L165V, der kan leve \pm 3 A til en belastning og er gode til strømforsyninger og mere elementær analog signalbehandling. Det er stiks vanskeligere hvis det drejer sig om lydgengivelse af høj kvalitet, for det er notorisk svært at bygge en integreret forstærker for lydgengivelse af god kvalitet; et par muligheder er TDA2040 og TDA7294, men udgangseffekten er begrænset til 50 W. De nævnte kredse har alle til fælles at de er opbygget som operationsforstærkere; blot for en højere strømstyrke.

Effektforstærkere fra 50 W og op følger i helt samme spor som operationsforstærkeren, blot er de nu diskret opbygget da spændingsniveauet er for højt til en effektiv integration på en chip, og den afsatte effekt er et problem, der ikke er let at tackle, som enhver computerinteresseret vil vide.

Konklusionen er at uanset om der skal designes analoge kredsløb med kommersielt tilgængelige operationsforstærkere, eller motorstyring med de integrerede effektforstærkere, eller højtydende effektforstærkere med diskrete komponenter, så er de alle operationsforstærkere. Teorien er helt den samme; så de følgende betragtninger i denne bog er alle fuldt ud relevante.

⁵⁸ Humphry Davy opdager i 1802 at en opvarmet, elektrisk leder udsender lys, og Warren de la Rue indbygger 1840 lederen i en lufttom glaskolbe hvorved glødelampen opfindes. Elektronemissionen opdages i 1873 af Frederick Guthrie og fører til elektronrøret med to elektroder, der derfor kaldes for en diode. John Ambrose Fleming bruger den ved radiomodtagelse da den kan ensrette et oscillérende signal, og 1902 indbygger han et gitter i dioden hvorved strømmen kan styres og trioden opfindes. Det leder til en eskalérende udvikling af radiokommunikationen, TV transmission og videnskab, og den skaber et nyt marked indenfor konsum.

3.2 En OpAmp

En OpAmp har to indgange og én udgang. Indgangene monitorerer potentiometret på hver sit sted i et kredsløb, og udgangen driver et tredje sted i kredsløbet som et resultat af hvad der er blevet målt. Det helt centrale princip for operationsforstærkeren er at:

En OpAmp placerer udgangen så de to indgange har samme potentiiale.

Dette enkle princip er nok til at definere den vigtigste komponent i den analoge elektronik, men det kan kun lade sig gøre, hvis operationsforstærkerens indgange og udgang er indenfor grænserne sat af effektforsyningen, og signalet ikke varierer så hurtigt at forstærkeren ikke kan følge med. Indtil videre går jeg ud fra at disse krav er opfyldt, men det er to meget vigtige forudsætninger, og da de ikke altid er opfyldt, vil jeg vende tilbage til problemstillingen senere i kapitlet.



Figur 97 – Den ideelle OpAmp vises til venstre med dens symbol og en analytisk model ses til højre. Fællesterminalen, der kaldes nul, stel, jord eller chassis (på engelsk *ground*), forefindes ikke, så en OpAmp er henvist til at benytte effektforsyningen som reference. Det meste litteratur afviger fra virkeligheden på dette punkt, og det er ret sjældent at forholdet overhovedet omtales.

Når de to indgange ligger på samme potentiale siger operationsforstærkeren at være i balance, og man kan tillade sig at forudsætte at den arbejder i sit lineære område. Hvis indgangene ikke er ens så er forstærkeren ikke i balance og den vil ændre spændingen på udgangen for at opnå balance. Det er så designerens opgave at sikre, at den retning forstærkeren flytter sin udgang er den rigtige for at gendanne balancesituationen. Hvis det ikke kan opnås vil forstærkerens udgangsspænding nå ud til enten den positive eller negative effektforsyning, eller hvad der nu sætter grænsen for udgangsspændingen. Forstærkeren er da overstyreret eller mættet (*saturated*) og den opfylder ikke længere den tilsigtede funktion; selvfølgelig med den undtagelse at overstyringen er tilsigtet.

De to indgange kan ikke nøjes med at måle spændingsdifferensen mellem to punkter i et kredsløb; de vil uundgåeligt påvirke kredsløbet. Det er en ret væsentlig ulempe ved en OpAmp, og det kan have afgørende betydning for om en aktuel OpAmp vil virke i en given opstilling og det problem analyseres side 129. Tilsvarende er udgangen ikke en ideel spændingskilde for der er en udgangsmodstand, og det giver problemer med kapacitiv belastning, som omhandles side 137. Problemer med forstærkning af højfrekvente signaler diskuteses side 134 og risiko for overstyring af forstærkeren ved stepsignaler behandles side 136. En OpAmp genererer sine egne signaler, hvilket forklares side 147, og den vil uundgåeligt forvrænge signalet, og det omtales side 166.

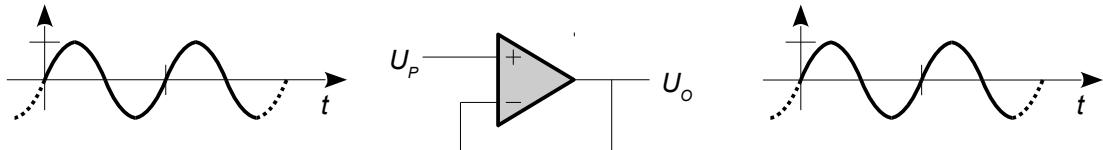
I de følgende afsnit vises brugen af en operationsforstærker i nogle helt almindelige opstillinger. Det er klart at der må være en eller anden form for effektforsyning, for ellers ville der jo ikke ske noget, men effektforsyningen er normalt underforstået i diagrammaterialet. Det skyldes at den sjældent er vigtigt for at forstå selve funktionen af kredsløbet. Den bliver ganske enkelt ikke tegnet, men i praksis er effektforsyningen vigtig og jeg vender tilbage til emnet på side 166.

Det anbefales at anvende et stelplan (et større kobberareal med forbindelse til nul) og desuden afkoble effektforsyningen ved hver eneste forbindelse til den positive eller negative effektforsyning direkte til stelplanet med 100 nF.

Indtil videre er en OpAmp et niveau i den abstraktion, man som designer må vænne sig til, for at drage fordel af dens elegante konstruktion. Der er selvfølgelig en lang række begrænsninger, med de mere væsentlige nævnt ovenfor, men man plejer at antage at operationsforstærkeren opfører sig som den ideelle operationsforstærker ville gøre det, og i lang udstrækning er det tilstrækkeligt for at lande helskinnet med et design.

3.3 Spændingsforstærkning

Det simplest mulige kredsløb indeholder kun operationsforstærkeren selv. Spændingen på den negative indgang er lig med spændingen på udgangen, og da forstærkeren søger at holde de to indgange på samme potentielle ($U_P = U_N$) vil udgangssignalet følge med indgangssignalet. Det kaldes for en spændingsfølger (*voltage follower*) eller en buffer.

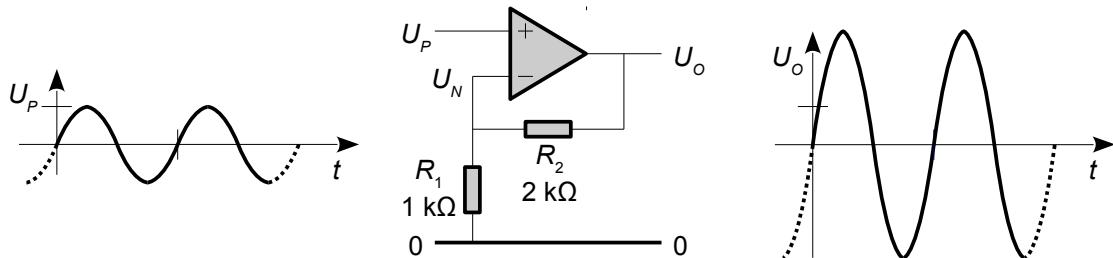


Figur 98 – En spændingsfølger kan adskille to netværk, så de ikke påvirker hinanden.

Ideelt set vil operationsforstærkeren ikke påvirke det indgangen måler på og udgangen antages at kunne leverer tilstrækkelig med strøm til at drive det efterfølgende kredsløb. Det betyder at en spændingsfølger kan isolere to dele af et kredsløb fra hinanden. Det er en ofte set kobling ved sammensætning af filtre og interface til sensorer, der ikke tåler hård belastning.

3.3.1 Ikke-inverterende forstærker

Med en spændingsdeler fra udgangen til den negative indgang skal udgangen svinge kraftigere end indgangen for at opnå balance, og det leder til en overordentligt populær konstruktion, da den kan forstærke det indkommende signal. Udgangssignalet har samme polaritet som signalet på indgangen, hvorfor den kaldes for en ikke-inverterende forstærker.



Figur 99 – En spændingsforstærker kan adskille to netværk og dertil forstærke signalet.

Udgangsspændingen U_O er nu givet af indgangsspændingen U_P og spændingsdelen ved R_1 og R_2 til følgende ligning, der nok er den mest benyttede relation overhovedet for en forstærker.

$$U_O = \left(1 + \frac{R_2}{R_1}\right) U_P \Rightarrow A_P = \frac{U_O}{U_I} = 1 + \frac{R_2}{R_1}$$

Med komponentværdierne fra figuren findes forstærkningen til $A_P = 3$. Med 2 V på indgangen vil udgangen stilles på 6 V. Hvis indgangssignalet er en sinus med ± 1 V så vil udgangen svinge med ± 3 V.

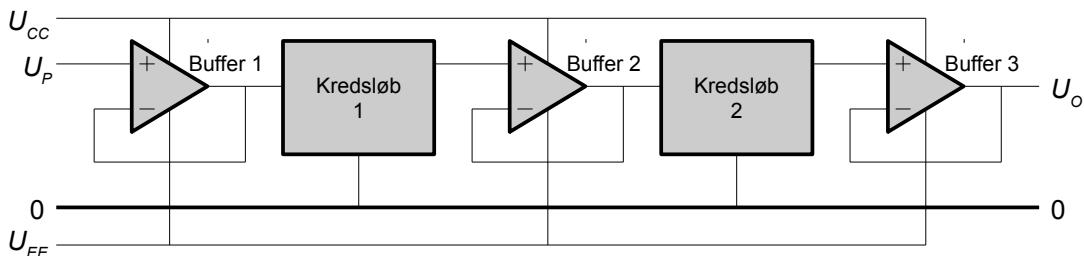
Det er en brøkdel af spændingen på udgangen, der føres tilbage til den negative indgang, og denne brøkdel er givet ved spændingsdelingen over R_1 og R_2 . Forstærkeren vil søge at holde de to indgange på samme potentielle.

$$\begin{aligned} U_N &= U_P \\ U_N &= \frac{R_1}{R_1 + R_2} U_O \Rightarrow U_O = \frac{R_1 + R_2}{R_1} U_P = \left(1 + \frac{R_2}{R_1}\right) U_P \end{aligned}$$

I det følgende præsenteres eksempler på anvendelsen af bufferen og den ikke-inverterende forstærker. Derefter introduceres den inverterende forstærker og en række kredsløbseksempler.

3.3.2 Buffer

Den klassiske anvendelse af en spændingsfølger er at sikre en lav modstandsværdi på det drevne punkt ved kredsløbets indgang, og en meget høj modstandsværdi ved kredsløbets udgang, så en belastning ikke får utilsigtet indvirkning på kredsløbet. Almindelig elektroteknik forudsætter at et kredsløb arbejder under disse forhold, så man står sig ved at inkludere det nødvendige antal spændingsfølgere (buffere) for at opnå den tilsigtede funktion. Den ikke-inverterende forstærker kan beskytte et kredsløb på samme måde som spændingsfølgeren og i tilgift forstærke signalet.

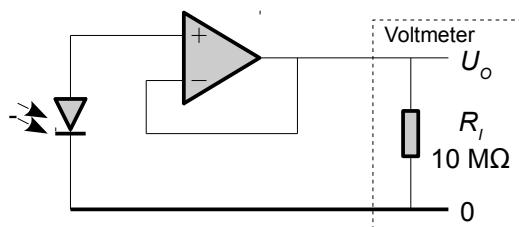


Figur 100 – En spændingsfølger driver indgangen til et kredsløb fra en lav modstandsværdi og sikrer imod en hård belastning af kredsløbets udgang.

Udgangssignalet skal holdes indenfor de grænser, der sættes af effektforsyningen, og det faktum at de fleste operationsforstærkere kan ikke nå helt ud til grænsen, men kræver et par volt mellem udgangen og effektforsyningen. Det er ikke det store problem ved $U_{CC} = 15 \text{ V}$ og $U_{EE} = -15 \text{ V}$, hvor udgangssignal kan forventes at være ubegrænset indtil $\pm 12 \text{ V}$, men ved effektforsyning fra en lavere værdi som $U_{CC} = 5 \text{ V}$ og $U_{EE} = 0$ skal der overvejes brug af en operationsforstærker af typen *rail-to-rail*, for ellers kan signalet blive begrænset og dermed forvrænget (se side 165). Dertil kommer at det maksimale indgangssignal skal være den samlede forstærkning mindre end det maksimale udgangssignal, så ved en effektforsyning på $\pm 15 \text{ V}$ og en forstærkning på 3 gange bliver det maksimale indgangssignal på $\pm 4 \text{ V}$ under de angivne forudsætninger.

3.3.2.1 Fotodiode

En fotodiode genererer en strøm som reaktion på det indfaldne lys på cirka 1 mA ved stærkt lys og ned til 10 nA ved netop synligt lys. Strømmen løber tilbage gennem dioden selv og ved stærkt lys er der cirka 600 mV over dioden. Spændingen vil ændres 60 mV per dekade af lysstyrken, så ved netop synligt lys er der 300 mV over dioden, og opstillingen har potentielle for at måle svagere lys end svarende til synligt niveau. En indgangsmotstanden R_i på $10 \text{ M}\Omega$ ved voltmeteret vil dog kun tillade en spænding på 0,1 V, så en spændingsfølger er nødvendig. Det er samtidigt et krav at den benyttede operationsforstærker belaster dioden med væsentligt mindre end 10 nA (se side 131).

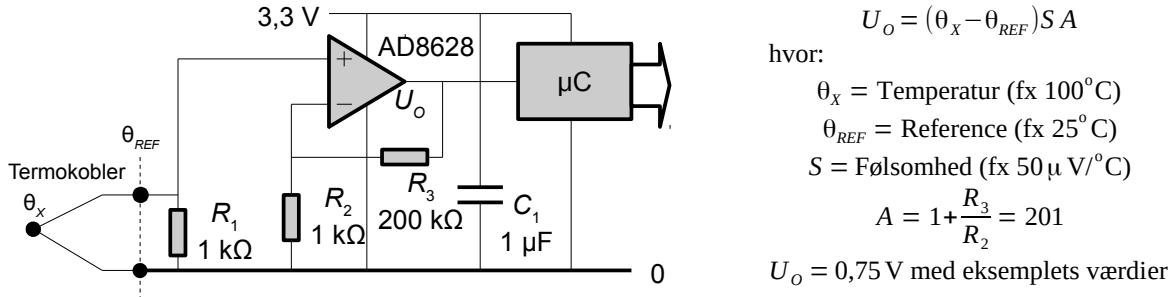


Figur 101 – En spændingsfølger anvendt som buffer for en fotodiode så spændingen måles uden at voltmeterets indre modstand belaster fotodioden ved at optage en strøm.

En anden anvendelse er ved en pH-probe, hvor spændingsforskellen er mellem to elektroder adskilt af en tynd membran af glas, så probens indre modstand er enorm. Her er der også en variation på 60 mV per dekade ændring i koncentrationen af brintioner.

3.3.3 Termokabler

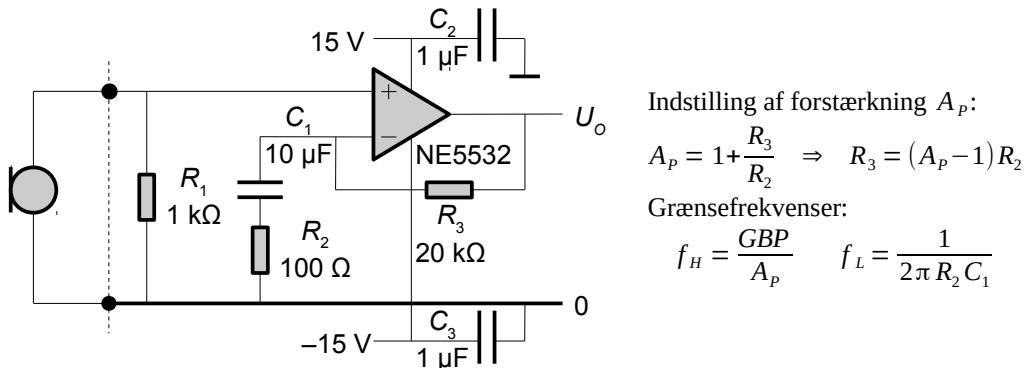
En ikke-inverterende forstærker kan løfte det svage signal fra en termokabler⁵⁹, op til et niveau hvor en A/D konverter kan omsætte det til et digitalt signal for brug i en mikroprocessor. En termokabler afgiver et meget svagt signal, omkring 50 μV per grad celsius, og det er væsentligt mindre end en A/D konverters opløsning⁶⁰, så der er behov for forstærkning, der løfter signalet op til 10 mV/ $^{\circ}\text{C}$. Modstanden R_1 holder indgangen på nul hvis termokoblen fjernes. Forstærkeren skal kunne arbejde med signaler nær ved nul og C_1 afkobler effektforsyningen (se side 166).



Figur 102 – En termokabler genererer en svag spænding, som skal forstærkes op for at kunne måles med en A/D konverter. Opstillingen måler positiv temperatur overfor referencepunktet.

3.3.3.1 Mikrofonforstærker

En mikrofon omsætter et akustisk signal til et tilsvarende elektrisk signal⁶¹, og mikrofonens signal er oftest under $\pm 10 \text{ mV}$. Signalet behøver forstærkning for at drive en effektforstærker eller en A/D konverter, og niveauet er her valgt til $\pm 1 \text{ V}$ ved et relativt kraftigt akustisk signal.



Figur 103 – En forstærker for en dynamisk mikrofon.

Signalet ledes til den positive indgang på operationsforstærkeren, og kondensator C_1 blokerer for DC så forstærkerens store offsetspænding så den ikke påvirker funktionen (se side 129). Kondensatoren bestemmer samtidigt den laveste frekvens, der kan overføres til 13 Hz, og den højeste frekvens er givet af forstærkeren og den indstillede forstærkning til 50 kHz (se side 134). Indgangen holdes nær nul ved R_1 for at sikre et korrekt potentiale på indgangen af operationsforstærkeren hvis mikrofonen udtages. Forstærkningen indstilles med R_2 og R_3 til 201 gange (se side 108). For at opnå et lavt støjniveau bør R_2 vælges til en lavere værdi end mikrofonen (se side 154).

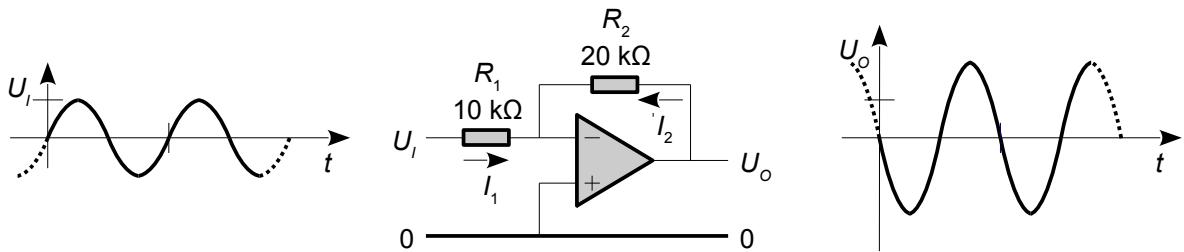
59 En termokabler består at to forskellige stykker metal, der er svejset sammen. Ved en temperaturforskæl mellem målepunktet og tilslutningsstedet vil der optræde en spændingsforskæl mellem metaltrådene. For en E-type termokabler er værdien 68 $\mu\text{V}/^{\circ}\text{C}$.

60 En typisk mikroprocessor har en 10 bit A/D konverter med 5 V måleområde. Spændingsdifferensen mellem de 1024 værdier, som A/D konverteren kan registrere, er derfor $U_{LSB} = (5 \text{ V})/(2^N - 1) = 4,9 \text{ mV}$.

61 En mikrofons signalamplitude er produktet af lydtrykkets amplitude, der måles i pascal, og mikrofonens følsomhed på 5 mV/Pa som en typisk værdi. Et lydtryk på 1 Pa svarer til 94 dB SPL, som er det niveau, der findes umiddelbart udenfor ens mund når der tales højt, men uden at råbe. I en afstand på 1 m fra den talende, eksempelvis ved en tilhørers øre, er lydtrykket faldet med 30 dB.

3.3.4 Inverterende forstærker

Hvis den positive indgang lægges på nul og signalet i stedet føres til den negative indgang opnås en inverterende forstærker. Ud over at invertere signalet, hvilket kan være en vigtig egenskab i sig selv, så er forskellen fra det tidligere trin at forstærkningen kan indstilles frit, der er ikke en nedre grænse af forstærkningen, så trinnet både forstærke eller dæmpe indgangssignalet blot ved valg af de to modstande.



Figur 104 – En inverterende forstærker holder den negative indgang nær ved nul, lidt i retning af en vippe fra en lejeplads. Signalet inverteres og opstillingen kan både forstærke og dæmpe et signal.

Udgangsspændingen U_O er givet ved indgangsspændingen U_I og de to modstande R_1 og R_2 som vist herunder og formlen udledes i eksemplet. Det negative fortegn betyder at udgangssignalet har modsat polaritet af indgangssignalet.

$$U_O = -\frac{R_2}{R_1}U_I \Rightarrow A_N = \frac{U_O}{U_I} = -\frac{R_2}{R_1}$$

Med illustrationens komponentværdier bliver forstærkningen $A_N = -2$.

Forstærkeren søger at holde den negative indgang på det samme potentiale som den positive indgang ($U_N = U_P$) og med den positive indgang på 0 V vil den negative indgang ligge meget nær ved 0 V, hvilket kaldes virtuelt nul (*virtual ground*). En analyse med Kirchhoff's knudelov udnytter at summen af de to strømme i modstandene R_1 og R_2 må give summen nul. Strømmen i R_1 er givet ved indgangsspændingen til U_I/R_1 og strømmen i R_2 beregnes på samme måde ved udgangsspændingen til U_O/R_2 .

$$I_1 + I_2 = 0 \Rightarrow \frac{U_I}{R_1} + \frac{U_O}{R_2} = 0$$

Leddene med U_O og U_I placeres på hver sin side af lighedstegnet.

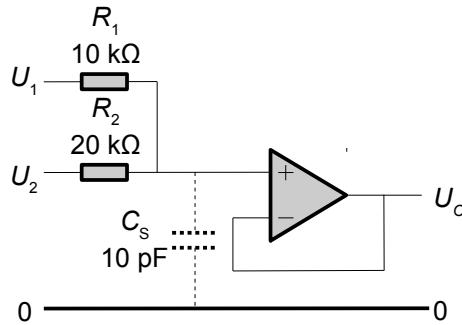
$$\frac{U_O}{R_2} = -\frac{U_I}{R_1} \Rightarrow U_O = -\frac{R_2}{R_1}U_I$$

Den inverterende forstærker har samme begrænsning som den ikke-inverterende med hensyn til at udgangsspændingen skal holdes indenfor effektforsyningens grænser, så ved en effektforsyning på ± 15 V kan der forventes et muligt udgangssignal på maksimalt ± 12 V. Ved en effektforsyning på +5 V og nul skal indgangssignalet være negativt fordi udgangen kun kan leverere en positiv værdi af udgangsspændingen og her er det et krav at forstærkeren er af typen *rail-to-rail*, for ellers bliver signalet blive kraftigt begrænset.

Der er ikke nogen begrænsning for hvor kraftig indgangssignalet må være, for det kan tilpasses ved valget af R_1 og R_2 , men indgangsmodstanden er lig med R_1 og det kan belaste det kredsløb, som driver signalet på indgangen. Forstærkeren er brugbar ved enkelt effektforsyning hvor indgangssignalet ligger lidt under nul, idet moderne, lavspændings operationsforstærkere tillader de to indgange at være -300 mV i forhold til nul. Det er selvfølgelig stadig et krav at udgangsspændingen er over nul og under den positive grænse sat af effektforsyningen.

3.3.5 Addition af signaler

To indgangssignaler kan adderes gennem et modstandsnetværk med R_1 og R_2 , som vist herunder hvor en ikke-inverterende forstærker benyttes som buffer af udgangssignalet. Det er en ofte set konstruktion, hvor den væsentligste anke er at kildeimpedansen fra hver af kilderne U_1 og U_2 skal være meget lille i forhold til modstandsværdien af R_1 og R_2 så det er svært at opnå stor præcision.



Resulterende udtryk:

$$U_o = \frac{R_2}{R_1+R_2} U_1 + \frac{R_1}{R_1+R_2} U_2$$

Grænsefrekvens:

$$f_H = \frac{1}{2\pi(R_1||R_2)C_s} \quad (f_H < GBP)$$

Ind- og udgangsmodstand:

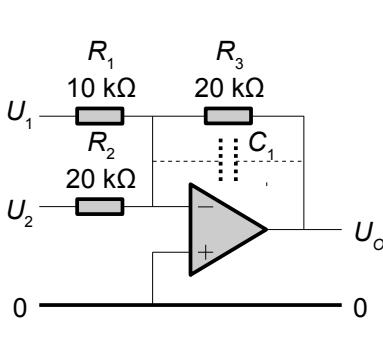
$$\begin{aligned} R_{I1} &\approx R_1 + R_2 & R_O &\approx 0 \\ R_{I2} &\approx R_1 + R_2 \end{aligned}$$

Figur 105 – To indgangssignaler kan adderes gennem modstande og en operationsforstærker kan bruges som buffer af signalet til udgangen.

Grænsefrekvensen bliver 2,4 MHz med de viste komponentværdier og spredningskapaciteten $C_s = 10\text{ pF}$, og det vil ofte være mindre end operationsforstærkerens båndbredde (GBP).

Der er en uundgåelig spredningskapacitet på grund af af komponenterne alle har en relativt kort afstand til nul, og dertil har operationsforstærkeren en vis indgangskapacitet. Det kan modelleres som en kondensator fra den positive indgang til nul på typisk 10 pF og konsekvensen er at den resulterende båndbredde let begrænses på grund af valget af for høje modstandsværdier.

Den inverterende forstærker har et såkaldt virtuelt nul ved den inverterende indgang, så den er velegnet til addition af flere signaler selv om signalets polaritet inverteres. Indgangsspændingerne U_1 og U_2 omdannes til strøm gennem R_1 og R_2 og strømmene balanceres af strømmen gennem tilbagekoblingen R_3 . Signalerne adderes som vist herunder for to indgange og metoden kan udvides til flere indgange efter behov, dog med en begrænsning i den opnåelige båndbredde på grund af operationsforstærkeren (se side 134). Opstillingen er populær når et signals DC niveau skal flyttes så det passer til indgangen på en A/D konverter, der normalt arbejder fra nul og op til effektforsyningen på typisk 5 V. Et AC signal på $\pm 1\text{ V}$ føres til U_1 og U_2 føres til $-2,5\text{ V}$. Signalet svinger nu med $\pm 2\text{ V}$ omkring 2,5 V.



Resulterende udtryk:

$$U_o = -\left(\frac{R_3}{R_1} U_1 + \frac{R_3}{R_2} U_2\right)$$

Uden C_1 : $f_H = \frac{GBP}{A_p}$ hvor $A_p = 1 + \frac{R_3}{R_1||R_2}$

Med C_1 : $f_H = \frac{1}{2\pi R_3 C_1}$ og $f_H < \frac{GBP}{A_p}$

Ind- og udgangsmodstand:

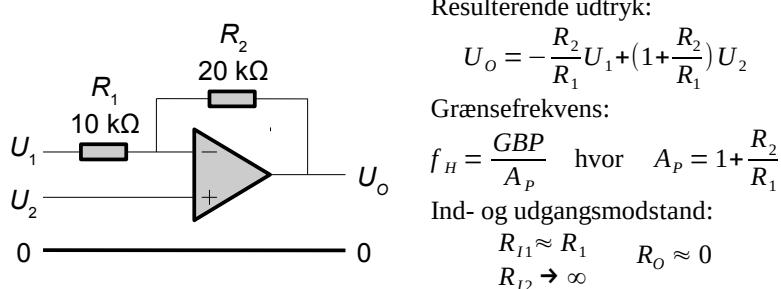
$$\begin{aligned} R_{I1} &\approx R_1 & R_O &\approx 0 \\ R_{I2} &\approx R_2 \end{aligned}$$

Figur 106 – En virtuel nul indgang kan bruges til at addere flere signaler uden gensidig påvirkning fra signalkildernes indre modstande.

Med de viste modstandsværdier bliver forstærkningen -2 fra U_1 til udgangen og -1 fra U_2 til udgangen. Med TL071 er $GBP = 3\text{ MHz}$ og den høje grænsefrekvens bliver $f_H = 750\text{ kHz}$ (-3 dB) hvis C_1 ikke er benyttet og med $C_1 = 100\text{ pF}$ bliver $f_H = 80\text{ kHz}$ (-3 dB).

3.3.6 Subtraktion af signaler

Operationsforstærkerens basale funktion er at trække to indgangssignaler fra hinanden så det vil ligge lige for at benytte den som subtraktor. Der vises herunder hvor signalet U_1 inverteres og forstærkes med $-R_2/R_1$, mens signalet U_2 forstærkes med $1 + R_2/R_1$. Den høje grænsefrekvens er givet ved forstærkerens GBP parameter og den indstillede forstærkning (se side 134).

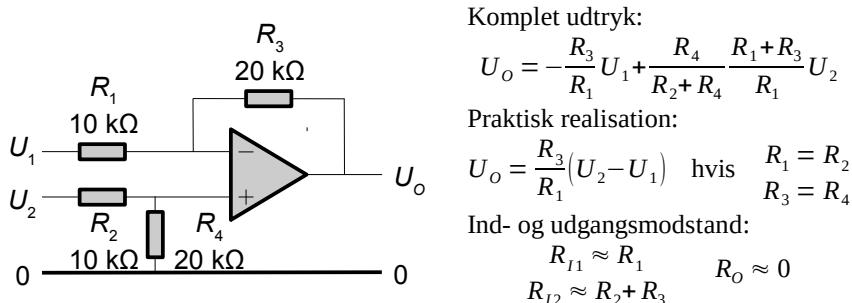


Figur 107 – En operationsforstærker kan bruges til at subtrahere to signaler.

En anvendelse kunne være at forstærke et signal U_2 op samtidigt med at det forskydes i positiv retning med en negativ spænding på U_1 . Hvis modstandene er ens vil et signal på U_2 på ± 1 V og med $-2,5$ V på U_1 vil det betyde at signalet svinger ± 1 V omkring et DC niveau på $2,5$ V.

3.3.7 Differensforstærker

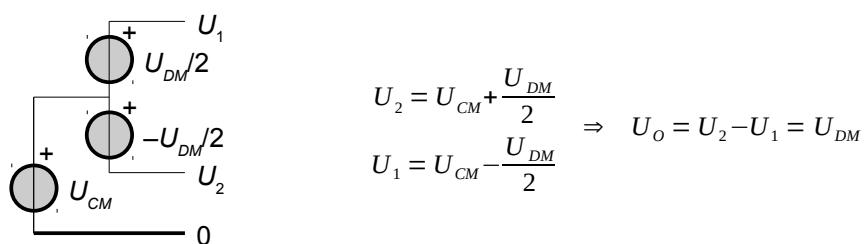
Der er meget ofte brug for at subtrahere to signaler så kun forskellen imellem dem optræder på udgangen. For at opnå denne subtraktionen er det nødvendigt at dæmpe signalet til den positive indgang svarende til forstærkningen, og det kan opnås med en spændingsdeler.



Figur 108 – Hvis modstandene tilpasses korrekt kan udgangens signal være en forstærket version af differensen mellem to indgangssignaler.

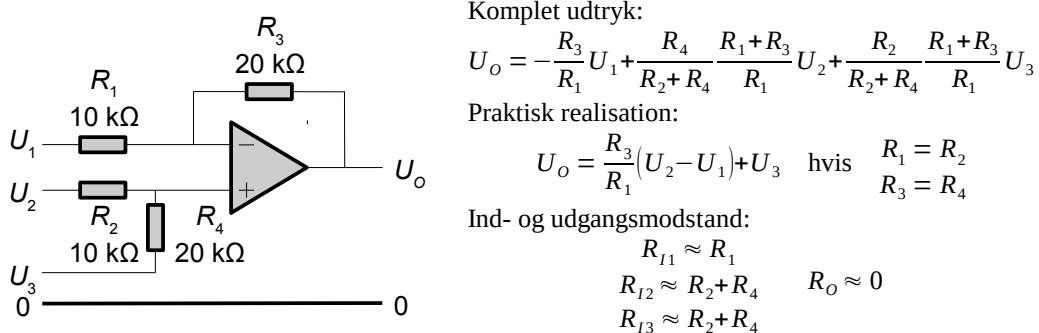
Med fire ens modstandsværdier bliver udgangssignalet en subtraktion med $U_o = U_2 - U_1$ så et eventuelt fælles signal fjernes og kun differensen mellem spændingerne benyttes. Det udnyttes indenfor lydproduktion og medicinsk teknologi til at undgå problemer med 50 Hz netbrum.

Indgangssignalet beskrives ved et nyttesignal U_{DM} (*DM, differential mode*) og et uønsket fælles signal for begge indgange U_{CM} (*CM, common mode*), der kan være DC eller en støjkilde. Udgangssignalet bliver lig med nyttesignalet og det fælles signal undertrykkes.



Operationsforstærker

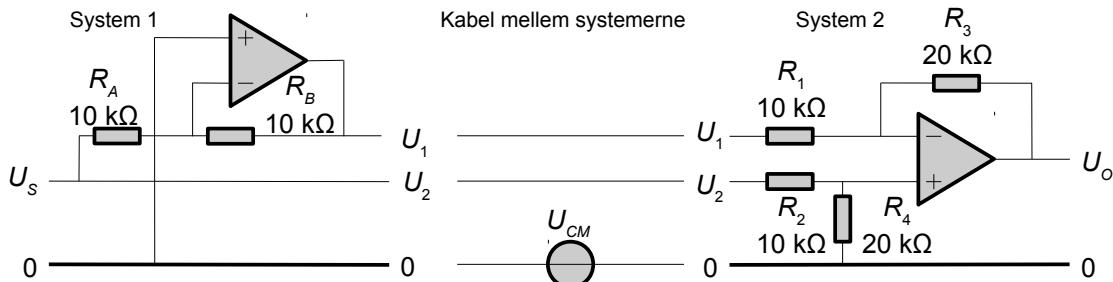
En interessant modifikation er at benytte den nye modstand R_4 som indgang for det giver U_3 som en referenceterminal for udgangen. Det betyder at en differens $U_2 - U_1$ på indgangen kan blive forstærket og flyttet til udgangen som differensen $U_o - U_3$.



Figur 109 – En differensforstærker overfører en differens på indgangen til en ny differens på udgangen overfor en referencespænding.

Med U_3 lagt til nul er der tale om et udgangssignal med nul som reference, og med en spænding på U_3 vil udgangssignalet svinge om denne referenceværdi. Det betyder at den enkle opstilling kan måle en spænding samt benyttes til at overføre et differentielt signal mellem to systemer med hver sin nul-reference⁶² uden at signalet belastes af lysnettets 50 Hz brum⁶³.

For at sende et signal fra system 1 til system 2 benyttes et kredsløb, der omdanner et signal U_s til et såkaldt komplementært udgangssignal hvor U_2 er lig med signalet ($U_2 = U_s$) og U_1 er en invertet udgave af indgangssignalet ($U_1 = -U_s$), samt et andet kredsløb hvor differensen uddrages og relateres til det nye systems nul.



Figur 110 – Analoge og digitale signaler kan overføres mellem to systemer hvor nul-referencen ikke er fælles, ved at dele signalet op i to komplementære signaler og benytte en differensforstærker. Et forstyrrende signal i forbindelsen kan opfattes som en forskel mellem de to nul-referencer og det vil derfor blive undertrykket.

Det er uden betydning om der skal overføres analoge eller digitale signaler⁶⁴ mellem systemerne, for de kan begge ødelægges af en spændingsforskell fra det ene systems nul til det andet systems nul. For at opstillingen kan virke skal spændingsforskellen mellem de to systemers nul være indenfor det lovlige arbejdsmønster for differensforstærkeren, og denne spændingsværdi skal tillade at der samtidigt overføres et signal. For overførsel af digitale 5 V signaler med 5 V spændingsforskell skal differensforstærkeren som minimum benytte +10 V og -5 V effektforsyning.

Kredsløbet anvendes ved lydtransmission i studier for at undgå interferens fra lysnettets 50 Hz og internt i integrerede kredse for at mindske indkobling af digital støj. Ved digital signaltransmission benyttes helt det samme princip ved standarderne RS422 og USB.

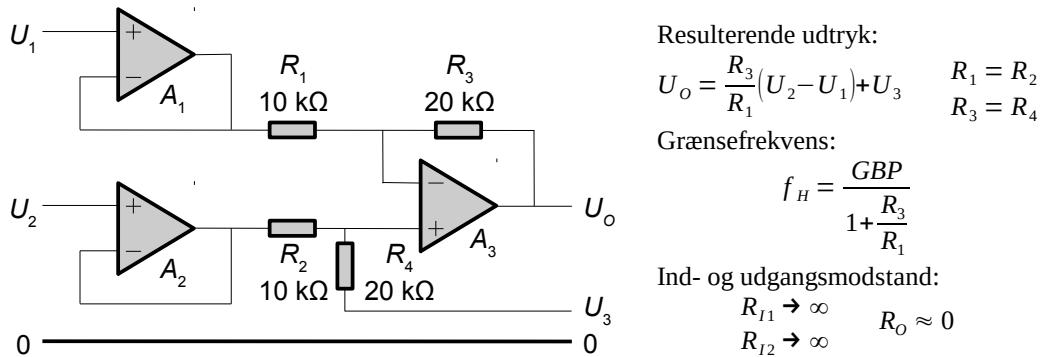
62 Situationen kan opstå hvis to systemer effektforsynes fra hver sin tilslutning til lysnettets 230 V.

63 Undertrykkelse af højfrekvent støj kræver en stor båndbredde af de benyttede operationsforstærkere.

64 Jeg har set det udnyttet i en digital båndoptager (omkring 1987) hvor alle printkort for digital signalbehandling kommunikerede data over en differentiel backplane. Bussen bestod derfor af 16 differentielle sendere og modtagere for en 16 bit ordlængde.

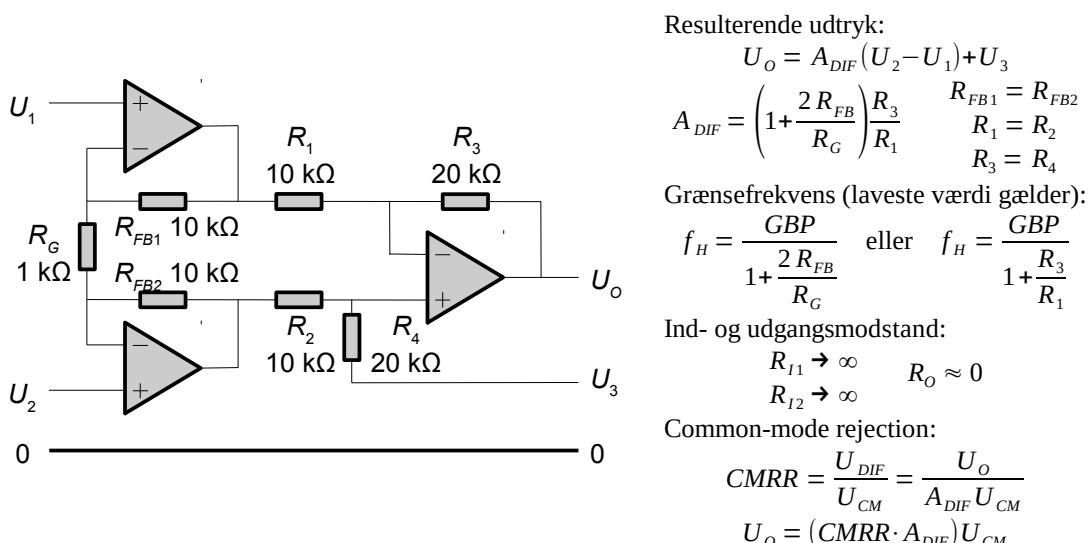
3.3.8 Instrumentationsforstærker

For at opnå en høj og ensartet indgangsmodstand for de to indgange er det normalt at supplere med et par buffere (A_1 og A_2 herunder). En mulighed er TL071-serien med JFET indgang, som giver en indgangsmodstand i mega-ohm niveau ved lave frekvenser. Operationsforstærker A_3 giver den indstillede forstærkning og er den betydnende komponent for opstillingens båndbredde.



Figur 111 – En instrumentationsforstærker med buffere på indgangene for høj indgangsmodstand.

De to buffere kan desuden udnyttes som forstærkere af nyttesignalet og opstillingen kaldes da en instrumentationsforstærker (*instrumentation amplifier*). Den er populær indenfor måleopgaver, hvor fx netbrum skal undertrykkes. Udfordringen er hvor godt komponenterne kan matches, for med blot 0,1 % ubalance vil et fællessignal U_{CM} kun dæmpes 60 dB. De integrerede instrumentationsforstærkere benytter derfor lasertrimmede modstande for at opnå en god dæmpning af fællessignalet. Det er også vigtigt at forbindelsen til U_3 har en indre modstand nær nul; for det påvirker balancen, og ved en værdi af R_4 på 20 k Ω skal den eksterne modstands værdi være under 20 Ω .



Figur 112 – En klassisk instrumentationsforstærker med tre operationsforstærkere, som her er vist med en indstillet forstærkning på 42 gange.

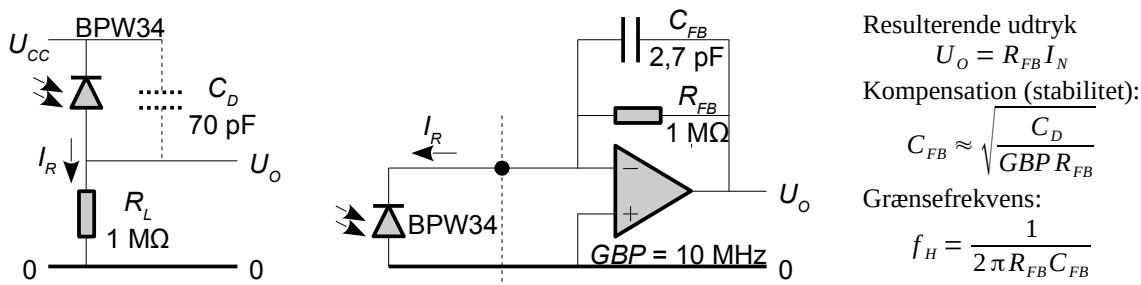
Undertrykkelse af forstyrrelsen U_{CM} oplyses som CMRR (Common-Mode Rejection Ratio), der er funktion af frekvensen og har den valgte forstærkning som parameter. Definitionen relaterer indgangssignalet $U_2 - U_1$ til forstyrrelsen U_{CM} , og det har den konsekvens at kurver over CMRR ser bedst ud ved høj forstærkning. Det er dog udgangens signal U_O , der føres videre i systemet, og en forøgelse af forstærkningen overfor indgangssignalet A_{DIF} vil derfor ikke ændre på udgangsspændingen som følge af forstyrrelsen U_{CM} .

3.4 Strømforstærkning

En del komponenter arbejder med strøm frem for spænding. En fotodiode omsætter eksempelvis lysets fotoner til en strøm af elektroner og sensorer baseret på piezoelektricitet kan også opfattes som strømkilder. Laserdioder kræver en styring af DC strømmen i dioden, og for at få en konstant amplitude af strømmen i en spole skal den strømstyrtes hvis frekvensen varieres.

3.4.1 Fotodiode

Strømmen fra en fotodiode kan omsættes til en spænding ved at bruge en modstand, men fordelen ved at inkludere en operationsforstærker er at diodens indre kapacitet kun i ringe grad påvirker den opnåelige båndbredde. Med de viste komponentværdier vil opstillingen til højre være 30 gange hurtigere end uden operationsforstærkeren.



Figur 113 – En fotodiode er en lysstyret strømkilde og operationsforstærkeren omsætter strømmen til en spænding uden at båndbredden bliver begrænset af diodens parallelkapacitet.

Strømmen dannes i diodens PN overgang og løber fra katode til anode; det er modsat normal strømretning. Fotodioder af silicium har typisk en omsætning på 0,6 A/W. Ved fuldt solskin er der cirka 1 kW/m² ved jordoverfladen og det lysfølsomme areal for fotodioden BPW34 er på 7,5 mm² så dioden modtager 7,5 mW og det giver en strøm på $I_R = 4,5 \text{ mA}$.

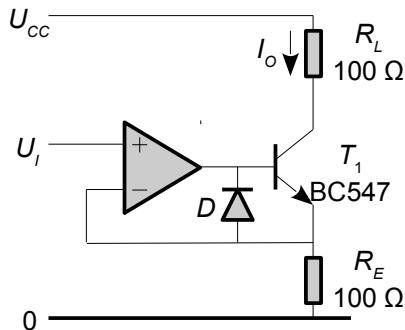
Ved en dæmpet belysning på 1 W/m², der er rigeligt til at læse en avis, giver dioden BPW34 en strøm på 4,5 μA og med en modstand på $R_L = 1 \text{ M}\Omega$ bliver udgangsspændingen 4,5 V. Diodens indre kapacitet er på 70 pF og det vil begrænse båndbredden til 2 kHz så kredsløbet er kun egnet som lysmåler og ikke til fx datatransmission. Med en operationsforstærker dirigeres strømmen gennem tilbagekoblingen R_{FB} , og den lave modstands værdi ved indgangens virtuelle nul kortslutter dioden så kapaciteten ikke påvirker båndbredden. En eksternt parallelkondensator C_{FB} må dog indskydes for stabilitet, for diodens kapacitet danner en pol med tilbagekoblingen (se side 137). Med en OpAmp med $GBP = 10 \text{ MHz}$ beregnes $C_{FB} = 2,7 \text{ pF}$ og den nye grænsefrekvens $f_H = 60 \text{ kHz}$ er en forbedring på 30 gange.

3.4.2 DC strømkilde

En DC strømkilde kan dannes med en transistor, der styres af en operationsforstærker. For at holde operationsforstærkeren i balance vil dens udgang tvinge basis af transistoren til den værdi, hvor spændingen over R_E er lig med indgangens U_I og derved er strømmen givet af Ohms lov. Indgangsspændingen kommer fra et eksternt kredsløb og kan eventuelt være styret af en mikroprocessor, så strømkilden er kan tændes og slukkes fra en logisk udgang, eller fra en D/A konverter, så den kan styres analogt. Det er muligt at benytte en PNP transistor for en strøm fra den positive effektforsyning til nul, men styresignalet ligger nu oppe ved effektforsyningen.

Ved høj strømstyrke skal man være opmærksom på at der afsættes effekt i transistoren, givet ved strøm gange spænding, og spændingen over transistoren er i illustrationen noget konservativt sat lig med effektforsyningens U_{CC} uanset spændingsfaldet over R_L og R_E . For den viste transistor er begrænsningen $P_{C MAX} = 300 \text{ mW}$ så strømmen bør holdes under 20 mA ved $U_{CC} = 15 \text{ V}$. Man står sig ved at holde sig under det maksimalt tilladelige, og jeg ville holde mig et godt stykke under en effekt på 100 mW ved en småsignaltransistor som BC547 for ikke at varme den for meget op.

Operationsforstærker



Udgangsstrøm:

$$I_o = \frac{U_i}{R_E} \quad \text{og} \quad U_i > 0,5 \text{ V}$$

Anbefalet strømområde:

$$1 \mu\text{A} < I_o < \frac{P_{C\ MAX}}{U_{CC}}$$

Anbefalet strømforsærkning:

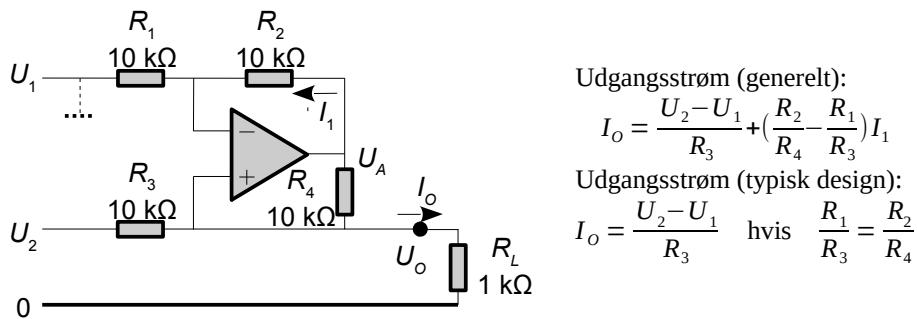
$$\beta > 100$$

Figur 114 – En transistor kan styres af en operationsforstærker til en præcis DC strømkilde, der kan styres af en spænding fra et eksternt kredsløb, som fx en mikroprocessor. Dioden beskytter transistorens basis-emitter strækning hvis operationsforstærkeren har en negativ effektforsyning.

Modstande har normalt $\pm 1\%$ nøjagtighed og kan købes til $\pm 0,1\%$ så tolerancen er ikke et problem. Derimod vil operationsforstærkeren introducere en usikkerhed i spændingen over modstanden på typisk $\pm 5 \text{ mV}$, så ved $U_i = 500 \text{ mV}$ er fejlen på $\pm 1\%$ (se side 129). For at reducere fejlkilden skal der vælges en OpAmp med omhu eller U_i skal holdes et passende stykke over nul og det leder så til anbefalingerne i illustrationen om en indgangsspænding på mindst $0,5 \text{ V}$. Transistorens strøm i kollektor er mindre end strømmen i emitter på grund af strømmen i basis, så strømforsærkningen bør være over 100. Ved meget lav strøm kan man være plaget af lækstrøm i transistoren og biasstrøm i operationsforstærkeren (se side 131). For at holde fejlen lav bør der ikke designes efter mindre end $1 \mu\text{A}$ ved 25°C eller $100 \mu\text{A}$ ved 100°C , med mindre databladet konsulteres først.

3.4.3 AC strømkilde

Der er ofte brug for en strømkilde, der kan leve strøm til en belastning med den ene terminal til nul, og det viste kredsløb kan tilmed leve strøm af begge polariteter. Opstillingen kendes under betegnelsen en Howland strømkilde og figurerer ofte i de populære magasiner og på internettet.



Udgangsstrøm (generelt):

$$I_o = \frac{U_2 - U_1}{R_3} + \left(\frac{R_2}{R_4} - \frac{R_1}{R_3} \right) I_1$$

Udgangsstrøm (typisk design):

$$I_o = \frac{U_2 - U_1}{R_3} \quad \text{hvis} \quad \frac{R_1}{R_3} = \frac{R_2}{R_4}$$

Figur 115 – En strømgenerator, der kan benyttes både ved AC og DC, men der stilles ret store krav til komponenternes præcision.

Funktionen ses ved at antage $U_1 = 0$ og $U_2 = 1 \text{ V}$ hvilket skal give en udgangsstrøm på $I_o = 100 \mu\text{A}$. Med en belastning på $R_L = 1 \text{ k}\Omega$ bliver udgangsspændingen $U_o = 0,1 \text{ V}$ så strømmen i R_3 bliver $90 \mu\text{A}$. Det vil løbe en strøm i R_1 på $I_1 = 10 \mu\text{A}$ så operationsforstærkerens udgang indstiller sig på $U_A = 0,2 \text{ V}$ og det giver en strøm i R_4 på $10 \mu\text{A}$, der sammen med de $90 \mu\text{A}$ fra R_3 giver udgangsstrømmen.

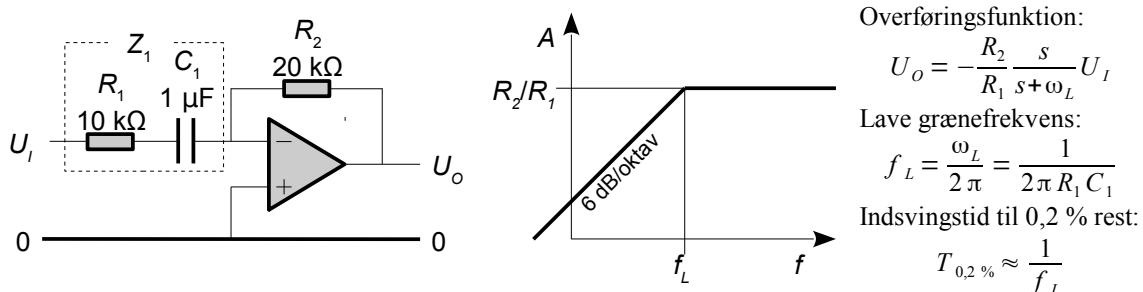
Operationsforstærkerens udgang U_A indstilles på to gange U_2 (med $U_1 = 0$), så der kan let opstå overstyring af operationsforstærkeren. En fejl i modstandene på $\pm 1\%$ vil kunne påvirke udgangsstrømmen med $\pm 4\%$ af den interne strøm, der med de viste komponentværdier er $I_1 = 250 \mu\text{A}$ ved 5 V over belastningsmodstanden. Det svarer til $\pm 10 \mu\text{A}$ eller en relativ fejl på $\pm 10\%$, så der er ikke tale om en præcisionsopstilling uden en justering af modstandsforholdene.

3.5 Filtrerende kredsløb

Operationsforstærkeren kan benyttes for realisering af de matematiske funktioner differentiation og integration, som hyppigt benyttes ved interface til sensorer, og i musikanlægget er korrektion af amplituderesponsen en vigtig del af lydoplevelsen. I dette afsnit vises hvordan disse funktioner realiseres. Dertil kommer behovet for at blokere for et DC niveau undervejs i en signalkæde eller dæmpe en højfrekvent interferens fra fx en mobiltelefon.

3.5.1 Højpasled

Den inverterende forstærker på side 111 kan forstærke et AC signal og samtidigt blokere for en DC spænding ved at indskyde en kondensator i serie med modstanden R_1 . Kondensatoren vil lades op til DC niveauet og efter cirka 6 tidskonstanter er udgangen faldet til ro og kun AC signalet passerer. Det giver en indsvingstid med en rest på 0,2 % af det oprindelige DC niveau.



Figur 116 – En kondensator blokkerer DC og ved tilstrækkeligt høje frekvenser er forstærkningen givet af modstandsforholdet (højpasfilter).

Med de viste værdier er tidskonstanten $R_1C_1 = 10\text{ ms}$ så den lave grænsefrekvens er $f_L = 16\text{ Hz}$ og opstillingen er i ro efter $T_\infty = 63\text{ ms}$. Forstærkningen af signaler noget over grænsefrekvensen er -2 .

Overføringsfunktionen ved at indsætte impedansen af serieforbindelsen af R_1 og C_1 i stedet for R_1 i den oprindelige ligning. Impedansen af en kondensator er $1/sC$, og det viste udtryk dannes ved at gange med s og dividere med R_1 i tæller og nævner. Forstærkningen beskrives som før ved faktor $-R_2/R_1$ og relationen er udvidet med et højpasled.

$$U_O = -\frac{R_2}{Z_1}U_I = -\frac{R_2}{R_1 + \frac{1}{sC_1}}U_I = -\frac{R_2}{R_1} \frac{s}{s + \omega_L}U_I \quad \text{hvor} \quad \omega_L = \frac{1}{R_1 C_1}$$

Hvis DC niveauet pludselig ændres vil der optræde et kortvarigt spændingsspring på udgangen, som derefter klinger ud med tidskonstanten R_1C_1 . Det vises herunder at et pludseligt skift i DC niveauet på U_{DC} vil blive forstærket med forstærkningen i den inverterende forstærker på $-R_2/R_1$ og at det derefter klinge ud med tidskonstanten R_1C_1 . Det er uden betydning om der samtidigt er et AC signal på indgangen, for det behandles som før, men udgangen bliver summen af de to signaler.

Til tidspunktet $t = 0$ påtrykkes det konstante niveau U_{DC} så indgangssignalet bliver U_{DC}/s efter Laplace-transformationen, og det udtryk ganges på overføringsfunktionen.

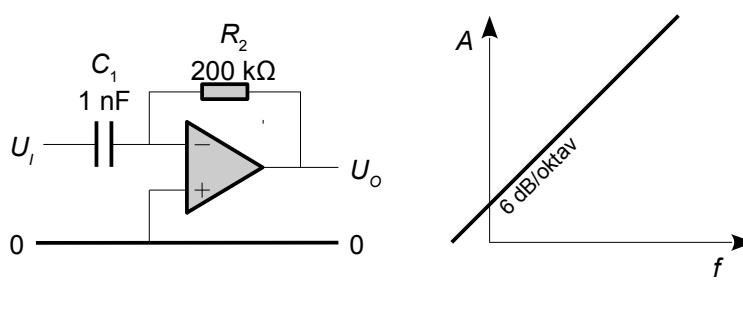
$$U_O = -\frac{R_2}{R_1} \frac{s}{s + \omega_L} \frac{U_{DC}}{s} = -\frac{R_2}{R_1} \frac{1}{s + \omega_L} U_{DC}$$

Første led er forstærkningen i trinnet og viser at en pludseligt tilsluttet spænding U_{DC} vil give en momentant stigende spænding fra udgangen på $-(R_2/R_1)U_{DC}$. Andet led kan findes i tabellen over Laplace transformationen som en eksponentialfunktion.

$$u_O(t) = -\frac{R_2}{R_1} \exp(-\omega_L t) U_{DC}$$

3.5.2 Differentiation

Hvis modstanden R_1 fjernes helt vil funktionen give en differentiation af indgangssignalet, og det er en af de klassiske analoge funktioner i den analoge computer fra tiden før 1950, dog bygget med radiorør og ikke med en halvlederbaseret operationsforstærker. I moderne elektronik mødes differentiationen ved interface til sensorer hvis signalamplitude aftager med stigende frekvens⁶⁵. For kredsløbet herunder differentieres signalet, så udgangens amplitude er proportional både med indgangssignalets amplitude og med frekvensen.



Overføringsfunktion (ideelt):

$$U_O = -s C_1 R_2 U_I$$

$$u_O = -C_1 R_2 \frac{du_I}{dt}$$

Overføringsfunktion ved en dominérende pol på ω_0 :

$$U_O = -s C_1 R_2 \frac{\omega_0}{\omega_0 + s} U_I$$

Ind- og udgangsimpedans:

$$Z_I \approx \frac{1}{s C_1} \quad Z_O \approx 0$$

Figur 117 – En differentierende opstilling hvor forstærkningen er proportional med frekvensen indtil operationsforstærkeren ikke længere har tilstrækkelig båndbredde.

Med de viste komponenter er kondensatorenens impedans $X_C = 1/2\pi f C_1 = 200 \text{ k}\Omega$ ved frekvensen 159 kHz. Ved denne frekvens er forstærkningen -1 og den aftager til $-0,001$ ved 159 Hz. Med et indgangssignal i form af en sinus vil udgangssignalet være en cosinus med samme frekvens og en ændret amplitude.

Problemet ved et differentierende trin er at kredsløbets forstærkning stiger med frekvensen, så en højfrekvent støj trænger igennem med fuld amplitude og vil let kunne genere funktionen. Den er derfor sjælden gæst i hovedparten af moderne elektronik, og man bruger den kun hvis påkrævet. Differentiationen stiller store krav til operationsforstærkerens båndbredde for dens åben-sløjfe forstærkning aftager ved stigende frekvens (se side 134). Endelig er der et problem hvis signalet til det differentierende trin kommer fra en lav impedans, for så kan en kraftig og højfrekvent amplitude ødelægge forstærkeren.

Det kan virke underligt at kondensatoren kan differentiere når den i det teoretiske afsnit beskrives ved en integration, men skyldes at spændingen over kondensatoren udtrykkes som integralet af strømmen, mens det er strømmen, der her skal benyttes som funktion af spændingen. Relationen mellem strøm og spænding findes ved at differentiere ligningen for kondensatoren, og resultatet bliver at strømmens amplitude er proportional med den hastighed amplituden af spændingen varierer med; det vil sige at strømmen vokser med frekvensen.

$$u_C = \frac{1}{C} \int i dt \Rightarrow \frac{du_C}{dt} = \frac{1}{C} i \Rightarrow i = C \frac{du_C}{dt}$$

Spændingen over kondensatoren er identisk med indgangssignalet u_I og strømmen løber gennem modstanden R_2 med positiv retning ind mod det virtuelle nul så udgangsspænding bliver.

$$u_O = -i R_2 = -C_1 R_2 \frac{du_C}{dt} = -C_1 R_2 \frac{du_I}{dt}$$

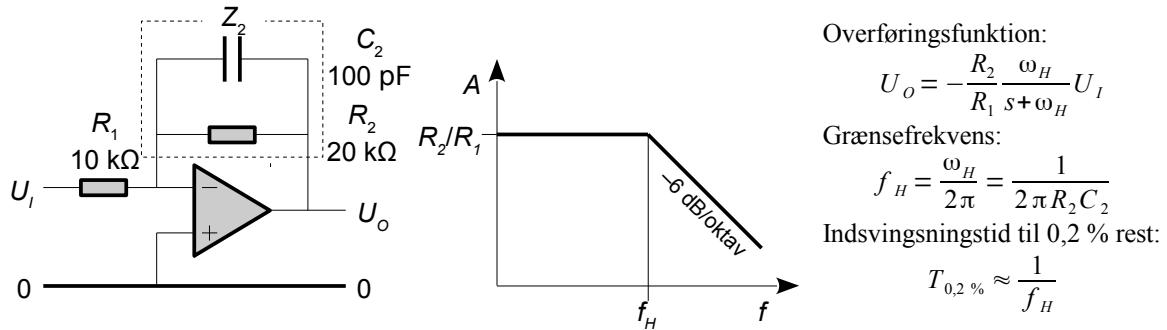
Udtrykt i Laplace notationen erstattes en differentiation af s ganget med den transformerede af indgangsspændingen, hvilket er $U(s)$.

$$U_O = -s C_1 R_2 U_I$$

⁶⁵ For en elektrodynamisk højttaler er lydtrykket inde i kabinet funktion af membranens udsving, mens lydtrykket foran højttaleren er relateret til accelerationen af membranen. Her skal der differentieres to gange for at bestemme lydtrykket.

3.5.3 Lavpasled

Med en kondensator i parallel med modstanden i tilbagekoblingen af en inverterende forstærker vil de høje frekvenser dæmpes. Det danner et lavpasfilter, og er en af de oftest anvendte koblinger både som en tilsigtet og utilsigtet begrænsning af båndbredden. Det er i mange anvendelser god latin at begrænse båndbredden for at undgå problemer med indstråling af højfrekvens fra en mobiltelefon, eller en utilsigtet modtagelse af radiotransmissioner. Amplituden er 3 dB nede ved grænsefrekvensen så ved fx tre seriekoblede forstærkere i signalkæden lægges grænsefrekvensen en oktav oppe, så amplituden kun dæmpes 1 dB per trin ved den ønskede høje grænsefrekvens.



Figur 118 – En inverterende forstærker arrangeret til at dæmpe høje frekvenser (lavpasfilter).

Med de viste komponentværdier dæmpes frekvenser med 3 dB ved 80 kHz og dæmpningen vil ved højere frekvenser gå imod $-6 \text{ dB}/\text{oktav}$, eller $-20 \text{ dB}/\text{dekade}$, som betyder at amplituden er omvendt proportional med frekvensen. Signaler under grænsefrekvensen vil passere med den programmerede forstærkning.

Overføringsfunktionen bestemmes ved at indsætte parallellforbindelsen af modstanden R_2 og kondensatoren C_2 i stedet for R_2 i formlen for forstærkningen ved den inverterende forstærker.

$$U_O = -\frac{Z_2}{R_1} U_I = -\frac{Z_R \| Z_C}{R_1} U_I = -\frac{R_2(1/sC_2)}{R_1(R+1/sC_2)} U_I = -\frac{R_2}{R_1} \frac{1}{1+sR_2C_2} U_I$$

Grænsefrekvensen er givet ved den frekvens hvor impedansen af kondensatoren er numerisk lig med værdien af modstanden.

$$R_2 = \left| \frac{1}{sC_2} \right| \Rightarrow R_2 = \frac{1}{\omega_H C_2} \Rightarrow f_H = \frac{\omega_H}{2\pi} = \frac{1}{2\pi R_2 C_2}$$

3.5.4 Integration

Ved en meget stor værdi af modstanden R_2 flytter grænsefrekvensen nedad så kondensatoren er aktiv i hovedparten af frekvensområdet. I det ekstreme tilfælde hvor modstanden overhovedet ikke er der, ændres forstærkeren til ren integration, hvor signalet aftager med $-6 \text{ dB}/\text{oktav}$. Det er dog en sandhed med en modifikation, for operationsforstærkerens åben-sløjfe forstærkning begrænser integrationen ved meget lave frekvenser.

Overføringsfunktionen bliver:

$$U_O = -\frac{1}{R_1/R_2 + sR_1C_2} U_I \xrightarrow{s \rightarrow \infty} -\frac{1}{sR_1C_2} U_I \Rightarrow u_o = -\frac{1}{R_1C_2} \int u_I dt$$

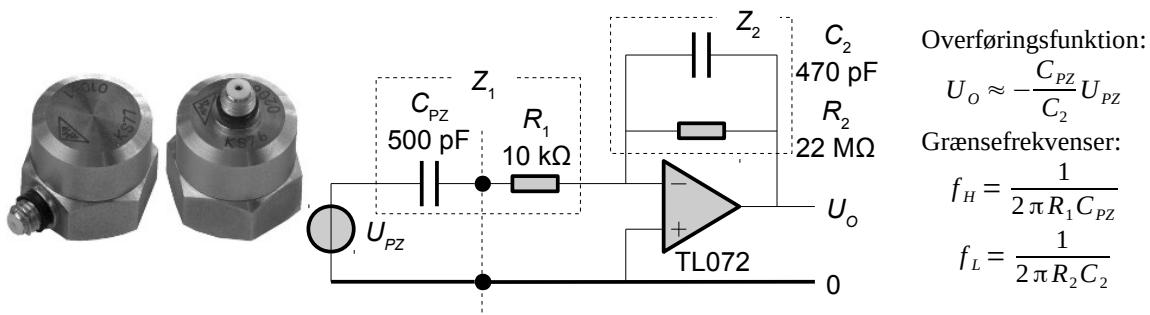
I praksis er modstanden R_2 påkrævet for at stabilisere forstærkeren overfor DC drift med mindre et eksternt kredsløb sikrer arbejdspunktet, for en nok så lille strøm i den negative indgang integreres op så udgangen glider imod enten den positive eller den negative forsyning hvor forstærkeren ikke længere kan fungere korrekt. I den analoge computer blev en kontakt brugt for at nulstille trinnet.

3.5.5 Piezo-elektrisk accelerometer

Et piezo-elektrisk krystal har den egenskab at det genererer en spænding når de udsættes for en mekanisk belastning. Spændingen kan være ganske betydelig, som for eksempel i en lighter hvor krystallet benyttes til at danne en gnist, men som regel er signalets amplitude mere moderat. De piezo-elektriske krystaller fandt anvendelse i ældre grammofoners pick-up hvor de omsatte rillens bevægelser til elektriske spændinger. De finder desuden anvendelse i visse diskanthøjtalere hvor den modsatte effekt udnyttes; nemlig at krystallet bøjes når det påtrykkes en spænding. Indenfor industrien anvendes de til måling af acceleration, hvor en masse bæres af krystallet, som derfor vil bøjes hvis sensoren bevæges.

Et krystal kan sjældent lede en elektrisk strøm, så det overrasker næppe at det modelleres som en kondensator dannet af de to tilslutninger og værdien er i pF området. For at måle spændingen fra det piezo-elektriske krystal kan man gå to veje. Den ene er at bygge en forstærker med en meget høj indgangsmodstand, og den anden vej er at belaste krystallet så hårdt at det er strømmen der måles. Det sidste kaldes for en ladningsforstærker (*charge amplifier*).

Det piezo-elektriske krystal kobles til en virtuel-nul indgang, hvilket effektivt kortslutter krystallet så strømmen bestemmes af krystallets spænding og impedansen af dets kondensator. Der dannes et differentierende led på grund af kondensatoren C_{PZ} og for at måle spændingen U_{PZ} fra krystallet er det nødvendigt at integrere signalet for at kompensere for differentiationen.



Figur 119 – Et piezo-elektrisk accelerometer kan være udformet så det er let at montere på en maskine for måling af vibration. Transduceren fungerer som en spændingsgenerator U_{PZ} i serie med en kondensator C_{PZ} , så ved belastning med en virtuel-nul indgang bliver signalet differentieret og forstærkeren må derfor integrere signalet for at genskabe transducerens spænding.

Med de viste komponenter er forstærkningen tæt på én, så følsomhed på 1 mV/ms^{-2} ændres ikke.

Et eksempel skal gives for interface til en sensor med 10 mV udgangsspænding for en acceleration på $g = 9,82 \text{ m/s}^2$. Med SI-enheder er følsomheden S for sensoren cirka 1 mV/ms^{-2} .

Overføringsfunktionen beregnes ud fra impedansen Z_1 af serieforbindelsen af det piezo-elektriske krystals kapacitet C_{PZ} og indgangsmodstanden R_1 på forstærkeren, samt af impedansen Z_2 for parallelforbindelsen af tilbagekoblingen ved R_2 og C_2 .

$$U_o = -\frac{Z_2}{Z_1} U_{PZ} \quad \text{hvor} \quad Z_2 = \frac{R_2(1/sC_2)}{R_2 + 1/sC_2}$$

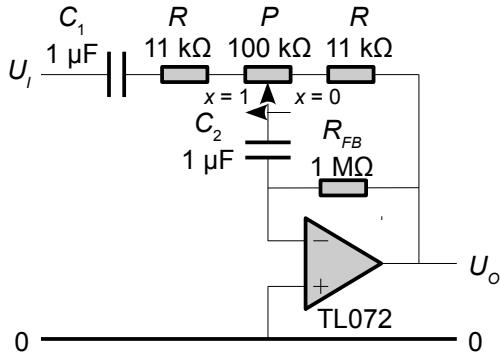
$$Z_1 = R_1 + 1/sC_{PZ}$$

Resultatet er et passivt båndpasfilter med grænsefrekvenserne $f_L = 15 \text{ Hz}$ og $f_H = 32 \text{ kHz}$. Indenfor dette område er udgangssignalet uafhængigt af frekvensen.

$$U_o = -\frac{sR_2 C_{PZ}}{(1+sR_1 C_{PZ})(1+sR_2 C_2)} U_{PZ} = -\frac{C_{PZ}}{C_2} U_{PZ} \quad \text{for} \quad \frac{1}{2\pi R_2 C_2} < f < \frac{1}{2\pi R_1 C_{PZ}}$$

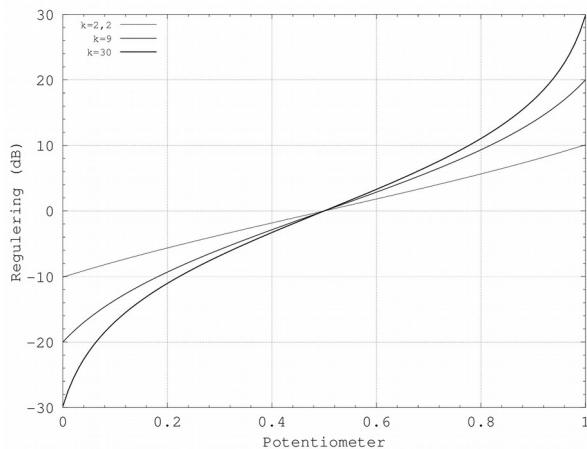
3.5.6 Variabel forstærkning

Den inverterende forstærkers forstærkning er givet ved forholdet mellem to modstande, og ved at indskyde et potentiometer kan forstærkningen varieres kontinuerligt. Reguleringen er tilnærmet logaritmisk fra -20 dB til 20 dB hvilket giver en behagelig oplevelse af drejningen på et potentiometer, der udmærket kan være digitalt med styring fra en mikroprocessor.



$$A = -\frac{xk+1}{(1-x)k+1}$$

$$k = P/R$$

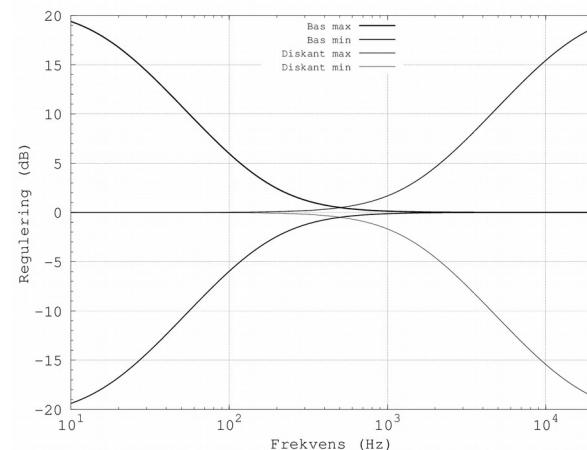
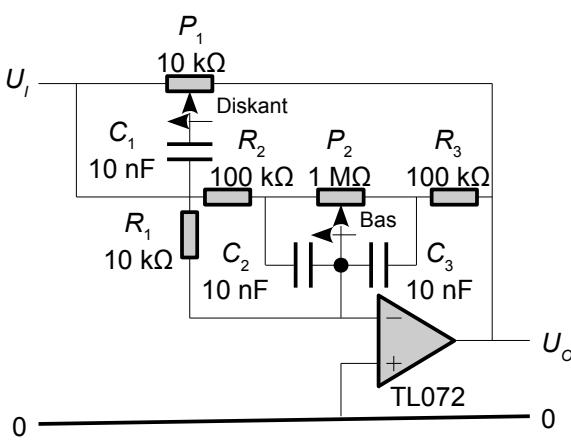


Figur 120 – En variabel forstærkning som funktion af et potentiometers drejning.

Kondensator C_1 og C_2 er DC blokering for at undgå en knasende lyd når potentiometeret drejes, og modstand R_{FB} giver operationsforstærkeren sin biasstrøm. Med potentiometeret i midten er de to grene ens og forstærkningen er 0 dB på nær belastningen fra R_{FB} , der giver $0,5$ dB fejl. Med potentiometret i $x = 1$ er forstærkningen 20 dB med $10R = P + R$, og den er -20 dB ved $x = 0$. Ved $x = 0,25$ er forstærkningen på $0,419$ ($-7,6$ dB) og ikke de forventede $0,316$ (-10 dB) så fejlen er her på $2,4$ dB. på $0,5$ dB.

3.5.7 Toneregulering

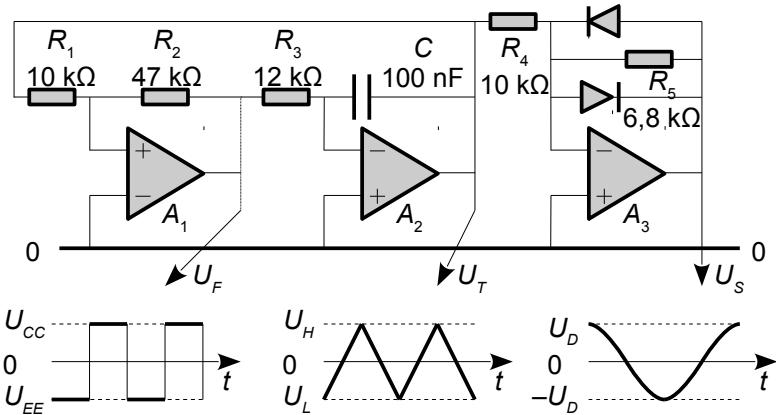
En variant over principippet giver separate bas- og diskantreguleringer. Ved høje frekvenser er P_2 kortsluttet af C_2 og C_3 så R_2 og R_3 indstiller forstærkningen til 0 dB ved lave frekvenser, mens de høje frekvenser varieres ved at R_1 og C_1 lægges i parallel over R_2 for at øge niveauet, eller R_3 for at sænke niveauet. Ved lave frekvenser er P_1 ikke aktiv på grund af den høje impedans af C_1 og P_2 indstiller niveauet lave frekvenser. Grænsefrekvensen for bassen er på 160 Hz, som indstilles ved $C_2 = C_3$ og $R_2 = R_3$, og flyttes nedad ved at øge C_2 . Grænsefrekvensen for diskanten indstilles ved R_1 og C_1 og til $1,6$ kHz og flyttes opad ved at mindske C_1 .



Figur 121 – En toneregulering med separat potentiometer for bas og diskant. En forbedring opnås ved at inkludere DC blokering fra styrkereguleringen.

3.6 Tonegeneratører

Mikroprocessoren har lettet genereringen af repetitive signaler, men der kan alligevel opstå behov for at danne svingninger udelukkende med analoge komponenter. Herunder vises hvordan firkant, savtak og sinus kan dannes med få komponenter.



Oscillatorfrekvens:

$$f_{osc} = \frac{R_2}{R_1 4 R_3 C}$$

Høje grænsefrekvens:

$$f_{osc} \ll \frac{SR_1}{2(U_{cc} - U_{ee})}$$

Spændingsamplitude:

$$U_H = -\frac{R_1}{R_2} U_{ee}$$

$$U_L = -\frac{R_1}{R_2} U_{cc}$$

$$U_D \approx 0,7 \text{ V}$$

Figur 122 – En tonegenerator bygget over en Schmitt-trigger, der med de viste komponentværdier vil svinge på 1 kHz med $\pm 5 \text{ V}$ firkant, $\pm 1 \text{ V}$ triangulært og $\pm 0,6 \text{ V}$ sinus ved en $\pm 5 \text{ V}$ effektforsyning.

Forstærker A_1 er en såkaldt Schmitt-trigger, der har to stabile tilstande; med udgangen høj eller lav (U_F på enten U_{cc} eller U_{ee}), og hvor tilstanden skifter som følge af niveauet fra udgangen af A_2 (U_T). Med udgangen lav skal den positive indgang på A_1 hæves over nul for at udgangen kan skifte høj, og det kræver at U_T når op til $U_H = 1 \text{ V}$ med de viste modstande og $\pm 5 \text{ V}$ fra effektforsyningen.

Forstærker A_2 er integrerende og med lavt niveau fra Schmitt-triggeren ($U_F = U_{ee}$) stiger udgangen med en hastighed givet af strømmen gennem R_3 og kondensatoren C . Når niveauet kommer over den værdi hvor Schmitt-triggerens udgang skifter høj (U_H) ændres funktionen så U_T integreres nedad mod den lave grænse (U_L). Det betyder at udgangen oscillerer mellem U_L og U_H med den indstillede frekvens f_{osc} .

Forstærker A_3 bruger dioder af typen 1N4148 til at begrænse trekantsvingningens spidsværdi så den tilnærmer en sinus, der med de viste komponentværdier giver en forvrængning på cirka 5 %. En diode af silicium begynder at lede ved $\pm 0,5 \text{ V}$ og strømmen stiger stærkt ved højere spænding over dioden, og for at virke som tilsiget skal R_4 indstilles til et udgangssignal på $\pm 0,7 \text{ V}$ uden dioderne. Hvis der ikke er behov for en sinus kan A_3 og de tilhørende komponenter blot fjernes.

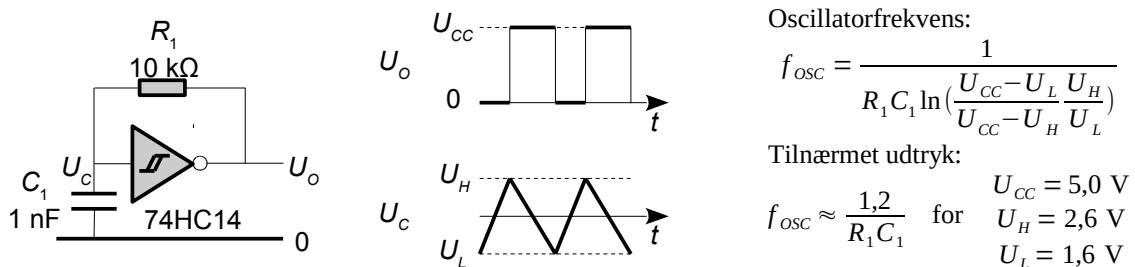
Den viste formel for oscillatorens frekvens gælder for operationsforstærkere af typen rail-to-rail for A_1 , hvor udgangen skifter næsten helt ud til effektforsyningens grænser, fx TLC274. Hvis det ikke er tilfældet vil strømmen gennem R_3 ikke nå op på det forudsatte og frekvensen bliver lavere.

Schmitt-triggeren behøver en vis tid for at skifte mellem yderstillingerne, og det giver en øvre grænse for den frekvens oscillatoren kan arbejde ved. Grænsen findes ud fra parameteren SR for den benyttede operationsforstærker for A_1 . Betegnelsen står for *slew-rate*, der angiver hvor hurtigt udgangen kan ændre sig. En typisk specifikation for en prisbillig OpAmp er omkring $1 \text{ V}/\mu\text{s}$, så den behøver $10 \mu\text{s}$ for at skifte fra den lave værdi til den høje værdi ved $\pm 5 \text{ V}$ effektforsyning. En hel periode behøver derfor $20 \mu\text{s}$ så en oscillationsfrekvens over 50 kHz er helt umulig, og hvis den acceptable fejl sættes til 10 % betyder det at perioden skal være mindst $200 \mu\text{s}$, som svarer til en øvre grænsefrekvens på 5 kHz.

En forbedring kan opnås ved at benytte en speciel OpAmp kaldet en *comparator* for A_1 . Den skifter meget hurtigt og med eksempelvis LM393 vil det tillade oscillatoren at arbejde på 100 kHz. For en højere frekvens er det desuden nødvendigt at finde en hurtigere operationsforstærker.

3.6.1 CMOS oscillator

Oscillatorer for høje frekvenser kræver meget hurtige operationsforstærkere og det er lettest og billigst at realisere ved brug af digitale kredse som fx den populære Schmitt-inverter 74HC14, der er bygget i CMOS teknologi (se senere i bogen). Komponenten stiller udgangen lav når indgangen kommer over den høje tærskelværdi ($U_H = 2,6$ V) og den skifter først høj når indgangen kommer under den lave tærskelværdi ($U_L = 1,6$ V). Med R_1 og C_1 dannes en oscillator, som ramper op og ned mellem de to tærskelværdier på samme måde som den tidligere opstilling. Resultatet er en firkant fra udgangen, men da tærskelværdierne ikke ligger symmetrisk omkring halv forsyning er pulsens høje og lave tider ikke lige lange.

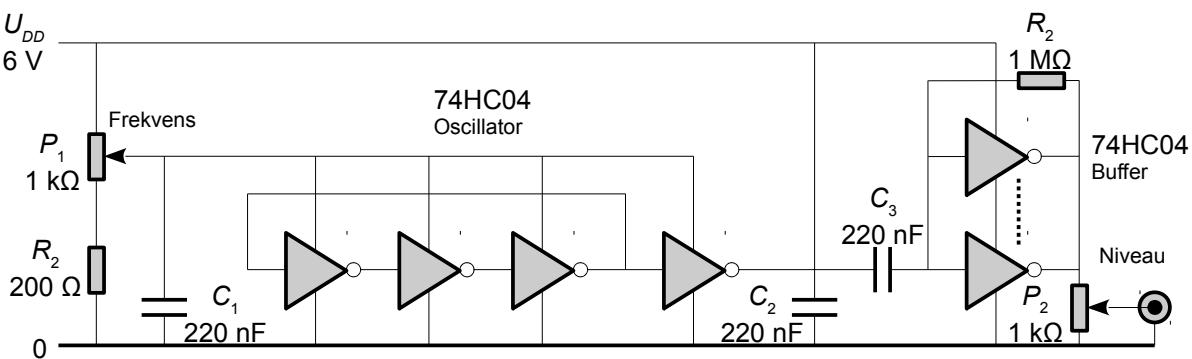


Figur 123 – En HC14 inverter med hysterese giver mulighed for en simpel oscillator, der kan benyttes til relativt høje frekvenser.

Den digitale kreds skifter mellem lav og høj på blot 15 ns så oscillatoren bruger 30 ns af perioden på at skifte udgangen, og hvis det accepteres at skiftet varer 10 % af hele perioden så bliver den øvre grænsefrekvens i omegnen af 3 MHz.

3.6.2 Spændingsstyret oscillator

De foregående oscillatorers frekvens er begrænset af den tid den aktive komponent er om at skifte mellem lav og høj, men det er muligt at gøre en dyd ud af en nødvendighed og basere oscillatoren på skiftetiden, nærmere bestemt komponentens *propagation delay*, t_d . Den enkleste måde er at koble tre invertere i en lukket sløjfe, for når sidste udgang går lav vil denne status løbe igennem kæden og ende med at tvinge den til at skifte høj. Kæden falder derfor aldrig til ro, men oscillerer på den frekvens, der er givet af den samlede tidsforsinkelse. For en CMOS digital kreds er propagation delay en funktion af effektforsyningen, så det er relativt enkelt at styre oscillatorenens frekvens med en variabel spænding i området fra 2 V til 6 V for CMOS komponenten 74HC04.

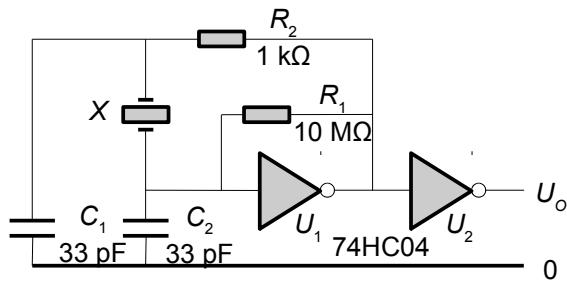


Figur 124 – En oscillator baseret på komponentens spændingsafhængige propagation delay.

Opstillingen har været brugt som en HF oscillator for EMC testning, hvor frekvensen blev varieret af P_1 fra cirka 10 til 40 MHz. Udgangsspændingen fra oscillatoren er givet af effektforsyningen, så en buffer blev indsatt på 6 V for at sikre en konstant amplitude, og alle seks invertere i kredsen var forbundet i parallel for at give en lav udgangsmodstand. Frekvensen reguleres på P_1 og udgangsniveauet på P_2 .

3.6.3 Krystaloscillator

Et krystal kan give en præcis indstilling af frekvensen, og denne type oscillatorer benyttes i mikroprocessorerne kredsløb. Modstanden R_1 sikrer at inverteren U_1 holdes i det aktive område, og kondensatorerne C_1 og C_2 går i resonans med selvinduktionen i krystallet X. Det betyder at oscillatoren svinger på krystallets parallelresonans, og det er normalt den specificerede frekvens for et krystal. For at undgå oscillation på en af krystallets overtoner kan modstanden R_2 inkluderes for at danne et lavpasfilter med C_1 så forstærkningen reduceres ved høje frekvenser.

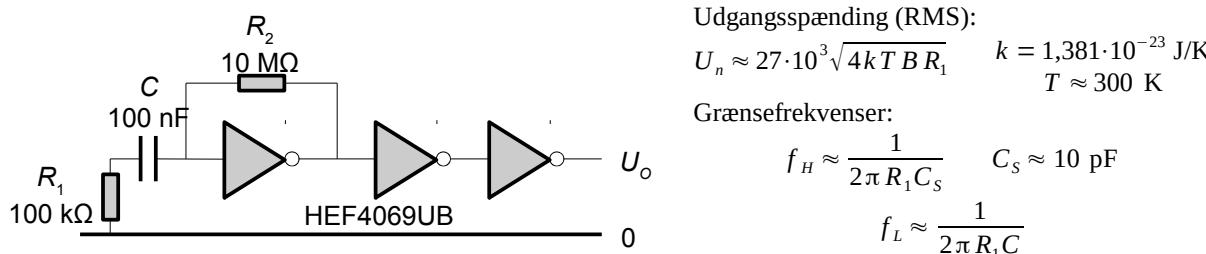


Figur 125 – En Colpitts-oscillator med et krystal for en præcis indstilling af frekvensen, og en buffer for at undgå påvirkning fra det kredsløb, der skal drives af oscillatoren.

Inverter U_2 er en buffer, der sikrer at udgangen ved U_o kan drive et tilsluttet kredsløb uden at det influerer på selve oscillatorens funktion, og dermed sikres en konstant frekvens. Udgangssignalet er en firkant, som er velegnet til digitale kredsløb. Det kan være nødvendigt at justere på de viste komponentværdier for at få en stabilt virkende funktion.

3.7 Støjgenerator

Den populære digitale kreds kan benyttes som inverterende forstærker, som vist ovenfor ved den krystalstyrede oscillator, men den kan også varetage normal analoge forstærkning hvis der vælges en inverter med kun et enkelt forstærkende trin, eksempelvis 4069UB. Den kan da opfattes som en særdeles bredbåndet, inverterende OpAmp med en åben-sløjfe forstærkning på cirka 30.



Figur 126 – En støjgenerator baseret på CMOS forstærkere.

Det anvendes her som en støjgenerator hvor den termiske støj fra en modstand forstærkes op til et brugbart niveau. Modstandens støj bestemmes af den viste formel hvor B er båndbredden af kredsløbet, der er givet af modstandsværdien og den kapacitet, som findes ved indgangen på CMOS komponenten og fra printudlægget; under ét betegnet med C_S . Antages en værdi på 10 pF bliver båndbredden på 160 kHz. Værdien af modstandens termiske støj bliver 16 μV og efter en forstærkning i de tre trin er niveauet nået op på omkring $U_{o,RMS} = 0,5 \text{ V}$. Værdien er afhængig af den valgte komponents parametre, den benyttede effektforsyning og belastningen af udgangen.

For at opstillingen kan fungere skal den skærmes fra ekstern støj, især 50 Hz fra lysnettet, så den bør bygges ind i et lukket metalkabinet og effektforsyningen skal afkobles. Modstanden R_1 skal være lille set i relation til indgangsmodstanden på den første CMOS forstærker; det er R_2 divideret med forstærkningen eller omkring 300 $\text{k}\Omega$, så R_1 bør ikke vælges til en højere værdi end den viste.

Note 3.

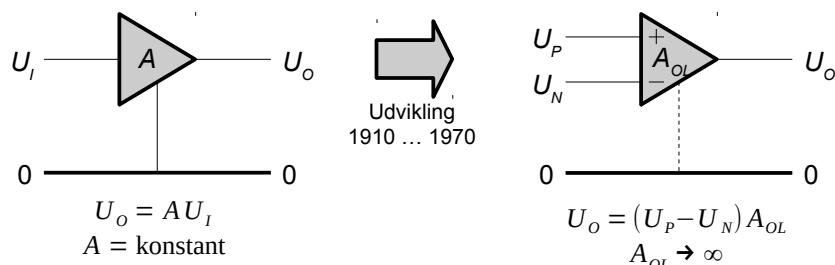
3.8 DC begrænsninger

Operationsforstærkeren har nogle begrænsninger, der skyldes dens opbygning, og de har betydning for hvor præcist den arbejder, hvor hurtig den er, og hvor mange "muskler" den har til at løfte en opgave. Det er fuldt ud muligt at designe velfungerende elektronik uden at tage andre forholdsregler end at være forsiktig, men det er bedre at være bekendt med begrænsningerne for at sikre sig at designet ikke har indbyggede svagheder.

De begrænsninger, der vil blive omtalt i det følgende angår evnen til at måle præcist (side 129) og klarlægge belastningen af måleobjektet (side 131), den begrænsede evne til at følge hurtige skift i indgangssignalet (side 134), den hastighed udgangen maksimalt kan bevæge sig med (side 136), ustabilitet (side 137), støjen operationsforstærkeren tilfører signalet (side 147) og det spændingsområde som ind- og udgang skal holdes indenfor (side 164). Ud over disse grænser kan nævnes at en OpAmp i plasthus kan arbejde indenfor et temperaturområde fra 0 til 70°C, mens militærspecificerede komponenter kan bruges ved et temperaturområdet fra -55°C til 125°C.

3.8.1 Åben-sløjfe forstærkning

Elektronikken var i starten af 1900-tallet baseret på radiorør og ved elektronikdesign var man især fokuseret på at få mest muligt forstærkning ud af den kostbare komponent. Den måde at designe på tog transistoren med sig, for også transistoren var en ganske dyr komponent kort efter 1947. Integrationen af mange transistorer på en enkelt chip fjernede gradvist den økonomiske hindring, og i løbet af decenniets tredje kvartal opstod en måde at designe forstærkere på, som udnyttede at det var nærmest gratis at øge antallet af transistorer.



Figur 127 – Et blokdiagram over en konventionel forstærker hvor udgangens spænding er givet ved indgangens spænding gange med forstærkningsfaktoren A og begge signaler refererer til nul. En moderne OpAmp har to indgange og det er differensen mellem de dem, der skal forstærkes op til udgangen. Her er kravet blot at dens forstærkning skal være "stør". Signalerne på ind- og udgang refererer ikke nødvendigvis til nul, for den terminal findes ikke tilgængelig på en moderne OpAmp; i stedet benyttes forsyningerne, hvilket ikke er helt uden komplikationer.

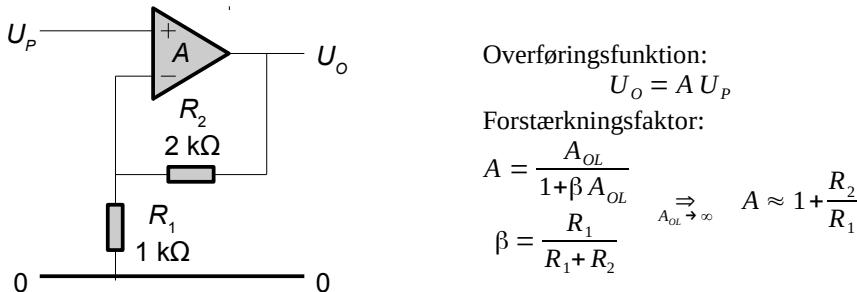
Modstande blev erstattet med strømkilder, der fyldte mindre på chippens areal, og det blev udnyttet at transistorerne var så godt som ens, hvorimod modstande var svære at integrere på en chip og tolerancen var elendig. Den ensartede opførsel af transistorerne tillod integration af en spændingsreference på chippen, så kredsløbet funktion blev uafhængig af forsyningsspændingen. Da der først var åbnet for mulighederne, opstod en helt anden måde at designe forstærkere på. Idéen var at tilbyde et så stort overskud af forstærkning, at kredsløbets funktion kunne fastlægges alene ved eksterne komponenter, hvorfor de integrerede modstande kun fik sekundær betydning. Diskrete modstande er i den forløbne tid blevet så stabile at en given forstærkning kunne indstilles meget præcist. Modstandstolerancen er i dag $\pm 1\%$ som standard, og $\pm 0,1\%$ er kun lidt dyrere, så man forventer ganske enkelt at forstærkningen kan indstilles tilsvarende præcist.

Designeren blev nu frigjort fra at skulle overveje enhver detalje om hvordan en given forstærkning skulle opnås, og til blot at forstærke signalet. Derved blev der frigjort overskud til koncentrere sig om væsentlige detaljer et andet sted i kredsløbet, og den moderne elektronik var født.

3.8.2 Tilbagekoblet forstærker

Den moderne operationsforstærker bruger differensen $U_P - U_N$ mellem de to indgange som bliver forstærket op med den såkaldte åben-sløjfe forstærkning A_{OL} (open-loop gain) til udgangens U_O . Det er forførende enkelt og fungerer overbevisende i praksis, men der er nogle knaster, som vil blive gennemgået i det følgende. Operationsforstærkere har en åben-sløjfe forstærkning fra 10^3 til 10^6 med hovedparten af de kommersielt tilgængelige komponenter i den høje ende af intervallet, så en forskel mellem de to indgange på $100 \mu V$ vil forstærkes op til $10 V$, og det er i langt de fleste tilfælde nær det maksimalt mulige før forstærkeren overstyrer. Det er uden betydning om de to indgange ligger på $0 V$ eller på $3 V$ da det kun er differensen, der forstærkes op.

En høj værdi af åben-sløjfe forstærkningen betyder at den negative indgang vil have så godt som samme potentielle som den positive indgang hvilket leder til det designprincip der blev postuleret på side 107 og som er bærende for alt arbejde med operationsforstærkere.



Figur 128 – Den ikke-inverterende forstærker. En brøkdel af udgangssignalet kobles tilbage til den negative indgang for at fastlægge forstærkningen til en værdi bestemt kun ved modstandenes værdier og ikke relateret til forstærkerens specifikationer.

I designet herover vil udgangen indstille sig på netop den værdi der bringer den negative indgang hen til det potentielle den positive indgang befinner sig på og forskellen er nogle få mikrovolt. Det er i langt hovedparten af al design så lidt at det kan ignoreres og indstilling af forstærkningen er meget præcis – i det mindste ved DC og lave frekvenser.

Med de viste komponentværdier forventes en forstærkning på 3 gange, og det vil praksis også vise at man får. Hvis R_2 ændres til $2 M\Omega$ bliver resultatet ikke en forstærkning på 2001 gange som man kunne forvente fra formlerne på side 108. Med en konservativ åben-sløjfe forstærkning på 10^5 er resultatet 1962 ganges forstærkning, der er 2 % under det forventede. Åben-sløjfe forstærkningen skal være meget større end den ønskede forstærkning. Som udgangspunkt skal A_{OL} være 100 gange højere end den ønskede forstærkning for at holde fejlen under 1 % så den typiske operationsforstærker kan anvendes for forstærkning af signaler på indtil 1000 gange (60 dB).

Udgangsspændingen er givet ved differensen mellem de to indgange som forstærkes op af åben-sløjfe forstærkningen. Ved tilbagekobling angiver β den brøkdel af udgangsspændingen som føres tilbage til den negative indgang.

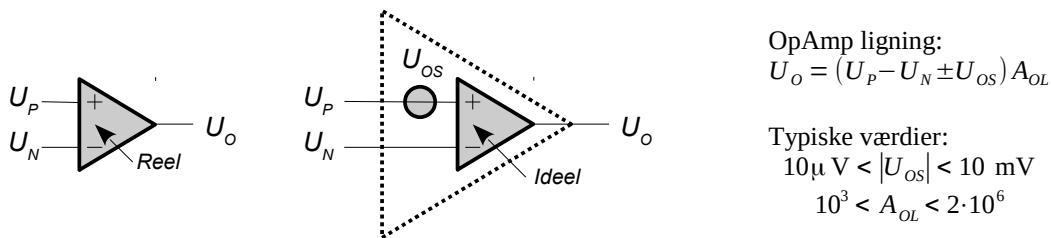
$$U_O = (U_P - U_N) A_{OL} \quad \text{hvor} \quad \beta = \frac{R_1}{R_1 + R_2} \quad \Rightarrow \quad U_O = \frac{A_{OL}}{1 + \beta A_{OL}} U_P$$

For at vise betydningen af åben-sløjfe forstærkningen deles udtrykket med A_{OL} i tæller og nævner, så åben-sløjfe forstærkningen kun optræder ét sted. For at kunne se bort fra ledet $1/A_{OL}$ skal A_{OL} være meget større end $1/\beta$ og da forstærkningen netop er givet ved $1/\beta$ bliver kravet $A_{OL} > 100 A_{CL}$ for en fejl under 1 % ved at bruge tilnærmelsen.

$$A = \frac{U_O}{U_P} = \frac{1}{1/A_{OL} + \beta} \xrightarrow{A_{OL} \rightarrow \infty} \frac{1}{\beta} \quad \Rightarrow \quad A = \frac{R_1 + R_2}{R_1} = 1 + \frac{R_2}{R_1}$$

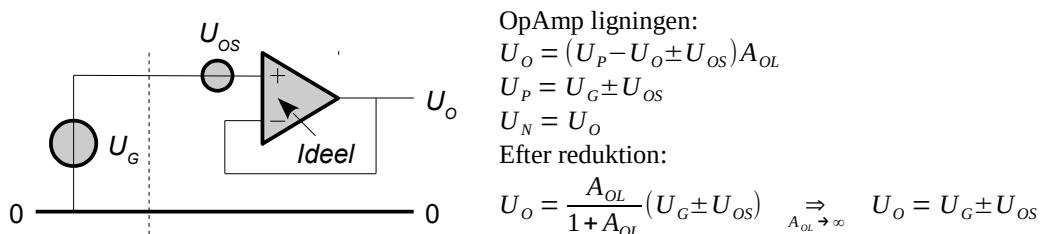
3.8.3 Indgangens fejfspænding

Indgangsdelen på operationsforstærkeren består af transistorer, der uddrager differensen mellem spændingsværdierne på den positive indgang og den negative indgang. Ideelt skal den samme spænding på de to indgange give en udgangsspænding nul, men der vil altid være en ubalance i selv den bedste operationsforstærker⁶⁶, så udgangsspændingen vil ikke indstilles på nul hvis de to indgange kortsluttes. Den praktiske betydning er at der altid vil optræde en fejl, som begrænser den nøjagtighed en operationsforstærker kan arbejde med. Om fejfspændingen så i praksis udgør et problem afhænger af den aktuelle situation, og det er emnet for dette og de følgende afsnit.



Figur 129 – Ubalance i indgangen indføres i modellen som en fejfspænding, der optræder som en forskel mellem indgangene for at forstærkeren er i balance. En DC model af den reelle forstærker beskrives ved en spændingskilde U_{OS} , der repræsenterer forstærkerens fejfspænding, og dertil en i øvrigt ideel forstærker.

For at kunne forholde sig kvantitativt til problemet beskrives en operationsforstærker ved en DC model bestående af en ideel operationsforstærker med en spændingskilde som repræsentant for fejfspændingen med symbolet U_{OS} (*input voltage offset, V_{OS}*). Fortegnet er ikke kendt, og det er den aktuelle værdi heller ikke; fabrikanten kan kun oplyse at værdien vil ligge indenfor et interval. Hvis en forstærker har specifikationen $U_{OS} = 5 \text{ mV}$ så er det eneste man ved, at værdien for en specifik OpAmp vil være et sted mellem -5 mV og $+5 \text{ mV}$. Den sandsynlige værdi er tæt på nul, men den kan i værste tilfælde nå ud til den angivne grænse, og det bør man tage hensyn til ved ethvert design af et elektronisk kredsløb.



Figur 130 – En spændingsfølger skal ideelt blot gengive indgangsspændingen fra generatoren på udgangen, men fejfspændingen giver en usikkerhed på udgangsspændingen.

Fejfspændingen er funktion af temperaturen, så oplyses ved en fast temperatur på 25°C , og ofte også for variationen over et temperaturområde⁶⁷. Fejfspændingen driver over tid, så den er ikke en fast værdi, men vil variere i løbet af produktets levetid⁶⁸. Dertil kommer at den momentant kan springe mellem forskellige værdier, hvilket specificeres som en spids-til-spids værdi for operationsforstærkere til krævende anvendelser⁶⁹, og det beskrives i støjafsnittet se side 162.

Den ikke-inverterende forstærker kan også benyttes til at isolere to kredsløb fra hinanden og den giver samtidigt en forstærkning af signalet. Et anvendelseseksempel er sensorer som pH-metre, med en høj indre modstand og behov for forstærkning.

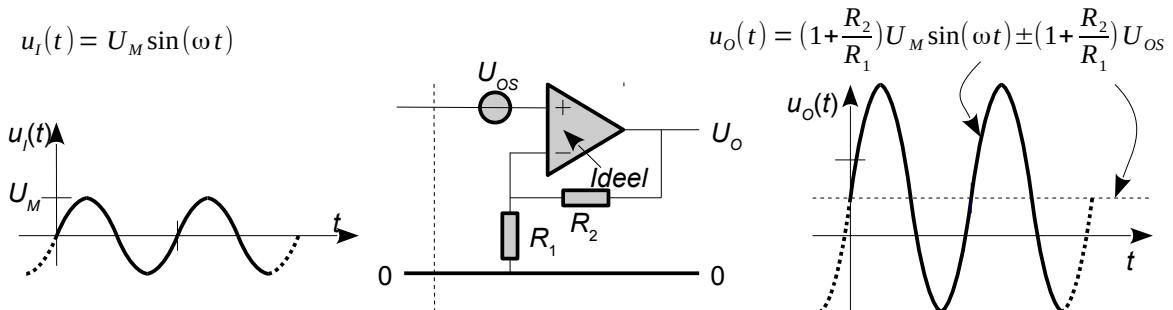
⁶⁶ OP27 oplyses til maksimalt $100 \mu\text{V}$. Værdien er omkring $\pm 5 \text{ mV}$ for gamle travere som uA741, LF356 og LM324, mens den kan nå ned på $\pm 10 \mu\text{V}$ for en moderne OpAmp med et internt korrektionskredsløb som ADA4051.

⁶⁷ OP27 oplyses til maksimalt $0,2 \mu\text{V}/^\circ\text{C}$.

⁶⁸ OP27 oplyses til maksimalt $2 \mu\text{V}/\text{måned}$.

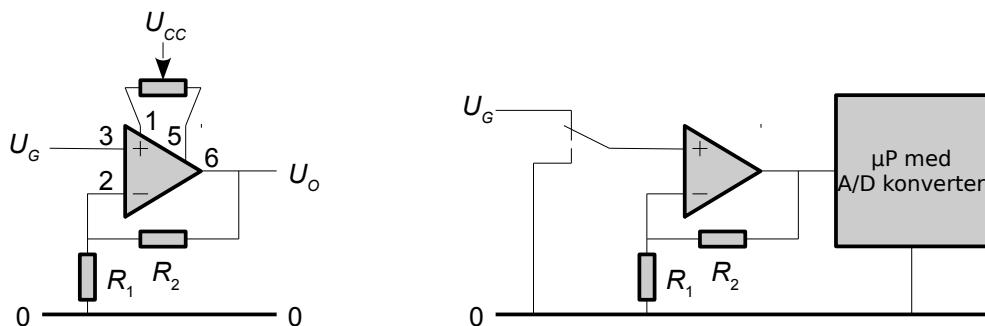
⁶⁹ OP27 oplyses til maksimalt 80 nV spids-til-spids indenfor et frekvensområde fra $0,1 \text{ Hz}$ til 10 Hz .

Hele idéen er at undersøge om problemet er så stort, at der er behov for en aktion, eller om man helt kan ignorere fejlpændingens indvirkning. For at bruge denne usikkerhed konstruktivt kan det omsættes til et krav, som eksempelvis et valg af en forstærker med en specificeret maksimal værdi af fejlpændingen. En anden mulighed er at sætte et krav til signalniveauet, eller man kan blive tvunget til at revurdere sit kredsløb.



Figur 131 – En forstærker vil også forstærke fejlpændingen op.

Økonomi kan diktere mange omveje til en løsning, for eksempel ved at chefen tvinger én til at finde en løsning med en billigere OpAmp. Man kan vælge at indføre et potentiometer for manuel kompensering af fejlpændingen, men den er ikke uden omkostninger et andet sted. Ud over prisen for den ekstra komponent, som dertil optager plads på kortet, så er der også en udgift i produktionen fordi en person skal foretage justeringen. Dertil kommer at uanset hvor omhyggelig personen er, så giver det en fejlmulighed, som så kræver en form for kontrol i produktionen. Endelig et et potentiometer ikke er en velset komponent på grund af en ringe mekanisk stabilitet. Trimning er en god måde til at reducere fejlpændingen i en forsøgsopstilling, men den er ikke god for en serieproduktion, hvor prisen på en OpAmp ofte kan forhandles ned ved storindkøb.



Figur 132 – Fejlpændingen kan justeres væk på mange forstærkere med en ekstern trimmer. En anden mulighed er at lade en mikroprocessor måle udgangsspændingen i en kontrolleret situation og derefter fjerne fejlpændingen i software.

En metode, der egner sig bedre til moderne kredsløb, bruger en mikroprocessor til at måle den aktuelle værdi af kredsløbets fejlpænding⁷⁰. En mulighed er at inkludere en analog omskifter i opstillingen, og så på passende tidspunkter i programmet kortsluttes forstærkerens indgang og dens udgangsspænding måles og gemmes i et register. Herefter kan fejlpændingen fjernes fra de efterfølgende målinger ved at subtrahere registerets værdi fra måledata.

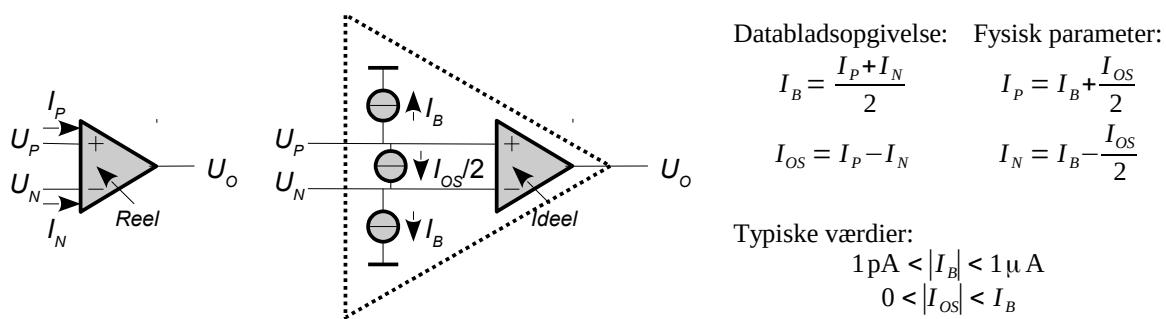
Fejlpændingen er den ene af to væsentlige fejlkilder ved DC måling, og den anden introduceres herunder. Afslutningsvis kombineres fejlkilderne til en fælles DC model for en OpAmp.

⁷⁰ En tilsvarende løsning findes integreret på operationsforstærkerens chip hvor der er indbygget et kredsløb for nulstilling af fejlpændingen; de *chopper stabilised amplifiers*, som fx AD8628 hvor fejlpændingen holdes under $\pm 1 \mu\text{V}$. En oscillator vil fra tid til anden afbryde indgangen for at måle fejlpændingen så der kan korrigeres for ubalance. Set udefra sker der intet, idet udgangens spænding fastholdes mens den interne kalibrering foregår. Signalforarbejdningen er kortvarigt afbrudt så den analoge båndbredde er begrænset på grund af samplingsprocessen.

3.8.4 Indgangsstrøm

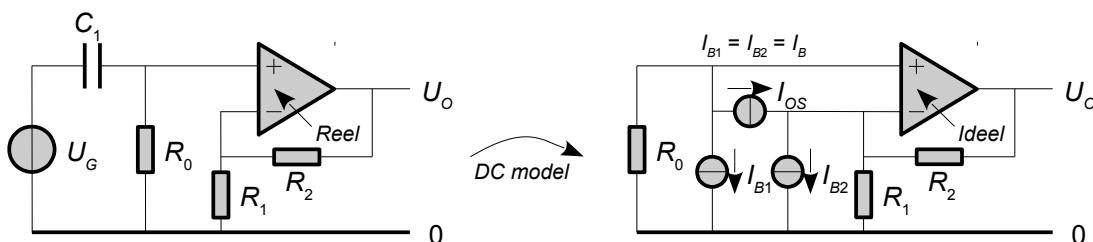
Operationsforstærkeren kræver at der løber en vis strøm i indgangen for at elektronikken fungerer, og strømmen belaster det kredsløb der måles på. Den kaldes for biasstrøm som en fordanskning af den engelske betegnelse, og værdien varierer meget mellem de forskellige typer; fra cirka 1 pA ved MOSFET og indtil 1 µA ved bipolære transistorer⁷¹. Strømretningen er som hovedregel kendt ved bipolære transistorer, men databladet bør konsulteres hvis det er vigtigt⁷².

For at analysere kredsløb for biasstrømmens betydning benyttes en DC model hvor en OpAmp beskrives som en ideel forstærker med strømkilder for biasstrømmen i de to indgange. Strømmen kaldes I_P i den positive indgang og I_N i den negative, men databladet OpAmp oplyser en gennemsnitlig værdi, kaldet I_B (*bias current*), og en ubalance kaldet I_{OS} (*bias offset current*). Den faktiske strøm i de to indgange er som udgangspunkt ens og lig med biasstrømmen I_B , men den faktiske værdi af strømmen har en usikkerhed, som beskrives ved en offsetstrøm I_{OS} . Der er ikke nogle gode danske betegnelser for disse strømme, derfor anvendes de fordanskede betegnelser.



Figur 133 – Der flyder en strøm i forstærkerens indgange og det repræsenteres analytisk ved strømkilder på indgangen af en i øvrigt ideel forstærker.

Illustrationen herunder viser en typisk forstærker, hvor indgangssignalet føres gennem en kondensator for at blokere for DC, så forstærkeren er kun beregnet til at forstærke AC signaler. DC forholdene kan dog være så påvirkede af biasstrømmen, at udgangens DC arbejdspunkt afviger for meget fra nul til at kunne opfylde den tiltænkte opgave, og den følgende analyse er derfor relevant selv for en AC forstærker. Formålet er at beregne hvor meget udgangens DC arbejdspunkt afviger fra det ønskede, der her er nul, og derefter vurderes om der er et problem. Ved DC analysen kan kondensatoren erstattes med en afbrydelse, for der løber ingen strøm når opstillingen er i ro.



Figur 134 – Til venstre vises en forstærker for AC signaler hvor udgangen skal stå på nul uden et indgangssignal. Til højre erstattes operationsforstærkeren med DC modellens strømkilder for en beregning af biasstrømmens indvirkning på DC arbejdspunktet.

⁷¹ Med bipolære transistorer er værdien cirka 100 nA (LM324, uA741), mens en støjsvag forstærker kan komme over 1 µA (NE5532, AD797). Med felt-effekt transistorer i indgangen er strømmen under 1 nA for J-FET typer (LF356, TL071) mens MOSFET og CMOS typer med et isolerende lag af siliciumoxyd mellem styreelektroden og den aktive del af transistoren når ned under 10 pA (TLC274). Den bipolare transistors biasstrøm er relativt uafhængig af temperaturen, hvormod felt-effekt transistorer typisk fordobler strømmen for hver 10°C stigning; så ved 100°C er biasstrømmen for en JFET øget til samme niveau som den bipolare transistor, og ved 150°C er biasstrømmen for en MOSFET sammenlignelig med den bipolare transistor.

⁷² OP27 benytter et indgangskredsløb for at holde biasstrømmen lav, og det giver en usikkerhed på strømretningen. Komponenten oplyses til en biasstrøm på maksimalt ±75 nA og en offsetstrøm på maksimalt ±80 nA.

Herunder vises en analyse, der inddrager alle strømkilder, og det vises at et hensigtsmæssigt valg af modstandsværdier kan eliminere bidraget fra biasstrømmen, men ikke fra offsetværdien.

Beregningen gennemføres ved anvendelse af superpositionsprincippet.

I_{B1} – Ved den positive indgang på operationsforstærkeren løber strømmen I_B gennem modstanden R_0 og genererer et spændingsfald på $-R_0 I_B$, hvor minusset skyldes at strømretningen er fra nul og op. Dette spændingsfald forstærkes med $1 + R_2/R_1$ og føres til udgangen.

I_{B2} – Ved den negative indgang på operationsforstærkeren løber strømmen I_B gennem modstanden R_2 og genererer et spændingsfald på udgangen på $R_2 I_B$. Det løber ingen strøm igennem R_1 fra strømkilden, idet den ideelle operationsforstærker holder den negative indgang på nul når alle øvrige kilder er nulstillet.

I_{OS} – Ved den positive indgang på operationsforstærkeren løber strømmen $I_{OS}/2$ gennem R_0 på samme måde som I_B og den genererer derfor et bidrag på $-R_0 I_{OS}$, der forstærkes op som før. Ved den negative indgang på operationsforstærkeren løber strømmen $I_{OS}/2$ gennem R_2 på samme måde som I_B og giver et spændingsfald på udgangen på $-R_2 I_{OS}/2$.

Den resulterende påvirkning ved udgangen er givet ved summen af de enkelte bidrag, og ligningen simplificeres ved at samle bidragene fra I_B og I_{OS} til hvert sit led.

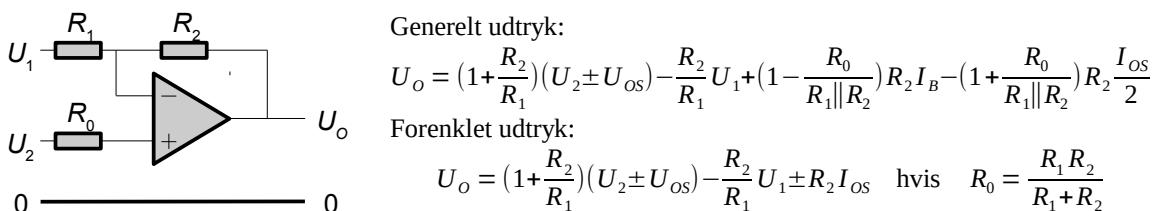
$$\begin{aligned} U_O &= \text{Bidrag fra } I_{B1} \quad \text{Bidrag fra } I_{B2} \quad \text{Bidrag fra } I_{OS} \\ &= -(1 + \frac{R_2}{R_1}) R_0 I_B \quad + R_2 I_B \quad -(1 + \frac{R_2}{R_1}) R_0 \frac{I_{OS}}{2} - R_2 \frac{I_{OS}}{2} \\ &\Rightarrow U_O = [R_2 - (1 + \frac{R_2}{R_1}) R_0] I_B - [R_2 + (1 + \frac{R_2}{R_1}) R_0] \frac{I_{OS}}{2} \end{aligned}$$

Første led kan sættes til nul hvis parallelværdien af R_1 og R_2 er lig med R_0 . Det indses ved at sætte den første parentes til nul og omforme udtrykket. I det tilfælde bliver betydningen fra I_{OS} på $R_2 I_{OS}$, idet andet led reduceres til $2R_2$.

$$R_2 - \frac{R_1 + R_2}{R_1} R_0 = 0 \Rightarrow R_2 [1 - \frac{R_1 + R_2}{R_1 R_2} R_0] = 0 \Rightarrow \frac{R_1 R_2}{R_1 + R_2} = R_0$$

3.8.5 DC model

Udtrykkene fra de foregående afsnit kan samles til et fælles udtryk for udgangsspændingen U_O for enten en ikke-inverterende forstærker ($U_1 = 0$) eller en inverterende forstærker ($U_2 = 0$) hvor påvirkningen fra biasstrømmen I_B kan undertrykkes ved at modstanden R_0 sættes lig parallelværdien af R_1 og R_2 . Påvirkningen fra fejlpændingen U_{OS} og offsetstrømmen I_{OS} kan ikke undertrykkes⁷³.



Figur 135 – Resulterende udtryk for operationsforstærkeren med fejlpændingen U_{OS} (offset voltage) og biasstrømmen I_B og bias offset I_{OS} , der løbet gennem de eksterne modstande.

Det er god designpraksis at kompensere for biasstrømmens indvirkning ved modstand R_0 for en OpAmp med bipolare transistorer. Ved MOSFET er der ikke brug for kompensering, men den kan alligevel være god i et udviklingsforløb, fordi den tillader designeren at bryde ind i kredsløbet uden at skulle overskære baner på et ellers færdigt printkort. Modstanden kan dog introducere støj fra en ekstern støjkilde.

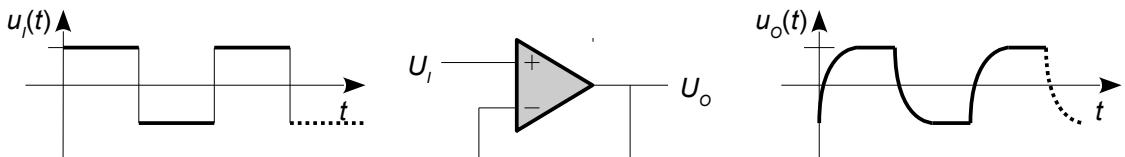
⁷³ Som regel har man en designmæssig udfordring hvis offsetstrømmen bidrager med et problem, og så var det bedre at genoverveje selve kredsløbet, frem for at slås med at få det til at virke stabilt over en produktion linje, og ikke mindst hos kunderne.

3.9 Båndbredde

Operationsforstærkeren er opbygget med en dominerende pol for at gøre den stabil i flertallet af kredsløb, men det har den konsekvens at der er grænser for hvor hurtig den kan følge et signal, hvor høje frekvenser den kan forstærke og hvor stor en kapacitet den kan tåle på ind- og udgang. Det teoretiske grundlag følger side 140 og herunder illustreres det med nogle eksempler.

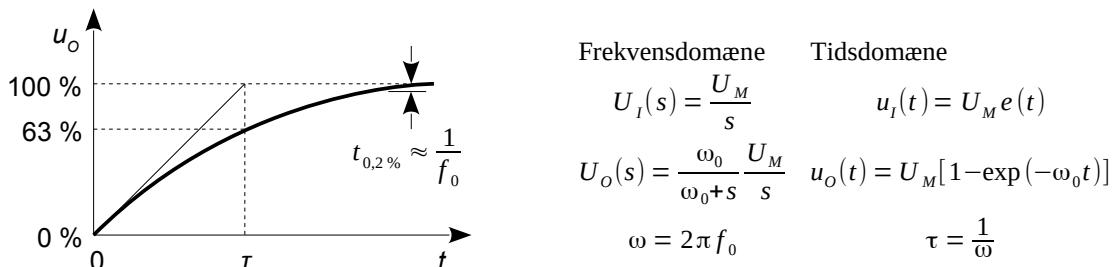
3.9.1 Steprespons

En pludselig ændring af indgangsspændingen til en operationsforstærker vil ikke resultere i en momentan ændring af udgangen til den nye position; i stedet vil udgangen roligt glide fra den hidtidige position til den nye værdi.



Figur 136 – Operationsforstærkeren kan ikke momentant springe op eller ned til en ny værdi af udgangens spænding, men vil glide på plads ved en eksponentialfunktion.

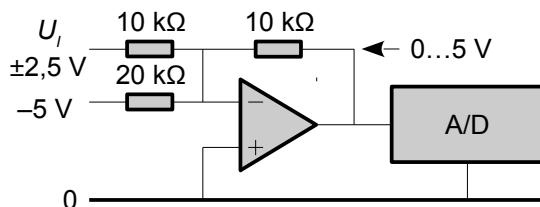
Relationen forløber eksponentielt, så den starter med en vis hastighed, som gradvist aftager mod nul nær ved den nye position, og det kan tolkes som om operationsforstærkeren var opbygget som et lavpasfilter med en enkelt pol. Det kan ses ved at bruge værktøjet fra Laplace transformationen, hvor reaktionen på en pludselig ændring af indgangssignalet kan forklares som tidssvaret fra et lavpasfilter med et step på indgangen.



Figur 137 – Operationsforstærkeren kan ikke momentant springe op eller ned til en ny værdi af udgangens spænding, men vil glide på plads ved en eksponentialfunktion.

Ved grænsefrekvensen $f_0 = 100$ kHz er tidskonstanten $\tau = 1,6$ μ s så udgangen i ro efter $t_{0,2\%} = 10$ μ s.

Hvis operationsforstærkeren skal drive et signal til en A/D konverter er kravet at fejlen i udgangens spænding skal være mindre end 1 LSB. Med N som antallet af bit skal udgangen være indenfor 2^{-N} af A/D konverterens område inden konverteringen kan startes. Beregningen herunder viser hvor lang tid, der kræves før fejlen er under 1 LSB, men man bør i praksis vente en del længere for at være sikker på at operationsforstærkerens udgangsspænding er stabil under konverteringen.



Rest efter tiden t_N :

$$\epsilon = \exp(-\omega_0 t_N) \Rightarrow t_N = \frac{1}{\omega_0} \ln(\frac{1}{\epsilon})$$

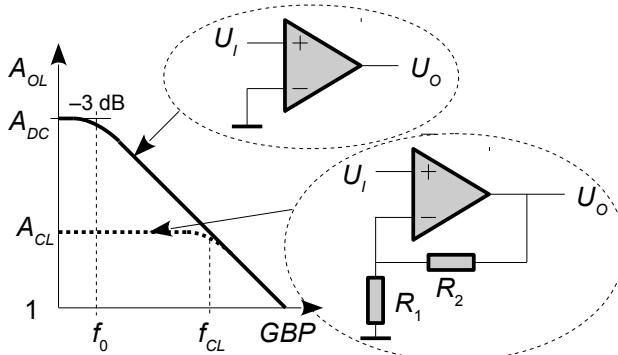
For 2^{-N} som rest:

$$t_N = \frac{N \ln(2)}{2\pi f_0} \approx \frac{N}{9f_0} \Rightarrow t_{18} \approx \frac{2}{f_0}$$

Figur 138 – Operationsforstærkeren benyttes ofte til at drive indgangen på en A/D konverter. Her er det en inverterende forstærker hvor indgangssignalets DC niveau forskydes fra nul til 2,5 V.

3.9.2 Dominerende pol

Operationsforstærkeren er opbygget som et lavpasfilter med stor forstærkning ved lave frekvenser, for at opnå stabilitet i flertallet af de opstillinger, den kan tænkes at skulle indgå i. Det betyder at åben-sløjfe forstærkningen A_{OL} aftager over en frekvens, der kaldes for den dominerende pol f_0 (dominating pole). Det er kun ved DC og lave frekvenser at operationsforstærkeren opfører sig som det er beskrevet i de foregående afsnit. Ved signaler med højere frekvens vil den miste noget af sit overskud af forstærkning, og den er derfor ikke i stand til at fungere lige så godt som ved lavere frekvenser. Det er en væsentlig begrænsning ved operationsforstærkeren, men den er indført for at sikre en stabil og veldokumenteret funktion, uanset hvad man byder den af udfordringer.



$$\begin{aligned} \text{Åben-sløjfe forstærkning: } A_{OL} &= \frac{\omega_0}{\omega_0 + s} A_{DC} \\ \text{DC forstærkning: } A_{DC} &= 1 \cdot 10^5 \dots 2 \cdot 10^6 \\ \text{Dominerende pol: } f_0 &= \frac{\omega_0}{2\pi} \\ \text{Gain-bandwidth product: } GBP &= A_{DC} f_0 \\ \text{Grænsefrekvens: } f_{CL} &= \frac{GBP}{A_{CL}} \\ \text{Forstærkning: } A_{CL} &= 1 + \frac{R_2}{R_1} \end{aligned}$$

Figur 139 – Åben-sløjfe forstærkningen for en forstærker med dominerende pol. Ved DC er værdien af forstærkningen A_{DC} og den aftager med $-6 \text{ dB}/\text{oktav}$ over den dominerende pol f_0 . Ved GBP er forstærkningen faldet til én (0 dB båndbredden).

Ved lave frekvenser er forstærkningen givet ved DC forstærkningen A_{DC} , der typisk er over 10^5 svarende til 100 dB, og som sikrer at selv med en udgangsspænding på $\pm 10 \text{ V}$ vil der være en differens mellem de to indgange på mindre end $\pm 100 \mu\text{V}$, hvilket i praksis kan ignoreres.

Over den dominerende pol aftager åben-sløjfe forstærkningen med $-6 \text{ dB}/\text{oktav}$ ($-20 \text{ dB}/\text{dekade}$) og polen ligger almindeligvis i området fra 1 Hz til 1 kHz, med en værdi under 10 Hz som det typiske for prisbillige operationsforstærkere. Ved en høj frekvens er åben-sløjfe forstærkningen reduceret til én gangs forstærkning (0 dB), og operationsforstærkeren kan ikke anvendes ved højere frekvenser. Denne frekvens kaldes for forstærkerens 0 dB båndbredde (*unity-gain bandwidth*) eller transitfrekvensen f_T . Den betegnes også med GBP , der står for *gain-bandwidth product*. Det er en praktisk betegnelse idet forstærkningen for frekvenser over polen vil aftage lige så meget som frekvensen stiger. Det er en af de vigtigste dataoplysninger for en operationsforstærker og den gives altid i databladet.

