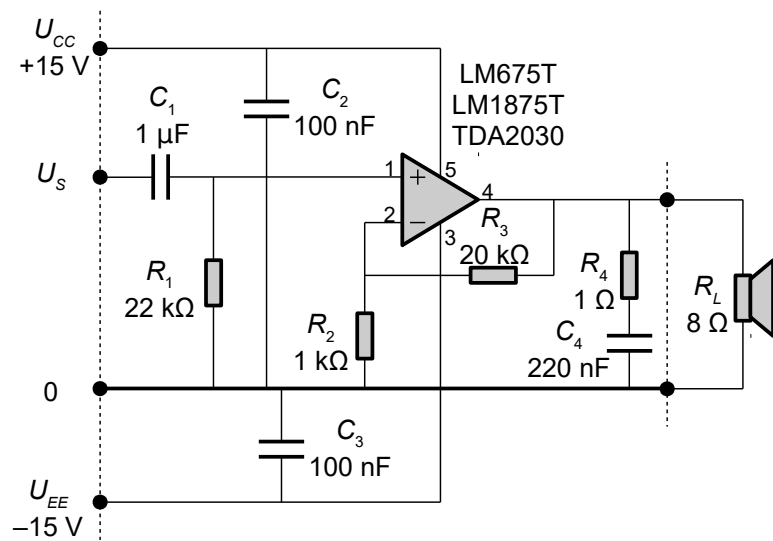



Analogteknik

T-005

Tore Skogberg

2017





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 Ingeniørhøjskolen i Århus for E3MSE mixed-signal elektronik og E4ASD analog systemdesign og som fra 2017 introducerer elektronikkomponenterne på første semester.

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 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 ved en CD og 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 designer analoge filtre med operationsforstærkere og i femte kapitel forklares om hvordan halvlederne og operationsforstærkeren er bygget. 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 billederne og læser billedteksten. Spring over tekstafsnit og formler. Når du derefter har skabt dig et overblik er det tid til at læse de dele af den ledsagende tekst som du har interesse i.

Bogens illustrationer har mange formler tilknyttet og kan sine steder virke som en designmanual, men det er dit ansvar at benytte oplysningerne korrekt. En kær lærer i folkeskolen sagde engang at *man må kun bruge en formel hvis man læser mindst fem linjers tekst inden og efter formelen.* Det råd er 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 behovet på Ingeniørhøjskolen. Siden er bogen blevet opdateret for at rette fejl og tilføje nyt materiale.

Min egen tilgang til elektronik startede i barndommen, men uden at være målrettet for jeg har blot været nysgerrig. Som dreng legede jeg med Märklin-tog og en hobbybog om elektriske kredsløb startede min interesse for elektronikken. Jeg så en annonce i et Anders And blad for Philips EE20 byggesættet der gav mulighed for at konstruere 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 tørst læste jeg billederne 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 en højttalers funktion og hvorfor nogle rum er bedre til lyd gengivelse 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 Ingeniørakademi 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 der har akkompagneret et kor ved flere koncerter. Fra 2006 er jeg civilingeniør i akustik fra Danmarks Tekniske Universitet og jeg har siden arbejdet med den elektroakustiske side af et headset, magnetisk opladning af batterier, interface til tryktransducere og test af software. Siden 1983 har jeg arbejdet med analog og digital elektronik, mikroprocessorer og dokumentation. September 2010 blev jeg ansat som vikar ved Ingeniørhøjskolen, februar 2011 som adjunkt og fra april 2013 er jeg lektor ved ingeniørhøjskolen.

En varm 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 mange diskussioner om elektronik og lyd og Arne Justesen (aju@ase.au.dk) for gode diskussioner om DC-DC konvertere og magnetiske systemer. Endelig en meget stor tak til de mange studerende der har stillet spørgsmål 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 og 10. Jeg har benyttet open-source programmet Octave og det professionelle MATLAB for modeltest. En del illustrationer har været en tur igennem GNOME og IrfanView billedredigeringerne. Jeg er gået over til at bruge islandsk tegnsætning da den giver et læsevenligt resultat og undgår den overflod af kommaer som kendetegner den grammatiske tegnsætning.

Bogen er trykt af Stakbogladen A/S.

Uddrag af bogens materiale må kun kopieres og benyttes i henhold til ophavsretsloven.

Jeg har også skrevet en bog om elektroakustik som du kan downloade fra min hjemmeside. Gå ind under "Publikationer" hvor Elektroakustik.pdf er tilgængelig.

www.torean.dk

Indholdsfortegnelse

1 ELEKTRONIKKOMPONENTER.....	7	3.7 Støj.....	181
1.1 Modstand – R.....	9	3.8 Sammenfatning.....	196
1.2 Konventioner og love.....	15	3.9 Spændingsområde.....	198
1.3 Diode – D.....	20	3.10 Overføringsfunktion.....	201
1.4 Transistor – T.....	23	3.11 Tilbagekobling.....	202
1.5 Kondensator – C.....	30	3.12 Opgaver.....	207
1.6 Spole – L.....	36	4 DESIGN.....	209
1.7 Resonans – LC.....	39	4.1 Oscillatorer.....	209
1.8 Elektricitet.....	42	4.2 Omsætning af fysiske størrelser.....	213
1.9 Ladning – Q.....	43	4.3 Korrektionskredsløb.....	220
1.10 Elektrisk felt – E.....	44	4.4 Ikke-lineære kredsløb.....	223
1.11 Magnetisk felt – B.....	50	4.5 Filtre.....	225
1.12 Magnetiserbar kerne.....	64	4.6 Designregler.....	246
1.13 Magnetiske kræfter.....	67	5 HALVLEDER.....	249
1.14 Signaler.....	69	5.1 Historisk oversigt.....	249
1.15 Opgaver.....	76	5.2 Dioden.....	250
2 ANALYSE.....	77	5.3 Lysdiode.....	267
2.1 Lineær algebra.....	78	5.4 Transistor ved DC.....	272
2.2 Eksempel med BJT.....	85	5.5 Transistor ved AC.....	286
2.3 Trigonometriske funktioner.....	86	5.6 Operationsforstærker.....	306
2.4 Logaritmiske funktioner.....	87	5.7 Specielle operationsforstærkere.....	321
2.5 Eksempel med MOSFET.....	88	5.8 Halvlederteori.....	324
2.6 Passive komponenter.....	89	5.9 Elektronrøret.....	336
2.7 Impedans.....	90	5.10 Opgaver.....	338
2.8 Komplekse tal.....	92	6 EFFEKTKREDSLØB.....	341
2.9 Spændingsdeling.....	95	6.1 Effektforstærker.....	341
2.10 Differentialregning.....	96	6.2 Klassifikation af udgangstrin.....	342
2.11 Taylor polynomier.....	103	6.3 Effektforsyning.....	363
2.12 Eulers identitet.....	106	6.4 Ensretter.....	363
2.13 Første ordens differentialligning.....	108	6.5 Spændingsregulator.....	367
2.14 Anden ordens differentialligning.....	113	6.6 Shuntregulator.....	368
2.15 Fourier serier.....	118	6.7 Serieregulator.....	374
2.16 Kompleks Fourier serie.....	120	6.8 DC-DC konverter.....	378
2.17 Fourier transformation.....	122	6.9 Opkonvertering.....	379
2.18 Diskret Fourier transformation.....	126	6.10 Nedkonvertering.....	382
2.19 Laplace transformation.....	130	6.11 Dimensionering.....	384
2.20 Analogier.....	146	6.12 Konverterens tab.....	386
2.21 Opgaver.....	149	7 APPENDIKS.....	387
3 OPERATIONSFORSTÆRKER.....	151	7.1 Konstanter.....	387
3.1 Lidt historie.....	152	7.2 Variable.....	387
3.2 En OpAmp.....	153	7.3 Forkortelser.....	388
3.3 Eksempler.....	154	7.4 Enhedssystemet.....	389
3.4 DC begrænsninger.....	165	7.5 Omregning.....	391
3.5 Båndbredde.....	170	7.6 Litteraturliste.....	392
3.6 Ind- og udgangsimpedans.....	178	8 INDEKS.....	393

1 ELEKTRONIKKOMPONENTER

Elektronikkens komponenter introduceres og det vises hvordan komponentværdierne kan beregnes fra de fysiske dimensioner med eksempler på anvendelse som temperatur- og kraftfølere. Ohms og Kirchhoffs love præsenteres som værktøj for analyse af elektroniske kredsløb. Kapitlet afsluttes med transformatoren og de magnetiske kræfter samt en introduktion til beregning af middelværdi og RMS.

Det er kutyme at opdele elektronikkens komponenterne i aktive og passive hvor de sidstnævnte hedder sådan fordi de tapper energi fra kredsløbet. Den logiske konklusion ville så være at den aktive komponent måtte tilføre energi til kredsløbet, men det gør den ikke. En aktiv komponent styrer hvor meget energi der skal flyde i et kredsløb så den er en regulator.¹



Figur 1 – De passive komponenter modstand, kondensator og spole (Wikipedia).

De passive komponenter er modstanden (*resistor*), kondensatoren (*capacitor*) og spolen (*coil*). En modstand omsætter elektrisk energi til varme mens de to sidste mere er nogle temporære lagre af energi. De kan dog afgive lidt energi til omgivelserne gennem elektriske og magnetiske felter og der kan også tages energi som varme.



Figur 2 – De aktive komponenter radorør, diode og transistor (Wikipedia).

Som aktive komponenter regnes radorøret, dioden og transistoren. De styrer eller regulerer en energistrøm i et kredsløb og muliggør opbygning af radiomodtagere og computere. Dioden har to tilledninger og er siden 1874 kendt for at lede strømmen i den ene retning samt spærre for den i modsat retning. Det er en vigtig egenskab, men dioden kan en del mere som temperaturmåling, spændingsregulering (*zener diode*), spændingsstyret kapacitet (*varicap diode*), oscillator for mikrobølgeområdet (*Gunn diode*), detektion af lys ved datatransmission (*photo diode*), produktion af elektrisk energi ved solpaneler (*solar cell*) og udsendelse af lys som LED.

Radorøret og transistorerne kan regulere energien i et kredsløb styret fra et spændingsniveau i en anden del. Radorøret introduceres i 1900 og starter den elektroniske udvikling, 1950 introduceres den bipolare transistor og 1980 overtager efterfølgeren MOSFET der i dag er dominerende.

¹ https://en.wikipedia.org/wiki/Electronic_component.

En diode kan udsende lys i form af en LED (*light-emitting diode*) ved brug af egnede materialer og finder i øjeblikket indpas i både biler og hjemmet hvor den erstatter energiforbrugende glødepærer og giver helt nye muligheder for belysning. Tidligere var dioden mest kendt for at udsende lys i det infrarøde spektrum og det fandt hurtigt anvendelse ved optisk kommunikation, men den moderne teknologi tillader lys der dækker hele spektret fra infrarød til ultraviolet.



Figur 3 – Lysdioder kan laves med forskellig farve og ved fleksibel montage tillader det brugeren at digte med ved montagen (Wikipedia).

De komponenter der tilfører energi til kredsløbet er alle omsættere af anden type af energi. Som et par eksempler frigør batteriet kemisk bundet energi til elektrisk energi og generatoren omsætter en mekanisk energi til elektrisk energi.



Figur 4 – Batterier omsætter kemisk bundet energi til elektrisk energi og en generator omsætter mekanisk energi til elektrisk energi (Wikipedia).

Termodynamikken siger at energi ikke kan opstå eller forsvinde. Produktion af elektrisk energi er derfor en omsætning af energi fra én form til en anden. Jorden rummer energi i sin varme kerne der blandt andet flytter rundt på de tektoniske plader, skaber jordskælv og giver vulkansk aktivitet. Det er muligt at føre lidt af den termiske energi op til overfladen og bruge den som en næsten uudtømmelig ressource til opvarmning og som gennem turbiner kan danne elektrisk energi.

Solens fusionsreaktor er i menneskets tidshorisont en stabil energikilde der sender kernepartikler og fotoner til jorden. Det kan udnyttes for opvarmning og som vindenergi, men fotonerne kan også modtages af lysfølsomme dioder og omsættes til elektrisk energi i solpaneler. Fusionsprocessen benytter omlagring af energi i atomets kerne og den resterende energi stråles ud til omgivelserne. Solen genererer derfor ikke energi ligeså lidt som jordens varme kerne gør det. Hvor energien oprindeligt kom fra kan jeg ikke svare på. Ofte nævnes "Big Bang" som et bud, men det besvarer ikke det centrale i spørgsmålet.

1.1 Modstand – R

Normalt er en modstand en komponent man køber til en specificeret værdi og de bestilles i ruller med nogle tusinde ens komponenter for automatisk bestyknings af et printkort.



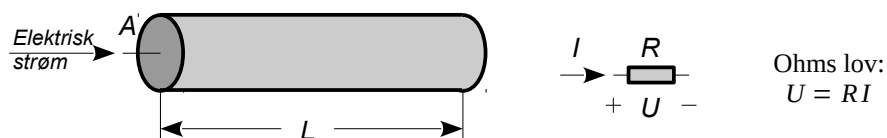
Figur 5 – SMD metalfilm indtil 0,5 W (Yaego). SFR25H af metalfilm på 0,5 W og AC01 trådviklet modstand på 1 W (Vishay). Modstand i TO-220 hus på 20 W for montage på køleplade (MQEC). Trådviklede modstande i cement for 10 til 100 W ved montage på en køleplade (Arcol).

Modstandsmaterialet kan være en tynd metalfilm der dampes på en keramisk basis hvilket er det almindelige både ved overflademontage SMD (*Surface-Mounted Devices*) som er vist ovenfor til venstre og for modstande med trånder. Modstandsmaterialet kan også være en tynd tråd der er viklet omkring et keramisk materiale og forseglet med lak eller cement. Denne type modstand kan tåle høj arbejdstemperatur og de kan optage ganske meget energi ved kortvarig overlast.² For at undgå viklingens selvinduktion benyttes to parallelle og tætliggende ledere hvor strømmen føres frem og tilbage (se side 53), men trods dette tiltag er de ikke egnede til højfrekvens.

De overflademonterede modstande fastholdes på kortet af loddetinnet. De må maksimalt belastes med 0,1 W da varmen især afledes til printkortet som kan blive skævvredet og opvarmningen af de nærmeste komponenter kan påvirke kredsløbets funktion. Modstande med trånder afgiver især effekten til omgivelserne ved varmestråling og konvektion og ved brug af afstandstag kan højden over printkortet øges så det modtager mindre varme. Ved stor effektafsættelse skrues modstanden fast til et køleprofil som afgiver den afsatte effekt til omgivelserne.

1.1.1 Ohms lov

Enhver elektrisk leder har frie elektroner der kan påvirkes af det elektriske felt fra en ekstern spændingskilde så elektronerne flyttes fra den negative pol mod den positive. Der vil løbe en elektrisk strøm og styrken er givet ved antallet af elektroner der passerer ind i (og ud af) lederen indenfor en given tid. Strømretningen blev fastlagt inden videnskaben havde fundet frem til at strømmen blev båret af negativt ladede elektroner og blev defineret som positiv med start 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.



Figur 6 – En leder beskrives ved en stang med længden L og et areal i tværsnittet på A . Det er uden betydning om tværsnittet er cirkulært, rektangulært eller andet.

Georg Simon Ohm formulerede 1827 en lov om denne sammenhæng og den udgør "grundloven" for elektricitet og elektronik. Spændingen U er den kraftkilde der bringer en strøm I af elektroner til at bevæge sig i en given retning med modstanden R som proportionalitet.

² Se fx <http://torean.dk/artikel/surge/surgetransient.html>.

1.1.2 Beregning af modstand

Forskellige materialer har forskellig mængde af frie elektroner per rumfang og de vil derfor lede en individuel strømstyrke når der lægges en spænding over enderne. Det betegnes med en konstant for det enkelte materiale kaldet for den specifikke modstand. En isolator har ingen frie elektroner og vil derfor ikke kunne lede en elektrisk strøm. For en leder vil den elektriske feltstyrke accelerere elektronerne op i fart, men ved kollision med de fastsiddende atomer og fejl i krystallet afgives der bevægelsesenergi i form af varme så hastigheden sænkes til en konstant gennemsnitlig værdi der er proportional med den eksterne spænding.

Modstandsværdien R beregnes fra længden L og arealet A af emnet samt materialets specifikke modstand ρ der også kaldes den specifikke resistivitet (electrical *resistivity*) med nogle vejledende værdier i tabellen herunder. De viste værdier gælder ved temperaturen 20°C .³

Materiale	ρ	α
Sølv	$15,9 \cdot 10^{-9} \Omega\text{m}$	$0,0038/^\circ\text{C}$
Kobber	$16,8 \cdot 10^{-9} \Omega\text{m}$	$0,0039/^\circ\text{C}$
Aluminium	$28,2 \cdot 10^{-9} \Omega\text{m}$	$0,0039/^\circ\text{C}$
Jern	$100 \cdot 10^{-9} \Omega\text{m}$	$0,005/^\circ\text{C}$
Platin	$106 \cdot 10^{-9} \Omega\text{m}$	$0,0039/^\circ\text{C}$
Konstantan	$490 \cdot 10^{-9} \Omega\text{m}$	$8 \cdot 10^{-6}/^\circ\text{C}$
Havvand	$0,2 \Omega\text{m}$	(ikke oplyst)
Kul (amorft)	$0,0006 \Omega\text{m}$	$-0,0005/^\circ\text{C}$
Germanium	$0,46 \Omega\text{m}$	$-0,048/^\circ\text{C}$
Silicium	$640 \Omega\text{m}$	$-0,075/^\circ\text{C}$

Modstandsværdi ved $\theta_0 = 0^\circ\text{C}$:

$$R_0 = \rho \frac{L}{A}$$

Modstandsværdi ved temperaturen θ :

$$R(\theta) = \rho \frac{L}{A} [1 + \alpha(\theta - \theta_0)]$$

Serie- og parallelforbindelse:

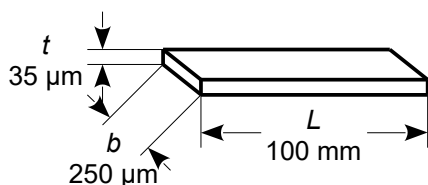
$$R_{\text{Serie}} = \rho \frac{L_1 + L_2}{A} = R_1 + R_2$$

$$R_{\text{Parallel}} = \rho \frac{L}{A_1 + A_2} = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}$$

Modstandens variation med temperaturen θ angives ved materialets temperaturkoefficient α og temperaturens afstand $(\theta - \theta_0)$ fra den temperatur θ_0 tabellen er angivet ved.

1.1.3 Eksempel

De elektriske forbindelser på et printkort laves af en tynd film af kobber med det ønskede mønster af baner som forbinder de elektroniske komponenter. Filmen er cirka $35 \mu\text{m}$ tyk og det er normalt med en bredde på $0,25 \text{ mm}$. Modstanden bliver $R_0 = 0,18 \Omega$ for $L = 100 \text{ mm}$ længde.



Modstandsværdi ved 20°C :

$$R_0 = \rho \frac{L}{A} = 15,9 \cdot 10^{-9} \Omega\text{m} \frac{100 \text{ mm}}{8,75 \cdot 10^{-9} \text{ m}^2} = 0,18 \Omega$$

$$A = tb = (35 \mu\text{m})(0,25 \text{ mm}) = 8,75 \cdot 10^{-9} \text{ m}^2$$

Korrektion for 50°C :

$$1 + \alpha(\theta - \theta_0) = 1 + (0,0039/^\circ\text{C}) \cdot (50^\circ\text{C} - 20^\circ\text{C}) = 1,117$$

Modstandsværdi ved 25°C :

$$R_{25} = R_0 \cdot 1,0975 = 0,20 \Omega$$

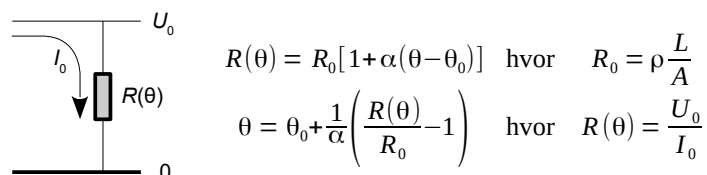
Figur 7 – Modstandsværdien beregnes for en kobberleder på et printkort.

Værdien gælder ved en temperatur på 20°C og den stiger med 11,7 % ved 50°C . Korrektionen på cirka $0,4 \%/^\circ\text{C}$ betyder at en elektrisk leders modstandsværdi er funktion af temperaturen og den kan derfor fungere som et termometer. Det er emnet for næste afsnit.

³ http://en.wikipedia.org/wiki/Electrical_resistivity_and_conductivity.

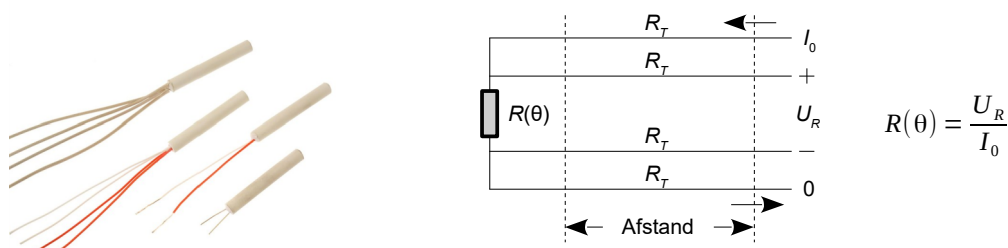
1.1.4 Platintermometer

Modstanden af en metaltråd stiger med temperaturen og det kan anvendes som termometer ved at løse ligningen $R(\theta)$ for temperaturen θ . En kendt strøm I_0 (typisk 1 mA) føres gennem metaltråden og spændingen U_0 måles med et voltmeter eller en mikroprocessors A/D konverter. For at bruge termometret skal man kende modstanden R_0 ved en referencetemperatur θ_0 og koefficienten α der fremgår af tabellen på forrige side.



Figur 8 – Et termometer dannet ved måling af en metaltråds modstand.

Modstanden varierer omtrent lineært med temperaturen. Kobber kan kun anvendes op til 150°C på grund af oxidation hvorimod platin kan anvendes fra –200°C til 850°C. For at beskytte metaltråden mod mekanisk overlast benytter man de indkapslede PT100 og PT1000 sensorer på henholdsvis 100 Ω og 1000 Ω ved 0°C. Deres modstandsværdi er tabellagt i 1°C trin og værdier mellem den enkelte trin kan findes ved interpolation.⁴



Figur 9 – Billede af PT100 temperaturfølere og en 4-tråds måling af lav modstandsværdi.

Måling af temperatur sker normalt et stykke væk fra måleapparatet så der vil være forholdsvis lange ledninger til platintermometret. Det betyder at den målte modstandsværdi består af PT100 temperaturfølerens modstand plus bidraget fra ledningerne. Man kan godt måle modstanden i de to kabler og trække bidraget fra, men temperaturen omkring ledningerne vil ikke være kendt så målingen bliver usikker. Derfor benyttes oftest en firetrådskobling hvor termometret forbindes til måleapparatet gennem fire ledninger. Det kaldes også for en Kelvinkobling. To af lederne bærer strømmen I_0 mens de to andre blot benyttes til at monitorere spændingen U_R uden en betydende strøm. Det tillader en stor afstand L mellem platintermometret og måleapparatet uden at det påvirker nøjagtigheden. Bemærk at måleapparatet skal måle en spændingsdifferens.⁵

Der optræder en mulighed for målefejl når metaltrådene skrues fast i måleapparatet hvor de fire metaltråde kommer i berøring med andre metaller. De elektrokemiske potentialer giver en svag spænding som adderes til måleværdien. For at undgå den fejlkilde kan man lade strømmen I_0 skifte retning og så måle igen. Et gennemsnit af to målinger vil fjerne de uønskede bidrag.

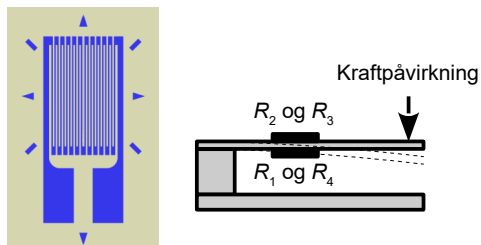
For at undgå indstråling af elektromagnetisk støj anbefales det at skærme lederne og skærmen forbindes til måleapparatets nul. Man skal undgå at forbinde skærmen til den apparatdel hvor temperaturen skal måles for ellers kan der dannes en sløjfe gennem en fælles returvej til jord der kan opfange elektromagnetiske støjsignaler.

⁴ <http://www.micropik.com/PDF/pt100.pdf>.

⁵ http://en.wikipedia.org/wiki/Resistance_thermometer.

1.1.5 Strain gauge

En metaltråd kan måle den mekaniske deformation af et emne hvis tråden limes fast på emnet og følger bevægelsen. Det kaldes en *strain gauge* modstand (der udtales "stræin gaitsj") som betyder en belastningsmåler, men da det er ikke en anerkendt betegnelse jeg vil bruge den engelske.⁶



Modstand af tråd:

$$R(\epsilon) = \rho \frac{(1+\epsilon)L}{A}$$

Areal vs. volumen:

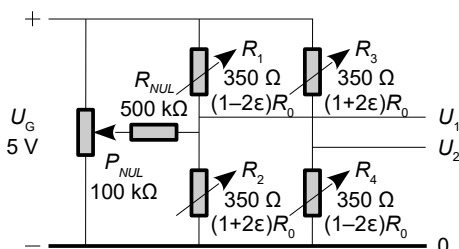
$$A = \frac{V}{(1+\epsilon)L}$$

$$\Rightarrow \text{Modstand: } R(\epsilon) \approx (1+2\epsilon)R_0$$

Figur 10 – Opbygningen af en vægt med strain gauge modstande hvor længden L ændres med ϵ . Den største modstandsvariation opnås i den lodrette retning med den viste opbygning. Arealet A reduceres så voluminet V af tråden forbliver konstant idet antallet af molekyler ikke ændres.⁷

Modstandsværdien er funktion af trådens længde og en strain gauge er ofte konstrueret ved at udfælde et tyndt metallag på en plastfolie som derefter kan limes fast til emnet. Modstanden vil ændres dobbelt så meget som længdeændringen da trådens tværsnit samtidigt reduceres. Det betyder at en relativ deformation med ϵ giver en modstandsændring på 2ϵ overfor værdien R_0 uden deformation ($\epsilon = 0$) hvis ϵ er lille ($\epsilon < 0,01$). En større ændring kan give en blivende deformation af tråden. Ved 1 % længdeændring vil en strain gauge på 350 Ω ændre modstandsværdien 7 Ω .

Den lave værdi af modstandsændringen kan måles ved en firetråds kobling som vist på forrige side eller ved en brokobling som vist herunder. De to grene arrangeres så spændingen fra den ene gren stiger (U_1 dannet af R_1 og R_2) mens den anden aftager (U_2 dannet af R_3 og R_4). Brokoblingen af fire ens strain gauge modstande betyder at temperaturfølsomheden kan balanceres ud.



$$U_1 = \frac{(1+2\epsilon)R_0}{2R_0} U_G \Rightarrow U_1 - U_2 = 2\epsilon U_G$$

$$U_2 = \frac{(1-2\epsilon)R_0}{2R_0} U_G$$

Figur 11 – Modstanden af en metaltråd stiger når den strækkes brøkdelen ϵ i længderetningen a og variationen forøges ved at arealet A mindskes hvis trådens samlede volumen V antages konstant.

Det vil normalt være påkrævet 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 til den ene gren fra et potentiometer anbragt mellem 5 V og nul og med en passende modstand til knudepunktet. Hver gren har en parallelmodstand på 175 Ω med de viste modstandsværdier så der skal bruges en strøm på $\pm 6 \mu A$ for at justere knudepunktet med ± 1 mV. Det kan opnås med et potentiometer på fx 100 k Ω og en seriemodstand på 500 k Ω . En alternativ metode er at måle spændingen med en mikroprocessors A/D konverter og efterfølgende trække værdien fra i software. En køkkenvægt bruger denne metode ved start samt når Tara-knappen aktiveres.

Udgangsspændingens værdi er funktion af den valgte effektforsyning. Med $U_G = 5$ V vil en ændring på ± 1 % i længderetningen svare til ± 100 mV fra broen. For en nøjagtig måling kræves en konstant spænding fra effektforsyningen der højest må variere ± 5 mV for at måle med $\pm 0,1$ % af fuld skala. For at måle med samme opløsning skal A/D konverteren være på mindst 12 bit.

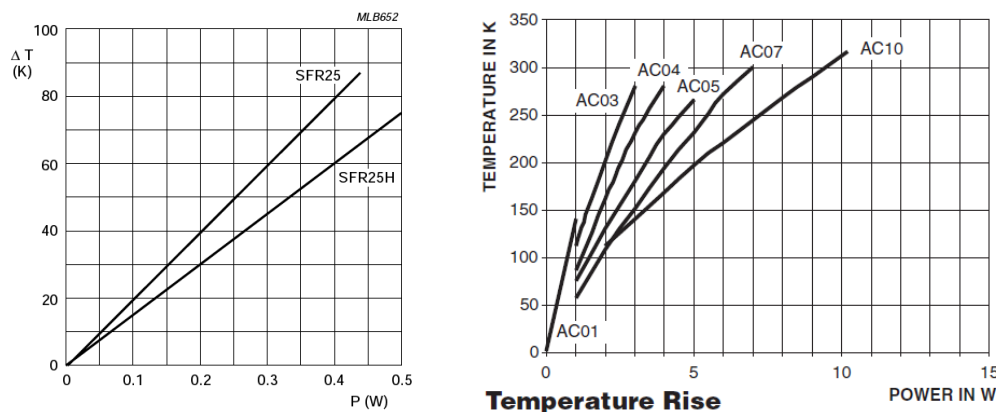
⁶ http://en.wikipedia.org/wiki/Strain_gauge.

⁷ Voluminet er ikke nødvendigvis konstant. Det dokumenteres i komponentens datablad som en afvigelse fra 2ϵ .

1.1.6 Effektafsættelse

Der afsættes energi i en modstand som følge af opbremsningen af elektronerne og det varmer modstanden op. Den afsatte effekt P er givet ved den afsatte energi per tidsenhed der beregnes som produktet af den aktuelle værdi af strøm I og spændingsfald U .

Ofte glemmer man at tage hensyn til den afsatte effekt i en komponent og det opdager man så ved at prototypen sender røgsignaler op når der over et tidsrum er afsat tilstrækkelig energi til at varme modstanden op til destruktion. Modstande til overflademontage er begrænset til en effekt på højst 0,1 W og lavere endnu for de mindste modstande. For modstande med trådbender (*leaded resistor*) er grænsen 0,5 W for metalfilm og 1 til 10 W for de mindre typer af trådviklede modstande.



Figur 12 – 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).

Et ofte overset problem er at modstandens overfladetemperatur er høj ved den tilladte belastning og det kan genere andre komponenter i nærheden ved at ændre på deres omgivelsestemperatur. En temperaturstigning vil også give en mekanisk belastning af et printkortet (skævvridning) på grund af dets udvidelseskoefficient. En belastning på 0,1 W ved en almindelig modstand med trådbender (SFR25) betyder en stigning på 20°C over omgivelsestemperaturen så modstanden vil opnå en driftstemperatur på 55°C der er nok til at man brænder fingrene ved berøring.

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 tilladelige effekt og levetiden reduceres. Indenfor mange produktkategorier kræves direkte en *derating* af arbejdsbetingelserne for at designet kan godkendes (fx militær og rumfart), men det er generelt en sund praksis aldrig at gå til kanten af hvad en komponent specificeres til at kunne klare uanset om det er spænding, strøm, effekt, temperatur eller mekaniske påvirkninger som vibration. Som en tommelfingerregel bør man holde sig under 50 % af det tilladte.⁸

Energi er effekt P afsat i løbet af tidsrummet T og energien kan også skrives som en ladning Q der føres gennem et spændingsfald U . Ladningen kan udtrykkes ved den strøm I der løber i et tidsrum og herfra følger formlerne for en effektafsættelse i en modstand.

$$\begin{aligned} W &= P T \\ W &= Q U \Rightarrow P = \frac{W}{T} = \frac{I T U}{T} = I U \Rightarrow \begin{aligned} P &= I^2 R \quad \text{da} \quad U = I R \\ P &= \frac{U^2}{R} \quad \text{da} \quad I = \frac{U}{R} \end{aligned} \\ Q &= I T \end{aligned}$$

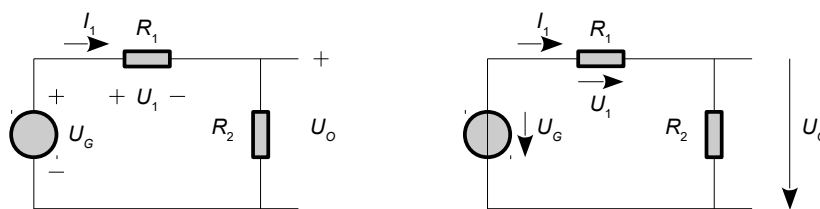
Når strømmen $I = 0,1$ A løber gennem en modstand på 50Ω så er spændingsfaldet $U = 5$ V og der afsættes en effekt på $P = I U = (0,1 \text{ A}) \cdot (5 \text{ V}) = 0,5$ W. En SFR25H modstand med trådbender vil opnå en driftstemperatur på 100°C ved 25°C omgivelsestemperatur.

8 Søg efter MTBF (*mean-time between failure*) og Arrhenius på nettet.

Note 1.

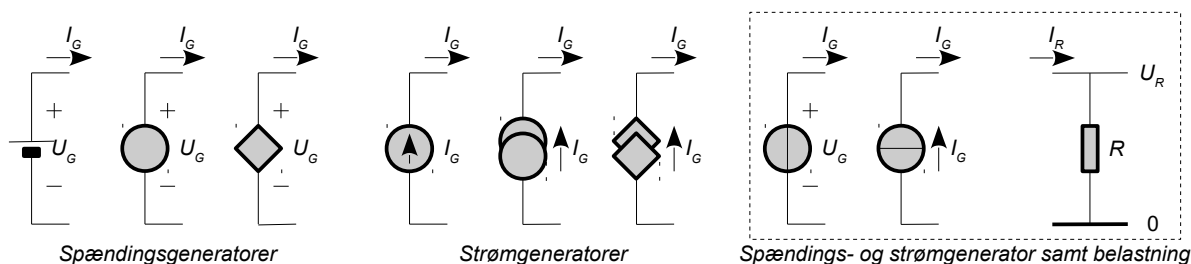
1.2 Konventioner og love

Praktisk brug af Ohms lov kræver en definition af *positiv retning* for spænding og strøm. Det er vedtaget at strømretningen er positiv når strømmen løber fra spændingskildens plus til minus.⁹



Figur 13 – 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 to eksempler. Til venstre ses den notation der er udbredt i engelsksproget 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 som IEC vedtog som international standard i 1970 og som mødes i tysksproget litteratur.¹⁰



Figur 14 – 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 lille 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 et signal et eller andet sted i kredsløbet. Tilsvarende er der flere bud på strømkilden som vist ved de midterste tre symboler. Yderst til højre vises de symboler som benyttes i denne bog.

En kilde noteres i denne bog med en cirkel. Ved en spændingskilde forbinder strengen de to poler som indikation af en lav indre modstand (0Ω). Ved en strømkilde indikerer strengen en afbrydelse mellem de to poler og derfor en høj indre modstand ($\infty \Omega$). Styrede kilder angives med relationen vist i betegnelsen for generatoren. En spænding der ikke er noteret med plus og minus regnes for positiv overfor et fælles nul-potentiale: 0, nul, stel, GND, ground og chassis.

Der er også en variation i litteraturen med hensyn til symbolet for spænding. Det internationale enhedssystem SI angiver symbolerne E , U og V for spænding. Se også appendiks.

1.2.1 Passiv fortegnskonvention

Strømmen fra en kilde løber fra minus til plus så afgiven effekt bliver negativ (afladning af batteri). Strømmen ved en belastning løber fra plus til minus så afsat effekt bliver positiv (opvarmning).

⁹ Det er modsat elektronens faktiske bevægelsesretning, men det har ingen praktisk betydning. Enkelte lærebøger har forsøgt at rette op på fejlen, men det har ikke slået an for det er svært at ændre på indarbejdede standarder.

¹⁰ Der findes endnu en notation hvor spændingen angives med pile i modsat retning. Den blev jeg selv undervist i omkring 1980, men jeg er ude af stand til at indse dens fordele.

1.2.2 Kirchhoffs love

En god udvidelse af Ohms lov er Kirchhoffs love der er et særdeles effektivt værktøj for at skabe en fast procedure for at angribe et problem og med lidt øvelse vil analysen føre sikkert i mål.

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 lignedes 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 der også kaldes KCL for *Kirchhoffs current law*.

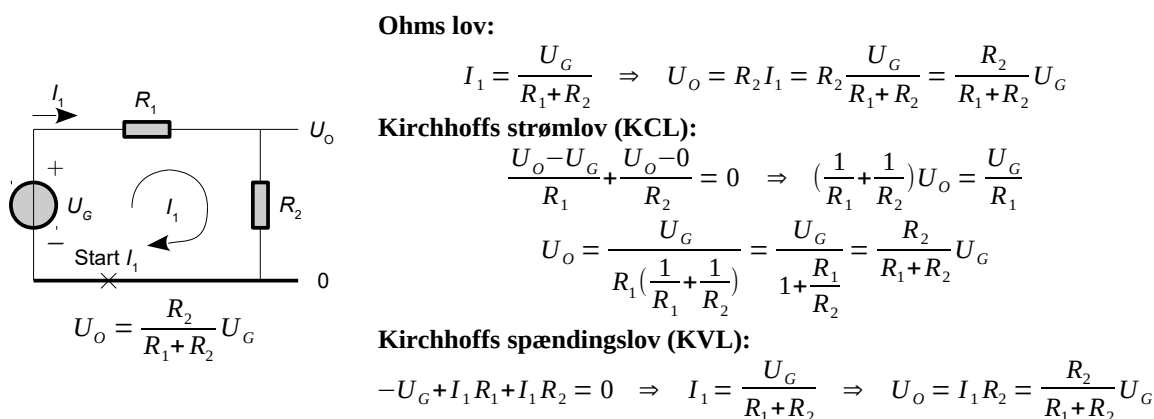
$$\text{Kirchhoffs strømlov (KCL): } \sum_{n=1}^N I_n = 0$$

Ved en maske, løkke eller sløjfe følger man en tur rundt i kredsløbet langs med de serieforbundne 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 der også kaldes KVL for *Kirchhoffs voltage law*.

$$\text{Kirchhoffs spændingslov (KVL): } \sum_{n=1}^N U_n = 0$$

1.2.3 Eksempel med spændingsdeler

Opstillingen herunder kaldes en spændingsdeler og den er tit anvendt i elektronikken. Både Ohms lov og Kirchhoffs to love skal benyttes for at finde udgangsspændingen.



Figur 15 – Analyse af en spændingsdeler.

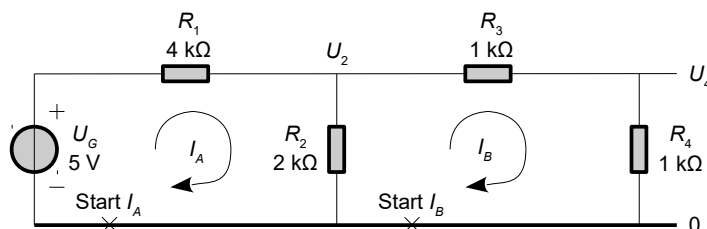
Ohms lov udnytter at modstandene R_1 og R_2 er i serie (se side 10) og at spændingen U_G driver strømmen I_1 gennem dem. Spændingen over R_2 bestemmes ligeledes fra Ohms lov.

Kirchhoffs strømlov (KCL) vises her med strøm væk fra knuden ved U_O . Ligningen løses ved at samle led med fælles variabel (U_O og U_G) og derefter isolere det eftersøgte led.

Kirchhoffs spændingslov (KVL) bruges ved at vælge en positiv strømretning for strømmen I_1 samt et udgangspunkt vist ved et kryds. Man følger den valgte omløbsretning og noterer spændingsfaldene indtil man er tilbage ved udgangspunktet. Spændingsfald noteres med fortegn så generatoren er $-U_G$ og modstandene er $R_1 I_1$ og $R_2 I_1$ i strømmens retning.

1.2.4 Større netværk

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 simuleringsværktøj af SPICE typen selv om det kun giver resultatets talværdi og ikke et analytisk udtryk.



Figur 16 – 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 repræsentere ganske mange systemer og ikke kun et elektroniske kredsløb. Her bestemmes strømmen I_B i modstanden R_4 ved maskeloven.

Der opstilles en ligning 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.

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

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 .

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

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.

1.2.5 Superposition

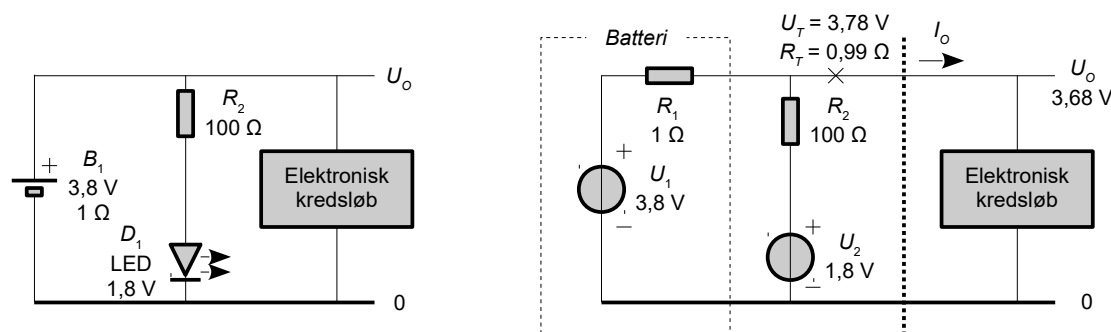
En af de vigtigste regler indenfor analyse er superposition der oversat til almindelig dansk betyder at forskellige bidrag ikke gensidigt påvirker hinanden. Reglen har sin bund i en antagelse om linearitet der er et bærende fundament indenfor naturvidenskab.

Linearitet betyder at hvis en påvirkning giver en vis respons så vil en fordobling af påvirkningen fordoble responsen og hvis påvirkningen reduceres til nul så vil responsen også blive nul. Det kan lyde ret trivielt 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. Addition af magnetiske felter i vakuum opfylder normalt kravet, men ved et magnetiserbart materiale som fx jern er lineariteten kompromitteret på grund af mætning og remanens. Indenfor elektronik antages alle komponenter at opføre sig lineært. Man ignorerer ganske enkelt at alle komponentens værdier er funktion af strøm og spænding, af temperaturen, af mekanisk påvirkning og 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 og det er en direkte konsekvens af linearitet. 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.

1.2.6 Eksempel på superposition

Superposition skal illustreres for kredsløbet herunder hvor batteriet B_1 på 3,8 V med $1\ \Omega$ som indre modstand skal drive en lysdiode (LED) der skal have 20 mA og et elektronisk kredsløb der trækker 100 mA. Opgaven er at bestemme spændingen U_o til kredsløbet.



Figur 17 – Et eksempel med et Li-ion batteri, der driver en lysdiode og et elektronisk kredsløb. Til højre vises en model af kredsløbet hvor dioden repræsenteres af en ideel spændingskilde.

Batteriet beskrives her ved en spændingskilde på $U_1 = 3,8\text{ V}$ og en indre modstand på $R_1 = 1\ \Omega$. Spændingen over lysdioden regnes for konstant på $U_2 = 1,8\text{ V}$ når dioden lyser. Modstanden R_2 har til formål at give lysdioden en strøm på omtrent 20 mA. I første del af analysen betragtes kun den del af kredsløbet der ligger til venstre for det elektroniske kredsløb. Det udnytter en antagelse af linearitet så belastningen fra kredsløbet kommer i betragtning senere og vi har derfor $I_o = 0$.

Kredsløbet kan opfattes som en spændingskilde på $U_T = 3,78\text{ V}$ med en indre modstand givet ved de parallellforbundne modstande R_1 og R_2 på $R_T = 0,99\ \Omega$. Udgangen skal levere strømmen I_o til belastningen så tomgangsspændingen U_T aftager med $(0,99\ \Omega)(0,1\text{ A}) = 0,099\text{ V}$ til 3,68 V.

Der er to spændingskilder U_1 og U_2 som giver spændingen U_T ved superposition af kilderne. Det betyder at første bidrag findes ved at nulstille U_2 og beregne bidraget fra U_1 og derefter findes andet bidrag ved at nulstille U_1 og beregne bidraget fra U_2 . Ved at udnytte antagelsen om linearitet kan de to bidrag blot adderes til tomgangsspændingen U_T der angiver spændingen fra knuden ved R_1 og R_2 (vist med et kryds) hvis der ikke løber strøm i udgangen ($I_o = 0$).

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

Den indre modstand fra netværket findes ved at kortslutte udgangen hvorved der vil løbe en strøm I_T og den indre modstand kan beregnes efter Ohms lov som tomgangsspændingen delt med kortslutningsstrømmen. Resultatet er de to modstande R_1 og R_2 i parallel.

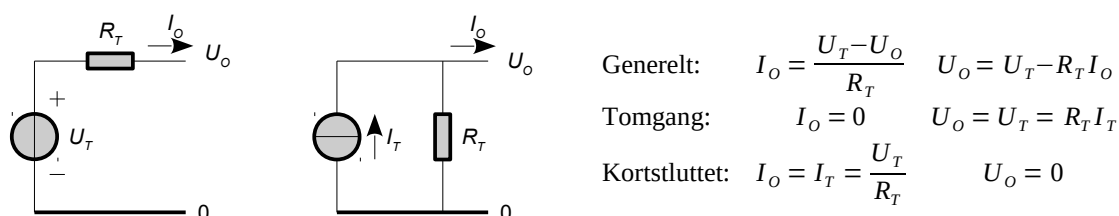
$$\begin{aligned} I_1 &= \frac{U_1}{R_1} \quad \text{og} \quad I_2 = \frac{U_2}{R_2} \Rightarrow I_T = I_1 + I_2 = \frac{U_1 R_2 + U_2 R_1}{R_1 R_2} \\ R_T &= \frac{U_T}{I_T} = \frac{\frac{R_2 U_1 + R_1 U_2}{R_1 + R_2}}{\frac{U_1 R_2 + U_2 R_1}{R_1 R_2}} = \frac{R_2 U_1 + R_1 U_2}{U_1 R_2 + U_2 R_1} = \frac{R_1 R_2}{R_1 + R_2} \end{aligned}$$

Strømmen til belastningen løber gennem R_T og sænker udgangsspændingen.

$$U_o = U_T - R_T I_o$$

1.2.7 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 oprindeligt Hermann von Helmholtz (1853), men det betegnes normalt et Thevenin ækvivalent efter Leon Charles Thevenin som genfandt relationen i 1883. Et Thevenin ækvivalent er således en spændingskilde U_T i serie med en modstand R_T der også kaldes for den indre modstand. Det er en særdeles populær model for batterier, sensorer og elektroniske kredsløb som fx effektforstærkere.



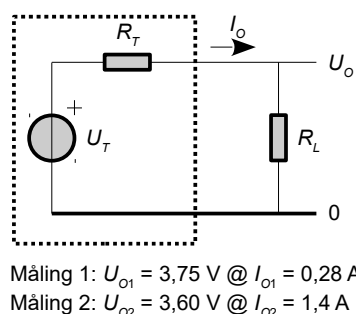
Figur 18 – To modeller for ethvert lineært elektrisk kredsløb. Thevenin ækvivalentet til venstre og det tilsvarende Norton ækvivalent til højre. Første ligning giver udgangsstrømmen og -spændingen i det generelle tilfælde. Anden ligning giver udgangsspændingen når udgangen er afbrudt og der ikke løber en strøm. Sidste ligning 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 sig frem til 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 for Mayer-Norton ækvivalentet eller blot et Norton ækvivalent. Det kan benyttes til at forenkle analysearbejdet.

1.2.8 Eksempel med batteri

Kredsløbets tomgangsspænding U_T og indre modstand R_T kan beregnes fra to målinger. Herunder er data hentet fra producentens datablad for et Li-ion batteri ved to oplyste belastningssituationer. Et fysisk kredsløb er sjældent lineært så beregningen bør udføres ved et strømniveau der ligger nær ved den påtænkte anvendelse. Den geometriske middelværdi af I_{O1} og I_{O2} giver $I_O = 0,6 \text{ A}$ hvor den beregnede udgangsspænding bliver $U_O = 3,70 \text{ V}$.¹¹



Udgangsspænding U_O versus I_O :

$$\text{Måling 1: } U_{O1} = U_T - R_T I_{O1}$$

$$\text{Måling 2: } U_{O2} = U_T - R_T I_{O2}$$

Subtraktion:

$$U_{O1} - U_{O2} = R_T (I_{O2} - I_{O1})$$

Indre modstand og tomgangsspænding:

$$R_T = \frac{U_{O1} - U_{O2}}{I_{O2} - I_{O1}} = \frac{3,75 \text{ V} - 3,60 \text{ V}}{1,4 \text{ A} - 0,28 \text{ A}} = 0,13 \Omega$$

$$U_T = U_{O1} + R_T I_{O1} = 3,75 \text{ V} + (0,13 \Omega) \cdot (0,28 \text{ A}) = 3,79 \text{ V}$$

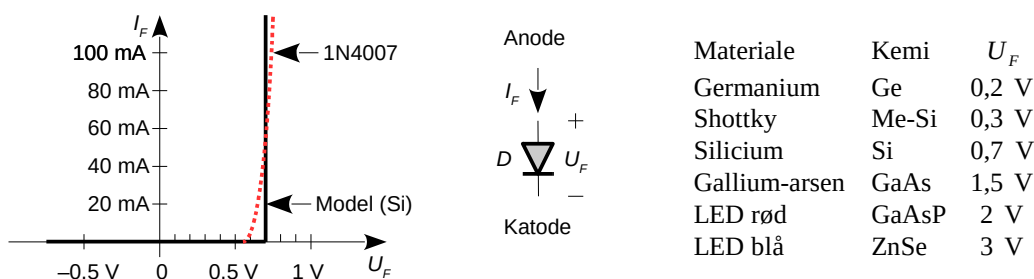
Figur 19 – Den indre modstand R_T og for et Li-ion batteri kan bestemmes ved at foretage to målinger af udgangsspændingen U_O ved to forskellige belastningsstrømme I_O .

Kredsløbets udgangsspænding U_O er funktion af belastningsmodstanden R_L hvor strømmen I_O kan beregnes fra Ohms lov. Der opskrives to ligninger for U_O som funktion af I_O (indeks 1 og 2). De to ligninger trækkes fra hinanden for at eliminere U_T der giver en formel for R_T . Herefter kan U_T beregnes ved at omrokere den ene af ligningerne.

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

1.3 Diode – D

En diode kan lede strømmen i den ene retning mens den spærrer i den anden så den kan ensrette en vekselspænding hvilket finder anvendelse i radiomodtagere, måleudstyr og effektelektronik. Den ensrettende virkning blev opdaget 1874 af Ferdinand Braun ved en systematisk undersøgelse af krystallers elektriske egenskaber, men teorien for dioden forblev ukendt indtil opdagelsen af transistoren nødvendiggjorde en forklaring som vises i femte kapitel. Navnet "diode" er inspireret af græsk hvor *di* betyder 2 og *ode* betyder vej for dioden har netop to veje ind til krystallet.^{12,13}



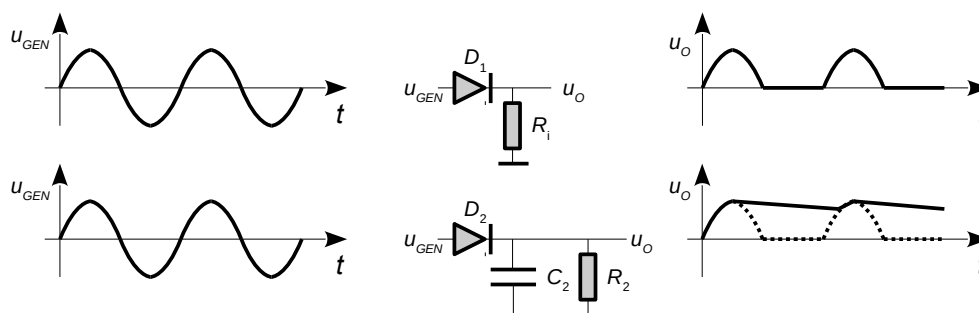
Figur 20 – Diodens idealiserede karakteristik (sort) sammenlignet med en reel diode (stiplet), dens symbol og vejledende værdier for ledespændingsfaldet for udvalgte teknologier.

Diodens symbol viser den vej strømmen kan løbe og det skal sandsynligvis også vise hvordan de første dioder blev fremstillet ved en spids metaltråd der blev presset mod et krystal.

- **I lederretningen** ligger der et spændingsfald U_F over dioden som afhænger af materialet og strømstyrken med nogle vejledende talværdier i tabellen. De fleste kender værdien 0,7 V for en diode af silicium. En Shottky-diode er opbygget af metal mod silicium og har et lavere spændingsfald. En lysdiode kaldes også LED (*Light-Emitting Diode*) og har højere værdi af spændingsfaldet da der kræves mere energi for at lysets bølgelængde bliver synlig.
- **I spærreretningen** løber der en minimal strøm gennem dioden. Den er normalt så ringe at man blot kan ignorere strømmen og regne dioden for en afbrydelse. Diodens lækstrøm I_R (reverse current eller leakage current) varierer fra 1 pA til mere end 1 μ A.

1.3.1 Ensretning

Indskydes en diode i et kredsløb vil den kun tillade strømmen at løbe i én retning så et signal der oscillerer omkring nul vil blive beskåret til kun at bestå af den positive del hvis dioden vender som vist ved D_1 i illustrationen herunder. Modstanden R_1 er nødvendig for at optage den svage strøm i spærreretningen så udgangsspændingen bliver tæt på nul når dioden spærrer og en kondensator kan inkluderes for at fastholde topspændingen.



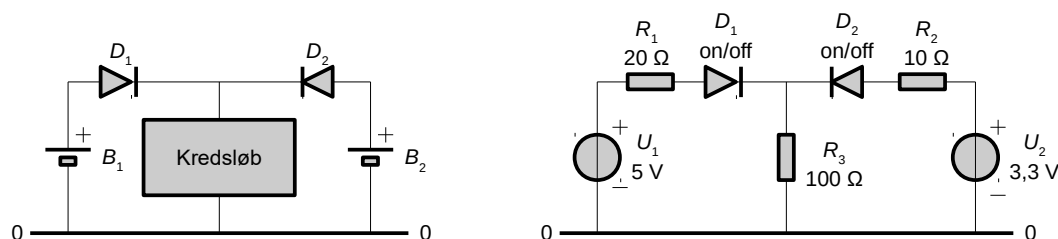
Figur 21 – Dioden tillader kun strømmen at løbe i den ene retning og spærrer i den anden. Med en kondensator kan spidsspændingen fastholdes i et kort tidsrum.

¹² <http://www.computerhistory.org/siliconengine/>.

¹³ <https://en.wikipedia.org/wiki/Diode#History>.

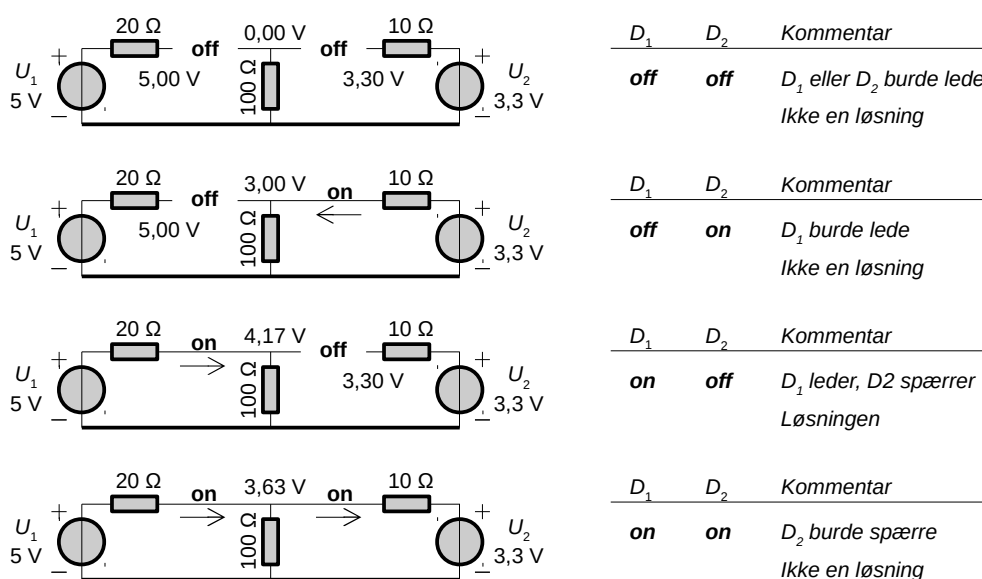
1.3.2 Kredsløbsanalyse

Når dioden indgår i et kredsløb skal man først hitte ud af om den er ledende eller spærrende for at kunne analysere kredsløbet. En metode er at dele kredsløbet op i mulighederne og så fravælge de kredsløb hvor der er en konflikt. Med to dioder er der i alt fire muligheder hvor de tre er i konflikt og kun én mulighed giver løsningen. Princippet skal vises ved en klassisk situation hvor et elektronisk kredsløb effektforsynes fra to batterier (B_1 og B_2) gennem hver sin diode (D_1 og D_2). Opgaven er at finde det batteri der driver kredsløbet. Batterierne beskrives ved tomgangsspændingen (U_1 og U_2) og den indre modstand (R_1 og R_2) og kredsløbet repræsenteres ved en modstand (R_3).



Figur 22 – Et kredsløb effektforsynes fra to batterier med en diode i hver gren.

For ikke at gøre analysen unødigt kompliceret repræsenteres dioderne af enten en kortslutning når dioden leder en strøm (dioden er **on**) eller en afbrydelse når der ikke løber en strøm (**off**). Hvis det senere viser sig at være for groft kan man vælge at forbedre analysen. Reglerne er at **on** skal have strømmen i diodens retning for at validere antagelsen og at **off** skal give en negativ spænding.

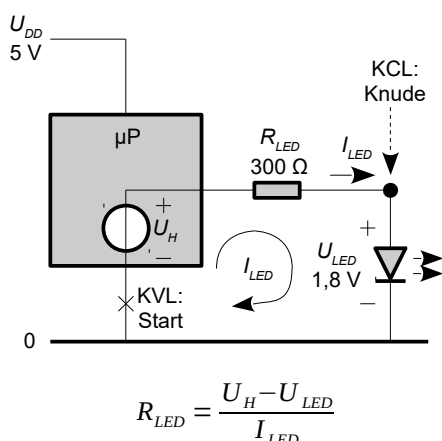


Figur 23 – Analyse af et kredsløb med to dioder repræsenteret ved de fire mulige kombinationer.

For den fundne løsning kan man nu tage hensyn til diodens spændingsfald så den forventede spænding over R_3 bliver 3,47 V ved en siliciumdiode (beregnet som 4,17 V minus 0,7 V).

1.3.3 Eksempel med lysdiode

En mikroprocessor skal drive en lysdiode og modstanden R_{LED} skal beregnes. En digital udgang kan ikke nå helt op til effektforsyningsens U_{DD} på 5 V når der skal leveres strøm til en belastning så det antages at spændingen er $U_H = 4,8$ V. Spændingsfaldet over en lysdiode er $U_{LED} = 1,8$ V ved en strøm på $I_{LED} = 10$ mA. Herunder vises tre måder for at finde formelen for modstanden.



Kredsløbsidentifikation

Spændingen U_H skal matche en serieforbindelse

$$U_H = R_{LED} I_{LED} + U_{LED} \Rightarrow R_{LED} = \frac{U_H - U_{LED}}{I_{LED}}$$

Kirchhoffs strømlov (KCL)

Positiv retning er væk fra knuden ved U_{LED}

$$\frac{U_{LED} - U_H}{R_{LED}} + I_{LED} = 0 \Rightarrow R_{LED} = \frac{U_H - U_{LED}}{I_{LED}}$$

Kirchhoffs spændingslov (KVL)

Positiv retning for I_{LED} er med uret

$$-U_H + R_{LED} I_{LED} + U_{LED} = 0 \Rightarrow R_{LED} = \frac{U_H - U_{LED}}{I_{LED}}$$

Figur 24 – En mikroprocessor skal tænde en lysdiode med HØJ på udgangen.

Kredsløbsidentifikation viser at udgangsspændingen U_H skal svare til summen af to bidrag: (1) spændingen over modstanden på $R_{LED} I_{LED}$ og (2) spændingen over dioden på U_{LED} .

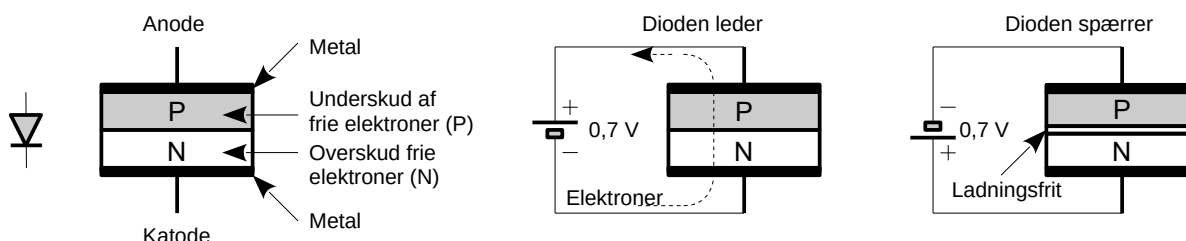
Kirchhoffs strømlov udføres ved knudepunktet ved lysdioden hvor spændingen er U_{LED} .

Strømmen regnes for positiv når den løber væk fra knudepunktet og den består af en strøm gennem R_{LED} og en strøm I_{LED} gennem lysdioden.

Kirchhoffs spændingslov vælges positiv strømretning som vist ved strømpilen for I_{LED} og det vælges at starte ved krydset. Spændingskilden U_H mødes ved minus og forlades ved plus så den noteres med negativt fortegn. Spændingen over modstanden er $R_{LED} I_{LED}$ og spændingen over lysdioden er 1,8 V.

1.3.4 Diodens opbygning

En diode består af to halvledere med forskellige egenskaber der er bragt i elektrisk kontakt. Det ene materiale har overskud af frie elektroner som kan lede en elektrisk strøm og det kaldes for N. Det andet materiale har underskud af valenselektroner og kaldes for P da en manglende elektron kan opfattes som et positivt "hul". Atomerne kan snuppe en elektron fra et naboatom og det kan stå for en strømtransport ved at elektronerne springer mellem hullerne i valenslaget.



Figur 25 – Diodens opbygning med et P og et N halvledermateriale.

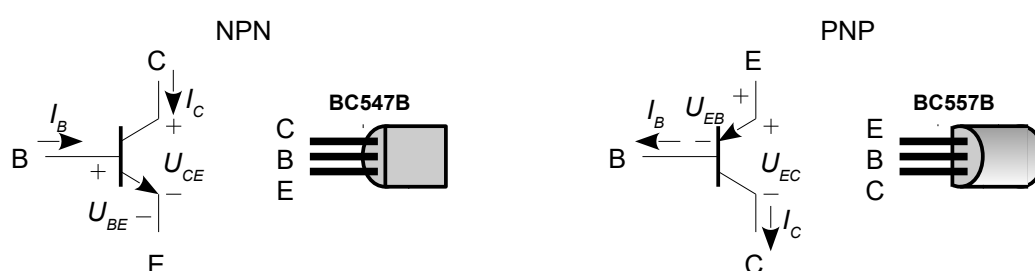
Et batteri med plus til P og minus til N skubber elektroner ind i N som rekombinerer med hullerne i P og føres tilbage til batteriet så *dioden leder en strøm*. Vendes batteriet om trækkes der elektroner ud af N og huller ud af P og det skaber et ladningsfrit område i grænselaget så *dioden spærrer*.

1.4 Transistor – T

Transistoren blev opfundet 1947 da Bardeen og Brattain fra Bell Labs arbejdede med et patent fra 1925. De presse tynde guldtråde ned mod et krystal af germanium og ved kort indbyrdes afstand opnåede de forstærkning af et signal. De havde konstrueret den bipolare transistor som Shockley året efter udviklede teorien bag. Herunder følger en introduktion til først den bipolare transistor der var elektronikkens foretrukne byggesten indtil 1980 hvor felt-effekt transistoren tog føringen og i dag er dominerende indenfor computere og anden teknologi.

1.4.1 Bipolar transistor

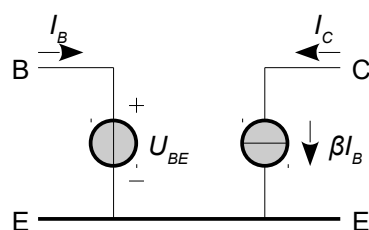
En bipolar transistor har tre tilledninger med hver sin funktion som vil blive uddybet i det følgende. Tilledningerne kaldes for emitter E (*emitter*), basis B (*base*) og kollektor C (*collector*). Den aktive del af en transistor fylder langt mindre end en kvadratmillimeter så den indstøbes i plastic med de tre tilledninger ført ud.



Figur 26 – De bipolare transistorer af typen NPN (venstre) og PNP (højre).

Transistoren leveres i to polariteter som kaldes for NPN og PNP hvor betegnelsen antyder at der er tre halvledere i en sandwich-konstruktion. En PN overgang er en diode så der er en diode mellem basis og emitter og den er vist som en pil. Der er også en diode fra basis til kollektor, men da basis er fælles for de to dioder og er meget tynd kan transistoren ikke opfattes som to dioder ryg mod ryg selv om mange bøger og referencer på internettet skriver dette. Så er det bedre at betegne den som "halvanden diode" for funktionen af kollektor-basis strækningen er helt og aldeles styret af forholdene ved basis-emitter dioden og det uddybes i bogens femte kapitel.

En populær model for en bipolar transistor er at strømmen I_C i kollektor er lig med strømmen I_B i basis ganget med strømforstærkningen β (beta) der også kaldes for h_{FE} i litteraturen. Talværdien for strømforstærkningen (*current gain*) oplyses altid i databladet.



Model af transistor:

$$I_C = \beta I_B$$

$$U_{BE} \approx 0,7 \text{ V}$$

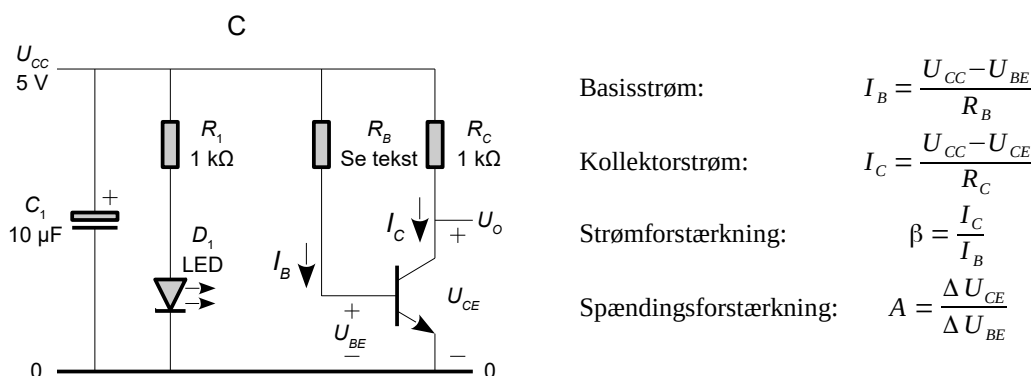
$\beta = 100 \dots 1000$ småsignaltransistorer
 $\beta = 10 \dots 100$ effekttransistorer

Figur 27 – Model af den bipolare transistor af typen NPN. Emmitter er normalt en fælles terminal for indgangssiden ved basis-emitter og udgangssiden ved kollektor-emitter. Transistoren ses dog også anvendt med basis eller kollektor som den fælles terminal mellem ind- og udgang.

NPN og PNP transistorerne er komplementære så fortegnet for strøm og spænding ombyttes. For begge typer løber der en strøm i kollektor-emitter som funktion af forholdet ved basis-emitter så en transistor kan styre en strøm i et kredsløb som følge af forholdene et andet sted i kredsløbet. Ved en NPN løber strømmen ind ved kollektor og ud ved emitter og det er omvendt for en PNP.

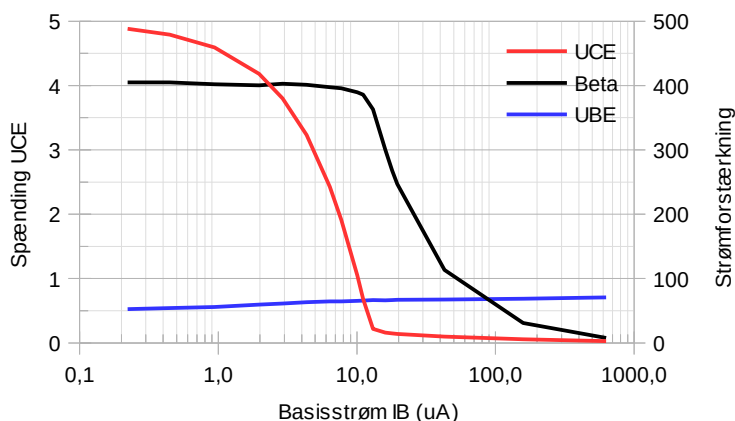
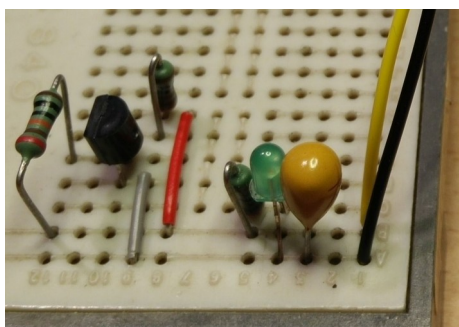
1.4.2 Måleopstilling

For at illustrere transistorens funktion blev følgende kredsløb bygget. Transistoren er af typen NPN og strømmen I_B til basis leveres via modstand R_B . Strømmen I_C i kollektor løber gennem R_C så udgangen trækkes ned ved stigende strømstyrke. Ved måleserien var R_B fra 6,8 k Ω til 20 M Ω og for hver værdi blev målt spændingen U_{BE} over basis-emitter og U_{CE} over kollektor-emitter samt effektforsyningens U_{CC} hvorefter strømforstærkningen β blev beregnet.



Figur 28 – Måleopstilling med en NPN transistor BC547B.

Yderst til venstre i diagrammet (og til højre i fotografiet herunder) vises tilslutningen til en 5 V effektforsyning (gul og sort ledning), dertil en tantal-elektrolyt kondensator C_1 for stabil funktion og en lysdiode D_1 for visuel indikation af de 5 V hvor modstand R_1 begrænser strømmen til 3,5 mA.



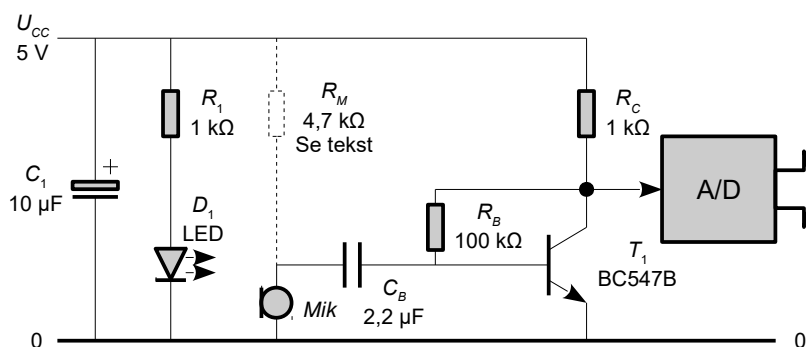
Figur 29 – Billede af måleopstillingen og måleresultaterne. Den vandrette akse er basisstrømmen I_B som indstilles af R_B og den lodrette akse til venstre viser U_{CE} (rød) og U_{BE} (blå). Den lodrette akse til højre viser strømforstærkningen β (sort).

Plottet viser spændingerne ved transistorens kollektor og basis samt strømforstærkningen. Det ses at strømforstærkningen er ret konstant på 400 for den anvendte transistor. Det passer godt med databladets opgivelse på 200 ... 450 ved $I_C = 2$ mA og 25°C. At strømforstærkningen ser ud til at aftage for I_B over 12 μ A skyldes at strømmen i kollektor nu er givet af modstanden R_C til $I_C = 5$ mA. Transistoren er næsten kortsluttet hvilket kaldes for mætning (*saturation*) og relationen $I_C = \beta I_B$ er ikke gyldig da U_{CE} ikke kan blive mindre end de målte 30 mV ved $I_B = 0,5$ mA.

Spændingen over basis-emitter er $U_{BE} = 0,648$ V ved $I_B = 8$ μ A hvor kollektor er på $U_{CE} = 1,8$ V og $U_{BE} = 0,656$ V ved $I_B = 10$ μ A hvor $U_{CE} = 1$ V. Det giver en spændingsforstærkning på $A = -100$ fra U_{BE} til U_{CE} . Transistoren er derfor i stand til at forstærke en svag AC signalspænding der føres ind på basis gennem en kondensator for at blokere for DC strømmen til basis.

1.4.3 Mikrofonforstærker

Transistoren kan forstærke et svagt signal fra en mikrofon op til et niveau der kan monitoreres med en A/D konverter. En dynamisk mikrofon tilsluttes til basis gennem kondensator C_B der blokerer for DC så transistoren får den planlagte strøm i basis. Med en elektret mikrofon skal den indbyggede felt-effekt transistor i kapslen have 0,5 mA tilført og det står modstanden R_M for.



Dimensionering:

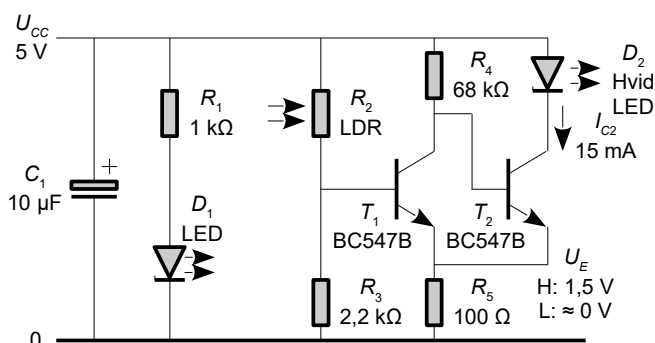
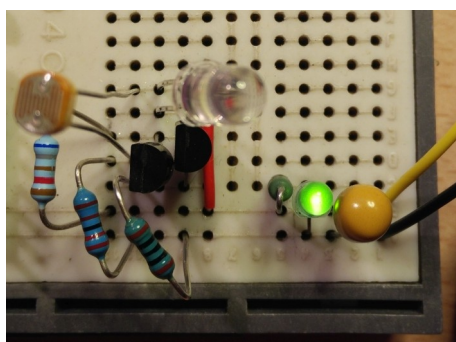
1. Antag en værdi af $U_{CE} = U_O$
2. Antag værdien: $U_{BE} = 0,7 \text{ V}$
3. Beregn: $I_B = \frac{U_{CE} - U_{BE}}{R_B}$
4. Beregn: $I_C = \beta I_B$
5. Beregn: $U_{CE} = U_{CC} - R_C I_C$
6. Gentag indtil U_{CE} er fundet

Figur 30 – En mikrofonns svage signal kan forstærkes op. Modstanden R_M skal kun monteres ved en elektret mikrofon hvor nul forbindes til den loddeterminale der er tilsluttet mikrofonens metalhus.

Modstanden R_B er her ført til kollektor frem for til effektforsyningen for at give en mere stabil værdi af udgangsspændingen end det tidligere kredsløb kan give. Selv om strømforstærkningen varierer fra 200 til 450 vil udgangen antage en sikker værdi i området 1,4 ... 2,2 V. Det forstærkede signal vil optræde som variationer omkring dette DC niveau der kan fjernes i mikroprocessoren.