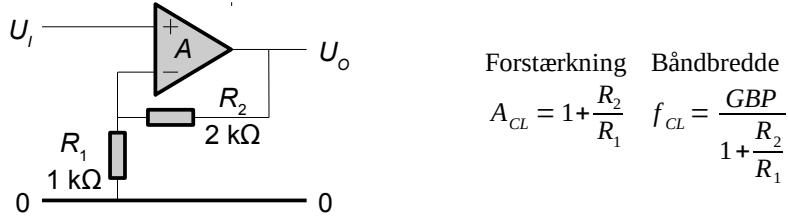
Idet produktet af indstillet forstærkning A_{CL} og båndbredde f_{CL} er konstant for frekvenser over f_0 , kan man beregne båndbredden ved en givet værdi af forstærkningen, hvilket er overordentligt praktisk ved planlægning af et nyt design. Man sammenholder simpelthen den ønskede båndbredde med det, der er muligt for en given operationsforstærker, og kan derved afgøre om den vil kunne løfte den aktuelle opgave, eller om man skal søge efter en anden operationsforstærker.

Som et eksempel benyttes en operationsforstærker med $GBP = 3 \text{ MHz}$, der er en almindelig specifikation for prisbillige kredse. Hvis den skal benyttes som spændingsfølger kan den klare signaler indtil 3 MHz hvor dens forstærkning er reduceret fra 0 dB til -3 dB . Skal operationsforstærkeren have en forstærkning på 100 vil den mulige båndbredde blive 30 kHz ved -3 dB grænsefrekvensen.

På tilsvarende vis kan man beregne den mulige værdi af forstærkningen ved en krævet båndbredde for en givet operationsforstærker.

Hvis det er krævet at et signal skal forstærkes med en båndbredde på 100 kHz (-3 dB) vil den højest mulige forstærkning blive 30.

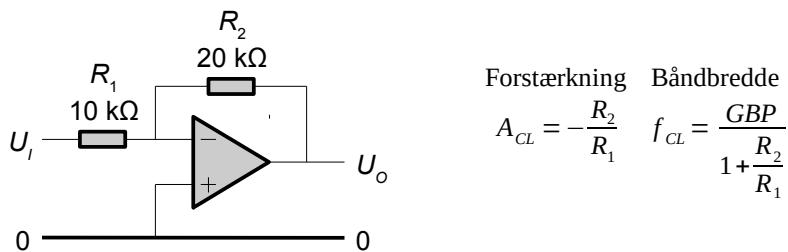
Ved lav værdi af forstærkningen er der en forskel på om man benytter en ikke-inverterende eller en inverterende forstærker, men man kan normalt se bort fra det. For den ikke-inverterende type er den opnåelige båndbredde *GBP* parameteren for operationsforstærkeren divideret med den værdi af forstærkning man vælger ved modstandene i tilbagekoblingen.



Figur 140 – En ikke-inverterende forstærker har en endelig båndbredde, der er givet ved den indstillede forstærkning.

Med de viste komponentværdier er forstærkningen $A_{CL} = 3$ gange og med $GBP = 3 \text{ MHz}$ vil forstærkeren kunne fungere op til $f_{CL} = 1 \text{ MHz}$ ved -3 dB punktet.

For en inverterende forstærker er beregningen baseret på samme formel, selv om indstillingen af forstærkning benytter en anden formel. Den inverterende forstærker indstilles naturligvis til en forstærkning af signalet på $-R_2/R_1$, men set fra operationsforstærkerens synspunkt er det tilbagekoblede signal stadigvæk $R_1/(R_1 + R_2)$ gange udgangssignalet, så forstærkeren synes den arbejder på fuldstændig samme måde som ved den ikke-inverterende forstærker.



Figur 141 – En inverterende forstærker har en endelig båndbredde, der er givet ved den indstillede forstærkning.

Med de viste komponentværdier er signalforstærkningen $A_{CL} = -2$, men da sløjfe-forstærkningen er 3 vil en forstærker med $GBP = 3 \text{ MHz}$ kunne fungere op til 1 MHz ved -3 dB punktet.

Overføringsfunktionen for den ikke-inverterende forstærker kan beregnes fra definitionen af åben-sløjfe forstærkningen ved indsættelse i udtrykket for en forstærker med en tilbagekobling gennem de to modstande R_1 og R_2 , der fører en andel på β af udgangssignalet tilbage til den inverterende indgang. Resultatet er et nyt lavpasfilter hvor polen er flyttet op i frekvens. Normalt er β større end 0,001 og A_{DC} er større end 105, så deres produkt er mindst 100 med almindelige operationsforstærkere, og dermed vindes ganske meget i båndbredde.

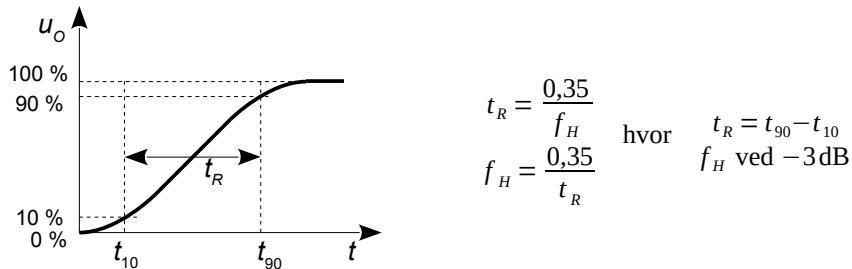
$$A_{CL} = \frac{A_{OL}}{1 + \beta A_{OL}} \Rightarrow A_{CL} = \frac{\frac{\omega_0}{\omega_0 + s} A_{DC}}{1 + \beta \frac{\omega_0}{\omega_0 + s} A_{DC}} = \frac{\omega_0 A_{DC}}{(1 + \beta A_{DC}) \omega_0 + s} = \frac{A_{DC}}{1 + \beta A_{DC}} \frac{(1 + \beta A_{DC}) \omega_0}{(1 + \beta A_{DC}) \omega_0 + s}$$

$$\beta = \frac{R_1}{R_1 + R_2}$$

Faktoren $1 + \beta A_{DC}$ kaldes for tilbagekoblingsfaktoren (*feedback factor*), og udgør det overskud, som forstærkeren bruger til at korrigere sin udgangsspænding med, og det påvirker en lang række af parametre. Tilbagekoblingsfaktoren vil reducere den harmoniske forvrængning med $1 + \beta A_{DC}$, støj indkoblet fra effektforsyningen vil blive dæmpet på samme måde og udgangsmodstanden fra kredsløbet reduceres tilsvarende.

3.9.3 Måling af båndbredde

Ved aflæsning af stepresponsen på et oscilloskop kan det være svært at se hvor kurven starter, og det giver en usikkerhed i bestemmelsen af båndbredden. Der findes en anden målemetode hvor man mäter signalets stigetid (rise time). Den er defineret som tiden fra signalet krydser 10 % af slutværdien og indtil den krydser 90 % af slutværdien.



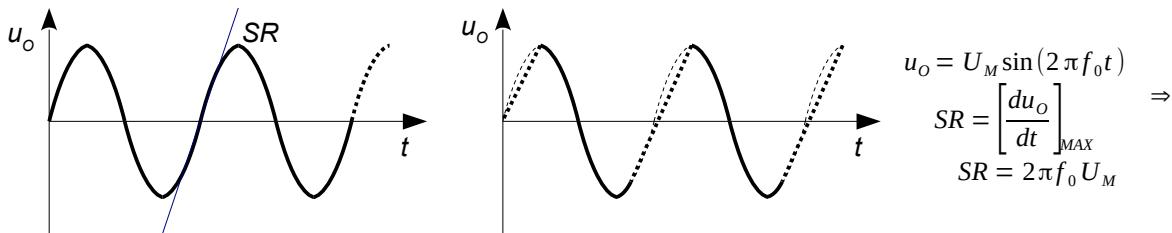
Figur 142 – Stigetiden af et signal er relateret til den høje grænsefrekvens f_H . Formlen er vist for et første-ordens lavpasfilter, men stigetiden benyttes normalt uden hensyn til filtrets orden.

Metoden benyttes især ved større kredsløb med mange operationsforstærkere hvor det med sikkerhed ikke er muligt at beskrives overføringsfunktionen som et første-ordens lavpasfilter. Ved højere ordens filtre vil kurven starte med vandret tangent, hvilket gør det meget svært at definere det tidspunkt hvor forløbet begynder; derimod er start- og slutniveauerne lette at definere. Mange af de lidt ældre oscilloskoper har direkte angivet hvordan signalet skal indstilles på skærmen for at aflæse de to punkter. På moderne, digitale oscilloskoper kan man benytte cursor-muligheden eller helt enkelt bede oscilloskopet om at beregne stigetiden.

Det er klart at en formel udledt for et første-ordens filter ikke umiddelbart giver et gyldigt resultat ved højere orden, men metoden er meget anvendt i praksis; den er enkel og resultatet er i praksis ganske brugbart om end man ikke skal regne med for mange decimaler.

3.9.4 Slew rate

Forudsætningen for beregningen af rise time er at operationsforstærkeren ikke overstyres, for det kan let forekomme på grund af dens *slew rate* specifikation. Operationsforstærkerens udgang er begrænset af det interne design til at flytte sig med en maksimum fart, som det vises i femte kapitel om operationsforstærkerens opbygning. Værdien er omkring 1 V/μs for prisbillige operationsforstærkere og op til 1000 V/μs for de hurtige versioner. Hvis indgangssignalet skifter abrupt med en høj spændingsværdi vil operationsforstærkeren blive overstyret og så falder forudsætningen for de ovenstående beregninger.



Figur 143 – Udgangens hastighed er bestemt ved spændingsændringen og tiden som ændringen sker indenfor. Forstærkeren har en øvre værdi af denne hastighed, som ikke kan overskrides. Forsøges at ændre udgangen hurtigere vil den blot stige eller falde med konstant hastighed.

Parameteren er vigtig ved analog signalbehandling af lydkilder.

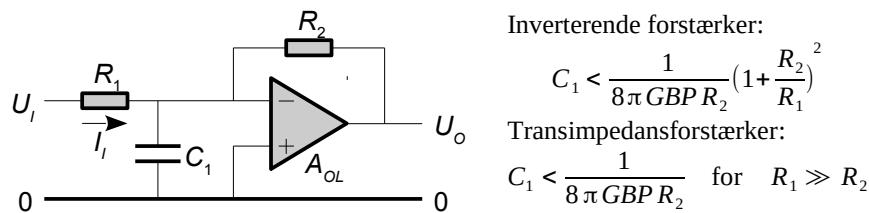
For et signal med en amplitude på 10 V er den maksimalt tilladelige frekvens på 16 kHz for forstærkeren uA741 med specifikationen $SR = 1 \text{ V}/\mu\text{s}$ kan alene af den årsag ikke anvendes til krævende lydformål.

3.9.5 Stabilitet

Operationsforstærkeren er stabil i så vidt muligt alle opstillinger, og man behøver ikke normalt bekymre sig om ustabilitet ved en typisk applikation. Stabiliteten kan dog blive kompromitteret i visse situationer, hvor ind- eller udgang er belastet af en kondensator. Det kan være en transimpedansforstærker for omsætning af strømmen fra en fotodiode⁷⁴ eller et accelerometer⁷⁵ til en spænding, en forstærker på et printkort⁷⁶, eller signalet føres frem via et skærmkabel⁷⁷ til den inverterende indgang eller fra en buffers udgang til et eksternt system. Disse situationer analyseres herunder, og efterfølgende gives en mere generel introduktion til kernen i problemet med stabilitet.

3.9.6 Kapacitet på indgang

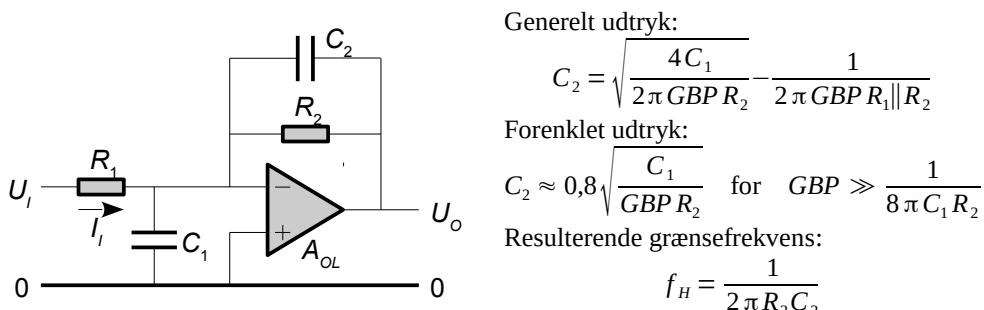
Tilbagekoblingsnetværket danner en pol med kapaciteten på indgangen, og reducerer stabiliteten så der opstår oversving ved en steppåvirkning. Den højest tilladelige værdi af denne kapacitet afhænger af operationsforstærkerens GBP og modstandsniveauet, og beregnes som vist herunder. Grænsen for stabilitet er her defineret som reelle rødder i nævnerens polynomium. For en inverterende forstærker føres signalet ind ved U_i . En transimpedansforstærker får strømmen ført ind som I_i og behøver ikke R_1 .



Figur 144 – En inverterende forstærker kan blive ustabil hvis modstandsniveauet er for højt, og en transimpedansforstærker benyttes ofte med fotodioder hvis indre kapacitet giver ustabilitet.

For R_1 uendelig stor, $R_2 = 1 \text{ M}\Omega$ og en operationsforstærker med $GBP = 10 \text{ MHz}$ kan der accepteres en kondensator over indgangen på 4 fF , hvilket er det samme som "ingen kondensator".

Det er muligt at acceptere en kapacitet C_1 på indgangen med kondensator C_2 i tilbagekoblingen. Prisen for den ubetinget stabilitet er en reduktion i signalets båndbredde.



Figur 145 – Kapaciteten på indgangen giver et uønsket negativt phasedej, men det kan modvirkes af et positivt phasedej gennem en parallelkoblet kapacitet over tilbagekoblingen.

For R_1 uendelig stor, $R_2 = 1 \text{ M}\Omega$ og en operationsforstærker med $GBP = 10 \text{ MHz}$ kan en kapacitet over indgangen på $C_1 = 330 \text{ pF}$ accepteres ved $C_2 = 4,6 \text{ pF}$. Båndbredden af signalet bliver $f_H = 35 \text{ kHz}$.

Beregningseksemplerne forklarer hvorfor det er populært at sætte en lille kondensator i parallel med tilbagekoblingens modstand. Hvis det i praksis viser sig at være en utilstrækkelig løsning, kan

⁷⁴ Fotodioder har et stort areal for at kunne modtage lysenergien, og det medfører en kapacitet fra 100 pF til 1 nF .

⁷⁵ Et accelerometre baseret på den piezoelektriske effekt er opbygget med et smalt krystal mellem to ledende flader.

⁷⁶ Kapaciteten fra en lederbane over et stelplan er af størrelsesordenen 1 pF/cm .

⁷⁷ Det typiske skærmkabel repræsenterer en kapacitet på 100 pF for hver meter kabel.

man prøve at isolere kapaciteten C_1 fra operationsforstærkerens indgang ved at tilslutte den gennem en modstand, frem for direkte til den negative indgang. En værdi af på 100Ω har vist sig effektiv.

Herunder gennemføres analysen for begge kredsløb med en kondensator over indgangen til operationsforstærkeren.

Overføringsfunktionen findes ved at opstille Kirchhoffs strømlov for den negative indgang U_N på operationsforstærkeren.

$$\frac{U_N - U_I}{R_1} + \frac{U_N}{1/sC_1} + \frac{U_N - U_O}{R_2} + \frac{U_N - U_O}{1/sC_2} = 0$$

Det er muligt at eliminere U_N fra ligningen ved at bruge definitionsligningen for en operationsforstærker. Åben-sløjfe forstærkningen er defineret som et lavpasfilter, men kan tilnærmes ved en integration, der er gyldig ved høje frekvenser (over nogle kilohertz).

$$\begin{aligned} U_O &= (U_P - U_N) A_{OL} \\ U_P &= 0 \\ A_{OL} &= \frac{\omega_0}{\omega_0 + s} A_{DC} = \frac{2\pi GBP}{\omega_0 + s} \end{aligned} \Rightarrow \begin{aligned} U_O &= -A_{OL} U_N \\ &\approx -\frac{2\pi GBP}{s} U_N \end{aligned} \Rightarrow U_N \approx -\frac{s}{2\pi GBP} U_O$$

Indsættelse og reduktion viser at forstærkningen som ventet består af faktoren $-R_2/R_1$ for den inverterende forstærker, og dertil et lavpasfilter af anden orden, der viser at der er mulighed for oscillation i stepresponsen.

$$\frac{U_O}{U_I} = -\frac{R_2}{R_1} \frac{\frac{2\pi GBP}{R_2 C_1}}{\frac{2\pi GBP}{R_2 C_1} + \left(\frac{R_1 + R_2}{R_1 R_2 C_1} + R_2 C_2 \frac{2\pi GBP}{R_2 C_1} \right) s + s^2}$$

Lavpasfiltrets grænsefrekvens $\omega_1 = 2\pi f_1$ og dæmpningsfaktor d kan nu defineres.

$$\frac{U_O}{U_I} = -\frac{R_2}{R_1} \frac{\omega_1^2}{\omega_1^2 + 2d\omega_1 s + s^2} \Rightarrow \begin{aligned} \omega_1^2 &= \frac{2\pi GBP}{R_2 C_1} \\ 2d\omega_0 &= \frac{R_1 + R_2}{R_1 R_2 C_1} + \frac{C_2}{C_1} 2\pi GBP \end{aligned}$$

Ubetinget stabilitet opnås for $d = 1$. For situationen uden tilbagekoblingskondensator ($C_2 = 0$) falder andet led ud i tælleren og da en lille værdi af R_1 gør det lettere at opfylde kravet vil en worst case tankegang kunne bruges til at simplificere udtrykket. Når C_1 er kendt, og ikke kan accepteres på grund af ustabilitet, kan den nødvendige værdi af C_2 beregnes som vist.

$$d = \frac{\frac{R_1 + R_2}{R_1 R_2 C_1} + \frac{C_2}{C_1} 2\pi GBP}{2\omega_1} = 1 \stackrel{C_2 = 0}{\Rightarrow} C_1 < \frac{R_2}{8\pi GBP(R_1 || R_2)^2} = \frac{1}{8\pi GBP R_2} \left(1 + \frac{R_2}{R_1}\right)^2 \stackrel{C_2 > 0}{\Rightarrow} C_2 = \sqrt{\frac{2C_1}{\pi GBP R_2}} - \frac{R_1 + R_2}{2\pi GBP R_1 R_2}$$

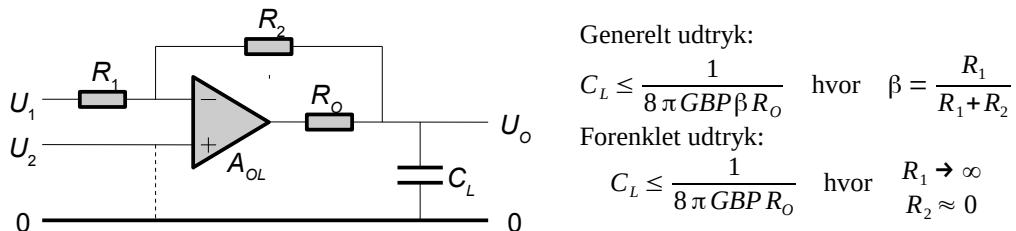
Sidste led i udtrykket for C_2 kan ignoreres når følgende ulighed er opfyldt. Eksempelvis vil det for $C_1 = 10 \text{ pF}$ og $R_2 = 1 \text{ M}\Omega$ kræve at GBP er væsentligt større end 4 kHz .

$$\sqrt{\frac{2C_1}{\pi GBP R_2}} \gg \frac{R_1 + R_2}{2\pi GBP R_1 R_2} \Rightarrow GBP \gg \frac{1}{8\pi C_1 R_2}$$

Analysen er gyldig for en inverterende forstærker og den tilsvarende ikke-inverterende, hvor R_1 er ført til nul og signalet tilsluttes den positive indgang på operationsforstærkeren.

3.9.7 Kapacitet på udgang

En operationsforstærker skal ofte drive et langt kabel hvorved det belastes af kapaciteten, der danner en pol med operationsforstærkerens udgangsmodstand r_o . Det påvirker stabiliteten på samme måde som ved foregående kredsløb. Modstandsværdien er typisk 10 til 50 Ω for en OpAmp og der skal normalt en kondensator til på mere end 100 pF til for at påvirke stabiliteten. Modstanden R_o ses indført for at flytte kondensatoren væk fra operationsforstærkeres udgang, og tilbagekoblingen flyttes med for at undgå spændingsdeling, men opstillingen er alligevel ustabil.



Figur 146 – Kapaciteten på udgangen giver en pol med operationsforstærkerens indre modstand.

Med forstærkeren koblet som spændingsfølger ($R_1 = \infty$ og $R_2 = 0 \Omega$) er $\beta = 1$. Med en udgangsmodstand på $R_o = 30 \Omega$ og båndbredden $GBP = 10 \text{ MHz}$ bør udgangen belastes med højest $C_L = 133 \text{ pF}$.

Beregningen viser at en operationsforstærker kan gøres stabil ved at inkludere forstærkning gennem de to modstande R_1 og R_2 , men det er et tveægget sværd, for med $R_2 > 0$ vil der være en pol ved indgangen og derved introduceres en uønsket fasedrejning i sløjfe-forstærkningen.

Overføringsfunktionen findes ved Kirchhoff's strømlov for operationsforstærkerens negative indgang. Det antage at modstanden R_2 ikke udgør en belastning af lavpasfiltret dannet af operationsforstærkerens indre modstand r_o og kondensator C_L . Spændingen i knudepunktet U_N kan udledes fra operationsforstærkerens definitionsligning, som vist i det tidligere eksempel, men nu skal polen fra udgangen også inkluderes i udtrykket.

$$\frac{U_N - U_I}{R_1} + \frac{U_N - U_O}{R_2} = 0 \quad \text{og} \quad U_O = -\frac{\omega_L}{\omega_L + s} \frac{2\pi GBP}{s} U_N \quad \text{hvor} \quad \omega_L = \frac{1}{R_o C_L}$$

Efter reduktion findes overføringsfunktionen, der beskriver en inverterende forstærker med et anden-ordens lavpasfilter.

$$\frac{U_O}{U_I} = -\frac{R_2}{R_1} \frac{\beta 2\pi GBP \omega_L}{\beta 2\pi GBP \omega_L + \omega_L s + s^2} \quad \text{hvor} \quad \beta = \frac{R_1}{R_1 + R_2}$$

Anden-ordens relationen skrives på standardformen hvorved grænsefrekvensen $\omega_0 = 2\pi f_0$ og dæmpningen d kan identificeres.

$$\frac{U_O}{U_I} = -\frac{R_2}{R_1} \frac{\omega_0^2}{\omega_0^2 + 2d\omega_0 s + s^2} \Rightarrow \frac{\omega_0^2}{2d\omega_0} = \omega_L$$

For ubetinget stabilitet kræves en dæmpning på én.

$$d = \frac{\omega_L}{2\omega_0} \geq 1 \Rightarrow \frac{\omega_L}{2\sqrt{\beta 2\pi GBP \omega_L}} \geq 1 \Rightarrow \frac{\omega_L}{\beta 2\pi GBP} \geq 4$$

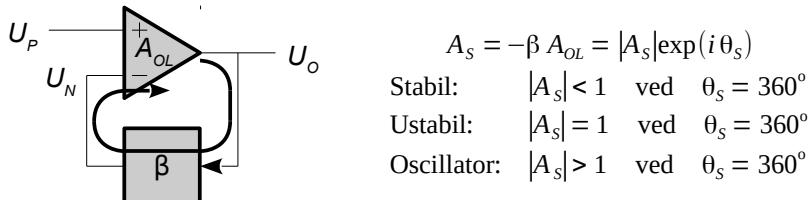
Herfra bestemmes kravet til grænsefrekvensen af polen (hvor $\omega_L = 2\pi f_L$).

$$f_L \geq 4\beta GBP \Rightarrow \frac{1}{2\pi R_o C_L} \geq 4\beta GBP \Rightarrow C_L \leq \frac{1}{8\pi GBP \beta R_o}$$

Parameteren β kan ikke komme over én, som derfor kan benyttes for at beregne den laveste værdi af den højest tilladelige kondensator på udgangen.

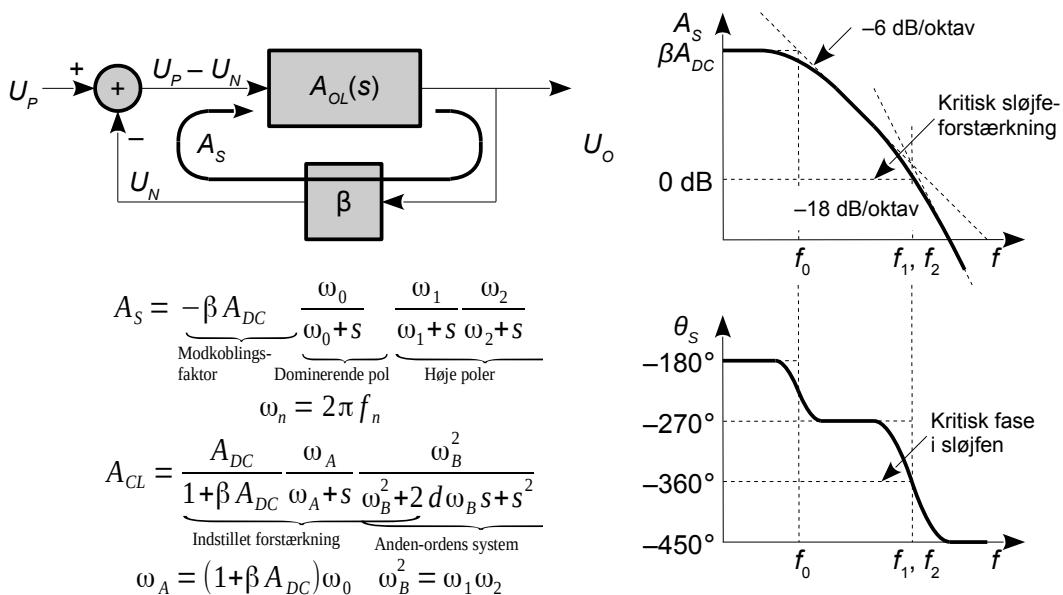
3.9.8 Fasemargin

En operationsforstærker arbejder normalt i frekvensområdet over den dominerende pol, så forstærkningen er omvendt proportional med frekvensen og fasedrejet er tæt på -90° . Det tilbagekoblede signal benytter den inverterende indgang, der yderligere drejer signalet -180° (inversion), så der mangler kun -90° for at det tilbagekoblede signal kan understøtte en oscillation. Betingelsen for stabilitet blev givet af Nyquist som kravet om at forstærkningen rundt i sløjfen skal være under én ved den frekvens hvor fasedrejet totalt er 360° . En forstærker, der opfylder Nyquist kriteriet vil ikke oscillere, men den kan være længe om at falde til ro efter en påvirkning.



Figur 147 – Sløjfe-forstærkningen er et mål for hvor stabil operationsforstærkeren er i et givet kredsløb, og stabilitet kræver at værdien er under én inden fasen nå op på 360° .

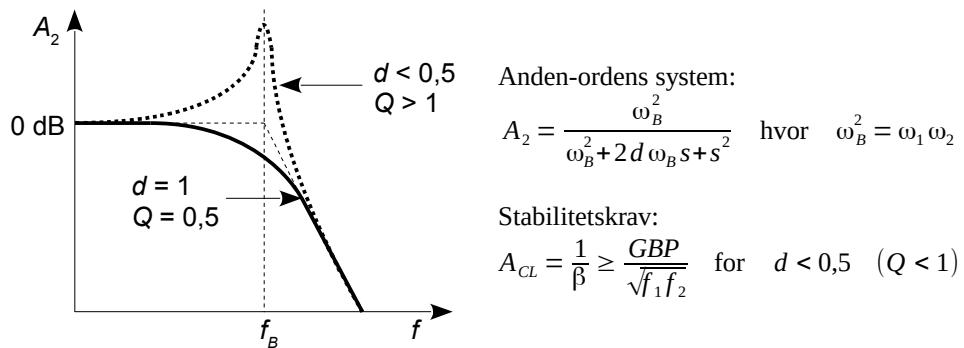
Forholdene beskrives gennem sløjfe-forstærkningen A_S , der er tilbagekoblingens andel β af udgangssignalet gange med åben-sløjfe forstærkningen A_{OL} . Den giver et godt billede af de betingelser som operationsforstærkeren arbejder under. Herunder vises en OpAmp med dominerende pol f_0 og to høje poler f_1 og f_2 , der kan være interne poler i den benyttede OpAmp, eller de kan stamme fra det eksterne kredsløb. Fasedrejet erkonstant på -270° fra den dominerende pol og forstærkerens inversion, indtil de høje poler bidrager med et fasedrej, og ved -360° skal sløjfe-forstærkningen være reduceret til under én for stabilitet.



Figur 148 – Sløjfe-forstærkningen angiver hvordan operationsforstærkeren opfatter omgivelserne og kan bruges til at beregne hvorvidt den er stabil eller vil oscillere. Den beregnes ved at bryde sløjfen og opstille ligningen fra den ene side af bruddet til den anden. Sløjfe-forstærkningen benytter f_0 for den dominerende pol (under 1 kHz), og polerne f_1 og f_2 repræsenterer de højtliggende poler i operationsforstærkeren (over 1 MHz) eller poler i tilbagekoblingens netværk.

Det resulterende udtryk for overføringsfunktionen fra ind- til udgang vises i eksemplet herunder at kunne skrives som forstærkning og båndbredde af den tilbagekoblede operationsforstærker med en dominerende pol, og dertil et anden-ordens system, der angiver muligheden for ustabilitet.

Operationsforstærkeren kan ikke dimensioneres som ubetinget stabil med højtliggende poler, så man må nøjes med mindre. Der opnås et marginalt stabilt system (45° fasemargin) hvis operationsforstærkeren er indstillet til en forstærkning A_{CL} over en mindste værdi⁷⁸, der er givet ved 0 dB båndbredden GBP og det geometriske gennemsnit af de høje poler. Resultatet benyttes i kapitel fem for at kunne placere den dominerende pol korrekt i en operations- eller effektforstærker.



Figur 149 – Ustabilitet ved den tilbagekoblede forstærker vil vise sig som en resonans ved en given frekvens, hvor systemets poler giver et lukket-sløjfe phasedrej nær ved 360° .

En fasemargin på 90° er en ønskesituations, der kun kan nås hvis tilbagekoblingens modstande holdes på lave værdier og monteres tæt på operationsforstærkeren, samt hvis den benyttede OpAmp opfører sig "pænt" i den høje ende af frekvensspektret. Det er som en tommelfingerregel tilstrækkeligt med en fasemargin på 45° hvor der kun er et lille oversving i stepresponsen. Hvis opstillingens fasemargin kommer ned under 45° bør designet revideres.

Herunder udledes kravet.

For at indlede analysen ganges de enkelte led sammen til følgende udtryk.

$$A_{OL} = \frac{\omega_0}{\omega_0 + s} \frac{\omega_1}{\omega_1 + s} \frac{\omega_2}{\omega_2 + s} A_{DC} \quad \text{hvor} \quad f_n = \frac{\omega_n}{2\pi} \Rightarrow$$

$$A_{OL} = \frac{\omega_0 \omega_1 \omega_2 A_{DC}}{\omega_0 \omega_1 \omega_2 + (\omega_0 \omega_1 + \omega_0 \omega_2 + \omega_1 \omega_2) s + (\omega_0 + \omega_1 + \omega_2) s^2 + s^3}$$

Definitionen indføres i udtrykket for den tilbagekoblede forstærker.

$$A = \frac{A_{OL}}{1 + \beta A_{OL}} = \frac{\omega_0 \omega_1 \omega_2 A_{DC}}{(1 + \beta A_{DC}) \omega_0 \omega_1 \omega_2 + (\omega_0 \omega_1 + \omega_0 \omega_2 + \omega_1 \omega_2) s + (\omega_0 + \omega_1 + \omega_2) s^2 + s^3}$$

Konstanten i nævnerens polynomium definerer grænsefrekvensen f_P for kredsløbet.

$$\omega_P^3 = (1 + \beta A_{DC}) \omega_0 \omega_1 \omega_2 \Rightarrow f_P = \sqrt[3]{(1 + \beta A_{DC}) f_0 f_1 f_2}$$

Tælleren bringes til at være ω_P^3 ved at gange og dividere med $1 + \beta A_{DC}$. Det giver et udtryk med forstærkningen indstillet ved β og A_{DC} , som her kaldet A_0 . Det udnyttes derefter at den dominerende pol er meget lavere end de to høje poler, så ω_0 kan ignoreres ved s og s^2 .

$$A \approx A_0 \frac{\omega_P^3}{\omega_P^3 + \omega_1 \omega_2 s + (\omega_1 + \omega_2) s^2 + s^3}, \quad A_0 = \frac{A_{DC}}{1 + \beta A_{DC}}, \quad \omega_0 \ll \sqrt{\omega_1 \omega_2}$$

Overføringsfunktionen beskrives ved et første-ordens led A_1 og et anden-ordens led A_2 .

$$A_3 \stackrel{\text{def}}{=} A_0 A_1 A_2 = A_0 \frac{\omega_A}{\omega_A + s} \frac{\omega_B^2}{\omega_B^2 + 2d\omega_B s + s^2}, \quad \omega_A = (1 + \beta A_{DC}) \omega_0 \quad \omega_B^2 = \omega_1 \omega_2$$

⁷⁸ For LF357 er mindste forstærkning 5 gange og for LMH6624 er den 10 gange. En effektforstærker designes normalt til en forstærkning på 20 gange og behøver derfor ikke være stabil ved en forstærkning på kun én gang.

Det giver to nye poler, hvor ω_A repræsenterer den tilbagekoblede forstærkers båndbredde og ω_B står for det nye anden-ordens leds grænsefrekvens, der her skal defineres som det geometriske gennemsnit mellem ω_1 og ω_2 hvilket vil vise sig hensigtsmæssigt. Det nye udtryks to led ganges sammen, for at finde relationerne imellem parametrene.

$$A_3 = A_0 \frac{\omega_A \omega_B^2}{\omega_A \omega_B^2 + (2d \omega_A \omega_B + \omega_B^2)s + (\omega_A + 2d \omega_B)s^2 + s^3}$$

Sammenlignes de to udtryk ses at $\omega_A \omega_B^2$ svarer til ω_p^3 og anden-ordens leddets dæmpning d kan findes fra koefficienten til s_2 (koefficienten til s giver blot $d = 0$).

$$\omega_A + 2d\omega_B = \omega_1 + \omega_2 \Rightarrow d = \frac{\omega_1 + \omega_2 - \omega_A}{2\sqrt{\omega_1 \omega_2}} \approx 1$$

For en meget lav værdi af ω_A sammenlignet med ω_1 og ω_2 er dæmpningen givet af forholdet mellem det aritmetiske gennemsnit af de to høje poler og det geometriske gennemsnit af de samme poler. Hvis polerne ligger nogenlunde lige højt vil dæmpningen være tæt på $d = 1$ og systemet er stabilt. Hvis de to højliggende poler har stor indbyrdes afstand bliver resultatet blot en større dæmpning, og dermed en endnu større grad af stabilitet; sagt anderledes så bliver forstærkeren mere sløv i den høje ende, og den vil glide langsomt på plads uden oversving efter en påvirkning. Problemet kommer når det geometriske gennemsnit af de to høje poler ikke er i sikker afstand fra den tilbagekoblede forstærkers grænsefrekvens.

Første-ordens polen ω_A er givet ved tilbagekoblingens $1 + \beta A_{DC}$ og den dominerende pols frekvens ω_0 , så det er ikke sikkert at den ligger lavt i forhold til de to høje poler. Det er ikke muligt at opnå ubetinget stabilitet med to højliggende poler ($d = 1$ eller tilsvarende $Q = 0,5$), for den dominerende pol alene giver en fasemargin på 90° , så der sigtes efter $d \geq 0,5$ hvor der vil være et mindre oversving ($Q < 1$). Det antages igen at de to høje poler er nogenlunde ens, så forholdet mellem de aritmetiske og geometriske middelværdier er nær ved én.

$$d = \frac{\omega_1 + \omega_2 - \omega_A}{2\sqrt{\omega_1 \omega_2}} \geq 0,5 \Rightarrow 1 - \frac{(1 + \beta A_{DC})\omega_0}{2\sqrt{\omega_1 \omega_2}} \geq 0,5 \Rightarrow \frac{(1 + \beta A_{DC})\omega_0}{2\sqrt{\omega_1 \omega_2}} \leq 0,5$$

For at opnå et mere brugbart udtryk udnyttes at DC forstærkningen er så stor at $1 + \beta A_{DC}$ kan erstattes med βA_{DC} , og det leder til et krav mellem den indstillede forstærkning A_{CL} og forstærkerens specifikation for $GBP = A_{DC}f_0$ samt de to højliggende poler.

$$\frac{\beta A_{DC} f_0}{\sqrt{f_1 f_2}} \leq 1 \Rightarrow \beta A_{DC} f_0 \leq \sqrt{f_1 f_2} \Rightarrow A_{CL} = \frac{1}{\beta} \geq \frac{GBP}{\sqrt{f_1 f_2}}$$

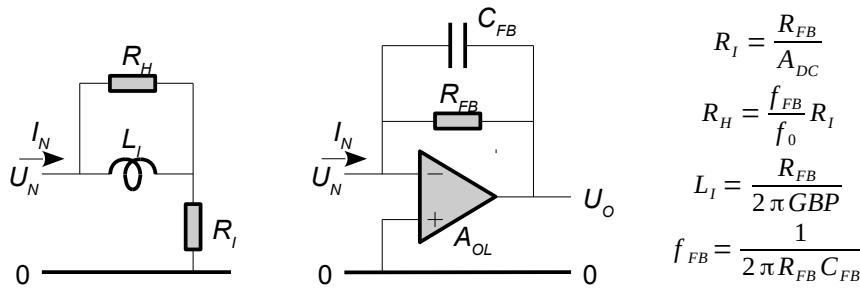
Hermed afsluttes analysen af båndbredde og stabilitet. De efterfølgende afsnit beskæftiger sig med ind- og udgangsimpedansen fra en tilbagekoblet operationsforstærker med dominerende pol, samt de begrænsninger dens støjkilder og effektforsyningen giver.

3.10 Ind- og udgangsimpedans

I det indledende afsnit er det forudsat at den ikke-inverterende forstærker har en meget høj indgangsmodstand, at den inverterende forstærker holder den negative indgang på nul, og at udgangen kan regnes for en ideel spændingskilde. Praksis er dog ikke helt så ideel, og det skyldes den endelige værdi af åben-sløjfe forstærkningen A_{OL} og den dominerende pol f_0 .

3.10.1 Inverterende forstærker

Indgangsimpedansen ved den virtuelle jord er givet ved impedansen af tilbagekoblingen divideret med åben-sløjfe forstærkningen A_{OL} . Ved meget lave frekvenser er værdien under 1Ω hvis tilbagekoblingsmodstanden er under cirka $100 \text{ k}\Omega$ og operationsforstærkerens DC forstærkning A_{DC} er på mindst 10^5 . I den aktive del af frekvensområdet er operationsforstærkeren integrerende, og indgangsimpedansen bliver nu en differentieret version af tilbagekoblingen (sZ_{FB}). En kondensator i tilbagekoblingen ($Z_{FB} = 1/sC_{FB}$) giver en reel værdi over indgangen, hvilket er en modstand, og ved en modstand i tilbagekoblingen ($Z_{FB} = R_{FB}$) er værdien proportional med frekvensen, hvilket er en selvinduktion i parallel over indgangen, der dog dæmpes at en modstandsværdi.



Figur 150 – En almindelig opstilling med begrænsning af båndbredden vil have en kompleks indgangsimpedans, der er induktiv i et frekvensområde.

Med $R_{FB} = 1 \text{ M}\Omega$ og en forstærker med $A_{DC} = 10^6$ og $G_BP = 1 \text{ MHz}$ vil der være en indgangsmodstand bestående af $R_I = 1 \Omega$ i serie med selvinduktionen $L_I = 0,16 \text{ H}$. Selvinduktionens reaktans vil nå 10Ω allerede ved 100 Hz så indgangen er ikke virtuel nul ret langt op i frekvens.

Åben-sløjfe forstærkningen udtrykkes først ved DC forstærkningen A_{DC} og en dominerende pol $f_0 = \omega_0/2\pi$, der almindeligvis er under 100 Hz . Indført i udtrykket beregnes nu en reel og en imaginær del af impedansen.

$$Z_I = \frac{R_{FB}}{A_{OL}} \quad \text{og} \quad A_{OL} = \frac{\omega_0}{\omega_0 + s} A_{DC} \quad \Rightarrow \quad Z_I = \frac{R_{FB}}{A_{DC}} \frac{\omega_0 + s}{\omega_0} = \frac{R_{FB}}{A_{DC}} \left(1 + \frac{s}{\omega_0}\right)$$

Den reelle del er indgangsmodstanden R_I . Den imaginære del ses at være proportional med frekvensen og fungerer som impedansen af en selvinduktion, der her kaldes L_I og defineres ved brug af operationsforstærkerens $G_BP = f_0 A_{DC}$.

$$Z_I = \frac{R_{FB} + s \frac{R_{FB}}{A_{DC} \omega_0}}{A_{DC}} = R_I + sL_I \quad \Rightarrow \quad R_I = \frac{R_{FB}}{A_{DC}}, \quad L_I = \frac{R_{FB}}{2\pi G_BP}$$

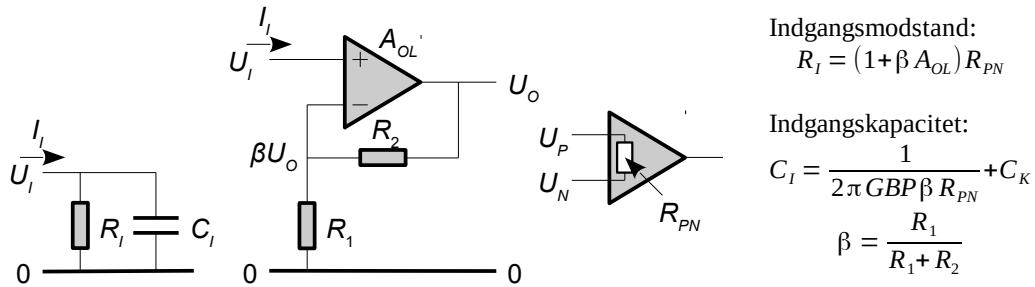
En parallelforbindelse af en modstand og en kondensator i tilbagekoblingen er hyppigt anvendt for båndbegrensning. Indgangsimpedansen bestemmes ved at indsætte parallelforbindelsen af R_{FB} og $1/sC_{FB}$ i udtrykket Z_{FB}/A_{OL} , og A_{OL} beskrives ved A_{DC} og den polen ω_0 .

$$Z_I = \frac{\omega_0 + s}{A_{DC} \omega_0} \frac{R_{FB}/sC_{FB}}{R_{FB} + 1/sC_{FB}} = \frac{R_{FB}}{A_{DC}} \frac{\omega_0 + s}{\omega_0} \frac{1/R_{FB} C_{FB}}{s + 1/R_{FB} C_{FB}} = \frac{R_{FB}}{A_{DC}} \frac{\omega_0 + s}{\omega_0} \frac{\omega_{FB}}{s + \omega_{FB}}$$

Ved høj grænsefrekvens af f_{FB} er indgangsimpedansen en selvinduktion med værdien L_I indtil frekvensen når op på grænsefrekvensen for R_{FB} og C_{FB} , hvorover værdien bliver reel (R_H).

3.10.2 Ikke-inverterende forstærker

For den ikke-inverterende forstærker er indgangsimpedansen givet ved operationsforstærkerens indre modstand R_{PN} mellem de to indgange og dertil produktet βA_{OL} , der angiver hvor kraftigt der er tilbagekoblet. En operationsforstærker har en indgangsmodstand langt over 1 MΩ selv ved en relativt lav værdi af R_{PN} , så indgangsmodstanden R_I vil oftest kunne ignoreres. Den dominerende pol repræsenterer en omvendt proportionalitet mellem indgangsmodstanden og frekvensen, hvilket kan beskrives som indgangens parallelle kapacitet C_I , der ligger i parallel med den specificerede værdi for operationsforstærkerens kobling til omgivelserne C_K , der er omkring 5 pF.



Figur 151 – Indgangsimpedansen på en ikke-inverterende forstærker er høj på grund af en positiv tilbagekobling, der effektivt set holder spændingsfaldet over modstanden mellem de to indgange på en OpAmp nær ved nul så strømmen i indgangen også bliver nær nul.

Med $GBP = 1 \text{ MHz}$, $\beta = 1$, $A_{OL} = 10^5$ og $R_{PN} = 10 \text{ k}\Omega$ er indgangsmodstanden $R_I = 1 \text{ G}\Omega$ og kapaciteten over indgangen er $C_I = 16 \text{ pF}$ så impedansen er tilbage på $10 \text{ k}\Omega$ ved 1 MHz.

Indgangstrinnet på en operationsforstærker kan repræsenteres ved en modstandsværdi mellem de to indgange (R_{PN}), med en værdi startende ved $10 \text{ k}\Omega$ for bipolare operationsforstærkere med høj strøm i indgangstrinnet, og stigende mod $10^{12} \Omega$ for felt-effekt transistorer.

Indgangsimpedansen beregnes ved at føre en strøm I_I ind på indgangen og beregne spændingen på indgangen. Strømmen løber igennem modstanden R_{PN} mellem de to indgange hvor den giver et spændingsfald på $R_{PN}I_I$ og indgangsspændingen bygges op som summen af dette spændingsfald og spændingen på den negative indgang. Strømmen ind i U_P vil løbe ud af U_N til tilbagekoblingens netværk, men det antages at R_1 er så lille i forhold til R_{PN} at bidraget kan ignoreres. Spændingen på den negative indgang er givet af udgangens spænding og tilbagekoblingen, beskrevet ved β .

$$U_I = R_{PN} I_I + U_N = R_{PN} I_I + \beta U_O$$

Udgangsspændingen er givet af differensen U_P og U_N gange med åben-sløjfe forstærkningen A_{OL} og da differensen er givet ved $R_{PN}I_I$ haves et bekvemt udtryk for udgangssignalet.

$$U_O = (U_P - U_N) A_{OL} = R_{PN} I_I A_{OL}$$

Indgangsspændingen kan nu udtrykkes ved kendte størrelser og indgangsmodstanden findes ved at dividere med strømmen i indgangen.

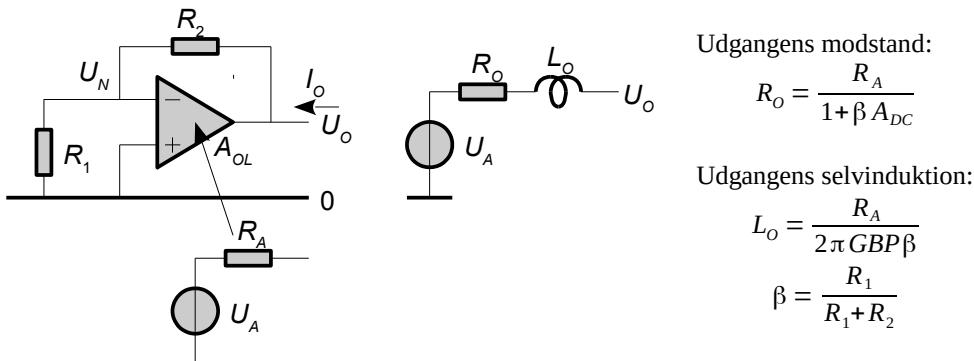
$$R_I = \frac{U_I}{I_I} = (1 + \beta A_{OL}) R_{PN}$$

Indsættes udtrykket for A_{OL} med en dominerende pol ses at indgangsimpedansen falder med stigende frekvens svarende til at der sidder en kondensator i parallel over indgangen.

$$Z_I = (1 + \beta \frac{\omega_0}{\omega_0 + s} A_{DC}) R_{PN} \approx \beta \frac{\omega_0}{s} A_{DC} R_{PN} = \frac{1}{s C_I} \Rightarrow C_I = \frac{1}{2\pi G B P \beta R_{PN}}$$

3.10.3 Udgangsimpedans

En operationsforstærker har en fysisk modstand i udgangen for at kunne måle udgangsstrømmen i forbindelse med et internt beskyttelseskredsløb. Værdien af udgangsmodstanden er 10 til 50 Ω hvortil kommer en dynamisk modstandsværdi fra udgangstrinnets transistorer. En tilbagekobling udnytter operationsforstærkerens store åben-sløjfe forstærkning til at reducere den effektive værdi af udgangsmodstanden til nær ved nul, men den begrænsede forstærkning ved høje frekvenser vil delvist opnæve korrektionen, og det kan ses som en induktiv udgangsimpedans.



Figur 152 – Udgangsimpedansen på en forstærker reduceres på grund af tilbagekoblingen, men den dominerende pol i operationsforstærkeren gør den resulterende impedans induktiv.

Ved $R_A = 50 \Omega$, $\beta = 1$, $A_{DC} = 10^5$ og $GBP = 1 \text{ MHz}$ er udgangsmodstanden $R_O = 5 \text{ m}\Omega$ ved $\beta = 0,1$. Den ækvivalente selvinduktion er $L_O = 80 \mu\text{H}$ ved $\beta = 0,1$, så impedansen når $2\pi f L_O = 500 \Omega$ ved 1 MHz.

Beregning af udgangsimpedansen starter med ligningen for forstærkerens udgangsspænding, dog benyttes U_A for forstærkerens indre spændingskilde, da U_O står for udgangsspændingen. Den positive indgang ligger på nul, men beregningen vil passe både for den inverterende forstærker med et indgangssignal tilført ved R_1 samt for den ikke-inverterende forstærker med et signal tilført ved den positive indgang. Den negative indgang modtager brøkdelen β af udgangsspændingen.

$$U_A = (U_P - U_N) A_{OL} = -U_N A_{OL} = -\beta U_O A_{OL}$$

Udgangsimpedansen findes ved at lede en ekstern strøm I_O ind i udgangen og finde den spænding som udgangen stiller sig på. Det antages at værdien af $R_1 + R_2$ er så stor i sammenligning med R_A at strømmen I_O kun løber igennem R_O . Udgangsspændingen kan beregnes som U_A plus spændingen over udgangsmodstanden.

$$U_O = R_A I_O + U_A = R_A I_O - \beta A_{OL} U_O \Rightarrow U_O (1 + \beta A_{OL}) = R_A I_O$$

Udgangsmodstanden bliver reduceret med tilbagekoblingen udtrykt ved $1 + \beta A_{OL}$.

$$Z_O = \frac{U_O}{I_O} = \frac{R_A}{1 + \beta A_{OL}}$$

Indsættes udtrykket for åben-sløjfe forstærkningen med en dominerende pol ses at impedansen stiger med frekvensen fra ovenstående værdi indtil R_A , der nås ved en frekvens på $1 + \beta A_{DC}$ gange den dominerende pol. Det udnyttes at $GBP = A_{OL} f_0$ og resultatet gengiver den reelle del som R_O samt en serieforbundet induktion.

$$Z_O = \frac{R_A}{1 + \beta \frac{\omega_0}{\omega_0 + s} A_{DC}} = R_A \frac{\omega_0 + s}{(1 + \beta A_{DC}) \omega_0 + s} \underset{\substack{s \gg \omega_0 \\ \beta A_{DC} \gg 1}}{\approx} R_A \frac{s}{\beta \omega_0 A_{DC}} \Rightarrow L_O = \frac{R_A}{2\pi GBP\beta}$$

Operationsforstærker

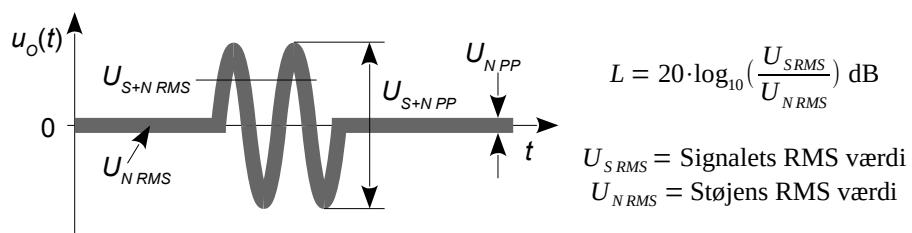
Note 4.

3.11 Støj

I det følgende introduceres støjen og dens matematik. Hvis du søger en hurtig vej til beregning af støj i et elektronisk kredsløb så spring frem til side 152 eller til sammenfatningen side 162.

Variationen i et hørbart signal som tale og musik beskrives ved dets dynamik, der er lig forholdet mellem den kraftigste og svageste amplitude. For et menneske er signalets niveauet opad begrænset ved smertegrænsen og nedad ved høretærsklen⁷⁹ til et område på 140 dB svarende til en variation i signalamplituden på 10⁷:1. Normalt vil det svage signal dog være maskeret af lyden fra andre mennesker i samme rum, en tændt PC på skrivebordet eller trafikken udenfor vinduet. Når lyden er forstyrrende kaldes den for støj.

Der gælder tilsvarende for de elektriske signaler i målekredsløb; at signalet kan være så svagt at det ikke kan detekteres på grund af den omgivende støj. Det sætter den nedre grænse ved måling af svage signaler fra termokoblere og piezoelektriske transducere, og ved datakommunikation kan støj påvirke de digitale mønstre så information kan gå tabt.



Figur 153 – Ethvert signal består af det ønskede signal overlejret en evigt fluktuerende støj.

Der er flere kilder til støj i elektroniske kredsløb. Der genereres en termisk støj i enhver elektrisk leder som skyldes elektronernes vibration på grund af omgivelserstemperaturen, og der genereres støj fra elektrolytens ladning, idet strømmen bæres af et heltaligt antal elektroner så strømmen er kvantificeret. Disse to kilder er veldokumenterede og kan behandles relativt enkelt. Andre støjkilder er omrøkering af elektrostatiske ladninger i halvlederen og ændringer i halvlederens krystalstruktur, der begge giver den såkaldte popcorn-støj, som dominerer ved lave frekvenser og beskrives under betegnelsen 1/f-støj. Dertil kan eksterne støjkilder koble elektromagnetisk til elektronikken, som fx den knurrende lyd i fra en tændt mobiltelefon i nærheden af en PC højtaler, eller gnister fra en lysnetkontakt. Komponenterne kan have en piezoelektrisk effekt, der omsætter en ekstern mekanisk påvirkning til en spænding. Endelig kan indgangssignalet selv være støjforurenset, eksempelvis en mikrofonoptagelse af en koncert med publikumsstøj. Støj er uønskede signaler og inkluderer virkningen af en DC fejfspænding og biasstrøm, selv om man normalt vil analysere de to situationer hver for sig. DC fejkilderne vil variere over tid og vil derfor opfattes som lavfrekvent støj.

Signalets signal-støj forhold er defineret ved forholdet mellem signalets RMS niveau U_S og støjens RMS niveau U_N , og det er normalt at udtrykke det i decibel. Det maksimale signal-støj forhold for et elektronisk kredsløb er givet ved overstyringsgrænsen delt med støjgulvet. Et signal er overstyret, når det begrænses (og dermed forvrænges) ved et forsøg på at hæve niveauet yderligere, og støjgulvet er det niveau hvorunder signalet mistes, hvis signalets niveau forsøges sänket. Som det blev vist i første kapitel kræver bestemmelse af RMS værdien integration over lang tid, så vi har forladt den trygge tilværelse med harmoniske signaler med en veldefineret amplitude, og må nu begrænse ambitionen til at betragte et signals middeleffekt over et tidsrum.

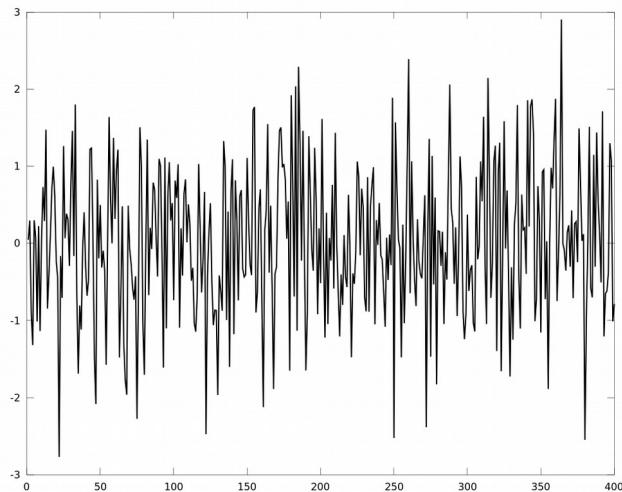
I en matematisk beskrivelse er støj en sum af et meget stort antal uafhængige signaler. Et støjsignal beskrives ved at indeholde en vis middeleffekt ved enhver båndbredde, og i analogi med synligt lys er støjen hvid (*white noise*) hvis der er en konstant effekt uanset hvor båndbredde ligger. Ved dominans af lave frekvenser kaldes støjen for rosa eller lyserød støj (*pink noise*).

⁷⁹ Smertegrænsen ligger cirka 140 dB over høretærsklen 0 dB, der er defineret som 20 µPa på grundlag af måling af et meget stort antal unge og raske mennesker. Lydtryk over smertetærsklen kan forårsage blivende ødelæggelse af høresansen, og vi vil forsøge at beskytte os imod smerten ved høje lydtryk. Et niveau under høretærsklen kan ikke erkendes. Ved omgivende støj optræder der en maskering af de svageste signaler, så vi ikke kan erkende signaler væsentligt under den omgivende støj. Et eksempel kunne være en samtale mellem to mennesker, der går langs en strand, hvor havets brusen overdøver de svage passager.

3.11.1 Støjens matematik

På grund af støjens tilfældighed kan matematikken virke lidt underlig, så det er vigtigt at notere sig, at støjens middeleffekt er det eneste man kan udtrykke sig nogenlunde sikkert om. Det betyder at støjberegningen ikke giver præcise resultater, men det lykkelige er at det er yderst sjældent at man har behov for at beregne støjens niveau med større præcision. Det er ofte tilstrækkeligt at vide om støjgulvet er "lavt nok", så i det følgende vil det blive antaget at en præcision på $\pm 20\%$ er nok, og det svarer til at signal-støj forholdet blot skal beregnes med en præcision på ± 2 dB.

Støj kan ikke beskrives ved en formel, sådan som fx en harmonisk svingning kan beskrives ved en sinus-funktion⁸⁰. For at kunne analysere støjens niveau benyttes et statistisk værktøj for at estime re middelværdien af støjens effekt; det vil sige støjspaændingens RMS værdi. Støjsignalet måles over et tidsrum og beskrives ved to parametre: middelværdien signalet svinger omkring og varian sen (signalets RMS værdi) der er et mål for hvor kraftige udsvingene er omkring middelværdien. Som et eksempel vises herunder et støjsignal hvor en funktion beregner tilfældige værdier, der opfylder nogle betingelser for tilfældigheden svarende til standardfordelingen (se senere). Der er egentlig tale om enkelte punkter langs med den horisontale akse, men af hensyn til synligheden er disse punkter forbundet med streger, så det ligner udskriften på et oscilloskop.



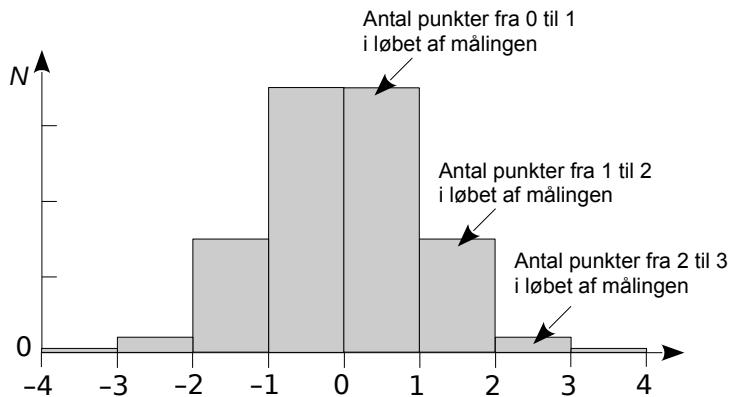
Figur 154 – Støjens er her afbildet ved hjælp af et MATLAB program hvor `randn(1:400)` giver 400 normalfordelte værdier med variansen én og middelværdien nul.

Ved at betragte billedet kan man se at signalet svinger op og ned omkring en middelværdi, der i eksemplet er nul, samt at spidserne af signalet når ud til cirka ± 3 . Hvis forløbet strækker sig over et længere forløb vil spidserne af signalet nå længere ud, men der vil være langt mellem de ekstreme værdier. Det ses også at der både er hurtige skift mellem store positive og negative værdier samt at der er nogle mere rolige forløb; det indikerer at der er involveret både høje og lave frekvenser.

For at beskrive dette uregelmæssige forløb kan man måle antallet af punkter, der rammer indenfor et interval fra 0 til 1, derefter antallet af punkter fra 1 til 2 og så videre. Når måleserien stoppes kan man angive de optalte værdier i et histogram som skitseret herunder. Søjernes bredde angiver et måleområde og højden angiver antallet af målepunkter, der faldt indenfor området. Forløbet er i en ideel verden symmetrisk omkring nul, men der skal mange målepunkter til før det opnås.

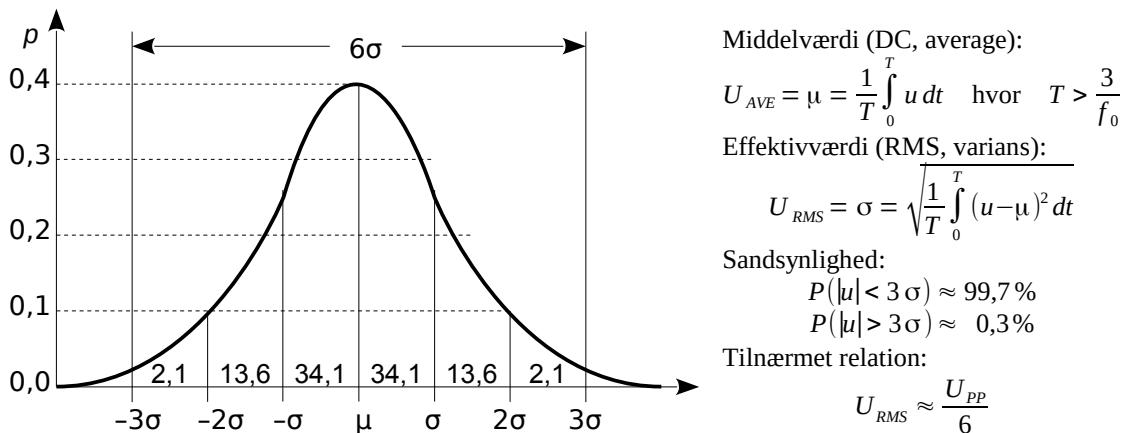
⁸⁰ Der findes algoritmer, som gennemløber en binær sekvens, der tolket som et analogt signal opfører sig som støj. Det kaldes for *pseudo-random noise*, og danner af *maximum-length sequence* (MLS), og har betydning ved måling af blandt andet en højtalers frekvensrespons og kryptering af en datatransmission. Fordelen er at et MLS signal både opfører sig som et støjsignal ved en given båndbrede og er deterministisk. Forskellen fra et "ægte" støjsignal er at forløbet er kendt og kan repeteres.

Hvis målepunkternes antal i hver søjle divideres med det samlede antal punkter i hele måleserien så opnås et mål for hvor hyppigt en måleværdi optræder indenfor et givet interval, og hvis arealet af flere søjler adderes findes et mål for sandsynligheden for at møde en måleværdi indenfor det tilsvarende interval på den vandrette akse.



Figur 155 – Støjens spændingsværdi udtrykkes ved hyppigheden for at den målte værdi er indenfor et givet interval i løbet af det tidsrum der måles over.

Arealet af søjlen angiver sandsynligheden for at måleværdien ligger inden for det interval søjlen repræsenterer. Øges antallet af intervaller, der samtidigt bliver mindre, går billedegradvist imod en kontinuerlig kurve kaldet for en standardfordeling (*standard distribution*). Sandsynligheden for at møde et signal indenfor et interval er givet ved arealet under kurven indenfor intervallet, og er her illustreret ved procentværdier for områder langs den vandrette akse i spring af variansen σ .



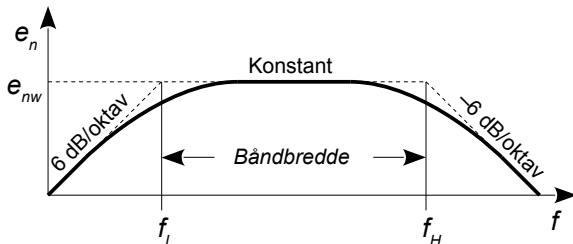
Figur 156 – Standardfordelingen med middelværdien μ , der er lig med DC værdien som et signal svinger omkring, og variansen σ , der er lig med RMS værdien. Ordinaten angiver tætheden, og for at finde sandsynligheden skal der integreres over et interval på abscissen. Ved integration over intervaller lig med variansen bliver procentsatsen som vist i figuren (afrundet til én decimal).

Sandsynligheden er på 34 % for at et signal ligger fra middelværdien og ét standardafvigelse væk i hver retning så sandsynligheden er 68 % for at signalet ligger indenfor intervallet $\pm U_{RMS}$. Sandsynligheden stiger til 95 % for at signalet ligger indenfor $\pm 2 U_{RMS}$ og 99,7 % ved $\pm 3 U_{RMS}$. Det er derfor meget sandsynligt at signalet indenfor et tidsrum er mindre end tre gange RMS værdien, og det er lidet sandsynligt (0,3 %) at det i samme tidsrum kommer udenfor grænserne.

Formlerne angiver at støjens skal følges over et tidsrum T , og da støjsignalet er uregelmæssigt skal tiden være lang i forhold til perioden af den laveste frekvens f_L i målingen, som vist i første kapitel. I praksis begrænses tidsrummet af nøjagtigheden og den acceptable fejl, og ved at vente i 1 sekund svarer det til at den lave grænsefrekvens bliver på 3 Hz ved 10 % fejlmargin.

3.11.2 Termisk støj

Det er umuligt at bygge et elektronisk kredsløb uden støj, for støjen genereres af de elektroner, der forefindes i materialet som enhver elektrisk leder er opbygget af. Elektronerne er i en konstant bevægelse på grund af temperaturen, og der optræder derfor en evigt fluktuerende omfordeling af ladning i lederen. Det betyder at selv en metaltråd genererer støj. Formeludtrykket er givet i figuren herunder for spændingsstøjtætheden e_n fra en modstand R , og den tilsvarende strømstøjtæthed i_n når modstanden kortsluttes. Støjens middeleffekt er uafhængig af i hvilken ende af frekvensspektret der måles; det kaldes for vid støj (*white noise*), så RMS værdien af spændingsstøjen U_n , eller RMS værdien strømstøjen I_n , er givet af båndbredden B , der benyttes ved målingen.



Termisk støjtæthed:

$$e_{nR} = \sqrt{4kTR}$$

$$i_{nR} = \sqrt{\frac{4kT}{R}} \quad \text{hvor } T \approx 300 \text{ K}$$

Støjspænding og støjstrøm (RMS):

$$U_n = e_n \sqrt{B} \quad \text{hvor } B = f_H - f_L$$

$$I_n = i_n \sqrt{B}$$

Additon af støjkilder:

$$U_{TOT} = \sqrt{U_{n1}^2 + U_{n2}^2 + \dots + U_{nN}^2}$$

Figur 157 – Støjen dækker et frekvensområde, defineret ved den lave grænsefrekvens f_L , der er nul eller højere, og den høje grænsefrekvens f_H , der opad begrænset af operationsforstærkeren og det ydre kredsløb. Båndbredden defineres ved de frekvenser hvor niveauet er reduceret med 3 dB.

Her er $k = 1,381 \cdot 10^{-23}$ J/K Boltzmanns konstant, $T = 273 \text{ K} + \theta$ er den absolutte temperatur på 300 K ved en stuetemperatur omkring $\theta = 23^\circ\text{C}$, og R er modstandsværdien. Ved $R = 1 \text{ k}\Omega$ bliver $e_{nR} = 4 \text{ nV}/\sqrt{\text{Hz}}$.

Ved lydoptagelse er dynamiske mikrofoner populære og de har en ofte indre modstand omkring 100 Ω , så støjens RMS værdi fra mikrofonen bliver 0,2 μV indenfor en båndbredde på 20 kHz. Ved et lydsignal fra mikrofonen⁸¹ på 1 mV bliver signal-støj forholdet på 75 dB.

Støjen kan reduceres ved at sænke temperaturen, og det anvendes fx ved radioastronomi, men der kræves en reduktion fra stuetemperatur ned til -123°C for blot 3 dB støjreduktion (fra 300 K til 150 K), så det er en ret kostbar metode. Støjen er nul ved det absolute nulpunkt, men en halvleders funktion ophører før denne temperatur nås, så det er ikke muligt at lave en støjfri forstærker med halvledere.

Det er almindeligt at opgive støjspecifikationen for en komponent indenfor et frekvensområde, der er relateret til den aktuelle anvendelse. Ved måling af støjen benyttes et RMS voltmeter med et indbygget båndpasfilter, og den målte værdi rapporteres som en støjtæthed (*noise density*) ved at dividere støjens middeleffekt med båndbredden. Det giver enheden watt per hertz (W/Hz). Den rapporterede værdi er en spændingsstøjtæthed e_n (*voltage noise density*), der får den usædvanlige enhed volt per kvadratrod hertz (V/ $\sqrt{\text{Hz}}$). På en tilsvarende måde defineres en strømstøjtæthed i_n med enheden ampere per kvadratrod hertz på basis af en måling af RMS strømstøjen I_n .

Det er uden betydning hvor båndbreddens grænser er placeret, kun differensen mellem frekvensbåndets grænser er af betydning. Det er således uden betydning om området går fra 0 til 1 Hz eller fra 1000 til 1001 Hz. Formlen angiver ingen øvre grænse af båndbredden, men af kvantemekaniske årsager er den øvre begrænsning i THz området.

Hvis to eller flere støjsignaler skal føres til det samme sted, så er det middeleffekterne, der skal adderes. Effekten kan udtrykkes ved kvadratet på RMS værdien af spændingen for det enkelte signal (og tilsvarende for strømmen), og det leder til den viste formel for addition af støjbidrag. Den har stor betydning idet den kan simplificere beregningen ganske væsentligt.

⁸¹ Mikrofoner har almindeligvis en følsomhed på 1 ... 10 mV/Pa hvor lydtrykket 1 Pa svarer til 94 dB SPL, så de 1 mV er et ret typisk niveau fra en mikrofon i nærheden af en sanger eller et musikinstrument. Støjen kan let give problemer ved kvalitetsoptagelse af musik, for allerede ved 16 bit er det teoretiske signal-støj forhold på 96 dB.

Støjtætheden udtrykkes ved middeleffekten P_n målt indenfor en båndbredde B , og relationen til spænding og strøm beregnes som før ved at erstatte middeleffekten med de tilsvarende udtryk for spænding og strøm.

$$\text{Spændingsstøjtæthed: } \frac{P_n}{B} = \frac{U_n^2}{RB} = \frac{e_n^2}{R} \Rightarrow e_n = \frac{U_n}{\sqrt{B}}$$

$$\text{Strømstøjtæthed: } \frac{P_n}{B} = \frac{I_n^2 R}{B} = i_n^2 R \Rightarrow i_n = \frac{I_n}{\sqrt{B}}$$

For at beregne RMS værdien af en støjspænding (eller strøm) når den tilsvarende støjtæthed er kendt løses ligningerne for den søgte variabel.

$$\text{RMS spænding: } U_n = e_n \sqrt{B}$$

$$\text{RMS strøm: } I_n = i_n \sqrt{B}$$

For at beregne den resulterende spændingsstøj ved addition af støjen fra flere kilder benyttes middeleffekten, der kan skrives som U^2/R . Derefter gange med modstanden på begge sider af lighedstegnene og den resulterende spænding findes ved at tage kvadratroden.

$$P_{TOT} = P_1 + P_2 + \dots \Rightarrow \frac{U_{TOT}^2}{R} = \frac{U_1^2}{R} + \frac{U_2^2}{R} + \dots \Rightarrow U_{TOT}^2 = U_1^2 + U_2^2 + \dots \Rightarrow U_{TOT} = \sqrt{U_1^2 + U_2^2 + \dots}$$

For to støjsignaler U_1 og U_2 hvor $U_2 = 0,5 U_1$ beregnes summen ved indsættelse, og fejlen ved at ignorere den svage støjkilde er 12 %, som er lidt mindre end 1 dB.

$$U_{TOT} = \sqrt{U_1^2 + (0,5 U_1)^2} = \sqrt{(1+0,25) U_1^2} = 1,12 U_1$$

Ved addition af mange lige kraftige støjkilder hæves det samlede niveau med kvadratroden af antallet; det er et problem ved analoge mixerpulse med mange ens indgangskanaler.

$$U_{TOT} = \sqrt{U_n^2 + U_n^2 + \dots} = \sqrt{N} U_n$$

Addition af et større antal støjkilder af samme styrke mødes i mikserpulte hvor mange signaler skal samles, og det gælder både for analoge og digitale signaler, samt for lydsignaler og video. Det resulterende støjguld bliver hævet med kvadratroden af antallet af kanaler hvis de alle har samme støjniveau. For en mikserpult med 64 kanaler vil det resulterende støjniveau hæves 8 gange, og det svarer til en reduktion på 18 dB i det mulige signal-støj forhold.

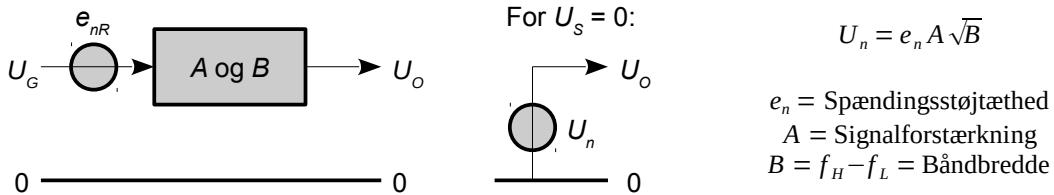
En tilsvarende relation mødes indenfor kommunikation med en fjern satellit. Her er signalet så svagt at det drukner i støjen. For at få signalet ned fra satellitten sendes det samme signal om og om igen uden nogen form for ændring. Ved addition af mange ens transmissioner vil signalets amplitude øges med antallet af transmissioner, mens støjgulvet kun adderes med kvadratroden af antallet fordi støjsignalet er nyt for hver transmission. Signal-støj forholdet bliver dermed forøget med kvadratroden af antallet af transmissioner. Samme metode benyttes ved måling af hjernebølger som følge af en visuel stimulering, eller måling af emission af lyd fra det indre øre som reaktion på et eksternt lydsignal, og her kommer antallet let over 1000.

Et af de store bidrag indenfor matematik har vist at et meget stort antal af uafhængige hændelser tilsammen giver en standardfordeling. Det er en af ingenørernes argumenter for at bruge netop dette værktøj til analyse af forhold hvor der er mange bidragydere, der ikke er indbyrdes koblet.

Operationsforstærkerens støj er relativt kompliceret at beregne, men det er almindeligvis fuldt tilstrækkeligt at betragte dens støj for hvid og benytte de viste formler for at kunne gennemføre en realistisk analyse. I det følgende beskrives en støjmodel for ethvert elektronisk kredsløb, hvor den termiske støjs niveau beregnes for modstande og operationsforstærkere. Det tillader beregning af om signal/støj-forholdet er tilstrækkeligt for at løfte en given opgave. Det er fuldt ud muligt at forfine beregningen, men det har for det meste kun akademisk interesse.

3.11.3 Beregningseksempler

En model af et elektronisk kredsløb vil meget ofte have støjkilden placeret på indgangen som en spændingsstøjtæthed e_n i serie med det tilsigtede signal, og kredsløbet beskrives ved parametrene A for forstærkningen i pasbåndet og B for bredden af det frekvensområde, der kan passere.



Figur 158 – En støjkilde e_{nR} på indgangen af et elektronisk kredsløb bliver forøget med kredsløbets forstærkning i pasbåndet A og frekvensbegrænses til båndbredden B_n af kredsløbet. På udgangen kan støjen opfattes som kommende fra en spændingsgenerator med en given RMS værdi.

Støjen forstærkes lige så meget som det signal kredsløbet bygges for at forstærke (U_s) og det eneste, der efterfølgende kan gøres for at sænke støjens niveau er at reducere båndbredden. Der er en kvadratrod inde i billedet så båndbredden skal reduceres 100 gange for at nå en reduktion i støjens værdi på 10 gange. Det er ikke i almindelighed en attraktiv metode at reducere båndbredden, for forstærkeren skal normalt tillade et bestemt frekvensområde at passere og derved er båndbredden fikseret af applikationen.

Ved forstærkning af svage signaler kan man forsøge at hæve signalets amplitude så signal-støj forholdet bliver størst muligt. Ved en lydoptagelse kan man fx bringe mikrofonen tæt på signalkilden, ved en radiomodtager kan man benytte en antennen med flere direktorer, og ved optisk signaltransmission må man overveje muligheden for at øge senderens effekt så signalets amplitude bliver tilstrækkelig stor.

En operationsforstærkers støj kan ikke umiddelbart beregnes, så i stedet benyttes databladets specifikation af støjen. Herunder vises nogle eksempler for tre modstandsværdier (e_{nR}) og fire typiske operationsforstærkere (e_{nA}).

Komponent	Værdi	Spændingsstøjtæthed
Modstand	10Ω	$e_{nR} = 0,4 \text{ nV}/\sqrt{\text{Hz}}$
Modstand	$1 \text{ k}\Omega$	$e_{nR} = 4 \text{ nV}/\sqrt{\text{Hz}}$
Modstand	$100 \text{ k}\Omega$	$e_{nR} = 40 \text{ nV}/\sqrt{\text{Hz}}$
OpAmp BJT støjsvag	AD797	$e_{nA} = 0,9 \text{ nV}/\sqrt{\text{Hz}}$
OpAmp BJT almindelig	NE5532	$e_{nA} = 5 \text{ nV}/\sqrt{\text{Hz}}$
OpAmp JFET almindelig	LF356	$e_{nA} = 12 \text{ nV}/\sqrt{\text{Hz}}$
OpAmp Micropower	OP291	$e_{nA} = 65 \text{ nV}/\sqrt{\text{Hz}}$

Støjberegning regnes for en svær disciplin, men det glemmes ofte at beregningen ikke skal være nøjagtig på tredje decimal for at kunne bruges. Det er almindeligvis fuldt tilstrækkeligt at vide om et signal-støj forhold er "stort nok" til den givne applikation⁸² og det er da rigeligt med en præcision omkring de 5 dB. Hvis der regnes worst-case, og signal-støj forholdet alligevel er stort nok, så er der jo ingen grund til at forbedre beregningen.

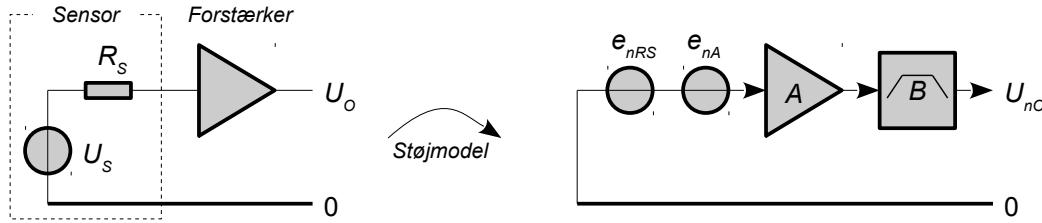
De følgende afsnit viser hvordan støj kan beregnes for nogle kredsløb.

⁸² Ved lydoptagelse kræves ofte et signal/støj-forhold af størrelsesordenen 80 dB, mens måling af et signals frekvens kan nøjes med et signal/støj-forhold omkring 0 dB.

3.11.4 Temperaturmåling

En meget hyppig situation er forstærkningen af et signal fra en sensor, for det dækker en række af opgaver, som temperaturmåling med PT100 eller et termoelement, mekanisk påvirkning med en strain gauge, radiomodtagelse fra en antenné og lydoptagelse med en mikrofon.

Sensoren afgiver en spænding U_s , som skal forstærkes op med forstærkningen A til udgangen U_o og forstærkeren begrænser samtidigt båndbredden til B . Disse parametre udgør den tilsigtede funktion ved opstillingen, men både sensor og forstærker genererer støj, som tilsammen danner et støjguld, og det er her formålet at estimere hvor dette uønskede niveau af egenstøj befinner sig.



Figur 159 – Signalet fra en sensor skal forstærkes op i en forstærker. Sensorens indre modstand R_s genererer spændingsstøjtætheden e_{nR} og forstærkeren genererer spændingsstøjtætheden e_{nA} . Der er en indstillet forstærkning på A og forstærkeren begrænser desuden båndbredden til B .

Første trin er at omforme diagrammet til en model, der fokuserer på støjkilderne og de parametre, der har betydning for beregning af den resulterende RMS værdi af støjen. Sensorens støjkilde vil her blive beskrevet ved dens indre modstand R_s , der genererer termisk støj, og forstærkeren vil blive beskrevet ved dens støjtæthed e_{nA} , der skal hentes fra databladet. De to støjkilder føres gennem forstærkeren til udgangen, så de forstærkes begge med faktor A og de begrænses begge af båndbredden B . Bruges superposition er proceduren den samme for begge støjkilder, og de skal RMS adderes ved udgangen.

Herunder vises hvordan algoritmen benyttes ved måling af temperatur med en termokabler. Det er ønsket at udgangsspændingen stiger med 10 mV per grad celsius, så udgangens støj skal være en del mindre for at det giver mening. Beregningen viser en RMS værdi af støjen på 0,4 mV så kravet er opfyldt med et signal-støj forhold på 30 dB. Det ses også at den valgte operationsforstærker dominerer i støjberegningen, så en forbedring kan være at vælge en mindre støjende model, eller at reducere båndbredden, for den er rigeligt høj for en temperaturmåling.

Et termoelement af K-typen giver en spænding på $41 \mu\text{V}/^\circ\text{C}$ og det ønskes omsat til $10 \text{ mV}/^\circ\text{C}$ så der skal forstærkes med $A = 244$ gange. Termoelementets indre modstand er $R_s = 10 \Omega$ som en antagelse. Operationsforstærkeren LF356 oplyses til en egenstøj på $e_{nA} = 12 \text{ nV}/\sqrt{\text{Hz}}$ og $GBP = 5 \text{ MHz}$. Ved den indstillede forstærkning bliver båndbredden $B = 20,5 \text{ kHz}$.

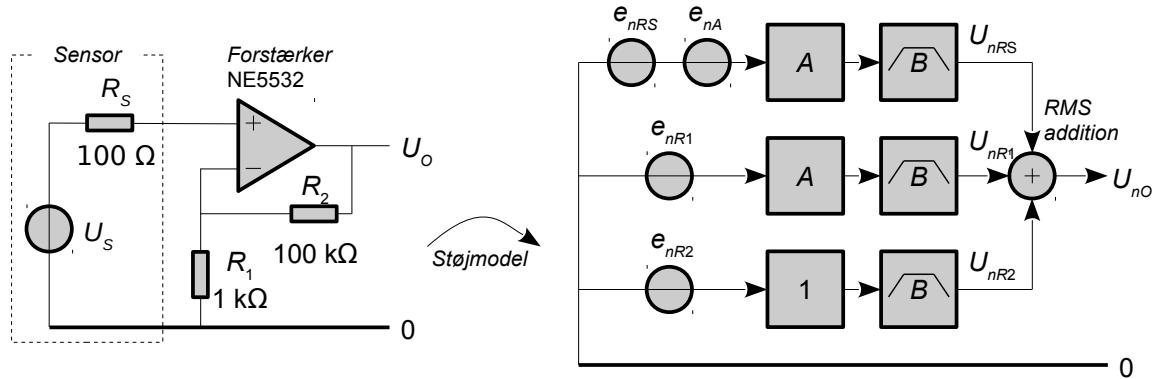
$$\begin{aligned} e_{nRs} &= \sqrt{4 k T R_s} = 0,4 \text{ nV}/\sqrt{\text{Hz}} \\ \text{Modstand: } U_{nRs} &= e_{nRs} A \sqrt{B} & e_{nA} &= \text{Datablad} = 12 \text{ nV}/\sqrt{\text{Hz}} \\ \text{Forstærker: } U_{nA} &= e_{nA} A \sqrt{B} & \text{hvor } A &= \frac{U_o}{U_i} = \frac{10 \text{ mV}/^\circ\text{C}}{41 \mu\text{V}/^\circ\text{C}} = 244 \\ \text{Total støj: } U_{no} &= \sqrt{U_{nRs}^2 + U_{nA}^2} & B &= \frac{GBP}{A} = \frac{5 \text{ MHz}}{244} = 20,5 \text{ kHz} \end{aligned}$$

Data indsættes og giver en RMS værdi af støjen på 0,4 mV, der viser at operationsforstærkeren dominerer i støjberegningen.

$$\begin{aligned} U_{nRs} &= (0,4 \text{ nV}/\sqrt{\text{Hz}}) \cdot 244 \cdot \sqrt{20,5 \text{ kHz}} = 0,014 \text{ mV} \\ U_{nA} &= (12 \text{ nV}/\sqrt{\text{Hz}}) \cdot 244 \cdot \sqrt{20,5 \text{ kHz}} = 0,419 \text{ mV} \\ U_{no} &= \sqrt{(0,014 \text{ mV})^2 + (0,419 \text{ mV})^2} = 0,419 \text{ mV} \end{aligned}$$

3.11.5 Mikrofonforstærker

Ved støjkritiske applikationer kan man ikke ignorere de modstande, der indstiller forstærkningen, så det følgende eksempel viser hvordan støjanalysen kan udvides til at tage dem med. Støjkilderne har ikke den samme forstærkning frem til udgangen, så det er nødvendigt at bestemme den aktuelle værdi af forstærkningen for hver af støjkilderne. Derefter føres signalerne frem til udgangen hvor de RMS adderes. Første eksempel viser at LF356 er ret støjende så her vælges NE5532.



Figur 160 – Mikrofonen, operationsforstærkeren og modstandene er alle støjkilder. Støjkilderne forstærkes, båndbreddebegrenses og overføres til udgangen hvor de RMS adderes.

Forstærkningen af signal og støj er $A = 101$ med de viste modstande⁸³, og støjen fra R_2 forstærkes ikke op.

Mikrofonens indre modstand er $R_S = 100 \Omega$, dens signal er $U_S = 10 \text{ mV}$, der skal løftes op til $U_o = 1 \text{ V}$ hvortil kræves $A = 100$. NE5532 har $e_{nA} = 5 \text{ nV}/\sqrt{\text{Hz}}$ og $GBP = 5 \text{ MHz}$.

Modstand R_S	$U_{nRS} = e_{nRS} A \sqrt{B}$	$e_{nRS} = \sqrt{4 k T R_S} = 1,3 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_1	$U_{nR1} = e_{nR1} A \sqrt{B}$	$e_{nR1} = \sqrt{4 k T R_1} = 4,1 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_2	$U_{nR2} = e_{nR2} \sqrt{B}$	$e_{nR2} = \sqrt{4 k T R_2} = 40,7 \text{ nV}/\sqrt{\text{Hz}}$
Forstærker	$U_{nA} = e_{nA} A \sqrt{B}$	$e_{nA} = \text{Datablad} = 5 \text{ nV}/\sqrt{\text{Hz}}$
Total støj	$U_{no} = \sqrt{U_{nRS}^2 + U_{nA}^2 + U_{nR1}^2 + U_{nR2}^2}$	$A = 1 + \frac{R_2}{R_1} = 1 + \frac{10 \text{ k}\Omega}{100 \Omega} = 101$
Signal/støj	$SNR = 20 \cdot \log_{10} \left(\frac{U_o}{U_{no}} \right) \text{ dB}$	$B = \frac{GBP}{A} = \frac{10 \text{ MHz}}{101} = 99 \text{ kHz}$

Den samlede RMS værdi af støjen bliver $0,147 \text{ mV}$ og signal/støj-forholdet bliver $76,6 \text{ dB}$, der er i underkanten for en krævende lydoptagelse.

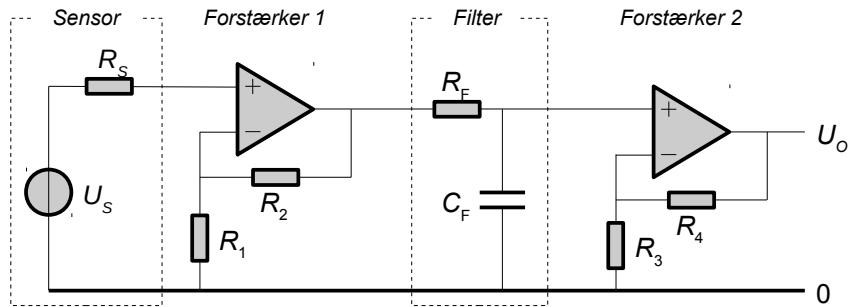
$$\begin{aligned}
 U_{nS} &= (1,3 \text{ nV}/\sqrt{\text{Hz}}) \cdot 101 \cdot \sqrt{99 \text{ kHz}} &= 0,041 \text{ mV} \\
 U_{nR1} &= (4,1 \text{ nV}/\sqrt{\text{Hz}}) \cdot 101 \cdot \sqrt{99 \text{ kHz}} &= 0,130 \text{ mV} \\
 U_{nR2} &= (40,7 \text{ nV}/\sqrt{\text{Hz}}) \cdot 1 \cdot \sqrt{99 \text{ kHz}} &= 0,013 \text{ mV} \\
 U_{nA} &= (5 \text{ nV}/\sqrt{\text{Hz}}) \cdot 101 \cdot \sqrt{99 \text{ kHz}} &= 0,159 \text{ mV} \\
 U_{no} &= \sqrt{(0,029)^2 + (0,029)^2 + (0,003)^2 + (0,112)^2} &= 0,210 \text{ mV} \\
 SNR &= 20 \cdot \log_{10} \left(\frac{1 \text{ V}}{0,147 \text{ mV}} \right) \text{ dB} &= 73,6 \text{ dB}
 \end{aligned}$$

Det anbefales generelt at modstanden R_1 ikke bør være større end signalkildens indre modstand i støjsvage forstærkere. Operationsforstærkeren bør heller ikke støje mere end R_1 , hvilket er tæt på at være opfyldt i dette design.

⁸³ Det er uden betydning for støjanalysen at støjen fra R_1 kun forstærkes 100 gange, idet forstærkeren her er inverterende. Undtagelsen er ved en indstillet forstærkning på nogle få gange, men det anvendes så godt som aldrig ved støjsvage kredsløb.

3.11.6 Forstærker med filter

En forstærker med et filter for reduktion af båndbredden benyttes meget ofte ved interface til sensorer, enten for at reducere støj ved en lav båndbredde eller for at modificere den aktuelle sensors overføringsfunktion⁸⁴. Signalet forstærkes op med forstærkningen indstillet ved R_1 og R_2 , båndbegrænses af filtret ved R_F og C_F for at blive forstærket med forstærkningen indstillet ved R_3 og R_4 . Det forudsættes at båndbredden af begge forstærkere er større end filtrets grænsefrekvens da det ellers ikke vil have nogen væsentlig funktion.



Figur 161 – Forstærker med et passivt filter for en forsætlig båndbegrænsning.

Støjen fra sensor, første forstærker og modstandene R_1 og R_2 bliver forstærket op med A_1 og A_2 og båndbredden begrænses til B_F . Støjen fra anden forstærker og modstandene R_3 og R_4 bliver forstærket med A_2 og båndbredden begrænses til B_2 . Udgangens støj er ikke hvid, for godt halvdelen af støjeffekten har en båndbredde på 10 Hz mens resten af støjen dækker op til 50 kHz. Støjberegningen viser kun hvor kraftig støjen er som en samlet RMS værdi af spændingen.

Beregning skal illustreres med en opstilling hvor et accelerometers signal skal forstærkes og om sættes fra acceleration til hastighed gennem integration af signalet. Den samlede forstærkning sættes til 100 ved frekvensen 1 kHz, og det vælges at filtret ved R_F og C_F dæmper signalet lige så meget som den første forstærker hæver signalet, derfor vælges filtret til 10 Hz og det kræver en forstærkning i første trin på 100. Det antages at udgangssignalet amplitude er 1 V.

Forstærkningen i hvert trin vælges til $A_1 = A_2 = 101$, der kan realiseres med $R_1 = R_3 = 1 \text{ k}\Omega$ samt $R_2 = R_4 = 100 \text{ k}\Omega$. Grænsefrekvensen er $B_F = 10 \text{ Hz}$ ved filtret, der realiseres med $R_F = 160 \text{ k}\Omega$ og $C_F = 100 \text{ nF}$. Som OpAmp vælges en LF356 type med data $e_{nA} = 12 \text{ nV}/\sqrt{\text{Hz}}$ og $GBP = 5 \text{ MHz}$. Accelerometeret antages at være en modstand på $10 \text{ k}\Omega$.

$$\text{Modstand } R_S \quad U_{nRS} = e_{nRS} A_1 A_2 \sqrt{B_F} = 0,42 \text{ mV}$$

$$\text{Forstærker 1} \quad U_{nA1} = e_{nA1} A_1 A_2 \sqrt{B_F} = 0,39 \text{ mV}$$

$$\text{Modstand } R_1 \quad U_{nR1} = e_{nR1} A_1 A_2 \sqrt{B_F} = 0,13 \text{ mV}$$

$$\text{Modstand } R_2 \quad U_{nR2} = e_{nR2} A_2 \sqrt{B_F} = 0,01 \text{ mV}$$

$$\text{Modstand } R_F \quad U_{nRF} = e_{nRF} A_2 \sqrt{B_F} = 0,02 \text{ mV}$$

$$\text{Forstærker 2} \quad U_{nA2} = e_{nA2} A_2 \sqrt{B_2} = 0,27 \text{ mV}$$

$$\text{Modstand } R_3 \quad U_{nR3} = e_{nR3} A_2 \sqrt{B_2} = 0,09 \text{ mV}$$

$$\text{Modstand } R_4 \quad U_{nR4} = e_{nR4} \sqrt{B_2} = 0,01 \text{ mV}$$

$$\text{Total støj} \quad U_{nO} = \sqrt{\sum_{\text{Alle}} U_n^2} = 0,65 \text{ mV}$$

$$\text{Signal/støj} \quad SNR = 20 \cdot \log_{10} \left(\frac{U_O}{U_{nO}} \right) = 64 \text{ dB}$$

$$e_{nRS} = \sqrt{4kT R_S} = 13 \text{ nV}/\sqrt{\text{Hz}}$$

$$e_{nRF} = \sqrt{4kT R_F} = 51 \text{ nV}/\sqrt{\text{Hz}}$$

$$e_{nR1} = e_{nR3} = \sqrt{4kT R_1} = 4 \text{ nV}/\sqrt{\text{Hz}}$$

$$e_{nR2} = e_{nR4} = \sqrt{4kT R_2} = 41 \text{ nV}/\sqrt{\text{Hz}}$$

$$e_{nA1} = e_{nA2} = \text{Datablad} = 12 \text{ nV}/\sqrt{\text{Hz}}$$

$$\text{hvor} \quad A_1 = 1 + \frac{R_2}{R_1} = 101$$

$$A_2 = 1 + \frac{R_4}{R_3} = 101$$

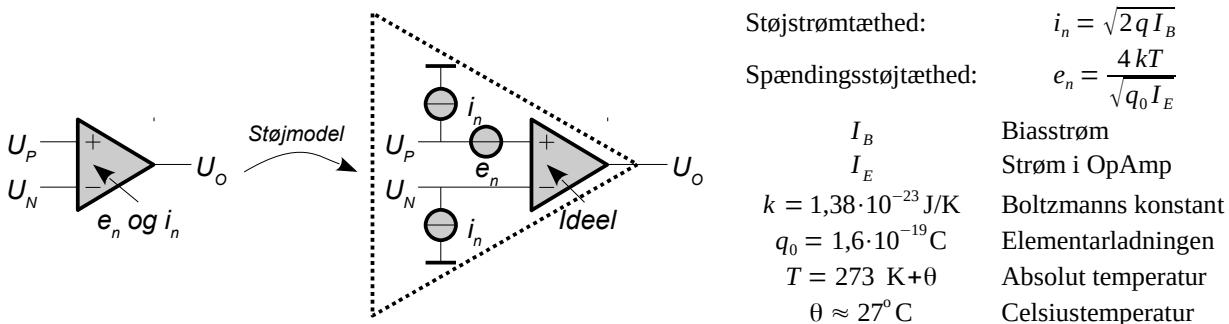
$$B_1 = B_2 = \frac{GBP}{A} = 49,5 \text{ kHz} > B_F$$

$$B_F = \frac{1}{2\pi R_F C_F} = 10 \text{ Hz}$$

⁸⁴ Ved en meget lav grænsefrekvens er kredsløbet integrerende og kan fx omsætte acceleration til hastighed.

3.11.7 Operationsforstærkerens støjmodel

Operationsforstærkeren er kun tilnærmelsesvist beskrevet ved spændingsstøjtætheden. For det første er der en strømstøj fra biasstrømmen i indgangene, og for det andet er støjen fra en operationsforstærker ikke hvid, men vokser ved lave frekvenser, og de to forhold beskrives i de følgende afsnit. Herunder vises en generel model, hvor støjen fra en operationsforstærker beskrives ved en spændingsstøjtæthed e_n og to strømstøjtætheder i_n . Som ved de øvrige modeller repræsenteres en OpAmp ved en ideel operationsforstærker med eksterne fejlkilder.



Figur 162 – Støjen i en forstærker repræsenteres af en spændingsgenerator i den ene indgang og en strømstøjgenerator for hver indgang.

Operationsforstærkerens biasstrøm I_B er en forudsætning for at transistorerne i indgangen virker. Strømmen opbygges af elementarladninger q_0 som ladning per tid. Da kun et heltalligt antal elementarladninger kommer igennem i et givet tidsrum vil der være en usikkerhed på strømstyrken, og denne usikkerhed optræder som en støj, der kaldes for haglstøj (shot noise) og som overlejer biasstrømmen⁸⁵. Støjstrømtætheden i_n kan opfattes som en AC støjstrøm, der løber igennem de eksterne komponenter i kredsløbet hvor den giver anledning til en AC spænding, som derefter kan forstærkes op på linje med spændingsstøjtætheden. Værdien kan beregnes fra biasstrømmen.

En OpAmp med bipolære transistorer har typisk $I_B = 100 \text{ nA}$ så støjstrømtætheden bliver $i_n = 200 \text{ fA}/\text{Hz}$.
En OpAmp med MOSFET transistorer har typisk $I_B = 10 \text{ pA}$ så støjstrømtætheden bliver $i_n = 2 \text{ fA}/\text{Hz}$.

Spændingsstøjtætheden for en operationsforstærker er givet ved den strøm, der løber i kollektor (eller drain) i de interne transistorer, og det vises i femte kapitel at denne strøm kan omsættes til en ækvivalent spændingsstøjkilde ved indgangen når de interne forhold er kendt.

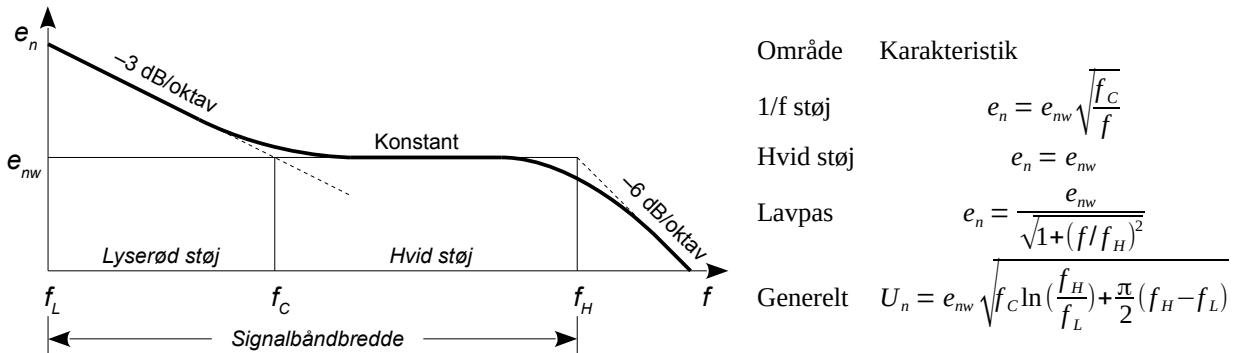
En OpAmp med strømmen $I_E = 100 \mu\text{A}$ i indgangen giver en spændingsstøjtæthed på $e_n = 4 \text{ nA}/\text{Hz}$.

Analysearbejdet er meget i stil med DC modellen for en operationsforstærker, og støjmodellen har meget til fælles med modellen for beregning DC fejl. Det er derfor muligt at benytte de samme kredsløbsligninger blot støjbidragene RMS adderes. Det er ikke muligt at kompensere for støjkilernes bidrag, som det var muligt for biasstrømmen, og der er ingen offsetstrøm i støjmodellen.

Strømstøjen er almindeligvis ikke noget større problem ved operationsforstærkere med JFET eller MOSFET i indgangen, da biasstrømmen er meget lav, og det er årsagen til at megen litteratur anbefaler denne type OpAmp for signalkilder med høj indre modstand. En OpAmp med bipolære transistorer i indgangen har ofte en betydelig biasstrøm, så strømstøjtætheden er også stor, men det er til gengæld muligt at fremstille bipolære operationsforstærkere med meget lave værdier af spændingsstøjtætheden, og den kan derfor være det rette valg ved lav kildemodstand. Grænsen mellem høj og lav kildemodstand kan som en tommelfingerregel sættes til $100 \text{ k}\Omega$, men det aktuelle niveau af støjen bør tjekkes før designet fastlægges.

⁸⁵ Begrebet haglstøj blev introduceret af Walter Shottky i 1918 ved studier af elektronstrømmen i et radiorør. Den statistiske beskrivelse benytter en Poisson fordeling, der også kan beskrive forløbet ved kast af en mønt. Ved det store antal elektroner, der er involveret, går Poisson fordelingen imod standardfordelingen. Haglstøjen er ikke funktion af temperaturen.

Støjeffekten er ved lave frekvenser omvendt proportional med frekvensen⁸⁶ så enheden bliver watt per hertz (W/Hz). Støjens kaldes for $1/f$ støj på grund af afhængigheden af frekvensen, og hvis støj-en udtrykkes som spændingsstøjtæthed aftager niveauet med kvadratroden af frekvensen, så hældningen er $-3 \text{ dB}/\text{oktav}$ eller $-10 \text{ dB}/\text{dekade}$, og har den populære betegnelse lyserød støj eller rosa støj (*pink noise*). Ved lave frekvenser er $1/f$ støj dominerende over den hvide støj, og overgangsfrekvensen f_c (*corner frequency*) opgives i databladet eller vises i en kurve for de støjsvage operationsforstærkere.



Figur 163 – Støjtætheden fra en forstærker aftager med frekvensen ($1/f$ støj) ved lave frekvenser, den er konstant for højere frekvenser og aftager over en grænse sat af forstærkerens båndbrede.

Den lave grænsefrekvens f_L er enten defineret af et højpasled, fx en DC blokering, eller af den tid T man observerer systemet over ($f_L > 3/T$). Ved DC forstærkere er frekvensen i teorien nul, men ved observation over et tidsrum på fx $T = 1 \text{ s}$ svarer det til en grænsefrekvens på cirka 3 Hz.

Overgangsfrekvensen f_c er opgivet i databladet og ligger oftest i området fra 100 Hz til 10 kHz, med den lave værdi for støjsvage operationsforstærkere.

Den høje grænsefrekvens f_H sættes af forstærkeren eller af det eksterne kredsløb, fx 100 MHz ved et analogt oscilloskop; 7 MHz for videosignaler; 96 kHz for lydoptagelser; 22 kHz for digital lyd (CD); 15 kHz for en FM radios båndbredde og måling af acceleration; 3,5 kHz for AM radio og telefoni; 200 Hz for hjerte-målinger (EKG) og 1 Hz for måling af temperatur.

For at kunne beskrive RMS værdien af støjspændingen skal forløbet integreres fra nul til uendelig, men det giver et problem, for ved frekvensen nul er bidraget uendeligt. Løsningen på problemet er at måletiden må være endelig, for en frekvens på nul svarer til at observationstiden er uendelig, og resultatet skulle gerne være fundet inden man selv dør af alderdom⁸⁷. Som vist i første kapitel svarer måletiden T til en grænsefrekvens på $3/T$ ved 10 % usikkerhed, så i de fleste tilfælde vil grænsefrekvensen være mindst 3 Hz da man sjældent venter mere end et sekund på målingen. Undtagelserne kan eksempelvis være en måling af temperatur med en termokabler, eller en måling af meget svage strømme i en integreret kreds, men uanset applikationen vil den lave grænsefrekvens være forskellig fra nul for man er nød til at afslutte målingen.

Resultatet af integrationen er vist i illustrationen som RMS værdien U_n af en spændingsstøjtæthed, der aftager med kvadratroden af frekvensen mod niveauet e_{nw} for frekvenser over f_c hvor den hvide støj dominerer, og som mod højere frekvenser aftager svarende til operationsforstærkerens båndbredde. Udtrykket under kvadratroden kaldes for den ækvivalente støjbåndbredde.

For NE5532 er $e_{nw} = 5 \text{ nV}/\text{Hz}$ og $f_c = 30 \text{ Hz}$. RMS værdien af spændingsstøjen bliver $U_{nA} = 1,4 \mu\text{V}$ ved grænsefrekvenserne $f_L = 1 \text{ Hz}$ og $f_H = 49,5 \text{ kHz}$. Mikrofonforstærkeren på side 154 forstærker $A = 244$ så udgangens RMS støj bliver $U_{no} = 0,34 \text{ mV}$ hvor den tidligere beregning gav $0,15 \text{ mV}$ (ændring på 7 dB).

Herunder gennemføres integrationen over de tre frekvensbånd.

86 Mekanismen bag $1/f$ -støj er ukendt, men formodes at være relateret til omrøkering af ladningskoncentrationer, strømme i overfladen af halvlederen og pludselige ændringer i krystalstrukturen.

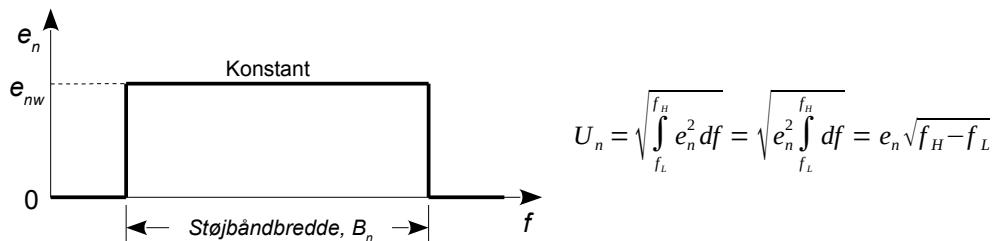
87 Ifølge engelske Wikipedia er der udført måling af $1/f$ -støj over flere uger, som bekræfter at forløbet er som beskrevet.

For at beregne en RMS værdi af støjspændingen U_n startes med at finde et udtryk for den samlede støjeffekt P_n fra kredsløbets udgang. Effekttætheden beregnes fra p_n , der beskrives ved en givet frekvens f . Støjeffekten i et smalt område df beregnes som $p_n df$ og den samlede effekt er givet ved at addere bidragene over hele det mulige frekvensområde. Det svarer til en integration over alle frekvenser; det er fra frekvensen nul til uendelig.

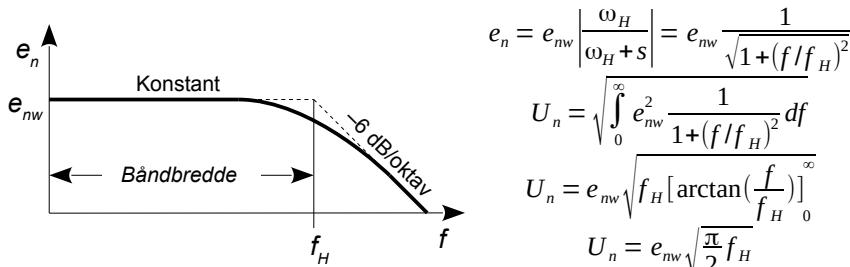
$$P_n = \int_0^\infty p_n df = \int_0^\infty \frac{e_n^2}{R} df = \frac{U_n^2}{R} \Rightarrow U_n^2 = \int_0^\infty e_n^2 df \Rightarrow U_n = \sqrt{\int_0^\infty e_n^2 df}$$

Effekttætheden p_n er givet ved kvadratet på spændingstætheden e_n delt med en modstand R , og resultatet af integrationen kaldes for U_n , som repræsenterer den RMS spænding, der giver samme effekt som støjen. Modstanden R er medtaget for at tydeliggøre relationen mellem effekt og spænding; den antages at være den samme under hele beregningen idet måleopstillingen er invariant, så den absolute værdi af modstanden har ingen betydning.

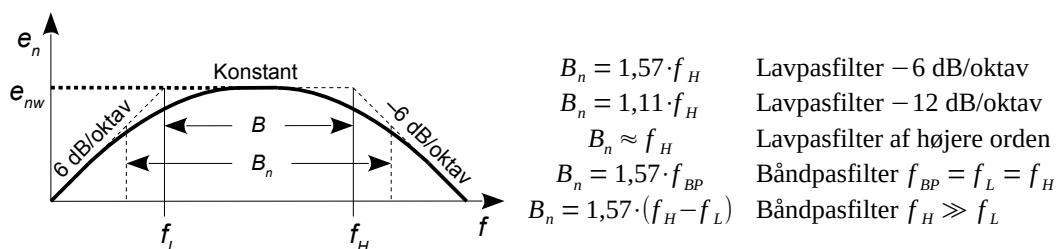
I midterområdet, hvor støjtætheden e_n er konstant, vil integrationen løbe fra f_L til f_H . Værdi af e_n er konstant og kan sættes udenfor, så en integration over frekvensområdet giver kvadratroden af båndbredden $f_H - f_L$. Det er den relation, der er vist tidligere i afsnittet (side 150).



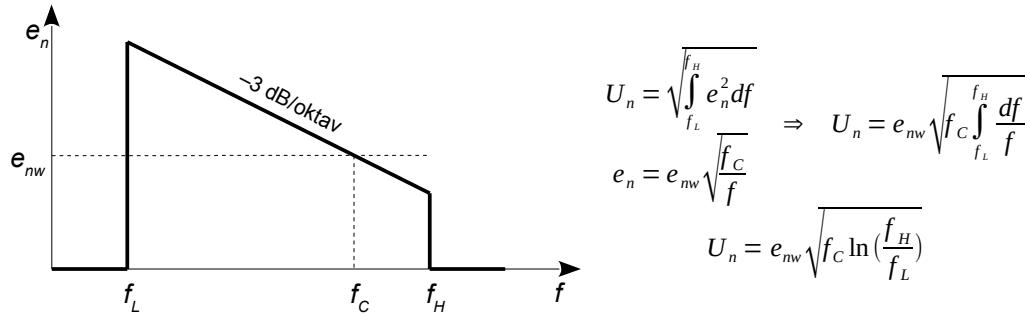
Båndbredden er ikke så skarp som vist i illustrationen for et såkaldt *brick-wall filter*, men som det skal vises er det en ganske god tilnærmelse. Støjspændingen bestemmes herunder for et første-ordens lavpasfilter, der kan repræsentere en operationsforstærkers båndbredde eller lavpasfiltret fra side 155. Støjtætheden e_{nw} på indgangen af filtret antages konstant og den effektive værdi af spændingsstøjen viser at der kommer et bidrag over -3 dB grænsefrekvensen.



Der kan gennemføres en lignende integration for et båndpasfilter og resultatet bliver $\pi/2$ gange differensen mellem filtrets -3 dB grænsefrekvenser. Herunder vises støjbåndbredden B_n for en række udvalgt filtre. Korrektionen er 2 dB ved første-ordens filtre, den er reduceret til 1 dB ved anden-ordens filtre, og mindre endnu for filtre af højere-orden. Det er vanskeligt at beregne støj med mindre end et par decibel usikkerhed, hvilket er argumentet for at ignorere korrektionen og blot regne båndbredden som givet ved -3 dB grænserne.



I det lave område, hvor $1/f$ -støjen dominerer, kan spændingsstøjtætheden e_n udtrykkes som kvadratroden af f_c/f for at gengive hældningen på $-3 \text{ dB}/\text{oktav}$ og den effektive værdi af spændingsstøjen beregnes ved at integrere spændingsstøjtætheden over området fra en lav grænse f_L til en høj grænse f_H . Resultatet er et udtryk baseret på en logaritme, hvis betydning bedst kan vises ved et eksempel. Med $f_L = 1 \text{ Hz}$, $f_c = 1 \text{ kHz}$ og $f_H = 10 \text{ kHz}$ samt $e_{nw} = 10 \text{ nV}/\sqrt{\text{Hz}}$ findes støjen til $U_n = 1 \mu\text{V}$. Den beregnede værdi stiger kun til $U_n = 1,3 \mu\text{V}$ ved den ekstremt lave grænsefrekvens på $f_L = 1 \text{ mHz}$, der kræver en måletid på knapt en time.



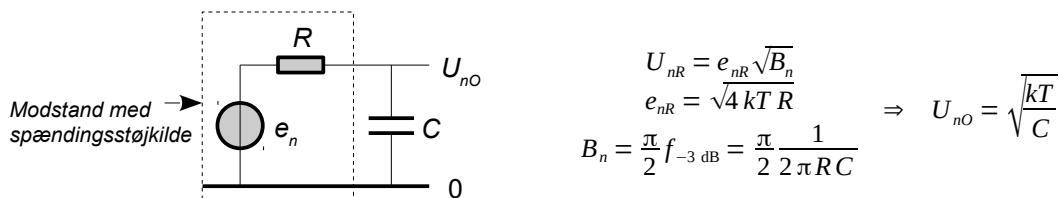
Konklusionen på beregningerne kan samles til et fælles udtryk for RMS værdien af operationsforstærkerens egenstøj, når den er karakteriseret ved spændingsstøjtætheden e_{nw} , der er en hvid støj og gælder for området hvor forstærkningen er konstant, og grænsen mellem $1/f$ -støj og den hvide støj er f_c . Operationsforstærkeren er indstillet til en spændingsforstærkning på A , og båndbredden er indstillet til f_L i den lave ende og f_H i den høje ende. Ved en DC forstærker kan man som regel sætte $f_L = 1 \text{ Hz}$ uden større betydning for beregningens præcision.

$$U_n = e_{nw} A \sqrt{f_c \ln\left(\frac{f_H}{f_L}\right) + \frac{\pi}{2}(f_H - f_L)}$$

Flertallet af anvendelser har et veldefineret frekvensområde, der giver mulighed for at stille et kvalificeret krav til den benyttede operationsforstærkers grænsefrekvens f_c for overgangen fra $1/f$ støj til hvid støj. Eksempelvis har mellemfrekvensen i en AM radiomodtager en båndbredde på 9 kHz centreret på 455 kHz, og ved en FM radiomodtager er båndbredden 300 kHz centreret på 10,7 MHz. Ved analog videotransmission over et kabel benyttes området fra 25 Hz til 7 MHz, en telefons omfang er fra 300 Hz til 3,5 kHz og måling af et hjertets signaler benytter fra 0,1 Hz til 150 Hz.

3.11.8 Støjen fra et RC led

En modstand genererer en spændingsstøjtæthed som er givet af modstandsværdien, og hvis der hægtes en kondensator i parallel over modstanden vil støjen båndbegrænses af det dannede filter mellem modstand og kondensator. Det kan vises at RMS værdien at støjens spænding alene er funktion af kondensatoren. Modstanden giver ganske vist mere støj fra sig når dens værdi stiger, men samtidigt falder den resulterende støjbåndbredde.

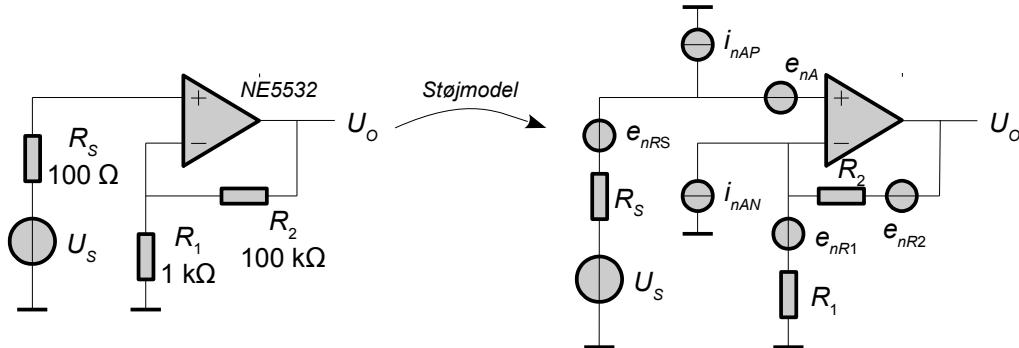


Figur 164 – Støjen fra en paralleltildelingen af en modstand og en kondensator afhænger kun af kondensatoren. Selv om RMS værdien af støjspændingen er konstant så er støjens båndbredde afhængig af den resulterende grænsefrekvens skabt af R og C .

Et RC led med en kondensator på 1 nF giver en RMS spænding på 2 μV indenfor den aktuelle støjbåndbredde, der med eksempelvis 100 $\text{k}\Omega$ og 1 nF er på 2,5 kHz.

3.11.9 Eksempel

Herunder illustreres brugen af modellen med et eksempel for en ikke-inverterende forstærker, der er identisk med mikrofonforstærkeren på side 154 og med forstærker 2 i kredsløbet på side 155. Der er to modstande i tilbagekoblingen, som indstiller forstærkning og båndbredde (R_1 og R_2), og en signalkilde med en indre modstand R_s . Beregningen af den resulterende støj på udgangen vises herunder med operationsforstærkeren NE5532, der er bygget med bipolare transistorer, og derfor har en relativt kraftig strømstøjtæthed.



Figur 165 – En ikke-inverterende forstærker kan med hensyn til støj opfattes som vist til højre, hvor alle støjkilder er tegnet ind og forstærkeren nu regnes for ideal.

Støjbiderne for NE5532 hentes fra databladet, modstandenes støjbidrag beregnes ved 300 K. Kredsløbets båndbredde beregnes ud fra GBP specifikationen og den indstillede forstærkning A. Den ækvivalente støjbåndbredde beregnes for modstandene til B_n , og for forstærkeren afviger båndbredden stort set ikke på grund af den lave værdi af $1/f$ støjen. Den lave grænsefrekvens sættes til 1 Hz. I den positive indgang løber strømstøjen gennem R_s og forstærkes med A. I den negative indgang løber strømstrømmen udelukkende gennem R_2 .

$$\text{OpAmp NE5532: } e_n = 5 \text{ nV}/\sqrt{\text{Hz}}, \quad f_c = 30 \text{ Hz}, \quad i_n = 0,7 \text{ pA}/\sqrt{\text{Hz}}, \quad GBP = 10 \text{ MHz}$$

$$\text{Forstærkning: } A = 1 + \frac{R_2}{R_1} = 101,$$

$$\text{Støjbåndbredde: } f_H = \frac{GBP}{A} = 100 \text{ kHz} \Rightarrow B_n = \frac{\pi}{2}(f_H - f_L) = 157 \text{ kHz}$$

$$\text{Modstandsstøj: } U_{nRS} = A \sqrt{4kTR_s B_n} = 0,051 \text{ mV}$$

$$U_{nR1} = A \sqrt{4kTR_1 B_n} = 0,163 \text{ mV}$$

$$U_{nR2} = \sqrt{4kTR_2 B_n} = 0,016 \text{ mV}$$

$$\text{OpAmp spændingsstøj } U_{neA} = Ae_{nA}\sqrt{B_n} = 0,200 \text{ mV}$$

$$\text{OpAmp strømstøj } U_{niRS} = AR_s i_n \sqrt{B_n} = 0,003 \text{ mV}$$

$$U_{niR2} = R_2 i_n \sqrt{B_n} = 0,028 \text{ mV}$$

$$\text{Udgangens støj } U_{no} = \sqrt{U_{nRS}^2 + U_{nR1}^2 + U_{nR2}^2 + U_{neA}^2 + U_{niRS}^2 + U_{niR2}^2} = 0,265 \text{ mV}$$

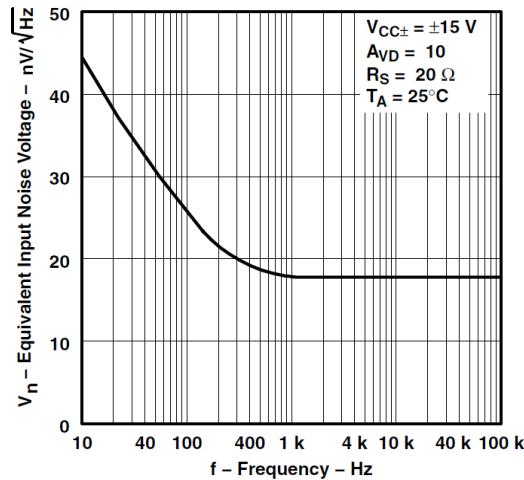
$$\text{Signal/støj-forhold } SNR = 20 \cdot \log_{10} \left(\frac{U_o}{U_{no}} \right) \text{ dB} = 71,5 \text{ dB}$$

Ved beregningen for kredsløbet på side 154 var RMS værdien af støjen på 0,210 mV, og den korrigerede værdi er cirka 2 dB større, alene på grund af den større båndbredde i analysen.

Den primære støjkilde er spændingsstøjtætheden fra operationsforstærkeren, med det næstmest betydende bidrag fra modstanden R_1 , så en forbedring ville sætte ind ved at finde en mere støjsvag forstærker og derefter reducere værdien af R_1 . Det ses også at strømstøjtætheden ikke giver et betydnende bidrag på grund af de lave modstandsværdier.

3.11.10 Databladet

Støjtæthederne oplyses traditionelt ved 1 kHz uanset om det er i området for hvid støj eller ej, så det er problemfyldt at bruge data direkte for en støjkritisk opstilling. For de støjsvage typer vises en kurve over spændingsstøjtætheden som funktion af frekvensen, og undertiden ligeledes en kurve for strømstøjtætheden, mens de mere typiske forstærkere blot har en talværdi ved 1 kHz. Det viste eksempel for TL071 repræsenterer de populære TL072 og TL074, og grænsefrekvensen mellem de to støjområder er cirka 300 Hz.

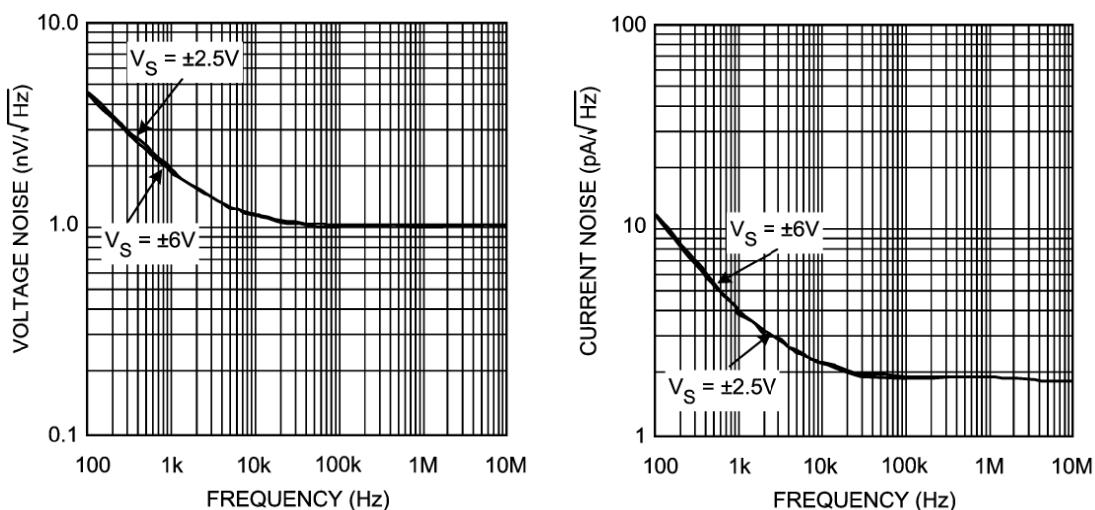


Forstærker	e_n $\text{nV}/\sqrt{\text{Hz}}$	i_n $\text{pA}/\sqrt{\text{Hz}}$	I_{CC} mA
LMH6624	1	2,3	11
AD797	0,9	2	8,2
NE5532	5	0,7	8
OP27	3,2	0,4	2,8
TL071	18	0,01	1,4
OP262	9,5	0,4	0,5
TLC271	25 ... 68	—	0,02 ... 0,2
OP293	65	0,05	0,015
ADA4051	95	0,1	0,01

Figur 166 – Støjtætheden for TL071-serien som funktion af frekvensen (venstre), og eksempler på specifikationerne for nogle udvalgte operationsforstærkere, hvor det ses at en støjsvag operationsforstærker har et relativt højt strømforbrug (højre).

LMH6624 er en bredbåndet forstærker (GBP på 1,5 GHz), som kræver en forstærkning på minimum 10 gange af hensyn til stabilitet. TLC271 er en CMOS forstærker med justerbar bias, der påvirker egenstøj, båndbredde og effektforbrug. ADA4051 er en såkaldt chopper stabiliseret forstærker, hvor en intern clock sampler DC fejl og derefter subtraherer disse fejl til en specifikation på typisk $\pm 2 \mu\text{V}$.

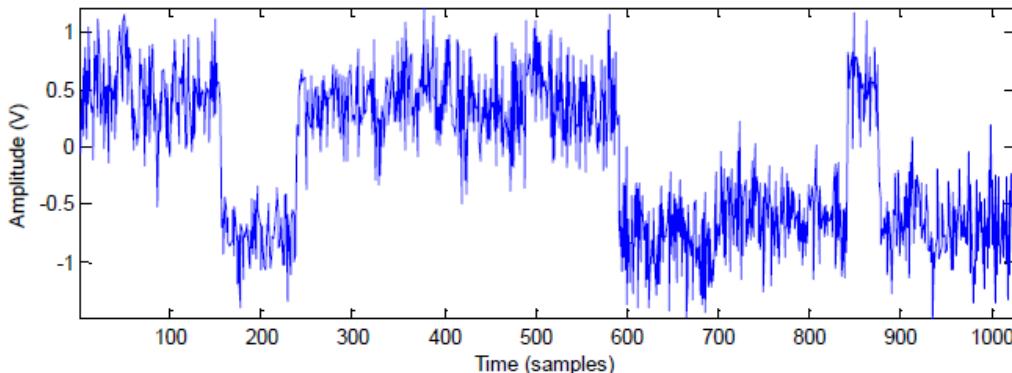
Støjsvage forstærkere kræver en høj værdi af strømmen i indgangens differentialtrin (se senere), så de kan ikke laves i *micro power* typer med et strømforbrug i mikroampere klassen. Det er vist i tabellen ved det typiske strømforbrug for et antal operationsforstærkere.



Figur 167 – Støjtætheden (spænding og strøm) for LMH6624 og LMH6626, der er enkelt og dobbelt forstærker i samme hus.

3.11.11 Popcorn støj

En noget speciel form for støj kaldes Popcorn støj (*popcorn noise*), og hvor der gives flere forklaringer: fx opbygning og løsladelse af ladninger i defekter i halvledermaterialet eller overfladen af krystallet, samt diffusion af forureningspartikler og pludselig ændring af krystalfejl (dislokation). Støjen optræder som en momentan ændring af spændingen på indgangen af en forstærker. Tidsskalaen for de pludelige ændringer kan være fra millisekunder til sekunder, og amplituden kan i de mest ekstreme tilfælde være så høj som 100 μ V.



Figur 168 – Eksempel på popcorn støj i en OpAmp med en ukendt indstilling af forstærkningen⁸⁸.

Popcorn støjen er problematisk ved forstærkning af meget små signaler, og ved lydbehandling hvor den høres som majskorn, der popper i en mikrobølgeovn. Den er specificeret for de bedre forstærkere i form af en garanteret spids-spids værdi af støjen fra 0,1 til 10 Hz. For den støjsvage operationsforstærker AD797 er værdien garanteret indenfor 50 nV spids-spids for det angivne frekvensområde. Parameteren verificeres som en del af testforløbet i produktionen, og da den nedre grænsefrekvens fordrer en testtid på flere sekunder er det en overordentlig kostbar test i et automatiseret produktionsforløb.

3.12 Sammenfatning

Modstande genererer støj på grund af elektronernes termisk forårsagede vibration, og det giver en evigt varierende spænding mellem modstandens ledninger. Strømmen bæres af et endeligt antal elektroner per tidsenhed og det giver en støj i_n , der er givet ved strømstyrken I_{DC} . Strømstøjen aftager med kvadratrodten af strømmens DC værdi, så den bliver relativt generende ved lav strømstyrke. Den er betydende ved operationsforstærkers biasstrøm, når denne strøm løber gennem eksterne modstande i kredsløbet og giver anledning til et spændingsfald.

Spændingsstøjtæthed (modstand):	$e_{nR} = \sqrt{4kT R}$	Boltzmanns konstant: $k = 1,38 \cdot 10^{-23}$ J/K
Spændingsstøjtæthed (OpAmp):	$e_{nA} \approx \frac{4kT}{\sqrt{q_0 I_E}}$	Elementarladningen: $q_0 = 1,60 \cdot 10^{-19}$ C
Strømstøjtæthed (OpAmp):	$i_{nA} = \sqrt{2q_0 I_B}$	Absolut temperatur: $T \approx 300$ K
RMS støjspænding (OpAmp):	$U_n = e_n A \sqrt{B_n}$	Forstærkning: A
Støjbåndbredde (tilnærmet):	$B_n = f_H - f_L$	Biasstrøm (OpAmp): I_B
Støjbåndbredde (OpAmp):	$B_n = f_C \ln\left(\frac{f_H}{f_L}\right) + \frac{\pi}{2}(f_H - f_L)$	Høj grænsefrekvens: f_H
		Lav grænsefrekvens: f_L
		Grænse for 1/f støj: f_C

Supplerende information om støj i operationsforstærkere er tilgængelig fra internettet⁸⁹.

88 http://www.advsolned.com/example_popcorn_noise.html.

89 http://www.analog.com/static/imported-files/rarely_asked_questions/moreInfo_raq_opAmpNoise2.html.
http://en.wikipedia.org/wiki/Johnson%20Nyquist_noise.

Din ekspert i
signalbehandlingsløsninger
og kommunikationsinterfaces

PERFORMANCE
MADE
SMARTER

TEMPERATUR | EX-BARRIERERER | KOMMUNIKATIONSINTERFACES | MULTIFUNKTIONEL | ISOLATION | DISPLAYS

Hos PR electronics sætter vi altid spørgsmålstegn ved konventionerne inden for signalbehandling, så procesingeniører overalt kan kommunikere med, overvåge og forbinde målepunkter med deres processtyringssystemer - effektivt og omkostningsbesparende. Vi styrer os af principper som enkelhed, fokus, mod og dygtighed og har fokus på at videreforske vores kernekompetence inden for innovativ højpræcisionsteknologi med lavt energiforbrug.

For at være på forkant med den nyeste teknologi har vi brug for den nyeste viden, og vi er løbende på udkig efter talentfulde personer, som er indstillet på at bygge bro mellem uddannelse og en faglig karriere, og som er klar til at begynde at opdage deres styrke gennem udfordrende jobs og projekter. Vi anser vores samarbejde med Ingeniørhøjskolen Aarhus Universitet som et særdeles vigtigt element i vores innovationsdrevne vækst, så gør som Rene og Kevin – bliv praktikant i en global innovativ virksomhed.

"PR electronics er en virksomhed, der fokuserer meget på den ansatte... og derfor også på mig som praktikant. Jeg har arbejdet på lige fod med alle andre og haft mulighed for at komme med mine egne indsparker..."

Rene Arendt Sørensen
Praktikant, Technology Department

"Jeg har fået en masse ny viden og er blevet klogere på, hvordan det er at arbejde sammen med andre ingeniører og det at arbejde i projektgrupper. Jeg er blevet mere klar over, hvad det vil sige at arbejde som ingeniør..."

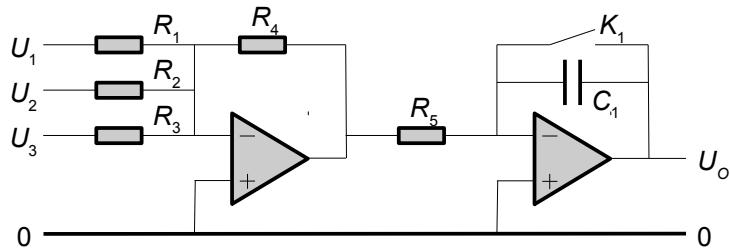
Kevin Grooters
Praktikant, Technology Department

Besøg <http://www.prelectronics.dk/career> og se, hvordan du kan blive en del af PERFORMANCE MADE SMARTER.

PR
electronics

3.12.1 Spændingsområde

De tidlige operationsforstærkere, typisk repræsenteret ved uA741, var designet til at arbejde med en effektforsyning på ± 15 V, hvilket ikke var tilfældigt valgt, for en meget væsentlig grund til at de integrererede forstærkere overhovedet blev udviklet var for anvendelse i analoge computere. Der var godt nok et voksende kommersIELT marked, men det var videnskabelige og militære anvendelser, der var den økonomiske drivkraft. Ved at begrænse det analoge signalområde til ± 10 V plus en reserve, kunne den målte spænding relativt let omsættes til en dermed proportionel talværdi, der i matematikken kunne repræsentere intervallet ± 1 .



Figur 169 – Simpel analog computer, der adderer tre signaler og foretager en integration.

De grundlæggende matematiske operationer er addition og integration. Kredsløbet udfører en addition (og inversion) af tre spændinger, der herefter integreres (og inverteres) over det tidsrum τ hvor kontakten K_1 er afbrudt. Ligningen er en byggeblok for løsning af differentialligninger, som der er mange af i de videnskabelige discipliner indenfor matematik, mekanik, radarelektronik, kemi og rumfart.

$$U_o = \frac{1}{T_1} \int_0^\tau (\alpha_1 U_1 + \alpha_2 U_2 + \alpha_3 U_3) dt, \quad T_1 = R_5 C_1, \quad \alpha_n = \frac{R_4}{R_n}, \quad -10 \text{ V} \leq U_o \leq 10 \text{ V}$$

3.12.2 Indgangens spændingsområde

På indgangen skal det modtagne signal omsættes til strømme i operationsforstærkerens interne kredsløb, og signaler kan ikke trækkes langt udenfor effektforsyningens grænser før uønskede (parasitiske) funktioner i siliciummet bliver aktiveret. Det er almindeligt at beskrive dette krav som det spændingsområde de to indgange skal holdes indenfor: *common-mode range, CMR*.

$$U_{CMR\ min} \leq (U_P, U_N) \leq U_{CMR\ max}$$

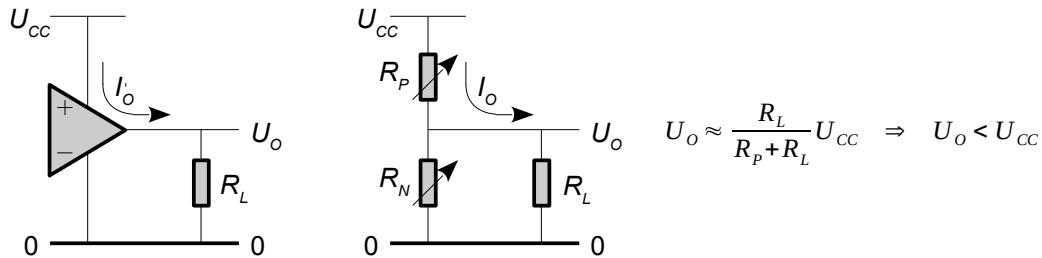
Området fra den laveste til den højeste spænding er en funktion af effektforsyningen, så de to grænseværdier er relateret til de arbejdsbetingelser man byder forstærkeren. Ved ± 15 V var det i gamle dage nok at holde indgangene indenfor ± 12 V for at være på den sikre side, men ved de moderne kredsløb for brug ved 3,3 V er det bydende nødvendigt at studere databladet.

Det er yderst sjældent at indgangsspændingen må gå mere end nogle få tiendedele volt under den negative linje; LM324 var den første OpAmp som tillod at arbejde med $-0,3$ V under nul, og kun få forstærkere kan klare en spænding over den øvre grænse fra effektforsyningen (LF356 og TL071).

Går man udenfor databladets grænser kan der ske særlige ting, som at udgangssignalets polaritet inverteres. En meget ubehagelig egenskab ved integrererede kredsløb er den destruktive *ch-up*, der skyldes en parasitisk funktion i siliciummet, hvorved kredsløbet omdannes til en kortslutning mellem effektforsyningens linjer. Strømmen begrænses kun af det eksterne kredsløb og det vil lede til afbrændte kredsløb. Denne ret uønskede egenskab plagede i lang tid især digitale kredsløb, men selv om der er ved at være styr på det skal man undgå at trække indgangene ud over de tilladte grænser, og det er normalt effektforsyningens grænser.

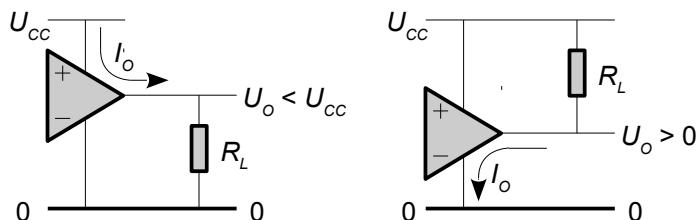
3.12.3 Udgangens spændingsområde

Operationsforstærkeren benytter transistorer til at stille udgangen på den ønskede spændingen og samtidigt at levere den krævede udgangsstrøm. Udgangstrinnet er opbygget som *push pull*, hvor der er to variable komponenter, der er vist som modstandene R_P og R_N i diagrammet herunder. Ved belastning af udgangen skal der løbe strøm i udgangens transistorer og der må nødvendigvis være et spændingsfald over dem. Udgangens spænding er derfor begrænset af effektforsyningen og kan aldrig levere et signal, der ligger udenfor effektforsyningens grænser.



Figur 170 – Udgangen kan opfattes som bestående af to variable modstande, der er styret af forstærkerens elektronik til at placere udgangen på den ønskede spænding, og samtidig skal der leveres den nødvendige strøm til belastningen.

Udgangstrinnet kan konstrueres på mange måder og det betyder at der kan være ganske stor forskel imellem operationsforstærkere. Med bipolare transistorer er det muligt at nå 50 mV fra effektforsyningen ved strækningen kollektor i udgangen, mens 600 mV er grænsen hvis basis-emitter er involveret. Med to transistorer i kaskade kan spændingsfaldet blive 1,2 V, som fx ses i LM324. Dertil kommer beskyttelse af forstærkeren imod overbelastning. Derved indskydes en modstand i serie med udgangen som giver et spændingsfald på 600 mV i tillæg til det allerede beskrevne. Med felt-effekt transistorer er udgangens funktion mere i retning af en modstand, så det er muligt at nå tæt på grænsen ved lave strømværdier. Uanset designet betyder det at man skal studere databladet grundigt hvis det er vigtigt at nå tæt på effektforsyningens grænser.



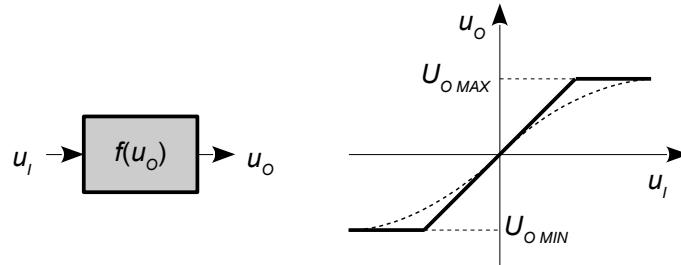
Figur 171 – Udgangen skal enten levere strøm til belastningen eller optage en strøm fra den, og der vil ligge et spændingsfald over udgangstrinnet, så det er ikke muligt at drive udgangen helt op til effektforsyningen eller helt ned til nul.

3.12.4 Udgangens strømområde

Udgangen skal levere strøm til belastningen eller den skal optage strøm fra den. Det stiller nogle krav til den strøm udgangen skal kunne arbejde ved, og det er ikke givet at den aktuelle forstærker er i stand til det. Strømmen er begrænset af det areal transistorerne optager på siliciummet og den temperatur de kan tåle at arbejde med. Almindelige forstærkere kan klare op til ± 25 mA mens de kraftige effektforstærkere som LM675T og L165V kan levere ± 3 A. Større strømværdi findes ved hybride kredsløb, hvor de integrerede komponenter er suppleret med diskrete effekttransistorer. For en endnu større strøm er det nødvendigt at konstruere en diskret opbygget effektforstærker med effekttransistorer, og det emne omhandles i kapitel fem.

3.12.5 Forvrængning

Operationsforstærkerens åben-sløjfe forstærkning A_{OL} antages normalt at være konstant ved en given frekvens, men den er desværre funktion af den øjeblikkelige værdi af udgangsspændingen. Det har den betydning at udgangssignalet ikke blot er en forstærket kopi af indgangssignalet, men en forvrænget kopi. Operationsforstærkerens åben-sløjfe forstærkning er størst ved svage signaler og reduceres ved store udsving fra middelværdien af signalet. Krumningen i karakteristikken tilfører anden og tredje harmoniske til signalet, som beskrevet i andet kapitel ved Fourier serierne.



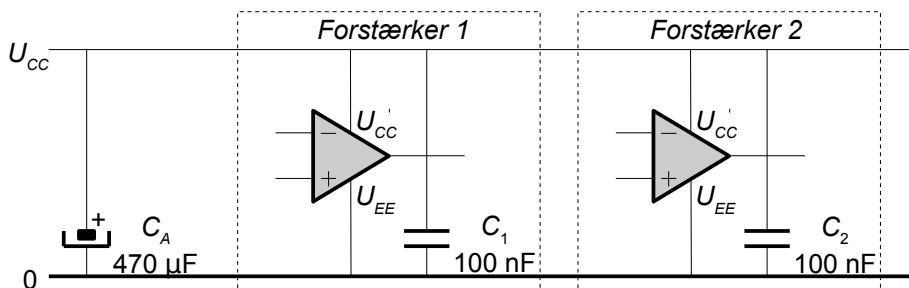
Figur 172 – Udgangssignalet er funktion af spændingsværdien på udgangen (stiplet linje). Den resulterende overføringsfunktion kan gøres mere lineær ved tilbagekobling, men den er begrænset af effektforsyningen og operationsforstærkerens konstruktion (fuld linje).

Forvrængningen udtrykkes som en totalharmonisk forvrængning THD, hvor amplituden af de harmoniske relateres til første harmoniske som en procentsats som vist i førest kapitel. Ved måling på testpersoner citeres ofte en nedre grænse på 0,25 % ved frekvenser midt i det hørbare område og ved forholdsvis højt lydtryk. I praksis er almindelige forbrugere forbløffende tolerante selv ved en ret voldsom harmonisk forvrængning, mens trænede lyttere med "gyldne ører" påstår at kunne høre en forvrængning, som modsætter sig måling. Påstandene har ikke kunne eftervises⁹⁰, men der er sandsynligvis mere mellem himmel og jord end givet ved en måling af THD. Der findes af samme årsag forskellige definitioner for hvorpå forvrængning kan provokes og måles.

Forvrængning opfattes normalt som noget uønsket, der skal undgås, men forvrængning kan i visse situationer være ønskværdig, som eksempel guitaristens *fuzz box*, hvor to dioder i tilbagekoblingen definerer signalets amplitude. Andre eksempler er reguleringeskredsløb, hvor kraftige signaler begrænses for ikke at overstyre elektronikken, og måling af en tones frekvens ændres ikke af overstyringen, men de efterfølgende kredsløb får et mere veldefineret signal at arbejde med.

3.13 Effektforsyning

Operationsforstærkeren har brug for en ekstern effektforsyning for at virke, for al elektronik består i en regulering af den elektriske strøm leveret fra en ekstern energikilde så som batteri eller netdel.



Figur 173 – Afkobling af forstærkeren er så godt som altid krævet for et stabilt resultat. En hyppigt set metode benytter en keramisk kondensator tæt på forstærkeren for afkobling af høje frekvenser, og dertil en fælles elektrolyt for hele kortet, som holder impedansen nede ved lavere frekvenser.

⁹⁰ Det er sjældent at personer med såkaldt "gyldne ører" stiller op til videnskabelige testforsøg. Man kan ikke undrage sig den tanke at de måske er klar over hvad de kan høre, og at de ikke ønsker det eftervist.

Effektforsyningen skal føres til operationsforstærkerens interne kredsløb, så der er behov for et par terminaler for at strømmen kan komme ind i forstærkeren og ud af den igen. Her i bogen benyttes U_{CC} som betegnelsen for den positive terminal og U_{EE} for den negative. De to betegnelser har deres oprindelse fra de første kommercielle operationsforstærkere hvor der kun blev benyttet transistorer af typen NPN med kollektor (*collector*) forbundet til plus (U_{CC} eller V_{CC}) og emitter til minus (U_{EE} eller V_{EE}). Fordoblingen af index til CC og EE benyttes ved effektforsyning, og det ses også ved de digitale komponenter og i mikrocomputeren, hvor den positive terminal er U_{DD} (eller V_{DD}) med D stående for *drain* og den negative terminal er U_{SS} (eller V_{SS}) med S for *source*.

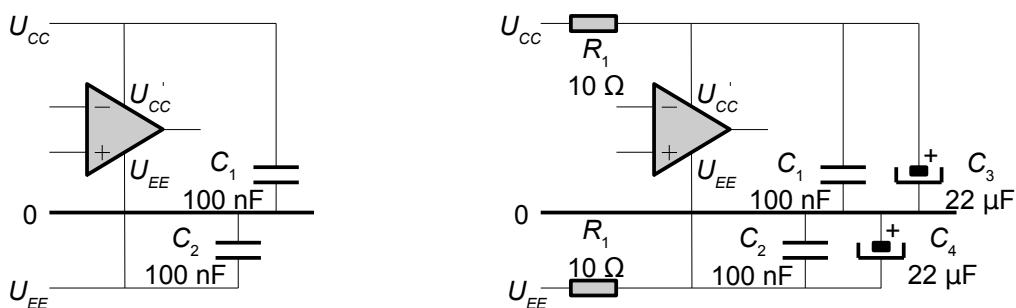
En typisk effektforsyning har 48 V for mikrofoners Phantom Power eller 24 V for industrielt udstyr som en PLC eller relæstyring. Analog elektronik har kørt fra ± 15 V eller ± 12 V og fra 9 V ved guitarpedaler. Mikrocomputeren har trukket elektronikken ned til først 5 V for at kunne benytte en fælles effektforsyning, og den analoge elektronik er fulgt med ned med 3,3 V og 1,8 V som de populære niveauer. Elektronik med bipolare transistorer kan arbejde ved 0,9 V og ved felteffekt transistorer kan effektforsyningen nå under 0,5 V.

Det er normalt at underforstå tilslutning af effektforsyningen hvorfor de kun meget sjældent vises i oversigtsdiagrammer. Ved komplette diagrammer er det normalt at samle effektforsyningen i det ene hjørne af diagrammet sammen med kondensatorerne for afkobling (*bypass capacitors*).

3.13.1 Afkobling

Vigtigheden af afkobling af effektforsyningen kommer som en overraskelse for mange, for de ses ikke i analytisk arbejde eller ved SPICE simulering af kredsløbet, og meget ofte virker de første opstillinger i laboratoriet som man forventer.

Operationsforstærkeren har ikke en terminal for nul (GND) så effektforsyningerne benyttes som reference for udgangsspændingen. Det betyder at variation i spændingsniveauet på en af effektforsyningens linjer vil slå igennem til udgangen. Tilbagekoblingen sikrer at udgangens spænding er korrekt i forhold til den valgte reference, hvilket som regel er nul, men det virker kun godt ved relativt lave frekvenser hvor der er en stor åben-sløjfe forstærkning. Mod høje frekvenser bliver forstærkningen ringe som følge af den dominerende pol, så operationsforstærkeren kan ikke længere nå at korrigere udgangsspændingen. Operationsforstærkeren har brug for stabile værdier af effektforsyningen, i det mindste ved høje frekvenser, og det opnås ved at have en lav impedans mellem effektforsyningens linjer og nulpunktet.



Figur 174 – Ved dobbelt forsyning bør hver af linjerne afkobles og det anbefales at de keramiske kondensatorer monteres tæt på operationsforstærkeren. For krævende analoge opstillinger ses undertiden en mere omfattende afkobling hvor hver forstærker desuden er isoleret fra støjen fra de øvrige forstærkere med en seriemarkant.

En lav impedans ved høje frekvenser kan tilbydes af kondensatorer, og det er almindelig praksis at benytte kondensatorer i området 100 til 470 μF , der sikrer at impedansen er tilstrækkelig lav, det vil sige under 1 Ω , ved frekvenser fra omkring 1 kHz og højere. Så store kondensatorer er af typen tantal eller aluminium elektrolyt, og de er ikke gode ved høje frekvenser. For at fastholde en god

afkobling ved høje frekvenser benyttes derfor i vid udstrækning keramiske kondensatorer i området 47 til 220 nF, som holder impedansen i effektforsyningen under 1 Ω fra cirka 1 MHz og op.

Som vist i første kapitel er der en selvinduktion associeret med enhver leder, og den er funktion af hvor stor arealet er for den sløjfe som kondensator, lederbaner og operationsforstærker udgør. For at holde selvinduktionen lav skal afkoblingen være så tæt på operationsforstærkeren som muligt, og det er lettest at opfylde hvis den keramiske kondensator monteres i umiddelbar nærhed af operationsforstærkeren. Følgelig er det anbefaelsesværdigt at montere en keramisk kondensator ved hver eneste operationsforstærker og ved hver eneste linje fra effektforsyningen⁹¹.

Det er ikke kun operationsforstærkerens begrænsede evne til at korrigere udgangen, der gør det nødvendigt med afkobling. Den interne opbygning af operationsforstærkerens kredsløb medfører problemer med stabilitet hvis der ikke er en tilstrækkelig lav impedans i effektforsyningen. Det er især et problem ved meget hurtige operationsforstærkere og design med lav forvrængning.

Når alle operationsforstærkere effektforsynes fra de samme linjer er der mulighed for at den ene kobler støj ind i effektforsyningen, som så forstyrre en anden operationsforstærker. Det kan derfor være en idé at isolere de mere følsomme kredsløb fra effektforsyningen med seriemodstande og en lokal afkobling med både keramiske kondensatorer og elektrolyt kondensatorer. Der er lidt forskel mellem de anbefalinger som man kan finde i databladene, men de ligger normalt omkring en modstand på 10 Ω og en elektrolyt på 22 µF. Grænsefrekvensen bliver lidt under 1 kHz som er i god overensstemmelse med den typiske operationsforstærkers dominerende pol.

3.13.2 Stelplan

Fra design af radio og hurtige digitale kredsløb er det god praksis at benytte et gennemgående kobberlag på printkortet ud over en afkobling af hver enkelt komponent. Det kaldes for et stelplan. Det kan lyde som luksus, men det har sin grund i erfaring med hurtige kredsløb. Uanset om det er en digital komponent eller en analog forstærker, så har den brug for en lav impedans ved de høje frekvenser, der er involveret, og de kan let nå over 1 GHz for en digital komponent med skiftetider i sub-nanosekund området.

Problemet med afkobling ved høje frekvenser er at selv om selvinduktionen kan holdes relativt lav, så er det svært at få den under 1 nH, der er en typisk værdi for overflademonterede komponenter, og det svarer til at impedansen kommer over 1 Ω for frekvenser over 100 MHz. Et stelplan giver en lav selvinduktion i den leder, der forbinder afkoblingen til komponentens nul og giver derfor en reduktion i den samlede længde af lederbane.

Et stelplan har dertil den gode egenskab at det skærmer for elektriske felter, og det kan være meget attraktivt når et følsomt kredsløb skal beskyttes imod indstrålet støj.

Indenfor den digitale verden er visse designere fortalere for at effektforsyningen også føres via kobberplaner. Der er god fornuft i forslaget, for serieinduktionen er meget lav i den brede leder, men kapaciteten mellem to planer er meget begrænset⁹² så det er ofte nødvendigt at supplere med keramiske kondensatorer. Det er sjældent at se forsyningsplaner i ved analog elektronik, blandt andet fordi det er ganske kostbart at afsætte tre kobberplaner til nul og effektforsyning.

⁹¹ Lærebøger og designartikler udviser meget store forskelle i holdningen hertil og mange bud har præg af en religiøs overbevisning, mere end en klar fysisk begrundelse for det aktuelle valg. Som elektronikudvikler har jeg naturligvis tillagt mig nogle vaner, som har vist sig at fungere, men jeg har ikke opbygget det samme kredsløb i flere varianter bare for at studere virkningen af afkoblingen. Jeg er med andre ord lige så præget af min lidt tilfældige erfaring som enhver anden elektronikudvikler. Det kan også formuleres som at nok er afkobling vigtig, men det er ikke altid vigtigt at den udføres på en bestemt måde.

⁹² For et europakort med målene 100 mm gange 160 mm, en dielektrisk konstant på 4,7 for materialet FR4 samt en antaget afstand mellem lagene på 0,1 mm bliver værdien under 10 nF ved brug af formlen fra første kapitel.

3.14 Opgaver

1 – Side 106 nævnes operationsforstærkeren LM324. Benyt din browser til at finde databladet og søg efter de steder hvor de nævnte parametre fremgår.

2 – Side 109 angives at et voltmeter med $10 \text{ M}\Omega$ indre modstand næsten kortslutter dioden. Beregn strømmen, der løber i voltmeterets indgang ved den angivne spænding og find den specifikation som en operationsforstærker skal have for højest at optage 10 % af diodens strøm ved svagt lys.

3 – Side 108 vises en ikke-inverterende forstærker med en forstærkning på 3 gange. Hvis der benyttes en LM324 med en effektforsyning på $\pm 15 \text{ V}$ skal det vises at indgangssignalet ikke må overstige $\pm 4,5 \text{ V}$. Bemærk at den angivne operationsforstærker ikke kan nå helt ud til effektforsyningens grænser.

4 – Side 111 er vist en inverterende forstærker. Signalet føres ind gennem modstand R_1 , som belaster kilden. Vis at indgangsstrømmen ved 1 V er på $100 \text{ }\mu\text{A}$.

5 – Mikrofonforstærkeren på side 110 for dynamisk mikrofon vil forstærke både AC og DC. Man kan tilføjes en kondensator i serie med R_3 for at reducere forstærkningen ved DC til én gang. Vis at der derved introduceres en lav grænsefrekvens givet ved $1/2\pi R_3 C_3$ og vis at der skal benyttes en kondensator på $122 \text{ }\mu\text{F}$ for at ramme samme grænsefrekvens som ved C_1 .

6 – Elektret mikrofonen på side 112 er specificeret til en følsomhed på typisk 5 mV/Pa ved $R_D = 2 \text{ k}\Omega$. Det antages at mikrofonen er en strømgenerator, og det skal vises at det svarer til en specifikation på $2,5 \text{ }\mu\text{A/Pa}$. Mikrofonens udgang ved R_D er belastet med R_L og vis at det ikke påvirker den oplyste følsomhed af mikrofonen. Vis endelig at grænsefrekvensen er bestemt til $1/2\pi C(R_D + R_L)$.

7 – Vis at for kredsløbene på side 112 vil to signaler med henholdsvis $U_1 = 1 \text{ V}$ og $U_2 = 2 \text{ V}$ give $U_o = 1,33 \text{ V}$ med de viste modstandsværdier for det første og $U_o = -4 \text{ V}$ for det andet kredsløb. Antag værdierne $C_S = C_1 = 10 \text{ pF}$ og vis at grænsefrekvensen bliver $f_H = 2,4 \text{ MHz}$ for det første kredsløb og $f_H = 800 \text{ kHz}$ for det andet. En LM324 har en specifikation på $GBP = 1 \text{ MHz}$ og vis at den vil reducere båndbredden ved begge kredsløb.

8 – Formlerne for kredsløbene på side 113 skal eftervises ved superposition hvor indgangssignalerne analyseres ét for ét. For at bestemme første led i formlen nulstilles U_2 og begge kredsløb bliver til inverterende forstærkere overfor U_1 . For at bestemme andet led skal U_1 nulstilles og forstærkningen overfor U_2 bestemmes.

9 – For AC strømkilden på side 117 skal formlen for udgangsstrømmen eftervises. Start med at bestemme udtrykket for I_1 og bestem herfra udtryk for U_A og U_o . Brug Kirchhoff for at opstille en ligning for udgangsstrømmen I_o og indsæt heri de fundne udtryk for U_A og U_o .

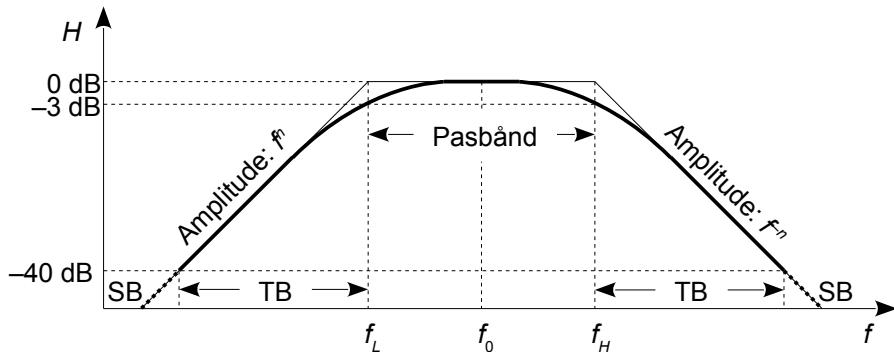
Operationsforstærker

Note 5.

4 FILTRE

Filtrering af et signal betyder almindeligvis at begrænse båndbredden, så som at fjerne DC ved at bruge et højpasfilter, eller at dæmpe højfrekvent støj ved et lavpasfilter. Nogle typiske eksempler er at give højtalere optimale arbejdsbetegnelser ved at begrænse båndbredden til den enkelte enhed samt at justere balancen mellem dybe og høje frekvenser i et lydanlæg. I en radio bruges passive filtre for tilpasning af en antennes impedans til elektronikken, til at udvælge radiostationen og undertrykke nabostationer samt til detektion af signalet. Ved de ældre lydmedier som grammofon og tonebånd blev frekvensresponsen korrigeret før og efter optagelsen for at dæmpe støj, og et stopfilter kan dæmpe uønskede signaler, som fx 50 Hz netbrum. En transducers signal kan kræve differentiering eller integration, og et signals fase kan korrigeres for at udligne en tidsforskydning, hvilket udnyttes ved højtalersystemer, langdistance telefoni og lydeffekter.

Filtre beskrives ved overføringsfunktionen, hvor der ofte fokuseres på amplitudens variation med frekvensen, fx som illustreret herunder for et båndpasfilter, men der kan også være interesse for andre parametre, som fasens variation med frekvensen eller stepresponsens udvikling over tid. Filtre af typen lavpas, båndpas og højpas beskrives ved grænsefrekvensen hvor filtrets amplitude skifter fra at være en funktion af frekvensen til at være tilnærmelsesvist konstant. Et selektivt filter kan desuden svinge villigt med på et signals frekvenskomponent (resonans).



Figur 175 – Amplituden af et båndpasfilter vist i dobbelt-logaritmisk afbildning (Bode plot, tynd streg) og den aktuelle amplitude (tyk streg). Grænsefrekvensen ved -3 dB er normalt markør for hvortil pasbåndet går. Områder indikeret med TB står for transmissionsbåndet hvor dæmpningen gradvist går imod asymptotens værdi, og med SB for stopbåndet, der arbitraert er vist ved -40 dB svarende til et niveau på højest 1 % af det oprindelige signal.

Det illustrerede filter er af typen båndpas da det tillader et signal at passere hvis frekvensen er i pasbåndet fra den lave grænsefrekvens f_L til den høje grænsefrekvens f_H hvor signalet er 3 dB under det maksimale niveau, der ofte er ved centerfrekvensen f_0 . Under den lave grænsefrekvens og over den høje er signalet i transmissionsbåndet fra pas- til stopbånd. Niveauet ved overgangen til stopbåndet har ikke en fast definition, men afhænger af formålet med filtret.

Filtrene opdeles efter overføringsfunktionens orden og karakteren af dens poler. Det simplest mulige filter er af første orden (én pol), som har polen placeret på den reelle akse i den venstre halvplan. Filter kan have højere orden, og ordnen er givet ved eksponenten til s i nævnerens polynomium. Ved filter af anden orden er polynomiet poler enten reelle, i hvilket fald det er muligt at splitte overføringsfunktionen op i to første-ordens blokke, eller de er komplekst konjugerede, og filtret vil beskrives ved dæmpningsfaktoren⁹³ d eller godheden $Q = 1/2d$.

Filtre af højere orden kan altid opdeles i blokke af anden orden, og for filter af ulige orden tilmed en enkelt blok af første orden. Det er følgelig kun nødvendigt at behandle designet af første-ordens og anden-ordens filtere.

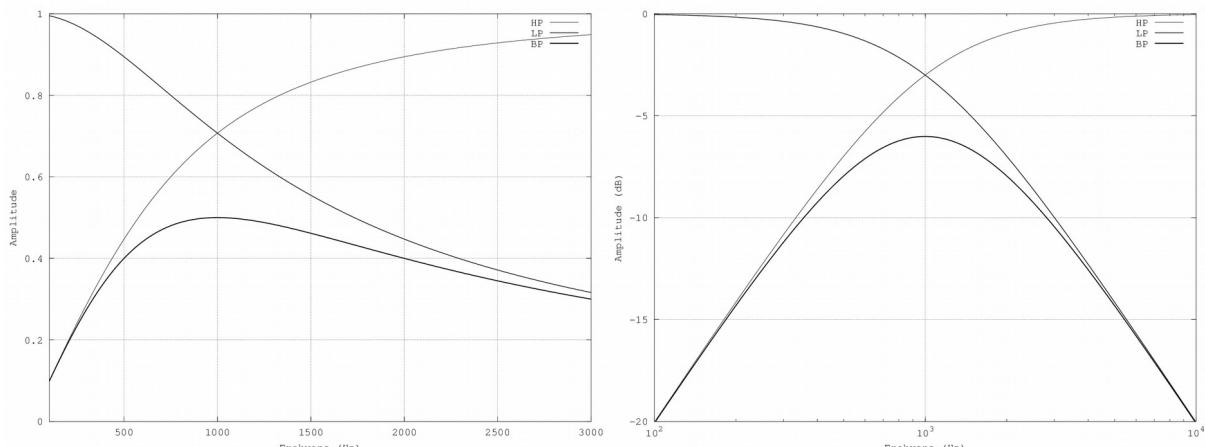
⁹³ En del kilder benytter det græske bogstav zeta (ζ) for dæmpningsfaktoren.

4.1 Filtre af første orden

Der er kun to muligheder for filtre af første orden; det er henholdsvis højpas og lavpas, hvor polen angiver grænsefrekvensen hvor niveauet er reduceret til -3 dB i forhold til pasbåndet, svarende til at halvdelens af signaleffekten er mistet. Amplituden er proportionel med frekvensen ($6\text{ dB}/\text{oktav}$) for højpas ved frekvenser mindst en oktav under polen, og for lavpas er amplituden omvendt proportionelt med frekvensen ($-6\text{ dB}/\text{oktav}$) fra godt en oktav over polen. Stejlheden er moderat og svarer til en radios tonekontrol med bas- eller diskant drejet helt ned, så der slipper et betydeligt signal igennem udenfor pasbåndet; amplituden af et signal i en dekades afstand fra grænsefrekvensen er kun⁹⁴ reduceret med 20 dB .

Filtertype	Overføringsfunktion	Lavfrekvens	Højfrekvens
Højpas	$H_{HP} = \frac{s}{\omega_L + s}$	$H_{HP} \underset{ s < \omega_L}{\approx} \frac{s}{\omega_L}$	$H_{HP} \underset{ s > \omega_L}{\approx} 1$
Lavpas	$H_{LP} = \frac{\omega_H}{\omega_H + s}$	$H_{LP} \underset{ s < \omega_H}{\approx} 1$	$H_{LP} \underset{ s > \omega_H}{\approx} \frac{\omega_H}{s}$

Overføringsfunktionens nævner er et polynomium af første orden. Ved lave frekvenser er s mindre end polen, hvorved nævneren bliver stort set konstant, og ved høje frekvenser dominerer s . Ved et højpasfilter er tællerens frekvensproportional, så amplituden bliver proportional med frekvensen ved lave frekvenser og konstant for frekvenser over grænsefrekvensen. Tælleren er konstant ved et lavpasfilter så amplituden bliver konstant ved lave frekvenser og aftager mod høje frekvenser.



Figur 176 – Amplituden plottet for første-ordens filtre af typen højpas (HP) og lavpas (LP), samt for båndpas dannet af de to filtre i kæde (BP). Til venstre vises et plot med lineære akser og til højre et plot med logaritmiske akser. Grænsefrekvensen er begge 1 kHz hvor HP og LP er 3 dB nede.

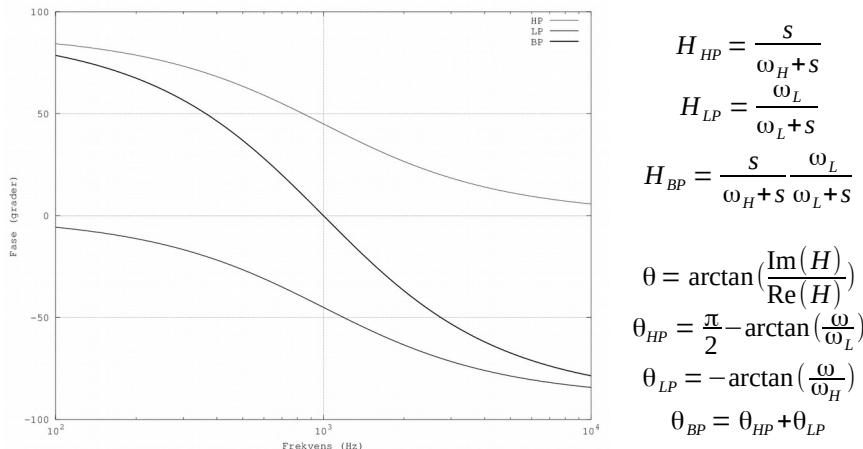
Amplituden kan bestemmes analytisk som den numeriske værdi af overføringsfunktionen, der er givet ved forholdet mellem amplituden af tæller og nævner. Tællerens amplitude er proportional med frekvensen for højpas og konstant for lavpas. Nævnerens amplitude består af to led, det reelle led ω_L eller ω_H og det imaginære led $s = j\omega$, og amplituden af nævneren er givet ved RMS addition.

$$\begin{aligned} \text{Højpas: } A_{HP} &= \left| \frac{s}{\omega_L + s} \right| = \frac{\omega}{\sqrt{\omega_L^2 + \omega^2}} = \frac{1}{\sqrt{1 + (f_L/f)^2}} & \text{hvor } \omega = 2\pi f \\ \text{Lavpas: } A_{LP} &= \left| \frac{\omega_H}{\omega_H + s} \right| = \frac{\omega_H}{\sqrt{\omega_H^2 + \omega^2}} = \frac{1}{\sqrt{1 + (f/f_H)^2}} & \omega_L = 2\pi f_L \\ && \omega_H = 2\pi f_H \end{aligned}$$

⁹⁴ Subjektivt svarer 10 dB reduktion til en halvering af lydniveauet, så ved -20 dB er det subjektive lydniveau kun reduceret til en fjerdedel. Filtrer er i mange henseender for svagt til at kunne dæmpe et støjsignal tilstrækkeligt.

Filtre

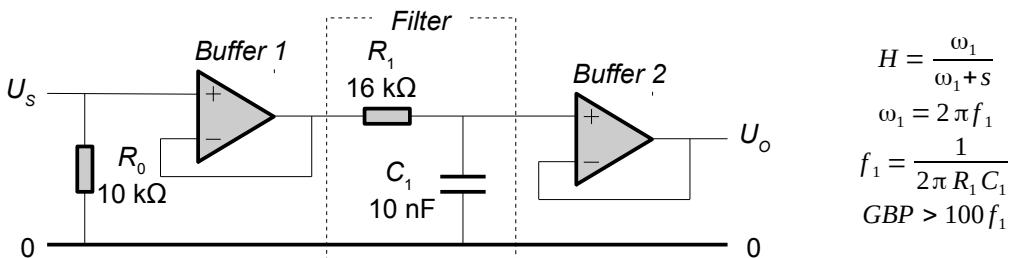
Der er en fase relateret til responsen og for første-ordens filtre er asymptoten 90° når amplituden stiger proportionelt med frekvens (signalet differentieres), den er 0° ved en konstant amplitude, og den er -90° når amplituden aftager med frekvensen (signalet integreres). I nærheden af polen (grænsefrekvensen) vil fasen forløbe blødt fra den ene asymptote til den anden, som det vises i illustrationen herunder for højpas, lavpas og båndpas, der dannes af både højpas og lavpas.



Figur 177 – Fasen plottet for første-ordens filtre af typen højpas (HP) og lavpas (LP), samt båndpas dannet af de to filtre i kæde (BP). Grænsefrekvensen er valgt til 1 kHz hvor fasen er drejet -45° .

4.1.1 Ikke-inverterende lavpasfilter

Et første-ordens lavpasfilter vises herunder med selve filtret i den stipede boks. Der er indskudt en buffer både før og efter filtret for at gøre funktionen uafhængig af det kredsløb, der omgiver filtret. Grænsefrekvensen er givet af RC produktet, så den lave udgangsmodstand fra Buffer 1 er garant for at modstanden R_1 ikke utilsigtet forøges af impedansen af det drivende kredsløb ved U_S . Buffer 2 sikrer tilsvarende at udgangen fra filtret ikke belastes af det tilsluttede kredsløb ved U_O .



Figur 178 – Et første-ordens lavpasfilter med buffere så kredsløbet ikke påvirkes af kredsløbet til venstre eller af kredsløbet til højre. Grænsefrekvensen er 1 kHz med de viste værdier.

4.1.2 Almindelige designregler

Værdien af R_1 vælges oftest i intervallet fra 1 k Ω til 100 k Ω . Buffer 1 skal drive strøm gennem R_1 og ved høje frekvenser er C_1 nærmest en kortslutning så modstanden belaster bufferen. Ved et signal med 10 V amplitude vil 1 k Ω kræve ± 10 mA og det er ret meget for en operationsforstærker.

For at opnå mindre end 1 % påvirkning af grænsefrekvensen skal udgangsmodstanden fra Buffer 1 holdes under $R_1/100 = 10 \Omega$ og det vil en spændingsfølger normalt kunne tilbyde. Tilsvarende skal indgangsmodstanden på Buffer 2 være over $100 R_1 = 10 \text{ M}\Omega$ for under 1 % påvirkning og det vil normalt også kunne opfyldes. Problemet optræder først ved så høje frekvenser at den benyttede forstærker ikke har nok åben-sløjfe forstærkning til at garantere de krævede værdier og det giver et krav til forstærkerens GBP specifikation, der som en tommelfingerregel mindst bør være $100 f_1$.

Filtre

Det er desuden nødvendigt at overveje konsekvensen af kapaciteten fra det omgivende kredsløb. En tommelfingerregel siger at der må påregnes en kapacitet fra printkortets baner på cirka 1 pF per centimeter alene fra den fysiske realisation af filtret og den kapacitet ligger i parallel med C_1 . Dertil kommer at Buffer 2 har en indgangskapacitet på typisk 2 pF og dertil en kapacitet af samme størrelsesorden på grund af operationsforstærkerens dominerende pol (se tredje kapitel). Den totale værdi af kondensatoren vil derfor øges med cirka 5 pF på grund af den fysiske realisering og det betyder at C_1 ikke bør vælges mindre end 500 pF hvis fejlen herfra skal holdes under 1 %. Det er fuldt ud muligt at benytte en lavere værdi af kondensatoren men formlerne vil næppe passe.

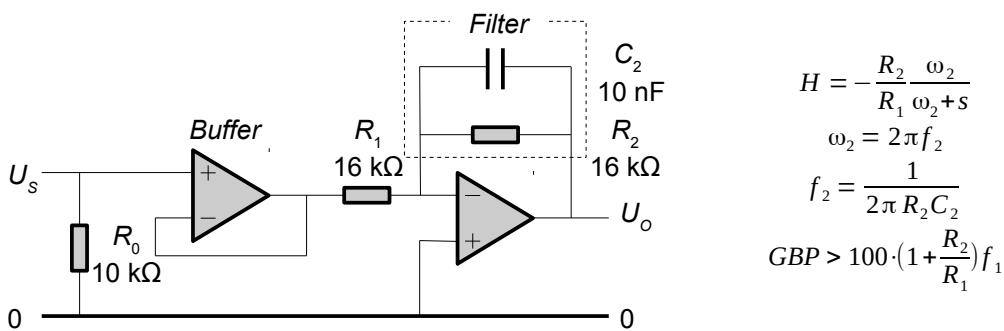
Der er en øvre grænse for kondensatorens værdi i omegnen af 10 μF for de gode keramiske og metalfilm typer og de elektrolytiske kondensatorer bør helt undgås ved signalbearbejdning. Disse krav vil som en tommelfingerregel holde grænsefrekvensen f_1 i området fra 0,2 Hz til 300 kHz.

Modstanden R_0 har det formål at sikre en DC forbindelse til nul. Det er vigtigt når signalet kommer via en kondensator eller hvis indgangen kan være afbrudt. Uden en DC forbindelse vil indgangens DC potentielle ikke være veldefineret og spændingen fra Buffer 1 kan drive væk fra nul og vil ende ved en af effektforsyningens grænser. Værdien skal vælges så lav at biasstrømmen for Buffer 1 ikke giver anledning til et betydnende DC niveau på indgangen. Med en biasstrøm på $I_B = 100 \text{ nA}$ vil den viste $R_0 = 10 \text{ k}\Omega$ give en spænding på 1 mV og det vil almindeligvis være acceptabelt.

I tillæg til disse designregler bør ethvert analogt og digitalt kredsløb benytte afkobling af samtlige effektforsyningens linjer. Min professionelle erfaring er at en keramisk kondensator på 220 nF bør monteres tæt på hver af operationsforstærkerens effektforsyningslinjer (både Buffer 1 og 2) samt at der uden undtagelse benyttes et printkort med et fuldt dækende stelplan.

4.1.3 Inverterende lavpasfilter med forstærkning

En populær realisation af første-ordens filtret vises herunder, hvor en inverterende forstærker har båndbredden begrænset at en kondensator i parallel i tilbagekoblingen. Grænsefrekvensen f_1 er givet ved den frekvens hvor impedansen af kondensatoren er lig med modstandens værdi. Det er ikke nødvendigt med en buffer for at sikre grænsefrekvensen, men den kan være påkrævet for at opnå den indstillede værdi af forstærkningen, idet den nu indstilles af R_1 .



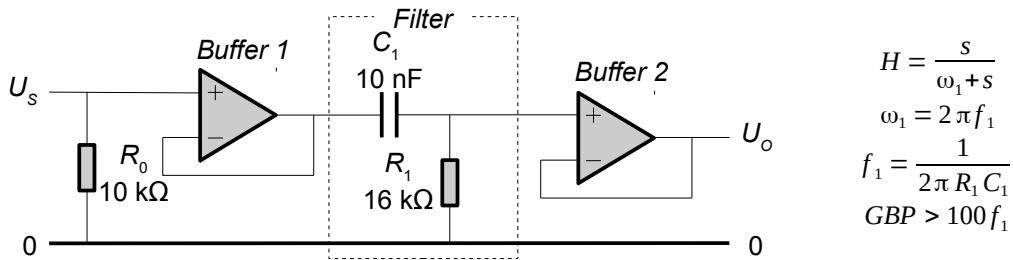
Figur 179 – Et første-ordens lavpasfilter med en inverterende forstærker. Bufferen kan være påkrævet for at opnå den ønskede forstærkning. Grænsefrekvensen er 1 kHz med de viste værdier.

En attraktiv egenskab er at der kan opnås både forstærkning eller dæmpning af signalet ved at vælge en passende værdi for R_1 . Signalet inverteres hvilket alt efter temperament og den aktuelle opgave kan være en fordel eller en ulempe.

Opstillingen er meget udbredt for der er så godt som altid inkluderet et lavpasfilter som vist i de almindelige kredsløb for lydbehandling. Grænsefrekvensen er 50 kHz eller højere. Det reducerer detektion af radiostationers transmission og den generende brummen fra mobiltelefoner og det vil almindeligvis forbedre forstærkerens stabilitet overfor utilsigted kapacitet fra printkortets stelplan.

4.1.4 Ikke-inverterende højpasfilter

Et højpasfilter realiseres som vist herunder hvor R_1 og C_1 er frekvensbestemmende. De to buffere sikrer igen optimale betingelser for filtret men er kun nødvendige hvis kredsløbet omkring filtret ikke kan holde en lav udgangsmodstand og en høj indgangsmodstand. Der gælder de samme krav til de benyttede operationsforstærkere som før selv om belastningen fra indgangskapaciteten ikke er et stort problem. Vær opmærksom på at den første buffer belastes med R_1 indenfor det aktive frekvensområde over grænsefrekvensen så bufferen skal kunne levere den nødvendige strøm fra udgangen. Anbefalingerne fra afsnittet på side 173 er derfor også gyldige her.



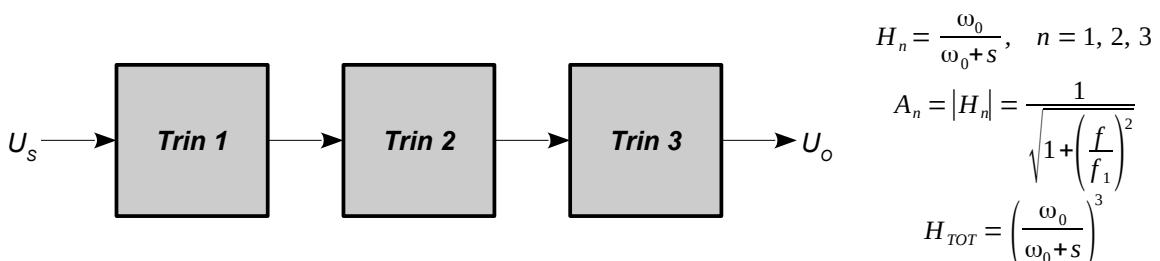
Figur 180 – Et første-ordens højpasfilter med buffere så kredsløbet ikke påvirkes af kredsløbet til venstre eller til højre. Grænsefrekvensen er 1 kHz med de viste værdier.

Også denne opstillingen er populær i blandt andet lydverdenen, for højpasfiltret kan blokere for et DC signal og dermed sikre at kredsløbet efter filtret er uafhængigt af DC niveauet før filtret.

Med en grænsefrekvens på 10 Hz vil C_1 blive på 1 μ F med den viste modstandsværdi. Til ukritiske formål vil en keramisk, overflademonteret kondensator med X7R dielektrikum være mest attraktiv ud fra pladskrav og økonomi, mens mere krævende formål vil have fordel af en plastfilm type. Den er fri for mikrofoni⁹⁵ og kan leveres uden dielektrisk absorption⁹⁶ der kan være et problem ved brug i industrielle applikationer og krævende lydformål.

4.1.5 Kaskade af filtre

Megen elektronik inkluderer mere end et enkelt lavpas- eller højpasfilter og meget ofte vil de have nogenlunde ens grænsefrekvens. Det betyder at den samlede dæmpning ved grænsefrekvenserne bliver væsentligt kraftigere end de forventede -3 dB. Hvis der eksempelvis er tre ens lavpasfiltre med $f_1 = 50$ kHz i kaskade vil de tilsammen give -9 dB ved frekvensen f_1 så den effektive værdi af filtrene er en afskæring ved cirka 25 kHz hvor hvert af filtrene dæmper med 1 dB.



Figur 181 – Et større system kan indeholde mange båndpasfiltre i kaskade og den resulterende respons kan blive degraderet af den successive begrænsning af frekvensområdet.

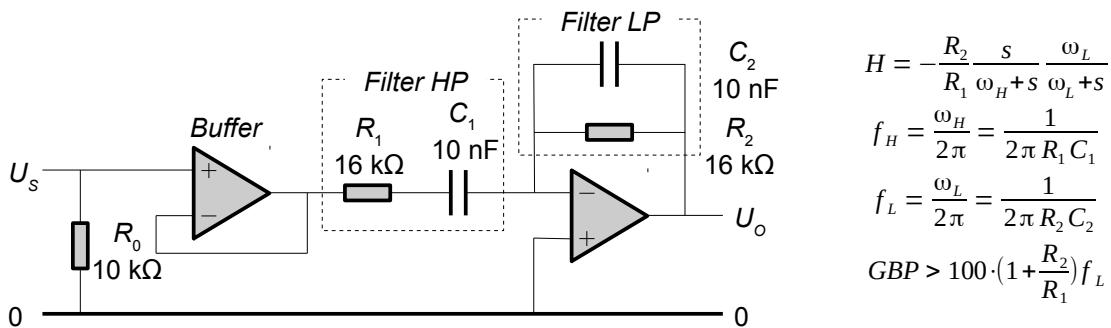
For en analog mikserpult i et lydstudie er der mange trin i signalkæden, som alle påvirker det samlede resultat så det er nødvendigt at overveje konsekvensen af at sammenkoble delene.

⁹⁵ Kondensatorens dielektrikum er piezoelektrisk og genererer en spænding når den udsættes for chok og rystelser.

⁹⁶ Dielektrisk absorption er en slags hukommelse, hvor en kondensators ladning vil ændre sig over et tidsrum som følge af situationen før en op- eller afladning påbegyndes, hvorefter kondensatoren overlades til sig selv. Det er et betydeligt problem i opstillinger hvor kondensatoren anvendes til at fastholde en DC værdi.

4.1.6 Inverterende båndpasfilter

Filtret herunder består af et højpasfilter i kaskade med et lavpasfilter og det giver et båndpasfilter. Det kan diskuteres om et båndpasfilter kan være af første-orden for overføringsfunktionen har to poler men det vil her blive opfattet som to første-ordens filtre i kaskade. En buffer sikrer at den indstillede forstærkning opnås og at højpasfiltrets grænsefrekvens f_L bliver som forventet.



Figur 182 – Et båndpasfilter med en inverterende forstærker. Bufferen kan være påkrævet for at opnå den ønskede forstærkning. Grænsefrekvensen er 1 kHz med de viste værdier.

Forstærkningen i pasbåndet indstilles gennem forholdet mellem R_2 og R_1 men kan ikke vælges helt frit idet begge modstande også skal varetage filtrering. Problemet med filtre i kaskade ses tydeligt fra kravet om at lavpasfiltrets grænsefrekvens f_H skal ligge langt over højpasfiltrets f_L for at undgå at de to filtre påvirker hinanden.

Båndpasfilter	Lavfrekvens	Mellemfrekvens	Højfrekvens
$H = \frac{s}{\omega_L + s} \frac{\omega_H}{\omega_H + s}$	$H \underset{ s < \omega_L}{\approx} \frac{s}{\omega_L}$	$H \underset{\substack{ s > \omega_L \\ s < \omega_H}}{\approx} 1$	$H \underset{ s > \omega_H}{\approx} \frac{\omega_H}{s}$
$\omega_L \ll \omega_H$	+ 6 dB/oktav + 90°	Pasbånd (0 dB) 0°	- 6 dB/oktav - 90°

Der er intet galt ved at vælge $f_H = f_L$ for det betyder blot at forstærkningen i pasbåndet kun når op på -6 dB af den indstillede værdi ved toppunktet. Der er derimod intet vundet ved at vælge en grænsefrekvens af højpasfiltret som er højere end lavpasfiltret for det giver blot en endnu større dæmpning i pasbåndet uden i øvrigt at ændre på filtreringen.

4.1.7 Fasedrejningsfilter

Hvis signalet fra lavpasfiltret forstærkes med faktor to og indgangssignalet derefter trækkes fra vil resultatet blive en overføringsfunktion der på engelsk går under betegnelsen *all-pass filter* fordi der ikke fjernes noget fra indgangssignalet. Jeg vil her kalde det for et fasedrejningsfilter, for derved at fokusere på det filtret essentielt foretager sig og derefter viser jeg hvordan det kan benyttes til at forsinke et signal i tidsdomænet.

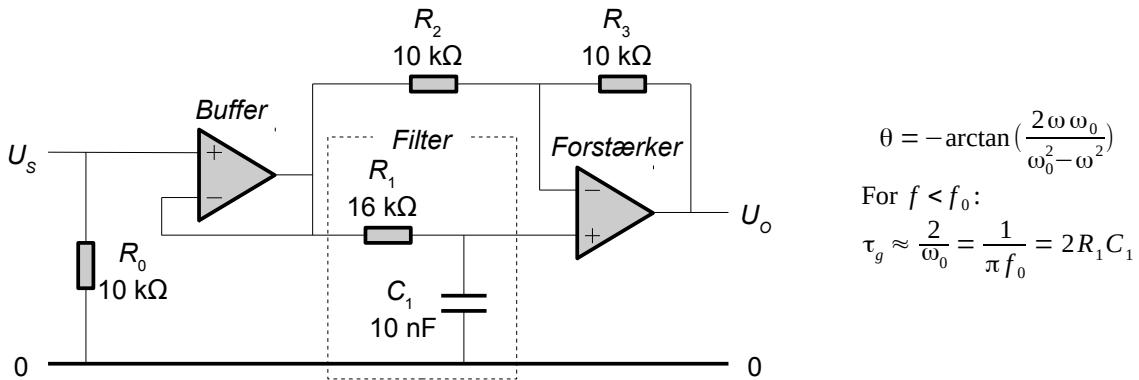
Et fasedrejningsfilter kan konstrueres som vist herunder hvor grænsefrekvensen gives af R_1 og C_1 . Ved lave frekvenser er kondensatoren ikke aktiv idet impedansen er stor i sammenligning med R_1 så forstærkeren er en ikke-inverterende spændingsfølger. Ved høje frekvenser er impedansen af kondensatoren lav så forstærkeren er inverterende. Følgelig vil fasen dreje fra nær 0° ved lave frekvenser til -180° ved høje frekvenser⁹⁷.

Bufferen i indgangen har samme funktion som nævnt tidligere men kan undværes når signalkilden har en tilstrækkeligt lav impedans og der skal altid være en DC forbindelse til nul.

97 Det er med vilje at en inverterende forstærker ikke betegnes med +180°, for det ville svare til et filter hvor signalet blev skudt frem i tid hvilket er fysisk umuligt. Det kan dog af og til være praktisk at se på en fasedrejning som om den skubber signalet frem i tid.

Filtre

Den frekvensafhængige fase er tilnærmelsesvis en lineær funktion af frekvensen for frekvenser under f_0 hvilket gør at den kan relateres til en tidsforskydning. Så hvis en sinussvingning med en frekvens på f påtrykkes filtret bliver resultatet en forsinkelse af svingningen. Tidsforsinkelsen er ikke voldsom, med $f_0 = 1 \text{ kHz}$ er den på $\tau_g = 0,3 \text{ ms}$, men det er alligevel nok til at finde praktisk anvendelse. Fasedrejningsfiltret ses anvendt i telefonsystemer til at korrigere for lydens forskellige løbetider ved lave og høje frekvenser i lange kabler og i højtalersystemer til at korrigere for en uønsket tidsforsinkelse på grund af højtalernes placering på forpladen⁹⁸.



$$\theta = -\arctan\left(\frac{2\omega\omega_0}{\omega_0^2 - \omega^2}\right)$$

For $f < f_0$:

$$\tau_g \approx \frac{2}{\omega_0} = \frac{1}{\pi f_0} = 2R_1C_1$$

Figur 183 – Et phasedrejningsfilter påvirker ikke amplituden. Fasen drejes 0° ved lave frekvenser, -90° ved grænsefrekvensen og -180° ved høje frekvenser. Grænsefrekvensen er 1 kHz med de viste komponentværdier hvor $\theta = -90^\circ$ og modstandene R_2 og R_3 skal være ens.

Overføringsfunktionen udledes herunder. Den kan skrives som en lineær fase (tidsforsinkelse) hvis arctan funktionen beskrives ved det første led af Taylor-serien. Simplifikationen er god for $x < 1$ hvor fejlen er under 20 % så den kan ikke anvendes for frekvenser over f_0 .

Overføringsfunktionen kan opskrives som følger.

$$H = 2H_{LP} - 1 = 2\frac{\omega_0}{\omega_0 + s} - 1 = \frac{\omega_0 - s}{\omega_0 + s} \quad \text{hvor} \quad \omega_0 = \frac{1}{RC}$$

Amplituden af overføringsfunktionen er konstant (den er én), så signalets amplitude påvirkes ikke. Det ses ved at beregne "længden" af tæller og nævner efter Pythagoras.

$$|H| = \left| \frac{\omega_0 - j\omega}{\omega_0 + j\omega} \right| = \frac{\sqrt{\omega_0^2 + \omega^2}}{\sqrt{\omega_0^2 + \omega^2}} = 1$$

Fasen af overføringsfunktionen beregnes af forholdet mellem den imaginære del og den reelle del så der er behov for at dele overføringsfunktionen op i disse to dele. Det opnås ved at gange med det komplekst konjugerede til nævneren i både tæller og nævner.

$$H = \frac{\omega_0 - j\omega}{\omega_0 + j\omega} = \frac{\omega_0 - j\omega}{\omega_0 + j\omega} \frac{\omega_0 - j\omega}{\omega_0 - j\omega} = \frac{\omega_0^2 - \omega^2}{\omega_0^2 + \omega^2} - j \frac{2\omega\omega_0}{\omega_0^2 + \omega^2}$$

Fasen kan nu beregnes. Den er tæt på 0° ved lave frekvenser hvor den imaginære del er lille, den er -90° ved grænsefrekvensen hvor argumentet til arctan går imod uendelig og den går imod -180° ved høje frekvenser idet tællerne igen går imod nul.

$$\theta = \arctan\left(\frac{\text{Im}(H)}{\text{Re}(H)}\right) = -\arctan\left(\frac{2\omega\omega_0}{\omega_0^2 - \omega^2}\right) = -\arctan\left(\frac{2x}{1-x^2}\right) \quad \text{hvor} \quad x = \frac{\omega}{\omega_0}$$

98 Der kan foretages en "elektrisk flytning" af en højttalerenhed tilbage fra forpladen, så signalet bringes i fase med en anden højttalerenhed. Hvis korrektionen baseres på $f_0 = 10 \text{ kHz}$ er $\tau_g = 30 \mu\text{s}$, og med lydens hastighed på $c = 340 \text{ m/s}$ vil det svare til en justering på $c\tau_g \approx 10 \text{ mm}$. Hvis der er behov for en større værdi af tidsforsinkelsen kan flere phasedrejningsled seriekobles.

Taylor-serien for arctan vises herunder.

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots \Rightarrow \arctan(x) \approx x \quad \text{for } |x| < 1$$

Brugt på udtrykket for fasen betyder det at ω^2 i nævneren kan ignoreres, fordi frekvensen antages lavere end grænsefrekvensen, og fasen bliver tilnærmelsesvis proportionel med frekvensen. Ved sammenligning med udtrykket for sinussvingningen ses at konstanten τ_g kan udtrykkes ved to gange filtrets tidskonstant, så relationen er brugbar fra 0° til -90° .

$$\theta = -\arctan(2\frac{\omega}{\omega_0}) \approx -2\frac{\omega}{\omega_0} \Rightarrow \theta \approx -2\frac{2\pi f}{\omega_0} = -2\pi(2RC)f = -2\pi\tau_g f \Rightarrow \tau_g = 2RC$$

I den teoretiske litteratur er tidsforsinkelsen givet af den afledte af fasen med hensyn til frekvensen, og resultatet kaldes for gruppetidsforsinkelsen (*group delay*). Begrænses analysen til området under grænsefrekvensen, hvor simplifikationen er gyldig, findes helt det samme udtryk som før.

$$\tau_g \stackrel{\text{def}}{=} -\frac{d\theta}{d\omega} \Rightarrow \tau_g \approx -\frac{1}{\omega_0} \frac{d(-\arctan(2x))}{dx} = \frac{2}{\omega_0} = 2RC$$

Med en sinussvingning på indgangen af filtret bliver resultatet en tidsforsinkelse τ_g der normalt kaldes for gruppeløbetiden (*group delay*).

$$u(t) = U_m \sin(2\pi t + \theta) \Rightarrow u(t) = U_m \sin(2\pi f(t - \tau_g))$$

$$\theta = -2\pi\tau_g f$$

Filtre

Note 6.

4.2 Filtre af anden orden

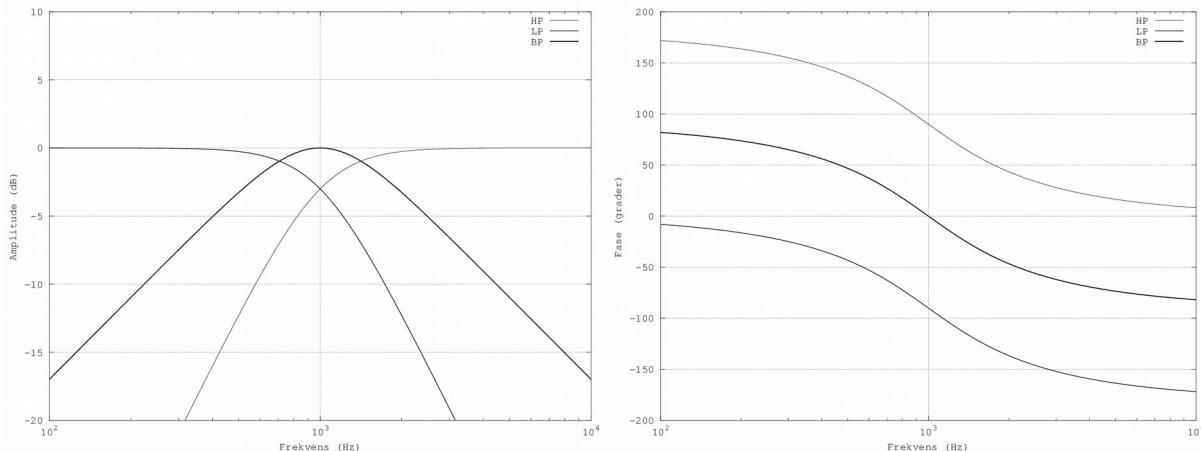
Et anden-ordens filter har to poler, og de er enten reelle eller komplekst konjugerede. I det første tilfælde kan filtret opfattes som bestående af to første-ordens filtre i kaskade, så det bringer intet nyt stof på banen. I det andet tilfælde skabes et filter med mulighed for at påvirke responsen i omegnen af grænsefrekvensen, hvilket giver nogle væsentlige muligheder.

$$\begin{aligned} \text{Højpas: } H_{HP} &= \frac{s^2}{\omega_0^2 + 2d\omega_0 s + s^2} & f_d &= \sqrt{1+d^2} f_0 & f_0 &= \frac{\omega_0}{2\pi} \\ \text{Båndpas: } H_{BP} &= \frac{2d\omega_0 s}{\omega_0^2 + 2d\omega_0 s + s^2} & f_d &= f_0 & 0 < d \leq 1 & Q = \frac{1}{2d} \\ \text{Lavpas: } H_{LP} &= \frac{\omega_0^2}{\omega_0^2 + 2d\omega_0 s + s^2} & f_d &= \sqrt{1-d^2} f_0 & A_0 &= 20 \log_{10}(Q) \end{aligned}$$

Højpasfiltret dæmper signalet med 12 dB/oktav for frekvenser under grænsefrekvensen f_0 , og lavpasfiltret dæmper med -12 dB/oktav for frekvenser over grænsefrekvensen. Båndpasfiltret afskærer med 6 dB/oktav under og -6 dB/oktav over filtrets grænsefrekvens.

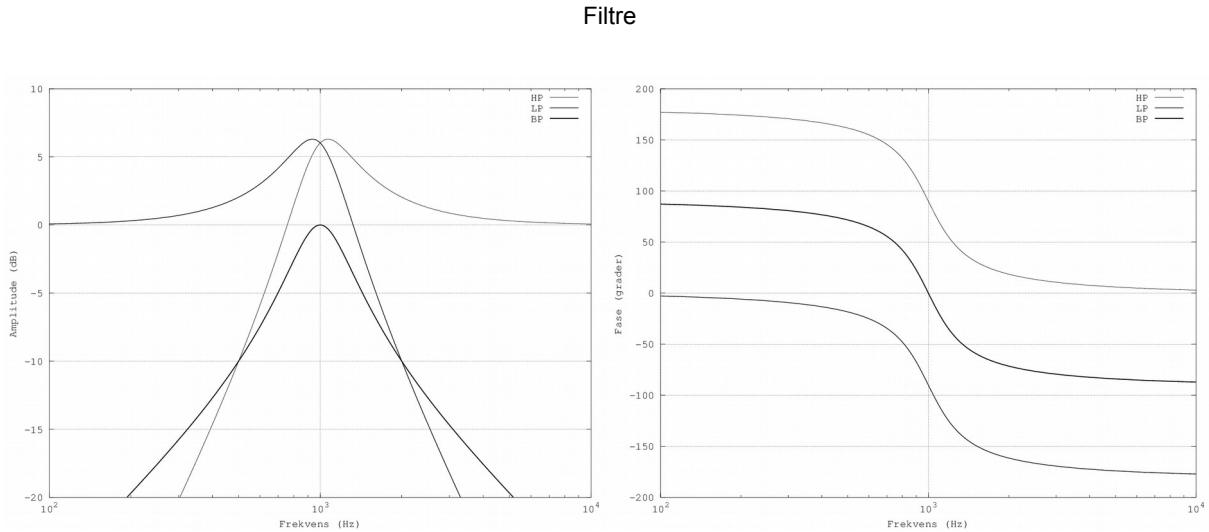
Den karakteristiske frekvens f_0 angiver filtrets asymptotiske grænsefrekvens, altså den frekvens hvor asymptoterne for s gående mod nul og for s gående mod uendelig krydser hinanden. Det er den udæmpede resonansfrekvens f_0 . Frekvensen som filtret svinger på ved en påvirkning kaldes for den dæmpede resonansfrekvens f_d . Niveauet ved den udæmpede resonansfrekvens er givet ved A_0 , men amplituden er højest ved den dæmpede resonansfrekvens.

Niveauet i området omkring grænsefrekvensen styres af en ny parameter, der kaldes for dæmpningen d eller ζ , eller tilsvarende for filtrets godhed Q . Herunder vises responsen for situationen hvor $d = Q = 1/\sqrt{2} = 0,71$, som almindeligvis vil blive betegnet som et maksimalt fladt filter (maximum flat filter), eller et filter af Butterworth typen, hvor niveauet er -3 dB ved grænsefrekvensen.



Figur 184 – Anden-ordens filtre med responsen for højpas (HP, tynd streg), lavpas (LP, mellem streg) og båndpas (BP, tyk streg). Alle filtre er indstillet til dæmpningen $d = 0,707$ (godheden $Q = 0,707$) og med en grænsefrekvens på 1 kHz.

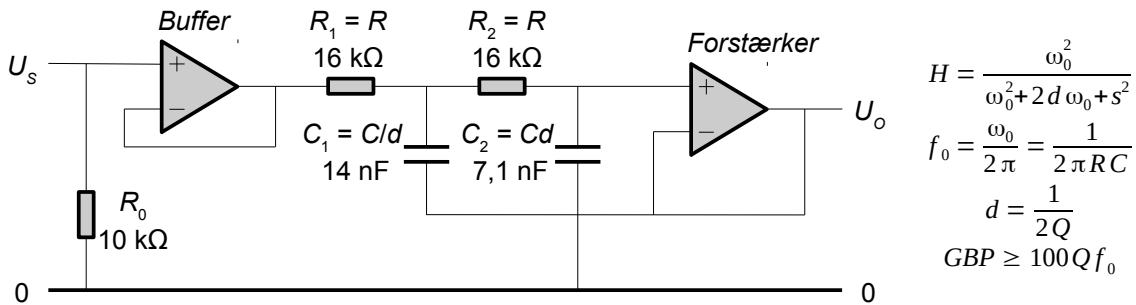
Responsen vises herunder for $d = 0,25$ ($Q = 2$). Niveauet er 6 dB ved $f_0 = 1$ kHz svarende til en forstærkning af signalet på to gange. Toppunktet ligger ved $0,97 f_0$ for lavpasfiltret og ved $1,03 f_0$ for højpasfiltret, og ved stigende godhed (aftagende dæmpning) rykker de to frekvenser hen imod f_0 . Fasen forløber stejlere i overgangsområdet end ved det første eksempel, så filtret holder fasen nærmere den asymptotiske værdi ved aftagende dæmpningen (stigende godhed).



Figur 185 – Anden-ordens filtre med dæmpningen $d = 0,25$ svarende til godheden $Q = 2$ og med en grænsefrekvens på 1 kHz. Alle tre filtre hæver niveauet ved grænsefrekvensen med en værdi, der er givet ved godheden (her to gange).

4.2.1 Ikke-inverterende lavpasfilter

Lavpasfiltret herunder er især kendtegnet ved en DC forstærkning på én og et minimalt forbrug af komponenter. Et typisk design benytter ens modstande og kondensatorerne indstilles da efter den ønskede værdi af godheden Q (dæmpningen d). Med to ens kondensatorer er $d = 1$ ($Q = 0,5$) og filtret opfører sig som om det var to første-ordens led i kaskade. Et opte mødt design benytter en dæmpning på 0,71 for at opnå den såkaldte Butterworth karakteristik, der er vist herunder, og hvor amplituden er jævn med -3 dB ved grænsefrekvensen.



Figur 186 – Et anden-orden lavpasfilter med en buffer så kredsløbet ikke påvirkes af impedansen fra kredsløbet til venstre. Grænsefrekvensen er 1 kHz og godheden er 0,71 med de viste værdier.

Filtret kan selvfølgelig bruge individuelle værdier af både modstande og kondensatorer, men det er stadig forholdet mellem kondensatorerne, der indstiller godheden, for den er givet ved forholdet mellem den geometriske og aritmetiske middelværdi af modstandene.

$$f_0 = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}} \quad \text{og} \quad Q = \frac{\sqrt{R_1 R_2}}{R_1 + R_2} \sqrt{\frac{C_1}{C_2}}$$

Det kan eventuelt udnyttes så frekvensen kan ændres ved R_1 (eller R_2) med lille påvirkning af godheden. Ændres R_1 fra $R/4$ til $4R$ (16 gange), så ændres grænsefrekvensen fra $f_0/2$ til $2f_0$ (4 gange), mens godheden kun ændres over området fra 0 til -20% .

Den viste buffer i indgangen har samme funktion som ved de tidligere viste filtre; at sikre et design med den forventede grænsefrekvens og godhed. Bufferen er ikke nødvendig, hvis det foranstillede kredsløb har en tilstrækkelig lav udgangsmodstand, men det er bydende nødvendigt med en DC forbindelse til nul, eller til det referenceniveau som opstillingen skal fungere ved. Der gælder de

samme bemærkninger til komponenterne som før, så modstandene bør være mindst $1\text{ k}\Omega$ for at holde belastningen af forstærkeren nede, og kondensatorer under 500 pF bør ikke anvendes hvis præcisionen er vigtig. Det betyder at filtrets grænsefrekvens højest bør være 150 kHz , med mindre man kan nøjes med en ringere præcision. Ved stor værdi af godheden bør forstærkeren have et ret betydeligt overskud af forstærkning for at kunne løfte opgaven.

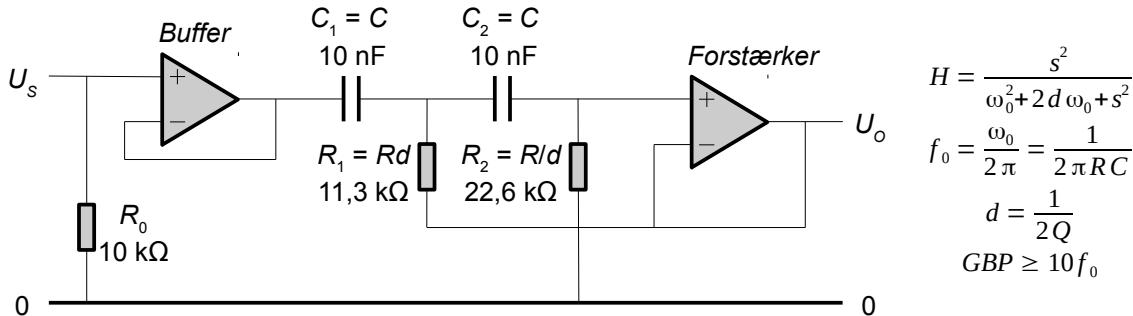
Valg af komponentværdier bør som udgangspunkt ligge i området fra $1\text{ k}\Omega$ til $100\text{ k}\Omega$ både for modstandene og for impedansen af kondensatorerne. Det sikrer at operationsforstærkerens udgang kan drive strømmen ind i R_1 og at indgangsmodstanden ikke belaster C_2 for meget.

Operationsforstærkeren opfattes som en ideel komponent ved de analytiske udtryk, så for at opnå den ønskede karakteristik skal dens båndbredde være så stor at der kan ses bort fra den. Det vil sige at filtrets grænsefrekvens bør indstilles til højest 1% af operationsforstærkerens GBP , og dertil kommer at der nær ved grænsefrekvensen desuden skal forstærkes med godheden Q . Som en tommelfingerregel skal forstærkerens båndbredde derfor vælges til mindst 100 gange med den ønskede godhed gange den valgte grænsefrekvens.

4.2.2 Ikke-inverterende højpasfilter

Filtret kan ændres til højpas ved at ombytte modstande og kondensatorer. Der gælder samme krav til komponentværdierne som ved lavpasfiltret.

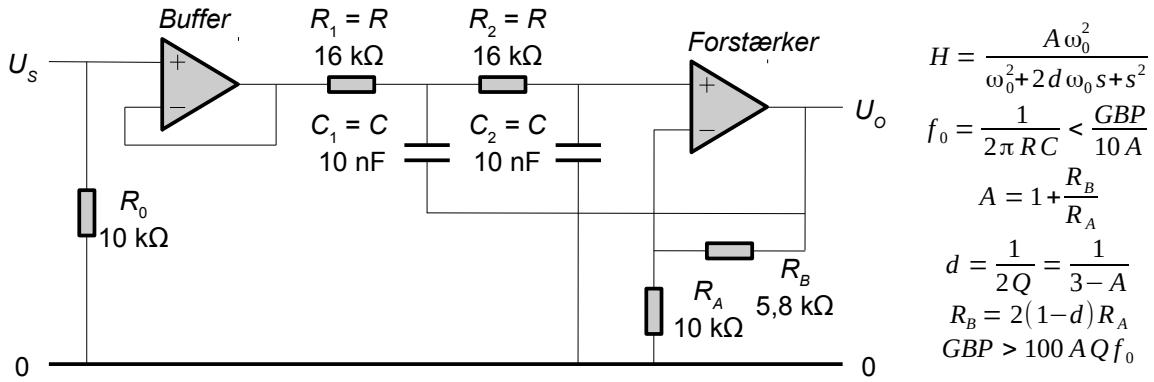
Et højpasfilter lader et signal passere når frekvensen er over grænsefrekvensen, og lavere frekvenser dæmpes progressivt med faldende frekvens, og DC blokeres. Ved høje frekvenser vil operationsforstærkeren virke som et første-ordens lavpasfilter på grund af den endelige båndbredde udtrykt ved operationsforstærkerens GBP specifikation. Det betyder i praksis at overføringsfunktionen for et højpasfilter bliver et asymmetrisk båndpasfilter.



Figur 187 – Et anden-orden højpasfilter med en buffer så kredsløbet ikke påvirkes af impedansen fra kredsløbet til venstre. Grænsefrekvensen er 1 kHz og godheden er $0,71$ med de viste værdier.

4.2.3 Ikke-inverterende lavpasfilter med forstærkning

Ved at gøre forstærkningen i operationsforstærkeren variabel kan godheden indstilles separat med to modstande uden at påvirke grænsefrekvensen, og det tillader brug af parvist ens komponenter. Det betyder dog også at signalet bliver forstærket ved passage gennem filtret. De tidligere noter om modstande og kondensatorer samt forstærkerens *GBP* specifikation er stadig gyldige.

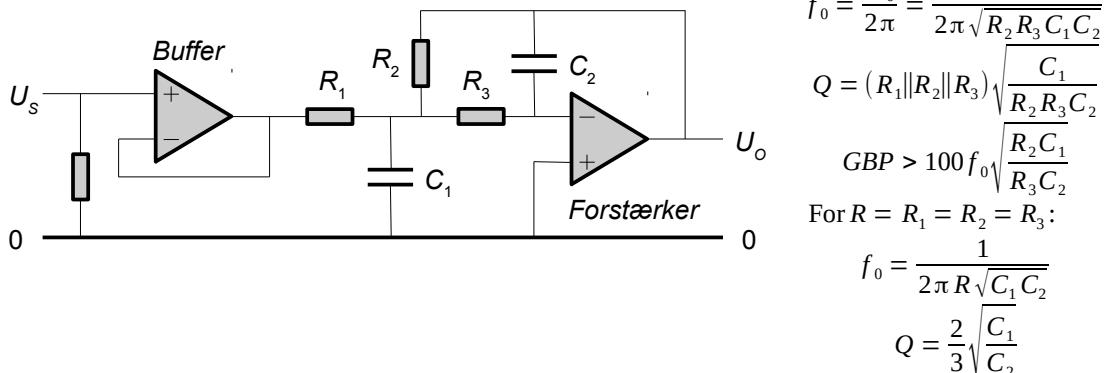


Figur 188 – Et anden-orden lavpasfilter med en buffer i indgangen og en variabel forstærkning A for tilpasning af dæmpningen. Grænsefrekvensen er 1 kHz og godheden er 0,71 med de viste værdier.

Operationsforstærkeren indgår som et aktivt element i filtrets funktion, og dens båndbredde vil optræde som et første-ordens lavpasfilter i kaskade med det ønskede anden-ordens filter. Den resulterende overføringsfunktion bliver af tredje orden, så det er kun ved lave frekvenser at det kan forsvarer at kalde det for et anden-ordens filter.

4.2.4 Inverterende lavpasfilter med forstærkning

Filtret inverterer signalet i pasbåndet og den opnåede forstærkning i pasbåndet indstilles ved to modstande. Den positive indgang ligger på nul, så der er ikke noget signal af betydning mellem forstærkerens indgange og nul (intet common-mode signal), og det bør holde den harmoniske forvrængning nede.

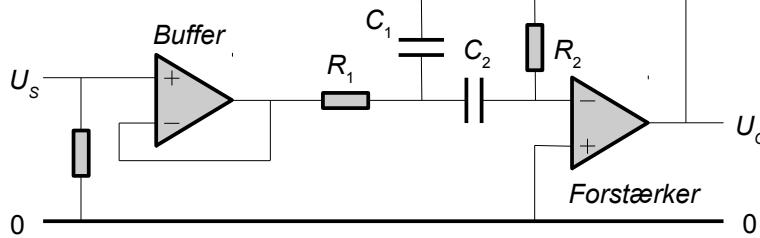


Figur 189 – Et lavpasfilter bygget over en inverterende forstærker.

Filtret er ganske populært trods det lidt mere komplicerede design. Filtret er ret ufølsomt for højfrekvente støjsignaler, idet det indkommende signal dæmpes af R_1 og C_1 inden det ledes til forstærkeren.

4.2.5 Inverterende båndpasfilter med forstærkning

Et båndpasfilter kan opbygges som vist herunder, hvor fire komponenter benyttes til at indstille resonansfrekvensen f_0 , forstærkning ved resonans A_0 og godheden Q .



$$H = -A_0 \frac{2d\omega_0 s}{\omega_0^2 + 2d\omega_0 s + s^2}$$

$$A_0 = Q^2 \left(1 + \frac{C_2}{C_1}\right)$$

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}}$$

$$Q = \sqrt{\frac{R_2}{R_1}} \sqrt{\frac{C_1 C_2}{C_1 + C_2}}$$

$$GBP \geq 20Q^2 f_0$$

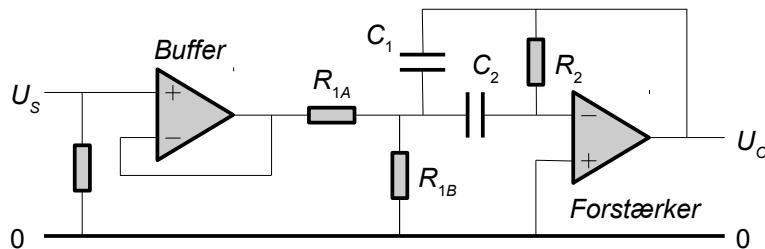
$$\text{For: } C = C_1 = C_2:$$

$$Q \approx \frac{1}{2} \sqrt{\frac{R_2}{R_1}}$$

$$f_0 = \frac{1}{4\pi Q R_1 C}$$

Figur 190 – Et båndpasfilter bygget over en inverterende forstærker.

På grund af at de tre vigtige parametrene f_0 , Q og A_0 er afhængige af de samme fire komponenter kan det være vanskeligt at finde ud af hvordan filtret bedst designes. Udtrykkene er derfor ofte simplificeret ved at antage ens værdier af kondensatorerne; det fungerer udmærket ved en moderat værdi af godheden. For en godhed på én skal forholdet mellem modstandene være fire gange, og ved en godhed på 10 skal R_2 være 400 gange R_1 , hvilket i det fleste tilfælde angiver en øvre værdi af godheden, idet modstandsværdierne ellers let bliver for ekstreme; modstande bør som et udgangspunkt være at finde i området fra 1 kΩ til 1 MΩ. Forstærkningen ved resonans er $A_0 = 2Q^2$ hvis det antages at de to kondensatorer er ens, så den benyttede forstærker skal have en GBP specifikation, der giver tilstrækkelig forstærkning ved den ønskede værdi af resonansen. Almindeligvis anbefales det at benytte $GBP > 20 \cdot Q^2 \cdot f_0$, for at få formlerne til at passe nogenlunde.



$$A_0 = \frac{R_{1A}}{R_{1A} + R_{1B}} Q^2 \left(1 + \frac{C_2}{C_1}\right)$$

$$R_1 = \frac{R_{1A} R_{1B}}{R_{1A} + R_{1B}}$$

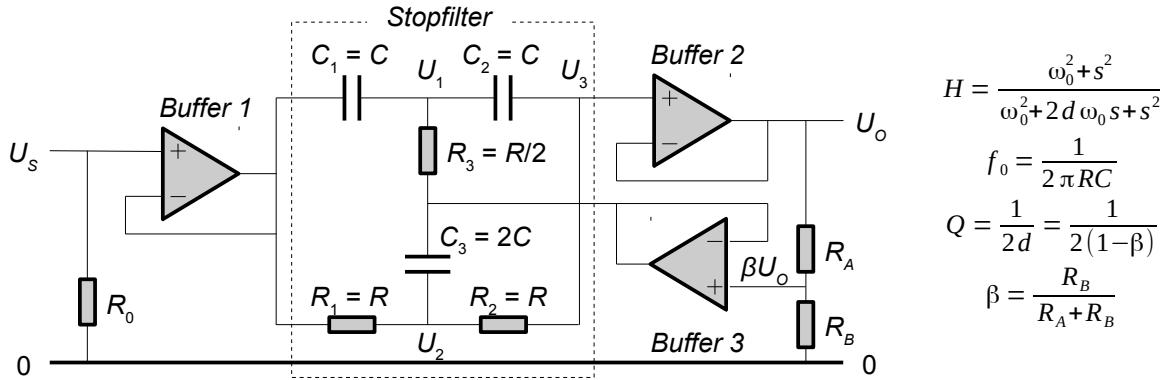
Figur 191 – Et båndpasfilter bygget over en inverterende forstærker. Spændingsdelingen med de to modstande i indgangen reducerer den samlede forstærkning og resonansfrekvensen indstilles nu med de to modstande i parallel. Kravet til forstærkerens GBP specifikation er ikke ændret.

Den store værdi af forstærkningen ved resonans kan mindskes ved at dæmpe indgangssignalet. Det foretages helt enkelt ved en opdeling af R_1 i to modstande, hvor parallelværdien benyttes som modstanden R_1 og forstærkningen reduceres ved spændingsdelingen mellem dem⁹⁹. Ved stor spændingsdeling ($R_{1B} \ll R_{1A}$) vil forstærkningen ved resonans være nærmest uafhængig af R_{1B} , så modstanden R_{1A} kan designes alene for at opnå en given forstærkning ved resonans. En vigtig detalje er dog, at selv om forstærkningen A_0 reduceres, så vil forstærkeren stadig skulle præstere det samme som før, så der er ikke opnået noget overfor den krævede GBP specifikation.

⁹⁹ Det kan indses ved at danne Thevenin-ækvivalentet for indgangssignalet U_S og de to modstande R_{1A} og R_{1B} .

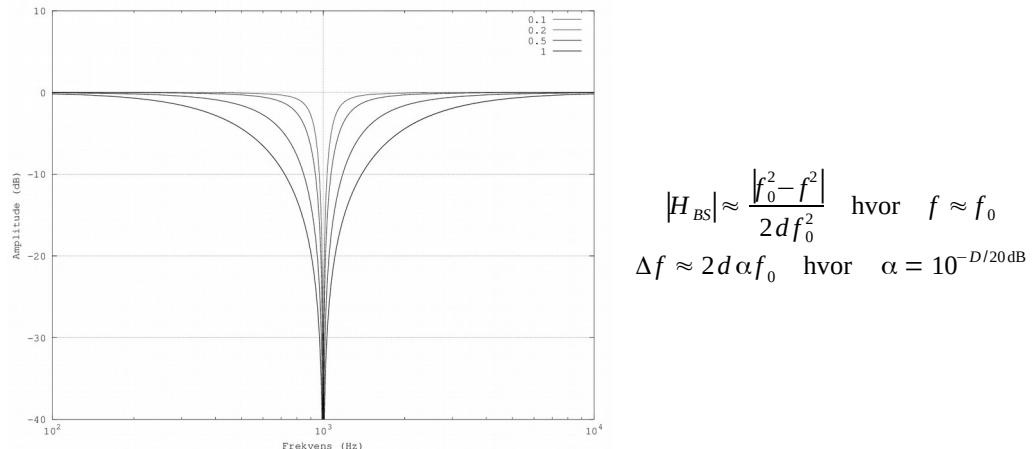
4.2.6 Stopfilter

Det er muligt at undertrykke et signal ved en given frekvens, hvilket kan være attraktivt, hvis der er en genevirning fra 50 Hz brum i et måleapparat, eller for måling af den harmoniske forvrængning fra et kredsløb ved at undertrykke første harmoniske og kun lade forvrængningen passere. Det kaldes for et stopfilter (*notch filter*), og den mest populære form ses herunder. Filtret dannes af to T-formede led og kaldes derfor et *Twin-T filter* i engelsk litteratur.



Figur 192 – Et stopfilter med tilbagekobling gennem potentiometeret P , hvor brøkdelen β af udgangssignalet føres tilbage til fællespunktet af stopfiltret for styring af stopfiltrets bredde.

Filtret kan i teorien undertrykke en enkelt frekvens komplet, men det er mere interessant at det kan dæmpe et frekvensbånd Δf , der kan indstilles ved tilbagekobling af en brøkdel af udgangssignalet, givet ved modstandene R_A og R_B , og som illustreres herunder med dæmpningen d som parameter.



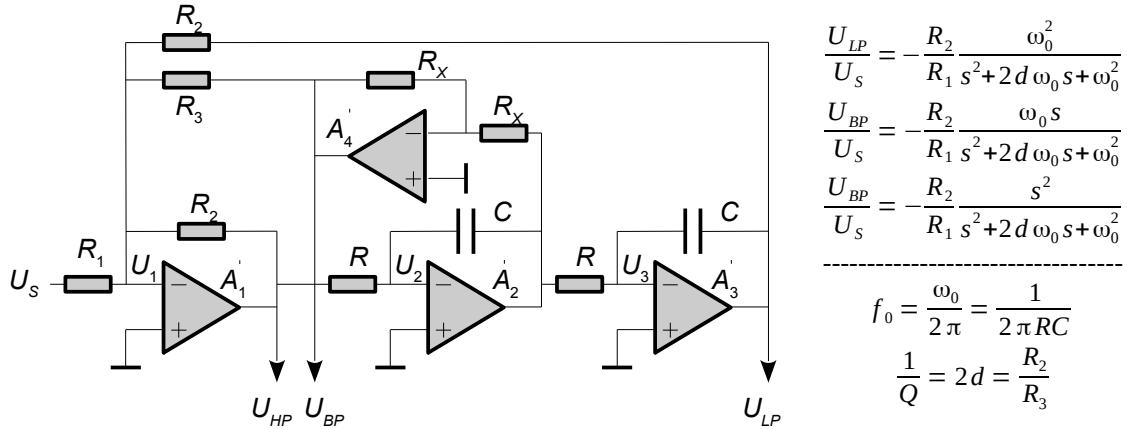
Figur 193 – Dæmpningen for et stopfilter ved 1000 Hz som funktion af dæmpningen fra $d = 0,1$ (smalt bånd) til $d = 1$ (bredt bånd med $\beta = 0$).

Det er ikke tilrådeligt at spare Buffer 3 væk for filtret skal fødes fra en meget lav impedans for at teorien holder. Værdien $d = 1$ svarer til at R_3 og C_3 er ført direkte til nul (uden buffer); en variant, der figurerer flittigt på internettet.

Bredden af stopbåndet kan udtrykkes ved parameteren α , der er givet ved den dæmpningen. For dæmpning af et signal ved en given frekvens til $D = -40$ dB eller derunder bliver $\alpha = 0,01$. Med en centerfrekvens på $f_0 = 1$ kHz er værdien af stopbåndet på $\Delta f = 20$ Hz som strækker sig fra 990 Hz til 1010 Hz, hvis dæmpningen er sat til $d = 1$. Bemærk at dæmpningen udenfor stopbåndet i dette tilfælde er betydelig og dækker langt ud til hver side af centerfrekvensen.

4.2.7 State-variable filter

Litteraturen har mange referencer til det såkaldte *state-variable filter*, der ikke har et dansk navn. Filtret benytter to gange integration for at realiserer højpas, båndpas og lavpas, hvor parametrene for resonans f_0 og godhed Q kan indstilles uafhængigt af hinanden.



$$\frac{U_{LP}}{U_S} = -\frac{R_2}{R_1} \frac{\omega_0^2}{s^2 + 2d\omega_0 s + \omega_0^2}$$

$$\frac{U_{BP}}{U_S} = -\frac{R_2}{R_1} \frac{\omega_0 s}{s^2 + 2d\omega_0 s + \omega_0^2}$$

$$\frac{U_{BP}}{U_S} = -\frac{R_2}{R_1} \frac{s^2}{s^2 + 2d\omega_0 s + \omega_0^2}$$

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi RC}$$

$$\frac{1}{Q} = 2d = \frac{R_2}{R_3}$$

Figur 194 – Et state-variable filter tilbyder tre udgange for lavpas, båndpas og højpas. Alle filtre har samme grænsefrekvens og godhed, der kan indstilles uden gensidig påvirkning.

Et anden-ordens stopfilter kan dannes ved at addere U_{LP} og U_{HP} . Filtret ses ofte realiseret med tre operationsforstærkere, men den viste form har de positive indgange på nul så der undgås problemer med common-mode, der opnås et enklere udtryk for indstilling af filtrets godhed, og udgangen for båndpas får den samme polaritet som udgangene for højpas og lavpas. Bemærk at forstærkningen for båndpas er funktion af den indstillede godhed, så en summering af alle udgangene giver ikke en identitet.

En analyse af filtret starter med at opskrive knudepunktsgligningerne U_1 , U_2 og U_3 for de tre inverterende indgange på operationsforstærkerne A_1 , A_2 og A_3 . Bemærk at A_4 kun har til funktion at invertere signalet. Det udnyttes at de inverterende indgange er virtuel nul.

$$\begin{aligned} \frac{U_1 - U_S}{R_1} + \frac{U_1 - U_{HP}}{R_2} + \frac{U_1 - (-U_{BP})}{R_3} + \frac{U_1 - U_{LP}}{R_2} &= 0 & 0 &= \frac{U_S}{R_1} + \frac{U_{HP}}{R_2} - \frac{U_{BP}}{R_3} + \frac{U_{LP}}{R_2} \\ \frac{U_2 - U_{HP}}{R} + \frac{U_2 - U_{BP}}{1/sC} &= 0 & \Rightarrow & 0 = \frac{U_{HP}}{R} + sCU_{BP} \\ \frac{U_3 - U_{BP}}{R} + \frac{U_3 - U_{LP}}{1/sC} &= 0 & 0 &= \frac{U_{BP}}{R} + sCU_{LP} \end{aligned}$$

Lavpasfunktionen isoleres ved at isolere U_{BP} fra sidste ligning og U_{HP} fra midterste ligning og udnytte at resultaterne kun afhænger af U_{LP} . Båndpas og højpas udledes på lignende vis.

$$\begin{aligned} U_{BP} &= -sRCU_{LP} \quad \text{og} \quad U_{HP} = -sRCU_{BP} = (sRC)^2 U_{LP} \\ U_S &= -\frac{R_1}{R_2} (sRC)^2 U_{LP} - \frac{R_1}{R_3} sRCU_{LP} - \frac{R_1}{R_2} U_{LP} \\ \frac{U_{LP}}{U_S} &= -\frac{1}{\frac{R_1}{R_2}(sRC)^2 + \frac{R_1}{R_3}sRC + \frac{R_1}{R_2}} = -\frac{R_2}{R_1} \frac{\frac{1}{(RC)^2}}{s^2 + \frac{R_2}{R_3}\frac{s}{RC} + \frac{1}{(RC)^2}} \\ \frac{U_{LP}}{U_S} &= -\frac{R_2}{R_1} \frac{\omega_0^2}{s^2 + 2d\omega_0 s + \omega_0^2} \quad \text{hvor} \quad \omega_0 = \frac{1}{RC} \\ &\quad 2d = \frac{1}{Q} = \frac{R_2}{R_3} \end{aligned}$$

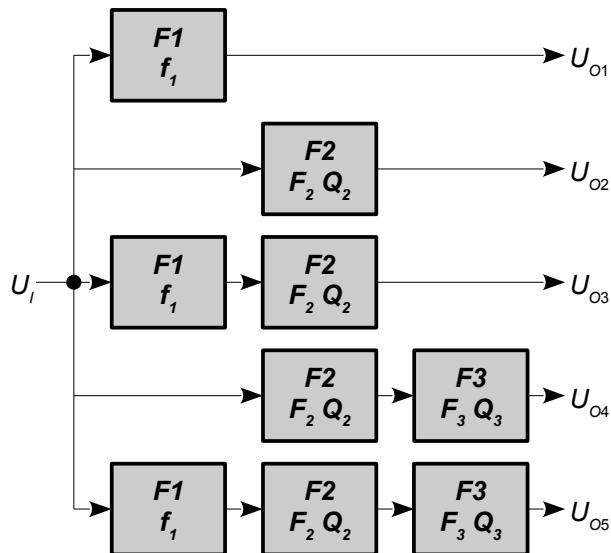
4.3 Filtre af højere orden

Som nævnt i kapitlet om analyse kan ethvert lav-, bånd og højpasfilter af højere orden dannes ved at sammensætte et antal blokke af første og anden ordens filtre, så skarpe filtre kan dannes ud fra en kaskade af de allerede behandlede filtre. Det er vist herunder ved første-ordens blokken $F1$ og anden-ordens blokkene $F2$ og $F3$ for filtre af stigende orden. Overføringsfunktionen bestemmes ved valg af de angivne parametre for hver blok, som det vises i de følgende afsnit for Butterworth og Chebychev filtertyperne.

Et femte-ordens lavpasfilter består af blokkene, der leder til U_{o5} i illustrationen og herunder vises overføringsfunktionen. Filret består af en første-ordens blok $F1$ med grænsefrekvensen $f_1 = \omega_1/2\pi$ som eneste parameter og dertil to anden-ordens blokke $F2$ og $F3$, med parametrene $f_2 = \omega_2/2\pi$ og Q_2 for den ene blok, respektivt $f_3 = \omega_3/2\pi$ og Q_3 for den anden blok. Rækkefølgen af blokkene er valgfri og kan tilpasses aktuelle behov.

$$H_{LP5} = \frac{\omega_1}{\omega_1 + s} \frac{\omega_2^2}{\omega_2^2 + 2d_2\omega_2 + s^2} \frac{\omega_3^2}{\omega_3^2 + 2d_3\omega_3 + s^2} \quad \text{hvor} \quad d_n = \frac{1}{2Q_n}$$

Overføringsfunktionen kan omformes til højpas, som det vises lidt senere, og dertil kommer en del valgmuligheder. For båndpasfiltre kan responsen være symmetrisk eller ej, og der kan inkluderes nulpunkter ved at modificere tællerlen. Emnet kan let vokse ud over alle grænser, så jeg vil nøjes med at gennemgå nogle ganske få muligheder.



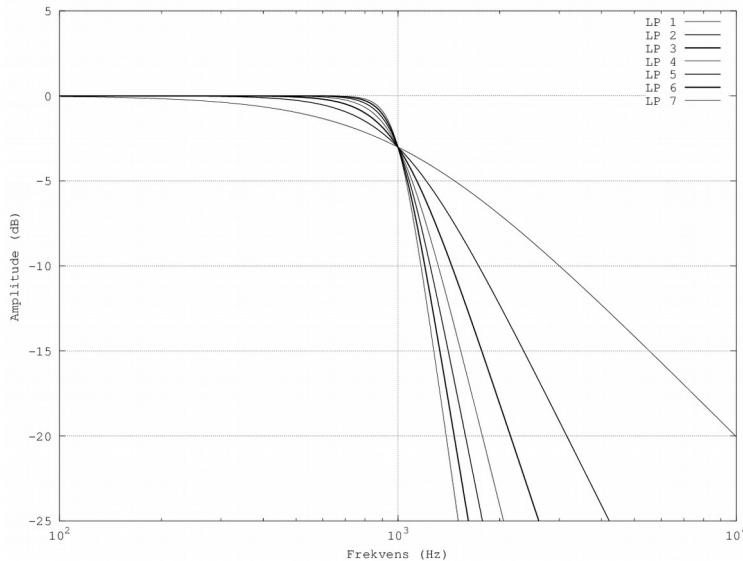
Figur 195 – Filtre kan opbygges af en vekslen mellem filterblokke af første eller anden orden.
Første-ordens blokken kun optræder ved ulige filterorden og har kun grænsefrekvensen f_1 som parameter. Anden-ordens blokken optræder det antal gange, der er påkrævet for at nå op på den ønskede orden og har to parametre, grænsefrekvensen f_2 og godheden Q_2 for $F2$ blokken.

Alle filtre af samme orden vil mod høje frekvenser have samme stejlhed i afskæringen; det er kun forløbet i omegnen af grænsefrekvensen, som man har mulighed for at påvirke. Det indses ved at lade s gå imod uendelig i den ovenstående formel, for herved kan nævneren simplificeres til s^5 , og med konstant tæller er afskæringen $-30 \text{ dB}/\text{oktav}$ uanset de indgående parametres værdier.

$$H_{LP5} \xrightarrow[s \rightarrow 0]{} 1 \quad \text{og} \quad H_{LP5} \xrightarrow[s \rightarrow \infty]{} \frac{\omega_1 \omega_2 \omega_3}{s^5} \Rightarrow \omega_0^5 = \omega_1 \omega_2^2 \omega_3^2 \Rightarrow f_0 = \frac{1}{2\pi} \sqrt[5]{\omega_1 \omega_2^2 \omega_3^2}$$

4.3.1 Butterworth filter

Det mest populære af alle filertyper kaldes Butterworth, efter den matematiker, der formulerede den bagvedliggende teori. Amplituden varierer efter følgende enkle relation, hvor f_0 er filtrets –3 dB grænsefrekvens, og n er filtrets orden, der i øvrigt er identisk med antallet af kondensatorer i filtret. Ved grænsefrekvensen er $f = f_0$ og amplituden bliver $A = 0,707$ uanset den valgte filterordenen, så karakteristikken går altid igennem –3 dB.



$$A = \frac{1}{\sqrt{1 + \left(\frac{f}{f_0}\right)^{2n}}}$$

$$n \geq \frac{\ln\left(\frac{1}{A^2} - 1\right)}{2 \ln\left(\frac{f}{f_0}\right)}$$

Figur 196 – Butterworth filtrets amplitude går igennem –3 dB ved grænsefrekvensen f_0 uanset filtrets orden (venstre). Amplituden A er defineret ved filtrets orden n og den nødvendige orden for en given dæmpning ved frekvensen f kan beregnes (højre).

For filtre med en orden indtil $n = 7$ vises sammenhængen herunder¹⁰⁰ mellem grænsefrekvensen af det enkelte led (f_1, f_2, \dots) og godheden af anden-ordens sektionen (Q_2, Q_3, \dots). Alle sektioner har samme grænsefrekvens ved Butterworth filtre så kun godheden skal justeres.

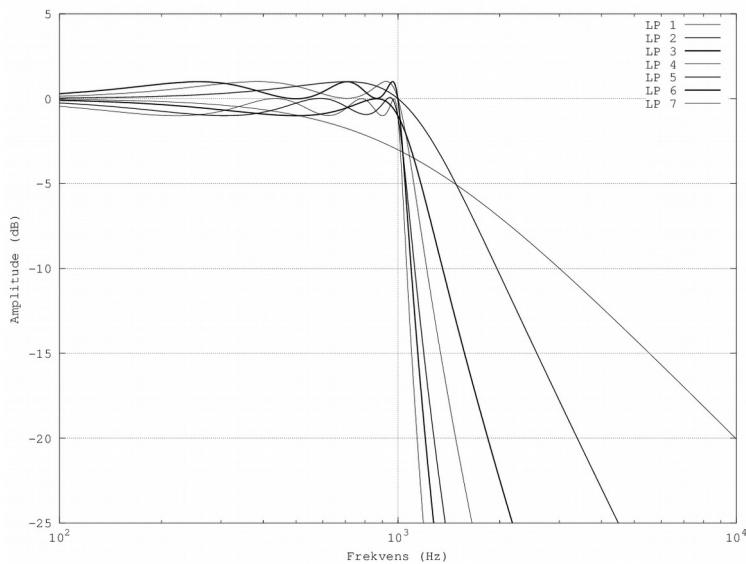
n	$F1$		$F2$ Anden orden		$F3$ Anden orden		$F4$ anden orden		A_2
	f_1/f_0	f_2/f_0	Q_2	f_3/f_0	Q_3	f_4/f_0	Q_4		
1	1,000								-7,0 dB
2		1,000	0,707						-12,3 dB
3	1,000	1,000	1,000						-18,1 dB
4		1,000	0,541	1,000	1,306				-24,1 dB
5	1,000	1,000	0,618	1,000	1,620				-30,1 dB
6		1,000	0,518	1,000	0,707	1,000	1,932		-36,1 dB
7	1,000	1,000	0,555	1,000	0,802	1,000	2,247		-42,1 dB

Sidste kolonne (A_2) angiver filtrets dæmpning en oktav over grænsefrekvensen, og det ses at den skrø asymptote (ordenen gange –6 dB/oktav) stort set er korrekt allerede ved første-ordens filtret; og det er netop baggrunden for Bode plottets succes.

¹⁰⁰ Sammensat af forskellige kilder på nettet og korrigert for diverse skrivefejl. Illustrationerne benytter tabellernes talværdier.

4.3.2 Chebychev filter

Afskæringen umiddelbart over grænsefrekvensen kan gøres mere stejl mod at acceptere en ujævn amplitude i pasbåndet (*pass band ripple*), men mod stigende frekvens vil asymptotens hældning være som filtrets orden angiver. Sammenlignet med Butterworth betyder det at filtret er bedre til at dæmpe signaler umiddelbart over grænsefrekvensen. Eksempelvis vil 30 dB dæmpning af signaler en oktav over grænsefrekvensen kræve et femte-ordens filter af typen Butterworth, mens opgaven klares af et fjerde-ordens Chebychev filter; prisen er en svag oscillation i pasbåndet.



Figur 197 – Chebychev filtrets amplitude vil for lige orden svinge mellem 0 dB og 1 dB med 0 dB ved grænsefrekvensen. For ulige orden vil amplituden svinge fra 0 dB og ned til –1 dB med –1 dB ved grænsefrekvensen. Undtagelsen er første-ordens filtret, der opfører sig som Butterworth.

Ripplens amplitude er en ekstra parameter for Chebychev filtret, og jeg har valgt at tabellen kun skal indeholde værdier for 1 dB spids-spids værdi af oscillationen; det kaldes "1 dB Chebychev". Bemærk at filtret med lige orden oscillerer over den vandrette 0 dB asymptote, og at filter med ulige orden oscillerer under asymptoten. Det eneste filter, der ikke holder sig indenfor det specificerede område på 1 dB spids-spids er første-ordens filtret; for det er identisk med Butterworth filtret.

n	F1		F2 Anden orden		F3 Anden orden		F4 anden orden		A₂	ΔA_{B-C}
	f₁/f₀	f₂/f₀	Q₂	f₃/f₀	Q₃	f₄/f₀	Q₄	A₂		
1	1,000							-7,0 dB	0,0 dB	
2		1,050	0,957					-11,4 dB	-0,9 dB	
3	0,494	0,997	2,017					-22,5 dB	4,4 dB	
4		0,529	0,785	0,993	3,559			-33,9 dB	9,8 dB	
5	0,289	0,655	1,399	0,991	5,554			-45,3 dB	15,2 dB	
6		0,353	0,761	0,747	2,198	0,995	8,001	-56,7 dB	20,6 dB	
7	0,205	0,480	1,300	0,808	3,156	0,996	10,901	-68,2 dB	26,1 dB	

Kolonne A_2 angiver filtrets dæmpning en oktav over grænsefrekvensen og ΔA_{B-C} angiver hvor meget Chebychev filtret dæmper mere end Butterworth filtret ved denne frekvens.

4.3.3 Højpasfilter

Overføringsfunktionen kan transformeres til højpas ved at substituere s/ω_1 med ω_1/s for såvel første-ordens blokken og for anden-ordens blokkene, som det skal vises herunder. Det ses at nævnerens polynomium ikke ændres ved transformationen, så tabelmaterialet kan også benyttes for design af højpasfiltre.

For første-ordens filtret vises omregningen lavpas til højpas. Begge filtre har grænsefrekvensen $f_1 = \omega_1/2\pi$. Leddet s/ω_1 erstattes med ω_1/s og resulterer i udtrykket for et højpasfilter.

$$H_{LP1} = \frac{\omega_1}{\omega_1 + s} = \frac{1}{1 + \frac{s}{\omega_1}} \Rightarrow H_{HP1} = \frac{1}{1 + \frac{\omega_1}{s}} = \frac{s}{\omega_1 + s}$$

For anden-ordens filtret gennemføres den tilsvarende beregning.

$$H_{LP2} = \frac{\omega_2^2}{\omega_2^2 + \frac{1}{Q} \omega_2 s + s^2} = \frac{1}{1 + \frac{1}{Q} \frac{s}{\omega_2} + \left(\frac{s}{\omega_2}\right)^2} \Rightarrow H_{HP2} = \frac{1}{1 + \frac{1}{Q} \frac{\omega_2}{s} + \left(\frac{\omega_2}{s}\right)^2} = \frac{s^2}{\omega_2^2 + \frac{1}{Q} \omega_2 s + s^2}$$

4.3.4 Filterdesign

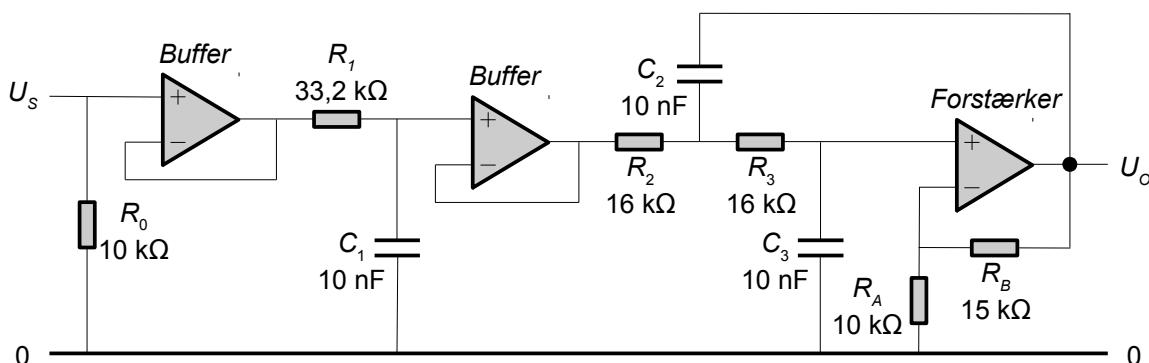
Ved design af filtre er grænsefrekvensen almindeligvis givet ud fra den påtænkte anvendelse, men den nødvendige filterorden og filtertype kan være svære at fastlægge, og dertil kommer hvilken af de viste opbygninger, der bør foretrækkes. I det følgende vil jeg gennemgå designet af et filter, der skal dæmpe signaler over 1 kHz med mindst 20 dB i én oktav afstand over grænsefrekvensen. Det er umiddelbart let at skalere filtret til en anden frekvens ved blot at ændre kondensatorerne.

4.3.5 Valg af filtertype og orden

Som udgangspunkt vil jeg anbefale Butterworth med mindre den støjtere afskæring for Chebychev er vigtig, fx for at reducere ordnen. Det begrundes i den lavere værdi af godheden i anden-ordens sektionerne i Butterworth, der gør det mere sandsynligt at designet kan gennemføre med succes, og dertil at fasedrejningen i området nær ved grænsefrekvensen er mindre stejl hvorved filtrets gruppeløbetid (*group delay*) varierer mindre. Kravet om 20 dB dæmpning ved 2 kHz giver her en filterorden på $n = 4$ for Butterworth og $n = 3$ for Chebychev. Det vælges at designe et Chebychev tredje-ordens filter og fra tabellen aflæses parametrene.

Første-ordens sektion F1: $f_1 = 0,494 \cdot f_g = 494$ Hz

Anden-ordens sektion F2: $f_2 = 0,997 \cdot f_g = 997$ Hz $Q_2 = 2,017$



Figur 198 – Et Chebychev lavpasfilter med 1 kHz grænsefrekvens og 1 dB variation i pasbåndet. Afskæringen er af tredje orden, det vil sige -18 dB/oktav (-60 dB/dekade), men filtret dæmper med hele 22,5 dB ved en oktav over grænsefrekvensen.

Det er uden betydning for overføringsfunktionen om første-ordens filtret placeres først eller sidst i kaskaden. Ved filtre af højere orden kan der dog være et problem med at placere anden-ordens sektioner med høj Q-værdi forrest da det kan give overstyring selv ved en moderat amplitude.

Det ikke-inverterende filter er det simpleste og varianten med forstærkning giver mulighed for at benytte ens værdier af modstandene og ens værdi af kondensatorerne, hvilket kan være en fordel når komponenterne købes, for modstande kan ofte bestilles i E96 serien, mens kondensatorer meget ofte er begrænset til E12 serien. Der findes ikke formler, som direkte beregner alle komponenternes værdier; formlerne angiver som regel kun en relation til produktet eller forholdet mellem to komponenter, så det er nødvendigt at vælge den ene komponents værdi for at beregne den anden.

4.3.6 Valg af passive komponenter

Designet gennemføres oftest ved at vælge værdien af kondensatorerne til en "rimelig" værdi og så finde de krævede modstandsværdier for at ramme grænsefrekvensen. Som en probe for hvad der er rimeligt kan man sigte efter en impedans i omegnen af $1 \text{ k}\Omega$ ved grænsefrekvensen og derefter vælge en kondensator så impedansen ligger i intervallet fra $1 \text{ k}\Omega$ til $100 \text{ k}\Omega$ hvilket bør sikre at operationsforstærkeren arbejder optimalt. Det giver her en kondensator på $15,9 \text{ nF}$ og det vælges at benytte 10 nF i alle tre sektioner (svarende til en impedans på cirka $16 \text{ k}\Omega$).

$$\begin{aligned} X_C &= \frac{1}{2\pi f_0 C} \Rightarrow C = \frac{1}{2\pi f_0 X_C} \quad \text{hvor } 1 \text{ k}\Omega \leq X_C \leq 100 \text{ k}\Omega \\ R_1 &= \frac{1}{2\pi f_1 C} = 32,2 \text{ kohm} \quad \text{for } \begin{array}{l} C_1 = 10 \text{ nF} \\ f_1 = 494 \text{ Hz} \end{array} \\ R_2 = R_3 &= \frac{1}{2\pi f_2 C} = 16,0 \text{ kohm} \quad \text{for } \begin{array}{l} C_2 = C_3 = 10 \text{ nF} \\ f_2 = 997 \text{ Hz} \end{array} \\ R_B &= 2(1-d)R_A \\ d &= \frac{1}{2Q_1} = 0,248 \quad \Rightarrow \quad R_B = 15,0 \text{ kohm} \quad \text{for } R_A = 10,0 \text{ kohm} \end{aligned}$$

Grænsefrekvensen f_1 giver modstanden for første-ordens sektionen. Grænsefrekvensen f_2 giver de frekvensbestemmende modstande for anden-ordens sektionen. Godheden Q_2 indstilles ved de to modstande R_A og R_B , hvor den ene modstand må vælges for at man kan finde en værdi for den anden. Her vælges en modstandsværdi på $R_A = 10 \text{ k}\Omega$.

4.3.7 Valg af operationsforstærkere

Den valgte operationsforstærker bør have en GBP specifikation på 100 gange grænsefrekvensen gange forstærkningen gange godheden, med mindre man kan være klar til at acceptere en afvigelse fra den beregnede karakteristik. I det aktuelle tilfælde kræves en operationsforstærker med en båndbredde på $\text{GBP} > 300 \text{ kHz}$ og det tillader brug af næsten alt, der kan købes.

$$\text{GBP} > 100 A Q f_0 \Rightarrow \text{GBP} > 303 \text{ kHz} \quad \text{for } \begin{array}{l} A = 1,5 \\ Q = 2,017 \\ f_0 = 1 \text{ kHz} \end{array}$$

Dertil kommer krav til indgangsimpedans og -kapacitet, for ikke at belaste de komponenter, der definerer grænsefrekvensen, og at udgangens skal kunne drive den krævede strøm. Ved høje frekvenser er C_2 nærmest en kortslutning, så strømmen gennem R_2 kan beregnes til signalets amplitude delt med modstandens værdi, og det skal ligge indenfor operationsforstærkerens data. Her vil jeg normalt regne meget konservativt, det vil sige med en sikkerhedsmargin på 10 gange, og for de fleste operationsforstærkere giver det den fornævnte optimale værdi omkring $10 \text{ k}\Omega$.

Det er tidligere nævnt at værdien af kondensatorerne helst skal være over 500 pF så det danner en praktisk nedre grænse for valget. Det kan måske være acceptabelt ned til cirka 50 pF , hvor en præcision af størrelsesordenen 10 % kan forventes.

Filtre

Note 7.

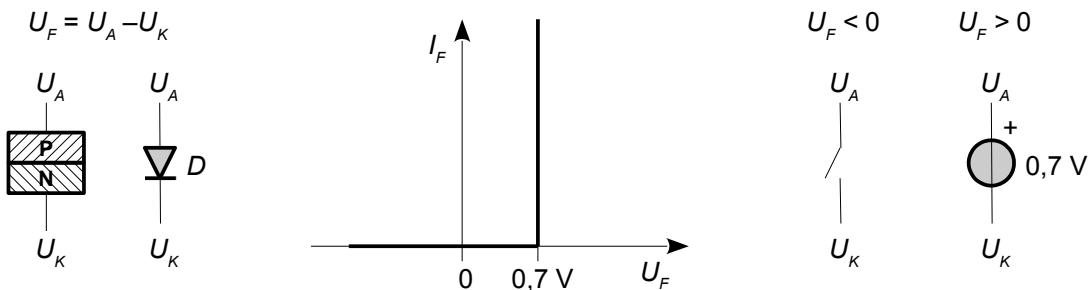
5 HALVLEDER

Ensretteren kan dannes af to ledende materialer der bringes i elektrisk kontakt. Den var i brug længe før elektronrøret blev opfundet for Braun opdagede i 1876 at krystaller af blyulfid kunne ensrette en strøm og Marconi brugte det til modtagelse af trådløs telegrafi fra 1895. Ensretteren var dog mekanisk ustabil så da Fleming opdagede at Edisons glødelampe kunne ensrette vekselstrøm ved at tilføje en elektrode og DeForest udviklede det til en triode i 1906 var elektronrøret skabt. Det dannede fundamentet for udvikling af radio- og TV transmission, lydgengivelse og computere. Lilienfeld patenterer forløberen til MOSFET i 1925, men teknologien var for krævende så gennembruddet måtte vente til 1947 hvor Bardeen, Brattain og Shockley udvikler den bipolære transistor ved Bell Labs. Dens lille størrelse, lave effektforbrug og lange levetid gav elektronrøret konkurrence som fx den transportable transistorradio fra 1954. Den integrerede kreds lanceres af Fairchild i 1960 og MOSFET teknologien introduceres i 1964 af RCA. Den første single-chip mikroprocessor sender Intel på markedet i 1971 og den når hobbyisterne få år senere. Det nye marked accelererer udviklingen af den personlige computer, internettet og mobiltelefonen, men der er stadig brug for at forstærke det analoge signal fra en termokabler og for at drive en højttaler.

Formålet med dette kapitel er at vise hvordan en operationsforstærker kan opbygges ved brug af dioder og transistorer og det udbygges til en effektforstærker i næste kapitel. Fælles for alle typer af halvledere er dioden, som introduceres herunder, derefter beskrives den bipolære transistor BJT fra side 218, felt-effekt transistorerne JFET og MOSFET starter side 227 og den typiske opbygning af en operationsforstærker beskrives fra side 253.

5.1 Dioden

Den grundlæggende egenskab ved dioder og transistorer er PN overgangen der i det væsentlige kun tillader en elektrisk strøm at løbe i én retning. Når strømmen løber er der et spændingsfald over PN overgangen som kun i ringe grad reflekterer strømmens styrke. En populær model er at spændingsfaldet har en konstant værdi på 0,7 V og det passer godt ved en moderat strømstyrke, men som det vil fremgå af det følgende er billedet en del mere nuanceret.



Figur 199 – Simpel model af dioden. Dioden er afbrudt når anodens spænding er lavere end katodens spænding og det behandles analytisk ved en afbrudt kontakt. Dioden repræsenteres derimod af en spændingsforskel på 0,7 V når anodens spænding er højere end katoden og der løber en strøm.

På grund af spændingsfaldet over dioden i lederetningen afsætter den en effekt i dioden og det er en af tabsmekanismene i den populære DC-DC konverter.

Ved en spænding i modsat retning vil dioden spærre. Selv om dioder af silicium er ganske gode isolatorer så vil der løbe en svag strøm af størrelsesordenen $I_F = -1 \text{ nA}$ dog stærkt afhængig af materialets type, fysisk størrelse og temperatur. Det er diodens lækstrøm (*leakage current*) og den afhænger stort set ikke af spændingen over dioden, men den varierer meget med temperaturen.

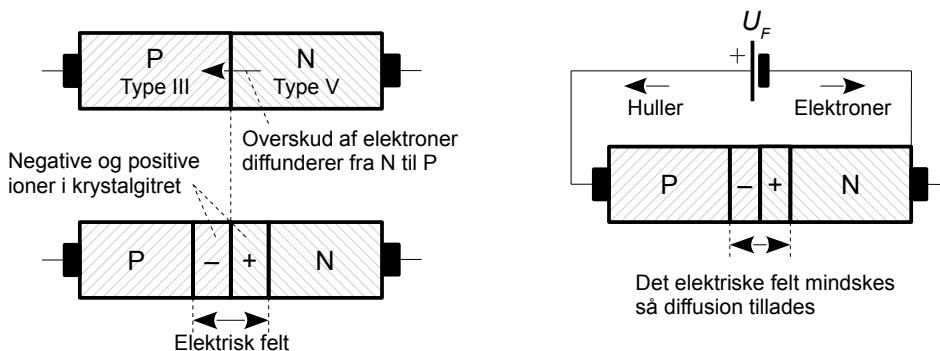
Der er tradition for at regne strømmen for positiv i den retning der er emnet for brugen i en given applikation og det vises ved symbolet med I_F for lederetningen (*forward*) og I_R for spærretretningen (*reverse*). Der gælder tilsvarende at spændingsfaldet normalt angives som positivt i lederetningen som U_F og i spærretretningen som U_R .

5.1.1 Diodens funktion

Silicium er det næstmest hyppigt forekommende grundstof i jordskorpen, men det er bundet til ilt som SiO_2 i form af kvarts og en energikrævende proces er nødvendig for at rense ilten væk. Det var blandt andet grunden til at de første transistorer kort efter 1950 blev fremstillet af germanium der kan renses alene ved zonesmelting. Før det blev selen og kobberoxyd brugt til ensretning af vekselstrøm og blysvulfid til detektion af radiofoni. Dette afsnit giver en kortfattet introduktion til PN overgangens funktion og for en mere uddybende teori henvises til side 271.

Silicium er af type IV der har fire valens-elektroner i den yderste skal. Ved en temperatur under smeltepunktet på 1420°C er atomerne arrangeret i et gitter hvor ethvert atom er forbundet til fire andre atomer. Der er kun få frie elektroner i krystallet ved stuetemperatur så silicium er elektrisk isolerende. Det kan ændres ved at udskifte siliciumatomer med fremmede atomer af type III der kun har tre elektroner i den yderste skal (fx bor) eller type V med fem elektroner (fx kvælstof). Uanset typen af forurening er halvlederen udadtil elektrisk neutral for der er det samme antal af negative elektroner enten i skallerne eller fritsvævende som der er positive ladninger i kernerne.

Ved forurening med et type V atom vil den overskydende elektron blive slået fri af omgivelsernes temperatur og den svæver derefter frit rundt i krystallet og kan lede en elektrisk strøm. Det kaldes for en N halvleder da der er et overskud af frie elektroner. Tilsvarende vil et type III atom mangle en elektron i valensbindingen hvilket kaldes for et "hul". Type III atomet kan snuppe en elektron fra et nabootom af silicium som derved mangler en elektron og hullet er flyttet i valensbåndet blot i den modsatte retning af elektronen. Hullet opfattes som en positiv ladningsbærer der kan deltag i strømtransporten på linje med elektronen og et overskud af flytbare huller giver en P halvleder.



Figur 200 – Diodens N halvleder har et overskud af frie elektroner og P halvlederen har et overskud af flytbare huller. Diffusionen af elektroner danner et elektrisk felt som modvirker diffusionen. En ekstern spændingskilde kan ændre på det elektriske felt og derved styre strømmen i dioden.