1.4.4 Lysdetektor

En lysfølsom modstand LDR (*Light-Dependent Resistor*) ændrer sin modstand som funktion af det lys der falder på den. Ved stærkt lys er modstanden lav og ved svagt lys er modstanden høj. Her benyttes den til at slukke for transistor T_1 når lysniveauet bliver lavt hvorved transistor T_2 tænder en hvid lysdiode D_2 så der kan skabes lidt lys om aftenen på et ellers mørkelagt område.



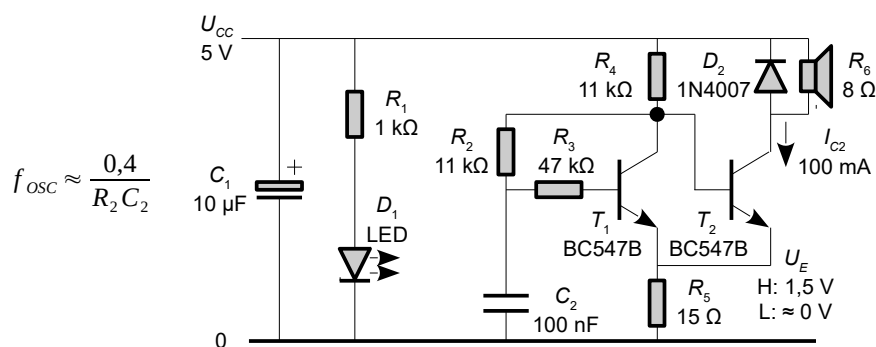
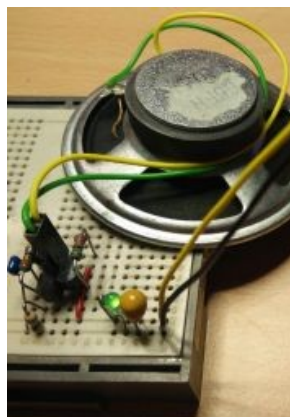
Figur 31 – En lysfølsom modstand reagerer på omgivelsernes lysniveau og tænder for en lysdiode når lysniveauet falder under et givet niveau.

De to transistorer er koblet gennem den fælles emitter for at give en stabil omskiftning mellem lys og mørke så kredsløbet ikke tænder og slukker blot lysniveauet ændres en lille smule. Når T_2 er OFF skal basis af T_1 ned under 0,7 V for at tænde for lysdioden og strømmen I_{C2} hæver den fælles emitter U_E til 1,5 V så basis af T_1 nu skal op på 2,2 V at slukke lysdioden igen.

Strømmen i lysdioden indstilles på cirka 15 mA af strømmen i R_4 på 40 µA når T_1 er OFF og dertil af strømforstærkningen β for T_2 på cirka 400.

1.4.5 Lydgiver

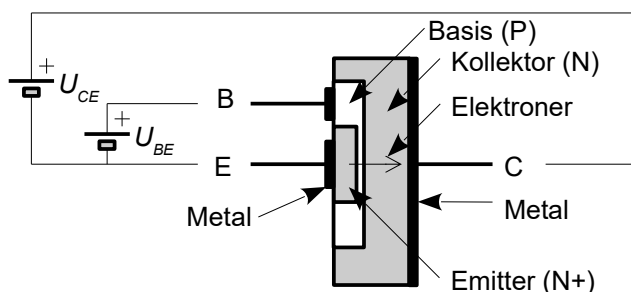
Opstillingen kan ændres til en alarm der hyler med en konstant frekvens på cirka 350 Hz med de viste komponentværdier. Frekvensen kan ændres ved R_2 der bør være fra 1 kΩ til 1 MΩ og ved at skifte kondensatoren. R_4 og R_5 er indstillet så strømmen i kollektor er på 100 mA.



Figur 32 – En oscillator driver en højttaler med en konstant hyletone.

1.4.6 Transistorens opbygning

En bipolar transistor består af tre lag af halvledere der kan arrangeres som NPN eller PNP. Her er vist opbygningen af en NPN transistor som BC547B. Halvlederen er set fra siden og vil ofte være rotationssymmetrisk. Kollektor er ved typer i metalhus fastgjort til huset og de to øvrige elektroder forbindes til metallagene der pådampes i fremstillingsprocessen og hvor ledningerne svejses til metallet ved en højfrekvent vibration der kortvarigt opvarmer tråden til smeltning.



Figur 33 – Opbygning af en NPN transistor som fx BC547B med et billede af en lignende transistor BC109C hvis metalhus er åbnet så den lille chip ses.¹⁴

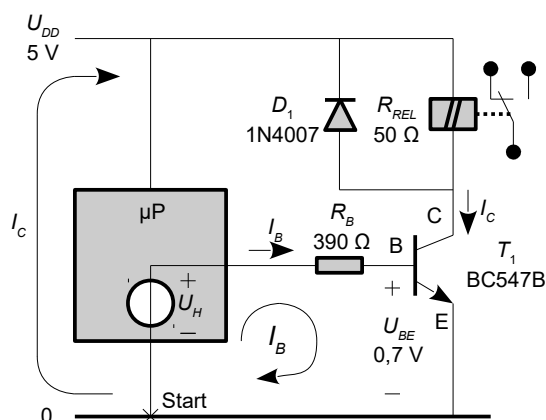
Ved en NPN transistor har emitteren mange frie elektroner (N+). Emitteren grænser op til et tyndt lag af en halvleder der har et underskud af elektroner og derved et overskud af positive huller (P). Kollektor har et overskud af elektroner, men færre end emitteren (N).

For at få elektronerne i emitter til at bevæge sig imod basis behøves en spændingskilde med plus til basis og minus til emitter med en værdi på omkring 0,7 V. Basis har mange huller i valenslaget og nogle af elektronerne vil "falde i et hul" hvilket kaldes rekombination og elektronen føres nu til basis som en strøm af valenselektroner. Det store flertal af elektroner vil dog fortsætte til kollektor da basislaget er så tyndt at kun lille procentdel af elektronerne når at rekombinere i den korte tid de er om at komme igennem basislaget. Strømmen i kollektor bliver derved mange gange større end strømmen i basis og det beskrives ved β .

¹⁴ <http://www.laserpointers.co.uk/shopcontent.asp?type=laserotech>.

1.4.7 Eksempel med relæ

Som fortsættelse af eksemplet fra side 22 vises hvordan en mikroprocessors udgang kan drive et relæ på 5 V med 50 Ω spolemodstand som er for hård en belastning for en typisk mikroprocessor. Transistor T_1 inkluderes for at drive strømmen i relæets spole da den kan trække en større strøm i kollektor end strømmen i basis. Når transistoren som her skal anvendes som en kontakt er det tilstrækkeligt med en strøm i basis på 10 % af strømmen i kollektor for at spændingsfaldet over kollektor-emitter bliver så tæt på nul at transistoren nærmest bliver en kortslutning. En transistor kræver dertil et spændingsfald fra basis til emitter på $U_{BE} = 0,7$ V når den er aktiv.



Positiv retning vælges med uret.

Strøm i kollektor:

$$-U_{DD} + R_{REL} I_C = 0 \Rightarrow I_C = \frac{U_{DD}}{R_{REL}} = 100 \text{ mA}$$

Strøm i basis (transistor ON):

$$I_B = \frac{I_C}{10} = 10 \text{ mA}$$

Modstand ved basis:

$$-U_H + R_B I_B + U_{BE} = 0 \Rightarrow R_B = \frac{U_H - U_{BE}}{I_B} = 410 \text{ } \Omega$$

Figur 34 – En mikroprocessor driver et relæ med en transistor som kontakt.

Når transistoren er ON er kollektor næsten kortsluttet til emitter så vi bruger $U_{CE} = 0$. Kirchhoffs spændingslov for sløjfen med relæet består derved af spændingskilden $-U_{DD}$ og relæets R_{REL} som giver strømmen I_C i transistorens kollektor.

For at transistoren går helt ON skal strømmen i basis I_B som en tommelfingerregel være 10 % af strømmen i kollektor. Mikroprocessoren driver I_B til transistoren gennem R_B og det kredsløb svarer til kredsløbet for lysdioden.

Modstanden i basis R_B beregnes til 410 Ω , men det er ikke en tilgængelig værdi. Modstande kan kun købes i fastlagte værdier som kaldes for E-række værdier så modstanden må nedjusteres til nærmeste tilgængelige værdi og det er 390 Ω .

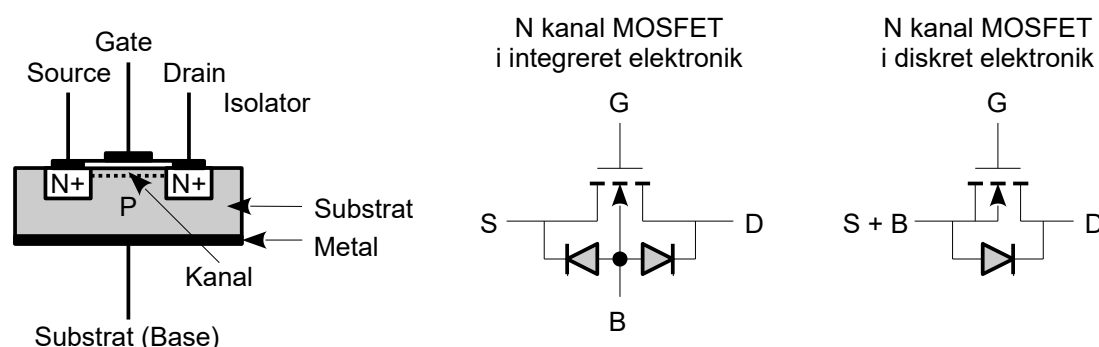
Der er ikke taget stilling til den transistor der skal benyttes selv om designet tager sit udgangspunkt i analysen af en BC547B fra den forrige side. Hovedparten af tilgængelige NPN transistorer kan forventes at fungere. Den største forskel vil ligge i strømforstærkningen, men når der er taget hensyn til den aktuelle transistors værdi bør den kunne fungere i opstillingen. Den skal naturligvis kunne tåle at trække en strøm på 100 mA.

Dioden D_1 over relæets spole er normalt i spærreretning, men beskytter transistoren mod de ødelæggende spændingsspidser når relæet skiftes fra ON til OFF (se side 52).

1.4.8 Felt-effekt transistor

Elektronik er i dag totalt domineret af felt-effekt transistoren der tog over efter 1970 og skubbede den bipolare transistor ud i skyggen. Idéen til felt-effekt transistoren skyldes Lilienfeld der i 1925 tog patent på den komponent vi i dag vil kalde en JFET (*Junction Field-Effect Transistor*) og som udviklede sig til den moderne MOSFET (*Metal-Oxide Silicon Felt-Effekt Transistor*).¹⁵

En MOSFET består af et substrat der i illustrationen er en P halvleder og dertil to mindre områder med N+ halvledere (source og drain) som er adskilt af en kort strækning der kaldes for en kanal. Over denne kanal er en elektrode placeret (gate), men elektrisk isoleret fra source, drain og substrat af et tyndt lag siliciumoxyd. Den tilsigtede funktion for transistoren er at lede en strøm af elektroner fra source til drain styret af spændingen på gate.



Figur 35 – Konstruktionen af en MOSFET, dens diagramsymbol og realisering af en diskret transistor med tre tilslutninger.

De to N+ halvledere og substratet danner to dioder der vender hver sin vej så der er ingen forbindelse mellem source og drain via P halvlederen. Det kan der ændres på ved at lægge en positiv spænding på gate i forhold til substratet. En positiv spænding vil skubbe substratets huller væk umiddelbart under gate og det skaber en laddningsfri zone. Hvis spændingen bliver stor nok trækkes elektroner ind i området fra enten source eller drain så kanalen skifter polaritet og der dannes en elektrisk kanal mellem source og drain. Kanals bredde styres af spændingen mellem gate og substrat og dermed styres modstandsværdien på samme måde som beskrevet side 9.

Når spændingen over source og drain øges vil strømme blive kraftigere, men ved tilstrækkelig høj spænding over transistoren vil den ikke længere være en spændingsstyret modstand. Funktionen går imod en spændingsstyret strømregulator hvor strømmen i transistoren kun er styret af spændingen over gate-source og ikke af spændingen over drain-source.

Den diskrete MOSFET har kun tre tilledninger så substratet er internt kortsluttet til den terminal der føres ud som source. Det betyder at dioden mellem drain og substrat nu ligger i parallel over transistorens source og drain og af samme årsag kan en diskret MOSFET ikke ombytte source og drain uden at dioden bringes i lederetning.

Den beskrevne komponent kaldes for en N kanal MOSFET da kanalen består af elektroner. Den komplementære P kanal MOSFET kan dannes af et N substrat med to P+ områder for source og drain. Der skal nu en negativ spænding til for at åbne kanalen. Det er relativt sjældent at se en P kanal MOSFET i diskret elektronik, men den benyttes flittigt i integreret elektronik og ikke mindst i de digitale CMOS kredse der for alvor vandt indpas efter 1980.

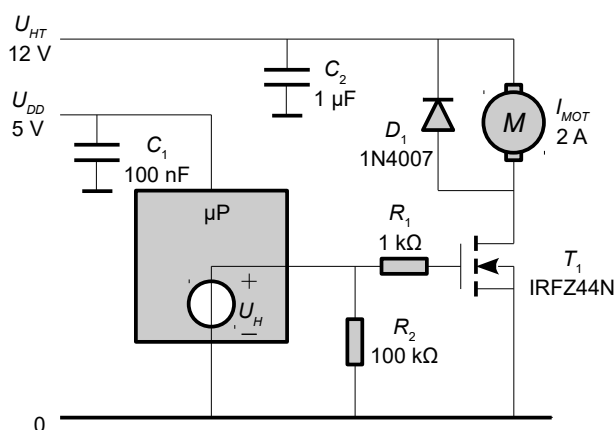
Som for dioden og den bipolare transistor er der en uddybende gennemgang i femte kapitel.

¹⁵ <http://www.computerhistory.org/siliconengine/field-effect-semiconductor-device-concepts-patented/>

1.4.9 Eksempel med en motor

Herunder vises en typisk situation hvor en MOSFET benyttes til at drive en mindre DC motor. Her er både spænding og strøm udenfor en mikroprocessors kapacitet.

Når mikroprocessorens udgang går høj bringes kanalen i T_1 til at lede og hvis transistorens ON modstand er tilstrækkelig lav vil den fungere som en kontakt der går ON og får motoren til at køre. Den viste transistor vil kunne klare nogle få ampere hvis mikroprocessoren giver 5 V som logisk høj, men hvis den kun giver 3,3 V er det ikke sikkert at transistoren overhovedet går ON.



Figur 36 – En MOSFET anvendt som kontakt for at styre en lille DC motor.

Dioden D_1 skal inkluderes for at fange de stærke spændingsspidser som en DC motor kan give når den kører og kondensator C_2 er nødvendig for at optage energien herfra. For at undgå udstråling af kraftig radiostøj bør C_2 monteres tæt på både T_1 og motoren. Kondensator C_1 er nødvendig for en stabilt fungerende mikroprocessor og det anbefales at følge databladet for den korrekte måde at afkoble dens effektforsyning.

Modstandene R_1 og R_2 er indskudt for at håndtere fejlsituationer. R_1 har til formål at beskytte den digitale udgang fra mikroprocessoren hvis der sker en fejl ved MOSFET og motor. En fejl der bringer gate op på den positive forsyning på 12 V vil øjeblikkeligt brænde mikroprocessoren af, men med R_1 indsat vil strømmen begrænses til cirka 5 mA som kan håndteres af det indbyggede beskyttelseskredsløb i mikroprocessoren.

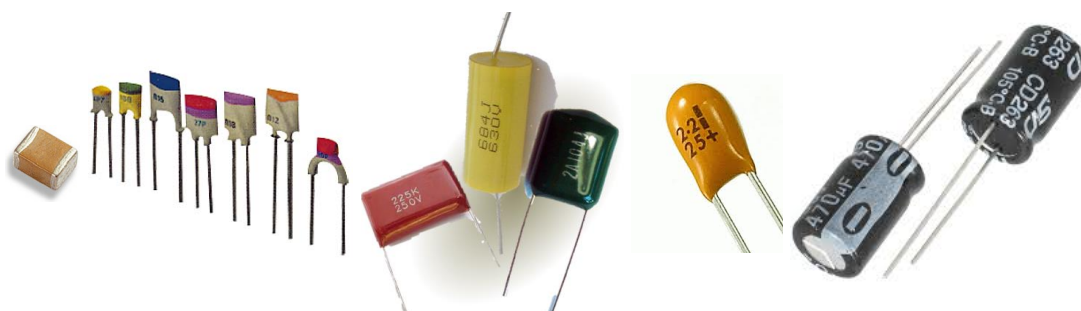
Modstand R_2 holder gate på 0 V og dermed transistoren OFF hvis mikroprocessorens udgang ikke er aktiv. Det vil være tilfældet efter tilslag af effektforsyningen hvor porten ikke er initialiseret samt hvis mikroprocessoren sættes i power down som fx ved en SLEEP kommando. I de situationer vil transistoren blot gå OFF og motoren kører ikke.

Modstand R_1 kan være problematisk hvis motoren skal styres med et PWM signal da den vil forlænge skiftetiden mellem ON-OFF og OFF-ON. Det skyldes at en MOSFET har en ret høj værdi af kapaciteten mellem gate og source. Hvis det er et problem kan man reducere værdien af modstanden, koble en kondensator på cirka 10 nF i parallel med modstanden eller helt udelade modstanden (den skal da kortsluttes).

1.5 Kondensator – C

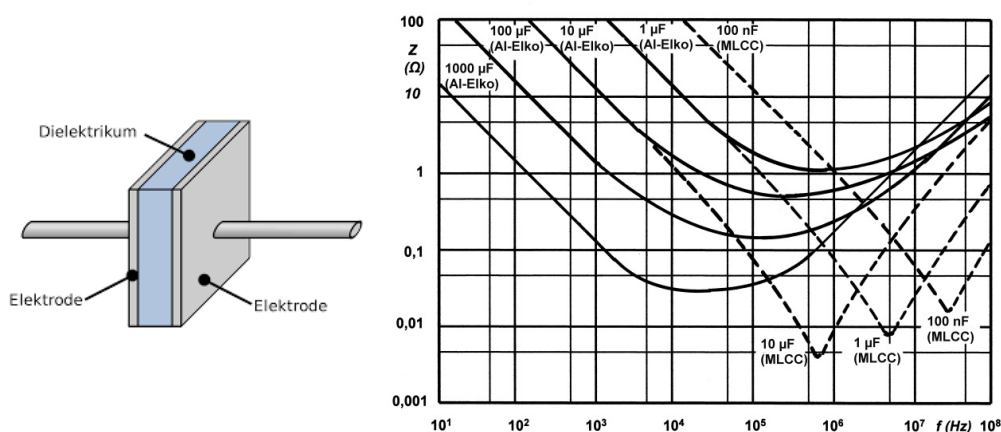
Enr kondensator kan opmagasinere energi i det elektriske felt mellem to plader der ligger tæt på hinanden kun adskilt af et tyndt lag af luft, keramik, plast eller andet. Energien kan tappes på et senere tidspunkt så der er en vis lighed med et genopladeligt batteri, men teorien er dog helt anderledes for et batteri lagrer energien kemisk hvor kondensatoren benytter et elektrisk felt.

Der er mange varianter som den følgende oversigt viser og forskellen ligger primært i det anvendte materiale mellem de to plader. Varianterne er keramiske kondensatorer, plastfilm kondensatorer og elektrolytiske kondensatorer. De to første tåler spænding i vilkårlig retning hvorimod den sidste er polariseret og skal vendes korrekt for ikke at blive ødelagt.¹⁶



Figur 37 – Til venstre keramiske kondensator for overflademontage (SMD) og med ledninger. Derefter plastfilm kondensatorer. Tilhøjre først en tantal elektrolyt og derefter et par aluminium elektrolyt kondensatorer. Bemærk at polariteten er markeret med et plus eller et minus.

Kondensatorer anvendes for blokering af langsomt varierende signaler, for transmission af hurtigt varierende signaler og som et lokalt energireservoir. I radiokredsløb anvendes kondensatoren med en spole i form af afstemte kredse (se side 39) der udvælger den ønskede radiostation.



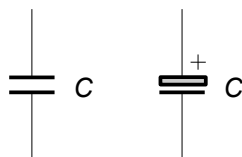
Figur 38 – Opbygning med to metalplader adskilt af et dielektrikum. Kurven til højre viser en kondensators modstandsværdi (impedans) som funktion af frekvensen. Al-Elko står for aluminium elektrolyt-kondensator og MLCC står for Multi-Layer Ceramic Capacitor (Wikipedia).

Den brede vifte af anvendelsesområder skyldes at kondensatoren opfører sig som en modstand overfor vekselstrøm hvor modstandsværdien aftager med stigende frekvens. Det betegnes som kondensatorens impedans Z (*impedance*). Ved jævnstrøm er kondensatoren afbrudt, ved lave frekvenser er impedansen høj og den aftager med stigende frekvens. Ved høje frekvenser når den et minimum ved den ækvivalente seriemodstand ESR (*equivalent series resistance*) og ved endnu højere frekvenser stiger impedansen på grund af selvinduktionen i kondensator og tilledning.

¹⁶ https://en.wikipedia.org/wiki/Capacitor_types.

1.5.1 Kondensatortyper

Der er mange måder at bygge kondensatorer på, men overordnet kan de placeres i to grupper for henholdsvis de polariserede og de ikke-polariserede. Førstnævnte kan kun tåle spænding med en given polaritet og må derfor ikke sættes forkert i kredsløbet hvorimod de ikke-polariserede har den samme funktion uanset om spændingen skifter fortegn.



Figur 39 – Kondensatorens elektriske symbol ligger tæt op ad den mekaniske opbygning med to plader i en vis afstand. Symbolet til venstre indikerer en symmetrisk kondensator der kan vende vilkårligt og typisk er med keramisk eller metalfilm som dielektrikum. Symbolet til højre indikerer en kondensator som skal have den rette polarisering og det gælder elektrolyt-kondensatorer.

Inden for hver af de to grupper er der igen en række underdelinger som skyldes det materiale der er mellem de to plader. Den mest enkle kondensator har luft mellem pladerne, men det benyttes sjældent da kapaciteten bliver lav. Normalt er der et keramisk materiale eller en plastfilm mellem pladerne. Det skyldes dels at det letter produktionen og dels fordi det øger kapaciteten. De mest populære typer er keramik og plastfilm der så igen kan opdeles i et system af undergrupper alt efter den anvendte type af keramik og plast.

De polariserede kondensatorer benytter en metaloxid som det isolerende materiale mellem de to plader. Den ene plade udgøres så af den tynde metalfolie mens den anden består af en elektrolyt. Det er en væske med frie ioner så strømmen kan ledes hen til kondensatoren. Fordelen ved denne konstruktion er muligheden for en meget høj kapacitet, men der er en halvlederfunktion mellem oxyden og elektrolytten som betyder at konstruktionen er en kondensator ved den ene polaritet mens den er en slags diode ved den modsatte polaritet. Det betyder at kondensatoren tager skade af at blive vendt forkert. Denne type kondensator har meget dårlige egenskaber ved højfrekvens og benyttes derfor til mindre kritiske funktioner som effektforsyning, energireservoir og afkobling. De anvendte metaller er aluminium og tantal.

1.5.2 Keramisk kondensator

En keramisk kondensator er opbygget med flere lag af keramik som dielektrikum og metalfilm for den elektriske forbindelse. Kapacitetsområdet dækker fra 1 pF til 10 μ F og det er en populær komponent i industrien da den egner sig for automatisk montering (*SMD, surface-mounted device*), den har lave tab og er generelt associeret med god kvalitet for pengene. Keramiske kondensatorer opdeles i klasser for nøjagtighed, temperaturområde og andre egenskaber.¹⁷

Klasse 1 (NP0) dækker fra 1 pF til 100 nF med god nøjagtighed ($\pm 2\%$) og en relativt lav værdi af temperaturkoefficienten (± 30 ppm/ $^{\circ}$ C).

Klasse 2 (X7R) dækker op til 10 μ F med ringere nøjagtighed. De kan desuden lide af mikrofoni der er en piezoelektrisk effekt hvor en mekanisk påvirkning giver en spænding over kondensatoren. De benyttes til mindre kritiske opgaver som signaloverførsel, filtrering, afkobling af effektforsyninger og blokering af jævnstrøm, men bør undgås ved forstærkning af meget svage signaler og som den centrale komponent i sample/hold-kredsløb.

Klasse 3 går op til 100 μ F, men kapaciteten varierer med spændingen og aftager over tid, tabene er høje og tolerancen er dårlig. Til gengæld fylder komponenten mindre end de to første klasser. Det anvendte dielektrikum er ikke-lineært så kondensatoren vil forvrænge et analogt signal.

¹⁷ http://en.wikipedia.org/wiki/Ceramic_capacitor.

1.5.3 Plastfilm kondensator

En plastfilm kondensator består af metalbelagte plastfolier der rulles sammen. De har større fysisk dimension og højere pris end tilsvarende keramiske kondensatorer. Der er mange varianter som dækker området fra 1 nF til 100 μ F. De har generelt god stabilitet og lave tab samt specifikke anvendelsesområder for signalbehandling, sample-and-hold, støjdæmpning og sikkerhed.¹⁸

Polykarbonat har lav dielektrisk absorption (ingen "hukommelse") så de benyttes til krævende formål som fx sample/hold-kredsløb. Temperaturkoefficienten er på ± 80 ppm/ $^{\circ}$ C og de tåler 125 $^{\circ}$ 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 $^{\circ}$ C.

1.5.4 Elektrolytisk kondensator

En elektrolytisk kondensator benytter en oxid af aluminium eller tantal som dielektrikum mellem selve metallet og en elektrolyt som den anden elektrode. Kapaciteten rækker fra 1 μ F til 10 mF og specielle typer når over 100 F (*SuperCap*). Kondensatoren tåler kun spænding i den ene retning så den skal vendes korrekt. Tolerance og temperaturkoefficient er meget ringe.¹⁹

Polariteten angives på komponentens hus, men der er selvfølgelig ikke konsensus. Ved aluminium elektrolyt markeres plus ved en indsnævring eller en streg ved minus. Ved tantal elektrolyt angives plus ved en streg. Hvis denne markering mangler vil plus være "til højre når teksten kan læses" og med benene nedad.

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 da de kan tåle strømstød. Den ækvivalente seriemodstand er høj ved lave kapacitetsværdier, men er lav for 100 μ F eller højere. Lækstrømmen er høj efter lang tids lagring, den aftager når komponenten har været i brug, men den bliver aldrig lav. Tolerancen er dårlig med typisk +50/–20 % og kan nå +100/–50 %. De leveres i temperaturkategorier på 85 $^{\circ}$ C og 105 $^{\circ}$ C og den nedre driftstemperatur er cirka –5 $^{\circ}$ C. Der tillades en svag negativ spænding på nogle få volt.

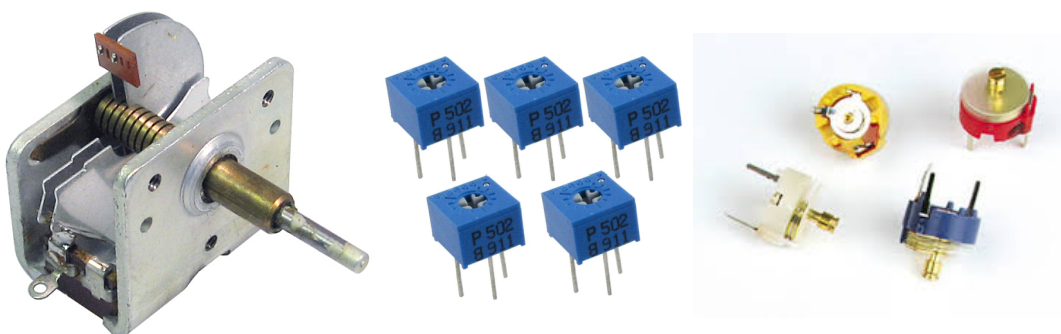
Tantal elektrolytkondensatorer dækker området fra 1 μ F til 100 μ F og er fysisk mindre end ved tilsvarende kapacitetsværdier ved aluminium. Tolerancen er bedre end for elektrolytkondensatorer af aluminium og den ækvivalente seriemodstand er lavere. Elektrolytkondensatorer af tantal tåler ikke ripplestrøm så de kan ikke benyttes i DC-DC effektforsyninger og de tåler heller ikke omvendt polarisering.

¹⁸ http://en.wikipedia.org/wiki/Film_capacitor.

¹⁹ http://en.wikipedia.org/wiki/Electrolytic_capacitor.

1.5.5 Variabel kondensator

Kapaciteten kan varieres hvis de to sæt af plader kan forskydes i forhold til hinanden. Det var en vigtig komponent i radioens barndom hvor radioens afstemte kredse og oscillator kunne indstilles til at modtage den ønskede station. Den variable kondensator benytter luft som dielektrikum, men mere moderne variable kondensatorer benytter en plastfilm for at tillade relativt høj værdi af den mulige kapacitet i små dimensioner. Kapaciteten er normalt i området 2 til 200 pF med en variation på højest 1:10 så en mulighed er en variabel kondensator på 5 til 50 pF.



Figur 40 – Variable kondensatorer (Wikipedia).

I tilgift til den mekaniske udgave kan en diode i spærreretningen benyttes som spændingsstyret kapacitet hvilket benyttes i næsten alle moderne radiomodtagere hvor en mikroprocessor står for kanalvalget. Det uddybes i femte kapitel.

1.5.6 Parasitisk kapacitet

Der vil altid være en kapacitet mellem to elektriske ledere. Værdien er givet af ledernes areal og afstanden mellem dem som det vises i næste afsnit. Det betyder i praksis at det er svært at komme væsentligt under 1 pF mellem to lederbaner i et kredsløb.

Signaltransmission over et kabel vil normalt give en kapacitet mellem lederne af størrelsesordenen 50 pF/m. Det betyder at det er vanskeligt at føre et signal væk fra et elektronisk kredsløb uden at belaste det med 50 pF og det kan have stor betydning for funktionen.

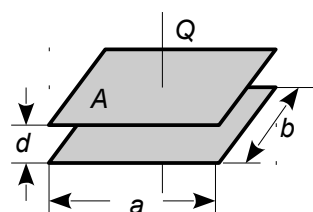
Internt i integrerede kredse kan man nå ned under 1 pF, men de små dimensioner tillader ikke at der ligger en høj spænding over kondensatoren uden et ødelæggende gennemslag på grund af det elektriske felts styrke (se side 44).

Dioder og transistorer fra side 7 vil i spærreretning opføre sig som en kondensator hvis værdi kan være fra 1 pF til 100 pF. Kapaciteten er spændingsafhængig og det udnyttes i radiokredsløb hvor en mikroprocessor kan skifte mellem de mulige radiokanaler ved at variere spændingen over komponenten.

En MOSFET benytter en kondensator som styreterminal for strømmen i en såkaldt kanal. Det betyder at indgangen på en MOSFET er en kapacitet. Værdien kan nå over 1 nF hvilket belaster det kredsløb der skal styre transistoren. Der skal leveres en strøm for at lade kondensatoren op så transistoren bliver aktiv og kondensatoren skal igen aflades når transistoren skal afbrydes. Det vil sætte en grænse for hvor hurtigt transistoren kan bringes til at skifte mellem aktiv og afbrudt.

1.5.7 Beregning af kapacitet

En kondensator består af to ledere med overfladeareal A i afstand d og med et materiale mellem de to plader. Materialet kaldes for et dielektrikum og beskrives ved den relative permittivitet ϵ_r der er defineret til én for vakuum og for andre materialer findes værdien i litteraturen eller på nettet. Den elektriske konstant ϵ_0 er en af de fundamentale konstanter indenfor SI.²⁰



$$\text{Ladning: } Q = C U$$

$$\text{Elektrisk felt: } E = \frac{U}{d}$$

$$\text{Energi: } W = \frac{1}{2} C U^2$$

$$\text{Kapacitet: } C = \frac{\epsilon_r \epsilon_0 A}{d} \quad A = ab \quad \begin{matrix} d \ll a \\ d \ll b \end{matrix}$$

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

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

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

$\epsilon_r =$	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 41 – Symbolet for en kondensator repræsenterer opbygningen med to parallelle streger.

Hvis en kondensator forbindes til et batteri vil der flyde elektroner fra minuspolen til den ene plade og det elektriske felt vil skubbe elektroner væk fra den anden plade som løber til batteriets pluspol. Der løber derfor en strøm i kondensatoren indtil det elektriske felt mellem de to plader er givet ved batteriets spænding divideret med afstanden mellem pladerne. Kondensatorens ladningen er et mål for hvor mange elektroner der er vandret ind til den ene plade og ud fra den anden.

Når kondensatoren er opladt indeholder den en energi som kan overføres til et andet sted i et elektronisk kredsløb. Det finder anvendelse som et lokalt reservoir i elektroniske opstillinger hvor en kondensator tæt på kredsløbet kan levere strøm når der er behov for det. Samme funktion udnyttes i effektforsyninger hvor en kondensator i det ene øjeblik oplades med energi og i det næste øjeblik afgiver energi. Sammen med en spole danner kondensatoren en svingningskreds hvor energien skubbes frem og tilbage mellem kondensator og spole. Det benyttes ved de fleste mikroprocessors clock-oscillator og i radiomodtagere.

En kondensators kapacitet måles i farad (F) og de værdier du vil møde dækker fra 1 pF (10^{-12} F) til 1 F. Lavere værdier mødes i integrerede kredsløb hvor der arbejdes med 1 fF (10^{-15} F) som laveste niveau. Området fra 1 nF (10^{-9} F) til 1 μ F (10^{-6} F) benyttes flittigt i analog elektronik for filtrering af enhver art. Værdier over 1 μ F benyttes næsten udelukkende som afkobling (lokalt energireservoir) og fra omkring 100 μ F og op bliver komponenten fysisk stor.

En serieforbindelse af to kondensatorer med samme areal, men forskellige afstande d_1 og d_2 vil svare til at afstandene adderes. En parallelforbindelse 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 parallelforbindelse af kondensatorer der er modsat formlerne for modstande.

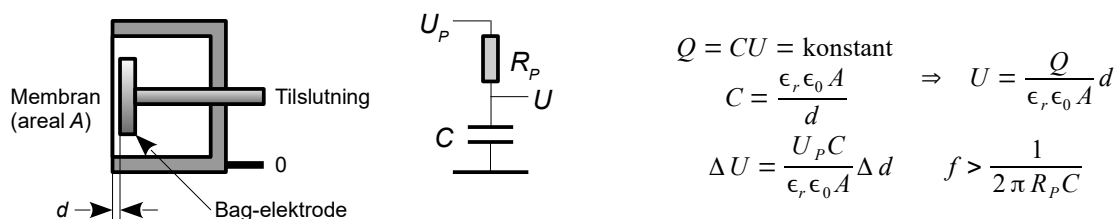
1.5.8 Eksempel

Elektronik bygges på printkort af glasfiber FR4 med $\epsilon_r = 4,7$ og en tykkelse på $d = 1,6$ mm som de typiske parametre. Hvis begge sider af kortet forsynes med et dækkende kobberareal vil de to planer danne en kondensator. Det populære Europakort måler 160 mm gange 100 mm og med kobberplaner på begge sider bliver arealet $A = 16 \cdot 10^{-3} \text{ m}^2$ og kapaciteten findes til $C \approx 400 \text{ pF}$.

²⁰ <https://sv.wikipedia.org/wiki/Permittivitet>.

1.5.9 Kondensatormikrofon

En populær mikrofontype er baseret på kondensatorens princip. Den ene plade er en metalfolie som lydets tryksvingninger sætter i bevægelse og den anden elektrode er fast. Kapaciteten er dermed funktion af trykvariationen og den kan omsætte til en spænding på to måder. Den kan indgå som del af en svingningskreds hvor en kapacitetsændring omsættes til en frekvensændring der igen omsættes til spænding ved velkendt teknik fra radiomodtagere. En anden mulighed er at give mikrofonen en fast ladning idet spænding og kapacitet da omvendt proportionale. Kapaciteten er dog kun 5 til 20 pF så en forstærker indbygges normalt som en integreret del af mikrofonen.

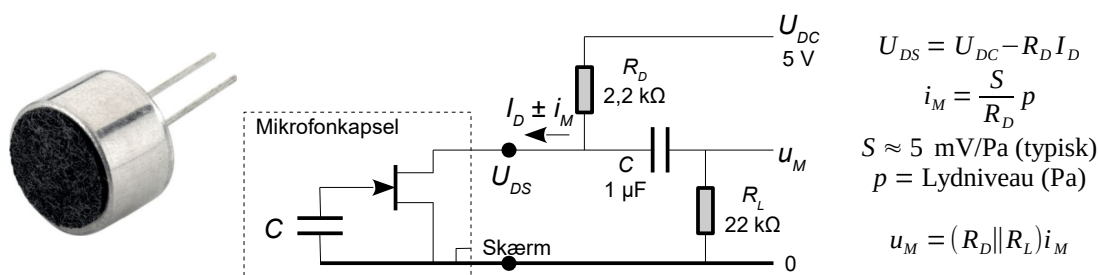


Figur 42 – En elektrostatiske mikrofon er opbygget med en tynd metalfolie som membran og en fast plade som bag-elektrode. Kondensatoren lades op fra en ekstern spænding gennem en stor modstand og spændingen over kondensatoren ændres som funktion af kapacitetsændringerne.

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

Der er to metoder for at give mikrofonens kondensator en fast værdi af ladningen. Ved en ekstern polariseringsspænding U_P gennem modstanden R_P får kondensatoren en ladning. Mikrofonens følsomhed bliver proportionel med polariseringsspændingen og en lav grænsefrekvens kræver en modstand på mindst 200 M Ω for at kunne gengive det hørbare frekvensområde.²¹

En anden metode benytter et dielektrikum mellem membran og bag-elektrode hvor ladningen er "infrosset" i materialet. Det er den såkaldte elektretmikrofon hvor en felteffekt transistor fungerer som buffer og mikrofontypen produceres i store serier til lav pris. Egenstøjen er relativt kraftig, men der bygges elektretmikrofoner for lyd- og målebrug med generelt gode data.^{22,23}



Figur 43 – En elektretmikrofon er en med "fastfrosset" ladning og indbygget felteffekt transistor som buffer. Der kræves en ekstern effektforsyning på 5 ... 12 V på grund af transistoren.

Transistoren trækker en strøm på $I_D < 0,5 \text{ mA}$ der leveres fra en effektforsyning U_{DC} gennem R_D . Signalet er en strøm i_M der optræder som variationer omkring DC arbejds punktet I_D for felteffekt-transistorens drain. Det bliver til en spændingsvariation u_M over R_D der signalteknisk er i parallel med R_L . Databladets angiver følsomheden ved "udgangsmodstanden" 2 k Ω til $S = 5 \text{ mV/Pa}$ som typisk værdi. Signalets spændingsamplitude er da $u_M = 5 \text{ mV}$ ved lydniveauet $p = 1 \text{ Pa}$ der er relativt kraftigt. Det svarer til en amplitude af signalet i transistorens drain på $i_M = 2,5 \text{ }\mu\text{A/Pa}$.²⁴

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

22 MCE-400 fra Monacor koster 20 kr. Følsomheden er 7,9 mV/Pa $\pm 3 \text{ dB}$ ved 2,2 k Ω og den dækker 50 Hz til 16 kHz.

23 MCE-4000 fra Monacor med Phantom Powering. Følsomheden er 5,6 mV/Pa $\pm 1 \text{ dB}$ og den dækker 30 Hz til 18 kHz.

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

1.6 Spole – L

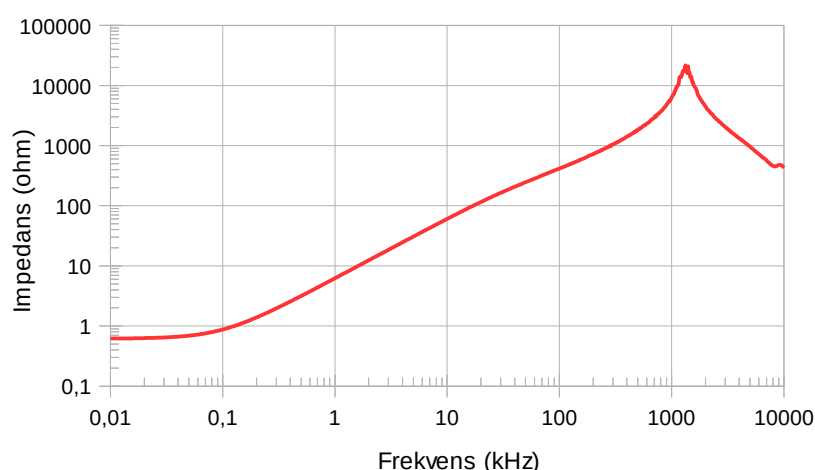
Den tredje og sidste af de passive komponenter er selvinduktions-spolen (*self-induction coil*) der i daglig tale blot kaldes for en spole (*inductor* eller *coil*). Ligesom kondensatoren kan den oplagre energi, men det sker gennem opbygning af et magnetisk felt som kræver at der løber en strøm i lederen. Man kan derfor ikke momentant starte eller stoppe strømmen i en spole for energien skal først opbygges eller afgives. Hvis man forsøger at afbryde forbindelsen til en spole hvor der løber en strøm vil den reagere med en gnist hvorved energien afgives. Det udnyttes i benzindrevne biler til at antænde gasblandingen.



Figur 44 – Små spoler for printkort, en spole uden kerne, to spoler med stavformet kerne og en toroid spole hvor den ringformede kerne danner en lukket sløjfe for magnetfeltet (Wikipedia).

Spolens uvillighed overfor ændring i strømstyrken skyldes det forhold at det tager tid at ændre på energien. Strømmen må derfor vokse fra nul i et behersket tempo til slutværdien og tilsvarende skal strømmen derefter reguleres ned til nul ved at afsætte den opmagasinerede energi i det ydre kredsløb. Spolen er derfor en modsætning til kondensatoren hvor energien oplagres ved at der løber strøm i et vist tidsrum som får spændingen til at vokse fra nul op til slutværdien. Dette modsætningsforhold er af stor vigtighed for elektroniske kredsløb med kondensatorer og spoler og uddybes senere i bogen.

En spoles selvinduktion måles i henry (H) og du vil møde værdier fra 10 nH for én enkelt vinding til omkring 10 mH for en spole med mange vindinger. Højere værdi af selvinduktionen kan opnås ved at inkludere en magnetisk kerne hvor selvinduktionen kan nå over 10 H.



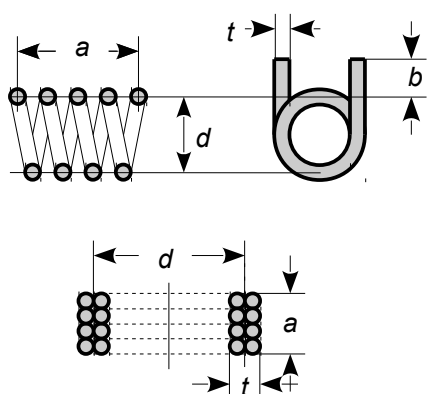
Figur 45 – Målt impedans af en 1 mH spole som funktion af frekvensen.

Ligesom kondensatoren kan spolen beskrives som en modstand overfor vekselstrøm. Det kaldes også her for en impedans. Ved meget lave frekvenser er spolens impedans lig med modstanden af den tråd spolen er viklet af. Ved højere frekvenser er impedansen proportional med frekvensen indtil et maksimum hvor spolens selvinduktion danner resonans med kapaciteten mellem spolens vindinger som vist side 39. Derover aftager impedansen igen da kapaciteten nu dominerer.

1.6.1 Beregning af selvinduktion

For en spole uden magnetiserbart materiale i nærheden beregnes selvinduktionen ved de følgende tre trin for en spole viklet i cylindrisk form. Formlerne kan benyttes for andre tværsnit hvis diameteren vælges som svarende til det samme areal af tværsnittet.

1. Formlen L_1 skyldes det magnetiske felt i lederen og gælder for den totale længde tråd i spole og tilledning. Værdien er 50 nH/m op til 20 kHz og aftager mod nul derofter.
2. Formlen L_2 skyldes det magnetiske felt gennem tilslutningens areal og inkluderer derfor også det kredsløb som spolen indgår i.
3. Formlen L_3 skyldes spolens udformning og jeg viser to populære formler for beregning.



Selvinduktion:	$L = L_1 + L_2 + L_3$	
Tråden selv:	$L_1 = \frac{\mu_0}{8\pi} l$	$f < 20 \text{ kHz}$
Samlet længde tråd:	$l = 2b + N\pi d$	
Magnetisk konstant:	$\mu_0 = 4\pi \cdot 10^{-7} \text{ H/m}$	
Tilslutning:	$L_2 = \frac{\mu_0 b}{\pi} \ln\left(\frac{2d}{t}\right)$	
Wheeler:	$L_{3W} = \frac{\pi \mu_0 N^2 d^2}{4a + 0,45d}$	$a > 0,4d$
Kort spole:	$L_{3K} = \frac{\mu_0 N^2 d}{2} \ln\left(\frac{d}{a}\right)$	$t < a < 0,1d$

Figur 46 – Formler for en cylinderformet spole.

For en spole med viklingen i ét lag benyttes Wheelers formel der påstås at være indenfor $\pm 0,35\%$ hvilket jeg dog tager forbehold overfor. Hvis faktor 0,45 erstattes med 0,42 kan formen bruges ned til en længde på 10 % af diameteren ved en nøjagtighed på $\pm 10\%$ (se side 56). For en kort spole benyttes sidste formel for en længde på højst 10 % af diameteren. Nøjagtigheden er ukendt, men den giver et fornuftigt estimat. Formlen udledes på side 60.

Herunder giver jeg to eksempler på brugen af formlerne. De rammer godt, men min erfaring er at nøjagtigheden ikke kan forventes bedre end $\pm 10\%$ og at der let opstår overraskelser.



Cylinderspole med $N = 25$ vindinger, diameteren $d = 32,5 \text{ mm}$, ydre længde $a = 12,5 \text{ mm}$ og tilledningerne er $b = 0,2 \text{ m}$ lange. Trådlængden bliver $N\pi d = 2,6 \text{ m}$ plus $0,4 \text{ m}$ fra tilledningen til en længde på $l = 3,0 \text{ m}$ så $L_1 = 0,15 \mu\text{H}$. Trådtykkelsen er $t = 0,25 \text{ mm}$ og trådene er løst snoet med $d = 5 \text{ mm}$ afstand så $L_2 = 0,3 \mu\text{H}$. Forholdet $a/d = 0,38$ er for lavt til Wheelers formel, men hvis det ignoreres findes $L_{3W} = 24,0 \mu\text{H}$ og estimatet bliver $L = 24,45 \mu\text{H}$. To ens spoler var af producenten målt til $24,2 \mu\text{H}$ og $24,5 \mu\text{H}$.



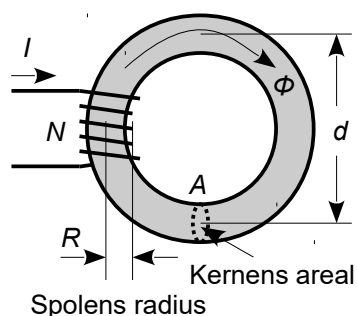
Kort spole med $N = 4$ vindinger, diameteren $d = 180 \text{ mm}$ fra centrum til centrum, trådtykkelsen er $t = 0,5 \text{ mm}$ og isolationen er $0,5 \text{ mm}$ tyk så beviklingens længde er $a = 5 \text{ mm}$. Med $b = 1 \text{ m}$ for tilledningen bliver trådlængden $l = 4,3 \text{ m}$ og derfor $L_1 = 0,22 \mu\text{H}$. Tilledningen var løst viklet så $d = 5 \text{ mm}$ og $L_2 = 1,20 \mu\text{H}$. Formlen for kort spole giver $L_{3K} = 6,48 \mu\text{H}$ så estimatet bliver $L = 7,90 \mu\text{H}$. Spolen er målt til netop denne værdi på en målebro ved 10 kHz .

1.6.2 Spole med kerne

For at øge selvinduktionen kan man forsyne den med en kerne af et magnetiserbart materiale. Det er materialer med en indbygget magnetisk dipol ved elektronernes rotation som kaldes for et spin. Som udgangspunkt peger de magnetiske dipoler i tilfældig retning så nettoresultatet bliver nul set udefra. Med et eksternt magnetisk felt fra spolen benyttes energi til at vippe nogle af dipolerne så de peger i samme retning som det eksterne felt og det giver en feltforstærkende virkning. Der nås et maksimum når alle dipoler er drejet i samme retning (se side 64). De benyttede materialer er jern (Fe), kobolt (Co) og nikkel (Ni) samt en del af undergruppen lanthaniderne.^{25,26}

Den feltforstærkende virkning udtrykkes ved den relative permeabilitet μ_r der er én for umagnetiske materialer som luft, plast, træ og en del keramik og større end én for ferromagnetiske materialer. Dertil kommer en relation til hvordan kernen er opbygget og den feltforstærkende virkning kan nå over 10^4 gange hvis kernen danner en lukket magnetisk sløjfe overfor det magnetiske felt.

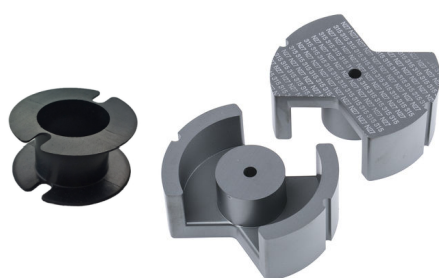
Ved lave frekvenser benyttes jern som kerne og oftest i form af tynde lameller der er isoleret fra hinanden for at undgå tab fra hvirvelstrømme. De benyttes til nettransformatorer både ved høj- og lavspænding, til analog signaltransmission og i drosselspoler. Se mere på side 64.



$$L = \frac{\mu_r \mu_0 N^2 A}{\pi d}$$

Materiale	μ_r
Mu-metal	20000
Permalloy	8000
Ferrit	16...1000
Stål	100
Vacuum	1

Ved højere frekvenser foretrækkes en kerne af ferrit der er et jernoxid (Fe_2O_3) i legering med nikkel, zink og mangan. Det kan støbes i den ønskede form og blandingsforholdet bestemmer funktionen med hensyn til feltforstærkning og brugbart frekvensområde med mere. Materialets magnetiske funktion repræsenteres i databladet ved parameteren A_L der angiver den opnåede selvinduktion ved én vinding så selvinduktionen er N^2 gange større for N vindinger.²⁷



$$L = N^2 A_L$$

$A_L = \text{Databladsopgivelse}$

Figur 47 – En spoleform af plast og en ferritkerne bestående af to ens skaller (RS, Epcos N27).

Udformningen af ferritkernen sikrer et minimum af spredning af det magnetiske felt så mange spoler kan anbringes tæt på hinanden med minimal kobling indbyrdes mellem spolerne og en minimal udstråling af energi til omgivelserne.

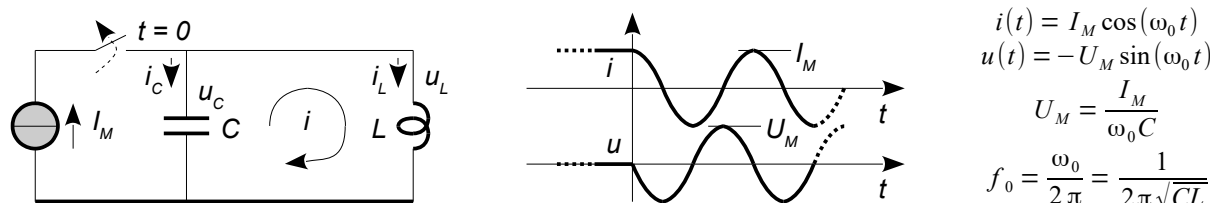
²⁵ <https://da.wikipedia.org/wiki/Magnetisme>.

²⁶ https://en.wikipedia.org/wiki/Magnetic_core.

²⁷ https://en.wikipedia.org/wiki/Ferrite_core.

1.7 Resonans – LC

De to reaktive komponenter kondensatoren og spolen er flittigt benyttet i radiokredsløb da de tilsammen har den vigtige egenskab at kunne danne en svingningskreds der kan svinge villigt indenfor et snævert frekvensområde omkring en centerfrekvens.



Figur 48 – En sammenkobling af en kondensator og en spole danner en svingningskreds.

En svingningskreds vises i illustrationen som en kondensator og spole i parallel. Strømkilden I_M driver en strøm i spolen og da modstanden i en ideel spole er nul vil spændingsfaldet over de to komponenter være nul. Tid tiden $t = 0$ afbrydes strømkilden hvorefter kondensator og spole er overladt til sig selv. Strømmen fra spolen kan ikke blot falde til nul (se side 52) så den bliver ved med at løbe efter at kontakten har afbrudt for strømkilden og kondensatoren bliver derved opladt mens spolens magnetfelt aftager. Når kondensatorens spænding er maksimal er strømmen faldet til nul og hele energien i spolens magnetfelt er overført til kondensatorens elektriske felt. Derefter starter den modsatte proces hvor kondensatoren aflades og genopbygger energien i spolen.

Det kaldes for resonans. Spænding og strøm er 90° ude af fase så den ene er maksimal når den anden er nul. I et mekanisk system svarer det til at bevægelsesenergi veksler med potentiel energi. Fx et standerurs pendul der enten er i ro i en yderstilling eller har fart på forbi ligevægt.

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 47).

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

Ved differentiation fjernes integrationen og relationen bliver en anden-ordens differentialligning. Strømmen ses at skulle 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 $t = 0$.

$$i = I_M \cos(\omega_0 t) \Rightarrow \frac{di}{dt} = \frac{d\{I_M \cos(\omega_0 t)\}}{dt} = -\omega I_M \sin(\omega_0 t) \\ \frac{d^2 i}{dt^2} = \frac{d\{-\omega I_M \sin(\omega_0 t)\}}{dt} = -\omega^2 I_M \cos(\omega_0 t)$$

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

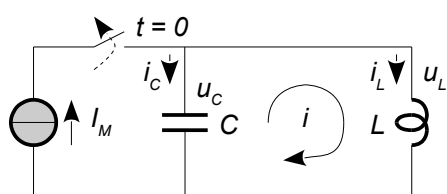
$$I_M \cos(\omega_0 t) = LC \omega_0^2 I_M \cos(\omega_0 t) \Rightarrow 1 = LC \omega_0^2 \Rightarrow \omega_0 = \frac{1}{\sqrt{LC}}$$

Strømmamplituden I_M er kendt og spændingen u_C over opstillingen som funktion af tiden kan findes ved indsættelse af strømmens forløb i definitionen for kondensatorens spænding.

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

1.7.1 Energi

Kondensatoren opmagasinerer energi ved at akkumulere ladning og spolen opmagasinerer energi ved en ladning i bevægelse. Ved at forbinde de to komponenter kan energien veksle mellem at elektronerne er oplagret eller i bevægelse. Det er den samme energimængde der flyttes frem og tilbage mellem de to komponenter så den totale energi i kredsløbet er til ethvert tidspunkt lig med summen af energierne i de to komponenter.



Energi i en kondensator:

$$w_C(t) = \frac{1}{2} C u^2(t) \Rightarrow W_{C\text{MAX}} = \frac{1}{2} C U_M^2$$

Energi i en spole:

$$w_L(t) = \frac{1}{2} L i^2(t) \Rightarrow W_{L\text{MAX}} = \frac{1}{2} L I_M^2$$

Energi i kredsløbet:

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

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

Figur 49 – Energien svinger frem og tilbage mellem kondensatoren og spolen i kredsløbet, og den samlede energi er givet ved strømmens eller spændingens maksimum.

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 hvorved der tilføres ladningen dq . Den samlede potentielle energi findes ved integration af dW over spændingsændringen fra $u = 0$ til $u = U_M$.

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

Spolens energi opbygges af strømmen i spolen der løber gennem et spændingsfald forårsaget af ændringshastigheden (se Faradays induktionslov på side 52). Ladningsændringen dq er givet ved at strømmen i løber over tidsintervallet dt . Den oplagrede kinetiske energi findes ved at integrere over strømændringen fra $i = 0$ til $i = I_M$.

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

1.7.2 Eksempler

For at få en ide om betydningen af energi i elektroniske kredsløb følger et par eksempler.

Ved elektrostatisk udladning ESD (*electrostatic discharge*) oplades en kondensator på 100 pF til en spænding på 8 kV inden den aflades som del af EMC direktivets test af robusthed overfor brug. Kondensatoren indeholder en energi på 3 mJ hvilket kan mærkes som et solidt elektrisk stød og det er nok til at beskadige elektroniske kredsløb. Derfor indpakkes kredsløbene som regel i en beskyttende kasse.^{28,29,30}

Et relæ aktiveres ved at lægge en spænding over relæets spole som så trækker i kontakterne. Når relæet skal deaktiveres skal strømmen i spolen standses og det kræver omtanke. Ved et relæ med en selvinduktion på 0,1 H vil en strøm på 10 mA svare til en oplagret energi på 5 μJ og selv om det ikke lyder af meget kan det beskadige den transistor som driver strømmen i spolen. Spændingen over spolen stiger indtil der findes en afladningsvej og hvis det er via transistorens kollektor så vil energien blive afsat i et mikroskopisk lille område af transistorens krystal. Opvarmningen kan smelte den aktive del af halvlederen og ødelægge transistoren.

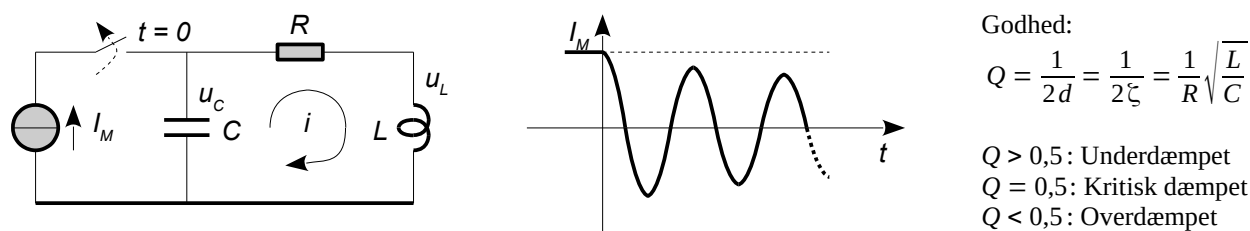
²⁸ http://www.elmac.co.uk/pdfs/ESD_considerations.pdf.

²⁹ https://en.wikipedia.org/wiki/Electromagnetic_compatibility.

³⁰ http://cache.nxp.com/documents/application_note/AN10853.pdf?pspll=1.

1.7.3 Dæmpet svingning

Energiudvekslingen mellem kondensatoren og spolen er ikke tabsfri så svingningen vil klinge ud i takt med at energien omsættes til varme. Tabene repræsenteres ved modstanden R . Strømmens amplitude vil gradvist blive reduceret mod nul med en fast brøkdelen for hver periode og det leder til en eksponentielt aftagende amplitude. Et mål for dæmpningen er godheden Q (*quality factor*) der også beskrives ved dæmpningsfaktoren d eller ζ . For kredsløb med en høj godhed angiver antallet af perioder af svingningen inden amplituden er aftaget til nær ved nul.^{31,32}

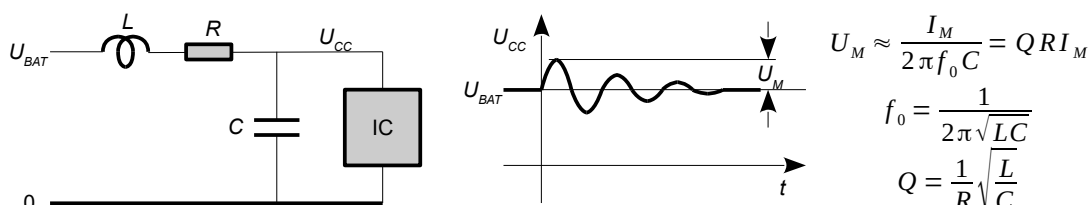


Figur 50 – Amplituden af svingning vil aftage eksponentielt mod nul på grund af tabene.

Alle LC kredsløb er dæmpede da der er en modstandsværdi i spolen som tapper energi og begge komponenter udstråler elektromagnetisk energi til omgivelserne.

1.7.4 Afkobling af effektforsyningen

Kredsløbet herunder repræsenterer en hyppigt mødt situation hvor den elektroniske komponent IC har en kondensator som lokal afkobling og får sin effektforsyning via en kobberbane på et printkort. Banen repræsenterer en selvinduktion L så der dannes en svingningskreds med afkoblingens kondensator C og svingningen dæmpes af modstanden R 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 da den resulterer i en variation i komponentens arbejdsbetingelse.



Figur 51 – En analog eller digital komponent er effektforsynet gennem en kobberbane på et print.

En kort leder har en selvinduktion af størrelsesordenen 1 nH/mm så en bane på 100 mm vil kunne nå en selvinduktion på 100 nH og med 0,25 mm bredde og 35 μ m tykkelse af kobberbanen er DC modstanden på $R = 0,2 \Omega$ (Se første kapitel). Med $C = 100$ nF bliver resonansens frekvens på $f_0 = 1,6$ MHz. En pludselig strømændring på $I_M = 0,1$ A giver en startværdi af amplituden på $U_M \approx \pm 0,1$ V hvor fortegnet er positivt ved en strømreduktion. Godheden er $Q = 10$ så svingningen klinger ud efter cirka 6 μ s.

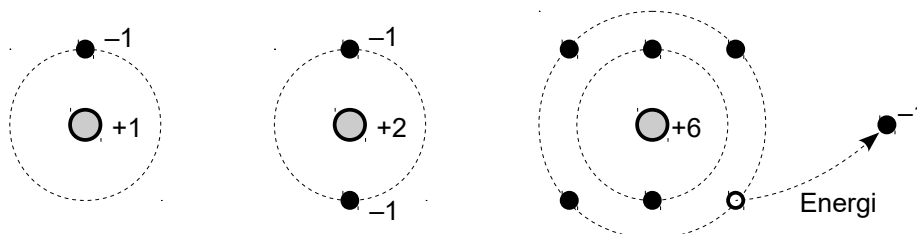
Se også fjerde kapitel hvor dette emne uddybes.

31 Amplituden aftager med $\exp(-\zeta\omega_0 t) = \exp(-\pi f_0 t / Q)$ til 4 % af begyndelsesniveauet efter Q gange perioden.

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

1.8 Elektricitet

Elektricitet skyldes elektronerne³³ 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 52 – 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 elektroner tegnes som en partikel i en cirkelformet bane om kernen og størrelsesforholdet er også 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 af temperaturen idet varme er udtrykt ved elektronerne's bevægelsesenergi og de vibrerer desto kraftigere når det er varmt. Vibrationens amplitude er statistisk betinget og ikke bare en pæn oscillation omkring en ligevægt. Elektron efter elektron vil få energi nok til at smutte væk fra atomet og de flyder så rundt indtil de bliver opfanget af et andet atom der mangler en elektron. 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 og den er båret af de frie elektroner.

Der kan derfor konkluderes at en elektrisk leder har mange frie elektroner og at en isolator har få eller slet ingen frie elektroner.

De frie elektroner kan accelereres op i fart og opnå energi nok til at slå andre elektroner løs ved kollision med krystallets atomer som så igen accelereres af det elektriske felt. Det kaldes for en lavineeffekt (*avalanche effect*) og det udnyttes til følsomme lysmodtagere ved fotomultiplikatorer, til spændingsregulering ved zenerdioder og til dannelse af smalle og energirige pulser.

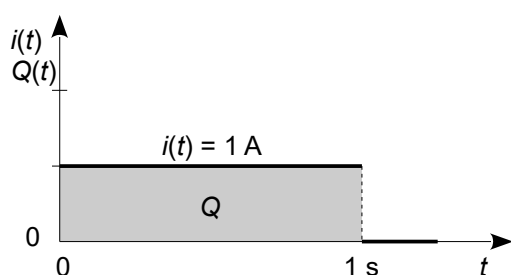
Ved tilstrækkelig høj temperatur kan elektroner nær ved overfladen opnå nok energi til at blive slynget helt ud af metallet. Det blev udnyttet i 1900-tallets rørdioder hvor katoden blev opvarmet til glødning så den afgav frie elektroner som derefter blev tiltrukket af den positive anode.

³³ http://www.denstoredanske.dk/It%2c_teknik_og_naturvidenskab/Fysik/Elektrontransport_i_faste_stoffer/elektrisk_felt

1.9 Ladning – Q

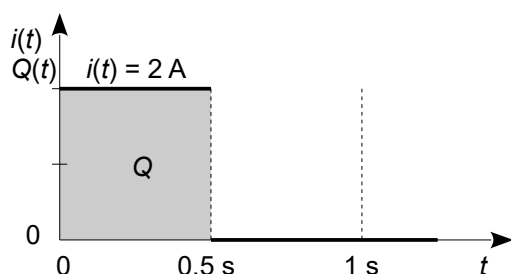
Hver elektron har en fast værdi af den elektriske ladning, elementarladningen, der beskrives ved symbolet q_0 . En samling af elementarladninger udgør en ladning (*charge*) der beskrives ved symbolet Q for en statisk ladning og $q(t)$ eller blot q for en tidsvarierende ladning. Ladning måles i coulomb (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 $6,242 \cdot 10^{18}$ elektroner.^{34,35}

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 tænkt (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.



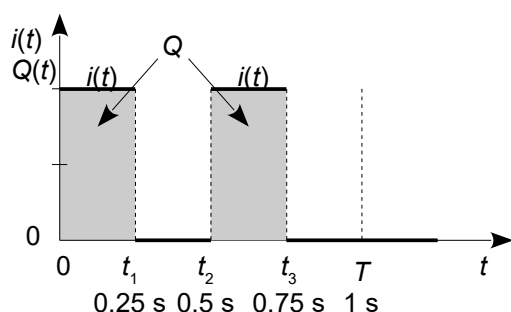
Konstant strøm på $I = 1 \text{ A}$ i tidsrummet $T = 1 \text{ s}$
Den samlede ladning er $Q = 1 \text{ C}$

$$Q = \int_0^T i(t) dt = \int_0^T I dt = I \int_0^T dt = IT = 1 \text{ A} \cdot 1 \text{ s} = 1 \text{ C}$$



Konstant strøm på $I = 2 \text{ A}$ i tidsrummet $T = 0,5 \text{ s}$
Den samlede ladning er 1 C

$$Q = \int_0^T i(t) dt = \int_0^T I dt = I \int_0^T dt = IT = 2 \text{ A} \cdot 0,5 \text{ s} = 1 \text{ C}$$



Konstant strøm på $I = 2 \text{ A}$ i to forløb på hver 0,25 s
Den samlede ladning er 1 C

$$Q = \int_0^T i(t) dt = I \int_0^{t_1} dt + I \int_{t_2}^{t_3} dt \Rightarrow$$

$$Q = I \cdot [t_1 - 0] + I \cdot [t_3 - t_2] = 2 \text{ A} \cdot 0,25 \text{ s} + 2 \text{ A} \cdot 0,25 \text{ s} = 1 \text{ C}$$

Figur 53 – 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.

34 Andre symboler er: q , q^- , e , e_0 og e^- .

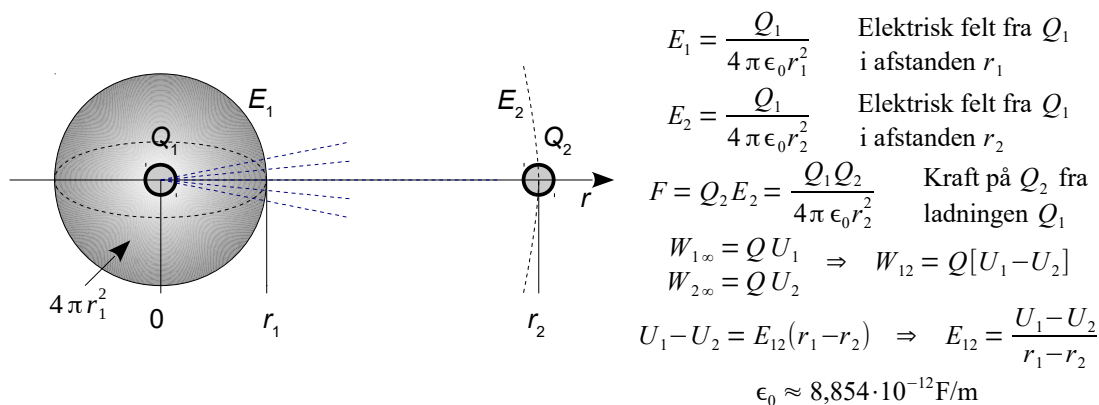
35 Se fx http://en.wikipedia.org/wiki/Electric_charge.

1.10 Elektrisk felt – E

En ladning består af et antal 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 tilnærmelse i atmosfærisk luft. Feltet bevæger sig med lysets fart så man kan normalt antage at feltet blot følger med når værdien af en ladning ændres eller flyttes.



Figur 54 – 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.

Den elektriske feltstyrke E får enheden volt per meter (V/m).

Når det elektriske felt eksisterer i et materiale er der en øvre grænse for hvad atomerne kan klare uden at elektroner trækkes fri. For atmosfærisk luft er grænsen på 3 kV ved en afstand på 1 mm som giver feltstyrken til $E = 3 \text{ MV/m}$. Hvis det elektriske felt bliver kraftigere kan der løbe en strøm på grund af de frigjorte elektroner (gnist eller lysbue). Man vil oftest benytte den konservative værdi på 1 kV/mm som svarer til $E = 1 \text{ MV/m}$ for at opnå en sikkerhedsmargin.

For at afbryde spændingen i højspændingsnettet skal der benyttes en mekanisk afstand på 0,6 m for sikkert at kunne spærre for strømmen i et 400 kV net.

Ioniseringen udnyttes i en såkaldt korona-højtaler hvor en høj spænding får en strøm til at løbe i luften og strømstyrken varieres i takt med det signal der skal gøres hørbart.

Feltstyrken i en halvleder kan blive meget høj på grund af de små afstande mellem de enkelte dele i den integrerede kreds. Grænsen for elektrisk feltstyrke er 30 MV/m for silicium og det nås let ved moderne integrationstæthed. Ved 100 nm afstand mellem de spændingsførende dele vil en forskel på 3 V nå op på grænseværdien.

Formlerne i illustrationen på forrige side kommenteres herunder.

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 . Kraftens retning er langs med r -aksen i figuren så en positiv værdi vil søge at skubbe ladning Q_2 væk.

Kraften er givet ved at samle udtrykkene og det leder til Coulombs lov. 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_1Q_2 er positivt og det betyder at Q_1 og Q_2 har samme fortegn. Kraften udgør derimod en tiltrækning hvis fortegnet af produktet er negativt og de to ladninger har forskellige fortegn.³⁶

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 hvorved der opbygges potentiel energi som senere kan indvindes igen. Som eksempel er en kondensator i stand til at opmagasinere potentiel elektrisk energi i sit elektriske felt mellem pladerne.

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 (positiv) eller om der tabes energi (negativ). 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 ($V = J/C$).

Det er ikke muligt at angive værdien af potentialet da definitionen er baseret på en ændring af ladningens position. Vi kan kun referere til ændringen i potentialet fra én position til en anden. Derfor er enhver angivelse af et potentiale underforstået som differensen mellem det aktuelle potentiale og en references potentiale. Referencen kaldes nul, stel, GND, ground eller chassis.

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 som dr står for 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 formlens 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.

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

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 $\cos(\theta) = 0$.

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

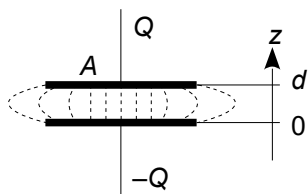
1.10.1 Kapacitet

Med en elektrisk spændingsforskel mellem to ledere dannes der et elektrisk felt mellem dem. For at føre en ladning fra den ene leder til den anden kræves tilførsel af energi så der kan oplagres en potentiel energi i det elektrostatiske felt. Det er muligt at trække energien ud igen ved at danne en forbindelse fra den ene plade til den anden. Der vil da løbe en strøm indtil al energien er tappet.

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 Q og spændingen U som 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 der kaldes farad ($F = C/V = \text{As/V}$). Herunder vises en beregning af kapaciteten for to plader i en relativ kort afstand. Forbeholdet for afstanden skyldes at feltet ikke kun består af parallelle feltlinjer for det buer udad ved kanten. Det kaldes for randfeltet.

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 permittivitet.³⁷



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

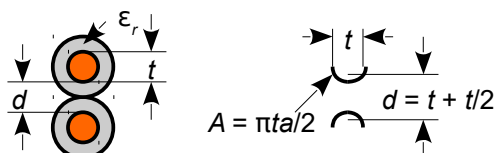
$$U = \frac{W}{Q} = \int_0^d \frac{QE}{Q} dz \quad \Rightarrow \quad C = \frac{\epsilon_r \epsilon_0 A}{d}$$

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 hvorved 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 .

1.10.2 Kapaciteten mellem to parallelle ledere

Et kabel med to parallelle ledere benyttes for at overføre energi fra lysnettet, fra en effektforstærker til en højttaler og for transmission af digitale data som ved et USB kabel. For det typiske kabel med to ledere kan kapaciteten mellem lederne beregnes til cirka 20 pF per meter længde så et 10 m langt kabel svarer til en samlet kapacitet på 200 pF mellem de to ledere.

Lederene tilnærmes ved to parallelle planer hvis overfladeareal sættes til det halve af lederens omkreds ($\pi t/2$) gange lederens længde a . Afstanden mellem de to arealer sættes til $d = t$ så fladernes krumning ignoreres. Det antages at feltet hovedsageligt løber indenfor isolationen med den effektive værdi på $\epsilon_r = 2,1$. Ved at dividere med a opnås beregning af kapaciteten per længde af kablet. Beregningen benytter kun den del af feltet der løber direkte mellem de to halvdele af lederens metal så randfeltet ignoreres og beregningen er kun en tilnærmelse.

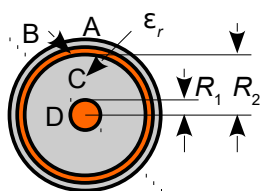
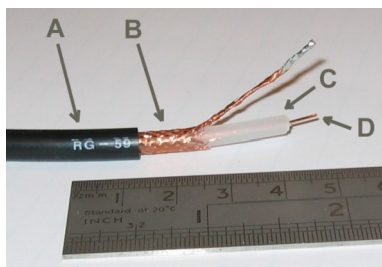


$$C \approx \frac{\epsilon_r \epsilon_0 \frac{\pi t}{2} a}{t + \frac{t}{2}} \quad \Rightarrow \quad \frac{C}{a} \approx \frac{\epsilon_r \epsilon_0 \pi}{3} \quad \epsilon_r \approx 2,1 \quad 20 \text{ pF/m}$$

³⁷ 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>

1.10.3 Kapaciteten ved et skærmerkabel

Et populært kabel består af en elektrisk leder omgivet af en koncentrisk leder der kaldes for en skærm (*shield*). Et skærmerkabel anvendes til at overføre svage eller højfrekvente signaler fra mikrofon, TV-antenne og Ethernet. Det elektriske felt er begrænset til det indre af kablet med den ydre leder som afskærmning overfor et forstyrrende eksternt felt.



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

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

$$\epsilon_r = 2,1$$

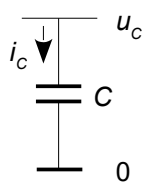
Figur 55 – Skærmerkablet består af en isolerende kappe A, skærmen B der danner den ydre leder, en isolator C og den indre leder D. I formelen er R_1 trådens ydre radius og R_2 er skærmens indre radius (fra Wikipedia). Forholdet på 3,2:1 mellem radierne svarer til et såkaldt RG58 kabel.

Den elektriske ladning er ens for skærmen og den indre leder, men arealet af de to leders overflader er forskellige. For at omgå det problem analyseres området mellem den indre og ydre leder som 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 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 formelen for serieforbundne kondensatorer der kan udtrykkes mere hensigtsmæssigt ved at benytte det reciprokke af kapaciteterne.

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

1.10.4 Kondensatorligningen

Der er nu materiale nok til at udlede en overordentlig praktisk formel for kondensatorer. På side 43 er ladningen Q givet ved strømmen I der løber i tidsrummet T som $Q = IT$ og kapaciteten er også givet ved $Q = CU$. De to udtryk beskriver begge kondensatorens ladning og de kan samles til et fælles udtryk der i denne bog vil blive kaldt for kondensatorligningen.