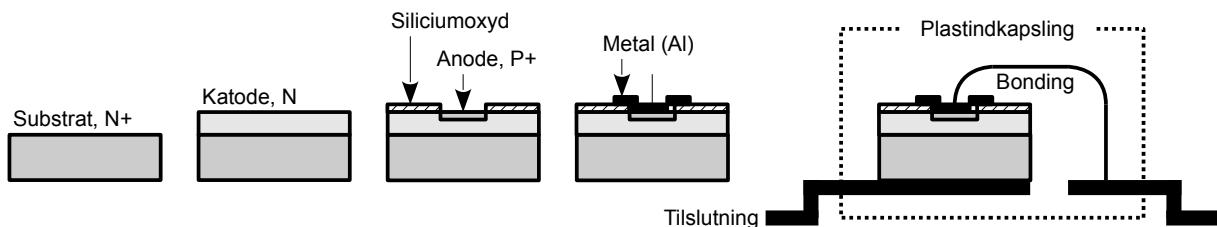
Når P og N lagene bringes i kontakt vil de frie elektroner i nærheden af PN overgangen diffundere i retning af P hvor der er et underskud af elektroner. Elektronen efterlader et type V atom der mangler en elektron og det bliver til en positiv ion i krystalgitret. Den frie elektron vil i P halvlederen møde et type III atom der kan optage elektronen og danne en negativ ion. Diffusionen opbygger derved et elektrisk felt over PN overgangen som bliver kraftigere efterhånden som der diffunderer elektroner mod P. Feltet modvirker diffusionen som klinger ud og der opnås en balance mellem diffusionen ogfeltet. Området mellem N og P er tømt for ladningsbærere og kaldes for *depletion layer*.

Hvis en ekstern spænding påtrykkes med plus ved P og minus ved N vil det tilføre frie elektroner til N halvlederen som rekombinerer med ionerne ved PN overgangen og mindsker det elektriske felt. Derved kan elektroner diffundere gennem PN overgangen og rekombinere med hullerne i P. Det skaber et underskud af huller i P som gendannes ved at elektroner løber til batteriets positive pol. *Der løber en elektrisk strøm i dioden.* Hvis det eksterne batteri vendes om vil det derimod trække elektroner ud fra N hvorved det elektriske felt øges og blokerer for en strøm. *Dioden spærre.*

5.1.2 Fremstilling

De første halvledere blev produceret ved at placere en perle af ét halvledermateriale på en basis af et andet og derefter varme konstruktionen op så grænseområderne smelte sammen. Det var ret enkelt at fremstille dioder på denne måde, men det var meget vanskeligt at styre geometrien for en produktion af transistorer. Det ledte til udvikling af planarteknologien der i dag danner grundlaget for produktion af dioder, transistorer og integrerede kredsløb i stor skala.

Først fremstilles et basismateriale fra et bad af smeltet silicium. Et krystal af rent silicium dypes og trækkes langsomt op hvorved siliciumatomer udfældes på overfladen og opbygger et næsten perfekt krystal. Ved tilsætning af type III eller V under processen doteres materialet til en halvleder af type P eller N med henholdsvis underskud eller overskud af elektroner. Derefter opbygges et lag af siliciumoxyd på halvlederen overflade. Et lysfølsomt lag påføres og blyses gennem en maske og derefter kan dele af laget ætses væk (fotolitografi). En gas af type III eller V tilsættes ved en temperatur på 600 til 800°C hvorved atomerne diffunderer ind gennem åbningerne og erstatter enkelte af siliciumatomerne. Det kalds for dotering og kan tilmed give fejl i krystalstrukturen. Ved dotering med basismaterialets forurening dannes et kraftigt doted område ($P+$ eller $N+$) der er velegnet for en elektrisk tilslutning og ved kraftig dotering med modsat polaritet overdøves den oprindelige dotering så materialet skifter funktion mellem P og N.



Figur 201 – Dioden fremstilles af smeltet silicium med kraftig dotering ($N+$) hvorpå et tyndt lag med svag dotering bygges (N). Ved en fotolitografisk proces doteres en del af N laget om med en kraftig dotering ($P+$) hvorved PN overgangen dannes. Et metallag skaffer elektrisk kontakt til anoden. Ved fremstilling af transistorer og integrerede kredsløb benyttes samme metode blot doteres der flere gange og metallag benyttes til at danne interne forbindelser mellem komponenterne.

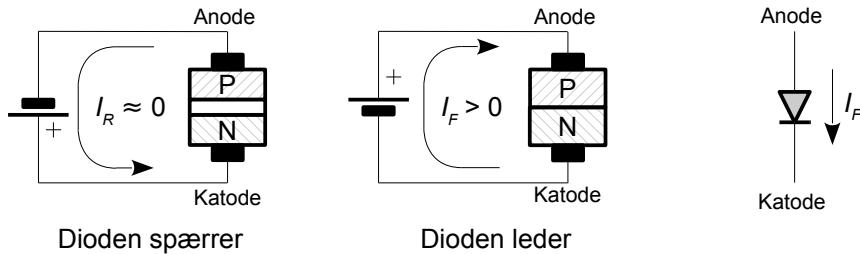
Selv om den dotedede halvleder er en relativt god leder så er modstanden for høj til at man kan se bort fra den. Derfor blev processen forfinet med et *epitaxial layer*; en kraftigt doted halvleder der udgør halvlederen basismateriale (*substrate*). Det kraftigt dotede lag ($N+$ i illustrationen) har til funktion at være en god elektrisk leder, men det har ikke en selvstændig funktion som halvleder. Alle komponenterne opbygges på en skive (*wafer*) med en diameter der er vokset fra 4 cm til over 40 cm gennem halvlederen udvikling. Der kan produceres mange tusinde helt ens komponenter ved den fotolitografiske proces. Komponenterne adskilles ved at skiven ridses med en diamant og knækkes i små stykker langs ridserne på samme måde som man skærer en glasrude til.

Komponenternes elektriske tilslutninger forbindes til eksterne terminaler ved svejsning af en tynd sølvtråd (*bonding*) og de små stykker støbes i plast eller keramik dels for at opnå god mekanisk stabilitet, dels for at blokere for luftens oxyderende ilt og dels for at blokere for lys da det kan slå elektroner løs i PN overgangen og derved producere en strøm i dioden. Den fotoelektriske effekt bliver dog udnyttet ved fotodioder og solpaneler.

Ud over dotering kan halvlederen udsættes for et bombardement med ioner af den ønskede type. Indtrængningsdybden er 10 nm til 1 µm så det er bedst egnet tæt ved overfladen af halvlederen. Der kan opnås stor koncentration af forurenningen, men den høje energi som ionerne accelereres op til kan give betydelige skader på halvlederen krystal. Det formodes at være en af årsagerne til den såkaldte 1/f-støj i halvledere.

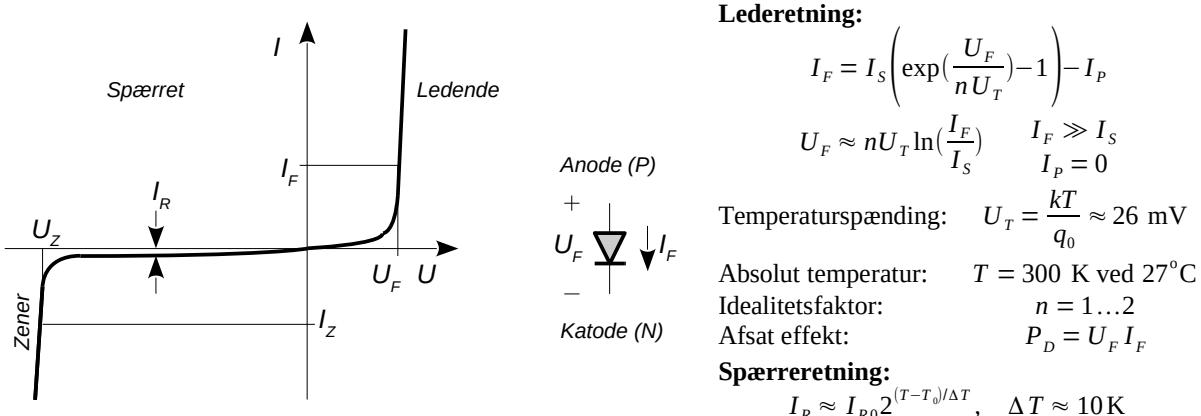
5.1.3 Diodens I-U karakteristik

I en PN overgang styres strømmen i dioden af en ekstern spændingskilde der forskyder balancen mellem diffusionen og det elektriske felt så der slipper ladningsbærere gennem PN overgangen. Hvis batteriet forbindes omvendt vil huller og elektroner trækkes væk fra PN overgangen hvilket vil forøge det elektriske felts styrke så diffusionen standses og der ikke løber nogen strøm bortset fra de relativt få elektroner som omgivelserne vedPN overgangen slår løs i halvlederen.



Figur 202 – Med dioden forspændt i spærretretningen øges bredden af PN overgangen og der løber ingen betydnende strøm i dioden. Med dioden forspændt i ledertretningen reduceres det elektriske felt over PN overgangen og diffusionsstrømmen får lov til at løbe. Den elektriske strøm bæres af huller i valensbåndet, der løber fra batteriets pluspol og ind mod PN overgangen, og af elektroner i ledningsbåndet, der løber fra batteriets minuspol og ind mod PN overgangens anden side.

Diodens strøm varierer eksponentielt med spændingen over dioden og det gælder også for den bipolare transistor der introduceres side 218. Det teoretiske fundament præsenteres side 271.



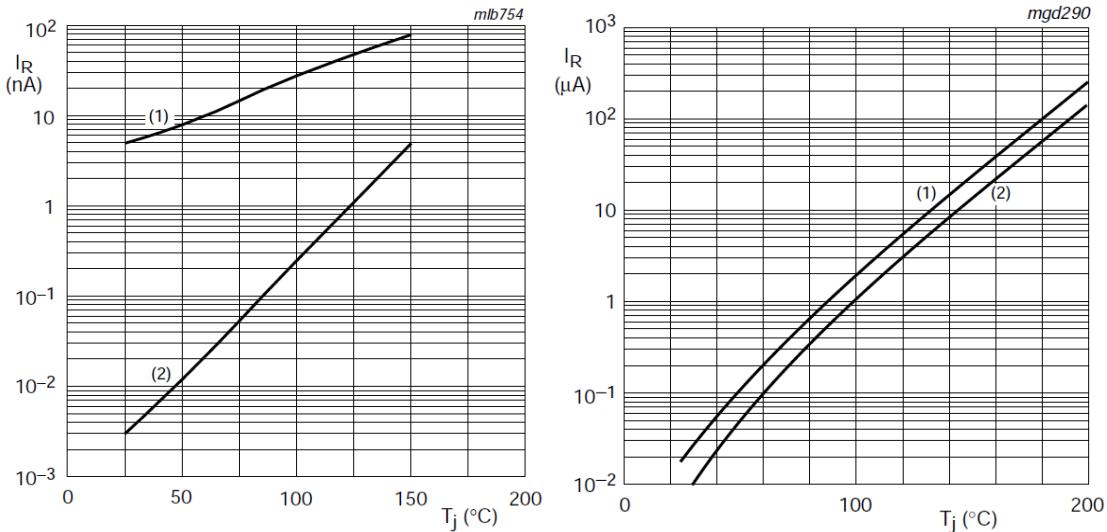
Figur 203 – Diodens karakteristik $I(U)$ har tre vigtige områder. I det ledende område varierer diodens strøm eksponentielt med diodens spænding. I det spærrende område er diodens strøm meget lille, men varierer eksponentielt med temperaturen. I zener området vil diodens strøm vokse voldsomt når værdien af spændingen over dioden overskrider en grænse givet af fremstillingsprocessen.

I ligningen er I_F strømmen i diodens ledertretning fra anode til katode (*forward current*) som funktion af spændingen U_F (*forward voltage*). Relationen er eksponentiel og strømmen vokser en dekade ved en ændring af spændingen over dioden på cirka 60 mV for $n = 1$. Relationen vil almindeligvis holde over et område på seks dekader fra 10 nA til 10 mA for småsignaldioder, men databladet bør konsulteres hvis det er kritisk for anvendelsen. Ved en stor strømstyrke afviger relationen fra eksponentialfunktionen på grund af modstandsværdien i halvlederen og det uddybes side 206.

I spærretretningen er strømmen I_R givet af diffusionsstrømmen for U_F negativ (mætningsstrømmen) plus de elektron-hulpar som temperaturen genererer i PN overgangen, af strømmen i krystallets overflade og strømmen i diodens indkapsling hvis yderside er utsat for omgivelsernes forurening. Det kaldes under ét for diodens lækstrøm I_R . Strømmen er voldsomt afhængig af temperaturen og omtales i næste afsnit.

5.1.4 Lækstrøm

Strømmen i spærretretningen I_R (*reverse current*) kaldes også lækstrøm (*leakage current*) og er en funktion af konstruktionen. Eksemplerne viser 3 pA for BAS116 og 10 nA for 1N4148 der begge er af silicium. En germaniumdiode som AA119 har 1 μ A og Shottky dioden 1N5819 når op på 1 mA.



Figur 204 – Mætningsstrøm i siliciumdioderne BAS116 til venstre og 1N4148 til højre (NXP, Philips).
For begge kurver er (1) maksimum og (2) typisk ved 75 V spænding.

Data opgives ved 25°C og værdien er stærkt temperaturafhængig med fordobling for hver 10°C så værdien øges med faktor 1000 for en temperaturstigning på 100°C. Det kan modelleres ved I_{R0} der er strømstyrken ved temperaturen T_0 på 25°C og en eksponentiel relation til temperaturen. Det skal dog bemærkes at det ikke udgør en eksakt relation.

5.1.5 Temperaturspænding

U_T kaldes temperaturspændingen (*temperature voltage*) og er på 26 mV ved stuetemperatur. Den er givet af Boltzmans konstant k , den absolutte temperatur T og elektronens ladning q_0 . Variationen er 0,33 %/°C ved stuetemperatur på grund af proportionaliteten med den absolutte temperatur. Temperaturspændingen bliver introduceret på side 271.

5.1.6 Idealitetsfaktor

Parameteren n kaldes for diodens idealitetsfaktor (*ideality factor*) der er et kompromis mellem de forskellige mekanismer der styrer strømmen i dioden. For en siliciumdiode er diffusionsstrømmen ofte dominerende ($n = 1$). Elektron-hul rekombination i PN overgangen (*space recombination*) og strømmen i krystallets overflade er især betyldende ved lav strøm ($n = 2$) og forklarer hvorfor en transistors strømforsærkning aftager ved meget lav strømstyrke.¹⁰¹

Værdien er omkring 1,6 for 1N4148 [DS-151 samt egne målinger] mens Shottky-dioder har værdier fra 1,0 til 1,1 [CC-241]. Megen teoretisk litteratur ser helt bort fra idealitetsfaktoren.

5.1.7 Fotostrøm

Parameteren I_P repræsenterer den strøm som absorption af en foton giver anledning til når den slår en elektron løs i PN overgangen. Elektronen tiltrækkes af den positive ladning fra ionerne i N halvlederen og føres til katoden som en overskydende elektron. Det er grunden til at en fotodiode har strømmen flydende modsat normal strømretning. En strøm af fotoner udgør en effektkilde P_λ og det er denne funktion der er central for brugen i fotooptik og solpaneler (se side 214).

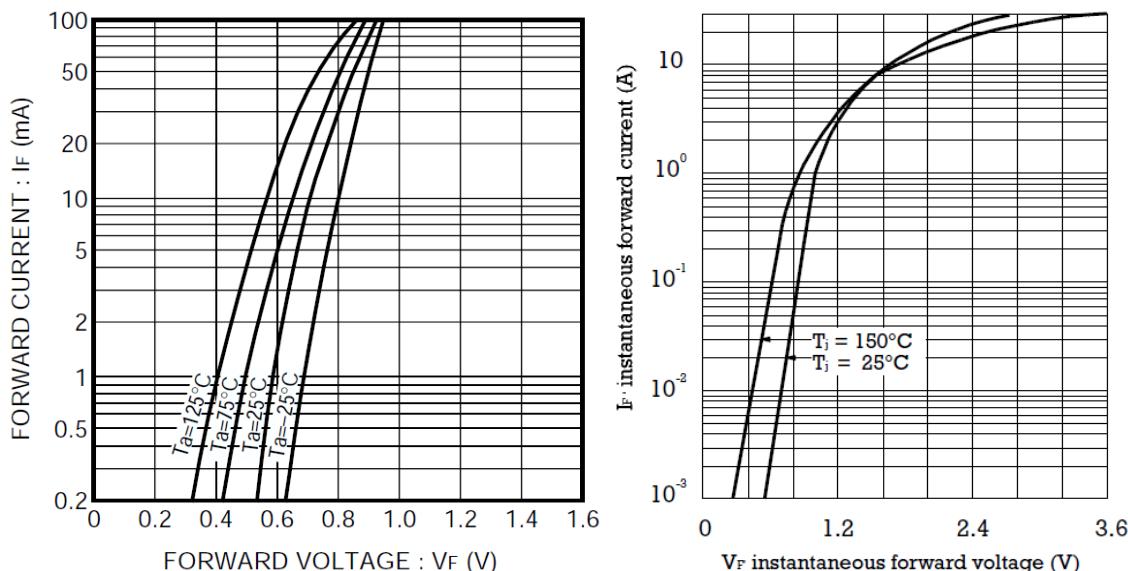
¹⁰¹En del af litteraturen benytter det græske bogstav η (eta).

5.1.8 Spændingsrelation

Den eksponentielle relation mellem strømmen I_F igennem dioden og spændingen U_F over dioden kan løses så spændingsfaldet bliver en logaritmisk funktion af strømmen. Spændingen over dioden vil stige med 60 mV når strømmen øges en dekade for $n = 1$ stigende til 120 mV/dekade for $n = 2$. Den absolute spændingsværdi varierer med teknologien som det vises side 199. Den logaritmiske egenskab er meget præcis og jeg skal senere vise eksempler baseret på den.

Herunder vises den publicerede karakteristik for de to populære dioder 1N4148, der udmærker sig ved at dukke op alle vegne hvor der er brug for "en diode", og 1N4007 der tåler en højere værdi af både strømmen i lederetningen og spændingen spærreretningen før sammenbrud. Dioder for høj strømværdi har et stort areal ved PN overgangen og kan derfor opfattes som parallelle smådioder der deler strømmen.

Den eksponentielle karakteristik skal ideelt set give en ret linje ved logaritmisk afbildning og det ses at passe under et vist strømniveau. Ved høje strømværdier krummer kurverne mod højre på grund af modstanden i halvlederen som giver et ekstra bidrag til spændingen over dioden. Som det venstre eksempel viser er 1N4148 omrent logaritmisk op til 10 mA og 1N4001 til 1 A. Internettet er rig på data for andre dioder.¹⁰²



Figur 205 – Relation mellem strøm og spændingsfald for småsignaldioden 1N4148 (Rohm), samt de lidt kraftigere dioder fra 1N4000-serien hvor 1N4001 kan klare 50 V i spærreretningen og 1N4007 op til 1600 V (Fagor). Den eksponentielle relation er gyldig i det område hvor linjerne er rette så begrensningen indtræder ved 10 mA for 1N4148 og 500 mA for 1N4001. Temperaturen ses som en forskydning mod venstre ved stigende temperatur. Kurverne krydser ved høj værdi af strømmen hvor modstanden bliver mest betydnende for dens værdi stiger med temperaturen.

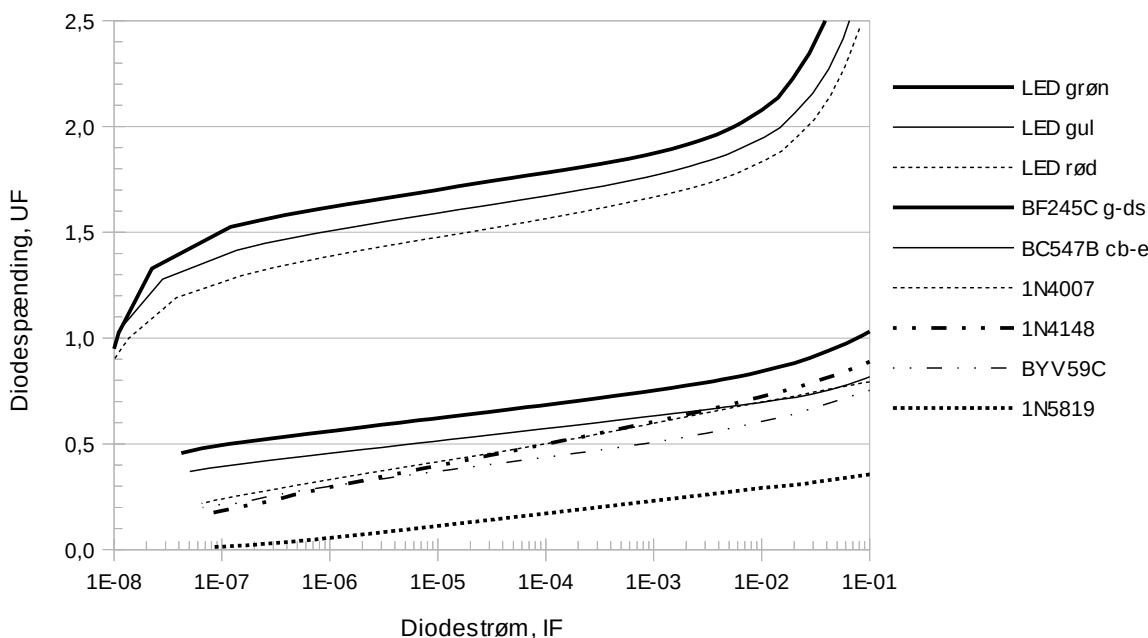
Der ses en relation til temperaturen ved den horisontale forskydning af kurverne. Spændingsfaldet aftager ved stigende temperatur for fastholdt strømværdi hvilket kan udtrykkes som større villighed til at lede strømmen ved stigende temperatur. For dioden 1N4148 indikerer kurven at spændingen aftager 180 mV for 100°C stigning så temperaturkoefficienten TC (*temperature coefficient*) bliver på $-1,8 \text{ mV}^\circ\text{C}$ og for 1N4007 er værdien på $-2,5 \text{ mV}^\circ\text{C}$. Intervallet afrundes ofte til $-2 \text{ mV}^\circ\text{C}$ og gælder for både dioder og transistorer.

Dioden kan bruges som et termometer ved at give den en fast strøm gennem en modstand og så måle diodens spænding. Der kræves dog en kalibrering fx med isvand og kogende vand.

¹⁰²Katalogfirmaer: Cypax, ELFA, Ericsson, Farnell og RS (Radio Shack).

Producenter: Diodes, Fagor, Fairchild, IRF, Motorola, NXP, On-Semi, Philips, Rohm, Signetics, Telefunken, Vishay, Zetex.

Herunder vises en måling på ni dioder. De tre øverste er lysdioder hvor det ses at den grønne lysdiode har højere spændingsfald end den røde fordi der kræves mere energi til at accelerere elektronerne op i fart. Ledespændingsfaldet er 1,8 ... 2,1 V ved en strøm på 10 mA. Materialet er type III (aluminium, gallium og indium) og type V (fosfor og arsen) da en type III-V halveder er bedre til at danne lys end silicium. Hældningen er omkring 90 mV/dekade så idealitetsfaktoren findes til $n \approx 1,5$. Kurverne krummer opad ved en strøm over 1 mA på grund af halvlederens modstand der beregnes til cirka 10 Ω . Afbøjningen under 100 nA tilskrives en usikkerhed ved bestemmelse af strømstyrken.



Figur 206 – Spændingsfaldet over et antal dioder som funktion af strømmen.

Målingen inkludere ikke dioder af germanium, men de vil ligge omrent som Shottky-dioden. Det vil sige et lavt ledespændingsfald og en høj lækstrøm for en diode i spærretretningen.

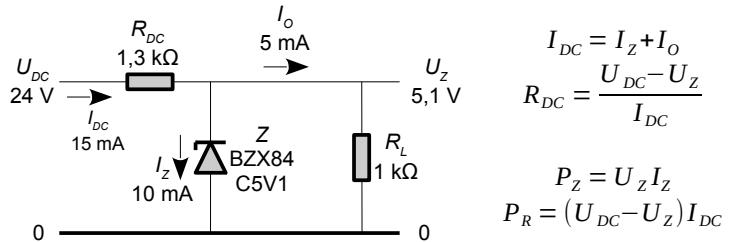
De følgende kurver er for halvledere af silicium hvor de to første er for diodekoblede transistorer. Øverst en JFET med drain forbundet til source ($n \approx 1,1$) og derefter en BJT med kollektor til basis ($n = 1,0$). Den beregnede værdi af halvlederens modstand var af størrelsesordenen 1 Ω for begge transistorer hvilket indikerer at det er småsignaltransistorer. Den højere værdi af ledespændingen for den målte JFET svarer til at den har en lav mætningsstrøm der blev målt til 2 fA for BF245C. Mætningsstrømmen blev fundet til 20 fA for BC547B der til gengæld har en lavere ledespænding.

De næste kurver refererer til de to populære dioder 1N4007 ($n \approx 1,5$) og 1N4148 ($n \approx 1,7$) samt dioden BYV59C ($n \approx 1,2$). Den beregnede værdi af halvledernes modstand var omkring 1 Ω og det halve for 1N4007. Mætningsstrømmene for dioderne blev beregnet til 1 ... 10 nA der er en del højere end for transistorerne til gengæld er ledespændingsfaldet generelt lavere ved den samme strømstyrke hvilket dog kun gælder under cirka 1 mA da idealitetsfaktoren er højere.

Den sidste diode er 1N5819 ($n \approx 1,0$) som er en Shottky-diode og altså består af et metallag mod en siliciumhalvleder og den har den laveste af de beregnede modstandsværdier på cirka 0,1 Ω . Den lave værdi af ledespænding og modstand forklarer hvorfor Shottky-dioder er populære valg for DC-DC konvertere hvor effekttabet er af primær betydning. Det lave spændingsfald vil også betyde at mætningsstrømmen skal være høj og den blev målt til 0,1 mA.

5.1.9 Zenerdiode

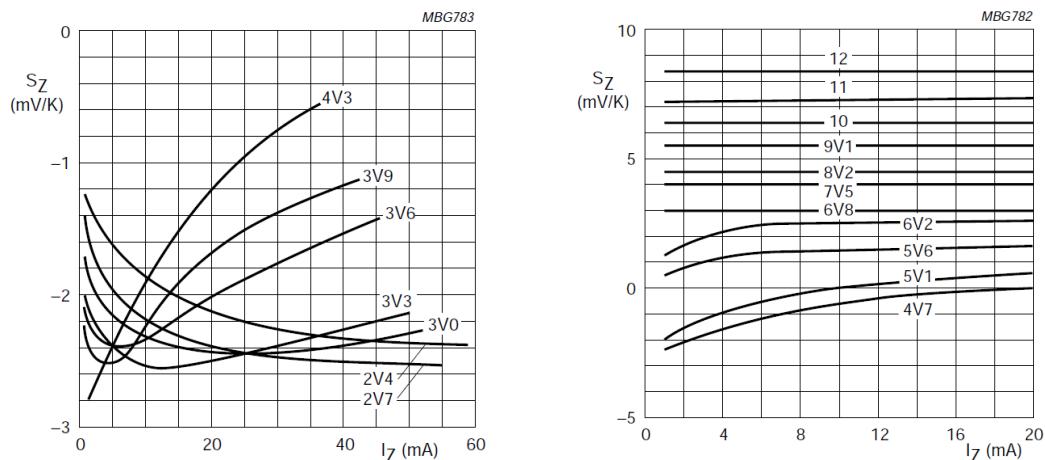
I spærreretningen er lækstrømmen lav indtil spændingen bliver så stor at PN overgangen bryder sammen og der løber en strøm I_z i spærreretningen. Det er ikke destruktivt blot effektafsættelsen holdes vel under databladets specifikation på 0,25 W for en typisk småsignaldiode. Værdien af spændingen kan designes fra 3 V til flere kilovolt ved at styre halvlederens doteringsprofil¹⁰³.



Figur 207 – En zenerdiode benyttes til stabilisering af en indgangsspænding med en modstand indsat for at styre diodens strøm så effektkravene overholdes.

Eksemplet nedsætter 24 V til 5,1 V med databladets specifikation på $I_z = 10 \text{ mA}$ for zenerdioden og dertil med $I_o = 5 \text{ mA}$ til en ekstern belastning. Der afsættes 0,05 W i zenerdioden og 0,3 W i modstanden. Udgangsstrømmen stiger til 18 mA hvis udgangen kortsluttes og modstanden belastes da med 0,44 W. Tilsvarende stiger effekten i zenerdioden til 0,08 W hvis udgangen ikke belastes (R_L fjernet).

Dioder designet for denne anvendelse har typisk en tolerance på $\pm 5\%$ af spændingsværdien U_z ved en specificeret strømværdi, men de kan leveres med en tolerance på $\pm 2\%$. Zenerdioden var en populær komponent indtil de integrerede kredse slog igennem med prisbillige alternativer som fx 7800-serien af spændingsregulatorer med en tolerance på $\pm 3\%$. Spændingsreferencer leveres også med en tolerance på $\pm 0,1\%$ for LM4051A eller $\pm 0,02\%$ for ADR4520.



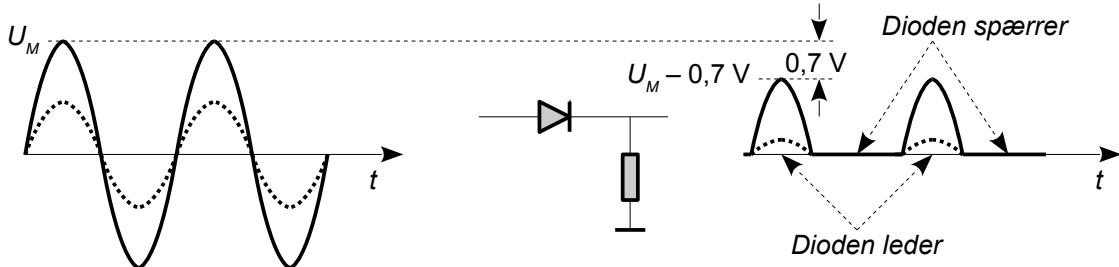
Figur 208 – Temperaturkoefficienten for dioden BZX84 serien er funktion af strømmen i dioden og skifter fortegn ved en spændingsværdi omkring 5 V.

Temperaturkoefficienten af zenerdioden er funktion af spænding og strøm og ved cirka 5 V er variationen stort set uafhængig af temperaturen ved en passende valgt strømstyrke. Transistoren har et tilsvarende sammenbrud ved de to PN overgange mellem kollektor-basis og emitter-basis og sidstnævnte bryder sammen allerede ved cirka 6 V.

¹⁰³Der er to årsager til sammenbruddet med den fælles betegnelse zenerdiode. Med en skarp ladningsprofil vil der være en tynd barriere ved PN overgangen (under 1 μm) så den elektriske feltstyrke bliver høj selv ved en lav spænding i spærreretningen. Det elektriske felt kan trække elektroner løs fra valensbåndet i N halvlederen og føre dem til ledningsbåndet. Det kaldes en tunneleffekt (tunneling). Ved en langstrakt ladningsprofil vil elektroner der er blevet frigjort af temperaturen accelereres op af det elektriske felt og ved sammenstød med atomerne i krystallet kan andre elektroner frigøres som derefter accelereres op og så videre. Det kaldes for en lavineeffekt (avalanche) og er dominerende ved højere spænding.

5.1.10 Ensretning

I illustrationen herunder ledes en spændingssvingning igennem en diode hvis udgang belastes af en modstand til nul. Hvis indgangssignalet er positivt vil udgangen følge med indgangssignalet i den positive halvperiode dog med en forskel på cirka 0,7 V.

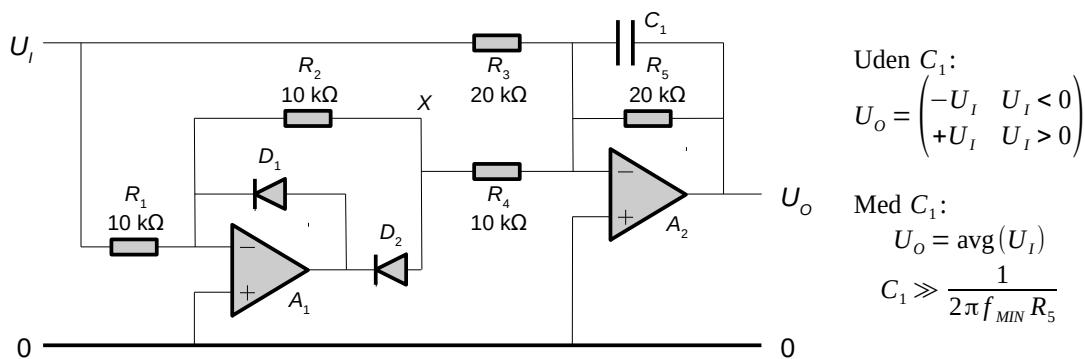


Figur 209 – Halvbølge ensretning af et vekselsignal giver en pulserende jævnstrøm for der kan kun løbe strøm i diodens lederetning så udgangsspændingen kan ikke blive negativ.

Ensretteren blev omkring 1900 benyttet ved modtagelse af amplitudemoduleret radiotransmission (AM, *amplitude modulated*) hvor det lavfrekvente signal fra tale, musik, video eller digitale data er moduleret ind på en højfrekvent svingning der kaldes for bærebølge (*carrier*). Amplituden ændres i takt med det lavfrekvente signal og det resulterende signal sendes fra en antennen til modtageren der modtage den valgte bærebølge gennem et selektivt filter. Dioden gendanner den oprindelige modulation ved at fjerne den ene halvdel af det højfrekvente signal idet middelværdien af signalet repræsenterer det oprindelige signal. Efter 1960 overgår radiotransmissionen til FM der benytter en frekvensmodulation af bærebølgen mens TV-transmissionen fortsætter med AM indtil 2009.

5.1.11 Absolutværdi

Spændingsfaldet over dioden er generende ved brug i analog signalbehandling som fx beregning af et signals absolutværdi og den kan undgås ved at benytte en operationsforstærker for at danne en såkaldt "ideel diode". Det vises herunder hvor A_1 modvirker spændingsfaldet over dioden D_2 og A_2 beregner absolutværdien af signalet så der opnås en helbølge-ensretning af signalet. Med en kondensator i parallel over R_5 lavpasfiltreres udgangssignalet. Operationsforstærkeren sætter en øvre grænse for det brugbare frekvensområdet til et sted i kilohertz-området.

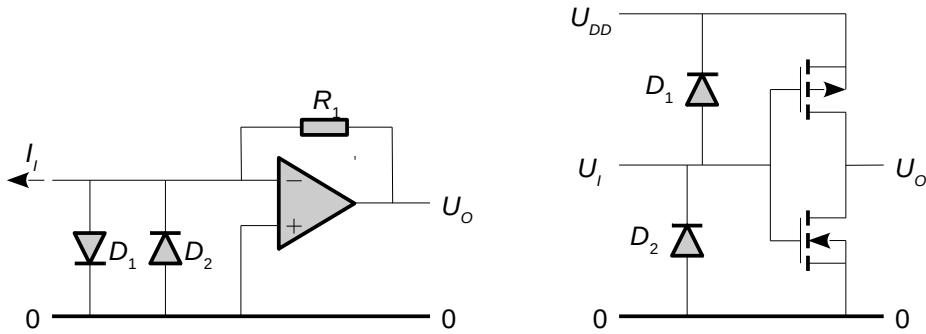


Figur 210 – Helbølge ensretning af et analogt signal for brug ved signalbehandling. Kondensator C_1 benyttes til at midle signalet og kan fx vælges til 100 gange under den laveste frekvens.

Indgangssignalet U_I føres ad to veje til udgangen. Gennem R_3 er forstærkningen $-R_5/R_3$ der med de viste komponentværdier er på -1 . Gennem R_1 er forstærkningen R_2R_5/R_1R_4 der her er 2 gange. Ved negativt indgangssignal går udgangen fra forstærker A_1 positiv indtil D_1 bliver ledende og D_2 spærre. Udgangen er en positiv kopi af det negative indgangssignal. Ved et positivt indgangssignal vil A_1 være en inverterende forstærker ved punktet X så de to forstærkningsmuligheder adderes til en forstærkning på 1 . Udgangen er altid positiv så signalet fra opstillingen er absolutværdien af indgangssignalet.

5.1.12 Beskyttelseskredsløb

Dioden kan beskytte en følsom indgang og to typiske anvendelser skal vises. I det første skal en forstærker måle en strøm så der er direkte forbindelse til den negative indgang og dermed ingen beskyttelse imod fejl. To antiparallele dioder kan begrænse spændingen på indgangen til $\pm 0,7$ V. Diodernes mætningsstrøm kan påvirke funktionen og kapaciteten i dioden kan give ustabilitet.

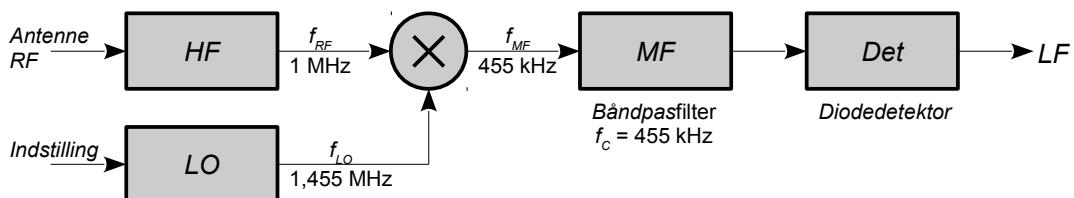


Figur 211 – Diodernes beskytter en følsom indgang imod overspænding (venstre), og hindrer en ødelæggelse af de højimpedante indgange på en CMOS inverter (højre).

De populære CMOS kredsløb¹⁰⁴ ødelægges af overspænding og benytter to dioder for at holde indgangen over $-0,7$ V og under $U_{DD} + 0,7$ V. En strøm i dioden D_1 vil løbe til U_{DD} der skal kunne optage strømmen for at beskyttelsen ikke bliver illusorisk (det er en klassisk designfejl).

5.1.13 Radiomodtagere

En vigtig komponent i radiomodtagere er mikseren (*mixer*) der bruges til at omsætte bærebølgen fra antennen til en lavere frekvens der er lettere at håndtere. Mikseren foretager en multiplikation af det modtagne højfrekvente signal *RF* (*radio frequency*) og radiomodtagerens egen oscillator *LO* (*local oscillator*). På udgangen af mikseren, der vist som en cirkel med et gange tegn, optræder der et system af frekvenser hvor det er differensen mellem f_{RF} og f_{LO} der er interessant. I en AM radio er mellemfrekvensen almindeligvis på 455 kHz så lokaloscillatoren indstilles til at svinge 455 kHz højere end den frekvens der skal modtages.



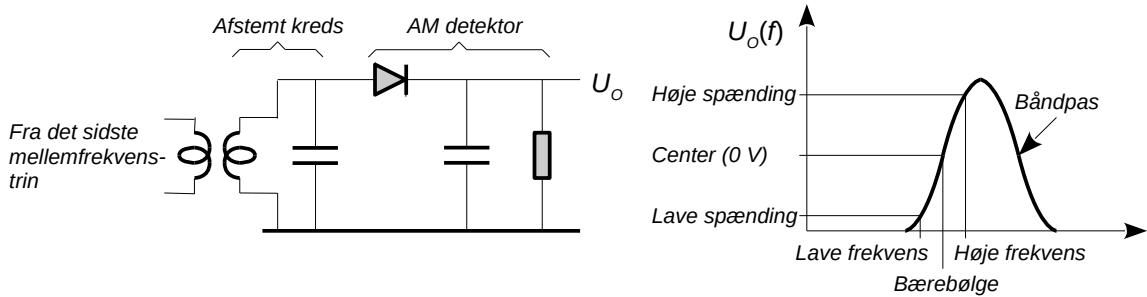
Figur 212 – En radiomodtager benytter dioden i en mikser for at omsætte et højfrekvensignal til en lavere frekvens, der kan forstærkes op og efter detektion føres lavfrekvenssignalet til udgangen.

Radiokanalen udvælges ved at justere på lokaloscillatoren idet kun den kanal hvis differens fra lokaloscillatoren ligger indenfor MF pasbåndet vil blive forstærket op og detekteret. Der optræder en spejlfrekvens for radiosendere der ligger 455 kHz over lokaloscillatoren og de må dæmpes af et filter i HF delen da de ellers vil blive modtaget og detekteret på linje med den ønskede kanal.

En moderne FM radiomodtagere benytter en mellemfrekvens på 10,7 MHz og en satellitmodtagere benytter flere lokaloscillatorer for trinvis nedsætning af frekvensen. Detektionen af det modtagne signal kan benytte den viste ensretter for amplitudemodulerede transmissioner (AM) hvor det ensrettede signals middelværdi repræsenterer modulationens signal. Ved en frekvensmoduleret transmission (FM) må frekvensændringen først omsættes til en tilsvarende amplitudemodulation.

¹⁰⁴Denne beskrivelse dækker hovedparten af CMOS4000 og 74HC-familierne og lavspændingsvarianter, alle mikroprocessorer og de programmerbare logiske komponenter, samt flertallet af analoge integrerede kredse.

En af mange metoder for at ændre en frekvensændring til en spændingsændring er at benytte den skrælle flanke på et afstemt filter. For at undertrykke enhver relation til signalets amplitude, hvilket ville optræde som en interferens, benyttes en så stor en forstærkning i mellemfrekvensen at signalet fra sidste trin overstyres og begrænses til en fast amplitude.

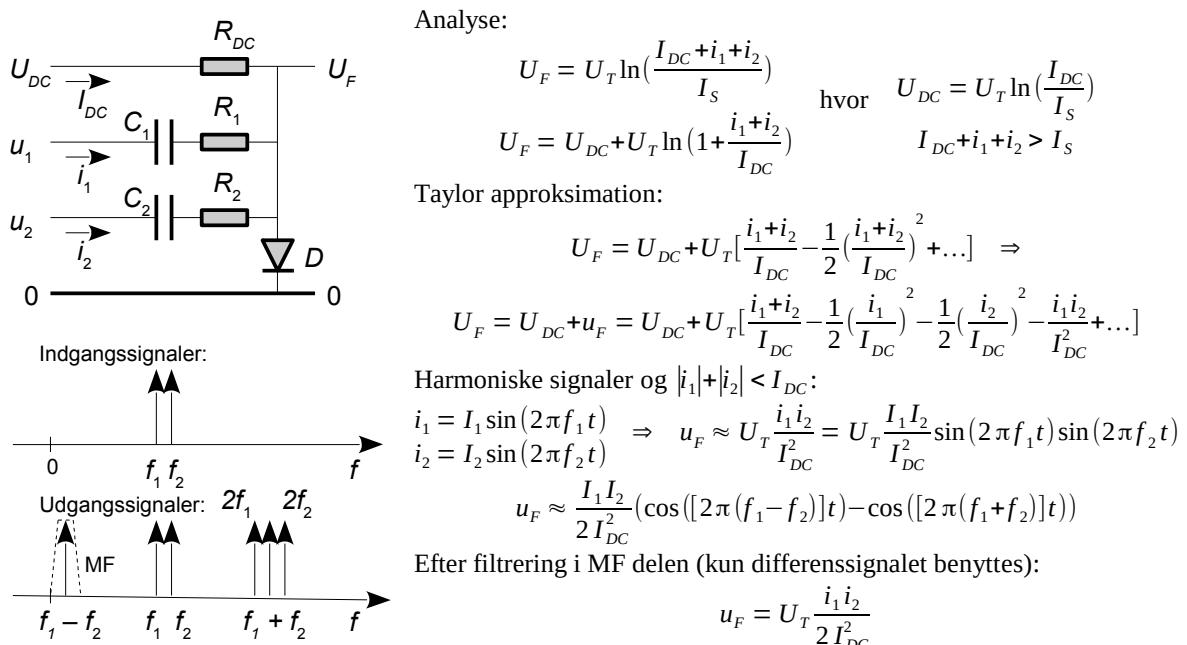


Figur 213 – I en FM modtager omsættes en frekvensændring i forhold til bærebølgen til en analog spændingsændring som derefter kan omformes til et lavfrekvensignal ved ensretning.

Krumningen af filtrets flange giver harmoniske forvrængning. Det kan reduceres med et såkaldt *stagger-tuned* filter der består af to svingningskredse hvor de to centerfrekvenser ligger tæt, men ikke er ens. Det giver en vendetangente i det skrælle område og derved en mere symmetrisk kurve.

5.1.14 Mikser

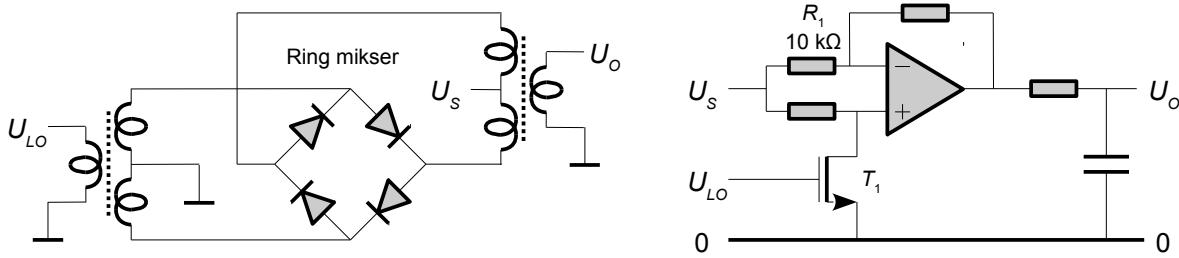
Frekvensomsætningen fra den modtagne bærebølge til mellemfrekvensens centerfrekvens har brug for en mikser. Herunder vises en mikser baseret på en diode. Indgangssignalerne er AC strømmene i_1 og i_2 og en DC strøm I_{DC} for at bringe dioden D til det ønskede arbejdspunkt.



Figur 214 – I en mikser benyttes diodens logaritmiske karakteristik til at gange to signaler sammen hvilket genererer summen og differensen af signalerne samt dertil de to oprindelige signaler og den dobbelte frekvens af dem. En Taylor-serie benyttes til at simplificere udtrykket for logaritmen og indgangssignalerne er harmoniske (sinus). Båndpasfiltret MF udvælger frekvensen $f_1 - f_2$.

Med 1 V amplitude for u_1 og u_2 og $R_1 = R_2 = 1 \text{ k}\Omega$ bliver de to indgangssignaler på $i_1 = i_2 = 1 \text{ mA}$ og ved en DC strøm på $I_{DC} = 5 \text{ mA}$ dæmpes differensens signal til 20 mV (-34 dB). De oprindelige signaler dæmpes til 200 mV og signalerne på den dobbelte frekvens dæmpes til 20 mV.

Dioden kan benyttes som mikser ved meget høje frekvenser, men indskudstabet er stort. Som beregningen viser er tabet omkring 34 dB så det er nødvendigt at forstærke det højfrekvente signal ganske meget inden mikseren benyttes.

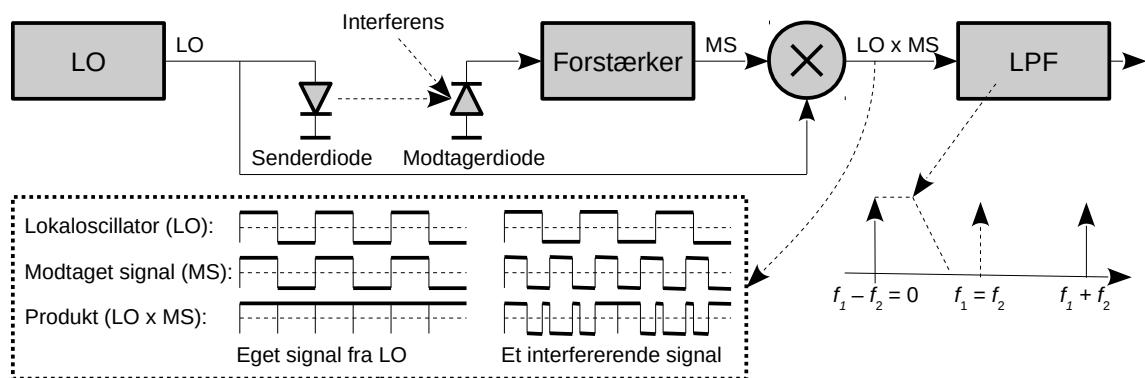


Figur 215 – En dobbeltdobbelteret mikser for højfrekvente signaler ved radiomodtagelse (venstre) og en analog mikser med et lavpasfilter for brug i industriel elektronik (højre).

En bedre konstruktion er den balanceret mikser som vises herover i to meget forskellige udgaver der dog er hinanden lig i funktion. De højfrekvente signaler kan være i gigahertz-området for det venstre kredsløb, men kun i kilohertz-området for det højre kredsløb. Lokaloscillatorenens U_{LO} vil for begge kredsløb omskifte udgangen mellem U_s og $-U_s$ så indgangssignalet bliver ganget med en firkant. Udgangssignalet består af sum- og differenssignalerne mellem U_s og U_{LO} .

5.1.15 Optiske sensorer

For brug i industrielle sensorer benyttes lokaloscillatoren LO til at styre en lysdiode der udsender lyspulser. Det modtagne signal fra fotodioden vil indeholde det pulserende lys fra senderen samt interfererende signaler. For sensorens eget signal giver differensen et konstant niveau hvorimod en forstyrrende interferens vil oscillere med en frekvens på differensen mellem de to sensorers LO frekvenser og kan undertrykkes med et lavpasfilter. Detektion af signalet kan herefter foretages ved at spændingen fra lavpasfiltret sammenlignes med et DC niveau.



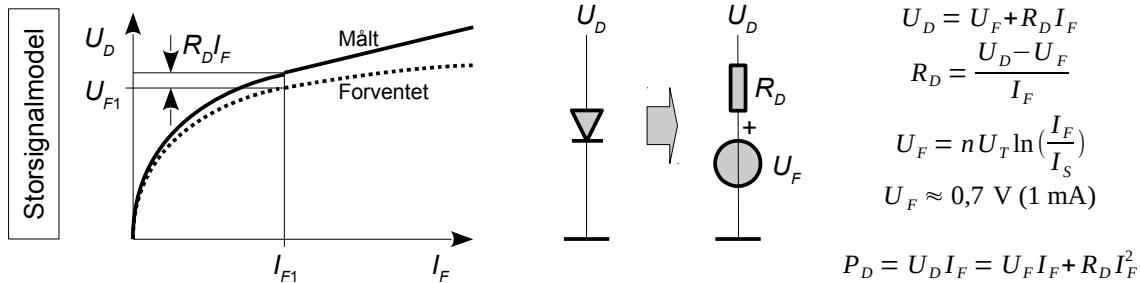
Figur 216 – Den balancerede mikser er meget anvendt for at opnå immunitet overfor forstyrrende interferens. Signalet fra lavpasfiltret LPF er en DC værdi hvis det modtagne signal er lig med lokaloscillatorenens signal. Ved en forskel i frekvens vil differenssignalet fra mikseren svinge med frekvensforskellen mellem LO og den interferende kilde og det fjernes med et lavpasfilter.

Med lokaloscillatoren indstillet til 10 kHz og lavpasfiltret til 100 Hz kan forstyrrende signaler i større afstand end 100 Hz fra sensorens eget signal dæmpes af filtret. Med et anden-ordens filter opnås 40 dB dæmpning ved en frekvensafstand på 1 kHz så andre sensorer kan operere under 9 kHz eller over 11 kHz med en forstyrrelse reduceret til cirka 1 %. Sensorens egen lokaloscillator vil give en DC spænding efter lavpasfiltret der er proportional med styrken af det modtagne signal.

Note 8.

5.1.16 Storsignal – DC seriemodstand

Halvlederen har en DC modstandsværdi og ved måling på dioden ses det som et spændingsfald i ledernetningen der er større end det forventede fra den logaritmiske relation. Det modelleres ved den forventede diodespænding U_F plus et tillæg fra diodemodstanden på $R_D I_F$. Det benyttes fx ved beregning af effekttabet i en effektforsyningens ensretter og for dioden i en DC-DC konverter.



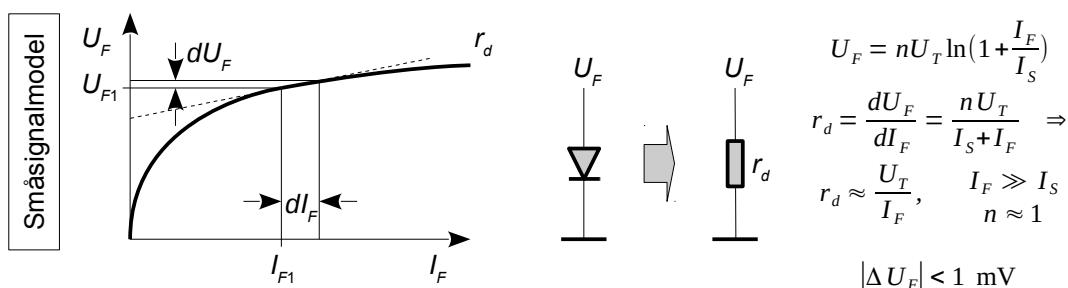
Figur 217 – Diodens seriemodstand skyldes ledningsevnen af halvlederenes materiale og en model af dioden kan inkludere en modstand.

En siliciumdiode måles til spændingen $U_D = 1,0$ V ved strømmen $I_F = 1$ A og modstanden af halvlederen skal estimeres. For en siliciumdiode er spændingen fx 0,7 V ved 1 mA og det korrigeres med 60 mA per dekade så den forventede spænding bliver $U_F = 0,88$ V ved $I_F = 1$ A. Spændingsfaldet over halvlederenes modstand bliver 0,12 V så seriemodstanden bliver $R_D = 0,12 \Omega$.

I en del litteratur benyttes tangenten til arbejdspunktet som krydser ordinaten i en specifik værdi, men der er ikke den store forskel på de forskellige metoder der alle blot giver en estimeret værdi af halvledermodstanden. Sagt anderledes så har modstanden R_D ikke en konstant værdi.

5.1.17 Småsignal – Dynamisk modstand

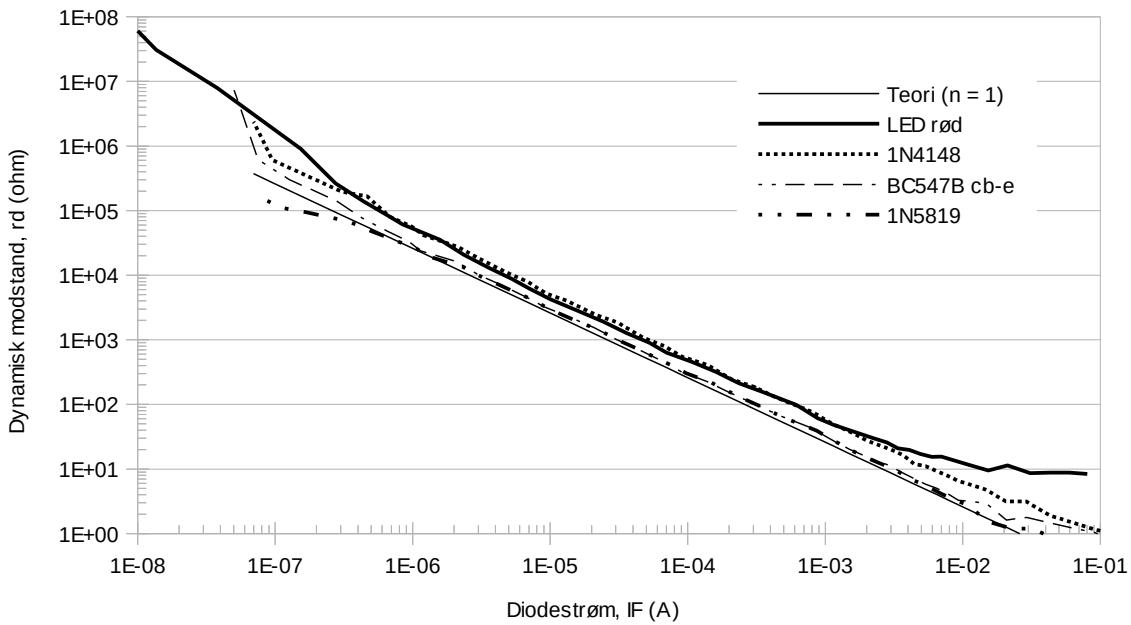
Ved lav strømstyrke hvor halvlederenes modstand R_D ikke har en større betydning er spændingen over dioden givet af den logaritmiske funktion. Ved små variationer i strømmen kan dioden ses som en strømstyret modstand der kan varieres fra 10Ω til $300 \text{ k}\Omega$ ved en strømstyrke fra 3 mA til 100 nA. Det kaldes for diodens dynamiske modstand r_d (*dynamic resistance*) i arbejdspunktet I_F . Det kan kun anvendes ved signaler med en amplitude under cirka 10 mV. Det var meget brugt i ældre elektronik for automatisk styrkeregulering (se side 207) hvor der i dag vil blive foretrukket anvendelse af et digitalt potentiometer eller en multiplicerende DAC.



Figur 218 – Diodens dynamisk modstand skyldes den logaritmiske relation mellem spænding og strøm. Spændingen kan ved små strømændringer beskrives ved en modstandsværdi der er givet ved tangenthældningen til kurven i arbejdspunktet.

For en typisk diode kan man passende antage $n = 1$ og sætte temperaturen til $T = 300$ K (27°C). Den dynamiske modstand bliver derved $r_d = 26 \Omega$ ved $I_F = 1$ mA og aftager med stigende DC strømstyrke.

Den målte værdi af den dynamiske modstand r_d er vist herunder for tre dioder samt en transistor koblet som diode med kollektor kortsluttet til basis.

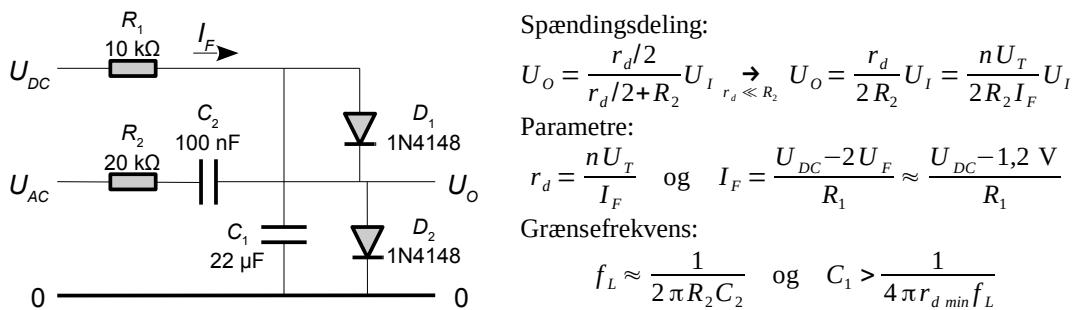


Figur 219 – Målte værdier af dynamisk modstand for en lysdiode (LED), en siliciumdiode (1N4148), en diode-koblet transistor (BC547B) og en Shottky diode (1N5819).

Formlen for den dynamiske modstand er vist som "Teori". Over 10 mA bliver DC halvledermodstanden dominerende for lysdioden så målingen går imod en konstant og under 200 nA er målenøjagtigheden årsagen. Idealitetsfaktoren ses som en lodret forskydning i forhold til teorien. Bemærk, at trods de store forskelle i diodernes halvledermateriale (type III-V, silicium og metal-halvleder) er kurverne stort set ens.

5.1.18 Lydstyrkeregulering

Den dynamiske modstand kan anvendes til strømregulering af forskellige parametre i et analogt kredsløb. Herunder vises et eksempel hvor amplituden af et AC signal kan varieres ved at ændre på værdien af en DC spænding. Der benyttes to dioder for at reducere forvrængningen af signalet hvilket tillader en signalamplitude på omkring 10 mV ved udgangen. Strømmen i dioderne gives af spændingskilden U_{DC} , modstanden R_1 og spændingsfaldet over serieforbindelsen af de to dioder. AC signalet kobles ind via R_2 hvor C_2 sikrer at strømmen i dioderne ikke påvirkes. På grund af C_1 er de to dioder i parallel overfor et AC signal så dioderne arbejder i parallel.

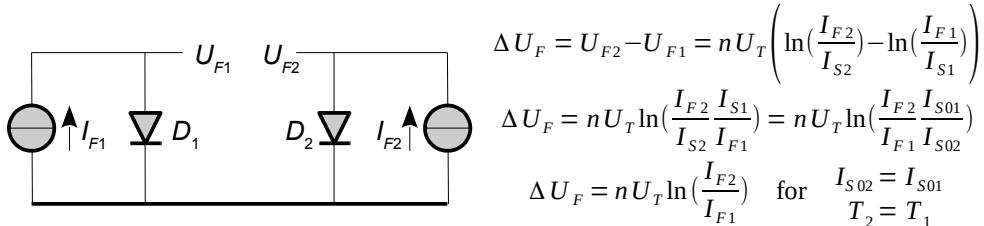


Figur 220 – Spændingsstyring af et signals amplitude ved brug af diodens dynamiske modstand. En DC spænding bestemmer diodens strøm og de to dioder giver en symmetrisk begrænsning.

Den dynamiske modstand er $r_d = 42 \Omega$ ved $I_F = 1 \text{ mA}$ for 1N4148 ($n = 1,7$) så signalet dæmpes 53 dB. Dæmpningen aftager til 15 dB ved $10 \mu\text{A}$, hvor det maksimale indgangssignal er aftaget til $\pm 200 \text{ mV}$. Reguleringen har en tidskonstant omkring 0,2 s og signalets grænsefrekvens er 80 Hz.

5.1.19 Analog signalbehandling

Hovedproblemet ved anvendelse af en diode til analog signalbehandling er mætningsstrømmens afhængighed af temperaturen. Som vist tidligere er den dynamiske modstand helt uafhængig af mætningsstrømmen på grund af differentiationen der effektivt set trækker to funktionsværdier fra hinanden. En løsning er derfor at studere virkningen af en differens mellem spændingsfaldet for to dioder der arbejder med hver sin strømværdi. Det leder til en metode der bruges i de integrerede kredsløb både for opnåelse af et stabilt arbejdspunkt og for måling af temperatur.



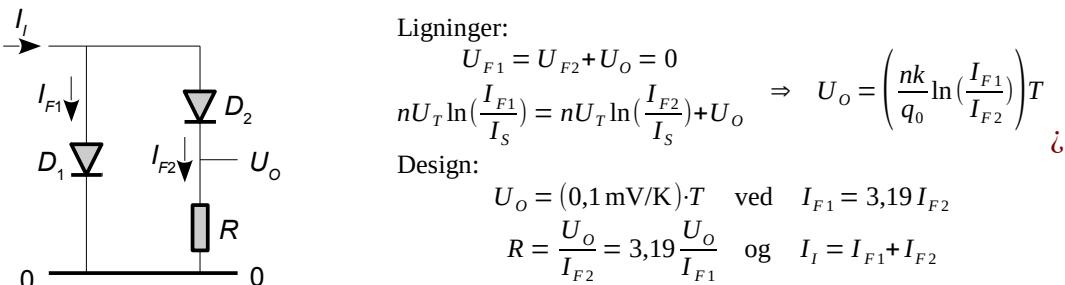
Figur 221 – Differensen mellem spændingen over to dioder er primært givet ved forholdet mellem de to strømstyrker og kun i ringe grad af de øvrige parametre.

Ved et forhold på 2 mellem strømværdierne for to ens dioder er differensen 18 mV og ved et forhold på 10 er differensen 60 mV ($n = 1$). Bemærk at relationen gælder for alle dioder (Si, Ge, GaAs) og uanset om de er sammensat af PN halvledere eller er dannet ved en metal-halvleder overgang som Shottky-dioder.

De to dioder arbejder ved hver sin strømværdi og differensen mellem deres spændinger bliver til en differens mellem to logaritmer hvilket kan skrives som logaritmen til forholdet mellem de to dioders strømme. Det har den attraktive betydning at mætningsstrømmen udgår når dioderne er ens. For to tilfældige dioder kan man ikke sige andet om forholdet mellem mætningsstrømmene end at det er tilnærmelsesvis konstant, men for to dioder på den samme siliciumskive er faktoren tæt på én og det åbner for brug af dioden til signalbehandling. En mikroprocessor kan benyttes til at omskifte den samme diode mellem to strømværdier og så går mætningsstrømmen ud.

5.1.20 Termometer

Enkel måde at trække to dioders spænding fra hinanden vises herunder hvor spændingen U_o er en funktion af den absolute temperatur. Konstanten er $nk/q_0 = 86,2 \mu\text{V}/^\circ\text{C}$ for $n = 1$ så udgangens spænding stiger 0,1 mV per grad celsius hvis logaritmen indstilles til 1,16. Det kræver at forholdet mellem strømværdierne er $e^{1,16} = 3,19$ så strømmene skal holdes på $I_{F1} = 3,19 I_{F2}$.



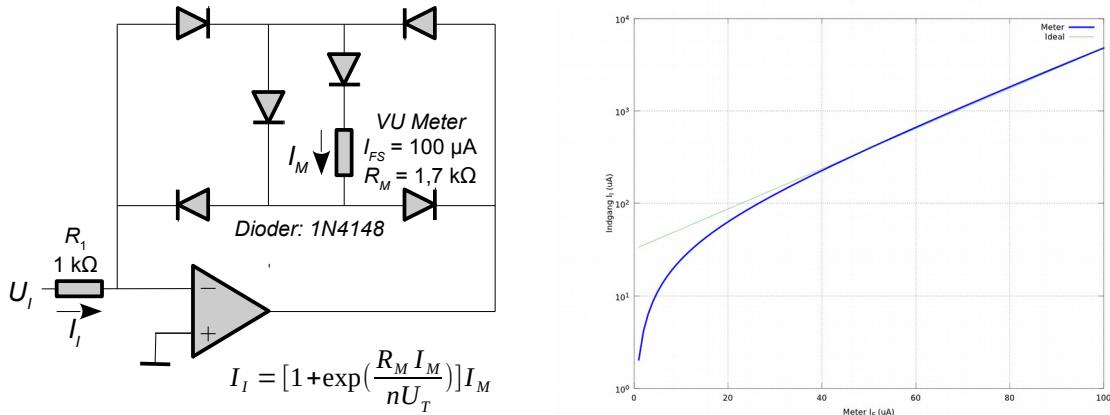
Figur 222 – Et termometer kan opbygges ved at lade to dioder arbejde ved hver sit strømniveau. Som dioder foreslås to BC547 benyttet i diodekobling.

Med $I_{F1} = 1 \text{ mA}$ bliver $I_{F2} = 0,31 \text{ mA}$ og $R = 83 \Omega$, vil strømmen ved indgangen være $I_I = I_{F1} + I_{F2} = 1,31 \text{ mA}$. Udgangsspændingen bliver $U_o = 30 \text{ mV}$ ved $T = 300 \text{ K}$ (27°C) og den ændres med $0,1 \text{ mV}/^\circ\text{C}$.

Opstillingen er ikke brugbar over et stort temperaturområde. Det skyldes at en stigende spænding på udgangen kræver en stigende strøm i diode D_2 , men hvis indgangens strøm holdes fast på I_I vil det medføre at strømmen i D_1 aftager og så holder relationen mellem I_{F1} og I_{F2} ikke.

5.1.21 Pseudo-logaritmisk måleinstrument

Indenfor lydverdenen er VU-metre¹⁰⁵ populære, men deres dynamikområde er kun 20 dB. Den viste opstillingen kan vise 50 dB på et instrument. En helbølgeensretter er bygget over en inverterende forstærker og fire dioder¹⁰⁶.



Figur 223 – Strømmen i instrumentet varierer omtrent logaritmisk med indgangsspændingen.

Kredsløbet er designet til at bruge 1N4148 ($n = 1,7$), men det vil fungere med de fleste dioder. Ved fuld udstyring på 100 μA ligger der 170 mV over R_M og strømmen i indgangen når 5 mA ved det kraftigste indgangssignal. Temperaturkoefficienten af drejespoleinstrumentets kobbertråd (0,33 %/K) kompenseres af diodernes tilsvarende temperaturkoefficient.

5.1.22 Geometrisk gennemsnit

Analoge kredsløb for løsning af matematiske beregninger med logaritmer og kvadratrødder bruges ikke i større omfang i dag for det varetages langt bedre af mikroprocessoren, men jeg har været ude for at en mikroprocessor ikke kunne indbygges i en sensor og en løsning derfor var ønsket¹⁰⁷.

Herunder vises hvordan den geometriske middelværdi af to signaler kan dannes. Den aritmetiske middelværdi af to signaler er let at beregne for det kræver blot to modstande, men det er mere kompliceret at beregne den geometriske middelværdi der er kvadratroden af produktet af to signaler, men det kan klares ved at udnytte diodens logaritmiske relation.

Indgangssignalerne kaldes x_1 og x_2 og første trin udnytter at eksponentialfunktionen til logaritmen giver signalet selv. Andet trin er at omskrive kvadratroden ved relationen $\ln(x^{0.5}) = 0.5 \cdot \ln(x)$, og det tredje trin er at logaritmen til et produkt lig med summen af logaritmen til hvert af signalerne. Det oversætter en multiplikation til en addition som enkelt kan klares af to modstande.

$$\sqrt{x_1 x_2} = \exp(\ln \sqrt{x_1 x_2}) = \exp\left(\frac{1}{2} \ln(x_1 x_2)\right) = \exp\left(\frac{\ln(x_1) + \ln(x_2)}{2}\right)$$

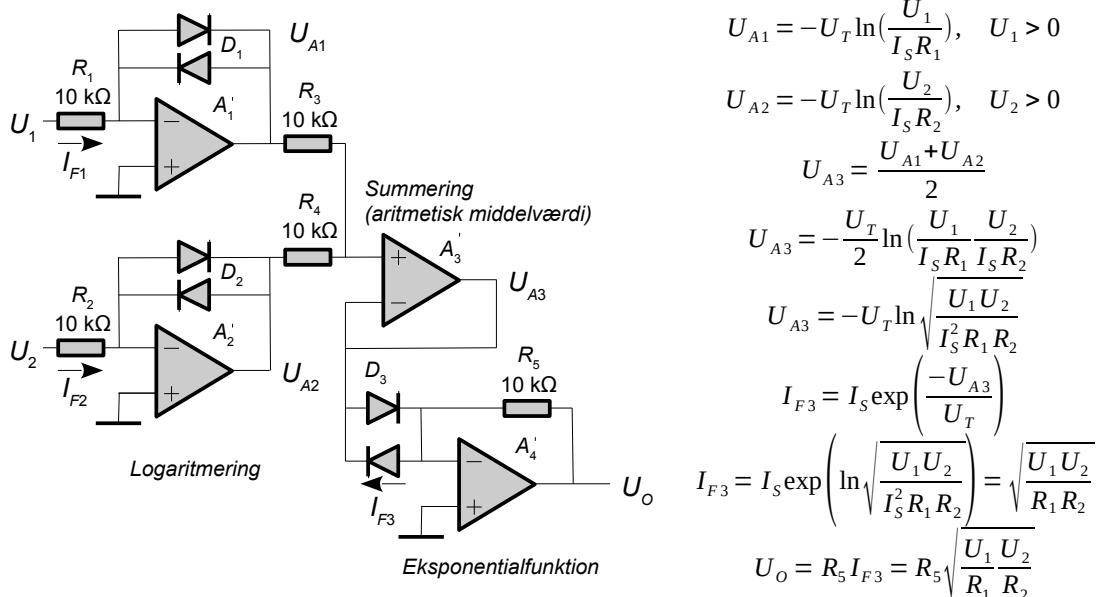
I kredsløbet herunder skal ligningen læses fra højre mod venstre for at følge funktionen. Først omsættes de to indgangsspændinger U_1 og U_2 til strømmene I_{F1} og I_{F2} ved R_1 og R_2 og de ledes gennem dioderne D_1 og D_2 for at få logaritmen til signalerne hvilket er $\ln(x_1)$ og $\ln(x_2)$ ovenfor. Den halve sum dannes ved modstandene R_3 og R_4 så U_{A3} er logaritmen til kvadratroden af produktet af indgangssignalerne.

105 Det er et drejespoleinstrument med indbygget ensretter så instrumentet slår ud alt efter hvor kraftigt signalet niveauet er. Området er fra +3 dB og ned til -20 dB og i den professionelle verden er der krav til instrumentets dæmpning (reaktionstid).

106 Se *Circuit Idea* i "Electronics World", July/August 1996, side 615.

107 Jeg benyttede et lignende kredsløb professionelt omkring 1995 til normering af en analog vektor i et apparat hvor der ikke var mulighed for en mikroprocessor. Ved normering beregnes længden af en vektor (x, y, z) som kvadratroden af $x^2 + y^2 + z^2$ og de enkelte koordinater skal så divideres med denne værdi for normering. Det fungerede godt i den aktuelle applikation, men dioderne skal være meget nær ens så der blev benyttet en integreret kreds med det nødvendige antal dioder integreret på samme chip.

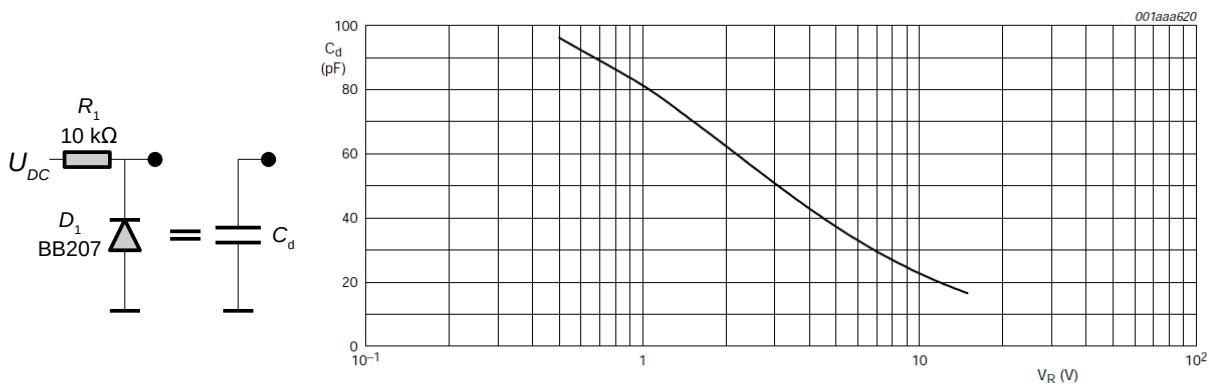
Spændingen U_{A3} konverteres til strøm I_{F3} af eksponentialfunktionen ved D_3 og føres gennem modstanden R_5 så udgangen bliver en spænding. For ideelle dioder vil I_S være den samme og den kan forkortes væk. Der benyttes ekstra dioder undervejs for en veldefineret opførsel ved negative indgangssignaler. Ved U_1 eller U_2 negativt opnås division med den negative indgang og ved to negative indgangssignaler vil begge signaler være reciprokke.



Figur 224 – Beregning af den geometriske middelværdi af to analoge signaler ved brug af dioder for logaritmering af signalerne. Division ved negativt indgangssignal er dog forkert tæt ved nul.

5.1.23 Varicap

Den spærrende PN overgang kan opfattes som to ladninger der er adskilt af et ganske tyndt lag. Det er selve definitionen på en kondensator så den spærreforspændte diode er karakteriseret af en kapacitet i spærreretningen. Værdien er ikke ret høj for en småsignaldiode (4 pF for 1N4148), men dioden kan konstrueres så værdien bliver betydnende og da tykkelsen af PN overgangen er funktion af spændingen over dioden kan kapaciteten ændres ved at ændre spændingen.



Figur 225 – Kapacitetsområde for dioden BB207 (NXP).

Dioder af denne type finder udbredt anvendelse i radiomodtagere hvor en mikroprocessor styrer en oscillators frekvens gennem en digital-til-analog konverter. Radiomodtageren kan da omskiftes mellem forskellige kanaler som kan være brugerens foretrukne radiostationer. Kapacitetsdioden er populær i faselåste kredsløb (*phase-locked loop*) hvor en stabil, krystalstyret oscillators frekvens kan omsættes til en anden frekvens der har samme præcision som krystaloscillatoren.

Note 9.

5.2 Lysdiode

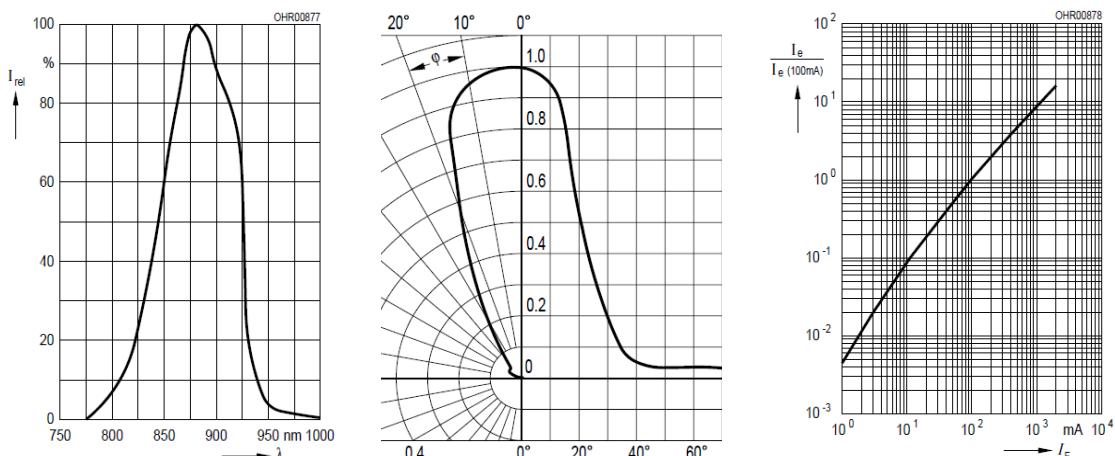
Når strømmen løber i halvlederen vil en elektron accelereres af spændingsfaldet, og den kan få nok energi til at slå en elektron fra et af atomerne op til en højere bane omkring kernen. På et eller andet tidspunkt falder elektronen tilbage igen, og den frigør da den oplagrede energi i form af en foton, der alt efter halvledermaterialet vil svare til en bølgelængde i det infrarøde, synlig eller ultraviolette område. Tilsvarende kan lysenergi slå en elektron fri af halvlederen, og hvis det sker i PN overgangen vil det elektriske felt trække elektronen mod de positive ioner i N halvlederen.

5.2.1 Lysemitterende diode, LED

Dioder af III-V typen er mest velegnede for lysudsendelse med høj effektivitet, men også dioder af silicium udsender lys, dog i den infrarøde ende og med lav effektivitet. Bølgelængden er relateret til spændingen over dioden i lederretningen, for elektronerne skal opnå en vis energi ved acceleration gennem spændingsfaldet, og de violette og ultraviolette lysdioder kræver en relativt høj energi¹⁰⁸.

Farve:	Infrarød	Rød	Gul	Grøn	Blå	Violet	Ultraviolet
Bølgelængde (nm):	> 760	610 – 760	570 – 590	500 – 570	450 – 500	400 – 450	210
Spænding U_F (V):	< 1,6	1,6 – 2,0	2,0 – 2,2	1,9 – 2,5	2,5 – 3,7	2,8 – 4,0	3,1 – 4,4
Eksempler:	GaAs	GaAsP	AlGaInP	AlGaP	ZnSe	InGaN	AlN

Et eksempel vises herunder for den infrarøde diode SFH485 af GaAlAs, der giver maksimalt lys ved 880 nm. Lysudsendelsen dækker fra 840 nm til 925 nm ved 50 % af maksimumværdien. Den midterste kurve viser at dioden ikke udsender lyset ligeligt i alle retninger for den vil kun udsende lysenergi indenfor $\pm 20^\circ$ fra diodens akseretning. Til højre vises strålingsintensiteten som en funktion af strømmen i dioden og der er en næsten lineær sammenhæng. Dioden kan arbejde kontinuerligt på 100 mA ved 25°C og der tillades kortvarige pulser op til 1 A ved 20 μ s. Dioden vil gå ON eller OFF indenfor 0,5 μ s regnet mellem 10 og 90 % niveauerne.



Figur 226 – Den spektrale fordeling af lysudsendelsen (venstre), retningskarakteristikken (midte), og den relativt udsendte effektæthed som funktion af strømmen (højre) for SFH485 (Osram).

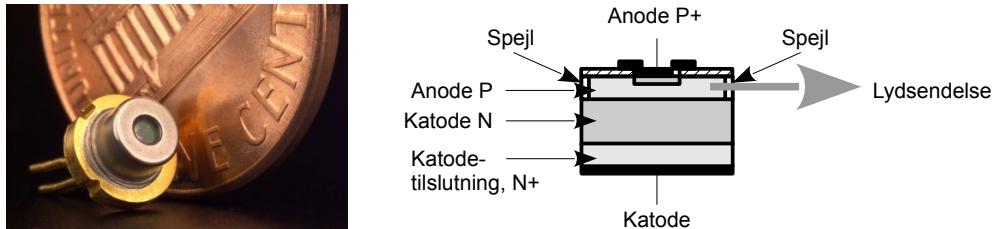
Strålingsintensiteten for SFH485 er 25 ... 160 mW/sr ved 100 mA hvor sr står for steradian. Det betyder at dioden belyser en flade med 25 ... 160 mW/m² i 1 m afstand¹⁰⁹. For omregning til en anden afstand benyttes afstandsreglen der angiver en reduktion på kvadratet af forøgelsen af afstanden. Diodens strålingsintensitet bliver derfor 1 ... 6,4 mW/m² i 5 m afstand. For at relatere til noget mere kendt svarer kraftigt sollys på en skyfri sommerdag til 1 kW/m² så effektætheden fra lysdioden er ikke farlig for øjet i modsætning til den koncentrerede lysstråle fra en laserdiode.

108 http://en.wikipedia.org/wiki/Light-emitting_diode.

109 En steradian er defineret som det vinkelrum, der afgrænsner arealet 1 m² af overfladen af en kugle med radius 1 m. En kugle med radius 1 m har et areal på 12,56 m², så 1 sr udgør 0,0795 af vinkelrummets 4 π .

5.2.2 Laserdiode

Einstein beskrev i 1917 muligheden for stimuleret lysudsendelse og ti år senere digtede Tolstoj en roman om et panservåben baseret på "Dødsstrålen". Einsteins påstand blev sandsynliggjort 1939 af Fabrikant og Townes opnåede forstærkning af mikrobølger i 1953, men det var først i 1960 at Maiman konstruerede en fungerende laser baseret på et krystal af rubin¹¹⁰.



Figur 227 – En laserdiode opbygger et kraftigt lysniveau ved at reflektere lyset mellem indbyggede spejle. Derved optimeres den samtidige udsendelse af fotoner, der er grunden til laserens virkning.

Efter sin fremkomst var laseren en parentes indenfor videnskaben for den havde intet umiddelbart anvendelsesområde. Det blev populært sagt at "laseren var svaret, men spørgsmålet manglede". Anvendelserne er i dag bredt favnende. Ved høje effekter benyttes en CO₂ laser til at skære i stål og i mindre skala benytter kirurgen en laser til at skære i væv og samtidigt dermed koagulere de små blodårer så blødningen mindskes. Ved datakommunikation sendes lyspulser fra en laser ind i plast- og glastråde og ved grundforskningen benytter laseren for nedkøling af atomer. Vi benytter i dagligdagen laseren til at hente data ud af en DVD og håndværkeren måler afstande med nogle få millimeters præcision med et håndholdt apparat fra et byggefirma.

Laserdioder kan i skrivende stund konstrueres fra 405 nm ved en InGaN halvleder der giver blåt lys som bruges i Blue-ray HD DVD afspillere og op til 3330 nm der dannes ved en GaInAsSb halvleder og det er langt inde i det infrarøde bølgelængdeområde.

En laser benytter stimuleret emission af lys der er en kvantemekanisk effekt. En anslået elektron vil af sig selv falde tilbage til udgangspositionen under udsendelse af energi, men den kan bringes til at falde tilbage samtidigt med at en anden foton passerer. Det betyder at få fotoner kan udløse en lavine af fotoner. For at opnå virkningen bygges en laserdiode med spejle på to parallelle flader som lyset sendes frem og tilbage mellem. For en lysfrekvens hvor afstanden mellem spejlene er et multiplu af en bølgelængde vil der opstå resonans og et kraftigt lysniveau kan opbygges i laseren. Det ene spejl er delvist gennemsigtigt så en del af den stående bølge kan passere som laserlys og fotonerne har da samme frekvens og svinger i takt¹¹¹.

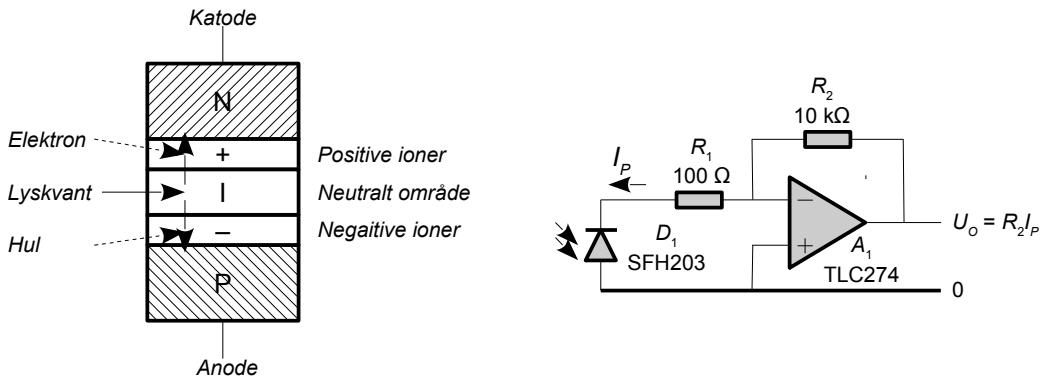
Laseren muliggør optiske systemer hvor lyset deles i to stråler ved endnu et delvist reflekterende spejl, og de to stråler kan igen blandes efter at have passeret forskellige emner eller strækninger, og de vil nu interferere. Hvis den ene lysstråle er forsinket et helt antal bølgelængder i forhold til den anden vil de to stråler adderes, mens de vil subtraheres hvis den ene lysstråle er bagefter med et ulige antal halve bølgelængder. Lysstyrken varierer som en cosinus-funktion til fasen mellem de to lysstråler, og det gør det muligt at bygge optiske systemer, hvor en variation i afstand kan måles med en opløsning på en brøkdel af lysets bølgelængde; det er en opløsning på omkring 1 nm.

110 <http://www.nbi.ku.dk/hhh/fysikkens/spinoff/laseren/>.

111 http://en.wikipedia.org/wiki/Laser_diode.

5.2.3 Fotodiode

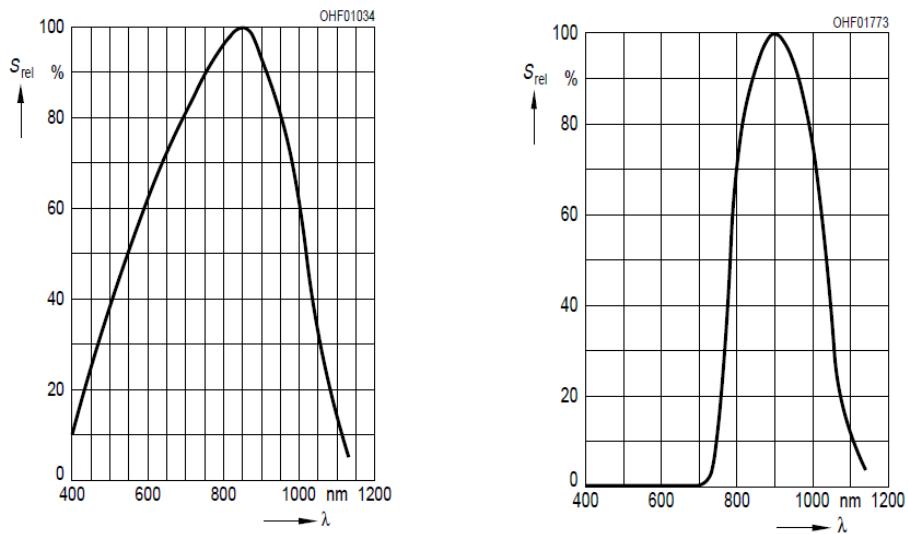
Alle halvledere er lysfølsomme og fotodioden er designet det til at omsætte lys til elektrisk strøm. Anvendelsesområderne omfatter datakommunikation, fjernmåling af temperatur og omsætning af solens lys til elektrisk energi ved solpaneler. Kort beskrevet kan et lyskvant slå en elektron fri fra et af siliciumatomerne og det elektriske felt i PN overgangen trækker derefter elektronen mod de positive ioner i N-halvlederens grænselag. Elektronerne løber ud ved katoden, men strømretningen er defineret modsat elektronens retning så strømmen løber diodens pileretning.



Figur 228 – I en fotodiode kan en foton slå en elektron fri fra siliciumatomet så der skabes både en elektron og et hul. Den elektriske feltstyrke i PN overgangen tiltrækker elektronen mod katoden og hullet med anoden og en operationsforstærker kan omsætte strømmen til en spænding.

En diode beregnet for lysdetektion har et stort halvlederområde for at optimere lysoptagelsen og dette område er ofte uden dotering hvilket kaldes for en PIN-diode hvor I står for *intrinsic* der er det oprindelige materiale uden dotering. En silicium fotodiode reagerer på elektromagnetisk stråling fra cirka 400 nm op til cirka 1100 nm der er et stykke inde i det infrarøde område. Øjet er følsomt fra 400 nm (violet) til 800 nm (rød).

Herunder vises responsen¹¹² for fotodioden SFH203.

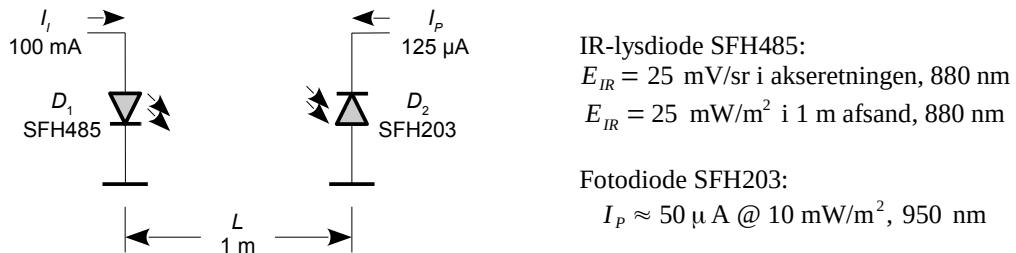


Figur 229 – Den spektrale følsomhed for SFH203 uden daglysfilter (venstre) og SFH203F med et filter der dæmper den synlige del af spektret (højre). Dioden er mest følsom omkring 850 nm der ligger på grænsen mellem den røde farve og det infrarøde spektrum (Osram).

¹¹² Ved et belysning fra en glødelampe vil dioden SFH203 give cirka 80 μA ved 1000 lux svarende til lyset under en bordlampe og ved direkte sollys på cirka 10⁵ lux er strømmen på omkring 8 mA. Dioden kan leveres med et filter der dæmper den synlige del af lysets spektrum, men bemærk at glødelamper som fx halogen har et betydeligt niveau af varmestraaling i det infrarøde område.

Dioden vil producere en strøm på $I_P = 50 \mu\text{A}$ ved energitætheden 10 W/m^2 og 950 nm . Den maksimale følsomhed er dog ved 850 nm hvor dioden kan forventes at give mest strøm så for et korrekt resultat skal spektret for IR-lymdioden ganges med den spektrale følsomhed for fotodioden SFH203 og integreres over det fælles bølgelængdeområde. Mindre kan dog gøre det hvis man kan leve med en betydelig usikkerhed for beregningen.

Den tidligere viste IR-lydiode SFH485 har en effekttæthed på mindst 25 mW/m^2 i 1 m afstand ved strømstyrken 100 mA . Fotodioden SFH203 er specificeret ved 10 mW/m^2 , men belysningsniveauet med SFH485 ved den angivne strøm er 2,5 gange højere så fotodioden vil producere en strøm på $125 \mu\text{A}$. Som nævnt er beregningen ikke præcis da både sender- og modtagerdiode specificeres til et større område end -50% til $+100\%$ for hver diode.

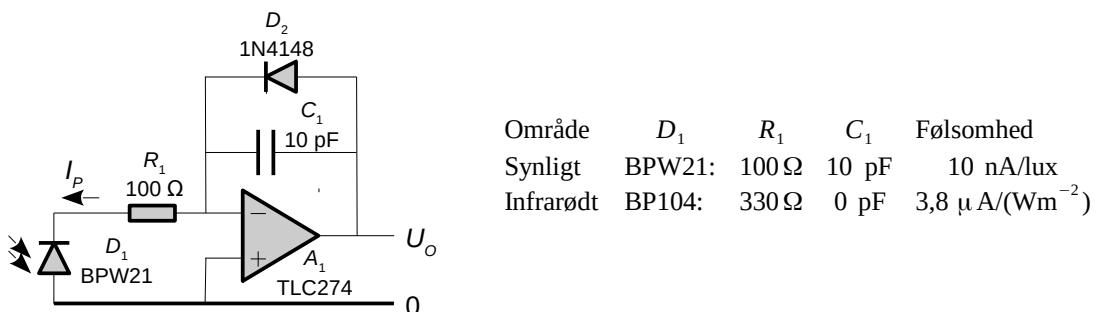


Figur 230 – Lyset fra den infrarøde senderdiode SFH485 modtages af fotodioden SFH203 i en afstand på 1 m .

Effekttæheten specificeres i steradianer så den reduceres med kvadratet på en ændring i afstanden. Det betyder at en fordobling af afstanden reducerer effekttæheten til en fjerdedel.

5.2.4 Lysmåler

Kredsløbet herunder har jeg brugt i en håndholdt lysmåler der var i stand til at vise et stort område på et drejespoleinstrument uden behov for omskiftning mellem måleområder¹¹³.



Figur 231 – Lysmåler, hvor den modtagne lyseffekt omsættes i en strøm og logaritmieres. De viste dioder svarer til det synlige lysspektrum med BPW21 og det nær-infrarøde med BP104.

1N4148 benyttes som logaritmisk konverter der klemmer et område på fem dekader ind på en enkelt skala. Spændingen på udgangen går med BPW21 fra 350 mV ved netop synligt lys (1 lux), til 650 mV ved direkte sollys (10^5 lux). Det giver en variation på cirka 60 mV/dekade . Med dioden BP104 måles infrarødt lys fra $850 \dots 1050 \text{ nm}$. Modstanden R_1 og kondensatoren C_1 kan have behov for justering så opstillingen bliver stabil med den anvendte operationsforstærker.

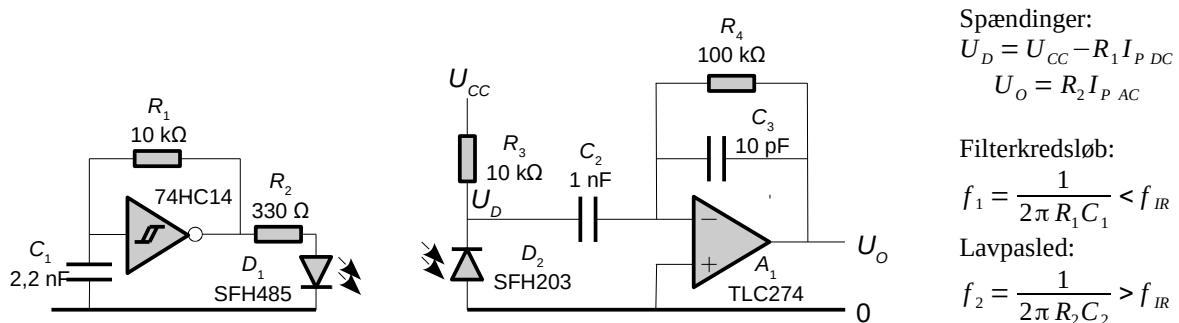
Hvis dioden 1N4148 erstattes med en modstand omsættes strømmen lineært til spænding. Det er en meget udbredt måde at bygge en lysmodtager på og det er idéen bag det næste eksempel.

113 Min artikel "Optoelectronics by Design" publiceret i Electronics World and Wireless World, May 1994, side 365.

5.2.5 Lysdetektor

Et populært kredsløb indenfor den industrielle elektronik fjerner indvirkningen fra omgivelserne ved at benytte en pulset IR-lysdiode. Ideen er at omgivelsernes lys, der ofte kaldes for DC lys, normalt varierer så langsomt at det kan fjernes med et højpasfilter.

Herunder dannes det pulsede lys fra en oscillator på 50 kHz som driver 10 mA til en IR-lysdiode. En fotodiode modtager den pulsede strøm og genererer en firkant på udgangen med den samme frekvens og en amplitude som er funktion af den modtagne signalstyrke.



Figur 232 – Lysdetektor hvor de modtagne lyspulse opfanges og omsættes til spænding.

Fotodiodens DC strøm optages af R_3 mens AC strømmen fra den pulsede IR-lysdiode føres videre til operationsforstærkeren gennem C_2 og giver et bidrag over R_4 . For at kunne klare omgivelsernes belysningsniveau skal R_3 designes så spændingen U_D ikke kommer så langt under nul at dioden D_2 bliver ledende. Man bør designe efter at holde niveauet over nul.

Grænsefrekvensen af højpasfiltret dannet af R_3 og C_2 skal tillade pulsen at nå helskindet igennem til operationsforstærkeren hvilket kan antages opfyldt når grænsefrekvensen er tre gange under oscillatorenens frekvens. Båndbredden ved lavpasfiltret dannet af R_4 og C_3 skal være høj nok til ikke at dæmpe pulsen for meget og det kan antages opfyldt når grænsefrekvensen er mindst tre gange oscillatorenens frekvens. Dertil har C_2 som opgave at stabiliserer operationsforstærkeren på grund af fotodiodens indre kapacitet.

Ved praktisk design bør man være opmærksom på at støj fra effektforsyningens U_{CC} kobles direkte til den inverterende indgang på operationsforstærkeren og at den forstærkes med R_4/R_3 . Det er derfor nødvendigt at afkoble effektforsyningen for at holde højfrekvens støj nede i niveau. Som første step bør man montere en kondensator på 47 nF til 220 nF mellem U_{CC} og nul og hvis det ikke er nok kan man supplere med 470 μF og måske også indskyde 10 Ω for at lavpasfiltrere effektforsyningen.

Ved industriel brug er der ofte flere sensorer i brug samtidigt. Det kan for eksempel være udstyr for en vaskehæld, en garagedør eller et transportbånd. For at adskille de pulsede signaler fra de forskellige sensorer benyttes individuelle frekvenser for hver sensor og der kan diskrimineres mellem de enkelte sensors ved en mikser og et lavpasfilter som korrelere det modtagne signal med oscillatoren der pulser IR-lysdioden (se side 204).

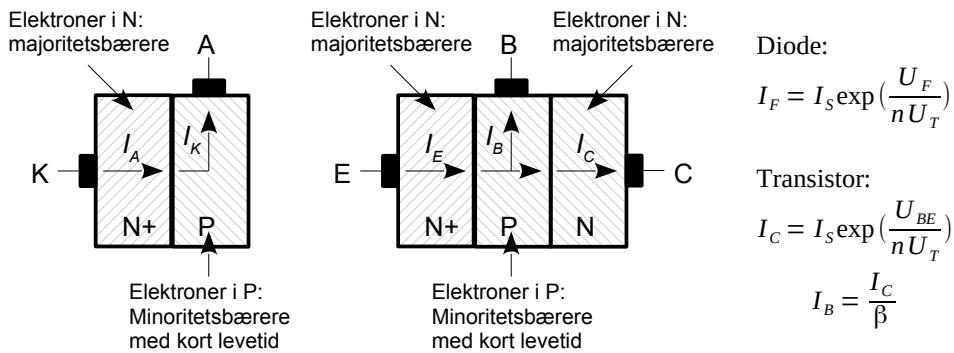
Note 10.

5.3 Transistor ved DC

Den bipolære transistor er en udbygning af dioden ved tilføjelse af endnu et lag halvleder som opsamler de udsendte elektroner fra katoden. Det betyder at den bipolære transistor bliver en "trebenet diode" hvor spændingen over diodens anode og katode (basis og emitter) varetager styringen af strømmen i dioden (kollektor og emitter). Dioden indgår i en JFET felteffekt transistor hvor den gennem et elektrisk felt i den spærreforspændte diode styrer strømmen i kanalen og i en MOSFET hvor diodens relation beskriver funktionen ved svag strømstyrke (*sub threshold*).

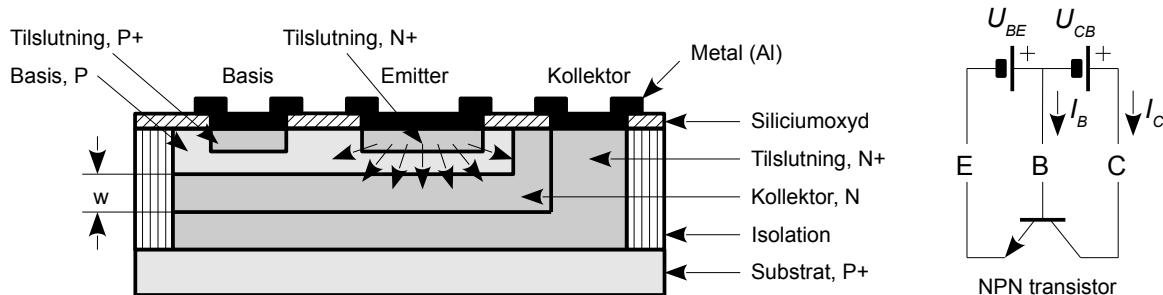
5.3.1 Den bipolære transistor

I dioden udsendes elektronerne fra en kraftigt doteret katode (N+) til en svagere doteret anode (P). En ekstern spændingskilde mindsker det indbyggede potentiale og muliggør elektrondiffusion over PN overgangen. Efter en kort strækning i P halvlederen rekombineres alle elektroner med huller så det eksterne batteri sender det samme antal elektroner ind ved katoden som huller ved anoden for at opretholde ladningskoncentrationerne. Indtrængningsdybden er i mikrometerområdet og angiver den strækning hvor 63 % af elektronerne rekombineres med huller. Ved ti gange dybden kan alle elektroner derfor antages rekombineret med huller. Det er denne træghed i rekombinationen der er helt afgørende for at forstå transistorens funktion.



Figur 233 – Den bipolære transistor fremstilles ud fra en diode ved at tilføje endnu et lag.
Eksemplet viser en NPN transistor med emitter (N+), basis (P) og kollektor (N). Elektronerne bevæger sig i den viste retning og basislaget skal være tyndt for at opnå den ønskede funktion.

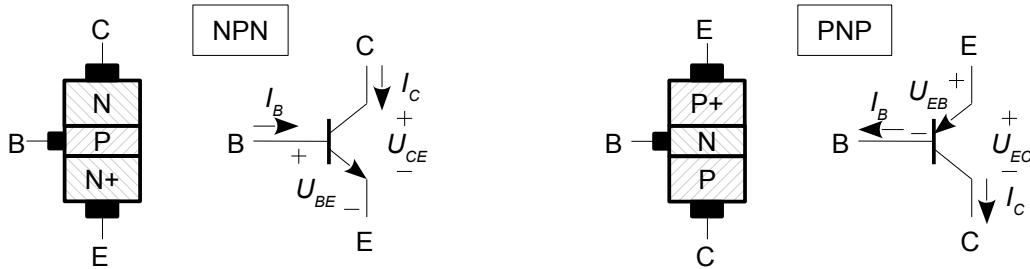
I en transistor udvides dioden i basis-emitter med endnu en N halvleder som danner kollektoren. Dioden opfører sig som en almindelig PN overgang hvor en ekstern spænding over basis-emitter styrer diffusionen af elektroner fra emitter til basis, men basislaget er væsentligt tyndere end indtrængningsdybden så hovedparten af elektronerne når at diffundere videre til kollektor. Kun de ganske få elektroner der "falder i et hul" i basislaget danner en basisstrøm.



Figur 234 – Den bipolære epitaxial-transistor består af to PN overgange hvor P-lagets tykkelse w er så ringe at elektronerne fra emitteren kan diffundere igennem basis til kollektor med minimalt tab fra rekombination i basis. Tilslutningerne benytter stærkt doterede områder for at opnå en god elektrisk kontakt (P+ og N+) mens basis og kollektor er svagt doterede (P og N).

5.3.2 NPN og PNP transistor

Der er to måder PN overgangene kan arrangeres på. Den netop beskrevne type kaldes for NPN da den benytter to negativt doterede lag omkring en positivt doteret halvleder. En PNP benytter den modsatte polaritet af lagene. Teorien bag NPN og PNP er den samme blot erstattes emissionen af elektroner fra emitteren i en NPN med en emission af huller i en PNP. Begge typer transistor er i brug ved de integrerede kredse og diskret opbyggede effektforstærkere.

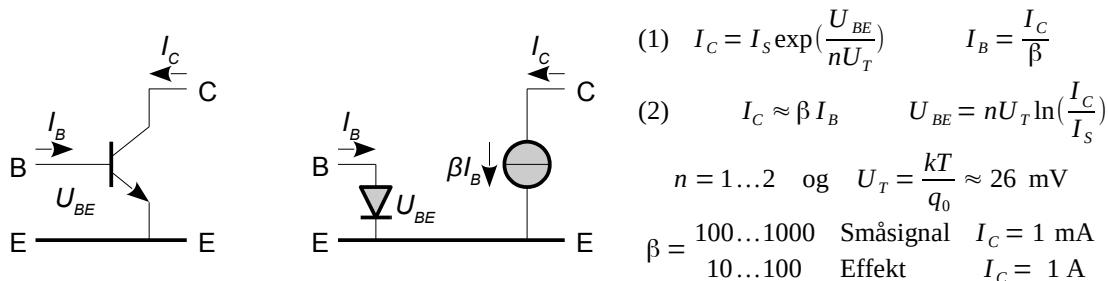


Figur 235 – To typer bipolære transistorer. Pilen i symbolot angiver strømretningen (modsat elektronernes bevægelsesretning) og viser samtidig den elektrode der kaldes emitter.

I praksis er der ikke den store forskel imellem de to typer ud over strømmens retning og et lidt højere spændingsfald over basis-emitter for en PNP ved den samme strømstyrke som for en NPN. De to typer kaldes derfor for komplementære transistorer. En PNP transistor er i principippet mindre egnet til højfrevens for diffusionsstrømmen bæres af huller der har en op til fire gange lavere mobilitet end elektroner så ladningstransporten foregår lidt langsommere.

5.3.3 DC model

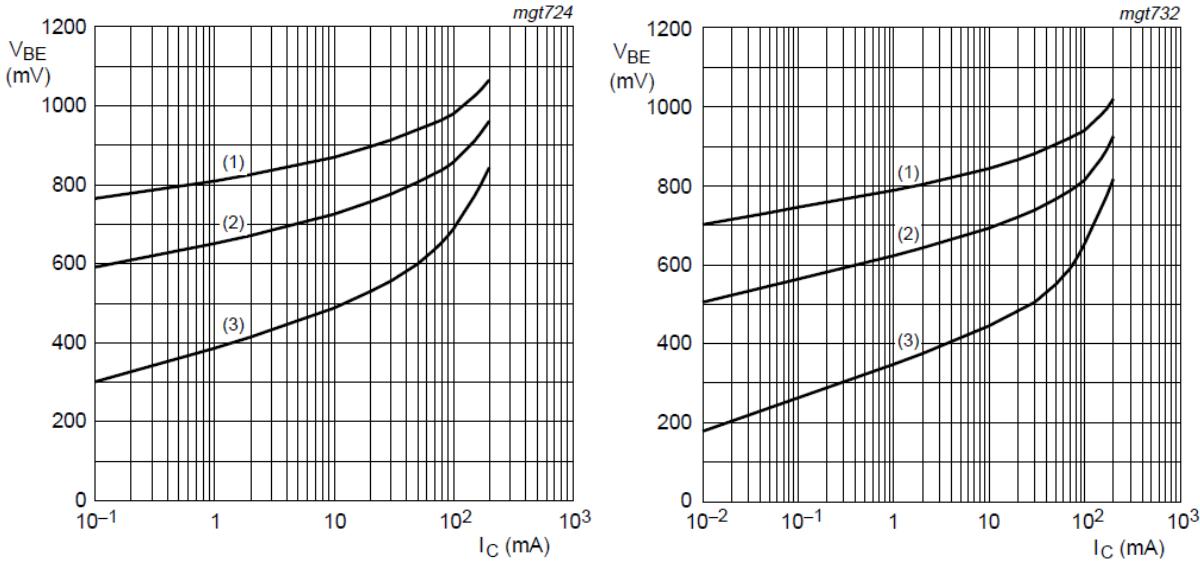
En transistor er en to-port så der er en indgangsport ved basis og emitter og en udgangsport ved kollektor og emitter. Strømmen i transistorens kollektor I_C beskrives med samme ligning som for dioden, dog nu med strømmen løbende i kollektor-emitter med den styrende spænding U_{BE} over dioden i basis-emitter. Strømmen kan derfor reguleres i ét kredsløb som funktion af spændingen i et andet og kun emitteren er fælles for de to kredsløb. Strømmen i basis I_B fra rekombinationen i det tynde basislag betyder at for hver elektron der rekombineres med et hul er der β gange så mange elektroner som når frem til kollektor så basisstrømmen er relativt lav.



Figur 236 – En model for den bipolare transistor med tilhørende matematisk fortolkning som en spændingsstyret strømkilde (1) eller en strømstyret strømkilde (2).

Strømmen i kollektor kan udtrykkes som en funktion af strømmen i basis med transistorens strømforsærkning β som proportionaliteten mellem de to. Spændingen over basis-emitter bliver nu en funktion af strømniveauet så tommelfingerreglen om 0,7 V over basis-emitter er en tilnærmelse. Denne model er velegnet til en DC analyse af et kredsløb som det skal vises på side 222.

Spændingsfaldet over basis-emitter dioden vises herunder for den populære transistor BC547 der også findes som BC847 for overflademontage. Som det ses stiger strømmen i kollektor-emitter efter en eksponentialfunktion ved en lineært stigende spænding over basis-emitter.



Figur 237 – Spændingsfaldet over basis-emitter for BC547A (venstre) og BC547C (højre). Tallene refererer til temperaturen med (1) for -55°C , (2) for 25°C og (3) for 150°C (NXP).

At der er tale om en småsignaltransistor ses ved at karakteristikken krummer opad allerede ved en strømstyrke over 10 mA. Ved højere strømstyrke bliver elektrontrætheden så høj at der ikke kan ses bort fra den gensidige påvirkning mellem elektronerne. Det er *current crowding* som betyder at der skal relativt mere spænding til for at drive strømmen i kollektor. Under denne grænse varierer spændingen over basis-emitter med 60 mV for en dekade ændring i strømmen (ved 25°C).

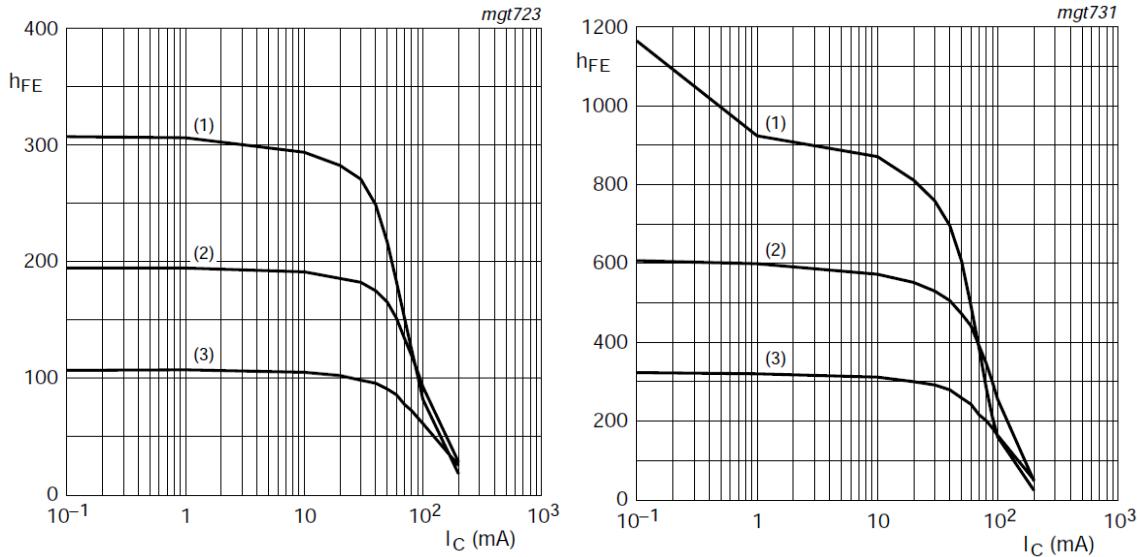
Mætningsstrømmen I_s (*saturation current*) afhænger af PN overgangens geometri og doteringens profil og strømmen fordobles for hver 10°C temperaturstigning. Værdien er cirka 10 fA for BC547 ved 25°C , men det er ikke en værdi man kan finde i databladet. Mætningsstrømmen optræder som faktor til den eksponentielle relation og sammen med idealitetsfaktoren n giver den de to parametre som behøves for at beskrive relationen mellem spændingen over basis-emitter U_{BE} og strømmen i kollektor I_C . En lidt fræk tolkning er at I_s og n blot er parametre for en *curve-fitting* til transistorens funktion så den eksponentielle relation kommer til at passe. Der er dog en solid teoretisk baggrund for de to parametre.

Transistoren er temperaturfølsom på grund af mætningsstrømmen og temperaturspændingen og det udtrykkes under ét ved transistorens temperaturkoefficient. Ved en strøm i kollektor på 1 mA kan det aflæses fra kurverne at spændingen over basis-emitter aftager med 0,43 V ved en ændring i temperaturen fra -55°C til 150°C og det kan omregnes til temperaturkoefficienten $-2,1 \text{ mV}/^{\circ}\text{C}$ der udgør en typisk værdi for dioder og transistorer ved dette strømniveau. Ved 100 μA er ændringen 0,47 V så temperaturkoefficienten bliver på $-2,3 \text{ mV}/^{\circ}\text{C}$ og ved 10 μA er den oppe på $-2,5 \text{ mV}/^{\circ}\text{C}$.

Transistoren temperaturkoefficient kan betragtes fra to synsvinkler. Hvis strømmen i kollektor skal holdes konstant så skal spændingen over dioden i basis-emitter reduceres med godt $2 \text{ mV}/^{\circ}\text{C}$. Det er den udfordring man står med ved DC koblede kredsløb der udgør hovedparten af al elektronik i dag og løsningen er at benytte to komponenter der arrangeres så de gensidigt kompenserer for hinanden. Den anden synsvinkel er at ved en konstant spænding over basis-emitter opnås en eksponentielt stigende strøm i kollektor ved stigende temperatur. Det udnyttes i fx 7800-serien af spændingsregulatorer hvor en transistor lukker ned ved så hård belastning at chip-temperaturen kommer over 150°C .

Idealitetsfaktoren n beskriver et kompromis mellem de mekanismer der styrer strømmen i en transistor og den sættes til én i megen litteratur, men som ved dioden er værdien relateret til transistorens konstruktion. Mine egne målinger viser at BC547 har en værdi tæt på én.

Transistorenstrømforstærkningen β er af historiske årsager opdelt i tre selekteringer med A for den laveste værdi, B for den mellemste og C for den højeste. Det skyldes den store variation i denne parameter som en konsekvens af at den skyldes den uundgåelige rekombinationen af elektroner og huller i basis. Uden en selektering er variationen mellem yderværdierne almindeligvis 1:8, men ved at vælge en af selekteringerne begrænser variationen til 1:2. Strømforstærkningen afhænger af arbejdspunktet og temperaturen som illustrationen herunder viser. Det er en tilnærmelse at beskrive den som konstant hvilket dog hyppigt antages i den populærvidenskabelige litteratur.

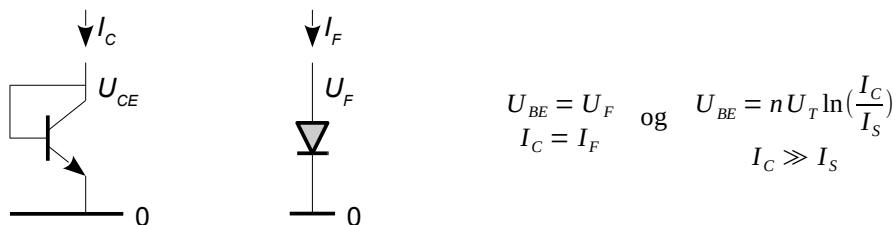


Figur 238 – Strømforstærkningen for BC547 og BC847 i A selektionen til venstre og C selektionen til højre. Tallene refererer til temperaturen med (1) for 150°C , (2) for 25°C og (3) for -55°C (NXP). Det viste knæk under 1 mA for BC547C er ikke kommenteret i databladet og har ingen fysisk årsag.

Temperaturkoefficienten af strømforstærkningen ses at være af størrelsesordenen $1\text{ }%/^{\circ}\text{C}$ så den sædvanlige antagelse om en konstant værdi af strømforstærkningen holder ikke og man er nød til at tage hensyn til det så et design ikke bliver alt for temperaturfølsomt.

5.3.4 Diodekobling

Transistoren kan kobles som en diode ved at kortslutte basis og kollektor. Den vil opføre sig helt som en diode. Diodekoblingen er en sjælden gæst i diskret analog elektronik, men den er flittigt benyttet i integrerede kredse da den kan fremstilles i samme proces som transistorerne. En vigtig anvendelse er ved strømspejle som omtales side 258.

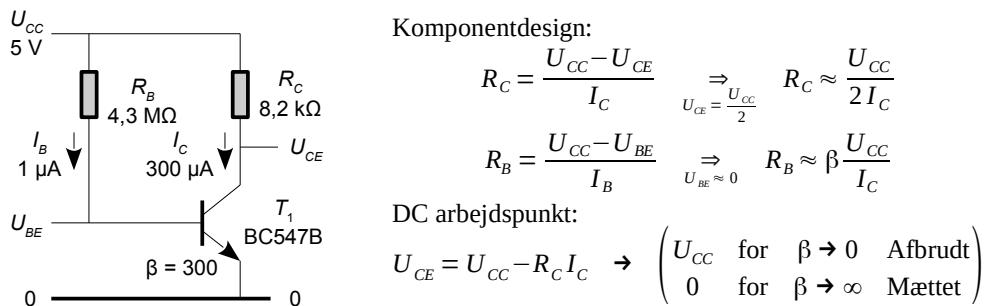


Figur 239 – En diode-koblet transistor opfører sig fuldstændig som en diode.

Den diodekoblede transistor bryder sammen ved cirka 6 V modsat polaritet mod mindst 30 V for en almindelig diode. Sammenbruddet er ikke destruktivt ved moderat strøm, men det rapporteres at transistoren eksponentielle karakteristik påvirkes af et sammenbrud. Kredsløb for eksponentiel og logaritmisk konvertering samt operationsforstærkere med lav egenstøj eller fejlspænding vil derfor ofte indeholde en diode over basis-emitter strækningen for at forhindre et sammenbrud.

5.3.5 DC arbejdspunkt

Opstillingen herunder fremhæves i mange elektronikbøger som et typisk eksempel på et design af en forstærker med en transistor. Som det vil fremgå af det følgende har jeg mine forbehold overfor den udlægning. Jeg har aldrig set den anvendt professionelt og forklaringen er den enkle at det er et ualmindeligt dårligt design der let kan forbedres.

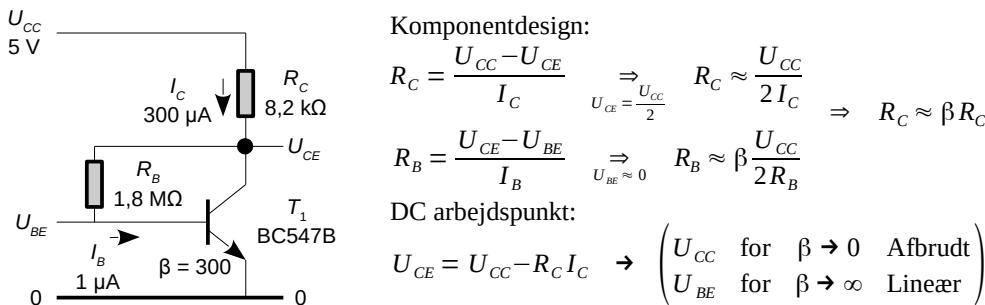


Figur 240 – En simpel forstærker der har problemer med usikkerheden ved strømforstærkningen. Signalet der skal forstærkes kobles ind på basis gennem en kondensator for at blokere for DC og det forstærkede signal udtages fra kollektor.

Tanken er at modstanden R_B i basis dimensioneres så strømmen i kollektor på $I_C = \beta I_B$ sætter U_{CE} til den ønskede værdi. Et typisk valg er halvdelen af effektforsyningen for at give mulighed for den størst mulige amplitude af signalet. Det er sjeldent mulig at udmåle strømforstærkningen så man er henvist til databladets angivelse der for BC547B giver $\beta = 200 \dots 450$ ved strømmen $I_C = 2 \text{ mA}$ og $\beta = 150$ typisk ved $I_C = 10 \mu\text{A}$. Spørgsmålet er så hvilken værdi der gælder ved det aktuelle kredsløb med strømstyrken $I_C = 300 \mu\text{A}$ og det spørgsmål kan desværre ikke besvares.

Hvis strømforstærkningen bliver for stor vil transistoren trække kollektor ned mod nul hvor den kortsluttes og en mulig AC forstærkning mistes. Hvis designet er baseret på en typisk værdi af strømforstærkningen (her $\beta = 300$) så bliver strømmen i kollektor fra -33% til $+50\%$ af planlagt og udgangsspændingen er et sted fra $3,4 \text{ V}$ ned til $1,3 \text{ V}$. Det er en variation på $2,1 \text{ V}$ på grund af usikkerheden i strømforstærkningen. Det kan måske fungere, men det tillader ikke en usikkerhed med temperaturen. En stigning fra 25°C til 75°C vil hæve strømmen i kollektor med cirka 50% og det kan bringe transistoren i mætning (se side 223).

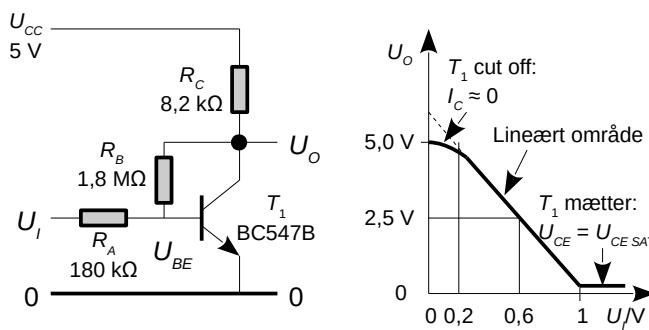
Et bedre design opnås ved at hente basisstrømmen fra kollektor. Ved samme strømforstærkning som ovenfor vil spændingen på kollektor variere fra $2,9 \text{ V}$ ned til $2,1 \text{ V}$ der giver $0,8 \text{ V}$ usikkerhed. Nok så væsentligt er det at designet garanterer at transistorens arbejdspunkt aldrig kan nå ned på nul så den vil altid være aktiv. Selv om strømforstærkningen går imod uendelig vil kollektoren ikke komme under cirka $0,6 \text{ V}$ så der vil mindst være $\pm 0,5 \text{ V}$ til rådighed for signalet.



Figur 241 – En kobling hvor strømforstærkningen har mindre betydning for DC arbejdspunktet som følge af tilbagekoblingen fra kollektor til basis. Forstærkeren er fælles emitter (**common emitter**) fordi emitter er fælles for ind- og udgang.

5.3.6 DC forstærkning

Transistoren kan fungere som DC forstærker. En indgangsspænding U_I omsættes til en strøm ved modstand R_A og adderes til strømmen fra R_B for derved at variere spændingen U_O på kollektor. Når indgangssignalet hæves vil strømmen i kollektor stige og udgangsspændingen falder.



Basisstrøm:

$$I_B = \frac{U_I - U_{BE}}{R_A} + \frac{U_O - U_{BE}}{R_B}$$

Udgangsspænding:

$$U_O = U_{CC} - R_C \beta I_B$$

$$U_O = \frac{\frac{R_B}{\beta R_C} U_{CC} + (1 + \frac{R_B}{R_A}) U_{BE} - \frac{R_B}{R_A} U_I}{1 + \frac{R_B}{\beta R_C}}$$

$$U_O \approx 5.9 \text{ V} - 5.8 U_I$$

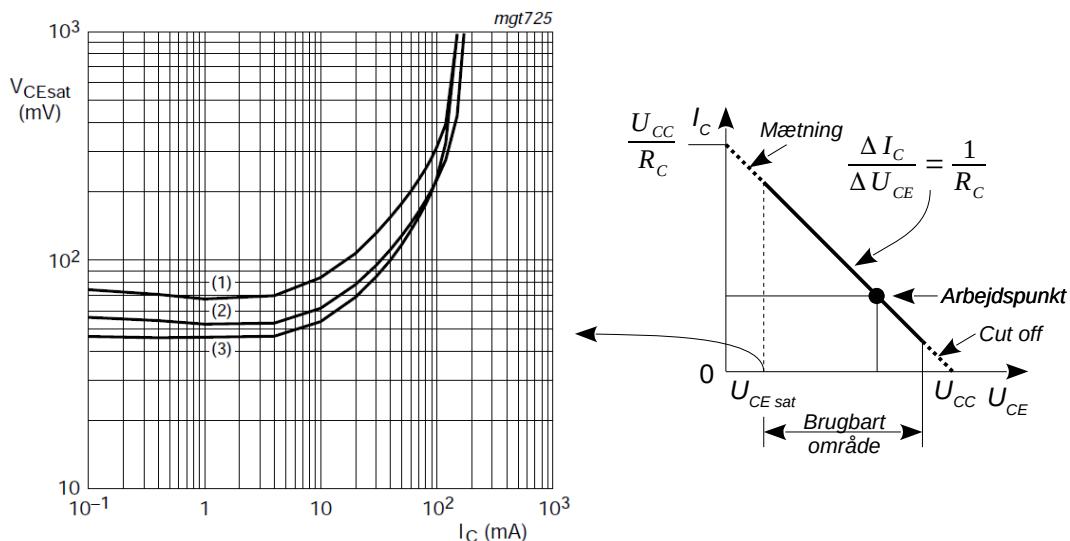
Figur 242 – En transistor kan forstærke et DC indgangssignal.

Udtrykket er ikke blandt de simpleste, men med komponenternes værdier indsatt giver det en ret linje med hældningen -5.8 . Det betyder at forstærkningen er mindre end forholdet R_B/R_A og det skyldes nævnerens værdi på 1.73 . Beregningen er kun gyldig ved moderate værdier af R_A for hvis modstanden sættes til nul angiver formlen en forstærkning på uendelig. Træerne vokser dog ikke ind i himlen og forstærkningen har en øvre værdi på omkring 100 som vist i afsnittet på side 233.

Den faktiske karakteristik runder af ved en kollektor nær ved 5 V hvor strømmen i kollektor bliver nær nul og hverken β eller U_{BE} er konstante. Udgangsspændingen kan ikke nå helt ned på nul da transistoren mætter når kollektors potentielle er tæt på emitterens.

5.3.7 Mættet transistor

Transistoren kan ikke styre forløbet når strømmen er begrænset af det ydre kredsløb. Transistoren siges da at være mættet (saturated) og relationen $I_C = \beta I_B$ holder ikke mere.



Figur 243 – Tallene repræsenterer temperaturen med (1) for 150°C , (2) for 25°C , og (3) for -55°C (BC547, NXP). Til højre ses strømmen i kollektor som funktion af spændingen på kollektor.

Transistoren er aktiv så længe U_{CE} er over cirka 0.1 V ved det strømområde som transistoren er designet til (cirka 1 V for effekttransistorer).

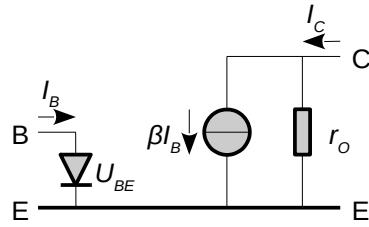
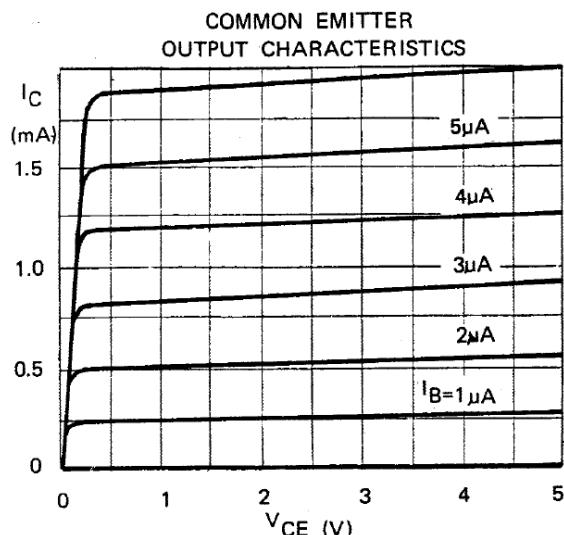
Kurven viser at transistorens aktive område ophører ved en kollektor-emitter spænding på 50 mV ved 1 mA og 25°C idet spændingen ikke kan blive lavere. For en strøm over 10 mA stiger den med strømmen og en mættet BC547 kan derfor beskrives ved en spændingskilde på 50 mV i serie med en modstand på 5 Ω. Det er typiske værdier for småsignaltransistorer.

Det er almindeligt at databladene specificere spændingen over transistorens kollektor-emitter ved en basisstrøm på $I_B = I_C/20$, men det betyder ikke at der er en strømforstærkning på 20 gange ved mætning. Det betyder blot at transistoren får rigeligt med basisstrøm. Transistoren svarer dermed til en kortslutning hvor strømmen er styret af de eksterne komponenter og ikke af transistoren. En konsekvens af mætning er at transistoren reagerer langsomt¹¹⁴.

Når transistoren ikke er mættet kan den arbejde med en spænding på kollektor der er laver end basis for så vidt at strømmen i basis-kollektor dioden ikke er problematisk for applikationen. Det udnyttes rutinemæssigt i eksponentielle og logaritmiske konvertere og ved operationsforstærkere for en effektforsyning i nærheden af 1 V.

5.3.8 Udgangskarakteristik

Strømmen i transistorens kollektor er så godt som uafhængig af spændingen over transistoren når spændingen over kollektor-emitter er stor nok. Som kurven herunder viser kan transistoren fungere ned til cirka 50 mV over kollektor-emitter ved en strøm under 2 mA. Det betyder at dioden i basis-kollektor er ledende forspændt, men ikke trækker en betydende strøm.



Parallelmodstand for $U_{CE} > 0,1$ V:

Aflæsning af kurve for $I_B = 6 \mu A$:

$$r_o = \frac{dU_{CE}}{dI_C} = \frac{5 \text{ V} - 0,5 \text{ V}}{2 \text{ mA} - 1,88 \text{ mA}} = 38 \text{ k}\Omega$$

Figur 244 – Udgangskarakteristikken for BC107 familien¹¹⁵ viser strømmen i kollektor som funktion af spændingen over kollektor-emitter med strømmen i basis som parameter (Micro Electronics).

Når transistoren er aktiv vil strømmen i kollektor kun afhængig ganske lidt af spændingen over kollektor-emitter og det kan tolkes som om der var en modstand i parallel med transistorens udgang. I de fleste situationer er modstandens værdi så høj at man kan se bort fra den og blot antage at transistoren kan beskrives analytisk som en ideel strømgenerator.

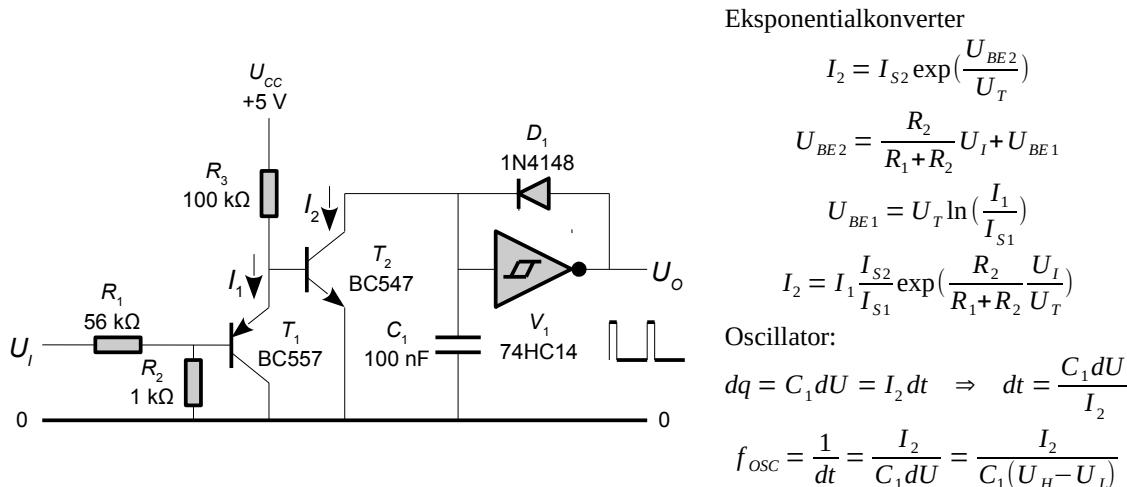
114 Den ældre digitalteknik benyttede både afbrudt og mættet som de to binære tilstande med den omkostning at kredsløbene var langsommere sammenlignet med dagens standard. TTL teknologien (Transistor-Transistor Logic) benytter mættede transistorer for at opnå veldefinerede logiske niveauer. Det tager dog tid at trække en transistor ud af mætning for der er et overskud af elektroner i dioden mellem basis og emitter og det skal fjernes før transistoren kan slukkes. Hastigheden ved TTL var begrænset til 25 MHz. En variation baseret på Shottky dioder udnytter at dioden i basis-kollektor bringes i ledende retning ved lav værdi på udgangen så transistoren optager sin egen basisstrøm og ikke mættes. Det blev udnyttet til at opnå samme hastighed som ved TTL, men ved et lavere effektforbrug. Ved ECL teknologien (Emitter-Coupled Logic) er transistorene altid i det lineære område og hastigheden kom over 100 MHz. Mikroprocessorer med MOSFET og CMOS teknologier er også baseret på mættede transistorer, men en felt-effekt transistor har ikke en problematisk rumladning og er derfor hurtigere end de bipolare. Hastigheden kan nå op i GHz området.

115 Databladet angives at passe for BC107-BC109, BC167-BC169, BC237-BC239 og BC317-BC319 og som fodnoten side 247 viser kan data forventes at passe med BC547-BC549.

5.3.9 Eksempel på en eksponentielt spændingsstyret oscillator

Den analoge musiksynthesizer blev introduceret omkring 1920, og blev brugt af en del klassiske komponister, men det var først omkring 1970 at den slog igennem ved jazz og rock. Et betydende musikinstrument er Minimoog hvor tonegeneratorer og filter blev styret af elektriske spændinger med en eksponentiel karakteristik på 1 oktav/volt. Det vil sige at frekvensen steg til det dobbelt ved én volt spændingsstigning. Det musiske interval en oktav svarer til frekvensforholdet 2:1 og de mellemliggende trin på 1/12 volt svarer til de tolv halvtoneptrin i den ligesværende temperatur som hovedparten af vestlig musik har anvendt siden 1800-tallet¹¹⁶.

Herunder vises et eksempel på en eksponentialkonverter der driver en oscillator. Der er ikke tale om en præcisionsoscillator for en musiksynthesizer, men eksemplet indeholder de basale dele. For en bedre oscillator anbefales det at søge på nettet efter diagrammerne for en af de populære musiksynthesizere selv om man i dag nok vil foretrække en DSP for tonegenerering¹¹⁷.



Figur 245 – En eksponentialkonverter, der omsætter en spændingsændring på indgangen på 1 V til en fordobling af strømmen i kollektor. Udgangen kan drive en frekvensdelerkæde som HEF4040.

Oscillatoren består af en Schmitt-trigger koblet som relaxations-oscillator hvilket blot betyder at dens udgang "står og hviler" på nul volt i det meste af tiden mens en kondensator aflades af en konstant strøm. Schmitt-triggeren er kun aktiv i det korte tidsrum fra spændingen er nået under den lave tærskelværdi U_L og udgangen går høj til kondensatoren er opladt til den høje tærskel U_H . Frekvensen kan beregnes af kondensatorligningen ved C_1 for strømmen I_2 og differensen mellem de to tærskelværdier (for 74HC14 er $U_H - U_L = 0,85$ V).

Strømmen genereres af transistor T_2 hvor transistor T_1 buffer indgangssignalet og kompenserer for temperaturens indvirkning på basis-emitter spændingsfaldet. Da transistor T_2 spændingsstyres stiger strømmen eksponentielt med styresignalet. Spændingsdelen med R_1 og R_2 dæmper et indgangssignal på 1 V til 18 mV der svarer til en fordobling af strømmen i kollektor på T_2 .

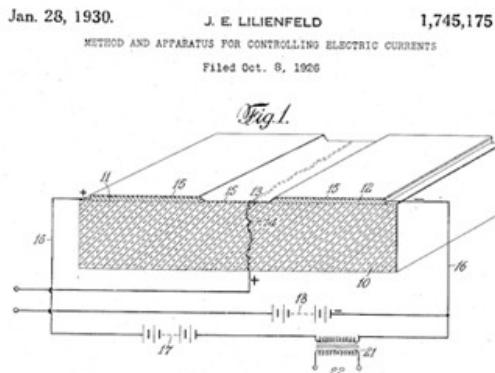
Som det ses af analysen er strømmen I_2 til oscillatoren en skalering af strømmen I_1 der er givet af forsyningsspændingen U_{CC} og modstanden R_3 til $44 \mu\text{A}$ med diagrammets komponentværdier. De to transistorers mætningsstrømme optræder som en fast faktor og er ikke funktion af temperaturen. Antages transistorerne ens (på nær forskellen i polaritet) bliver forholdet på én og oscillatorenens frekvens kan estimeres til $f_{osc} \approx 500$ Hz ved $I_2 = 44 \mu\text{A}$ og nul volt på indgangen. Frekvensen varieres eksponentielt fra 31 Hz ved $U_I = -4$ V op til 8 kHz ved $U_I = 4$ V. Her er strømmen nået op på 0,7 mA og der vil afsættes en effekt i T_2 på små 2 mW som varmer transistoren $0,4^\circ\text{C}$ op og det er nok til at frekvensen driver hørbart.

116 http://en.wikipedia.org/wiki/Analog_synthesizer.

117 Arp, Moog, Roland, Oberheim, Yamaha.

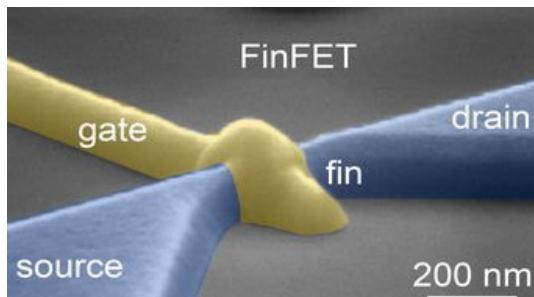
5.3.10 Felteffekt transistoren

Den felteffekt transistor som Lilienfeld præsenterer i 1925 genopstår i 1960 som en JFET og det leder cirka ti år senere til den populære MOSFET der danner fundamentet for moderne analog og digital signalbehandling^{118,119}.



Figur 246 – Julius Edgar Lilienfeld patenterer 1925 en konstruktion baseret på kobbersulfid der kan styre en elektronstrøm.

Felteffekt transistorens navn angiver at det er et elektrisk felt der styrer elektronstrømmen og det var netop Lilienfelds idé at erstatte elektronrørets sårbare glaskolbe med en robust konstruktion der var fri for opvarmning fra en kortlivet glødetråd. Idéen var at ledningsevnen af en kanal blev styret af en spænding.



Figur 247 – I en FinFET er dimensionerne er så små at den mest effektivt kan styres ved at lade elektroden dække flere sider af kanalen.

Den moderne udgave af Lilienfeldts transistor er en MOSFET hvor strømmen i halvlederens kanal styres af et elektrisk felt uden at involvere en PN overgang. Der er delte meninger om Lilienfeld fik transistoren til at virke og det leder til en interessant debat om emnet. Kommentaren herunder refererer desuden til de problemer Bell Lab havde med at navigere udenom patentet¹²⁰.

"[Lilienfeld] created his non-tube device around 1923, with one foot in Canada and the other in the USA, and the date of his Canadian patent application was October 1925. Later American patents followed, which should have been well known to the Bell Labs patent office. Lilienfeld demonstrated his remarkable tubeless radio receiver on many occasions, but God help a fellow who at that time threatened the reign of the tube."

- Harry E. Stockman Sercolab, 1981 letter to Wireless World.

På trods af den store betydning vil omtalen blive kort, men det skyldes kun at en stor del af mekanismerne allerede er blevet introduceret ved den bipolare transistor. Den følgende gennemgang vil derfor fokusere på det der er specielt ved "fætteren".

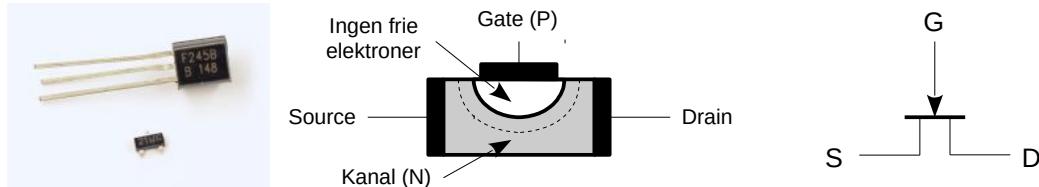
118 http://ethw.org/A_Very_Early_Conception_of_a_Solid_State_Device.

119 www.ece.umass.edu/ece/ece344/MOSFET_talk.ppt.

120 <http://www.edn.com/electronics-blogs/edn-moments/4422371/Lilienfeld-patents-field-effect-transistor>.

5.3.11 JFET

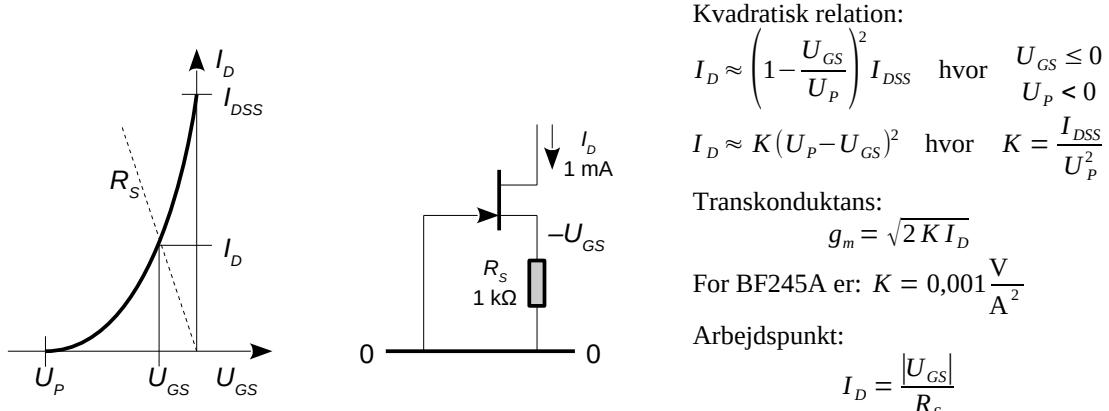
Den tidligste felteffekt transistor var en JFET der står for *junction field-effect transistor*. Den har tre elektroder der benævnes *gate* for styreelektroden, *source* for strømkilden og *drain* for opsamlingen af dem. Kanalen mellem source og drain er i mange tilfælde symmetrisk og de to elektroder kan da ombyttes. Hvis det ikke er tilfældet tegnes pilen nærmest ved source.



Figur 248 – En junction felteffekt transistor (JFET) er dannet ved en kanal mellem de to elektroder (Source og Drain) hvis effektive tværsnit kan påvirkes ved et elektrisk felt fra en ydre spænding påtrykt den midterste elektrode (Gate). Til højre vises symboler for en JFET af typen N-kanal.

Gate er dannet ved diffusion af en PN overgang til kanalen så der er to JFET typer alt efter om kanalen er dannet af P eller N type halvleder. Strømmen styres af det elektriske felt over den spærrede diode så der en kapacitet mellem gate og transistorens kanal på omkring 5 pF for en almindelig småsignaltransistor. Dioden er fremstillet af silicium så den vil opføre sig som enhver anden siliciumdiode med cirka 0,6 V i lederretningen. Det er dog usædvanligt at benytte dioden i lederretningen for det er idéen at strømmen skal styres af et elektrisk felt og ikke gennem injektion af ladningsbærere som i den bipolære transistor. Der flyder kun en svag strøm i spærretretningen som for BF245 er under 10 pA for den typiske specifikation¹²¹.

Alle transistorer er spændingsstyrede strømregulatorer og ved en JFET er kanalen ledende hvis der ikke foretages noget for at lukke af for strømmen. Da dioden skal være spærrende forspændet betyder det at gate skal være negativ i forhold til source og drain for at skrue ned for strømmen (modsat for en P kanal JFET). Det kaldes på engelsk for en *depletion mode* transistor da der skal en gate-spænding til for at kanalen bliver tømt (*depleted*) for elektroner.

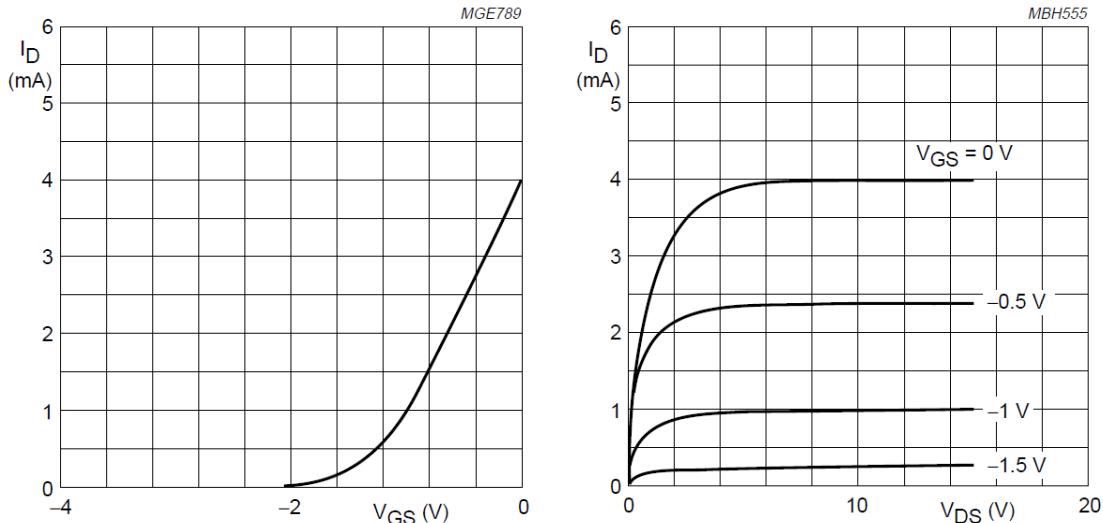


Figur 249 – En JFET anvendt som en strømgenerator med strømmen indstillet af en modstand.

En klassisk opstilling med JFET benytter *auto-biasing* hvor en modstand i source giver en negativ gate-source spænding. Strømmen i drain indstilles så spændingen på source svarer til et punkt på transistorens karakteristik hvor gate-source spændingen bliver lige præcis så negativ at det svarer til strømmen i drain. Opstillingen kan benyttes som strømgenerator og den danner grundlag for opbygning af en AC forstærker som vist på side 248.

Herunder vises en JFET karakteristik eksemplificeret ved BF245A.

121 En JFET var tidligere populær som diode med lav lækstrøm hvor man kortsluttede drain og source og benyttede gate som anode.



Figur 250 – Karakteristik for BF245A der er en N-kanal JFET. Gate-source spændingen er negativ så transistoren indstilles til 1 mA med gate på cirka -1 V overfor source. Kurven til højre viser at en felteffekt transistor er en spændingsstyret modstand ved lav spænding og strømbegrænsende for en drain-source spænding over et par volt.

BF245 får i tre selektioner. B selektionen har værdierne $U_P = -4$ V (for $I_D = 0$) og $I_{DSS} = 10$ mA (for $U_{GS} = 0$). Værdierne er $U_P = -8$ V og $I_{DSS} = 17$ mA for C-selektionen. Som ved den bipolære transistor er der en stor variation i parametrene så strømmen I_{DSS} kan variere mere end 1:2 indenfor selektionen og spændingen for pinch off varierer tilsvarende adskillige volt. Aflæst fra udgangskarakteristikken ses at for $U_{GS} = -1$ V er $I_D = 0,5$ mA ved $U_{DS} = 0,2$ V svarende til at transistoren er en modstand på $R_{DS} = 400 \Omega$ mens $U_{GS} = -0,5$ V giver $I_D = 1$ mA ved $U_{DS} = 0,1$ V svarende til $R_{DS} = 100 \Omega$.

Til højre visers illustrationen udgangskarakteristikken for BF245 der er strømmen i drain som en funktion af drain-source spændingen med gate-source spændingen som parameter. Ved $U_{GS} = 0$ er gate og source kortsluttet og strømmen vokser med spændingen indtil 4 mA der opnås ved 6 V. Ved højere spænding er strømmen konstant hvilket svarer til en strømregulator. Ved en spænding under cirka 0,1 V er der tilnærmedesvis proportionalitet mellem spænding og strøm og det svarer til en modstand. En JFET kan derfor anvendes som en spændingsstyret modstand.

Den matematiske relation mellem strømmen I_D og spændingen U_{GS} er ikke så veldefineret som for den bipolære transistor. Den tilnærmes som regel med en kvadratisk relation hvor U_P er den værdi hvor strømmen bliver nul som følge af at kanalen klemmes sammen. Det er transistorens *pinch-off voltage* og I_{DSS} er kanalens mætningsstrøm (*saturation current*) når der ikke er nogen spænding på gate-source strækningen ($U_{GS} = 0$). I integrerede kredse benyttes en JFET af og til med kortsluttet gate og source som en strømkilde hvor strømstyrken besluttes på designtidspunktet.

Felteffekt transistorens funktion er ikke baseret på rekombination af elektroner og huller i kanalen, som ved den bipolære transistors diode i emitter-basis, så en JFET er relativt hurtig med en -3 dB specifikation på 700 MHz for signalforstærkningen med BF245. Den er populær i radiokredsløb i UHF båndet på grund af gode støjegenskaber, men den almindelige JFET er langsomt på vej ud som følge af den succes der tilfaldet den moderne MOSFET^{122,123}.

I en N kanal JFET er ladningsbæreren elektroner i modsætning til en P kanal JFET hvor strømmen er båret af huller. Elektronerne bevæger sig i ledningsbåndet hvorimod hullerne er elektroner der hopper rundt mellem positionerne for atomernes valenselektroner. Det betyder at en P kanal JFET som udgangspunkt er lidt langommere end en ellers tilsvarende N kanal JFET på samme måde som der gælder for NPN versus PNP.

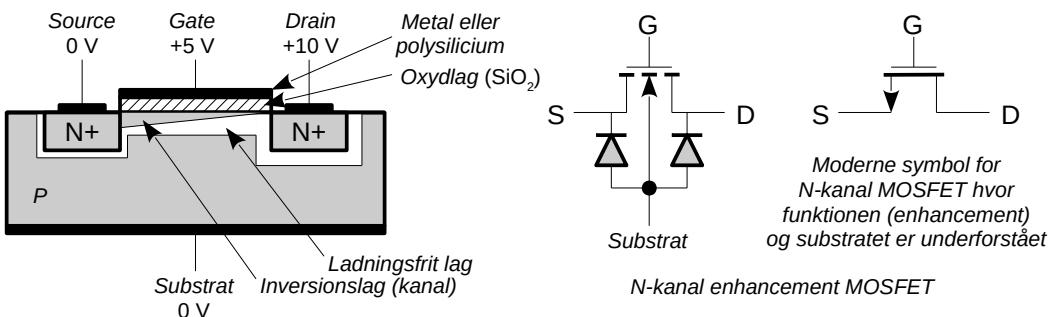
122BF245 (2N3819) fra omkring 1970 er specificeret til en afskæringsfrekvens på 700 MHz ved -3 dB.

123BF245 udgår i TO-92 på grund af vigende afsætning, men den produceres fremover som BF545 for overlademontage.

5.3.12 MOSFET

Den moderne udgave af felteffekt transistoren opbygger gate-elektroden som et metallag med et tyndt lag siliciumoxyd som isolator mellem gate og kanal. Det giver en MOSFET hvor betegnelsen står for *metal-oxide silicon field-effect transistor*. Uden en diode i gate spærres for DC strøm under alle forhold og isolationen er så god at strømmen i gate er under 1 pA så indgangsmodstanden er ekstremt høj. Tykkelsen af oxydlaget er omkring 1 μm så feltstyrken mellem gate og kanal kan give et gennemslag som vil ødelægge transistoren. Spændingen fremgår af databladet og skal som en tommelfingerregel holdes indenfor ± 20 V. Gate dannes af aluminium eller polysilicium hvilket er et amorf siliciumlag der kan dannes ovenpå siliciumoxyd og som er mere stabilt end aluminium. Transistoren benytter ikke rekombination mellem elektroner og huller så den er relativt hurtig.

Den mest populære MOSFET er *enhancement mode* transistoren der ved N-kanal kræver positiv spænding på gate for at trække strøm. Omskiftningen mellem ON og OFF kan programmeres ved geometrien og doteringen og er cirka 2 V for styring direkte fra en mikroprocessor eller lavere ved lavspændingselektronik. Værdien kaldes for *threshold voltage* $U_{GS(th)}$ og den er opgivet ved en forholdsvis lav strøm, men der er store variationer mellem databladene. For at åbne helt op for en større strøm skal der almindeligvis være 5 V mellem gate og source.



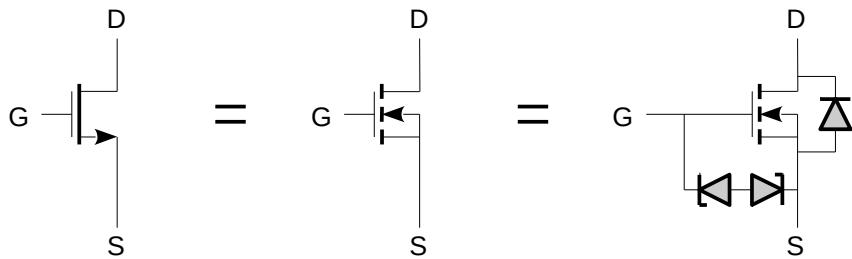
Figur 251 – En N-kanal enhancement MOSFET skaber kanalen fra drain til source ved en positiv spænding på gate i forhold til drain gennem forskydning af ladningen i et tyndt område lige under elektroden for gate.

En P-kanal MOSFET skabes ved at bytte om på N og P ved fremstillingen. Det er nu hullet der er aktiv som ladningsbærer i kanalen hvis gate-source spændingen er tilstrækkelig negativ.

En MOSFET opbygges fra et substrat (*substrate* eller *bulk*) hvor drain og source dannes som to stærkt doterede områder for at opnå god elektrisk kontakt med et overskud af termisk genererede ladningsbærere. De to områder danner dioder med substratet og de er spærrende forspændt ved normal brug. Når gate er positiv overfor substratet skabes et ladningsfrit lag umiddelbart under den isolerende siliciumoxyd hvor hullerne fra P halvlederen skubbes væk af det elektriske felt. Ved en tilstrækkelig høj spænding tiltrækkes overskydende elektroner fra source (og drain) og laget skifter polaritet. Det kaldes da for inversionslaget (*inversion layer*). De frie elektroner danner en kanal mellem source og drain som kan lede en strøm. Spændingen på gate styrer antallet af frie elektroner i inversionslaget og den styrer derfor også modstanden af kanalen. En MOSFET udgør derfor en modstand ved lav drain-source spænding.

Når drain er positiv overfor source vil spændingsforskellen mellem gate og kanal kun være høj når ved source og den aftager fra source til drain så inversionslaget aftager i tykkelse mod drain. Hvis U_{DS} overstiger $U_{GS} - U_{GS(th)}$ kan inversionslaget ikke dannes for den del af kanalen der er nærmest ved drain og kanalen bliver klemt sammen (*pinch off*). Strømmen af elektroner diffunderer gennem den ellers ladningsfri del af halvlederens kanal drevet af det elektriske felt fra drain. Strømmens styrke er givet af den resterende kanal i inversionslaget nær ved source og den afsnørede del af kanalen er nu blot et spændingsfald som accelererer elektronerne, men der løber ikke en større strøm af den grund. Transistoren er her en spændingsstyret strømregulering.

For en MOSFET med kun tre tilledninger er substratet kortsluttet til source så der ligger en diode over transistoren med katoden til drain. Med positiv spænding på drain er dioden spærrende, men hvis polariteten vendes bliver dioden ledende. Det udnyttes i effektforstærkere og H-broer for at lede strømmen fra induktive belastninger og producenten oplyser derfor grænsedata for dioden. For brug som analog kontakt er dioden et problem når substratet ikke kan forbines mere negativt end source og drain og man må enten benytte to MOSFET i anti-serie med fx drain mod drain eller finde et par transistorer med adgang til substratet. Som et eksempel HEF4007.



Figur 252 – De populære symboler for en N-kanal enhancement MOSFET og de indbyggede dioder for beskyttelse imod elektrostatisk afladning. Dertil den uundgåelige diode over drain og source.

En MOSFET er følsom overfor elektrostatisk afladning for der brændes let hul i det tynde oxydlag under gate. Ved introduktionen af MOSFET blev den leveret med en kortslutningsbøjle der kunne fjernes efter montage i printet. Det er ikke nødvendigt i dag hvor der er indbygget beskyttelse for de forhold som transistoren udsættes for ved produktion af elektroniske kredsløb. Transistoren kan dog stadig beskadiges ved håndtering, men nu er det dioderne i beskyttelseskredsløbet der tager skade. Dioderne kan godt blive skadet uden at kortslutte og derved har transistoren en risiko for at svigte på et senere tidspunkt. Derfor bør man ikke berøre benene på transistorer der skal forlade en produktion. Det gælder i øvrigt generelt ved al elektronik og ikke kun for MOSFET.

Beskyttelsesdioderne over gate og source betyder at indgangsstrømmen til gate primært skyldes beskyttelseskredsløbets lækstrøm. Strømmen i gate er langt under 1 pA for siliciumoxyd er en god isolator, men de publicerede data ligger ved 1 nA af nævnte årsag. Det betyder også at strømmen vokser eksponentielt med temperaturen med den velkendte fordobling for hver 10°C stigning.

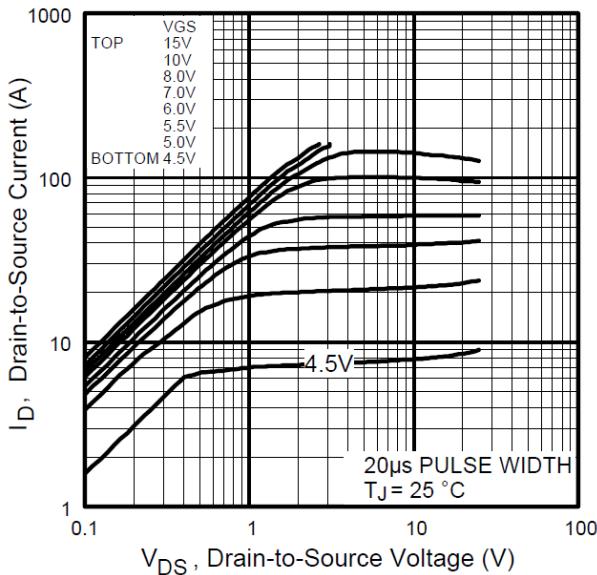
Der er en betydende kapacitet mellem gate og kanal som følge af at styringen benytter et elektrisk felt skabt mellem to adskilte ladninger. For en effekttransistor kan værdien let nå op på 1 nF og det stiller krav til driverkredsløbet hvis transistoren skal skifte hurtigt mellem ON og OFF.

Relationen mellem strømmen i drain og spændingen over gate-source minder meget om en JFET med de forskelle at pinch-off spændingen kaldes for threshold voltage $U_{GS(th)}$ og at værdien nu kan være både negativ og positiv for henholdsvis depletion og enhancement mode transistorer¹²⁴.

Ved lav drain-source spænding er strømmen i transistorens drain-source strækning proportional med drain-source spændingen når gate-source spændingen er over pinch off. Det er transistorens modstandsområde der er en vigtig parameter for en MOSFET til H-broer og DC-DC konvertere hvor den lave ON modstand reducerer effektabet.

For en spænding på gate-source under pinch off vil der stadig løbe en strøm i drain-source og denne strøm varierer eksponentielt efter samme formel som for den bipolare transistor med en hældning på 70 ... 100 mV/dekade så idealitetsfaktoren er større end én. Strømmen kaldes for en *sub-threshold current* og værdien I_{D0} for $U_{GS} = U_{GS(th)}$ er størst for MOSFET med lav værdi af pinch off. Det er et problem for en MOSFET anvendt som kontakt for den lukker ikke helt af i OFF. Ved analog signalbehandling bliver signalet ikke helt afbrudt og ved digitale kredsløb er effektforbruget ikke nul selv om den digitale komponent burde være afbrudt.

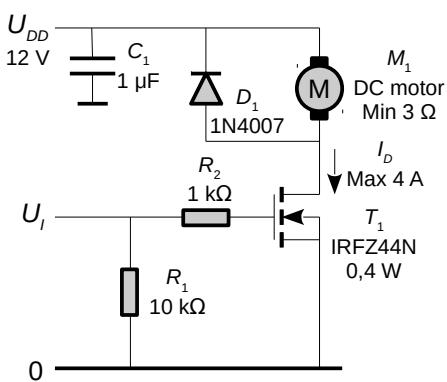
124 Se også: <http://en.wikipedia.org/wiki/MOSFET>.



Figur 253 – Udgangskarakteristikken for IRFZ44N specifiseret til $U_{GD(th)} = 2 \dots 4$ V for $I_D = 0,25$ mA.

Med $U_{GS} = 5$ V er strømmen I_D proportional med U_{DS} for en spænding indtil 0,5 V. Strømmen er $I_D = 15$ A ved $U_{DS} = 0,5$ V så modstandsværdien bliver $R_{DS} = U_{DS}/I_D = 0,03$ Ω. Ved $U_{DS} > 2$ V er transkonduktansen på $g_m = \Delta I_D / \Delta U_{GS} = 34$ S omkring 28 A hvor I_D ændres fra 20 til 37 A ved en stigning i U_{GS} fra 5 V til 5,5 V. Tilsvarende findes $g_m = 26$ S omkring 13 A hvor I_D ændres fra 7 til 20 A for U_{GS} fra 4,5 til 5 V.

Herunder vises et eksempel på en MOSFET anvendt for at starte og slukke en DC motor styret fra en mikroprocessor. En MOSFET kan kun opfattes som en ON/OFF kontakt hvis spændingen over drain-source holdes lav i ON stillingen. Karakteristikken for IRFZ44N viser at strømmen kan være indtil 4 A med 0,1 V over drain-source ved 5 V logisk høj på indgangen. Der afsættes 0,4 W som varmer transistoren op til 50°C ved normal omgivelsestemperatur når motoren kører.



Figur 254 – MOSFET transistoren IRFZ44N som kontakt for en motor styret fra en mikroprocessor.

Dioden D_1 beskytter transistoren mod høje spændinger fra motorens kommutator. Dioden skal også benyttes hvis opstillingen skal drive et relæs spole eller en lampe gennem lange ledninger. Modstand R_1 afbryder for transistoren hvis der ikke er forbindelse til mikroprocessoren eller dens udgange er afbrudt som typisk er tilfældet i power down og lige efter tilslag af effektforsyningen. Modstanden R_2 har primært til formål at beskytte mikroprocessoren ved en fejl i kredsløbet fx en defekt MOSFET der kunne trække mikroprocessorens udgang op til 12 V. Kondensator C_1 er et lokalt energireservoir som reducerer den udstrålede støj fra DC motoren. Den skal monteres tæt på både MOSFET og motor for at reducere den elektromagnetisk udstrålede støj og den indre modstand skal være lav så en aluminium elektrolytkondensator bør undgås.

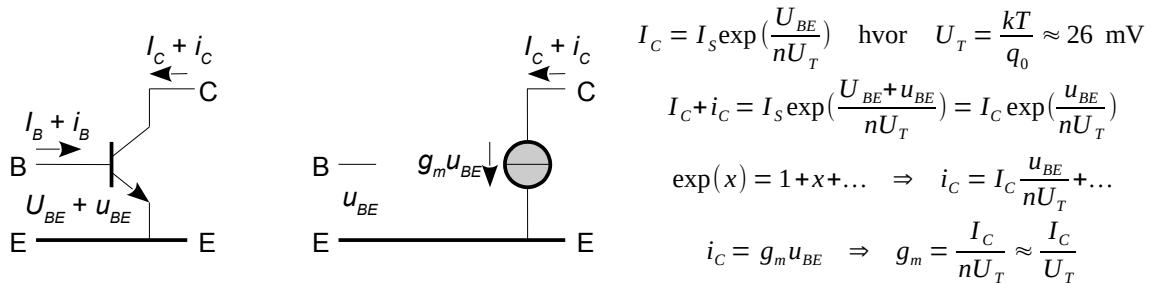
Note 11.

5.4 Transistor ved AC

Den følgende fremstilling antager at transistoren er bragt til et fornuftigt DC arbejdspunkt og vi er kun interesseret i hvordan AC strømme og spændinger varierer omkring arbejdspunktet. Set over lang tid vil signalet blot svinge med en konstant middelværdi lig med arbejdspunktet. Det vil også blive antaget at kredsløbet er lineært og det tillader superposition at blive brugt til at adskille DC og AC forholdene¹²⁵.

5.4.1 Simpel AC model

Karakteristikken for en transistor viser at strømmen i kollektor stort ikke afhænger af spændingen over kollektor-emitter (se side 224) og strømmen i transistorens kollektor kan derfor beskrives ved spændingen over basis-emitter efter eksponentialfunktionen fra side 219. For at blive brugbar som AC analyse beskrives den øjeblikkelige værdi af strøm og spænding ved summen af en DC værdi for arbejdspunktet (I_B , I_C og U_{BE}) og en AC værdi der repræsenterer signalet (i_B , i_C og u_{BE}).

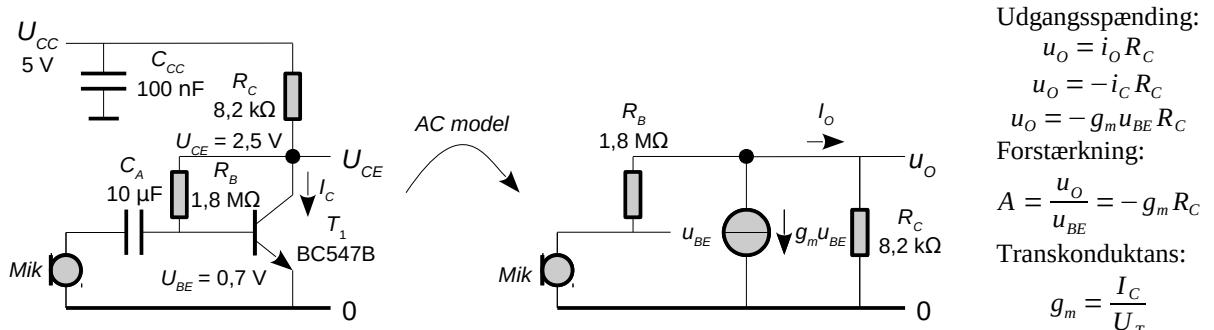


Figur 255 – En småsignalmODEL af den bipolære transistor hvor den øjeblikkelige værdi af spænding og strøm opfattes som et DC arbejdspunkt (U_{BE} og I_C) overlejret af en AC komponent (u_{BE} og i_C) der kan repræsentere et harmonisk signals amplitude eller et ikke-periodisk signal.

Proportionaliteten mellem AC strømmen i kollektor i_C som funktion af spændingen u_{BE} over dioden i basis-emitter er transistorens transkonduktans g_m der har dimension af ledningsevne (S, siemens). Det er den ubetinget vigtigste parameter i en AC model.

5.4.2 Mikrofonforstærker

Modellen skal illustreres med en mikrofonforstærker der løfter et svagt signal fra en dynamisk mikrofon op til et niveau som fx kan kobles til en mikroprocessors A/D konverter.



Figur 256 – En mikrofons signal skal forstærkes op for brug i et efterfølgende apparatur.

En mikroprocessor arbejder typisk på 5 V så det vælges som effektforsyning. Strømmen i kollektor I_C ligger ofte fra 50 μ A til 2 mA ved småforstærkere og her vælges at benytte 0,3 mA. For at kunne

¹²⁵Betegnelserne DC og AC står for *direct current* og *alternating current*, men benyttes generelt for signaler der ikke er konstante over lang tid. Her er "lang tid" ofte omkring 1 s, men ved højfrekvens hvor AC signalet er over 1 MHz, der svarer til en periode under 1 μ s, kan et langsomt varierende signal der opretholder sin værdi i måske 1 ms opfattes som DC.

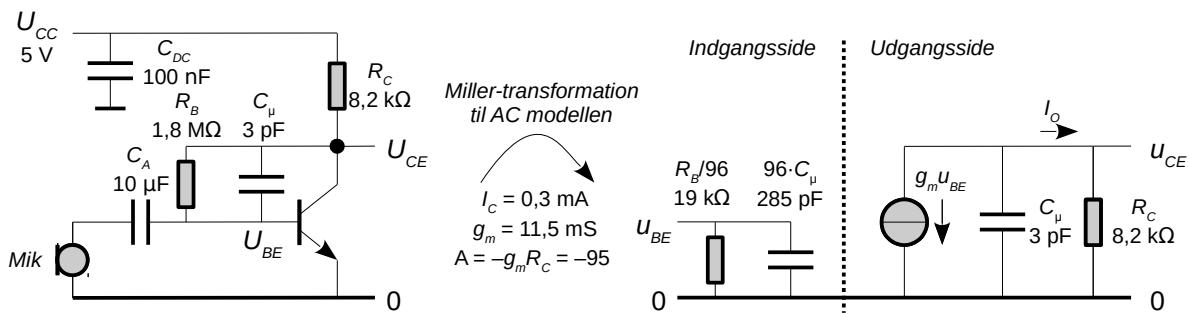
opnå maksimal amplitude på udgangen placeres DC værdien af kollektor-emitter ved halvdelen af effektforsyningens værdi ($U_{CE} = 2,5$ V) og derved kan modstanden i kollektor bestemmes fra Ohms lov. For at indstille værdien til det ønskede designes modstanden i basis ud fra basisstrømmen på $I_B = I_C/\beta = 1 \mu\text{A}$ ved en antaget værdi af strømforstærkningen på $\beta = 300$. Det giver $R_B = 1,8 \text{ M}\Omega$ hvis spændingen over basis-emitter antages til $U_{BE} = 0,7$ V. Den faktiske værdi af U_{CE} vil variere mellem 1,9 og 3,4 V på grund af usikkerheden ved strømforstærkningen.

Kredsløbet transformeres til et AC ækvivalent ved at nulstille alt der har med DC arbejdspunktet at gøre. DC spændingskilderne kortsluttes så vi har $U_{CC} = 0$ og modstanden R_C i kollektor går til nul i det ækvivalente AC diagram. Kondensator C_A er indskudt for at blokere for DC strøm så den ikke påvirker DC arbejdspunktet og den kan enten kortsluttes, som det er gjort her, eller den kan indgå i kredsløbet for en analyse af den lave grænsefrekvens.

Udgangsstrømmen løber modsat strømretningen i transistoren så forstærkeren inverterer signalet. Strømmen $g_m u_{BE}$ giver et spændingsfald over R_C der er udgangsspændingens amplitude. Egentlig bør modstanden R_B regnes som parallel med R_C , men det introducerer kun en ubetydelig ændring. En vigtig konklusion er at forstærkningen er $-g_m R_C$ for den størrelse vil dukke op igen og igen.

5.4.3 Miller transformation

Transistoren har en intern kapacitet C_μ i den spærrende diode i kollektor-basis og den begrænsrer den opnåelige bådbredde. Tilbagekoblingen gennem R_B og C_{CB} besværliggør analytisk arbejde da det er nødvendigt at opstille kredsløbsligninger for at studere virkningen. Kreative hjerner har dog udtænkt en metode til at lette arbejdet så forbindelsen mellem ind- og udgangsportene kan brydes og portene kan analyseres som separate kredsløb. Det kaldes for Miller transformationen hvor det fine navn ikke dækker over en dybere matematik, men kun lidt kredsløbsforståelse.



Figur 257 – Mikrofonforstærkeren med transistorens kapacitet C_μ i kollektor indtegnet i diagrammet.

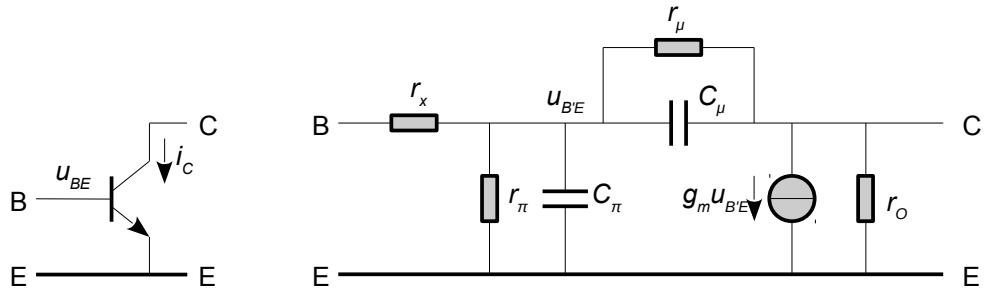
Strømmen i kollektor er $I_C = 0,3 \text{ mA}$ så transkonduktansen er $g_m = I_C/U_T = 11,5 \text{ mS}$ og med $R_C = 8,2 \text{ k}\Omega$ som belastning bliver forstærkningen $A = -g_m R_C = -95$.

Transistoren har forstærkningen $-g_m R_C$ der i eksemplet er -95 gange. Det betyder at et signal på basis med amplituden $u_{BE} = 1 \text{ mV}$ bliver overført til udgangen som et signal på $u_{CE} = -95 \text{ mV}$. Som en konsekvens heraf er der et spændingsfald over både modstanden R_B og kondensatoren C_μ med en amplitude på 96 mV . Følgelig virker impedansen af de to komponenter som en 96 gange lavere værdi ved indgangen. Set fra udgangen svinger kollektor kraftigt mens basis ligger næsten stille så her kan R_B og C_μ opfattes som liggende i parallel over udgangen. Forstærkeren præsenterer en belastning af mikrofonen på $19 \text{ k}\Omega$ i parallel med 285 pF der er nogenlunde reel op til 30 kHz .

På udgangssiden ligger C_μ i parallel med R_B og giver en dæmpning over 6 MHz . Det kræver dog at udgangen ikke belastes yderligere hvilket i praksis er ret usandsynligt for en AD konverter eller en efterfølgende forstærker vil belaste med mindst 10 pF . Hvis udgangen belastes med 50 pF falder grænsen til 400 kHz så denne type forstærker er meget følsom for belastningen på udgangen.

5.4.4 Hybrid-pi model

En mere udbygget småsignalmodel kaldes for *hybrid-pi* fordi diagrammet minder om det græske tegn pi (π). Modellen beskriver detaljer ved den bipolære transistor med fire modstande og to kondensatorer i tillæg til strømkilden og er her vist for en NPN transistor, men den gælder også for en PNP. Det viste materiale med en transkonduktans og kapaciteter kan også beskrive felt-effekt transistorer, CMOS analoge forstærkere og radiorør selv om teknologierne er forskellige.



Figur 258 – Hybrid-pi modellen består af en indgangsside med r_x for modstanden af det tynde lag i basis, r_π for den dynamiske indgangsmodstand og C_π for kapaciteten i basis-emitter. Dertil r_μ for tilbagekoblingen fra udgangen med C_μ for kapaciteten af dioden i kollektor. Udgangen består af en strømgenerator $g_m u_{BE}$ i parallel med r_o som model af strømgeneratorenens spændingsafhængighed.

$$\begin{aligned}
 r_\pi &\approx \frac{\beta}{g_m} & C_\pi &\approx \frac{g_m}{2\pi f_T} & r_x &\approx 100 \Omega & r_\mu &\approx 10 M\Omega & g_m &= \frac{I_C}{nU_T} & r_o &= \frac{U_A}{I_C} = \frac{1}{h_{OE}} \\
 &&&& C_\mu &\approx 1 \dots 5 pF &&& && r_o &\approx 35 k\Omega \\
 \beta &\approx 100 \dots 800 & f_T &\approx \frac{(300 \text{ MHz}) I_C}{10 \text{ mA}} & U_T &= \frac{kT}{q_0} \approx 26 \text{ mV} & & & U_A &\approx 70 \text{ V} & h_{OE} &\approx 30 \mu\text{s} \\
 && @ I_C < 10 \text{ mA} && && && & @ I_C = 2 \text{ mA} &
 \end{aligned}$$

Modellen kan virke lidt overvældende ved det første møde, men der er sjældent brug for samtlige parametre på én gang. Almindeligvis kan den reduceres til nogle ganske få af parametrene og et eksempel er AC modellen fra den indledende analyse hvor strømgeneratoren $g_m u_{BE}$ er den eneste parameter i brug og den efterfølgende analyse hvor C_μ tilføjes for at beregne båndbredden.

Der er så meget fælles gods fra transistor til transistor at hovedparten af det analytiske arbejde kan foretages inden man beslutter sig endeligt for den transistor der skal anvendes.

De enkelte parametre i hybrid-pi modellen har en meget direkte relation til den måde den bipolære transistor er opbygget. Der er andre modeller for hvordan transistoren opfører sig og de adskiller sig fra hinanden ved den måde parametrene relateres til arbejdspunktet. For h-parametermodellen vises parametrene som tabelværdier med et sæt kurver for at omregne til et andet arbejdspunkt end de 2 mA fra standarden. Den model introduceres på side 247. Dertil s-parametermodellen der er god ved høje frekvenser hvor hybrid-pi må give op hvilket er over f_T . Her beskrives en transistor som en to-port med komplekse impedanser for transmission og refleksion af signaler.

5.4.5 Basislagets modstand – r_x

Modstanden fra basis-terminalen og ind til den aktive del af basis kan være betydnende for basis er under 10 μm bred for en småsignaltransistor. Der vil typisk optræde en modstand på 100 Ω i serie med indgangen hvor den tilsvarende værdi er omkring 1 Ω for en effektransistor. I modsætning til de øvrige modstande i hybrid-pi modellen er r_x en reel modstand og ikke blot hældningen af en kurve. Som en følge heraf er der termisk støj fra modstanden og transistorer med lav værdi af r_x vil optræde som støjsvage transistorer. Værdien er ikke opgivet i databladet så man må i stedet bruge støjtallet NF (noise figure) for at finde en støjsvag transistor. Det emne behandles side 240.

5.4.6 Transkonduktans – g_m

AC strømmen i kollektor er givet ved AC spændingen over basis-emitter og proportionaliteten er transistorens transkonduktans (*transconductance*) med symbollet g_m hvor index kommer fra det engelske ord *mutual* stående for gensidig. Modellen har to porte (ind- og udgang) som forbides af transkonduktansen. Strømmen i_c er en AC strøm ligesom spændingen u_{BE} er en AC spænding og de er oscillationer omkring DC arbejdspunkterne. Transkonduktansen er en småsignalparameter og som vist herunder kræver det at amplituden af signalet over basis-emitter er under 1 mV.

Transistorens eksponentielle relation danner fundamentet for modellen.

$$I_c = I_s \exp\left(\frac{U_{BE}}{nU_T}\right), \quad I_c \gg I_s, \quad U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$$

Relationen mellem spændingen over basis-emitter og strømmen i kollektor defineres som den strømtilvækst i_c der opnås ved at øge spændingen med u_{BE} . Eksponentialfunktionen til en sum kan omskrives ved hjælp af regnereglen $\exp(x+y) = \exp(x) \cdot \exp(y)$. Derved bliver første led lig med DC arbejdspunktet I_c mens andet led udelukkende afhænger af signalets amplitude u_{BE} .

$$I_c + i_c = I_s \exp\left(\frac{U_{BE} + u_{BE}}{nU_T}\right) = I_c \exp\left(\frac{u_{BE}}{nU_T}\right) \Rightarrow i_c = I_c [\exp\left(\frac{u_{BE}}{nU_T}\right) - 1]$$

Eksponentialfunktionen kan udtrykkes som en sum af led fra Taylor rækkeudviklingen.

$$\exp(x) = 1 + x + \frac{x^2}{2} + \dots$$

En simplifikation med $\exp(x) \approx 1 + x$ er tilladelig hvis x er så lille at leddene af højere orden er meget små. Når det er tilfældet kan leddet med x benyttes som definition af transkonduktansen.

$$i_c = I_c \left[\frac{u_{BE}}{nU_T} + \frac{1}{2} \left(\frac{u_{BE}}{nU_T} \right)^2 + \dots \right] \xrightarrow{u_{BE} \rightarrow 0} i_c \approx \frac{I_c}{nU_T} u_{BE} \Rightarrow g_m \stackrel{\text{def}}{=} \frac{dI_c}{dU_{BE}} = \frac{i_c}{u_{BE}} \approx \frac{I_c}{nU_T}$$

For at beregne den harmoniske forvrængning udtrykkes signalet som en harmonisk svingning.

$$u_{BE}(t) = U_M \sin(\omega t), \quad \omega = 2\pi f$$

Indsættelse giver to nye led hvor det ene viser at DC arbejdspunktet bliver forskudt hvilket ikke normalt giver problemer og dertil et led med den dobbelte frekvens.

$$i_c(t) = I_c \left(\frac{1}{4} \left(\frac{U_M}{nU_T} \right)^2 + \frac{U_M}{nU_T} \sin(\omega t) + \frac{1}{4} \left(\frac{U_M}{nU_T} \right)^2 \cos(2\omega t) + \dots \right)$$

Leddet med den dobbelte frekvens repræsenterer et nyt signal og forvrængningen kaldes for den anden-harmoniske forvrængning (*second-order distortion*) da den indfører et signal på den anden-harmoniske plads. Forvrængningen beregnes ved amplituden af anden-ordens ledet i forhold til amplituden af første-ordens ledet.

$$D_2 = \frac{U_M}{4nU_T} \Rightarrow D_2 \approx 1 \% \frac{U_M}{\text{mV}}$$

Som en tommelfingerregel betyder det at forvrængningen bliver 1 % ved en signalamplitude på 1 mV. Tilnærmelsen forudsætter $n = 1$ og $U_T = 25$ mV. Forvrængningen kan regnes for hørbar¹²⁶ ved én procent så amplituden bør holdes under 1 mV.

¹²⁶Der er videnskabelig begrundelse for at angive 0,25 % som den nedre grænse for hørbar forvrængning med trænede lyttere, men det er de færreste, almindelige mennesker der er i stand til at høre en forvrængning under 1 %. I mange tilfælde vil der endda være en fordel ved en vis grad af forvrængning som Harry F. Olson påviste allerede i midten af 1900 tallet.

5.4.7 Indgangsmodstand – r_π

Modellens indgangsmodstand r_π skyldes den logaritmiske relation mellem spænding og strøm for transistorens indgangsside og den beregnes på samme måde som dynamiske modstand overfor den lille signalværdi u_{BE} dog med den forskel, at strømmen i basis er mindre end strømmen i kollektor på grund af relationen $I_C = \beta I_B$.

$$r_\pi = \frac{dU_{BE}}{dI_B} = \beta \frac{dU_{BE}}{dI_C} = \frac{\beta}{g_m} = \frac{\beta U_T}{I_C}$$

For en transistor med strømforsærkningen $\beta = 300$ og arbejdspunktet $I_C = 2 \text{ mA}$ er værdien $r_\pi \approx 4 \text{ k}\Omega$.

5.4.8 Indgangskapaciteten – C_π

Transistoren udviser en træghed overfor en pludselig ændring i basis-emitter spændingen før der er en reaktion i strømmen i kollektor. Det ses især ved brug af transistoren som kontakt hvor der kræves en vis tid for at tænde eller slukke for strømmen. Ved analog signalbehandling ses det som en begrænsning i transistorens egenskaber ved høje frekvenser. Hybrid-pi modelen repræsenterer denne træghed ved C_{jBE} fra ladningen i transistorens basis-emitter diode (*junction capacitance*) og C_{dBE} fra de faste ioner i PN overgangen (*depletion capacitance*) der er årsagen til det indbyggede potentielle. Den første af disse kapaciteter er funktion af strømmen og er normalt dominerende hvorimod den anden primært er givet ved fremstillingsprocessen.

$$C_\pi = C_{jBE} + C_{dBE} \quad D_{jBE} \gg C_{dBE} \quad C_\pi \approx \frac{g_m}{2\pi f_T}$$

Transistorens transitfrekvens f_T angiver den frekvens hvor strømforsærkningen er reduceret til én og oplyses altid i databladet. Værdien er ikke en konstant, men er proportional med strømmen i kollektor ved strømniveau under cirka 10 mA for småsignaltransistorer som BC547 og lignende så kondensatorens værdi er omrent konstant på cirka 150 pF under denne grænse (se side 238).

En kapacitet er defineret som ændringen i ladningen som følge af en ændring i spændingen. Ladningen er givet ved en faktor kaldet forward transient time τ_F , der måles eksperimentelt for transistoren. Ladningen Q_{BE} er givet ved $\tau_F I_C$ ifølge definitionen på ladning (tid gange strøm). En ændring beskrives ved differentiation hvor dI_C/dU_{BE} angiver transkonduktansen. En tidskonstant er relateret til -3dB frekvensen f_T gennem $\tau_F = 1/2\pi f_T$ hvilket introducerer transistortiden.

$$C_{jBE} = \frac{dQ_{BE}}{dU_{BE}} = \frac{d(\tau_F I_C)}{dU_{BE}} = \tau_F g_m = \frac{g_m}{2\pi f_T}$$

5.4.9 Kollektorkapaciteten – C_μ

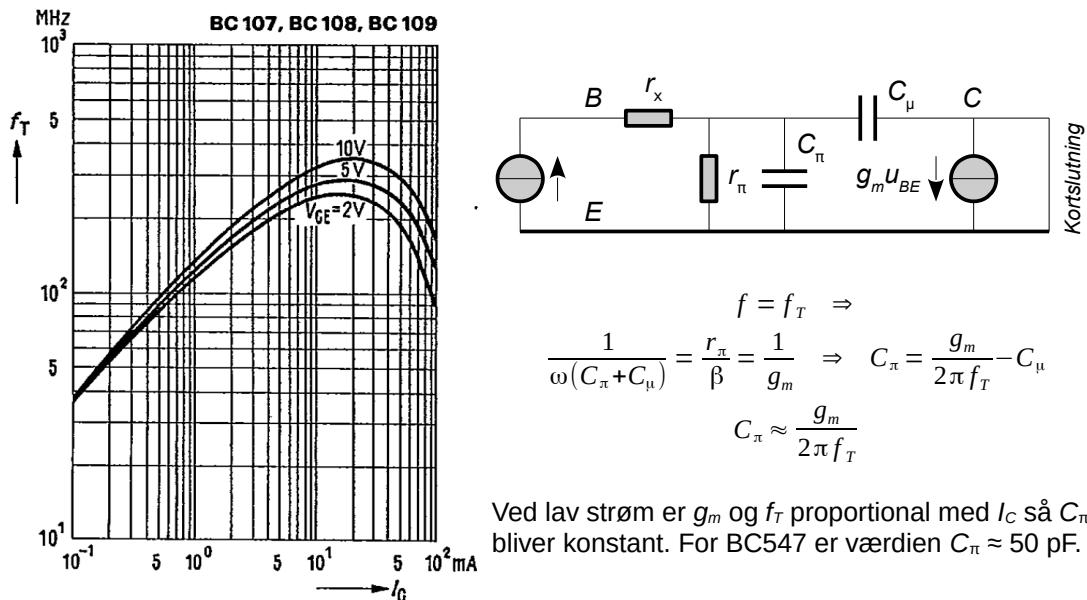
Normalt er dioden i kollektor-basis spærrende hvilket betyder at PN overgangen har to ladninger som adskilt af PN overgangen (*depletion layer*). Det er selve definitionen på en kondensator og den giver kapaciteten af kollektor-basis strækningen. Kapaciteten har fået betegnelsen C_μ og har en værdi på cirka 3 pF for typiske småsignaltransistorer.

Værdien er ikke konstant, men er en funktion af spændingen over kollektor-basis. Ved en forøget spændingsforskæft vil den ladningsfri strækning i PN overgangens grænse udvides og det reducerer kapaciteten. Det kaldes for Early-effekten efter den videnskabsmand der beskrev effekten og den er helt på linje med den spændingsafhængige kapacitet for en varicapdiode fra side 210.

Kondensatorenen begrænser båndbredden af en forstærker idet den sidder som en tilbagekobling fra udgang til udgang. I operations- og effektforstærkere er spændingsafhængigheden et problem for lineariteten da det kan lede til forvrængning. Indgangskapaciteten påvirkes også af C_μ fordi den kan Miller-transformeres til indgangen hvor den bliver forøget med åben-sløjfe forstærkningen.

5.4.10 Transitfrekvens – f_T

Transistorens funktion ved høje frekvenser defineres ved transitfrekvensen (*transition frequency*) der er defineret som den frekvens hvor strømforstærkningen er faldet til én. Den måles med en kortsluttet udgang så de to kondensatorer C_π og C_μ ligger i parallel. Strømmen i basis deles mellem kondensatorerne og r_π og strømforstærkningen når ned på én når impedansen af C_π og C_C har den samme værdi som indgangsmodstanden delt med DC værdien af strømforstærkningen β .



Figur 259 – Transitfrekvensen for typiske småsignaltransistorer som fx BC547 (Siemens).

Transitfrekvensen er funktion med strømmen for en strømstyrke under cirka 10 mA så en transistor mister båndbredde ved reduceret strømstyrke. Det ses også at båndbredden aftager ved en endnu højere strømværdi hvor tætheden af ladningsbærere bliver høj i PN overgangen.

AC spændingen over basis-emitter u_{BE} er givet af strømmen i basis i_B gange parallelforbindelsen af r_π og impedansen af C_π og C_μ i parallel. Kapaciteten C_π består af diffusionskapaciteten C_{dBE} og rumladningskapaciteten $\tau_F g_m$ hvor τ_F er transistorens *forward transition time* der er en slags reaktionstid som databladet angiver i form af transitfrekvensen f_T .

$$\begin{aligned} u_{BE} &= \frac{i_B}{1/r_\pi + j\omega(C_\pi + C_\mu)} \Rightarrow \beta = \frac{i_C}{i_B} = \frac{1}{\frac{1}{g_m r_\pi} + \frac{j\omega(C_{dBE} + C_\mu)}{g_m} + j\omega\tau_F} \\ i_C &= g_m u_{BE} \\ C_\pi &= C_{dBE} + \tau_F g_m \end{aligned}$$

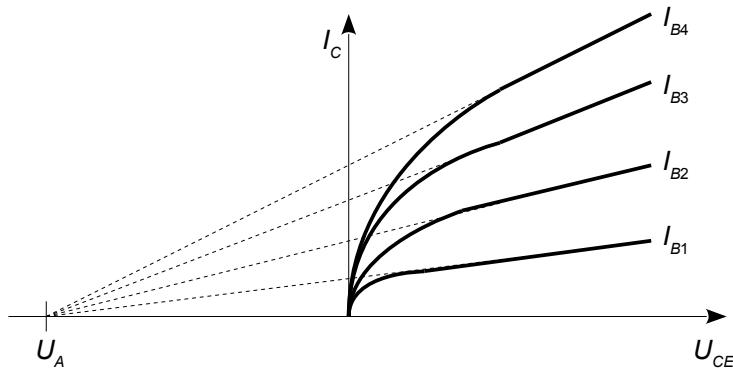
Ved DC er $\omega = 0$ så DC strømforstærkningen defineres som $\beta = g_m r_\pi$ og derved er $1/g_m r_\pi$ en lille værdi ligesom strømforstærkningen normalt er høj. Ved en tilstrækkelig høj frekvens bliver de to andre led større end $1/g_m r_\pi$ og reducerer strømforstærkningen. Transitfrekvensen bestemmes ved at udnytte at $g_m = I_C/U_T$ og at en lav strømværdi i kollektor giver en høj værdi af første led i nævneren.

$$\frac{\omega_T(C_{dBE} + C_\mu)U_T}{I_C} + \omega_T\tau_F = 1 \Rightarrow f_T = \frac{1}{2\pi(C_{dBE} + C_\mu)U_T + 2\pi\tau_F} \stackrel{I_C \text{ lav}}{\Rightarrow} f_T \approx \frac{I_C}{2\pi(C_{dBE} + C_\mu)U_T}$$

Transitfrekvensen er $f_T \approx 300$ MHz ved $I_C \approx 20$ mA og det betyder at $\tau_F = 1/2\pi f_T \approx 0,5$ ns. Ved lav strømværdi findes $C_{dBE} \approx I_C/2\pi f_T U_T = 50$ pF for $I_C = 1$ mA hvis C_μ ignoreres (omkring 3 pF).

5.4.11 Udgangsmodstanden – r_o

Strømmen i transistoren kollektor regnes almindeligvis som uafhængig af spændingsfaldet over kollektor-emitter strækningen, men strømmen i kollektor stiger med stigende spænding og det kan modelleres som en modstand r_o i parallel med udgangens strømkilde. Udgangskarakteristikken for strømmen i kollektor som funktion af spændingen over kollektor og emitter med strømmen i basis som parameter vises herunder i sin idealiserede udgave. Det kan vises at hældningen af kurverne kan repræsenteres ved et fælles skæringspunkt der kaldes transistoren Early voltage U_A .



Definition:

$$r_o = \frac{dU_{CE}}{dI_C}$$

Early voltage og h-parameter:

$$r_o \approx \frac{|U_A|}{I_C} \Rightarrow U_A = r_o I_C = \frac{I_C}{h_{oe}}$$

$$r_o = \frac{1}{h_{oe}}$$

Figur 260 – Udgangskarakteristikken for transistoren viser en relation mellem strømmen i kollektor og spændingen over kollektor-emitter. Hældningen i arbejdspunktet er i litteraturen specificeret som Early voltage U_A .

Teorien forklarer udgangsmodstanden ved en ændring af tykkelsen af basislaget som funktion af spændingen over dioden i kollektor-basis. Ved stigende spænding øges det ladningsfrie område i PN overgangen og derved reduceres basislagets tykkelse så sandsynligheden øges for at der slipper ladningsbærere igennem basis uden rekombination. Databladene angiver ikke værdien, men den kan findes som h-parametermodellens h_{oe} (se side 247) og fra SPICE simulatorer¹²⁷.

De viste formler i illustrationen er ment som en hjælp ved et analytisk arbejde, men de skal ikke opfattes som den ultimative sandhed. Ved sammenligning med h-parametermodellen ses blandt andet at udgangsadmittansen h_{oe} ikke er direkte proportional med strømmen så den viste relation mellem r_o og h_{oe} holder ikke til et nærmere eftersyn. Formlerne kan dog være brugbare for et overslag af hvilken værdi udgangsmodstanden har i en given opstilling, men de må ikke benyttes ukritisk. Se også beregningen på side 224.

For BC107, der er ækvivalent med BC547, er Early voltage fundet til $U_A = -96$ V så udgangsmodstanden bliver $r_o = 48$ kΩ ved 2 mA. De opgivne værdier for h-parameteren er $h_{oe} = 18$ µS for BC107A stigende til $h_{oe} = 60$ µS for BC107C. Udgangsmodstanden varierer tilsvarende fra $r_o = 56$ kΩ ned til 17 kΩ og dækker den beregnede værdi.

Det er normalt at ignorere udgangsmodstanden ved analytisk arbejde for en modstand i kollektor er normalt væsentligt mindre end transistoren udgangsmodstand. Undtagelsen er specifikke tiltag for at hæve værdien af modstanden forstærkningen maksimeres. Det kan være brug af en strømgenerator eller en ekstra transistor for boot strapping hvor et eksempel vises på side 243.

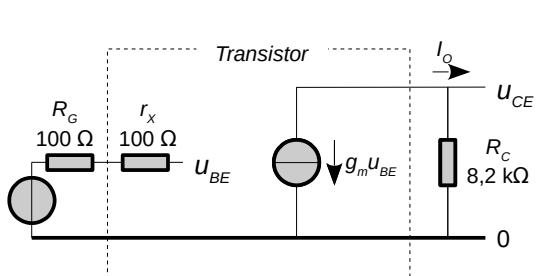
En strømgenerator benyttes i den integrerede kreds LM3900 hvor forstærkningen er 2800 gange.

For BC547 ved $I_C = 2$ mA er transkonduktansen $g_m = 77$ mS og med $r_o = 17 \dots 56$ kΩ er den maksimalt mulige forstærkning med denne transistor 1300 ... 4300 gange. Ved måling i laboratoriet har jeg dog ikke kunnet opnå mere end cirka 800 ganges forstærkning ved brug af boot strapping metoden.

¹²⁷Ved Multisim placeres transistoren på arbejdsfladen. Man dobbelt-klikker på komponenten og åbner for "Edit model" hvor SPICE parametrene er opført.

5.4.12 Egenstøj

Forstærkeren er ikke støjfri for der genereres både termisk støj fra modstandene og haglstøj fra strømmene i basis og kollektor. Den termiske støj e_{nR} stammer fra signalkildens modstand R_G og modstanden r_x i basis. Dertil kommer haglstøj i_{nB} fra strømmen I_B i basis som løber gennem R_G og r_x og genererer støjen e_{nB} ved indgangen. Endelig genereres en haglstøj i_{nC} fra strømmen I_C i kollektor som løber gennem den eksterne modstand R_C og giver endnu et støjbidrag. Denne støj henføres til basis som e_{nC} ved division med forstærkningen $g_m R_C$.



(1) Termisk støj: $e_{nR} = \sqrt{4kT(R_G + r_X)}$

(2) Haglstøj fra I_B : $i_{nB} = \sqrt{2q_0 I_B} = \sqrt{2q_0 \frac{I_C}{\beta}}$

(3) Haglstøj fra I_C : $e_{nB} = (R_G + r_X)i_{nB}$
 $i_{nC} = \sqrt{2q_0 I_C}$
 $e_{nC} = \frac{R_C i_{nC}}{g_m R_C} = U_T \sqrt{\frac{2q_0}{I_C}}$

Samlet indgangsstøj $e_{nTOT} = \sqrt{e_{nR}^2 + e_{nB}^2 + e_{nC}^2}$

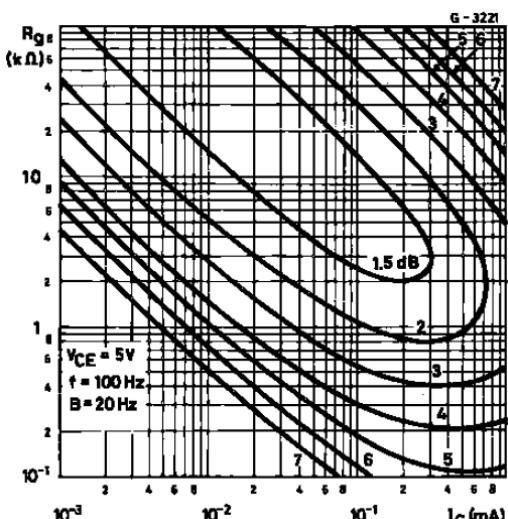
Figur 261 – Støjbidragene er signalkildens indre modstand R_G og transistorens modstand i basis r_X der begge er termiske støjkilder og dertil haglstøjen fra strømmen i basis og i kollektor. Alle bidrag bliver henført til indgangen ved forstærkningen ved $g_m R_C$ der antager at udgangen er ubelastet.

For mikrofonforstærkeren fra side 233 er spændingsstøjtætheden fra modstandene på $e_{nR} = 1,8 \text{ nV}/\sqrt{\text{Hz}}$. Transistoren arbejder ved $I_C = 0,3 \text{ mA}$ så $g_m = 11,5 \text{ mS}$ og med $\beta = 300$ bliver $I_B = 1 \mu\text{A}$. Haglstrøjen fra strømmen i basis bliver $i_{nB} = 0,56 \text{ pA}/\sqrt{\text{Hz}}$ så spændingsstøjtætheden herfra er $e_{nB} = 0,11 \text{ nV}/\sqrt{\text{Hz}}$. For strømmen i kollektor beregnes $i_{nC} = 9,8 \text{ pA}/\sqrt{\text{Hz}}$ så spændingsstøjtætheden ved undgangen er $80 \text{ nV}/\sqrt{\text{Hz}}$ der henføres til basis ved forstærkningen $q_m R_C = 95$ til $e_{nC} = 0,85 \text{ nV}/\sqrt{\text{Hz}}$. Totale støj: $e_{nT} = 2,0 \text{ nV}/\sqrt{\text{Hz}}$.

Den ækvivalente indgangsstøj er 2 nV/√Hz og er domineret af de to modstandes termiske støj. Til sammenligning er en ekstremt lav støjende operationsforstærker som fx AD797 på 0,9 nV/√Hz.

5.4.13 Støjtal

Modstanden r_x er ikke normalt opgivet i databladet til gengæld oplyses transistorens støjtal NF (*noise figure*) undertiden. Den angiver det antal decibel hvormed transistoren forringør det mulige signal/støj forhold fra indgangens SNR_G , der alene består af termisk støj fra modstanden R_G , og ned til SNR_K for hele kredsløbet. Estimering af r_x er enkel hvis modstandsstøjen dominerer.



$$\begin{aligned}
 & \text{Definition af støjtallet:} \\
 NF &= 20 \cdot \log_{10} \left(\frac{SNR_G}{SNR_K} \right) \text{ dB} = 20 \cdot \log_{10} \left(\frac{U_G/U_{nRG}}{U_G/U_{n(RG+RX)}} \right) \text{ dB} \\
 \Rightarrow \frac{NF}{\text{dB}} &= 20 \cdot \log_{10} \left(\frac{U_{n(RG+RX)}}{U_{nRG}} \right) = 20 \cdot \log_{10} \left(\frac{e_{n(RG+RX)}}{e_{nRG}} \right) \\
 \Rightarrow \frac{NF}{\text{dB}} &= 20 \cdot \log_{10} \left(\frac{\sqrt{4kT(R_G+r_x)}}{\sqrt{4kTR_G}} \right) = 10 \cdot \log_{10} \left(1 + \frac{r_x}{R_G} \right) \\
 \Rightarrow r_x &= R_G \left(10^{\frac{NF}{10 \text{ dB}}} - 1 \right)
 \end{aligned}$$

For $I_C < 500 \mu A$ og $R_C = 110 \Omega$ aflæses:

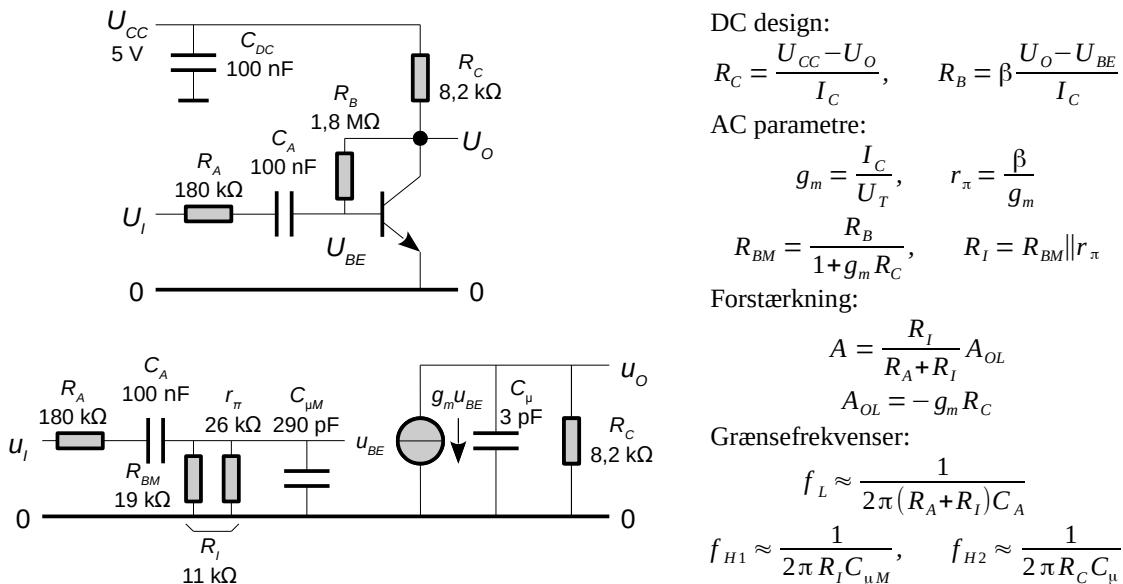
$$NF = 5 \text{ dB} \Rightarrow r_y = 110 \Omega \cdot \left(10^{\frac{5 \text{ dB}}{10 \text{ dB}}} - 1\right) = 240 \Omega$$

Figur 262 – Støjtallet NF for BC109 (BC549) der er funktion af generatorens modstand og strømmen i kollektor (SGS Thompson)

5.4.14 Inverterende forstærker

Den inverterende forstærker kan opfattes som en operationsforstærker hvor signalkilden føres ind via R_A og med tilbagekobling via R_B . En forstærkningen på $-R_B/R_A$ kunne derfor forventes, men både åben-sløjfe forstærkningen og indgangsmodstanden ved basis af transistoren er for lave til at målet kan nås. Problemet blev omtalt side 223 og uddybes herunder.

Åben-sløjfe forstærkning er $A_{OL} = -g_m R_C = -95$ så basismodstanden R_B kan Miller-transformeres til indgangen som $R_{BM} = R_B/96 = 19 \text{ k}\Omega$. Indgangsmodstanden på transistoren beregnes til $r_\pi = 26 \text{ k}\Omega$. De to i parallel giver $R_I = 11 \text{ k}\Omega$ så indgangssignalet bliver dæmpet med faktor 0,058 på grund af spændingsdeleren ved R_A , R_{BM} og r_π og forstærkes derefter med A_{OL} til en forstærkning på $-5,4$. Det er i god overensstemmelse med det tidligere estimat for forstærkningen på $-5,8$.



Figur 263 – Miller-transformering af tilbagekoblingen viser at i indgangsmodstanden og åben-sløjfe forstærkningen begge er for lave til at forstærkningen kan indstilles alene ved R_A og R_B .

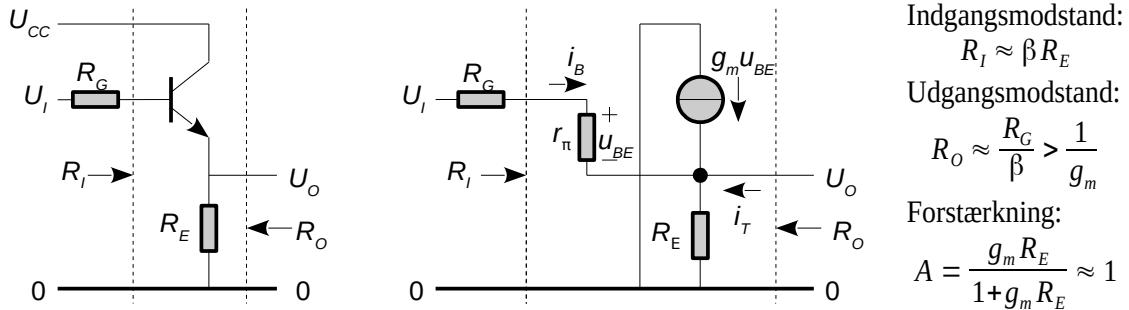
Med $U_{CC} = 5 \text{ V}$, $U_O = 2,5 \text{ V}$ og $I_C = 0,3 \text{ mA}$ bliver $R_C = 8,2 \text{ k}\Omega$ og for $\beta = 300$ er $R_B = 1,8 \text{ M}\Omega$. Transistorens transkonduktans beregnes til $g_m = 11,5 \text{ mS}$ så åben-sløjfe forstærkningen bliver $A_{OL} = -94$. Modstanden i basis Miller-transformeres til $R_{BM} = 19 \text{ k}\Omega$ og transistorens indre modstand er $r_\pi = 26 \text{ k}\Omega$ så de to i parallel giver $R_I = 11 \text{ k}\Omega$. Den resulterende spændingsforstærkning beregnes som en spændingsdeler ved R_A og R_I samt åben-sløjfe forstærkningen og findes til $A = -5,4$.

Kondensator C_A har primært til formål at blokere for DC så arbejdspunktet bliver som forventet. Den har en lav impedans ved høje frekvenser så den ikke påvirker funktionen, men impedansen stiger ved lave frekvenser og den vil tage over når kondensatorens impedans $1/2\pi f C_A$ bliver stor nok. Grænsefrekvensens -3 dB punkt f_L nås når serieforbindelsen af R_A og R_I bliver lig med impedansen af kondensatoren og strømmen i indgangskredsen reduceres til 71 %. Det sker ved en frekvens på 8 Hz og mod lavere frekvenser aftager forstærkningen med 6 dB per oktav.

Mod høje frekvenser bliver de to kondensatorer fra hybrid-pi modellen betydelige og vil dæmpe forstærkningen for frekvenser over f_{H1} og f_{H2} . Den første skyldes indgangens lavpasfilter ved $C_{\mu M}$ og R_I . Parallelforbindelsen med R_A ignoreres da den er meget større end R_I . En Miller-transformering af tilbagekoblingens kondensator giverer $C_{\mu M} = 95C_\mu = 290 \text{ pF}$ så grænsen ved indgangstrinnet er på 50 kHz (-3 dB). Ved udgangen er C_μ i parallel med modstanden i kollektor og giver en grænse lidt over 6 MHz. Det kunne virke som om det er den første af de to som betyder noget, men f_{H2} er funktion af belastningen fra det efterfølgende kredsløb. Ved 400 pF er de to grænsefrekvenser ens. Det svarer til nogle få meter kabel tilsluttet udgangen.

5.4.15 Buffer

Ligesom en operationsforstærker ofte benyttes som buffer mellem to kredsløb så de ikke påvirker hinanden kan transistoren isolere et kredsløb fra et andet. Den betegnes da som en emitterfølger (*emitter follower*) eller fælles kollektor (*common collector*) idet kollektoren er fælles elektrode for indgang og udgang. Udgangens DC potentielle er forskudt med cirka 0,7 V i forhold til indgangen og der er ikke fuldstændig adskillelse mellem ind- og udgang for indgangsmodstanden er funktion af belastningen på udgangen og vice versa.



Figur 264 – En buffer har en høj indgangsmodstand og lav udgangsmmodstand. Udgangen er 0,7 V forskudt i forhold til indgangens DC niveau og forstærkningen er lidt under én.

Indgangsmodstanden regnes som strømforstærkningen gange modstanden i emitter og er typisk over 100 kΩ for småsignaltransistorer. For en MOSFET er indgangsmodstanden en kondensator. En BJT effektransistor med 8 Ω belastning og β = 25 giver en indgangsmodstand på 200 Ω, men med to emitterfølgere i kaskade (*Darlington pair*, se næste kapitel) bliver den over 20 kΩ.

Udgangsmmodstanden er givet af signalkildens indre modstand divideret med strømforstærkningen og kommer sjældent over 100 Ω. Som minimum er udgangsmmodstanden $1/g_m$ der ofte kaldes for transistorens emittermodstand r_E og ved en strøm på $I_C = 1$ mA er den $r_E \approx 1/g_m = 25$ Ω. For en MOSFET gælder også at $R_O = 1/g_m$ hvor transkonduktansen dog beregnes som vist side 248.

Emitterfølgerens spændingsforstærkning er lidt mindre end én idet der er en spændingsdeling mellem udgangsmmodstanden R_O og emittermodstanden R_E .

Indgangsmmodstanden på basis findes ved at beregne indgangsspændingen U_I som følge af strømmen i_B i basis og danne forholdet mellem de to.

$$\begin{aligned} U_I &= r_\pi i_B + u_E \\ u_E &= R_E(i_B + i_C) = (1+\beta) R_E i_B \end{aligned} \Rightarrow R_I = \frac{U_I}{i_B} = r_\pi + (1+\beta) R_E$$

Udgangsmmodstanden findes med en teststrøm i_T påtrykt udgangen og med indgangen på nul. Det antages at r_π er større end generatorens indre modstand R_G samt at β er stor.

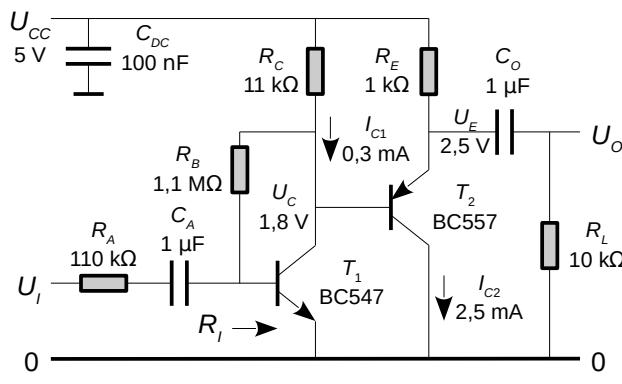
$$\begin{aligned} U_O &= R_E[i_B + i_C + i_T] \\ i_B &= -\frac{U_O}{R_G + r_\pi} \end{aligned} \Rightarrow R_O = \frac{U_O}{i_T} = \frac{(R_G + r_\pi) R_E}{R_G + r_\pi + (1+\beta) R_E} \approx \frac{R_G + r_\pi}{1+\beta} \approx \frac{1}{g_m}$$

Forstærkningen bestemmes ved Kirchhoff på knudepunktet ved emitteren. Det sidste trin i udledningen herunder antager at strømforstærkningen kan regnes for stor.

$$\begin{aligned} \frac{U_O - U_I}{R_G + r_{pi}} + \frac{U_O}{R_E} - g_m u_{BE} &= 0 \\ u_{BE} &= -\frac{r_\pi}{R_G + r_\pi} (U_O - U_I) \end{aligned} \Rightarrow A = \frac{U_O}{U_I} = \frac{1 + g_m r_\pi}{1 + g_m r_\pi + \frac{R_G + r_\pi}{R_E}} = \frac{1}{1 + \frac{R_G}{(1+\beta) R_E} + \frac{r_\pi}{(1+\beta) R_E}} \approx \frac{g_m R_E}{1 + g_m R_E}$$

5.4.16 Forstærker med buffer

Forstærkeren kan suppleres med en buffer for at gøre forstærkningen mindre afhængig af det efterfølgende trins belastning der her repræsenteres ved R_L . DC niveauet på udgangen ligger en diodes spændingsfald højere end kollektor af T_1 så spændingen på kollektor af T_1 er sänket og det øgede spændingsfald over R_C hæver åben-sløjfe forstærkningen i T_1 .



Designprocedure:

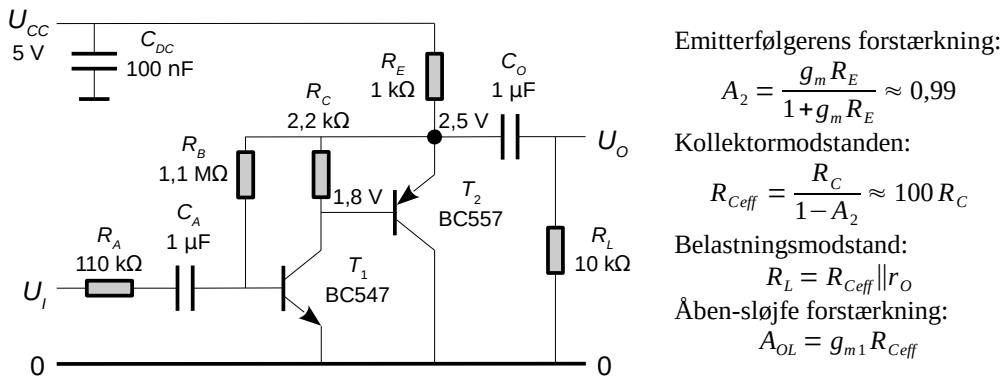
1. Vælg forsyningen U_{CC} .
2. Vælg strømmene I_{C1} og I_{C2} .
3. Aflæs β_1 og β_2 fra datablad.
4. $R_E = U_{CC}/2I_{C2}$.
5. $U_C = U_{CC}/2 - U_{BE2}$.
6. $R_C = (U_{CC} - U_C)/I_{C1}$.
7. $R_B = \beta_1(U_C - U_{BE1})/I_{C1}$.
8. $g_{m1} = I_{C1}/U_T$ og $g_{m2} = I_{C2}/U_T$.
9. $A_{OL} = g_{m1}R_C$.
10. $r_{\pi1} = \beta_1/g_{m1}$ og $r_{\pi2} = \beta_2/g_{m2}$.
11. $R_I = r_{\pi} \parallel R_B/A_{OL}$.
12. $R_O = (R_C + r_{\pi2})/A_{OL}$.

Figur 265 – BJT inverterende forstærker med buffer for at tillade belastning af udgangen.

Indgangsmodstanden på basis af T_2 er $\beta_2 R_E = 270 \text{ k}\Omega$ ved en strømforstærkning på $\beta_2 = 300$ så den belaster forstærkeren minimalt og åben-sløjfe forstærkningen er $g_{m1} R_C = 127$. Modstanden R_B Miller-transformeres til indgangen som $R_B/A_{OL} = 8,7 \text{ k}\Omega$ og ligger i parallel med $r_{\pi} = 26 \text{ k}\Omega$ til en resulterende indgangsmodstand ved basis på $R_I = 6,5 \text{ k}\Omega$. Indgangssignalet dæmpes af R_A og R_I med faktor 0,059 og forstærkes med åben-sløjfe forstærkningen til -7,4. Udgangsmodstanden fra T_1 er modstanden R_C så udgangsmodstanden fra T_2 estimeres til $R_O = R_C/\beta_2 + 1/g_{m2} = 47 \Omega$.

5.4.17 Forstærker med boot strap

Bufferen kan øge den effektive værdi af modstanden i kollektor på T_1 så forstærkningen øges og ultimativt kun begrænses af transistorens udgangsmodstand. Modstanden R_C er placeret mellem basis og emitter på T_2 så spændingsfaldet over modstanden er næsten konstant. Bufferen har en forstærkning på 0,99 med de viste komponentværdier så AC strømmen reduceres til nær ved en hundrededel og modstanden virker tilsvarende 100 gange forstørret. Det kaldes for *boot strap* og teknikken benyttes både i integrerede kredsløb og diskret opbyggede effektforstærkere.



Emitterfølgerens forstærkning:

$$A_2 = \frac{g_m R_E}{1 + g_m R_E} \approx 0,99$$

Kollektormodstanden:

$$R_{Ceff} = \frac{R_C}{1 - A_2} \approx 100 R_C$$

Belastningsmodstand:

$$R_L = R_{Ceff} \parallel r_o$$

Åben-sløjfe forstærkning:

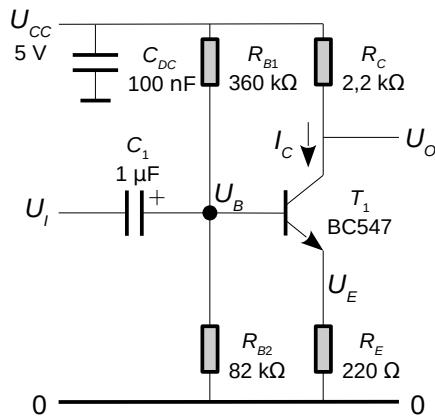
$$A_{OL} = g_{m1} R_{Ceff}$$

Figur 266 – En inverterende forstærker med buffer for mest mulig forstærkning.

Åben-sløjfe forstærkningen er stort set kun funktion af udgangsmodstanden fra T_1 og med en værdi omkring 50 $\text{k}\Omega$ bliver forstærkningen på 600 gange. Den resulterende forstærkning ved tilbagekobling gennem R_B bliver -9,1.

5.4.18 Forstærker med lokal tilbagekobling

Den inverterende forstærker kan udbygges med en modstand i emitter som giver tilbagekobling fra udgangen da strømmen i emitter og kollektor er næsten ens. Det øger indgangsmodstanden ved basis af transistoren, mindske forvrængningen og giver en veldefineret forstærkning der ofte ses anført som $A = R_C/R_E$ i den populære litteratur, men uden kommentar om nøjagtigheden. Det var en populær opstillingen fra omkring 1960 hvor transistoren tog over fra elektronrøret, men viger i takt med at operationsforstærkeren bliver dominerende indenfor analog signalbehandling.



Forstærkning:

$$A = -\frac{g_m R_E}{1 + g_m R_E} \frac{R_C}{R_E}$$

Tilnærmede udtryk:

$$A = -\frac{R_C}{R_E} \quad \text{for } R_E \gg \frac{1}{g_m}$$

$$A = -g_m R_C \quad \text{for } R_E \ll \frac{1}{g_m}$$

Indgangsmodstand:

$$R_I = R_{B1} \parallel R_{B2} \parallel (r_\pi + \beta R_E)$$

Udgangsmodstand:

$$R_O = r_O \parallel R_C \approx R_C$$

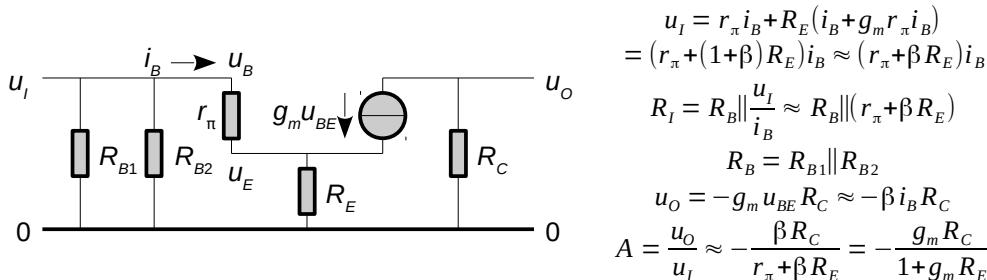
Figur 267 – En inverterende forstærker med tilbagekobling fra modstanden i emitter.

De viste komponenter giver $U_B = 0,93$ V så $U_E = 0,23$ V hvorefter $I_C = 1$ mA, $g_m = 38,5$ mS og $r_\pi = 7,8$ kΩ for $\beta = 300$. Forstærkningen er 8,9 gange og indgangsmodstanden bliver 35 kΩ.

Jeg vil ikke anbefale designet for nye produkter. Problemet er at ind- og udgangene benytter hver sin reference da indgangen refererer til nul og udgangen til effektforsyningen. Et støjsignal på effektforsyningen kobles direkte ind på udgangen samt gennem spændingsdelerens R_{B1} til basis. Strømmen fra udgangen kobler desuden tilbage på effektforsyningen og kan derved generere andre kredsløb. Det giver en kompliceret kobling mellem flere trin i et kredsløb og det regnes for årsagen til den meget irriterende *motor boating* der viser sig som en svag brummen der spontant opstår og forsvinder igen.

Opstillingen beskrives i en del af populærlitteraturen som værende uden tilbagekobling, men det er helt forkert selv om der ikke er en modstand fra kollektor til basis. Udgangssignalet er en strøm i kollektor og en strømvariation vil producere en tilsvarende spændingsændring over modstanden i emitter. Transistoren styres af spændingsfaldet over basis-emitter og det dannes af differensen mellem indgangsspændingen og spændingen over modstanden i emitter.

Analysen benytter indgangsmodstanden $r_\pi = \beta/g_m$ og strømkilden $g_m u_{BE}$ fra hybrid-pi modellen. Spændingen over basis-emitter skrives som $u_{BE} = r_\pi i_B$ og det udnyttes at strømforstærkningen er stor så der ikke er betydnende forskel på β og $1 + \beta$.

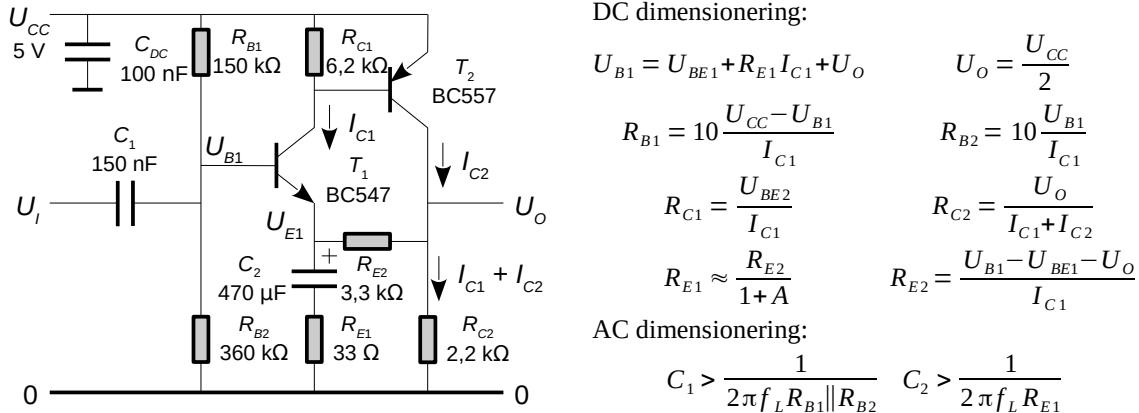


$$\begin{aligned} u_I &= r_\pi i_B + R_E(i_B + g_m r_\pi i_B) \\ &= (r_\pi + (1 + \beta) R_E) i_B \approx (r_\pi + \beta R_E) i_B \\ R_I &= R_B \parallel \frac{u_I}{i_B} \approx R_B \parallel (r_\pi + \beta R_E) \\ R_B &= R_{B1} \parallel R_{B2} \\ u_O &= -g_m u_{BE} R_C \approx -\beta i_B R_C \\ A &= \frac{u_O}{u_I} \approx -\frac{\beta R_C}{r_\pi + \beta R_E} = -\frac{g_m R_C}{1 + g_m R_E} \end{aligned}$$

Figur 268 – Analysen af forstærkeren tager udgangspunkt i strøm og spænding ved basis.

5.4.19 To-trins forstærker

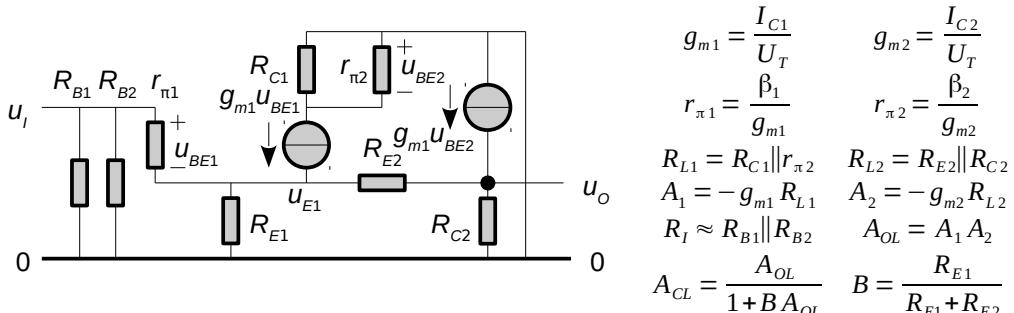
Forstærkeren med lokal tilbagekobling kan udvides med en komplementær transistor for at øge åben-sløjfe forstærkningen. Det leder til en konstruktion der var meget populær til lydformål frem til omkring 1980 hvor operationsforstærkere af tilstrækkelig god kvalitet var tilgængelige.



Figur 269 – BJT som ikke-inverterende forstærker med tilbagekobling over emitter.

En mulig designprocedure gennemgås herunder. De to kondensatorer har som formål at blokere for DC strøm og skal være store nok til at lade frekvenser over den lave grænsefrekvens passere. Forstærkningen fastsættes af modstandene R_{E1} og R_{E2} samt åben-sløjfe forstærkningen A_{OL} . Som ved den tidligere forstærker kan der indkobles støj fra effektforsyningen så en afkobling ved R_{B1} er vigtig og den var ofte realiseret ved at dele R_{B1} i to og indskyde en kondensator til filtrering.

Designet af opstillingen starter med valg af effektforsyningen (her $U_{CC} = 5$ V) og af strømstyrken i transistorerne. De er $I_{C1} = 100 \mu A$ for at få en lav værdi af strømmen i basis på første transistor og $I_{C2} = 1 \text{ mA}$ for en lav udgangsmodstand fra forstærkeren. Strømforstærkningen er antaget til $\beta_1 = 250$ og $\beta_2 = 500$. Udgangsspændingen stilles på halvdelen af effektforsyningen ($U_O = 2,5$ V) og der forventes kun et mindre spændingsfald over R_{E2} så basis af T_1 lægges på $U_{B1} = 3,5$ V. Spændingsdeleren R_{B1} og R_{B2} designes for 10 % af strømmen i I_{C1} så strømmen i basis kun får minimal indvirkning. De to modstande i kollektor indstilles til det ønskede strømniveau hvor R_{C2} skal optage begge transistorers strømme. Modstanden R_{E2} kan fastlægges ud fra ønsket om et spændingsfald på 1 V over U_{BE1} og R_{E2} . Modstanden R_{E1} gives af den ønskede forstærkning A_{CL} . Kondensatorerne skal blokere for DC og definerer den lave grænsefrekvens (her $f_L = 10$ Hz). Det er normalt nok at antage en indgangsmodstand på R_{B1} i parallel med R_{B2} og runde C_1 op.

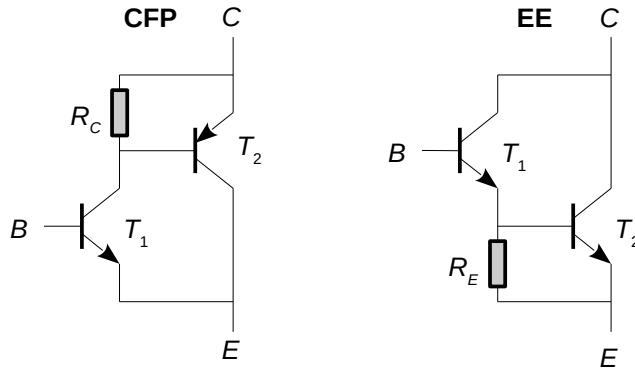


Figur 270 – Analyse af den ikke-inverterende forstærker.

Med $g_{m1} = 3,8 \text{ mS}$ og $g_{m2} = 38 \text{ mS}$ bliver $r_{\pi 1} = 65 \text{ k}\Omega$ og $r_{\pi 2} = 13 \text{ k}\Omega$, som i parallel med R_{C1} giver en belastning af første transistor på $R_{L1} = 4,2 \text{ k}\Omega$. Åben-sløjfe forstærkning bliver $A_1 = -14,6$ ved brug af det tidligere afsnits formler hvor $g_m R_E = 0,13$. Belastningen af andet trins udgang er på $R_{L2} = 1,3 \text{ k}\Omega$ så åben-sløjfe forstærkning bliver $A_2 = -50,7$. Åben-sløjfe forstærkningen for hele kredsløbet bliver $A_{OL} = 743$ og forstærkningen indstilles af R_{E1} og R_{E2} til $A_{CL} = 89$.

5.4.20 Forstærket emitterfølger

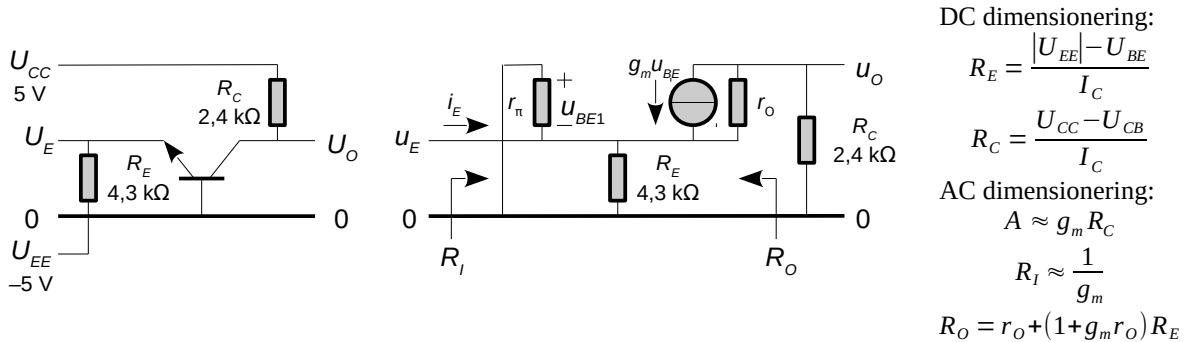
En forstærket emitterfølger med komplementære transistorer (CFP, complementary feedback pair) kan benyttes til at drive en stor strøm med en lav indgangsstrøm. Kredslobet minder om meget om den dobbelte emitterfølger (EE, Darlington), men der er kun en enkelt basis-emitter strækning fra indgang til udgang. De har begge strømforstærkningen $\beta_1\beta_2$, og modstanden R_C eller R_E indstiller strømmen i kollektor af T_1 samt forbedrer hastigheden hvormed T_2 kan afbrydes.



Figur 271 – En forstærket emitterfølger (CFP) og en dobbelt emitterfølger (EE).

5.4.21 Basisjordet transistor

Ved den basisjordede transistor udgør basis den fælles terminal for indgangen ved emitter og udgangen ved kollektor og den kaldes også fælles basis (common base). Den er populær ved højfrekvens hvor kapaciteten mellem kollektor og basis ikke kobler tilbage til indgangen. Den har den samme høje forstærkning som ved fælles emitter, men indgangsmodstanden er meget lav.



Figur 272 – En basisjordet transistor har stor forstærkning og lav indgangsmodstand. Fordelen er at kapaciteten i kollektor ikke føres tilbage til indgangen, men til nul.

Spændingsforstærkningen A findes ved at opstille Kirchhoffs ligning for knuden ved kollektor og det udnyttes at $u_{BE} = -u_E$ samt at R_C er mindre end transistorens udgangsmodstand r_o .

$$g_m u_{BE} + \frac{u_O - u_E}{r_o} + \frac{u_O}{R_C} = 0 \Rightarrow (g_m + \frac{1}{r_o}) u_E = (\frac{1}{r_o} + \frac{1}{R_C}) u_O \Rightarrow A = \frac{u_O}{u_E} = \frac{g_m + 1/r_o}{1/r_o + 1/R_C} \approx g_m R_C$$

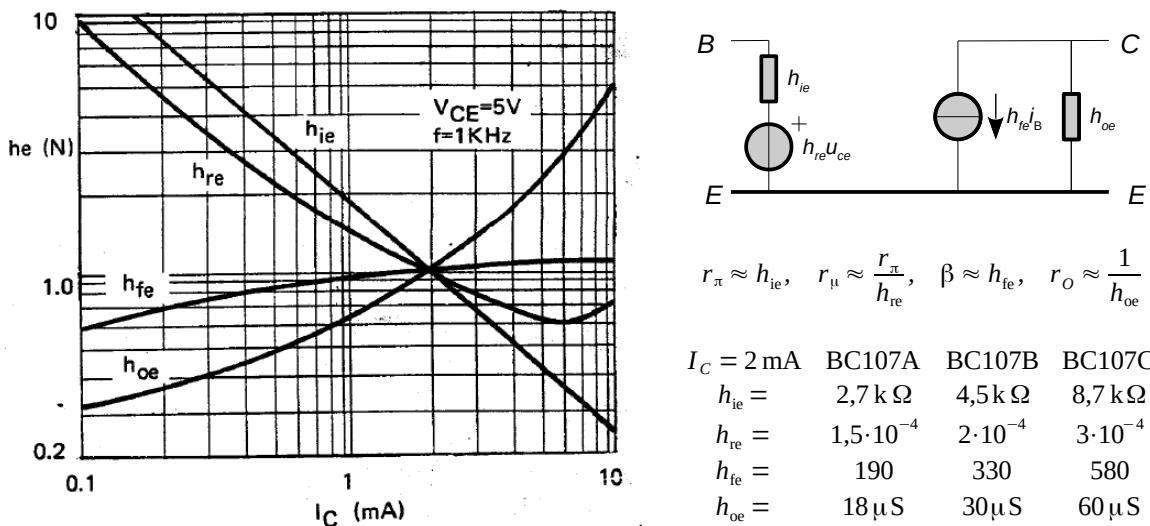
Indgangsmodstanden R_I findes ved at opstille Kirchhoffs ligning for knuden ved emitter. Den kan normalt tilnærmes ved U_T/I_C der er identisk med diodens dynamiske modstand på $R_I = 26 \Omega$ for en strøm i kollektor på $I_C = 1 \text{ mA}$.

$$-i_E + \frac{u_E}{r_\pi} + \frac{u_E}{R_E} - g_m u_{BE} = 0 \Rightarrow (\frac{1}{r_\pi} + \frac{1}{R_E} + g_m) u_E = i_E \Rightarrow R_I = \frac{u_E}{i_E} = \frac{1}{1/r_\pi + 1/R_E + g_m} \approx \frac{1}{g_m}$$

På samme måde bestemmes udgangsmodstanden R_O .

5.4.22 Publicerede parametre

Transistorens datablad oplyste tidligere h-parameter modellens parametre, men i dag er det kun få leverandører der yder denne service. Den bipolære transistor beskrives ved indgangsmodstanden h_{ie} , en tilbagevirkning fra udgangens spænding til indgangen på h_{re} , strømforstærkningen h_{fe} og en udgangsadmittans h_{oe} (reciprok modstand). Alle data publiceres som fælles-emitter så index er "e". Data oplyses ved en strøm på 2 mA i kollektor. For at omregne til en anden værdi suppleres data med et sæt kurver der viser variationen i forhold til strømmen.



Figur 273 – Et eksempel på typiske værdier for h-parameter modellen.

Relationen mellem de to modellers parametre er vist for transistoren BC107 der hedder BC547 for indkapsling i plasthus og BC847 for overflademontage¹²⁸.

Kurverne krydser ved 2 mA der er arbejdspunktet for tabellens data så eksempelvis udgangsadmittansen ændres fra $h_{oe} = 30 \mu\text{S}$ ved $I_c = 2 \text{ mA}$ ($r_o = 33 \text{ k}\Omega$) til $h_{oe} = 0,4 \cdot 30 \mu\text{S} = 12 \mu\text{S}$ ved $I_c = 0,3 \text{ mA}$ ($r_o = 83 \text{ k}\Omega$).

Ved $I_c = 2 \text{ mA}$ og en strømforstærkning på $\beta = 190$ er indgangsmodstanden på $r_\pi = 2,5 \text{ k}\Omega$ efter hybrid-pi og der aflæses $h_{ie} = 2,7 \text{ k}\Omega$ efter h-parameter modellen så de to parametre er i god overensstemmelse.

Udgangsmodenstanden r_o kan estimeres fra h-parameter modellens værdier hvor de viste data giver værdien $r_o = 17 \dots 56 \text{ k}\Omega$ ved $I_c = 2 \text{ mA}$ stigende til $76 \dots 250 \text{ k}\Omega$ ved $I_c = 0,1 \text{ mA}$. Det geometriske gennemsnit af disse værdier er på $65 \text{ k}\Omega$ ved $I_c = 0,4 \text{ mA}$

$$r_o = \frac{U_A}{I_c} = \frac{1}{h_{oe}} \Rightarrow U_A \approx \frac{I_c}{h_{oe}}$$

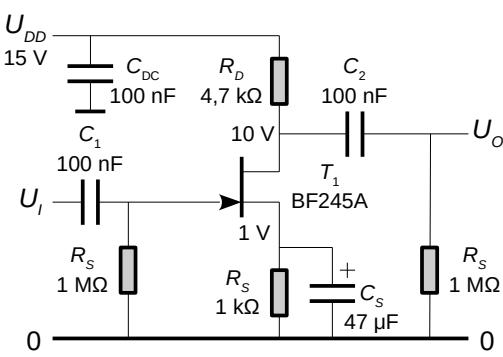
Værdien specificeres i hybrid-pi modellen ved transistorens Early-voltage værdi U_A der ganske vist ikke opgives i databladet, men kan findes i SPICE analysatorens modelbibliotek. Multisim oplyser parameteren til 80 V der svarer til 40 kΩ ved $I_c = 2 \text{ mA}$. Værdien passer med beregningen side 224, men jeg har set publicerede kurver der overhovedet ikke passer med beregninger. Med andre ord er udgangsmodenstanden en usikker parameter.

Som en brugbar tomelfingerregel kan man antage $r_o = 50 \text{ k}\Omega$ ved et lavt strømniveau.

128 Ifølge <http://en.wikipedia.org/wiki/BC548> er BC108 den oprindelige betegnelser for en NPN epitaxial transistor-serie som blev introduceret 1966 og blev populær i Europa. BC107 kan klare højere spændinger end BC108 (45 V versus 30 V) og BC109 har en højere strømforstærkning samt et lavere støjniveau. Transistoren er indkapslet i et hus af metal og keramik og er senere lanceret i plastversionerne BC547, BC548 og BC549 med de komplementære varianter som BC557, BC558 og BC559. Andre betegnelser er BC147, BC148 og BC149, med de tilsvarende PNP transistorer som BC157, BC158 og BC159. Overflademonterede varianter fås som BC846 (60 V), BC847, BC848 og BC849, med de komplementære transistorer BC856, BC857, BC858 og BC859.

5.4.23 JFET forstærker

Den spændingsstyrede strømgenerator kan udbygges til en spændingsforstærker ved at lede strømmen igennem en modstand hvorover den genererer et spændingsfald. Et AC signal på gate vil modulere strømmen i transistoren og dermed ændre på udgangsspændingen. C_1 og C_2 er DC blokeringer og AC signalforstærkningen er 3 uden kondensator C_S og 9 med. Båndbredden er opad begrænset af R_D til 1 MHz hvis udgangen belastes med 33 pF. Støj fra effektforsyningen kan kobles ind gennem modstanden i drain så afkobling er vigtig og C_{DC} udgør et minimum.



Designprocedure:

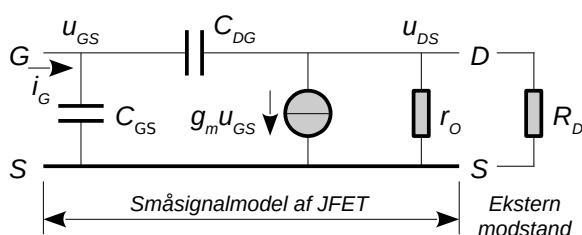
1. Vælg U_{DD} høj nok for strømregulering.
2. Vælg strømmen i drain, I_D .
3. Aflæs gate-source spændingen, U_{GS} .
4. Beregn $R_S = U_{GS}/I_D$.
5. Vælg udgangsspændingen U_D .
6. Beregn $R_D = (U_{DD} - U_D)/I_D$.
7. Beregn g_m fra formlen på næste side.
8. Beregn $A = g_m R_D / (1 + g_m R_S)$, eller ...
9. Beregn $A = g_m R_D$ for brug af C_S og ...
10. Vælg lave grænsefrekvens f_L .
11. Beregn $C_S > A_{OL}/2\pi f_L R_S$.

Figur 274 – En JFET anvendt som spændingsforstærker med en forstærkning på højest 9 gange.

Der løber en svag DC strøm i gate (5 pA for BF245 ved 25°C) og strømmen fordobles for hver cirka 10°C stigning så ved 125°C er strømmen oppe på 5 nA. Det har ingen betydning for den viste opstilling hvor det blot giver et spændingsfald på 5 mV over R_S .

5.4.24 JFET småsignalmodel

Transistorens småsignalmodel ligner hybrid-pi modellen for den bipolære transistor trods forskelle i opbygning. Det er især definitionen af transkonduktansen der danner forskellen og dertil kommer en indgangsmodstand i GΩ området som almindeligvis blot ignoreres. Der er en kapacitet mellem gate og source fra den spærrende diode i gate og tilsvarende mellem gate og drain og de er ansvarlige for en lav værdi af indgangsimpedansen ved høje frekvenser.



Overføringsfunktion:

$$A = -\frac{g_m r_o}{1+sC_{DG}r_o}, \quad f_H = \frac{1}{2\pi C_{DG} r_o}$$

Indgangsimpedans:

$$Z_I = \frac{U_{GS}}{i_G} = \frac{1}{s(C_{GS} + (1+g_m r_o)C_{DS})}$$

Transkonduktans:

$$i_D \approx g_m u_{GS} \quad \text{hvor} \quad g_m = 2 \frac{I_{DSS}}{|U_p|} \left(1 - \frac{U_{GS}}{U_p}\right)$$

$$g_m \approx \frac{I_{DSS}}{|U_p|}, \quad U_{GS} = \frac{U_p}{2}$$

Typiske parametre (BF245):

$$C_{GS} \approx 3 \text{ pF} \quad \text{og} \quad C_{DG} \approx 0,7 \text{ pF}$$

$$r_o \approx 50 \text{ k}\Omega \quad \text{ved} \quad 0,25 I_{DSS}$$

$$r_o \approx 25 \text{ k}\Omega \quad \text{ved} \quad I_{DSS}$$

Figur 275 – En model af en felt-effekt transistor for småsignalanalyse.

En BF245A med $I_{DSS} = 4 \text{ mA}$ og $U_p = -2 \text{ V}$ vil ved $U_{GS} = -1 \text{ V}$ have en strøm i drain på $I_D = 2 \text{ mA}$ og en transkonduktans på $g_m = 2 \text{ mS}$ så forstærkningen bliver $-9,4$ ved $R_D = 4,7 \text{ k}\Omega$. En BJT har ved samme strøm $g_m = 80 \text{ mS}$ så den vil i principippet kunne give en spændingsforstærkning på -376 .

I databladet betegnes g_m med g_{fs} hvor "fs" står for *forward common source* og kapaciteterne har betegnelserne C_{is} og C_{rs} frem for C_π og C_μ . Der findes litteratur hvor kapaciteterne ret logisk kaldes for C_{GS} og C_{GD} . For en opstilling som på forrige side ses det at R_D er i parallel med r_o så formlerne kan bruges ved at erstatte r_o med parallelværdien af R_D og r_o eller blot ved at benytte R_D .

Transkonduktansen findes ved at beregne tilvæksten i arbejdspunktet I_D som en følge af en tilvækst i U_{GS} hvilket svarer til differentiation. Her benyttes kæderegralen: $dy/dx = (dy/dz)(dz/dx)$. For at nå frem til udtrykket skal det indsættes at U_P er negativ.

$$g_m = \frac{dI_D}{dU_{GS}} = \frac{d}{dU_{GS}} \left(1 - \frac{U_{GS}}{U_P} \right)^2 I_{DSS} = 2 I_{DSS} \frac{d(1-U_{GS}/U_P)^2}{d(1-U_{GS}/U_P)} \frac{1}{U_P} = -2 \frac{I_{DSS}}{U_P} \left(1 - \frac{U_{GS}}{U_P} \right)$$

Udgangsmodstanden er givet ved hældningen af karakteristikken i arbejdspunktet. For BF245A findes af databladet en strømvariation på $\Delta I_D = 0,1$ mA omkring arbejdspunktet $I_D = 2,4$ mA ved ændringen ΔU_{DS} fra 5 V til 10 V så udgangsmodstanden bliver $r_o \approx 50$ k Ω .

$$r_o = \frac{dU_{DS}}{dI_D} \approx \frac{\Delta U_{DS}}{\Delta I_D} = \frac{U_{DS2} - U_{DS1}}{I_{D2} - I_{D1}}$$

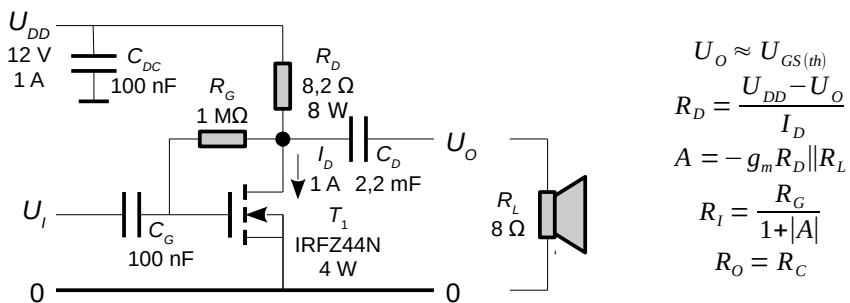
Med en spændingskilde på gate og Kirchhoff på drain kan overføringsfunktionen bestemmes.

$$\frac{U_{DS} - U_{GS}}{1/sC_{DG}} + g_m U_{GS} + \frac{U_{DS}}{r_o} = 0 \Rightarrow \frac{U_{DS}}{U_{GS}} = \frac{sC_{DG} - g_m r_o}{1 + sC_{DG} r_o} \approx -\frac{g_m r_o}{1 + sC_{DG} r_o} \quad \text{for } f \ll \frac{g_m r_o}{2\pi C_{DG}}$$

Forstærkningen er $g_m r_o = 100$ for $g_m = 2$ mS og $r_o = 50$ k Ω så værdien af ledetet sC_{DG} er uden betydning. Grænsefrekvensen ved nævnerens polynomium er 3 MHz med $C_{DG} = 1$ pF.

5.4.25 MOSFET forstærker

Det er muligt at benytte en enhancement MOSFET som en analog forstærker hvis spændingen på gate placeres der hvor transistoren trække den ønskede strømstyrke. Ved at forbinde gate til drain gennem en modstand vil spændingen på gate være lig med spændingen på drain for strømmen i gate er så lille at der ikke noget betydnende spændingsfald over R_G . DC arbejdspunktets spænding indstilles derfor ved valget af transistoren og strømmen fastsættes ved valg af modstanden R_D . En modstand fra gate til source kan bruges til at justere på arbejdspunktet.



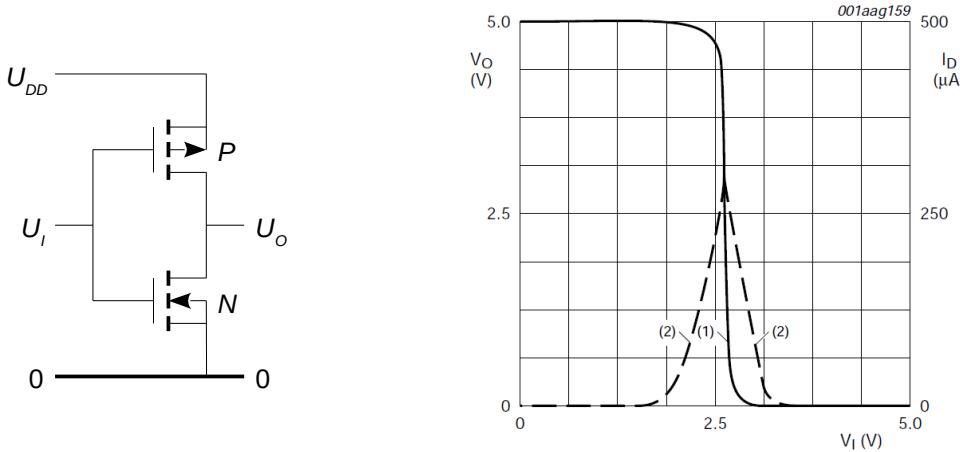
Figur 276 – En MOSFET effektforstærker der kan levere indtil 1 W til højttaleren, men da der ikke er en tilbagekobling er forvrængningen givet af transistorens udgangskarakteristik.

Den viste opstilling vil være i balance ved en spænding over drain-source på 4 V hvor transistoren trækker en strøm i drain på 1 A. Ved et analogt signal på indgangen varieres strømmen i drain og differensen fra arbejdspunktet drives gennem højttaleren. Det er et eksempel på en primitiv klasse A effektforstærker der kan afgive op til 1 W. Bemærk at der afsættes 4 W i transistoren og 8 W i modstanden så de har begge brug for køling. Transistoren har en transkonduktans¹²⁹ på 6 S ved dette arbejdspunkt så spændingsforstærkningen estimeres til 25 gange.

¹²⁹Kurverne på side 231 viser at transistorens parameter findes til $K = 9$ A/V². For en drain strøm på $I_D = 1$ A bestemmes $U_{GS} = 3,85$ V hvorved transkonduktansen beregnes til $g_m = 6$ S.

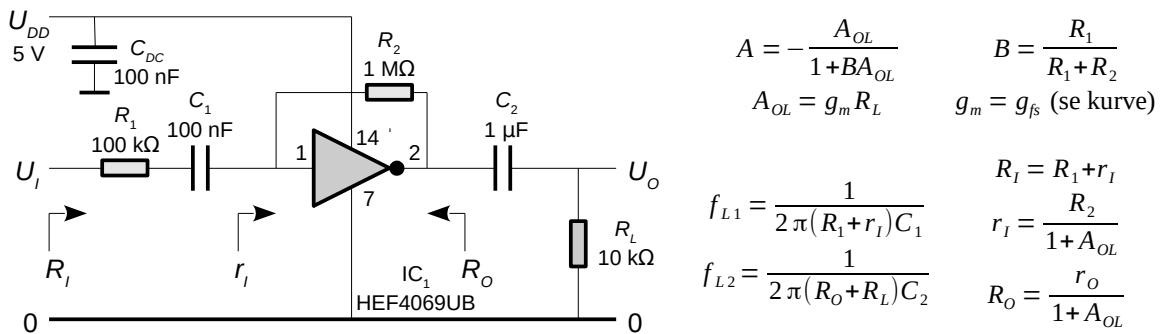
5.4.26 CMOS forstærker

To MOSFET, henholdsvis en P kanal og en N kanal, danner et komplementært par der er velegent som en digital inverter og skabte gennembrud for digitale kredsløb med lavt statisk effektforbrug. Teknologien kaldes for CMOS stående for *complementary* MOSFET. De to transistorer er begge af typen enhancement mode så der er altid én transistor ON og én transistor OFF og da de sidder i serie er serieforbindelsen altid afbrudt. Undtagelsen er ved skift mellem de logiske niveauer hvor begge transistorer kortvarigt er aktive og der løber en strøm.



Figur 277 – En CMOS inverter består af en P kanal transistor og en N kanal transistor hvor kun den ene transistor normalt er ledende. Strømforbruget er minimalt ved digitalt indgangssignal, men det stiger drastisk i området hvor udgangen skifter mellem lav og høj (NXP).

Der er normalt tre invertere i kaskade for at opnå et veldefineret skift mellem lav og høj, men der leveres CMOS invertere med kun et enkelt par af transistorer og de er gode til analoge formål. De kaldes for *unbuffered* og arbejdspunktet indstilles ved at holde ind- og udgang på samme niveau der herunder effektueres ved modstanden R_2 . Det trin anvendes i mikroprocessorer som analog forstærker i forbindelse med et krystal for at producere en stabil clock (se tredje kapitel). Med seks invertere i en DIL14 pakke og et samlet strømforbrug under 2 mA ved 5 V kan der bygges kredsløb for analog forstærkning, filtrering og oscillatorer.