Kondensatorligning:

$$Q = C U = I T$$

... på differensform:

$$\Delta Q = C \Delta U = I \Delta T$$

... på differentialform:

$$dq = C du = i dt$$

Strømmen $i(t)$:

$$i = \frac{dq}{dt} = C \frac{du}{dt}$$

Spændingen $u(t)$:

$$du = \frac{i dt}{C} \Rightarrow u = \frac{1}{C} \int_0^T i dt + U_0$$

Figur 56 – Ladningen på en kondensator er givet ved potentialet og den tid en strøm har løbet.

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.

Næst efter elektronikkens *grundlov* som Ohms lov udgør er det her *formlen* for kondensatorer. Den er rigtig god ved opstillinger hvor en kondensator skal op- eller aflades over et tidsrum. Du vil se formelen anvendt i megen litteratur uden at den overhovedet defineres. Det antages simpelthen at den er del af den ballast læseren er bekendt med.

Kondensatorens ladning ændres under op- eller afladning når en strøm I løber i tidsrummet T så det er ganske praktisk at beskrive tidsintervallet som ΔT eller dt hvor strømmen I løber og ændringerne i ladning og spænding betegnes tilsvarende som ΔQ og ΔU eller som dq og du . Der er ingen grundlæggende forskel fra den første ligning for det er mere et udtryk om konventioner.

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 i dette tidsrum med $\Delta U = I\Delta T/C = (100 \text{ mA})(10 \text{ ms})/(1000 \mu\text{F}) = 1 \text{ V}$.

1.10.5 Genopladeligt batteri

Kondensatorligningen kan anvendes ved batterier selv om ladningen her er kemisk bundet. Et batteris ladning måles almindeligvis i ampere timer (Ah) der er det samme som coulomb (C) på nær en faktor. Der er 3600 sekunder i hver time så vi har $1 \text{ Ah} = 3600 \text{ As} = 3600 \text{ C}$.

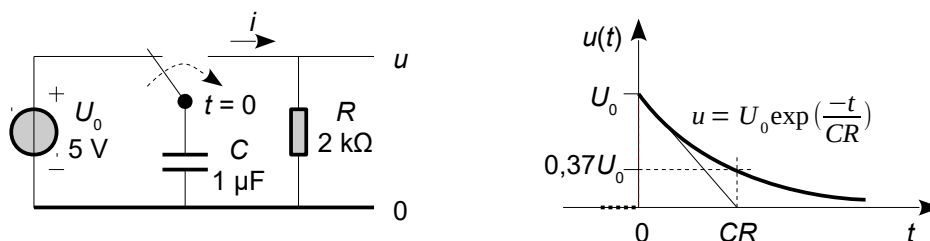
Et batteri oplades med en konstant strøm på $0,1 \text{ A}$ til ladningen $0,5 \text{ Ah}$ er opnået og opladningen behøver en tid på $T = Q/I = (0,5 \text{ Ah})/(0,1 \text{ A}) = 5 \text{ h} = 18\,000 \text{ s}$.

Ikke alle batterityper tåler fuld afladning. Et Li-Ion batteri er fuldt opladt ved $4,2 \text{ V}$ og afladning til en spænding under 3 V initierer en destruktiv proces. En del batterier indeholder derfor et kredsløb der afbryder ved underafladning. Ellers kan der opstå en eksplosion. Et blybatteri (bilbatteri) er mere robust og tåler fuld afladning, men det indeholder tre gange mindre ladning per vægt.³⁸

Der er en virkningsgrad knyttet til sekundærbatterier så det er kun $80 \dots 95 \%$ af energien der kan tappes fra batteriet efter opladning. Dertil har batterier en selvafladning der langsomt aflader det med cirka 10% per måned for Li-Ion så efter et års tid er to-tredjedele af ladningen gået tabt. Der mistes en tilsvarende cirka 10% af den oprindelige ladning for et typisk Li-ion batteri efter omkring 250 komplette op- og afladninger.^{39,40}

1.10.6 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 tidspunktet $t = 0$ og derefter glider asymptotisk mod nul for $t > 0$.



Figur 57 – En kondensator holdes opladet 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 med de viste komponenter er på $CR = 2 \text{ ms}$. Det betyder at i løbet af tidsrummet fra nul til $t = CR$ mindskes spændingen fra U_0 til $\exp(-1) \cdot U_0$ hvilket er $0,37 \cdot U_0$. Efter to tidskonstanter er spændingen $0,135 \cdot U_0$ og den aftager langsomt mod nul. I teorien når spændingen aldrig nul, men den er under 1% af det oprindelige niveau efter $t = 4,6 CR$ så man kan normalt regne en kondensator for afladet efter mindst fem gange tidskonstanten.

³⁸ <http://da.wikipedia.org/wiki/Energi>.

³⁹ http://en.wikipedia.org/wiki/Lithium-ion_battery.

⁴⁰ http://batteryuniversity.com/learn/article/how_to_prolong_lithium_based_batteries.

Spændingen over kondensatoren er givet af kondensatorligningen (se side 47) 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 idet 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 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. I modsat fald må man prøve med et modificeret 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 leddet $-CR$ giver et ét-tal hvorved konstanten B er defineret.

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

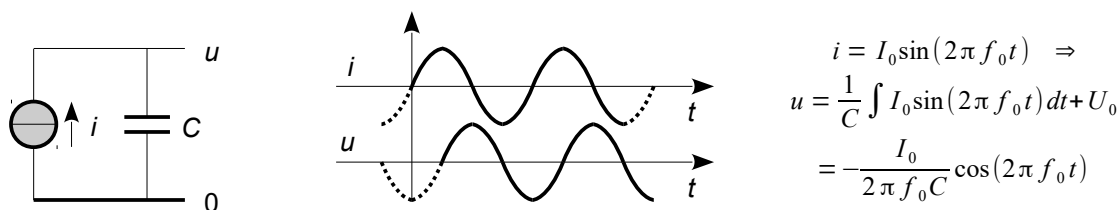
Til tiden $t = 0$ skal kondensatorens spænding være $u(0) = U_0$ hvilket giver konstanten A . Ved at løse ligningen for t kan man bestemme tiden for afladning til en given værdi. Som eksempel er der brug for tiden $t = 4,6 \cdot CR$ for afladning til 1 % af startværdien.

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

1.10.7 Oscillerende strøm

Når en strøm føres til en kondensator så lades den op og når strømmen er negativ så aflades den. Hvis strømmen er en svingning vil spændingen over kondensatoren også blive en svingning, men spændingen er ikke i fase med strømmen. Som vist side 47 skal strømmen integreres og det gennemgås herunder ved et harmonisk signal altså en svingning af typen sinus.

En strømgenerator giver herunder en harmonisk svingning med amplituden I_0 og frekvensen f_0 . Det antages at kondensatoren ikke har nogen ladning ved analysens start så $Q = 0$ og Integrationens konstant U_0 bliver nul. Resultatet er at kondensatorens spænding varierer med samme frekvens som strømmen, men fasen er 90° bagefter og amplituden aftager med stigende frekvens f_0 . Det leder i næste kapitel til en definition af begrebet impedans.

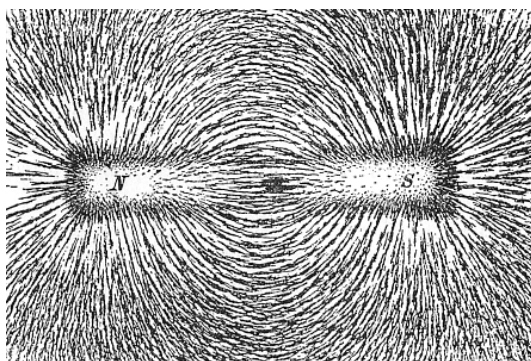


Figur 58 – En strømkilde driver en kondensator med en sinus-svingning. Spændingen over kondensatoren får samme form, men fasen er 90° bagud og amplituden afhænger af frekvensen.

Mange kredsløb baseres på integration. Eksemplerne er filterkredsløb der beskrives i tredje og fjerde kapitel samt måling af signalet fra sensorer hvor integration ofte er påkrævet for at få signalet bragt på en brugbar form. Kondensatoren er derfor en overordentlig vigtig komponent.

1.11 Magnetisk felt – B

Når elektronerne bevæger sig dannes der et magnetfelt. Det blev opdaget i 1820 af H. C. Ørsted der så at en elektrisk strøm kunne påvirke en kompasnål. Hans offentliggørelse startede et arbejde med deltagelse af Faraday, Ampere, Laplace, Biot og Svart for blot at nævne nogle få. Der kom snart mange anvendelser som telegrafene fra 1837 og elektriske maskiner fra 1860. J. C. Maxwell sammenfattede materialet til elektromagnetismen der blev publiceret 1873.⁴¹

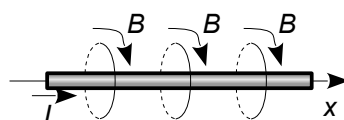


Figur 59 – Jernspåner i feltet fra en stangmagnet viser retningen af den magnetiske flux. Feltlinjerne går vedtægtsmæssigt fra nordpol til sydpol og metalspånerne samler sig som små magneter der klæber til hinanden med retning som små tangenter til linjerne (Wikipedia).

For en permanentmagnet løber feltlinjerne fra nordpol til sydpol udenfor magneten og de fortsætter inde i magneten fra sydpol til nordpol så de danner lukkede sløjfer og sløjferne krydser aldrig hinanden.

I lighed med det elektriske felt betegnes et magnetfelt 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 af elektromagnetismen så en ændring i feltet fra en elektrisk strøm i en ledning udbredes med lysets hastighed, men det vil jeg ikke komme nærmere ind på i denne bog.

Feltretningen fra nordpol til sydpol er en vedtægt helt som den elektriske strøms retning. Ligeledes har man besluttet at feltet højreskruer omkring en elektrisk leder. Det kan beskrives ved at en proptrækker skrues i strømmens retning. Griber man om lederen med højre hånds tommelfinger i strømmens retning angiver fingrene feltretningen (højrehåndsreglen). Det magnetiske felt er et hvirvelfelt og magnetfeltlinjerne danner lukkede sløjfer der aldrig starter eller ender ved en kilde for der er ingen magnetiske monopoler.⁴²



Figur 60 – 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.

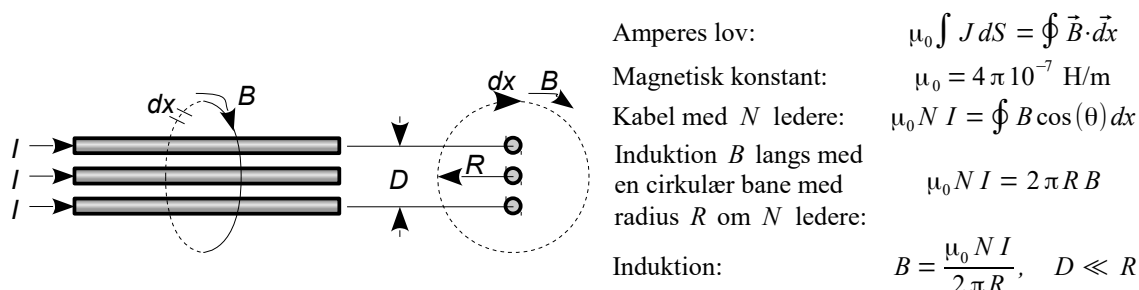
Et magnetfelts styrke betegnes ved den magnetiske induktion (*magnetic induction*) med B som symbol og enheden er tesla i SI systemet og defineret ved $T = \text{Wb/m}^2$ der igen kan udtrykkes som $T = \text{N/Am}$. Jordens eget magnetfelt er omkring $65 \mu\text{T}$ ved overfladen.

⁴¹ Brødrene Ramagnosi publicerede allerede i 1802 opdagelsen i italienske aviser. Senere genopdagede danskeren Ørsted i 1820 relationen mellem elektricitet og magnetisme og sendte en forskningsnote til samtidens fysikere. Fra 1861 til 1873 formulerede Maxwell elektromagnetismen som knyttede elektricitet og magnetisme sammen som én kraft. Lysets hastighed blev samtidigt fastlagt ved en naturkonstant ϵ for det elektriske system og en naturkonstant μ for det magnetiske som $c = 1/\sqrt{\epsilon\mu}$.

⁴² http://en.wikipedia.org/wiki/Magnetic_monopole.

1.11.1 Amperes lov

Den magnetiske induktion danner cirkulære baner omkring lederen med centrum på lederen og den magnetiske induktions værdi i et punkt kan findes ved Amperes lov. Loven angiver den magnetiske induktion langs med en lukket bane omkring lederen. Definitionen herunder er generel og vil blive simplificeret for brug i denne bog.



Figur 61 – Amperes lov for den magnetiske induktion B i afstanden R fra et kabelbunt med N ledere der hver bærer strømmen I . I afbildningen til højre løber strømmen ind mod papiret.

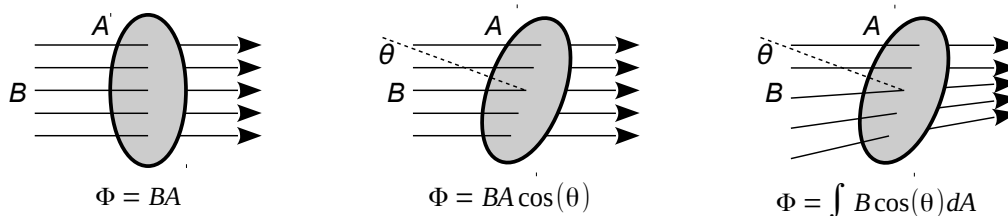
For tre lange ledere ($N = 3$) og en strømstyrke på $I = 1 \text{ A}$ i hver leder er den magnetiske induktion $B = 6 \mu\text{T}$ i afstanden $R = 0,1 \text{ m}$ hvilket er cirka 10 gange svagere end jordens magnetfelt.

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 lederens tværsnit giver det den samlede strøm uanset om strømtætheden J er konstant over arealet. Det første integral kan normalt skrives som $\mu_0 I$ hvor I angiver strømmen i lederen og μ_0 er den magnetiske konstant (Permeabilitet). Hvis løkken omfatter et bundt med N ledere hver med strømmen I så løses integralet til $\mu_0 NI$.

Det 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 dx for et linjeelementet der er en vektor og er givet ved tangenten til den valgte kurve. Hvis de to vektorer B og dx danner vinklen θ så bliver integralet B gange med cosinus til vinklen mellem vektor B og vektor dx . I stor afstand fra kabelbundtet kan løkken beskrives ved periferien af en cirkel med radius R og centrum i kabelbundtet. De to vektorer har samme retning så vinklen er nul og cosinus bliver én. Den samlede længde som dx integreres over bliver omkredsen på $2\pi R$.

1.11.2 Magnetisk flux

Den 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 ($\text{Wb} = \text{Vs}$). En magnetisk flux på 1 Wb gennem et areal på 1 m^2 giver en magnetiske induktion på 1 T .

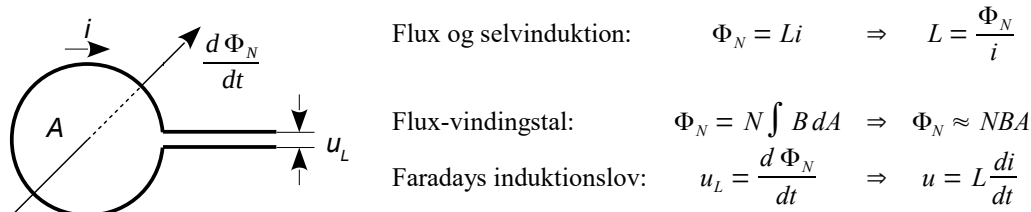


Figur 62 – Den magnetiske flux gennem et areal er givet ved den magnetiske induktion integreret over et areal under hensyn til variationen i den magnetiske induktion og vinklen med arealet.

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 vinklen mellem feltets retning og normalen til arealet. Hvis der ikke er en konstant værdi af induktionen skal den magnetiske flux beregnes ved integration.

1.11.3 Selvinduktion

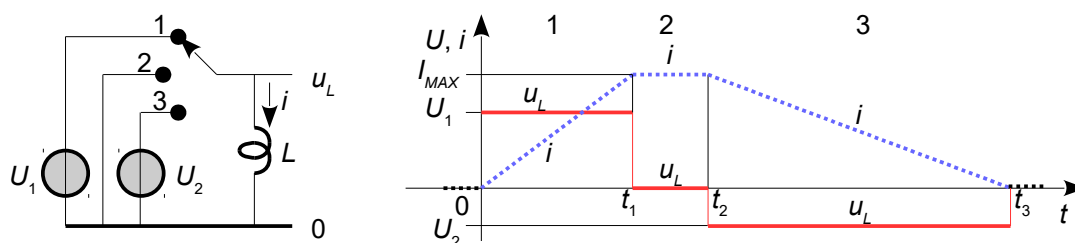
For en lukket elektrisk sløjfe defineres dens selvinduktion (*self induction*) som proportionaliteten mellem den magnetiske flux i arealet som lederen omslutter og strømmen i lederen. Den har fået symbolet L for at hædre Lenz og enheden er henry ($H = \text{Wb/A} = \text{Vs/A}$) for at hædre endnu en af elektromagnetismens mange bidragydere. For en spole med N tætliggende vindinger vil hver vinding bidrage med en flux og den samlede flux i spolens areal kaldes for flux-vindingstallet.



Figur 63 – En variation i den magnetiske flux gennem et areal giver anledning til en spænding over den leder der definerer den lukkede kurve.

Ifølge Faradays induktionslov vil en ændring i den magnetiske flux gennem et areal inducere en elektrisk spænding over den leder der afgrænser arealet. Hvis kredsen er kortsluttet genererer den inducerede spænding en strøm som skaber et modsat rettet magnetfelt der søger at modvirke ændringen i feltet så en magnet der falder gennem en kortsluttet løkke bremses.^{43,44}

Faradays induktionslov kan løses for strømmen som funktion af spændingen over spolens ender. Hvis spændingen over spolen ændres vil strømmen variere som det illustreres herunder. Den tykke røde streg er spændingen over spolen og den stiplede blå 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 ses bort fra tab.



Figur 64 – Strømmen i spolen er funktion af spændingen over den. Hvis spændingen er konstant vil strømmen ændres jævnt (retlinet). En momentan afbrydelse af strømmen kræver 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 $t = t_2$.

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

Med spolen forbundet til en positiv spændingskilde vil strømmen vokse lineært som $(U_1/L)t$ fra startværdien. Kortsluttes spolen vil strømmen forblive uændret (I_{MAX} i figurens eksempel 2). Med spolen forbundet til en negativ spændingskilde vil strømmen aftage lineært som $(U_2/L)t$.

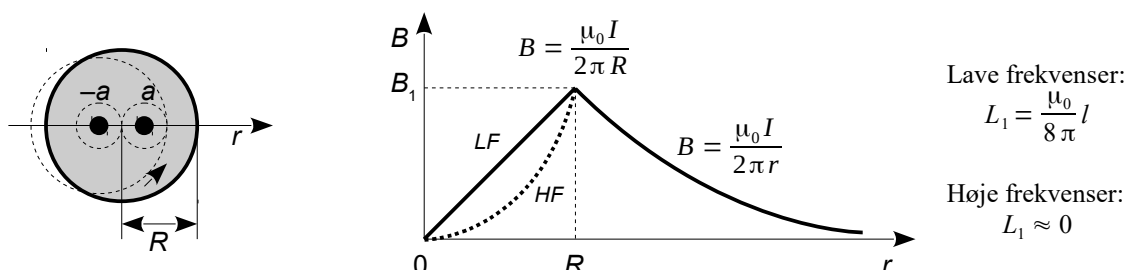
I det følgende vises nogle typiske eksempler på beregning af selvinduktionen.

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

44 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 faldet bremses og tiden forlænges på grund af den inducerede strøm i røret.

1.11.4 Enkeltleder – L_1

Der er et elektrisk felt inde i en ledning og det giver anledning til en selvinduktion på 50 nH for hver meter ledning. Ved en spole inkluderer det både tilledningen og spolens bevikling. Værdien gælder op til cirka 20 kHz for en tråddykkelse på 1 mm og aftager til nul ved cirka 1 MHz hvor strømmen stort set kun løber i overfladen og det indre af ledningen er næsten feltfri (se side 63).



Figur 65 – 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 lederens kerne og løber i et smalt område nær ved overfladen af tråden.

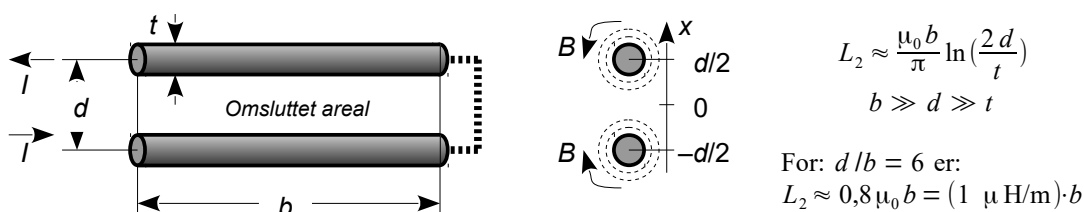
Strømmen i ledningen antages jævnt fordelt over hele lederens tværsnit og kan derfor opfattes som bestående af mange parallelle tråde hver med en brøkdel af strømmen. En given tråd og dens spejlbillede vil ligge i samme afstand fra centrum og felterne fra dette trådpar vil ophæve hinanden i centrum så lederens centrum er altid feltfri. Feltet vokser fra nul til det viste udtryk ved overfladen der er givet ved Amperes lov (se side 51) 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 ledningen ved a så det er kun tråden ved $-a$ og dermed kun halvdelen af strømmen der giver et bidrag til den magnetiske flux i ledningen⁴⁵. Den magnetiske induktion beregnes ved integration over et areal i ledningens længde l på $dA = l dr$.

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

1.11.5 Dobbeltleder – L_2

En elektrisk forbindelse med en signalvej og en returvej kan beskrives ved en dobbeltleder og det beskriver ikke blot tilslutningen til en spole, men også den elektriske installation i huset, telefon- og datakommunikation, ledningen til en højttaler og forbindelsen mellem komponenter på et printkort.



Figur 66 – To lange ledere i parallel omspænder et areal mellem de to ledere og derved dannes en selvinduktion. Det kan beskrive en almindelig dobbeltledning og signalvejen på et printkort.

Et kabel med afstanden $d = 6$ mm mellem lederne og tråddykkelsen $t = 1$ mm har selvinduktion $L_2 = 1 \mu\text{H}$ for $b = 1$ m længde. For 10 m længde er selvinduktion $L_2 = 10 \mu\text{H}$ og impedansen når 1Ω ved 16 kHz.

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

Her er b længden af lederen, d er afstanden mellem lederes centre og t er tråddykkelsen. Strømmen i lederne danner en magnetisk flux omkring hver leder og denne flux løber gennem det areal lederne danner hvilket giver en selvinduktion.⁴⁶

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 felterne fra de to ledere modsatrettet og feltet aftager med kvadratet på afstanden fra lederparret mens felterne i området mellem de to ledere har samme retning. Hvis de to ledere er lange i forhold til afstanden mellem dem ($b \gg d$) kan formelen for B side 51 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 $R = d/2 + x$ og tilsvarende er afstanden fra den øverste leder til punktet x givet ved $R = d/2 - x$. Feltet er i positionen x 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 $dA = bdx$ fra $x = -d/2 + t/2$ ved den nederste leder til $x = d/2 - t/2$ ved den øverste leder, hvor b 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].

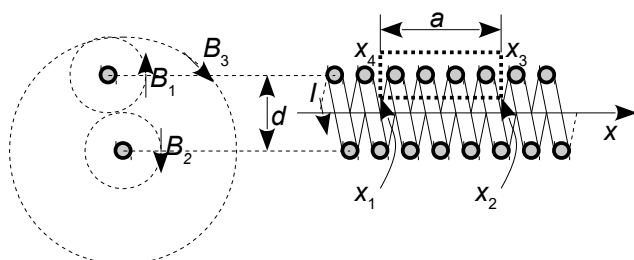
$$\Phi_N = N \int B_2 dA = -\frac{\mu_0 b d I}{2\pi} \int_{-\frac{d}{2} + \frac{t}{2}}^{\frac{d}{2} - \frac{t}{2}} \frac{dx}{x^2 - d^2/4} = -\frac{\mu_0 b d I}{2\pi} \left[\frac{1}{2(d/2)} \ln \left| \frac{x - d/2}{x + d/2} \right| \right]_{-\frac{d}{2} + \frac{t}{2}}^{\frac{d}{2} - \frac{t}{2}}$$

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 $-\ln(x) = \ln(1/x)$.

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

1.11.6 Udsnit af lang cylindrisk spole – L_U

Som introduktion til beregning af selvinduktionen for spoler vises et klassisk eksempel med lederen viklet i skrueform på en cylinder af ikke-magnetiserbart materiale. Formlen citeres ofte i lærebøger, gengives i formelsamlinger og florerer på internettet. Desværre er den ofte angivet som formelen for en spoles selvinduktion selv om den kun gælder for et udsnit af en uendeligt lang spole.



$$\begin{aligned} \text{Selvinduktion:} \quad L_U &= \frac{\mu_0 N^2 A}{a} = \frac{\pi}{4} \frac{\mu_0 N^2 d^2}{a} \\ \text{Generel formel:} \quad L_X &= \mu_0 N^2 d K \\ \text{Formfaktor:} \quad K &= \frac{\pi}{4} \frac{d}{a} \end{aligned}$$

Figur 67 – En udsnit af en meget lang spole der er viklet som en skrue. Udenfor spolen modvirker magnetfelterne hinanden, men inde i spolen har de den samme retning. Selvinduktionens formel gælder kun for et udsnit af spolen, men den er asymptotisk korrekt for meget lange spoler.

⁴⁶ I visse kredse indenfor lydengivelse 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 betydende som komponent. Det eneste der aftager ved at øge afstanden mellem lederne er kapaciteten mellem dem.

Hvis spolens længde er lig med radius vil formelen give en værdi på det dobbelte af det korrekte og fejlen bliver endnu større ved de helt korte spoler. Den følgende udledning er derfor kun ment som introduktion til en beregningsprocedure og de efterfølgende afsnit giver nogle bedre formler.⁴⁷

I spolens indre vil magnetfeltet fra hver snoning af lederen give et magnetfelt i samme retning som nabovindingerne så felterne understøtter hinanden (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 aftage med kvadratet på afstanden.

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 $\cos(\theta) = 0$. Strækningen fra x_3 til x_4 er så langt væk fra spolen at feltet antages at være nul. Der krydses nu tilbage til indersiden ved x_4 til x_1 , og igen er det vinkelret på feltet så vektorproduktet er nul. Der er derfor kun et bidrag fra det første stykke inde i spolen så integralet langs med den lukkede kurve giver aB . Da kurven omslutter N vindinger er den omsluttende 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 afgrænser 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.

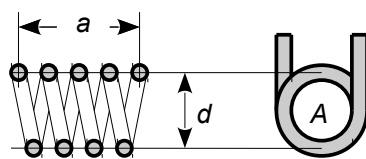
$$\Phi_N = N \int B dA = N B A$$

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.

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

1.11.7 Wheelers formel for en spole med ét lag – L_{W1}

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.



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

Generel formel: $L_X = K \mu_0 N^2 d$

Formfaktor: $K = \frac{\pi}{4} \frac{d}{a + 0,45d}$

Figur 68 – Wheelers formel for en spole viklet i ét lag med tynd tråd.

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.

Formlen angives at være korrekt indenfor $\pm 0,35$ % for spoler med en længde på mindst 40 % af diameteren, men det kræver at spolen vikles i et enkelt lag med meget tynd tråd så tykkelsen af beviklingen er nær ved nul. For en spole med flere lag vises en anden formel herunder. For korte spoler udvikles en formel på side 60. Der skal dog ikke forventes en nøjagtighed bedre end ± 10 % for en hjemmевiklet spole og det begrundes på side 61.

47 Se også: Thompson "Inductance Calculation Techniques" Part I og II, 1999 via linket: <http://www.thompsonrd.com/>.

1.11.8 Baggrund for Wheelers formel

Wheeler's formel for en spole i ét lag (se side 54) kan sandsynliggøres ud fra formelen for et udsnit af en spole ved nogle antagelser. For det første antages det at feltet udenfor spolens to ender ikke momentant falder til nul, men at det aftager lineært fra værdien inde i spolen til nul i en afstand lig med radius. Det giver en formel der er relativt tæt på Wheelers formel.

Feltet i spolens akseretning antages at aftage lineært til nul i afstanden R udenfor spolens to ender og feltet antages nul langs med spolens yderside. Integration langs med akse giver nu bidraget Ba fra det indre i spolen og to bidrag på $BR/2$ for hver ende af spolen. Divisor 2 skyldes at integrationen her beregner arealet af en trekant med højden R og bredden B .

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

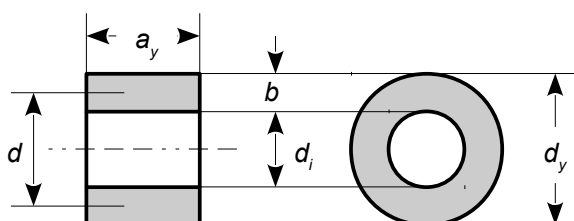
Derefter antages det at feltet inde i spolen ikke er helt oppe på den teoretiske værdi samt at feltet udenfor spolen ikke er nul. Her vælges en parameter α der formodes at være "lille" som relation for hvor meget feltet reduceres inden i spolen og hvor meget det er tilbage udenfor. Det har ingen klar fysisk begrundelse, men er en leg med formelen. For værdien $\alpha = 0,10$ findes Wheelers formel og for værdien $\alpha = 0,16$ findes en ganske brugbar formel for en cylindrisk spole der viser sig at være korrekt indenfor $\pm 10\%$ for en længde ned til 10% af diameteren.

Feltet i spolens indre antages til $(1-\alpha)B$ og til αB langs med ydersiden.

$$\begin{aligned} \mu_0 N I &= \oint B dx \\ \mu_0 N I &= a(1-\alpha)B + a\alpha B + R(1-\alpha)B \Rightarrow B = \frac{\mu_0 N I}{a+(1-\alpha)R} \Rightarrow L = \frac{\mu_0 N^2 A}{a+\frac{1-\alpha}{2}d} \Rightarrow L_{0,10} = \frac{\mu_0 N^2 A}{a+0,45d} \\ \mu_0 N I &= aB - a\alpha B + a\alpha B + RB - R\alpha B \\ \mu_0 N I &= [a+R(1-\alpha)]B \Rightarrow L_{0,16} = \frac{\mu_0 N^2 A}{a+0,42d} \end{aligned}$$

1.11.9 Wheelers formel for en spole med flere lag – L_{W2}

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

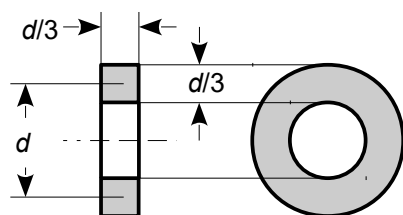


$$\begin{aligned} L_{W2} &= \frac{\pi}{4} \frac{8\mu_0 N^2 d^2}{9a+10b+3d} \\ d &= \frac{d_y + d_i}{2} \quad b = \frac{d_y - d_i}{2} \\ d_y &= D_y - t \quad d_i = D_i + t \quad a = a_y - t \end{aligned}$$

Figur 69 – Wheelers formel for en spole viklet i flere lag med ydre diameter D_y , indre diameter D_i og ydre længde a_y . Tråddykkelsen t kan måles med en skydelærer.

1.11.10 Brooks spole – L_B

Spolens DC modstand er proportional med længden af tråd så et optimalt design vil maksimere selvinduktionen for den kortest tråd. Det giver *Brooks coil* hvor formfaktoren er $K = 0,676$.

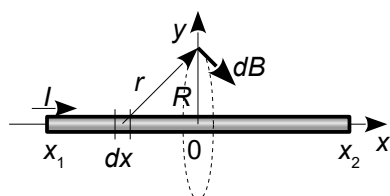


$$L_B = 0,676 \mu_0 N^2 d$$

Figur 70 – Brooks formel for en spole med lavest mulig DC modstand for en given selvinduktion.

1.11.11 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 analysere 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.



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

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

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

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

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

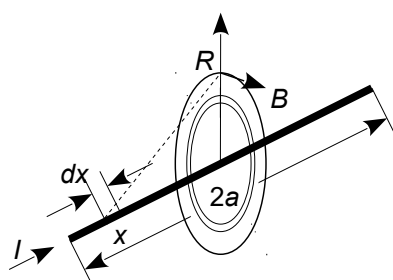
Figur 71 – Det magnetiske felt i et punkt kan beregnes ved at summere delbidrag fra lederen og det enkelte delbidrag beregnes efter Biot-Savarts lov.

Udtrykket hører ikke til de mest indbydende, men 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 leder med endelig længde er tæt på værdien som er givet ved Amperes lov for en uendelig lang leder.

1.11.12 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 så metoden efter Biot og Savart er velegnet for beregning af den magnetiske induktion ved korte ledningsstykker.



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

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

Figur 72 – 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 blot 10 % af lederens længde ($a/R = 5$) er feltet på 98 % af Amperes lov for en uendelig lang leder.

Det udledes herunder.

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 lederens midtpunkt for at

simplificere udtrykkene. Der tages ikke hensyn til at strømmen skal føres frem til og væk fra lederen for feltet herfra kan beregnes på samme måde. Linjestykket dx på lederen er placeret ved x og længden af vektor r beregnes af Pythagoras.

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

Magnetfeltets vektor $d\vec{B}$ 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 kvadratrods 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 $dx = R d\alpha$. 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 tælleren svarer det til at konstanten nu indeholder R i nævneren. Integrationen klares med et tabelopslag [RW-160].

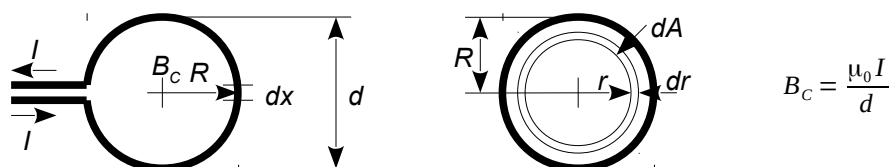
$$B = \frac{\mu_0 I}{4\pi} \int_{-a/R}^{a/R} \frac{R}{[R^2 \alpha^2 + R^2]^{3/2}} R d\alpha = \frac{\mu_0 I}{4\pi R} \int_{-a/R}^{a/R} \frac{1}{[\alpha^2 + 1]^{3/2}} d\alpha = \frac{\mu_0 I}{4\pi R} \left[\frac{\alpha}{\sqrt{\alpha^2 + 1}} \right]_{-a/R}^{a/R}$$

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

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

1.11.13 Feltet i centrum af en spole med én vinding

Hvis lederen bukkes til en cirkulær spole med en enkelt vinding vil den magnetiske induktion fra de mange delelementer dx adderes i cirkelns indre mens den magnetiske induktion i et punkt udenfor spolen vil udbalanceres på samme måde som ved spolen med to parallelle ledere på side 53. Den magnetiske induktion i centrum er π større end i afstanden R fra en lang leder (se side 51).



Figur 73 – Den magnetiske feltstyrke i centrum kan beregnes af Biot og Savarts lov.

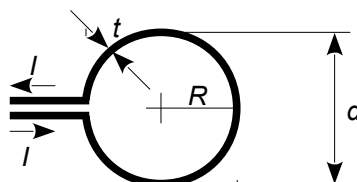
Den magnetiske induktion i centrum er givet ved Biot og Savarts lov som en sum af bidragene fra dx på cirkelperiferien. I afstanden $r = R$ fra centrum løber dx én gang rundt i cirklen svarende til en længde på $x = 2\pi R$. Vinklen mellem dx og r er $\theta = \pi/2$ (90°) så $\sin(\theta) = 1$.

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

Selvinduktionen kan estimeres hvis det antages at den magnetiske flux er konstant over arealet, men resultatet bliver for lavt da den magnetiske induktion stiger voldsomt tæt på tråden. Derfor udvikles et udtryk for den magnetiske induktion i arealet indenfor spolens afgrænsning.

1.11.14 Maxwells spole med én vinding

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. Litteraturen citerer ofte en bestemt formel for beregning af selvinduktionen for en cirkulær spole med en enkelt vinding som kaldes for Maxwells formel.⁴⁸



$$L_M = \frac{\mu_0 d}{2} \left[\ln\left(\frac{8d}{t}\right) - 2 \right] = \frac{\mu_0 d}{2} \ln\left(\frac{1,08d}{t}\right)$$

$$t < 0,03 d$$

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

Figur 74 – Selvinduktion af en spole med en enkelt vinding kaldes for Maxwells formel.

For en relativt stor løkke med diameteren $d = 200 \text{ mm}$ og trådtykkelsen $t = 1 \text{ mm}$ er $L_M = 675 \text{ nH}$.

Formlen gælder for strømmen løbende i overfladen af tråden som svarer til høje frekvenser. Ved lave frekvenser løber strømmen over hele lederens tværsnit og der skal kun subtraheres 1,75. Ved denne bog adderes selvinduktion fra tråden separat som L_1 og det ses at forskellen på 0,25 netop svarer til L_1 for en trådlængde lig med omkredsen af lederen.

$$L_1 = \frac{\mu_0}{8\pi} l = \frac{\mu_0}{8\pi} \pi d = \frac{\mu_0}{8} d = 0,25 \frac{\mu_0 d}{2}$$

Herunder vises et eksempel på hvordan en beregning af vindingens selvinduktion kan gribes an. Den magnetiske induktion er i centrum givet af Biot og Savart som vist øverst på denne side og i nærheden af tråden er den givet af Amperes lov som vist på side 51. De to udtryk kombineres til en formel for den magnetiske induktion ethvert sted i arealet indenfor spolens tråd. Ved hjælp af dette kan den magnetiske flux bestemmes ved integration over arealet og selvinduktion er da givet.

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. B_1 giver cirka 1/3 af B_C i centrum så B_C reduceres til cirka 2/3.

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

Den magnetiske induktion varierer med radius r . R er cirkelns 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 cirkelns 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 $1 - t/2R$.

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

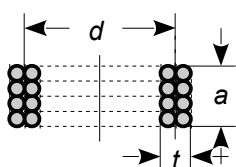
$$\Phi_N = \frac{\pi-1}{2R} \mu_0 I \left[r^2 \right]_0^{R-\frac{t}{2}} - \mu_0 I R \left[x + \ln |x-1| \right]_0^{1-\frac{t}{2R}}$$

Grænserne indsættes og det antages at trådtykkelsen er lille i forhold til cirkelns 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 i tælleren på $e^{0,071} = 1,073$ der er meget tæt på de 1,08 fra Maxwells formel. Udtrykket reduceres til det viste ved $d = 2R$.

$$\Phi_N = \mu_0 I R \left(\frac{\pi-1}{2} - 1 - \ln\left(\frac{t}{2R}\right) \right) = \mu_0 I R \left(0,071 + \ln\left(\frac{2R}{t}\right) \right) = \mu_0 I R \ln\left(\frac{2e^{0,071}R}{t}\right) = \mu_0 I R \ln\left(\frac{1,07d}{t}\right)$$

1.11.15 Kort spole – L_K

Formlen for L_M 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 vikling når gennem alle vindingernes arealer. Det er ikke realistisk alene på grund af trådens tykkelse så en del af feltet når ikke rundt om alle vindingerne. Konsekvensen er at den estimerede værdi bliver for høj, men det kan der korrigeres for ved at ignorere faktoren på 1,08 og erstatte trådtykkelsen med spolens dimension i længderetningen.



Selvinduktion: $L_K = \frac{\mu_0 N^2 d}{2} \ln\left(\frac{d}{a}\right) \quad t < a < \frac{d}{10}$

Generel formel: $L_K = K \mu_0 N^2 d$

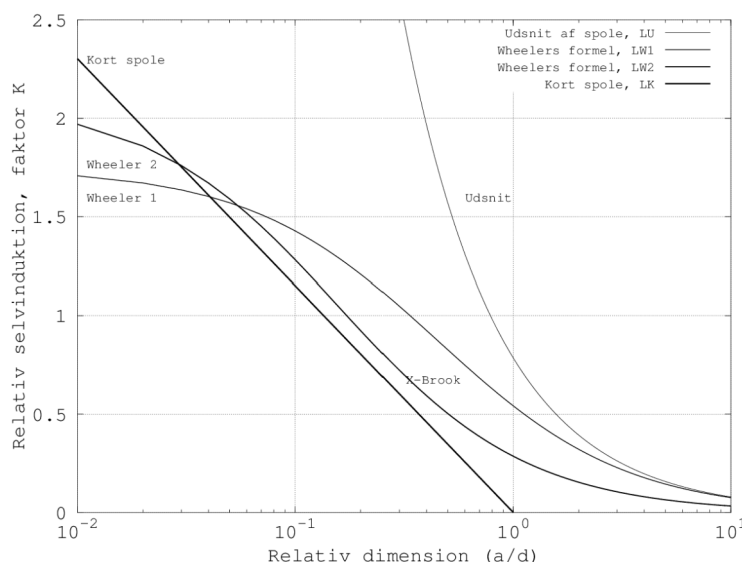
Formfaktor: $K = \frac{1}{2} \ln\left(\frac{d}{a}\right)$

Figur 75 – Formlen for en kort spole kan udvikles fra en skiveformet spole.

Som vist tidligere har jeg god erfaring med formelen for korte spoler, men det er ikke muligt at garantere en præcision i estimatet. En sammenligning mellem de viste formler som præsenteres herunder antyder at formelen ligger lidt lavere end de to formler fra Wheeler.

1.11.16 Sammenligning af formlerne

Formlerne dækker spoler med en længde der enten er meget længere end beviklingens radius eller meget kortere. Det betyder at der skal træffes et valg forud for beregning af selvinduktion for formlerne giver ikke samme værdi for samme spole og visse spoletyper kan slet ikke beregnes med det viste materiale. Wheelers formel L_{W1} er brugbar for spoler med en længde på mindst 40 % af diameteren og bevikling i ét lag mens Wheelers formel L_{W2} benyttes for spoler viklet i flere lag. For korte spoler kan L_K vælges når længden er under 10 % af diameteren. Brooks formel L_B er et bud på en spole hvor DC modstanden er minimal.



Generel formel:

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

$$K = 0,5 \dots 2 \text{ typisk}$$

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

Udsnit af spole:

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

Wheelers formler:

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

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

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 76 – Sammenligning mellem formlerne, der er normeret med $\mu_0 N^2 d$ og plottet som funktion af længden a i forhold til diameteren d . Brooks coil er vist som et kryds og passer med Wheelers formel L_{W2} for en tyk spole med $b = a$.

Som en slags "facitliste" for beregning af selvinduktion refereres ofte til et omfattende arbejde af Lorentz og Nagaoka der nærmest behandles med ærefrygt i mange referencer. 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 tilnærmelse alene ved at forsøge at sammenligne med Lorentz arbejde.

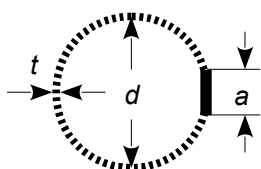
Materialet fra Lorentz og Nagaoka giver en for høj selvinduktion ved en kort spole idet et current sheet svarer til en tråd med tykkelsen nul. Her giver Maxwells formel et mere realistisk bud, men den tillader ikke flere vindinger så her er man på tynd is for den almindelige spole med 25 til 100 vindinger. Der er altså flere bud på en facitliste for spoler med realistiske dimensioner og der er desværre ikke en universel formel.

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 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 diameteren er det svært at undgå en deformation af spolen og det er endnu en mulig fejl af samme størrelsesorden. Anvendes en form til vikling af spolen skal man tage hensyn til både tråddykkelse og isolation og så kan en håndviklet spole ikke vikles lige så tæt som det måske var planlagt.

Min erfaring er at der ikke kan forventes en nøjagtighed bedre end $\pm 10\%$.

1.11.17 Selvinduktion af ledning

En populær tommelfingerregel siger at selvinduktionen er 1 μH per meter ledning. Jeg har set det demonstreret i praksis med et måleapparat, men selv om det tilsyneladende kan eftervises ved et eksperiment så er det forkert. Enhver ledning indgår i et lukket kredsløb så den målte selvinduktion skyldes et udsnit af den samlede løkke. Ved måling med et måleapparat indgår tilledningerne som en del af selve målingen, men måleapparatet er normalt indstillet til at vise en selvinduktion på nul med kortsluttede prober så det er kun ændringen fra kortsluttede prober til situationen med en kort ledningen indskudt der vises på instrumentets display.



$$\frac{L}{a} = \frac{L_M}{\pi d} = \frac{\mu_0}{2\pi} \ln\left(\frac{1,08 d}{t}\right) = 0,6 \dots 1,4 \mu\text{H/m}$$

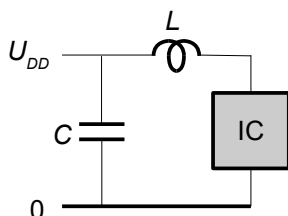
$$\Rightarrow \frac{L}{a} \approx 1 \mu\text{H/m}$$

Figur 77 – Tommelfingerreglen om 1 $\mu\text{H}/\text{m}$ for en kort ledning er god til overslagsberegning.

Med formelen for selvinduktionen af Maxwells spole kan selvinduktionen per meter beregnes ved at dividere L_M med omkredsen af spolen πd . Konstanten $\mu_0/2\pi$ er 0,2 $\mu\text{H}/\text{m}$ og da forholdet d/t mellem diameter og trådtykkelse ligger i området fra 20 til 1000 vil logaritmen give en værdi i det relativt begrænsede område fra 3 til 7. Selvinduktionen af en enkelt vinding er derfor fra 0,6 til 1,4 $\mu\text{H}/\text{m}$ med et gennemsnit på 1 $\mu\text{H}/\text{m} \pm 40\%$. Hvis måleopstillingen er af en størrelse, der kan ligge på et bord kan variationen fejlagtigt afskrives som en måleusikkerhed.

1.11.18 Eksempler

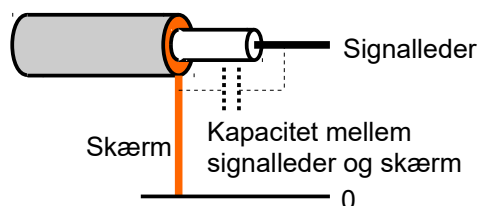
Selv om beregningen er usikker kan den dog give et fingerpeg om selvinduktionen af en ledning. Herunder vises en integreret kreds der modtager sin strøm gennem en kort ledning. Med 10 mm fra kondensatoren svarer det til 10 nH som har en impedans på 10 Ω ved 150 MHz. En digital komponent kræver en stabil effektforsyning så kredsløbet kan fejle ved pludselige strømtræk. En tommelfingerregel er at holde samtlige banelængder så korte som overhovedet muligt.



Lederbane: $a = 10 \text{ mm}$
 Serieinduktion: $L = 10 \text{ nH}$
 Reaktans: $X_L = 2\pi f L$
 $X_L = 10 \Omega$ ved 150 MHz

Figur 78 – En lederbane repræsenterer en selvinduktion så det er svært at undgå en impedans i serie med effektforsyningen.

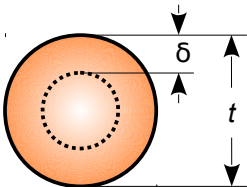
En lignende situation opstår når et skærmerkabel forbindes til et printkort ved at leder og skærm føres til skrueterminaler på printet. Impedansen ved skærmen tillader højfrekvente støjsignaler på skærmen at koble ind på signalvejen gennem kablets kapacitet. Løsningen er at skærmen skal fastgøres til et stelplan eller et metalkabinett med egnede konnektorer som SMA og BNC.



Figur 79 – Et skærmerkabel forbindes ofte til nul gennem en kort leder skabt ved at sno kablets skærm og skrue den fast til metalkabinettet. Det kaldes "pig tail" og er en klassisk fejlmontage.

1.11.19 Strømfortrængning

Ved jævnstrøm er strømmen jævnt fordelt over hele lederens tværsnit, men ved stigende frekvens fortrænges strømmen fra centrum af lederen og koncentrerer sig til et bånd ved overfladen af lederen. Det kaldes for strømfortrængning eller skineffekt. Det kan vises at 63 % af strømmen vil løbe indenfor et område kaldet indtrængningsdybden δ fra overfladen og ind i lederen som givet ved formelen herunder hvor lederens materialekonstant ρ kan aflæses på side 10.



$$\delta = \sqrt{\frac{\rho}{\pi \mu_0 f}} \Rightarrow \begin{array}{l} \delta_{\text{Cu}} = 2,1 \text{ mm} \\ \delta_{\text{Al}} = 2,7 \text{ mm} \\ \delta_{\text{Fe}} = 5,0 \text{ mm} \end{array} \quad \text{ved } f = 1 \text{ kHz}$$

Figur 80 – Feltfortrængning mindsker det effektive areal af lederen som fører strømmen.

For ledere af metal er indtrængningsdybden omkring 10 mm ved 50 Hz og det er noget man tager hensyn til ved et højspændingskabel. En stålkerne forøger trækstyrken af kablet og da der ikke løber en betydende strømstyrke i stålkerne betyder den ringe ledningsevne for stål kun et lille effekttab. For kabler til høj strømstyrke deles lederen i parallelle tråde for at udnytte materialet bedst muligt sammenlignet med et enkelt kabel med stort tværsnit.

Ved frekvenser i det hørbare område er indtrængningsdybden 0,5 mm eller højere så hele trådens tværsnit er aktiv for de typiske kabeldimensioner brugt i analog elektronik og højttalere og man behøver ikke normalt at tage hensyn til det.

Indtrængningsdybden er under 0,1 mm ved frekvenser over 1 MHz så i en radio vil strømmen løbe i trådens overflade og den reelle værdi af impedansen (modstandsværdien) stiger derfor. Ved høje effekter for radiotransmittere benyttes spoler med et tyndt lag sølv for at sænke effekttabet. Spolen kan tilmed være et rør med kølevæske inde i røret.

Havvand har en indtrængningsdybde på 0,2 m ved 1 MHz så radioforbindelse med en nedsænket undervandsbåd er vanskelig for feltet aftager med mere end 40 dB/m dybde. Ved en bærebølge på 100 Hz er indtrængningsdybden dog oppe på det mere brugbare 20 m, men der kan kun sendes en meget begrænset information ved den lave frekvens og undervandsbåden kan ikke svare på grund af antennens lave effektivitet og den krævede sendeeffekt.

Det kaldes for *ELF (Extremely Low Frequencies)* der dækker et frekvensområde fra 3 til 300 Hz. Antennen er væsentlig kortere end bølgelængden og ud over en elendig effektivitet vil den stråle i alle retninger så kommunikationen kan kun skjules ved brug af kryptering. Amerikanerne har arbejdet med transmission på 76 Hz (1968 til 2004) og russerne på 82 Hz hvor bølgelængden er omkring 3500 km. Indien vides også at kommunikere med undervandsbåde på denne måde.⁴⁹

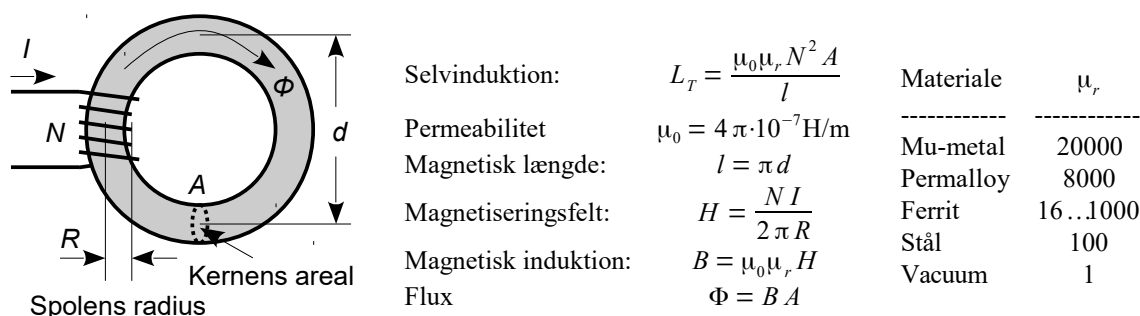
Amerikanerne benyttede frekvensmodulation hvor 72 Hz og 80 Hz svarer til 0 og 1 og hastigheden var under 1 bit/s. For at reducere fejl benyttes Reed-Solomon fejlkorrigerende kode og der kunne transmitteres 64 symboler som er nok til tal, punctuation og bogstaver og informationen krypteres. Signalet moduleres på en pseudo-tilfældig sekvens så det udsendte signal ligner støj. Den lave kommunikationshastighed tillader kun et begrænset antal kommandoer som primært har været at instruere en undervandsbåd om at gå til overfladen for almindelig radiokommunikation.⁵⁰

49 https://en.wikipedia.org/wiki/Communication_with_submarines.

50 <https://www.hep.wisc.edu/~prepost/ELF.pdf>.

1.12 Magnetiserbar kerne

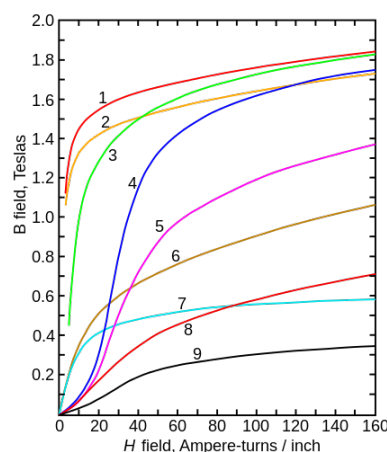
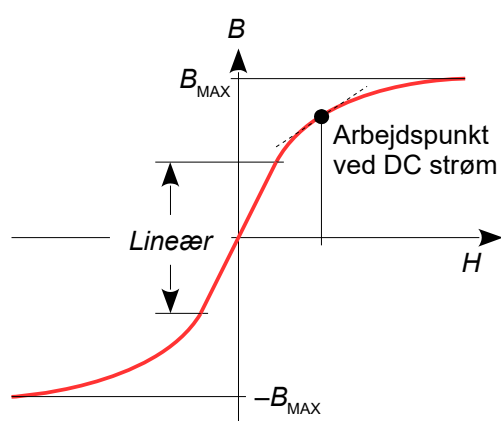
For at opnå en høj selvinduktion skal der anvendes fysisk store spoler med mange vindinger og det giver en stor DC modstand af tråden den vikles af. Selvinduktionen kan øges radikalt ved at benytte en kerne af et magnetiserbart materiale som fx jern og nikkel. Den feltforstærkende effekt beskrives ved konstanten μ_r der kaldes for den relative permeabilitet. Værdien er større end én for de magnetiserbare materialer, den er én for magnetisk neutrale materialer som luft, træ og plast mens eksempelvis grafit modsætter sig magnetisering og har en værdi under én.



Figur 81 – 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 et magnetfelt. Atomerne er arrangeret i domæner hvilket blev opdaget af Weiss i 1906 og indenfor et sådan 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 være en udbalancering så der ikke er et resulterende magnetfelt. Hvis der påtrykkes et eksternt magnetfelt (et H -felt) kan domænerne få energi nok til at skifte til samme retning som det eksterne felt og derved vil de ikke længere ophæve hinandens magnetfelter, men derimod virke feltforstærkende. Denne process virker indtil alle domæner peger i samme retning hvorefter feltet ikke forstærkes yderligere og kernen siges at være mættet (*saturated*).

Magnetiseringsfeltet H fra spolen beregnes fra Amperes lov på samme måde som B -feltet side 51. Enheden er A/m og feltet afhænger ikke af om der er en magnetiserbar kerne til stede i nærheden af spolen. Kernens feltforstærkende virkning skyldes at B -feltet i kernen forøger den resulterende magnetiske flux gennem spolen.

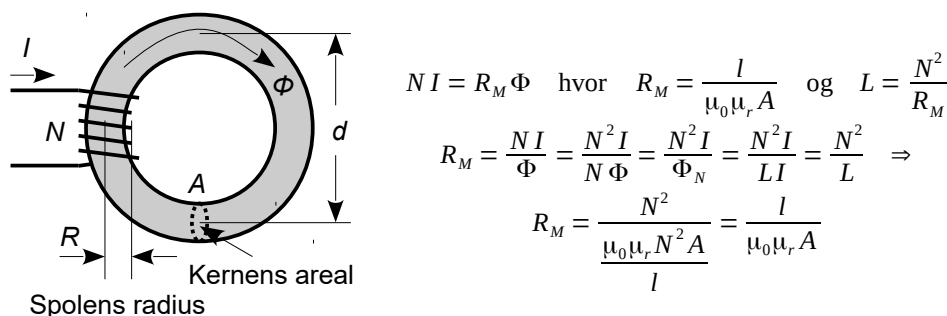


Figur 82 – Den magnetiske feltforstærkning er lineær ved moderat magnetisk induktion. Figur 1 til 6 viser forskellige typer stål, 7: Nikkel, 8: Kobalt og 9: Magnetit. Enheden for H -feltet er ampere per tomme (inch) hvor $1 \text{ A/inch} = 39 \text{ A/m}$. (Wikipedia)

Relationen mellem B og H felterne er lineær ved moderat magnetisering, men feltforstærkningen reduceres ved kraftig magnetisering. En spole der skal bære en stor DC strøm kan miste noget af sin selvinduktion fordi arbejdsunktet flytter fra det lineære område til krumningen.

1.12.1 Magnetisk modstand

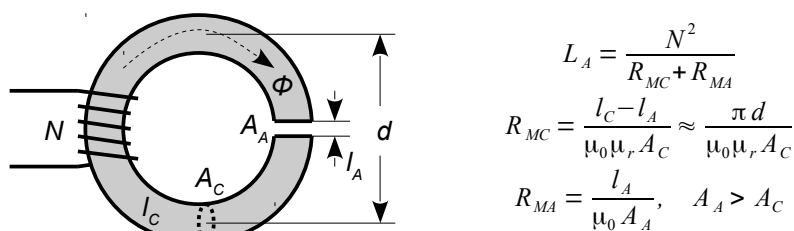
Den magnetiske modstand R_M som også kaldes reluktans introduceres 1888 af Oliver Heaviside. Hans idé var at produktet NI blev betragtet som en "motor" der driver den magnetiske flux Φ i de magnetiske kredsløb i analogi til strømmen i Ohms lov. Det magnetiske kredsløb beskrives ved en magnetisk modstand R_M der beregnes fra længden af det magnetiske materiale l som feltlinjerne følger, den feltforstærkende virkning af materialet μ_r og arealet A af materialets tværsnit. Herunder vises en beregning der anvender selvinduktionen L_U fra side 54.



Figur 83 – Et magnetisk system kan beskrives ved den magnetiske modstand.

1.12.2 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 da $\mu_r = 1$. Den magnetiske modstand består nu af en serieforbindelse med R_{MC} fra kernen og R_{MA} fra luftspalten. Selvinduktionen beregnes som vist herunder hvor index C står for core og A for air.



Figur 84 – Et kort luftgab indskyder en forholdsvis stor magnetisk modstand i kredsløbet.

For et materiale med $\mu_r = 1000$ vil en luftspalte på 0,1 % af kernens magnetiske vejlængde svare til reluktansen af kernen hvis forholdet A_A/A_C antages til én. Feltet spredes dog meget i luftspalten så et estimat baseret på $A_A = A_C$ vil kun være en grov tilnærmelse.

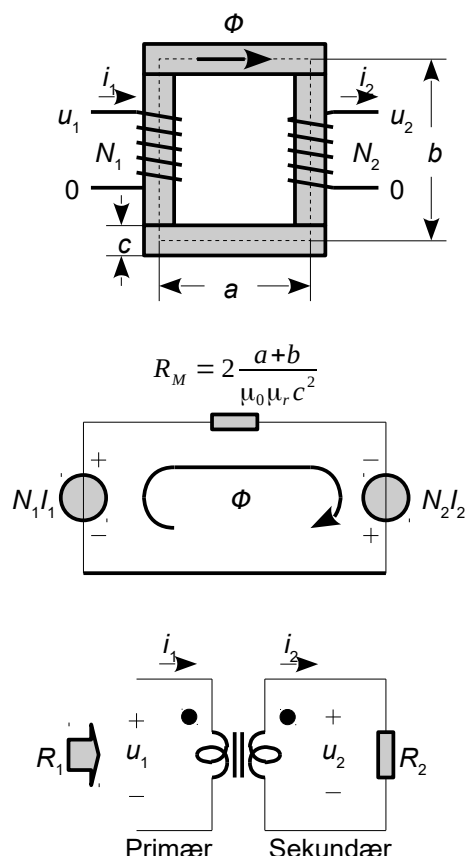
1.12.3 Hvirvelstrømme

Ved vekselstrøm dannes der hvirvelstrømme i kernen som giver et tab. For at undgå det udformes kernen med metallet i tynde strimler der arrangeres parallelt med B-feltet så der ikke kan dannes hvirvelstrømme i generende omfang. Metallet overfladebehandles (oxidation eller maling) så de enkelte strimler ikke bliver kortsluttet når kernen blades sammen.

Ved en ringkerne vikles det lange bånd til den toroide form og da båndene derved ligger en lille smule på skrå i forhold til den magnetiske flux introduceres en mindre luftspalte. Ved spoler med EI-kerne blades kernen sammen af metalplader standset i E-form og I-form hvor der skiftevis startes med E og I. Grænsen mellem E og I profilerne introducerer også en lille luftspalte. Et populært materiale ved høje frekvenser er ferrit hvor metallet er pulveriseret og opblandet i et keramisk materiale som kan støbes i den ønskede form. Da metalpartiklerne er uden elektrisk forbindelse kan tab fra hvirvelstrømmen holdes nede.

1.12.4 Transformator

Hvis det magnetiske system opbygges med flere spoler på samme kerne kan kredsløbsanalysens værktøj give den resulterende flux i kernen. Det vises herunder for et system med to spoler med vindingstallene N_1 og N_2 på en fælles kerne med arealet c^2 . Hvis spredningen af feltet ignoreres vil hele den magnetiske flux fra spole N_1 vil løbe igennem spole N_2 og ligeledes for den magnetiske flux fra spole N_2 . Det giver en relation mellem amplituden af spændingerne u_1 og u_2 over de to viklinger som kun er afhængig af antallet af vindinger N_1 og N_2 på de to spoler.



Magnetisk induktion i kernen:

$$B = \mu_0 \mu_r (H_1 + H_2) = \mu_0 \mu_r \frac{N_1 i_1 + N_2 i_2}{2\pi R} \ll B_{MAX}$$

Faradays lov:

$$u_1 = N_1 \frac{d\Phi}{dt} \quad \text{og} \quad u_2 = N_2 \frac{d\Phi}{dt}$$

Heraf omsætningsforholdet:

$$\frac{d\Phi}{dt} = \frac{u_1}{N_1} = \frac{u_2}{N_2} \Rightarrow \frac{u_1}{u_2} = \frac{N_1}{N_2}$$

Selvinduktion:

$$L_1 = \frac{\mu_0 \mu_r A N_1^2}{l} \quad \text{og} \quad L_2 = \frac{\mu_0 \mu_r A N_2^2}{l}$$

$$A = c^2 \quad \text{og} \quad l = 2(a+b)$$

Indgangsimpedans med ubelastet udgang:

$$Z_1 = 2\pi f L_1$$

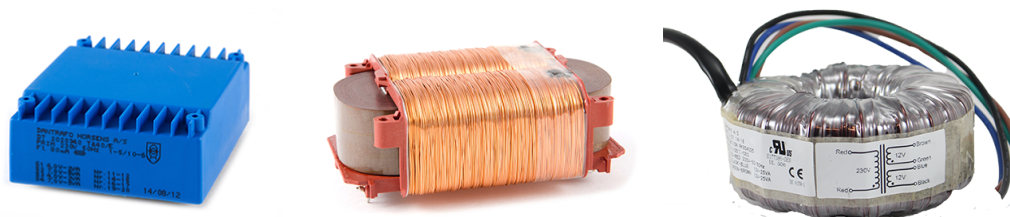
$$i_1 = \frac{u_1}{Z_1}$$

Indgangsmodstand med udgang belastet med R_2 :

$$R_1 = \frac{u_1}{i_1} = \frac{\frac{N_1}{N_2} u_2}{\frac{N_2}{N_1} i_2} = \left(\frac{N_1}{N_2}\right)^2 \frac{u_2}{i_2} = \left(\frac{N_1}{N_2}\right)^2 R_2$$

Figur 85 – To spoler på en fælles kerne kan overføre et AC signal fra den ene vinding til den anden vinding. Kernen symboliseres ved to parallelle streger som angiver en bladet kerne.

Kredsløbet udgør en transformator der kan omsætte (transformere) en spænding på indgangen til en spænding på udgangen med omsætningsforholdet N_2/N_1 der er givet alene ved antallet af vindinger på de to sider. Beregningen indeholder en differentiation fra Faradays induktionslov så det er ikke muligt at overføre jævnspænding. Som vist i illustrationen vil en belastning af udgangen med en modstand kunne ses fra indgangssiden som en transformeret værdi heraf.

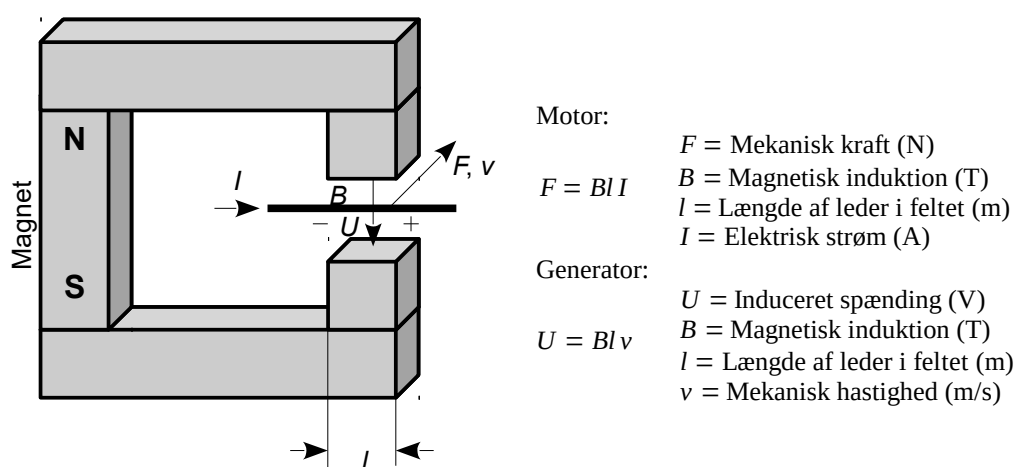


Figur 86 – Transformator for printmontage, med C-kerne og ringkerne (<http://www.dantrafo.dk/>).

1.13 Magnetiske kræfter

De magnetiske kræfter er overordentligt vigtige i vor moderne civilisation hvor elektriske maskiner omgiver os i alle applikationer. De benyttes i mobiltelefonens vibrator, tandbørster, barbermaskiner og ventilatorer dertil droner, biler og tog samt ved produktion af elektrisk strøm fra turbiner drevet af vand fra vandfald eller damp fra forbrænding af gas, olie eller affald.

Elektriske maskiner benytter en stationær del og en bevægelig del hvor energiomsætningen er baseret på magnetisk induktion. For at omsætte en strøm til mekanisk bevægelse skal strømmen ledes gennem et magnetfelt der giver en mekanisk kraft på lederen som så fører til en bevægelse gennem Newtons anden lov. Når lederen bevæges i magnetfeltet resulterer det i en elektrisk spænding over lederen. Det betyder derfor at når den elektriske strøm fører til en bevægelse så vil bevægelsen virke tilbage som et spændingsfald der kan reducere strømmen. En maskine kan af samme grund virke både som motor og som generator.⁵¹



Figur 87 – En elektrisk leder i et magnetfelt vil blive påvirket af en mekanisk kraft hvis der løber en strøm. Når lederen bevæges i magnetfeltet vil der induceres en spænding over lederen. Det antages at strømmen, den magnetiske induktion og bevægelsen danner et treretvinklet system.

Den enkleste form for motor er en elektrisk leder anbragt i det magnetiske felt fra en permanent magnet, men magneten kan udmærket være en spole der drives af en ekstern strømkilde. Det giver den frihed at den magnetiske induktion kan varieres.

Produktet af magnetfeltets induktion B og længden l af den del af lederen der befinder sig inde i magnetfeltet kaldes for kraftfaktoren Bl (*force factor*). For at opnå den bedste kraftoverførsel skal Bl produktet være så stort som muligt. Den magnetiske induktion kan være meget stor i metallet der leder feltet, men det er nødvendigt med et luftgab for at muliggøre bevægelsen og induktionen falder derved til en lavere værdi. Elektriske maskiner og højttalere benytter snævre tolerancer for designet af det nødvendige luftgab så den magnetiske induktion maksimeres.

Formlerne i illustrationen beskriver den ideelle situation hvor strømmen i lederen bevæger sig vinkelret på den magnetiske induktion. Kraften vil da have retning vinkelret på både strøm og spænding efter den såkaldte højrehåndsregel hvor tommelfinger peger i strømmens retning, pegefinger angiver den magnetiske induktions retning og langfinger viser kraftens retning. Det angives som et vektorprodukt ved $\vec{F} = \vec{B} \times \vec{I}$. Der gælder tilsvarende for den inducerede spænding at bevægelsen skal være vinkelret på både den magnetiske induktion og lederens retning.

Herunder følger et eksempel på den elektrodynamiske højttalers motor. Elektriske maskiner falder udenfor denne bogs område og vil ikke blive beskrevet yderligere.

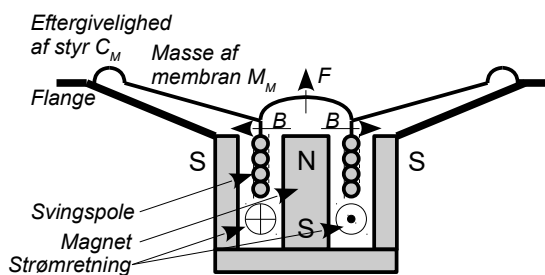
⁵¹ Se også: http://denstoredanske.dk/It,_teknik_og_naturvidenskab/Elektricitet/Elektriske_maskiner/elektriske_maskiner.

1.13.1 Den elektro-dynamiske højttaler

De to relationer for magnetiske kredsløb gør det muligt at forstå et kombineret elektrisk-mekanisk system som den elektro-dynamiske højttaler der vises i snit herunder. Afsnittet er af introducerende karakter og for en uddybende analyse henvises til litteraturen.^{52,53,54,55}

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 svingspole er monteret på membranen og befinder sig i et magnetisk felt 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 membran inkluderes i massen M_M . Der indgår flere komponenter i en højttaler, men det er tilstrækkeligt at koncentrere analysen om magnetsystemets kraftfaktor Bl og massen af det bevægelige system M_M som en introduktion.



$$F = Bl i = M_M \frac{dv}{dt} \Rightarrow v = \frac{Bl}{M_M} \int i dt$$

$$i = I_0 \sin(\omega t) \Rightarrow v = \frac{Bl I_0}{\omega M_M} \cos(\omega t)$$

$$\omega = 2\pi f$$

$$u = Blv = \frac{(Bl)^2 I_0}{\omega M_M} \cos(\omega t)$$

$$f > f_0 \quad f_0 = \frac{1}{2\pi \sqrt{M_M C_M}}$$

Figur 88 – Den elektro-dynamiske højttaler generere 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 formelen er gyldig.

Magnetfeltets styrke og længden af tråden på svingspolen udgør kraftfaktoren Bl der altid opgives i databladet og en typisk værdi er $Bl = 5 \text{ N/A}$ for en mindre højttaler hvor massen af membran og svingspole typisk er $M_M = 0,02 \text{ kg}$. Hvis der drives en strøm i svingspolen med amplituden $I_0 = 1 \text{ A}$ 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, men da accelerationen er konstant er lydtrykket det også.

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 eftergivlighed 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 eftergivligheden C_M udgør højttalerens resonansfrekvens f_0 der med eksemplets værdier er på 36 Hz .

52 Tore Skogberg "Elektroakustik", 2015 der kan hentes fra www.torean.dk.

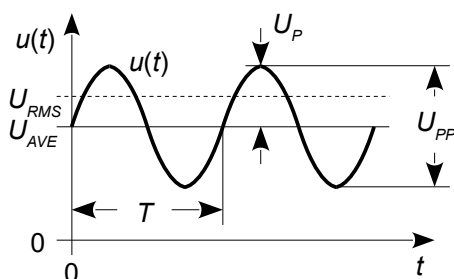
53 Knud Rasmussen "Højttalere", Danmarks Tekniske Universitet, Note 2109, 1996.

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

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

1.14 Signaler

Den harmoniske svingning (sinus) er den mest fundamentale svingning og danner basis for mere komplicerede svingninger der introduceres i næste kapitel. Svingningen beskrives ved frekvensen f_0 (frequency) som angiver hvor mange svingninger der gennemføres på ét sekund og topværdien U_P (amplitude). Signaler i et musikanlæg svinger omkring 0 V, men i apparater med mikroprocessor er der sjældent mulighed for negative spændinger og så benytter man i stedet en positiv DC værdi som signalet svinger omkring U_{DC} . Det kaldes også signalets middelværdi U_{AVE} (average value).



Harmonisk svingning: $u(t) = U_{DC} + U_P \sin(\omega_0 t)$

Vinkelhastighed: $\omega_0 = 2\pi f_0 = \frac{2\pi}{T}$

Amplitude: $U_P = \frac{U_{PP}}{2}$

AC del af svingning: $u_{AC}(t) = U_P \sin(\omega_0 t)$

DC del af svingning: $U_{AVE} = U_{DC}$

RMS uden DC: $U_{ACRMS} = \frac{U_P}{\sqrt{2}} = 0,71 U_P$

RMS med DC: $U_{DCRMS} = \sqrt{U_{DC}^2 + (0,71 U_P)^2}$

Figur 89 – En harmonisk svingning med definition af topværdi, effektivværdi med mere.

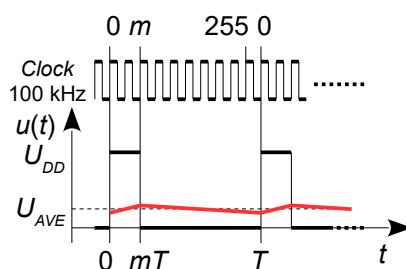
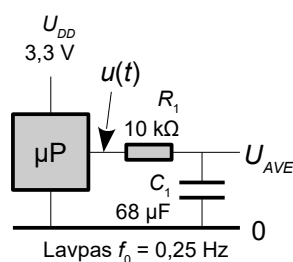
Signalets styrke betegnes med en effektivværdi U_{RMS} der angiver værdien af den DC spænding som ville afsætte den samme effekt i en belastning. For lysnettet er effektivværdien 230 V der svarer til en topværdi på 325 V så lysnettet svinger mellem -325 V som den nedre grænse og $+325$ V som den øvre. Vi har derfor $U_{AVE} = 0$ V, $U_{RMS} = 230$ V og $U_P = 325$ V. Akronymet RMS står for *root mean square* og svarer til 71 % af topværdien for en harmonisk svingning.

1.14.1 Middelværdi (AVE)

Et signals middelværdi (average) angiver den gennemsnitlige værdi af signalet over tiden T som kan være en enkelt periode af signalet, et helt antal perioder af signalet eller et tidsforløb der er "meget langt" i forhold til signalets periode hvilket uddybes lidt senere.

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

Det udnyttes i mikroprocessorer som en enkel måde at omforme en digital talværdi til en analog spænding gennem PWM der står for *pulse-width modulation*. Den digitale udgang skifter mellem HØJ og LAV over et tidsforløb og et analogt filter giver middelværdien af svingningen. Den digitale udgang er HØJ med værdien U_{DD} i m perioder af en *clock* og den er LAV resten af tiden. Perioden T udgør typisk 256 perioder af denne *clock*. Med *clock* på 100 kHz er $T = 2,56$ ms og lavpasfiltret skal have en tidskonstant $R_1 C_1$ på mindst $256 T$ for at udglatte signalet.



Tidsforløb:

$$u(t) = \begin{cases} U_{DD} & \text{for } 0 \leq t < mT \\ 0 & \text{for } mT \leq t < T \end{cases}$$

Udgangsspænding:

$$U_{AVE} = \frac{1}{T} \int_0^{mT} U_{DD} dt = m U_{DD}$$

Figur 90 – En mikroprocessor benytter en digital udgang til at danne en analog spænding ved brug af et eksternt lavpasfilter. Bemærk at spændingen U_{AVE} er proportional med både m og U_{DD} .

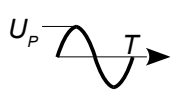
1.14.2 Effektivværdi (RMS)

Når der løber strøm i en modstand afsættes der effekt og modstanden varmes op. Det beskrives ved den effektive værdi af strømmen eller spændingen der vil afsætte den samme effekt som en DC strøm ville gøre. På engelsk kaldes det for RMS stående for *root mean square* som beskriver metoden i beregningen.

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

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

En sinus svingning med amplituden U_P er et typisk eksempel på beregning af effektivværdi. Kvadratet på en sinus kan ved tabelopslag skrives med cosinus til to gange vinklen. Integralet af cosinus giver nul over et helt antal perioder så er det kun den konstante faktor der resulterer i et bidrag. Med grænser indsat findes det velkendte udtryk hvor den effektive værdi er givet ved amplituden delt med kvadratroden af to som svarer til 71 % af amplituden.

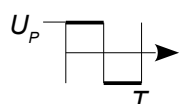


$$u(t) = U_P \sin(\omega t) \quad \omega = \frac{2\pi}{T}$$

$$U_{RMS} = \sqrt{\frac{1}{T} \int_0^T U_P^2 \sin^2(\omega t) dt} = \sqrt{\frac{1}{T} \int_0^T U_P^2 \frac{1 - \cos(2\omega t)}{2} dt}$$

$$U_{RMS} = \sqrt{\frac{1}{T} U_P^2 \left[\frac{1}{2} t \right]_0^T} = U_P \sqrt{\frac{1}{2}} = \frac{U_P}{\sqrt{2}} \Rightarrow U_{RMS} \approx 0,71 U_P$$

En firkantsvingning med amplituden U_P i halvdelen af intervallet og $-U_P$ afsætter lige så meget effekt som en DC spænding på U_P for det er lige meget hvilken vej strømmen løber.



$$u(t) = U_P \quad \text{for } 0 \leq t < T/2$$

$$u(t) = -U_P \quad \text{for } T/2 \leq t < T$$

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

1.14.3 Signal med DC og AC komponenter

Undertiden 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.

Signalet beskrives ved et DC niveau U_{DC} der er overlejret med et AC signal U_P . Udtrykket for $u(t)$ indsættes i formelen for beregning af RMS værdien.

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

Parentesen ganges ud og integrationen deles op.

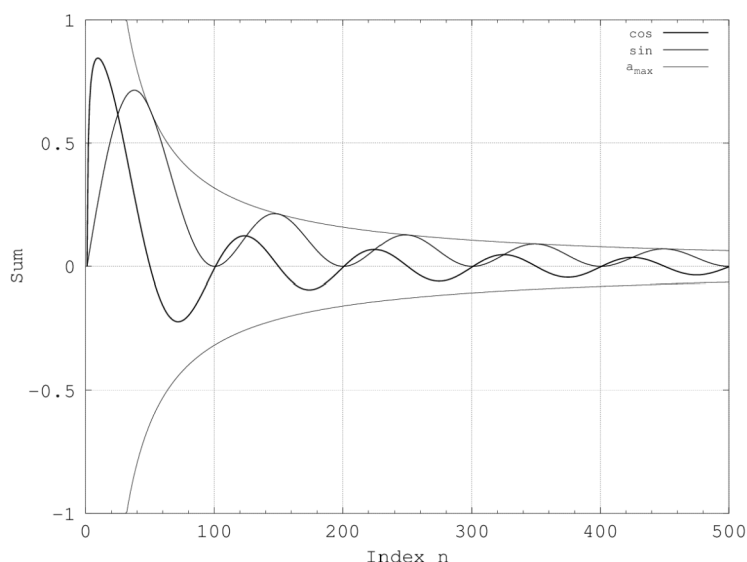
$$U_{RMS} = \sqrt{\frac{1}{T} \int_0^T U_{DC}^2 dt + \frac{1}{T} \int_0^T U_P^2 \sin^2(\omega t) dt + \frac{1}{T} \int_0^T 2 U_{DC} U_P \sin(\omega t) dt}$$

Første led er uproblematisk fordi spændingen U_{DC} ikke varierer over tid så integralet bliver U_{DC}^2 gange med tiden T og ved multiplikation med $1/T$ er der U_{DC}^2 tilbage. Andet led indeholder et integrale af kvadratet på en sinus der giver $1/2$ som kan skrives $(0,71 U_P)^2$ og altså er kvadratet på RMS værdien af U_P . Tredje led bliver nul fordi sinus integreres over en hel periode.

$$U_{RMS} = \sqrt{U_{DC}^2 + (0,71 U_P)^2}$$

1.14.4 Integrationstid

Eksemplerne fra de foregående afsnit benyttede integration over en periode af signalet, men ved praktisk brug skal man ofte måle på signaler hvis udseende først kendes efter endt dataopsamling. Man må så benytte sin viden om det kredsløb der måles på til at træffe en kvalificeret beslutning af hvor lang tid der skal måles over for at målefejlen bliver til at leve med. Det er ofte nok med fem perioder af den laveste frekvens.



Figur 91 – 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.

Kurverne herover viser beregning af RMS for sinus og cosinus der begge har 100 datapunkter for en periode af svingningen. Det ses at fejlen er stor efter én periode (index 0 til 100), men kommer under 10 % efter fem perioder (index 400 til 500). Det vil i mange situationer være tilfredsstillende med de fem perioder. For at nå under 1 % fejl skal der dog integreres over 32 perioder.

Der kan være mange frekvenser involveret så det er ikke givet at der overhovedet kan måles over et helt antal perioder. Dertil kommer at signalet kan variere over tid og at støjsignaler slet ikke har en periode. Hvis signalet har flere involverede svingninger kan man bruge den laveste frekvens.

Ved lydmåling er den lave grænse ofte på 10 Hz så her skal der integreres over mindst 0,5 s for en fejl på 10 % og mindst 3,2 sekunder hvis kravet er en præcision på 1 %.

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 kurven A_{max} for sinus. Ved at sætte et krav til A_{max} kan det nødvendige antal samples N beregnes. I udtrykket er $N_p = 100$ antallet af data i en periode af signalet og N er det samlede antal data.

$$A_{max} = \frac{N_p}{\pi N} \Rightarrow N = \frac{N_p}{\pi A_{max}}$$

N	Nødvendigt antal samples
N_p	Antal samples i én periode
A_{max}	Krævet fejlgrænse

For et krav om en nøjagtighed på 1 % skal der benyttes 32 gange så mange data som det periodiske signal opsamles over.

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

1.14.5 Misforståelsen RMS watt

Ved beregning af udgangseffekten fra en effektforstærker benyttes strømmens effektive værdi I_{RMS} ganget med spændingens effektive U_{RMS} og det ligger lige for at kalde produktet for RMS watt, men det er forkert. Det korrekte er angive middelværdien af effekten. Problemet er nok at det lyder mindre prangende end RMS watt som derfor benyttes bredt i brochuremateriale.

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

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

1.14.6 Numerisk beregning

Ved dataopsamlingsudstyr (*data acquisition system*) beregnes signalets middelværdi og effektive værdi ud fra data målt på diskrete tidspunkter. Signalet omsættes fra et tidskontinuerligt signal $u(t)$ til et tidsdiskontinuerligt u_n hvor signalet kun er kendt ved N diskrete værdier over tidsforløbet T .

$$\text{Dataopsamling: } u_n = u\left(\frac{nT}{N}\right), \quad n = 0, 1, 2, \dots, N-1$$

$$\text{Middelværdi: } U_{AVE} = \frac{1}{N} \sum_{n=0}^{N-1} u_n$$

$$\text{Effektivværdi: } U_{RMS} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} u_n^2}$$

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 u_n til tidspunkterne nT/N hvor n er et heltal. Det antages desuden at $u(t)$ varierer så langsomt indenfor tidsrummet T/N at det er tilladeligt at regne værdien for konstant indenfor intervallet mellem to måledata. Disse antagelser indsættes i definitionen af middelværdi og det udnyttes at integrationen kan splittes op i mindre stykker.

$$U_{AVE} = \frac{1}{T} \int_0^T u dt = \frac{1}{T} \left[\int_0^{T/N} u_0 dt + \int_{T/N}^{2T/N} u_1 dt + \dots + \int_{(N-1)T/N}^T u_{N-1} dt \right]$$

Integralet fra $t = 0$ til $t = T/N$ benytter den konstante værdi u_0 som integreres over tiden 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 indsat giver T/N . Følgelig er integralet at 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.

$$U_{AVE} = \frac{1}{T} \left[u_0 \frac{T}{N} + u_1 \frac{T}{N} + \dots + u_{N-1} \frac{T}{N} \right] \Rightarrow U_{AVE} = \frac{1}{N} \sum_{n=0}^{N-1} u_n$$

Helt tilsvarende beregnes den effektive værdi som følger.

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

1.14.7 Algoritme for middelværdi og RMS

For beregning af middelværdi og RMS ved en måleserie er det hensigtsmæssigt at dele opgaven op ved at beregne to talværdier U_{ADD} og U_{SQR} undervejs i måleserien og efter måleserien kan de to tal kombineres til beregning af middelværdien og den effektive værdi. Hver gang data er modtaget dannes en sum af måledata i variabelen U_{ADD} og en sum af kvadratet på måledata i variabelen U_{SQR} . Ved afslutning af måleserien beregnes de to parametre ved følgende enkle formler.

$$U_{AVG} = \frac{U_{ADD}}{N} \quad \text{og} \quad U_{RMS} = \sqrt{U_{SQR} - (U_{AVG})^2} \quad \text{hvor} \quad \begin{array}{l} U_{ADD} = \text{Sum af måledata} \\ U_{SQR} = \text{Sum af kvadratet på måledata} \end{array}$$

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

Formlen for RMS af et sammensat signal kvadreres og omroteres og kvadratroden tages på begge sider for at ændre formelen til beregning af RMS værdien af et AC signal uden DC.

$$\begin{array}{ll} \text{RMS med DC:} & U_{DC\,RMS} = \sqrt{U_{DC}^2 + U_{AC\,RMS}^2} \\ \text{RMS uden DC:} & U_{AC\,RMS} = \sqrt{U_{DC\,RMS}^2 - U_{DC}^2} \\ \text{Algoritme:} & U_{AC\,RMS} = \sqrt{\left(\frac{1}{N} \sum_{n=0}^{N-1} u_n^2\right) - \left(\frac{1}{N} \sum_{n=0}^{N-1} u_n\right)^2} \\ & U_{AC\,RMS} = \sqrt{U_{AVE}^2 - U_{DC}^2} \end{array}$$

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 sættes til det antal punkter der skal benyttes i algoritmen (4096). Derefter startes A/D konverteren og mikroprocessoren vil nu afvente data gennem interrupt.

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

Data læses fra A/D konverteren ind til variabel Vadc og variable Vavg (egentlig Vadd) og Vsqr opdateres 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 middelværdi og RMS.

```
StopConvert();
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();
```

1.14.8 Decibel

Elektronik og akustik gør udstrakt brug af signalers amplitude relateret til en reference som kan være indgangssignalet, udgangssignalet eller en underforstået værdi. Der dannes et forhold mellem signalets amplitude og referencen. Det bruges ofte for at beskrive menneskers opfattelse af tingenes tilstand. Et menneskes sanser reagerer på den relative ændring i en fysisk variabel og ikke på ændringen selv så en ændring på 10 % føles nogenlunde lige stor uanset om den fysiske ændring sker ud fra en lille eller stor påvirkning. Det kaldes for *Stevens' power law*.

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 der kan udtrykkes som kvadratet på spænding, strøm, etc. Enheden er 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 logaritmen til forholdet mellem to spændinger, strømme, lydtryk, etc.

Definition	Kommentar
$L = 10 \cdot \log_{10} \left(\frac{P}{P_0} \right) \text{ dB}$	Generel definition ved titals-logaritmen til et forhold mellem effekterne P og referencen P_0 .
$L = 10 \cdot \log_{10} \left(\frac{P}{P_0} \right) \text{ dBm}, \quad P_0 = 1 \text{ mW}$	Indenfor radiokommunikation benyttes "dBm" som absolut målestørrelse af udstrålet effekt med 0 dBm defineret som 1 mW.
$L = 20 \cdot \log_{10} \left(\frac{U}{U_0} \right) \text{ dB}$	Definition ved forholdet mellem to 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 styrke $P = P_0$ giver forholdet $P/P_0 = 1$ så logaritmen bliver nul ($L = 0 \text{ dB}$). En effektfordobling $P_1 = 2P_0$ svarer til 3 dB. Ved spændingsudtryk bliver det $U_1 = \sqrt{2} \cdot U_0 = 1,41 \cdot U_0$ der er 3 dB forøgelse i spændingens amplitude. En fordobling af spændingen giver 6 dB forøgelse der igen svarer til fire gange større effekt.

Ved brug af spænding frem for definitionens effekt udnyttes at relationen er $P = U^2/R$ og ved strøm $P = I^2 R$. Ved samme modstandsværdi for signal og reference bliver udtrykket 20 gange logaritmen til forholdet mellem spændingerne fordi $\log(x^2) = 2 \cdot \log(x)$.

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

Omregning fra decibel til den aktuelle amplitude findes ved først at dele med enheden, dernæst tage eksponentialfunktionen base 10 og til sidst gange med referencen. Som et eksempel en omregning fra lydtrykniveau L i decibel til lydtrykkets amplitude p i pascal.

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

Tabellen er ikke komplet og der er desværre en del uenighed om hvad det enkelte symbol står for så jeg giver nogle supplerende kommentarer herunder.

- Enheden dBm benyttes ved radiokommunikation for en udsendt effekt på 1 mW. En del referencer kalder enheden dBu for det samme og dBc angiver som 0 dB en reference til amplituden af en bærebølge (*carrier*).
- Indenfor lydformål står dBm for 1 mW effekt afsat i en belastning på 600 Ω . Det svarer til 0,775 V som er blevet en standard i lydverdenen uanset om det er brugt 600 Ω eller ej. Et referenceniveau i den professionelle lydverden er 4 dBm. Benyttes $U_0 = 0,775$ V ses det at svare til en amplitude af signalet på $U_0 \cdot 10^{4/20} = 1,23$ V.
- Selv om decibel er defineret fra et forhold mellem to størrelser så benyttes enheden også som en absolut angivelse af amplituden af signalet. Et eksempel er dBV der benytter en reference på 1 V så -6 dBV svarer til en signalamplitude på 0,5 V. Sjældnere ses en angivelse i forhold til 1 μ V som dB μ V eller blot dB μ hvor 1 dB μ er 1 μ V.
- Ved digital lydtransmission i fx Bluetooth headsets benyttes 0 dB for niveauet 3 dB under overstyringsgrænsen der er den maksimalt mulige amplitude af signalet. Hvis der benyttes en ordlængde på 8 bit er det maksimale udstyringsområde ± 127 og 0 dB bliver ± 90 for et harmonisk signal.

Målemetoden angives meget ofte i enheden hvilket er i modstrid med SI standarden der direkte siger at enheden ikke må afspejle målemetoden. Som vist ovenfor afspejler enheden netop den valgte reference og indenfor lydverdenen suppleres yderligere med brug af filtre. Et velkendt eksempel er dB(A) hvor lydtrykket refereres til 20 μ Pa, signalet filtreres gennem et A-vejningsfilter og RMS-værdien af signalet midles over et kort tidsrum. Så kan man vist ikke komme længere væk fra standarden, men decibel-enheden er slidt til ved praktisk brug og det er nok "som at kæmpe mod vindmøller" at forsøge at ændre på det forhold.

Der er en parallel til bit og byte i computerverdenen hvor informationsenheden ifølge SI er "bit". Reklameverdenen har skabt "byte" der ikke har nogen officiel definition, men normalt er på 8 bit. Dertil kilobyte der snart betyder 1.000 byte eller 1.024 byte og der er ikke enighed om hvorvidt præfix er k eller K eller sågar Ki. Ved datatransmission kan en hastighed opgives til "100 Mb" der kan være alt fra 100 Mbit/s og op til 105 Mbyte/s (da $1.024 \times 1.024 = 1.048.576$). Der er sågar problemer med præfix M som nogle referencer skriver som "meg".

1.15 Opgaver

1 – Benyt formlerne på side 43 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 $I_p = 1$ A og $T = 20$ ms ved brug af formlen fra det forrige spørgsmål. **Svar:** $Q_1 = 6,4$ mC.

3 – Brug en af formlerne side 44 til at beregne den kraft som ladningen Q_1 fra forrige spørgsmål vil påvirke en elektron med i afstanden $r_2 = 1$ μm . **Svar:** $F = 9,2$ nN.

4 – Brug formlerne side 9 til at beregne modstanden af en printkorts kobberbane og den afsatte effekt ved en strøm på $I = 2$ A. Banen har tykkelsen 35 μm , bredden 0,25 mm og længden 0,1 m. **Svar:** $R = 0,18$ Ω , $P = 0,71$ 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 13). 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 11) skal foretages ved en strøm på $I_0 = 1$ mA. Beregn modstanden af PT100 føleren ved $\theta = 100^\circ\text{C}$ og spændingsændringen ΔU over føleren ved formlerne side 9. **Svar:** $R_{100} = 129,3$ Ω og spændingsændringen er $\Delta U = 29,3$ mV.

7 – En strain gauge benyttes i en brokobling som vist på side 12. Påvirkningen af emnet ændrer strain gaugens længde med 1 %. Beregn ændringen i broens spænding. **Svar:** $U_1 - U_2 = 0,1$ 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 $U_T = 3,75$ V og $R_T = 0,13$ Ω . Beregn værdierne for I_T og R_T ved et Norton ækvivalent af batteriet. **Svar:** $I_T = 28,8$ A og $R_T = 0,13$ Ω .

11 – En kondensator på $C = 100$ μF er opladt til $U_0 = 5$ V, og oplades yderligere med $I = 1$ A over $\Delta T = 10$ μs . Beregn spændingen på kondensatoren efter endt opladning (se side 47). **Svar:** $U_0 = 5,1$ V.

12 – Kondensatoren aflades nu gennem en modstand på $R = 5,1$ k Ω . Beregn kredsløbets tidskonstant og hvor lang tid der går inden spændingen er aftaget til 51 mV. **Svar:** $RC = 0,51$ s og $t_{0,01} = 2,34$ s (se side 48).

13 – Elektret-kondensatormikrofonen fra side 35 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:** $R_D = 30$ k Ω og $S = 75$ mV/Pa.

14 – Udledningen af formlen for en kort spole på side 55 skal gennemføres som beskrevet i forklaringen. Sæt α til de angivne værdier (0,05 og 0,08).

15 – For Wheelers formel side 56 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 62 til tommelfingerværdien 1 μH per meter. Beregn værdien for en dobbeltleder på 10 m og sammenlign med beregningen på side 53 for dine egne værdier af parametrene d og t . **Svar:** Tommelfingerreglen giver 20 μH for to ledere og med $d = 10$ mm og $t = 1$ mm findes 12 μH .

17 – For en luftspole med 10 mm diameter og 10 vindinger med trådtykkelsen 0,1 mm, beregnes den magnetiske induktion i centrum ved strømstyrken 10 mA (se side 51). 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 64. **Svar:** $B = 4$ μT uden kerne og $B = 13$ mT med toroidkerne. **Tip:** Find den gennemsnitlige længde $l = 94$ mm, arealet af toroidens tværsnit $A = 78,5 \cdot 10^{-6}$ m², reluktansen af feltets vej $R_M = 955$ H⁻¹, den magnetiske flux $\Phi = 1,05 \cdot 10^{-6}$ A/H og den magnetiske induktion findes fra side 51 ved (igen) at dele med arealet af toroidens tværsnit.

18 – Beregn toroiden selvinduktion. **Svar:** $L = 105$ μH .

19 – Toroiden på forsynes med en luftspalte på 1 mm og selvinduktionen beregn (se side 65). **Svar:** $L = 99$ nH.

20 – LC kredsen på side 39 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 mH, så for at kunne benytte en 100 gange større værdi af kondensatoren (over 1 nF) bliver kravet til spolen $L < 25$ μH . Ved valget $L = 10$ μH findes $C = 2,5$ 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 41). Kobbertrådens har 0,1 mm diameter og den længste mulige tråd skal bestemmes (se side 9). **Svar:** $R < 1,3$ Ω , $a < 0,65$ m.

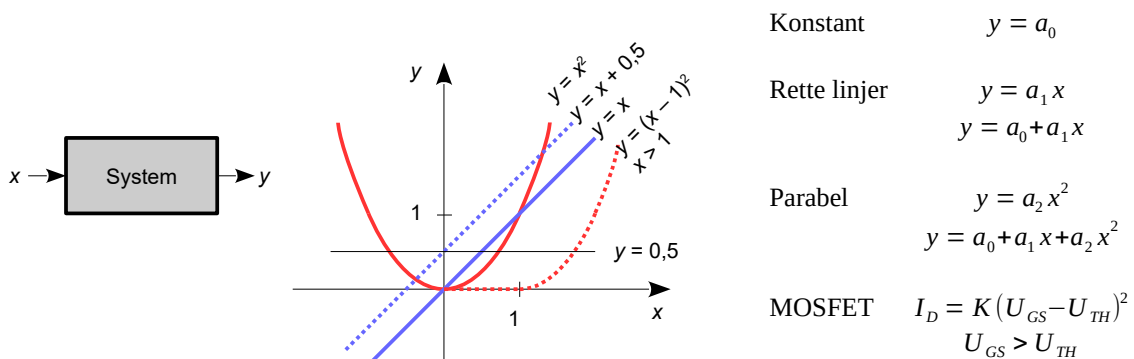
22 – Tråden vikles til en spole med diameteren 5 mm. Beregn det maksimalt antal vindinger, længden af en spole i to lag og spolens selvinduktion (se fx side 60). Beregn den nye værdi af kondensatoren for at fastholde frekvensen på 1 MHz. **Svar:** $N = 42$ vindinger, som rundes ned til $N = 40$. Længden bliver $a = 2$ mm, selvinduktionen bliver $L = 4,6$ μH , og der behøves $C = 5,5$ nF.

23 – Højttaleren på side 68 har en kraftfaktor på $Bl = 5$ Tm, massen af det bevægelige system er $M_M = 5$ g og styrenes eftergivelse er $C_M = 1$ mm/N. Beregn den frie resonans. Højttaleren drives ved 100 Hz med 1 A som strømamplitude. Beregn kraften F og hastigheden af membranen v . **Svar:** $f_0 = 71$ Hz, $F = 5$ N. $v = 1,6$ m/s.

2 ANALYSE

De passive komponenter skal repræsenteres ved deres komplekse impedanser for at beskrive hvordan de reagerer på oscillerende signaler. Vi får brug for Fourier serier og Laplace transformationen for at arbejde effektivt med signaler der abrupt skifter værdi. Endelig vises det at matematikkens værktøj kan benyttes på meget andet end elektriske kredsløb for der er et nært slægtskab med mekaniske, akustiske og termiske systemer som alle kan repræsenteres og analyseres ved elektronikkens værktøj som SPICE simulering.

Indenfor al videnskab opbygger vi matematiske modeller som beskriver hvordan et "system" vil opføre sig ved en given påvirkning. Vi kan betegne indgangssignalet x og systemets respons y og transformationen fra x til y kan da beskrives ved et matematisk udtryk. Den enkleste relation er en konstant der fx kan være en effektforsyning. Her vises det for $a_0 = 0,5$ med en tynd linje. Den rette linjes ligning beskriver en proportionalitet mellem to størrelser og er vist med en tyk linje hvor a_1 er den konstant der beskriver proportionaliteten. Et eksempel herpå er mediecentrets forstærker hvor amplituden af et signal øges så det kan drive en højttaler.



Figur 92 – Det grundlæggende værktøj for matematisk analyse.

Et kredsløb er lineært når fordobling i x fordobler y og når nulstilling af x medfører nulstilling af y . Et eksempel er Ohms lov $U = RI$ der angiver at spændingen U varierer med strømmen I og hvor R er proportionalitetens konstant.

Der er mulighed for at forskyde linjen i y -aksens retning ved at bruge konstanten a_0 som det vises i illustrationen med en stiptet linje hvor $a_0 = 0,5$. Relationen er ikke lineær for $x = 0$ giver $y = 0,5$ der er forskellig fra nul. Det er dog ikke det samme som at sige at indgangssignalet forvrænges for a_0 kan repræsentere en harmløs spændingsforskel i et elektrisk kredsløb hvilket er det almindelige. Det er mere interessant hvordan en variation i x (som $x + dx$) overføres til en variation i udgangen y (som $y + dy$) og kredsløbet kan opfattes som lineært hvis $dy = a_1 dx$ er opfyldt.

Parablen angiver en ikke-lineær sammenhæng der som regel mødes i et forsøg på at modellere et system med et matematisk udtryk. Hvis konstanten a_2 er lille vil anden-ordens leddet beskrive en lille afvigelse fra den rette linje og det viser da forvrængningen i det aktuelle kredsløb. Med sinus på indgangen $dx = \sin(\omega t)$ vil det ses i udgangssignalet som en skaleret version $dy = a_1 \sin(\omega t)$ plus en forvrængningskomponent der svinger på den dobbelte frekvens af indgangssignalet med a_2 som parameter. Dette emne uddybes side 118.

En anden-ordens relation mødes i en felleffekt transistor med strømmen i drain I_D som funktion af spændingen over gate-source U_{GS} . Det vises i illustrationen herover med stiptet linje. Det ses at der først kommer en strøm i transistoren når indgangssignalet er højt nok på grund af transistorens *threshold voltage* U_{TH} som i illustrationen er sat til én og U_{GS} tolkes som x . Konstanten K er givet af transistorens egenskaber og det behandles i femte kapitel.

2.1 Lineær algebra

Matematikken daterer sig tilbage til Babylon cirka 1800 år før Kristus hvor den retvinklede trekants sætning $a^2 + b^2 = c^2$ findes på nogle lertavler og 300 før Kristus introducerer Euklid bevisførelsen. Middelalderens fokus på talmystik udviklede matematisk stringens der blev vigtig i renæssancen hvor den objektive videnskab begyndte. Lidt før 1700 præsenterer Leibnitz differentialregningen og Newton grundlægger den moderne fysik. I dag løses mange problemer ved *finite-element* metoden der bygger på differentiaalligninger for beregning af eksempelvis betonkonstruktioners styrke.⁵⁶

Matematikens notation er slidt til ved praktisk brug hvilket kan oversættes til at der er en hel del underforstået viden i vores matematiske notation og de følgende afsnit introducerer disse meget væsentlige forudsætninger for brugen af de analytiske metoder.

2.1.1 Regneregler

Langt hovedparten af en ingeniørs opgaver kan repræsenteres ved eksemplerne på forrige side. Et system beskrives ved et konstant bidrag a_0 plus en forskydning ved a_1 plus noget ikke-lineært der beskrives ved a_2 . Problemet kan selvfølgelig være mere udbygget, men som oftest søger vi efter et simpelt analytisk udtryk.

Du forventes at være bekendt med de grundlæggende regneregler.⁵⁷

De grundlæggende regler dækker ombytning af rækkefølgen for sum og produkt, brugen af parenteser til at samle flere led til ét samt de love der afgør en beregnings rækkefølge:

For $ax^2 + b$ beregnes først kvadratet på x , dernæst ganges med a og til sidst adderes b .

Funktionsudtryk angives ved en betegnelse for funktionen (her f og g) og de uafhængige variable i parentes (her x og x, y). Resultatet af funktionen er en fysisk størrelse der er et tal gange med en enhed som fx $v = 50$ m/s.

Konstanter, variable og funktioner benyttes på samme måde med hensyn til lovene.

Udtryk kan omformes til noget nyttigt ved lovlige operationer uden at tabe information, men ikke alle operationer kan klare en enhed hos den variable fx logaritme og sinus.

Flere af de lovlige operationer stiller krav til både y og $f(x)$. Division med nul er ulovligt og variable skal være indenfor definitionsområdet som fx at logaritmen $\ln(x)$ kun er defineret for positive tal. Flertydighed ved kvadrering og kvadratrods skal respekteres: $2^2 = (-2)^2 = 4$.

Love

Kommutativ

$$x + y = y + x$$

Distributiv

$$xy = yx$$

$$x(y+z) = xy + xz$$

Hieraki

- (1) Potens og rod: x^2 og \sqrt{x}
 (2) Produkt og kvotient: xy og $\frac{x}{y}$
 (3) Sum og differens: $x + y$ og $x - y$

Operationer

Ligning

$$y = f(x)$$

$$z = g(x, y)$$

Addition

$$y + a = f(x) + a, \quad a = \text{konstant}$$

Multiplikation

$$by = bf(x), \quad b = \text{konstant}$$

$$b(y+a) = b(f(x)+a)$$

Eksponent

$$\exp(y) = \exp(f(x))$$

Differentiation

$$\frac{dy}{dx} = \frac{df(x)}{dx}$$

Integration

$$\int y dx = \int f(x) dx + a$$

Inversion

$$\frac{1}{y} = \frac{1}{f(x)}, \quad y, f(x) \neq 0$$

Logaritme

$$\ln(y) = \ln(f(x)), \quad y, f(x) > 0$$

Kvadrering

$$y^2 = [f(x)]^2, \quad y, f(x) \geq 0$$

Kvadratrods

$$\sqrt{y} = \sqrt{f(x)}, \quad y, f(x) \geq 0$$

Trigonometri

$$\sin(y) = \sin(f(x)), \quad 0 \leq y, f(x) \leq 2\pi$$

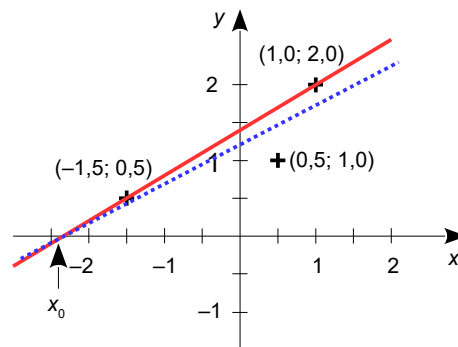
Regnereglerne danner grundlaget for resten af bogen.

⁵⁶ https://da.wikipedia.org/wiki/Matematikens_historie.

⁵⁷ http://www.gymnasiet.dk/media/2932/oversigt_algebraiske-regneregler.pdf.

2.1.2 Ret linje

En linje skal ofte bringes til at passe med nogle datapunkter fra en måling for at vise en tendens såsom hældningen a , skæringspunktet med y -aksen b eller skæringspunktet med x -aksen x_0 .



Ret linje

$$y = ax + b$$

Skæring med x -aksen ($y = 0$)

$$x_0 = \frac{-b}{a}$$

To datapunkter

$$\begin{aligned} y_1 &= ax_1 + b \\ y_2 &= ax_2 + b \end{aligned} \Rightarrow \begin{aligned} a &= \frac{y_1 - y_2}{x_1 - x_2} \\ b &= y_2 - ax_2 \end{aligned}$$

Mange datapunkter

$$\begin{aligned} a &= \frac{N \sum x_n y_n - \sum x_n \sum y_n}{N \sum x_n^2 - (\sum x_n)^2} \\ b &= \frac{\sum y_n - a \sum x_n}{N} \end{aligned}$$

Figur 93 – Ret linje gennem to punkter og lineær regression som bedste approksimation til en punktmængde.

En ret linje kan altid bringes til at skære gennem to punkter der ikke ligger samme sted. Ligningens konstanter a og b findes ved at opskrive ligningen to gange for de samhörende værdier hvor x og y kendes. Konstanterne a og b findes ved at trække ligningerne fra hinanden så b fjernes og vi kan beregne værdien af a som derefter benyttes til at finde b ved indsættelse i en af ligningerne.

Først benyttes to punkter $(x_1, y_1) = (-1,5, 0,5)$ og $(x_2, y_2) = (1,0, 2,0)$. Konstanterne a og b findes fra formelen og linjens skæringspunkt med x -aksen x_0 findes.

$$\begin{aligned} a &= \frac{0,5 - 2,0}{-1,5 - 1,0} = 0,6 \\ b &= 2,0 - 0,6 = 1,4 \\ x_0 &= \frac{-1,4}{0,6} = -2,33... \end{aligned}$$

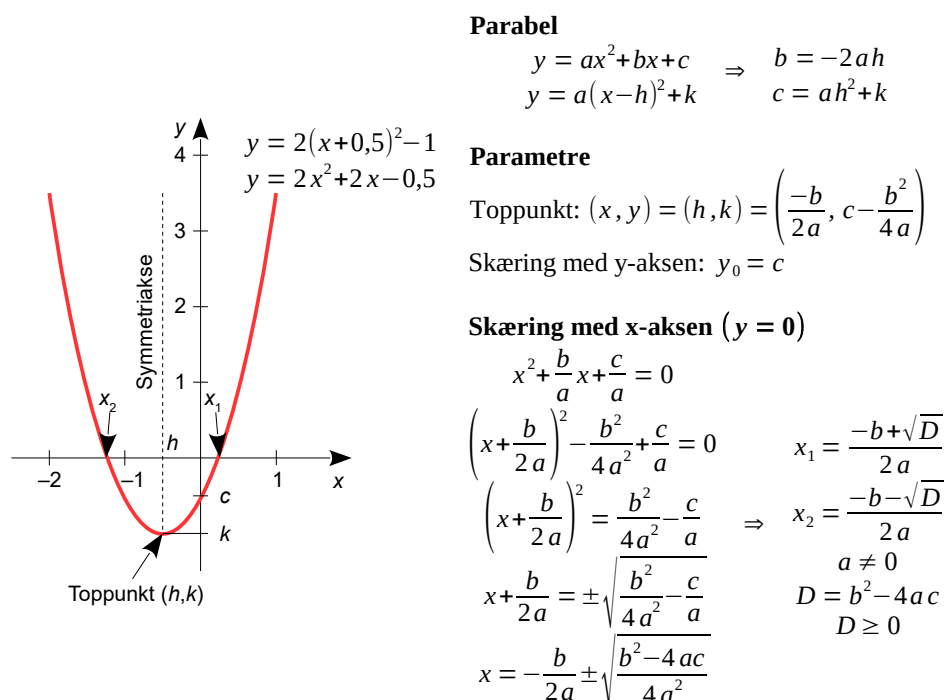
Ved flere punkter benyttes mindste kvadraters metode (*least squares fit*) der ikke skal udledes her. Metoden benytter det græske bogstav Σ (sigma) der betyder at data skal summeres. Det vil sige at Σx_n betyder $x_1 + x_2 + \dots + x_N$ og at $\Sigma x_n y_n$ betyder $x_1 y_1 + x_2 y_2 + \dots + x_N y_N$ hvor N er det totale antal af punkter. Anvendelsen vises herunder med et eksempel.

Med endnu et datapunkt $(x_3, y_3) = (0,5, 1,0)$ bestemmes koefficienterne ved mindste kvadraters metode hvor der opstilles fire summer. Desuden beregnes skæringspunktet x_0 med x -aksen på helt samme måde som for den rette linje gennem to punkter.

$$\begin{aligned} N &= 3 \\ \sum x_n y_n &= (-1,5)(0,5) + (1,0)(2,0) + (0,5)(1,0) = 1,75 \\ \sum x_n &= (-1,5) + (1,0) + (0,5) = 0,00 \\ \sum y_n &= (0,5) + (2,0) + (1,0) = 3,50 \\ \sum x_n^2 &= (-1,5)^2 + (1,0)^2 + (0,5)^2 = 3,50 \end{aligned} \Rightarrow \begin{aligned} a &= \frac{(3)(1,75) - (0,00)(3,5)}{(3)(3,50) - (0,00)^2} = 0,50 \\ b &= \frac{3,5 - (0,50)(0,00)}{3} = 1,166... \\ x_0 &= \frac{-1,166}{0,50} = -2,33... \end{aligned}$$

2.1.3 Parabel

Ved et kast med en sten vil stenen beskrive en bue der følger et udsnit af en parabel. Relationen blev fundet eksperimentelt af Galileo Galilei (1564 – 1642) og kaldes faldloven. Den blev senere udledt af Isaac Newton (1643 – 1727) baseret på opdagelsen af gravitation og differentialregning. Naturvidenskaben gør stor brug af parabler som beskrivelse af ikke blot bevægelse med konstant kraftpåvirkning, men også for at kunne beskrive en afvigelse fra en retlinet sammenhæng, for en simplificering af analytiske udtryk og for design af projektører og radioantenner.



Figur 94 – Parabel med definition af toppunktets koordinater og skæringspunkterne med akserne.

En parabel beskrives ved dens toppunkt (*vertex*) der ligger i punktet (h, k) , dens skæringspunkt med y-aksen som findes ved $x = 0$ og om hvorvidt den skærer x-aksen i punkterne x_1 og x_2 hvilket den gør hvis determinanten D er positiv eller nul. I sidste tilfælde er der kun én rod ($x_1 = x_2$). Dertil kommer om den har åbningen opad eller nedad der bestemmes af fortegnet for a med opad for positiv værdi. Symmetriaksen (*axis of symmetry*) er givet ved parameteren h .

For den viste parabel er toppunktet valgt til $(h, k) = (-0.5, -1)$ og konstanten til det kvadratiske led er valgt til $a = 2$ der er positiv så parablen krummer opad. Med de viste oplysninger kan parablens ligning opskrives.

$$\begin{aligned} y &= 2(x+0.5)^2 - 1 \\ y &= 2x^2 + 2x - 0.5 \\ D &= 2^2 - 4 \cdot 2 \cdot (-0.5) = 8 \end{aligned} \Rightarrow \begin{aligned} x_1 &= \frac{-2 + \sqrt{8}}{2 \cdot 2} = 0.2071... \\ x_2 &= \frac{-2 - \sqrt{8}}{2 \cdot 2} = -1.2071... \end{aligned} \Rightarrow h = \frac{x_1 + x_2}{2} = -0.5$$

Parablen skærer y-aksen ved $c = -0.5$ og dens skæringspunkter med x-aksen x_1 og x_2 beregnes fra de viste formler. De ligger begge to i samme afstand fra symmetriaksen ved $h = -0.5$.

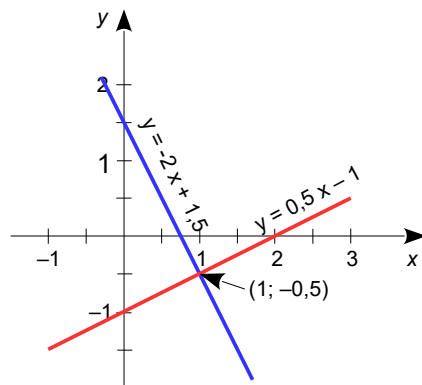
Parabler kan ligge ned så de er symmetriske om en akse der er parallel med x-aksen. Det er også muligt med skrå symmetriakse. At beskrive disse tilfælde ligger udenfor bogens sigte, men for den interesserede er det udførligt beskrevet i blandt andet Wikipedia.⁵⁸

58 <https://en.wikipedia.org/wiki/Parabola>.

2.1.4 Ligningssystem

En første-ordens ligning beskriver en ret linje i et xy-plot så med to første-ordens ligninger vil der beskrives to linjer. Et naturligt spørgsmål er så hvor de to linjer skærer hinanden. Ligningerne kan fx beskrive to processer hvor y stiger med x af én årsag (ligning 1) og falder med x af en anden årsag (ligning 2). Vi ønsker så at finde den værdi af x hvor begge processer giver lige store bidrag.

Herunder er vist to ligninger med x og y hvor skæringspunktet mellem de to rette linjer skal findes. Den første repræsentation har x og y på venstre side af lighedstegnet med en fast differens eller sum stående på højre side. Den anden repræsentation er i tråd med opfattelsen af ligningen som en funktion hvor y er den afhængige variabel.



Ligningssystem

$$\begin{aligned} (1): x - 2y &= 2 & \Rightarrow & y = 0,5x - 1 \\ (2): 4x + 2y &= 3 & & y = -2x + 1,5 \end{aligned}$$

Løsning ved substitution

$$\begin{aligned} (1a): x &= 2y + 2 \\ (1a) \rightarrow (2): y &= -2(2y + 2) + 1,5 \\ y &= -4y - 4 + 1,5 \\ 5y &= -2,5 & \Rightarrow & y = -0,5 \\ x &= 2 \cdot (-0,5) + 2 & \Rightarrow & x = 1 \end{aligned}$$

Løsning ved elimination

$$\begin{aligned} (2): y &= -2x + 1,5 \\ (1b): 4y &= 2x - 4 \\ (2) + (1b): 5y &= -2,5 & \Rightarrow & y = -0,5 \\ x &= 2 \cdot (-0,5) + 2 = 1 \end{aligned}$$

Figur 95 – Ligningssystem med to linjer der krydser hinanden.

Det er tre mulige veje for at finde det punkt hvor de to linjer skærer hinanden.

Ved **tegning af linjerne** i et xy-plot kan man aflæse skæringspunktet som det vises i illustrationen. Metoden er enkel, men aflæsning giver ikke en præcis værdi.

Ved **substitution** løses den ene ligning for x udtrykt ved y og dette udtryk indsættes i den anden ligning på den plads hvor der står x . Det giver en ligning i y som kan løses og den fundne værdi af y indsættes i den ene af ligningerne for at bestemme x .

Ved **elimination** ganger man den ene af ligningerne med en konstant så den ene af de to variable får samme talværdi men med modsat fortegn. Her er der $-2x$ i første ligning og $2x$ i den anden.

Ved at addere de to ligninger forsvinder x og rester forløber på samme måde som før.

2.1.5 Større ligningssystemer

Databehandling kan effektivt beskrives ved et system af ligninger. Det kan dreje sig om regulering af motorer, pumper og varmelegemer som følge af måling af tryk, temperatur og hastighed på forskellige steder i et anlæg. Et populært system er *state-variable* for regulering af et mekanisk system ud fra position og hastighed af den variable hvor reguleringen benytter et ligningssystem. Man kan også tolke billedbehandling og digitale filtre som et system af ligninger.

I denne bog vises systemer bestående af henholdsvis to ligninger med to variable og tre ligninger med tre variable. Det er en størrelse der kan håndteres manuelt. Ved større system af ligninger vil man altid benytte en computer med MATLAB, MathCad, Mable eller lignende.

Systemet beskrives ved en række afhængige variable (y_1, y_2, \dots) der er funktion af en række af uafhængige variable (x_1, x_2, \dots) med et større antal koefficienter a_{rs} hvor index angiver positionen i ligningssystemet med r for rækken og s for søjlen. I midten herunder vises helt det samme med de variable som søjlevektorer og koefficienterne som en matrix. Til højre vises en matrix-notation med de afhængigt variable som søjlevektoren Y , koefficienterne i matricen A og de uafhængigt variable som søjlevektoren X .⁵⁹

$$\begin{array}{l} y_1 = a_{11}x_1 + a_{12}x_2 \\ y_2 = a_{21}x_1 + a_{22}x_2 \end{array} \quad \text{versus} \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{versus} \quad Y = AX$$

$$\begin{array}{l} y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ y_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{array} \quad \text{versus} \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Formålet med et ligningssystem er at oversætte de givne x -værdier til de tilhørende y -værdier med koefficienterne som faktorer. Både x -værdier og koefficienter kan skifte værdi over tid. Ved digital signalbehandling af lyd opsamles lyden typisk med 44100 nye værdier hvert sekund så X -vektoren er 44100 elementer lang hvis den skal rumme ét sekunds dataopsamling. Y -vektoren bliver lige så lang og kan vise frekvensindholdet med 1 Hz spring mellem hvert element hvis ligningssystemet skal udføre en såkaldt FFT-analyse af lyden. Det uddybes senere i kapitlet (se side 126).

Matricen med koefficienter kan blive større end 1 million så det siger sig selv at det egner sig bedst for computere med en software algoritme eller ved en DSP chip hvor talbehandlingen kan afvikles direkte i hardware. Det er ikke noget man pusler med i hånden. Matrix-notationen har den fordel at den effektivt kan beskrive et vilkårligt stort system med en enkel notation.

En speciel matrice er identitetsmatricen I der har ét-taller i diagonalen og nul alle andre steder. Det betyder at y_1 er givet direkte af x_1 uden indvirkning fra de andre variable og tilsvarende for y_2 og y_3 .

$$\begin{array}{l} y_1 = x_1 \\ y_2 = x_2 \\ y_3 = x_3 \end{array} \quad \text{versus} \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{versus} \quad Y = IX$$

Det er en ganske brugbar form som vil blive benyttet i de følgende afsnit.

⁵⁹ Det er muligt at opbygge ligningssystemer med uens antal x og y variable, men det falder udenfor sigtet med denne bog. Jeg vil heller ikke komme ind på et systems robusthed overfor en variation i koefficienter eller variable.

2.1.6 Gauss elimination

Et ligningssystem kan bringes på identitetsmatricens form ved brug af regnereglerne fra side 78. Ligning 1 og 2 ganges med konstanterne $1/c_{11}$ og $1/c_{21}$ henholdsvis og efterfølgende subtraktion skaffer et nul i matricen for positionen a_{21} . Det gentages for at skaffe nul ved de andre positioner under diagonalen. Derefter skaffes nuller over diagonalen og om nødvendigt divideres for at reducere tallene til ét-taller i diagonalen. Løsningen er nu givet ved indholdet i Y-vektoren.

Metoden illustreres herunder, men det er meget sjældent at man gennemfører Gauss eliminationen i hånden. Det er relativt enkelt at opbygge en algoritme der reducerer et ligningssystem til $Y = IX$ og alle matematiske værktøjer har disse algoritmer indbygget.

Et lidt større ligningssystem med tre ligninger og tre ubekendte vises herunder.

Ligningssystem:

$$y_1 = 5x_1 - 5x_2 + 10x_3$$

$$y_2 = 2x_1 + 8x_3$$

$$y_3 = x_1 + x_2 + 5x_3$$

Løsningen eftersøges for:

$$(y_1 \ y_2 \ y_3) = (-25 \ 6 \ 9)$$

$$\begin{pmatrix} -25 \\ 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 5 & -5 & 10 \\ 2 & 0 & 8 \\ 1 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xRightarrow[(2) \times 1/2]{(1) \times 1/5} \begin{pmatrix} -5 \\ 3 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 0 & 4 \\ 1 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xRightarrow[(3)-(1)]{(2)-(1)} \begin{pmatrix} -5 \\ 8 \\ 8 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 1,5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xRightarrow{(3)-(2)} \begin{pmatrix} -5 \\ 8 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & -0,5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xRightarrow{(3) \times 4} \begin{pmatrix} -5 \\ 8 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xRightarrow[(3) \text{ times } -1/2]{(1)+(2)} \begin{pmatrix} -9 \\ 4 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xRightarrow{(1)+(2)} \begin{pmatrix} -5 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Hvis matricen har et nul i position a_{11} skal man ombytte to rækker. Det ændrer ikke på løsningen da man bare bytter om på rækkefølgen af de variable i vektorerne og matricen. Ligningssystemet er ikke ændret af ombytningen. Man kan også have fordel af at ombytte rækkerne hvis en talværdi i diagonalen er meget lille for at undgå problemer med regnenøjagtigheden.

2.1.7 Matrixregning

En matrice er en effektiv notation for en gruppe af tal hvis anvendelse er bredere end blot løsning af ligningssystemer. Matricer er en udvidelse af talbegrebet, men de opfører sig nogenlunde som de almindelige tal. Ganges en matrice med en fast faktor (en skalar) så ganges alle elementer i matricen med denne faktor. Matricer adderes eller subtraheres ligeledes element for element. Nogle specielle operationer er transponering hvor rækker og søjler ombyttes og determinanten der fortæller om ligningssystemets ligninger har en lineær afhængighed (determinanten er nul) eller om ligningssystemet ligninger er unikke (determinanten er forskellig fra nul).

Søjlevektor	$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
Rækkevektor	$X^T = (x_1 \ x_2)$
Vektorprodukt	$X^T X = (x_1 \ x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$
Matricer	$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ og $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$
Determinant	$\det(A) = a_{11}a_{22} - a_{21}a_{12}$
Transponering	$A^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$
Skalarprodukt	$kX = \begin{pmatrix} kx_1 \\ kx_2 \end{pmatrix}$
	$kA = \begin{pmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{pmatrix}$
Addition	$A+B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}$
Subtraktion	$A-B = \begin{pmatrix} a_{11}-b_{11} & a_{12}-b_{12} \\ a_{21}-b_{21} & a_{22}-b_{22} \end{pmatrix}$

Multiplikation af en matrix A med en vektor X giver AX der er en vektor med polynomier mens XA ikke kan defineres. Multiplikation af to matricer er kun muligt når antallet af rækker og søjler passer og for en uddybning heraf henvises til litteraturen. Her vises det for to kvadratiske matricer (2×2). Resultatet afhænger af rækkefølgen så AB er som hovedregel forskellig fra BA .

Multiplikation af matrix og vektor	$AX = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$
Multiplikation af matricer	$c_{ij} = \sum_{k=1}^N a_{ik} b_{kj} \quad AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$

Der findes ingen regneregler for division af matricer, men man kan definere en invers matrice A^{-1} der opfører sig som det reciprokke af et tal. Det betyder at i stedet for at dividere så må man gange med det inverse af matricen. Den inverse matrice kan findes på samme måde som ved Gauss elimination, men jeg anbefaler i at benytte MATLAB eller lignende.

Definition	$AA^{-1} = A^{-1}A = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
Invers matrice	$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$

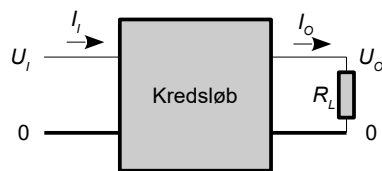
2.1.8 Ligningssystem

Et ligningssystem kan løses ved at benytte definitionen af det inverse af en matrice. Det ændrer ikke på løsningen til ligningssystemet at der ganges på begge sider med den samme matrice så det er tilladt at gange med A^{-1} på begge sider. Det giver vektor $A^{-1}Y$ på venstre side og $A^{-1}AX$ på højre side der kan simplificeres til IX og dermed står vektor X alene tilbage på højre side.

$$\begin{aligned} Y &= AX \\ A^{-1}Y &= A^{-1}AX \Rightarrow X = A^{-1}Y \\ A^{-1}Y &= IX \end{aligned}$$

2.2 Eksempel med BJT

Elektroniske kredsløb kan beskrives ved en 2×2 matrice H for spænding og strøm ved udgangen (U_o og I_o) som funktion af indgangen (U_i og I_i). Parameteren h_{11} er spændingsforstærkningen fra indgang til udgang, h_{22} er strømforstærkningen fra indgang til udgang, h_{12} er en modstand som omsætter indgangens strøm til udgangens spænding (når der ikke er spænding på indgangen) og h_{21} er en reciprok modstand som omsætter indgangens spænding til udgangens strøm (når der ikke løber strøm i indgangen).



$$Y = H X, \text{ hvor } Y = \begin{pmatrix} U_o \\ I_o \end{pmatrix} \quad H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \quad X = \begin{pmatrix} U_i \\ I_i \end{pmatrix}$$

$$\begin{aligned} U_o &= h_{11} U_i + h_{12} I_i \\ I_o &= h_{21} U_i + h_{22} I_i \end{aligned}$$

Parameterbestemmelse (talværdier for emitterfølger):

$$\begin{aligned} h_{11} &= \left. \frac{U_o}{U_i} \right|_{I_i=0} = 1 & h_{12} &= \left. \frac{U_o}{I_i} \right|_{U_i=0} = 0 \\ h_{21} &= \left. \frac{I_o}{U_i} \right|_{I_i=0} = 0 & h_{22} &= \left. \frac{I_o}{I_i} \right|_{U_i=0} = \beta \end{aligned}$$

Beregning af indgangsmodstand på basis af emitterfølger:

$$\begin{aligned} H &= \begin{pmatrix} 1 & 0 \\ 0 & \beta \end{pmatrix} \Rightarrow H^{-1} = \frac{1}{\beta} \begin{pmatrix} \beta & 0 \\ 0 & 1 \end{pmatrix} \\ X = H^{-1}Y &\Rightarrow \begin{pmatrix} U_i \\ I_i \end{pmatrix} = \frac{1}{\beta} \begin{pmatrix} \beta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} U_o \\ I_o \end{pmatrix} = \begin{pmatrix} U_o \\ I_o/\beta \end{pmatrix} \end{aligned}$$

$$Z_i = \frac{U_i}{I_i} = \frac{U_o}{I_o/\beta} = \beta \frac{U_o}{I_o} = \beta R_L$$

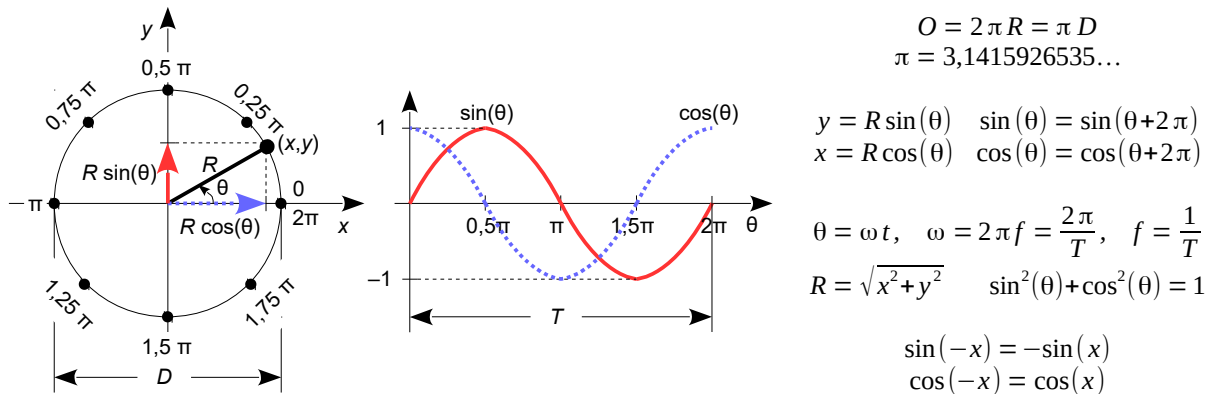
Figur 96 – Et kredsløb på matrix-form er her eksemplificeret ved en BJT som emitterfølger.

Illustrationen viser en transistor koblet som emitterfølger hvor udgangen følger indgangen ($h_{11} = 1$) på nær et spændingsfald på 0,7 V der her ignoreres og transistorens DC strøm I_{DC} ignoreres også så der analyseres kun AC signalbearbejdning. Udgangsspænding og -strøm er knyttet sammen ved Ohms lov som $U_o = R_L I_o$ og strømforstærkning er β . Kun parametrene h_{11} og h_{22} benyttes her så det antages at $h_{12} = h_{21} = 0$ vil give en brugbar beskrivelse og det viser sig at passe.

For at finde indgangsmodstanden på basis af transistoren skal ligningssystemet løses for de to indgangsvariable U_i og I_i så den inverse matrice for H skal bestemmes. Indgangsmodstanden kan herefter beregnes ved indgangsspændingen U_i delt med indgangsstrømmen I_i hvor resultatet er at indgangsmodstanden er βR_L som det også vises i femte kapitel, dog udledt på en anden måde.

2.3 Trigonometriske funktioner

Indenfor elektronik arbejder vi med svingninger og de beskrives elegant ved cirklen med ligningen: $x^2 + y^2 = R$. Et punkt på cirkelns periferi beskrives ved punktet (x,y) der defineres ved cosinus og sinus fra en vinkel som betegnes med det græske bogstav θ (theta). Når vinklen stiger fra nul til 2π er punktet ført hele vejen rundt langs periferien og er tilbage ved start. Punktet har tilbagelagt en strækning på cirkelns omkreds O der er 2π gange radius R eller π gange diameteren D . For en omdrejning i løbet af en periode på T er vinkelens hastighed $\omega = 2\pi/T$. Hvis frekvensen er $f = 1 \text{ kHz}$ bliver perioden $T = 1 \text{ ms}$ og vinkelhastigheden bliver $\omega = 6283 \text{ radianer per sekund}$.



Figur 97 – Cirklen med definition af de trigonometriske funktioner sinus og cosinus.

Afbildningen af punktet på x-aksen kaldes cosinus og afbildningen på y-aksen kaldes sinus og det definerer de to harmoniske svingninger cosinus og sinus. Disse to svingninger er fundamentale indenfor naturvidenskaben. Linjen fra $(0,0)$ til punktet (x,y) danner vinklen θ med x-aksen der er positiv i retning mod uret. Vinklen kan være funktion af tiden som $\theta = \omega t$ hvor t er tiden og kurverne bliver gentaget i det uendelige når vinklen θ når ud over 2π .

Nogle vigtige trigonometriske relationer [RW-127,128]:

$$\begin{aligned} \sin(\alpha + \beta) &= \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) \\ \sin(\alpha - \beta) &= \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta) \\ \cos(\alpha + \beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \cos(\alpha - \beta) &= \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) \end{aligned}$$

$$\sin(\alpha)\sin(\beta) = \frac{1}{2}[\cos(\alpha - \beta) - \cos(\alpha + \beta)]$$

$$\sin(\alpha)\cos(\beta) = \frac{1}{2}[\sin(\alpha - \beta) + \sin(\alpha + \beta)]$$

$$\cos(\alpha)\cos(\beta) = \frac{1}{2}[\cos(\alpha - \beta) + \cos(\alpha + \beta)]$$

$$\sin(2\alpha) = 2\sin(\alpha)\cos(\alpha)$$

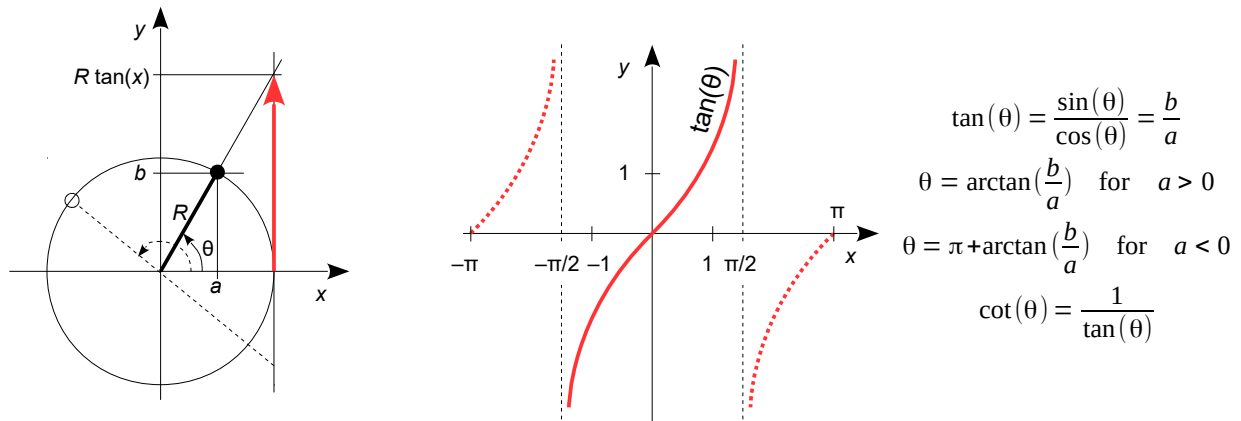
$$\cos(2\alpha) = 1 - 2\sin^2(\alpha)$$

$$\sin^2(\alpha) = \frac{1}{2}[1 - \cos(2\alpha)]$$

$$\cos^2(\alpha) = \frac{1}{2}[1 + \cos(2\alpha)]$$

$$\begin{aligned} a\cos(\alpha) + b\sin(\alpha) &= r\cos(\alpha - \phi) \\ a\sin(\alpha) + b\cos(\alpha) &= r\sin(\alpha + \phi) \end{aligned} \quad \text{hvor} \quad \begin{aligned} r &= \sqrt{a^2 + b^2} \\ \phi &= \arctan\left(\frac{b}{a}\right) \end{aligned} \quad \text{se næste side}$$

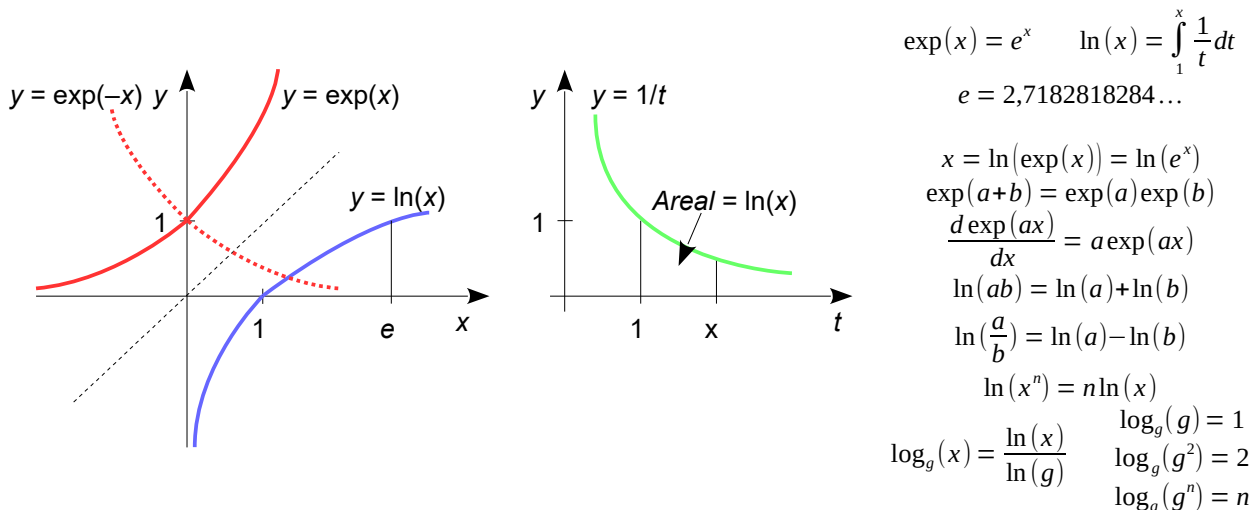
Relationen mellem siderne i en retvinklet trekant og den ene af vinklerne benytter tangens til vinklen udtrykt ved forholdet mellem den fjerne side og den hosliggende side. Den resterende vinkel findes fra den samlede vinkelsum på 180° . For en drejningsvinkel over 90° forlænges linjen til skæring med den lodrette linje gennem $(1, 0)$ og det giver en negativ værdi med den valgte retning for positiv vinkel. Bemærk at tangens er oscillerende og gentages med en periode på π .



Figur 98 – Cirklen med definition af den trigonometriske funktion tangens.

2.4 Logaritmiske funktioner

Mennesket opfatter ikke ændringer på en absolut skala, men derimod relativt. Der skal derfor være et referenceniveau som ændringen tager udgangspunkt fra.⁶⁰ Du kender givetvis decibel (dB) som angiver hvordan mennesket opfatter ændring i lydets styrke. Dels derfor, men også på grund af fundamentale fysiske forhold ender næsten al videnskab med logaritmiske udtryk.



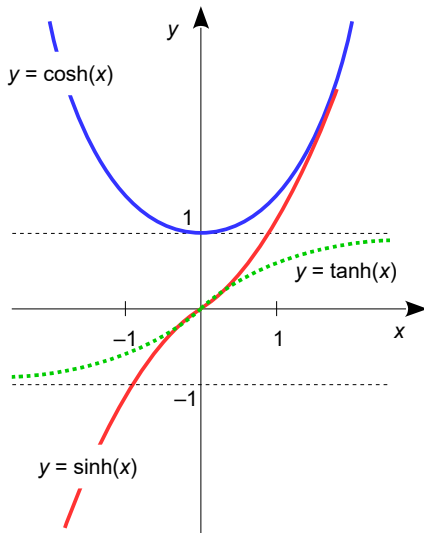
Figur 99 – Eksponentialfunktionen og logaritmefunktionen.

Illustrationen benytter e som grundtal hvilket kaldes den naturlige logaritmefunktion $\ln(x)$. En anden mulighed er grundtallet 10 som $\log_{10}(x)$ der benyttes ved decibel og pH-værdi. En logaritme er kun defineret for positive værdier af x og går imod minus uendelig for x gående mod nul uanset grundtallet og alle logaritmer giver værdien 0 for $x = 1$.

60 https://en.wikipedia.org/wiki/Stevens'_power_law.

2.4.1 Hyperbolske funktioner

De hyperbolske funktioner defineres ud fra hyperblen med ligningen $x^2 - y^2 = 1$ der giver udtryk med eksponentialfunktion. Navngivningen skyldes at de hyperbolske funktioner har egenskaber der er analoge med de trigonometriske funktioner selv om de ikke er oscillerende. De hyperbolske benyttes af og til ved løsning af differentialligninger og den hyperbolske tangens dukker op ved en beskrivelse af en operationsforstærkers indgangstrin.



$$\sinh(x) = \frac{\exp(x) - \exp(-x)}{2}$$

$$\cosh(x) = \frac{\exp(x) + \exp(-x)}{2}$$

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$$

$$\Rightarrow \begin{aligned} \cosh(x) + \sinh(x) &= \exp(x) \\ \cosh(x) - \sinh(x) &= \exp(-x) \end{aligned}$$

$$\sinh(x \pm y) = \sinh(x)\cosh(y) \pm \cosh(x)\sinh(y)$$

$$\cosh(x \pm y) = \cosh(x)\cosh(y) \pm \sinh(x)\sinh(y)$$

$$\sinh(x)\sinh(y) = \frac{1}{2}[\cosh(x+y) - \cosh(x-y)]$$

$$\sinh(x)\cosh(y) = \frac{1}{2}[\sinh(x+y) + \sinh(x-y)]$$

$$\cosh(x)\cosh(y) = \frac{1}{2}[\cosh(x+y) + \cosh(x-y)]$$

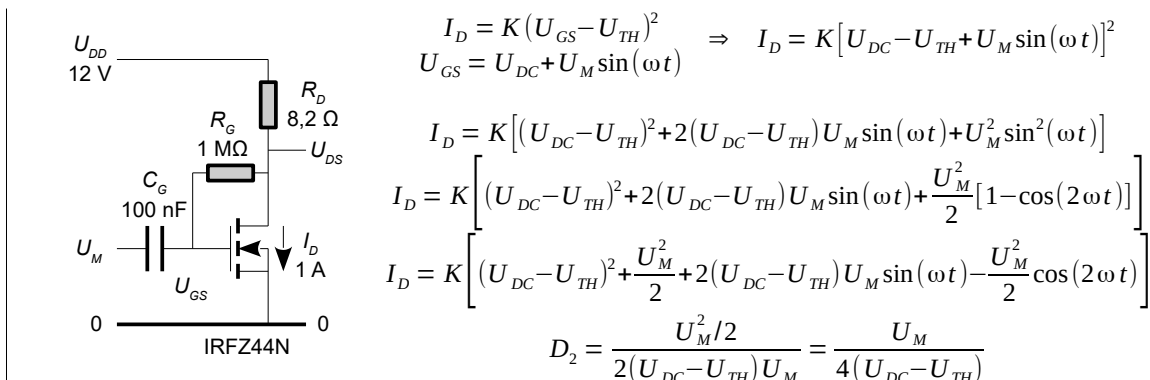
$$\sinh(-x) = -\sinh(x)$$

$$\cosh(-x) = \cosh(x)$$

Figur 100 – Hyperbolske funktioner.

2.5 Eksempel med MOSFET

For en MOSFET er der en kvadratisk relationen fra spændingen over gate-source U_{GS} til strømmen I_D i drain. Det betyder at en spændingssvingning på gate ikke blot resulterer i en strømsvingning i drain, men også i en forvrængningskomponent. Det kan analyseres ved brug af regneregler og en trigonometrisk funktion. Med en sinus på indgangen vil strømmen i drain bestå af tre led: en DC strøm, en sinus for det ønskede AC udgangssignal og en svingning på den dobbelte frekvens som repræsenterer en uønsket forvrængning af signalet. Forvrængningen kan estimeres som forholdet mellem amplituderne af det uønskede signal og af det ønskede signal.

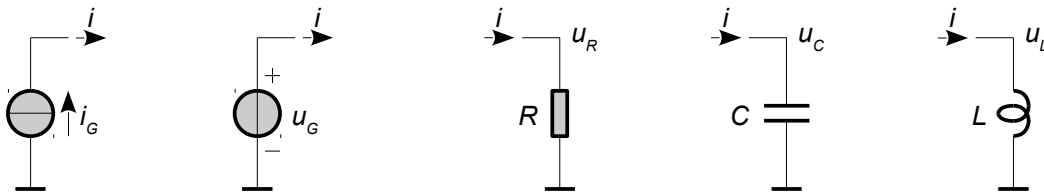


For IRFZ44N er $U_{TH} \approx 3,56$ V (*threshold voltage*) ifølge simulatoren NI Multisim og formelen gælder når $U_{GS} > U_{TH}$. Med de viste komponenter var $U_{DC} = U_{DS} = 3,80$ V og med amplituden $U_M = 0,1$ V er forvrængningen $D_2 = 10$ % ifølge formelen hvor en simulering på kredsløbet viste 8 %.

2.6 Passive komponenter

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 signal.⁶¹

Nedenfor vises symbolerne for to aktive kilder og de passive komponenter med den valgte notation for positiv spænding og strøm. De aktive kilder 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 og som angives ved det ledsagende symbol. For eksempel kan strømgeneratoren afhænge af spændingen et andet sted i kredsløbet.



Figur 101 – 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 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 $di/dt = 0$.

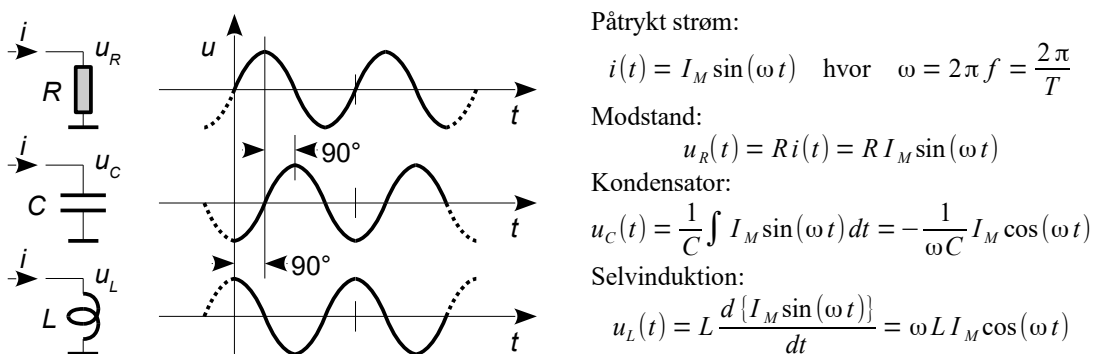
Indledningsvis vil de tre passive komponenter blive opfattet som ideelle. Det vil sige at funktionen kan beskrives matematisk eksakt med de viste ligninger. De virkelige komponenter er dog plaget af en række fejl og mangler og det vender jeg tilbage til i fjerde kapitel.

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

2.7 Impedans

Det blev illustreret i første kapitel at hvis man driver en kondensator med en oscillerende strøm vil der over komponenten dannes en oscillerende 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 oscillerende strøm $i(t)$. Den stiplede del af kurverne angiver at signalerne længe har oscilleret 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 102 – En strømkilde oscillerer som en sinus og driver en modstand R , en kondensator C og en spole L . Spændingen over komponenterne har forskellig fase.

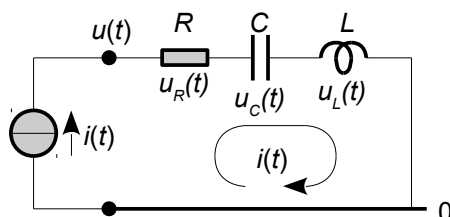
Parameteren $\omega = 2\pi f$ har dimensionen per sekund (s^{-1}) og angiver den hastighed hvormed fasen ωt øges; det er fasehastigheden der beskrives ved radianer per sekund (rad/s) og i en stor del af litteraturen lidt misvisende kaldes for vinkelfrekvensen (*angular frequency*).

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).



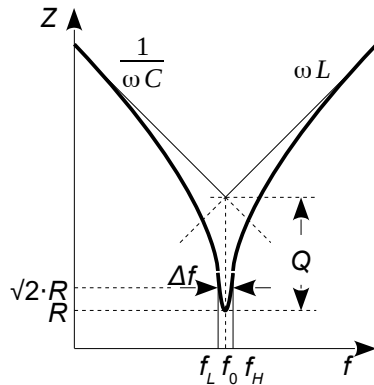
Spænding: $u(t) = u_R(t) + u_C(t) + u_L(t)$
 Drivende signal: $i(t) = I \sin(\omega t)$, $\omega = 2\pi f$ \Rightarrow

$$u(t) = RI \sin(\omega t) - \frac{1}{\omega C} I \cos(\omega t) + \omega L I \cos(\omega t)$$

$$u(t) = RI \sin(\omega t) + \left(\omega L - \frac{1}{\omega C} \right) I \cos(\omega t)$$

Figur 103 – 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 modificerer faser af det oscillerende 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.



$$\frac{\omega_0 L}{R} - \frac{1}{\omega_0 RC} = 0 \Rightarrow f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{LC}}$$

$$f \gg f_0 \quad Z \approx \omega L \quad \text{Fase: } 90^\circ$$

$$f = f_0 \quad Z = R \quad \text{Fase: } 0^\circ$$

$$f \ll f_0 \quad Z \approx 1/\omega C \quad \text{Fase: } -90^\circ$$

$$Q = \frac{U_C}{U_R} = \frac{I/\omega_0 C}{RI} = \frac{1}{\omega_0 RC} = \frac{\sqrt{LC}}{RC} = \frac{1}{R} \sqrt{\frac{L}{C}}$$

$$\Delta f = f_H - f_L = \frac{f_0}{Q}$$

Figur 104 – 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 kvadratroden af 2. For at finde grænsefrekvenserne kvadreres for at fjerne kvadratroden, der divideres med R^2 og det hele sættes på fælles brøkstreg. Kvadratroden tages på begge sider hvilket giver en fortegnssusikkerhed 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 værktøj og der deles med CR i både tæller og nævner for at introducere Q under kvadratroden. Det udnyttes at ω_1 skal være positiv og L skrives som $(L/C)^{0.5}$ gange $(LC)^{0.5}$ hvorved både Q og ω_0 indføres i udtrykket.

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

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.

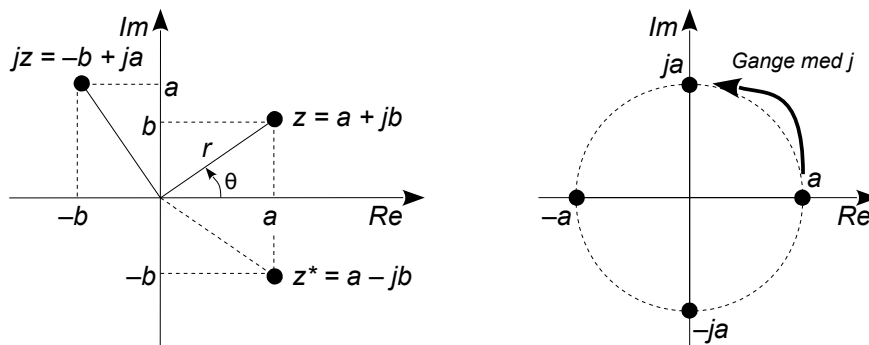
$$\Delta f = \frac{\omega_H - \omega_L}{2\pi} = \frac{\omega_0 (\sqrt{1 + 4Q^2} + 1) - (\sqrt{1 + 4Q^2} - 1)}{2Q} = \frac{f_0}{Q}$$

2.8 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 billedet.⁶²

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

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 horisontale 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 105 – 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. Multipliceres 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.

$$j^2 = -1$$

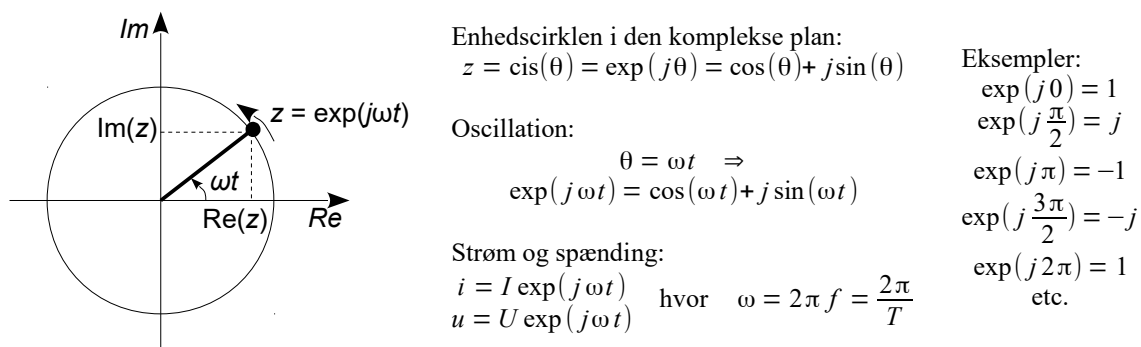
Definitionen skal bare accepteres for det kan ganske enkelt ikke forstås. Det er præcis derfor det hedder den imaginære enhed. Multipliceres $3j^2$ igen med j kommer vi frem til $3j^3 = -3j$ der nu peger nedad den imaginære akse og ved endnu en multiplikation med j er vi tilbage ved udgangspunktet på 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.