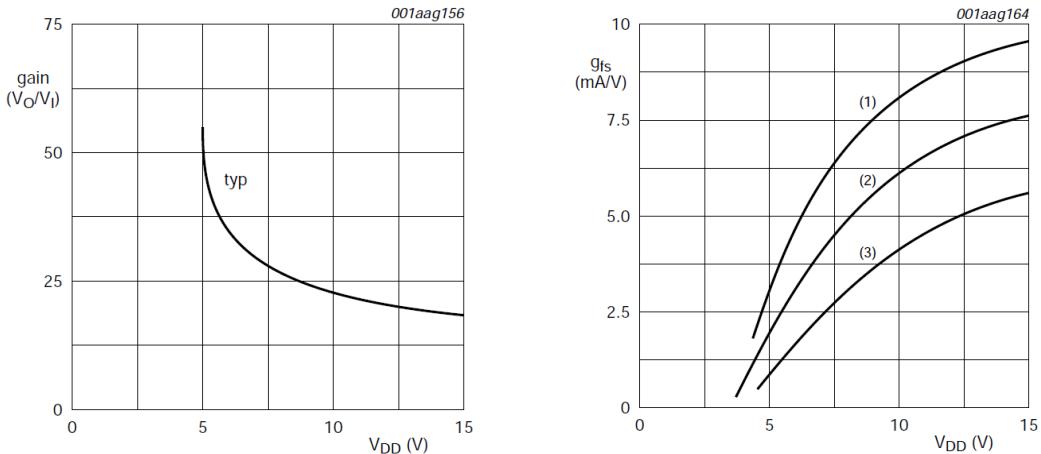


Figur 278 – En CMOS forstærker benytter et DC balancepunkt nær ved den halve effektforsyning.

Med de viste komponentværdier og $U_{DD} = 5$ V er $g_m = 2 \text{ mS}$ (se kurven på næste side) der giver $A_{OL} \approx 20$ på grund af belastningen fra R_L så forstærkningen bliver $A \approx -7$. Med $r_i \approx 50 \text{ k}\Omega$ bliver indgangsmodstanden på $R_I \approx 150 \text{ k}\Omega$. Udgangsmodstanden er $r_o \approx 3,5 \text{ k}\Omega$. Grænsefrekvenserne bliver $f_{L1} \approx 10 \text{ Hz}$ og $f_{L2} \approx 12 \text{ Hz}$.

Indgangssignalet kobles til forstærkeren gennem R_1 , forstærkningen indstilles sammen med R_2 og C_1 blokerer for DC strøm og definere den lave grænsefrekvens. På udgangen fjerner C_2 DC strøm til belastningsmodstanden R_L og danner endnu et højpasfilter. Åben-sløjfe forstærkningen er lav, omkring 20, så der skal en del trin til for at forstærke et svagt signal op til linjeniveau.

Transkonduktansen er vist herunder for en *hex inverter* HEF4069UB der indeholder i alt seks ens forstærkere. Kurverne gælder også for HEF4007 der indeholder tre komplementære par med adgang til flere af de individuelle elektroder. Kurverne er specifieret ved 5 V, men kredsen kan bruges fra 3 til 15 V hvor den nedre grænse er givet af at transkonduktansen bliver for lav og den øvre grænse er givet ved effektafsættelsen. Der er risiko for sammenbrud (*latch up*) ved en forsyningsspænding over 18 V.



Figur 279 – Åben-sløjfe forstærkning (venstre) og transkonduktans (højre) for HEF4069UB som funktion af effektforsyningen. Index er 1: maksimum, 2: typisk og 3: minimum (NXP).

Udgangsmodstanden er på $r_o \approx 25 \text{ k}\Omega$ så forstærkningen er i praksis defineret af den eksterne belastningsmodstand med mindre den ene CMOS forstærker driver den næste i en kæde.

Indgangskapaciteten er opgivet til 7,5 pF ved digital brug hvilket ignorerer tilbagekoblingen fra udgangen og det skyldes at forstærkeren arbejder med mættede transistorer i det digitale miljø. Antages tilbagekoblingen til det mindst mulige, der er omkring 2 pF, vil værdien forøges til den Miller-transformerede værdi på 40 pF (for $A_{OL} = 20$) til en værdi omkring 50 pF når kapaciteten mellem gate og forsyningerne medtages. Det svarer til en grænsefrekvens på 300 kHz hvis en CMOS udgang driver en anden CMOS indgang direkte uden at involvere andre komponenter.

Egenstøjen i en CMOS kan antages at bestå primært af haglstøj så med to transistorer og en effektforsyning på 5 V beregnes indgangens spændingsstøjtæthed til værdien $e_n = 17 \text{ nV}/\sqrt{\text{Hz}}$ hvilket svarer til støjen fra en 20 kΩ modstand.

Uden en belastning af udgangen er forstærkningen givet af produktet af transkonduktansen og udgangsmodstanden r_o og udgangsmodstanden er da givet ved de oplyste data.

$$A_{OL} = \frac{U_O}{U_I} = -g_{fs} r_o \Rightarrow r_o = \frac{|A_{OL}|}{g_{fs}}$$

Ved en tilbagekoblet forstærker reduceres udgangsmodstanden med tilbagekoblingsfaktoren.

$$R_o = \frac{r_o}{1 + B |A_{OL}|}, \quad B = \frac{R_1}{R_1 + R_2}$$

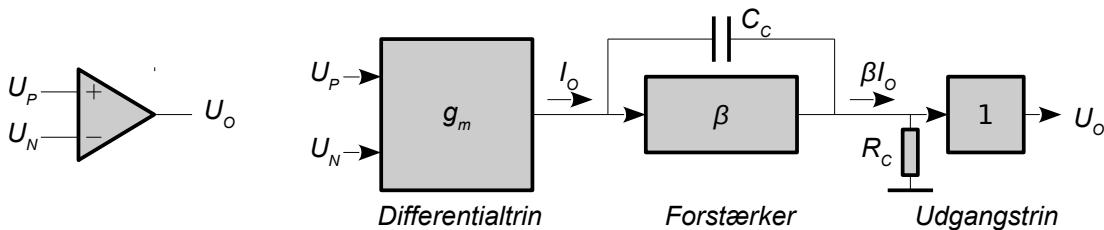
Støjstrømtægheden som følge af haglstøjen er givet af arbejdspunktets $I_D = 0,28 \text{ mA}$. Den giver et spændingsfald over udgangsmodstanden r_o og med to transistorer i kredsløbet forøges RMS værdien af støjspændingen med kvadratorden af to. Det giver spændingsstøjtægheden u_n og støjen kan henføres til indgangen som spændingsstøjtægheden e_n .

$$\begin{aligned} i_n &= \sqrt{2q_0 I_D} = 9,5 \text{ pA}/\sqrt{\text{Hz}} \\ u_n &= \sqrt{2} r_o i_n \approx 334 \text{ nV}/\sqrt{\text{Hz}} \end{aligned} \Rightarrow e_n = \frac{u_n}{A_{OL}} \approx 16,7 \text{ nV}/\sqrt{\text{Hz}}$$

Note 12.

5.5 Operationsforstærker

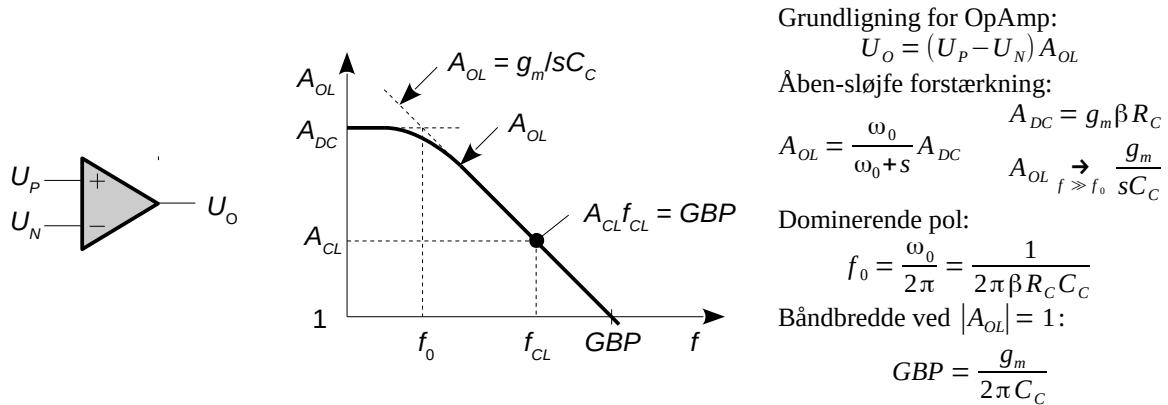
En OpAmp består af tre blokke. En spændingsdifferens ved indgangen omformes til en strøm ved første blok og forstærkes op i den midterste blok for at blive omsat til en spænding af en modstand. En buffer med en spændingsforstærkning på én driver en strøm til belastningen. Kondensatoren begrænser båndbredden og gør forstærkeren stabil, men er også med til at gøre den mere lineær.



Figur 280 – En OpAmp kan beskrives ved et differentielt indgangstrin der omsætter en spænding til en strøm ved transkonduktansen g_m og en strømforstærker β der danner en spænding over R_C .

Der er selvfølgelig variationer over temaet, men beskrivelsen dækker hovedparten af de populære operationsforstærkere samt de kredsløb der benyttes internt i integrerede kredse. I mange tilfælde er udgangsbufferen ikke påkrævet og kan udelades. Specielle operationsforstærkere som current-feedback (CFB), transkonduktansforstærkere (OTA) og transresistansforstærkere (Norton OpAmp) omtales på side 268.

En OpAmp er kendtegnet ved en meget høj forstærkning ved lave frekvenser A_{DC} og en pol f_0 der reducerer forstærkningen mod høje frekvenser. Ved lave frekvenser er forstærkningen givet ved indgangstrinnets transkonduktans g_m , det midterste trins strømforstærkning β og modstanden R_C . Ved høje frekvenser er forstærkningen givet ved transkonduktansen g_m og kondensatoren C_C og den aftager med stigende frekvens. Det fælles forløb fra lave til høje frekvenser kaldes åben-sløjfe forstærkningen A_{OL} og den dominerende pol f_0 , danner overgangen mellem de to områder.



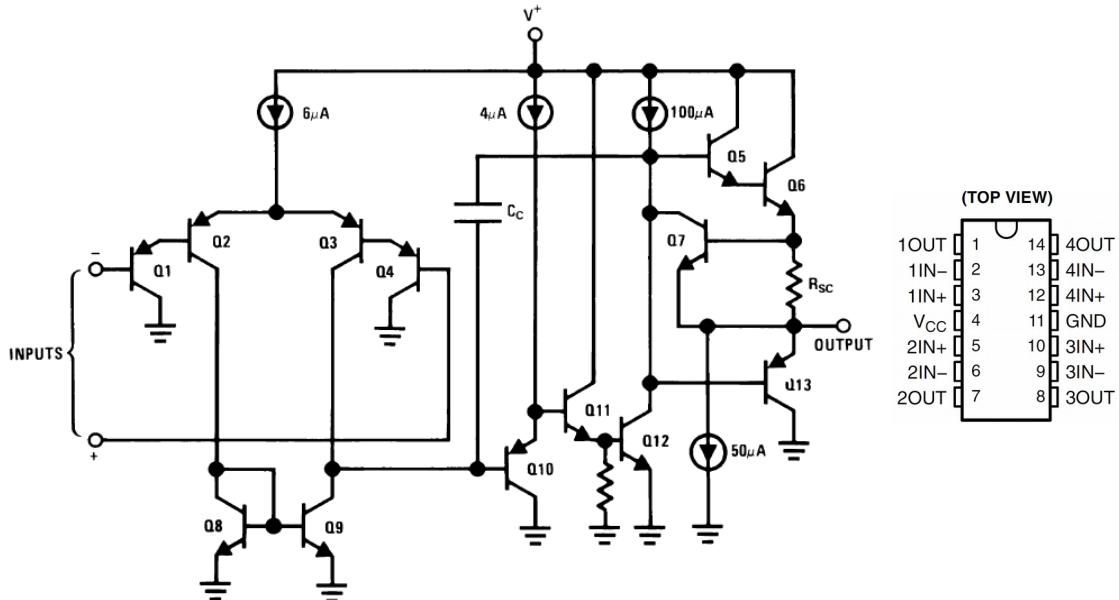
Figur 281 – Åben-sløjfe forstærkningen svarer til et lavpasfilter hvor grænsefrekvensen ofte ligger under den laveste frekvens for den pågældende applikation så forstærkeren er integrerende.

For frekvenser over den dominerende pol er forstærkningen omvendt proportionel med frekvensen og et specielt kendtegn ved en operationsforstærker er derfor at forstærkning gange båndbrede er konstant. Det kaldes for *GBP* (*gain-bandwidth product*). Forstærkningen er reduceret til 0 dB ved den frekvens og værdien angives altid i databladet. *GBP* kaldes også for *transitfrekvensen* f_T eller *unity-gain bandwidth* og værdien er udelukkende bestemt af g_m og C_C .

Fordelen ved dette design er at operationsforstærkeren er stabil ved almindelige kredsløb og man kan ofte udskyde det endelige valg af operationsforstærkeren til et senere tidspunkt.

5.5.1 LM324

Herunder ses diagrammet for en af de mest populære operationsforstærkere nogensinde. Det er den legendariske LM324 fra 1972 der har fire ens operationsforstærkere i en DIL14 pakke. Den findes også som LM2902 med lidt ringere data og som LM358 i en DIL8 pakke med to ens trin. LM324 består af de samme tre blokke som flertallet af operationsforstærker er bygget over og dertil et par elegante tricks der er værd at notere sig. Selv om det er over 40 år siden at den blev introduceret så kan den forbløffende nok tåle sammenligning med nyere kredsløb.



Figur 282 – Den populære LM324 med de væsentligste detaljer vist (National Semiconductor).

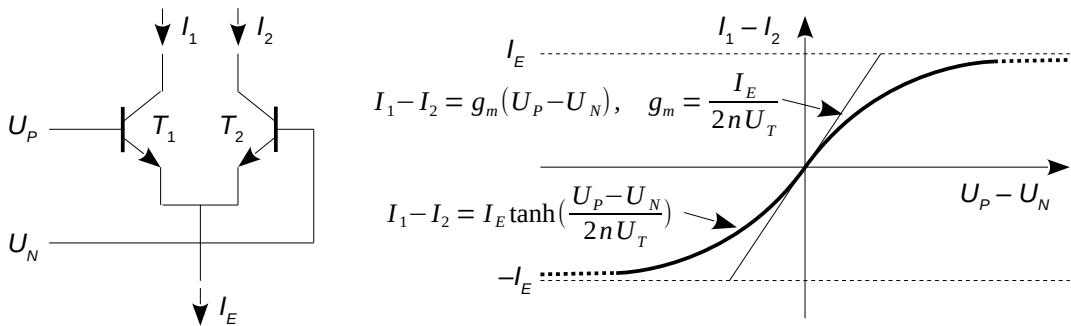
Indgangstrinnet består af differentialforstærkeren Q1 til Q4 med to PNP transistorer i hver gren der tillader at indgangen fungerer selv om transistorernes baser er under nul. Databladet tillader en indgangsspænding ned til $-0,3$ V. Indgangsstrømmen er 45 nA og varierer ikke med temperaturen så værdien kan være lavere end for MOSFET operationsforstærkere ved 125°C . Strømspejlet Q8 og Q9 omformer forskellen mellem de to trins strømme til en strømdifferens der ganges med β^3 fra Q10 til Q12 i kaskade. Strømmen omsættes til spænding ved udgangstrinnets indgangsmodstand så R_C fra side 253 er ikke en fysisk komponent, men er en funktion af den eksterne belastning. Kondensatoren C_C sætter den dominerende pol ved cirka 3 Hz og DC forstærkningen er 300.000 så GBP specifikationen bliver på 1 MHz der er en typisk værdi for operationsforstærkere.

Spændingen fra kollektor af Q12 har Q5 og Q6 som buffer for strøm ud af forstærkeren og Q13 for strøm ind i forstærkeren. Den maksimale strøm ud af operationsforstærkeren begrænses af Q7 så en kortslutning af udgangen ikke er destruktiv. Ved strøm ind er Q13 strømbegrænsende på grund af sin opbygning. Udgangen arbejder i klasse A for Q5 og Q6 holdes aktive af strømgeneratoren der er mærket $50\text{ }\mu\text{A}$ og Q13 er afbrudt da basis er positiv i forhold til emitter. Udgangstrinnet er derfor en dobbelt emitterfølger for en belastning på højest $\pm 50\text{ }\mu\text{A}$. Ved større udgangsstrøm skifter udgangstrinnet til klasse B hvor Q5 og Q6 leverer strømmen i de positive halvperioder og Q13 optager strømmen i de negative halvperioder. På grund af kompenseringen ved C_C kan kollektor af Q12 ikke flyttes hurtigere end $\pm 0,5\text{ V}/\mu\text{s}$ så signalet forvrænges ved brug af denne mulighed.

Selv om LM324 kan arbejde med indgangene under den negative linje og udgangen tæt på nul så er den ikke rail-to-rail for både ind- og udgang skal være $1,5$ V under den positive linje. LM324 kan arbejde fra 5 V eller fra en dobbelt forsyning på ± 15 V og det samlede strømforbrug er 2 mA .

5.5.2 Differentialtrin

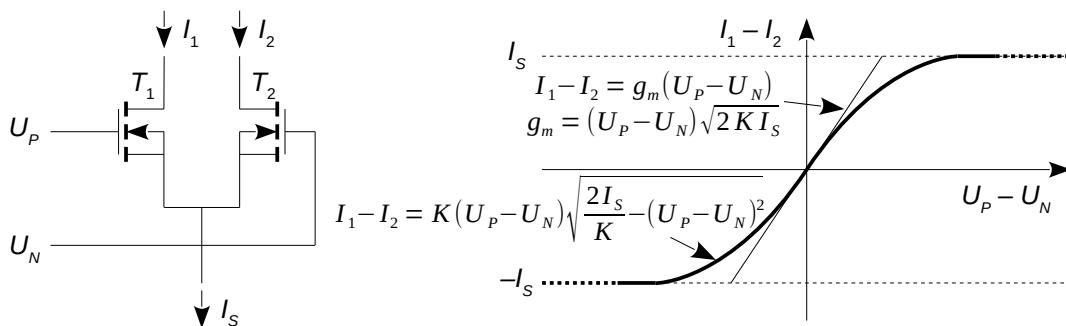
En operationsforstærker har to indgange med spændingen U_P på den positive indgang og U_N på den negative indgang. På udgangssiden løber der strømmen I_1 i transistor T_1 og I_2 i transistor T_2 . Differensen mellem de to strømme er funktion af differensen mellem de to indgangsspændinger og ved en svag udstyring er der en næsten lineær sammenhæng der kaldes for differentialtrinnets transkonduktans g_m (*transconductance*). En operationsforstærker vil normalt blive benyttet med tilbagekobling som sikrer at de to indgange ligger på meget nær det samme potentiale med kun en lille differens som forstærkes op til udgangssignalet. Med en DC forstærkning på typisk 10^5 vil en udgangsspænding fra operationsforstærkeren på 1 V blive resultatet af en forskel mellem de to indgange på $10 \mu\text{V}$ og det er så lidt at en kredsløbsanalyse normalt blot antager at $U_N = U_P$.



Figur 283 – Et differentialtrin kan opfattes som en strømfordeler. Emitterens strøm I_E er en konstant værdi, fastsat af det eksterne kredsløb og differentialtrinnet balancerer denne strøm mellem de to udgange I_1 og I_2 med fordelingen givet ved spændingsforskellen mellem de to indgange.

De to transistorer har emitter eller source koblet sammen så de deler strømmen I_E eller I_S imellem sig. Strømmen kommer fra et kredsløb der vil blive beskrevet på side 259.

Uden en differens mellem de to indgange er strømmen i de to transistorer ens så de deler den fælles strøm 50:50. I den situation er der ingen forskel mellem de to strømme så differensen er nul og i en ideel verden vil operationsforstærkerens udgang stille sig på 0 V. En lille differens mellem de to indgange vil forskyde transistorernes strømme fra 50:50 balancen så strømdifferensen ikke er nul. I den situation vil der sendes en nettostrøm til de efterfølgende kredsløb for at korrigere på udgangen så strømdifferensen reduceres til et minimum for at skabe en balancesituation.



Figur 284 – Et differentialtrin med MOSFET opfører sig stort set som ved de bipolære transistorer, dog er transkonduktansen væsentligt lavere end for BJT for samme strøm i emitter og source.

Ved en stor forskel mellem de to indgange tager den ene transistor broderparten mens den anden transistor ikke trækker nogen betydende strøm. Det giver en stor strømdifferens til de efterfølgende kredsløb som vil flytte udgangen med den størst mulige hastighed for at genoprette balancen. Ved en ekstrem spændingsdifferens hvor den ene transistor tager 100 % den fælles strøm I_E eller I_S vil strømdifferensen til det efterfølgende kredsløb aldrig kunne overstige $\pm I_E$ eller $\pm I_S$.

5.5.3 Overføringsfunktionen for bipolære transistorer

De to transistorer deler strømmen I_E fra det eksterne kredsløb så summen af de to strømme i kollektor er konstant. Det ignoreres at strømmen i basis ikke er nul, men fejlen er almindeligvis under 1 %. Uden et ydre signal ($U_P = U_N$) er arbejdspunktet for de to transistorer ens så de har den samme spænding over basis-emitter ($U_{BE1} = U_{BE2}$), og strømmen i kollektor er derfor også ens for de to transistorer nemlig $I_1 = I_2 = I_E/2$. Det er den balance som en operationsforstærker normalt holdes tæt på. Med en differens mellem de to baser som følge af et eksternt signal vil arbejdspunktet ændres så den ene transistor øger sin strømstyrke og den anden reducerer sin. Differensen $U_P - U_N$ kan beskrives som en tilvækst ved den ene transistor og en reduktion ved den anden.

$$U_{BE1} = U_{BE} + \frac{U_P - U_N}{2}, \quad U_{BE2} = U_{BE} - \frac{U_P - U_N}{2}, \quad I_E = I_1 + I_2$$

Strømmen i kollektor er givet ved den eksponentielle definition for den bipolare transistor og det vil i det følgende blive forudsat, at transistorerne er så tilpas ens at mæettingsstrømmen I_S kan regnes for den samme samt at transistorerne har samme arbejdstemperatur. Det er kun muligt at opfylde for integrerede kredsløb, men praksis viser at selv ved diskrete komponenter vil resultatet passe ganske godt med teorien. Ved at indsætte definitionen for U_{BE1} i udtrykket for strømmen i kollektor I_1 ses at det kan skrives som arbejdspunktet $I_E/2$ gange med et udtryk for den forskydning strømmen er utsat for.

$$I_1 = I_S \exp\left(\frac{U_{BE1}}{nU_T}\right) = I_S \exp\left(-\frac{\frac{U_P - U_N}{2}}{nU_T}\right) = I_S \exp\left(\frac{U_{BE}}{nU_T}\right) \exp\left(-\frac{U_P - U_N}{2nU_T}\right) = \frac{I_E}{2} \exp\left(\frac{U_P - U_N}{2nU_T}\right)$$

For transistor T_2 bliver resultatet det samme blot med negativt fortegn til differensen $U_P - U_N$ og det er nu muligt at definere forholdet mellem differensen og summen mellem strømmene. Det kan derefter udnyttes at summen er lig med I_E for at lede til den ønskede relation.

$$\frac{I_1 - I_2}{I_1 + I_2} = \frac{\exp\left(\frac{U_P - U_N}{2nU_T}\right) - \exp\left(-\frac{U_P - U_N}{2nU_T}\right)}{\exp\left(\frac{U_P - U_N}{2nU_T}\right) + \exp\left(-\frac{U_P - U_N}{2nU_T}\right)} \Rightarrow I_1 - I_2 = I_E \tanh\left(\frac{U_P - U_N}{2nU_T}\right)$$

En meget vigtig konklusion er, at den problematiske mæettingsstrøm I_S ikke indgår. For at lede frem til en småsignalanalyse benyttes en Taylor rækkeudvikling af den hyperboliske tangens.

$$\tanh(x) = x - \frac{x^3}{3} + \dots \Rightarrow I_1 - I_2 = I_E \left[\frac{U_P - U_N}{2nU_T} - \frac{1}{3} \left(\frac{U_P - U_N}{2nU_T} \right)^3 + \dots \right] \Rightarrow g_m = \frac{I_E}{2nU_T}$$

Første led angiver en lineær relation mellem strømdifferensen $I_1 - I_2$ og spændingsdifferensen $U_P - U_N$ hvilket er differentialtrinnets transkonduktans g_m . Andet led angiver en afvigelse fra den lineære relation ved forøget amplitude af indgangssignalet og det udtrykker forvrængningen i differentialtrinnet. For at definere en øvre grænse for det tilladelige signalniveau benyttes et indgangssignal som en sinus med amplituden U_M .

$$\begin{aligned} U_P - U_N &= U_M \sin(\omega t) \\ \sin^3(\theta) &= \frac{3}{4} \sin(\theta) - \frac{1}{4} \sin(3\theta) \end{aligned} \Rightarrow \frac{I_1 - I_2}{I_E} = \frac{U_M}{2nU_T} \sin(\omega t) - \frac{1}{3} \left[\frac{3}{4} \frac{U_M \sin(\omega t)}{2nU_T} - \frac{1}{4} \left(\frac{U_M}{2nU_T} \right)^3 \sin(3\omega t) \right]$$

Den tredje-harmoniske forvrængningen D_3 repræsenteres ved forholdet mellem amplituden til $\sin(3\omega t)$ og amplituden til $\sin(\omega t)$ og det ses at differentialtrinnet kan klare 18 mV for 1 % tredje harmoniske forvrængning.

$$D_3 = \frac{1}{12} \left(\frac{U_M}{2nU_T} \right)^2 \underset{n=1}{\approx} \frac{1}{48} \left(\frac{U_M}{U_T} \right)^2 \Rightarrow D_3 = 1\% \quad \text{ved} \quad U_M = U_T \sqrt{0,01 \cdot 48} = 18 \text{ mV}$$

5.5.4 Overføringsfunktionen for felt-effekt transistorer

S-kurven for en MOSFET ligner kurven for en bipolær transistor, men udledningen følger en anden vej da det er mere besværligt at tumle rundt med en kvadratisk relation.

Overføringskarakteristikken for JFET og MOSFET beskrives almindeligvis ved en kvadratisk relation der er vist herunder for transistor T_1 og tilsvarende for T_2 . Det gør det muligt at knytte et spændingsfald U_{GS1} over gate-source til en strøm i drain I_1 . Relationen er dog ikke så præcis som ved en BJT. Faktoren K er relateret til transistorens opbygning (se side 227 og 231).

$$I_1 = K(U_{GS} - U_{GS(th)})^2 \Rightarrow (U_{GS1} - U_{GS(th)})^2 = \frac{I_1}{K} \Rightarrow U_{GS1} = \sqrt{\frac{I_1}{K}} - U_{GS(th)}$$

Kirchhoff på indgangen viser at differensen mellem de to indgangsspændinger kan beskrives direkte ved en differens mellem de to transistorers gate-source spændinger.

$$-U_p + U_{GS1} - U_{GS2} + U_N = 0 \Rightarrow U_{GS1} - U_{GS2} = U_p - U_N$$

Vi kan nu opskrive udtrykket for spændingsdifferensen givet ved transistorernes strøm under den forudsætning at transistorerne er ens.

$$U_{GS1} - U_{GS2} = U_p - U_N = \sqrt{\frac{I_1}{K}} - \sqrt{\frac{I_2}{K}} \quad \text{hvor} \quad |U_p - U_N| \leq \sqrt{\frac{I_s}{K}}$$

Der kvadreres, ligningen omrokeres, der kvadreres igen og det udnyttes at $I_s = I_1 + I_2$.

$$\begin{aligned} (U_p - U_N)^2 &= \frac{I_1}{K} + \frac{I_2}{K} - 2\frac{\sqrt{I_1 I_2}}{K} \\ K(U_p - U_N)^2 - I_s &= -2\sqrt{I_1 I_2} \\ K^2(U_p - U_N)^4 + I_s^2 - 2K(U_p - U_N)^2 I_s &= 4I_1 I_2 \end{aligned}$$

Højre side kan skrives som en differens mellem to led hvor det ene er kvadratet på den fælles strøm i source og det andet er kvadratet på differensen mellem dem.

$$4I_1 I_2 = (I_1 + I_2)^2 - (I_1 - I_2)^2 \Rightarrow 4I_1 I_2 = I_s^2 - (I_1 - I_2)^2$$

Det er nu muligt at eliminere ledet med strømmen I_s i det tidligere udtryk så kun differensen ved udgangsstrømmen bliver tilbage.

$$(I_1 - I_2)^2 = 2K(U_p - U_N)^2 I_s - K^2(U_p - U_N)^4$$

Kvadratoden tages på begge sider og det giver den ønskede relation.

$$\begin{aligned} \sqrt{(I_1 - I_2)^2} &= \sqrt{2K(U_p - U_N)^2 I_s - K^2(U_p - U_N)^4} \\ I_1 - I_2 &= K(U_p - U_N) \sqrt{\frac{2I_s}{K} - (U_p - U_N)^2} \end{aligned}$$

Transkonduktansen kan bestemmes ved at føre faktoren med I_s og K udenfor kvadratoden.

$$I_1 - I_2 = K(U_p - U_N) \sqrt{\frac{2I_s}{K}} \sqrt{1 - \frac{K(U_p - U_N)^2}{2I_s}}$$

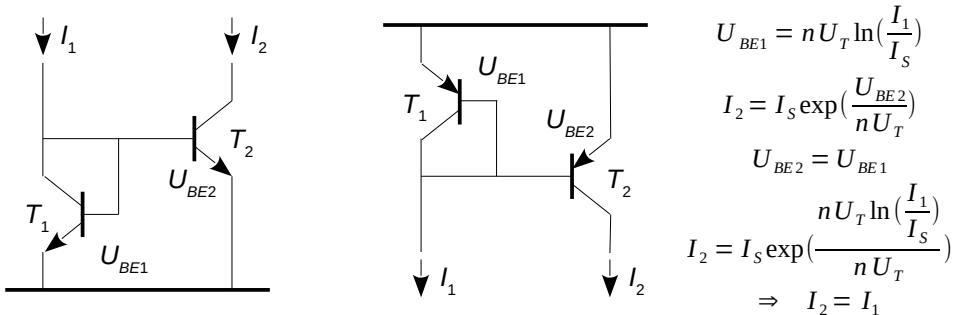
Ledet under kvadratoden er tæt på nul, idet de to transistorer deler strømmen I_s ligeligt og ved division med $U_p - U_N$ dannes et led der udtrykker ændringen i strømmen i drain som følge af en ændring i spændingen mellem indgangene og det giver transkonduktansen.

$$I_1 - I_2 \approx K(U_p - U_N) \sqrt{\frac{2I_s}{K}} \Rightarrow g_m = \frac{dI_D}{dU_{GS}} \approx \frac{\Delta I_D}{\Delta U_{GS}} = \frac{I_1 - I_2}{U_p - U_N} = \sqrt{2KI_s}$$

5.5.5 Strømspejl

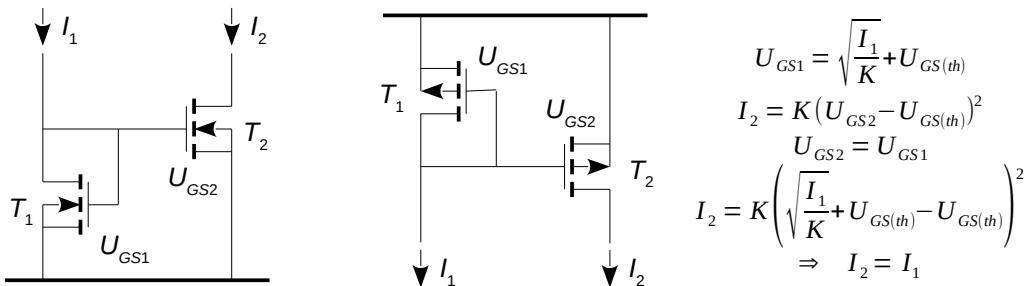
Differensen mellem de to strømme fra differentialtrinnet dannes af et strømspejl (*current mirror*) der herunder vises i sin mest enkle udførelse. Det er et populært trin i integrerede kredse hvor det benyttes både i operationsforstærkerens differentialtrin og i de øvrige kredsløb. Det ses også i kredsløb med blandet analog og digital elektronik og det benyttes i effektforstærkere for at fordele strømmen på samme måde som i operationsforstærkeren.

En strøm I_1 i den første transistor vil give en tilsvarende strøm I_2 i den anden transistor så det er muligt at bruge strømmen i ét kredsløb til at styre en strøm i et andet. Strømmen I_1 drives af det eksterne kredsløb gennem den diode-koblede transistor T_1 og spændingen over basis-emitter U_{BE1} (eller gate-source U_{GS1}) er derfor givet ved transistorens parametre. Transistor T_2 har den samme spænding over sin basis-emitter (gate-source) så strømmen i kollektor (drain) bliver den samme som strømmen i den diode-koblede transistor hvis de to transistorer er ens.



Figur 285 – Et strømspejl udnytter transistorernes ens karakteristik til at indstille strømmen i I_2 som er identisk med den strøm I_1 det eksterne kredsløb driver i T_1 .

Ved den bipolære transistor forudsætter beregningen, at der ikke løber nogen basisstrøm så det er ikke en eksakt relation, men den er tilstrækkelig til flertallet af anvendelser. Det er nærmest gratis at inkludere flere transistorer i en integreret kreds så der er mange bud i litteraturen på at forbedre kredsløbet, men det vil jeg ikke komme ind på her da det simple trin er nok til at forstå idéen.



Figur 286 – Et strømspejl med MOSFET opfører sig på samme måde som med BJT.

Relationen er gyldig både ved DC og ved analog signalbehandling der kan strække sig langt ind i MHz området som følge af de lave modstandsværdier der er involveret. Ved en BJT er den indre modstand i den diodekoblede transistor på $1/g_m = 26 \Omega$ ved $I_C = 1 \text{ mA}$. Kapaciteten C_π er typisk på 150 pF så polen ligger ved 40 MHz . For en MOSFET er den indre modstand højere, men de involverede kapaciteter er mindre så den resulterende værdi af polen flyttes ikke langt væk fra estimaten. Der er naturligvis endnu en pol ved strømspejlets udgang, men den er funktion af belastningen og den er i øvrigt emnet for operationsforstærkerens andet trin.

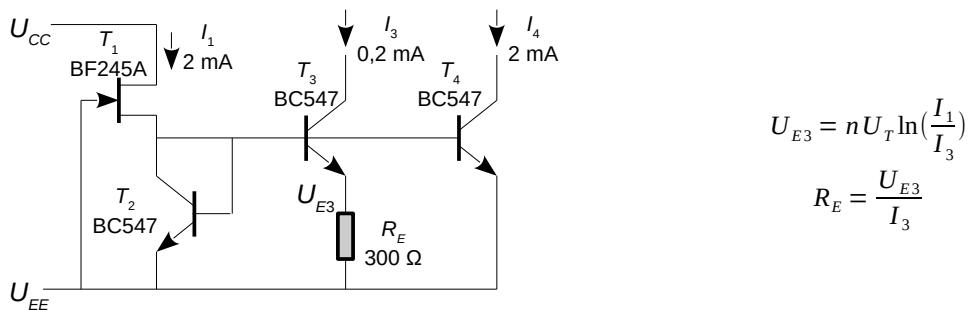
Ved BJT type PNP og P kanal MOSFET er ladningsbæreren huller hvis mobilitet er 2 til 4 gange lavere end elektronerne så de har en lavere grænsefrekvens. Den lavere mobilitet betyder dog blot at polen flytter ned i retning af 10 MHz og er derfor stadigvæk relativt højt oppe i frekvens.

5.5.6 Strømkilde

Strømkilder kan dannes af en modstand fra det aktuelle punkt i kredsløbet til effektforsyningen, men modstande egner sig ikke for integration på en siliciumchip. Der behøves ret meget areal for værdier i kiloohm-området, de genererer termisk støj og strømmen bliver funktion af spændingen over modstanden og dermed af den benyttede effektforsyning. Det sidste åbner for en uønsket indkobling af ekstern støj der er så stort et problem at det dokumenteres i databladet. I forbindelse med differentialtrinnet vil en modstand som strømkilde desuden modulere transkonduktansen hvis spændingen på de to indgange varieres. Det mindsker trinnets evne til at undertrykke støj fra en spænding på indgangene og det er en meget vigtig egenskab ved en operationsforstærker¹³⁰.

Stort set alle operationsforstærkere indeholder et kredsløb som holder de interne strømme relativt konstant¹³¹. En JFET er særlig velegnet som strømkilde og strømspejlet kan bruges til at fordele strømmen internt i kredsløbet og samtidigt tillader det en skalering af strømmen så værdien kan tilpasses de individuelle behov.

Et eksempel vises herunder hvor felt-effekt transistor T_1 har en negativ spænding på gate i forhold til source på 0,7 V og indstilles til at trække cirka 2 mA. Eksemplet forudsætter brug af BF245A og selv om det er usædvanligt at opbygge strømspejle med diskrete komponenter så kan det sagtens lade sig gøre. Med tilstrækkelig høj værdi af spændingen over drain-source vil kanalen mætte og det giver en konstant strøm til den diode-koblede transistor T_2 uanset effektforsyningen. T_2 danner et strømspejl med T_4 som derfor også trækker en strøm på omkring 2 mA.



Figur 287 – En felt-effekt transistor er en effektiv strømgenerator, og en modstand i strømspejlet kan programmere strømmen i udgangstransistoren til en lavere værdi.

Transistor T_3 styres også af spændingen over T_2 , men den inkluderer en modstand i sin emitter og det gør det muligt at reducere strømmen i T_3 overfor den strøm der løber i T_2 . Som det blev vist ved den bipolære transistor vil 60 mV reduktion af spændingen over basis-emitter reducere strømmen i kollektor med faktor 10. Derfor vil $R_E = 300 \Omega$ reducere strømmen til $I_3 = 0,2 \text{ mA}$. Denne modstand er af lav værdi og kan opbygges som metalfilm i den samme process der bruges til at forbinde de interne kredsløb i chippen. Modstande af metalfilm er stabile og relativt nøjagtige sammenlignet med alternativet der er at dotere et område af halvlederen til opgaven som en modstand. Derfor er denne type kredsløb anvendt i mange operationsforstærkere.

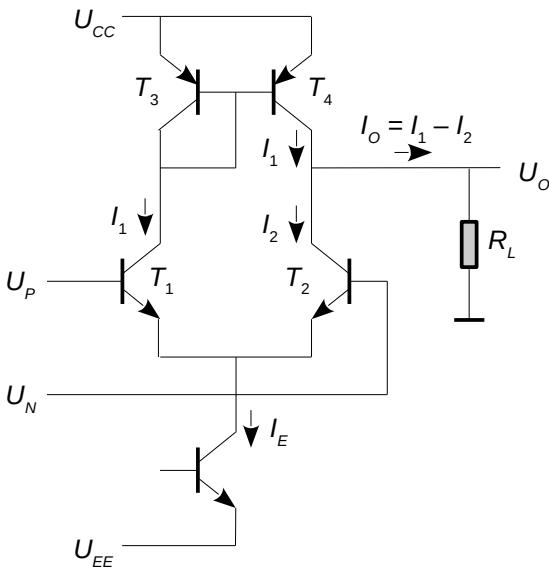
En ulempe ved opstillingen er at spændingen U_E har en positiv temperaturkoefficient på 0,35 %/K ved 60 mV som der ikke kompenseres for så for temperaturområdet -20°C til 85°C vil fejlen nå op i nærheden af $\pm 20\%$ i forhold til et design baseret på 30°C . En anden ulempe er at der ikke kan justeres på I_1 uden at relationen mellem I_3 og I_4 ændres fra det planlagte så designet skal vælges til et fast sæt af strømværdier.

130 CMRR (common-mode rejection ratio) angiver styrken af udgangssignalet fra operationsforstærkeren ved påvirkning af de to indgange med en harmonisk svining. Værdien opgives som en ækvivalent (differentiel) spændingskilde ved indgangen så der opnås en bedre specifikation ved stor forstærkning uden at det betyder en bedre undertrykkelse af signalet.

131 En undtagelse er LM386 low-voltage power amplifier der er tiltænkt ukritiske formål som en forstærker for højtalere og motorer.

5.5.7 Trin 1 – Differentialforstærker

Kombinationen af differentialtrinnet med et strømspejl giver en fleksibel byggeblok der findes i stort set enhver operationsforstærker samt flertallet af effektforstærkere og som indgår i utallige grene af analog signalbehandling. Den er kendetegnet ved at modtage en spændingsdifferens mellem de to indgange U_P og U_N og som respons herpå driver den en strøm I_o i udgangen. Proportionaliteten mellem strøm og spænding kaldes for trinnets transkonduktans g_m .



Overføringsfunktion:

$$U_O = (U_P - U_N) A_{OL}$$

Forstærkning:

$$A_{OL} = g_m R_L \Rightarrow A_{OL} = \frac{R_L}{2 U_T} I_E$$

Forstærkningen er én når:

$$A_{OL} = 1 \Rightarrow g_m = \frac{1}{R_L}$$

Ligninger:

$$I_O = g_m (U_P - U_N)$$

$$U_O = R_L I_O$$

$$g_m = \frac{I_E}{2 U_T}$$

$$U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$$

Figur 288 – Et differentialtrin plus strømspejl danner en fleksibel blok for analog signalbehandling der findes i både operationsforstærkere og effektforstærkere.

I diagrammet er ikke alle detaljer medtaget idet strømmen I_E blot er vist som en enkelt transistor med kollektor forbundet til den fælles emitter. Det kunne eksempelvis være T_3 fra strømkilden vist på forrige side mens simplere kredsløb blot benytter en modstand til U_{EE} i stedet for transistoren.

Det der for alvor gør blokken interessant ved analog signalbehandling er at transkonduktansen g_m er funktion af strømmen I_E uden at strømmen selv optræder i udgangssignalet. Det betyder at det analoge signal spændingsdifferensen $U_P - U_N$ ganges med det analoge signal strømmen I_E . Som kredsløbsdesigner kan trinnets transkonduktans derfor programmeres til opgaven ved at tilpasse strømmen I_E efter behov. Det ses udført i operationsforstærkeren TLC271 hvor designeren kan vælge mellem en stor forstærkning eller et lavt effektforbrug ved omskiftning af et af benene.

Differentialtrinnet forvrænger signalet ved mere end cirka 10 mV differentielt signal. En almindelig operationsforstærker benyttes altid med en ekstern tilbagekobling (typisk to modstande) så de to indgange holdes tæt på hinanden. Dermed er forvrængningen af mindre betydning når blot åbent-sløjfe forstærkningen er tilstrækkelig høj fordi tilbagekoblingen så holder differensen nær nul.

Det er muligt at korrigere for forvrængningen ved en mod-forvrængning af signalet inden det føres til indgangen. Det benyttes i de moderne transkonduktansforstærkere LM13600 og LM13700 som anvendes ved analog multiplikation hvor en tilbagekobling ikke er mulig. De indeholder dioder i indgangen for at varetage denne opgave¹³².

Differentialtrinnet anvendes bredt indenfor analog signalbehandling hvor Gilbert cellen (*Gilbert cell*) skal nævnes¹³³. Den anvendes som phasedetektor i faselåste oscillatorer og ved radiotransmission som en synkron detektor (mikser)¹³⁴.

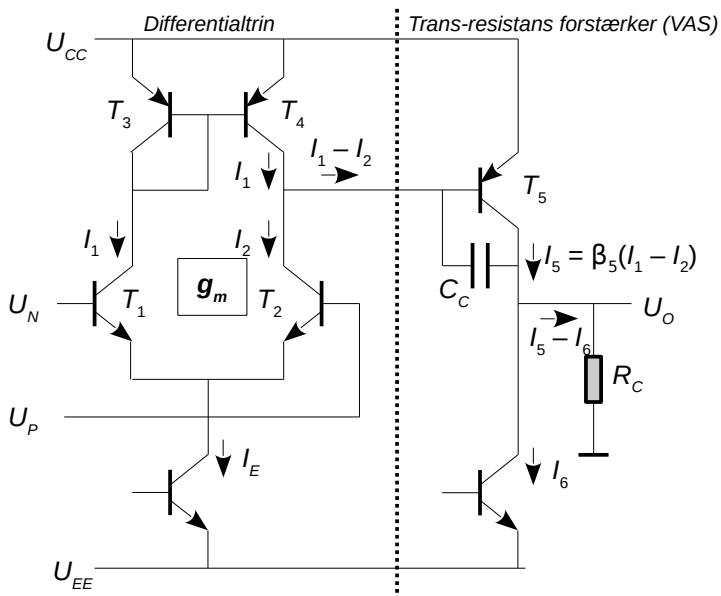
132 Se også min hjemmeside <http://www.torean.dk/artikel/publikationer.html> hvor du skal hente dokumentet "Gain Control".

133 Den blev opfundet af Howard Jones i 1963.

134 https://en.wikipedia.org/wiki/Gilbert_cell.

5.5.8 Trin 2 – Forstærker (VAS)

Vi er nu fremme ved operationsforstærkerens trin 2 der modtager strømmen fra differentialtrinnet, forstærker den op med strømforstærkningsfaktoren β og leder den igennem modstanden R_C hvor strømmen konverteres til en spænding. Det kaldes også for det spændingsforstærkende trin VAS (*voltage-amplification stage*) selv om det strengt taget er en trans-resistans forstærker med strøm ind og spænding ud. DC forstærkningen er givet af transkonduktansen g_m , forstærkerens β_5 for T_5 og modstanden R_C så trods det relativt komplicerede diagram er der kun få betydende parametre. Transistor T_5 har sin emitter ved den positive effektforsyning i figuren herunder så kollektoren kan stort set svinge frit mellem grænserne defineret ved effektforsyningen.



DC forstærkning:

$$A_{DC} = g_m \beta_5 R_C$$

Dominerende pol:

$$f_0 = \frac{1}{2\pi r_{\pi 5} C_{CM}} = \frac{1}{2\pi \beta_5 R_C C_C}$$

Slew rate:

$$SR = \frac{I_E}{C_C}$$

Relationer (M = Miller):

$$r_{\pi 5} = \frac{\beta_5}{g_{m5}} \quad C_{CM} = g_{m5} R_C C_C$$

$$g_m = \frac{I_E}{2 U_T} \quad g_{m5} = \frac{I_{C5}}{U_T}$$

Figur 289 – Kombinationen af et differentialtrin og et strømspejl giver en opstilling hvor indgangens differentielle spænding omsættes til en bipolær strøm på udgangen. En femte transistor kan give yderligere forstærkning af strømmen inden den føres til en belastningsmodstand og ender som en spænding. Signalinversionen ved T_5 betyder at den positive og negative indgang bytter plads.

Med $I_E = 200 \mu A$ er $g_m = 3,8 \text{ mS}$. Hvis T_5 har strømforstærkningen $\beta_5 = 500$ og belastningen er $R_C = 1 \text{ k}\Omega$ bliver DC forstærkningen $A_{DC} = 1900$ og den stiger mod 19000 hvis belastningen hæves til $10 \text{ k}\Omega$. Ved en højere modstandsværdi bliver forstærkningen dog begrænset af transistorens indre modstand.

For at opnå stor DC forstærkning kan strømmen I_E i differentialtrinnet emitter indstilles til en høj værdi for åben-sløjfe forstærkningen er proportional med denne parameter. Der kan også benyttes en transistor T_5 med høj strømforstærkning og i praksis benyttes et antal transistorer i kaskade for at gange deres strømforstærkninger sammen. Endelig kan værdien af R_C sættes op ved brug af en emitterfølger som det tredje trin og det behandles i næste afsnit.

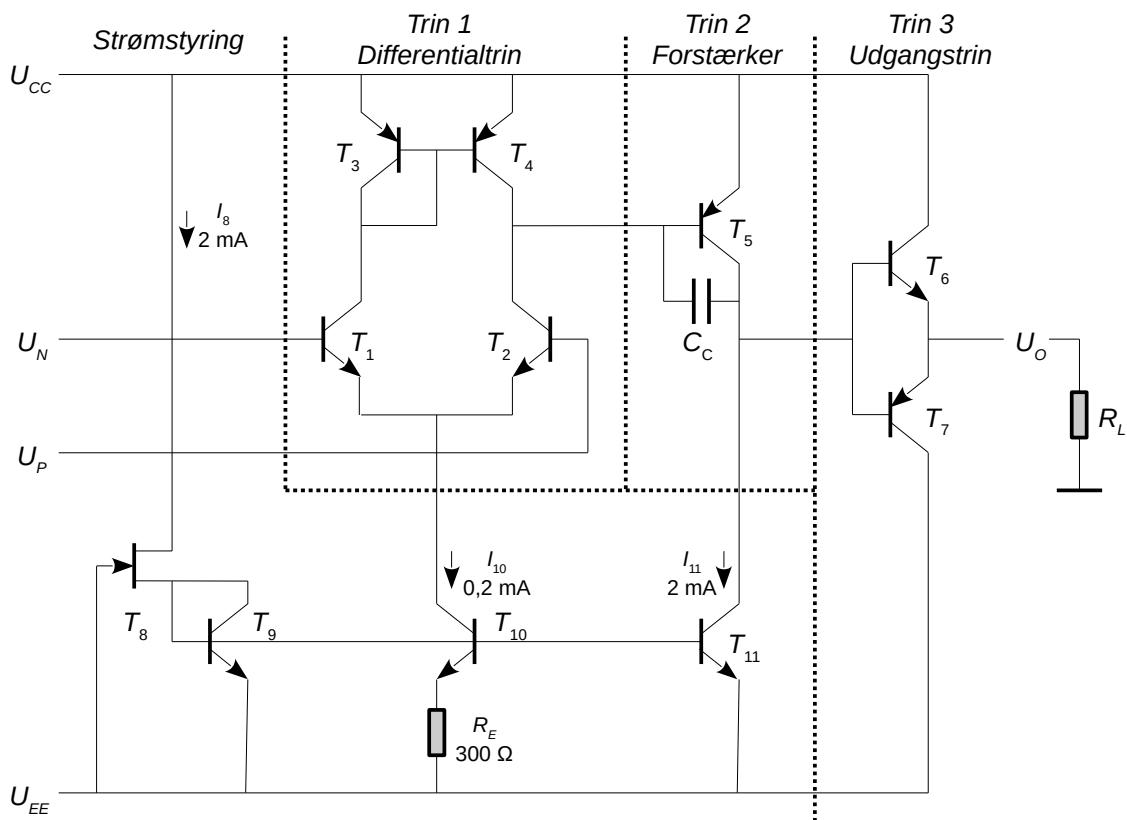
Basisstrømmen til T_5 kan kun løbe én vej og det kræver en ubalance i differentialtrinnet ($I_2 > I_1$) for at optage basisstrømmen fra T_5 og trinnet må derfor køre en smule skævt, men det har ikke større betydning i praksis da emitterstrømmen I_E er meget større end basisstrømmen. Med MOSFET som forstærkende komponent løber der ingen strøm i gate så her er balancen ikke påvirket.

Kondensator C_C sænker polen ved T_5 så den bliver dominerende og sikrer stabilitet. Fasedrejet bliver -90° for frekvenser over den dominerende pol og op indtil åben-sløjfe forstærkningen når ned på én uanset de højere liggende poler. For at flytte udgangens spænding skal kondensator C_C op- eller aflades og her er det vigtigt at $I_1 - I_2$ er en bipolar strøm. Der er dog også en bagside af medaljen for strømmen kan højest være $\pm I_E$ så der er en grænse for hvor hurtigt spændingen kan varieres. Det beskrives ved slew rate specifikationen der uddybes på side 265.

5.5.9 Trin 3 – Udgangstrin

Operationsforstærkere med to trin benyttes internt i større integrerede kredse hvor værdien af R_C er en kendt størrelse og man derfor kan forholde sig til at forstærkningen er funktion af den¹³⁵.

For en bredt anvendelig operationsforstærker skal belastningens betydning reduceres og en god mulighed er at inkludere én eller flere emitterfølgere som adskillelse mellem modstanden R_C og belastningen R_L . Herunder er det vist med både en NPN og en PNP hvor basis-emitter dioderne er koblet sammen så kun én af transistorerne kan være aktiv til et givet tidspunkt. Det fungerer som en symmetrisk emitterfølger der kan klare begge polariteter af signalet. I praksis er det ikke en god opstilling, men det problem udskydes til næste kapitel om effektorstærkere.



Figur 290 – Model af en OpAmp med differentialtrin T_1 til T_4 , forstærker T_5 og buffer ved T_6 og T_7 . En felt-effekt transistor T_8 giver en strøm der er uafhængig af forsyningsspændingen og som driver strømmen i differentialtrinnet og forstærkeren ved strømspejlene T_9 til T_{11} .

Indgangsmodstanden på basis af en emitterfølger er strømforstærkningen gange belastningen af udgangen. En belastning med $R_L = 100 \Omega$ giver derfor en indgangsmodstand på $50 \text{ k}\Omega$ ved en strømforstærkning på 500 i udgangstransistorerne. Det betyder at DC åben-sløjfe forstærkningen bestemmes af udgangsmodstanden fra T_5 og T_{11} i parallel med indgangsmodstanden og da de alle er af samme størrelsesorden vil forstærkningen kunne beregnes fra $R_C = 10 \text{ k}\Omega$ som et kompromis. Med de viste strømme i diagrammet er transkonduktansen af differentialtrinnet på $g_m = 3,8 \text{ mS}$ og hvis strømforstærkningen af T_5 antages til $\beta_5 = 500$ bliver åben-sløjfe forstærkningen $A_{OL} = 20000$. Værdien bliver uafhængig af belastningen på udgangen når $R_L > 100 \Omega$.

Der er flere eksempler på nettet¹³⁶. Operationsforstærkere baseret på en kraftig udvidelse af det viste tre-trins kredsløb markedsføres som moduler for kommersiel salg¹³⁷.

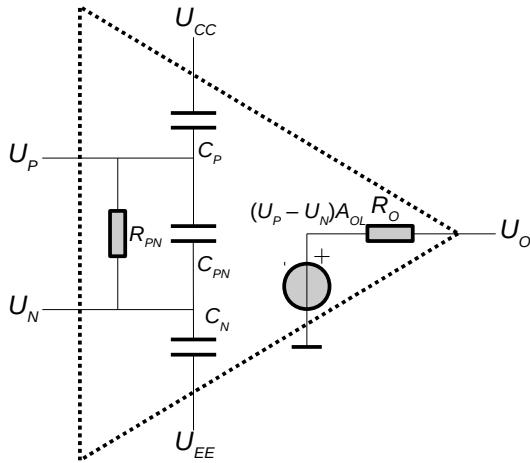
135 http://www.ijetae.com/files/Volume3Issue12/IJETAE_1213_17.pdf.

136 http://www.sg-acoustics.ch/analogue_audio/discrete_opamps/.

137 <http://www.johnhardyco.com/pdf/990.pdf>.

5.5.10 Ind- og udgangsimpedans

På nuværende tidspunkt er kredsløbet så godt defineret at det muliggør præsentationen af nogle få parametre for operationsforstærkeren der refererer til indgangstrinnet og udgangsmodstanden.



Differentiel indgangsmodstand:

$$R_{PN} \approx 2r_\pi = 2 \frac{\beta}{g_m} = \frac{4\beta U_T}{I_E}$$

Differentiel indgangskapacitet:

$$C_{PN} \approx \frac{C_\pi}{2} \approx \frac{1}{2} \frac{g_m}{2\pi f_T} = \frac{I_E}{8\pi U_T f_T}$$

$f_T \approx 10 \dots 100 \text{ MHz}$ typisk

Common-mode indgangskapacitet:

$$C_P \approx C_N \approx C_\mu$$

$C_\mu \approx 3 \text{ pF}$ typisk

Udgangsmodstand:

$$R_O \approx \frac{r_{O5} \| r_{O11}}{\beta_6} + \frac{1}{g_m} \approx \frac{r_O}{2\beta_6} + \frac{U_T}{I_O}$$

$R_O \approx 100 \Omega$

Figur 291 – Ind- og udgangstrinnet af en OpAmp består af en impedans mellem de to indgange en kapacitet fra hver af indgangene til referencepunktet samt en udgangsmodstand.

Indgangstrinnet består af to transistorer med fælles emitter så de to komponenter r_π og C_π fra hybrid-pi modellen er i serie og den resulterende impedans bliver $2r_\pi$ i parallel med $C_\pi/2$. Det giver en indgangsmodstand på $R_{PN} = 160 \text{ k}\Omega$ og en indgangskapacitet på $C_{PN} = 25 \text{ pF}$ ved en strøm i indgangstrinnet på $I_E = 0,2 \text{ mA}$ og med $\beta = 300$. De to værdier indikerer at indgangsimpedansen er reel op til cirka 50 kHz hvorover den aftager. Felteffekt transistorer vil have en væsentligt højere indgangsmodstand, typisk i $\text{G}\Omega$ området, og en lavere indgangskapacitet.

Kapaciteten i transistorernes kollektor-basis, eller drain-gate, er forbundet til effektforsyningen idet potentiometret ved kollektor eller drain af T_1 og T_2 er næsten konstant. I praksis regnes kapaciteten som en afkobling mod den fælles reference nul selv om den terminal ikke er findes for en normal operationsforstærker. Det er tilladeligt idet enhver anvendelse vil inkludere kondensatorer fra U_{CC} til nul samt fra U_{EE} til nul for afkobling af effektforsyningen. Værdien inkluderer ikke kapaciteten af den interne opbygning af operationsforstærkeren og heller ikke af banerne på printkortet. En mere realistisk værdi vil være omkring 5 pF fra hver indgang til nul når disse værdier inkluderes.

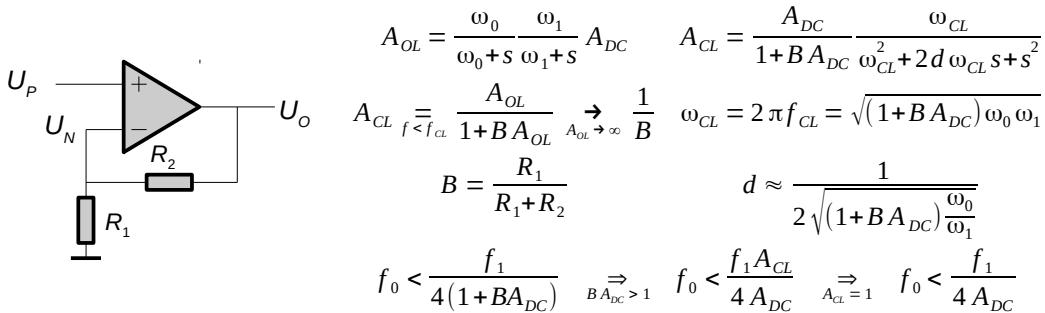
Udgangsmodstanden R_O er givet ved modstanden i punktet ved kollektor eller drain af T_5 divideret med værdien af strømforstærkningen for den aktive udgangstransistor (β_6 eller β_7). Modstanden r_o for T_5 og T_{11} kan beregnes fra h-parametermodellens h_{oe} eller transistorens Early Voltage som har en værdi i omegnen af 80 V for typiske bipolære transistorer så værdien er omkring 40 k Ω . Dertil skal lægge den indre modstand fra den aktive udgangstransistor som er $1/g_m$ og derfor er funktion af udgangsstrømmen. Benyttes $I_O = 0,5 \text{ mA}$ findes den indre modstand til 50 Ω og den samlede udgangsmodstand for operationsforstærkeren bliver af størrelsesordenen 100 Ω .

For en MOSFET kan udgangsmodstanden beregnes fra $1/g_m$ alene da r_π for en BJT ikke indgår i modellen for en felt-effekt transistor. Transkonduktansen er væsentligt mindre for en MOSFET end for en BJT ved samme strømstyrke så udgangsmodstanden bliver let over 1 k Ω .

Modellen viser en spændingskilde i udgangen med reference til nul hvilket er almindeligt ved det analytiske arbejde, men nul-terminalen findes ikke i en reel operationsforstærker. Udgangen må derfor referere til U_{CC} eller U_{EE} hvilket giver en mulig støjindkobling gennem effektforsyningen. Af samme årsag anbefales det altid at afkoble effektforsyningens linjer til nul gennem kondensatorer med en værdi på cirka 100 nF der har en impedans under 100 Ω for frekvenser over 10 kHz.

5.5.11 Stabilitet

Det er ikke tilstrækkeligt at beskrive en reel operationsforstærker ved den dominerende pol alene for der er altid en eller flere højtliggende poler. Ved tilbagekobling giver det en overføringsfunktion af mindst anden orden og det kan give problemer med stabilitet. For at gardere sig kan man vælge den dominerende pol f_0 så lavt at systemet bliver ubetinget stabilt. Det nås når fasedrejet fra den laveste af de høje poler er uden betydning inden forstærkningen er nået ned på én. Ved en høj pol på $f_1 = 1 \text{ MHz}$ og en DC forstærkning på $A_{DC} = 10^5$ vil $f_0 < 2,5 \text{ Hz}$ give ubetinget stabilitet.



Figur 292 – En tilbagekoblet operationsforstærker skal have en dominerende pol for stabilitet.

Overføringsfunktionen bestemmes ved at indsætte åben-sløjfe forstærkningen A_{OL} i definitionen for forstærkningen A_{CL} og efter reduktion ender det med en anden-ordens overføringsfunktion.

$$A_{CL} = \frac{\frac{\omega_0}{\omega_0+s} \frac{\omega_1}{\omega_1+s} A_{DC}}{1+B \frac{\omega_0}{\omega_0+s} \frac{\omega_1}{\omega_1+s} A_{DC}} = \frac{\omega_0\omega_1 A_{DC}}{(\omega_0+s)(\omega_1+s) + B\omega_0\omega_1 A_{DC}} = \frac{\omega_0\omega_1 A_{DC}}{\omega_0\omega_1 + (\omega_0+\omega_1)s + s^2 + B\omega_0\omega_1 A_{DC}}$$

Nævnerens konstanter samles til $(1+BA_{DC})\omega_0\omega_1$ der identificeres som kvadrateet på systemets grænsefrekvens ved $\omega_{CL} = 2\pi f_{CL}$. Overføringsfunktionen skrives med $(1+BA_{DC})\omega_0\omega_1$ i tælleren og giver et udtryk for DC forstærkningen A_{DC} og et anden-ordens lavpasfilter med dæmpningen d givet ved ledet af første-orden som koefficienten $2d\omega_{CL}$.

$$A_{CL} = \frac{A_{DC}}{1+BA_{DC}} \frac{(1+BA_{DC})\omega_0\omega_1}{(\omega_0+\omega_1)\omega_{CL}s + s^2} \Rightarrow$$

$$A_{CL} = \frac{A_{DC}}{1+BA_{DC}} \frac{\omega_{CL}}{\omega_{CL}^2 + 2d\omega_{CL}s + s^2} \quad \text{hvor} \quad \omega_{CL} = 2\pi f_{CL} = \sqrt{(1+BA_{DC})\omega_0\omega_1}$$

$$d = \frac{\omega_0+\omega_1}{2\omega_{CL}}$$

Dæmpningen ses at være relateret til sløjfeforstærkningen $1+BA_{DC}$ hvor B er givet af den indstillede forstærkning. Forholdet ω_0/ω_1 angiver hvor lav den dominerende pol er i forhold til den højere pol. Det udnyttes at den dominerende pol er meget lavere end den høje pol så tælleren kan simplificeres.

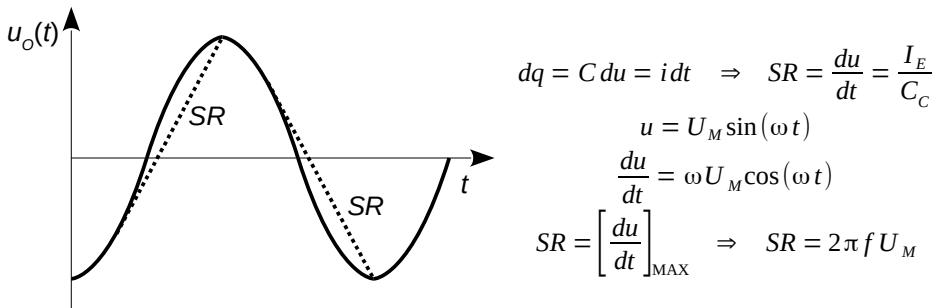
$$d = \frac{\omega_0+\omega_1}{2\sqrt{(1+BA_{DC})\omega_0\omega_1}} \approx \frac{\omega_1}{2\sqrt{(1+BA_{DC})\omega_0\omega_1}} = \frac{1}{2\sqrt{(1+BA_{DC})\frac{\omega_0}{\omega_1}}}$$

For at opnå en stabil kobling med operationsforstærkeren kræves det at dæmpningen er mindst én svarende til at den resulterende overføringsfunktion har to reelle poler. Det giver en binding for den dominerende pols frekvens f_0 overfor den høje pol f_1 . Det kan normalt antages at BA_{DC} er meget større end én og derved fremkommer den simplificerede formel i illustrationen.

$$d \approx \frac{1}{2\sqrt{(1+BA_{DC})\frac{\omega_0}{\omega_1}}} > 1 \Rightarrow 2\sqrt{(1+BA_{DC})\frac{\omega_0}{\omega_1}} < 1 \Rightarrow (1+BA_{DC})\frac{f_0}{f_1} < \frac{1}{4} \Rightarrow f_0 < \frac{f_1}{4(1+BA_{DC})}$$

5.5.12 Slew rate

Med kondensator C_C for at skabe en dominerende pol er der en grænse for hvor høje frekvenser udgangen kan gengive. Differentialtrinnet kan højest afgive en strøm på $\pm I_E$ fra udgangen af strømspejlet og denne strøm skal op- og aflade kondensatoren C_C . Følgelig er betydningen af kondensatoren at der er en grænse for hvor hurtigt operationsforstærkerens udgang kan flytte sig. Forsøges grænsen overskredet vil udgangen flytte sig med en konstant hastighed der uafhængigt af indgangssignalets amplitude hvilket kaldes for *slew rate*. Det er en spændingshastighed, men der ikke er en god dansk betegnelse og derfor benyttes det engelske slew rate.



Figur 293 – Den dominerende pol giver stabilitet, men begrænser også hvor hurtigt udgangen kan flyttes. Hvis værdien forsøges overskredet bliver signalet forvrænget.

Operationsforstærkerens slew rate kan beregnes ud fra kondensatorenligningen og med $I_E = 1 \text{ mA}$ ved differentialtrinnet og $C_C = 100 \text{ pF}$ findes $SR = 10 \text{ MV/s}$ der også skrives som $10 \text{ V/}\mu\text{s}$.

Specifikationen har megen fokus indenfor lydbehandling hvor det er den holdning at værdien skal være så høj som muligt dels for at kunne reproducere et musikstykke korrekt og dels for at den aktuelle operationsforstærker "lyder godt". Et musikstykke er frekvensbegrænset til cirka 20 kHz og ved et maksimalt spændingssving på $\pm 15 \text{ V}$ er den højest mulige hastighed på 2 MV/s.

Man vil normalt indkalkulere en sikkerhedsmargin så differensen mellem operationsforstærkerens to indgange holdes tæt på nul. Hele argumentationen for en operationsforstærker er jo at dens indgangstrin arbejder mest lineært når de to indgange er på næsten det samme potentielle og der er et forstærkningsoverskud til at forstærke en lille differens op til udgangssignalet.

Vi så tidligere at der er en harmonisk forvrængning på 1 % i indgangstrinnet ved cirka 20 mV amplitude af signalet mellem de to indgange. Ved den grænse vil der kun udnyttes omkring 20 % af differentialtrinnets udstyringsmulighed på $\pm I_E$ og specifikationen bør derfor være tilsvarende højere med 10 MV/s som et bud for at kunne håndtere et kraftigt og højfrekvent udgangssignal. Den værdi kan benyttes som et udgangspunkt ved design med operationsforstærkere hvor der benyttes en effektforsyning på $\pm 15 \text{ V}$ og ved effektforstærkere bør værdien mindst være det dobbelte på grund af det større spændingssving på udgangen.

Til sammenligning har operationsforstærkere som NE5534 og TL071 en specifikation på 13 MV/s og de har opnået bred accept indenfor lydbehandling.

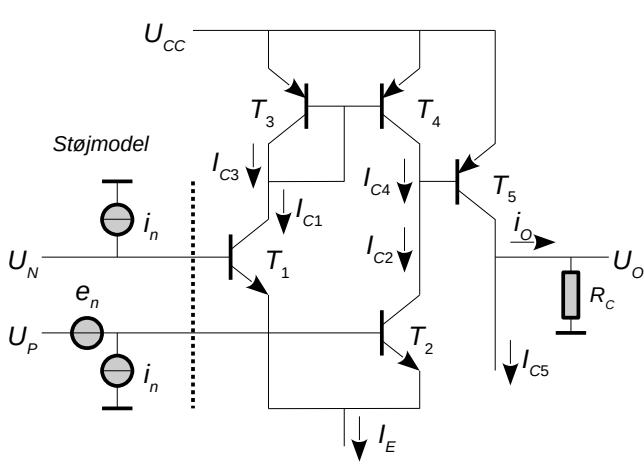
Denne introduktion til slew rate skyldes de debatter der fra tid til anden raser på internettet. Der fremsættes ofte ret fantastiske påstande om hvor stor en slew rate en forstærker behøver og det forsøges så skudt ned i debatspalterne med reference til hvor meget der er nødvendigt for at kunne gengive musik. Som mit eksempel viser er der en fornuftig grund til at overdimensionere og ikke blot udnytte operationsforstærkeren til sidste blodsdråbe.

Man bygger jo heller ikke en bro så den kun lige akkurat kan klare den typiske spidsbelastning.

5.5.13 Egenstøj

Egenstøjen fra en operationsforstærker skyldes især haglstøj (*shot noise*). Strømmen bæres af elektronener med en fast ladning og det giver støj fra kvantificeringen når elektron efter elektron slipper igennem en halvleders PN overgang. Den statistiske fluktuation omkring middelværdien af strømmen virker som en støjkilde. Der er selvfølgelig også bidrag af termisk støj fra kredsløbets modstande, men dem er der ikke mange af i en operationsforstærker dog med undtagelse af r_x ved de bipolare transistorer. Med cirka 100Ω for småsignaltransistor er spændingsstøjtætheden e_n af størrelsesordenen $1 \text{ nV}/\sqrt{\text{Hz}}$ og er kun betydnende for støjsvage operationsforstærkere.

Støjmodellen fra tredje kapitel benytter spændingsstøjtætheden e_n og to strømstøjtætheder i_n , der begge kan findes ud fra transistorernes haglstøj. Der er andre støjkilder i både BJT og MOSFET, men resultatet er ret tæt på det der kan måles.



$$\begin{aligned} \text{Haglstøj: } & i_n \stackrel{\text{def}}{=} \sqrt{2q_0 I_{DC}} \\ \text{Strømstøjtæthed: } & i_n = \sqrt{\frac{q_0 I_E}{\beta}} \quad \text{BJT} \\ & i_n = \sqrt{2q_0 I_G} \quad \text{MOSFET} \\ \text{Haglstøj til } T_5 \text{ basis: } & i_{nC} = \sqrt{i_{nC1}^2 + i_{nC2}^2 + i_{nC3}^2 + i_{nC4}^2} \\ & i_{nC} = 2\sqrt{q_0 I_E} \\ \text{Støjspænding: } & e_{nO} = 2\beta_5 R_C \sqrt{q_0 I_E} \\ \text{Forstærkning: } & A_{DC} = g_m \beta_5 R_C \\ \text{Spændingsstøjtæthed: } & e_n = \frac{2\sqrt{q_0 I_E}}{g_m} \\ & e_n = 4 U_T \sqrt{\frac{q_0}{I_E}} = \frac{4kT}{\sqrt{q_0 I_E}} \end{aligned}$$

Figur 294 – Den ækvivalente spændingsstøjtæthed for et differentialtrin med bipolare transistorer kan beregnes alene ud fra arbejdspunktet og antallet af aktive halvledere.

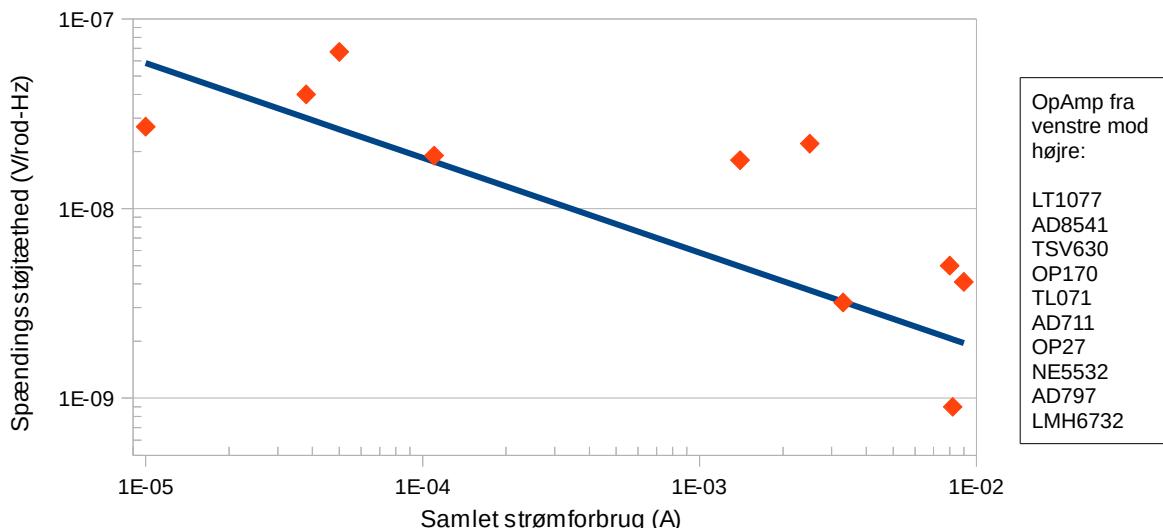
Strømstøjen i_n angiver strømmens fluktuation omkring den middelværdi I_{DC} der udgør strømmen i basis for den bipolare transistor. Basisstrømmen er $I_C/\beta = I_E/2\beta$, og med $I_E = 1 \text{ mA}$ beregnes en strømstøjtæthed på $i_n = 0,7 \text{ pA}/\sqrt{\text{Hz}}$ for en strømforstærkning på $\beta = 300$. Strømmen kan opfattes som adderet til en støjfri DC strøm i basis og strømstøjen vil derfor flyde igennem modstande i det eksterne kredsløb hvor den giver en spændingsstøjtæthed på indgangen som derefter vil blive forstærket op på normal vis.

For en MOSFET er strømmen i gate på nogle få picoampere og strømstøjtætheden bliver blot af størrelsesordenen $1 \text{ fA}/\sqrt{\text{Hz}}$. En operationsforstærker med JFET eller MOSFET er derfor velegnet som forstærker for en transducer med en høj indre modstand, men det opnåede signal/støj-forhold bliver alligevel ringe som følge af transducerens høje indre modstand der giver et stort bidrag af den termiske støj.

Spændingsstøjen e_n beregnes fra RMS summen af haglstrøm i kollektor (eller drain) af T_1 til T_4 der alle antages ens så værdien bliver kvadratroden af fire gange en af transistorernes haglstøj. Strømstøjen forstærkes op i T_5 med strømforstærkningsfaktoren β_5 og ledes til belastningen R_C hvor den giver spændingsstøjtætheden e_{nO} . Støjken henføres til indgangen som en ækvivalent støjkilde e_n ved at dividere med forstærkningen A_{DC} . For $I_E = 1 \text{ mA}$ bliver spændingsstøjtætheden $e_n = 1,3 \text{ nV}/\sqrt{\text{Hz}}$ svarende til en støjsvag OpAmp og man bør så overveje om den termiske støj fra r_x skal medregnes.

Til sammenligning har BIFET forstærkeren TL071 værdierne $e_n = 18 \text{ nV}/\sqrt{\text{Hz}}$ og $i_n = 0,01 \text{ pA}/\sqrt{\text{Hz}}$, OP27 har $e_n = 3,2 \text{ nV}/\sqrt{\text{Hz}}$ og $i_n = 0,4 \text{ pA}/\sqrt{\text{Hz}}$ og den støjsvage AD797 har $e_n = 0,9 \text{ nV}/\sqrt{\text{Hz}}$ og $i_n = 2 \text{ pA}/\sqrt{\text{Hz}}$.

Herunder vises den estimerede spændingsstøjtæthed e_n (blå linje) sammenlignet med publicerede data (orange punkter) som funktion af effektforsyningens strøm fra $10 \mu\text{A}$ til 9 mA . Det er antaget at der løber 3 % af det samlede strømförbrug i indgangstrinnet ved plotning af den blå linje som en værdi der gav rimelig visuel overensstemmelse med data.



Figur 295 – Den estimerede spændingsstøjtæthed e_n (blå linje) sammenlignet med opgivne data (orange punkter) som funktion af det typiske strømförbrug for hele komponenten uanset antallet af operationsforstärkere der vises i parentesen udfør den benyttede komponent.

De to operationsforstärkere der ligger under den blå linje har en strøm i indgangstrinnet på mere end 3 % af det samlede strømförbrug. For LT1077 er forklaringen at det samlede strømförbrug er på $10 \mu\text{A}$ og det giver ikke meget mening at skubbe indgangstrinnet ned i nanoampere-området. For AD797 er det primære mål at holde egenstøjen lav og det kræver at alle trin arbejder på sammenlignelige strømværdier hvilket vil sige $1/3$ af det samlede strømförbrug til hvert trin. De operationsforstärkere der ligger over den blå linje har en lavere strøm i indgangstrinnet end de benyttede 3 % af det samlede strømförbrug.

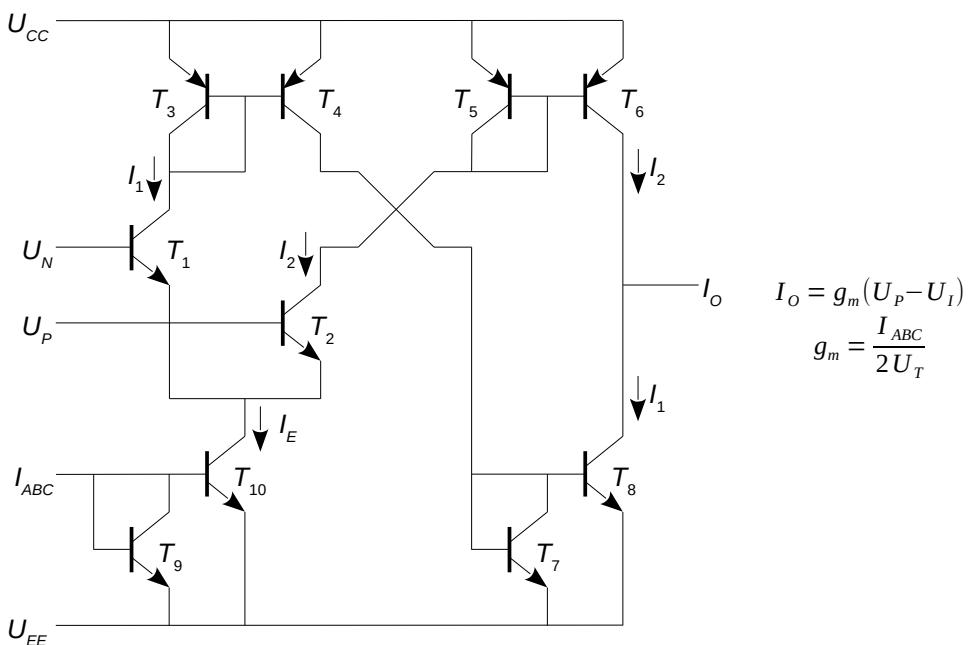
5.6 Specielle operationsforstærkere

En operationsforstærker kan bygges anderledes end det viste tema. Der findes andre måder at arrangere differentialtrinnet på eller det kan udvides til at bestå af to differentialtrin, der kan være to eller flere transistorer i VAS-trinnet for at øge strømforstærkningen og udgangen kan varieres med et andet arrangement af transistorerne eller helt udelades. Dertil kommer at transistorer som JFET og MOSFET giver chipdesigneren endnu en frihedsgrad. Opfindsomheden slutter ikke her for der kan også benyttes strømme ved ind- eller udgang eller lade strømmen i indgangstrinnet være en funktion af spændingsdifferensen på indgangen.

Dette afsnit afrunder introduktionen til operationsforstærkere med nogle af de specielle typer.

5.6.1 Transkonduktansforstærker – OTA

En BJT omsætter en spændingsdifferens mellem basis og emitter til en strøm i kollektor og for en MOSFET gælder tilsvarende at en spændingsdifferens mellem gate og source omsættes til en strøm i drain. Det kaldes for en transkonduktans og det er muligt at bygge en operationsforstærker der kan omsætte spændingsdifferensen til en bipolar strøm ved udgangen. Det kaldes for en OTA (*operational transconductance amplifier*) hvor LM13700 er den moderne udgave af den tidligere så populære CA8030 der vises herunder

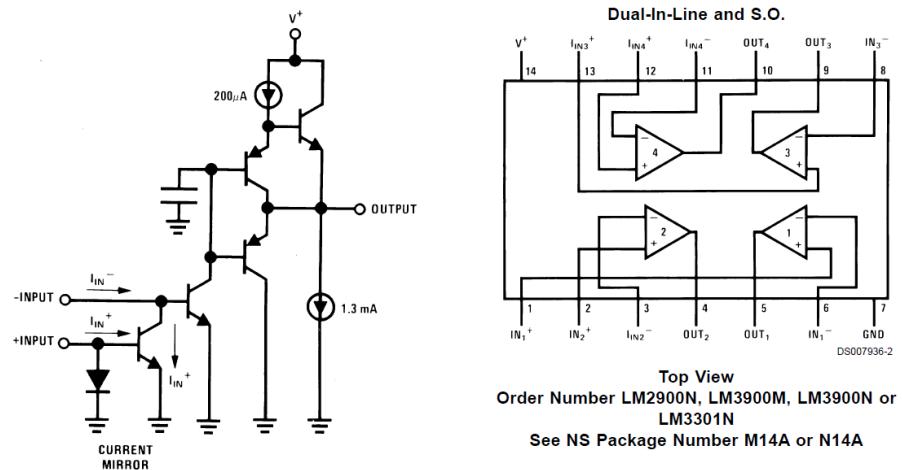


Figur 296 – I en OTA dannes differentialtrinnets strømdifferens ved udgangen og strømmen er dertil en funktion af strømmen I_{ABC} .

Ved at inkludere to ekstra strømspejle kan differensen mellem strømmene I_1 og I_2 i kollektor af de to transistorer dannes helt fremme ved udgangen. Et strømspejl styrer den fælles strøm I_E i de to transistorer i differentialtrinnet så den er lig med en ekstern strøm I_{ABC} og det gør det muligt at styre transkonduktansen over et område på mere end seks dekader. En variabel transkonduktans gør det muligt at styre forstærkningen i et kredsløb eller frekvensen af en oscillator eller et filter. Det blev udnyttet i de populære musik synthesizere fra omkring 1980 inden mikroprocessoren blev hurtig nok til digital signalbehandling. Dertil kommer opbygning af balancede mixere hvor det ene signal (indgangsspændingen $U_1 - U_2$) moduleres med det andet signal (strømmen I_{ABC}) eller udgangssignalet fra flere kredsløb kan samles i et fælles punkt med I_{ABC} brugt for en omskiftning mellem kanalerne. Der er mange andre applikationer i databladet for LM13700.

5.6.2 Transresistansforstærker

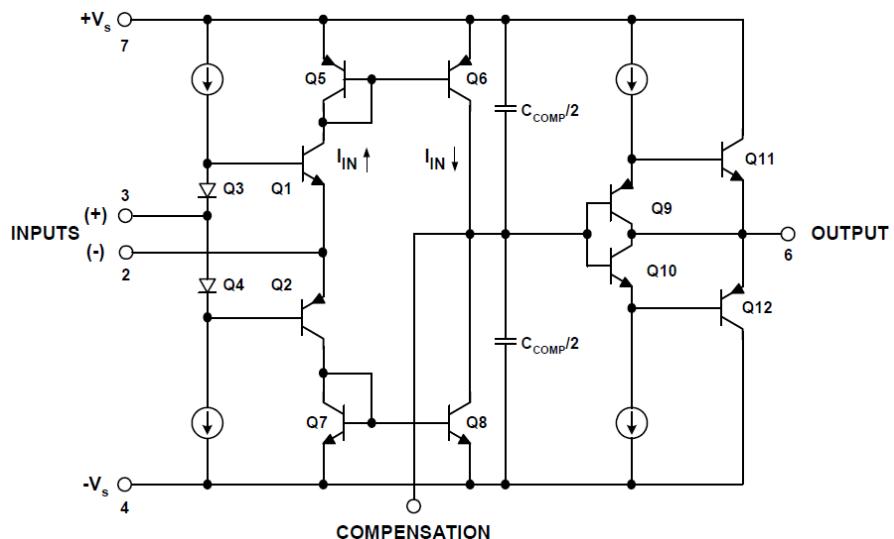
Et strømspejl på indgangen af en BJT giver transistoren en basisstrøm bestående af differensen mellem to indgangsstrømme og en buffer omsætter derefter kollektorens strøm til en spænding. Det kaldes for en *Norton amplifier* og den var overordentlig populær i en periode omkring 1980. Det er en alsidig byggeblok der kan varetage en lang række funktioner, men den voldte problemer for designere der ikke gad sætte sig ind i dens lidt specielle opbygning. Der er fire ens kredsløb i en DIL14-pakke og databladet over LM3900 indeholder et stort katalog af eksempler¹³⁸.



Figur 297 – En Norton-forstærker omsætter en strømdifferens til en spænding.

5.6.3 Current feedback

En almindelig operationsforstærker har to indgange med meget høj indgangsmodstand, men ved en current feedback operationsforstærker er indgangsmodstanden på den negative indgang lav (næsten nul). Det giver den unikke egenskab at man kan indstille operationsforstærkerens slew rate og båndbredde til den aktuelle opgave ved valget af de eksterne modstande. Det anbefales at studere databladet grundigt inden den benyttes¹³⁹.



Figur 298 – I en current-feedback amplifier (som den nu udgåede AD846) afgør strømmen i den inverterende indgang hvor hurtig operationsforstærkeren kan flytte udgangen (Analog Devices).

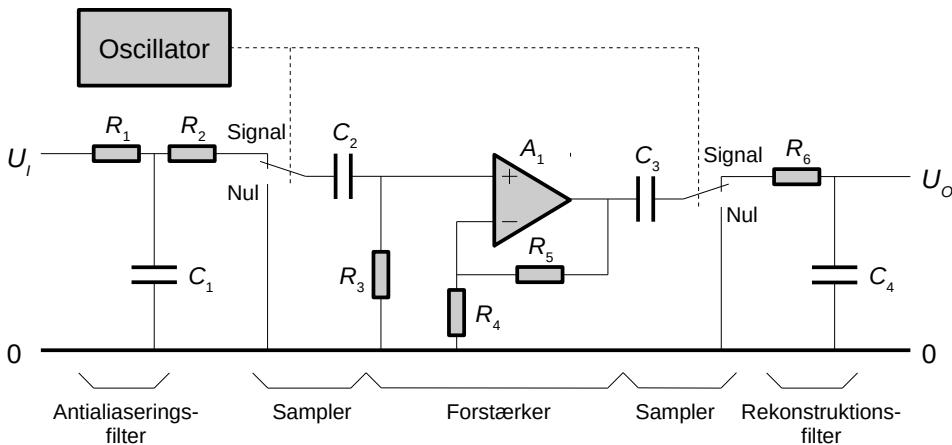
138 <http://www.ti.com/lit/ds/symlink/lm13700.pdf>.

139 <http://www.analog.com/static/imported-files/tutorials/MT-034.pdf>.

5.6.4 Chopperstabiliseret forstærker

Enhver forstærker har en DC forskydning af signalet på grund af indgangstrinnets fejlpænding og strøm (input offset voltage og input bias current). For almindelige operationsforstærkere er den typiske værdi af den samlede fejl af størrelsesordenen 1 mV så forstærkning af de svage signaler fra en strain gauge eller en termokabler kræver en justering af den samlede DC forskydning.

For en måling over nogle få minutter er det tilstrækkeligt at inkludere en justering hvor indgangen kortsluttes og udgangens fejlpænding måles af en mikroprocessor eller et potentiometer drejes manuelt indtil udgangen står på nul. Ud over besvaret med en nuljustering før hver måling så er det ikke brugbart ved måling over lang tid for enhver operationsforstærkers udgang vil langsomt flytte sig væk fra den korrekte værdi. Specifikationen findes i databladet for operationsforstærkere med lav fejlpænding og man kan så søge efter en operationsforstærker med tilstrækkelig lav drift. Hvis det da overhovedet er muligt at finde en med tilstrækkeligt gode specifikationer.



Figur 299 – I en chopper stabiliseret forstærker foretages justeringen af nulpunktet internt i komponenten ved en oscillator der kortvarigt kortslutter indgangen.

Problemet kan løses med en automatisk nuljustering styret af en oscillator og et eksempel vises ovenfor¹⁴⁰. En moderne *chopper stabilized amplifier* vil benytte en langt mere avanceret teknik¹⁴¹, men det viste trin giver en god introduktion til idéen.

- Med de to omskiftere i stilling "Nul" kortsluttes indgangen til nul og kondensator C_2 bliver opladt til en spænding givet ved indgangsstrømmen gennem R_3 . Udgangen er ligeledes kortsluttet til nul og kondensator C_3 oplades til DC værdien på udgangen som stammer fra spændingen over R_3 , fejlpændingen i A_1 og indgangsstrømmen i R_4 og R_5 .
- Med omskifterne i stilling "Signal" flyttes spændingsværdien på den positive indgang af A_1 til den aktuelle spænding ved indgangen og ændringen forstærkes op. Udgangsspændingen (efter C_3) flyttes tilsvarende fra nul til en forstærket version af ændringen på indgangen.

Der er nogle problemer med opstillingen som er rettet i de mere moderne chopperforstærkere. For det første mistes noget af forstærkningen ved udgangens filtrering af det choppede signal. For en præcisionsforstærkning er det derfor nødvendigt at kalibrere forstærkningen og sikre sig en stabil PWM fra oscillatoren. For det andet bliver signalet samplet af de to sample-and-hold kredsløb ved omskifterne. Det betyder at signalet på indgangen skal være begrænset til frekvenser under den halve oscillatorfrekvens og den opgave varetager filtret ved R_1 og C_1 så der kan kun forstærkes en langsomt varierende DC spænding.

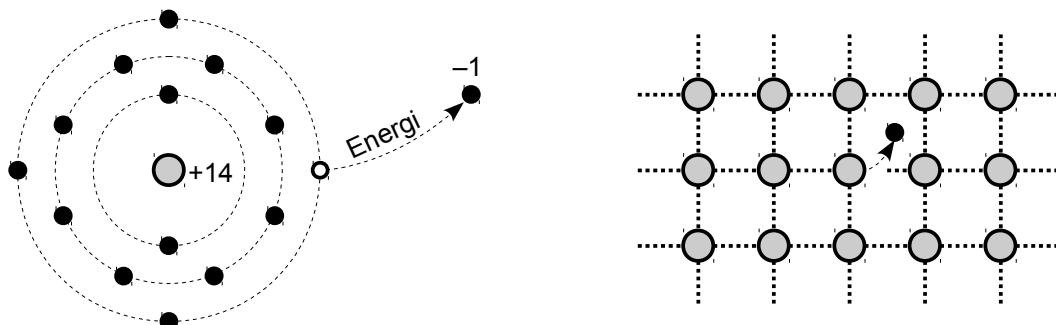
140 <http://www.analog.com/media/en/training-seminars/tutorials/MT-055.pdf>.

141 http://www.analog.com/media/en/technical-documentation/data-sheets/ADA4051-1_4051-2.pdf.

5.7 Halvlederteori

Atomets opbygning blev klarlagt af Niels Bohr i starten af 1900 tallet og beskriver atomet som en kerne omkranset af elektroner. Det simpleste atom er brint (H, hydrogen) med blot én elektron og dernæst af helium (He) der har to elektroner. Elektronerne har vedtægtsmæssigt negativ ladning og kernen har en tilsvarende positiv ladning så atomet er uadtilt neutralt. Elektronerne tiltrækkes af kernens positive ladning og holdes derved i en balance mellem en elektrostatisk tiltrækning og en kraft væk fra kernen på grund af elektronens rotation.

I de faste stoffer er atomerne arrangeret i en gitterkonstruktion med en fast rumlig orientering med det amorte glas som en af undtagelserne. Silicium har fire elektroner i den yderste skal og stoffet arrangeres så det enkelte atom er i tæt kontakt med fire andre atomer som de derved kan dele elektroner med (valenselektroner). Transistorer kan dog også fremstilles af materialer med tre og fem valenser fx AlAs, GaN, GaP, GaAs, AlGaAs og InP der anvendes til højfrekvens og lysdioder.



Figur 300 – Simplificeret atommodel med elektronerne omkring kernen i fastlagte baner og til højre en todimensional model af et krystalgitter. Bohrs atommodel viser hvordan elektronerne fordeles med 2 og 8 elektroner i de inderste skaller og med 8 elektroner i den yderste skal (to for helium) hvorved atomet er kemisk inaktivt. En elektron kan slås løs fra atomet hvis der tilføres tilstrækkelig energi og den fri elektron kan lede en elektrisk strøm indtil den igen opfanges af et andet atom der mangler en elektron. Her er vist opbygningen for silicium.

Atomets opbygning er beskrevet ved et antal skaller hvori der kan være et vist antal af elektroner. Hver enkelt af disse skaller repræsenterer et niveau af energi og da naturen tilstræber det lavest mulige energiniveau fyldes skallerne som hovedregel op indefra.

For et atom i lang afstand fra andre atomer er elektronens bane om kernen veldefineret og svarer til specifikke energiniveauer. For atomer arrangeret i et krystalgitter vil de yderste skaller derimod påvirkes af de andre atomers nærhed og i stedet for en skarp energiopdeling komprimeres deres skaller så de virker som et kontinuerligt energibånd der kaldes valensbåndet (*valence band*).

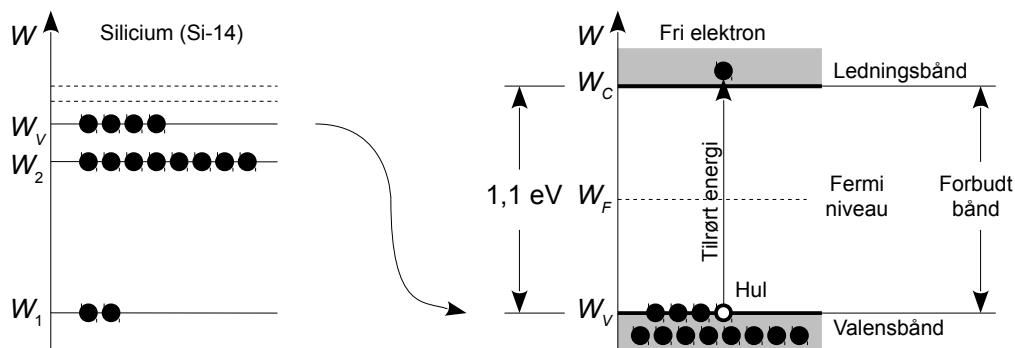
Det er muligt at flytte en elektron til et højere energiniveau ved at tilføje et kvantum af energi som kan komme fra omgivelsernes temperatur eller elektromagnetisk stråling (foton) og atomet siges da at være anslættet. Elektronen vil ikke til evig tid forblive i den højere bane, men vil før eller senere springe tilbage til en lavere position hvorved der frigøres energi i form af en foton. Tidsrummet inden dette henfald sker er ikke veldefineret, men derimod statistisk betinget og beskrives ved en middellevetid for et anslættet atom. Ved et stort antal anslæde atomer vil tætheden af fotoner aftage eksponentielt efter at den eksterne påvirkning et ophørt med fluorescens som et eksempel. Det er muligt at synkronisere henfaldet til at ske samtidigt med at andre atomer henfalder og det udnyttes i en laser med et populært eksempel ved laserdioden i en DVD afspiller.

For at slå elektronen fri af atomet kræves et minimum af energi der for silicium er på 1,125 eV¹⁴² hvorefter elektronen kan bevæge sig frit omkring i krystallet. Den siges at være i ledningsbåndet (*conduction band*) og et eksternt elektrisk felt kan påvirke elektronen så den bevæger sig.

¹⁴²En elektron modtager energien 1 eV hvis den accelereres af et spændingsfald på 1 V. Elektronen har ladningen $q_0 = 1,6 \cdot 10^{-19}$ C så en energi på 1 eV svarer til $1,6 \cdot 10^{-19}$ J.

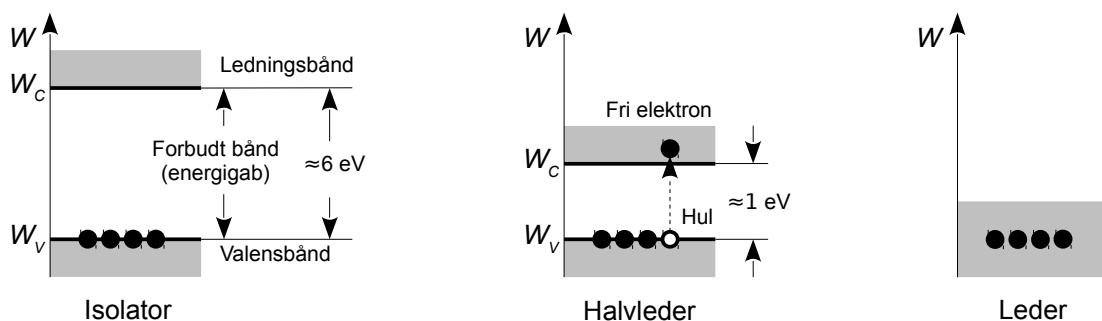
Halvleder

Afstanden mellem den øvre grænse af valensbåndet og den nedre grænse af ledningsbåndet har den sigende betegnelse "det forbudte bånd" (*forbidden gap*) for elektroner kan ikke opholde sig i dette energiområde. Enten har elektronen fået energi nok til at foretage springet eller også har den ikke fået tilført energi nok i hvilket tilfælde atomet ikke bliver anslået og energien ikke bliver absorberet. Det betyder at der er et energibånd mellem valensbåndet og ledningsbåndet hvor der ikke kan eksistere elektroner.



Figur 301 – Energimodel af grundstofet silicium der har 14 elektroner fordelt i tre skaller. Siliciums yderste skal har fire elektroner der står for den kemiske binding til andre atomer. En elektron kan slås fri fra den yderste skal hvis atomet får tilført tilstrækkelig energi som for silicium er 1,125 eV. De frie elektroner kan påvirkes af et elektrisk felt og bære en elektrisk strøm. En fri elektron falder tilbage til valensbåndet efter en vis tid, hvorved dens energi frigives i form af stråling.

Materialer opdeles almindeligvis i isolatorer og ledere med halvlederne placeret imellem disse to ydergrupper. For en isolator er den krævede energi for at passere det forbudte bånd mellem valensbåndet og ledningsbåndet så stor, at der stort set ikke vil være mulighed for at møde frie elektroner oppe i ledningsbåndet. Stoffet kan derfor ikke i væsentligt omfang lede den elektriske strøm. De typiske eksempler er keramik, plastmaterialer og atmosfærisk luft, men også tørt træ og kemisk rent vand er forholdsvis gode isolatorer.



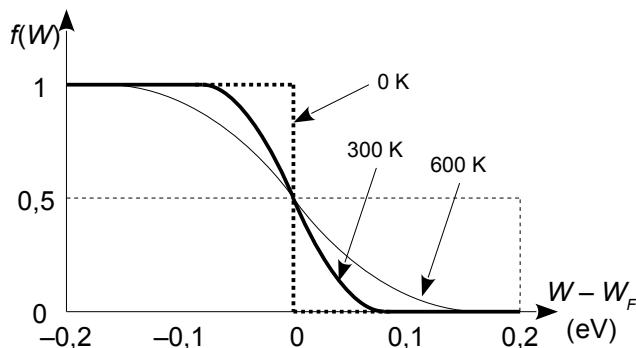
Figur 302 – Materialer kan opdeles i isolatorer, halvledere og ledere ud fra værdien af den energi der kræves for at elektroner kan passere det forbudte bånd og springe op til ledningsbåndet.

Valens- og ledningsbåndene er sammensmeltet for lederen og der er derfor altid frie elektroner til at bære den elektriske strøm. Metallerne adskilles ved hvor mange frie elektroner der kan deltage i en transport af ladning og dermed ved værdien af ledningsevnen. En kemisk forbindelse kan også være elektrisk ledende gennem dannelsen af positive og negative ioner der ligesom frie elektroner kan flyttes af et elektrisk felt og dermed lede en elektrisk strøm. Et eksempel er at vand H_2O bliver elektrisk ledende ved at tilsætte salt, $NaCl$, der opløses i vandet og skaber ionerne Na^+ og Cl^- .

Forskellen mellem en leder og en halvleder er at forurening af krystallet reducerer ledningsevnen for en leder så dens modstandsværdi forøges mens en forurening vil forbedre ledningsevnen for en halvleder så dens modstandsværdi aftager. Der er undtagelser fra denne regel, men dem vil jeg ikke komme ind på.

5.7.1 Den rene halvleder

Atomerne i krystallets gitter er meget små i det mindste i forhold til vores dagligdag. Der er mere end 10^{19} atomer i en terning af silicium med 1 mm sidelængde så halvlederteorien er baseret på sandsynligheden for hvor en elektron kan være og ikke en præcis viden om den enkelte partikel. Sandsynligheden for at møde en elektron ved et specifiseret energiniveau er givet ved Fermi-Dirac fordelingen $f(W)$ der kommer fra kvantemekanikken. Her er W elektronens energi (work) overfor et referenceniveau W_F kaldet Fermi-niveauet hvor fordelingen giver $f(W_F) = 0,5$. For brug i denne bog udnyttes det at der skal en hel del energi til for at passere det forbudte bånd så fordelingen kan simplificeres til en eksponentialfunktion.



Fermi-Dirac fordeling:

$$f(W) = \frac{1}{1 + \exp\left(\frac{W - W_F}{kT}\right)} \xrightarrow{W \gg W_F} \exp\left(-\frac{W - W_F}{kT}\right)$$

Elektron- og hultæthed:

$$n = \int_{W_C}^{\infty} \rho(W) f(W) dW$$

$$n = N_C \exp\left(-\frac{W_C - W_F}{kT}\right)$$

$$p = N_V \exp\left(-\frac{W_F - W_V}{kT}\right)$$

$$np = n^2 = N_C N_V \exp\left(-\frac{W_C - W_V}{kT}\right)$$

Materialekonstanter for silicium:

$$N_C \approx 2,84 \cdot 10^{25} \text{ m}^{-3}$$

$$N_V \approx 3,08 \cdot 10^{25} \text{ m}^{-3}$$

Figur 303 – Fermi-Dirac fordelingen $f(W)$ viser sandsynligheden for at møde en elektron med en energi W med temperaturen T som parameter (300 K ved 27°C). Parameteren W_F er Fermi niveauet der for den rene halvleder ligge midtvejs mellem valens- og ledningsbåndene og k er Boltzmanns konstant på $1,381 \cdot 10^{-23}$ J/K. Tætheden af elektroner n findes ved integration over energien for at nå til ledningsbåndet hvor $\rho(W)$ er tætheden af de mulige tilstande (udledes ikke her). Tætheden af huller p beregnes fra $1 - f(W)$. Tætheden af ladningsbærere np gives ved energigabet $E_G = W_C - W_V$ (det forbudte bånd) med N_C for ledningsbåndet og N_V for valensbåndet.

Med de viste talværdier er der $n = 1,7 \cdot 10^{16}$ frie elektroner per kubikmeter hvilket svarer til 17 millioner elektroner i en terning med 1 mm sidelængde (beregnet ved at sætte $W - W_F$ til halvdelen af 1,1 eV).

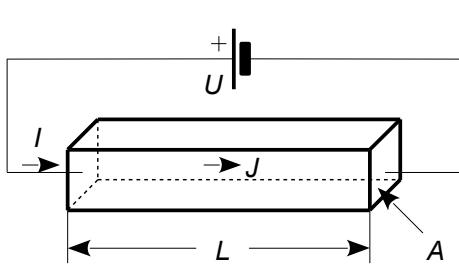
Den energi der mindst kræves for at slå en elektron fri at atomet og op til ledningsbåndet er givet ved differensen mellem energiniveauerne for ledningsbåndets W_C og valensbåndet W_V . For silicium kræves en energi på $W_G = 1,125$ eV og for germanium er det 0,663 eV. Der er derfor næsten ingen frie elektroner i siliciumkrystallet så rent silicium er en dårlig leder.

Atomerne vibrerer som et udtryk for atomets temperatur, men amplituden af vibrationen er ikke ens for alle atomer for den er statistisk betinget. Et fåtal af atomer vil fra tid til anden vibrere så kraftigt at de får nok energi til blive slynet op i ledningsbåndet. Når en elektron forlader atomet efterlades en tom plads med en positiv ladning hvilket kaldes for et hul. Hvis pladsen optages af en elektron fra et nabatom vil der nu være et nyt hul på en anden plads i atomet og man kan opfatte det som om det er hullet der har flyttet sig i modsat retning. Hullet kan opfattes som en bærer af elektrisk strøm i valensbåndet og den tillægges både ladning, masse og mobilitet.

De frie ladningsbærer leder til en koncentration af elektroner i ledningsbåndet på n og af huller i valensbåndet på p og de er begge i stand til at lede en elektrisk strøm. For at beregne antallet af ladningsbærere per rumfang af en halvleder benyttes antallet af mulige tilstande i ledningsbåndet N_C og i valensbåndet N_V .

5.7.2 Modstanden i halvlederen

En halvleder opfører sig som en modstand hvis værdi beregnes på samme måde som angivet for modstanden af en tråd i første kapitel. Det elektriske felt fra en ekstern spændingskilde accelererer elektronerne til en gennemsnitlig hastighed bestemt ved nedbremsningen fra kollision med atomer og fejl i krystallet. Den specifikke ledningsevne σ eller den specifikke modstandsværdi ρ er begge materialekonstanter der kan beregnes fra koncentrationen p og n af huller og elektroner og dertil af mobiliteten¹⁴³ μ_p og μ_n der udtrykker hvor villigt henholdsvis huller og elektroner flytter sig i et elektrisk felt.



$$\begin{aligned} J_p &= q_0 p \mu_p E & J &= J_p + J_p = \sigma E \\ J_N &= q_0 n \mu_N E & \sigma &= q_0 p \mu_p + q_0 n \mu_N \\ E &= U/L & I &= \sigma \frac{U}{L} A \Rightarrow U = \frac{L}{\sigma A} I \\ I &= JA \end{aligned}$$

$$\begin{aligned} R &= \frac{L}{\sigma A} = \rho \frac{L}{A} & \mu_p &\approx 0,048 \text{ m}^2/\text{Vs} \\ \rho &= \frac{1}{q_0(p\mu_p+n\mu_N)} & \mu_N &\approx 0,136 \text{ m}^2/\text{Vs} \end{aligned}$$

Figur 304 – Ledning af strøm i en halvleder med ensartet dotering. Længden er L , tværsnittets areal er A og der er et elektrisk felt på E på langs af halvlederen som følge af en påtrykt spænding U og modstandsværdien af halvlederen er R .

Hastigheden v beregnes ved en kendt elektriske feltstyrke E og mobiliteten μ som $v = \mu E$. Strømtætheden J er givet ved antallet af ladningsbærere der passerer gennem et tværsnit som $J = q_0 n v$ for elektronerne. Koncentrationen af ladningsbærerne for en ren halvleder er $p = n = n_i$ så den specifikke modstandsværdi beregnes til $\rho = 3,4 \cdot 10^3 \Omega \text{m}$. En terning med sidelængden $L = 1 \text{ mm}$ har modstanden $R = 3,4 \text{ M}\Omega$.

En ladning i bevægelse på grund af et elektrisk felt kaldes for driftsstrøm (*drift current*). Da antallet af ladningsbærere stiger med temperaturen vil halvlederen som udgangspunkt lede bedre ved en forhøjet temperatur, men mobiliteten aftager når tætheden bliver stor så relationen er ikke simpel. Som det ses af illustrationens talværdier er hullets mobilitet lavere end elektronens. En halvleder er derfor en ringere elektrisk leder når hullet står for den elektriske strømtransport frem for elektronen. For den bipolære transistor betyder det at en PNP transistor som udgangspunkt reagerer langsomt på et højfrekvent signal i sammenligning med en NPN idet PNP transistoren benytter huller for strømtransporten og de bevæger sig langsommere.

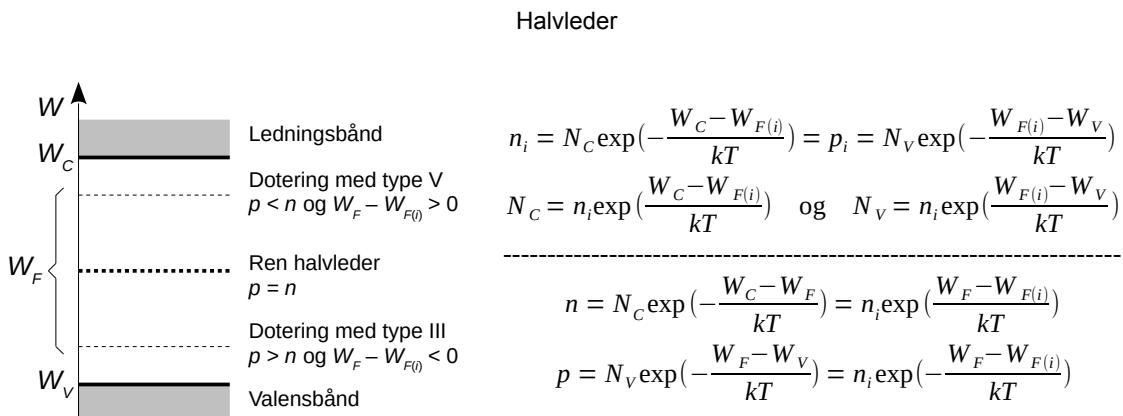
5.7.3 Den doterede halvleder

Det rene silicium har kun få elektroner og huller som bærere af den elektriske strøm, men ved at tilføje fremmede atomer til krystallet kan koncentrationen af huller eller elektroner forskydes og det forbedrer ledningsevnen radikalt. Processen kaldes for dotering og introduceres herunder.

Silicium er af type IV og doteringen benytter grundstoffer af type III eller V. Hvis et type V atom erstatter et siliciumatom i krystallet vil kun fire af materialets fem elektroner danne bindinger mens den overskydende elektron sidder løst i krystallets gitter. Den energi der skal til for at bringe en elektron fra valens- til ledningsbåndet er med 0,05 eV kun en brøkdel af den energi der skal til for at løsrive en elektron fra et siliciumatom så stort set alle elektroner flyttes til ledningsbåndet af omgivelsestemperaturen og det efterlader positive ioner på faste pladser i krystallet.

Ladningen er $q_0 N_D$ hvor N_D er antallet af frie elektroner per rumfang og det resulterende materiale bliver en N-type halvleder idet donormaterialet har tilført frie elektroner som kan bære strømmen. De mulige stoffer er kvælstof N, fosfor P, arsen As og antimon Sb.

¹⁴³Mobiliteten for huller og elektroner er forholdsvis konstant ved lav koncentration, men aftager over cirka 10^{16} cm^{-3} til en tiendedel af startværdien ved en koncentration på eller over 10^{19} cm^{-3} .



Figur 305 – Ved den rene halvleder (*intrinsic semiconductor*) er Fermi niveauet W_F placeret midt i det forbudte bånd mellem valens- og ledningsbåndene ($W_F = W_{F(i)}$) for termisk ligevægt. Dotering med type V giver et overskud af frie elektroner og Fermi niveauet hæves ($W_F = W_{F(d)}$) så differensen $W_C - W_F$ mindskes og antallet af elektroner øges. Tilsvarende øges differensen $W_F - W_V$ så antallet af huller reduceres. Dotering med type III har den modsatte virkning så Fermi niveauet W_F sænkes i retning af valensbåndet hvilket mindsker antallet af elektroner og øger antallet af huller. Det kan som vist udtrykkes ved forskydningen $W_F - W_{F(i)}$.

Med en forskydning af Fermi-niveauet på $W_F - W_{F(0)} = 0,1$ eV vil koncentrationen af elektroner øges fra det naturlige niveau på $n_i \approx 1 \cdot 10^{16} \text{ m}^{-3}$ med faktoren $\exp(3,86) \approx 50$ til $n \approx 50 \cdot 10^{16} \text{ m}^{-3}$. Tilsvarende aftager antallet af huller med faktoren $\exp(-3,86) \approx 1/50$ så elektronkoncentrationen dominerer. Den specifikke modstand bliver $\rho \approx 100 \Omega \text{m}$ så modstandsværdien af en terning med 1 mm sidelængde vil aftage fra de tidligere beregnede $3,4 \text{ M}\Omega$ hvor elektroner og huller begge varetog transport af strømmen til cirka $100 \text{ k}\Omega$ hvor der stort set kun er elektroner som ladningsbærere.

Hvis et type III atom erstatter et siliciumatom kan der kun dannes tre bindinger. For at type III materialet skal opnå en komplet skal med otte elektroner må den manglende elektron skaffes fra et af siliciumatomerne der derfor kommer til at mangel en elektron. Type III materialet kaldes for en accepter idet det optager en elektron og resultatet er dannelse af huller i valensbåndet. Hullet kan optage en elektron fra et naboatom hvilket svarer til at hullet flyttes, men hullets mobilitet er 2 til 4 gange lavere end for elektronen (langsommere) så den elektriske modstand i en P halvleder er som udgangspunkt højere end for en N halvleder. Donormulighederne omfatter bor B, aluminium Al, gallium Ga og indium In.

Når der er forurening med både type III og V atomer vil den største koncentration afgøre typen af halvleder og Fermi niveauet justeres så der opnås en ligelig ladningsfordeling for en halvleder. Da halvlederen uadtil er neutral skal antallet af elektroner n og de negative ioner fra accepter N_A modsvare antallet af positive huller p og de positive ioner fra donor N_D .

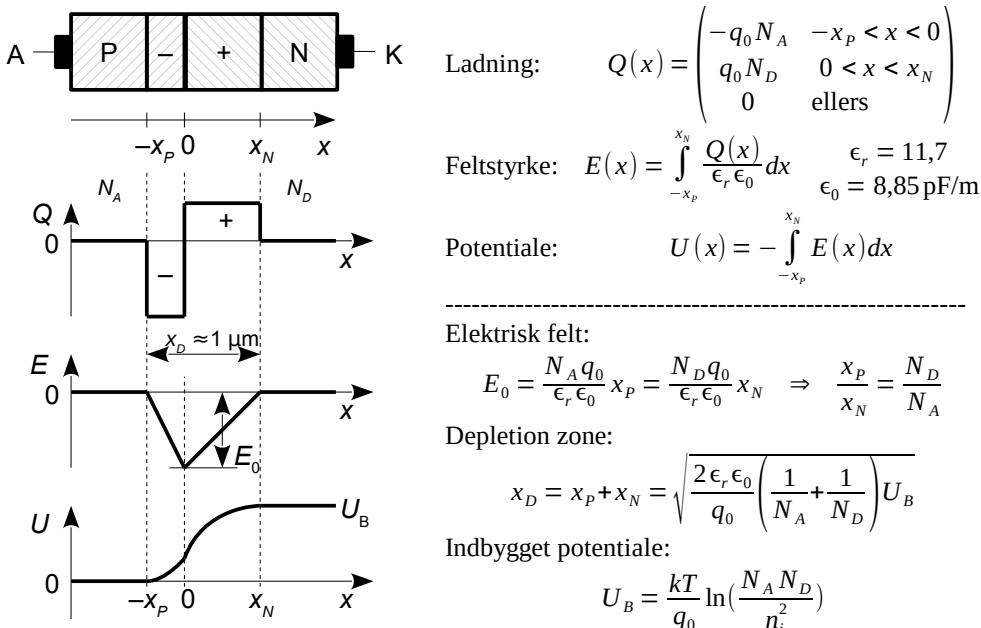
5.7.4 PN overgang

En diode består af to stykker silicium i elektrisk kontakt med hinanden. Det ene stykke er doteret med et antal accepter-atomer per rumfang på N_A der giver et underskud af elektroner og dermed et overskud af huller og det kaldes for en P halvleder. Tilsvarende vil donor-koncentrationen N_D give et overskud af elektroner og det kaldes for en N halvleder. Hver for sig er de to halvledere elektrisk neutrale for alle positive ladninger modsvares af en tilsvarende negativ ladning.



Figur 306 – Dioden konstrueres ud fra to lag af silicium hvor P materialet har underskud af elektroner og N materialet har overskud. Der dannes et ladningsfrit område mellem de to halvlederes PN overgang som styrer diodens funktion.

Når de to stykker bringes i elektrisk kontakt vil frie elektroner fra N halvlederen diffundere gennem overgangen mellem de to materialer og ind i P halvlederen. Den drivende kraft for diffusionen er en forskel mellem koncentrationen af elektroner i N og P halvlederne. Elektronerne rekombinerer med hullerne i P halvlederen mens de fastsiddende ioner nær ved PN overgangen forbliver på deres pladser. Derved opbygges negativ ladning i P laget og positiv ladning i N laget nær ved grænsen. Det elektriske felt herfra reducerer diffusionen så der ender med at være en balance mellem de to modsatte kræfter på elektronerne. Området ved PN overgangen er nu stort set tømt for frie ladningsbærere og det kaldes derfor *depletion layer* på engelsk.



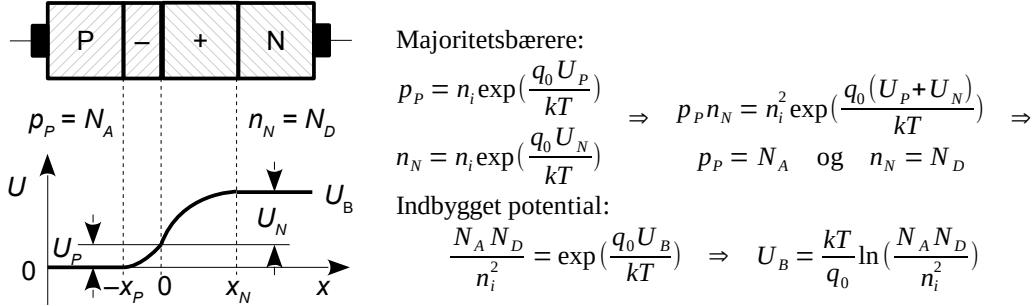
Figur 307 – Diodens interne forhold kan udledes fra ladningsfordelingen i de to halvledermaterialer og her antages en konstant doteringsprofil hvor N_A og N_D er uafhængige af positionen i materialet. Den elektriske feltstyrke $E(x)$ langs med x -aksen beregnes fra Gauss lov ved at integrere ladningen $Q(x)$ over x . Feltstyrken er nul i første og sidste del af halvlederen for halvlederen er elektrisk neutral uanset om den er ren eller doreret. Med en konstant koncentration af accepter og donor vil resultatet af integrationen blive en retlinet varierende feltstyrke hen over PN overgangen. Det elektriske felt kan opfattes som dannet af et indbygget potentiale U_B der findes ved at integrere endnu en gang. Bredden af depletion layer området x_D er givet ved doteringen med accepter- og donorkoncentrationer samt det indbyggede potentielle [CC-141].

Med $N_A = 2 \cdot 10^{21} \text{ m}^{-3}$, $N_D = 1 \cdot 10^{21} \text{ m}^{-3}$ og $n_i = 1 \cdot 10^{16} \text{ m}^{-3}$ er $U_B = 0,73 \text{ V}$ og $x_D = 1,2 \mu\text{m}$. Da $x_p/x_N = 0,5$ er det neutrale lags længder $x_p = 0,4 \mu\text{m}$ og $x_N = 0,8 \mu\text{m}$ og feltstyrkens maksimum bliver $E_0 = 10 \text{ V/m}$.

Længderne af områderne x_p og x_N er funktion af koncentrationen ved N_A og N_D og da der er lige mange positive og negative fastsiddende ioner på de to sider af PN overgangen bliver længderne funktion af koncentrationerne. Illustrationen viser den dobbelte koncentration af accepter så x_p er halvt så lang som x_N . Det er muligt at foretage en individuel beregning af længderne, men jeg har valgt kun at gengive formlen for PN overgangens længde x_D og den interessererde henvises til referencen for en uddybning [CC-138].

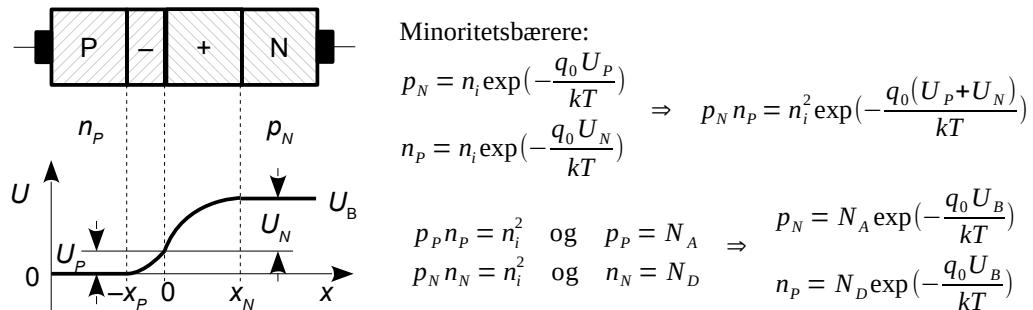
Beregning af potentialet $U(x)$ giver spændingsværdien U_P for P halvlederen ved PN overgangen og U_N for N halvlederen som vist i illustrationen herunder og de repræsenterer forskydninger i Fermi-niveauet i forhold til den rene halvleder. Koncentrationen af ladningsbærere beregnes af udtrykkene fra forrige side hvor n_i er den rene halvleders niveau af termisk genererede ladningsbærere. Forskydningen beskrives ved $q_0 U_P$ for P halvlederen og $q_0 U_N$ for N halvlederen. Summen af de to bidrag giver det indbyggede potentielle U_B (*build-in potential*) og som det vises i det følgende afsnit er det fundamentalt for diodens funktion.

Halvleder



Figur 308 – Det indbyggede potentielle kan beregnes fra koncentrationen af minoritetsbærere hvor det udnyttes at majoritetsbærernes koncentration er givet ved doteringen. I en P halvleder sænkes Fermi-niveauet ved dotering med $W_F - W_i = -q_0 U_P$ så tætheden af huller p_P øges til koncentrationen af acceptorer N_A . For en N halvleder hæves Fermi-niveauet med $W_F - W_i = q_0 U_N$ og det øger tætheden af elektroner til $n_N = N_D$. Ved at gange de to udtryk sammen dannes $U_B = U_P + U_N$ som argument for eksponentialfunktionen og ligningen kan løses for det indbyggede potential.

Det er nødvendigt at skelne mellem to forskellige typer af ladningsbærere: majoritetsbærere og minoritetsbærere (*majority carriers* og *minority carriers*). Strømmen i en P halvleder er båret af huller p_P der er majoritetsbærer hvormod elektronen n_P i P halvlederen er en minoritetsbærer der bevæger sig i "fjendeland" med en vis sandsynlighed for at rekombinere med et hul. Tilsvarende er elektronen n_N majoritetsbærer i N halvlederen og hullet p_N er minoritetsbærer.



Figur 309 – Koncentrationen af minoritetsbærere beregnes fra det indbyggede potentielle på en helt analog måde til den der blev benyttet ved beregning af majoritetsbærerne.

Koncentrationen af minoritetsbærere er lav, men den kan øges ved at reducere det indbyggede potential og det er præcis det der sker, når dioden udsættes for en spændingsdifferens med den positive terminal til anoden og den negative terminal til katoden.

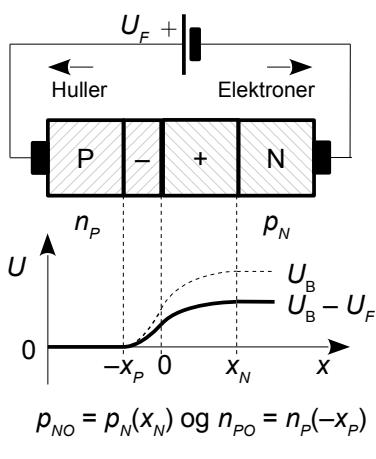
5.7.5 Strømmen i halvlederen

Med et eksternt batteri reduceres det elektriske felt hvilket åbner for en diffusionsstrøm gennem PN overgangen der ved diodens tilslutning kan måles som en elektrisk strøm. Herunder skal relationen bestemmes mellem strømmen I_F i dioden og den eksterne spænding U_F over dioden.

Det eksterne batteri sender huller ind i P halvlederen som bæres frem til PN overgangen hvor de rekombinerer med de negative ioner og reducerer koncentrationen. Tilsvarende sender batteriet elektroner ind i N halvlederen som bæres frem til PN overgangen hvor de rekombinerer med de positive ioner og reducerer koncentrationen. Det elektriske felt er givet ved antallet af fastsiddende ioner og vil derfor reduceres. Når feltet reduceres vil det indbyggede potentielle også reduceres og virkningen heraf indføres ved at U_B erstattes med $U_B - U_F$.

Reduktionen af det elektriske felt tillader en diffusion af elektroner fra N til P hvilket betyder at der løber en elektrisk strøm i dioden og der mangler nu blot at beregne strømstyrken I_F som funktion af det eksterne batteris spænding U_F .

Diffusionsstrømmen stiger fra niveauet givet af de termisk genererede minoritetsbærere p_{NO} og n_{PO} til en diffusionsstrøm forårsaget af det eksterne batteri. Det noteres i illustrationens ligninger som en diffusionstæthed p_N^D af huller i N halvlederen, og tilsvarende som en diffusionstæthed n_P^D af elektroner i P halvlederen. De repræsenterer begge minoritetsbærere der har kort gennemsnitlig levetid inden de forsvinder ved rekombination.



Minoritetsbærere:

$$p_N = N_A \exp\left(-\frac{q_0(U_B - U_F)}{kT}\right) \Rightarrow p_N = p_{NO} \exp\left(\frac{q_0 U_F}{kT}\right)$$

$$n_P = N_D \exp\left(-\frac{q_0(U_B - U_F)}{kT}\right) \Rightarrow n_P = n_{PO} \exp\left(\frac{q_0 U_F}{kT}\right)$$

Termisk genererede minoritetsbærere:

$$p_{NO} = N_A \exp\left(-\frac{q_0 U_B}{kT}\right) = \frac{n_i^2}{N_D} \quad \text{da} \quad \frac{N_A N_D}{n_i^2} = \exp\left(\frac{q_0 U_B}{kT}\right)$$

$$n_{PO} = N_D \exp\left(-\frac{q_0 U_B}{kT}\right) = \frac{n_i^2}{N_A}$$

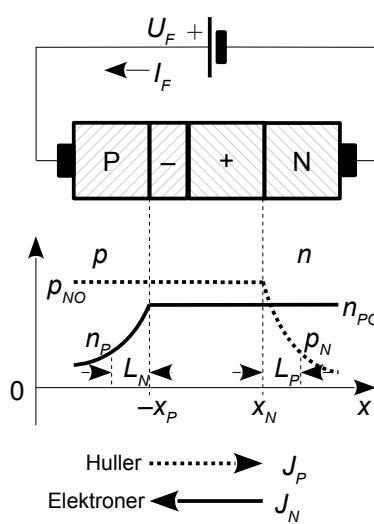
Diffusionsstrøm:

$$p_N^D = p_N - p_{NO} = \frac{n_i^2}{N_D} \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right]$$

$$n_P^D = n_P - n_{PO} = \frac{n_i^2}{N_A} \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right]$$

Figur 310 – Koncentrationen af de termisk genererede minoritetsbærere hæves af den eksterne spændingskilde fra et niveau givet af det indbyggede potentiiale og det kan igen udtrykkes ved den rene halvleders n_i samt doteringen med N_A og N_D der netop skaber det indbyggede potentiiale. Det eksterne batteri øger minoritetsbærerne gennem en diffusionsstrøm.