⁶² 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 og i 1799 noterede Wessel at det komplekse tal er et punkt i en talplan.

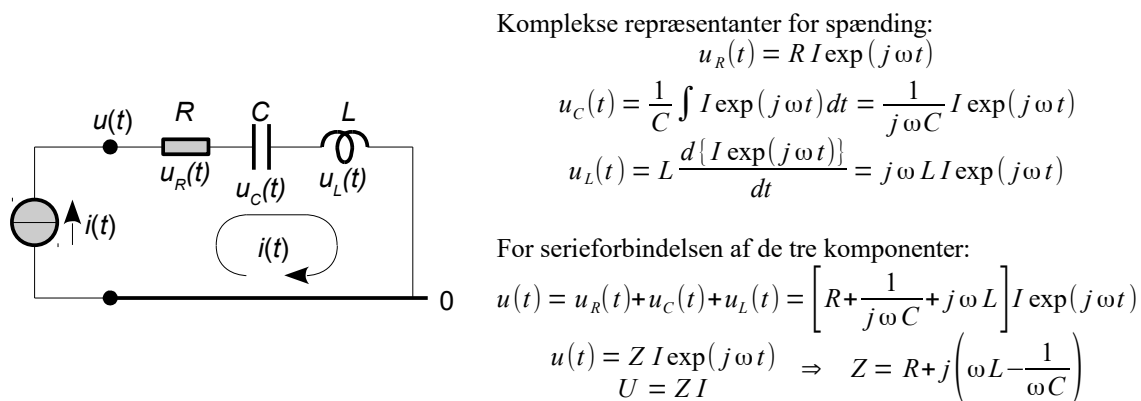
Punkterne på periferien af enhedscirklen er givet ved $x = \cos(\theta)$ og $y = \sin(\theta)$ så i den komplekse talplan skal cosinus pege ud af den reelle akse og sinus op af den imaginære akse. Det skrives ved funktionen $\text{cis}(\theta)$, der står for cosinus plus den imaginære enhed ganget med 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 106.

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 tiden.⁶³



Figur 106 – 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.

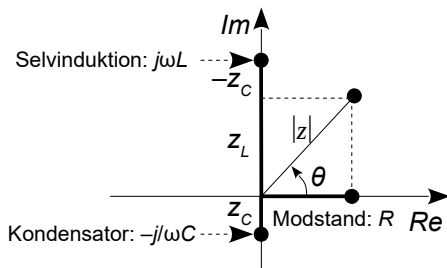


Figur 107 – Kredsløbet med en serieforbindelse af de tre komponenter R, C og L analyseres med en repræsentation ved de komplekse tal.

Forskellen fra det tidligere resultat med strømmen defineret som $i = I \sin(\omega t)$ er at alle led nu har den oscillerende 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 oscillerende 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.

⁶³ 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 reelt 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.



Modulus (længde):

$$|z| = \sqrt{(\operatorname{Re}(z))^2 + (\operatorname{Im}(z))^2} = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}$$

Argument (fase):

$$\theta = \arctan\left(\frac{\operatorname{Im}(z)}{\operatorname{Re}(z)}\right) = \arctan\left(\frac{\omega L - \frac{1}{\omega C}}{R}\right)$$

Figur 108 – 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 afbildes 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 afbildes 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 afbildes 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) repræsenteres af 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 sædvanen 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 på.⁶⁴

$$\begin{aligned} U &= Z I \\ u(t) &= U \exp(j\omega t) \\ i(t) &= I \exp(j\omega t) \\ \omega &= 2\pi f \end{aligned}$$

$$\begin{aligned} \text{Modstand:} \quad U_R &= R I \quad \Rightarrow \quad Z_R = \frac{U_R}{I} = R \\ \text{Kondensator:} \quad U_C &= \frac{1}{j\omega C} I \quad \Rightarrow \quad Z_C = \frac{U_C}{I} = \frac{1}{j\omega C} = -j \frac{1}{\omega C} \\ \text{Spole:} \quad U_L &= j\omega L I \quad \Rightarrow \quad Z_L = \frac{U_L}{I} = j\omega L \end{aligned}$$

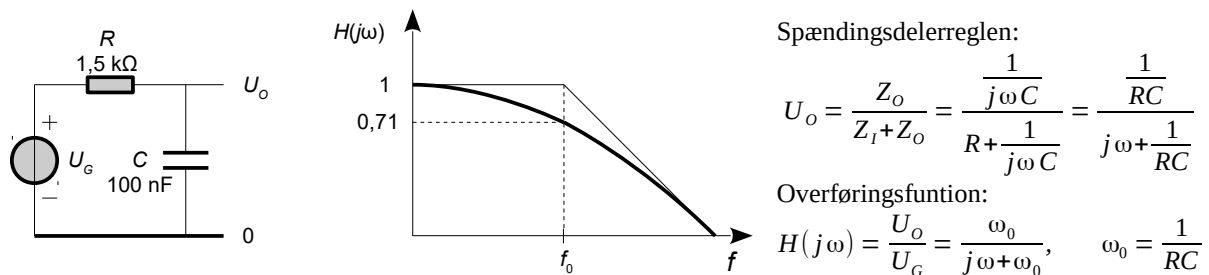
Den tidsafhængige relation gennem $\exp(j\omega t)$ er fjernet fra ligningerne hvilket er tilladeligt da vi kun beskæftiger os med den stationære tilstand for kredsløbet altså at svingningen altid har været der og at hver periode er identisk med den foregående. Denne definition af impedans sætter os 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 drejekondensator 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.

⁶⁴ Almindelige notationer: z , \hat{z} , Z , r/θ .

2.9 Spændingsdeling

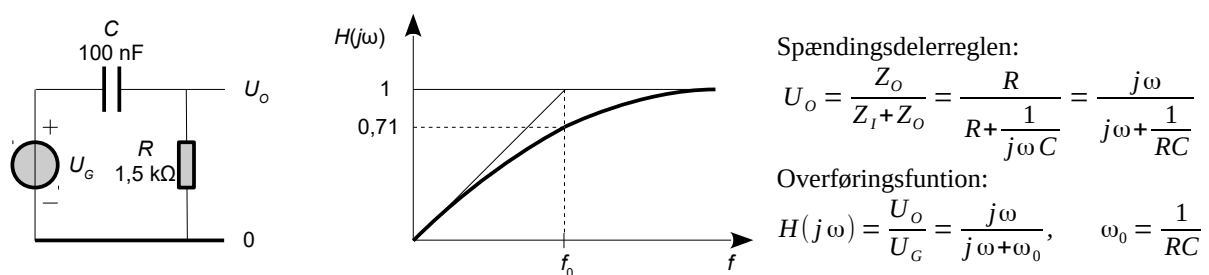
Mange kredsløb kan analyseres direkte ved spændingsdelerreglen og som et eksempel vises en analyse af filtret herunder. Her er Z_o belastningen af udgangen og Z_i er impedansen i indgangen. Der er kun en enkelt kondensator i udgangen som har impedansen $1/j\omega C$ og impedansen i indgangen er modstanden 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 109 – 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 set 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 110 – 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.

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

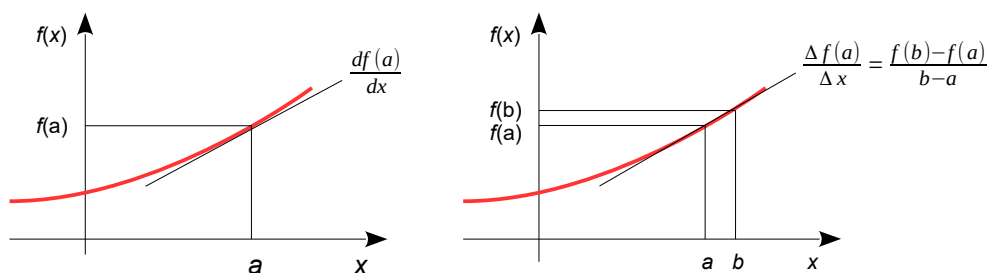
2.10 Differentialregning

Meget videnskabeligt arbejde beskæftiger sig med at opstille og løse differentiaalligninger 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 er ikke begrænset til svingninger for den kan også beskrive diffusion eller hvordan varmeenergi udbredes. Eksemplerne beskrives med den samme matematik som for strømmen i en kondensator der oplades gennem en modstand.

Differentiaalligningerne blev udviklet af Leibnitz i 1684 og det er hans notation der normalt benyttes. Newton udviklede omkring 1666 en tilsvarende matematik der dog først blev publiceret 1687.

2.10.1 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 111 – 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 akserne ligger langt fra hinanden vil resultatet blive en grov tilnærmelse 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:

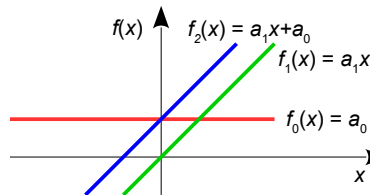
$$f'(x) = \frac{df(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

Hvis Δx gøres mindre og mindre vil beregningen gradvist ende med den korrekte værdi hvilket skrives med **lim** der står for *limes* og som betyder grænseværdien for dx gående imod nul. Jeg giver nogle eksempler i det følgende. Differentiation 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.

2.10.2 Eksempler på differentiation

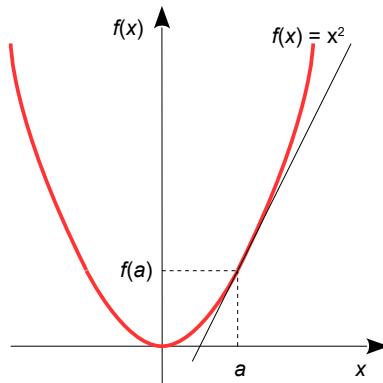
Herunder vises hvordan differentiationen gennemføres for rette linjer og parabler og derefter for eksponentialfunktionen og sinus. En mere udtømmende liste findes side 101.

Rette linjer har en hældning og differentieres til denne værdi (a_1 i figuren). Når linjen er vandret er hældningen nul ($a_1 = 0$) så en konstant differentieres til nul (a_0 i eksemplerne herunder).



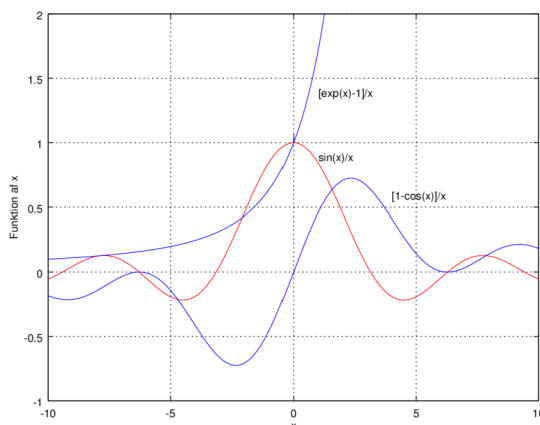
$$\begin{aligned}\frac{d\{a_0\}}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{a_0 - a_0}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{0}{\Delta x} = 0 \\ \frac{d\{a_1 x\}}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{a_1(x + \Delta x) - a_1 x}{\Delta x} = a_1 \\ \frac{d\{a_1 x + a_0\}}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{a_1(x + \Delta x) + a_0 - (a_1 x + a_0)}{\Delta x} = a_1\end{aligned}$$

Polynomier skal her repræsenteres ved parabeln 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 metoden kan udvides til højere orden.



$$\begin{aligned}\frac{d\{x^2\}}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} \\ \frac{d\{x^2\}}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{x^2 + 2x\Delta x + \Delta x^2 - x^2}{\Delta x} \\ \frac{d\{x^2\}}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{2x\Delta x + \Delta x^2}{\Delta x} \\ \frac{d\{x^2\}}{dx} &= \lim_{\Delta x \rightarrow 0} (2x + \Delta x) = 2x \\ \frac{d\{x^n\}}{dx} &= n x^{n-1}\end{aligned}$$

Eksponentielle og trigonometriske funktioner er lidt tricky at bestemme for $\sin(\Delta x)/\Delta x$ giver nul delt med nul når Δx går imod nul og det samme gør $\exp(\Delta x) - 1$ og $1 - \cos(\Delta x)$ delt med Δx . Det kan løses mere elegant end det jeg gør herunder hvor jeg blot plotter leddene og viser at $\sin(\Delta x)/\Delta x$ og $[\exp(\Delta x) - 1]/\Delta x$ går imod én for Δx gående imod nul mens $[1 - \cos(\Delta x)]/\Delta x$ går mod nul. En forklaring findes i Taylor serien for den pågældende funktion (se side 103) der viser at $\sin(x)$ og $\exp(x)$ har hældning én for x gående mod nul mens $\cos(x)$ har hældning nul.



$$\begin{aligned}\frac{d\{\exp(x)\}}{dx} &= \frac{\exp(x + \Delta x) - \exp(x)}{\Delta x} \\ &= \exp(x) \frac{\exp(\Delta x) - 1}{\Delta x} \xrightarrow{\Delta x \rightarrow 0} \exp(x)\end{aligned}$$

$$\begin{aligned}\frac{d\{\sin(x)\}}{dx} &= \frac{\sin(x + \Delta x) - \sin(x)}{\Delta x} = \\ &= \frac{\sin(x)\cos(\Delta x) + \cos(x)\sin(\Delta x) - \sin(x)}{\Delta x} \\ &= -\sin(x) \frac{1 - \cos(\Delta x)}{\Delta x} + \cos(x) \frac{\sin(\Delta x)}{\Delta x} \\ &\xrightarrow{\Delta x \rightarrow 0} \cos(x)\end{aligned}$$

2.10.3 Regneregler ved differentiation

Polynomier som $a_0 + a_1x + a_2x^2 + \dots$ er repræsenteret ved regnereglen for differentiation af x^n og benyttes for simplifikation af analytiske udtryk som det vises på side 103. Eksponentialfunktionen $\exp(x) = e^x$ differentieres til sig selv og de trigonometriske funktioner skifter frem og tilbage mellem sinus og cosinus med skiftende fortegn, men de forbliver svingninger med samme frekvens.

Potensfunktioner	$y = x^n \Rightarrow \frac{dy}{dx} = n x^{n-1}$
Eksponentialfunktioner	$y = \exp(ax) \Rightarrow \frac{dy}{dx} = a \exp(ax)$
Trigonometriske funktioner	$y = \sin(\omega t) \Rightarrow \frac{dy}{dt} = \omega \cos(\omega t)$
	$y = \cos(\omega t) \Rightarrow \frac{dy}{dt} = -\omega \sin(\omega t)$

Differentiation er lineær hvilket betyder at den afledte til en sum af led blot består af de enkelte led differentieret. En konstant faktor a deltager ikke i differentiationen og beholdes som en fast faktor. Funktioner med en konstant faktor til den variable af typen $f(ax)$ differentieres ved substitution hvor dx erstattes af dax og funktionen kan nu differentieres med hensyn til denne nye variabel ax . For at tillade substitutionen skal udtrykket ganges med a da dax står i nævneren.

$$\begin{aligned} \frac{d}{dx}\{f(x)+g(x)\} &= \frac{d}{dx}f(x)+\frac{d}{dx}g(x) \\ \frac{d}{dx}\{af(x)\} &= a\frac{d}{dx}\{f(x)\} \\ \frac{d}{dx}f(ax) &= a\frac{d}{dax}f(ax) \end{aligned}$$

Reglerne skal illustreres ved differentiation af en funktion bestående af både cosinus og sinus til en vinkel givet ved vinkelhastigheden ω og tiden t .

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

Substitution benyttes ved differentiation af funktioner der afhænger af andre funktioner. Det kaldes for kæderegele (chain rule).

$$y = f(g(x)) \Rightarrow \frac{dy}{dx} = \frac{df(g(x))}{dg(x)} \frac{dg(x)}{dx}$$

Det skal vises at sinus til den dobbelte vinkelhastighed differentieres til cosinus til den samme vinkelhastighed med amplituden 2ω . Det udnyttes at 2ω er konstant og at t differentieres til 1.

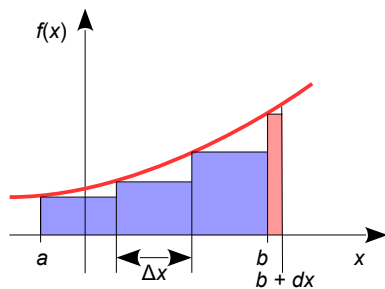
$$\begin{aligned} u &= \sin(v) \\ v &= 2\omega t \\ y = \sin(2\omega t) &\Rightarrow \frac{du}{dv} = \frac{d\sin(v)}{dv} = \cos(v) \Rightarrow \frac{dy}{dt} = \frac{du}{dv} \frac{dv}{dt} = \cos(v) 2\omega = 2\omega \cos(2\omega t) \\ \frac{dv}{dt} &= \frac{d(2\omega t)}{dt} = 2\omega \end{aligned}$$

Differentiation af et produkt af funktioner.

$$\frac{d}{dx}\{f(x)g(x)\} = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx}$$

2.10.4 Integration

Integration er opstået som bestemmelse af arealet under en funktion 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 stort og tilnærmelsen til det eksakte resultat vil være god. Påstanden er at for Δx gående mod nul vil arealet gå mod den eksakte værdi.⁶⁵



Summation af rektangler med højden $f(x)$ og bredden dx :

$$A_a^b = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} f(a+n\Delta x) \Delta x \quad \Delta x = \frac{b-a}{N}$$

Arealet gives af stamfunktionen:

$$A_a^b = \int_a^b f(x) dx = F(b) - F(a)$$

Stamfunktionen $F(x)$ til $f(x)$:

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

Figur 112 – Ved integration bestemmes arealet af et område defineret vertikalt ved funktionen $f(x)$ og den vandrette akse samt horisontalt ved et startpunkt a og et slutpunkt b .

Der findes en funktion 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*).⁶⁶

Nogle eksempler på de to regneregler vises på side 101.

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 differentialen $dF(x)$, altså ændringen i $F(x)$ ved en lille ændring i x , bestemmes ved at gange udtrykket ovenfor med dx på begge sider af lighedstegnet. Den modsatte proces af differentiation er integration og den kan defineres som den proces der danner funktionen $F(x)$ fra differentialen $dF(x)$.

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

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

⁶⁵ Noget tyder på at en romer arbejdede med integration for bestemmelse af voluminet af emner allerede omkring Kristi fødsel.

⁶⁶ http://en.wikipedia.org/wiki/Fundamental_theorem_of_calculus.

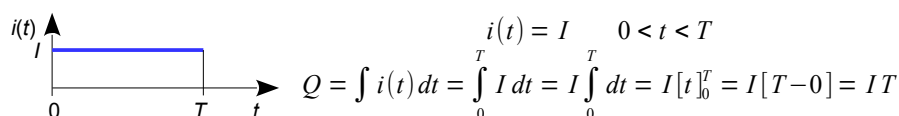
2.10.5 Eksempler på integration

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 $t = T$.

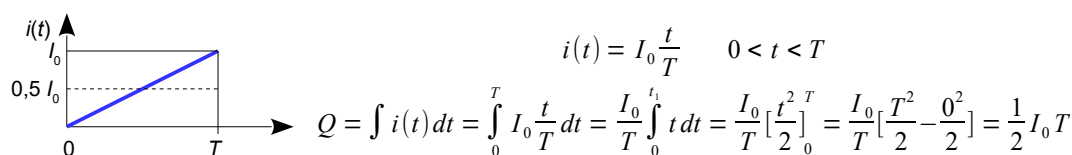
$$Q = \int_0^T i(t) dt$$

Hvis strømmen $i(t)$ er konstant kan man erstatte den med symbolet 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.

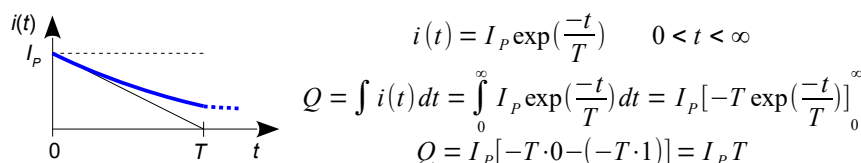
En konstant strømstyrke svarer til opladningen af et bilbatteri fra en ladestation. Strømmen antages konstant og forløber over et tidsrum på T . Resultatet er at ladningen Q er givet ved strømmens DC værdi I ganget med tidsrummets længde T .



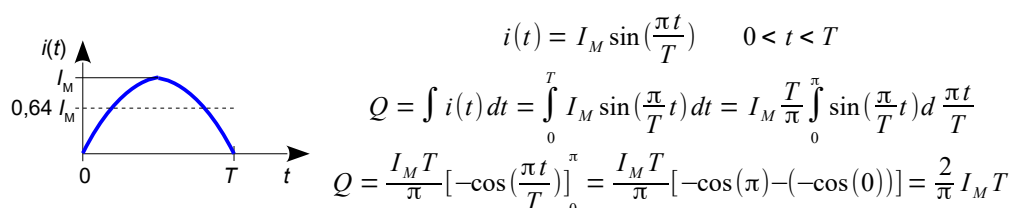
En lineært voksende strømstyrke mødes ved en DC-DC konverter (se sjette kapitel) og den resulterer i en ladning på halvdelen af strømmen I_0 gange med tidsrummet T .



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 princippet 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 $5T$) 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.



2.10.6 Regneregler ved integration

Ved integration findes en stamfunktion $F(x)$ til integranden $f(x)$. Ved differentiation af $F(x)$ findes funktionen $f(x)$ og da differentiation af $F(x) + C$ giver samme resultat er der en ubestemthed ved en integration som repræsenteres ved konstanten C . For integration over et område fra a til b findes et areal A under funktionen og resultatet er præcist uden en konstant fra integrationen.

$$\begin{aligned} F(x) &= \int f(x) dx + C \\ A &= \int_a^b f(x) dx = F(b) - F(a) \end{aligned}$$

Integration er en lineær operation så integralet af en sum beregnes individuelt for de enkelte led og en konstant faktor forbliver en fast faktor.

$$\begin{aligned} \int f(x) + g(x) dx &= \int f(x) dx + \int g(x) dx \\ \int a f(x) dx &= a \int f(x) dx \end{aligned}$$

Integration af et produkt af funktioner beregnes ved partiel integration (*integration by parts*) der er udledt fra regnereglen om differentiation af et produkt. Metoden er specielt attraktiv når den ene af funktionerne er en potensfunktion som efter et antal gange differentiation bliver til én.

$$\begin{aligned} \int f(x) g(x) dx &= f(x) G(x) - \int \frac{df(x)}{dx} G(x) dx \\ \int_a^b f(x) g(x) dx &= [f(x) G(x)]_a^b - \int_a^b \frac{df(x)}{dx} G(x) dx \end{aligned}$$

Funktionen $2x \cos(x)$ der består af funktionen $2x$ og funktionen $\cos(x)$. Det vælges at $f(x)$ skal repræsentere $2x$ idet den afledte af x er én og derfor simplificerer det tiloversblevne integral.

$$\begin{aligned} A &= \int 2x \cos(x) dx \\ f(x) = 2x &\Rightarrow \frac{df(x)}{dx} = 2 &\Rightarrow A = 2x \sin(x) - \int 2 \sin(x) dx \\ g(x) = \cos(x) &\Rightarrow G(x) = \sin(x) &\Rightarrow A = 2x \sin(x) - 2(-\cos(x)) + C \\ & &\Rightarrow A = 2x \sin(x) + 2 \cos(x) + C \end{aligned}$$

Ved en funktion af en anden funktion benyttes integration ved substitution.

$$\begin{aligned} \int_a^b f(g(x)) g'(x) dx &= \int_{g(a)}^{g(b)} f(t) dt \quad \text{hvor } t = g(x) \\ \int_a^b f(g(x)) g'(x) dx &= \int_{g(a)}^{g(b)} f(t) dt \end{aligned}$$

Variabel t skal være den indre funktion i den sammensatte funktion og herunder er det x^2 . Den nye variabel skal differentieres og ved at isolere dx opnås et udtryk med dt i stedet. Herefter indsættes udtrykket for dx , der integreres over variabel t og endelig genindsættes x^2 .

$$\begin{aligned} B &= \int x \exp(x^2) dx \\ t = x^2 &\Rightarrow \frac{dt}{dx} = 2x &\Rightarrow dx = \frac{1}{2x} dt \\ B &= \int x \exp(t) \frac{1}{2x} dt \\ B &= \frac{1}{2} \int \exp(t) dt \\ B &= \frac{1}{2} \exp(t) + C \\ B &= \frac{1}{2} \exp(x^2) + C \end{aligned}$$

2.10.7 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 kan noteres 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 den variable for 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	Integral
$f' = f'(x) = \frac{df(x)}{dx}$	$f = f(x)$	$F = F(x) = \int f(x) dx + c$ $F(b) - F(a) = \int_a^b f(x) dx$
$f' = ag' + bh'$	$f = ag + bh$	$F = aG + bH + c$
$f' = a[g'h + gh']$	$f = agh$	$F = a[gH - \int g'H dx]$
$f' = a \frac{g'h' - gh'}{g^2}$	$f = a \frac{g}{h}$	Benyt dekomposition for at reducere g/h til enkle brøker
$f' = ag'(h)h'$	$f = ag(h)$	Substituer $z = h(x)$, find dz/dx , isoler dx , indsæt og løs integralet
$f' = a$	$f = ax$	$F = \frac{ax^2}{2} + c$
$f' = anx^{n-1}$	$f = ax^n$	$F = \frac{a}{n+1} x^{n+1} + c \quad n \neq -1$
$f' = \frac{a}{2\sqrt{x}}$	$f = a\sqrt{x} = ax^{0.5}$	$F = \frac{a}{1.5} x^{1.5} + c$
$f' = a \frac{n}{x^{n+1}}$	$f = \frac{a}{x^n} = ax^{-n}$	$F = \ln x + c \quad n = 1$ $F = \frac{a}{(n-1)x^{n-1}} + c \quad n \neq 1$
$f' = ab \exp(bx)$	$f = a \exp(bx)$	$F = \frac{a}{b} \exp(bx) + c$
$f' = \frac{a}{x}$	$f = a \ln(bx)$	$F = ax \ln(bx) - x + c$
$f' = ab \cos(bx)$	$f = a \sin(bx)$	$f = -\frac{a}{b} \cos(bx) + c$
$f' = -ab \sin(bx)$	$f = a \cos(bx)$	$F = \frac{a}{b} \sin(bx) + c$
$f' = ab[1 + \tan^2(bx)] = \frac{ab}{\cos^2(bx)}$	$f = a \tan(bx)$	$F = -\frac{a}{b} \ln \cos(bx) + c$
$f' = ab \cosh(bx)$	$f = a \sinh(bx)$	$F = \frac{a}{b} \cosh(bx) + c$
$f' = ab \sinh(bx)$	$f = a \cosh(bx)$	$F = \frac{a}{b} \sinh(bx) + c$
$f' = ab[1 - \tanh^2(bx)] = \frac{ab}{\cosh^2(x)}$	$f = a \tanh(bx)$	$F = \frac{a}{b} \ln(\cosh(bx)) + c$

2.11 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 logaritmefunktionerne 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 $f(x)$ hvor $d^2 f(a)/dx^2$ betyder at funktionen $f(x)$ er differentieret to gange og derefter sættes $x = a$. Udtrykket $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ 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.

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

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). Det ses også at $\sin(x)/x$ fra side 97 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 $a = 0$.

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

2.11.1 Tabel over Taylor polynomier

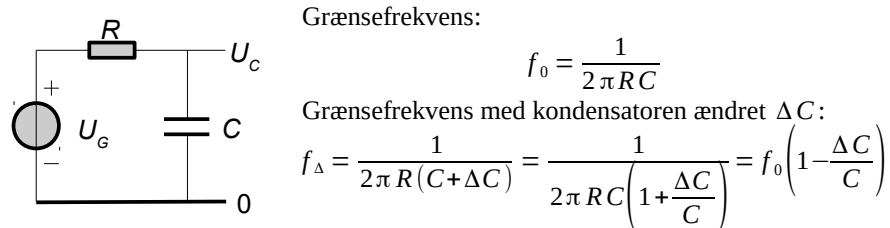
Herunder er variabel x ethvert reelt eller komplekst tal indenfor de angivne grænser. Bemærk at for intervallet $-1 < x < 1$ kan grænsernes værdier ikke benyttes da rækken ikke konvergerer mod en bestemt værdi. For $x = -1$ er eksempelvis logaritmen ikke defineret og for $x = 1$ kan de angivne rækker ikke konvergere.

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

En værdifuld egenskab ved Taylor polynomierne er at de forskellige funktioner ofte kan tilnærmes med de første få led af polynomiet hvis x er nær ved nul. Det kan give væsentlige fordele ved analytisk arbejde hvor en kompliceret relation kan simplificeres.

2.11.2 Eksempler på brug af Taylor polynomier

Et lavpasfilter består af en modstand R og en kondensator C og har grænsefrekvensen f_0 . Filtret er ikke særlig stejlt, men det kan være illustrativt at notere sig hvor meget grænsefrekvensen ændres med en ændring i eksempelvis kondensatorens værdi. Resultatet er at en stigning på 1 % sænker grænsefrekvensen med 1 %. Det kan vises ved at benytte Taylor-serien for $1/(1+x)$ hvor x er den relative ændring i kondensatoren.



Figur 113 – Et lavpasfilter analyseres for sensitivitet overfor ændring i komponentværdier.

Det overstående kaldes for en sensitivitsanalyse med sensitiviteten -1 . Et andet eksempel vises herunder hvor frekvensen f_0 af et LC led ændres ved en ændring i selvinduktionen på ΔL og i kapaciteten på ΔC . På grund af kvadratroden vil $+1$ % ændring i den ene af komponenterne give en frekvensændring på $-0,5$ % så sensitiviteten er $-0,5$. I beregningen herunder udnyttes det at ændringerne er så små at produktet $\Delta L \Delta C$ kan ignoreres.

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

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

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \quad f_{\Delta} \underset{\substack{\Delta L \ll L \\ \Delta C \ll C}}{\approx} \frac{1}{2\pi\sqrt{LC\left(1+\left[\frac{\Delta L}{L}+\frac{\Delta C}{C}\right]\right)}} \underset{x=\frac{\Delta L}{L}+\frac{\Delta C}{C}}{\stackrel{\text{Taylor, med}}{\approx}} \frac{1}{2\pi\sqrt{LC}}\left(1-\frac{1}{2}\left[\frac{\Delta L}{L}+\frac{\Delta C}{C}\right]\right)$$

Figur 114 – Et båndpasfilter analyseres for sensitivitet overfor ændring i komponentværdier.

I kapitel 5 udledes en relation for forstærkningen ved den bipolære transistor. Det vises at for små ændringer i spændingen over basis-emitter U_{BE} varierer strømmen i kollektor I_C tilnærmelsesvist lineært og det kaldes for transistorens transkonduktans. I det følgende beskrives ændringen i I_C ved en tilvækst i_C der skyldes en ændring i spændingen U_{BE} over basis-emitter på u_{BE} og derefter benyttes Taylor-serien for at simplificere $\exp(x)$.

Definitions ligning:

$$I_C = I_S \exp\left(\frac{U_{BE}}{nU_T}\right) \quad \text{hvor} \quad U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$$

Strømændring:

$$I_C + i_C = I_S \exp\left(\frac{U_{BE} + u_{BE}}{nU_T}\right) = I_S \exp\left(\frac{U_{BE}}{nU_T}\right) \exp\left(\frac{u_{BE}}{nU_T}\right) \Rightarrow$$

$$I_C + i_C = I_C \exp\left(\frac{u_{BE}}{nU_T}\right) \Rightarrow i_C = I_C \exp\left(\frac{u_{BE}}{nU_T}\right) - I_C$$

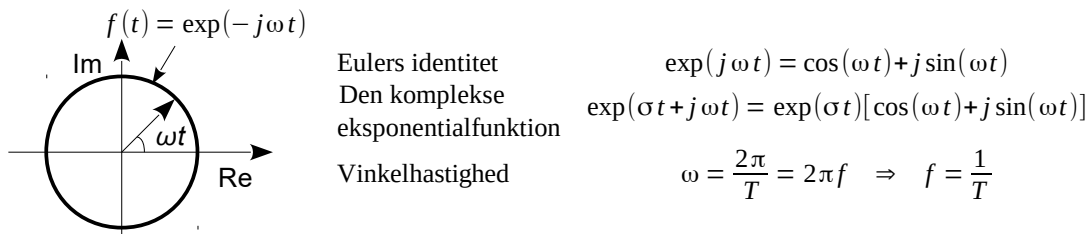
Taylor-serien for eksponentialfunktionen:

$$i_C = I_C \left(1 + \frac{u_{BE}}{nU_T} + \left(\frac{u_{BE}}{nU_T}\right)^2 + \dots\right) - I_C \quad u_{BE} \ll U_T \quad i_C \approx \frac{I_C}{nU_T} u_{BE}$$

Figur 115 – En bipolær transistor har en eksponentiel relation mellem spændingen over dens basis-emitter og strømmen i dens kollektor.

2.12 Eulers identitet

Den vigtigste relation indenfor de komplekse tals matematik er den identitet som Leonhard Euler fandt i midten af 1700-tallet. Den indeholder både eksponentialfunktionen og de trigonometriske funktioner hvilket letter arbejdet med at løse differentiaalligninger. Caspar Wessel viste cirka 1800 at Eulers identitet beskriver punkterne langs med en enhedscirkel i den komplekse talplan og det er netop hvad vi har brug for for at beskrive noget der svinger.



Figur 116 – Den komplekse eksponentialfunktion beskriver en cirkel i den komplekse talplan.

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 $\sigma t + j\omega t$ og ikke kun for de imaginære tal. Indenfor elektriske kredsløb er σ negativ så den komplekse eksponentialfunktion beskriver en svingning med vinkelhastigheden ω der dør ud med en hastighed givet ved σ . For en værdi på $\sigma = 0$ er resultatet en svingning med konstant amplitude hvilket er en oscillator. For $\sigma > 0$ vil amplituden vokse og det beskriver en oscillator der er ved at starte op.

At sinus og cosinus indgår i udtrykket for den komplekse eksponentialfunktion betyder at de to trigonometriske funktioner kan udtrykkes alene ved den komplekse eksponentialfunktion. Det betyder igen at løsningen til en ordinær differentiaalligning altid er en eksponentialfunktion.

Eulers identitet kan også benyttes til at definere de trigonometriske funktioner med et komplekst argument og det leder til de hyperbolske funktioner \cosh og \sinh . De er ikke oscillerende som \cos og \sin , men de opfører sig på mange måder som de trigonometriske funktioner.

Trigonometriske funktioner:

$$\cos(\theta) = \frac{\exp(j\theta) + \exp(-j\theta)}{2}$$

$$\sin(\theta) = \frac{\exp(j\theta) - \exp(-j\theta)}{2j}$$

Komplekse trigonometriske funktioner:

$$\cos(j\theta) = \frac{\exp(-\theta) + \exp(\theta)}{2} = \cosh(\theta)$$

$$\sin(j\theta) = \frac{\exp(-\theta) - \exp(\theta)}{2j} = j \sinh(\theta)$$

Herunder udledes de viste formler. Eulers identitet vil blive benyttet i de efterfølgende kapitler. Der vil dog ikke forekomme referencer til de komplekse trigonometriske funktioner 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. Det udnyttes at cosinus er en lige funktion $\cos(-\theta) = \cos(\theta)$ mens sinus er en ulige funktion $\sin(-\theta) = -\sin(\theta)$. Udtrykkene fremkommer ved at løse de to udtryk for henholdsvis cosinus og sinus.

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

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

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

Identiteten kan benyttes til at lette udledningen af de trigonometriske relationer. Som et eksempel vises hvordan kvadratet på sinus giver et udtryk med cosinus til det dobbelte af argumentet.

Kvadratet på sinus udtrykkes enkelt med Eulers identitet. Først opskrives udtrykket for sinus som derefter kvadreres. Efter reduktion ses at resultatet bliver en konstant minus cosinus til den dobbelte vinkel.

$$\sin^2(\theta) = \left(\frac{\exp(j\theta) - \exp(-j\theta)}{2j} \right)^2 = \frac{\exp(j2\theta) + \exp(-j2\theta) - 2\exp(0)}{-4} = \frac{1}{2}(1 - \cos(2\theta))$$

Denne relation benyttes ved beregning af forvrængningen i et elektronisk kredsløb hvor det undersøges hvor godt en antagelse om linearitet kan forventes at holde.

Euler fandt identiteten ved at sammenligne Taylor serierne for eksponentialfunktionen med et imaginært argument med serierne for cosinus og sinus med et reelt argument. Ved at samle leddene i reelle og imaginære dele findes Taylor-serierne for de to trigonometriske funktioner.

$$\begin{aligned} \exp(jx) &= 1 + jx - \frac{x^2}{2} - j\frac{x^3}{6} + \frac{x^4}{24} + j\frac{x^5}{120} - \dots \\ \exp(jx) &= \left(1 - \frac{x^2}{2} + \frac{x^4}{24} - \dots \right) + j \left(x - \frac{x^3}{6} + \frac{x^5}{120} - \dots \right) \\ \exp(jx) &= \cos(x) + j \sin(x) \end{aligned}$$

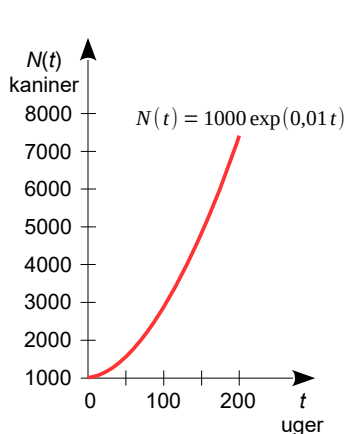
2.13 Første ordens differentialligning

En differentialligning beskriver de mekanismer der giver bevægelse så den kan vise hvordan kemiske processer udvikler sig over tid, hvordan mekaniske systemer svinger og hvordan et elektronisk kredsløb vil opføre sig efter en påvirkning. Differentialligninger arrangeres efter orden og jeg vil her kun beskrive første og anden orden da højere ordens systemer kan opsplittes i ligningssystemer af lavere orden.

Jeg vil introducere differentialligningen ved nogle eksempler.

Kaniner

Jo flere kaniner der er desto flere kaninunger kommer der. De vokser op og får selv unger så det samlede antal vokser. Lad os sige at der til et tidspunkt er $N = 1000$ kaniner på et område og at det tal forøges med en rate på $r = 0,01$ uge⁻¹. Tilvæksten dN/dt i antallet af kaniner i løbet af en uge er $Nr = 1000 \text{ kaniner} \cdot 0,01 \text{ uge}^{-1} = 10$ kaniner. En uge senere er der $N = 1010$ kaniner og antallet vokser så længe der er mad nok.⁶⁷



Differentialligning

Prototype på løsning

Differentiation

Indsættelse

Startværdi

Løsning

Fordobling

Indsættelse

Løsning

$$\frac{dN(t)}{dt} = N(t)r$$

$$N(t) = A \exp(Bt)$$

$$\frac{dN(t)}{dt} = A B \exp(Bt)$$

$$A B \exp(Bt) = A \exp(Bt)r \Rightarrow B = r$$

$$N(0) = A \exp(0) \Rightarrow A = N(0)$$

$$N(t) = N(0) \exp(rt)$$

$$N(t_2) = 2 N(0)$$

$$N(0) \exp(rt_2) = 2 N(0)$$

$$\exp(rt_2) = 2$$

$$t_2 = \frac{\ln(2)}{r} = \frac{\ln(2)}{0,01 \text{ uge}^{-1}} = 69 \text{ uger}$$

Ligningen angiver at den afledte af funktionen for N er lig med sig selv gange med r . Fra tabellen over afledte ser vi at eksponentialfunktionen besidder netop denne egenskab og derfor kan den være en prototype for en løsning. Her prøves med $A \exp(Bt)$ hvor A og B er nogle konstanter der skal findes for at verificere prototypen. Først differentieres udtrykket for N og ved indsættelse får vi defineret konstanten B som raten. For at fastlægge konstanten A benytter vi vores viden om at antallet til tiden $t = 0$ er på $N(0) = 1000$ kaniner.

Vi kan nu bruge ligningen til at forudsige hvor lang tid der går indtil antallet af kaniner er fordoblet. Det vil sige at vi søger efter tidspunktet t_2 for $N(t_2) = 2 N(0)$ og løsningen er 69 uger. Antallet vokser eksponentielt så antallet af kaniner fordobles igen efter samme tidsforløb hvis der ikke er andre forhold som gør sig gældende.

Eksemplet viser et generelt træk ved løsning af differentialligninger som irriterer mange ved første møde. Tilgangen til opgaven er at "indse" og ikke at benytte en "formel" for at løse opgaven. Vi har med andre ord forladt den trygge folkeskole og gymnasiet hvor løsningsmetoder blot skulle læres udenad for at bestå. En af mine matematiklærere beskrev løsning af differentialligninger som en kunstart i modsætning til det meste andet matematik. Den tilgang vil præge resten af dette kapitel for løsning af differentialligninger kræver både intuition og frækhed.

Der findes en generel metode til løsning af første-ordens differentialligninger som vises side 112.

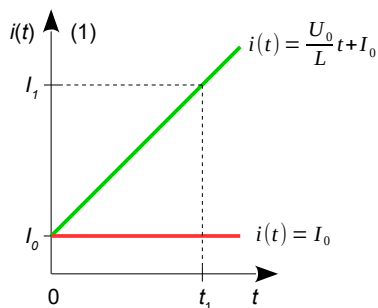
⁶⁷ <https://www.mathsisfun.com/calculus/differential-equations.html>.

Selvinduktionsspole

Spændingen over en spole som funktion af strømmen giver en første-ordens differentialligning. Med spændingen over spolen givet ved en generator $u(t)$ er strømmen $i(t)$ den ubekendte som skal findes. Ligningen siger at spændingen $u(t)$ over spolen til tidspunktet t er en funktion af *ændringshastigheden* i strømmen di/dt til det samme tidspunkt. Hastigheden er indenfor al fysik givet ved den afledte af variabelen med hensyn til tiden.

Hvis strømmen har en konstant værdi $i(t) = I_0$ så vil den afledte af strømmen være nul og det er en løsning til ligningen for $u(t) = 0$. Sagt med ord så vil strømmen i en (tabsfri) spole ikke ændre værdi hvis spændingen over spolen er nul. Når spændingen ikke er nul må vi finde et udtryk for strømmen. Herunder isoleres $di(t)$ ved at dividere på begge sider med L og gange med dt . Strømmen $i(t)$ findes ved integration på hver side af lighedstegnet idet integralet af di giver i .

Integrationen startes til tidspunktet $t = 0$ og stoppes $t = t_1$. Det er valgfrit hvornår man kalder et tidspunkt for $t = 0$. Man starter blot stopuret når det passer og stopper igen når målet er nået. Integrationskonstanten I_0 skyldes at enhver lineær forskydning af løsningen også er en løsning til differentialligningen (a_0 fra side 97). Det er den strøm der løb i spolen i det øjeblik vi startede analysen. Differentialligningen beskriver kun *ændringen* i strømmen fra I_0 og tiden t_1 frem.



Definitions ligning for en spole

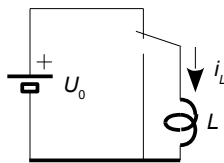
$$u(t) \stackrel{\text{def}}{=} L \frac{di(t)}{dt}$$

Bestemmelse af strømmen

$$di(t) = \frac{1}{L} u(t) dt$$

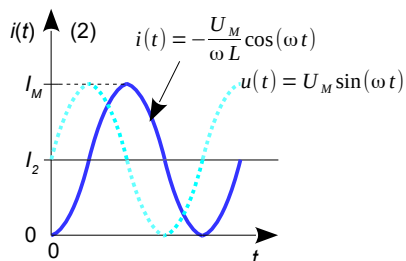
$$i(t) = \int_0^{t_1} di(t) = \int_0^{t_1} \frac{1}{L} u(t) dt$$

$$i(t) = \frac{1}{L} \int_0^{t_1} u(t) dt$$



(1) Konstant spænding: $u(t) = U_0$

$$i(t) = \frac{1}{L} \int_0^{t_1} U_0 dt = \frac{U_0}{L} \int_0^{t_1} dt = \frac{U_0}{L} t_1 + I_0$$



(2) Sinussvingning: $u(t) = U_0 \sin(\omega t)$, $\omega = 2\pi f$, $f = \frac{1}{T}$

$$i(t) = \frac{1}{L} \int_0^{t_1} U_0 \sin(\omega t) dt = \frac{U_0}{L} \int_0^{t_1} \sin(\omega t) dt$$

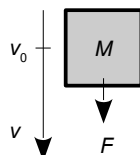
$$i(t) = -\frac{U_0}{\omega L} \cos(\omega t) + I_2$$

Med spolen tilsluttet over et batteri er spændingen konstant på $u(t) = U_0$ og strømmen vil vokse lineært over tid fra startværdien I_0 fra tiden før $t = 0$. Det beskriver situationen i en DC-DC konverter for det tidsrum hvor en transistor er ON og der opbygges energi i spolen ved at øge strømmen fra startværdien I_0 til slutværdien I_1 . Hvis spændingen til tidspunktet t_1 skifter fortegn til $u(t) = -U_0$ vil strømmen aftage lineært fra den høje værdi I_1 med hastigheden $-U_0/L$.

Sidste eksempel er en spole der til $t = 0$ tilkobles en oscillerende spændingskilde. Det kan som et eksempel være en transformator der tilkobles lysnettet ved nulgennemgang. Strømmen må vokse fra nul og startværdien I_2 skal derfor tilpasses så $i(0) = 0$ og det giver en strøm i spolen på $I_2 = U_0/\omega L$. Den strøm løber faktisk i en transformator i nogle få perioder efter indkobling for at klinge ud mod nul af årsager som den simple differentialligning ikke beskriver.

Frit fald

Eksemplet benytter Newtons anden lov for en faldende masse og viser en løsning med et polynomium. Newtons anden lov beskriver accelerationen dv/dt af en masse M der påvirkes med en kraft F fra jordens tiltrækning gennem tyngdeaccelerationen g . Hvis loddet slippes accelereres massen og hastigheden v beregnes til tidspunktet t . Løsningen for hastigheden benyttes derefter til at finde positionen x som funktion af tiden efter at loddet er sluppet ved at udnytte at hastigheden er defineret som den afledte af positionen.



$$\text{Kraft på massen } M \quad F = Mg, \quad g = 9,82 \text{ m/s}^2$$

$$\text{Newton II} \quad F = M \frac{dv}{dt}$$

$$\text{Differentialligning} \quad \frac{dv}{dt} = g \Rightarrow dv = g dt$$

$$\text{Hastighed} \quad v = \int dv = \int g dt = g \int dt = gt + v_0$$

$$\text{Position} \quad v = \frac{dx}{dt} \Rightarrow dx = v dt$$

$$x = \int dx = \int v dt = \int (v_0 + gt) dt$$

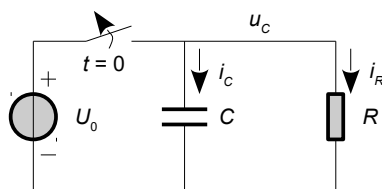
$$x = x_0 + v_0 t + \frac{1}{2} g t^2$$

Løsningen vil indeholde en konstant v_0 fordi løsningen altid kan forskydes med et konstant led der beskriver loddets hastighed i det øjeblik analysen startes. En mulighed er at loddet er i ro inden det slippes og så er begyndelseshastigheden nul ($v_0 = 0$).

Positionen x til tidspunktet t findes af definitions-ligningen $v = dx/dt$ ved at isolere dx og integrere. Igen optræder en integrationskonstant x_0 som angiver højden over jorden før vi slipper loddet.

Afladning af en kondensator

En kondensator C holdes opladt på spændingen U_0 gennem en kontakt og fra tidspunktet $t = 0$ aflades den gennem modstanden R . Ligningen kan løses ved at differentiere på begge sider for at fjerne integrationen. Løsningen formodes at være en eksponentialfunktion så en prototype på løsningen dannes og ved at differentiere og indsætte ses at det er en mulighed. Forløbet starter med $u(0) = U_0$ og slutter med en tømt kondensator som derved fastlægger konstanterne.



$$\text{Definition} \quad u_C(t) \stackrel{\text{def}}{=} \frac{1}{C} \int i_C(t) dt$$

$$\text{Strøm og fortegn} \quad i_R(t) = \frac{u_C(t)}{R} = -i_C(t)$$

$$\text{Differentialligning} \quad u_C(t) = \frac{1}{C} \int \frac{-u_C(t)}{R} dt$$

$$\text{Differentiation} \quad \frac{du_C(t)}{dt} = -\frac{1}{C} \frac{u_C(t)}{R}$$

$$\text{Prototype} \quad u_C(t) = U_1 \exp(at) + U_2$$

$$\text{Differentiation} \quad \frac{du_C(t)}{dt} = a U_1 \exp(at)$$

$$\text{Sammenligning} \quad a = -\frac{1}{CR}$$

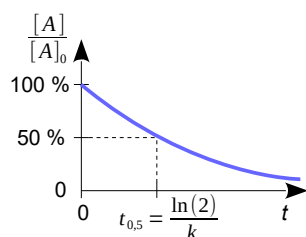
$$\text{Startværdi} \quad u(0) = U_0 \Rightarrow U_0 = U_1 + U_2$$

$$\text{Slutværdi} \quad u(\infty) = 0 \Rightarrow U_2 = 0$$

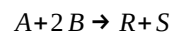
$$\text{Løsning} \quad u(t) = U_0 \exp\left(\frac{-t}{CR}\right)$$

Kemisk ligevægt

Kemiske reaktioner beskrives ved den hastighed som reaktionen forløbet med. Hastigheden v af processen udtrykkes ved ændringen i koncentrationen af en reagent. Det skrives som $[A]$ for koncentrationen af reagent A. Her vises en tænkt reaktion mellem en syre (*acid*) og en base der giver en reaktant og et salt. Processens hastighed beskrives ved k som er en konstant for en given koncentration af A og B. Det giver en første ordens differentialligning.⁶⁸



Eksempel på en reaktion



Reaktionshastighed $v \stackrel{\text{def}}{=} -\frac{d[A]}{dt} = -\frac{1}{2} \frac{d[B]}{dt} = \frac{d[R]}{dt} = \frac{d[S]}{dt}$

$$v = k[A][B]$$

Reaktionsligning

$$\frac{d[A]}{dt} = -k[A]$$

Løsning

$$[A] = [A]_0 \exp(-kt)$$

Halveringstid

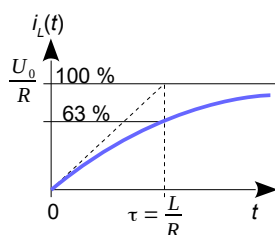
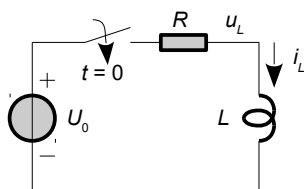
$$\frac{[A]}{[A]_0} = \exp(-kt) = \frac{1}{2}$$

$$-kt_{0,5} = \ln\left(\frac{1}{2}\right) = -\ln(2)$$

$$t_{0,5} = \frac{\ln(2)}{k}$$

Opladning af en reel spole

En spole har ud over selvinduktionen L en modstand R fra den tråd den er viklet af så der er en grænse for hvad strømmen kan vokse til. Ligningen følger helt samme procedure med at opstille en differentialligning og gætte på en løsning. Første ordens differentialligninger har eksponentialfunktionen som løsning og ved indsættelse af det viste udtryk for i_L i differentialligningen bestemmes konstanterne a , b og c .



Kirchhoff på kredsløbet

$$U_0 = Ri_L(t) + u_L(t)$$

Differentialligning

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

Prototype

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

Strømmen starter fra nul

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

Revision af prototype

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

Indsættelse

$$U_0 = Ra[1 - \exp(ct)] - L[ac \exp(ct)]$$

Strømmen starter fra nul

$$t = 0 \Rightarrow U_0 = -Lac \Rightarrow c = -\frac{U_0}{La}$$

Strømmen er begrænset

$$t \rightarrow \infty \Rightarrow U_0 = Ra \Rightarrow a = \frac{U_0}{R}$$

Krav: $c < 0$

Heraf sidste konstant

$$c = -\frac{R}{L}$$

Løsning

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

2.13.1 Konklusion på eksemplerne

Løsning af differentialligninger kræver oplysninger om hvordan forløbet starter og måske også om hvordan det slutter for selve differentialligningen beskriver kun det generelle om forløbet. Derfor er problemets startværdi (*initial value*) eller grænsebetingelse (*boundary condition*) vigtige for at man kan fastlægge løsningen i detaljer.

68 <http://logx.dk/keminoter.pdf>.

2.13.2 Generelt tilfælde

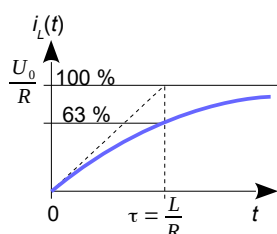
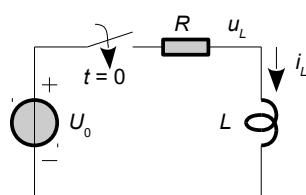
Jeg har hidtil beskæftiget mig med problemer hvor koefficienterne var konstante og givet ved de benyttede komponenter (modstand, kondensator og spole). Differentialligninger af første orden kan dog løses eksakt selv om koefficienten til $x(t)$ ikke er konstant og herunder vises hvordan. Der er brug for en integrationsfaktor $\mu(t)$ der beregnes ud fra faktoren $p(t)$. Det er uden betydning om $x(t)$ er funktion af tid, sted eller noget andet så variabelen t kan repræsentere enhver anden variabel.

Differentialligning	$\frac{dx(t)}{dt} + p(t)x(t) = f(t)$
Løsning	$x(t) = \frac{\int \mu(t)f(t)dt + c}{\mu(t)}$
Integrationsfaktor	$\mu(t) = \exp(\int p(t)dt)$

Udledningen af formelen er lidt tricky og skal ikke vises her, men interesserede kan læse artiklen som fodnoten henviser til. Den gennemgår det grundigt og i et forståeligt sprog.⁶⁹

Opladning af en reel spole

Som illustration af proceduren benyttes eksemplet fra forrige side med opladning af en reel spole bestående af selvinduktion L og modstand R . Faktoren $p(t)$ er her konstant, men det viser hvordan opgaven gennemføres og resultatet er selvfølgelig det samme som på forrige side.



Differentialligning	$L \frac{di(t)}{dt} + Ri(t) = U_0$
Rearrangering	$\frac{di(t)}{dt} + \frac{R}{L}i(t) = \frac{U_0}{L} \Rightarrow p(t) = \frac{R}{L}$
Integrationsfaktoren	$\mu(t) = \exp(\int \frac{R}{L}dt) = \exp(\frac{R}{L}t)$
Generel løsning	$i(t) = \frac{\int \exp(\frac{R}{L}t) \frac{U_0}{L} dt + c}{\exp(\frac{R}{L}t)} = \frac{\frac{U_0}{R} \exp(\frac{R}{L}t) + c}{\exp(\frac{R}{L}t)}$
Startværdi $i(0) = 0$	$c = -\frac{U_0}{R}$
Løsning	$i(t) = \frac{U_0}{R} \frac{\exp(\frac{R}{L}t) - 1}{\exp(\frac{R}{L}t)} = \frac{U_0}{R} \left[1 - \exp(-\frac{R}{L}t) \right]$

Hermed slutter eksemplerne på løsning af første-ordens differentialligninger.

⁶⁹ <http://tutorial.math.lamar.edu/Classes/DE/Linear.aspx>.

2.14 Anden ordens differentiaalligning

En anden-ordens differentiaalligning indeholder den to gange afledte af funktionen $x(t)$ og noteres som vist herunder. I en mekanisk sammenhæng beskriver det accelerationen af et emne, men det bruges også ved elektroniske kredsløb, elektromagnetisme, hydrodynamik, akustik og kemi. Det betyder at mange problemstillinger har samme differentiaalligning og derfor vil opføre sig på en sammenlignelig måde. Det uddybes sidst i kapitlet på side 146 for nogle udvalgte systemer.

$$\ddot{x}(t) = \frac{d}{dt} \left(\frac{dx(t)}{dt} \right) = \frac{d^2 x(t)}{dt^2}$$

Den anden afledte er herunder sat lig nul som eksempel på en anden-ordens differentiaalligning og det er så spørgsmålet hvad det giver af løsning. Ved at integrere på begge sider af lighedstegnet bliver den anden afledte af $x(t)$ på venstre side til den første afledte og højre sides nul integreres til en konstant a der hidrører fra ubestemtheden ved integrationen. Strengt taget har begge integraler en ubekendt konstant, men de er her samlet til en fælles konstant på højre side.

$$\begin{aligned} \frac{d^2 x(t)}{dt^2} = 0 & \quad \Rightarrow \quad \int \frac{d^2 x(t)}{dt^2} dt = \int 0 dt + a \\ & \quad \Rightarrow \quad \int \frac{d}{dt} \left(\frac{dx(t)}{dt} \right) dt = a \\ & \quad \Rightarrow \quad \int d \left(\frac{dx(t)}{dt} \right) = a \\ & \quad \Rightarrow \quad \frac{dx(t)}{dt} = a \\ & \quad \Rightarrow \quad \int \frac{dx(t)}{dt} dt = \int a dt + b \\ & \quad \Rightarrow \quad \int dx(t) = a \int dt + b \\ & \quad \Rightarrow \quad x(t) = at + b \end{aligned}$$

Ved endnu en integration vil den første-afledte af funktionen blive til funktionen $x(t)$ og konstanten a integreres til hældningen af en ret linje. En ny konstant b må introduceres fra ubestemtheden af integration nummer to. Resultatet bliver en ret linjes ligning. At det faktisk er en løsning ses ved at differentiere funktionen $x(t)$ to gange. Resultatet bliver nul.

$$x(t) = at + b \quad \xRightarrow{\text{Differentiation}} \quad \frac{dx(t)}{dt} = a \quad \xRightarrow{\text{Differentiation}} \quad \frac{d^2 x(t)}{dt^2} = 0$$

Vi ser at løsningen til en anden-ordens differentiaalligning introducerer to konstanter (a og b) som kræver yderligere information om opgaven for at kunne fastlægges. Det er en konsekvens af at en differentiaalligning kun beskriver det overordnede udseende af løsningen som i dette tilfælde er en ret linje, men uden detaljer om hældningskoefficient og skæring med y -aksen.

I en fysisk fortolkning beskriver differentiaalligningen Newtons første lov. En masse der ikke er udsat for en resulterende kraft vil bevæge sig med konstant hastighed eller ligge stille.

Faldloven

Hvis kraften F er jordens træk i massen M ved tyngdeaccelerationen g giver Newtons anden lov en anden-ordens differentiaalligning med positionen $x(t)$ som tidsafhængig variabel. Integrationen giver to konstanter med v_0 for hastigheden til tidspunktet $t = 0$ og x_0 for startpositionen.

$$\begin{aligned} F = M \frac{d^2 x(t)}{dt^2} & \Rightarrow \frac{d^2 x(t)}{dt^2} = g \\ F = Mg & \quad \Rightarrow \quad \int \frac{d^2 x(t)}{dt^2} dt = g \quad \Rightarrow \quad \int \frac{d}{dt} \left(\frac{dx(t)}{dt} \right) dt = g \\ & \quad \Rightarrow \quad \int d \left(\frac{dx(t)}{dt} \right) = g \quad \Rightarrow \quad \frac{dx(t)}{dt} = gt + v_0 \\ & \quad \Rightarrow \quad \int \frac{dx(t)}{dt} dt = \int (gt + v_0) dt \\ & \quad \Rightarrow \quad \int dx(t) = \frac{1}{2}gt^2 + v_0 t + x_0 \\ & \quad \Rightarrow \quad x(t) = \frac{1}{2}gt^2 + v_0 t + x_0 \end{aligned}$$

Det resultat blev også udledt i afsnittet om første-ordens differentiaalligninger. I første omgang var accelerationen den første afledte af hastigheden hvilket gav $gt + v_0$ og i anden omgang var hastigheden den første afledte af positionen og det gav den ovenstående parabel.

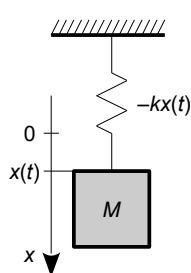
En klassiker indenfor anden-ordens differentiaalligninger vises herunder. Ligningen til venstre er tydeligvis en anden-ordens differentiaalligning med den dobbelt afledte af den variable $x(t)$ og den variable selv. Omformningen til højre viser at der søges en funktion med den egenskab at den to gange differentierede af funktionen skal give funktionen selv med modsat fortegn. Som det vides besidder de trigonometriske funktioner sinus og cosinus denne egenskab så det vil være naturligt at søge en løsning blandt disse funktioner.

$$\frac{d^2 x(t)}{dt^2} + x(t) = 0 \Rightarrow \frac{d^2 x(t)}{dt^2} = -x(t)$$

Det følgende eksempel viser at løsningen er en harmonisk svingning.

Den harmoniske oscillator

En masse ophængt i en fjeder kan beskrive en bils affjedringssystem, svingningskredse i de elektroniske kredsløb og højttalere monteret i et kabinet. Jordens tiltrækning vil trække i massen og strække fjederen som trækker tilbage med kraften $-kx_0$ indtil der optræder ligevægt. Denne position benyttes som udgangspunkt for analysen. Parameteren k er fjederkonstanten der måles i N/m så den viser hvor stor en kraft fjederen yder ved en given deformation. Det er her opgaven at beskrive hvad der sker hvis loddet trækkes x_M væk fra hvilepositionen og derefter slippes og overlades til sig selv.



Newton II

$$F = M \frac{d^2 x(t)}{dt^2}$$

Fjederkraft

$$F = -kx(t)$$

Differentiaalligning

$$M \frac{d^2 x(t)}{dt^2} = -kx(t)$$

Prototype

$$x(t) = x_M \sin(\omega t + \theta)$$

Differentiation

$$\frac{dx(t)}{dt} = x_M \omega \cos(\omega t + \theta)$$

$$\frac{d^2 x(t)}{dt^2} = -x_M \omega^2 \sin(\omega t + \theta)$$

Indsættelse

$$-M x_M \omega^2 \sin(\omega t) = -k x_M \sin(\omega t + \theta)$$

Reduktion

$$M \omega^2 = k$$

Svingning

$$\omega = \sqrt{\frac{k}{M}}$$

$$f_{osc} = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{k}{M}}$$

Startværdi

$$x(0) = x_M \Rightarrow \theta = 90^\circ$$

Løsning

$$x(t) = x_M \cos(\omega t)$$

Ved at sætte de to kraftudtryk lig med hinanden skabes en ligning i $x(t)$ der siger at den anden afledte af $x(t)$ skal være lig med sig selv med modsat fortegn og ganget med en konstant. Det vælges at basere en prototype på en sinus så $x(t) = x_M \sin(\omega t)$ hvor konstanten ω bestemmer svingningshastigheden og x_M bestemmer amplituden. Da fasen af den eftersøgte svingning ikke er kendt vælges at supplere med fasevinkel θ .

Analysen viser at antagelsen om en harmonisk svingning leder til en brugbar løsning om end startbetingelsen angiver at løsningen må være en cosinus.

Frekvensen af svingningen f_{osc} stiger når fjedren bliver stiv og falder når massen bliver stor.

2.14.1 Anden ordens homogen differentialligning

Den generelle form for en anden ordens differentialligning har to konstanter A og B der er kendt når opgaven er defineret. Der er ingen ydre påvirkning af kredsløbet så differentialligningen beskriver kun hvad det foretager sig efter en påvirkning der kan være at $x(t)$ er trukket væk fra ligevægt og analysen startes lige i det øjeblik der slippes eller at kredsløbet sættes i gang klokken $t = 0$ ved et kortvarigt stød.

Differentialligning	$\frac{d^2 x(t)}{dt^2} + A \frac{dx(t)}{dt} + Bx(t) = 0$
Karakterligning	$a^2 + Aa + B = 0$
Rødder	$a = \frac{-A \pm \sqrt{A^2 - 4B}}{2} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$
Løsning	$x_H(t) = X_1 \exp(a_1 t) + X_2 \exp(a_2 t)$

Denne type differentialligning har den komplekse eksponentialfunktion som løsning hvilket indsættes ved en prototype på løsningen af typen $\exp(at)$ der differentieres og indsættes. Det resulterer i en anden-ordens ligning der kaldes systemets karakterligning (*auxiliary equation*). Den har to rødder a_1 og a_2 så løsningen til differentialligningen består af to led. Enhver kombination af løsningerne er også en mulig løsning så de to led kan kombineres med hver sin vægtning X_1 og X_2 der først kan fastlægges når den aktuelle opgave er defineret.

Herunder vises opstilling og løsning af en differentialligning for et elektrisk kredsløb med RCL i en serieforbindelse der til tidspunktet $t = 0$ tilsluttes en spændingskilde med værdien U_0 .

Dæmpet svingning

Kirchhoffs maskelov for kredsløbet med en valgt retning af strømmen giver differentialligningen. Den indeholder både differentiation og integration, men integrationen kan fjernes ved at differentiere på begge sider af lighedstegnet. Det gør samtidigt ligningen homogen.

Kirchhoff (KVL) $-U_0 + u_R(t) + u_C(t) + u_L(t)$
 $-U_0 + Ri(t) + \frac{1}{C} \int_0^t i(t) dt + L \frac{di(t)}{dt} = 0$

Kredsløbsligning $U_0 = Ri(t) + \frac{1}{C} \int_0^t i(t) dt + L \frac{di(t)}{dt}$

Differentiation $0 = R \frac{di(t)}{dt} + \frac{1}{C} i(t) + L \frac{d^2 i(t)}{dt^2}$

Omrokering $\frac{d^2 i(t)}{dt^2} + \frac{R}{L} \frac{di(t)}{dt} + \frac{1}{LC} i(t) = 0$

Karakterligningen giver rødderne til prototypen. Svingningens frekvens er $f = \omega/2\pi = 1003$ Hz så tiden for en periode er tæt på 1 ms. Realdelen σ er negativ så svingningen dør ud og er reduceret med $\exp(-1) = e^{-1} = 0,37$ efter to perioder for de viste komponentværdier.

Karakterligning: $a^2 + \frac{R}{L}a + \frac{1}{LC} = 0$

Parameter: $a = \frac{1}{2} \left(-\frac{R}{L} \pm \sqrt{\left(\frac{R}{L}\right)^2 - \frac{4}{LC}} \right) \Rightarrow$ Talværdier $a = \sigma \pm j\omega = \begin{pmatrix} -500 \text{ s}^{-1} + j6304 \text{ s}^{-1} \\ -500 \text{ s}^{-1} - j6304 \text{ s}^{-1} \end{pmatrix}$

Løsning: $i(t) = I_1 \exp([\sigma + j\omega]t) + I_2 \exp([\sigma - j\omega]t)$
 $i(t) = \exp(\sigma t) [I_1 \exp(j\omega t) + I_2 \exp(-j\omega t)]$

For at bestemme I_1 og I_2 benyttes viden om kredsløbet til tidspunktet $t = 0$. Strømmen $i(t)$ kan kun ændres gradvist så den er nul ved start og det giver $I_2 = -I_1$. Som en konsekvens heraf kan de to komplekse eksponentialfunktioner kombineres til en sinus som vist side 106.

$$\begin{array}{ll} \text{Strøm til } t = 0: & i(0) = I_1 + I_2 = 0 \Rightarrow I_2 = -I_1 \\ \text{Løsning} & i(t) = I_1 \exp(\sigma t) [\exp(j\omega t) - \exp(-j\omega t)] \\ \text{Euler} & i(t) = 2j I_1 \exp(\sigma t) \sin(\omega t) \end{array}$$

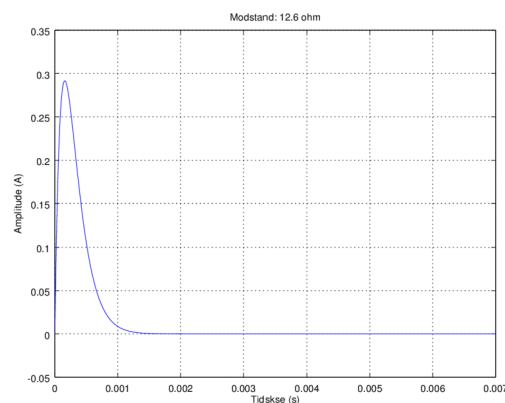
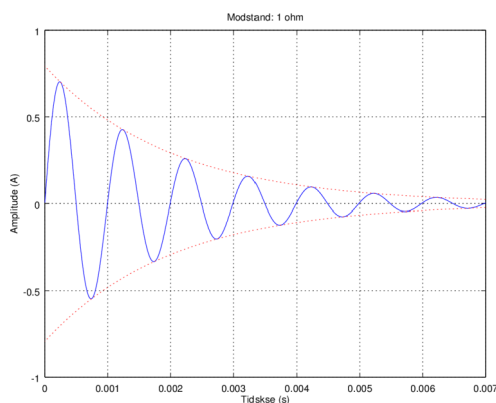
Værdien af I_1 findes ved starttidspunktet $t = 0$ som spændingen over spolen L der momentant får påtrykt spændingen U_0 . Spændingsfaldet over de to andre komponenter er nul da $i(0) = 0$.

$$\begin{array}{ll} \text{Hældning} & U_0 = L \frac{di(t)}{dt} = L \frac{d(2j I_1 \exp(\sigma t) \sin(\omega t))}{dt} \\ & U_0 = L 2j I_1 [\sigma \exp(\sigma t) \sin(\omega t) + \exp(\sigma t) \omega \cos(\omega t)] \\ \text{Startværdi } t = 0 & U_0 = 2j \omega L I_1 \Rightarrow I_1 = \frac{U_0}{2j \omega L} \\ \text{Løsning} & i(t) = \frac{U_0}{\omega L} \exp(\sigma t) \sin(\omega t) \end{array}$$

Løsningen er en dæmpet svingning. Grænsen til en aperiodisk løsning kræver reelle rødder og den findes hvor karakterligningens determinant bliver positiv. Det kræver $R > 12,65 \Omega$ og for så stor en værdi skal løsningen opskrives med to reelle eksponentialfunktioner.

$$\text{Reelle rødder: } \left(\frac{R}{L}\right)^2 > \frac{4}{LC} \Rightarrow \frac{R}{L} > \frac{2}{\sqrt{LC}} \Rightarrow \frac{R}{L} \sqrt{LC} > 2 \Rightarrow R > 2\sqrt{\frac{L}{C}} = 12,65 \Omega$$

Herunder vises strømmens variation over tid for svingningskredsen når den momentant tilsluttes en spændingskilde på 5 V. Til venstre vises reaktionen med de opgivne komponentværdier hvor det ses at svingningens amplitude aftager eksponentielt (stiplet kurve). Til højre er modstandens værdi øget så den er meget tæt på grænsen for aperiodisk dæmpning.



Svingningens maksimum findes ved at differentiere $i(t)$ og finde det første tidspunkt hvor hældningen er nul. Det er i første tilfælde forventet tæt på $t \approx 0,25$ ms hvor sinus er én.

$$\begin{array}{ll} \sigma \sin(\omega t) = -\omega \cos(\omega t) & \\ \frac{di(t)}{dt} = \frac{U_0}{\omega L} [\sigma \exp(\sigma t) \sin(\omega t) + \exp(\sigma t) \omega \cos(\omega t)] = 0 & \Rightarrow \tan(\omega t) = \frac{\sin(\omega t)}{\cos(\omega t)} = -\frac{\omega}{\sigma} \\ & t_{\max} = \frac{1}{\omega} \tan^{-1}\left(\frac{-\omega}{\sigma}\right) \\ R = 1 \Omega \quad \sigma = -500 \text{ rad/s} \quad \omega = 6304 \text{ rad/s} \quad t_{\max} = 237 \mu\text{s} \quad I_{\max} = 0,70 \text{ A} \\ R = 12,6 \Omega \quad \sigma = -6300 \text{ rad/s} \quad \omega = 557 \text{ rad/s} \quad t_{\max} = 158 \mu\text{s} \quad I_{\max} = 0,29 \text{ A} \end{array}$$

2.14.2 Anden ordens inhomogen differentialligning

En inhomogen differentialligning adskiller sig fra den homogene differentialligning ved at der er et drivende led som herunder kaldes $g(t)$. Det kan vises at den komplette løsning til en inhomogen differentialligning består af de to løsninger til den homogene differentialligning plus en såkaldt specifik løsning (*particular solution*) der giver $g(t)$ når den indsættes i ligningen. Opgaven kan derfor løses ved en to-trins proces. Første sættes $g(t) = 0$ for bestemmelse af den homogene differentiallignings løsning $x_H(t)$ og derefter adderes den specifikke løsning $x_P(t)$.⁷⁰

Differentialligning	$\frac{d^2 x(t)}{dt^2} + A \frac{dx(t)}{dt} + Bx(t) = g(t)$	
Løsning til homogene ligning	$x_H(t) = X_1 \exp(a_1 t) + X_2 \exp(a_2 t)$	
Løsning til differentialligning	$x(t) = x_H(t) + x_P(t)$	
	$g(t)$	$x_P(t)$
Prototype	$\exp(\alpha t)$ $\cos(\omega t), \sin(\omega t)$ t^n	$A \exp(\alpha t)$ $A \cos(\omega t) + B \sin(\omega t)$ $A_n t^n + A_{n-1} t^{n-1} + \dots + A_1 t + A_0$

Den specifikke løsning vil altid være at finde blandt en af tre mulige løsningstyper. Den er enten baseret på en eksponentialfunktion, de trigonometriske funktioner med samme frekvens som det drivende led eller et polynomium af samme grad som det drivende led. Ved påvirkning med en sinus vil det være bedst at prøve med en løsning med både cosinus og sinus for at inkludere muligheden for at fasen er ændret. Følgelig kan proben også skrives $A \cos(\omega t + \theta)$ da det giver samme type løsning (se side 86). Hvis det drivende led er en potensfunktion bør prototypen bruge alle lavere potenser.

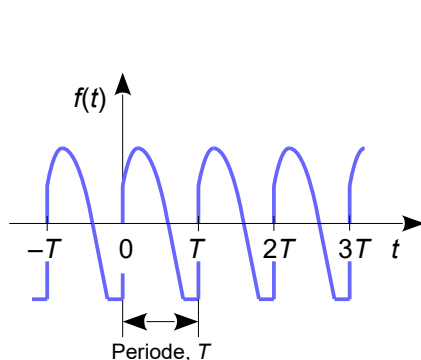
⁷⁰ <http://tutorial.math.lamar.edu/Classes/DE/NonhomogeneousDE.aspx>.

2.15 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 kendte.⁷¹

Fouriers arbejde prægede udviklingen indenfor matematikken også selv om han mødte en del modstand for den pæne sinus var ifølge Fourier i stand til at eftergøre diskontinuerte svingninger. Det er funktioner som springer abrupt mellem værdierne og det havde samtidens matematikere 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 ved digital signalbehandling hvor et signal skal filtreres eller analyseres for et bestemt indhold.