Den elektriske strøm i halvlederen bæres af en kombination af huller i valensbåndet og elektroner i ledningsbåndet. Hullerne vil fortsætte fra P halvlederen og ind i N halvlederens valensbånd og på tilsvarende vis vil elektronerne fortsætte ind i P halvlederens ledningsbånd så ladningsbærerne skifter status til minoritetsbærere. På den videre færd gennem halvlederen vil minoritetsbærerne forsvinde ved rekombination og de derved opbrugte huller og elektroner erstattes af nye huller fra batteriets positive pol og elektroner fra batteriets negative pol. Der løber derfor en strøm som følge af rekombinationen udenfor PN overgangen.



Diffusionsstrømmens koncentration:

$$p_N^D = \frac{n_i^2}{N_D} \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right] \exp\left(\frac{x_N - x}{L_p}\right), \quad x > x_N$$

$$n_P^D = \frac{n_i^2}{N_A} \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right] \exp\left(\frac{x_P + x}{L_N}\right), \quad x < -x_P$$

Strømtæthed:

$$J_P = \left(-q_0 D_p \frac{dp_N^D}{dx} \right)_{x=-x_P} = \frac{q_0 D_p n_i^2}{L_N N_A} \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right]$$

$$J_N = \left(q_0 D_N \frac{dn_P^D}{dx} \right)_{x=x_N} = \frac{q_0 D_N n_i^2}{L_P N_D} \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right]$$

Samlet strømtæthed:

$$J_F = J_P(x_N) + J_N(-x_P) \Rightarrow$$

$$J_F = J_S \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right], \quad J_S = q_0 \left(\frac{D_p n_i^2}{L_P N_D} + \frac{D_N n_i^2}{L_N N_A} \right)$$

Figur 311 – Diffusionsstrømmens koncentration af minoritetsbærere aftager eksponentielt udenfor PN overgangen og det beskrives ved diffusionslængden L for hver af de to områder og fortægnet fjernes som resultat af differentiationen.

Den samlede strømtæthed J i halvlederen er summen af tæthederne for elektroner og huller. For at beregne strømtæthen benyttes Fricks første lov der siger at antallet af ladningsbærere gennem en tænkt flade i et givet tidsrum er funktion af den rate hvormed koncentrationen ændres nær ved observationspunktet der her er de to grænser til PN området. Antallet af ladningsbærere per tid er derefter omsat til en elektrisk strøm ved at gange med elektronens ladning q_0 .

Det kan vises at koncentrationen af minoritetsbærer aftager eksponentielt med afstanden væk fra PN overgangen hvilket er diffusionslængderne L_P og L_N som defineres ved at koncentrationen er aftaget til e^{-1} af det oprindelige niveau (37 %). Det kan opfattes som en gennemsnitlig levetid τ_P eller τ_N for minoritetsbæreren (*minority carrier lifetime*) og efter en væsentligt længere tid må man antage at alle ladningsbærere er rekombineret. Koncentrationen af ladningsbærere er størst ved grænsen til PN overgangen så beregningen af strømtæthen benytter positionerne $-x_P$ og x_N .

Diodens strøm I_F som funktion af spændingen U_F og det indbyggede potentiale kan beregnes fra arealet A af halvlederen for et tværsnit vinkelret på strømmens retning i halvlederen.

	Diodens ligning	Indbygget potential
$+ \quad U_F \quad -$	$I_F = I_S [\exp(\frac{q_0 U_F}{kT}) - 1]$	$U_B = \frac{kT}{q_0} \ln(\frac{N_A N_D}{n_i^2})$
$I_S = q_0 A \left(\frac{D_P n_i^2}{L_P N_D} + \frac{D_N n_i^2}{L_N N_A} \right)$	$D_P = \mu_p kT / q_0$	$\mu_p \approx 0,048 \text{ m}^2/\text{Vs}$
$n_i = \sqrt{N_C N_V} \exp(-\frac{W_G}{2 kT})$	$D_N = \mu_n kT / q_0$	$\mu_n \approx 0,136 \text{ m}^2/\text{Vs}$
$W_G = W_C - W_V \approx 1,8 \cdot 10^{-19} \text{ J}$	$L_P = \sqrt{D_P \tau_P}$	$N_C \approx 2,84 \cdot 10^{25} \text{ m}^{-3}$
	$L_N = \sqrt{D_N \tau_N}$	$N_V \approx 3,08 \cdot 10^{25} \text{ m}^{-3}$
		$n_i \approx 1,07 \cdot 10^{16} \text{ m}^{-3}$
		$\tau \approx 1 \text{ ns...1 ms}$

Figur 312 – Strømmen I_F i diodens lederebane som funktion af den eksterne spænding U_F , med parametrene mætningsstrømmen I_S , elektronens ladning q_0 , Boltzmanns konstant k og den absolute temperatur T . De øvrige parametre er diffusionshastigheden D , elektronkoncentrationen n_i for den rene halvleder, diffusionslængden L og tætheden af ladningsbærere i valensbåndet N_C og i ledningsbåndet N_V . Mobilitten beskrives ved μ og den gennemsnitlige levetid ved τ . Energien W_G kræves for at passere det forbudte bånd. Talværdierne er for silicium ved 300 K (27°C)¹⁴⁴.

For silicium har vi $D_P = 1,2 \cdot 10^{-3} \text{ m}^2/\text{s}$ og $D_N = 3,5 \cdot 10^{-3} \text{ m}^2/\text{s}$ og med $\tau = 1 \mu\text{s}$ er $L_P = 35 \mu\text{m}$ og $L_N = 59 \mu\text{m}$. Med dotering til $N_A = 2 \cdot 10^{21} \text{ m}^{-3}$ og $N_D = 1 \cdot 10^{21} \text{ m}^{-3}$ bliver mætningsstrømmen $I_S = 1,2 \text{ pA}$ for en chip med sidelængden 1 μm hvor arealet er $A = 10^{-6} \text{ m}^2$. Det indbyggede potentiiale bliver på $U_B = 0,61 \text{ V}$.

5.7.6 Shottky-diode

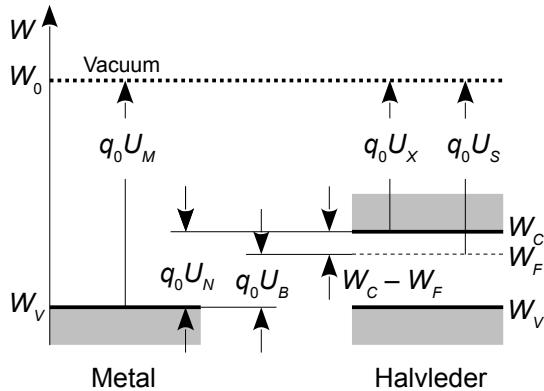
Det er ikke kun PN overgangen der kan fungere som diode for en metal-halvleder overgang giver også en ensrettende virkning¹⁴⁵. I modsætning til diodens PN overgang hvor ladningsbæreren går fra at være majoritetsbærende til minoritetsbærende så benytter Shottky-dioden kun én type af ladningsbærer som er majoritetsbærere (elektroner eller huller, men ikke dem begge). Fraværet af rekombination gør den hurtigere end en diode baseret på en PN overgang.

Der er mange frie elektroner i metallet sammenlignet med N halvlederen så de diffunderer fra metallet til halvlederen ligesom ved en PN overgang mellem to forskellige halvedere. Der er dog en anden mekanisme som giver en modsat rettet strøm af elektroner. Metallets elektroner har lav energi for der skal ikke meget termisk energi til for at slå en elektron fri. Halvlederens elektroner behøver derimod en højere energi for at blive slået op til ledningsbåndet så de søger mod metallet under afgivelse af energi og det efterlader positive ioner i N halvlederen nær ved overgangen mellem metal og halvleder og det skaber et elektrisk felt.

¹⁴⁴Data er fra referencen [CC] side 41, 45, 75, 89, 97, 101, 105, 155 og 156.

¹⁴⁵De første halvedere blev opbygget af krystallet blyulfid som en metalnål blev presset ned imod hvorefter man eksperimentelt havde fundet frem til at komponenten kunne ensrette et vekselsignal. Det var en vigtig komponent i de første radiomodtagere inden udviklingen radiorøret muliggjorde en mere stabil funktion af ensretteren.

Halvleder



Løsrivelsesenergi:

$$U_M = U_B + U_S \Rightarrow U_S = U_X + U_N - U_B$$

$$U_M = U_N + U_X$$

$$\Rightarrow U_S = U_X + \frac{W_C - W_F}{q_0}$$

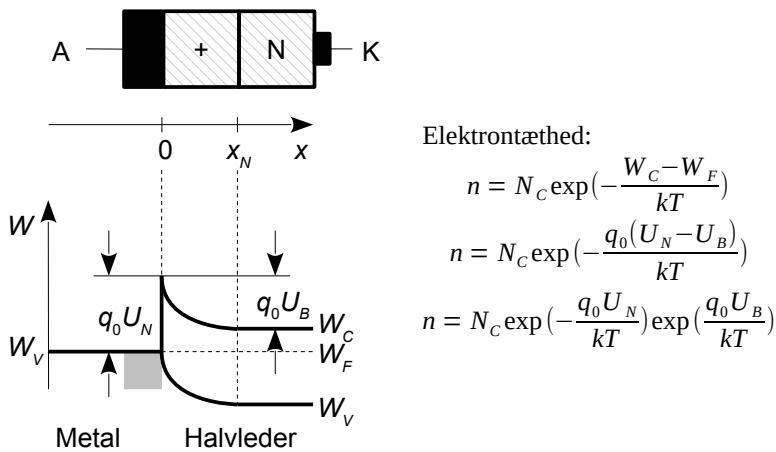
Indbygget potentiale:

$$U_B = U_M - U_S = U_M - U_X - \frac{W_C - W_F}{q_0}$$

$$U_B = U_N - \frac{W_C - W_F}{q_0} \Rightarrow U_N - U_B = \frac{W_C - W_F}{q_0}$$

Figur 313 – De aktuelle energiniveauer omfatter de potentielle energier q_0U_M for metallet og q_0U_X for halvlederen hvilket efterlader en differens q_0U_N der på engelsk kaldes for *barrier height*. På samme måde kan halvlederen beskrives ved energierne fra ledningsbåndet W_c og Fermi-niveauet W_F der giver det indbyggede potentielle U_B .

Ladningsprofilen vises herunder. Fermi-niveauet er konstant ved termisk ligevægt og påvirker ikke relationen mellem U_N og U_B så koncentrationen af elektroner kan beregnes ud fra materialets konstant N_C for de mulige tilstande i ledningsbåndet (se side 275).



Elektrontæthed:

$$n = N_C \exp\left(-\frac{W_C - W_F}{kT}\right)$$

$$n = N_C \exp\left(-\frac{q_0(U_N - U_B)}{kT}\right)$$

$$n = N_C \exp\left(-\frac{q_0U_N}{kT}\right) \exp\left(\frac{q_0U_B}{kT}\right)$$

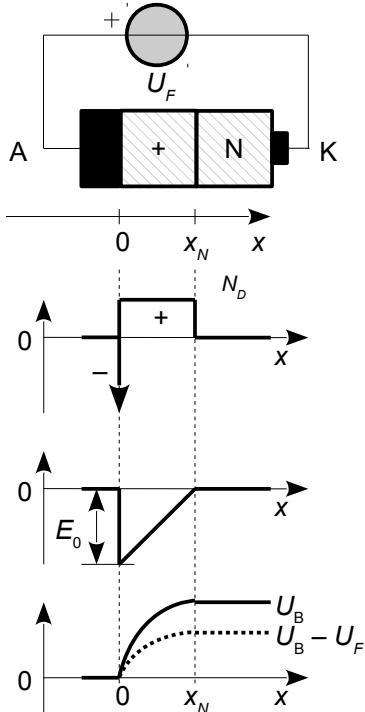
Figur 314 – Elektrontætheden bestemmes af to materialeparametre N_C og U_N samt det indbyggede potentielle U_B og for en halvleder i termisk ligevægt er Fermi-niveauet konstant.

Den elektriske strøm er relateret til antallet af ladningsbærere n der er givet ved de mulige tilstande N_C og differensen $W_C - W_F$ og det kan beskrives ved U_N der er en materialeparameter, og dertil det indbyggede potentielle U_B . Uden en ydre spændingskilde vil der ikke løbe en strøm så de to mekanismer er i balance og det udtrykkes i illustrationen herunder ved proportionaliteten C gange med koncentrationen af elektroner.

Med en ekstern spændingskilde U_F vil balancen forrykkes idet strømmen af elektroner fra halvlederen til metallet påvirkes gennem det elektriske felt hvorimod strømmen af elektroner fra metallet er konstant da det elektriske felt kun refererer til halvlederen.

Strømtætheden J_F omsættes til strøm I_F ved at gange med arealet A af overgangen fra metal til halvleder og det kan vises at produktet CN_C kan skrives som $A_R T^2$, hvor A_R er en konstant ved de givne materialer (*Richardson constant*) og T er den absolutte temperatur. Relationen er identisk med formlen for strømmen i en PN halvleder hvor de eneste forskelle er mætningsstrømmen I_S og det lavere spændingsfald i lederetningen U_F .

Halvleder



Strømtæthed i dioden:

$$J_F = J_{Si \rightarrow M} - J_{M \rightarrow Si}$$

Strømtæthed i metallet:

$$J_{Si \rightarrow M} = C N_C \exp\left(-\frac{q_0(U_B - U_F)}{kT}\right) \Rightarrow$$

Strømtæthed i halvlederen:

$$J_{M \rightarrow Si} = C N_C \exp\left(-\frac{q_0 U_B}{kT}\right)$$

Heraf:

$$J_F = C N_C \exp\left(-\frac{q_0 U_B}{kT}\right) \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right]$$

Strøm i dioden:

$$I_F = A J_F$$

$$I_F = I_S \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1 \right]$$

$$I_S = A A_R T^2 \exp\left(-\frac{q_0 U_B}{kT}\right)$$

$$A_R \approx \begin{cases} 130 \text{ A/K}^2 \text{ cm}^2 & P \\ 250 \text{ A/K}^2 \text{ cm}^2 & N \end{cases} \quad \text{Si ved } 20^\circ\text{C}$$

Figur 315 – En metal-halvleder overgang opfører sig næsten som en PN overgang og er i stand til at ensrette en vekselstrøm.

Man kan opfatte ligningen for I_F som funktion af U_F som en *curve-fitting* algoritme for dioden og med mætningsstrømmen I_S som parameter. Det betyder at hvis strømmen I_F i ledretningen skal fastholdes ved en lavere værdi af spændingen over dioden U_F så må mætningsstrømmen stige. Det er præcis hvad der sker ved at erstatter en almindelig diode med en Shottky-diode. Det lavere spændingsfald i ledretningen resulterer i en højere værdi af strømmen i spærretretningen som det vises herunder ved et taleksempel.

Hvis strømmen i diodens ledretning fastholdes på 1 mA så vil en siliciumdiode med et spændingsfald på 0,6 V have en mætningsstrøm på 100 fA og en Shottky-diode med et spændingsfald på 0,3 V vil have en mætningsstrøm på 10 nA. Det betyder at mætningsstrømmen øges med faktor 10^5 når spændingen over dioden reduceres med 0,3 V. Det er den velkendte relation med en spændingsændring på 60 mV for hver dekade ændring af strømmen.

Spændingen over dioden i ledretningen er mindre end spændingsfaldet for dioder af silicium så Shottky-dioden er populær i DC-DC konvertere hvor det lave spændingsfald giver mindre afsat effekt ved stor strømstyrke. Ulempen er at mætningsstrømmen er markant højere end for en diode af silicium og det kan give et stort statisk tab i konverteren når dioden burde spærre¹⁴⁶.

5.7.7 Elektrisk kontakt i en halvleder

Metal benyttes for at skabe elektrisk kontakt til lagene i en halvleder så der er en Shottky-diode i serie med enhver elektrisk tilslutning til halvlederens chip. Det gælder også for de interne baner i en chip der dannes af aluminium. For at undgå en uønsket ensretning af strømmen til og fra en diode eller transistor skal Shottky-dioden "ødelægges" så den ledet lige godt i begge retninger. En kraftig dotering af halvlederen sænker det indbyggede potentielle så der kun er et ganske ringe spændingsfald i ledretningen og mætningsstrømmen øges samtidigt så strømmen heller ikke bliver begrænset i spærretretningen. Det er grunden til at alle tilslutninger til dioder og transistorer sker til en kraftig dotering med enten type III materiale (P+) eller type IV materiale (N+).

¹⁴⁶Mætningsstrømmen er $I_S \leq 1 \text{ mA}$ for Shottky-dioden 1N5819.

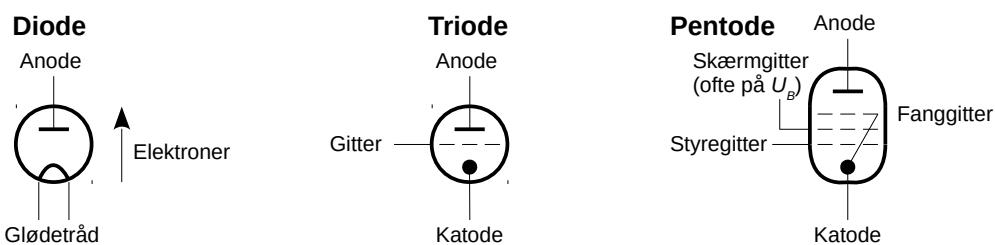
5.8 Elektronrøret

Mange vil nok mene at elektronrøret har udspillet sin rolle og at en omtale hører til ved museets udstillingsmonstre, men elektronrøret bruges i dag ved radiotransmission og røntgenstråler (*x-ray*), ved mikrobølgeovne og langtrækkende radar samt for videnskabelig brug ved studium af de høje effekters indvirkning på atomart niveau. Transistoren er kun velegnet op til nogle få kilowatt og ved højere effekt kan en sammenkobling af flere trin blive ret så upraktisk når der er behov for effekter på flere megawatt. Omtalen vil her begrænse sig til analog signalbehandling ved lave effekter og der sigtes kun efter en introduktion til elektronrøret.



Figur 316 – Elektronrør af forskellig størrelse hvor nogle få har anodens tilslutning i toppen af røret. Fælles for elektronrørene er indkapsling i en lufttom glaskolbe og en opvarmet katode for udsendelse af elektroner (Wikipedia).

Elektronrøret, der også kaldes for radiorør eller blot rør og på engelsk *vacuum tube* eller *valve*, har sit afsæt ved Edisons glødelampe hvor det var kendt at den glødende tråd udsendte elektroner og at de kunne opfanges af en elektrode med en positiv spænding overfor glødetråden (en anode). Med et negativt potentiale på anoden blev elektronerne frastødt så der ikke kunne løbe strøm. Komponenten kunne derfor ensrette en vekselspænding. Edison skal have kendt til funktionen, men han så ikke potentialet. Det var Fleming der i 1904 benyttede elektronrøret som ensretter for en stabil detektion af radiotransmission med skibe.



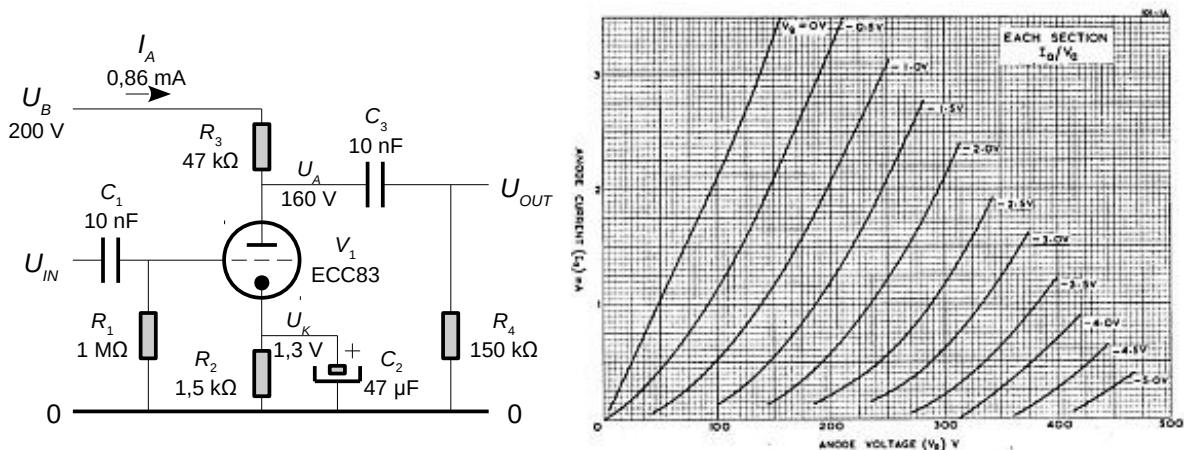
Figur 317 – Elektronrøret kan kun lede strømmen i én retning og er derfor egnet til ensretning. Med et gitter kan strømstyrken reguleres og flere gitre indføres for at forbedre funktionen. Flertallet af elektronrør har indirekte opvarmed katode så glødetråden er underforstået i diagrammet.

Omkring 1906 finder deForrest på at indbygge endnu en elektrode i røret. Det er et gitter der også kaldes for et styregitter og han bliver dermed i stand til at variere strømmen i elektronrøret; trioden er opfundet. Den videre udvikling indbygger glødetråden i et rør af et velegnet materiale (katoden) så produktionen af elektroner optimeres, men også for at forlænge glødetrådens levetid. Det giver det indirekte opvarmede elektronrør der bruges i det store flertal af konstruktioner.

Elektronerne accelereres inden de rammer anoden og de kan få energi nok til at slå elektroner fri fra anodens metal. Et fanggitter (*suppressor grid*) placeres tæt på anoden og forbundet elektrisk til katoden og det negative potentielle tvinger elektronerne tilbage til anoden. Der er en kapacitet fra anode til styregitret på trods af fanggitret og det begrænser funktionen ved høje frekvenser så et skærmgitter (*screen grid*) reducerer virkningen. Skærmgitteret er forbundet til en positiv spænding så det vil opfange en del af elektronerne fra katoden og bærer derfor en strøm. Elektronrøret har nu fem elektroder og kaldes for en pentode. Jeg giver et eksempel på anvendelsen i næste kapitel.

5.8.1 Forstærker

Opstillingen herunder viser en forstærker med en triode hvor indgangssignalet U_{IN} modulerer gitterets spænding så strømmen i anoden varierer og giver et spændingssving over modstanden R_3 i anoden der signalteknisk er i parallel med R_4 . Anodens DC strøm indstilles af elektronrørets parametre og R_2 i katoden på samme måde som en JFET. Modstanden R_1 holder gitret på 0 V og den negative gitterspænding opnås ved spændingsfaldet over R_2 . Forsyningsspændingen er typisk på $U_B = 200$ V, men kan være så lav som 12 V. Kondensator C_3 blokerer for DC og kondensator C_2 holder katoden fast overfor AC signaler så der undgås tilbagekobling (se side 244).



Figur 318 – En forstærker med autobias (Philips 1970) og rørets udgangskarakteristik for den ene halvdel af det dobbelte rør ECC83 (Mazda 1963). Der er ingen tilbagekobling så forstærkning og harmonisk forvrængning er givet af elektronrørets egenskaber.

Ifølge databladet kan den viste opstilling give 34 gange forstærkning hvilket uddybes herunder.

For en strøm i anoden på $I_A = 0,86$ mA ved $U_A = 160$ V på anoden findes spændingen over R_2 til $U_K = 1,3$ V. Strømmen stiger fra $I_{A1} = 0,65$ mA ved $U_{G1} = -1,5$ V til $I_{A2} = 1,3$ mA ved $U_{G2} = -1,0$ V. Det giver transkonduktansen $g_m = 1,3$ mS. Ved $U_G = -1,5$ V ændres strømmen fra $I_{A3} = 0,5$ mA ved $U_{A3} = 145$ V til $I_{A4} = 1$ mA ved $U_{A4} = 185$ V. Det giver den indre modstand $r_o = 80$ kΩ.

$$g_m = \frac{dI_A}{dU_G} \approx \frac{I_{A2} - I_{A1}}{U_{G2} - U_{G1}} \quad r_o = \frac{dU_A}{dI_A} \approx \frac{U_{A4} - U_{A3}}{I_{A4} - I_{A3}} \quad A = g_m R_L \quad R_L = r_o \| R_3 \| R_4$$

Med $R_3 = 47$ kΩ og $R_4 = 150$ kΩ samt den indre modstand r_o udgør den samlede belastning af anoden $R_L = 25$ kΩ og forstærkningen beregnes til $A = 32$ gange. Kapaciteten fra anode til gitter oplyses til 1,7 pF og Miller-transformeres til indgangen som 56 pF der belaster det foregående trin. Ved en tilsvarende belastning af udgangen bliver båndbredden 110 kHz.

Ulemperne ved elektronrøret er den høje spænding og glødetrådens effekt der med flere watt gør det svært at designe for batteridrift. Dertil kommer glødetrådens korte levetid og en sårbar kolbe af glas. Som fordel hævdtes det at røret er uovertruffen indenfor lydgengivelse^{147,148}.

147 Se fx <http://www.john-a-harper.com/tubes201/>.

148 Se fx <http://www.thevalvepage.com/index.shtml>.

5.9 Opgaver

1 – Vis af mæettingsstrømmen I_S kan udedes fra diodens formel på side 196 til udtrykket herunder, når I_P ignoreres og $\exp(U_F/U_T)$ er stor. Find værdien ved brug af kurven på side 198 ved $I_F = 1 \text{ mA}$, 25°C og $n = 1$. **Svar:** $I_S \approx 0,2 \text{ pA}$.

$$I_S = \frac{I_F}{\exp\left(\frac{U_F}{n U_T}\right) - 1}$$

2 – Vis at idealitetsfaktoren n kan findes ved at foretage to målinger af ledespændingen U_{F1} og U_{F2} ved to værdier af strømmen I_{F1} og I_{F2} . **Tip:** Benyt formlen for U_F og opskriv de to udtryk for U_{F1} og U_{F2} . Træk de to udtryk fra hinanden og isoler n fra ligningen.

$$n = \frac{U_{F1} - U_{F2}}{U_T \ln\left(\frac{I_{F1}}{I_{F2}}\right)}$$

3 – Benyt formlen for at bestemme idealitetsfaktoren for dioden 1N4148, hvis data vises side 198. **Svar:** $n \approx 1,6$ ved brug af $I_{F1} = 1 \text{ mA}$ og $I_{F2} = 2 \text{ mA}$ ved 25°C .

4 – Gentag beregningen af I_S med den fundne værdi for n . **Svar:** $I_S \approx 900 \text{ pA}$. Sammenlign med lækstrømmen vist på side 197 (typisk data ved 25°C og 75 V); de rapporterede data er væsentligt højere end I_S .

5 – Vis at spændingen over en diode ændres tilnærmet logaritmisk fra U_{F1} til U_{F2} ved en strømændring fra I_{F1} til I_{F2} som vist herunder.

$$U_{F2} - U_{F1} = n U_T \ln\left(\frac{I_{F2}}{I_{F1}}\right)$$

6 – For dioden 1N4148 ændres strømmen fra $I_{F1} = 1 \text{ mA}$ til $I_{F2} = 2 \text{ mA}$. Beregn ændringen i spændingen over dioden.

Svar: $U_{F2} - U_{F1} = 29 \text{ mV}$. For de fleste dioder ændres spændingen med 18 mV . Overvej grunden til den større ændring for 1N4148.

7 – For kurven på side 197 aflæses den typiske lækstrøm for dioden BAS116 ved 25°C . Find den temperatur hvor strømmen er vokset 1000 gange. **Svar:** 3 pA og 140°C .

8 – Beregn temperaturkoefficienten TC for 1N4148 på side 198 ved strømstyrkerne 1 mA og 10 mA og sammenlign data. **Svar:** $TC = -1,8 \text{ mV/}^\circ\text{C}$ og $-1,6 \text{ mV/}^\circ\text{C}$.

9 – En zenerdiode af typen BZX84 på $6,2 \text{ V}$ med data på side 200 forbindes i serie med en typisk siliciumdiode. Bestem det resulterende spændingsfald! **Svar:** $6,9 \text{ V}$. Hvilken strøm skal benyttes for at opnå en forsvindende lille temperaturkoefficient? **Svar:** $I_Z = 3 \text{ mA}$.

10 – For opstillingen til venstre på side 202 vil diodens strøm kunne genere funktionen. Det ønskes at måle en strøm på $1 \mu\text{A}$ indenfor $\pm 10 \text{ nA}$ ved 25°C . Kan 1N4148 anvendes? **Svar:** Fra kurverne side 197 ses at dioden maksimalt har en lækstrøm i spærretningen på 5 nA , og med dioden i ledernetningen er strømmen under dette niveau så længe spændingen over dioden i ledernetningen er under 70 mV , hvilket operationsforstærkeren sikrer, så svaret er "Ja".

11 – Dioden 1N4148 benyttes i strømmåleren fra side 202 ved en temperatur på 125°C . Aflæs diodens maksimale lækstrøm og afgør om opstillingen kan måle indenfor den ønskede nøjagtighed. **Svar:** $10 \mu\text{A}$ og "Nej".

12 – For opstillingen til højre på side 202 ledes 20 mA til indgangen. Overvej hvilken forsyningssgren som strømmen føres til, og om der vil kunne opstå et problem: **Svar:** Det er den positive effektforsyning, som modtager strømmen. Der skal være et eller andet i kredsløbet, som kan optage strømmen, ellers vil spændingsværdien stige.

13 – For mikseren på side 202 vil en radiosender med frekvensen 1 MHz blive modtaget ved en lokaloscillator på $1,455 \text{ MHz}$. Hvilken anden radiofrekvens vil også kunne modtages ved samme indstilling? **Svar:** $1,91 \text{ MHz}$.

14 – Der indsættes et båndpasfilter i HF blokken med en centerfrekvens på 1 MHz for at dæmpe den uønskede spejlfrekvens på $1,91 \text{ MHz}$. Hvilken omtrentlig godhed skal filtret have for at dæmpe spejlfrekvensen med 40 dB ?

Svar: En første-ordens afskæring giver 6 dB/oktav og spejlfrekvensen ligger $1,91$ over centerfrekvensen, så den dæmpes med $5,6 \text{ dB}$ hvis filtret repræsenteres ved dets Bode-plot. Filtrets godhed er lig med forholdet mellem niveauer ved centerfrekvensen for filtret og dets Bode-plot ved den samme frekvens (0 dB). Der mangler $34,4 \text{ dB}$, så godheden skal være $Q = 10^{\frac{34,4}{20}} = 52$ som et estimat.

15 – For mikseren på side 204 benyttes $f_{LO} = 10 \text{ kHz}$ og et interfererende signal modtages på 11 kHz . Beregn de to grundfrekvenser fra mikseren. **Svar:** 1 kHz og 21 kHz . Et første ordens lavpasfilter indskydes efter mikseren med en grænsefrekvens ved 100 Hz . Beregn dæmpningen af de to interfererende signaler. **Svar:** 1 kHz dæmpes 20 dB og 21 kHz dæmpes 46 dB .

16 – For dioden 1N4148 aflæses spændingsfaldet ved en strøm i ledernetningen på 1 mA fra figuren på side 198 ved en temperatur på 25°C . Beregn den forventede spænding over dioden ved 10 mA . Aflæs den typiske spænding ved denne strøm og beregn DC modstanden R_D . **Svar:** $0,58 \text{ V}$, $0,68 \text{ V}$, $0,70 \text{ V}$ og $R_D \approx 2 \Omega$.

17 – Gentag beregningen ved 100 mA . **Svar:** $R_D \approx 3 \Omega$.

18 – Ved styrkereguleringen side 207 er det maksimale AC niveau angivet til 200 mV ved en strøm på $10 \mu\text{A}$. Hvorfor? **Svar:** Ved dette AC niveau er signalets spidsstrøm lig med den valgte DC strøm, og forudsætningen om lav AC strøm i forhold til DC strømmen ikke opfyldt.

19 – Ved samme opstilling beregnes DC strømmen med en formel, hvor diodens spændingsfald antages til $0,6 \text{ V}$. Hvorfor benyttes ikke de sædvanlige $0,7 \text{ V}$? **Svar:** Der benyttes en strøm under 1 mA , hvor diodens spænding i ledernetningen er under $0,6 \text{ V}$, som derved er en bedre tilnærmelse end $0,7 \text{ V}$.

20 – Kapacitetsdioden BB207 side 210 skal benyttes i en LC oscillator, der skal kunne varieres fra 100 til 200 MHz ved en spænding fra 1 til 10 V . Beregn selvinduktionen for spolen i oscillatoren. **Svar:** $L = 32 \text{ nH}$.

21 – En blå lysdiode (se side 212) skal arbejde fra 5 V ved en strøm på 5 mA. Beregn værdien af den nødvendige modstand i serie med lysdioden. **Svar:** Fra 260 til 500 Ω .

22 – Lysdioden udskiftes med en rød LED uden at ændre andet. Hvad bliver diodens strøm? **Svar:** Fra 6 til 13 mA.

23 – Brug oplysningerne fra side 212 for den infrarøde lysdiode SFH485 til at finde båndbredden B af en mulig datatransmission, og hvis der skal sendes digitale data beregnes den maksimale bitrate. **Svar:** $B \approx 700$ MHz og 1,4 Mbit/s.

24 – Dioden SFH485 pulses ved 1 A. Beregn den typiske strålingsintensitet i 1 m afstand, der udtrykkes som den geometriske middelværdi af intervallets ydergrænser.

Svar: 0,63 W/m².

25 – Fotodioden SFH203 side 214 anbringes på aksen for SFH485 og det ignoreres at den maksimale følsomhed for de to dioder ikke er helt ens. Beregn fotostrømmen for SFH203. **Svar:** Følsomheden er 50 μ A ved 10 mW/m² og da den belyses med 630 mW/m² bliver $I_p \approx 3,15$ mA når SFH485 er tændt og 0 A når SFH485 er slukket.

26 – For lysdetektoren side 216 beregnes det kraftigste lys, der kan accepteres når spændingen over dioden skal forblive i spærretningen. **Svar:** Den maksimale strøm i fotodioden SFH203 er 0,5 mA hvor katoden når 0 V. Følsomheden er 50 μ A ved 10 mW/m² så belysningens maksimale intensitet bliver 100 mW/m² ved 850 nm.

27 – Lysdetektoren skal reagere på det pulsede signal fra IR dioden SFH485. Ved en amplitude på udgangen på 1 V skal der beregnes fotostrøm I_p , belysningsintensitet E , og den største afstand d_{MAX} mellem SFH485 og SFH203.

Svar: $I_p = 10 \mu$ A, $E = 2 \text{ mW/m}^2$ og SFH485 giver en intensitet på 25 mW/m² i 1 m afstand så intensiteten kan reduceres 12,5 gange svarende til at afstanden kan øges med 3,5 gange til 3,5 m.

28 – Fra illustrationen for BC547A på side 221 aflæses strømforstærkningen β_{25} ved 25°C og β_{150} ved 150°C, den procentuelle ændring $\Delta\beta$ og temperaturkoefficienten TC . **Svar:** $\beta_{25} = 190$, $\beta_{150} = 310$, $\Delta\beta = 63\%$ og $TC = 0,5\%/\text{C}$.

29 – Forstærkeren på side 222 øverst er dimensioneret for brug af en transistor med $\beta = 300$. Beregn spændingen på kollektor hvis transistoren erstattes af en type med $\beta = 900$ uden at foretage andre ændringer. **Svar:** Basisstrømmen er uændret på 1 μ A så kollektorstrømmen ville nå 0,9 mA, der ville danne et spændingsfald over R_C på 7,4 V, hvis der var spænding nok. Der er der ikke så transistoren vil mætte og den analoge funktion ophører.

30 – Vis at spændingen U_{CE} på kollektor af forstærkeren på side 222 nederst kan udtrykkes som vist herunder ved brug af formlerne i figuren, og beregn spændingen for en værdi på $\beta = 900$. **Svar:** $U_{CE} = 1,5$ V.

$$I_B = \frac{U_{CE} - U_{BE}}{R_B} = \frac{I_C}{\beta} = \frac{U_{CC} - U_{CE}}{\beta R_C} \Rightarrow U_{CE} = \frac{\frac{U_{CC} + U_{BE}}{\beta R_C} + \frac{1}{R_B}}{\frac{1}{\beta} + \frac{1}{R_C}}$$

31 – For forstærkeren side 223 skal indgangsspændingen U_i beregnes for en udgangsspænding på nul volt. **Svar:** Med $U_{CE} = 0$ skal kollektorens strøm være 600 μ A, så basis skal have strømmen 2 μ A. Basismodstanden aftager cirka 0,4 μ A, så der skal flyde 2,4 μ A i $R_A = 180$ k Ω og det giver en indgangsspænding på 1,13 V.

32 – For forstærkeren side 223 skal indgangsspændingen U_i beregnes for en udgangsspænding på 2,5 V. **Svar:** Der må i denne situation ikke løbe strøm i R_A så $U_i = 0,7$ V.

33 – For forstærkeren side 223 skal signalforstærkningen beregnes. **Svar:** Når en indgangsspænding varieres fra 0,7 til 1,13 V ændres udgangsspændingen fra 2,5 til 0 V, så forstærkningen bliver -5,8 gange.

34 – For forstærkeren side 223 kan udgangen U_{CE} ikke nå helt ned til 0 V. Ved brug af kurven side 223 beregnes den lavest mulige udgangsspænding. **Svar:** $U_{CE} > 55$ mV.

35 – For JFET strømkilden med BF245A side 227 skal modstanden R_S beregnes for en strøm på $I_D = 0,5$ mA. **Svar:** $U_{GS} = -1,3$ V og $R_S = 2,6$ k Ω .

36 – For samme strømkilde skal findes den laveste værdi af spændingen på drain. **Svar:** $U_D > 2$ V.

37 – Motorstyringen side 231 skal kunne drive en DC motor med et strømforbrug på 5 A. Beregn den laveste værdi af spændingen på gate. **Svar:** $U_{GS} > 5,5$ V.

38 – For AC modellen side 233 skal det vises at erstatning af U_{BE} med $U_{BE} + u_{BE}$ giver det angivne udtryk. Derefter erstattes eksponentiaffunktionen med Taylor seriens to første led og det vises at relationen bliver $i_C = g_m u_{BE}$ med transkonduktansen som $g_m = I_C/U_T$.

39 – Ved brug af Taylor serien for eksponentiaffunktionen fra andet kapitel skal det vises at AC modellen forudsætter en signalamplitude på cirka 1 mV for at anden-ordens ledet er under 1 % af første-ordens ledet. **Tip:** Indsæt værdierne 1 mV, 2 mV, ... og beregn ledernes værdier.

40 – For hybrid-pi modellen side 235 beregnes værdierne af parametrene for en BC547 ved $I_C = 0,3$ mA. Oplyste værdier: $n = 1$ og $\beta = 300$. **Svar:** $g_m = 11,5$ mS, $r_\pi = 26$ k Ω , $f_T = 9$ MHz, $C_\pi = 200$ pF, $r_o = 230$ k Ω .

41 – For opstillingen side 244 gennemføres en beregning af DC arbejdspunktet for $I_C = 1$ mA, β på ∞ og $U_C = U_{CC}/2$. Derefter sættes $\beta = 300$ og ændringen i U_B vises at være på 0,23 V. Hvilken betydning har den ændring?

42 – For opstillingen side 245 vises det at basisstrømmen til T_1 på 400 nA giver mindre end 50 mV ændring i potentialet på bases af transistoren.

43 – Vis at udgangsmodstanden R_o for den basisjordede forstærker side 246 kan udtrykkes som vist herunder, ved at injicere en strøm i_{TEST} på kollektor. **Tip:** Opskriv to ligninger for Kirchhoff på kollektor og emitter af transistoren. Tilføj en strøm i_{TEST} ind mod udgangen og ignorer R_C .

$$R_o = \frac{U_C}{i_{TEST}} = r_o + (1 + g_m R_E) R_E$$

44 – Vis at forstærkningen A og indgangsimpedansen Z , er som vist side 248. Opstil Kirchhoff for knudepunktet ved drain, og løs for U_{DS}/U_{GS} . Indsæt de typiske værdier for C_{DG} , g_m og r_o og sammenlign de to led i tælleren. Opstil derefter Kirchhoff for knudepunktet ved gate med en strøm I_G tilført indgangen, og løs for U_{GS}/I_G under forudsætning af at frekvensen er så lav at $sC_{DG}r_o$ kan ignoreres.

45 – Transkonduktansen g_m for differentialtrinnet side 257 udledes for MOSFET ved differenser hvor definitionen benytter differentiation i et arbejdspunkt. Vis at det leder til det samme resultat for arbejdspunktet hvor $U_P - U_N = 0$. Du får brug for differentiation af en sammensat funktion fg , hvor $(fg)' = fg + fg'$, og kædereglen $df/dx = (df/dy)(dy/dx)$.

46 – Med udgangspunkt i mikrofonforstærkeren side 234 skal det vises at grænsefrekvensen ved indgangen af T_5 på side 261 er som følger, når transistoren fødes fra en høj modstandsværdi:

$$f_G = \frac{1}{2\pi\beta R_C C_\mu}$$

47 – I en PN halvleder er de positive og negative områder adskilt af det neutrale lag (*depletion layer*), der udgør et isolerende lag af silicium med $\epsilon_R = 11,7$. Den fysiske udstrækning af halvlederen er $100 \mu\text{m} \times 100 \mu\text{m}$ og tykkelsen af laget er $d = 1 \mu\text{m}$. Med $\epsilon_0 = 8,85 \cdot 10^{-12} \text{ F/m}$ og udtrykket for en kondensator fra første kapitel skal det vises at kapaciteten mellem P og N lagene er cirka 1 pF .

48 – Med et forbudt bånd $E_c - E_v = 1,1 \text{ eV}$ skal det vises at en siliciumhalvleder har en naturlig (*intrinsic*) koncentration af ladningsbærere på $n_i \approx 1,7 \cdot 10^{16} \text{ m}^{-3}$ ved formlen for n_i^2 på side 273. Antag at Fermi niveauet ligger midt i det forbudte bånd og vis at koncentrationen af elektroner alene er på samme niveau, men nu ved brug af formlen for elektronernes tæthed n .

49 – Beregn koncentrationen af ladningsbærere ved en temperatur på 100°C og vis at modstanden af en terning af silicium med sidelængden 1 mm bliver $R = 31 \text{ k}\Omega$.

50 – Den gennemsnitlige hastighed er $v_{th} = 10^5 \text{ m/s}$ for elektroner ved almindelig stuetemperatur. Vis at den gennemsnitlige energi (udtrykt som $0,5 mv_{th}^2$) er omkring $4,6 \cdot 10^{-21} \text{ J}$ og sammenligne med den energi, der kræves for at elektronen løsrives fra atomet.

51 – Udfør integrationen for det elektriske felt E på side 276 og vis at E_0 bliver som vist for x_p ledet. Begrund hvorfor x_n ledet skal give samme resultat og vis derefter at der gælder at $x_n/x_p = N_a/N_d$.

52 – Sæt $U_p = 0$ og $U_n = U_B$ i formlerne på side 276 og gennemfør beregning af det indbyggede potentiale U_B .

53 – Beregn diffusionskonstanten med tabelldata side 278 for N siden af en halvleder og derefter diffusionslængden, for $\tau_D = 1 \mu\text{s}$. **Svar:** $D_N = 1,2 \text{ m}^{-3}$ og $L_N = 35 \mu\text{m}$.

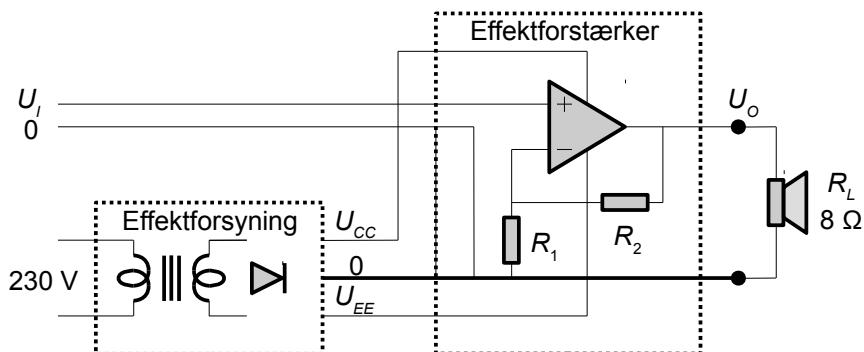
6 EFFEKTREDKREDLØB

Operationsforstærkeren er byggeblokken for analog signalbehandling, men udgangen kan sjældent klare at afgive mere end ± 30 mA så den kan lige akkurat levere effekt nok til en hovedtelefon og den kan slet ikke drive en højttaler eller en DC motor. Der findes effekt-operationsforstærkere som kan klare op til ± 3 A, men der er termiske problemer med integrationen af effekttransistorer på en silicium chip for tabseffekten skal skaffes bort. De fleste kender til problemerne med at køle mikrocomputeren i en PC og de integrerede effektkredsløb har samme udfordring at slås med. En løsning er at benytte diskrete transistorer der bedre kan klare en stor effektafsættelse og i dette kapitel vises hvordan man løser de udfordringer som de store spændinger, strømme og effekter medfører. Kapitlet indlædes med effektforstærkere for lydgengivelse via en højttaler, derefter følger de lineære spændingsregulatorer og til sidst introduceres DC-DC konverteren.

6.1 Effektforstærker

Effektforstærkere har været i brug siden starten af radiorørets æra omkring 1900 for at løfte et elektrisk signal op på et niveau hvor det kan drive en højttaler. Halvlederen tog over fra cirka 1960 og er i dag i stand til at levere en udgangseffekt på flere kilowatt. Ved høje niveauer benyttes en parallelkobling af effektforstærkere med transformatorer for at kombinere udgangssignalene, men det er stadig radiorøret, blandt andet i form af klystronen, der leverer den virkelig høje effekt ved radarsystemer og forskning i elementarpartikler med effekter på flere megawatt i korte pulser.

En effektforstærker vil i denne bog være synonym med en forstærker der er i stand til at drive en højttaler med en udgangseffekt i området op til 50 W afsat i en belastning på nominelt 8Ω . Der er ingen grundlæggende forskel fra effektforstærkere til højere effekter, men der kræves en hel del erfaring for at slippe helskinnet igennem designet.



Figur 319 – En effektforstærker fungerer på samme måde som en operationsforstærker, men den benyttes næsten udelukkende som en ikke-inverterende forstærker. Effektforsyningen regnes normalt som en integreret del af effektforstærkerens design.

Effektforstærkeren opfører sig som en operationsforstærker med tilbagekobling gennem R_1 og R_2 for indstilling af den ønskede signalforstærkning. Det betyder at det analytiske materiale fra de tidligere kapitler fuldt ud kan genbruges. En simplifikation er at en effektforstærker arbejder ikke-inverterende og indstilles til en fast forstærkning. Der er selvfølgelig undtagelser fra reglen, men de er meget sjældne da det kompromitterer stabiliteten.

De væsentligste forskelle fra design af en operationsforstærker er det højere effektniveau som nødvendiggør hensyn til termiske forhold, problemer med at udgangstrinnet ikke arbejder lineært hvilket kræver kredsløbstekniske hensyn og at effektforsyningen skal levere en så stor en effekt at den bør opfattes som en integreret del af designet. Jeg har derfor medtaget et afsnit om design af effektforsyninger startende med nettransformator, ensretter og udglatning på side 310, derefter spændingsregulering ved lineære serieregulatorer på side 321 samt de energieffektive DC-DC konvertere der introduceres på side 326.

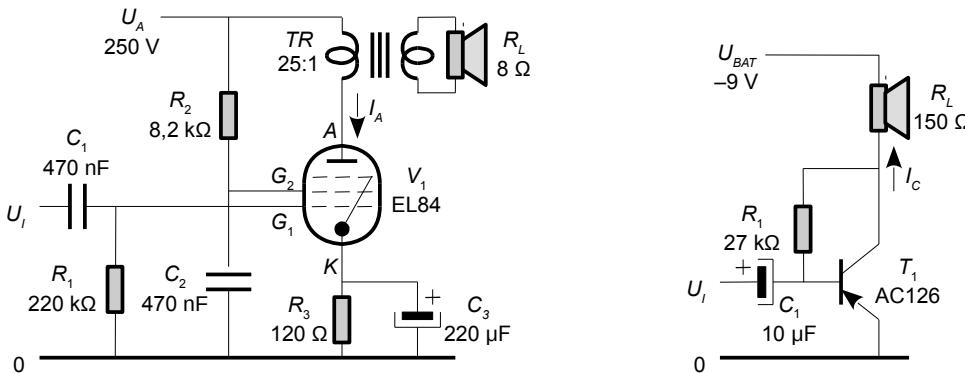
6.2 Klassifikation af udgangstrin

Udgangstrin opdeles i klasser, der betegnes med bogstaver for at beskrive specifikke egenskaber ved kredsløbet. Dette kapitel sigter mod effektforstærkere for musikgengivelse over højttalere, men radiotransmission med radiorør for var i brug allerede fra 1920 og det præger nomenklaturen.

De grundlæggende klasser er A, B, C og D som omhandles indgående i det følgende. Dertil klasse E og F der udelukkende er for højfrekvens samt klasse G og H der er rettet mod lydgengivelse af høj kvalitet og højt lydniveau, men uden den eskalerende effektafsættelse. Litteraturen er rig på en lang række af varianter som klasse S og T der begge er derivater af klasse D, *current dumping* der er en kombination af klasse A og B samt klasse XD der flytter et klasse B trins cross-over væk fra normalt lytteniveau.

6.2.1 Klasse A

Når noget er i klasse A vil det almindeligvis blive associeret med noget fremragende og indenfor lydteknologi hører en effektforstærker i klasse A da også til i den høje ende af prisskalaen, men klasse A er ikke en garanti for den højeste kvalitet som det skal vises med nogle eksempler på simple effektforstærkere.



Figur 320 – Eksempler på klasse A udgangstrin hvor der ikke benyttes tilbagekobling så lineariteten afhænger af den aktive komponent og forvrængningen når let op omkring 10 %.

Udgangstrinnet i ældre radiomodtagere benyttede radiorør og højttaleren var tilsluttet gennem en transformator for at undgå en DC strøm i svingspolen. I eksemplet arbejder pentoden EL84 ved forsyningens $U_A = 250$ V og anoden A indstilles til strømmen $I_A = 48$ mA ved modstanden R_3 så potentialet er $-5,5$ V på gitteret G_1 overfor katoden K . Skærmgitret G_2 reducerer koblingen fra anoden til G_1 og trækker 5 mA som benyttes til at give et spændingsfald over R_2 for et potentiale på 210 V. Gitter G_3 standser de elektroner der slås løs fra anoden og er internt forbundet til katoden. Der afsættes 12 W i røret hvortil kommer 5 W fra opvarmningen af katoden (glødetrådens kredsløb er ikke vist). Transformatoren omsætter R_L til en belastning på 5 k Ω ved anoden og udgangseffekten når maksimum på 5,7 W når anodens strøm varierer med $\pm 100\%$. Virkningsgraden er 33 % og den harmoniske forvrængning oplyses til 10 %.

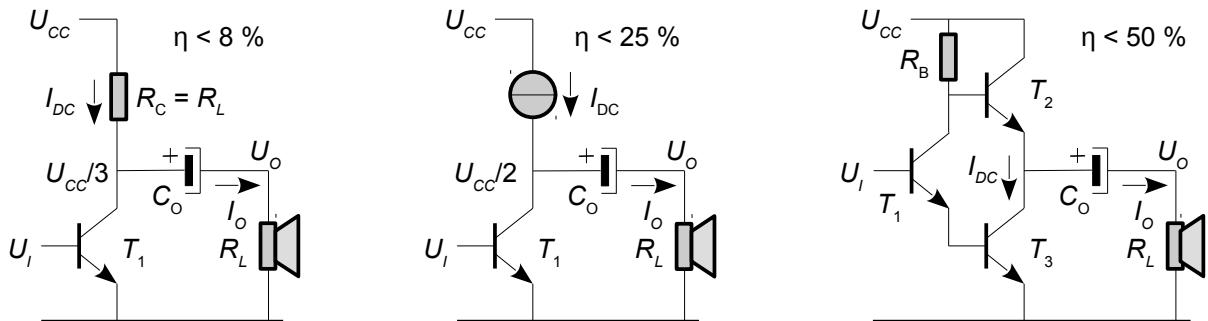
Det andet eksempel er fra elektronikbyggesættet Philips EE20 fra omkring 1965 hvor transistoren AC126 driver en 150 Ω højttaler uden transformator. Der løber en DC strøm i svingspolen hvilket afsætter effekt i den og membranen forskydes fra den neutrale position. En strømforsærkning på 100 får transistoren til at indstille sin strøm i kollektor på 20 mA så der optages 180 mW. Hvis strømmen svinger ± 20 mA om arbejdspunktet bliver RMS værdien 14 mA og der vil afsættes en AC effekt på 30 mW i højttaleren kan derfor ikke spille ret højt. Virkningsgraden på 17 % er lavere end for radiorøret da spændingen på kollektor ikke kan komme over 9 V. Den harmoniske forvrængning når 10 % allerede ved et indgangssignal på 10 mV på grund af transistorens eksponentielle karakteristik og her er udgangseffekten under halvdelen af det maksimale. Bemærk at en PNP transistor arbejder med negativ spænding fra effektforsyningen.

En klasse A effektforstærker har sine gode egenskaber fra det forhold at den aktive komponent primært arbejder ved små variationer omkring arbejdspunktet og den harmoniske forvrængning er lav af den simple grund at forvrængningen er proportional med signalniveauet. De to eksempler er derfor kun gode effektforstærkere ved moderat lytteniveau, men det er jo også det normale at en klasse A forstærker arbejder ved en procentdel af sin maksimalt udgangseffekt.

Effektforbruget over en periode af signalet er konstant ved klasse A så belastningen af effektforsyningen er uafhængig af effekten til belastningen. Virkningsgraden for en klasse A forstærker kan maksimalt nå 50 % så en 100 W forstærker vil have et konstant effekttab på minimum 200 W. Når signalniveauet er lavt bliver virkningsgraden nær ved 0 % og ved musikgengivelse er forstærkeren aldrig fuldt udnyttet så der afsættes altid en stor effekt i transistorerne. Den største udfordring ved klasse A er derfor at bringe effekten væk så transistorerne ikke brænder af. Selv om effektforbruget burde gøre klassifikationen uinteressant så arbejder enhver forstærker for små signaler i klasse A og man skal aktivt foretage sig noget for ikke at bygge klasse A.

Der er flere grundkoblinger som fra tid til anden dukker op på internettet og i elektronikmagasiner og tre af dem skitseres herunder med en kort introduktion til dimensioneringen. Diagrammerne vil ikke inkludere alle nødvendige komponenter fx mangler indstilling af basisstrøm.

Den første opstilling er en almindelig fælles-emitter forstærker hvor belastningen R_L tilsluttes via en kondensator så kun AC delen af udgangssignalet overføres. Strømmen i kollektor deles mellem R_C og R_L og med to ens modstande er virkningsgraden på blot 8 %. En forbedring erstatter R_C med en strømgenerator der fastsætter DC arbejdspunktet og den maksimale udgangsstrøm til I_{DC} og hæver virkningsgraden til maksimalt 25 %. Kondensatoren i udgangen kan undgås ved at benytte en dobbelt effektforsyning, men det kræver at U_{CE} svinger om $U_{CC}/2$. Den sidste opstilling benytter tre transistorer hvor T_1 moduler strømmen i T_2 og T_3 så strømmen i den ene øges med samme bidrag som den mindskes i den anden og det hæver virkningsgraden til det teoretisk maksimale for en klasse A effektforstærker på 50 %.



Figur 321 – Klasse A udgangstrin med en enkelt transistor hvor det midterste ofte dukker op i den populære litteratur. Til højre et mere udbygget push pull udgangstrin.

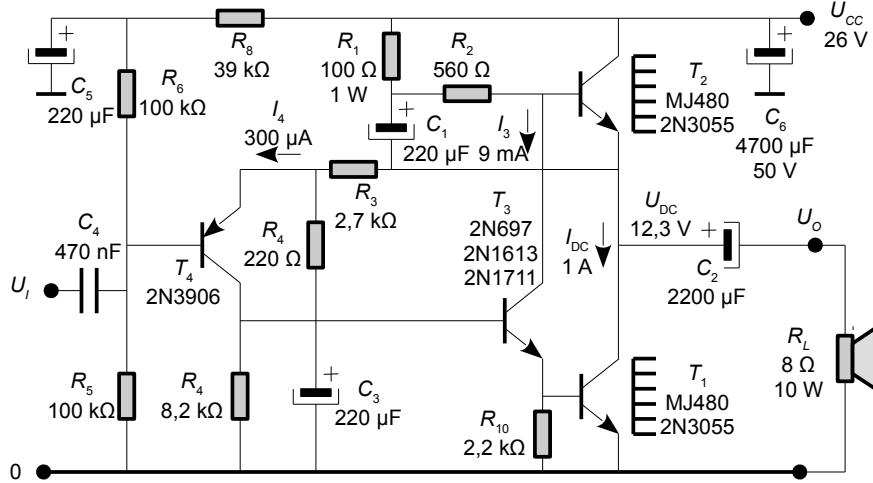
Udgangseffekten kan beregnes fra $P_O = U_{ORMS}^2/R_L = I_{ORMS}^2 \cdot R_L$ og for en udgangseffekt på $P_O = 1 W$ bliver amplituden $U_O = \pm(2 \cdot P_O \cdot R_L)^{0.5} = \pm4 V$ ved $R_L = 8 \Omega$ der giver spændingsfaldet til $U_{CE} = 4 V$ i hvile. Strømmedlingen mellem R_C og R_L giver amplituden af kollektorens strøm så der behøves $I_{DC} = 2 \cdot (2 \cdot P_O / R_L)^{0.5} = 1 A$ i kollektor i hvile. Effektforsyningen beregnes til $U_{CC} = R_C I_{DC} + U_{CE} = 12 V$ og dermed afgives effekten $P_{CC} = U_{CC} I_{DC} = 12 W$ fra batteriet så virkningsgraden bliver på $\eta = P_O / P_{CC} = 8 \%$.

Effektforsyningen kan her reduceres til spids-spids værdien af udgangens svingning på $U_{CC} = (2 \cdot P_O \cdot R_L)^{0.5} = 8 V$ ved en udgangseffekt på $P_O = 1 W$ og $R_L = 8 \Omega$. Udgangsstrømmen kan nå op på I_{DC} og DC arbejdspunktet bliver $I_{DC} = (2 \cdot P_O / R_L)^{0.5} = 0.5 A$. Den optagne effekt reduceres til $P_{CC} = U_{CC} I_{DC} = 4 W$ og virkningsgraden når den typiske værdi for de mindre klasse A effektforstærkere på $\eta = P_O / P_{CC} = 25 \%$.

Strømmens svingning er for hver transistor $\pm I_{DC}$ omkring DC arbejdspunktet, men udgangsstrømmen er givet ved differensen mellem T_2 og T_3 til værdien $\pm 2 \cdot I_{DC}$. DC arbejdspunktet bliver $I_{DC} = 0.5 \cdot (2 \cdot P_O / R_L)^{0.5} = 0.25 A$ ved $P_O = 1 W$ og $R_L = 8 \Omega$. Udgangsspændingens amplitude giver det samme krav til effektforsyningen som før på $U_{CC} = (2 \cdot P_O \cdot R_L)^{0.5} = 8 V$, men effektforsyningen afgiver nu kun en effekt på $P_{CC} = U_{CC} I_{DC} = 2 W$ så virkningsgraden er $\eta = P_O / P_{CC} = 50 \%$.

Beregningerne forudsætter at transistorerne kan svinge ned til 0 V over kollektor-emitter, hvilket helt ignorerer mætningsspændingsfaldet og der er ikke medtaget effekttab fra det øvrige kredsløb. Man bør designe konservativt med mindst 10 % højere værdier af U_{CC} og I_{DC} end som angivet ved designformlerne under illustrationen og et eksempel skal gives.

Herunder vises en klasse A effektorforstærker med *push pull* udgang der leverer 10 W til 8 Ω ved lav forvrængning. Kredsløbet blev bragt af John Linsley Hood i fagbladet Wireless World April 1969. Effektforsyningen var i artiklen seriereguleret for at undertrykke netbrum, men den del medtages ikke i diagrammet¹⁴⁹. Kredsløbet er af samme type som vist til højre i forrige figur.



Figur 322 – En klasse A forstærker for 10 W i 8 Ω med en harmonisk forvrængning på 0,05 %.

Forfatterens kommentar til kredsløbet var at funktionen ikke afhænger af transistorernes data så længe de kan klare arbejdsbetingelserne. Årsagen til dette fokus skyldes at radiorøret dannede referencen i midten af 1900-tallet og her var det vigtigt at vælge et egnet rør til et givet projekt.

For den specificerede udgangseffekt kræves en RMS værdi af udgangsspændingen på 9 V og det svarer til en spids-spids værdi på 25 V. Effekttransistorerne kræver mindst 0,5 V over kollektor og emitter for at fungere så den mindste værdi af effektforsyningen bliver 26 V. Udgangsstrømmen kan nå op på 1,6 A så DC arbejdspunktet for effekttransistorerne skal være på mindst 0,8 A. Man står sig dog ved at dimensionere efter en højere strøm da flertallet af højttalere har en lavere DC modstandsværdi end den nominelle værdi så der er behov for en DC strøm i omegnen af 1 A. Det betyder at der til stadighed afsættes 13 W i hver transistor og effekten skal bringes væk.

DC strømmen indstilles ved R_1 og R_2 der sender 18 mA ind i knuden mellem basis af T_2 og kollektor af T_3 som igen sender strømmen videre til basis af T_1 . Antages en ligelig fordeling af strømmen vil hver af effekttransistorerne modtage 9 mA så strømforstærkningen skal være over 100 for det planlagte arbejdspunkt. Strømforstærkningen er på mindst 30 og typisk 200 for MJ480 samt mindst 20 og typisk 70 for 2N3055 og det viser at denne type effektorforstærker enten bør bruge udmålte transistorer eller at der bør være en justeringsmulighed¹⁵⁰.

Kondensator C_1 holder spændingen over R_2 konstant uanset udgangsspændingen så strømmen i modstanden bliver konstant hvilket kaldes for *boot strap*. Bemærk at der afsættes 0,2 W i T_3 og op til det dobbelte under brug så den bør mindst kunne klare 0,8 W.

Udgangens DC potentiale fastsættes til halvdelen af effektforsyningen ved transistor T_4 hvis basis holdes på 10,9 V af modstandene R_5 , R_6 og R_8 med filtret ved R_8 og C_5 for at undertrykke brum fra effektforsyningen. Emitter af transistoren ligger på 11,5 V og udgangen ligger 0,8 V højere på grund af af R_3 og strømmen på 0,3 mA der indstilles af R_4 og basisstrømmen i T_3 . Forstærkningen af signalet er $1 + R_3/R_4 = 13$ gange så forstærkeren afgiver 10 W ved 0,7 V på indgangen.

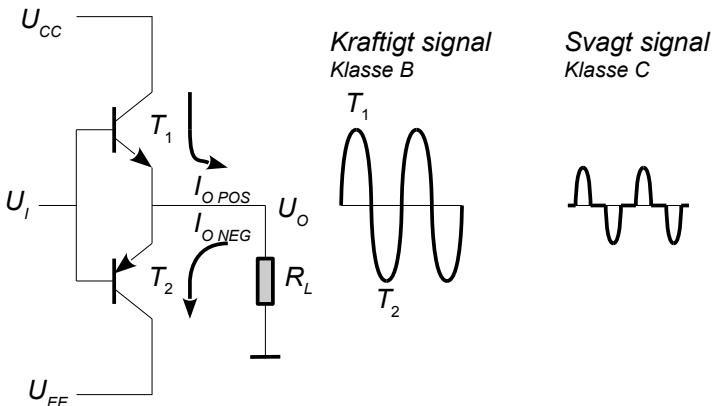
149 Artiklen kan blandt andet findes på: <http://www.keith-snook.info/wireless-world-magazine/wireless-world-articles.html>.

150 Da jeg i sin tid byggede en klasse A forstærker efter det viste diagram kunne man hos komponentleverandøren købe 2N3055 med en strømforstærkning udmålt til mindst 100, men kun hvis man lovede at bygge klasse A.

Note 13.

6.2.2 Klasse B

Den store og konstante effektafsættelse i en klasse A forstærker kan reduceres væsentligt ved at opdele signalet i positive og negative halvdele som derefter processers individuelt og adderes ved udgangen. En klasse B forstærker benytter to transistorer som skiftes til at lede i hver sin polaritet af signalet. I illustrationen herunder er T_1 aktiv når signalet er positivt hvor den leverer strøm fra den positive effektforsyning til belastningen mens T_2 er afbrudt. Tilsvarende er T_2 aktiv når signalet er negativt mens T_1 er afbrudt. Da kun den ene af de to transistorer er aktiv til et givet tidspunkt kan opstillingen opfattes som om den var en symmetrisk emitterfølger.



Figur 323 – Et klasse B udgangstrin har altid to aktive komponenter, der skiftes til at lede strømmen i den positive og negative halvperiode, men spændingen over basis-emitter giver en forvrængning af signalet ved nulgennemgang.

Effektafsættelsen holdes nede ved at der ikke er involveret en DC effekt når der ikke er et signal og det muliggjorde fra 1950 anvendelse af transistorer der ikke i samme grad som radiorøret tåler en stor effektafsættelse. Der afsættes maksimalt 40 % af den nominelle udgangseffekt i udgangen (se side 300) så en 100 W forstærker afsætter op til 20 W i hver udgangstransistor i modsætning til de konstante 100 W i hver udgangstransistor ved klasse A.

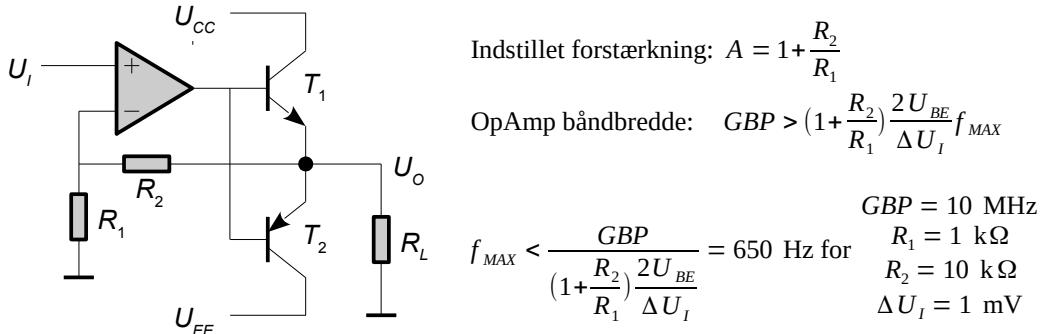
Der er sjældent behov for at afgive den fulde udgangseffekt over længere tid for et musikprogram består mest af svage signaler med få kraftige spidser og selv ved blot 10 dB til spidserne afsættes blot en tiendedel af fuld udgangseffekt lige inden forstærkeren overstyres. Fordelen ved klasse B er derfor at udgangstransistorerne kun opvarmes ganske lidt i sammenligning med klasse A.

Virkningsgraden kan teoretisk nå 78 % ved fuld udgangseffekt, hvilket meget litteratur nævner som et vigtigt punkt, men det er efter min mening langt vigtigere med det lave gennemsnitlige optag af effekt fra effektforsyningen.

Klasse B er i sin rene udformning en sjælden gæst indenfor lydgengivelse på grund af en særdeles problematisk harmonisk forvrængning der kaldes for *cross over*. En transistor kræver 0,7 V over basis-emitter for at trække en strøm af betydning så udgangen vil ikke reagere indenfor et interval på $\pm 0,7$ V omkring nul. Udgangen vil ligge på nul selv om indgangssignalet ændres.

Hvis der foretages en Fourier-analyse på signalet vil det vise kraftige harmoniske i den høje ende og dertil kommer en kraftig krydsmodulation mellem flere samtidige signaler. Begge artefakter er særdeles hørbare og vil i praksis give fænomenet "lyttetræthed" der prægede et så stort antal af de tidlige effektorstærkere med transistorer at det nedsættende blev betegnet med "transistorlyd" som en klar adskillelse fra de gode og velkendte effektorstærkere med radiorør.

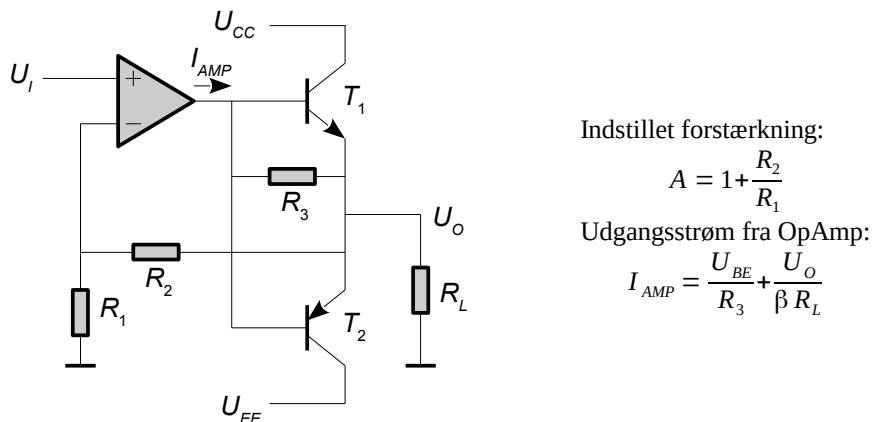
En korrektionsforstærker kan tvinge udgangen til at følge med indgangssignalet. Det er fint ved DC og lave frekvenser, men der skal skiftes meget hurtigt igennem de $\pm 0,7$ V hvor begge transistorer er afbrudt og det giver et meget højt krav til forstærkerens GBP for lydgengivelse. Desuden fjerner det ikke cross-over forvrængningen – problemet bliver reduceret, men det elimineres ikke.



Figur 324 – En OpAmp kan korrigere for dødbåndet på $\pm 0,7$ V, men opstillingen er kun brugbar ved DC og lave frekvenser for ukritiske opgaver som fx styring af DC motorer.

For et skift i indgangssignalet ΔU_I på 1 mV kræves en åben-sløjfe forstærkning på 1400 blot for at skifte mellem T_1 eller T_2 som aktiv. Den indstillede forstærkning er $1 + R_2/R_1$ og for en forstærkning på 10 gange skal åben-sløjfe forstærkningen være på 14000 ved den aktuelle frekvens. Ved en frekvens på 20 kHz skal korrektionsforstærkerens GBP være over 280 MHz – og der mistes stadigvæk ± 1 mV af indgangssignalet.

En forbedring er at inkludere en modstand mellem basis og emitter af udgangstransistorerne. Det driver strøm til belastningen selv om udgangstransistorerne ikke er aktive. Signalet fra udgangen af operationsforstærkeren dæmpes dog ved spændingsdeling mellem R_3 og R_L når transistorerne ikke er aktive og det giver en forvrængning der dog kan reduceres af tilbagekoblingen.



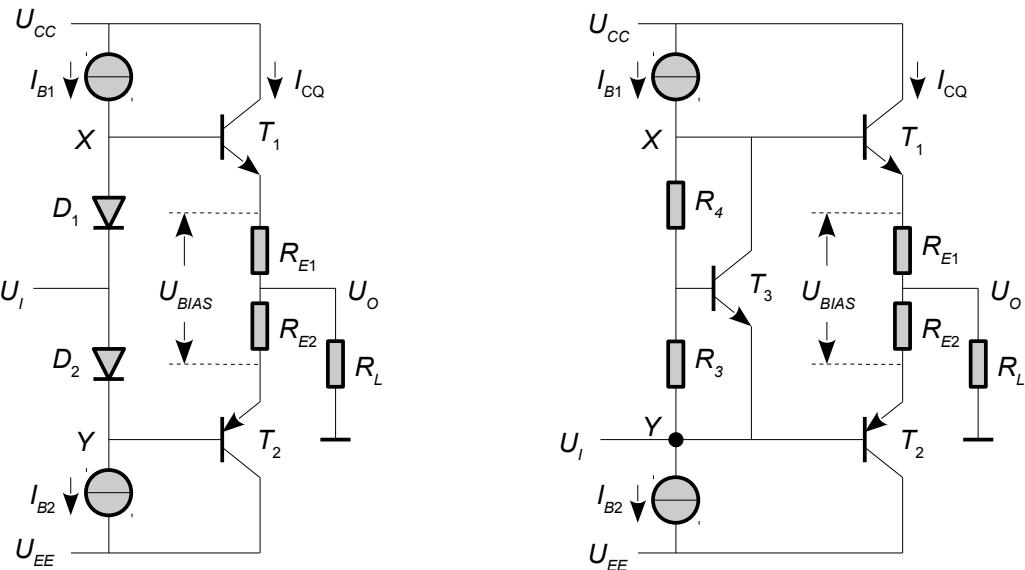
Figur 325 – Simpelt udgangstrin med korrektion for cross-over forvrængningen gennem brug af en operationsforstærker med tilbagekobling.

Metoden kan raffineres til et princip der kaldes for current dumping og som bruges i QUAD 405. Her udnyttes en klasse B effektforstærkers lave effekttab og problemet med det døde bånd klares af en klasse A forstærker. Et filter kombinerer de to udgange så klasse B forstærkeren kun arbejder med de lavfrekvente signaler (langsomm varierende) mens klasse A forstærkeren tager sig af de hurtigt varierende signaler. Jeg har ikke kendskab til om andre effektforstærkere benytter dette princip selv om patentet er udløbet for flere år siden¹⁵¹.

151 <http://www.keith-snook.info/wireless-world-magazine/Wireless-World-1975/Current%20Dumping%20Audio%20Amplifier%20DCD.pdf>.

6.2.3 Klasse AB

For at mindske problemerne med overgangen mellem transistorerne benytter effektforstærkere et kredsløb som holder begge udgangstransistorer aktive i området nær ved nul. Udgangstrinnet er nu i klasse A ved lavt signalniveau og skifter over til klasse B ved et kraftigere signal. Af forståelige årsager kaldes det for klasse AB og det udgør det absolut dominerende princip.



Figur 326 – To almindelig måder at kompensere for det døde bånd. Det udnyttes at både dioderne og transistoren har den samme temperaturkoefficient som udgangstransistorerne så de skal have termisk kontakt til kølepladen for at følge med temperaturændringerne. Indgangsmodstanden ved U_I er tilnærmelsesvist βR_L hvor β er strømforstærkningsfaktoren for transistorerne.

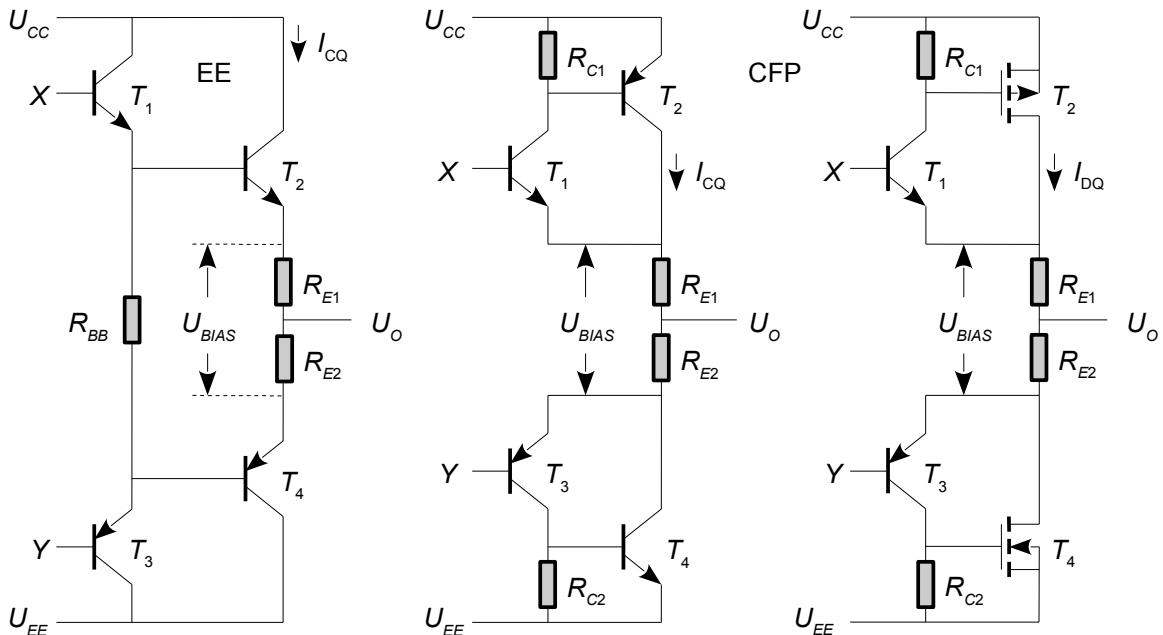
Ved det venstre kredsløb giver dioderne D_1 og D_2 en spænding på omkring 1,4 V der ideelt svarer til transistorernes basis-emitter strækninger plus spændingsfaldet over modstandene i emitter. Dioderne kompenserer for transistorernes temperaturkoefficient på cirka $-2,5 \text{ mV}^\circ\text{C}$ hvis de kan placeret så de følger temperaturen i udgangstransistorerne. Det er dog sværere end det lyder for dioderne er vanskelige at bringe i termisk kontakt med kølepladen og kølepladens temperatur er desuden en del lavere end udgangstransistorernes temperatur inde på chippen.

I det højre kredsløb er transistor T_3 en såkaldt forstærket diode hvor kollektor-emitter spændingen fastlægges af spændingsdelen gennem R_3 og R_4 . Det er en populær opstilling som kan gøres justerbar ved at indskyde et potentiometer ved R_3 for indstilling til netop det arbejdspunkt der giver den laveste forvrængning. Spændingen U_{BE} for T_3 er forstærket med faktor $1 + R_4/R_3$ og ligeledes gælder for temperaturkoefficienten.

Spændingen over R_{E1} og R_{E2} stiger med strømmen i transistorerne og det reducerer spændingen over basis-emitter strækningen som en stabiliserende tilbagekobling. Strømgeneratorerne I_{B1} og I_{B2} giver strømmen til basis udgangstransistorerne T_1 og T_2 og repræsenterer enten det forstærkende element (T_5 i forrige kapitel) eller en strømgenerator. I de ældre effektforstærkere blev den ene gren typisk realiseret ved boot-strap, som vist på side 290.

DC strømniveauet I_{CQ} (quiescent current) er på 20 til 50 mA for typiske klasse AB effektforstærkere og indstilles ved at holde en konstant værdi af hvilespændingsfaldet U_{BIAS} over R_{E1} og R_{E2} . Der er ingen "gylden værdi" for DC strømmen så det optimale arbejdspunkt må findes ved at måle den harmoniske forvrængning. Det indskyder desværre et testpunkt i produktionen, men det er svært at serieproducer klasse AB uden dette fordyrende testpunkt. Forvrængningen stiger voldsomt når strømmen er for lav (cross-over distortion) mens en for høj strøm primært er en termisk udfordring.

Effekttransistorer har en lav værdi af strømforstærkningen idet basislaget er relativt tykt og det stiller krav til forstærkeren (transistor T_5 fra det tidligere kapitel) som skal leverer en betydelig strøm til udgangstransistorernes basis. For en udgangsstrøm på 5 A er strømmen i basis på 100 mA ved en strømforstærkning på 50 så effektforstærkere benytter almindeligvis to transistorer i hver gren. Herunder vises tre af de mange muligheder hvor X og Y refererer til forrige side.



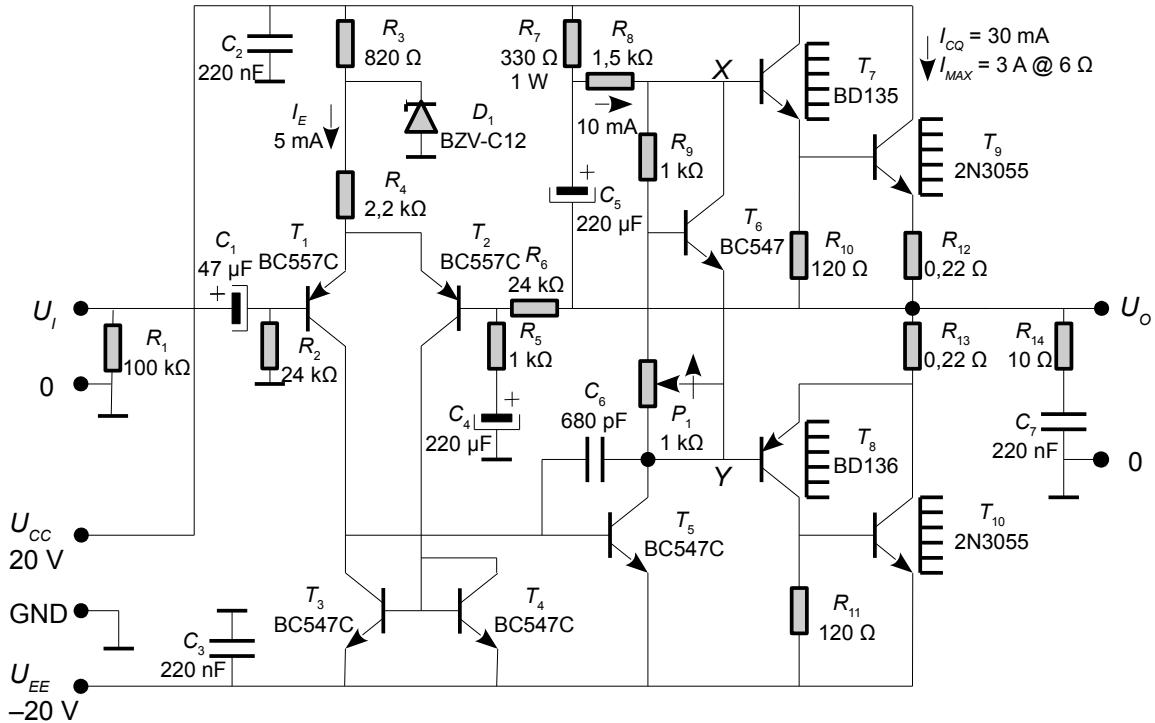
Figur 327 – Udgangstrinnet benytter stort set altid dobbelte transistorer for at hæve den effektive værdi af strømforstærkningen. Her er vist to klassiske udgangstrin med BJT og desuden et trin med en kombination af BJT og MOSFET.

Til venstre ses to dobbelte emitterfølgere (EE) der populært kaldes for et *Darlington pair* hvor T_1 forstærker strømmen til basis af T_2 og tilsvarende for T_3 og T_4 . Der er fire basis-emitter strækninger fra X til Y, men arbejdstemperaturen varierer mellem transistorerne så det er svært at holde en konstant værdi af strømmen I_{CQ} . Enkelte effekttransistorer leveres med dioder på siliciumchippen for en mere præcis tilbagekobling fra temperaturen. En god egenskab er at R_{BB} forbedrer funktionen ved høje frekvenser ved at potentiometret på basis kan trækkes forbi potentiometret på emitter så rumladningen i basis-emitter dioden hurtigere kan fjernes.

I midten ses en to-trins forstærker der danner en forstærket emitterfølger der også beskrives som *complementary feedback pair* (CFP). Drivertransistorerne T_1 og T_3 er ikke i samme grad som før følsom for variation i temperaturen fordi udgangstransistorernes temperatur ikke indvirker på spændingsfaldet mellem X og Y. Det er derfor lettere at holde DC strømmen konstant, men det betyder ikke at CFP er bedre end EE som vist af Douglas Self [DS-385]. Egenskaberne ved høje frekvenser er lidt ringere end for EE idet transistorerne i udgangen kun kan skiftes fra ON til OFF gennem tømning af rumladningen i basis-emitter via R_{C1} og R_{C2} . Det betyder at modstandenes værdier skal være relativt lave og det giver en høj belastning af drivertransistorerne.

Til højre præsenteres en variant med en MOSFET i udgangen. Der skal ligge cirka 5 V over R_{C1} og R_{C2} for at åbne udgangstransistorerne og det lægger en binding på effektforsyningen så mange MOSFET effektforstærkere benytter flere linjer i effektforsyningen for at holde det totale effektab nede. En MOSFET har ingen betydende strøm i gate, men den store kapacitet i indgangen skal op- og aflades for at transistoren kan skifte mellem ON og OFF. Det kræver en strøm på C_{dL}/dt og for en kapacitet på 5 nF med en ændringshastighed på 5 V/μs giver det 25 mA der skal leveres via T_1 og T_3 for at åbne for T_2 og T_4 og gennem modstandene R_{C1} og R_{C2} for at afbryde for dem.

Herunder kombineres materialet til prototypen på en effektforstærker for 20 W i 8 Ω. Udgangen har en dobbelt emitterfølger EE i den øvre gren og en forstærket emitterfølger CFP i den nedre gren. Det kaldes for *quasi complementary* og tillader brug af to ens NPN effekttransistorer. Det var ret udbredt før 1980 hvor effekttransistorer af typen PNP var meget ringere end NPN. Der kan opnås en lavere forvrængning med to gange EE eller to gange CFP [DS-151].



Figur 328 – Prototype på en klasse AB forstærker for 20 W i 8 Ω med quasi komplementær udgang.
Der er ikke beskyttelse mod kortslutning af udgangen.

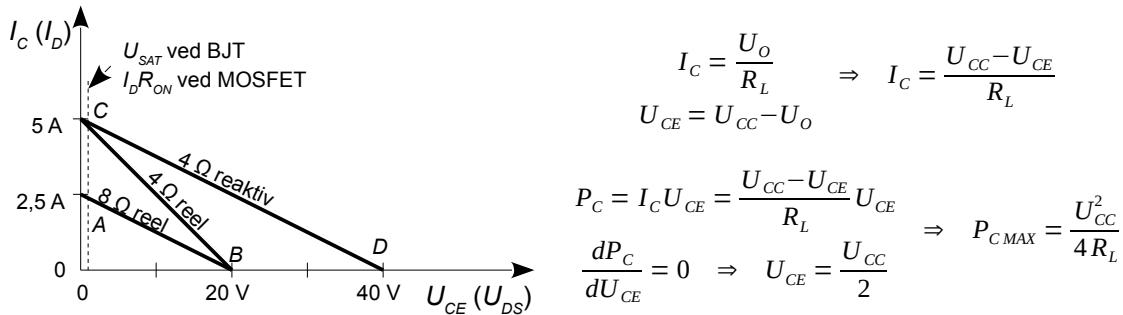
Designet starter ved udgangen for at fastlægge den nødvendige effektforsyning og kravet til driver-trinnet. For at a give 20 W i 8 Ω behøves en udgangsspænding på $U_{O\text{ RMS}} = 12,5 \text{ V}$ der er lig med 36 V spids-spids. For at have lidt overskud til indstilles effektforsyningen 10 % højere ved $\pm 20 \text{ V}$. Med en realistisk belastning på $R_L = 6 \Omega$ (se side 297) bliver strømmens spidsværdi 3 A og der leveres da $P_O = 26 \text{ W}$. Som vist på side 300 belastes udgangstransistorerne T_9 og T_{10} med 20 % hver (5,2 W). Da designet er klasse AB må der desuden inkluderes et tab fra DC bias i udgangstrinnet (0,6 W i hver gren). Med en strømforstærkning i udgangen på $\beta_9 = \beta_{10} = 30$ behøves en basisstrøm på 100 mA som leveres fra drivertransistorerne T_7 og T_8 og der afsættes op til 4 W i hver af dem ved fuld udstyring. Ved en strømforstærkning i driverne på $\beta_7 = \beta_8 = 25$ er der behov for 4 mA, men R_7 og R_8 er designet til 10 mA for at have et overskud. Strømmen i R_8 holdes konstant på grund af *boot strap* med C_5 der sikrer at spændingen over R_8 ikke varierer.

Med en forventet strømforstærkning i T_5 på $\beta_5 = 200$ er basisstrømmen 0,5 mA og indgangstrinnet bør have 10 gange højere strøm så den vælges til $I_E = 5 \text{ mA}$ der indstilles ved D_1 og R_4 . Fuld udstyring ved $U_{I\text{ RMS}} = 0,5 \text{ V}$ på indgangen opnås ved $A_{CL} = 1 + R_6/R_5 = 25$ og et valg af R_5 på 1 kΩ.

Et småsignaldesign beregner først transkonduktansen i $T_1 \dots T_4$ til $g_m = I_E/2U_T = 0,1 \text{ S}$. Åben-sløjfe forstærkningen beregnes fra indgangsmodstanden på udgangstrinnet ($R_C = \beta_7\beta_9R_L = 4,5 \text{ k}\Omega$) til en DC forstærkning på $A_{DC} = g_m\beta_5R_C = 90000$ ved $R_L = 6 \Omega$. Med et gæt på høje poler ved $f_H = 1 \text{ MHz}$ estimeres den dominerende pol til $f_0 = f_H/A_{DC} = 11 \text{ Hz}$. Den tilbagekoblede forstærkning er $A_{CL} = 25$ og hæver polen til $f_P = A_{CL}f_0 = 275 \text{ Hz}$. Det giver en kompensationen på $C_6 = 1/2\pi f_P\beta_5R_C = 640 \text{ pF}$. Forstærkerens GBP er på 1 MHz og det giver en båndbredde på 40 kHz ved -3 dB. Den opnåede slew rate bliver $SR = I_E/C_6 = 7 \text{ V}/\mu\text{s}$ der er tilstrækkelig for musikgengivelse.

6.2.4 Arbejdslinje

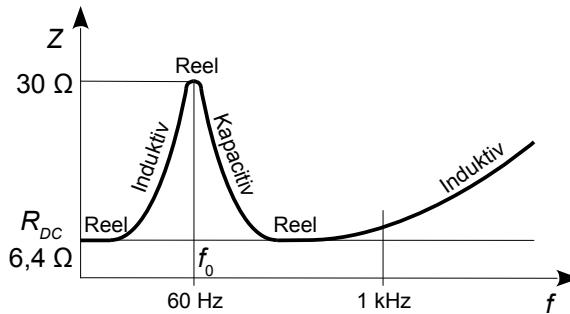
Belastningen af en effektforstærker regnes normalt for reel og der er da en lineær relation mellem udgangsspændingen U_O og strømmen I_O til belastningen og dermed I_C i udgangstransistoren. Det udtrykkes ved en arbejdslinje der viser sammenhængen mellem strømmen I_C og spændingen U_{CE} over transistorens kollektor-emitter eller tilsvarende I_D og U_{DS} for en MOSFET. Det giver en relation som går fra den maksimale strøm ved kortsluttet transistor til nul ved fuld forsyning over den. Det vises herunder som linjen fra A til B ved en belastning på $R_L = 8 \Omega$ og tilsvarende gælder for kurven fra C til D ved $R_L = 4 \Omega$. Talværdierne ved akserne er for 20 W effektforstærkeren på side 296.



Figur 329 – Arbejdslinjerne for en effektforstærker med 8Ω og 4Ω belastning samt med en induktiv belastning (4Ω reaktiv).

Transistoren kan ikke trække udgangen helt ud til effektforsyningen på grund af opbygningen med dobbelte samt mætringsspændingsfaldet på $U_{SAT} \approx 1$ V for en bipolær effekttransistor, eller kanalmodstanden $R_{ON} \approx 0,1 \Omega$ for en MOSFET, men det kan normalt ignoreres uden større betydning.

En højttaler er ikke alene beskrevet ved en modstand på 8Ω . Højttalerens svingspole har en DC modstand R_{DC} på typisk 80 % af den nominelle værdi (det er $6,4 \Omega$, men jeg har målt ned til $5,6 \Omega$) så strømmen er over 3 A ved maksimal udgangsspænding. Ved højttalerens mekaniske resonans vokser impedansen, og den falder tilbage til DC værdien ved middel frekvens, for igen at vokse mod høje frekvenser hvor svingspolens selvinduktion får betydning. Forløbet varierer mellem at være en kapacitiv, reel og induktiv belastning med tilsvarende variation i fasevinklen.



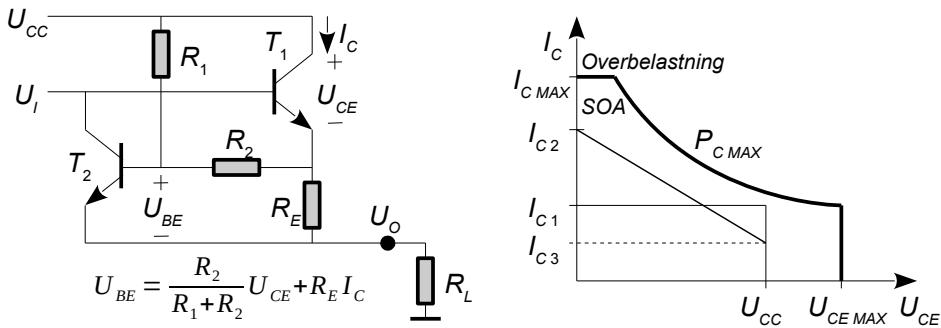
Figur 330 – Det typiske forløb af impedansen for en 8Ω bashøjttaler i et lukket kabinet.

For at inkludere en margin overfor den faktiske belastning for en effektforstærker på 8Ω designes for en reel belastning på 4Ω , hvilket betyder at 20 W forstærkeren skal kunne klare en spidseffekt i udgangstrinnet på 25 W, der kun mødes kortvarigt med et musiksignal. Den værst mulige reaktive værdi kan repræsenteres ved at designe efter den dobbelte værdi af spændingsfaldet for en reel belastning¹⁵² som vist ved arbejdslinjen fra C til D, der er worst case situationen [DS-446].

¹⁵²Der oplagres energi i en induktiv belastning når der løber en strøm, og strømmen vil kortvarigt blive ved med at løbe i samme retning selv om spændingen ændres over den induktive belastning. Hvis udgangen springer fra en høj værdi til nul vil strømmen i spolen trække udgangen ned under nul for at holde T_6 aktiv, og den får derved et større spændingsfald end 20 V over sig.

6.2.5 Safe-operating area, SOA

Modstanden R_E i emitter har til formål at stabilisere den indstillede strøm I_{CQ} overfor en variation i temperaturen, men den har endnu en vigtig funktion som del af et beskyttelseskredsløb, der gør det muligt at måle strømmen i udgangstransistoren, så den afsatte effekt altid holdes indenfor det tilladelige; transistorens *safe operating area*, SOA. Det er ikke nok at beskytte transistoren med en smeltesikring, for transistoren kan brænde af længe før sikringen når at afbryde. En smeltesikring kan dog være en god investering for at undgå følgeskader i effektforsyningen [DS-443].



Figur 331 – En ekstra transistor kan begrænse udgangsstrømmen ved overbelastning, så hverken den maksimalt afsatte effekt ikke overskrides uanset spændingen over transistoren.

En transistor skal beskyttes mod tre typer overlast: spænding, strøm og effekt. Den kan klare en spænding $U_{CE\ MAX}$ over kollektor og emitter uden sammenbrud på grund af den elektriske feltstyrke i PN overgangen. Den kan klare strømmen $I_{C\ MAX}$ uden problemer med strømtætheden i siliciummet, men ved større strøm er der risiko for punktopvarmning (sekundært sammenbrud) eller den interne tråd i emitteren kan smelte som en sikring. Endelig kan transistoren klare effekten $P_{C\ MAX}$ uden at opvarmes til destruktion. PN overgangen er dannet af en forurening med type III og V halvledere, og koncentrationsforskellen udlignes ved diffusion, der accelereres voldsomt med opvarmning (søg fx på nettet efter Arrhenius). Ved overskridelse af *Absolute Maximum Ratings* fra transistorens datablad risikeres alt fra en nedsat levetid til et momentant sammenbrud¹⁵³.

Transistor T_2 og modstandene R_1 , R_2 og R_E i figuren ovenfor danner et typisk beskyttelseskredsløb. Strømmen I_C giver en spænding over R_E og hvis den når over 0,7 V vil T_2 trække strøm og derved forhindre T_1 i at øge strømmen yderligere. Hvis effektorstærkeren skal overleve en kortslutning af udgangen til nul vil hele forsyningen på U_{CC} ligge over transistoren. Den afsatte effekt $P_C = I_{C1}U_{CC}$ skal også i den situation holdes under det maksimalt tilladelige $P_{C\ MAX}$ for transistoren.

2N3055 kan tåle en effektafsættelse på $P_C = 100$ W ved 50°C på monteringsfladen (køleprofilet). Ved en effektforsyning på $U_{CC} = 20$ V er den maksimalt tilladelige strøm i kollektor på $I_{C1} = P_C/U_{CC} = 5$ A. Benyttes $U_{BE} = 0,7$ V bliver den mindste værdi af modstanden i emitter $R_E = U_{BE}/I_{C1} = 0,14$ Ω.

Modstanden R_2 bør inkluderes som beskyttelse af T_2 mod for høj strømværdi, hvis transistor T_1 er lidt sløv i optrækket, men spændingen over modstanden må ikke genere funktionen, hvilket leder til en værdi omkring $R_2 = 100$ Ω.

Ved en maksimal strøm i kollektor af T_2 på 100 mA og en strømforstærkning på 100 bliver strømmen 1 mA i basis. Spændingsfaldet over R_2 andrager da 0,1 V og kortslutningsstrømmen hæves med 15 %.

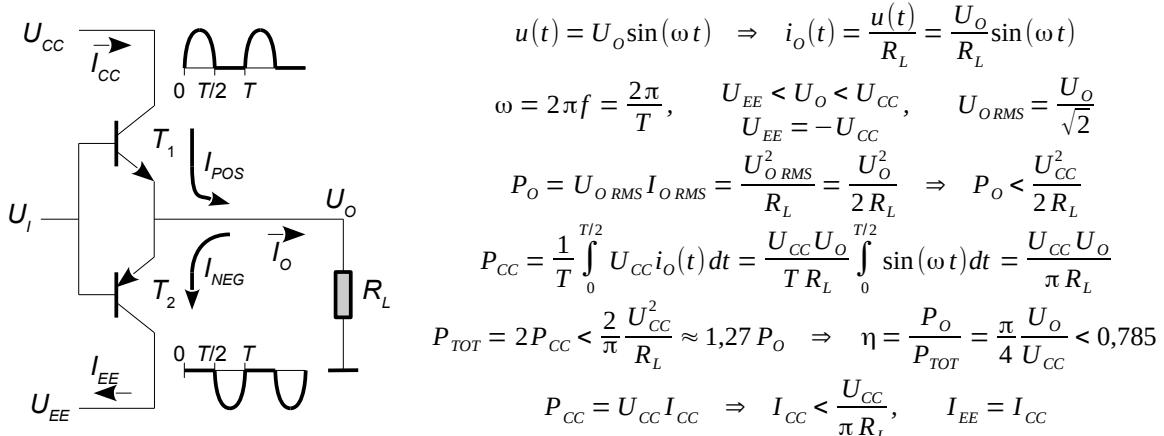
Ved høj spænding fra effektforsyningen kan det være svært at begrænse den afsatte effekt ved en kortslutning uden at påvirke den normale funktion. Strømmen er $I_{C2} \approx U_{CC}/R_L$ ved den højest mulige spænding på udgangen, hvor $U_{CE} \approx 0$. Ved at inkludere R_1 kan strømmen reduceres ved kortsluttet udgang til $I_{C3} < P_{C\ MAX}/U_{CC}$.

¹⁵³Ved et praktisk forsøg blev en småsignaltransistor BC847 (SOT23, 45 V og 0,3 W) udsat for 120 V over kollektor-emitter uden at den brød sammen. Den afsatte effekt blev indstillet til 2 W ved den høje spænding, men den fortsatte med at fungere selv om den flød rundt i smeltet loddetin. Ved at øge effekten til 2,5 W brændte den øjeblikkeligt sammen.

Note 14.

6.2.6 Effektberegning

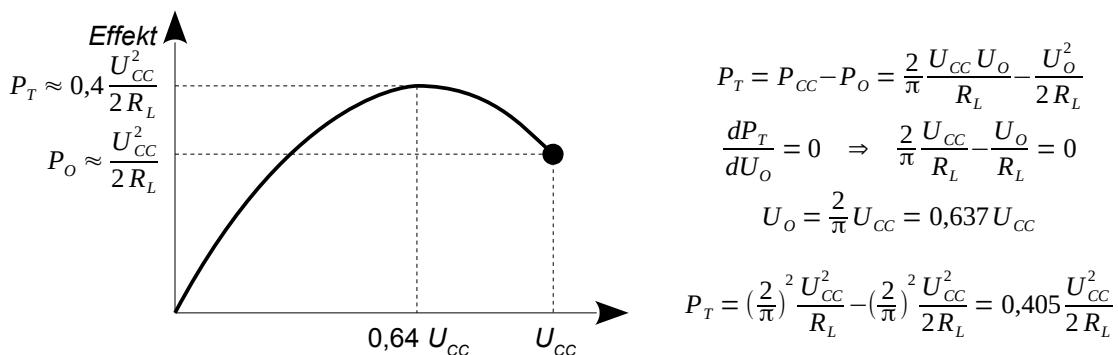
I dette afsnit beregnes kravet til effektforsyningen samt den afsatte effekt i udgangstrinnet. Effektforsyningen skal totalt levere P_{TOT} på knapt 30 % over udgangseffekten P_o ved fuld udstyring. Det kræver en middelværdi af strømmen I_{CC} på en tredjedel af spidsstrømmen til belastningen R_L og det gælder for hver gren af effektforsyningen.



Figur 332 – Udgangseffekten er begrænset til P_o af effektforsyningens spænding U_{CC} og samlet skal der kunne leveres 27 % mere end den maksimale udgangseffekt fra effektforsyningen.

For $P_o = 20$ W til $R_L = 8 \Omega$ er $U_o = 18$ V og $I_o = 2,2$ A så der kræves $U_{CC} = -U_{EE} > 18$ V og $I_{CC} = I_{EE} > 0,7$ A. Ved et design efter teoriens minimum på $U_{CC} = 18$ V skal effektforsyningen kunne leve $P_{TOT} = 25$ W.

Som en øvre grænse i analysen benyttes $U_o = U_{CC}$ og det giver den ofte citerede virkningsgrad for klasse B på 78 %. I praksis vil amplituden ikke nå op til effektforsyningen da transistorerne har et mætningsspændingsfald og dertil andre konstruktionsdetaljer så en klasse AB effektforstærker har næppe en virkningsgrad over 70 %. Det er dog af mindre betydning for det er langt vigtigere at den afsatte effekt er en funktion af udgangseffekten. Det betyder at klasse B ikke afsætter en stor effekt ved lavt udgangsniveau.



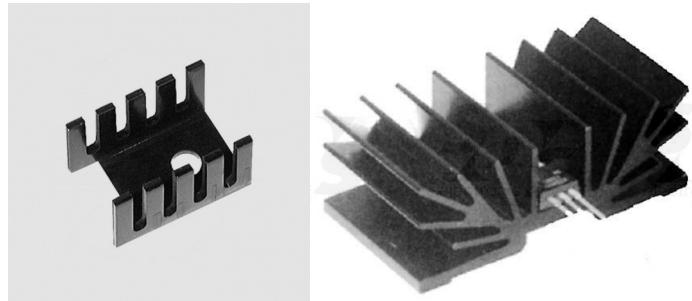
Figur 333 – Den afsatte effekt i udgangstransistorerne er tilsammen under 40 % af den maksimalt mulige udgangseffekt og værdien nås ved mindre end fuld udgangseffekt.

Med $U_o = U_{CC} = 20$ V og $R_L = 8 \Omega$ bliver den maksimale effektafsættelse i udgangstrinnet på $P_T = 8$ W, så hver transistor belastes med 4 W.

Den afsatte effekt i udgangstransistorerne P_T er den andel, der er mangler i effektregnskabet og den beskrives ved en parabel med toppunktet ved 64 % af spændingen for fuld udgangseffekt. Der afsættes derfor 40 % af fuld udgangseffekt ved den værst tænkelige amplitude af udgangssignalet. For en 100 W klasse B effektforstærker afsættes der $P_T = 40$ W så hver udgangstransistor skal optage 20 W hvor der til sammenligning afsættes fem gange så meget ved 100 W i klasse A.

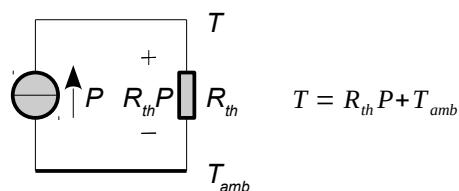
6.2.7 Termisk design

Udgangstransistorerne skal køles for ikke at blive for varme, og det gælder meget ofte også for drivertransistorerne. Hertil benyttes køleprofiler hvor de små typer blot er bukket aluminium mens de større profiler benytter extruderet aluminium som skæres ud i passende længde alt efter hvor gode de skal være til at fjerne varmen fra transistoren. To eksempler vises herunder for køling af transistorer med den populære hustype TO220.



Figur 334 – Køleprofiler (venstre) med målene $19,1 \times 13,2 \times 12,7 \text{ mm}^3$ og termisk modstand 27 K/W samt (højre) $22,5 \times 37,5 \times 54 \text{ mm}^3$ med termisk modstand $5,5 \text{ K/W}$ (Elektronik Lavpris).

En køleprofil beskrives normalt ved en termisk modstand og Ohms lov antages at gælde for det termiske system. Man siger at effektstrømmen P løber gennem den termiske modstand R_{th} og det danner et temperaturfald på $R_{th}P$ som hæver temperaturen T over omgivelsestemperaturen T_{amb} . Den termiske modstand kaldes undertiden for køleprofilets K-værdi og enheden er en stigning i temperatur per effekt. Det er $^{\circ}\text{C}/\text{W}$ eller K/W der her betyder det samme.



Figur 335 – Beregning af temperaturstigning ved en effektstrøm P i en termisk modstand.

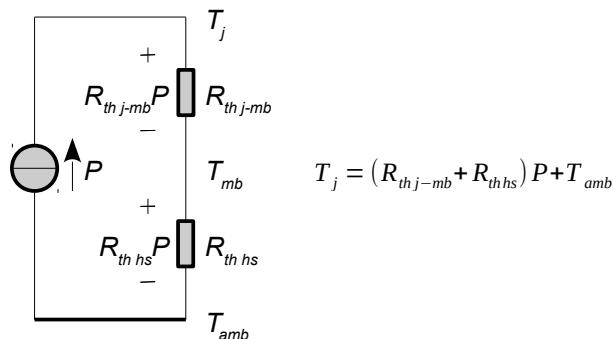
Med den store køleprofil på $R_{th} = 5,5 \text{ K/W}$ vil en afsat effekt på $P = 4 \text{ W}$ hæve temperaturen med 22°C og med en omgivelsestemperatur på $T_{amb} = 25^{\circ}\text{C}$ bliver temperaturen ved monteringsstedet på $T = 47^{\circ}\text{C}$.

Den termiske modstand af et køleprofil er almindeligvis fra 1 til $50^{\circ}\text{C}/\text{W}$. Det er muligt at opnå en lavere værdi, men køleprofilet bliver både stort og kostbart. En negativ værdi kan kun opnås ved en energikrævende varmetransport som fx ved et Peltier-element. En højere værdi er sjældent af interesse da det svarer til komponentens egen værdi af den termiske modstand til omgivelserne. Som et par eksempler er en MOSFET som IRFZ44N specificeret til $60^{\circ}\text{C}/\text{W}$ som den termiske modstand fra den interne chip til omgivelserne. En småsignaltransistor som BC547 har tilsvarende en værdi på $250^{\circ}\text{C}/\text{W}$ og kan selvfølgelig drage nytte af køling, men det er sikrest at benytte en type der er beregnet for effektafsættelse og som kan monteres på køleprofilet med en skrue eller en fjeder. Det er almindeligvis ikke helt let med en småsignaltransistor.

Varmen i transistoren udvikles i halvlederens PN overgang og varmen løber gennem den interne termiske modstand til monteringsfladen og derfra til omgivelserne gennem køleprofilet. For en TO220 er den termiske modstand $R_{th-j-mb}$ på 2 til $4^{\circ}\text{C}/\text{W}$ fra PN overgangen (*junction*) til monteringsflade (*mounting base*). Værdien opgives i databladet for effekttransistoren.

Hvis monteringsfladen kan holdes på 47°C vil en transistor med $4^{\circ}\text{C}/\text{W}$ fra halvlederens PN overgang til monteringsflade blive opvarmet til 63°C ved en afsat effekt på 4 W .

For at beregne temperaturen i transistorens halvleder (kaldet *junction* eller *chip*) kan beregningen udvides med den termiske modstand $R_{th\ j-mb}$ for transistoren. Den afsatte effekt P i transistoren vil løbe igennem en serieforbindelse af den indre termiske modstand i transistoren og den eksterne termiske modstand for køleprofilet, der her kaldes $R_{th\ hs}$ hvor hs står for *heat sink*.

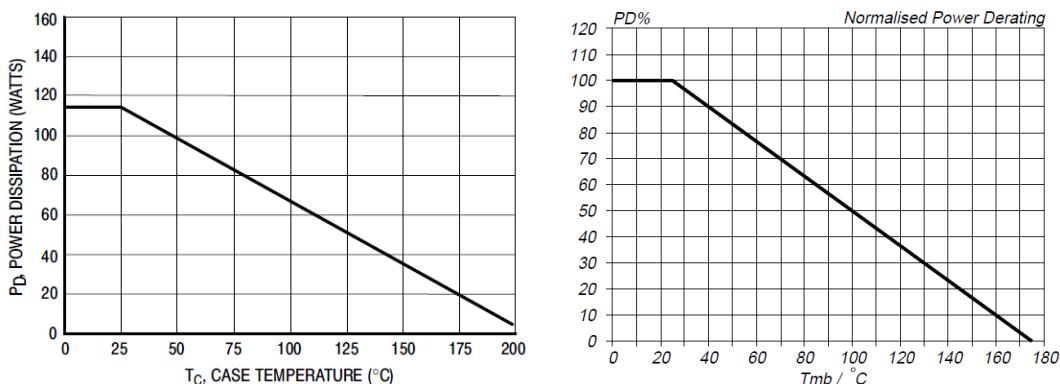


Figur 336 – Beregning af chippens temperaturstigning ved en transistor monteret på et køleprofil.

Med en termisk modstand i transistoren på $R_{th\ j-mb} = 4 \text{ K/W}$ og et køleprofil med $R_{th\ hs} = 5,5 \text{ K/W}$ vil en afsat effekt på $P = 4 \text{ W}$ hæve temperaturen i halvlederenes chip til $T_j = 63^\circ\text{C}$.

Der er en grænse for hvor varm en transistors halvleder må blive for ikke at reducere levetiden og den grænse ligger i intervallet fra 125°C til over 200°C afhængigt af typen. Transistoren brænder ikke af ved højere effekt end databladet angiver som den maksimalt tilladelige belastning ved korte pulser, men det er en farlig vej at gå for en designer af effektforstærkere, der jo ikke arbejder med kortvarige pulser, så det anbefales at overholde databladets statiske angivelser.

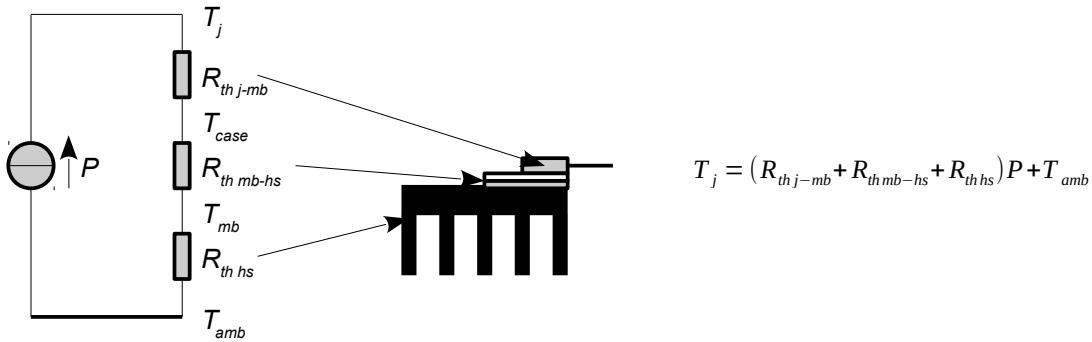
Herunder vises den maksimalt afsatte effekt som funktion af den temperatur køleprofilet giver ved monteringsstedet for to forskellige transistorer. Kurverne er typiske med en konstant effekt op til en temperatur på 25°C og med en værdi på 0 W ved en højere temperatur, der for 2N3055 er noget over 200°C og for IRFZ44N er på 175°C .



Figur 337 – Reduktion af afsat effekt som funktion af temperaturen ved køleprofilet. Til venstre for den gamle traver 2N3055 (ON semiconductor), og til højre for IRFZ44N (Philips) hvor 100 % svarer til en afsat effekt på 94 W.

For de viste transistorer vil en temperatur ved køleprofilet på 47°C tillade en afsat effekt på 100 W ved 2N3055 og 80 W for IRFZ44N (regnet som 85% af 94 W).

Det er ofte nødvendigt at isolere transistoren fra køleprofilet med en isolationsskive af glimmer eller silicone og det medfører en temperaturforskæl mellem transistorens hus og køleprofilets overflade. Værdien afhænger blandt andet af isolationsskiven med en typisk værdi i omegnen af 2°C/W . Den kan inkluderes i den termiske model som endnu en termisk modstand $R_{th\ mb-hs}$ med den værdi som producenten angiver i databladet.

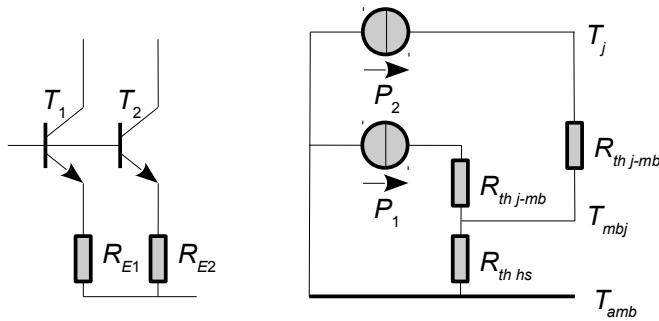


Figur 338 – Beregning af temperaturstigning ved en transistor monteret med isolationsskive.

Der er en række af antagelser ved det termiske design som man kunne tage hensyn til, men det er som regel tilstrækkeligt med den viste metode. Målet bør ikke være at nå en specifik temperatur ved en given effektafsættelse, men at designe konservativt. Det er bedst at designe med en stor margin præcis som man gør ved et design af en bro hvor der kan opstå meget store belastninger ved en uheldig kombination af køretøjer eller ved usædvanlige vejrforhold.

Hvis køleprofilet sidder udenpå kabinetet vil der være nogle krav til den temperatur som kunden kan forventes at acceptere. Det vil være i omegnen af 50°C for ved en højere temperatur på køleprofilets overflade vil brugeren kunne risikere en forbrænding. Hvis transistoren kan berøres vil der tilsvarende gælde at dens overfladetemperatur ikke må være ret meget højere. Temperaturen inde i selve halvlederen er næppe mere end 50°C højere end køleprofilet og derved bliver designkravet en chip-temperatur i omegnen af 100°C.

Hvis køleprofilet sidder indvendigt i kabinetet kan temperaturen være højere uden at kunne skade brugeren, men transistorens temperatur kan blive for høj hvis køleprofilet ikke må belaste designet for hårdt økonomisk og derfor vælges mindre end optimalt. Her kan løsningen være at forbinde to eller flere transistorer i parallel. Den afsatte effekt i køleprofilet bliver det samme som ved en enkelt transistor, men temperaturen i PN overgangen i den enkelte transistor bliver noget lavere da effekten i hver transistor reduceres¹⁵⁴.



Figur 339 – Beregning af temperaturstigning ved to transistorer i parallel.

Parallelkobling af effektransistorer ses anvendt i effektforstærkere til høj effekt og for at sikre en nogenlunde ligelig fordeling af effekten benyttes modstande i emitter på typisk 0,1 Ω. Modstanden mindske afhængigheden af temperaturfølsomheden for spændingen over basis-emitter.

"For meget køling har endnu ikke skadet en transistor" som en bog for selvbyggere engang skrev og det bør enhver designer af effektforstærkere skrive sig bag øret.

¹⁵⁴Jeg har engang designet en 600 W forstærker for industriel anvendelse hvor der var fire transistorer i parallel i hver gren. Det var ikke kun et ønske om at reducere den termiske belastning, men også et ønske om at inkludere en god margin. Det viste sig at være en god idé for belastningen af forstærkeren var forkert specificeret, hvilket fordoblede effektafsættelsen, og dertil kom at den valgte transistor pludselig blev svær at få fat i og måtte erstattes af en anden type med en større termisk modstand.

6.2.8 Beregning af termisk modstand

Den afsatte effekt P i transistoren ledes ofte til køleprofilet gennem en isolationsskive eller der er benyttet en aluminiumsvinkel som leder varmen fra en eller flere transistorer til samme køleprofil. Det indskyder en termisk modstand R_{th} og en temperaturstigning ΔT . Den termiske modstand kan beregnes fra en materialekonstant λ (den specifikke varmeledning), længden L i den retning som varmen ledes og arealet A af det tværsnit som varmen passerer gennem [AL-7].

$$\Delta T = R_{th} P \quad R_{th} = \frac{L}{\lambda A}$$

λ_{Kobber}	\approx	385 W/Km
$\lambda_{\text{Aluminium}}$	\approx	205 W/Km
λ_{Jern}	\approx	45 W/Km
λ_{Plast}	\approx	0,3 W/Km

En isolationsskive af plast har tykkelsen $L = 0,1$ mm og tværsnittet er 10 mm x 15 mm så den termiske modstand bliver $R_{th} = 2,2$ K/W. En lignende værdi findes for de noget tykkere silicone-isolationsskiver.

Det er sjældent at man dimensionerer sin egen køleprofil, men det kan være nødvendigt at benytte noget eksisterende som en metalplade i et kabinet. Herunder introduceres en beregningsmetode hvor det antages at køleprofilet overalt har samme temperatur og som et beregningseksempel benyttes den store køleprofil fra side 301.

Varmen fra køleprofilet afgives til omgivelserne gennem varmestråling og konvektion. Effekten P_R afgives som infrarød elektromagnetisk stråling til omgivelserne med temperaturen T_{AMB} . Dertil vil effekten P_K afgives ved konvektion hvor luften tæt på køleprofilet opvarmes, stiger op og trækker kold luft til. Den termiske modstand af køleprofilet $R_{th HS}$ (HS for *heat sink*) kan beregnes når overfladetemperaturen T_{HS} er kendt, men den indgår også i de tilhørende beregningsudtryk så der skal bruges en iterativ proces. Man starter med at gætte på en temperatur T_{HS} og beregner de afgivne effekter ved stråling og konvention. Hvis summen af de to effekter er mindre end den afsatte effekt P må T_{HS} hæves og beregningen gentages. Det er tilstrækkeligt med to eller tre beregningstrin.

$$R_{th HS} = \frac{\Delta T}{P} = \frac{T_{HS} - T_{AMB}}{P_R + P_K}, \quad P = P_R + P_K$$

Med $T_{HS} = 50^\circ\text{C}$ ved overfladen af et køleprofil, en omgivelsestemperatur på $T_{AMB} = 25^\circ\text{C}$ og en samlet effektafsættelse på $P = 1,3 \text{ W} + 3,0 \text{ W} = 4,3 \text{ W}$ (se eksemplerne nedenfor) beregnes $R_{th HS} = 5,8 \text{ K/W}$ der skal sammenlignes med specifikationen på 5,5 K/W for det store køleprofil.

Varmestråling betegnes med P_R for den overførte effekt hvor emissionskonstanten e angiver den betydning overfladen har, σ er Stefan-Boltzmanns konstant og A_R er det projicerede areal hvor køleprofilet betragtes som en kasse uden detaljerne ved køleribberne. Dertil T_K for køleprofilets temperatur i kelvin og T_{AMB} for omgivelsestemperaturen [AL-21 og Wikipedia].

$$P_R = e \sigma A_R (T_{HS}^4 - T_{AMB}^4) \quad \begin{aligned} e_{\text{Maling}} &\approx 0,9 \\ e_{\text{Poleret}} &\approx 0,1 \end{aligned} \quad \sigma = 5,67 \cdot 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$$

For det store køleprofil er de ydre mål 22,5 x 37,5 x 54 mm³ så $A_R = 0,0082 \text{ m}^2$. Ved 50°C på køleprofilets overflade ($T_{HS} = 323 \text{ K}$) og en omgivelsestemperatur på 25°C ($T_{AMB} = 298 \text{ K}$) afgives $P_R = 1,3 \text{ W}$ for en målet overflade der er antaget som repræsentativ for sort eloxering.

Konvektion betegnes med P_K for effekten overført til luften tæt på køleprofilet. Den afsatte effekt er givet ved det samlede areal A_K luften passerer langs med køleprofilets lodrette højde L . Værdierne indsættes i SI-enheder og effekten P_K fås i watt. Formlen er stærkt simplificeret [AL-16].

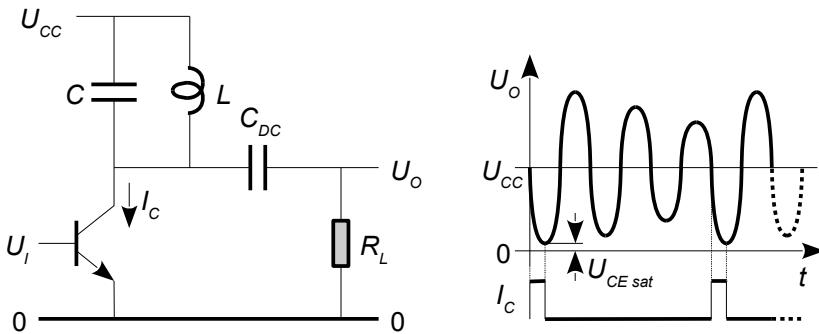
$$P_K \approx 1,2 \frac{A_K}{L^{0,25}} (T_K - T_{AMB})^{1,25}$$

For det store køleprofil er der 14 nogenlunde ens finner og dermed i alt 28 overflader og med et gæt på målene 20 x 37,5 mm² som gennemsnit for finnerne bliver arealet $A_R = 0,02 \text{ m}^2$ og højden er $L = 37,5 \text{ mm}$. Ved en overfladetemperatur på 50°C og 25°C ved omgivelserne bliver den afgivne effekt $P_K = 3,0 \text{ W}$.

Note 15.

6.2.9 Klasse C

Dette princip anvendes primært indenfor radiotransmission hvor udgangssignalet er en højfrekvent svingning der enten føres til en antenné eller et efterfølgende kredsløb. Svingningen dannes i en LC svingningskreds og transistorens formål er at opretholde svingningen ved at give et skub på det rette tidspunkt. Den aktive komponent i klasse C trækker strøm i mindre end 50 % af tiden og det er selve definitionen på klasse C der typisk har en virkningsgrad omkring 85 %. Hvis transistoren er aktiv i 50 % af tiden er det en én-transistor klasse B, over 50 % af perioden er det klasse AB og når transistoren er aktiv i hele perioden arbejder den i klasse A [KBR-472].



Figur 340 – Et klasse C udgangstrin er aktivt i mindre end 50 % af perioden, og anvendes fx i radiosendere, hvor en LC svingningskreds får tilført effekt i korte pulser for at opveje den afgivne effekt til en antenné. Udgangsfrekvensen kan være en af de harmoniske til indgangssignalet.

I illustrationen svinger signalet på kollektor omkring den positive effektforsyning U_{cc} og spids-spids værdien kan komme tæt på to gange effektforsyningen. Transistoren begrænser den negative del af svingningen, som for en BJT nås ved mætningsspændingsfaldet $U_{CE\text{ sat}}$ på cirka 0,5 V, mens en MOSFET har en indbygget diode i drain-source og det giver en begrænsning ved 0,7 V under nul. Kondensator C_{DC} blokerer for DC strøm til belastningen.

Forvrængningen af indgangssignalet er af mindre betydning ved højfrekvens da transistoren driver en afstemt kreds som dæmper overtonerne i kollektorens strøm. Ved radiotransmission efterfølges svingningskredsen af et afstemt filter for yderligere reduktion af de harmoniske så internationale standarder overholdes.

Styresignalet kan være et analogt signal, hvorved strømmen i kollektor er funktion af amplituden på indgangen, eller strømmen i kollektor kan være konstant med effekten styret af et PWM signal.

En ændring af oscillationens frekvens eller amplitude kan anvendes til at modulere tale, musik eller digital information ind på en bærebølge. Frekvensen kan ændres i forhold til den naturlige frekvens f_0 af LC svingningskredsen med godheden Q som den bestemmende faktor. Båndbredden af en LC svingningskreds er $B = f_0/Q$ så en oscillator på 100 MHz kan trækkes 500 kHz til hver side ved en reduktion på 3 dB i amplitude for en godhed på $Q = 100$. Amplituden af svingningen kan på tilsvarende vis ændres hvis det sker langsomt i forhold til båndbredden.

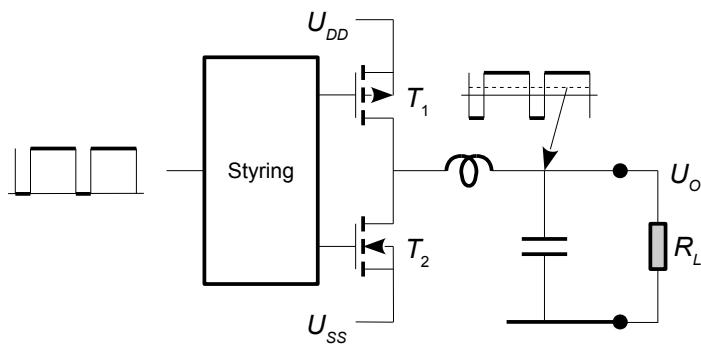
Klasse C kan anvendes til multiplikation af frekvensen fra et krystal. Svingningskredsen LC vil her afstemmes til anden eller tredje harmoniske af det pulserende signal så kredsen oscillerer på en højere frekvens end det indkommende signal.

Klasse C er sjælden indenfor lavfrekvens hvor oscillation på en fast frekvens ikke er centralt, men den kan mødes ved opretholdelse af svingningen i et pendul for et pynteur. En klasse B forstærker uden biaskredsløb kan opfattes som en symmetrisk udgave af klasse C for udgangstransistorerne er hver især aktive i mindre end 50 % af tiden. Tilsvarende gælder i et vist omfang for klasse D selv om grundidéen her er markant anderledes.

6.2.10 Klasse D

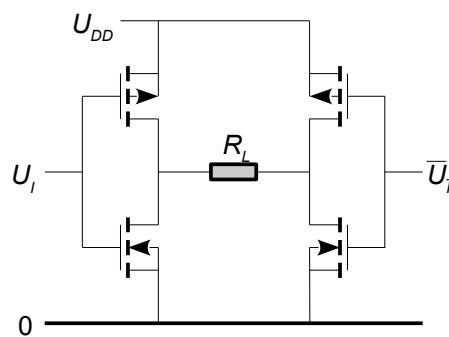
Princippet blev beskrevet af Baxandall i 1959 og Osborne i 1968 som en effektiv måde at generere en sinussvingning med høj effekt for radiotransmission og er senere udnyttet ved effektforstærkere til lydgengivelse. Det er i teorien muligt at opnå 100 % virkningsgrad, men en praktisk opstilling vil typisk kun kunne nå over 90 % da den øvrige elektronik skal tælles med i det samlede forbrug.

Transistoren anvendes her som en kontakt der skifter mellem kortsluttet og afbrudt: ON og OFF. Ved en kortsluttet transistor afsættes en minimal effekt da spændingen over transistoren er lav selv om der løber en stor strøm og med afbrudt transistor afsættes der ideelt set ingen effekt. Der er to transistorer i udgangen som skiftes til at være ON eller OFF så udgangssignalet skifter mellem de to ydergrænser givet af effektforsyningen. Skiftefrekvensen er over 100 kHz for lydgengivelse og gerne over 500 kHz hvis transistorerne er hurtige nok. Den integrerede elektronik i headsets og høreapparater kan have en lavere skiftefrekvens for at reducere effektforbruget ved skiftene hvor de interne kapaciteter i transistorerne skal op- og aflades.