En Fourier serie efterligner en svingning der fx kan være et musikinstruments 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 periode 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 $f_1, 2f_1, 3f_1, \dots$ med individuelle vægte ved koefficienterne a_n og b_n . Serien går imod $f(t)$ når antallet af svingninger n går imod uendelig.



Funktion:

$$f(t) = f(t+mT), \quad m = 0, \pm 1, \pm 2 \dots, \quad f_1 = \frac{1}{T}$$

Fourier serie:

$$f(t) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(2\pi n f_1 t) + b_n \sin(2\pi n f_1 t)]$$

$$a_0 = \frac{1}{T} \int_{\tau}^{\tau+T} f(t) dt$$

$$a_n = \frac{2}{T} \int_{\tau}^{\tau+T} f(t) \cos(2\pi n f_1 t) dt \quad c_n = \sqrt{a_n^2 + b_n^2}$$

$$b_n = \frac{2}{T} \int_{\tau}^{\tau+T} f(t) \sin(2\pi n f_1 t) dt \quad n = 1, 2, 3, \dots$$

Figur 117 – Svingningen skal være periodisk og alle perioder skal være identiske.

Koefficienten a_0 er middelværdien af svingningen og angiver en eventuel DC komponent. De to koefficienter a_n og b_n bestemmes som vist i illustrationen hvor en integration fra τ til $\tau + T$ blot betyder at starttidspunktet for integrationen kan vælges frit. Det er kun krævet at der integreres over en hel periode. Alle integraler kan derfor foretages hvor det er lettest at slippe igennem og det er typisk fra 0 til T hvor $\tau = 0$ eller fra $-T/2$ til $T/2$ hvor $\tau = -T/2$. Amplituden af den n 'te harmoniske c_n er som regel vigtigere end faseinformationen.

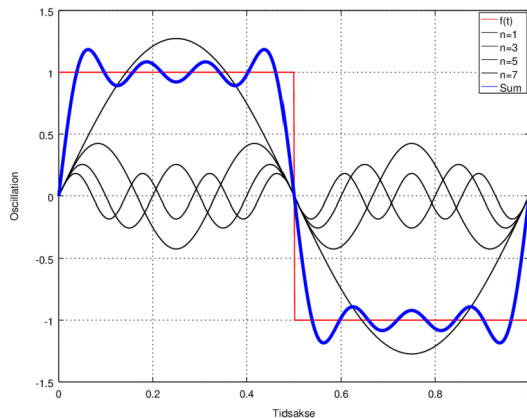
Teorien bag bestemmelse af koefficienterne uddybes på side 121.

En konklusion på Fouriers arbejde er at de analyser vi laver på elektriske kredsløb kan bruges på alle former for komplicerede signaler blot de 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 også karakterisere et elektronisk system ved brug af en harmonisk svingning og derefter beregne hvordan det vil fungere ved et diskontinuerligt signal som en firkantsvingning.

⁷¹ Det er ikke kun sinus og cosinus der har denne egenskab. Andre muligheder er Bessel, Chebychev og Walsh funktioner.

2.15.1 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 funktioner.⁷²



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

$$f(t) = \sum_{n \text{ ulige}} \frac{4}{n\pi} \sin(n 2\pi \frac{t}{T})$$

Harmoniske:

$$n = 1 \quad \frac{4}{\pi} = 1,2732 \quad +2,1 \text{ dB}$$

$$n = 3 \quad \frac{4}{3\pi} = 0,4244 \quad -7,4 \text{ dB}$$

$$n = 5 \quad \frac{4}{5\pi} = 0,2546 \quad -11,9 \text{ dB}$$

Figur 118 – 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) idet 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, ...) 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 $f(t)\cos(2\pi n f_1 t)$ 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.

$$a_n = \frac{2}{T} \left(\int_0^{T/2} \cos\left(\frac{n 2\pi}{T} t\right) dt - \int_{T/2}^T \cos\left(\frac{n 2\pi}{T} t\right) dt \right) = 0 \Rightarrow a_n = 0, \quad a_0 = 0, \quad n = 1, 2, 3, \dots$$

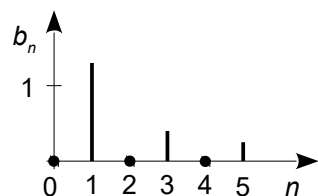
Signalets symmetri medfører desuden at alle lige harmoniske for b_n er nul. Det udnyttes her at $\cos(n\pi)$ svinger med ± 1 for lige og ulige værdi af index n og at $\cos(2n\pi)$ er lig med én for alle værdier af n .

$$b_n = \frac{2}{T} \left(\int_0^{T/2} \sin\left(\frac{n 2\pi}{T} t\right) dt - \int_{T/2}^T \sin\left(\frac{n 2\pi}{T} t\right) dt \right)$$

$$b_n = \frac{1}{n\pi} \left(\left[-\cos\left(\frac{n 2\pi}{T} t\right) \right]_0^{T/2} - \left[-\cos\left(\frac{n 2\pi}{T} t\right) \right]_{T/2}^T \right)$$

$$b_n = \frac{1}{n\pi} ([-\cos(n\pi) + 1] - [-\cos(2n\pi) + \cos(n\pi)])$$

$$b_n = \frac{4}{n\pi} (1 - (-1)^n) \Rightarrow b_n = \begin{cases} 0 & n = 2, 4, 6, \dots \\ 4/n\pi & n = 1, 3, 5, \dots \end{cases}$$



⁷² 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.

2.16 Kompleks Fourier serie

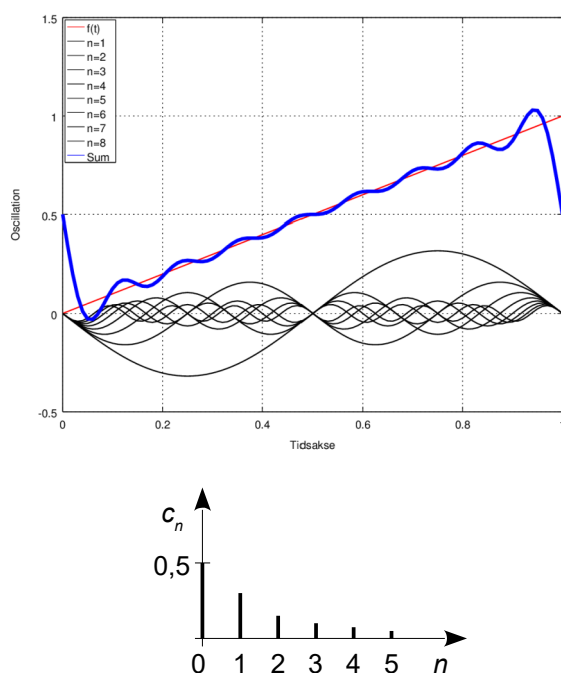
En hensigtsmæssig omformning af Fourier serien benytter de komplekse repræsentanter for sinus og cosinus ved brug af Eulers identitet fra side 106 der indsættes i formlerne for a_n og b_n og koefficienterne samles til en kompleks amplitude defineret som $c_n = (a_n - jb_n)/2$. Efter lidt arbejde kan den komplekse Fourier serie skrives som vist herunder. Bemærk at der er to måder at skrive serien på. Udtrykkene til højre beregner amplituden (modulus) og fassen af den n 'te harmoniske.

$$\begin{aligned} \text{Metode 1:} \quad & f(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_1 t) \\ \text{Metode 2:} \quad & f(t) = c_0 + \sum_{n=1}^{\infty} (c_n \exp(j2\pi n f_1 t) + c_{-n} \exp(-j2\pi n f_1 t)) \\ \text{Koefficienter:} \quad & c_n = \frac{1}{T} \int_{\alpha}^{\alpha+T} f(t) \exp(-j2\pi n f_1 t) dt \end{aligned} \quad \begin{aligned} f_1 &= \frac{1}{T} \\ |c_n| &= \sqrt{(\operatorname{Re}(c_n))^2 + (\operatorname{Im}(c_n))^2} \\ \theta_n &= \arctan\left(\frac{\operatorname{Im}(c_n)}{\operatorname{Re}(c_n)}\right) \end{aligned}$$

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

2.16.1 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.



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:

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

Figur 119 – 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.

2.16.2 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 fx $\cos(2\pi n f_1 t)$ ved at gange funktionen med svingningen og integrere over en hel periode. Hvis svingningen optræder som del af funktionen vil $\cos(2\pi n f_1 t)$ blive kvadreret og er dermed positiv så integrationen giver et bidrag og 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.

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

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 118.

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

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:

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

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 eksponentialfunksions 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.

$$\begin{aligned}\int_0^T f(t) \exp(-j 2\pi m f_1 t) dt &= \int_0^T \left(\sum_{n=-\infty}^{\infty} c_n \exp(j 2\pi (n-m) f_1 t) \right) dt \\ \int_0^T f(t) \exp(-j 2\pi m f_1 t) dt &= \sum_{n=-\infty}^{\infty} \left(c_n \int_0^T \exp(j 2\pi (n-m) f_1 t) dt \right) \\ m = n &\Rightarrow \int_0^T f(t) \exp(-j 2\pi n f_1 t) dt = c_n \int_0^T dt = c_n T \Rightarrow c_n = \frac{1}{T} \int_0^T f(t) \exp(-j 2\pi n f_1 t) dt \\ m \neq n &\Rightarrow \int_0^T f(t) \exp(-j 2\pi n f_1 t) dt = 0\end{aligned}$$

2.17 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 kontinuert spektrum. Den komplekse koefficient c_n bliver til en funktion af frekvensen og kaldes for $F(f)$.

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

Faktoren $1/T$ foran integralet indgår ikke i Fourier transformationen så dimensionen ændres. Hvis $f(t)$ er en spændingskilde har den Fourier transformerede dimension af volt gange sekund.

En del litteratur erstatter frekvensen f med vinkelhastigheden $\omega = 2\pi f$ der medfører division med 2π ved den inverse Fourier-transformation idet $d\omega = 2\pi df$. Visse matematiske værker "fordeler" denne konstant ligeligt mellem transformationens to dele ved at gange med $1/\sqrt{2\pi}$ begge steder. Det er en misforståelse, men den er meget udbredt og adskillige tabelværker har denne konstant inkluderet.

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.

2.17.1 Rektangulær puls

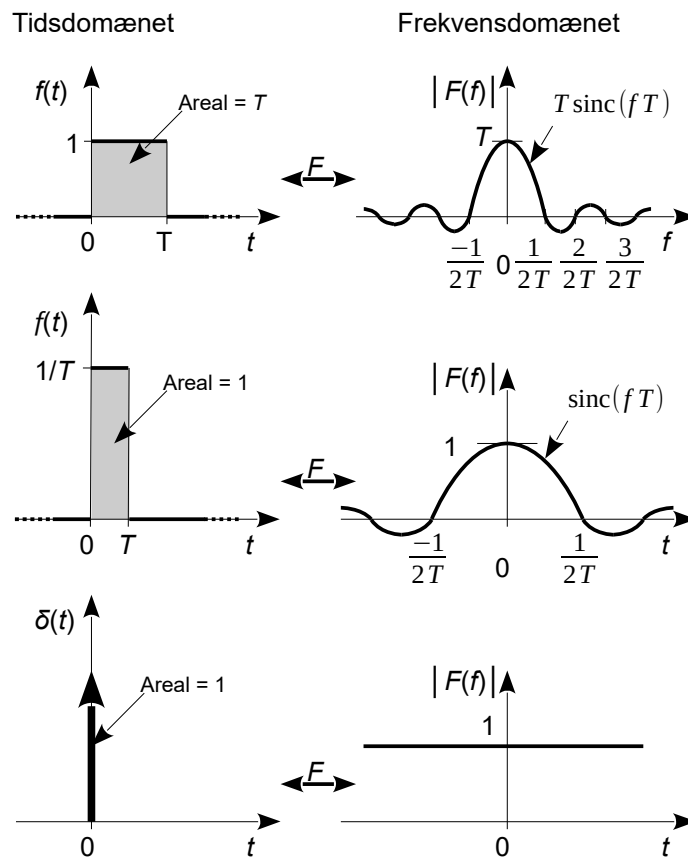
Som eksempel på Fourier transformationen vises herunder resultatet for en rektangulær puls $f(t)$ der er nul for al negativ tid, der til tidspunktet $t = 0$ springer til værdien én hvor den forbliver til $t = T$ for så at springe tilbage til nul. Det er en ofte benyttet puls i videnskabelig teori og 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 97, 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 $\text{sinc}(x) = \sin(x)/x$ 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 nær ved 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 $1/T = 500$ 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.

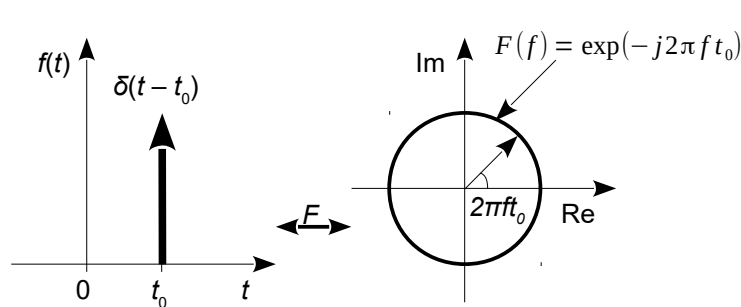


Figur 120 – En rektangulær puls har mest signalenergi ved lave frekvenser mens energien ved de høje frekvenser aftager i styrke og er nul ved de specifikke frekvenser $f = n/2T$ hvor n er et helt tal.

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 imod uendelig mens arealet forbliver én. Det kaldes en Dirac puls der benyttes indenfor teoretisk fysik og ved digital signalbehandling.

2.17.2 Dirac puls

En Dirac puls $\delta(t)$ "fyrrer" når argumentet er nul og 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 fyrrer 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 en kompleks oscillation i frekvensdomænet.



Figur 121 – Dirac pulsen fyrrer når argumentet er nul. Fourier transformationen giver en kompleks eksponentialfunktion hvor amplituden er én mens fasen roterer med en hastighed givet ved t_0 .

Rektangulær puls med amplituden 1:

$$f(t) = \begin{cases} 1 & 0 < t < T \\ 0 & \text{ellers} \end{cases}$$

Fourier transformation:

$$F(f) = \int_{-\infty}^{\infty} f(t) \exp(-j2\pi f t) dt$$

$$F(f) = \int_0^T \exp(-j2\pi f t) dt$$

$$F(f) = \frac{1}{j2\pi f} [\exp(-j2\pi f t)]_0^T$$

$$F(f) = \frac{1}{j2\pi f} [\exp(-j2\pi f T) - 1]$$

$$F(f) = \frac{\sin(2\pi f T)}{2\pi f} + j \frac{1 - \cos(2\pi f T)}{2\pi f}$$

Modulus af $F(f)$:

$$|F(f)| = \sqrt{\text{Re}^2\{H(f)\} + \text{Im}^2\{H(f)\}}$$

$$|F(f)| = T \frac{\sin(\pi f T)}{\pi f T}$$

$$|F(f)| = T \text{sinc}(fT)$$

Modulus af $F(f)$ ved amplituden $1/T$:

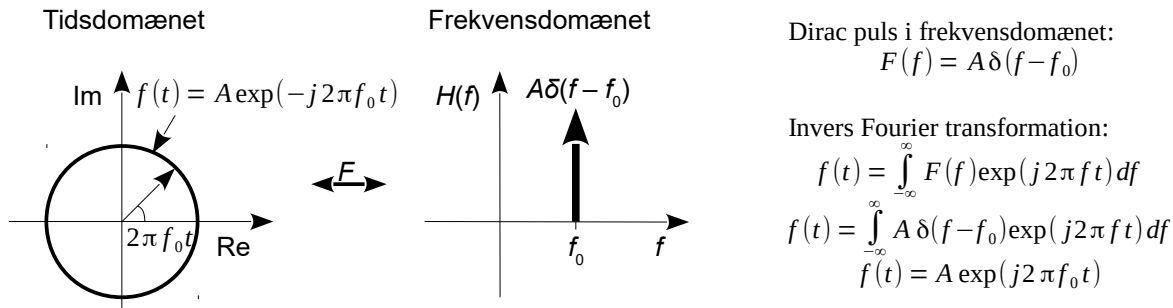
$$|F(f)| = \text{sinc}(fT)$$

Modulus af $F(f)$ for $T \rightarrow 0$:

$$|F(f)| = 1$$

2.17.3 Kompleks harmonisk svingning

En kompleks eksponentialfunktion i tidsdomænet bliver Fourier transformeret 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 .



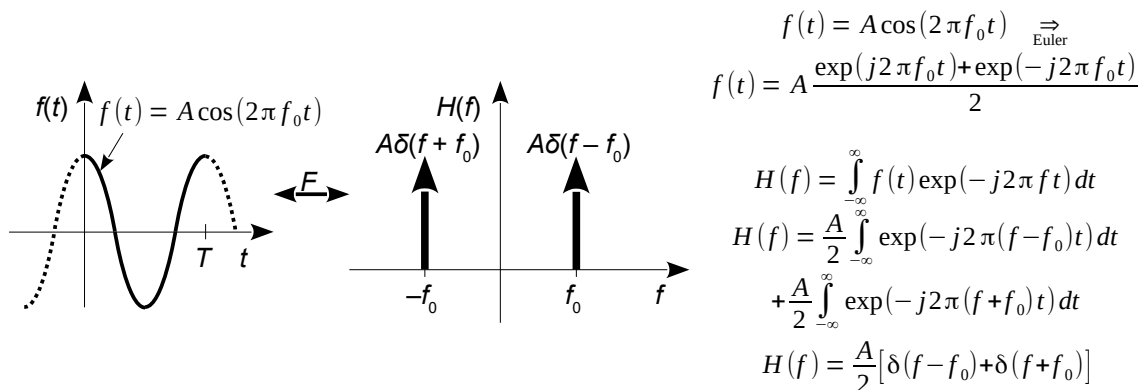
Figur 122 – En harmonisk oscillation transformeres til en Dirac puls i frekvensdomænet.

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 119 og 120. I frekvensdomænet optræder disse svingninger som en serie af Dirac pulser placeret ved de harmoniske frekvenser og med individuel vægtning.

2.17.4 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.



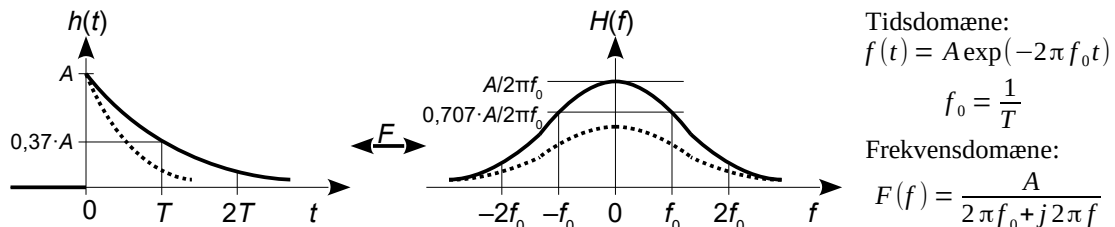
Figur 123 – En cosinus transformeres til to Dirac pulser i frekvensdomænet.

2.17.5 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. Der gælder det samme med negativ frekvens som derfor må opfattes som endnu et lag i den matematiske abstraktion fra det fysiske.

2.17.6 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 124 – 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 dB/oktav.

Frekvensspektret er tosidigt hvilket vil sige at Fourier-transformationen benytter både positive og negative frekvenser. Det har primært akademisk interesse så for en ingeniørmæssig 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 faser for de to frekvensområder.

Lyden af signalet er et "klik". Båndbredden er givet ved den hastighed hvormed udgangssignalet glider på plads. Grænsefrekvensen (-3 dB) nås når frekvensen er $f = f_0 = 1/T$ hvor nævneren af polynomiet bliver $(1 + j)2\pi f_0$ og lyden er kraftigst når T er en lang tid.

Ved høje frekvenser ($f > f_0$) er niveauet af den Fourier transformerende omvendt proportional med frekvensen. Det ses ved at udtrykket kan simplificeres til $F_{HF}(f) = A/2\pi f$ for høje frekvenser så amplituden er aftagende ved stigende frekvens over grænsefrekvensen.

Ved lave frekvenser ($f < f_0$) er den Fourier transformerendes 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.

$$f(t) = \begin{cases} A \exp(-2\pi f_0 t) & t \geq 0 \\ 0 & t < 0 \end{cases} \quad f_0 = \frac{1}{T}$$

$$F(f) = A \int_0^{\infty} \exp(-2\pi f_0 t) \exp(-j 2\pi f t) dt$$

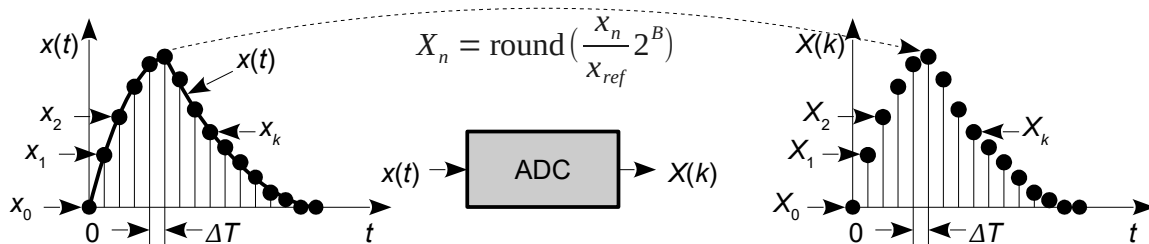
$$F(f) = A \int_0^{\infty} \exp(-[2\pi f_0 + j 2\pi f] t) dt$$

$$F(f) = -\frac{A}{2\pi f_0 + j 2\pi f} [\exp(-[2\pi f_0 + j 2\pi f] t)]_0^{\infty}$$

$$F(f) = \frac{A}{2\pi f_0 + j 2\pi f}$$

2.18 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 125 – 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 bipolære 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 fundamentale frekvens f_1 er givet ved det krav at netop én periode af det oscillerende 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 svingningerne 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:	$f_{\max} = Nf_1 < \frac{f_s}{2}$	$N < \frac{K}{2}$	f_1 Mindste frekvensopløsning
Frekvens for dataopsamling:	$f_s = \frac{K}{T} = K f_1$		f_s Opsamlingsfrekvens
			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 tidligere 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.

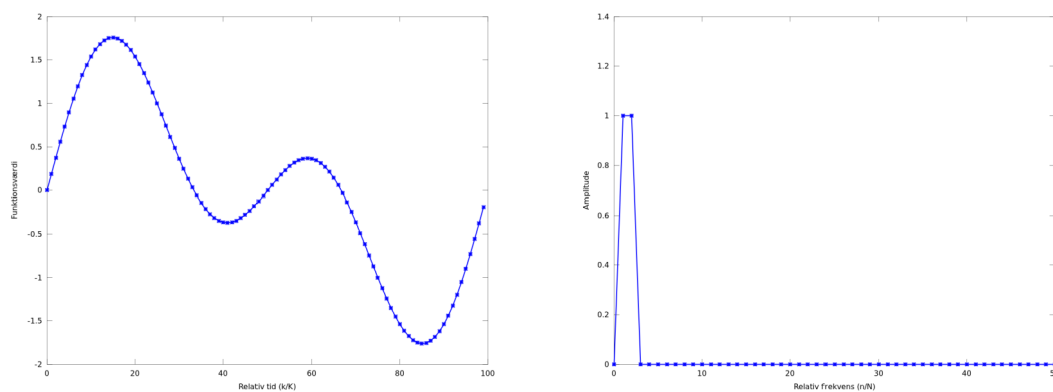
$$a_n = \frac{2}{T} \sum_{k=0}^{K-1} \int_{kT}^{(k+1)T} h(k\Delta T) \cos(2\pi \frac{n}{T} k\Delta T) dt \Rightarrow a_n = \frac{2}{K} \sum_{k=0}^{K-1} h(k\Delta T) \cos(2\pi \frac{nk}{K})$$

Et systems respons udtrykt som amplitude og fase som funktion af frekvensen beregnes som vist herunder hvor det bedste resultat for faserne findes hvis arctan kan bringes til at dække over alle fire kvadranter i den komplekse plan. Den funktion hedder Atan2 i mange programmeringssprog.

$$|c_n| = \sqrt{a_n^2 + b_n^2} \quad \text{og} \quad \theta_n = \arctan\left(\frac{b_n}{a_n}\right)$$

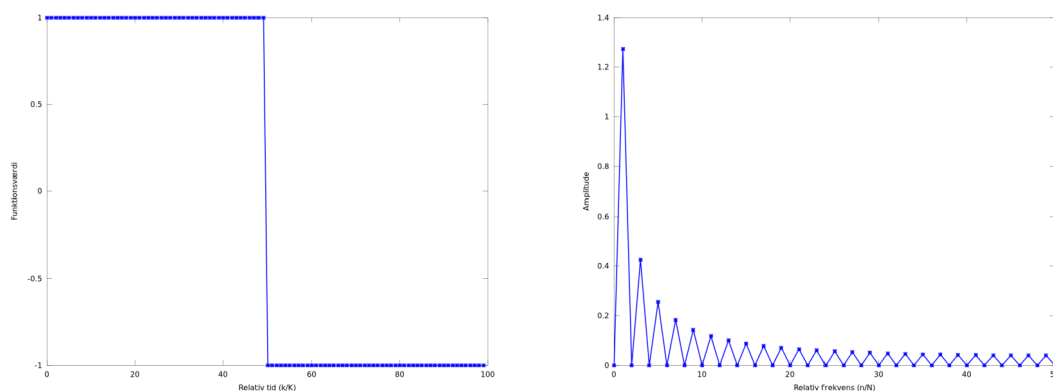
Algoritmen skal illustreres ved to signaler hvor den første består af to harmoniske svingninger på henholdsvis repetitionsfrekvensen f_1 og to gange denne frekvens med samme amplitude for hver svingning så det beregnede spektrum har derfor markering ved de relative frekvenser 1 og 2 samt nul for alle 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 firsntsvingning 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 126 – To svingninger med samme amplitude og frekvenserne f_1 og $2f_1$ giver kun uddata for de tilsvarende frekvenspunkter 1 og 2.

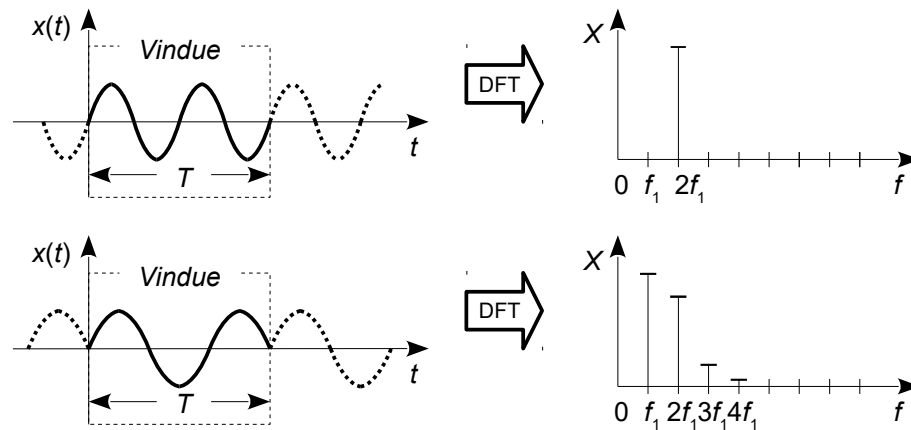
Den symmetriske firkantsvingning er analyseret tidligere i kapitlet og har amplituden af første harmoniske ved $4/\pi = 1,27$ og der er kun koefficienter for ulige frekvenser som det ses herunder. Niveauet ved 0 Hz repræsenterer DC niveauet i svingningen og er her lig med nul.



Figur 127 – 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.



Figur 128 – 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 nuller hvilket på engelsk kaldes for *zero padding*. Idéen er at gøre perioden T længere så repetitionsfrekvensen bliver lavere, men det tilføjer 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-løkken for funktionsværdierne. De to løkker kan derved reduceret til en enkelt for-løkke fordi koefficienterne $a(n)$ 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:K)=-1; % Firkant uden DC.
% SPEKTRUM.
n=0:N; a(n+1)=0; b(n+1)=0; % Forbered frekvens.
for k=0:K-1 % Løb data igennem.
    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; % Korrigør 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.

2.19 Laplace transformation

Pierre-Simon Laplace (1749 – 1827) var en fransk matematiker og astronom der lagde grunden til det vi nu kalder for Laplace transformationen. Han har bidraget væsentligt indenfor flere områder af matematik og naturvidenskab og var den første til at indse muligheden af at stjerner kunne være så store og deres gravitation så voldsom at lys ikke kunne undslippe. Laplace fulgte 1785 i sporet af Euler og Lagrange med studier af hvordan differentialligninger kunne løses. Han opdagede at ved integration af hele differentialligningen blev resultatet lettere at løse.⁷³

Samtiden havde ikke nytte heraf og denne del af Laplaces arbejde blev glemt. Omkring midten af 1800-tallet var et af de store problemer at forstå hvorfor telegrafstationens signaler blev forvansket undervejs til modtagerstationen. Forklaringen blev givet af Oliver Heaviside der havde opbygget sin egen matematik baseret på Fourier transformationen og ved hjælp af den fandt han frem til en løsning på problemet. Det viste sig at være meget lig det Laplace havde fundet frem til og herefter blev Laplace transformationen genopdaget og videreudviklet.

2.19.1 Introduktion

Laplace transformationen tager en tidsafhængig funktion $f(t)$ og omformer den til en funktion $F(s)$ af den komplekse variabel s . Det er vigtigt at forstå at $F(s)$ indeholder helt den samme information om funktionen som $f(t)$ så der mistes intet ved transformationen.

$$F(s) = \int_0^{\infty} f(t) \exp(-st) dt \quad \text{hvor} \quad f(t) = 0 \quad \text{for} \quad t < 0$$

$$s = \sigma + j\omega \quad \text{hvor} \quad \omega = 2\pi f$$

Den Laplace transformerede funktion $F(s)$ får enheden s/rad (sekund per radian) gange med enheden af funktionen $f(t)$. Det betyder at strøm og spænding ikke bare transformeres til strøm og spænding, men til en *abstraktion* fra strøm og spænding.

Hvis du sammenligner med Fourier transformationen fra side 122 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 Fourier transformationen. Det er med andre ord et særdeles effektivt værktøj.

Den komplekse variabel s består af realdelen σ og den imaginære del ω . At s indeholder $j\omega$ giver et fingerpeg om at hele værktøjet fra de forrige afsnit om impedans kan benyttes.

Det er en forenkling blot at beskrive s som en frekvensvariabel for det er et punkt i den komplekse talplan, det er matematisk set en *operator* der udfører en differentiation og analysen er ikke begrænset til noget oscillerende.

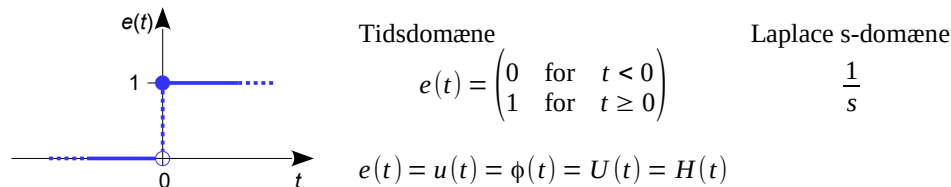
Du vil på de følgende sider opleve en matematik der virker meget fremmedartet. Mit råd er at være åben overfor det nye for jeg garanterer at tågen vil lette og at solen kommer til at skinne igen.

Jeg vil starte med at illustrere hvordan Laplace transformationen håndterer indkobling af en spændingskilde på et kredsløb. Derefter viser jeg hvordan nogle kurveformer kan opbygges for påvirkning af det elektroniske kredsløb.

⁷³ https://en.wikipedia.org/wiki/Pierre-Simon_Laplace.

2.19.2 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 antager én af to diskrete værdier afhængig af tidsvariablen. For negativ tid er funktionens værdi nul og ellers er den én. Jeg benytter symbolet $e(t)$, men kært barn har mange navne og jeg lister en del af de mulige repræsentationer.



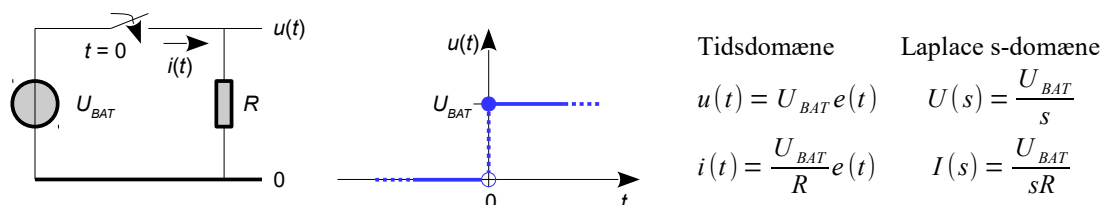
Figur 129 – Stepfunktionen repræsenterer en kontakt der sluttet til et givet tidspunkt.

Funktionen bruges til at beskrive en påvirkning der indkobles til tidspunktet $t = 0$ og derefter kan man så studere hvad der sker i kredsløbet.

Funktionen $e(t)$ er nul inden integrationen starter så alt hvad der behøves er at erstatte $f(t)$ i Laplace transformationen med et ét-tal og integrere over tid. Argumentet er komplekst ($-st$), men det følger de sædvanlige regneregler for integration.

$$I(s) = \int_0^{\infty} e(t) \exp(-st) dt = \int_0^{\infty} \exp(-st) dt = -\frac{1}{s} [\exp(-st)]_0^{\infty} = -\frac{1}{s} [0 - 1] = \frac{1}{s}$$

I eksemplet herunder benyttes en kontakt for at indkoble spændingskilden U_{BAT} til en modstand R og det sker til tidspunktet $t = 0$. Spændingskilden U_{BAT} optræder som en konstant faktor til $1/s$ og det gør strømmen U_{BAT}/R også.

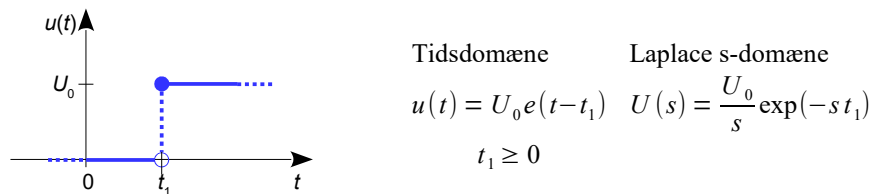


Figur 130 – Momentan indkobling af et batteri på et kredsløb.

Udtrykket kan virke mystisk, men Laplace transformationen leder til nogle pudsige udtryk. Lidt fornuft er der dog i det for s repræsenterer en frekvens gennem variabelen $\omega = 2\pi f$. Relationen fra frekvens til tid er $f = 1/T$ så en lav frekvens svarer til en lang periode. Værdien af $I(s)$ vokser over alle grænser for s gående imod nul hvilket svarer til uendelig 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. En lav værdi af s svarer til lave frekvenser og det giver kun mening når 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.

2.19.3 Tidsforsinkelse

Et enhedsstep behøver ikke at skifte til tidspunktet $t = 0$. En forsinkelse på tiden t_1 i indkoblingen skrives $e(t - t_1)$ idet enhedssteppet først skifter når differensen i parenteser ændrer fortegn. Det eneste nye i udtrykket er multiplikationen med $\exp(-st_1)$ der giver en ændring i faser.



Figur 131 – Et step kan forskydes i tid og konsekvensen er en frekvensafhængig fase.

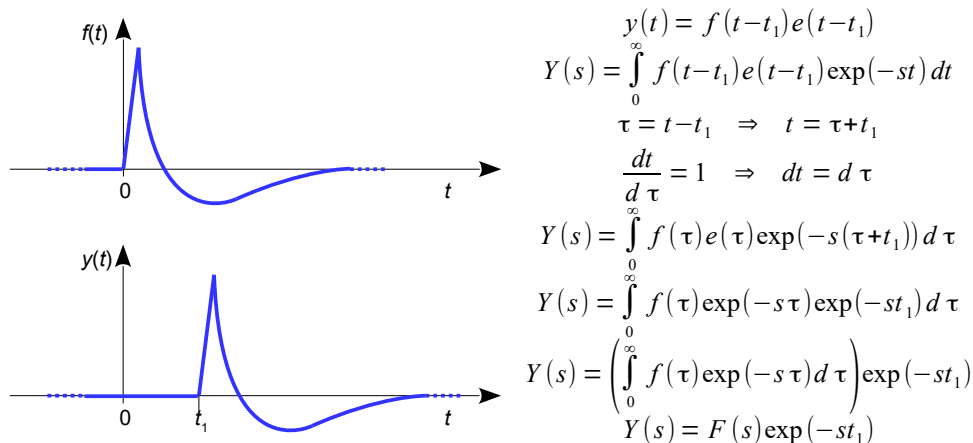
Laplace transformationen af $e(t)$ kaldes herunder $f(t)$ og beregnes som før ved indsættelse. Det udnyttes at $e(t - t_1)$ er nul for $t < t_1$ så integrationen giver først bidrag fra tidspunktet t_1 og op.

$$U(s) = \int_0^{\infty} U_0 e(t - t_1) \exp(-st) dt = U_0 \int_{t_1}^{\infty} \exp(-st) dt$$

$$U(s) = \frac{U_0}{-s} [\exp(-st)]_{t_1}^{\infty} = \frac{U_0}{-s} [0 - \exp(-st_1)] = \frac{U_0}{s} \exp(-st_1)$$

Hvis man ignorerer realdelen af s er eksponentialfunktionens amplitude én og faseren er $-2\pi ft$. Overfor et oscillerende signal med frekvensen $f = 1/t_1$ betyder det at faseren ændres med 2π for et tidsforløb fra 0 til t_1 . Ved højere frekvens vokser faseren hurtigere svarende til at signalet kan nå at gennemføre flere perioder indenfor et givet tidsrum. En tidsforsinkelse beskrives derfor som en negativ fase der varierer proportionalt med frekvensen.

Enhver funktion $f(t)$ kan indkobles forsinket med start til tidspunktet t_1 hvilket vil sige at funktionen er nul indtil det nye startpunkt. Funktionen skifter ikke udseende, men flyttes blot i tid.

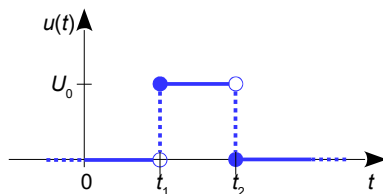


Figur 132 – En vilkårlig funktion kan forsinkes ved at Laplace transformere funktionen og gange med $\exp(-st_1)$ der svarer til en fasedrejning.

Ved Laplace transformationen benyttes substitution af tidsvariablen så $t - t_1$ skrives som τ . Funktionen starter først når enhedssteppet $e(\tau)$ skifter til én så Laplace transformationen kan starte fra $\tau = 0$. Eksponentialfunktionen opdeles i to ved $\exp(a + b) = \exp(a)\exp(b)$ og da det ene led er en konstant kan det sættes udenfor integrationen. Resultatet genkendes som en Laplace transformation af den oprindelige funktion blot med den nye variabel τ hvilket er uden betydning.

2.19.4 Tidsvindue

To stepfunktioner kan kombineres til et praktisk værktøj hvor den første stepfunktion åbner for tidsvinduet klokken $t = t_1$ og den anden stepfunktion lukker igen klokken $t = t_2$. Den firkantede parentes i tidsdomænet er 1 i tidsrummet fra t_1 til t_2 og 0 udenfor dette tidsrum.



Tidsdomæne

$$u(t) = U_0[e(t-t_1) - e(t-t_2)]$$

$$0 \leq t_1 \leq t_2$$

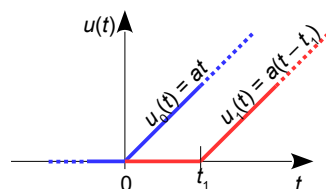
Laplace s-domæne

$$U(s) = \frac{U_0}{s}[\exp(-st_1) - \exp(-st_2)]$$

Figur 133 – Stepfunktionen kan benyttes til at tænde og slukke for et batteri med spændingen U_0 .

2.19.5 Rampe

En spænding eller strøm der vokser over alle grænser er ikke realistisk, men det kan være en god model for et kort øjeblik af et elektronisk kredsløb hvor noget sættes i gang. For eksempel ved en DC-DC konverter hvor en transistor går ON og strømmen i en spole vokser lineært.



Tidsdomæne

$$u_0(t) = at e(t)$$

$$u_1(t) = a(t-t_1)e(t-t_1)$$

Laplace s-domæne

$$U_0(s) = \frac{a}{s^2}$$

$$U_1(s) = \frac{a}{s^2} \exp(-st_1)$$

Figur 134 – En rampe der starter klokken nul og stiger med hældningen a (fx ampere per sekund).

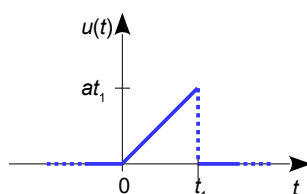
Udledningen benytter partiel integration med $f(t) = t$ og $g(t) = \exp(-st)$. Det første led bliver nul for t gående imod uendelig idet $\exp(-st) = 1/\exp(st)$ går mod nul med stigende orden.

$$U(s) = \int_0^{\infty} a t e(t) \exp(-st) dt = a \int_0^{\infty} t \exp(-st) dt \Rightarrow U(s) = a \left[\frac{t \exp(-st)}{-s} \right]_0^{\infty} - a \int_0^{\infty} \frac{\exp(-st)}{-s} dt$$

$$\int f(t) g(t) dt = f(t) G(t) - \int f'(t) G(t) dt \Rightarrow U(s) = \frac{a}{s} \left[\frac{\exp(-st)}{-s} \right]_0^{\infty} = \frac{a}{s^2}$$

2.19.6 Udsnit af rampe

Benyttes tidsvinduet på rampen kan den ændres til kun at eksistere i tidsrummet fra 0 til t_1 for derefter at springe tilbage til nul. Hvis sidste led i $U(s)$ fjernes forbliver funktionen på at_1 efter t_1 .



Tidsdomæne

$$u(t) = at[e(t) - e(t-t_1)]$$

Laplace s-domæne

$$U(s) = \frac{a}{s^2}[1 - \exp(-st_1)] - \frac{at_1}{s} \exp(-st_1)$$

Figur 135 – En rampe ganget med et tidsvindue.

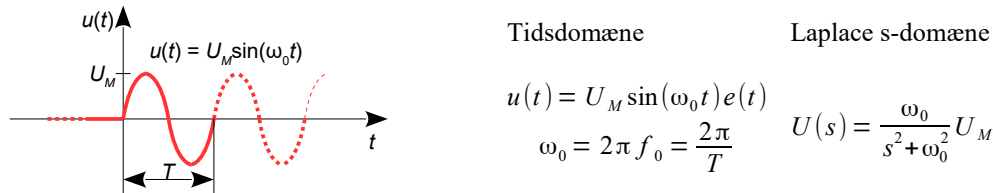
$$u(t) = at[e(t) - e(t-t_1)] = at e(t) - at e(t-t_1)$$

$$u(t) = at e(t) - a(t-t_1+t_1)e(t-t_1) = at e(t) - a(t-t_1)e(t-t_1) - at_1 e(t-t_1)$$

$$U(s) = \frac{a}{s^2} - \frac{a}{s^2} \exp(-st_1) - \frac{at_1}{s} \exp(-st_1) = \frac{a}{s^2}[1 - \exp(-st_1)] - \frac{at_1}{s} \exp(-st_1)$$

2.19.7 Oscillerende funktioner

Laplace transformationen kan håndtere oscillerende funktioner som det skal vises for en sinus der starter til tidspunktet nul og derefter fortsætter med at oscillere.



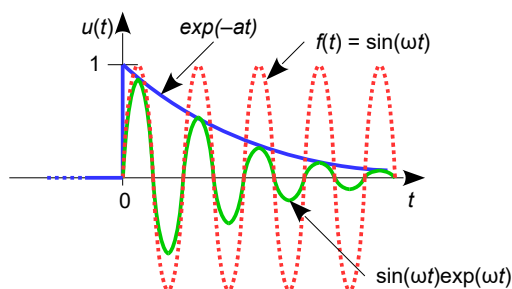
Figur 136 – En sinus der starter til tidspunktet nul og oscillerer i al tid derefter.

Udledningen benytter Eulers formel til at omskrive sinus til den komplekse eksponentialfunktion. Definitionen $s = \sigma + j\omega$ benyttes for at finde funktionsværdien for tiden gående mod uendelig. Eksponentialfunktionen til den imaginære del er begrænset til ± 1 når tiden vokser mens den reelle del af eksponentialfunktionen går imod nul. Herefter sættes på fælles brøk og reduceres.

$$\begin{aligned}
 U(s) &= \int_0^{\infty} U_M \sin(\omega_0 t) \exp(-st) dt \\
 \sin(\omega_0 t) &= \frac{\exp(j\omega_0 t) - \exp(-j\omega_0 t)}{2j} \Rightarrow U(s) = U_M \int_0^{\infty} \frac{\exp(j\omega_0 t) - \exp(-j\omega_0 t)}{2j} \exp(-st) dt \\
 U(s) &= \frac{U_M}{2j} \int_0^{\infty} \exp(j\omega_0 t) \exp(-st) dt - \frac{U_M}{2j} \int_0^{\infty} \exp(-j\omega_0 t) \exp(-st) dt \\
 U(s) &= \frac{U_M}{2j} \int_0^{\infty} \exp((j\omega_0 - s)t) dt - \frac{U_M}{2j} \int_0^{\infty} \exp((-j\omega_0 - s)t) dt \\
 U(s) &= U_M \left[\frac{\exp(j(\omega_0 - \omega)) \exp(-\sigma t)}{2j(j\omega_0 - s)} \right]_0^{\infty} - U_M \left[\frac{\exp(-j(\omega_0 + \omega)) \exp(-\sigma t)}{2j(-j\omega_0 - s)} \right]_0^{\infty} \\
 U(s) &= -\frac{U_M}{2j(j\omega_0 - s)} - \frac{U_M}{2j(-j\omega_0 - s)} = \frac{j\omega_0 + s + j\omega_0 - s}{2j(j\omega_0 - s)(-j\omega_0 - s)} U_M = \frac{\omega_0}{s^2 + \omega_0^2} U_M
 \end{aligned}$$

2.19.8 Dæmpning af en funktion

En funktion kan bringes til at klinge ud ved at gange med $\exp(-at)$ hvor a er en positiv konstant. Det transformerede udtryk ses at være modificeret alene ved at s ændres til $s + a$.



Tidsdomæne	Laplace s-domæne
$u(t) = f(t) \exp(-at)$	$U(s) = F(s+a)$
$a > 0$	

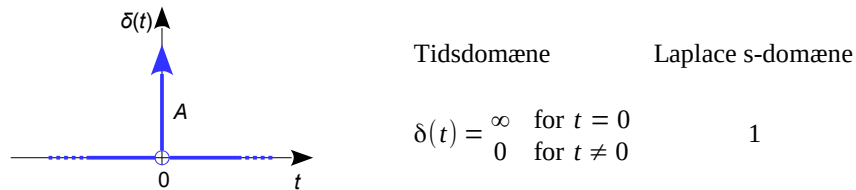
Figur 137 – Dæmpning af en funktion er her illustreret ved en sinus.

Udledningen benytter substitutionen $p = s + a$ for at identificere Laplace transformationen $F(p)$.

$$\begin{aligned}
 u(t) &= f(t) \exp(-at) \\
 U(s) &= \int_0^{\infty} f(t) \exp(-at) \exp(-st) dt = \int_0^{\infty} f(t) \exp(-(s+a)t) dt \\
 \Rightarrow_{p=s+a} U(s) &= \int_0^{\infty} f(t) \exp(-pt) dt = F(p) = F(s+a)
 \end{aligned}$$

2.19.9 Impuls

En speciel form for signal er impulsen der angiver en pludselig påvirkning som fx et slag med en hammer mod et søm eller et stød mellem to kugler i et spil billard. Den kaldes også for Dirac delta efter Paul Dirac (1902 – 1984) der var en af bidragsyderne ved kvantemekanikken. Den angiver at til tidspunktet $t = 0$ tilføres et system et kvantum af energi hvorefter systemet overlades til sig selv igen. Impulsen tegnes som en pil der angiver tidspunktet hvor impulsen "fyrrer".



Figur 138 – En deltaimpuls eksisterer kun i et punkt og defineres ved arealet på én. For en anden værdi angives den som et symbol ved siden af deltaimpulsen.

Den Laplace transformerede findes ved at indsætte impulsen på funktionens plads og indse at integrationen kun giver et resultat ved $t = 0$ hvor impulsen fyrrer og eksponentialfunktionen er én. Det udnytter at impulsen er defineret som det vises efterfølgende (under "Sampling").

$$F(s) = \int_0^{\infty} \delta(t) \exp(-st) dt = 1$$

Det er det resultat alle kendte tabelværker angiver uden yderligere diskussion af dimensionen for impulsen $\delta(t)$. Nogle fora på internettet adresserer dette problem med den konklusion at dimensionen må være det reciprokke af variabelen. Her er det således det reciprokke af tid.

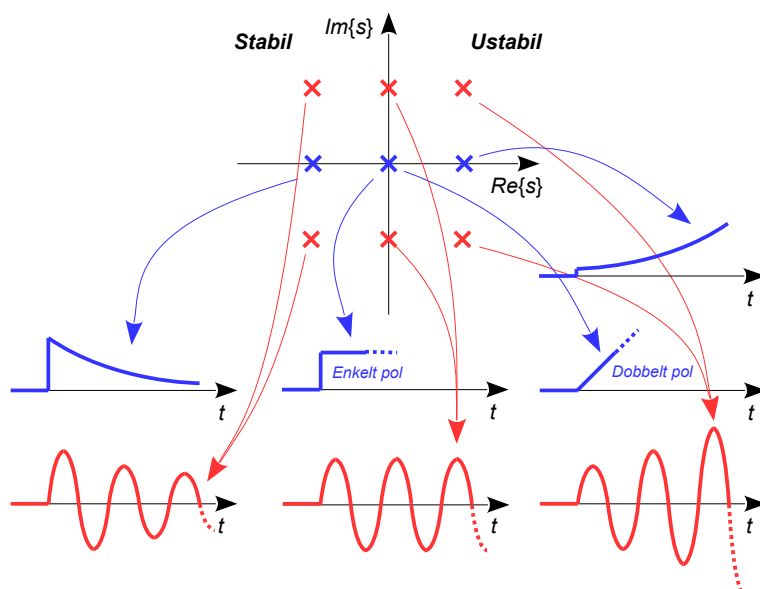
Impulsen kan lidt løst beskrives som "uendelig høj i punktet nul og nul udenfor". Dens areal bliver derfor højden uendelig gange bredden nul og det passer ikke godt med den gængse matematik så den defineres ved et integral. Impulsen er symmetrisk omkring nul og kan skaleres ved $\delta(at)$ der svarer til at dens areal ændres. Vigtigere er det at delta impulsen sammen med en funktion $f(t)$ kan udtage en bestemt funktionsværdi (*sampling*) hvilket benyttes indenfor digital signalbehandling og at et enhedsstep kan beskrives ved integralet af en impuls.⁷⁴

Definition	$\int_{-\infty}^{\infty} \delta(t) dt \stackrel{\text{def}}{=} 1$
Lige symmetri	$\delta(-t) = \delta(t)$
Skalering	$\int_{-\infty}^{\infty} a \delta(t) dt = \int_{-\infty}^{\infty} \delta(u) \frac{du}{ a } = \frac{1}{ a }, \quad a \neq 0$
	$\delta(at) = \frac{1}{ a } \delta(t)$
Sampling	$\int_{-\infty}^{\infty} f(t) \delta(t) dt = f(0)$
	$\int_{-\infty}^{\infty} f(t) \delta(t-t_1) dt = f(t_1)$
Enhedsstep	$e(t) = \int_{-\infty}^t \delta(\tau) d\tau$

⁷⁴ https://en.wikipedia.org/wiki/Dirac_delta_function.

2.19.10 Poler

Et vigtigt forhold ved Laplace transformationen er placeringen af funktionens poler i s-planen. Det er den værdi af s hvor funktionsværdien går imod uendelig og det markeres med et kryds i den komplekse s-plan. Alle udtryk har s i nævneren så polen ligger der hvor nævneren bliver nul.



Figur 139 – Funktionerne kan arrangeres i tre grupper. Poler i venstre halvplan giver stabilitet og sikrer at en påvirkning vil dø ud. Poler andre steder indikerer ustabilitet.

For stabil funktion skal polerne ligge i venstre halvplan hvor $s < 0$. Poler i venstre halvplan vil altid være resultatet med en passiv konstruktion. Med to komplekst konjugerede poler vil funktionen oscillere med aftagende amplitude efterhånden som tiden går.

Med en enkelt pol i $(0,0)$ opnås en konstant funktionsværdi som ved et enhedsstep, mens en dobbeltpol samme sted vil få funktionsværdien til at eskalere som ved en rampe. Med to komplekst konjugerede poler på den imaginære akse har vi en oscillator med konstant amplitude.

Poler i højre halvplan hvor $s > 0$ vil få funktionsværdien til at eskalere. Det er ikke en ønskværdig situation med mindre der er tale om en oscillator som er ved at starte. Når amplituden er stor nok må "noget" gribe ind og flytte polerne ind på den imaginære akse. Hvis der ikke er et styrende kredsløb vil effektforsyningen sætte grænsen for hvor stor funktionsværdien kan blive.

2.19.11 Differentiation og integration

For at løse differentiaalligninger er der brug for at kende den Laplace transformere den afledte af en funktion såvel som integralet af en funktion. Den tidsafledte af variabel $f(t)$ transformeres til $sF(s)$ og integralet af variabel $f(t)$ transformeres til $1/s$ gange med det Laplace transformerede.

Differentiation	$\frac{df(t)}{dt}$	$\Leftrightarrow_{\text{Laplace}}$	$sF(s) - f(0)$
Integration	$\int_0^t f(\tau) d\tau$	$\Leftrightarrow_{\text{Laplace}}$	$\frac{F(s)}{s}$

Med en kendt startværdi er $f(0)$ funktionsværdien ved starttidspunktet ellers sætter man værdien til nul for den kan tilføjes senere. Som et eksempel kan en spole have en vis strøm løbende inden analysen starter som det vil være tilfældet ved analyse af DC-DC konverteres funktion.

Differentiation udledes ved partiel integration der er gengivet herunder med g for eksponentialfunktionen og h som repræsentant for den tidsafledte f' idet denne skal integreres i begge trin og derved bliver til $f(t)$ som kan Laplace transformeres til $F(s)$.

$$F(s) = \int_0^{\infty} f'(t) \exp(-st) dt = \int_0^{\infty} \exp(-st) f'(t) dt \Rightarrow F(s) = [\exp(-st) f(t)]_0^{\infty} - \int_0^{\infty} [-s \exp(-st)] f(t) dt$$

$$\int g(x) h(x) dx = g(x) H(x) - \int g'(x) H(x) dx$$

$$F(s) = 0 - f(0) + sF(s)$$

Spændingen over en spole er selvinduktionen L gange med strømAendringen di/dt . Spændingen $u_L(t)$ transformeres til $U_L(s)$ og strømmen til $I(s)$. Selvinduktionen L er konstant og deltager ikke i transformationen. Sidste led med $i(0)$ er den værdi strømmen i spolen har til starttidspunktet $i(0)$.

$$u_L(t) = L \frac{di(t)}{dt} \Leftrightarrow_{\text{Laplace}} U_L(s) = sL I(s) - Li(0)$$

Integration udledes ved at indsætte integralet af funktionen i Laplace transformationen med den tidsafhængige variabel ændret til τ for at holde de variable adskilt. Derefter udnyttes det at man kan bytte om på rækkefølgen af integrationerne da de har hver sin variable og den ene af integrationerne derfor er konstant når den anden varierer. Den inderste integration ses nu at give $1/s$ der er en konstant ved integration over tid.

$$F(s) = \int_{t=0}^{\infty} \left(\int_{\tau=0}^t f(\tau) d\tau \right) \exp(-st) dt = \int_{\tau=0}^{\infty} f(\tau) \left(\int_{t=\tau}^{\infty} \exp(-st) dt \right) d\tau$$

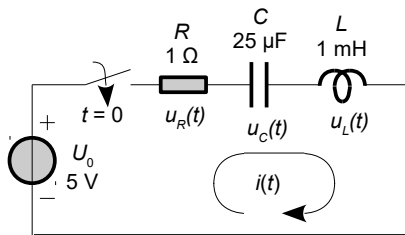
$$F(s) = \int_{\tau=0}^{\infty} f(\tau) \frac{1}{s} d\tau = \frac{1}{s} \int_{\tau=0}^{\infty} f(\tau) d\tau = \frac{1}{s} F(s)$$

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 $t = 0$.

$$u_C(t) = \frac{1}{C} \int_0^t i(\tau) d\tau \Leftrightarrow_{\text{Laplace}} U_C(s) = \frac{1}{sC} I(s)$$

2.19.12 Kredsløbsanalyse

Værktøjet skal nu demonstreres med kredsløbet herunder hvor en spændingskilde U_0 aktiveres til tidspunktet $t = 0$ og driver en strøm $i(t)$ i serieforbindelsen af en modstand R , en kondensator C og en spole L . Det er formålet at beregne strømmens funktionsudtryk og analysen starter med at opskrive Kirchhoffs maskelov. Det er det samme kredsløb som blev analyseret på side 115.



Kirchhoff (KVL)	$-U_0 + u_R(t) + u_C(t) + u_L(t) = 0$
Kredsløbsligning	$U_0 = Ri(t) + \frac{1}{C} \int_0^t i(\tau) d\tau + L \frac{di(t)}{dt}$
Laplace s-domæne	$\frac{U_0}{s} = RI(s) + \frac{1}{sC} I(s) + sLI(s)$
Isolation af $I(s)$	$I(s) = \frac{U_0}{s(R + \frac{1}{sC} + sL)}$
Reduktion	$I(s) = \frac{1}{s^2 + s\frac{R}{L} + \frac{1}{LC}} \frac{U_0}{L}$
Standardform	$I(s) = K \frac{\omega_0}{s^2 + 2d\omega_0 s + \omega_0^2} U_0$
Identifikation	$\omega_0^2 = \frac{1}{LC} \Rightarrow \omega_0 = \frac{1}{\sqrt{LC}}$
	$2d\omega_0 = \frac{R}{L} \Rightarrow d = \frac{R}{2\omega_0 L} = \frac{R}{2} \sqrt{\frac{C}{L}}$
	$K\omega_0 = \frac{1}{L} \Rightarrow K = \frac{1}{\omega_0 L} = \sqrt{\frac{C}{L}}$
Resulterende ligning	$I(s) = \sqrt{\frac{C}{L}} \frac{\omega_0}{s^2 + 2d\omega_0 s + \omega_0^2} U_0$
Klargøring for tabel	$I(s) = \frac{1}{\sqrt{1-d^2}} \sqrt{\frac{C}{L}} \frac{\sqrt{1-d^2}\omega_0}{s^2 + 2d\omega_0 s + \omega_0^2} U_0$
Tidsdomæne	$i(t) = \frac{K}{\sqrt{1-d^2}} \sin(\sqrt{1-d^2}\omega_0 t) \exp(-d\omega_0 t)$
Talværdier	$\omega_0 = 6325 \text{ rad/s} \quad d = 0,0791$ $f_0 = \frac{\omega_0}{2\pi} = 1007 \text{ Hz} \quad K = 0,1581$

Figur 140 – En seriekreds tilsluttet klokken $t = 0$ til en spændingskilde og det er ønsket at finde strømmens forløb.

Spændingskilden Laplace transformeres ved en stepfunktion til U_0/s , strømmen $i(t)$ til $I(s)$ så spændingen over modstanden bliver $RI(s)$, kondensatoren integrerer strømmen så spændingen bliver $I(s)/sC$ og spolen differentierer strømmen så spændingen bliver $sLI(s)$. Standardformen for et anden-ordens polynomium benyttes til at beregne den frie resonans ω_0 , dæmpningen d samt faktoren K for den resulterende ligning for strømmen $I(s)$.

For at transformere udtrykket tilbage til tidsdomænet ved tabellen over Laplace transformationspar skal ligningen indeholde $(1 - d^2)^{0,5}$ i tælleren og for ikke at ændre på resultatet skal der også deles med den. Oscillationens frekvens sænkes fra f_0 uden dæmpning til $f_{osc} = (1 - d^2)^{0,5} f_0 = 1003 \text{ Hz}$ på grund af dæmpningen. Faktoren til $K/(1 - d^2)^{0,5} = 0,1586 \Omega^{-1}$ giver en startamplitude på $0,7931 \text{ A}$ så ved tidspunktet $t_{max} = 237 \mu\text{s}$ bliver første maksimum på $I_{max} = 0,70 \text{ A}$ som var forventet.

2.19.13 Tabel over Laplace transformationen

A og B samt a og b er konstanter der kan være reelle eller komplekse og kan indeholde en enhed. Funktionerne $f(t)$, $g(t)$ og $h(t)$ er fysisk realisable og $e(t)$ er et enhedsstep der er én for $t \geq 0$ og nul ellers. Funktionerne $F(s)$, $G(s)$ og $H(s)$ er de tilsvarende Laplace transformerede udtryk.

	Funktion i tidsdomænet	Funktion i s-domænet	Kommentar
1	$h(t) = f(t) \cdot e(t)$	$H(s) = \int_0^{\infty} h(t) \exp(-st) dt$ $s = \sigma + j\omega, \quad \omega = 2\pi f$	Laplace transformation.
2	$h(t) = [A f(t) + B g(t)] \cdot e(t)$	$H(s) = A F(s) + B G(s)$	Linearitet = transformation af hvert led for sig.
3	$h(t) = A \frac{d[f(t)]}{dt} \cdot e(t)$	$H(s) = s A F(s) - A f(0)$	Differentiation i tid med startværdi $f(0)$.
4	$h(t) = A \int_0^t f(\tau) d\tau \cdot e(t)$	$H(s) = A \frac{f(s)}{s}$	Integration over tid.
5	$h(t) = f(t-T) \cdot e(t-T)$	$H(s) = F(s) \exp(-sT)$	Tidsforskydning.
6	$h(t) = f(t) \exp(-at) \cdot e(t)$	$H(s) = F(s+a)$	Dæmpet funktion.
7	$h(t) = A \cdot \delta(t)$	$H(s) = A$	Diracpuls til $t = 0$.
8	$h(t) = A \cdot e(t)$	$H(s) = A \frac{1}{s}$	Enhedsstep til $t = 0$.
9	$h(t) = A t \cdot e(t)$	$H(s) = A \frac{1}{s^2}$	Indkobling af rampe til $t = 0$.
10	$h(t) = A \exp(-at) \cdot e(t)$	$H(s) = A \frac{1}{s+a}$	Ekspontielt aftagende fx et RC-led med $a = 1/CR$.
11	$h(t) = A[1 - \exp(-at)] \cdot e(t)$	$H(s) = A \frac{a}{s(s+a)}$	Ekspontiel mod asymptote fx et RC-led med $a = 1/CR$.
12	$h(t) = A t \exp(-at) \cdot e(t)$	$H(s) = A \frac{1}{(s+a)^2}$	Anden-ordens funktion med dobbeltpol ($a > 0$).
13	$h(t) = A[1 - at] \exp(-at) \cdot e(t)$	$H(s) = A \frac{s}{(s+a)^2}$	Anden-ordens funktion med nul og dobbeltpol ($a > 0$).
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 funktion med to poler ($a, b > 0$).
15	$h(t) = \frac{A}{a-b} [a \exp(-at) - b \exp(-bt)] \cdot e(t)$	$H(s) = A \frac{s}{(s+a)(s+b)}$	Anden-ordens funktion med nul og to poler ($a, b > 0$).
16	$h(t) = A \left[\frac{1}{ba} + \frac{\exp(-at)}{a(a-b)} - \frac{\exp(-bt)}{b(b-a)} \right] \cdot e(t)$	$H(s) = A \frac{1}{s(s+a)(s+b)}$	tredje-ordens funktion med to poler ($a, b > 0$) og pol i nul.
17	$h(t) = A \sin(\omega_0 t) \cdot e(t)$	$H(s) = A \frac{\omega_0}{s^2 + \omega_0^2}$	Sinus.
18	$h(t) = A \cos(\omega_0 t) \cdot e(t)$	$H(s) = A \frac{s}{s^2 + \omega_0^2}$	Cosinus.

Tabellen fortsætter på næste side.

	Funktion i tidsdomænet	Funktion i s-domænet	Kommentar
19	$h(t) = A \sin(\omega_d t) \exp(-at) \cdot e(t)$	$H(s) = A \frac{\omega_d}{(s+a)^2 + \omega_d^2}$	Dæmpet sinus $a > 0$.
20	$h(t) = A \sin(\sqrt{1-d^2} \omega_0 t) \exp(-d \omega_0 t) \cdot e(t)$	$H(s) = A \frac{\sqrt{1-d^2} \omega_0}{s^2 + 2d \omega_0 s + \omega_0^2}$	Dæmpet sinus $0 < d < 1$.
21	$h(t) = A \cos(\omega_d t) \exp(-at) \cdot e(t)$	$H(s) = A \frac{s+a}{(s+a)^2 + \omega_d^2}$	Dæmpet cosinus $a > 0$.
22	$h(t) = A \cos(\sqrt{1-d^2} \omega_0 t) \exp(-d \omega_0 t) \cdot e(t)$	$H(s) = A \frac{s + \sqrt{1-d^2} \omega_0}{s^2 + 2d \omega_0 s + \omega_0^2}$	Dæmpet cosinus $0 < d < 1$.

Se også referencerne RW og NA samt internettet.

De to udtryk med kvadratrødder kan virke overraskende så det første udtryk med sinus udledes herunder. Resultatet er at den dæmpede sinus svinger på en frekvens der er lidt lavere end den frie resonans for systemet uden dæmpning ($a = 0$). Korrektionen er funktion af dæmpningen a der også kan udtrykkes ved dæmpningsfaktoren d og den frie resonans ω_0 som $a = d\omega_0$. For en værdi af dæmpningsfaktoren på $d = 0,1$ bliver korrektionen af svingningens frekvens på $-0,5\%$.

Transformationspar
for dæmpet sinus
(version 1)

$$h(t) = A \sin(\omega_d t) \exp(-at) \cdot e(t)$$

$$H(s) = A \frac{\omega_d}{(s+a)^2 + \omega_d^2} = A \frac{\omega_d}{s^2 + 2as + a^2 + \omega_d^2}$$

Reference for $s = 0$

$$H(0) = A \frac{\omega_d}{a^2 + \omega_d^2} \Rightarrow A = \frac{a^2 + \omega_d^2}{\omega_d} H(0)$$

Standardform

$$H(s) = K \frac{\omega_0}{s^2 + 2d\omega_0 s + \omega_0^2}$$

Reference for $s = 0$

$$H(0) = \frac{K}{\omega_0} \Rightarrow A = \frac{a^2 + \omega_d^2}{\omega_d \omega_0} K$$

$$\omega_0^2 = a^2 + \omega_d^2 \Rightarrow \omega_0 = \sqrt{a^2 + \omega_d^2}$$

Sammenligning

$$2d\omega_0 = 2a \Rightarrow \omega_0 = \frac{a}{d} \Rightarrow a = d\omega_0$$

Konstant

$$A = \frac{\omega_0}{\omega_d} K \Rightarrow K = \frac{\omega_d}{\omega_0} A$$

Dæmpning

$$d = \frac{a}{\omega_0} = \frac{a}{\sqrt{a^2 + \omega_d^2}} \Rightarrow d^2 = \frac{a^2}{a^2 + \omega_d^2}$$

$$d^2(a^2 + \omega_d^2) = a^2 \Rightarrow d^2 \omega_d^2 = (1 - d^2) a^2$$

Dæmpet svingning

$$\omega_d^2 = \frac{(1 - d^2) a^2}{d^2} \Rightarrow \omega_d = \frac{a}{d} \sqrt{1 - d^2} = \sqrt{1 - d^2} \omega_0$$

Konstant

$$K = \sqrt{1 - d^2} A$$

Transformationspar
for dæmpet sinus
(version 2)

$$h(t) = A \sin(\sqrt{1 - d^2} \omega_0 t) \exp(-d \omega_0 t) \cdot e(t)$$

$$H(s) = A \frac{\sqrt{1 - d^2} \omega_0}{s^2 + 2d \omega_0 s + \omega_0^2}$$

Med udgangspunkt i udledningen for sinus kan transformationsparret for cosinus valideres alene ved inspektion.

2.19.14 Start- og slutværdi

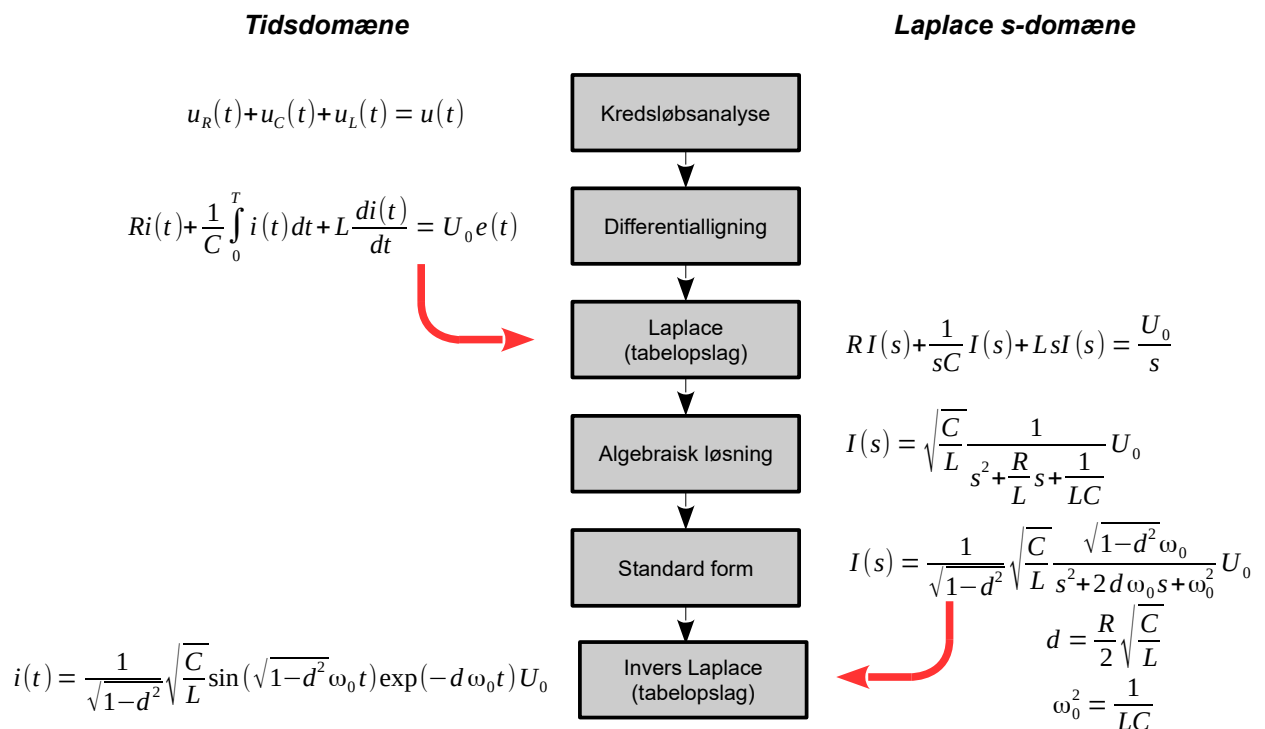
Ved startværdisætningen (*initial value theorem*) findes tidsfunktionens værdi i startøjeblikket ved at gange s på $F(s)$ og lade s gå imod uendelig. Ved slutværdisætningen (*final value theorem*) findes tidsfunktionens værdi når der er gået så lang tid at kredsløbet er faldet til ro ved at gange s på $F(s)$ og benytte $s = 0$. Forudsætningen er at polerne i $F(s)$ ligger i venstre halvplan så reglen gælder ikke for oscillerende funktioner.

Startværdisætningen:	$\lim_{t \rightarrow 0} \{f(t)\} = \lim_{s \rightarrow \infty} \{sF(s)\}$
Slutværdisætningen:	$\lim_{t \rightarrow \infty} \{f(t)\} = \lim_{s \rightarrow 0} \{sF(s)\}$

Sætningerne vil ikke blive udledt, men den interesserede kan finde beviserne på Wikipedia.^{75,76}

2.19.15 Transformation tilbage til tidsdomænet

Laplace-transformationen er velegnet til løsning af selv ret komplicerede kredsløb. Det er ofte et ønske at finde svaret i tidsdomænet og det involverer en invers Laplace transformation som de færreste har lyst til at give sig i kast med. I stedet benyttes en tabel. Kredsløbets differentialligning transformeres fra tidsdomænet til Laplace s -domænet hvor det er enkelt at løse ligningen ved at isolere den variable. Gennem et tabelopslag findes det tilhørende transformationspar og så er opgaven gennemført uden at involvere en invers Laplace transformation.



Figur 141 – Laplace-transformationen flytter løsningen af differentialligninger til frekvensdomænet hvor differentiation er erstattet af multiplikation med s og tilsvarende er integration erstattet af $1/s$. Ved brug af tabelopslag transformeres udtrykket tilbage til tidsdomænet.

De ledsagende formler svarer til eksemplerne side 115 og 138.

⁷⁵ https://en.wikipedia.org/wiki/Initial_value_theorem.

⁷⁶ https://en.wikipedia.org/wiki/Final_value_theorem.

2.19.16 Partialbrøk

Overføringsfunktionen i Laplace s-domænet giver ofte en brøk mellem to polynomier og det er sjældent tilgængeligt i tabelværkerne. Der findes en invers Laplace transformation, men selv de mere avancerede tekstbøger undgår den på grund af en meget vanskelig integration. Der er derfor behov for at opdele en brøk i mindre bidder som kan transformeres tilbage til tidsdomænet. Vi har følgelig behov for at bringe udtrykket på en enklere form hvor man kan slå op i en tabel for at finde det tilhørende tidssvar.

Først skal nulpunkterne i nævneren findes og det er gennemgået på side 80. De kaldes her for x_1 og x_2 .

Brøken opdeles i to led med konstanterne A og B i tælleren og hver sit nulpunkt i nævneren.

De to konstanter findes ved at gange med nævnerens faktorer på begge sider. Ved at sætte $x = x_1$ kan ledet med B elimineres og A kan bestemmes. På tilsvarende vis findes B .

$$\begin{aligned}\frac{1}{ax^2+bx+c} &= \frac{1}{a(x-x_1)(x-x_2)} \\ \frac{1}{a(x-x_1)(x-x_2)} &= \frac{A}{x-x_1} + \frac{B}{x-x_2} \\ \frac{1}{a} &= A(x-x_2) + B(x-x_1) \\ x = x_1 &\Rightarrow A = \frac{1}{a(x_1-x_2)} \\ x = x_2 &\Rightarrow B = \frac{1}{a(x_2-x_1)}\end{aligned}$$

Som eksempel benyttes parablens ligning fra før hvor nulpunkterne er kendte. Ved at sætte det resulterende led på fælles brøkstreg ses at det faktisk er meget lig med det oprindelige udtryk. Det viste eksempel ender med en tæller der er tæt på $\frac{1}{2}$ som ønsket.

Ligning og rodbestemmelse:

$$\begin{aligned}\frac{1}{2x^2+2x-0,5} &= \frac{1}{2(x-x_1)(x-x_2)} \\ x_1 &= 0,207 \\ x_2 &= -1,207\end{aligned}$$

Dekomposition:

$$\begin{aligned}\frac{1}{2x^2+2x-0,5} &= \frac{1}{2(x-0,207)(x+1,207)} = \frac{A}{x-0,207} + \frac{B}{x+1,207} \\ A &= \frac{1}{2(0,207-(-1,207))} = 0,354 \\ B &= \frac{1}{2(-1,207-0,207)} = -0,354\end{aligned}$$

Kontrol:

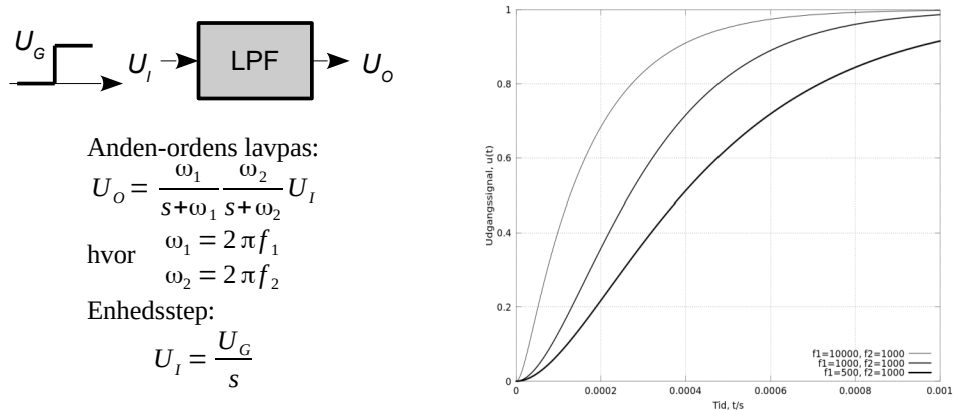
$$\frac{0,354}{x-0,207} - \frac{0,354}{x+1,207} = \frac{0,354(x+1,207) - 0,354(x-0,207)}{(x-0,207)(x+1,207)} = \frac{0,501}{(x-0,207)(x+1,207)}$$

Der findes nogle generelle regneregler som listes herunder. For alle udtryk er det nævnerens polynomium der bestemmer hvilken metode der skal anvendes. Det kan dog kræves en del kreativitet for at finde frem til konstanterne [RW-120].

$\frac{1}{(x-x_1)(x-x_2)} = \frac{A}{x-x_1} + \frac{B}{x-x_2}$ $\frac{1}{(x-x_1)^2} = \frac{A}{x-x_1} + \frac{B}{(x-x_1)^2}$ $\frac{1}{(x-x_1)(x-x_2)^2} = \frac{A}{x-x_1} + \frac{B}{x-x_2} + \frac{C}{(x-x_2)^2}$

2.19.17 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 142 – 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 begge er én.

$$u_O(t) = \left[1 + \frac{\omega_2}{\omega_1 - \omega_2} \exp(-\omega_1 t) + \frac{\omega_1}{\omega_2 - \omega_1} \exp(-\omega_2 t) \right] \cdot U_G e(t)$$

At tidssvaret fra filtret starter fra nul kan ses ved at bruge startværdisætningen fra side 137.

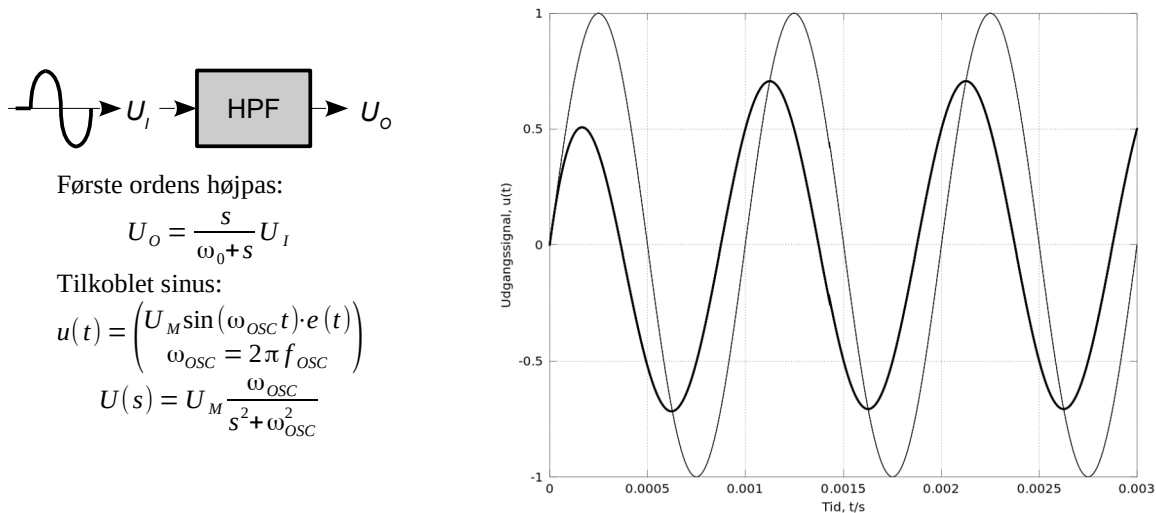
$$u_O(0) = \lim_{s \rightarrow \infty} \left\{ s \frac{\omega_1}{s + \omega_1} \frac{\omega_2}{s + \omega_2} \frac{U_G}{s} \right\} = \lim_{s \rightarrow \infty} \left\{ \frac{\omega_1}{s + \omega_1} \frac{\omega_2}{s + \omega_2} U_G \right\} = 0$$

Tilsvarende kan det ses fra slutværditeoremet at funktionen efter tilstrækkelig lang tid vil være nået op på slutværdien U_G .

$$u_O(0) = \lim_{s \rightarrow 0} \left\{ s \frac{\omega_1}{s + \omega_1} \frac{\omega_2}{s + \omega_2} \frac{U_G}{s} \right\} = \lim_{s \rightarrow 0} \left\{ \frac{\omega_1}{s + \omega_1} \frac{\omega_2}{s + \omega_2} U_G \right\} = U_G$$

2.19.18 Indsvingning

En anden vigtig situation er en pludselig tilslutning af et oscillerende 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.



Figur 143 – 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 leddet og et polynomium med to konstanter $Bs + C$ for andet-ordens leddet. Opgaven er nu at finde de tre konstanter A , B og C .

$$U(s) = U_M \frac{s \omega_{OSC}}{(s + \omega_0)(s^2 + \omega_{OSC}^2)} = U_M \left[\frac{A}{s + \omega_0} + \frac{Bs + C}{s^2 + \omega_{OSC}^2} \right] \quad \text{og} \quad A, B, C = \text{konstanter}$$

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.

$$s \omega_{OSC} = (A + B)s^2 + (B\omega_0 + C)s + A\omega_{OSC}^2 + C\omega_0 \Rightarrow \begin{pmatrix} A + B = 0 \\ B\omega_0 + C = \omega_{OSC} \\ A\omega_{OSC}^2 + C\omega_0 = 0 \end{pmatrix}$$

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 .

$$\begin{pmatrix} A = -B \\ C = -B\omega_0 + \omega_{OSC} \\ C\omega_0 = B\omega_{OSC}^2 \end{pmatrix} \Rightarrow B = \frac{\omega_{OSC}\omega_0}{\omega_{OSC}^2 + \omega_0^2} \Rightarrow C = \frac{\omega_{OSC}^3}{\omega_{OSC}^2 + \omega_0^2}$$

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 koefficienten 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 $0,71 \cdot \sin(\omega_{OSC} t + 45^\circ)$.

2.19.19 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 variabelen 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 harmoniskes svingninger med de komplekse eksponentialfunktioner. 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.

$$M(f) = |H(f)| = \left| \frac{\omega_0}{j2\pi f + \omega_0} \right| = \left| \frac{f_0}{jf + f_0} \right| = \frac{f_0}{\sqrt{f^2 + f_0^2}} = \frac{1}{\sqrt{1 + (f/f_0)^2}}$$

2.20 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	Spænding	Strøm	Kapacitet	Modstand	Selvinduktion
Impedans: $V/A = \Omega$	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	Kraft	Hastighed	Fjeder	Modstand	Masse
Impedans: $N/ms^{-1} = kg/s$	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	Tryk	Vol.hast.	Fjeder	Modstand	Masse
Impedans: $Nm^{-2}/m^3s^{-1} = kg/m^4s$	$Pa = N/m^2$	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	Temperatur	Effekt	Kapacitet	Modstand	(Ingen)
Impedans: K/W	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 selvinduktion og $1/sC$ for kondensatoren.

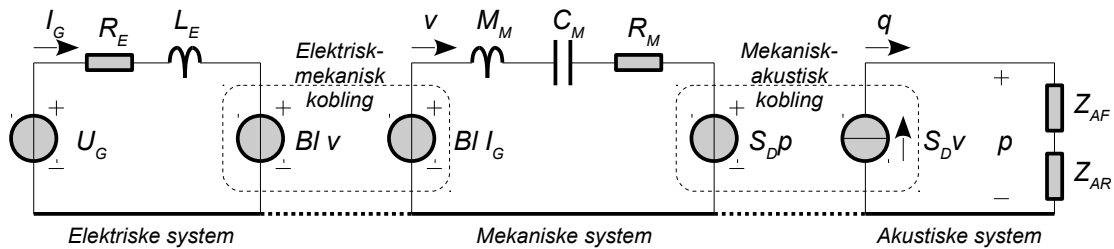
Det mekaniske system benytter kraft F og hastighed v for massen M , modstanden R og fjedren F . De tre komponenter beskrives ved de samme ligninger som for det elektriske systems komponenter så en masse kan repræsenteres ved en spole og fjedrens eftergivelse C (compliance) kan repræsenteres ved en kondensator. Newtons anden lov $F = Mdv/dt$ giver relationen mellem kraft og masse og positionen x kan 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 så relationen er ikke eksakt.

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 akustisk 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 højttalerens 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 inertie findes, men den mødes kun i ekstremt sjældne tilfælde så som 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.

2.20.1 Elektro-dynamisk højttaler

I første kapitel blev den elektro-dynamiske højttaler introduceret og det blev vist at membranens hastighed er 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.⁷⁷



Figur 144 – 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 eftergivelse 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.

2.20.2 Elektriske 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 8 Ω højttaler har $R_E \approx 6 \Omega$. Selvinduktionen af svingspolen er $L_E \approx 1$ mH for en bashøjttaler og 100 μ 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.

$$-U_G + R_E I_G + sL_E I_G + Bl v = 0 \Rightarrow I_G = \frac{U_G - Bl v}{R_E + sL_E} \quad \text{og} \quad f_s = \frac{R_E}{2\pi L_E}$$

For at lette bestemmelsen af membranens hastighed v ignoreres svingspolens selvinduktion som med eksempelvis $R_E = 6 \Omega$ og $L_E = 1$ mH giver øvre grænse for modellen på $f_s < 1$ kHz.

2.20.3 Mekaniske 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 eftergivelsen C_M for en højttaler i et kabinet.

⁷⁷ Modellen opbygges i min bog Elektroakustik der kan downloades i PDF fra min hjemmeside www.torean.dk under Publikationer.

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 leddet $s^2 M_M C_M$ 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 $(Bl)^2/R_E$ der derfor er bestemmende hvordan højttaleren opfører sig omkring f_0 , og det er detaljeret i første kapitel.

$$\text{Generelt} \quad v = \frac{sC_M}{s^2 M_M C_M + sC_M [R_M + (Bl)^2/R_E] + 1} Bl \frac{U_G}{R_E} \quad \text{hvor} \quad L_S = 0 \quad \text{og} \quad p = 0$$

$$f_0 < f < f_s \quad v \approx \frac{Bl}{s M_M} \frac{U_G}{R_E} \quad f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi \sqrt{M_M C_M}}$$

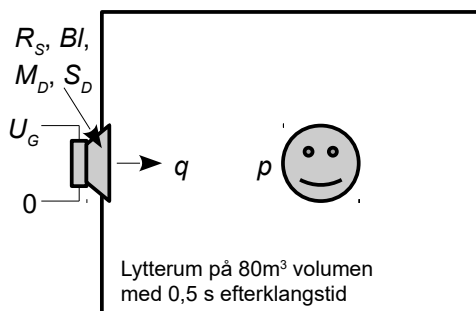
Med højttalerens værdier $M_M = 0,02$ kg og $C_M = 1,4$ mm/N findes resonansen til $f_0 = 30$ Hz.

2.20.4 Akustiske 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 145 – 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 $v_{RMS} = 0,93$ m/s ved $f = 40$ Hz for en højttaler med radius 100 mm ($S_D = 31,4 \cdot 10^{-3}$ m²), og amplituden af membranens bevægelse bliver på $x_{MAX} = 5$ 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 lydtryk. 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 ± 10 mm. Selv om man kunne så ville lydengivelsen lide under forvrængning på grund af Dopplereffekten.⁷⁸

⁷⁸ 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 $P_A = 15$ mW når rummet er på 80 m³ 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.

2.21 Opgaver

Impedans

1 – Benyt formlerne på side 91 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 seriemodstand på $R = 10 \Omega$. Beregn godheden Q og båndbredden Δf . **Svar:** $f_0 = 5 \text{ kHz}$, $Q = 3,2$ og $\Delta f = 1,6 \text{ kHz}$.

Komplekse tal

2 – Tegn et koordinatsystem for den komplekse talplan som vist på side 92. 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 side 92 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:** $z_1/z_2 = -j$

Differentiation

9 – Brug eksemplet på side 97 til at vise at $f(x) = 4x^2$ giver $f'(x) = 8x$ ved differentiation og at konstanten 4 derfor ikke deltager i differentiationen.

10 – Brug eksemplet på side 97 til at vise at $f(x) = x^3$ giver $f'(x) = 3x^2$ ved differentiation.

11 – Brug eksemplet på side 97 til at vise at $f(x) = \cos(x)$ giver $f'(x) = -\sin(x)$ ved differentiation.

12 – Vis at differentialkvotienten af $f(x) = \exp(\omega t)$ bliver til $f'(x) = \omega \exp(\omega t)$ hvor ω er en konstant.

13 – Brug eksemplet side 100 for halvbølge sinus til at vise at integralet over cosinus fra 0 til π er nul.

14 – Brug tabellen side 101 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:** $c = 0$.

Integration

15 – Brug tabellen side 101 til at vise at $f(t) = \exp(\omega t)$ hvor

ω er en konstant integreres til $F(t) = \exp(\omega t)/\omega$.

16 – 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 101 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. Hvad er strømstyrken vokset til ved $t_{ON} = 2t_{ON}$? **Svar:** $i_L = 100 \text{ mA}$.

$$u_L = L \frac{di_L}{dt} \Rightarrow i_L = \frac{U}{L} t_{ON} + I_0 = 50 \text{ mA}$$

17 – Fakultetsregning introduceres side 103 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 $5! = 120$.

Taylor serier

18 – Taylor serierne beregnes ved fortsat differentiation af et polynomium. Beregn alle afledte af $f(x) = a + bx + cx^2$. **Svar:** $f'(x) = b + 2cx$, $f''(x) = 2c$, $f^{(3)}(x) = 0$ og alle højere forbliver på nul: $f^{(n)}(x) = 0$ for $n > 3$.

19 – Brug tabellen side 104 til at vise at $[1 - \cos(x)]/x$ bliver nul for $x = 0$.

20 – En sammenkobling af spolen L og kondensatoren C danner resonans ved frekvensen f_1 som er givet herunder. Brug tabellen side 104 til at vise at en beskedent 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 formelen 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 106 til at udlede udtrykket for sinus.

22 – Brug Eulers identitet side 106 til at eftervises den følgende formel for produktet af to sinus-funktioner.

$$\sin(x)\sin(y) = \frac{\cos(x+y) - \cos(x-y)}{2}$$

Differentialligning

23 – Vis at en homogen anden-ordens differentialligning har eksponentialfunktionen som løsning. Som prototype foreslås $x(t) = a \exp(\omega t)$ der differentieres to gange og indsættes. Vis at det kan reduceres til karakterligningen.

Fourier serie

24 – Koefficienterne for en Fourier serie skal beregnes efter formlerne side 118 for en firkantet puls der er én fra nul til τ og nul for resten af perioden på T . Amplituden af harmoniske a_n beregnes som vist og opgaven er at vise udtrykkene for koefficienterne a_n og b_n samt at beregne amplituden c_n af de første seks harmoniske for $\tau/T = 1/3$.

Analyse

$$f(t) = \begin{cases} 1 & 0 \leq t \leq \tau \\ 0 & \tau < t < T \end{cases} \quad \text{periode: } T$$

$$a_n = \frac{\sin(2\pi n\tau/T)}{\pi n} \quad c_n = \sqrt{a_n^2 + b_n^2} = \left| \frac{2}{\pi n} \sin\left(\frac{\pi n\tau}{T}\right) \right|$$

$$b_n = \frac{1 - \cos(2\pi n\tau/T)}{\pi n}$$

Svar: $a_0 = 0,333$, $c_1 = 0,551$, $c_2 = 0,276$, $c_3 = 0$, $c_4 = 0,138$, $c_5 = 0,110$, $c_6 = 0$.

25 – Vis at modulus af $H(f)$ side 123 giver sinc-funktionen.

Du får brug for definitionen af modulus side 92 og et par regneregler for trigonometriske funktioner.

Laplace transformation

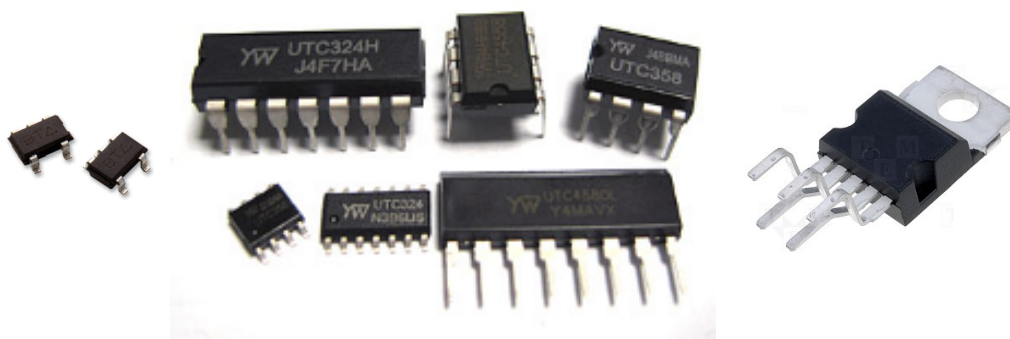
26 – Find udtrykket for strømmen $I(t)$ hvis modstanden side 131 bliver erstattet af en ideel spole med impedansen sL . **Svar:** $I(s) = U_0/s^2L$.

27 – Vis at tidsvinduet side 133 giver de viste formler ved indsætning i definitionen af Laplace transformationen.

3 OPERATIONSFORSTÆRKER

Dette kapitel omhandler en universelt anvendelig komponent der kan programmeres til at realisere en matematisk operation – en operationsforstærker eller OpAmp (*operational amplifier*). Den kan forstærke spændinger i μV området fra en termokobler eller strømme i nA området fra lysfølsomme dioder, den kan tilpasse en transducers signal ved matematiske operationer som differentiation og integration, den kan undertrykke brum og støj ved signaltransmission, den kan regulere tonebalancen i et lydanlæg eller varetage ikke-lineær signalbehandling som ensretning, logaritmering og begrænsning. Der findes varianter for brug i en radio, for videosignaler i et distributionsanlæg eller for at drive en DC motor eller en højttaler.

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. Der er nogle krav som den anvendte operationsforstærker skal opfylde for at kunne løfte en given opgave, men når kravene først er defineret bør enhver kompatibel forstærker kunne fungere i opstillingen. Det er denne fleksibilitet der er årsagen til operationsforstærkerens store succes.



Figur 00146 – Yderst til venstre SOT23-5 pakker (Small Outline Transistor) der måler et par millimeter på hver led. Øverst i midten vises DIL pakker (Dual In-Line) med 2,54 mm benafstand (1/10 tomme), derunder SOIC (Small Outline Integrated Circuit) med 1,27 mm afstand og en SIL (Single In-Line) der blandt andet bruges til små effektforstærkere. Til højre vises en TO220 variant med 5 ben der rummer en 30 W effektforstærker.

Søger man i producenternes hjemmesider findes et helt overvældende udbud af valgmuligheder der dels viser hvor voldsom udbredelsen er blevet siden operationsforstærkerens introduktion kort efter 1960 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 overflødhedshorn af valgmuligheder, men for den nytilkomne er den barokke overflod nok mere skræmmende end tillokkende. For at komme ind i det fine selskab skal du jo starte et eller andet sted og mit råd er at gribe 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.⁷⁹

Der er mange opfattelser af hvorvidt en operationsforstærker er egnet til en given opgave eller om det var bedre med en diskret tilgang. Indenfor lyd gengivelse var det tidligere holdningen at undgå komponenten. Det skyldes dels dårlige erfaringer med de tidlige operationsforstærkere der ikke alle var velegnede til krævende analog signalbehandling og dels konservatisme efter devisen at "selvgjort er velgjort". Operationsforstærkeren fremstilles i dag for krævende signalbehandling, for brug ved frekvenser over 100 MHz og for opbygning af effektforstærkere hvor en strømstyrke indtil $\pm 5\text{ A}$ er mulig i de relativt små TO220 pakker. Spændinger over $\pm 25\text{ V}$ er dog en udfordring.

⁷⁹ Analog Devices, Linear Technology, Maxim Semiconductors, Motorola, National Instruments, On-Semi, Philips (NXP).

3.1 Lidt historie

Design af elektroniske kredsløb er siden 1900-tallets start drejet fra at overveje opbygningen i alle detaljer til at vælge en integreret komponent fra producentens datablad. Det er radikal nytænkning og paradigmeskiftet sker omkring 1970. Mange udviklere fortsætter dog med at designe sine egne kredsløb til den aktuelle opgave for det er 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. Lidt efter lidt accepteres databladene dog og overskuddet bruges på at udvikling af mere komplekse opstillinger eller på en bedre kvalitet af produktet.

Radiorøret opfindes 1902 og er 20 år senere kommercielt interessant med verdensomspændende radiotransmissioner og opbygning af et marked hos den voksende mellemklasse. Julius Edgar Lilienfeld patenterer 1925 en halvlederkomponent med samme funktion som radiorøret, dog uden en sårbar glaskolbe og kortlivet glødetråd, men den slår ikke an grund af kommercielle interesser i radiorøret. Alan Blumlein patenterer 1936 en balanceret kobling med to radiorør der er forløberne for den moderne operationsforstærker og udnyttes i analoge computere. Den bipolære transistor patenteres 1947 af John Bardeen og Walter Brattain ved Bell lab og muliggør fra 1960 integrerede kredsløb med digitale funktioner og 1964 introducerer Fairchild operationsforstærkeren uA702. Lilienfelds transistor dukker op igen som JFET og siden MOSFET og danner forudsætningen for mikroprocessoren med Intel 4004 fra 1971.^{80,81,82,83}

Batteriforsyning er interessant og LM324 fra 1972 kan arbejde fra 5 V ved 1,5 mA i forbrug. Dens indgange kan arbejde ved spændinger under nul og udgangen tæt på nul hvilket foregriber den moderne forstærker med rail-to-rail som ADA4051 der arbejder fra 1,8 V med 13 μ A i forbrug. Der blev i stigende grad brugt felteffekt transistorer som TL071 fra 1978 der har JFET i indgangen og en indgangsstrøm i picoampere området. Den blev populær på grund af et fornuftigt kompromis mellem kvalitet og pris og senere kom TLC271 der udelukkende består af MOSFET. Højtydende operationsforstærkere som NE5534 fra 1980 og AD797 fra 1992 fandt vej til den hurtigt voksende og meget krævende lydverden på grund af lav støj og forvrængning.

De beskrevne komponenter har til fælles at de kun kan levere cirka ± 20 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 levere op til ± 3 A og er gode til strømforsyninger og elementær analog signalbehandling. Det er staks vanskeligere hvis det drejer sig om lyd gengivelse af høj kvalitet for det er notorisk svært at bygge en integreret effektforstærker. Nogle muligheder er LM3886, TDA2040 og TDA7294 der benyttes i musikanlæg, HDTV og højttalere med indbygget effektforstærker. De nævnte komponenter er alle opbygget som operationsforstærkere blot for et højere spændings- og strømniveau.

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 kommercielt tilgængelige operationsforstærkere eller motorstyring med integrerede effektforstærkere eller lyd gengivelse af høj kvalitet med diskret opbyggede effektforstærkere så er de alle operationsforstærkere. Teorien er helt den samme så de følgende betragtninger i denne bog dækker alle applikationer.

80 Humphry Davy opdager i 1802 at en opvarmet, elektrisk leder udsender lys og da Warren de la Rue 1840 indbygger lederen i en lufttom glaskolbe opfindes glødelampen og Edison når omkring 1880 frem til et fungerende produkt. Elektronemissionen opdages i 1873 af Frederick Guthrie og fører til elektronrøret med to elektroder der kaldes for en diode. John Ambrose Fleming bruger dioden ved radiomodtagelse da den kan ensrette et signal og 1902 indbygger han et gitter i dioden hvorved strømmen kan styres og trioden opfindes. Det leder til en eskalerende udvikling af radiokommunikationen, TV transmission, videnskab og hobby.

81 https://en.wikipedia.org/wiki/Julius_Edgar_Lilienfeld.

82 https://en.wikipedia.org/wiki/Analog_computer.

83 <https://www.calvin.edu/~pribeiro/courses/engr332/Handouts/ho18opamp.pdf>.

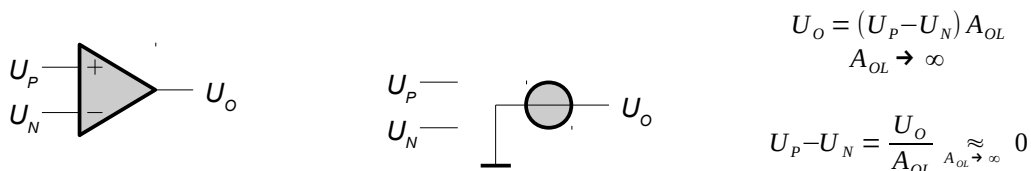
3.2 En OpAmp

En OpAmp har to indgange og én udgang. Indgangene monitorerer potentialet på hver sit sted i et kredsløb og udgangen driver et tredje sted i kredsløbet som et resultat af hvad der er blevet målt.

En OpAmp placerer udgangen så de to indgange har samme potentiale.

Når de to indgange ligger på samme potentiale siges operationsforstærkeren at være i balance og den arbejder i sit lineære område. Hvis de to indgange ikke er ens er forstærkeren ikke i balance og den vil ændre spændingen på udgangen for at genoprette 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 den positive eller negative effektforsyning eller hvad der nu sætter grænsen for udgangsspændingen. Forstærkeren er da overstyret eller mættet (*saturated*) og relationen gælder ikke.

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 147 – Diagramsymbolet for en OpAmp vises til venstre og den analytiske 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.

Elektronikken var i starten af 1900-tallet baseret på radorør 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 kostbar komponent ved sin introduktion. Integrationen af mange transistorer på en enkelt chip fjernede gradvist den økonomiske hindring og ledte til en ny måde at designe forstærkere på som udnyttede at det var nærmest gratis at øge antallet af transistorer.

- Fra side 154 og frem vises kredsløb med operationsforstærkere for forstærkning, filtrering og tonegeneration. Fra side 165 skiftes fokus mod begrænsninger ved DC og side 170 bliver signalerne oscillerende. Fra side 181 analyseres operationsforstærkerens egenstøj.
- 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, men den er vigtig og emnet uddybes side Error: Reference source not found.

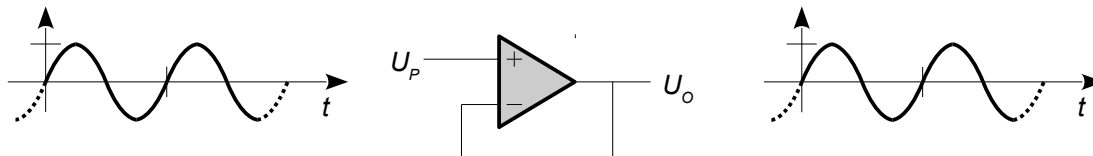
Indtil videre er en OpAmp et niveau i den abstraktion man som designer må vænne sig til for at drage fordel af dens elegante konstruktion. Der er selvfølgelig en lang række begrænsninger med de mere væsentlige nævnt ovenfor, men man plejer at antage at operationsforstærkeren opfører sig som den ideelle operationsforstærker ville gøre det og i lang udstrækning er det tilstrækkeligt for at lande helskinnet med et design.

3.3 Eksempler

Her følger kredsløbseksempler med operationsforstærkere ledsaget af en kort forklaring. For designeksempler henvises til fjerde kapitel.

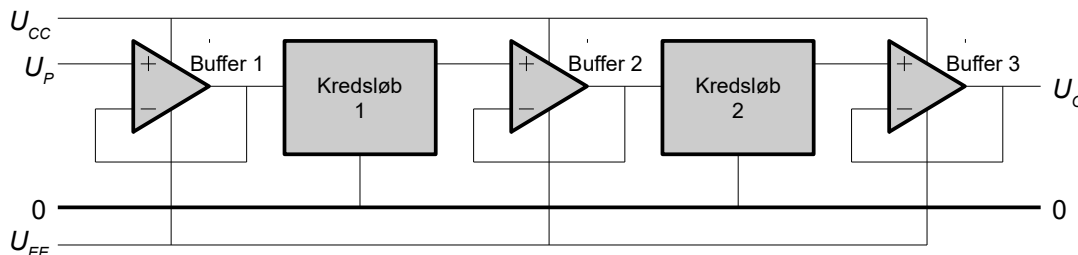
3.3.1 Buffer

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 potentiale ($U_P = U_N$) vil udgangssignalet følge med indgangssignalet. Det kaldes for en spændingsfølger (*voltage follower*) eller en buffer.



Figur 148 – 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 ved interface til sensorer der ikke tåler hård belastning. Almindelig elektroteknik forudsætter at et kredsløb arbejder under disse forhold så man står sig ved at inkludere det nødvendige antal buffere.



Figur 149 – En spændingsfølger driver indgangen til et kredsløb fra en lav modstandsværdi og sikrer imod en hård belastning af kredsløbets udgang.

Både ind- og udgangssignalerne skal holdes indenfor de grænser der sættes af effektforsyningen og belastningen må ikke være så hård at udgangen ikke kan levere strømmen. Desuden kræver de fleste operationsforstærkere et par volt mellem udgangen og effektforsyningen for at fungere korrekt.

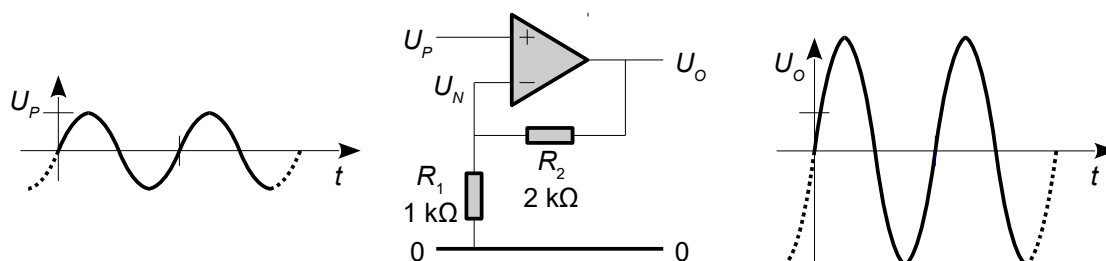
De færreste operationsforstærkere kan arbejde med indgangsspændinger udenfor grænserne sat af effektforsyningen. Undtagelserne er de gamle BIFET serier TL06x, TL07x, TL08x og LF35x der kan arbejde med spændinger over den positive effektforsyning og LM324 der kan arbejde med spændinger ned til $-0,3$ V under den negative effektforsyning (typisk GND). Moderne komponenter for enkelt effektforsyning (5 V og GND) kan som regel klare indgangsspændinger ned til $-0,3$ V.

Ved effektforsyning fra $U_{CC} = 15$ V og $U_{EE} = -15$ V vil udgangssignalet højst kunne være ± 13 V for et typisk operationsforstærker og det er ofte tilstrækkeligt.

Ved $U_{CC} = 5$ V og $U_{EE} = 0$ skal der overvejes en *rail-to-rail* operationsforstærker for ellers blive signalet let begrænset og vil dermed blive forvrænget (se side 199).

3.3.2 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 der kan forstærke det indkommende signal. Udgangssignalet har samme polaritet som indgangens signal hvorfor den kaldes for en ikke-inverterende forstærker (*non-inverting amplifier*).



Figur 150 – En spændingsforstærker kan adskille to netværk og dertil forstærke signalet. Med de viste komponentværdier indstilles forstærkningen til $A_P = 3$. Med 2 V på indgangen vil udgangen stilles på 6 V. Hvis indgangssignalet er en sinus med ± 1 V vil udgangen svinge med ± 3 V.

Udgangsspændingen U_O er nu givet af indgangsspændingen U_P og spændingsdeleren ved R_1 og R_2 til følgende ligning der nok er den mest benyttede relation overhovedet for en forstærker.

$$U_O = \left(1 + \frac{R_2}{R_1}\right) U_P \Rightarrow A_P = \frac{U_O}{U_P} = 1 + \frac{R_2}{R_1}$$

Ved en maksimal udgangsspænding på ± 12 V og en forstærkning på 3 gange begrænses det maksimale indgangssignal til ± 4 V for at undgå overstyring.

Bemærk at forstærkerens offset også forstærkes og at det vil flytte udgangen væk fra den ønskede position. Med en typisk specifikation på 2 mV for prisbillige operationsforstærkere vil det give en fejl på udgangen som er ± 6 mV med de viste komponenter. Det emne gennemgås side 166.

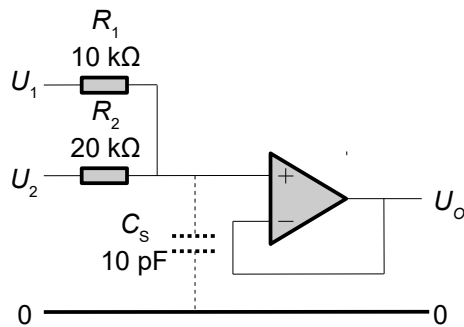
Det er en brøkdel af spændingen på udgangen der føres tilbage til den negative indgang og denne brøkdel er givet ved spændingsdelingen over R_1 og R_2 .

$$U_N = U_P$$

$$U_N = \frac{R_1}{R_1 + R_2} U_O \Rightarrow U_O = \frac{R_1 + R_2}{R_1} U_P = \left(1 + \frac{R_2}{R_1}\right) U_P$$

3.3.3 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. Det er en ofte set konstruktion hvor den væsentligste anke er at kildeimpedansen fra hver af kilderne U_1 og U_2 skal være lav i forhold til modstandsværdien af R_1 og R_2 for at opnå stor præcision (se side 160).



Resulterende udtryk:

$$U_o = \frac{R_2}{R_1 + R_2} U_1 + \frac{R_1}{R_1 + R_2} U_2$$

Grænsefrekvens:

$$f_H = \frac{1}{2\pi(R_1 \parallel R_2)C_s} \quad (f_H < GBP)$$

Ind- og udgangsmodstand:

$$R_{I1} \approx R_1 + R_2 \quad R_O \approx 0$$

$$R_{I2} \approx R_1 + R_2$$

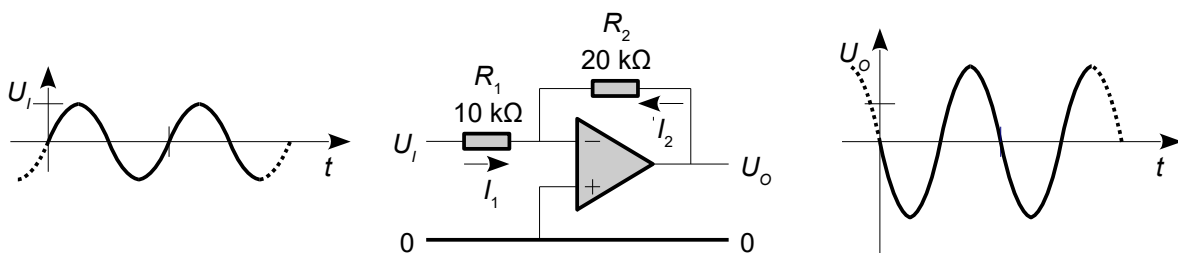
Figur 151 – To indgangssignaler kan adderes gennem modstande og en operationsforstærker kan bruges som buffer af signalet til udgangen.

Grænsefrekvensen bliver 2,4 MHz med de viste komponentværdier og spredningskapaciteten $C_s = 10$ pF.

Der er en uundgåelig spredningskapacitet på grund af at alle komponenter kobler 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 begrænses af valget af modstandsværdier.

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. Invertering af signalet kan være en vigtig egenskab i sig selv, men i modsætning til den ikke-inverterende forstærker kan forstærkningen 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 152 – 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.

Med illustrationens komponentværdier bliver forstærkningen $A_N = -2$.

Udgangsspændingen U_o er givet ved indgangsspændingen U_i og de to modstande R_1 og R_2 som vist herunder. Det negative fortegn betyder at udgangssignalet har den modsatte polaritet af indgangssignalet.

$$U_o = -\frac{R_2}{R_1} U_i \Rightarrow A_N = \frac{U_o}{U_i} = -\frac{R_2}{R_1}$$

Der er ikke nogen begrænsning for hvor kraftig indgangssignalet må være for det kan tilpasses ved valget af modstandsværdier, men indgangsmodstanden er lig med R_1 og kan belaste det kredsløb som driver signalet på indgangen. Forstærkeren er ofte brugbar ved et indgangssignalet lidt under nul da moderne operationsforstærkere ofte tillader de to indgange at være $-0,3\text{ V}$ i forhold til den negative gren af effektforsyningen (tjek altid med databladet). Det er selvfølgelig stadig et krav at udgangsspændingen er indenfor grænserne sat af effektforsyningen.

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 Kirchhoffs knudelov udnytter at summen af de to strømme i modstandene R_1 og R_2 må give 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 .

$$I_1 + I_2 = 0 \Rightarrow \frac{U_I}{R_1} + \frac{U_O}{R_2} = 0$$

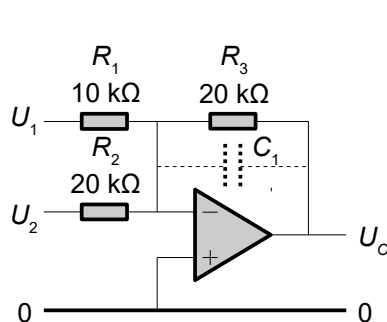
Leddene med U_O og U_I placeres på hver sin side af lighedstegnet.

$$\frac{U_O}{R_2} = -\frac{U_I}{R_1} \Rightarrow U_O = -\frac{R_2}{R_1}U_I$$

Den inverterende forstærker har samme begrænsning som den ikke-inverterende med hensyn til at udgangsspændingen skal holdes indenfor effektforsyningens grænser med en vis margin. Ved en effektforsyning på $\pm 15\text{ V}$ kan der forventes et muligt udgangssignal på maksimalt $\pm 12\text{ V}$. Ved en effektforsyning på $+5\text{ V}$ og nul skal indgangssignalet være negativt fordi udgangen kun kan levere en positiv værdi af udgangsspændingen.

3.3.5 Addition og inversion af signaler

Den inverterende forstærker har et såkaldt virtuelt nul ved den inverterende indgang så den er velegnet til addition af flere signaler. 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 171).



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 : $f_H = \frac{1}{2\pi R_3 C_1}$ hvis $f_H < \frac{GBP}{A_p}$

Ind- og udgangsmodstand:

$$\begin{aligned} R_{I1} &\approx R_1 & R_O &\approx 0 \\ R_{I2} &\approx R_2 \end{aligned}$$

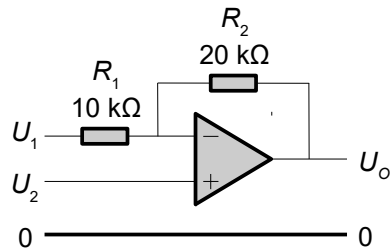
Figur 153 – En virtuel nul indgang kan bruges til at addere flere signaler uden gensidig påvirkning fra signalkildernes indre modstande.

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 båndbredden $f_H = 80\text{ kHz}$ (-3 dB).

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 en positiv referencespænding der typisk er 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 $2,5\text{ V}$.

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 $1 + R_2/R_1$.



Resulterende udtryk:

$$U_O = -\frac{R_2}{R_1}U_1 + \left(1 + \frac{R_2}{R_1}\right)U_2$$

Grænsefrekvens:

$$f_H = \frac{GBP}{A_p} \quad \text{hvor} \quad A_p = 1 + \frac{R_2}{R_1}$$

Ind- og udgangsmodstand:

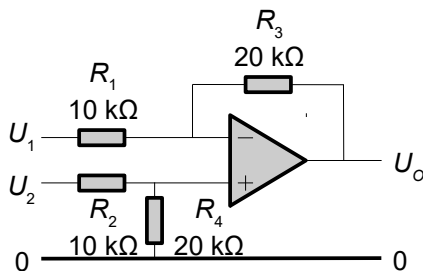
$$R_{I1} \approx R_1 \quad R_O \approx 0 \\ R_{I2} \rightarrow \infty$$

Figur 154 – En operationsforstærker kan bruges til at subtrahere to signaler.

De to signaler forstærkes forskelligt så det er ikke kun differensen der optræder på udgangen, men derimod en kombination af de to indgangssignaler med hver sin forstærkning.

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 korrekt subtraktion er det nødvendigt at dæmpe signalet til den positive indgang svarende til forstærkningen og det opnås med en spændingsdelers.



Komplet udtryk:

$$U_O = -\frac{R_3}{R_1}U_1 + \frac{R_4}{R_2 + R_4} \frac{R_1 + R_3}{R_1}U_2$$

Praktisk realisation:

$$U_O = \frac{R_3}{R_1}(U_2 - U_1) \quad \text{hvis} \quad \begin{matrix} R_1 = R_2 \\ R_3 = R_4 \end{matrix}$$

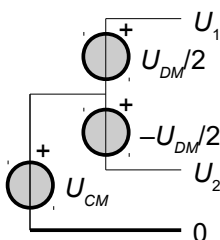
Ind- og udgangsmodstand:

$$R_{I1} \approx R_1 \quad R_O \approx 0 \\ R_{I2} \approx R_2 + R_3$$

Figur 155 – Hvis modstandene tilpasses korrekt kan udgangens signal være en forstærket version af differensen mellem to indgangssignaler.

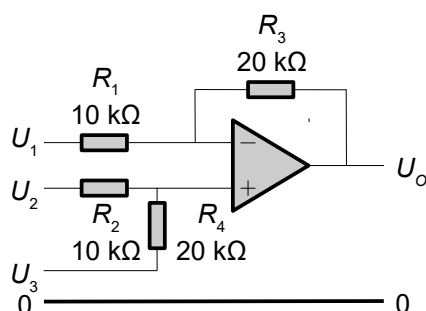
Med fire ens modstandsværdier bliver udgangssignalet en subtraktion med $U_O = U_2 - U_1$ så et eventuelt fælles signal fjernes og kun differensen mellem spændingerne benyttes. Det udnyttes indenfor lydproduktion og medicinsk teknologi til at undgå problemer med 50 Hz netbrum.

Indgangssignalet beskrives ved et nyttesignal U_{DM} (DM, differential mode) og et uønsket fælles signal U_{CM} (CM, common mode) der kan være DC spænding eller en ekstern støjkilde. Udgangssignalet bliver lig med nyttesignalet og det fælles signal undertrykkes.



$$\begin{aligned} U_2 &= U_{CM} + \frac{U_{DM}}{2} \\ U_1 &= U_{CM} - \frac{U_{DM}}{2} \end{aligned} \Rightarrow U_O = U_2 - U_1 = U_{DM}$$

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 $U_O - U_3$.



Komplet udtryk:

$$U_O = -\frac{R_3}{R_1}U_1 + \frac{R_4}{R_2+R_4}\frac{R_1+R_3}{R_1}U_2 + \frac{R_2}{R_2+R_4}\frac{R_1+R_3}{R_1}U_3$$

Praktisk realisation:

$$U_O = \frac{R_3}{R_1}(U_2 - U_1) + U_3 \quad \text{hvis} \quad \begin{matrix} R_1 = R_2 \\ R_3 = R_4 \end{matrix}$$

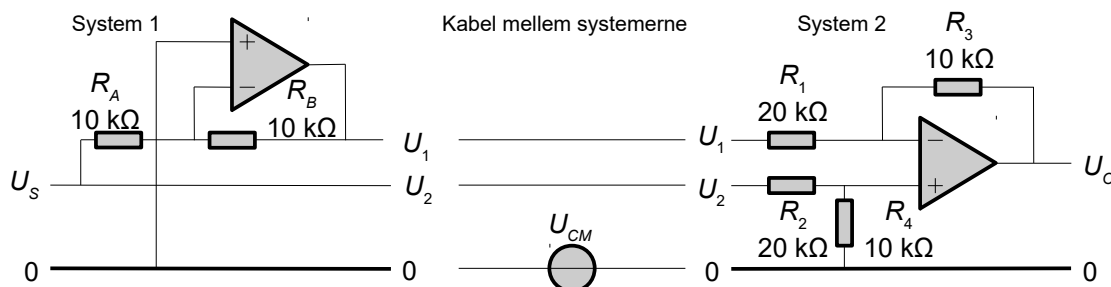
Ind- og udgangsmodstand:

$$\begin{matrix} R_{I1} \approx R_1 \\ R_{I2} \approx R_2 + R_4 \\ R_{I3} \approx R_2 + R_4 \end{matrix} \quad R_O \approx 0$$

Figur 156 – 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. Med en spænding på U_3 vil udgangssignalet svinge om denne værdi som reference. Det betyder at den enkle opstilling kan måle en spændingsdifferens samt overføre målingen mellem to systemer med hver sit nul uden at signalet forurenes med fx lysnettets 50 Hz brum.^{84,85}

For at sende et signal fra system 1 til system 2 benyttes et kredsløb der omdanner et signal U_S til et komplementært udgangssignal hvor $U_2 = U_S$ og $U_1 = -U_S$ og differensen bliver $U_2 - U_1 = 2U_S$. Det komplementære signal sendes til System 2 hvor differensen uddrages, forstærkningen er en halv så $U_O = U_S$ og signalet relateres til det nye systems nul.



Figur 157 – 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 bliver dæmpet.

For at opstillingen kan virke skal spændingsforskellen mellem de to systemers nul U_{CM} være sikkert indenfor operationsforstærkerens effektforsyning og der skal være plads til signalet. Det vil ofte lede til ± 15 V effektforsyning. Operationsforstærkerens båndbredde skal være høj for at kunne dæmpe højfrekvente støjsignaler og evnen til at undertrykke støjsignalet specificeres gennem parameteren $CMRR$ der står for *Common-Mode Rejection Ratio*.⁸⁶

Kredsløbet anvendes ved lydtransmission i studier for at undgå brum fra lysnettets 50 Hz og ved svage signaler fra en mikrofon føres både U_1 og U_2 til signalkilden for at undgå stelplanets støj. Transmission af digitale signaler benytter samme princip ved standarderne RS422 og USB.

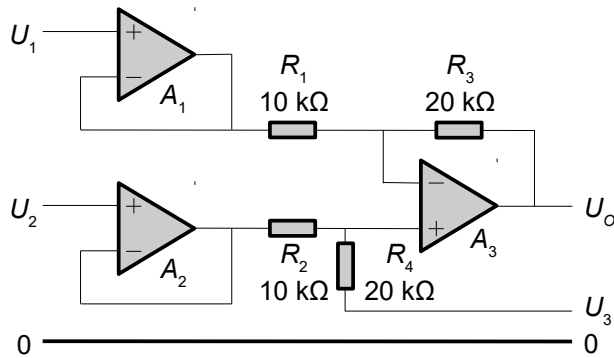
84 Situationen kan opstå hvis to systemer effektforsynes fra hver sin tilslutning til lysnettets 230 V. Vær dog opmærksom på at både System 1 og System 2 skal have en effektforsyning hvor de to nul-potentialer holdes indenfor nogle få volt fra hinanden.

85 Jeg har set det udnyttet i en digital båndoptager omkring 1985 hvor alle printkort for digital signalbehandling kommunikerede data over en differential backplane. Bussen bestod derfor af 16 differentielle sendere og modtagere for en 16 bit ordlængde.

86 https://en.wikipedia.org/wiki/Common-mode_rejection_ratio.

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. Operationsforstærker A_3 giver den indstillede forstærkning og er den betydende komponent for opstillingens båndbredde.



Resulterende udtryk:

$$U_o = \frac{R_3}{R_1}(U_2 - U_1) + U_3 \quad \begin{matrix} R_1 = R_2 \\ R_3 = R_4 \end{matrix}$$

Grænsefrekvens:

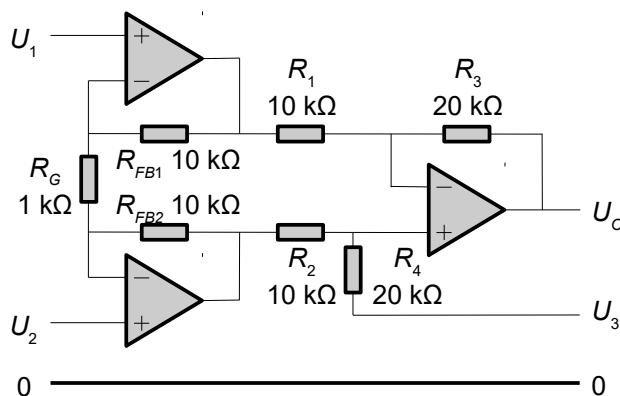
$$f_H = \frac{GBP}{1 + \frac{R_3}{R_1}}$$

Ind- og udgangsmodstand:

$$\begin{matrix} R_{I1} \rightarrow \infty \\ R_{I2} \rightarrow \infty \end{matrix} \quad R_o \approx 0$$

Figur 158 – 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 udgaver af kredsløbet 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 ellers påvirkes balancen. Ved en værdi af R_4 på 20 kΩ skal den eksterne modstandsværdi være under 20 Ω.



Resulterende udtryk:

$$U_o = A_{DIF}(U_2 - U_1) + U_3$$

$$A_{DIF} = \left(1 + \frac{2R_{FB}}{R_G}\right) \frac{R_3}{R_1} \quad \begin{matrix} R_{FB1} = R_{FB2} \\ R_1 = R_2 \\ R_3 = R_4 \end{matrix}$$

Grænsefrekvens (laveste værdi gælder):

$$f_H = \frac{GBP}{1 + \frac{2R_{FB}}{R_G}} \quad \text{eller} \quad f_H = \frac{GBP}{1 + \frac{R_3}{R_1}}$$

Ind- og udgangsmodstand:

$$\begin{matrix} R_{I1} \rightarrow \infty \\ R_{I2} \rightarrow \infty \end{matrix} \quad R_o \approx 0$$

Common-mode rejection:

$$CMRR = \frac{U_{DIF}}{U_{CM}} = \frac{U_o}{A_{DIF} U_{CM}}$$

$$U_o = (CMRR \cdot A_{DIF}) U_{CM}$$

Figur 159 – En klassisk instrumentationsforstærker med tre operationsforstærkere som her er vist med en indstillet forstærkning på 42 gange.

Undertrykkelse af forstyrrelser oplyses i databladet 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 så man er nødt til at beregne amplituden af forstyrrelsen ved udgangen i den givne situation.⁸⁷

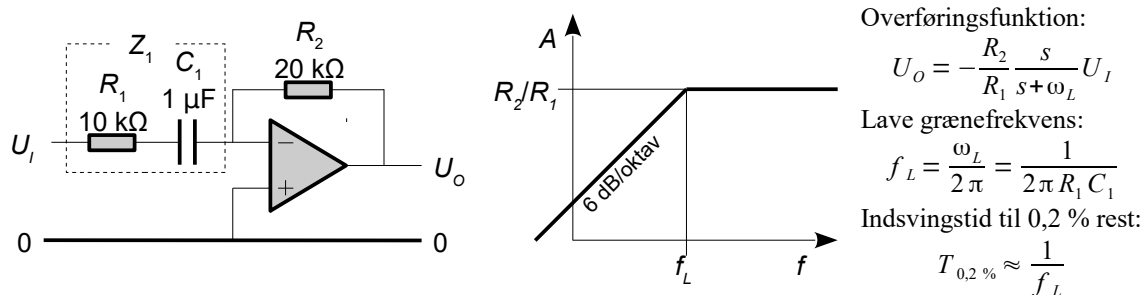
⁸⁷ https://en.wikipedia.org/wiki/Common-mode_rejection_ratio.

3.3.9 Filtrende kredsløb

Operationsforstærkeren kan benyttes for realisering af de matematiske funktioner differentiation og integration som hyppigt benyttes ved interface til sensorer. Dertil kommer behovet for at blokere for et DC niveau undervejs i en signalkæde eller dæmpe en højfrekvent interferens fra mobiltelefonen. Herunder vises hvordan funktionerne kan realiseres og der er flere eksempler i fjerde kapitel.

3.3.10 Højpasled

Den inverterende forstærker på side 156 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. Efter denne indsvingning er der en rest på 0,2 % af det oprindelige DC niveau.



Figur 160 – En kondensator blokerer DC og ved tilstrækkeligt høje frekvenser er forstærkningen givet af modstandsforholdet (højpasfilter).

Med de viste værdier er tidskonstanten $R_1 C_1 = 10$ ms så den lave grænsfrekvens er $f_L = 16$ Hz og opstillingen er i ro efter $T_\infty = 63$ ms. Forstærkningen af signaler noget over grænsfrekvensen 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 er nu suppleret 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 $R_1 C_1$. 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 $R_1 C_1$. 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.

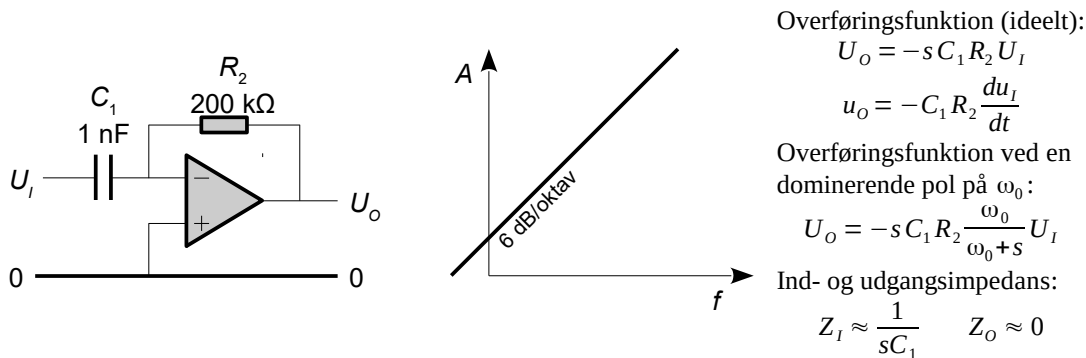
$$U_o = -\frac{R_2}{R_1} \frac{s}{s + \omega_L} \frac{U_{DC}}{s} = -\frac{R_2}{R_1} \frac{1}{s + \omega_L} U_{DC}$$

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.

$$u_o(t) = -\frac{R_2}{R_1} \exp(-\omega_L t) U_{DC}$$

3.3.11 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. I den moderne elektronik mødes differentiationen ved interface til sensorer hvis signalamplitude aftager med stigende frekvens. For kredsløbet herunder differentieres signalet, så udgangens amplitude er proportional både med indgangssignalets amplitude og med frekvensen.⁸⁸



Figur 161 – En differentierende opstilling hvor forstærkningen er proportional med frekvensen indtil operationsforstærkeren ikke længere har tilstrækkelig båndbredde.

Med de viste komponenter er kondensatorens 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 171). 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 det 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.

$$u_O = -i R_2 = -C_1 R_2 \frac{du_C}{dt} = -C_1 R_2 \frac{du_I}{dt}$$

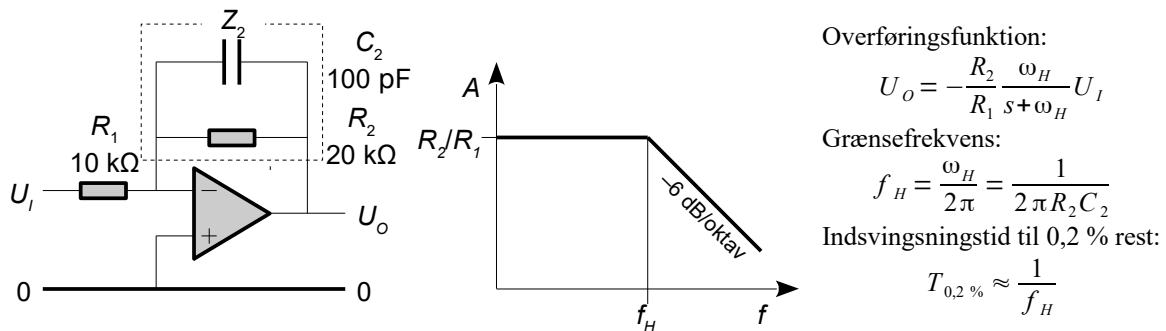
Udtrykt i Laplace notationen erstattes en differentiation af s ganget med den transformerede af indgangsspændingen hvilket er $U_I(s)$.

$$U_O = -s C_1 R_2 U_I$$

⁸⁸ 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.

3.3.12 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 højfrekvent støj fra en mobiltelefon og en utilsigtet modtagelse af radiotransmissioner. Amplituden er 3 dB nede ved grænsefrekvensen f_H så ved tre seriekoblede forstærkere i signalkæden lægges grænsefrekvensen en oktav oppe så amplituden kun dæmpes 1 dB per trin ved den ønskede høje grænsefrekvens.



Figur 162 – 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 dB/oktav (-20 dB/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.

$$U_o = -\frac{Z_2}{R_1} U_i = -\frac{Z_R \parallel Z_C}{R_1} U_i = -\frac{R_2(1/sC_2)}{R_1(R + 1/sC_2)} U_i = -\frac{R_2}{R_1} \frac{1}{1 + sR_2C_2} U_i$$

Grænsefrekvensen er givet ved den frekvens hvor impedansen af kondensatoren er numerisk lig med værdien af modstanden.

$$R_2 = \left| \frac{1}{sC_2} \right| \Rightarrow R_2 = \frac{1}{\omega_H C_2} \Rightarrow f_H = \frac{\omega_H}{2\pi} = \frac{1}{2\pi R_2 C_2}$$

3.3.13 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 dB/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 arbejds punktet 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

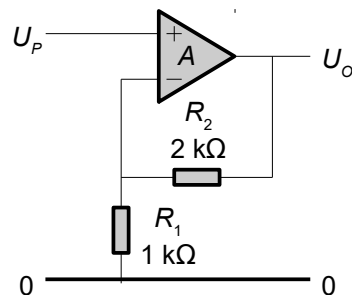
Note 2.

3.4 DC begrænsninger

Operationsforstærkeren har nogle begrænsninger der har betydning for hvor præcist den arbejder. Det er fuldt ud muligt at designe velfungerende elektronik uden at tage andre forholdsregler end at være forsigtig, men det er bedre at være bekendt med begrænsningerne for at sikre sig at ens design ikke har svagheder.

3.4.1 Tilbagekoblet forstærker

En operationsforstærker forstærker differensen $U_P - U_N$ mellem de to indgange med dens åben-sløjfe forstærkning A_{OL} (open-loop gain) til udgangens U_O . Værdien af A_{OL} er omkring 10^6 så en forskel mellem de to indgange på $10 \mu\text{V}$ forstærkes op til 10 V ved udgangen. De to indgange kan ligge 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 næsten det samme potentiale som den positive indgang hvilket leder til designprincippet fra side 153 og som er selve fundamentet for alt praktisk arbejde med operationsforstærkere.



Overføringsfunktion:

$$U_O = A U_P$$

Forstærkningsfaktor:

$$A = \frac{A_{OL}}{1 + B A_{OL}}$$

$$A_{OL} \approx 10^6 \quad A_{OL} \rightarrow \infty$$

$$B = \frac{R_1}{R_1 + R_2}$$

$$A \approx 1 + \frac{R_2}{R_1}$$

$$A_{OL} > 1000 A$$

Figur 163 – Den ikke-inverterende forstærker. En brøkdel af udgangssignalet kobles tilbage til den negative indgang for at fastlægge forstærkningen bestemt alene ved modstandenes værdier og uden betydende relation til forstærkerens specifikationer.

I designet herover vil udgangen indstille sig på netop den værdi der bringer den negative indgang hen til potentialet for den positive indgang. Med spændingsdelingen B mellem udgangen og den negative indgang betyder det at udgangssignalet skal forstærkes med $1/B$ for at opnå balance.

Med de viste komponentværdier forventes en forstærkning på 3 gange og praksis vil man nå det indenfor $-0,0003 \%$ så det er i høj grad modstandenes præcision der bestemmer. Ved en højere forstærkning bliver fejlen større og åben-sløjfe forstærkningen A_{OL} skal som en tommelfingerregel være mindst 1000 gange højere end den ønskede forstærkning for at holde fejlen under $0,1 \%$. En typisk OpAmp kan derfor benyttes til en DC forstærkning på højest 1000 gange (60 dB).

Udgangsspændingen er givet ved differensen mellem de to indgange som forstærkes op af åben-sløjfe forstærkningen. Ved tilbagkobling angiver B den brøkdel af udgangsspændingen som føres tilbage til den negative indgang.

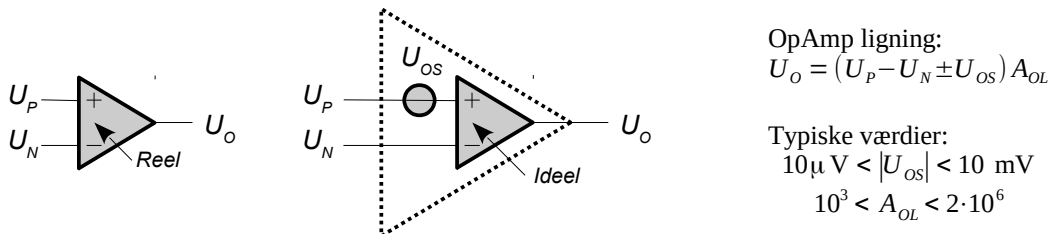
$$U_O = (U_P - U_N) A_{OL} \quad \text{hvor} \quad \begin{aligned} U_N &= B U_O \\ B &= \frac{R_1}{R_1 + R_2} \end{aligned} \Rightarrow U_O = \frac{A_{OL}}{1 + B A_{OL}} U_P$$

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 leddet $1/A_{OL}$ skal $1/A_{OL}$ være meget mindre end B og da forstærkningen netop er givet ved $1/B$ bliver kravet $A_{OL} > 100 A$ for en fejl under 1% ved at bruge tilnærmelsen.

$$A = \frac{U_O}{U_P} = \frac{1}{1/A_{OL} + B} \xrightarrow{A_{OL} \rightarrow \infty} \frac{1}{B} \Rightarrow A = \frac{R_1 + R_2}{R_1} = 1 + \frac{R_2}{R_1}$$

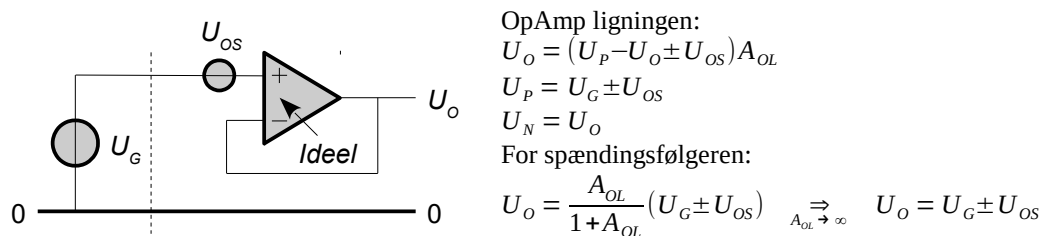
3.4.2 Indgangens fejlspænding

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 operationsforstærker så udgangsspændingen vil ikke blive nul hvis indgangene kortsluttes. Den praktiske betydning er operationsforstærkerens præcision påvirkes. Om det så i praksis udgør et problem afhænger af den aktuelle situation.



Figur 164 – Ubalance i indgangen indføres i modellen som en fejlspænding 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 fejlspænding 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 fejlspændingen med symbolet U_{OS} (*input voltage offset*). Fortegnet er ikke kendt og det er den aktuelle værdi heller ikke. Det eneste fabrikanten oplyser er at værdien ligger indenfor et interval. Hvis en forstærker har specifikationen $U_{OS} = 5 \text{ mV}$ så er det eneste man ved at værdien 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 design af et elektronisk kredsløb. Tilgængelige operationsforstærkere har meget stor variation i netop denne parameter og der er opnået store fremskridt gennem tiden.⁸⁹



Figur 165 – En spændingsfølger skal ideelt blot gengive indgangsspændingen fra generatoren på udgangen, men fejlspændingen giver en usikkerhed på udgangsspændingen.

Fejlspændingen er funktion af temperaturen og oplyses derfor ved en fast temperatur på 25°C . For krævende brug specificeres også en variationen over et temperaturområde. Fejlspændingen driver over tid så den er ikke en fast værdi, men vil variere i løbet af produktets levetid. Dertil kommer at den momentant kan springe mellem forskellige værdier hvilket specificeres som en spids-til-spids værdi for en OpAmp til krævende anvendelser (se popcorn støj på side 196).^{90,91,92}

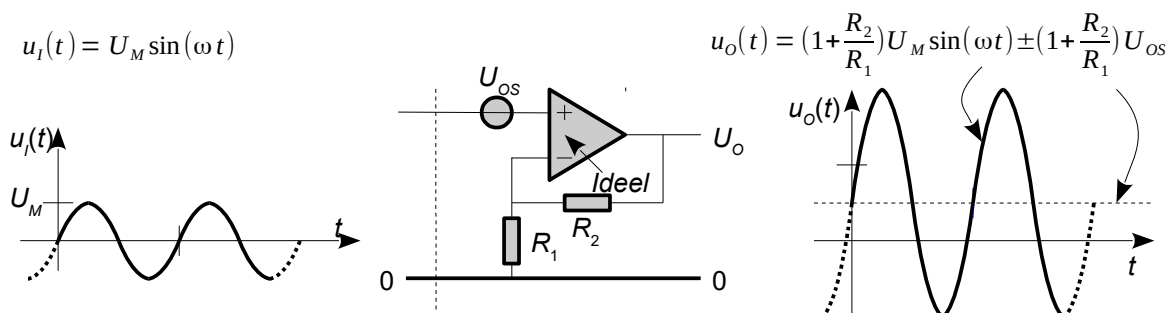
⁸⁹ Værdien er omkring $\pm 5 \text{ mV}$ for gamle travere som uA741, LM324 og LF356. Den nyere OpAmp OP27 har maksimalt $100 \mu\text{V}$. Man kan nå længere ned med et internt korrektionskredsløb på chippen: fx $\pm 10 \mu\text{V}$ for ADA4051 og $\pm 1 \mu\text{V}$ for AD8628.

⁹⁰ OP27 oplyses til en temperaturvariation på maksimalt $0,2 \mu\text{V}/^\circ\text{C}$.

⁹¹ OP27 oplyses til en drift på maksimalt $2 \mu\text{V}/\text{måned}$.

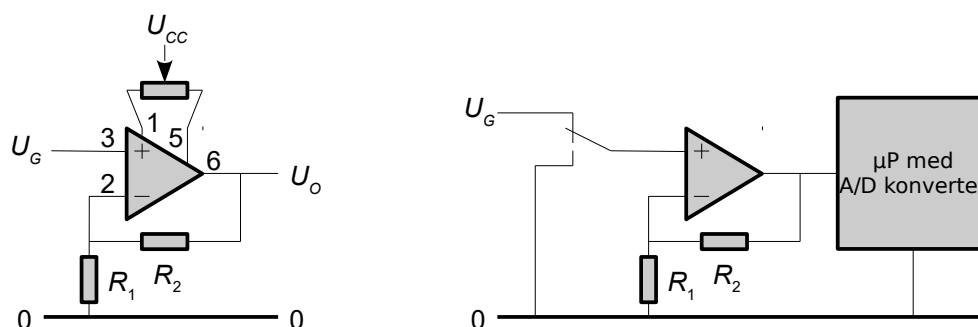
⁹² OP27 oplyses til en popcorn støj maksimalt 80 nV spids-spids indenfor et frekvensområde fra $0,1 \text{ Hz}$ til 10 Hz .

Hele idéen er at undersøge om problemet er så stort at der er behov for en aktion eller om det er så ringe at man helt kan ignorere fejlspændingen. 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 fejlspændingen. En anden mulighed er at sætte et krav til signalniveauet eller man kan blive tvunget til at revurdere sit kredsløb.



Figur 166 – En forstærker vil forstærke både signal og fejlspænding.

Man kan indføre et potentiometer for manuel trimning af fejlspændingen, men det er bestemt ikke uden omkostninger. For det første koster den ekstra komponent penge og den optager plads på printkortet. Dertil er der en udgift i produktionen fordi en tekniker skal forestå trimningen og uanset hvor omhyggelig personen er så giver det en fejlmulighed som kræver en kontrol et senere sted i produktionsforløbet. Endelig et potentiometer ikke en velset komponent på grund af en ringe mekanisk stabilitet. Trimning kan benyttes i en forsøgsopstilling, men den er ikke anvendelig for en serieproduktion hvor prisen på en bedre OpAmp kan være en mere attraktiv løsning.



Figur 167 – Fejlspændingen 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 fejlspændingen i software.

En metode der egner sig bedre til moderne kredsløb bruger en mikroprocessor til at måle den aktuelle værdi af kredsløbets fejlspænding. En analog omskifter skal inkluderes i opstillingen og på passende tidspunkter i programmet kortsluttes forstærkerens indgang og dens udgangsspænding måles og gemmes i et register. Fejlspændingen fjernes fra de efterfølgende målinger ved at subtrahere registrets værdi fra måledata. For at håndtere drift i parameteren gentages målingen med et passende tidsinterval.⁹³

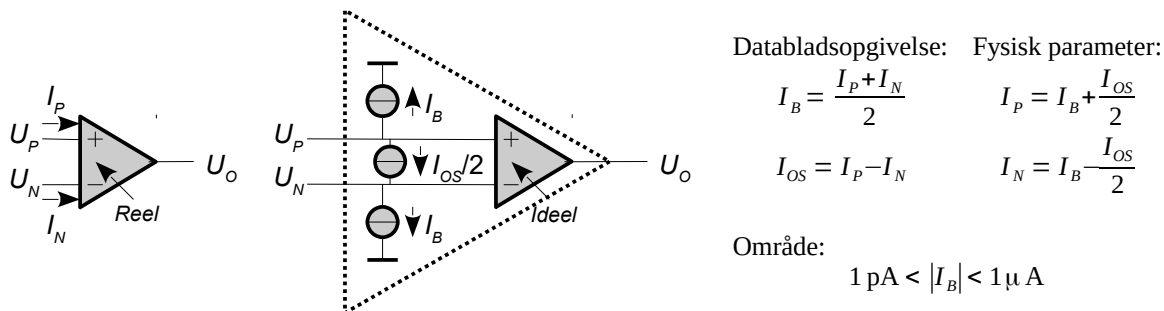
Fejlspændingen er den ene af to væsentlige fejlkilder ved DC måling og den anden introduceres herunder. Afslutningsvis kombineres fejlkilderne til en fælles DC model for en OpAmp.

93 En tilsvarende løsning findes integreret på operationsforstærkerens chip hvor der er indbygget et kredsløb for nulstilling af fejlspændingen. Det er de *chopper stabilised amplifiers* som fx AD8628 hvor fejlspændingen holdes under $\pm 1 \mu V$. En oscillator vil fra tid til anden afbryde indgangen for at måle fejlspændingen 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.

3.4.3 Indgangsstrøm

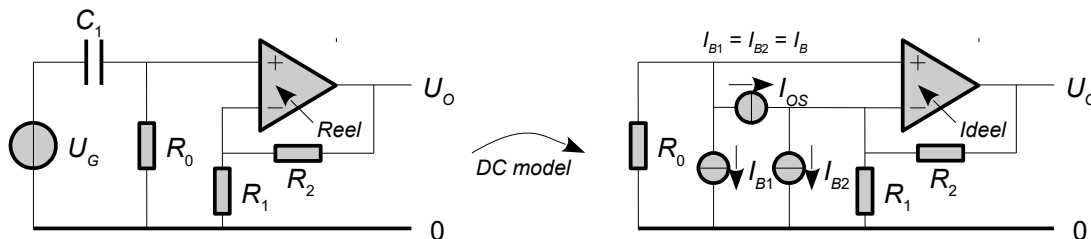
Operationsforstærkeren kræver at der løber en vis strøm i indgangen for at elektronikken fungerer og strømmen belaster det kredsløb der måles på. Det kaldes for biasstrøm som en fordanskning af den engelske betegnelse og værdien varierer mellem de forskellige typer. Værdien er cirka 1 pA ved MOSFET og kan nå op på 1 μ A ved bipolære transistorer. Strømretningen er som hovedregel kendt ved bipolære transistorer, men databladet bør konsulteres hvis det er vigtigt.^{94,95}

For at analysere kredsløb for biasstrømmens betydning benyttes en DC model hvor en OpAmp beskrives som en ideel forstærker med strømkilder i de to indgange. Strømmen kaldes I_P i den positive indgang og I_N i den negative. Databladet oplyses kun en gennemsnitlig værdi kaldet I_B (*bias current*) og en differens kaldet I_{OS} (*bias offset current*). Der er ikke nogle gode danske betegnelser for disse strømme så jeg benytter de engelske.



Figur 168 – 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 forstærker for et analogt signal hvor indgangssignalet føres via en kondensator for at blokere for DC. Forstærkeren skal kun forstærke AC signaler. Udgangens DC værdi skal være nul, men biasstrømmen kan påvirke det så meget at det indvirker på den tilsigtede funktion. Der er ikke nogen fast grænse for værdien afhænger af den aktuelle opgave.



Figur 169 – 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 arbejds punktet.

Formålet med en analyse er at beregne hvor meget udgangens DC arbejds punkt afviger fra nul og derefter kan man så vurdere om der er et problem.

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.

94 Med bipolære transistorer er værdien cirka 100 nA (LM324, uA741) og en støjsvag forstærker kan nå over 1 μ A (NE5532, AD797). Med fetteffekt 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). Værdien er relativt uafhængig af temperaturen for de bipolære transistorer hvorimod fetteffekt transistorer typisk fordobler strømmen for hver 10°C stigning. Ved 100°C er biasstrømmen for en JFET almindeligvis øget til samme niveau som den bipolære transistor og ved 150°C er det også tilfældet for en MOSFET.

95 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.

Beregningen gennemføres ved anvendelse af superpositionsprincippet. Det betyder at alle kilder nulstilles (strømkilden afbrydes) og én efter én aktiveres kilden og resultatet beregnes. Når alle bidrag er kendt adderes de med hensyn til fortegn.

I_{B1} – Ved den positive indgang på operationsforstærkeren løber strømmen I_B gennem 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 R_2 og genererer et spændingsfald på udgangen på $R_2 I_B$. Det løber ingen strøm igennem R_1 fra strømkilden da 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 $-R_0 I_{OS}/2$ 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. Ligningen simplificeres ved at samle bidragene fra I_B og I_{OS} til hvert sit led.

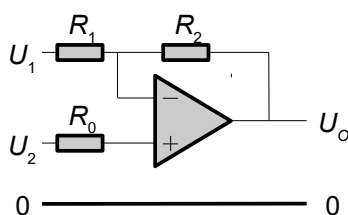
$$\begin{aligned} & \begin{array}{ccc} \text{Bidrag fra } I_{B1} & \text{Bidrag fra } I_{B2} & \text{Bidrag fra } I_{OS} \\ U_O = & -(1+\frac{R_2}{R_1})R_0 I_B & +R_2 I_B & -(1+\frac{R_2}{R_1})R_0 \frac{I_{OS}}{2} - R_2 \frac{I_{OS}}{2} \\ \Rightarrow U_O = & [R_2 - (1+\frac{R_2}{R_1})R_0] I_B - [R_2 + (1+\frac{R_2}{R_1})R_0] \frac{I_{OS}}{2} \end{array} \end{aligned}$$

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 $2R_2$.

$$R_2 - \frac{R_1 + R_2}{R_1} R_0 = 0 \Rightarrow R_2 [1 - \frac{R_1 + R_2}{R_1 R_2} R_0] = 0 \Rightarrow \frac{R_1 R_2}{R_1 + R_2} = R_0$$

3.4.4 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 ($U_2 = 0$). Bidraget fra biasstrømmen I_B kan reduceres ved at sætte modstanden R_0 lig med parallelværdien af R_1 og R_2 . Bidragene fra fejlspændingen U_{OS} og offsetstrømmen I_{OS} kan ikke undertrykkes.⁹⁶



Generelt udtryk:

$$U_O = (1 + \frac{R_2}{R_1})(U_2 \pm U_{OS}) - \frac{R_2}{R_1} U_1 + (1 - \frac{R_0}{R_1 \parallel R_2}) R_2 I_B - (1 + \frac{R_0}{R_1 \parallel R_2}) R_2 \frac{I_{OS}}{2}$$

Forenklet udtryk:

$$U_O = (1 + \frac{R_2}{R_1})(U_2 \pm U_{OS}) - \frac{R_2}{R_1} U_1 \pm R_2 I_{OS} \quad \text{hvis} \quad R_0 = \frac{R_1 R_2}{R_1 + R_2}$$

Figur 170 – Resulterende udtryk for operationsforstærkeren med fejlspændingen 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 bipolære 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.