Figur 341 – Et klasse D udgangstrin benytter transistoren som kontant og MOSFET foretrækkes på grund af dens korte skiftetid.

Udgangssignalet skiftes mellem den positive og negative effektforsyning (U_{DD} og U_{SS}) og der er et lavpasfilter i udgangen som kun tillader middelværdien af signalet at passere. Spændingen på udgangens er givet ved forholdet mellem den tid signalet er højt og den samlede tid for en periode. Det kaldes for PWM, *pulse width modulation*, og princippet er særdeles populært ved headsets og høreapparater hvor effektforbruget skal være lavt for at opnå mest mulig brugstid med et batteri samt ved PA anlæg hvor udgiften til køling af elektronikken kan holdes nede selv om der afgives en effekt på flere kilowatt.



Figur 342 – En brokobling giver en højere udgangseffekt og er populær i høreapparater og headsets med lav effektforsyning.

Et af problemerne ved klasse D er at transistorerne ikke samtidigt må være ON for så løber der en meget stor strøm i dem, men transistorerne har forskellige skiftetider for ON og OFF.

Udgangsspændingen er funktion af U_{DD} og U_{SS} så enhver støj på effektforsyningen slår direkte igennem til udgangen og det er nødvendigt at benytte en stabiliseret effektforsyning eller korrigere signalets PWM så der kompenseres for brum og støj. Der er derfor behov for en tilbagekobling der kan undertrykke støjen og dermed er vi tilbage ved problematikken med klasse B hvor en OpAmp skal kompensere for problemerne i udgangstrinnet. Brum fra ensretteren (se side 310) kan kompenseres ved *feed forward* hvor effektforsyningens spænding føres direkte til PWM kredsløbet uden at involvere tilbagekoblingens kredsløb.

Der er en del hybride klasse D forstærkere på vej indenfor kvalitets lydgengivelse og klasse D er givetvis fremtiden indenfor effektforstærkere.

6.2.11 Klasse E og F

Klasse E dækker over designprincipper indenfor radiotransmission hvor klasse C og D benyttes med et afstemt filter så udgangstrin, filter og antenne kombineres til ét design. Der er så vidt vides ingen definition af klasse F [DS-33].

6.2.12 Klasse G og H

I begge tilfælde er det et kredsløbstekniske tiltag for at reducere effekttabet. Ved klasse G benyttes to eller flere effektforsyninger i både den positive og negative gren og udgangstrinnet er designet til at trække energien fra den laveste af forsyningerne hvilket giver det laveste effekttab [DS-33].

Ved klasse H benyttes to klasse D effektforstærkere til at forsyne en almindelig forstærker med U_{CC} og U_{EE} der følger med signalets spidsværdi. Det betyder at udgangstrinnet kan være klasse A hvor man kan drage fordel af den gode linearitet uden at det store effekttab bliver et problem [DS-35].

6.2.13 Klasse XD

Cross-over displacement benytter en almindelig klasse AB effektforstærker hvor en strømkilde trækker cross-over punktet væk fra det følsomme område nær nul [DS-330]. Det betyder i praksis at den ene af udgangstransistorerne kører i klasse A og at udgangstrinnet kun skifter til klasse B ved signalets spidser. Princippet er benyttet i den integrerede operationsforstærker LM324, men det er alligevel lykkedes for Douglas Self at patentere princippet.

Effektkredsløb

Note 16.

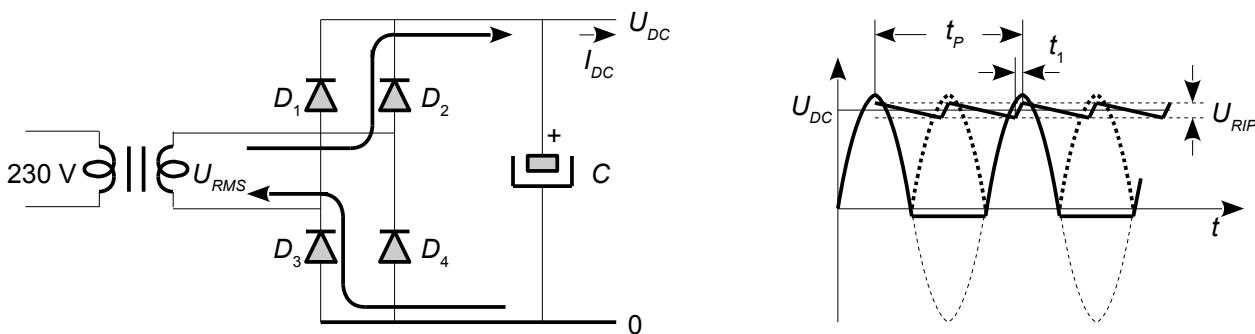
6.3 Effektforsyning

Elektronikken skal effektforsynes for at kunne virke og de umiddelbart foreliggende muligheder er at tage effekten fra bynettet på 230 V eller at benytte et batteri og det er det første af disse emner der skal behandles i det følgende. Andre muligheder er effektforsyning fra solen ved solpaneler, fra temperaturforskelle gennem brug af et Seebeck element (også kaldet for Peltier element) samt fra mekanisk arbejde så som vibration eller bøjning hvor den piezoelektriske effekt er brugbar, men disse emner falder alle udenfor rammerne af denne bog.

Bynettets 230 V vekselspænding¹⁵⁵ anvendes i stor udstrækning til en effektforsyning og det vises i dette afsnit. Derefter følger metoder til stabilisering af DC spændingen og afsnittet slutter med en analyse af hvordan en DC-DC konverter er opbygget. Som nævnt i bogens indledning er emnet lydreproduktion så omtalen af DC-DC konverteren vil kun berøre batteriforsyнет apparater så de kraftige omformere for vindmøller og elektrisk togdrift vil slet ikke blive omtalt. Kapitlet er derfor af orienterende karakter mere end det er en designmanual.

6.4 Ensretter

Det er muligt at ensrette bynettet direkte, men da den oplyste spænding er en RMS værdi betyder det en resulterende DC spænding på 320 V samt at opstillingen bliver berøringsfarlig. Omsætning til lavere spænding benytter en transformator hvis udgangsspænding ensrettes og udglattes¹⁵⁶.



Figur 343 – Effektforsyning med transformator, broensretter og elektrolyt kondensator, samt den pulserende DC spænding som følge af belastningsstrømmen.

De fire dioder D_1 til D_4 udgør en broensretter der også kaldes for en Graetz kobling.

Princippet i opstillingen er skitseret ved de to pile der viser strømretningen når den øverste leder fra transformatorens sekundær er positiv i forhold til den nederste. Dioderne D_2 og D_3 er forspændt i lederretningen mens dioderne D_1 og D_4 er spærrende og ved modsat polaritet ombyttes rollerne. De ledende dioder oplader en kondensator til en spændingsværdi tæt på den maksimale værdi af AC spændingen. Spændingen er normalt en sinus så den maksimale værdi er $\sqrt{2}$ gange med RMS værdien hvorfra skal trækkes to seriekoblede dioders spændingsfald¹⁵⁷.

$$U_{MAX} \approx 1,4 \cdot U_{RMS} - 1,4 \text{ V}$$

Med 18 V fra sekundæren af en transformator bliver spændingen efter ensretning på 24 V.

¹⁵⁵Efter harmoniseringen i slutningen af sidste årtusind er spændingen gradvist ændret fra det tidligere benyttede 220 V til 230 V og en direkte måling af spændingen fra en stikkontakt viser at harmoniseringen i det store og hele er gennemført i Danmark. England benyttede tidligere 245 V og har tilsluttet sig harmoniseringen. Det er almindeligt at kræve at forsyning fra bynet skal kunne klare området op til 250 V. Amerika benytter en spænding på 115 V ved 60 Hz og i Japan er spændingen 100 V. Togdrift benytter 25 kV ved 50 Hz, eller 15 kV ved 16,7 Hz gennem en roterende omsætter eller 1,5 kV DC (kilde: DSB).

¹⁵⁶De populære effektforsyninger for PC og småelektronik benytter ensretning af de 230 V hvorefter en DC-DC konverter varetager nedtransformering og galvanisk isolation så lavspændingen ikke er berøringsfarlig.

¹⁵⁷Undertiden ses en installation hvor transformatorens sekundær er på 24 V i den tro at så kommer der 24 V ud fra ensretteren. Efter ensretning vil elektrolyt kondensatoren dog nå op på 32 V der er 30 % over det forventede.

Formlen skal ikke tages for bogstaveligt og det gælder i øvrigt for alle formler i dette kapitel. For det første er indgangsspændingen kun lovet indenfor $\pm 10\%$ i Europa og ved nedtransformering bliver usikkerheden større på grund af transformatorens produktionstolerance. Spændingen på sekundæren er dertil funktion af belastningen og diodernes spændingsfald er kun cirka 0,7 V om end usikkerheden herfra er relativt ringe.

6.4.1 Ripplespænding

Dioderne trække kondensatorens spænding op på værdien U_{MAX} og efter en opladning leverer kondensatoren strømmen til belastningen så den aflades inden den næste puls fra broensretteren igen oplader kondensatoren. Denne svingning kaldes for ripplespændingen U_{RIP} hvor betegnelsen er hentet fra det engelske *ripple voltage* da der ikke er en god dansk betegnelse.

$$U_{RIP} \approx \frac{I_{DC} t_p}{2C} \quad \text{og} \quad U_{MIN} \approx U_{MAX} - U_{RIP}$$

DC værdien som ripplespændingen oscillerer omkring er $U_{RIP}/2$ lavere end maksimalværdien, men det er mere interessant at angive den laveste DC spænding der nås lige før kondensatoren igen bliver opladt. Det er den laveste spænding der er kritisk for om det efterfølgende kredsløb kan udføre sin mission. Det er $U_{MAX} - U_{RIP}$.

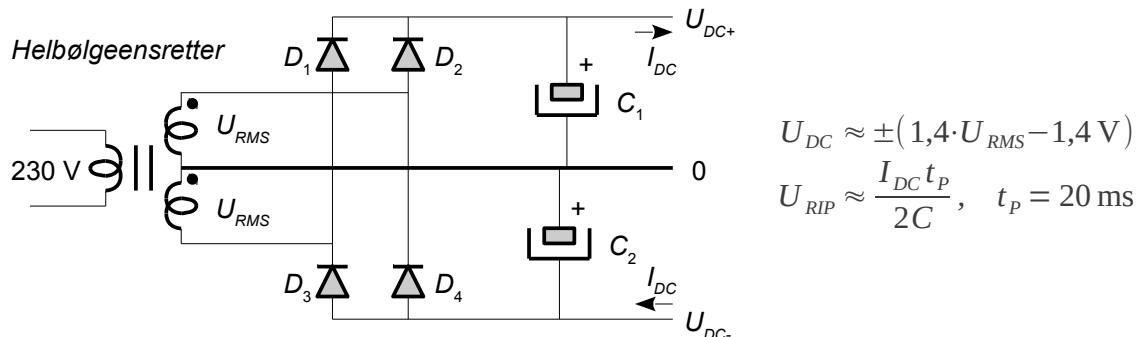
Kondensatorligningen kan benyttes til at beregne hvor meget spændingen falder i perioden med afladning. Det simpleste resultat nås hvis det antages at belastningsstrømmen er konstant over hele den analyserede periode hvilket i praksis er en glimrende approksimation.

$$\begin{aligned} dQ &= C dU = i dt \Rightarrow dU = \frac{i dt}{C} = \frac{I_{DC} t_p}{2C}, \quad t_p = 20 \text{ ms} \\ i &= I_{DC} \end{aligned}$$

Det antages at frekvensen er 50 Hz svarende til en periode på 20 ms, men selve analysen er generel og kan benyttes ved en anden frekvens som fx de 400 Hz i et fly.

6.4.2 Dobbelte effektforsyninger

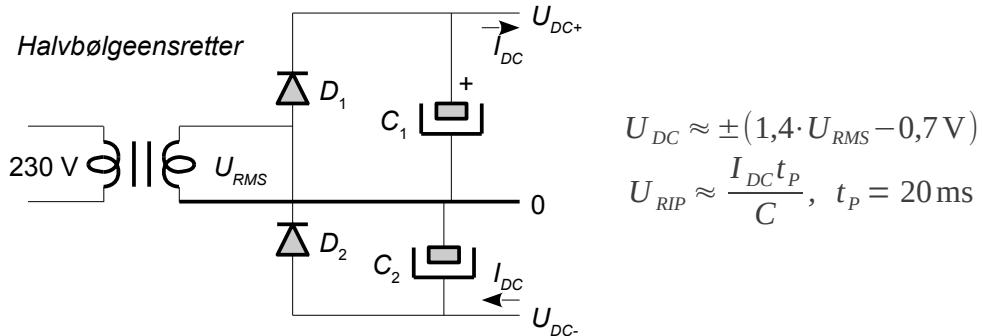
Den viste ensretter giver en effektforsyning med en enkelt spænding ligesom et batteri og som alt efter temperament kan være en positiv eller negativ værdi overfor nul. Meget elektronik har dog brug for to effektforsyninger som fx ± 15 V til analog elektronik eller ± 25 V til effektforstærkere og de kan dannes med en minimal ændring af kredsløbet.



Figur 344 – En transformator skal ofte give to komplementære spændinger der som regel belastes lige hårdt og transformatoren leveres da med to ens sekundærviklinger.

Dioderne i ensretterens bro kan være individuelle dioder og en ofte benyttet komponent er 1N4007 ved lave effekter eller broen kan fås som en færdig komponent alene specificeret ved strømstyrke og spændingsniveau. Sekundærens to prikker viser viklingernes polaritet. Det svarer til et batteris plusterminal og transformatoren skal kobles som vist for at fungere korrekt.

Transformatorer med kun en enkelt vinding kan udmærket danne to komplementære spændinger som vist herunder. Der er kun en enkelt diode i hver udgang så spændingstabet er mindre, men til gengæld oplades kondensatorerne kun ved 50 Hz. Ripplen bliver dobbelt så stor sammenlignet med helbølgeensretteren ved den samme forbrugsstrøm.

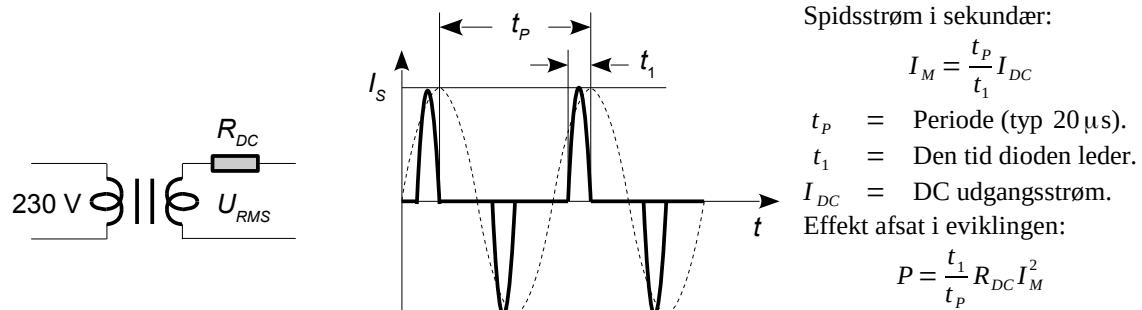


Figur 345 – En transformator med en enkelt vinding kan forsyne elektronik med både en positiv og en negativ forsyning, men de to forsyningslinjer bliver kun halvbølgeensrettet.

Der er som antydet en vis usikkerhed ved de indgående parametre så design af en effektforsyning er mere baseret på erfaring end formler. Det er almindeligt at man som udvikler finder frem til en underleverandør af transformatorer som man kommunikerer godt med og efter et par indledende forsøg rammer de efterfølgende design som regel rigtigt i første forsøg.

6.4.3 Effektforhold

Den afgivne effekt fra en effektforsyning kan ikke overstige den tilførte effekt minus tabene for ellers havde vi opfundet en evighedsmaskine¹⁵⁸,



Figur 346 – Strømmen i transformatorens sekundær løber kun i korte intervalle når dioderne er åbne, så strømmens topværdi bliver høj for at kunne leve forbrugsstrømmen.

Den pulserende strøm løber igennem transformatorens DC modstand R_{DC} fra bevirkelsen og det kan give anledning til en betragtelig effektafsættelse i transformatoren. Som en tommelfingerregel kan man antage at dioderne er aktive i cirka 25 % af tiden.

¹⁵⁸En 50 W transformator, der afgiver 25 V efter ensretning vil derfor højest kunne belastes med 2 A på udgangen. Hvis der er to udgange på hver 25 V kan hver af dem maksimalt belastes med 1 A.

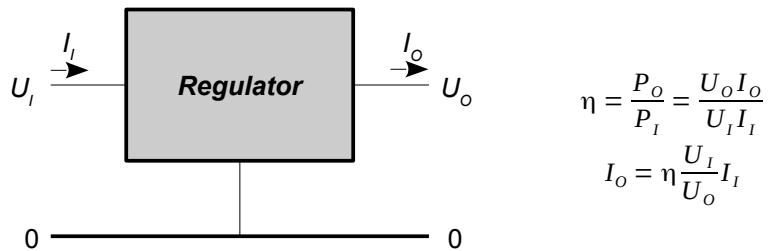
Note 17.

6.5 Spændingsregulator

Ensretterens DC spænding er tilstrækkelig til effektforstærkere for lydgengivelse via højtalere og ved styring af DC motorer. For brug ved analog signalbehandling og digitale kredsløb er der behov for en mere veldefineret spænding. Eksemplerne inkluderer 1,8 V og 3,3 V for mikroprocessorer og radiokredsløb, 5 V for digitale kredsløb og analog interface i mikrocomputersystemer, ± 15 V for analoge kredsløb, 24 V for PLC interface og 48 V for mikrofonforsyning (*Phantom power*).

Der er flere muligheder for at opbygge stabile spændingsforsyninger hvor der skal gennemgås en shuntregulator, en serieregulator og en DC-DC konverter. Listen over muligheder er dog en del længere, men de mere specielle konstruktioner som ladningspumpe, solpanelets power-tracking og frekvenskonvertere er ikke medtaget.

For at beskrive hvor meget af den tilførte effekt der kommer forbrugeren til gode benyttes et mål kaldet for regulatorens virkningsgrad (*efficiency*). Det angiver hvor meget af den tilførte effekt P_I der bliver overført til udgangen som afgiven effekten P_O . Det ideelle er at hele indgangseffekten kan overføres til belastningen på udgangen, men det nås ikke i praksis og den resterende effekt bliver afsat som varme og må skaffes væk fra regulatoren.



Figur 347 – En spændingsregulator modtager en DC spænding på indgangen og omsætter den til en DC spænding på udgangen.

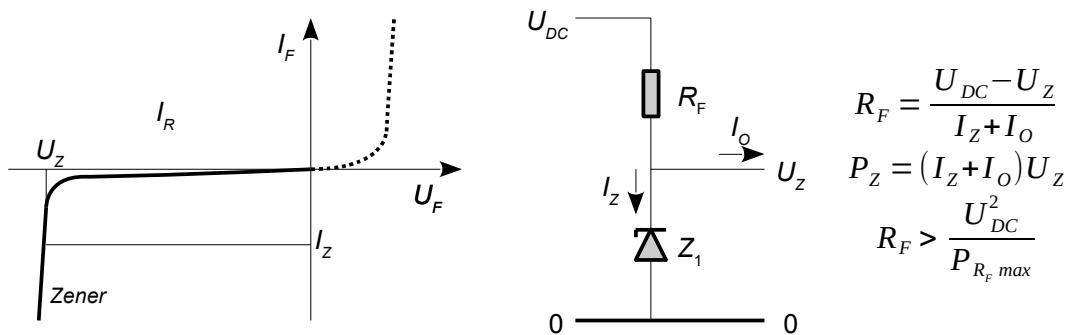
I det følgende gennemgås tre typer af spændingsregulatorer.

1. En shuntregulator har en meget ringe effektivitet da både udgangsstrøm og -spænding er lavere end indgangssidens ($U_O < U_I$ og $I_O < I_I$) så det er sjeldent at virkningsgraden kan komme ret meget over 25 %.
2. Ved en serieregulator løber strømmen så godt som direkte til belastningen ($I_O \approx I_I$), men da udgangsspændingen altid er lavere end indgangsspændingen ($U_O < U_I$) er virkningsgraden oftest under 75 %.
3. En DC-DC konverter kan både hæve og sænke udgangsspændingen og virkningsgraden kan nå over 95 %. Den er derfor populær ved omsætning af store effektorer samt hvor det er vigtigt at et batteri lever så længe som muligt, men omkostningen er et mere kompliceret design.

Den følgende omtale vil koncentrere sig om reguleringer med positiv udgangsspænding, men der er ofte brug for regulering af en negativ spænding og det følger helt den samme rute.

6.6 Shuntregulator

I forrige kapitel blev zenerdioden introduceret og det blev vist hvordan en spændingsregulering kan opbygges med en modstand og en zenerdiode. Det var tidligere en populær måde at skaffe en stabil spænding, men i dag vil man foretrække integrerede kredsløb der benyttes på samme måde¹⁵⁹. En alternativ brug af zenerdioden er at beskytte imod overspænding, og der findes specialtyper som er udviklet til at optage energi ved transient overbelastning.



Figur 348 – En zenerdiode kan anvendes til at give en konstant spænding eller som beskyttelse af en indgang der ikke kan tåle spændinger over en given værdi.

Zenerdioder anvendes med en modstand R_F der begrænser strømmen i zenerdioden til I_Z og som desuden skal leve strømmen I_O til belastningen. Den anbefalede strøm i zenerdioden kan findes i databladet og ligger ofte ved $I_Z = 5 \text{ mA}$ for smådioder mens integrerede kredsløb kan opretholde den specificerede præcision ved mindre end $100 \mu\text{A}$.

Komponenterne skal kunne klare den afsatte effekt der er 100 mW for den overflademonterede SMD modstand 0603, 500 mW for SFR16 med trådender og 250 mW for zenerdioden BZX84. Dioden skal optage hele strømmen fra modstanden når belastningen ikke trækker strøm ($I_O = 0$) og den resulterende effektafsættelse P_Z skal også her være indenfor det tilladelige. Man står sig ved at være tilbageholdende med at presse sit design for en stigning i temperaturen påvirker både udgangsspændingen og levetiden af komponenterne¹⁶⁰.

En shuntregulator vil reducere brum og støj fra effektforsyningen ved den dynamiske modstand R_Z der er defineret som spændingsvariationen over dioden som følge af en strømvariation i dioden på helt samme måde som for en diode i lederetningen. Forskellen er at den dynamiske modstand kun kan findes i databladet. Specifikationen kan bruges til at beregne hvor stabil udgangsspændingen er overfor ændringer i effektforsyningens spænding eller belastningsstrømmen.

$$\Delta U_Z = \frac{R_Z}{R_F + R_Z} \Delta U_{DC}$$

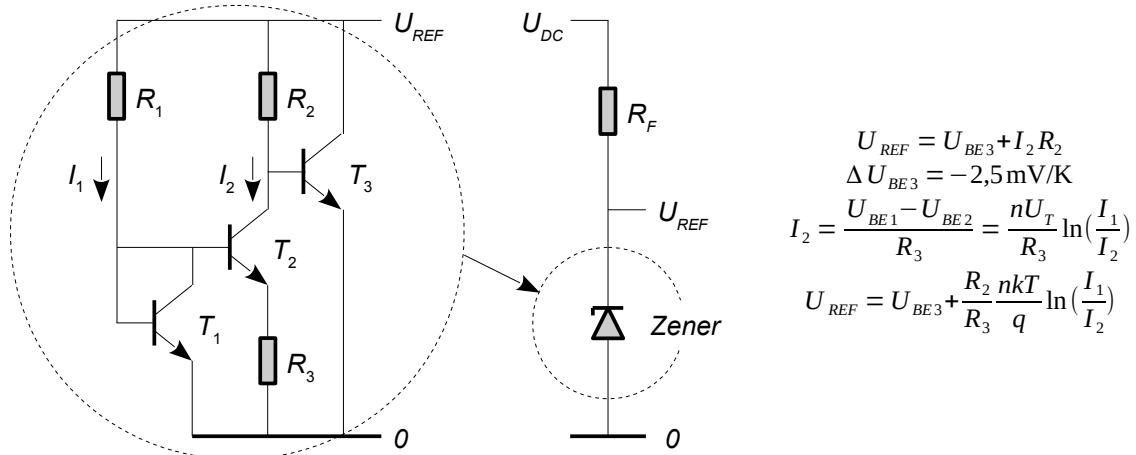
Med $R_F = 3 \text{ k}\Omega$ og $R_Z = 10 \Omega$ vil en ændring af effektforsyningens spænding med $\Delta U_{DC} = 1 \text{ V}$ give en variation af udgangsspændingen på $U_{Z RIP} = 3,3 \text{ mV}$.

159 Integrerede kredsløb som LM129, LM368, MC1400 og LT1021 kan fungere som en zenerdiode med en initial præcision fra $\pm 5 \%$ ned til $\pm 0,05 \%$, men de er ikke så hurtige som zenerdioder til at opfange en transient.

160 Man bør dog som hovedregel aldrig afsætte mere end 50 % af den maksimalt tilladelige effekt, og helst ikke over 25 % (derating), for at holde belastningen på et rimeligt niveau af både komponenten selv, men også printkort og øvrige komponenter i nærheden. Ved hård belastning vil temperaturen stige og det reducerer levetiden (søg efter "Arrhenius equation").

6.6.1 Referencespænding

Integrerede kredsløb har overtaget zenerdiodens rolle når der kræves en stor præcision og dette afsnit introducerer tankegangen bag den populære *band-gap reference*. 1971 publicerede Widlar nedenstående kredsløb der kan anvendes på samme måde som en zenerdiode.



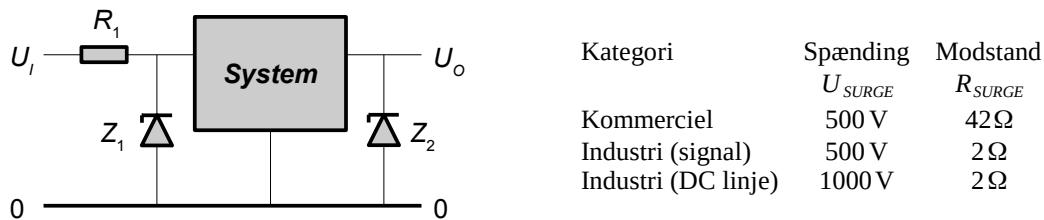
Figur 349 – Et kredsløb med nogenlunde samme funktion som en zenerdiode. Det kræver dog at transistorerne er meget ens og at de har samme temperatur.

Referencespændingen er givet ved basis-emitter spændingsfaldet over udgangstransistoren T_3 plus spændingsfaldet over modstanden R_2 i kollektor af T_2 . Hvis de to første transistorer er ens, men arbejder ved forskellige strømme i kollektor, vil spændingsforskellen over R_3 give strømmen i T_2 og ved en dekades forskel mellem strømmene ($I_1 = 10I_2$) er differensen cirka 60 mV med en positiv temperaturkoefficient på 0,33 %/K. For at opstillingen bliver uafhængig af temperaturen skal det modsvares af den negative temperaturkoefficient af transistor T_3 på cirka -2,5 mV/K. T_2 skal derfor forstærke temperaturkoefficienten på 0,2 mV/K med en faktor 12 som derved fastlægger forholdet R_2/R_3 .

En ændring af strømmen I_Z skal optages af T_3 hvilket ændrer dens basis-emitter spænding med circa 20 mV ved en fordobling og det svarer til en indre modstand på 20 Ω ved $I_Z = 1$ mA.

6.6.2 Beskyttelseskredsløb

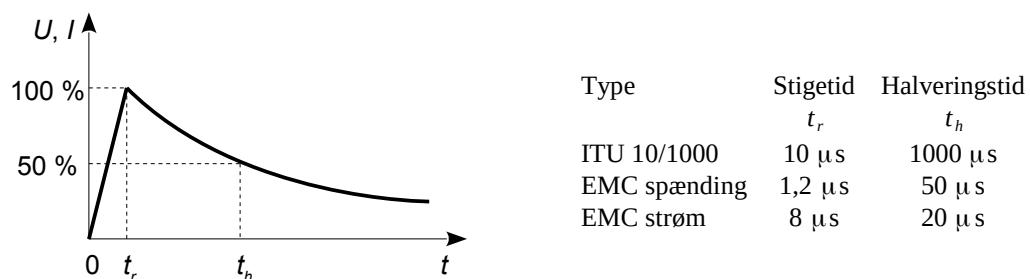
Zenerdioden er en vigtig komponent ved beskyttelse af et apparat eller et systems tilslutninger til det omgivende kredsløb mod overspænding både fra fejlbetjening og fra kortvarige transienter. Zenerdiodens spænding U_Z står nu for den højeste værdi signalet må antage uden at beskadige kredsløbet og ved højere niveau begrænses spændingen på ind til apparatet. Strømmen fra det ydre kredsløb må begrænses af en seriemarkstand eller af det ydre kredsløb selv.



Figur 350 – Beskyttelse af ind- og udgang imod transient overspænding. Strømmen i zenerdioden kan begrænses af en markstand eller alene af det ydre kredsløb (testgeneratorens indre markstand). Tabellen viser hovedtrækkene i EMC direktivets krav for *Surge Transient* og er kun vejledende.

Dioden er ikke-ledende ved tilsigtet brug og bliver kun aktiv ved en overbelastning. Zenerdiodens lækstrøm og kapaciteten i spærretreningen vil kunne påvirke den normale funktion og bør tages i betragtning ved planlægning af hvordan systemet interagerer med omgivelserne.

For at beskytte imod en langvarig overbelastning skal den afsatte effekt i markstand og diode ligge indenfor det tilladelige for komponenterne. Langvarig overbelastning ses ved sensorer for montage af elektrikere og private hvor en fejlagtig brug kan ødelægge sensorens elektronik og derfor give produktet et dårligt ry. Internt i et apparat kan man beskytte en følsom og kostbar elektronik mod et sammenbrud i effektforsyningen.



Figur 351 – Definitionen af den puls, som et beskyttelseskredsløb skal testes med. Telefonindustrien (ITU) har defineret en såkaldt 10/1000 puls der i dag benyttes til at specificere hvor hårdt dioden kan belastes. EMC direktivet har defineret en spændingspuls der benyttes ved test af radio, TV og andet elektronisk udstyr og en strømpuls for brug ved industriel elektronik.

Beskyttelse imod kortvarige transiente fik betydning efter 1996, hvor EMC direktivet trådte i kraft. Der blev her indført et krav om at visse linjer på apparater og systemer skal kunne overleve den såkaldte lyntransient (Surge Transient). Det er en kortvarig, men kraftig overspænding der skal emulere situationen når et lyn slår ned i nærheden af en forsyningslinje og gennem en magnetisk kobling giver anledning til en transient der udbreder sig gennem elforsyningen til forbrugerne. Det forventes at transienten kobler mellem parallelførte ledere i kabelbakkerne og har ledt til en række krav til signal- og forsyningelinjer for kommerciel og industriel elektronik. Uanset om transienten er en reel trussel eller ej så er det min erfaring at kravet i betydelig grad har hævet robustheden af elektronikindustriens produkter.

Der er forskellige specifikationer for linjer til effektforsyning samt analoge og digitale linjer og der er flere kategorier så jeg vil her koncentrere mig om det mere generelle i beskyttelsesmetoden.

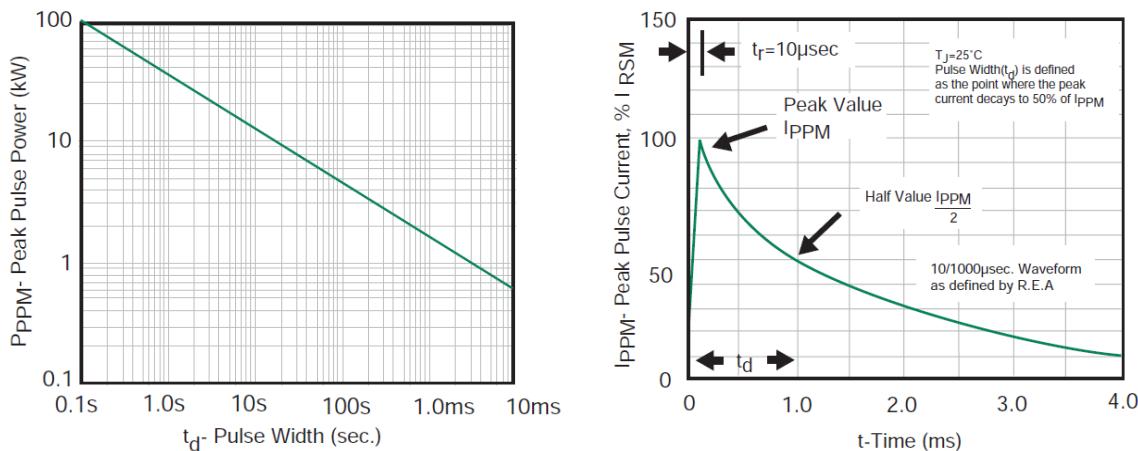
Lyntransienten er defineret som en kortvarig, ikke-repeteret puls der stiger til maksimum i løbet af kort tid og derefter klinger eksponentielt ud. Pulsen ligner meget en eksisterende transient fra telefonstandarden ITU, men forløbet er meget hurtigere og der er en specifikation for både spændings- og strømpulser, men jeg vil kun komme ind på spændingspulsen i denne bog. Lyntransienten vil blive eksemplificeret ved topværdien $U_{SURGE} = 500$ V og den indre modstand fra generatoren på $R_{SURGE} = 42 \Omega$. Det svarer til kravene for kommerciel elektronik. Ved en direkte kortslutning af generatoren er den maksimale værdi af strømmen i nærheden af 25 A så testen er ikke det rene barnemad. Ved industriel elektronik er kravet meget skrappere med $R_{SURGE} = 2 \Omega$.

Den almindelige zenerdiode kan benyttes hvis den forsynes med en tilstrækkelig stor modstand, men der er udviklet en speciel komponent for beskyttelse imod transient overlast. Den har en mere veldefineret spænding, en mindre lækstrøm end den typiske zenerdiode og den kan optage en ret så betydelig energi. Dertil kommer at den vil begå selvmord ved stor overlast idet den smelter sammen og kortslutter hvorev den beskytter det efterfølgende kredsløb selv om der så skal en reparation til for at udskifte komponenterne. Denne type zenerdiode kaldes for *TranZorb*, *TranSil* eller lignende¹⁶¹ og diodens zenerspænding kaldes for dens *clamp voltage*.

Zenerdioderne til absorbering af transient energi fås i flere effektklasser med et strømforløb som defineret ved ITU standardens 10/1000 puls og de tåler gentagne pulser hvis der er mindst 10 s mellem hver puls. Populære effektklasser er fra 150 W til 5 kW. Her vil effektklassen blive betegnet med P_{PPM} stående for *peak pulse maximum power*.

6.6.3 Dioden

Ved dimensionering af en beskyttelse imod transient overbelastning afsættes energi i zenerdioden i det korte tidsrum transienten varer. Det står i modsætning til det mere almindelige princip hvor der er balance mellem den afsatte elektriske effekt i en komponent og den afgivne termiske effekt til omgivelserne ved blandt andet varmestrålning. Ved transient brug oplagres energien kortvarigt i komponenten for derefter at diffundere væk fra det aktive område ved diodens PN overgang. En vigtig specifikation er derfor zenerdiogens evne til momentant at oplagre energi.



Figur 352 – Databladet for dioden 1.5KE der kan optage en effekt med topværdien 1,5 kW hvis strømmen aftager fra maksimum med 1 ms halveringstid (Littlefuse).

Ved EMC direktivets lyntransient, der er meget kortere end ITU standardens transient, vil energien blive afleveret i et meget kort tidsrum. Det har den konsekvens at dioden kun kan klare at optage 20 % af den energi dioden kunne klare ved ITU transienten fordi der er mindre tid til at energien kan komme væk fra det aktive område i halvlederen mens transientens energi afsættes.

¹⁶¹ Ud over zenerdioder er der mulighed for at anvende en varistor, en slags spændingsafhængig modstand, som er populær på 230 V siden af strømforsyninger, og gasudladningsrør, der kan optage en betydelig energi, men som reagerer langsommere. Disse komponenter vil ikke blive omtalt her, men de dimensioneres efter en lignende procedure.

Som en følge heraf er energien, udtrykt ved produktet af effekt og tid, ikke konstant. Det ses herover for dioden 1.5KE der kan klare en effektafsættelse på 1,5 kW ved 10/1000 pulsen. Ved EMC direktivets 20 gange kortere transient på 50 µs kan den optage 6 kW, men den afsatte energi aftager. Data er verificeret eksperimentelt ved måling på dioder med en *surge generator*¹⁶².

Strømmen i zenerdioden skal begrænses til topværdien I_{PPM} givet ved diodens clamp voltage (U_C) under den transiente overbelastning samt den effekt zenerdioden kan optage kortvarigt (P_{PPM}) og som aflæses fra figuren. For 1.5KE er strømmen 180 A for en clamp voltage på 33 V.

$$P_{PPM} = U_C I_{PPM} \Rightarrow I_{PPM} = \frac{P_{PPM}}{U_C}$$

Det giver en enkel mulighed for at beregne den nødvendige seriemodstand for begrænsning af strømmen så databladets specifikation overholdes. Modstanden består af serieforbindelsen af den indre modstand på $R_{SURGE} = 42$ fra EMC direktivets testgenerator og R_1 og herfra kan den nødvendige seriemodstand R_1 beregnes. Hvis beregningen giver en negativ værdi betyder det blot at der ikke er behov for en ekstern modstand.

$$R_1 \geq \frac{U_{SURGE} - U_C}{I_{PPM}} - R_{SURGE}$$

6.6.4 Modstanden

Ved at bruge den viste relation for I_{PPM} vil udtrykket for beregning af seriemodstanden nu bestå af EMC testgeneratorens parametre U_{SURGE} og R_{SURGE} samt zenerdiodens effektklasse P_{PPM} og dens clamp voltage U_C . Det er antaget at zenerdiodens U_C er lille i sammenligning med U_{SURGE} og fejlen derved er mindre end 5 % for $U_C < 50$ V. Når relationen giver en negativ værdi som den laveste grænse er der ikke behov for en seriemodstand og ved at sætte modstandsværdien til nul ses at en modstand ikke er nødvendig ved de mindste 150 W dioder når clamp voltage er under 12 V.

$$R_1 \geq \frac{U_{SURGE} U_C}{P_{PPM}} - R_{SURGE} \quad R_1 = 0 \quad U_C \leq \frac{R_{SURGE} P_{PPM}}{U_{SURGE}}$$

Modstanden belastes hårdt så man må bruge en modstand specielt for absorbering af transient energi. Herunder vises data for to trådkviklede modstande¹⁶³ AC01 for 1 W nominel belastning og AC10 for 10 W. Som det ses kan de begge tåle 1000 gange den nominelle effektafsættelse hvis pulsen har en varighed under 400 µs hvilket opfyldes af EMC direktivets lyntransient.

Pulsens effektive varighed findes ved at sætte arealet under en rektangulær puls med en varighed på t_p lig med arealet under den eksponentielt hendøende transient hvor opstartens korte forløb er ignoreret da det kun drejer sig om at finde en enkel relation mellem de to pulser.

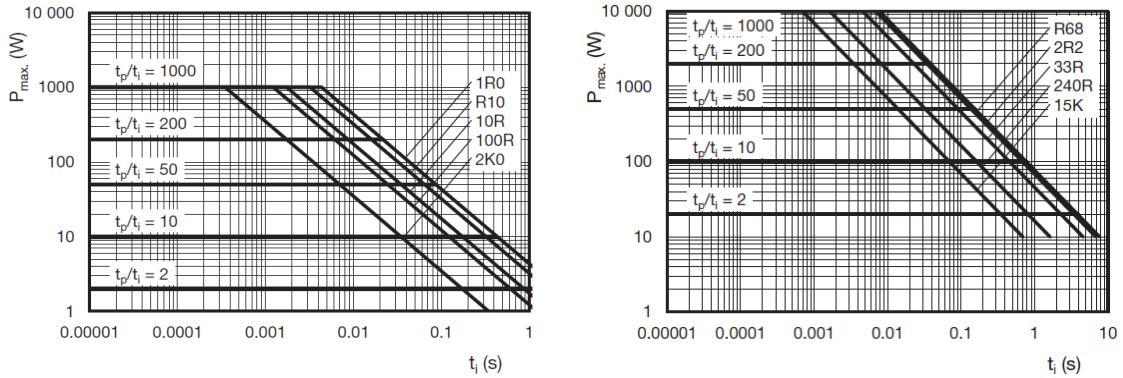
$$\int_0^{t_p} dt = \int_0^{\infty} \exp(-\ln(2) \frac{t}{t_h}) dt \Rightarrow t_p = \frac{1}{\ln(2)} t_h \Rightarrow t_p = 1,4 t_h$$

Resultatet er at en rektangulær puls skal regnes som $t_p = 70$ µs for EMC direktivets lyntransient og her kan AC01 klare 1 kW og AC10 kan klare 10 kW hvis der er tilsvarende er tusinde gange pulsens bredde til næste puls ankommer så belastningen må først gentages efter 70 ms hvilket ikke er et problem.

162 Dioderne BZD27C, BZT03C, BZW04P, P6KE og 1.5KE (150 W til 1500 W) blev udsat for EMC direktivets lyntransient ved stigende strømstyrke indtil de brød sammen. Dioderne kom fra forskellige producenter, men resultatet var homogen, og angav at dioderne ved EMC direktivets lyntransient kunne optage 22 % af energien fra ITU transienten. Publiceret i EMC Avisen nummer 15, 1995.

163 AC01 er 4,3 mm i diameter og længden er 11 mm, og AC10 har diameteren 8 mm og længden 44 mm, hvortil kommer trådenderne. Modstanden dannes af en metaltråd, der vikles omkring en keramisk kerne. Effektbelastning er opgivet ved 40°C og overfladens temperatur stiger med 100°C for AC01 og 320°C for AC10 ved fuld belastning (Vishay).

Effektkredsløb



Figur 353 – Trådviklede modstande kan optage en betydelig effekt hvis transienten er tilstrækkelig kort og for mindre end cirka 400 μ s kan de viste modstande optage 1000 gange deres nominelle specifikation. Modstandene er AC01 til venstre og AC10 til højre (Vishay).

Topeffekten i modstanden R_1 er givet ved spændingsfaldet over modstanden og hvis der ses bort fra at zenerdioden optager lidt af effekten kan modstanden dimensioneres inden der er truffet en endelig beslutning om hvilken spænding zenerdioden skal arbejde ved.

$$P_1 = \frac{(U_{SURGE} - U_C)^2}{R_1} \quad U_C \ll U_{SURGE} \quad P_1 \approx \frac{U_{SURGE}^2}{R_1}$$

Ved $U_{SURGE} = 500$ V bliver effekten $P_1 = 2,5$ kW ved $R_1 = 100 \Omega$ hvilket eksempelvis kan klares af AC10.

6.6.5 Designprocedure

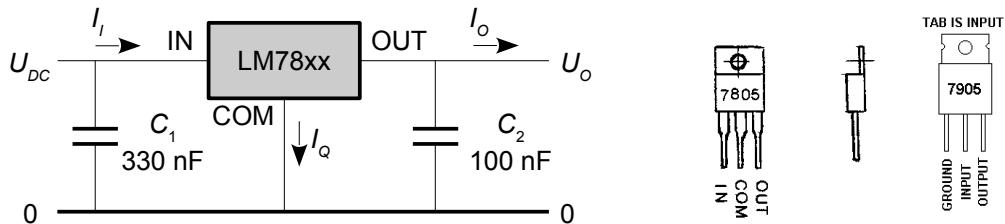
Følgende algoritme kan benyttes som vejledning for et design.

1. *Vælg effektklassen for zenerdioden (P_{PPM}) der typisk er fra 150 W til 1,5 kW. Valget er ofte betinget af pladsen optaget på printkortet og prisen på zenerdioden. Der kan være en mulig afbalancering ved at vælge en højere effektklasse for derved at undgå at inkludere en seriemodstand i designet.*
2. *Beslut den maksimale værdi den pågældende linje må antage; det er clamp voltage (U_C). Det er nødvendigt at lægge værdien noget højere end signalets niveau for at undgå en for stor påvirkning fra diodens lækstrøm.*
3. *Beregn seriemodstandens værdi fra formlen for R_1 . Hvis formlen angiver en negativ værdi skal der ikke benyttes en seriemodstand.*
4. *Beregn modstandens effektklasse fra formlen for P_1 og vælg en modstand der kan løfte opgaven.*

Et eksempel kunne være at benytte en diode der kan klare $P_{PPM} = 150$ W for at beskytte en linje beregnet for 24 V. Der vælges en zenerdiode med en clamp voltage på $U_C = 33$ V og derved kan seriemodstanden beregnes til $R_1 = 68 \Omega$. Topeffekten bliver 3,7 kW så en AC10 kan løfte opgaven med god margin. Modstanden er dog fysisk stor, meget større end en 150 W zenerdiode, og hvis der i stedet blev valgt en diode med $P_{PPM} = 400$ W vil seriemodstanden kunne undværes.

6.7 Serieregulator

Effektforbruget er højt ved shuntregulatoren og kan reduceres ved brug af en serieregulator, hvor den optagne strøm stort set er den samme som udgangsstrømmen til belastningen, og den afsatte effekt er derfor tilnærmelsesvis proportional med udgangsstrømmen. De populære regulatorer af 7800-typen kan klare op til cirka 1,5 A og findes for både positive og negative spændinger. De kan tåle kortslutning af udgangen, og hvis komponenten bliver for varm som følge af overbelastningen, så afbryder den for udgangen og køler af. Udvalget af serieregulatorer er så stort, og prisen så lav, at det simpelthen ikke kan betale sig at bygge en serieregulator selv¹⁶⁴.

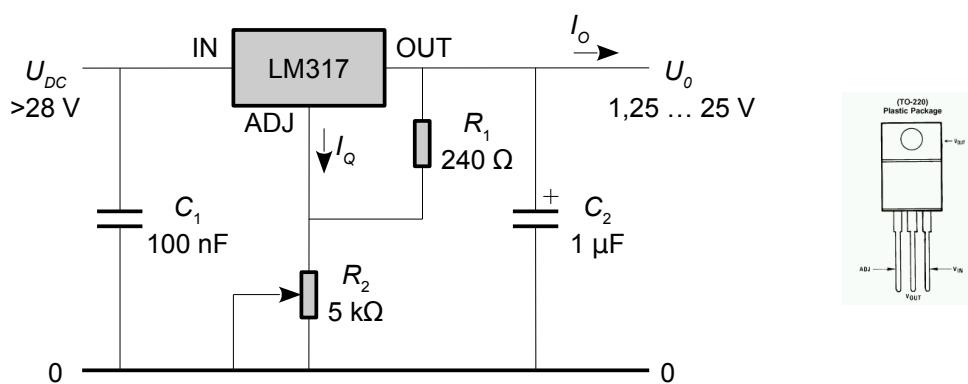


Figur 354 – Den typiske opstilling med en serieregulator nedsætter spændingen fra en ustabil effektforsyning til en stabil og kortslutningssikker forsyningslinje. De viste kondensatorer er krævet af hensyn til stabilitet og vil almindeligvis fungere, men det anbefales kraftigt at studere databladet for den valgte komponent. Bemærk at 7800 og 7900 har forskellige benspecifikation.

Serieregulatorer sikrer at udgangsspændingen ligger på en fast værdi over referencebenet, der kaldes *com* eller *ground* til højre i illustrationen. For 78-typerne er det fra 5 V og op til 24 V, med højest 35 V som indgangsspænding, dog 40 V for 7824. For 79-typerne er spændingen negativ. Tolerancen er i dag $\pm 2\%$ som standard, men der findes komponenter med $\pm 0,05\%$. Hvis COM hæves med en diode eller zenerdiode så følger udgangen med op, så 7805 og 1N4007 giver en udgangsspænding på 5,7 V. Udgangsspændingen kan ændres ved to modstande (se herunder), men strømmen I_Q kan nå 8 mA ved 78-typen, og afkoblingen bør inkludere COM for stabilitet.

$$U_o = \left(1 + \frac{R_2}{R_1}\right) U_{REF} + I_Q R_2 \quad \text{og} \quad U_{DC} > U_o + 2V$$

Serieregulatoren LM317 har 1,25 V mellem udgang og reference, og strømmen I_Q på typisk 50 μ A ($<100 \mu$ A) gør den særlig velegnet til at justere udgangsspændingen med to faste modstande. Den interne forbrugssstrøm føres til udgangen, så strømmen i OUT bør ikke komme under 10 mA.

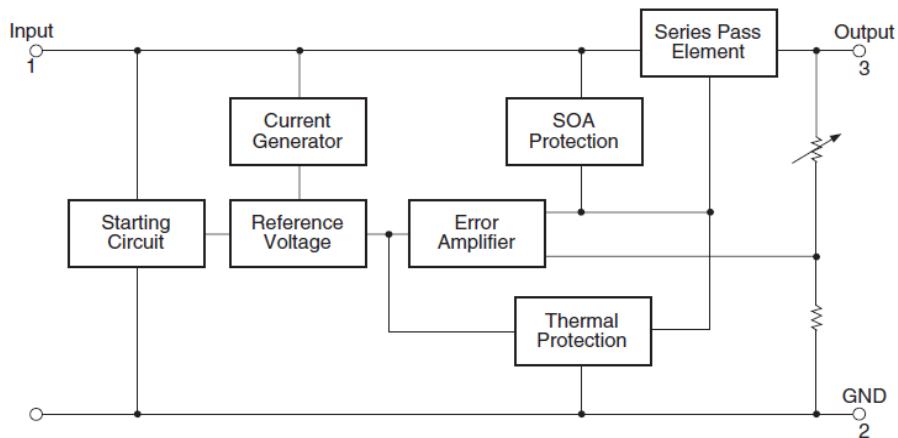


Figur 355 – Variabel udgangsspænding med LM317 ifølge en application note fra National Semiconductor. Bemærk at komponenten har en anden benforbindelse end LM7800 og LM7900.

¹⁶⁴Der findes et righoldigt udvalg af forslag på internettet, men jeg vil kraftigt anbefale at du benytter 7800-serien. Der er mange fejl og mangler ved de publicerede forslag, fx brænder flertallet af opstillinger af ved en kortslutning af udgangen.

6.7.1 En 7800-type serieregulator

En serieregulator er en integreret komponent, og som eksemplet herunder viser indeholder den de allerede nævnte blokke. Serieelementet øverst til højre er den transistor, som udgangsstrømmen til forbrugerne løber igennem. Det er en NPN transistor med emitteren til udgangen, og den styres af en forstærker, der sammenligner en neddelt version af udgangsspændingen med en indbygget reference, der ofte er på 1,25 V, men ved 7800 er den opbygget noget anderledes. Blokken "SOA Protection" sikrer at NPN transistoren ikke kommer udenfor det sikre strøm- og spændingsområde (*Safe Operating Area*), og den begrænser udgangsstrømmen hvis transistorens specifikation bliver overskredet. Den afsatte effekt varmer hele chippen op, og ved en temperatur over 150°C vil "Thermal Protection" ganske enkelt slukke for udgangsstrømmen for at beskytte chippen.



Figur 356 – Den interne opbygning af en 78-type serieregulator (National Semiconductor).

Herunder vises et detaljeret diagram over en 78-type serieregulator fra National Semiconductor. En sammenligning med andre kilder viser væsentlige forskelle i diagrammet, så der kan være forskelle mellem den følgende gennemgang af kredsløbet og til en aktuelt indkøbt komponent.

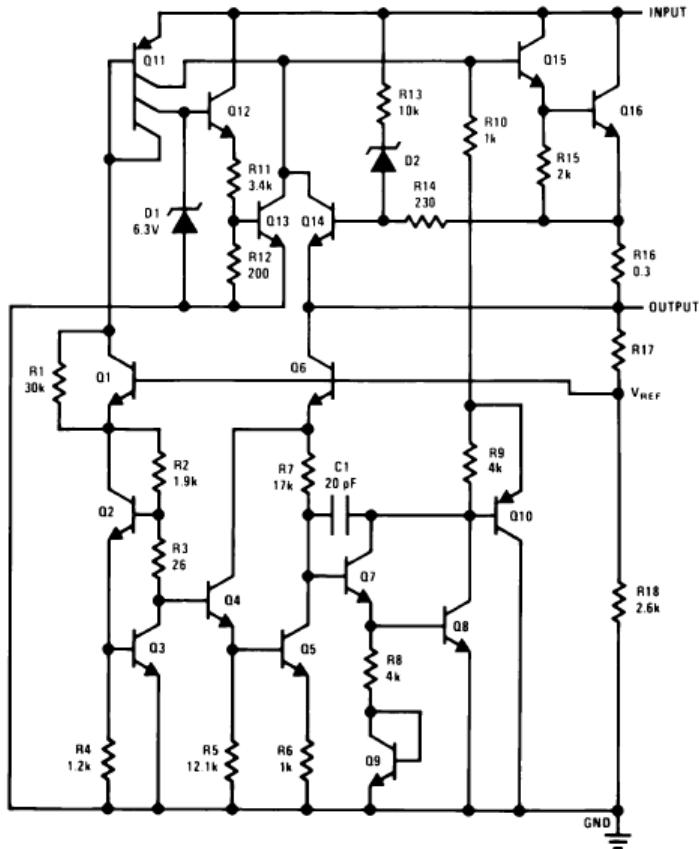
Serieelementet Q16 skal bære en ret stor strøm, så den har Q15 som strømforstærker, og basis af Q15 ligger cirka 1,4 V over den indstillede udgangsspænding, hvilket er en væsentlig grund til at databladet foreskriver at der skal ligge mindst 2 V fra ind- til udgang.

Strømmen i Q16 giver et spændingsfald over R16, der med $0,3 \Omega$ er 0,6 V ved 2 A. Hvis strømmen stiger over den værdi bliver Q14 aktiv og reducerer basisstrømmen til Q15, så udgangsstrømmen ikke stiger yderligere. Komponenterne D2, R13 og R14 hæver spændingen på basis af Q14 ved stor indgangsspænding, så strømbegrænsningen sætter ind ved en lavere strøm. Det udgør SOA beskyttelsen for Q16, fx hvis udgangen kortsluttes ved en høj værdi af indgangsspændingen.

Q11 er en strømgenerator (strømspejl) med to udgange, og den drives af referencekredsløbet med cirka 2 mA. Q11 giver strøm til basis af Q15 og zenerdioden D1 på 6,3 V, som neddeles af Q12, R11 og R12 til 0,3 V, der driver basis af Q13. Transistoren kræver 0,7 V for at være aktiv, så den er afbrudt i normalsituationen, men dens temperaturkoefficient på $-2,5 \text{ mV}/\text{°C}$ gør, at den træder i aktion ved cirka 150°C, hvor den tager basisstrømmen fra udgangstransistoren og forhindrer yderligere opvarmning af chippen.

Udgangsspændingen neddeles af R17 og R18 og føres til fejlforsærkeren, der er en inverterende forstærker bestående af Q7 til Q10. Indgangen er ved basis af Q7 og udgangen er emitter af Q10. Der er en dominerende pol givet ved C1, så man kan også sige at fejlforsærkeren integrerer den strøm, der måtte løbe til indgangen. Basis af Q7 ligger på 1,2 V så den neddelte spænding driver en strøm gennem Q6 og R7, og Q5 trækker en konstant strøm ud af indgangen. Hele regulatoren er derfor i balance når strømmen i R7 matcher strømmen i Q5.

Effektkredsløb



Figur 357 – Detaljeret kig ind i en 78-type serieregulator (National Semiconductor).

Q2 og Q3 danner en dobbelt diode, idet basis af Q2 ligger to spændingsfald over GND. Q3 holder strømmen i Q2 fast på $500 \mu\text{A}$ og strømmen i Q3 er omtrent det dobbelte (se senere). Spændingen fra den dobbelte diode reduceres med 20 mV af R3 og driver Q4 og Q5. De arbejder begge ved en strøm på en tiendedel af strømmen i henholdsvis Q2 og Q3, så basis-emitter spændingsfaldet for Q4 og Q5 er 60 mV lavere end ved Q2 og Q3. Følgelig ligger der 100 mV over R6 og strømmen i Q5 bliver $100 \mu\text{A}$. Det kan nu konkluderes at spændingsfaldet over R7 bliver $1,7 \text{ V}$ og derfor vil referencespændingen blive på $3,5 \text{ V}$ på grund af basis-emitter spændingsfaldet ved Q6. Strømmen i Q3 kan nu beregnes til $850 \mu\text{A}$ og spændingsfaldet over R3 bliver de anførte 20 mV .

Temperaturkoefficienten for en diode er $-2,5 \text{ mV}/^\circ\text{C}$, og for en fastholdt udgangsspænding kræves det at spændingsfaldet over R7 stiger med $7,5 \text{ mV}/^\circ\text{C}$ på grund af transistorerne Q6, Q7 og Q8. Strømmen fra Q5 skal stige $0,4 \mu\text{A}/^\circ\text{C}$, så spændingen over R6 skal stige $0,4 \text{ mV}/^\circ\text{C}$, og det bliver realiseret ved at Q4 og Q5 arbejder ved en væsentligt lavere strøm end Q2 og Q3.

R1 er nødvendig for at starte kredsløbet. Uden R1 vil Q1 ikke trække strøm, fordi udgangen ligger på nul og derfor ikke starter de tilhørende kredsløb op. Derved ville Q11 ikke få den basisstrøm, der er nødvendig for at åbne for Q16. Strømmen i modstanden varierer med indgangsspændingen fra cirka $0,1 \text{ mA}$ ved 7 V indgangsspænding til omkring 1 mA ved 35 V . Strømmen i Q2 og Q3 er dog tilnærmelsesvis konstant da spændingen ved emitter af Q1 stiger en lille smule ved høj indgangsspænding så strømmen fra Q1 derfor falder tilsvarende. Når R1 ikke er ført til nul, men til kollektor af Q2 og Q3, kan årsagen være, at variationen i spændingsfaldet over R3 kan modvirke en ændring af udgangsspændingen og derved lette arbejdet for fejlforsærkeren; det er dog kun spekulation.

Shuntregulatoren har et konstant effektforbrug på $U_{DC}I_O$ uanset om der trækkes strøm ud af den eller ej, og det gør den kun egnet til simple småopstillinger. Serieregulatoren har et lavt internt for-

Effektkredsløb

brug på 8 mA for 78-typerne og mindre for LM317, så effekttabet er stort set proportionelt med udgangsstrømmen I_o .

$$P_D = U_{DC} I_Q + (U_{DC} - U_o) I_o$$

For LM7805 med $U_o = 5$ V og $I_Q < 8$ mA er effekttabet $P_D = 2,2$ W ved indgangsspændingen $U_{DC} = 25$ V og en udgangsstrøm på $I_o = 100$ mA. Af effektafsættelsen kommer de 0,2 W fra det faste effekttab.

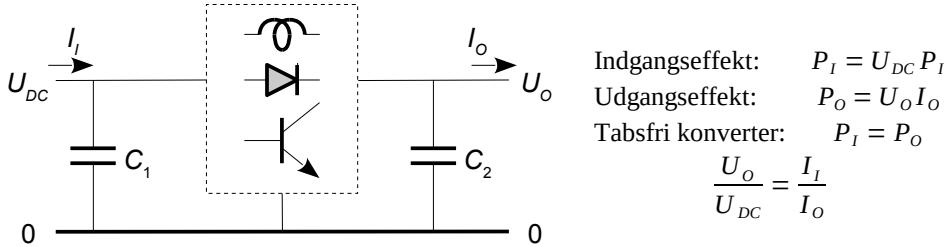
Opvarmningen af komponenten består af to bidrag, hvor det ene er det faste tab fra strømmen I_Q i COM for 78-typerne eller ADJ for LM317-typen, og det andet er det variable tab fra strømmen til udgangen, som løber igennem spændingsfaldet fra indgangsspændingen U_{DC} til udgangens U_o .

Note 18.

6.8 DC-DC konverter

En serieregulator taber en del effekt når der er stor forskel mellem effektforsyningens spænding og den ønskede udgangsspænding og her er en DC-DC konverter (*switch-mode converter*) attraktiv da virkningsgraden er høj. Prisen er et mere kompliceret kredsløb og elektromagnetisk støj.

Kort sagt benytter en DC-DC konverter muligheden for temporært at oplagre energi i en spoles magnetfelt for kort efter at sende energien videre til opplagring i en kondensators elektriske felt. Frem for de lineære kredsløb i en serieregulator er diode og transistor nu kontakter som styrer energitransporten.



Figur 358 – DC konverteren benytter en spole til at optage energi fra indgangen i et kort tidsrum hvorefter energien overføres til udgangen i det efterfølgende tidsrum. Den proces gentages i en evindelighed styret af en transistor og en diode der arbejder som kontakter. Transistoren kan være bipolær eller felteffekt. Kondensatorerne mindsker strømstødene på ind- og udgang.

En DC-DC konverter anvendes ved batteriforsynet udstyr hvor batteriet skal holde længe som fx headsets og høreapparater samt ved autonomt udstyr med radiokommunikation. Den er desuden vigtig ved omsætning af store effekter som ved effektforsyning af computerchippen i en PC.

Virkningsgraden kan komme over 95 % hvilket er umuligt ved serieregulatoren, konverteren kan både designes for hæve eller sænke spændingen, den kan udvides til at skabe isolation mellem ind- og udgang som er anvendt i udstyr tilsluttet bynettet og den kan konstrueres for at leve flere forskellige udgangsspændinger.

Ved at benytte en forholdsvis høj frekvens ved omskiftningen kan der benytte små værdier af både spole og kondensator så en DC-DC konverter optager mindre plads end en nettransformator med samme effekt. Desværre betyder den høje frekvens også at der udstråles elektromagnetisk støj så der er ofte et behov for at isolere konverteren fra resten af opstillingen fx ved at indbygge den i et omsluttende metalkabinet og ved at kræve filtre for at dæmpe ledningsbåret støj.

De følgende to grundprincipper vil blive omtalt i det følgende.

Opkonvertering (*boost converter*) benyttes eksempelvis ved effektforsyning fra en enkelt battericelle¹⁶⁵ hvor batteriets 1,5 V omsættes til eksempelvis 3,3 V for en mikrocomputer. En bilradio kan tilsvarende omsætte 12 V fra et blybatteri til en højere spænding for brug ved bilradioens effektforstærkere.

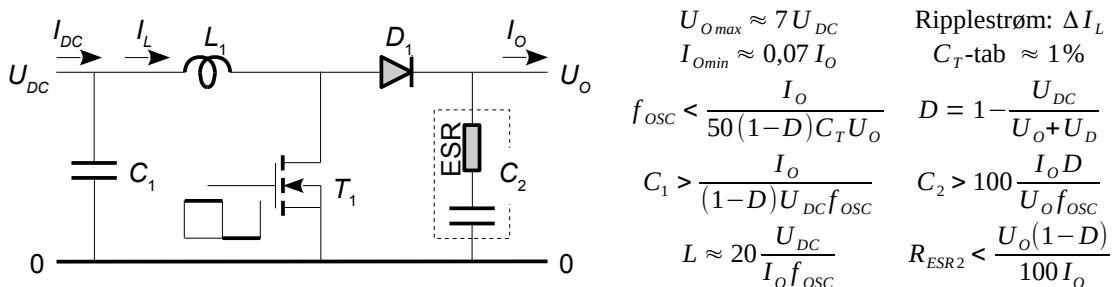
Nedkonvertering (*buck converter*) benyttes i headsets og høreapparater hvor levetiden af et Li-Ion batteri skal være så lang som muligt samt i en PC hvor effektforsyningens 12 V lokalt omsættes til processorens arbejdsspænding på 1,7 V da en strømstyrke på 50 A ikke er velegnet til intern distribution gennem kabler, stik og printbaner.

Der er ikke den store forskel på omsætning mellem to forskellige DC niveauer og så omsætning mellem DC og AC, som ved de undersøiske kabler, eller omsætning mellem to AC niveauer med hver sin frekvens som benyttes i elforsyningen, men det ligger udenfor rammerne af denne bog hvor DC-DC konverteren blot skal introduceres.

¹⁶⁵Cypress PSOC komponenter kan startes op fra en enkelt celle på blot 0,9 V ved en indbygget DC-DC konverter, og konverteren kan følge ned til omkring 0,5 V før systemet lukkes ned.

6.9 Opkonvertering

Den følgende opstilling danner en klassisk DC-DC konverter med højere udgangsspænding end indgangsspændingen (*boost converter*). Hvis transistoren er OFF vil spolen lede strøm gennem dioden til udgangen så udgangsspændingen trækkes op til forsyningsspændingen minus diodens spændingsfald inden konverteren er startet. Med transistoren ON er spolen forbundet direkte over effektforsyningen U_{DC} så strømmen i spolen vokser lineært i tidsrummet t_{ON} . Med transistoren OFF vil strømmen fra spolen løbe gennem dioden til udgangen hvorved kondensatoren oplades over tidsrummet t_{OFF} og samtidigt aftager strømmen i spolen til startværdien hvor transistoren går ON. Strømmen i spolen svinger med ΔI_L omkring middelværdien I_L og kondensator C_1 udjævner den oscillerende strøm så indgangsstrømmen I_{DC} bliver tilnærmedesvis konstant.



Figur 359 – Udgangsspændingen U_o hæves over indgangsspændingen U_{DC} givet ved parameter D .

Svingningen i strøm og spænding omkring den nominelle værdi (*ripple*) er ΔI_L . Oscillatorens frekvens f_{osc} vælges af designeren fra en maksimalt tilladelig værdi. Dioden bør være af typen Shottky for at holde spændingstabten nede. Udgangens kondensator C_2 skal være af en god type med lav ESR. Kondensator C_T repræsenterer den samlede kapacitet af transistor, diode og spole. Udgangsspændingen kan ikke hæves over en grænse sat af komponenternes parametre og udgangens strøm må ikke komme under et minimum for at undgå en stigning i den indstillede spænding.

Omskiftningen mellem ON og OFF styres af en oscillator der typisk svinger på mindst 100 kHz for at holde den fysiske størrelse nede på spolen og kondensatorerne. Designet af oscillatoren er ikke omhandlet for det er lettest at benytte en PWM udgang fra en mikroprocessor der så også kan stå for en regulering af konverteren som korrektion for en ændring i belastningen af udgangen.

Udgangens spændingsværdi indstilles ved den tid transistoren er ON i relation til perioden af den frekvens f_{osc} oscillatoren arbejder på. Det kaldes for konverterens *duty cycle* D .

$$D = \frac{t_{ON}}{t_{ON} + t_{OFF}} = t_{ON} f_{osc}$$

For at designe en DC-DC konverter skal der træffes en række valg og herunder gives et eksempel på hvordan et design kan gribes an. De benyttede formler udledes side og som det vil ses af udledningen er der meget stor frihed med hensyn til valget af komponentværdier, men der er også nogle begrænsninger, hvor den væsentligste nok er at en DC-DC konverter stiller krav om at udgangsstrømmen ikke kommer under en givet værdi.

Opkonvertering fra $U_{DC} = 3,3$ V til $U_o = 5$ V ved en strøm på $I_o = 100$ mA.

Mindste belastningsstrøm bliver 7 mA, så udgangen bør ikke være ubelastet. Dioden vælges til en Shottky med et spændingsfald på $U_D = 0,3$ V og oscillatoren skal give $D = 0,38$. Transistoren antages at have en indre kapacitet på $C_T = 1$ nF så den maksimale frekvens er 640 kHz og der vælges $f_{osc} = 100$ kHz. De to kondensatorer bliver $C_1 > 490$ nF og $C_2 > 7,6$ μ F, hvor der vælges to ens kondensatorer på 10 μ F. Den ækvivalente seriemarkant skal være $R_{ESR} < 0,3 \Omega$ for C_2 , og det bliver let at opfylde. Spolen findes til $L = 6,6$ mH, som afrundes til 10 mH.

6.9.1 Analyse

Når en DC-DC konverter arbejder skifter den mellem to faser hvor den enten er ON eller OFF og det reflekteres i analysen der opdeles tilsvarende. Det forudsættes at der løber en konstant strøm til belastningen hvilket kaldes for den stationære tilstand (*stationary condition*). Dermed ses der bort fra situationen lige efter at DC-DC konverteren er startet. Opstartsfasen kan være vigtig, men den er uden betydning når den stationære tilstand først er nået. Med transistoren ON vokser strømmen i spolen lineært med ΔI_{ON} som er givet af værdien af effektforsyningens spænding U_{DC} , spolens selvinduktion L_1 og den tid t_{ON} som transistoren har været ON. Det er uden betydning hvor stor strømmen er i spolen ved start af analysen for den beskriver kun strømmens ændring. ON-fasen slutter med en strøm I_2 i spolen som danner startbetingelsen for den efterfølgende fase hvor transistoren er OFF. Her aftager strømmen med et bidrag der tilsvarende kaldes for ΔI_{OFF} og ender ved startværdien I_1 for ON-fasen. Herefter vil hele forløbet blot gentages i det uendelige.

Transistor ON. Transistoren forbinder spolens højre side til nul så dioden spærre og blokerer for kredsløbet til højre der derfor skal analyseres separat. Kirchhoffs maskelov benyttes for effektforsyningen U_{DC} og spændingen over spolen u_L og der vælges positiv omløbsretning med uret. Spændingen over spolen udtrykkes ved definitionsligningen $u_L = L_1 \frac{di_L}{dt}$.

$$-U_{DC} + u_L = 0 \Rightarrow L_1 \frac{di_L}{dt} = U_{DC} \Rightarrow di_L = \frac{U_{DC}}{L_1} dt$$

Strømmen i spolen vokser fra en eller anden startværdi I_1 til slutværdien I_2 hvor transistoren går OFF og værdien bestemmes ved integration over tidsrummet t_{ON} .

$$i_L(t_{ON}) = \int_0^{t_{ON}} \frac{U_{DC}}{L_1} dt + I_1 \Rightarrow I_2 = i_L(t_{ON}) = \frac{U_{DC}}{L_1} t_{ON} + I_1$$

Slutværdien er derfor givet ved startværdien plus et bidrag ΔI_{ON} der er givet ved den tid t_{ON} hvor transistoren har været ON samt parametrene U_{DC} og L .

$$\Delta I_{ON} = \frac{U_{DC}}{L} t_{ON}$$

Transistor OFF. Transistoren er nu afbrudt så strømmen i spolen ledes gennem dioden D_1 til udgangen hvor den lader kondensatoren op. For at simplificere beregningen antages det at udgangens spænding ikke ændres hvilket kan tillades da ændringen skal være lille. Kirchhoffs maskelov inkluderer effektforsyningen U_{DC} , spolen L_1 , dioden D_1 og udgangsspændingen U_o . Dioden beskrives her ved et konstant spændingsfald på U_D på 0,7 V for siliciumdioder og 0,3 V for Shottky dioder. Analysen foregår på samme måde som før ved at indsætte definitionen for u_L og gennemføre en integration over tidsrummet t_{OFF} hvor integrationskonstanten kaldes I_2 da den netop danner startbetingelsen for OFF-fasen.

$$-U_{DC} + u_L + U_D + U_o = 0 \Rightarrow L_1 \frac{di_L}{dt} = U_{DC} - U_D - U_o \Rightarrow i_L(t_{OFF}) = -\frac{U_o + U_D - U_{DC}}{L} t_{OFF} + I_2$$

Strømmen i spolen aftage lineært fra startværdien I_2 til slutværdien $i_L(t_{OFF}) = I_1$ der danner startværdien for den efterfølgende fase hvor transistoren er ON.

$$\Delta I_{OFF} = -\frac{U_o + U_D - U_{DC}}{L} t_{OFF}$$

Med forudsætningen om at der løber en konstant strøm I_0 til belastningen vil strømmen i spolen svinge op og ned omkring en fast værdi som er middelværdien af strømmen i spolen. Den samlede ændring i spolens strøm er nul for en hel periode hvor transistoren har været både ON og OFF for ellers vil middelværdien af strømmen i spolen ændres.

$$\Delta I_{ON} + \Delta I_{OFF} = 0$$

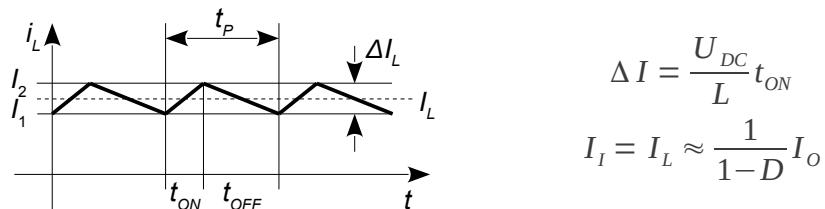
Resultatet er at udgangsspændingen U_o kan udtrykkes som funktion af indgangsspændingen U_{DC} og parameteren D . Bemærk at hverken spolens selvinduktion, de to kondensatorer eller den frekvens som oscillatoren arbejde på indgår i ligningen. Det betyder at man har stor frihed til at vælge disse parametre ud fra andre designrelaterede krav.

$$U_o = \frac{U_{DC}}{1-D} - U_D \quad \text{hvor} \quad D = \frac{t_{ON}}{t_{ON}+t_{OFF}} = t_{ON} f_{osc}$$

Hvis transistoren er ON i halvdelen af tiden og OFF i den resterende tid bliver $D = 0,5$ og det medfører en fordobling af udgangsspændingen på nær diodens spændingsfald. Blandt andet af den årsag benyttes ofte en Shottky diode hvor spændingsfaldet kun er af størrelsesordenen 0,3 V.

Formlen antyder at udgangsspændingen vokser over alle grænser for $D = 1$, men tab i opstillingen vil modvirke dette. Udgangsspændingen kan dog blive høj nok til at beskadige et tilsluttet kredsløb eller dioden D_1 hvis den maksimale værdi af spændingen over den overskrides. Formlen viser ikke at der skal aftages en mindste værdi af udgangsstrømmen. Hvis det ikke er tilfældet bliver spolen ved med at pumpe ladning i kondensator C_2 indtil "et eller andet" bliver ledende og begrænser opladningen. Der er følgelig en minimum belastningsstrøm $I_{o\ min}$ med mindre D styres ved at måle udgangsspændingen. Det er det normale, men det kredsløb skal ikke beskrives her.

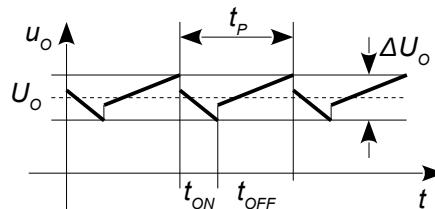
Det resulterende forløb i den stationære tilstand vises herunder for tre perioder ud af et forløb med et meget stort antal ens perioder. Den stiplede linje står for middelværdien af strømmen i spolen I_L og er lig med den strøm der optages fra effektforsyningen U_{DC} .



Figur 360 – Strømmen i spolen svinger i den stationære tilstand omkring en middelværdi I_L med en spids-spids værdi ΔI_l der er styret af hvor lang tid transistoren er ON og OFF.

Ignoreres tabene i opstillingen vil effekten optaget fra indgangen være lig med effekten til belastningen og det giver den viste relation for indgangens strømværdi.

Ideelt skal DC-DC konverteren afgive en fast DC værdi, men den er overlejret en støj bestående af en kombination af det trekantede forløb fra op- og afladning af kondensatoren og et firkantet forløb fra skiftet i strømretning på grund af kondensatorens ækvivalente seriemedstand ESR (se senere).

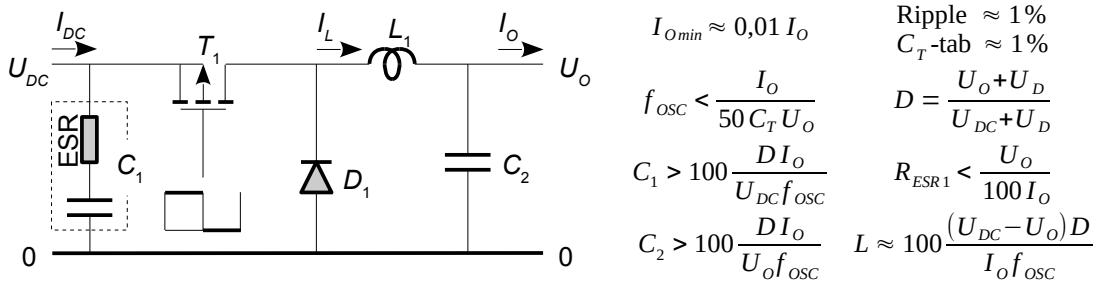


Figur 361 – Den resulterende spændingsprofil på udgangen af konverteren består af lineære segmenter fra op- og afladning af kondensatoren, samt spændingsspring fra den skiftende strømretning i kondensatorens ækvivalente modstand.

En DC-DC konverter kræver brug af gode kondensatorer for de belastes af en pulserende strøm. Ved opkonvertering er belastningen især hård for C_2 som skal have en lav værdi af ESR for at holde effektafsættelsen lav. Som dielektrikum benyttes altid en aluminium elektrolyt mens tantal bør undgås da de ikke kan klare den pulserende strøm.

6.10 Nedkonvertering

Omrokering af spole, transistor og diode skaber en DC-DC konverter hvor udgangsspændingen er lavere end indgangsspændingen (*buck converter*). Transistoren vil ofte være en P kanal MOSFET på grund af gode egenskaber som kontakt, mens IGBT er populær ved meget høje effekter.

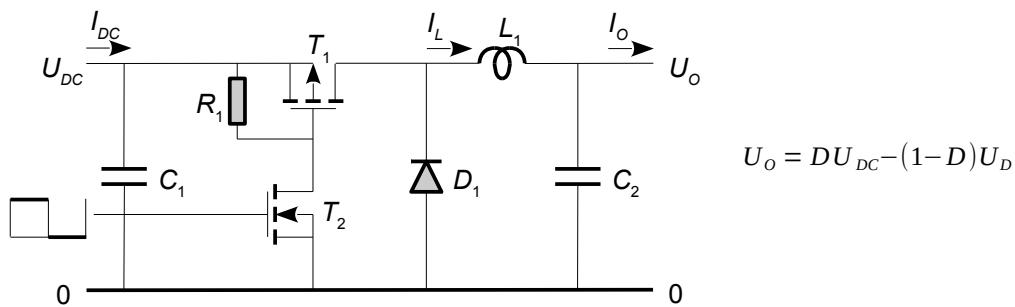


Figur 362 – Udgangsspændingen U_o sænkes til D gange indgangsspændingen U_{DC} og spolens strøm er lig med belastningsstrømmen.

Swingningen i strøm og spænding omkring den nominelle værdi (*ripple*) er fastsat til 1 %. Oscillatorens frekvens f_{osc} vælges af designeren fra en maksimalt tilladelig værdi. Dioden bør være af typen Shottky for at holde spændingstabet nede, og indgangens kondensator skal være af en god type. Den første formel angiver at konverteren ikke kan leve mindre end 1 % af den nominelle udgangsstrøm. Kondensator C_T repræsenterer den samlede kapacitet af transistor, diode og spole. Kondensator C_1 dæmper den støj, der optræder på indgangssiden, mens C_2 blot reducerer ripplen på udgangen.

Ved opstart af kredsløbet inden PWM oscillatoren er indstillet er det vigtigt at gate er på samme niveau som source for ellers vil udgangsspændingen trækkes op til indgangsspændingen.

Source er på den positive værdi af effektforsyningen og gate skal trækkes helt op til U_{DC} for at få en MOSFET til at afbryde. Hvis konverteren benytter samme effektforsyning som en mikroprocessor kan den drives direkte fra PWM udgangen ellers må der inkluderes et kredsløb for at drive gate korrekt og et eksempel vises herunder hvor T_2 driver gate på T_1 .



Figur 363 – Udgangsspændingen U_o sænkes til D gange indgangsspændingen U_{DC} og spolens strøm er lig med belastningsstrømmen.

Kredsløbet ser enkelt ud, men styringen af gate på T_1 er ikke optimal for der er ret betydende indre kapaciteter i en MOSFET. De udgør ganske vist ikke noget problem når T_2 går ON for den kan let oplade disse kapaciteter, men når T_2 går OFF skal kapaciteterne aflades gennem R_1 og det kan tage for lang tid. Med en modstand på 1 kΩ og 1 nF fra de indre kapaciteter bliver tidskonstanten på 1 µs og T_1 afbryder tilsvarende langsomt. Det vil være et tab når transistoren både leder en strøm og har et stort spændingsfald over sig. Den afsatte energi kan fordeles over hele perioden som en middeleffekt og den vil udgøre et effekttab der kun kan reduceres ved at vælge en lav frekvens. En løsning er at inkludere et klasse B trin med to bipolare transistorer (NPN og PNP) mellem drain for T_2 og gate for T_1 eller at benytte en special IC for dette formål. Det vil også reducere tabet fra effekten afsat i R_1 når T_2 er ON.

6.10.1 Analyse

Forløbet er det samme som før med kredsløbsligninger for hver af faserne. Med transistoren ON spærre dioden og spolen lades op ved spændingsforskellen mellem ind- og udgangene. Med transistoren OFF skifter spændingen over spolen fortegn da magnetfeltet nu skal reduceres så spolens venstre side svinger nedad indtil dioden bringes i ledende retning og spændingsfaldet over spolen er nu lig med udgangsspændingen på nær dioden.

Der benyttes positiv retning med uret og maskerne inkluderer indgangsspændingen U_{DC} med negativt fortegn da retningen er modsat positiv retning, spolens spænding er Ldi_L/dt og på udgangen er spændingen U_o .

$$\begin{aligned} \text{ON: } -U_{DC} + L \frac{di_L}{dt} + U_o &= 0 & \Rightarrow \Delta I_{ON} = \frac{U_{DC} - U_o}{L} t_{ON} \\ \text{OFF: } U_D + L \frac{di_L}{dt} + U_o &= 0 & \Delta I_{OFF} = -\frac{U_D + U_o}{L} t_{OFF} \end{aligned}$$

For den stationære tilstand gælder krav om balance for strømændringerne med transistoren ON og OFF. Efter indsættelse og reduktion ses det at spændingsfaldet dioden betyder lidt mindre for denne type konverter som følge af faktoren $1 - D$. Konverterens udgangsspænding er nu proportionel med parameteren D så spændingen kan reguleres lineært fra 0 til U_{DC} uden risiko for at spændingen pludseligt stiger.

$$U_o = D U_{DC} - (1 - D) U_D$$

Formlen siger dog intet om hvordan konverteren opfører sig umiddelbart efter start. Hvis T_1 er aktiv efter opstart vil det svare til $D = 1$ hvor udgangsspændingen hurtigt bliver maksimal.

6.11 Dimensionering

Analysen af forløbets to faser gav nogle relationer som kan bringes til at give et bud på frekvensen af oscillatoren, spolens størrelse og kondensatorernes værdier. Det er nødvendigt at beslutte en værdi for variationen ΔI_L af strømmen i spolen for den har stor betydning for DC-DC konverterens funktion. Hvis værdien af ΔI_L er stor kan spolens selvinduktion reduceres, men udgangsstrømmens mindst tilladelige værdi bliver til gengæld høj så belastningen kan ikke variere helt så frit.

6.11.1 Oscillatorens frekvens

Skiftet mellem ON og OFF kræver at spændingen over transistoren skifter mellem nul og en høj spændingsværdi. Derved op- og aflades kapaciteterne ved transistorens drain så der gemmes en ladning som senere smides væk og det udgør et af DC-DC konverterens mange tabsmekanismer. Her er inkluderet den parasitiske kapacitet ved spolen, diodens kapacitet fra dens spærrelag og desuden de indre kapaciteter for drain-gate og drain-source for en MOSFET og tilsvarende for en BJT. Det hele adderes til C_T som dog ikke behøver at være kendt med en større præcision inden designet starter for den skal kun benyttes til at angive en højeste værdi af oscillatorens frekvens. Værdien er typisk i området 1 ... 10 nF, men man kan altid vende tilbage og gentage beregningen når første bud på DC-DC konverteren er kendt og der er et godt bud på komponenterne.

Type	Kapacitet	Oplagret energi:	Gennemsnitligt effekttab over en periode:
MOSFET	$C_T = C_{DG} + C_{DS} + C_{Diode} + C_{Spole}$	$W_C = \frac{1}{2} C_T U_X^2$	$P_C = \frac{W_C}{t_P} = \frac{1}{2} C_T U_X^2 f_{osc}$ hvor $\frac{U_{X_{op}}}{U_{X_{ned}}} = \frac{U_O}{U_{DC}}$
Bipolær	$C_T = C_{CB} + C_{CE} + C_{Diode} + C_{Spole}$		

Energien W_C der tilføres kondensatoren under opladning fra nul til spændingen U_X er givet ved en formel fra elektrostatikken. For at benytte den som udtryk for et effekttab omregnes energien til en middeleffekt over en periode ved at dividere med den tid en periode varer ($t_P = 1/f_{osc}$) idet effekt er energi per tid. Dernæst kan man så beslutte hvor stor tabseffekten må være. Her er grænsen valgt til 1 % af den effekt som DC-DC konverteren skal aflevere.

Ved at indføre kravet på højest 1 % af $U_O I_O$ som tabt effekt kan den højest tilladelige frekvens f_{osc} bestemmes.

$$P_C < \frac{U_O I_O}{100} \Rightarrow \frac{1}{2} C_T U_X^2 f_{osc} < \frac{U_O I_O}{100} \Rightarrow f_{osc} < \frac{U_O I_O}{50 C_T U_X^2} \text{ hvor } \frac{U_{X_{op}}}{U_{X_{ned}}} = \frac{U_O}{U_{DC}}$$

For en afgiven effekt til belastningen på 1 W ved $U_X = 5$ V og med $C_T = 1$ nF bliver $f_{osc} < 400$ kHz.

Ved nedkonvertering er $U_X = U_O$ så U_X erstattes blot med U_O og det tillader U_O i tælleren at blive forkortet væk. Derimod vil en opkonvertering skulle bruge $U_X = U_{DC}$ og da $U_O = U_{DC}/(1 - D)$ bliver relationen $U_X = (1 - D)U_O$.

Type	Oscillatofrekvens
Op	$f_{osc} < \frac{I_O}{50(1-D)^2 C_T U_O}$
Ned	$f_{osc} < \frac{I_O}{50 C_T U_O}$

For op-konvertering fra $U_{DC} = 5$ V til $U_O = 15$ V bliver $D = 0,68$. Ved $I_O = 100$ mA og med $C_T = 1$ nF bliver den højest tilladelige værdi af f_{osc} på 1,3 MHz.

For ned-konvertering til $U_O = 5$ V ved $I_O = 100$ mA og $C_T = 1$ nF bliver maksimum af f_{osc} på 400 kHz.

Der er intet til hinder for at benytte en højere værdi af oscillatorens frekvens, men tabene må så forventes tilsvarende højere end estimatets 1 % af udgangseffekten.

6.11.2 Spolens selvinduktion

Med en kendt frekvens kan spolen bestemmes ud fra kravet om hvor meget strømmen i spolen skal ændres i løbet af en periode. Herunder gengives formlerne for ændringen i perioden hvor transistoren er ON og det kan omsættes til oscillatorens frekvens ved $t_{ON} = D/f_{osc}$.

Type	Strømvariation = ΔI_L	Selvinduktion
Op	$\Delta I_{ON} = \frac{U_{DC}}{L_1} t_{ON}$	$L_1 = \frac{DU_{DC}}{\Delta I_L f_{osc}}$
Ned	$\Delta I_{ON} = \frac{U_{DC} - U_o}{L_1} t_{ON}$	$L_1 = \frac{D(U_{DC} - U_o)}{\Delta I_L f_{osc}}$

For en op-konvertering fra $U_{DC} = 5$ V til $U_o = 15$ V bliver $D = 0,68$. Med $\Delta I_o = 10$ mA og $f_{osc} = 1,3$ MHz bliver spolen $L = 250$ μ H.

For ned-konvertering fra $U_{DC} = 10$ V til $U_o = 5$ V bliver $D = 0,5$. Med $\Delta I_o = 10$ mA og $f_{osc} = 400$ kHz bliver spolen $L = 625$ μ H.

6.11.3 Kondensatorer

Udgangen fra op-konverteren er afbrudt i ON-fasen og kondensator C_2 skal her levere strømmen til udgangen. Kondensatorens mindste størrelse kan bestemmes af kondensatorligningen når der er truffet et valg om hvor meget udgangens spænding må variere (ripple) U_{RIP} . Der gælder ikke en helt tilsvarende relation for ned-konverteren da strømmen fra spolen er tilnærmedesvis konstant, men her vil indgangens strøm skifte mellem nul og udgangsstrømmens værdi. For at holde støjen fra konverteren på et rimeligt niveau dimensioneres kondensatoren så spændingen ændres med U_{RIP} for denne strømvariation.

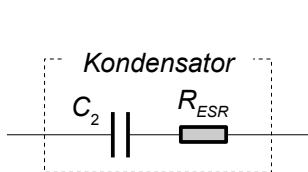
$$dq = C du = i dt \Rightarrow C > \frac{i dt}{du} = \frac{I_o t_{ON}}{U_{RIP}} = \frac{D I_o}{U_{RIP} f_{osc}}$$

For en op-konverter med $I_o = 100$ mA, $D = 0,68$ og $U_{RIP} = 10$ mV beregnes $C_2 > 5$ μ F ved $f_{osc} = 1,3$ MHz.
For en ned-konverter med $I_o = 100$ mA, $D = 0,5$ og $U_{RIP} = 10$ mV beregnes $C_1 > 12$ μ F ved $f_{osc} = 400$ kHz.

6.11.4 Kondensatorens ESR

Kondensatoren har en ækvivalent seriemarkstand ESR (equivalent series resistor) der optræder som en seriemarkstand. Den repræsenterer de samlede tab i kondensatoren. Værdien er funktion af både temperatur og strømstyrke samt kondensatorens konstruktion så den specificeres i databladet ved en given belastningsform. For brug ved denne indledende analyse antages den dog at være en konstant og reel modstand. Den bør ikke være ret meget over 0,1 Ω for anvendelse i en DC-DC konverter.

Type	Strømripple	Spændingsripple	Kondensatorens ESR
Op:	$I_{RIP} = \frac{U_o}{U_{DC}} I_o$	$U_{RIP} = R_{ESR2} I_{RIP}$	$R_{ESR2} = \frac{U_{RIP}}{I_o} \frac{U_{DC}}{U_o}$
Ned:	$I_{RIP} = I_o$	$U_{RIP} = R_{ESR1} I_{RIP}$	$R_{ESR1} = \frac{U_{RIP}}{I_o}$



Figur 364 – Den ækvivalente seriemarkstand (ESR) giver et spændingsfald over kondensatoren som funktion af strømmen i kondensatoren.

Ved op-konverteren giver kondensatoren i udgangen anledning til en ripplespænding fra dens ækvivalente seriemarkstand R_{ESR2} . Strømmen i kondensatoren svinger mellem den positive værdi på $I_L - I_o$ ved opladning og den negative værdi $-I_o$ ved afladning så den samlede ændring af spændingsfaldet over kondensatoren bliver I_L gange med R_{ESR} . Strømmen i spolen er lig med indgangsstrømmen som er relateret til udgangsstrømmen gennem $I_L = (U_o/U_{DC})I_o$ efter side 326. Indgangens kondensator er mindre kritisk da strømmen i spolen ikke momentant skifter retning.

Ved ned-konverteren giver kondensatoren i indgangen anledning til en en ripplespænding fra dens ækvivalente seriemodstand R_{ESR1} . Strømmen i kondensatoren varierer mellem nul når transistoren er OFF og I_o når transistoren er ON. Med en ekstern effektforsyning kan strømstødene udjævnes, men hvis det ignoreres kan en værdi estimeres på samme måde som vist ovenfor.

For op-konvertering fra $U_{DC} = 5$ V til $U_o = 15$ V med $U_{RIP} = 10$ mV og $I_o = 100$ mA kræves $R_{ESR} < 0,03 \Omega$.

For ned-konvertering fra $U_{DC} = 10$ V til $U_o = 5$ V med $U_{RIP} = 10$ mV og $I_o = 100$ mA kræves $R_{ESR} < 0,1 \Omega$.

6.11.5 Mindste belastning

Forudsætningen for at en DC-DC konverter kan fungere er at der løber en strøm i spolen som ikke når ned på nul i løbet af perioden ellers siges konverteren at være i *discontinuous mode* hvor analysens forudsætninger ikke gælder. Det betyder at der skal aftages en minimum strøm fra konverteren og det niveau er givet ved strømændringen ΔI_L i spolen. For op-konverteren er spolens middelstrøm lig med indgangsstrømmen, og den er igen lig med $(U_o/U_{DC})I_o$ som vist side 326, så spolens middelstrøm må ikke være mindre end U_o/U_{DC} gange med ΔI_L .

Type	Udgangsstrøm
Op	$I_o > \frac{U_o}{U_{DC}} \Delta I_L$
Ned	$I_o > I_L$

For ned-konverteren er middelværdien af spolens strøm lig med udgangsstrømmen så spolens strøm må ikke være mindre end ΔI_L .

En DC-DC konverter kan godt køre ved mindre udgangsstrøm, men det kræver at parameteren D kan reguleres ned i nærheden af nul og måske at transistoren ikke går ON i én eller flere perioder. En del af de kommersielt tilgængelige kredse for opbygning af DC-DC konvertere tilbyder mulighed for at arbejde i denne tilstand under betegnelser som fx *pulse skipping*. Konverterens frekvens bliver sænket idet én eller flere perioder helt overspringes og det kan give en lavfrekvent støj som kan være meget svær at dæmpe og som kan give hørbare gener i lydudstyr.

6.12 Konverterens tab

Fremstillingen har indtil nu benyttet ideelle komponenter, men det er relativt enkelt at medtage de mere åbenlyse tabskilder i analysen. Her vil jeg illustrere det ved modstandsværdierne af spolen, dioden og transistoren samt den mistede energi ved op- og afladning af parasitiske kapaciteter. Værdierne kan findes i databladet, måske med undtagelse af diodens seriemodstand, som i givet fald må findes ved måling på dioden, hvor man plotter den målte karakteristik og sammenligner med den forventede.

Det er ikke nødvendigt at opstille ligningerne påny for at aflure konverterens tab, for det er almindeligvis tilstrækkeligt med en idé om hvor store tabene er.

Type	Transistor ON	Transistor OFF	Modstandstab	Indre modstand
Op	$(R_L + R_T)I_L^2$	$(R_L + R_D)I_L^2$	$P_R = R_{T_{pp}}I_o^2$	$R_{T_{op}} = \frac{R_L + D R_T + (1-D) R_D}{(1-D)^2}$
Ned	$(R_T + R_L)I_L^2$	$(R_D + R_L)I_L^2$	$P_R = R_{T_{ned}}I_o^2$	$R_{T_{ned}} = R_L + D R_T + (1-D) R_D$

I tabellen står D for den brøkdel af tiden hvor transistoren er ON og $(1 - D)$ den brøkdel hvor transistoren er OFF. Modstandene R_L er spolens DC modstand, R_T er transistorens ON modstand og R_D for diodens seriemodstand, dertil er I_L middelværdien af strømmen i spolen og I_o er udgangens strøm.

Tabellen beregner tabene i hver af de to faser som effekten udtrykt ved modstandsværdien gange med kvadratet på strømmen i modstanden. Det kan udtrykkes ved en effektiv modstandsværdi R_T , der er den indre modstand i konverteren (Thevenin modstanden). I udtrykket for opkonvertering er

Effektkredsløb

der omregnet fra strømmen i spolen I_L til strømmen i udgangen I_o og heraf kvadratet på $1 - D$.

Note 19.

7 APPENDIKS

Her er samlet information af referencekarakter.

7.1 Konstanter

Symbol	Værdi	Definition	Kommentar
c_0	$299,79 \cdot 10^6$ m/s	$c_0^2 = 1/\epsilon_0\mu_0$	Lysets hastighed i vacuum
c	345 m/s		Lydens hastighed ved jordoverfladen
ϵ_0	$8,854 \cdot 10^{-12}$ F/m	$\epsilon_0 = 1/\mu_0 c_0^2$	Elektrisk konstant
ϵ_r			Relativ elektrisk konstant, oftest ≥ 1
k	$1,381 \cdot 10^{-23}$ J/K		Boltzmanns konstant
μ_0	$4\pi \cdot 10^{-7}$ H/m	Eksakt	Magnetisk konstant
m_0	$9,109 \cdot 10^{-31}$ kg		Elektronens masse
q_0	$1,604 \cdot 10^{-19}$ C		Elementarledningen, fundamental enhed

7.2 Variable

Symbol	Enhed	Definition	Kommentar
A	m^2		Areal
A	-		Forstærkning
α	-		Generel talværdi
B	T	$T = \text{Wb}/m^2$	Magnetisk induktion
B	-		Tilbagekoblingsfaktor
B	Hz		Båndbredde, støjbåndbredde
β	-	$\beta = I_C/I_B$	Transistorens strømforstærkning
C	F		Kapacitet
d	m		Diameter, afstand
d	-	$d = 1/2Q$	Dæmpningsfaktor
δ	m		Indtrængningsdybde
E	Vs^{-1}		Elektrisk feltstyrke
e_n	$V/\sqrt{\text{Hz}}$		Spændingens støjtæthed
F	N		Mekanisk kraft
Φ	Wb = Vs		Magnetisk flux
f	Hz		Frekvens (svingning per sekund)
G	$S = \Omega^{-1}$		Ledningsevne
g_m	$S = \Omega^{-1}$		Transkonduktans
H	-		Overføringsfunktion, funktion af s
h	-		Tidsrespons, funktion af t
I, i	A		Elektrisk strøm

i_n	A/Hz	Strømmens støjtæthed
J	A/m ²	Strømtæthed
L	H = Wb/A = Vs/A	Selvinduktion
L	dB	20 log10(p_1/p_{REF}) Lydniveau, $p_{REF} = 20 \mu\text{Pa}$
l	m	Trådlængde
N, n	-	Antal
ω	s ⁻¹	$\omega = 2\pi f$ Vinkelhastighed
P, p	W	$p = ui$ Effekt
Q, q	C	Elektrisk ladning
Q	-	Godhedsfaktor
R	$\Omega = V/A$	Modstand, resistans
R, r, a, b	m	Radius, afstand
ρ	Ωm	Specifik resistans
s	s ⁻¹	Laplace operator
T, t	s	Tid
T	K	Absolut temperatur
ΔT	K, °C	$\Delta T = T_2 - T_1$ Temperaturdifferens, 1 K = 1°C
t	m	Trådtynkelse
Δt	s	$\Delta t = t_2 - t_1$ Tidsinterval
θ	radian	Vinkel
U, u	V	Spændingsforskel
W	J	Energi (work) = effekt integreret over tid
x, y, z	(ikke oplyst)	Aksevariabel, generel variabel
Z	Ω	Elektrisk impedans

7.3 Forkortelser

Symbol	Definition	Kommentar
AC	Alternating current	Vekselstrøm eller vekselspænding
AVE	Average	Middelværdi
DC	Direct current	Jævnstrøm eller jævnspænding
GBP	Gain-Bandwidth Product	
μP	Microprocessor	
RMS	Root Mean Square	Effektivværdi
SNR	Signal to Noise Ratio	Signal-støj forhold
SR	Slew Rate	Hældningskoefficient (maksimal værdi)
THD	Total Harmonic Distortion	Samlet harmonisk forvrængning

7.4 Enhedssystemet

I denne bog anvendes det internationale enhedssystem, *Système international d'unités*, der i daglig tale blot kaldes SI, og som er implementeret i hele verden undtaget USA.

7.4.1 Historie

Systemet tager sit udgangspunkt i den franske revolution (1795 til 1812), da der opstod et behov for nye enheder, som ikke havde relationer til den afsatte adel. Længdeenheden meter (m) blev defineret som en ti milliontedel af afstanden fra nordpolen til ækvator, og masseenheden gram (g) defineres som massen af vand med et volumen på en milliontedel af en kubikmeter ved 0°C, der blev repræsenteret ved et lod af metal på 1000 g indtil den 1799 blev erstattet med et platinlod på 1 kilogram, med det nye præfix kilo. Dertil den allerede kendte tidsenhed sekund (s).

De europæiske lande gik 1875 sammen om at varetage udviklingen af et fælles målesystem baseret på det franske system, og ved den officielle introduktion af SI enhedssystemet var der tilføjet strømstyrken ampere, temperaturrenheden kelvin og lysenheden candela. I 1971 blev stofenheden mol tilføjet. I samme periode er længdeenheden redefineret som et bestemt antal bølgelængder af en karakteristisk lyskilde, og tidsenheden er redefineret ved frekvensen af lyset fra en bestemt lyskilde, så enhederne for tid og afstand er kombineret i en fælles definition. Der pågår arbejde med at redefinere de øvrige enheder ud fra basale fysiske konstanter, og det forventes at dette arbejde afsluttes inden 2015.

7.4.2 Enhed

Følgende grundenheder er defineret.

Enhed for ...	Symbol	Navn	Definition	Årstaal
Længde	m	meter	1/299 792 458 af den afstand som lys tilbagelægger på 1 sekund i vacuum,	1983
Masse	kg	kilogram	Standardlod i Paris	1889
Tid	s	sekund	Varigheden af 9 192 631 770 perioder af lys fra cæsium 133.	1967
Strømstyrke	A	ampere	Kraftpåvirkning på $2 \cdot 10^{-7}$ N/m mellem to parallelle tråde i 1 m afstand.	1946
Temperatur	K	kelvin	Absolut nul er 0 K = -273,15°C og vands triplepunkt er 273,16 K = 0,01 °C.	1967
Stofmængde	mol	mol	Samme antal stofpartikler som indeholdt i 0,012 kg af carbon 12.	1967
Lysstyrke	cd	candela	Lysintensiteten i en given retning ved frekvensen $540 \cdot 10^{12}$ Hz (grønt lys), der har strålingsintensiteten 1/683 W/sr.	1979

Enhedssymbolet skrives med lodret skrift og enhedens navn staves fuldt ud med små bogstaver (m, meter). Den eneste undtagelse er grammatiske krav i et givet sprogområde, som fx stort begyndelsesbogstav først i en sætning. Enheden må ikke flertalbøjes. Symbolet for en enhed opkaldt efter en forsker skrives med et stort begyndelsesbogstav (H, henry), og hvis der er sammenfald med en anden enhed benyttes et stort bogstav efterfulgt af et eller flere små bogstaver (W for watt og Wb for weber).

Ud over de syv grundenheder er der defineret en række afledte enheder, der alle kan føres tilbage

til de syv grundenheder.

Som et eksempel er enheden Ω (ohm) for modstandsværdi defineret som $\Omega = V/A$. Enheden for spænding er $V = J/C$, enheden for energi er $J = Ns$ og enheden for ladning er $C = As$, så det er muligt at skrive enheden for modstandsværdi som $\Omega = (Ns/As)/A = N/A^2$. Endelig er enheden for kraft $N = \text{kg m s}^{-2}$ så enheden for modstand bliver $\Omega = \text{kg m s}^{-2} A^{-2}$.

En enhed er et symbol og ikke en forkortelse, så der skal ikke punktum efter et enhedssymbol.

En typisk fejl er at sætte et punktum efter kr for møntenheden kroner; en stavekontrol vil meget ofte kræve et punktum her, hvilket er i modstrid med SI enhedssystemet.

7.4.3 Præfix

Et præfix skrives ligesom enheder med lodret skrift, de sættes altid foran enheden og SI systemet tillader ikke sammenstilling af enheder, så μ tillades ikke som erstatning for p (pico). Listen vist herunder viser ikke definitionen af centi, deci, deca og hecto, som SI anbefaler at undgå.

Præfix	atto	femto	pico	nano	mikro	milli	kilo	mega	giga	tera	peta	exa
Symbol	a	f	p	n	μ	m	k	M	G	T	P	E
Faktor	10^{-18}	10^{-15}	10^{-12}	10^{-9}	10^{-6}	10^{-3}	10^3	10^6	10^9	10^{12}	10^{15}	10^{18}

7.4.4 Symbol

SI enhedssystemet kræver at variable, parametre og konstanter skrives med kursiv og talværdier skrives med lodret skrift. Matricer skrives undertiden med en fed og lodret skrift.

7.4.5 Fysisk størrelse

J. C. Maxwell foreslog i midten af 1800-tallet at en fysisk størrelse skulle opfattes som en talværdi gange med en enhed og det benyttes konsekvent i SI enhedssystemet. Hvis den fysiske størrelse er x kan man betegne talværdien med $\{x\}$ og enheden med $[x]$, og den fysiske størrelse er derfor produktet af de to.

$$\begin{array}{ccc} \text{Symbol} & \text{Talværdi} & \text{Enhed} \\ x & = & \{x\} \cdot [x] \end{array}$$

Man kan foretage en matematisk beregning på både talværdi og enhed. Som et eksempel er farten 36 km/h lig med 10 m/s.

$$V = 36 \frac{\text{km}}{\text{h}} = 36 \frac{1000 \text{ m}}{3600 \text{ s}} = 10 \frac{\text{m}}{\text{s}}$$

En praktisk konsekvens heraf er at man kan tjekke en beregning for fejl ved at undersøge om enheden er korrekt. Hvis der er fejl i enheden er beregningen ganske enkelt forkert udført. Det er absolut anbefalelsesværdigt at lære relationen mellem de mest almindelige enheder indenfor ens faggruppe, så man allerede undervejs i en beregning kan se om der er noget galt.

En ofte set beregning er at benytte en matematisk funktion på en fysisk størrelse, hvilket ikke giver nogen mening hvis funktionen indvirker på enheden ikke kan defineres. Som eksempel benyttes hastigheden $V = 36 \text{ km/h}$ hvor beregningen giver $\log(V) = 1,56$ og ved $V = 10 \text{ m/s}$ giver beregningen $\log(V) = 1,00$. De to resultater burde være ens hvilket viser at man ikke kan se bort fra enheden.

7.5 Omregning

Det er muligt at repræsentere de forskellige enhedsværdier grundenhederne. I tidligere litteratur blev det refererer til i form af MKSA stående for meter, kilogram, sekund og ampere. Herunder vises hvordan de mest relevante enheder for nærværende bog er defineret.

Parameter	Enhed	Definition	Relation	MKSA
Længde	m	Grundenhed		m
Masse	kg	Grundenhed		kg
Tid	s	Grundenhed		s
Strøm	A	Grundenhed		A
Ladning	C	As		s A
Hastighed	m/s			$m s^{-1}$
Acceleration	m/s^2			$m s^{-2}$
Kraft	N	$kg m s^{-2}$		$kg m s^{-2}$
Energi	J	Nm	$kg m s^{-2} \cdot m$	$kg m^2 s^{-2}$
Effekt	J/s	Nm/s	$kg m^2 s^{-2} / s$	$kg m^2 s^{-3}$
Spænding	V	J/C	$kg m^2 s^{-2} / A s$	$kg m^2 s^{-3} A^{-1}$
Modstand	Ω	V/A	$kg m^2 s^{-3} A^{-1} / A$	$kg m^2 s^{-3} A^{-2}$
Kapacitet	F	C/V	$A s / kg m^2 s^{-3} A^{-1}$	$kg^{-1} m^{-2} s^4 A^2$
Magnetisk flux	Wb	Vs	$kg m^2 s^{-3} A^{-1} \cdot s$	$kg m^2 s^{-2} A^{-1}$
Selvinduktion	H	Wb/A	$kg m^2 s^{-2} A^{-1} / A$	$kg m^2 s^{-2} A^{-2}$

7.6 Litteraturliste

- AL A. Laursen *Varmetransmission i elektronik*, Elektronikcentralen 1970.
- CC H. Craig Casey, Jr. *Devices for Integrated Circuits*, John Wiley & Sons, 1999.
- DS Douglas Self *Audio Power Amplifier Design handbook*, Focal Press, 2009.
- ER E. G. Richardson *Sound*, Edward Arnold & Co, second edition, 1935.
- HJ K. G. Hansen, Niels Jonassen *Elektricitet, Magnetisme, Atomfysik*, Statens trykningskontor, Un 04-352, 1967, Genoptryk 1972.
- JH Jean Hiraga *Rörförstärkare*, Editions frequences, Tessekart Media Förlags, 1986.
- KR Knud Rasmussen *Højtalere*, Note 2109, Institut for akustisk teknologi, 1996.
- LB Leo L. Beranek *Acoustics*, Acoustical Society of America, 1993.
- M *Reference Manual of Transistor Circuits*, Mullard Ltd., 1960.
- MH Jacob Millman, Christos C. Halkias *Integrated Electronics: Analog and Digital Circuits and Systems*, International Student Edition, McGraw-Hill, 1972.
- ML W. Marshall Leach, Jr. *Introduction to Electroacoustics & Audio Amplifier Design*, Third Edition, Kendall / Hunt, 2003.
- NA Nakhlé H. Asmar *Partial Differential Equations*, Second Edition, Prentice Hall, 2005.
- RW Lennart Råde og Bertil Westergren *Mathematics Handbook for Science and Engineering*, femte udgave, Studentlitteratur, 2004.
- SF Sergio Franco *Design with Operational Amplifiers and Analog Integrated Circuits*, Third Edition, McGraw Hill, 2002.
- TØ Tore Øvensen *Halvlederteknik*, anden udgave, Teknisk forlag, 1981.

8 INDEKS

0	
0 dB båndbredde.....	134
1	
1/f støj.....	157
10/1000 puls.....	317
3	
3 dB grænsefrekvens.....	172
4	
4-tråds måling.....	13
455 kHz.....	202
A	
A-vejning.....	48
Absolute Maximum Ratings.....	298
Afkobling.....	167, 216
Aliasing.....	85f.
AM.....	201
Ampere timer.....	24
Amperes lov.....	28
Analog signalbehandling.....	208
Anders And.....	4
Annonce.....	4
IDA.....	2
PR electronics.....	163
Anode.....	282
Antiparallele dioder.....	202
Arbejdslinje.....	297
Arrhenius.....	298
Atan2.....	86
B	
Balance.....	107
Basisjordet.....	246
Batteri.....	24
Beskyttelsesdiode.....	230
Beskyttelseskredsløb.....	298
Biasstrøm.....	131, 156
Bifilar vikling.....	12
Bin.....	86
Binær repræsentation.....	85
Biot og Savarts lov.....	34
Bly sulfid.....	194
Bode plot.....	171
Bonding.....	195
Boost converter.....	326f.
Boot strap.....	243, 290, 296
Brick-wall filter.....	158
Buck converter.....	326, 330
Buffer.....	109, 242
Butterworth.....	181
Bølgelængde.....	212
Båndbredde.....	61
C	
CFP.....	295
Chopper stabilised amplifier.....	130
Chorus.....	96
cis(θ).....	63
Clamp voltage.....	318
CMOS.....	250
4069UB.....	125
74HC04.....	124
74HC14.....	124
Forstærker.....	250
HEF4007.....	230, 251
HEF4069UB.....	251
CMRR.....	115, 259
Comparator.....	123
Cross-over forvrængning.....	99, 292
Current dumping.....	293
D	
Darlington pair.....	242, 295
DC lys.....	216
DC-DC konverter.....	326
Discontinuous mode.....	334
Nedkonvertering.....	330
Opkonvertering.....	327
Decibel.....	47
Depletion layer.....	194, 276
Derating.....	13
Designprocedure.....	
BJT forstærker.....	243
JFET forstærker.....	248
Designregler.....	174
Dielektrisk absorption.....	25, 175
Dielektrisk konstant.....	20
Differentialkvotient.....	66
Differentiation.....	
Brøk af funktioner.....	76
Funktion.....	66
Diffusion.....	194
Diffusionslængde.....	279
Diode.....	193, 275
1N4007.....	198
1N4148.....	197f., 207
1N5819.....	197, 207, 281
AA119.....	197
BAS116.....	197
BB207.....	210
BF245.....	227
BP104.....	215
BPW21.....	215
BPW34.....	116
BYV59C.....	199
LED.....	207
SFH203.....	214
SFH485.....	212
Diodemikser.....	204
Diodemodstand.....	206
Diodespænding.....	206
Diskret Fourier transformation.....	87
E	
Early voltage.....	239
ECL.....	224
EE.....	295
Effektforsyning.....	167
Effektivværdi.....	49
Eksponentialefunktion.....	63
Elektretmikrofon.....	26
Elektrisk felt.....	9
Elektrisk konstant.....	9
Elektrisk strøm.....	8
Elektro-dynamisk højtalere.....	46, 102
Elektronrør.....	282
EMC direktivet.....	317
Emitterfølger.....	242
Emittermodstand.....	242
Energi.....	13
Enhed.....	339
ampere.....	8
ampere per kvadratrod hertz.....	150
bel.....	47
coulomb.....	8
dBm, dBu, dBV.....	48
farad.....	20
henry.....	29
neper.....	47
tesla.....	27
volt per kvadratrod hertz.....	150
weber.....	28
Epitaxial layer.....	195
Eulers identitet.....	75, 79
F	
Fakultet.....	73
Faradays induktionslov.....	29
Fasedrejningsfilter.....	176
Fasehastighed.....	60
Feed back.....	97
Feed forward.....	308
Fejfspænding.....	129
Filter.....	171
Butterworth.....	180, 188
Chebychev.....	189
Design.....	190
Første orden.....	172
Højere orden.....	187
Højpas transformation.....	190
Firetrådskobling.....	13
Fluorescens.....	271
Flux-vindingstal.....	29
FM.....	201
Formler.....	3

Indeks

DC-DC konverter.....	330	Grundtone.....	78	Komplementært signal.....	114
Decibel.....	47	Gruppetidsforsinkelsen.....	178	Kondensator.....	20, 25
Diodespænding.....	198	Grænseværdi.....	61	Elektrolyt.....	25, 310
Diodestrøm.....	196, 279	H		Film.....	25
Effektivværdi.....	49	h-parameter model.....	247	Keramisk.....	25
Egenstøj.....	266	Haglstøj.....	156, 266	NP0.....	25
Elektroner per rumfang.....	274	Harmonisk forvrængning.....	99	X7R.....	25
Energi.....	43	Cross over.....	292	Kondensatorligningen.....	265
Energi (spole).....	43	Radiorør.....	288	Kondensatormikrofon.....	26
Fermi-Dirac fordeling.....	273	Harmonisk række.....	78	Konstanter.....	337
Forstærker med BJT.....	243f.	HDTV.....	77	Boltzmann.....	273
Forvrængning.....	256	Heaviside.....	40, 89	Elementarladning.....	8
Godhed.....	44	Howland strømkilde.....	117	Elementarladningen.....	271
Halvledermodstand.....	274	Hvid støj.....	147, 150	Permeabilitet.....	28
Hybrid-pi model.....	235	Hybrid-pi model.....	235	Permittivitet.....	20
Indbygget potentiale.....	276	Hybride kredsløb.....	165	Specifik modstand.....	11, 274
Kapacitet.....	20	Højspændingskabel.....	32	Vacuumpermittivitet.....	9
Kondensatorligning.....	22	Højtaler.....	46	Konvektion.....	304
Kraft (magnetisk).....	45	I		Kraft.....	45
Magnetisk induktion.....	28, 36	Idealitetsfaktor.....	197	Kraftfaktor.....	45, 102
Middelværdi.....	49	BC547.....	220	Kredsløb.....	3
Minoritetsbærere.....	277	IEC.....	15	78-type.....	323
Modstand.....	11	IGBT.....	330	Addition af signaler.....	112
RMS algoritme.....	53	Impedans.....	61, 65	Beskyttelsesdioder.....	202
Sammenfatning.....	55	Indgangskapacitet.....	112	BJT forstærker.....	245
Selvinduktion.....	30	BJT.....	237	Båndpas.....	176
Brooks coil.....	34	OpAmp.....	263	Båndpasfilter.....	184
Dobbeltleder.....	31	Indgangsmodstand.....	111	Chebychev lavpasfilter.....	190
Ledning.....	30, 38	BJT.....	237	Differentialtrin.....	255
Maxwell.....	37	MOSFET.....	229	Effektforsyning.....	310
Skiveformet.....	36	OpAmp.....	263	Elektretmikrofon.....	26
Toroid med luftspalte.....	41	Indre modstand.....	19	Elektrostatisk højtaler.....	26
Wheeler (enkeltag).....	33	Indtrængningsdybde.....	32, 218	Fasedrejning.....	177
Wheeler (flere lag).....	34	Instrumentationsforstærker.....	115	Forstærker med BJT.....	233, 243
Slew rate.....	265	Integration.....	68	Forstærket emitterfølger.....	246
Stabilitet.....	264	Isolator.....	10	Geometrisk middelværdi.....	210
Strømfotrængning.....	32	ITU.....	317f.	Halvbølge ensretning.....	201
Temperaturspænding.....	197	J		Helbølgeensretter.....	201
VAS.....	261	JFET.....	227	Højpas, første orden.....	175
Forstærket diode.....	294	K		Højpasfilter.....	118
Forvrængning.....	99, 236	Kapacitet.....	20	Højpasfilter, anden orden.....	182
Fotodiode.....	109, 116	JFET gate.....	227	JFET strømgenerator.....	227, 248
Fotolitografi.....	195	Skaermkabel.....	21	Klasse C udgangstrin.....	306
Foton.....	271	Tolederkabel.....	21	Lavpasfilter.....	120
Fourier.....	77	Katalogfirmær.....	198	Lavpasfilter, anden orden.....	183
DFT.....	87	Katode.....	282	Lavpasfilter, første orden.....	174
Transformation.....	81	Kelvinkobling.....	13	LM317 serieregulator.....	321
FR4.....	20	Keramisk kondensator.....	25	Lysdetektor.....	216
Frekvensbånd.....	61	Kirchhoffs love.....	16	Lysmåler.....	215
Frekvensområde.....	159	Kirchhoffs strømlov.....	16	Mikrofonforstærker.....	110, 233
Fuzz box.....	166	Kirchhofs spændingslov.....	16	Mikser.....	203
Fælles basis.....	246	Klasse A.....	288	MOSFET effektforstærker.....	249
Fælles emitter.....	222	Klasse AB.....	294	MOSFET kontakt.....	231
Fælles kollektor.....	242	Klasse B.....	292	OpAmp.....	262
G		Klasse C.....	306	PT100 temperaturføler.....	13
Gasudladning.....	318	Klasse D.....	308	Radiomodtager.....	202
GBP.....	134, 253	Kobberoxyd.....	194	Serieregulator.....	321
Geometrisk middelværdi.....	209	Kompleks Fourier serie.....	79	Spændingsstyring.....	207
Germanium.....	194	Komplekse tal.....	62	Stopfilter.....	185
Gibbs fænomen.....	78	Komplementær transistor.....	219	Strain gauge (brokobling).....	14
Gitter.....	282			Strømspejl.....	258
Godhed.....	61, 180			Termometer.....	208
Graetz kobling.....	310			Udgangstrin.....	295
				VU-meter.....	209

Kvarts.....	194	Nyquist.....	52, 86
		Stabilitet.....	140
L			
Ladning.....	8	Octave.....	88
Ladningsforstærker.....	121	Offsetstrøm.....	131
Laplace.....	89	Ohms lov.....	11
Impedans.....	91	OpAmp.....	105, 107, 253
Transformationspar.....	92	4069UB.....	125
Laser.....	213, 271	AD797.....	106, 131, 162, 240, 266
Latch-up.....	164	AD846.....	269
LC led.....	74	AD8628.....	130
LC svingningskreds.....	306	ADA4051.....	106, 129, 161
Ledningsbånd.....	271	CA8030.....	268
Li-ion batteri.....	19	Chopperforstærker.....	270
Lige funktion.....	75, 78	L165V.....	106, 165
Ligesvævende temperatur.....	225	LF356.....	129, 131, 164
Limes.....	66	LF357.....	141
Linearitet.....	17	LM13700.....	268
Lokaloscillator.....	204	LM2902.....	254
Lorentz.....	45	LM324.....	106, 129, 131, 164, 254
Luftspalte.....	41	LM358.....	254
Lyntransient.....	317	LM386.....	259
Lyserød støj.....	147, 157	LM3900.....	239, 269
Lyttegræthed.....	292	LM393.....	123
Lækstrøm.....	193, 197	LM675T.....	106, 165
M			
Magnetfelt.....	27	LMH6624.....	141, 161
Magnetisk induktion.....	27	NE5532.....	131, 154
Centrum.....	36	NE5534.....	106, 265
Kort leder.....	35	Norton amplifier.....	269
Magnetisk modstand.....	40	OP27.....	129, 131, 266
Majoritetsbærer.....	277	TDA2040.....	106
Maksimalt fladt filter.....	180	TDA7294.....	106
MATLAB.....	88	TL071.....	106, 115, 131, 161, 164,
Maxwell.....	27	265	
Mayer-Norton ækvivalent.....	19	TLC271.....	106, 161, 260
Mekanisk kraft.....	9	TLC274.....	123, 131
Middellevetid.....	271	uA741.....	106, 129, 131, 136, 164
Middelværdi.....	49		
Mikrofoni.....	175		
Mikser.....		Operationsforstærker.....	105
Balanceret.....	204	Udgangsmodstand.....	139
Diode.....	203		
Gilbert celle.....	260		
Miller transformation.....	234		
Minimoog.....	225		
Minoritetsbærer.....	277		
Mobilitet.....	219, 274		
Modstand.....	10		
0603.....	315		
AC01.....	12, 319		
SFR16.....	315		
Modulus.....	79		
Motor boating.....	244		
Mættet.....	107		
Måling af.....			
mekanisk deformation.....	14		
Spænding.....	114		
strøm.....	202		
temperatur, PT100.....	13		
N			
Norton ækvivalent.....	19		
		Peltier element.....	310
		Pentode.....	283, 288
		pH-probe.....	109
		Phantom Power.....	26, 167
		Phasing.....	96
		Philips EE20.....	4, 288
		Piezo-elektrisk krystal.....	121
		Planarteknologi.....	195
		PLC.....	167
		Polynomium.....	67
		Popcorn støj.....	162
		Potentiale.....	10
		Potentiel energi.....	9, 43
		Producenter.....	198
		Propagation delay.....	124
		Præfix.....	340
		PT100.....	13
		Push pull.....	165, 289f.
		PWM.....	306f.
		Q	
		Q.....	61
		QUAD 405.....	293
		R	
		Radiorør.....	7, 106, 282, 288
		ECC83.....	283
		EL84.....	288
		Radiotransmission.....	201
		Reluktans.....	40
		Resistivitet.....	11
		Resonans.....	42
		Resonansfrekvens.....	61, 103
		Ringkerne.....	40
		Ripplespændingen.....	311
		RMS.....	49
		Root mean square.....	49
		Rosa støj.....	147, 157
		RS422.....	114
		Rør.....	282
		S	
		Sampling.....	85
		Sampling frequency.....	85
		Samplingfrekvens.....	86
		Schmitt-trigger.....	123
		Seebeck element.....	310
		Sekundærbatteri.....	24
		Selen.....	194
		Selvinduktion.....	29
		Kinetisk energi.....	43
		Lang spole.....	32
		Sensitivetsanalyse.....	74
		Shannon.....	52, 86
		Shottky-diode.....	279
		SI enhedssystem.....	339
		Silicium.....	194
		Sinc.....	81
		Slew rate.....	136, 261, 265
		Slutværdisætningen.....	91
		Sløjfefforstærkningen.....	140
		SMD.....	12
		Småsignalmodel.....	
		BJT.....	235
		JFET.....	248
		SOA.....	298
		Sollys.....	212
		Solpanel.....	214
		Sound Pressure Level.....	48
		Specifik ledningsevne.....	274
		Specifik modstand.....	274
		Specifik varmeledning.....	304
		Spejlfrekvens.....	202
		Spektrum.....	86
		Spændingsforskæl.....	10
		Spændingsfølger.....	108
		Spændingsreference.....	
		ADR4520.....	200
		LM4051A.....	200
		Spændingsstøjæthed.....	150
		Stagger-tuned filter.....	203
		Standardfordeling.....	149
		Startbetingelse.....	71
		Startværdisætningen.....	91

Indeks

State-variable filter.....	186	Harmonisk forvrængning.....	236	Udgangsmodstand.....	
Stationær tilstand.....	60, 64, 328	Kapacitet.....	174	BJT.....	224
Stelplan.....	168	Kapacitet til omgivelser.....	100	CMOS.....	251
Steradian.....	212	Lav støj.....	156	Hybrid-pi model.....	239
Stigetid.....	136	Lækstrøm i elektrolyt.....	100	JFET.....	249
Strain gauge.....	14	Parasitiske komponenter.....	100	Operationsforstærker.....	145
Strømforstærkning.....	219, 221	R, C og L.....	55	Uændelig baffel.....	102
Strømkilde.....	116, 259	Sampling.....	86	Ulige funktion.....	75, 78
Strømretning.....	10	Selvinduktion af leder.....	100	Undervandsbåd.....	32
Strømtæthed.....	28	Stelplan.....	174	USB.....	114
Strålingsintensitet.....	212	Udgangsmodstand, BJT.....	247		
Støjtal.....	235, 240	Transformator.....	40	V	
Støjtæthed.....	150	Transistortrekven.....	237	Valens.....	194
Subtraktor.....	113	Transistor.....		Valensbånd.....	271
Superposition.....	17	2N3055.....	290	Valenselektron.....	7, 271
Surge Transient.....	317	AC126.....	288	Varistor.....	318
Svingningskreds.....	42	BC547.....	301	Varmestråling.....	304
Svingspole.....	46	IRFZ44N.....	301	Vinkelhastighed.....	63, 75
Switch-mode converter.....	326	MJ480.....	290	Virkningsgrad.....	314
T		TO220.....	301	Radiorør.....	288
Taylor polynomium.....	73	Transistor.....		Virtuelt nul.....	111f.
Temperaturkoefficient.....	11	2N3819.....	228	Volt.....	10
1N4148.....	198	BC547.....	219, 221, 223	Volumehastighed.....	103
Bipolær transistor.....	220	BC547B.....	207	W	
Kondensator.....	25	BC549.....	240	Wireless World.....	290
Strømforstærkning.....	221	BC847.....	219, 221	Z	
Transistor.....	219	BF245.....	228	Zener.....	200
Zenerdiode.....	200	BF245A.....	228	1.5KE.....	318
Temperaturområde.....	127	BF545.....	228	BZX84.....	315
Termisk modstand.....	301	JFET auto-biasing.....	227	LM129.....	315
Termoelement.....	153	Maksimal forstærkning.....	239	LM368.....	315
Termokabler.....	110	Mætning.....	223	LT1021.....	315
Termometer.....	198	Transitfrekvens.....	134, 238, 253	MC1400.....	315
THD.....	166	Transkonduktans.....		Widlar kredsløb.....	316
Thevenin ækvivalent.....	18f.	BJT.....	233, 236, 256	Zero padding.....	88
Thevenin-modstanden.....	19	CMOS.....	251	Æ	
Tidskonstant.....	23	Differentialtrin, BJT.....	255	Ækvivalent seriemodstand.....	333
Tilbagekoppling.....	97, 244	ECC83.....	283	Ækvivalent støjbåndbredde.....	157
Tilbagekoblingsfaktor.....	97, 135	JFET.....	248	Ø	
Tommelfingerregel.....		JFET og MOSFET.....	257	Ørsted.....	27
Afkobling.....	174	LM13700.....	268	Å	
Effektafsættelse.....	13	MOSFET.....	249	Åben-sløjfe forstærkning.....	128
Fasemargin.....	141	TranZorb.....	318		
Gate sammenbrud.....	229	Triode.....	282		
GBP specifikation.....	173	TTL.....	224		
Grænsefrekvensen.....	174	U			