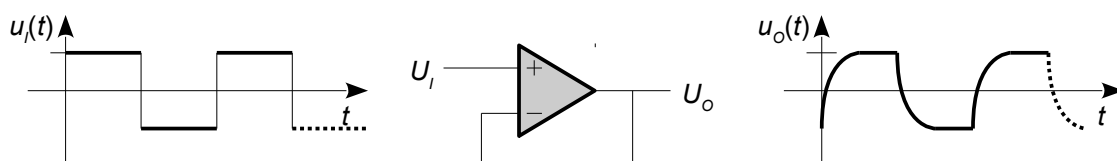
⁹⁶ Som regel har man en designmæssig udfordring hvis offsetstrømmen bidrager med et problem og så er det bedre at genoverveje selve kredsløbet. Ellers kommer man til at slås med at få det til at virke i en produktionslinje og ikke mindst hos kunderne.

3.5 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 177 og herunder illustreres det med nogle eksempler.

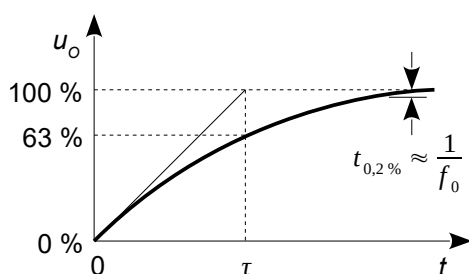
3.5.1 Steprespons

En pludselig ændring af indgangsspændingen til en operationsforstærker vil ikke resultere i en momentan ændring af udgangen til den nye position. Udgangen vil roligt glide fra den hidtidige position til den nye værdi.



Figur 171 – 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 når spændingstrinnet ikke er for voldsomt, og det kan tolkes som om operationsforstærkeren var et lavpasfilter med en enkelt pol. Det kan ses ved brug af værktøjet fra Laplace transformationen.



Frekvensdomæne

Tidsdomæne

$$U_I(s) = \frac{U_M}{s}$$

$$u_I(t) = U_M e(t)$$

$$U_O(s) = \frac{\omega_0}{\omega_0 + s} \frac{U_M}{s}$$

$$u_O(t) = U_M [1 - \exp(-\omega_0 t)]$$

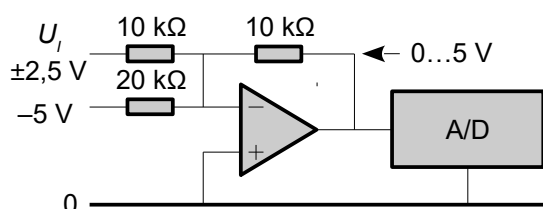
$$\omega_0 = 2\pi f_0$$

$$\tau = \frac{1}{\omega_0}$$

Figur 172 – 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 nå 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 :

$$\epsilon = \exp(-\omega_0 t_N) \Rightarrow t_N = \frac{1}{\omega_0} \ln\left(\frac{1}{\epsilon}\right)$$

For $\epsilon = 2^{-N}$:

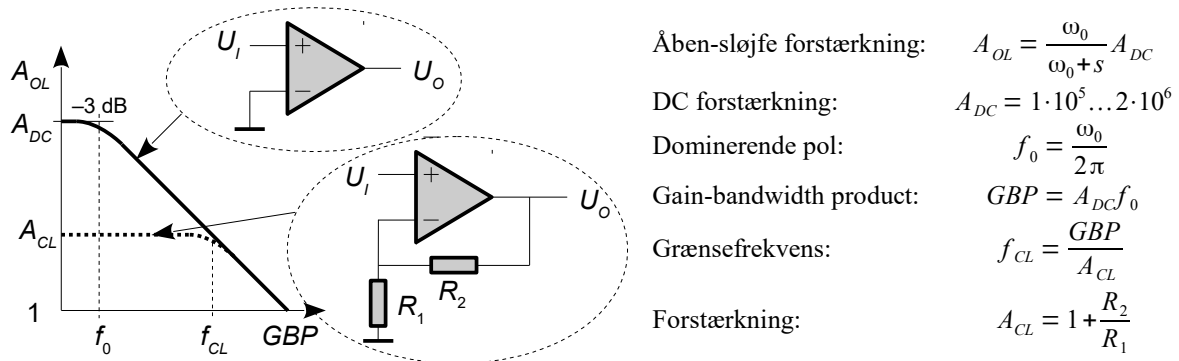
$$t_N = \frac{N \ln(2)}{2\pi f_0} \approx \frac{N}{9 f_0} \Rightarrow_{N=18} t_{18} \approx \frac{2}{f_0}$$

Figur 173 – 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.

Med en 18 bit konverter ($N = 18$) vil en båndbredde på $f_0 = 1$ MHz give en settling time på $t_{18} = 2$ μ s.

3.5.2 Dominerende pol

Operationsforstærkeren er opbygget som et lavpasfilter med stor forstærkning ved lave frekvenser for at opnå stabilitet i flertallet af opstillinger. Det betyder at åben-sløjfe forstærkningen A_{OL} aftager over en frekvens der kaldes for den dominerende pol f_0 (*dominating pole*). Det er et fælles træk for alle operationsforstærkere og betyder at forstærkerens overskud aftager mod høje frekvenser så en OpAmp er ikke i stand til at fungere lige godt ved alle frekvenser.



Figur 174 – Åben-sløjfe forstærkningen for en forstærker med dominerende pol. Ved DC er værdien af forstærkningen A_{DC} og den aftager med -6 dB/oktav over den dominerende pol f_0 . Ved GBP er forstærkningen faldet til én (0 dB båndbredden).

Ved DC og lave frekvenser er forstærkningen givet ved A_{DC} der typisk er 10^6 og som sikrer at selv med en udgangsspænding på ± 10 V vil der være en differens mellem de to indgange på mindre end ± 10 μ V hvilket i praksis kan ignoreres ved DC applikationer.

Over den dominerende pol aftager åben-sløjfe forstærkningen med -6 dB/oktav (-20 dB/dekade). 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) som også kaldes 0 dB båndbredde (*unity-gain bandwidth*) eller 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.⁹⁷

Produktet af indstillet forstærkning A_{CL} og båndbredde f_{CL} er konstant for frekvenser over f_0 så man kan let beregne båndbredden ved en givet værdi af forstærkningen. Man dividerer simpelthen GBP specifikationen med den ønskede forstærkning. Tilsvarende kan man beregne den højest mulige forstærkning for en ønsket båndbredde ved at dividere GBP specifikationen med værdien af den ønskede båndbredde.

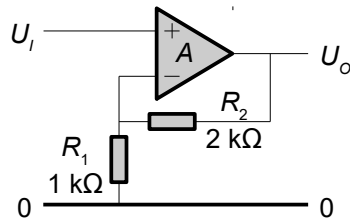
Som et eksempel benyttes en operationsforstærker med $GBP = 3$ MHz. Hvis den skal bruges som spændingsfølger kan den klare signaler indtil 3 MHz hvor dens forstærkning er reduceret fra 0 dB til -3 dB. Skal operationsforstærkeren have en forstærkning på 100 vil den mulige båndbredde blive 30 kHz ved -3 dB grænsefrekvensen.

⁹⁷ Det ses også benævnt transitfrekvensen f_t .

3.5.3 Forstærkning versus båndbredde

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** forstærker er båndbredden givet ved GBP for operationsforstærkeren divideret med den værdi af forstærkning man vælger ved modstandene i tilbagekoblingen.

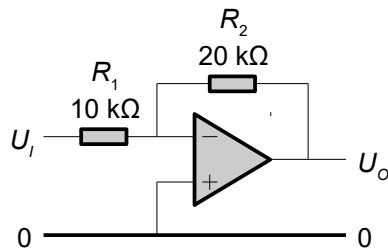


Forstærkning	Båndbredde
$A_{CL} = 1 + \frac{R_2}{R_1}$	$f_{CL} = \frac{GBP}{1 + \frac{R_2}{R_1}}$

Figur 175 – 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 $GBP = 3$ MHz vil forstærkeren kunne fungere op til $f_{CL} = 1$ MHz ved -3 dB punktet.

For den **inverterende** forstærker skal beregningen af båndbredde benytte samme formel selv om forstærkningen af signalet er $-R_2/R_1$ for operationsforstærkerens modtager også her $R_1/(R_1 + R_2)$ gange udgangssignalet til den inverterende indgang. Operationsforstærkeren kan ikke se forskel på om den bruges som inverterende eller ikke-inverterende forstærker.



Forstærkning	Båndbredde
$A_{CL} = -\frac{R_2}{R_1}$	$f_{CL} = \frac{GBP}{1 + \frac{R_2}{R_1}}$

Figur 176 – 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 $GBP = 3$ 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 B af udgangssignalet tilbage til den inverterende indgang. Resultatet er et nyt lavpasfilter hvor polen er flyttet op i frekvens. Normalt er B større end 0,001 og A_{DC} er omkring 10^6 så deres produkt er mindst 1000 med almindelige operationsforstærkere og dermed vindes der ganske meget i båndbredde.

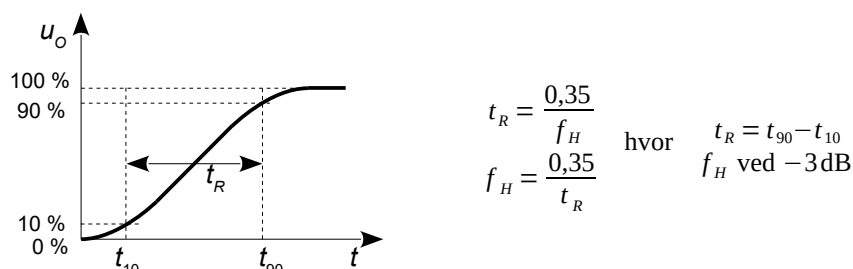
$$A_{CL} = \frac{A_{OL}}{1 + B A_{OL}} \Rightarrow A_{CL} = \frac{\frac{\omega_0}{\omega_0 + s} A_{DC}}{1 + B \frac{\omega_0}{\omega_0 + s} A_{DC}} = \frac{\omega_0 A_{DC}}{(1 + B A_{DC}) \omega_0 + s} = \frac{A_{DC}}{1 + B A_{DC}} \frac{(1 + B A_{DC}) \omega_0}{(1 + B A_{DC}) \omega_0 + s}$$

$$B = \frac{R_1}{R_1 + R_2}$$

Faktoren $1 + B A_{DC}$ kaldes for tilbagekoblingsfaktoren (*feedback factor*) og angiver det overskud som forstærkeren bruger til at korrigere sin udgangsspænding med. Tilbagekoblingsfaktoren vil reducere den harmoniske forvrængning med $1 + B A_{DC}$, støj indkoblet fra effektforsyningen vil blive dæmpet på samme måde og udgangsmotstanden fra kredsløbet reduceres tilsvarende.

3.5.4 Måling af båndbredde

Ved aflæsning af stepresponsen på et oscilloskop kan det være svært at se hvor reaktionen på et step starter og det giver en usikkerhed i bestemmelsen af båndbredden som benytter hældningen af tangenten til starttidspunktet. Der findes en anden målemetode hvor man bestemmer signalets stigitid (*rise time*). Den er defineret som tiden fra signalet krydser 10 % af slutværdien og indtil den krydser 90 % af slutværdien.



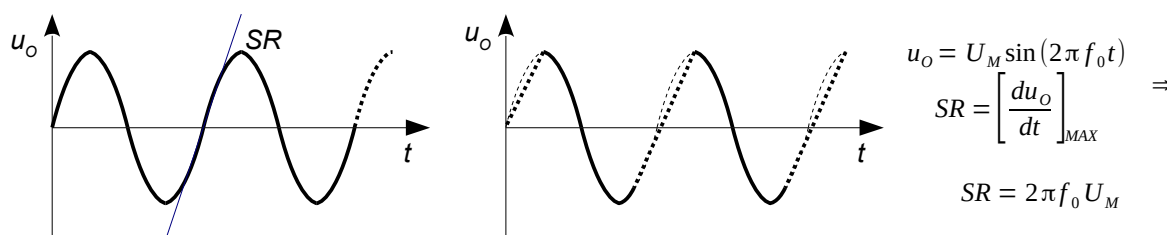
Figur 177 – Stigitiden af et signal er relateret til den høje grænsefrekvens f_H . Formlen er vist for et første-ordens lavpasfilter, men beregningen benyttes normalt uden hensyn til filtrets orden.

Metoden benyttes ofte ved større kredsløb med mange operationsforstærkere hvor det ikke er muligt at beskrives overføringsfunktionen som et første-ordens lavpasfilter. Højere ordens filtre vil starte med vandret tangent hvilket gør det meget svært at definere starttidspunktet. Derimod er start- og slutniveauerne lette at definere. Metoden er så populær at de ældre oscilloskoper har angivet hvordan signalet skal indstilles på skærmen for at lette aflæsningen af stigitiden. De digitale oscilloskoper kan endda indstilles så de selv beregner stigitiden.

Det er klart at en formel udledt for et første-ordens filter ikke umiddelbart giver et gyldigt resultat ved højere orden, men metoden anvendes alligevel ofte. Den er enkel og resultatet er i praksis ganske brugbart om end man ikke skal regne med for mange decimaler.

3.5.5 Slew rate

Forudsætningen for beregningen af rise time er at operationsforstærkeren ikke overstyres og 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/ μ s for en prisbillig OpAmp og op til 1000 V/ μ s for de hurtige versioner. Hvis indgangen skifter abrupt med en høj spændingsværdi vil operationsforstærkeren blive overstyret og så falder forudsætningen for beregningerne.



Figur 178 – 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.

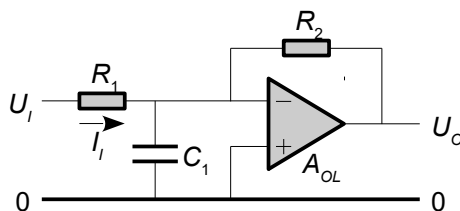
Parameteren er vigtig ved analog signalbehandling af lydkilder. For et harmonisk signal med en amplitude på 10 V er den maksimalt tilladelige frekvens på 16 kHz for forstærkeren uA741 med specifikationen $SR = 1 \text{ V}/\mu\text{s}$ og det er for lidt til krævende formål.

3.5.6 Stabilitet

At en operationsforstærker er stabil vil sige at udgangen reagerer som et første-ordens filter på en ændring ved indgangen. Et pludseligt skift på indgangen vil få udgangen til roligt at glide hen til sin nye position. Det er tilfældet i hovedparten af opstillinger og man behøver ikke normalt at bekymre sig om det. Stabiliteten kompromitteres dog hvis ind- eller udgang belastet af en kondensator som resulterer en pol i overføringsfunktionen. Eksemplerne er en fotodiode eller et accelerometer på en inverterende indgang og et skærmbord der forbinder to systemer. For en generel behandling af stabilitet henvises til side 203 og herunder analyseres problemerne med kapaciteter.^{98,99,100}

3.5.7 Kapacitet på indgang

Tilbagekoblingen R_2C_1 danner en pol i tilbagekoblingen. Den reducerer stabiliteten så der kan opstå oversving ved en steppåvirkning. Den tilladte værdi af kapaciteten afhænger af GBP for operationsforstærkeren og modstandsniveauet som vist herunder. Grænsen for stabilitet er reelle rødder i nævnerens polynomium. For en inverterende forstærker føres signalet ind ved U_i via R_1 mens en transimpedansforstærker får strømmen I_i ført direkte ind og R_1 behøves ikke.



Inverterende forstærker:

$$C_1 < \frac{1}{8\pi GBP R_2} \left(1 + \frac{R_2}{R_1}\right)^2$$

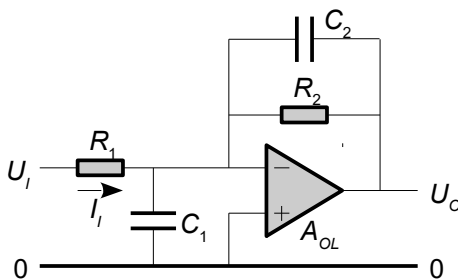
Transimpedansforstærker:

$$C_1 < \frac{1}{8\pi GBP R_2} \quad \text{for } R_1 \gg R_2$$

Figur 179 – En inverterende forstærker kan blive ustabil hvis modstandsniveauet er for højt overfor spredningskapaciteten C_1 på typisk 10 pF eller kondensatoren skyldes signalgiveren.

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 ikke kan realiseres så der må forventes ustabilitet.

Det er muligt at acceptere en kapacitet C_1 på indgangen med kondensator C_2 i tilbagekoblingen. Prisen for stabilitet er en reduktion i signalbehandlingens båndbredde f_H .



Krav til C_1 når C_2 ikke er monteret:

$$C_1 < \frac{1}{8\pi GBP R_2} \left(1 + \frac{R_2}{R_1}\right)^2$$

Stabilisation ved C_2 når C_1 er for stor:

$$C_2 \approx \sqrt{\frac{C_1}{GBP R_2}}$$

Resulterende grænsefrekvens:

$$f_H = \frac{1}{2\pi R_2 C_2}$$

Figur 180 – Kapaciteten på indgangen giver et uønsket negativt fasedrej, men det kan modvirkes af et positivt fasedrej gennem en parallelkoblet kapacitet over tilbagekoblingen.

For R_1 uendelig stor, $R_2 = 1 \text{ M}\Omega$ og $GBP = 10 \text{ MHz}$ kan en kapacitet på $C_1 = 330 \text{ pF}$ accepteres ved at introducere $C_2 = 4,6 \text{ pF}$. Båndbredden af signalet begrænses dog til $f_H = 35 \text{ kHz}$.

Det er populært at sætte en kondensator C_2 i parallel med tilbagekoblingens modstand for at opnå tørre tolerance overfor en kondensator på indgangen. Det er også muligt at isolere kapaciteten C_1 fra operationsforstærkerens inverterende indgang med en modstand hvor 100Ω kan være nok.

98 Fotodioder har et stort areal for at kunne modtage lysenergien og det medfører en kapacitet i omegnen af 100 pF.

99 Et piezoelektrisk accelerometre er opbygget som et smalt krystal mellem to ledende flader der giver cirka 500 pF.

100 Det typiske skærmbord repræsenterer en kapacitet på 100 pF for hver meter kabel.

Herunder gennemføres analysen for begge kredsløb med kondensator C_1 over indgangen til operationsforstærkeren og C_2 i tilbagekoblingen. Fokus er rettet mod overføringsfunktionen fra U_i til U_o der med to poler kan beskrives ved en resonansfrekvens $f_1 = \omega_1/2\pi$ og en dæmpning d . For at være ubetinget stabil skal dæmpningen være mindst én hvilket svarer til to reelle poler.

Overføringsfunktionen findes ved at opstille Kirchhoffs strømlov for den negative indgang U_N på operationsforstærkeren.

$$\frac{U_N - U_i}{R_1} + \frac{U_N}{1/sC_1} + \frac{U_N - U_o}{R_2} + \frac{U_N - U_o}{1/sC_2} = 0$$

Den interne knude ved U_N elimineres ved at bruge operationsforstærkerens definitions ligning hvor den positive indgang U_P er på nul. Åben-sløjfe forstærkningen A_{OL} kan tilnærmes med en integration for frekvenser over nogle få kilohertz.

$$\begin{aligned} U_o &= (U_P - U_N) A_{OL} \Rightarrow U_o = -A_{OL} U_N \\ U_P &= 0 \\ A_{OL} &= \frac{\omega_0}{\omega_0 + s} A_{DC} = \frac{2\pi GBP}{\omega_0 + s} \approx \frac{2\pi GBP}{s} \Rightarrow U_N \approx -\frac{s}{2\pi GBP} U_o \end{aligned}$$

Indsættelse og reduktion viser at forstærkningen som ventet består af faktoren $-R_2/R_1$ for en inverterende forstærker og dertil et lavpasfilter af anden orden.

$$\frac{U_o}{U_i} = -\frac{R_2}{R_1} \frac{\frac{2\pi GBP}{R_2 C_1}}{\frac{2\pi GBP}{R_2 C_1} + \left(\frac{R_1 + R_2}{R_1 R_2 C_1} + R_2 C_2 \frac{2\pi GBP}{R_2 C_1} \right) s + s^2}$$

Lavpasfiltrets karakteristiske parameter $\omega_1 = 2\pi f_1$ og dæmpningsfaktor d kan identificeres ved at sammenligne med definitionen af et anden-ordens lavpasfilter.

$$\begin{aligned} \frac{U_o}{U_i} &= -\frac{R_2}{R_1} \frac{\omega_1^2}{\omega_1^2 + 2d\omega_1 s + s^2} \Rightarrow \omega_1^2 = \frac{2\pi GBP}{R_2 C_1} \\ 2d\omega_0 &= \frac{R_1 + R_2}{R_1 R_2 C_1} + \frac{C_2}{C_1} 2\pi GBP \end{aligned}$$

Ubetinget stabilitet opnås for $d \geq 1$. For situationen uden tilbagekoblingskondensator ($C_2 = 0$) falder andet led ud og der kan defineres en højeste værdi af kondensatoren ved indgangen hvor parentesen kan sættes til én for en forstærker med strømkilde som indgangssignal. Når C_1 er for stor til at kunne ignoreres kan den nødvendige værdi af C_2 beregnes.

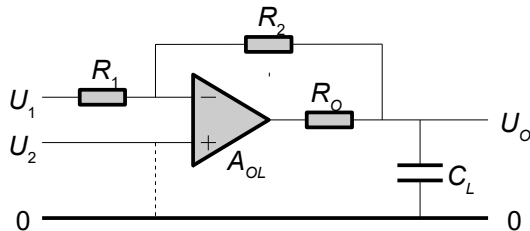
$$\begin{aligned} d = \frac{\frac{R_1 + R_2}{R_1 R_2 C_1} + \frac{C_2}{C_1} 2\pi GBP}{2\omega_1} \geq 1 \quad \Rightarrow_{C_2=0} \quad C_1 \leq \frac{R_2}{8\pi GBP (R_1 \parallel R_2)^2} = \frac{1}{8\pi GBP R_2} \left(1 + \frac{R_2}{R_1} \right)^2 \\ \Rightarrow_{C_2>0} \quad C_2 \geq \sqrt{\frac{2C_1}{\pi GBP R_2} - \frac{R_1 + R_2}{2\pi GBP R_1 R_2}} \approx \sqrt{\frac{2C_1}{\pi GBP R_2}} \end{aligned}$$

Sidste led i udtrykket for C_2 kan ignoreres når følgende ulighed er opfyldt. Eksempelvis vil det for $C_1 = 10$ pF og $R_2 = 1$ MΩ kræve at GBP er væsentligt større end 4 kHz hvilket i praksis altid vil være opfyldt.

$$\sqrt{\frac{2C_1}{\pi GBP R_2}} \gg \frac{R_1 + R_2}{2\pi GBP R_1 R_2} \Rightarrow GBP \gg \frac{1}{8\pi C_1 R_2}$$

3.5.8 Kapacitet på udgang

En operationsforstærker skal ofte drive et langt kabel hvorved det belastes af kablets kapacitet der danner en pol med operationsforstærkerens udgangsmodstand r_o . Det påvirker stabiliteten på samme måde som ved det foregående kredsløb. Modstandsværdien af r_o er på 10 til 100 Ω for en OpAmp og der skal normalt en kondensator på mere end 100 pF til for at påvirke stabiliteten.



Generelt udtryk:

$$C_L \leq \frac{1}{8\pi GBP B R_o} \quad \text{hvor} \quad B = \frac{R_1}{R_1 + R_2}$$

Forenklet udtryk:

$$C_L \leq \frac{1}{8\pi GBP R_o} \quad \text{hvor} \quad \begin{matrix} R_1 \rightarrow \infty \\ R_2 \approx 0 \end{matrix}$$

Figur 181 – 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 højst $C_L = 133 \text{ pF}$.

Polen kan undgås ved at indskyde en modstand på 10 til 100 Ω som placeres efter det punkt som R_2 tager signalet fra. Det fjerner kondensatorens pol fra tilbagelkoblingen. En anden mulighed er at inkludere forstærkning gennem de to modstande R_1 og R_2 . Det er dog et tveægget sværd for med $R_2 > 0$ vil der være en pol ved indgangen på grund af spredningskapaciteten (C_1 fra side 174).

Overføringsfunktionen findes ved Kirchhoffs strømlov for operationsforstærkerens negative indgang. Det antages 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.

$$\frac{U_N - U_I}{R_1} + \frac{U_N - U_O}{R_2} = 0 \quad \text{og} \quad U_O = -\frac{\omega_L}{\omega_L + s} \frac{2\pi GBP}{s} U_N \quad \text{hvor} \quad \omega_L = \frac{1}{R_o C_L}$$

Efter reduktion findes overføringsfunktionen der beskriver en inverterende forstærker med et anden-ordens lavpasfilter.

$$\frac{U_O}{U_I} = -\frac{R_2}{R_1} \frac{B 2\pi GBP \omega_L}{B 2\pi GBP \omega_L + \omega_L s + s^2} \quad \text{hvor} \quad B = \frac{R_1}{R_1 + R_2}$$

Anden-ordens relationen skrives på standardformen hvorved grænsefrekvensen $\omega_0 = 2\pi f_0$ og dæmpningen d kan identificeres.

$$\frac{U_O}{U_I} = -\frac{R_2}{R_1} \frac{\omega_0^2}{\omega_0^2 + 2d\omega_0 s + s^2} \Rightarrow \begin{matrix} \omega_0^2 = B 2\pi GBP \omega_L \\ 2d\omega_0 = \omega_L \end{matrix}$$

For ubetinget stabilitet kræves en dæmpning på mindst én.

$$d = \frac{\omega_L}{2\omega_0} \geq 1 \Rightarrow \frac{\omega_L}{2\sqrt{\beta 2\pi GBP \omega_L}} \geq 1 \Rightarrow \frac{\omega_L}{B 2\pi GBP} \geq 4$$

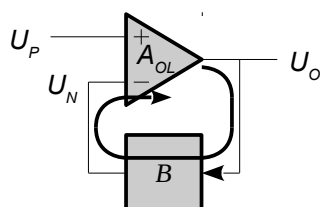
Herfra bestemmes kravet til grænsefrekvensen af polen hvor $\omega_L = 2\pi f_L$.

$$f_L \geq 4 B GBP \Rightarrow \frac{1}{2\pi R_o C_L} \geq 4 B GBP \Rightarrow C_L \leq \frac{1}{8\pi GBP B R_o}$$

Parameteren B kan ikke komme over én så formlen kan benyttes for at beregne den lavest tilladte værdi af kondensatoren på udgangen ved at sætte $B = 1$.

3.5.9 Fasemargin

Ved frekvensområdet over den dominerende pol er der omvendt proportionalitet mellem frekvens og forstærkning 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 signalet fra tilbakoblingens netværk med amplituden BU_O kan understøtte en oscillation.



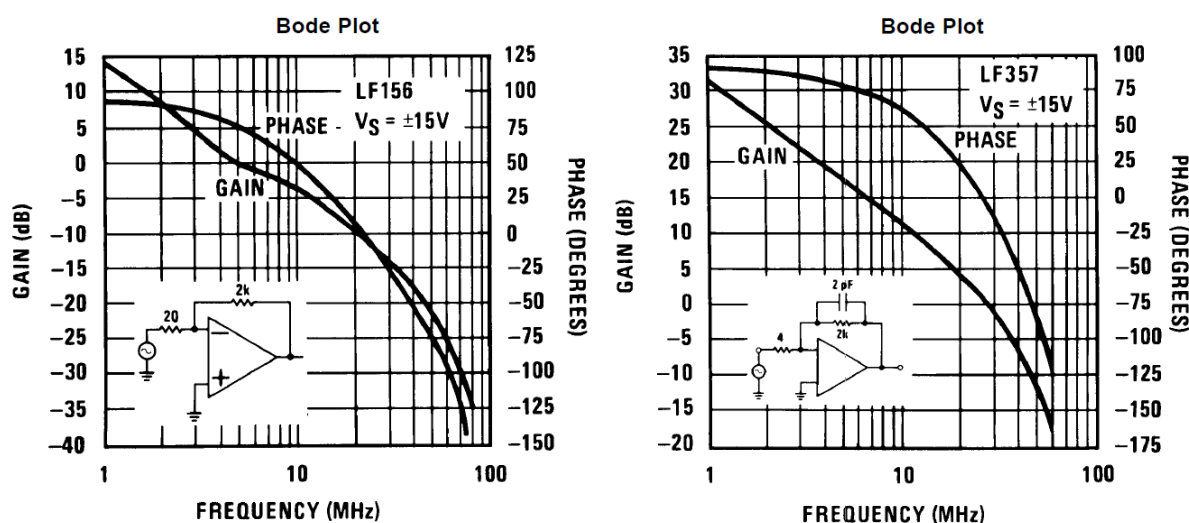
$$A_s = -BA_{OL} = |A_s| \exp(j\theta_s)$$

Stabil:	$ A_s < 1$	ved $\theta_s = 360^\circ$
Ustabil:	$ A_s = 1$	ved $\theta_s = 360^\circ$
Oscillator:	$ A_s > 1$	ved $\theta_s = 360^\circ$

Figur 182 – 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° .

Betingelsen for stabilitet blev givet af Nyquist som en forstærkning A_s under én ved den frekvens hvor fasedrejet er 360° . En forstærker der opfylder Nyquist kriteriet vil ikke oscillere, men den kan være længe om at falde til ro efter en påvirkning. Stabilitet oplyses ofte som en fasemargin (*phase margin*) der er den rest af fasedrejning der mangler for at nå 360° ved en forstærkning på 0 dB. Som en tommelfingerregel bør fasemarginen være mindst 45° .

Operationsforstærkeres åben-sløjfe forstærkning og fasemargin vises herunder for et par af de gamle travere LF356 og LF357 med JFET i indgangen.¹⁰¹



Figur 183 – Åben-sløjfe-forstærkning og fasemargin for to operationsforstærkere (Texas).

For LF356 er $GBP = 5$ MHz og med en indstillet forstærkning på 100 gange falder forstærkningen af over 50 kHz hvor fasemargin som forventet er 90° . Ved 5 MHz er forstærkningen aftaget til én og fasemargin er reduceret til 75° så de interne poler har tilføjet 15° , men da forstærkningen ved højere frekvenser er under én er det uden betydning.

For LF357 er $GBP = 20$ MHz og med en indstillet forstærkning på 500 gange falder forstærkningen af over 40 kHz hvor fasemargin som forventet er 90° . Ved 25 MHz er dens fasemargin 0° mens forstærkningen er lidt over én (1 dB). Det er et tegn på at forstærkeren ikke kan tilbagekobles til en forstærkning på én uden at oscillere. LF357 er da også specificeret til $A_{CL} > 5$.

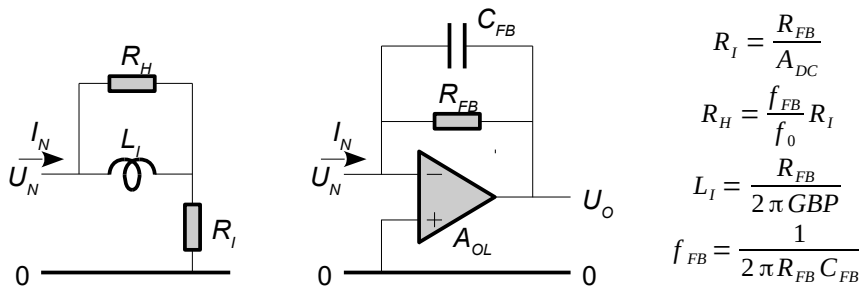
¹⁰¹Kurven skriver LF156 der er en udgave for militært brug. Det vil sige at den er indkapslet i keramik i modsætning til LF356 der sælges i et plasthus for det kommercielle marked med temperaturområdet 0 til 70° . Data gælder for LF156, LF256 og LF356 og tilsvarende gælder for den hurtigere LF157, LF257 og LF357.

3.6 Ind- og udgangsimpedans

Det forudsættes normalt at den ikke-inverterende forstærker har en så høj indgangsmodstand at den kan ignoreres, at den inverterende forstærkers virtuelle nul har en indgangsmodstand 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 .

3.6.1 Inverterende forstærker

Indgangsimpedansen ved den virtuelle jord er givet ved impedansen af tilbagkoblingen divideret med åben-sløjfe forstærkningen A_{OL} . Ved lave frekvenser er værdien under $1\ \Omega$ ved moderat værdi af tilbagkoblingens modstandsværdi og en DC forstærkning A_{DC} på en million. I den aktive del af frekvensområdet er operationsforstærkeren dog integrerende og indgangsimpedansen bliver en differentieret version af tilbagkoblingen ($Z_{FB} = 1/sC_{FB}$). En kondensator i tilbagkoblingen ($Z_{FB} = 1/sC_{FB}$) giver en reel værdi over indgangen og ved en modstand i tilbagkoblingen ($Z_{FB} = R_{FB}$) er værdien proportional med frekvensen hvilket er en selvinduktion i parallel over indgangen.



Figur 184 – En almindelig opstilling med begrænsning af båndbredden vil have en kompleks indgangsimpedans der er induktiv i et frekvensområde.

Med $R_{FB} = 1\ \text{M}\Omega$, en DC forstærkning på $A_{DC} = 10^6$ og $GBP = 1\ \text{MHz}$ 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 $10\ \Omega$ allerede ved $100\ \text{Hz}$ så indgangen er ikke virtuel nul ret langt op i frekvens.

Åben-sløjfe forstærkningen udtrykkes herunder ved DC forstærkningen A_{DC} og en dominerende pol $f_0 = \omega_0/2\pi$ der almindeligvis er omkring $100\ \text{Hz}$. Indført i udtrykket beregnes nu en reel og en imaginær del af impedansen.

$$Z_I = \frac{R_{FB}}{A_{OL}} \quad \text{og} \quad A_{OL} = \frac{\omega_0}{\omega_0 + s} A_{DC} \Rightarrow Z_I = \frac{R_{FB}}{A_{DC}} \frac{\omega_0 + s}{\omega_0} = \frac{R_{FB}}{A_{DC}} \left(1 + \frac{s}{\omega_0}\right)$$

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 $GBP = f_0 A_{DC}$.

$$Z_I = \frac{R_{FB}}{A_{DC}} + s \frac{R_{FB}}{A_{DC} \omega_0} = R_I + sL_I \Rightarrow R_I = \frac{R_{FB}}{A_{DC}}, \quad L_I = \frac{R_{FB}}{2\pi GBP}$$

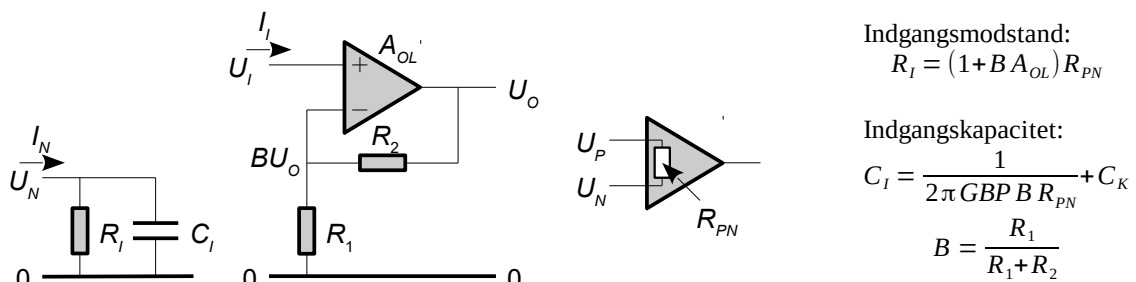
En parallellforbindelse af en modstand og en kondensator i tilbagkoblingen er hyppigt anvendt for båndbegrænsning. Indgangsimpedansen bestemmes ved at indsætte parallellforbindelsen af R_{FB} og $1/sC_{FB}$ i udtrykket Z_{FB}/A_{OL} og A_{OL} beskrives igen ved A_{DC} og polen ved ω_0 .

$$Z_I = \frac{\omega_0 + s}{A_{DC} \omega_0} \frac{R_{FB}/sC_{FB}}{R_{FB} + 1/sC_{FB}} = \frac{R_{FB}}{A_{DC}} \frac{\omega_0 + s}{\omega_0} \frac{1/R_{FB} C_{FB}}{s + 1/R_{FB} C_{FB}} = \frac{R_{FB}}{A_{DC}} \frac{\omega_0 + s}{\omega_0} \frac{\omega_{FB}}{s + \omega_{FB}}$$

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).

3.6.2 Ikke-inverterende forstærker

For den ikke-inverterende forstærker er indgangsimpedansen givet ved operationsforstærkerens indre modstand R_{PN} mellem de to indgange og dertil produktet BA_{OL} der angiver hvor kraftigt der er tilbagekoblet. En operationsforstærker har en indgangsmodstand langt over 1 M Ω selv ved en lav værdi af R_{PN} så indgangsmodstanden R_I vil oftest kunne ignoreres. Over den dominerende pol er der omvendt proportionalitet med frekvensen hvilket kan beskrives som en kapacitet C_I i parallel over indgangen. Dertil kommer den uundgåelige kobling til omgivelserne C_K på typisk 5 pF.



Figur 185 – 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 nær ved nul så strømmen i indgangen også bliver nær ved nul.

Med $GBP = 1$ MHz, $B = 1$, $A_{OL} = 10^6$ og $R_{PN} = 10$ k Ω er indgangsmodstanden $R_I \approx 10$ G Ω og kapaciteten over indgangen er $C_I = 16$ pF så impedansen er tilbage på 10 k Ω ved 1 MHz.

Indgangstrinnet på en operationsforstærker kan repræsenteres ved en modstandsværdi mellem de to indgange (R_{PN}) med en værdi startende ved 10 k Ω for bipolære operationsforstærkere med høj strøm i indgangstrinnet og stigende mod 10¹² Ω for felt-effekt transistorer.

Indgangsimpedansen beregnes ved at føre en strøm I_I ind på indgangen og finde 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 brøkdelen B .

$$U_I = R_{PN} I_I + U_N = R_{PN} I_I + B U_O$$

Udgangsspændingen er givet af differensen U_P og U_N gange med åben-sløjfe forstærkningen A_{OL} og da differensen er givet ved $R_{PN}I_I$ gives 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.

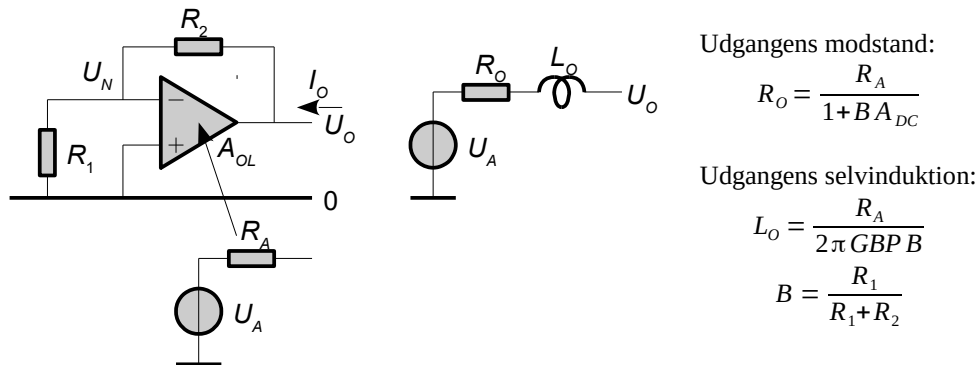
$$R_I = \frac{U_I}{I_I} = (1 + B A_{OL}) R_{PN}$$

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.

$$Z_I = (1 + B \frac{\omega_0}{\omega_0 + s} A_{DC}) R_{PN} \approx B \frac{\omega_0}{s} A_{DC} R_{PN} = \frac{1}{s C_I} \Rightarrow C_I = \frac{1}{2\pi GBP B R_{PN}}$$

3.6.3 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 mod høje frekvenser vil delvist ophæve korrektionen og det kan ses som en induktiv udgangsimpedans.



Figur 186 – Udgangsimpedansen på en forstærker reduceres på grund af tilbagekoblingen, men den dominerende pol i operationsforstærkeren gør den resulterende impedans induktiv.

Ved $R_A = 50 \Omega$, $B = 1$, $A_{DC} = 10^5$ 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 $B = 0,1$ 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 B af udgangsspændingen.

$$U_A = (U_P - U_N) A_{OL} = -U_N A_{OL} = -B U_O A_{OL}$$

Udgangsimpedansen findes ved at lede en 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.

$$U_O = R_A I_O + U_A = R_A I_O - B A_{OL} U_O \Rightarrow U_O (1 + B A_{OL}) = R_A I_O$$

Udgangsmodstanden bliver reduceret med tilbagekoblingen udtrykt ved $1 + B A_{OL}$.

$$Z_O = \frac{U_O}{I_O} = \frac{R_A}{1 + B A_{OL}}$$

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 $1 + B A_{DC}$ 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.

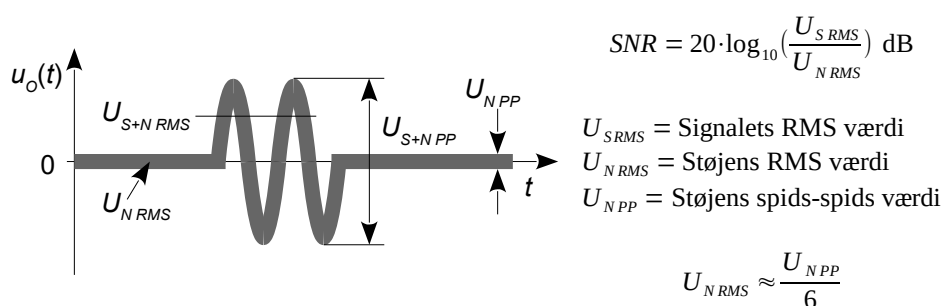
$$Z_O = \frac{R_A}{1 + B \frac{\omega_0}{\omega_0 + s} A_{DC}} = R_A \frac{\omega_0 + s}{(1 + B A_{DC}) \omega_0 + s} \stackrel{s \gg \omega_0}{\approx} R_A \frac{s}{B A_{DC} \omega_0} \Rightarrow L_O = \frac{R_A}{2\pi GBP B}$$

3.7 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 186 eller til sammenfatningen side 196.

Støj er uønskede signaler som generer os mennesker. Det kan maskere et ønsket signal så vi mister indholdet af en samtale eller nattesøvnen forstyrres. Ved et digitalt transmissionssystem vil støj kunne ødelægge data og ved måling vil støjen gøre målingen usikker.

Som ingeniører har vi brug for at sætte talværdier på begreberne. Et signals signal-støj forhold *SNR* (*signal-to-noise ratio*) er forholdet mellem signalets og støjens RMS niveauer og udtrykkes i decibel (dB). Signalets dynamikområde er forholdet mellem de kraftigst og svagest mulige RMS niveauer. For et menneske giver smertegrænsen og høretærsklen et dynamikområde på $10^7:1$ eller 140 dB mellem det højeste og laveste lydniveau. Almindelig tale har et lydtryk omkring 60 dB over høretærsklen målt ved ørekanalen og med støj fra omgivelserne på 40 dB er signal-støj forholdet på 20 dB der er tilstrækkeligt til at forstå tale næsten fejlfrit.^{102,103}



Figur 187 – Ethvert signal består af det ønskede signal overlejret en evigt fluktuerende støj.

Der er flere kilder til støj i elektriske kredsløb. Der genereres en termisk støj i enhver elektrisk leder på grund af omgivelsestemperaturen der får elektronerne til at sitre hvilket giver skiftende fordeling af ladning i lederen. Der genereres støj i en halvleder ved at strømmens bæres af et endeligt antal elektroner. Andre støjkilder er omrokering af elektrostatiske ladninger i halvlederen og ændringer i krystalstrukturen som begge dominerer ved lave frekvenser og kaldes $1/f$ -støj eller rosa støj.

Eksterne støjkilder kan koble elektromagnetisk til elektronikken som den knurrende lyd i fra en tændt mobiltelefon i nærheden af en PC højttaler eller gnister fra en lyskontakt. Dertil kan de elektroniske komponenter have en piezoelektrisk effekt der omsætter en mekanisk vibration til spænding. Endelig kan signalet selv indeholde støj som fx en optagelse af en koncert med publikumsstøj. Støj er uønskede signaler og inkluderer en langsom drift i offset spænding og biasstrøm selv om man normalt vil analysere de to situationer hver for sig.

Bestemmelse af RMS værdien kræver måling over tid da signalet er evigt fluktuerende. Vi skal derfor betragte signalet over et tidsrum der inkluderer mange perioder af signalet og for de typiske situationer er ét sekund tilstrækkeligt (se første kapitel). Det giver en middelværdi ved en given båndbredde i måleapparatet og i analogi med synligt lys kaldes det for hvid støj (*white noise*) hvis der er en konstant middelværdi ved en given båndbredde uanset hvor centerfrekvensen ligger.

102 Smertegrænsen ligger cirka 140 dB over høretærsklens 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 højere 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.

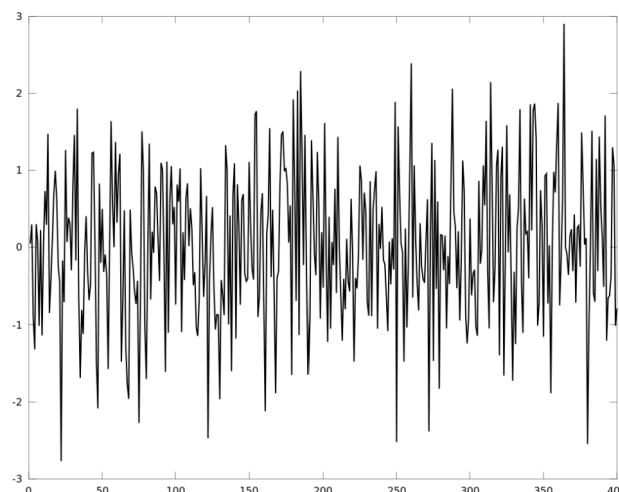
103 Se <http://www.torean.dk/artikel/Elektroakustik.pdf> for en uddybning.

3.7.1 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-funktion. For at analysere støjens niveau benyttes et statistisk værktøj for at estimere nogle parametre der beskriver støjen. Støjsignalets spænding måles over et tidsrum og beskrives ved middelværdien (*mean value*) og spredningen (*spread*). Middelværdien er spændingen som støjen svinger omkring og spredningen er et mål for hvor kraftige udsving fra middelværdien der kan forventes. Spredningen er i ingeniørtermer signalets RMS værdi.

Herunder vises et støjsignal hvor en matematisk funktion beregner tilfældige værdier der opfylder nogle betingelser for tilfældigheden som svarer til standardfordelingen (se senere). Der er egentligt 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.¹⁰⁴



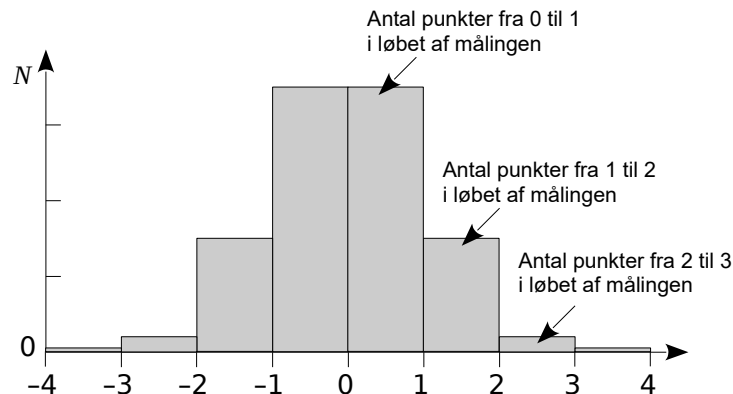
Figur 188 – Støjen er her afbildet ved hjælp af et MATLAB program hvor `randn(1:400)` giver 400 normalfordelte værdier med variansen én og middelværdien nul.

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 . Spidserne af signalet kan godt nå længere ud, men der vil være langt mellem disse 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.

¹⁰⁴Der findes algoritmer som gennemløber en binær sekvens der tolket som et analogt signal opfører sig som støj. Det kaldes for *pseudo-random noise* og dannes af *maximum-length sequence* (MLS). Det 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 båndbredde og er deterministisk. Forskellen fra et "ægte" støjsignal er at forløbet er kendt og kan repeteres.

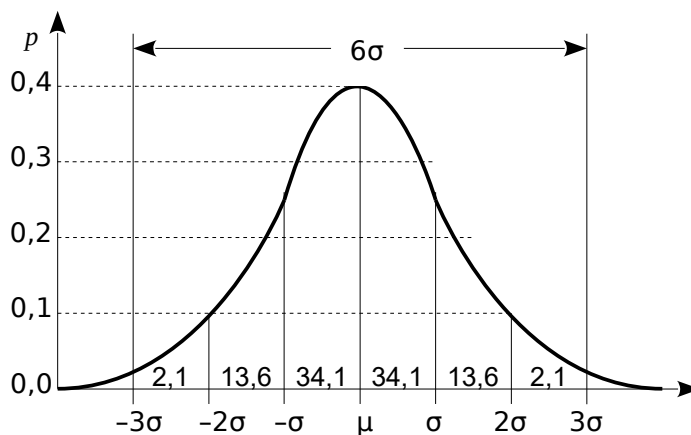
For at beskrive dette uregelmæssige forløb kan man måle antallet af punkter der rammer indenfor et interval fra 0 til 1 derefter antallet af punkter fra 1 til 2 og så videre. Når måleserien stoppes kan man angive de optalte værdier i et histogram som skitseret herunder. Søjlerne bredde angiver et måleområde 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.

Hvis målepunkternes antal i hver søjle divideres med det samlede antal punkter i hele måleserien 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 189 – 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 smallere går billedet gradvist imod en kontinuert 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 spredningen σ .



Middelværdi (DC, average):

$$U_{AVE} = \mu = \frac{1}{T} \int_0^T u dt \quad \text{hvor} \quad T > \frac{3}{f_0}$$

Effektivværdi (spredning, RMS):

$$U_{RMS} = \sigma = \sqrt{\frac{1}{T} \int_0^T (u - \mu)^2 dt}$$

Sandsynlighed:

$$P(|u| < 3\sigma) \approx 99,7 \%$$

$$P(|u| > 3\sigma) \approx 0,3 \%$$

Tilnærmet relation:

$$U_{RMS} \approx \frac{U_{PP}}{6}$$

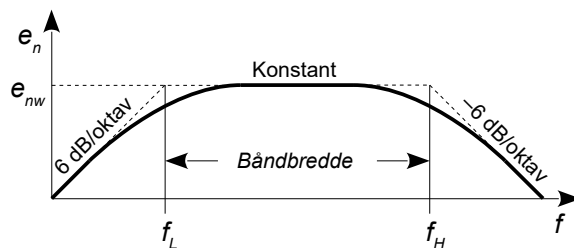
Figur 190 – Standardfordelingen med middelværdien μ der er lig med DC værdien som et signal svinger omkring. Spredningen σ er lig med RMS værdien. Ordinaten angiver tætheden og for at finde sandsynligheden skal der integreres over et interval på abscissen. Ved integration over intervaller lig med variansen bliver procentsatsen som vist i figuren (afrundet til én decimal).

Sandsynligheden er 34 % for at et signal ligger indenfor én standardafvigelse væk fra middelværdien så sandsynligheden er 68 % for at signalet ligger i intervallet $\pm U_{RMS}$. Sandsynligheden stiger til 95 % for at signalet ligger indenfor $\pm 2 U_{RMS}$ og til 99,7 % ved $\pm 3 U_{RMS}$ og er derfor nær nul for mere end $\pm 3 U_{RMS}$.

3.7.2 Støjkilder

Det er umuligt at bygge et elektronisk kredsløb uden støj for støjen genereres af de elektroner der bærer strømmen i enhver elektrisk leder. Elektronerne vibrerer konstant på grund af temperaturen og der optræder derfor en evigt fluktuerende omfordeling af ladning som kortvarigt producerer en spændingsforskel over en leders to ende. Elektronernes vibration er funktion af temperaturen og svinger kraftigere ved stigende temperatur hvor den fluktuerende spænding giver et større bidrag. Det kaldes for termisk støj (*thermal noise* eller *Johnson noise*).¹⁰⁵

Formlen i illustration giver spændings-støjtætheden e_n fra en modstand med værdien R og ved temperaturen T der er den absolutte temperatur i kelvin (K). For omgivelsestemperaturen 27°C er det 300 K og en modstand på 1 kΩ genererer her en støjtæthed på $e_n = 4 \text{ nV}/\sqrt{\text{Hz}}$.



Termisk støj:

$$e_n = \sqrt{4kTR} \quad \text{hvor} \quad k = 1,38 \cdot 10^{-23} \text{ J/C} \\ T = T_{\text{Celsius}} + 273 \text{ K}$$

Haglstøj:

$$i_n = \sqrt{2q_0 I_{DC}} \quad \text{hvor} \quad q_0 = 1,6 \cdot 10^{-19} \text{ C}$$

Støjspænding og støjstrøm (RMS):

$$U_n = e_n \sqrt{B} \quad \text{hvor} \quad B = f_H - f_L \\ I_n = i_n \sqrt{B}$$

Addition af støjkilder:

$$U_{nTOT} = \sqrt{U_{n1}^2 + U_{n2}^2 + \dots + U_{nN}^2}$$

Figur 191 – 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.

Elektronerne giver anledning til endnu en støjmekanisme på grund af deres ladning. Når en strøm løber i en transistor er strømmens styrke givet ved det gennemsnitlige antal elektroner der løber indenfor et tidsrum. Elektronens ladning er konstant så værdien af strømmen vil indenfor et kort tidsrum vil fluktuere overfor den gennemsnitlige værdi der måles over lang tid. Det kan opfattes som en kombination af en konstant strøm I_{DC} og en fluktuerende strøm i_n der er strømtætheden og værdien beregnes som vist i illustrationen. Det kaldes for haglstøj (*shot noise*). Formlen gælder for halvledere, men ikke for strømmen i en elektrisk ledning.^{106,107}

En strømstyrke I_{DC} på 1 mA vil generere en strømtæthed på $i_n = 18 \text{ pA}/\sqrt{\text{Hz}}$. Hvis strømmen løber gennem en modstand på 1 kΩ vil det generere en spændings-støjtæthed på $18 \text{ nV}/\sqrt{\text{Hz}}$. Der kan derfor være mange støjbidrag i en elektronisk opstilling.

Støjens værdi er uafhængig af hvor i frekvensspektret der måles ved både termisk støj og haglstøj. Det kaldes for hvid støj (*white noise*). RMS værdien af støjen er kun funktion af den båndbredde B der benyttes ved målingen og man omregner fra støjtæthed til en RMS værdi ved at gange e_n eller i_n med kvadratroden af båndbredden der normalt kan antages til kredsløbets –3 dB båndbredde. For en operationsforstærker er forholdene mere komplicerede og det behandles side 190.

Det er uden betydning hvor båndbreddens grænser er placeret for det er kun differensen mellem grænserne af frekvensbåndet der 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 den er i THz området af kvantemekaniske årsager.

¹⁰⁵ https://en.wikipedia.org/wiki/Johnson%E2%80%93Nyquist_noise.

¹⁰⁶ https://en.wikipedia.org/wiki/Shot_noise.

¹⁰⁷ Begrebet haglstøj blev introduceret af Walter Shottky i 1918 ved studier af elektronstrømmen i et radorør.

3.7.3 Beregning af støjbidrag

Der er sædvanligvis flere støjklender i en forstærker og de kan alle føres frem til udgangen hvor de RMS adderes for en beregning af den totale støjspænding. Beregning af støj tillader en opdeling i dominerende og ikke-dominerende støjbidrag hvor de sidste blot ignoreres så beregningen kan simplificeres.

Støjtætheden beregnes fra en middeleffekt P_n der er målt indenfor båndbredden B . Det kan relateres til spænding- eller strømtæthed ved at benytte udtrykkene for effekten af en spænding eller en strøm. RMS værdien af en støjspænding U_n (eller strøm I_n) kan tilsvarende bestemmes når støjtætheden e_n (eller i_n) og båndbredden B af transmissionskanalen er kendt.

$$\begin{aligned} \text{Spændingsstøjtæthed: } \frac{P_n}{B} &= \frac{U_n^2}{RB} = \frac{e_n^2}{R} \Rightarrow e_n = \frac{U_n}{\sqrt{B}} \Rightarrow U_n = e_n \sqrt{B} \\ \text{Strømsstøjtæthed: } \frac{P_n}{B} &= \frac{I_n^2 R}{B} = i_n^2 R \Rightarrow I_n = i_n \sqrt{B} \Rightarrow i_n = \frac{I_n}{\sqrt{B}} \end{aligned}$$

Hvis to eller flere støjsignaler skal føres til det samme sted så er det middeleffekterne der skal adderes og det leder til den såkaldte RMS addition af støjklender. Akronymet RMS står for præcis det der sker nemlig *root mean square*. Det leder til en meget praktisk regel for addition af støj idet støjbidrag 6 dB under et givet niveau ikke har nogen større betydning og kan ignoreres. Reglerne kommer i spil ved beregning af støjen i forstærkere hvor der er flere støjklender.

Addition af støj fra flere klender benytter middeleffekten af støjen P_{TOT} der også kan skrives som U_{TOT}^2/R hvorfra U_{TOT} kan bestemmes.

$$P_{TOT} = P_1 + P_2 + \dots \Rightarrow \frac{U_{TOT}^2}{R} = \frac{U_1^2}{R} + \frac{U_2^2}{R} + \dots \Rightarrow U_{TOT}^2 = U_1^2 + U_2^2 + \dots \Rightarrow U_{TOT} = \sqrt{U_1^2 + U_2^2 + \dots}$$

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. Det betyder at svage støjklender trygt kan ignoreres ved et estimat.

$$U_{TOT} = \sqrt{U_1^2 + (0,5 U_1)^2} = \sqrt{(1 + 0,25) U_1^2} = 1,12 U_1$$

Addition af mange støjklender hæver det samlede støjniveau med kvadratroden af antallet hvis støjklenderne er lige kraftige. Det kan benyttes ved både analoge og digitale mixerpulte for det gælder også for støjen fra afrundingen i de digitale algoritmer. Ved 64 kanaler vil støjniveauet hæves 8 gange som svarer til en reduktion i signal-støj forholdet på 18 dB.

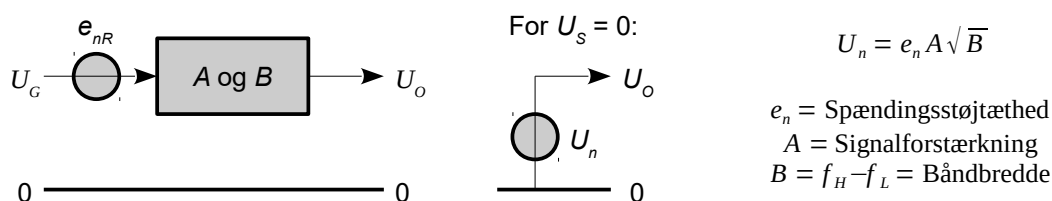
$$U_{TOT} = \sqrt{U_1^2 + U_1^2 + \dots} = \sqrt{N} U_1$$

Reglen for addition kan udnyttes til at øge signal/støj-forholdet ved kommunikation med en fjern satellit. Det samme signal sendes 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 støjen hele tiden skifter. Det forøger SNR værdien 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 mellemøret som reaktion på et eksternt lydssignal hvor der summeres over mindst 1000 testpulser.

Operationsforstærkerens støj er relativt kompliceret at beregne, men det er ofte fuldt tilstrækkeligt at betragte dens støj for hvid og bruge de viste formler for en analyse. I det følgende beskrives en støjmodel for elektroniske kredsløb hvor støjens niveau beregnes og føres frem til udgangen. Det tillader vurdering af om signal-støj forholdet for et givet kredsløb er tilstrækkeligt til opgaven.

3.7.4 Beregningseksempler

En model af et elektronisk kredsløb vil meget ofte have støjkilden placeret på indgangen som en spændingsstøjtæthed e_n i serie med indgangssignalet og kredsløbet beskrives ved forstærkning A og båndbredden B for det frekvensområde der kan passere.



Figur 192 – En støjkilde e_{nR} på indgangen af et elektronisk kredsløb bliver forøget med kredsløbets forstærkning i pasbåndet A og frekvensbegrænses til båndbredden B 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 U_S kredsløbet bygges for at forstærke og det eneste der efterfølgende kan gøres for at sænke støjens niveau er at reducere båndbredden. Der er en kvadratrods inde i billedet så båndbredden skal reduceres 100 gange for at nå en reduktion i støjens værdi på 10 gange. Det er kun en acceptabel metode når signalet varierer så langsomt at den lave båndbredde ikke i sig selv udgør en begrænsning. Det kan være tilfældet ved måling af temperatur med PT100 eller termokobler og måske ved en vægt med strain gauges.

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 kilden, ved en radiomodtager kan man benytte en antenne med flere direktorer så signalet bliver kraftigere og ved optisk signaltransmission må man overveje at øge senderens effekt eller fokusere lyset så signalets amplitude bliver tilstrækkelig ved modtageren.

En operationsforstærkers støjtæthed e_{nA} kan ikke beregnes som støjen fra en modstand så i stedet benyttes databladets specifikation af støjtætheden. Herunder vises eksempler af støjtætheden for tre modstande e_{nR} og fire operationsforstærkere e_{nA} .

Komponent	Værdi	Spændingsstøjtæthed	Formel
Modstand	10 Ω	$e_{nR} = 0,4 \text{ nV}/\sqrt{\text{Hz}}$	$e_n = \sqrt{4kTR}$
Modstand	1 k Ω	$e_{nR} = 4 \text{ nV}/\sqrt{\text{Hz}}$	
Modstand	100 k Ω	$e_{nR} = 40 \text{ nV}/\sqrt{\text{Hz}}$	
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 applikation 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.¹⁰⁸

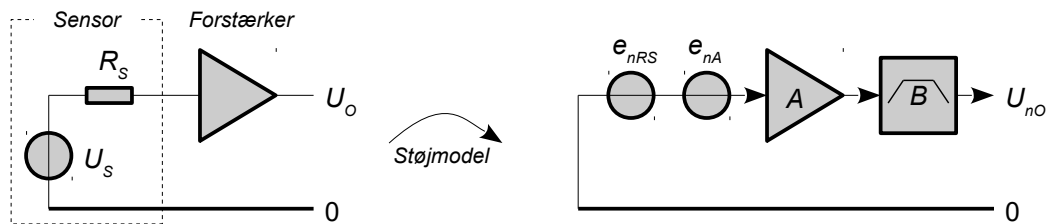
De følgende afsnit viser hvordan støj kan beregnes for nogle kredsløb.

¹⁰⁸Ved lydoptagelse kræves ofte et signal/støj-forhold på mere end 80 dB mens måling af et signals frekvens kan nøjes med et signal/støj-forhold omkring 0 dB.

3.7.5 Forstærkning af svage signaler

Forstærkningen af et svagt signal fra en sensor kan repræsentere en lang række af opgaver som temperaturmåling med PT100 eller et termoelement, mekanisk påvirkning med strain gauge eller accelerometer, radiomodtagelse fra en antenne 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, men både sensor og forstærker genererer støj som tilsammen danner et støjgulv og det er her formålet at estimere hvor dette uønskede niveau af egenstøj befinder sig.



Figur 193 – 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 støjklenderne og de parametre der har betydning for beregning af RMS værdien af støjen. Sensoren vil her blive beskrevet ved dens indre modstand R_s der genererer termisk støj. Forstærkeren beskrives ved dens støjtæthed der er specificeret i databladet. De to støjklender sidder i serie så bidragene forstærkes begge med faktor A og begrænses af båndbredden B . Proceduren den samme for begge støjklender og reglen om superposition tillader en RMS addition af bidragene ved udgangen.

Algoritmen skal eksemplificeres ved måling af temperatur med en termokobler. 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 god margin. Det ses også at operationsforstærkeren 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 mV/ $^\circ\text{C}$ så der skal forstærkes med $A = 244$ gange. Termoelementets indre modstand er $R_s = 10 \Omega$. Operationsforstærkeren LF356 oplyses til en egenstøj på $e_{nA} = 12 \text{ nV}/\sqrt{\text{Hz}}$ og $\text{GBP} = 5 \text{ MHz}$. Ved den indstillede forstærkning bliver båndbredden $B = 20,5 \text{ kHz}$.

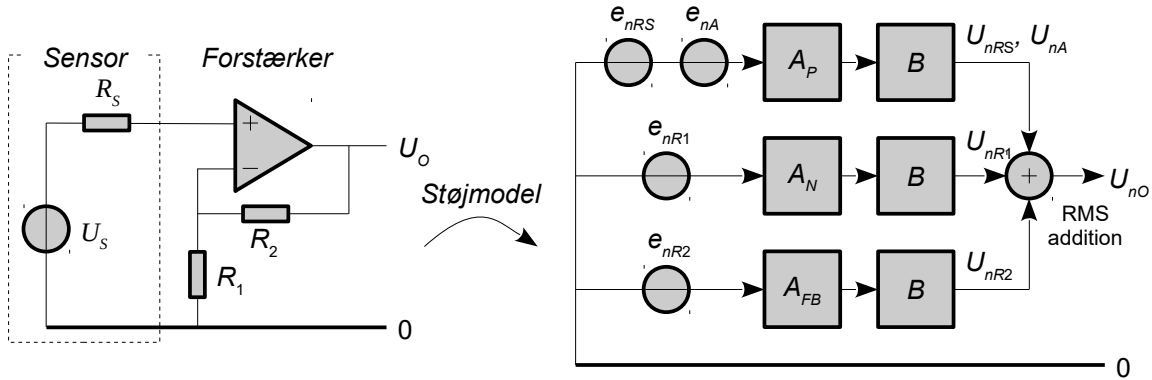
$$\begin{aligned}
 \text{Modstand: } U_{nRS} &= e_{nRS} A \sqrt{B} & e_{nRS} &= \sqrt{4kTR_s} = 0,4 \text{ nV}/\sqrt{\text{Hz}} \\
 \text{Forstærker: } U_{nA} &= e_{nA} A \sqrt{B} & e_{nA} &= \text{Datablad} = 12 \text{ nV}/\sqrt{\text{Hz}} \\
 \text{Total støj: } U_{no} &= \sqrt{U_{nRS}^2 + U_{nA}^2} & \text{hvor } A &= \frac{U_o}{U_i} = \frac{10 \text{ mV}/^\circ\text{C}}{41 \mu\text{V}/^\circ\text{C}} = 244 \\
 & & B &= \frac{\text{GBP}}{A} = \frac{5 \text{ MHz}}{244} = 20,5 \text{ kHz}
 \end{aligned}$$

Data indsættes og giver en RMS værdi af støjen på udgangen til 0,4 mV og viser samtidigt at operationsforstærkeren dominerer i støjberegningen.

$$\begin{aligned}
 U_{nRS} &= (0,4 \text{ nV}/\sqrt{\text{Hz}}) \cdot 244 \cdot \sqrt{20,5 \text{ kHz}} = 0,014 \text{ mV} \\
 U_{nA} &= (12 \text{ nV}/\sqrt{\text{Hz}}) \cdot 244 \cdot \sqrt{20,5 \text{ kHz}} = 0,419 \text{ mV} \\
 U_{no} &= \sqrt{(0,014 \text{ mV})^2 + (0,419 \text{ mV})^2} = 0,419 \text{ mV}
 \end{aligned}$$

3.7.6 Forstærker med flere støjklider

Ved støjkritiske applikationer kan man ikke ignorere de modstande der indstiller forstærkningen så det følgende eksempel viser hvordan støjanalysen kan udvides. Støjkliderne har ikke den samme forstærkning frem til udgangen så det er nødvendigt at bestemme forstærkningen for hver enkelt støjkilde. Derefter føres signalerne frem til udgangen hvor de RMS adderes. Første eksempel viste at LF356 er ret støjende så her vælges den støjsvage NE5532.



Figur 194 – Sensoren, operationsforstærkeren og modstandene er alle støjklider. Støjkliderne forstærkes, båndbredden begrænses og bidragene RMS adderes ved udgangen.

Forstærkningen af signalet og støj fra R_S er $A_P = 1 + R_2/R_1$ med de viste modstande. Støjen fra R_1 er $A_N = -R_2/R_1$ da støjkliden sidder i serie med R_1 . Støjen fra R_2 forstærkes ikke op som det vises se side 168 så $A_{FB} = 1$. Analysen illustrerer tommelfingerreglen om at modstanden R_1 ikke bør være større end signalkildens indre modstand R_S ved støjsvage kredsløb.

Operationsforstærkeren bør om muligt ikke støje mere end R_S så man bør overveje den ekstremt støjsvage AD797.

Sensorens indre modstand er her $R_S = 100 \Omega$. Signalet $U_S = 10 \text{ mV}$ skal løftes op til $U_O = 1 \text{ V}$ hvortil kræves $A = 100$. NE5532 har $e_{nA} = 5 \text{ nV}/\sqrt{\text{Hz}}$ og $GBP = 5 \text{ MHz}$.

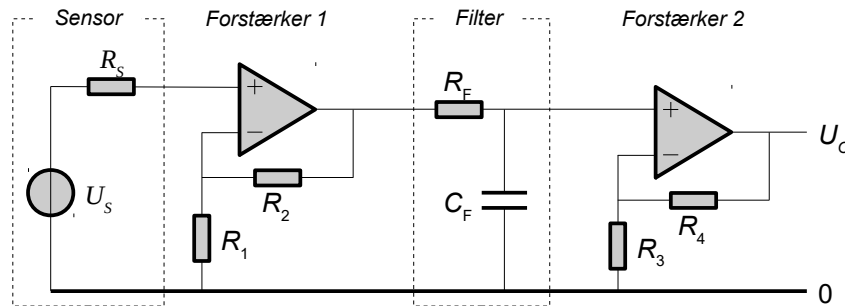
Modstand R_S	$U_{nRS} = e_{nRS} A_P \sqrt{B}$	$e_{nRS} = \sqrt{4kTR_S} = 1,3 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_1	$U_{nR1} = e_{nR1} A_N \sqrt{B}$	$e_{nR1} = \sqrt{4kTR_1} = 4,1 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_2	$U_{nR2} = e_{nR2} \sqrt{B}$	$e_{nR2} = \sqrt{4kTR_2} = 40,7 \text{ nV}/\sqrt{\text{Hz}}$
Forstærker	$U_{nA} = e_{nA} A_P \sqrt{B}$	$e_{nA} = \text{Datablad} = 5 \text{ nV}/\sqrt{\text{Hz}}$
Total støj	$U_{nO} = \sqrt{U_{nRS}^2 + U_{nA}^2 + U_{nR1}^2 + U_{nR2}^2}$	hvor $A_P = 1 + \frac{R_2}{R_1} = 101$ $A_N = -\frac{R_2}{R_1} = -100$
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 mV og signal/støj-forholdet bliver 73,6 dB der er i underkanten for en krævende lydoptagelse, men rigeligt for en temperaturmåling.

$$\begin{aligned}
 U_{nS} &= (1,3 \text{ nV}/\sqrt{\text{Hz}}) \cdot 101 \cdot \sqrt{99 \text{ kHz}} = 0,041 \text{ mV} \\
 U_{nR1} &= (4,1 \text{ nV}/\sqrt{\text{Hz}}) \cdot 100 \cdot \sqrt{99 \text{ kHz}} = 0,129 \text{ mV} \\
 U_{nR2} &= (40,7 \text{ nV}/\sqrt{\text{Hz}}) \cdot 1 \cdot \sqrt{99 \text{ kHz}} = 0,013 \text{ mV} \\
 U_{nA} &= (5 \text{ nV}/\sqrt{\text{Hz}}) \cdot 101 \cdot \sqrt{99 \text{ kHz}} = 0,159 \text{ mV} \\
 U_{nO} &= \sqrt{(0,041)^2 + (0,129)^2 + (0,013)^2 + (0,159)^2} = 0,210 \text{ mV} \\
 SNR &= 20 \cdot \log_{10} \left(\frac{1 \text{ V}}{0,210 \text{ mV}} \right) \text{ dB} = 73,6 \text{ dB}
 \end{aligned}$$

3.7.7 Forstærker med filter

En forstærker med et filter for reduktion af båndbredden benyttes meget ofte for at reducere støj eller for at modificere den aktuelle sensors overføringsfunktion.¹⁰⁹



Figur 195 – Forstærker med et passivt filter for en forsætlig båndbegrænsning.

Signalet forstærkes op med forstærkningen indstillet ved R_1 og R_2 , båndbegrænses af filtret ved R_F og C_F og forstærkes yderligere indstilles ved R_3 og R_4 . Det forudsættes at båndbredden af begge forstærkere er større end filtrets grænsefrekvens.

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 forstærkes med A_2 og båndbredden begrænses til B_2 . Udgangens støj er ikke hvid for en del af støjen har en båndbredde på 10 Hz mens resten af støjen strækker op til 50 kHz. Beregningen 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 omsættes 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. Af den grund vælges filtret til 10 Hz som dæmper 100 gange ved 1 kHz. Det antages at udgangssignalet amplitude er 1 V.

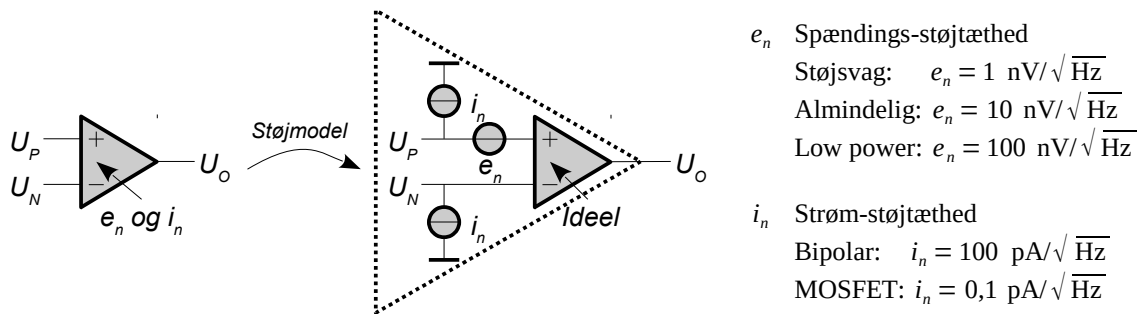
Forstærkningen i hvert trin vælges til $A_1 = A_2 = 101$ der kan realiseres med $R_1 = R_3 = 1 \text{ k}\Omega$ samt $R_2 = R_4 = 100 \text{ k}\Omega$. Grænsefrekvensen er $B_F = 10 \text{ Hz}$ ved filtret der realiseres med $R_F = 160 \text{ k}\Omega$ og $C_F = 100 \text{ nF}$. Som OpAmp vælges en LF356 type med $e_{nA} = 12 \text{ nV}/\sqrt{\text{Hz}}$ og $GBP = 5 \text{ MHz}$. Accelerometeret antages at kunne repræsenteres ved en modstand på $R_S = 10 \text{ k}\Omega$.

Modstand R_S	$U_{nRS} = e_{nRS} A_1 A_2 \sqrt{B_F} = 0,42 \text{ mV}$	$e_{nRS} = \sqrt{4kTR_S} = 13 \text{ nV}/\sqrt{\text{Hz}}$
Forstærker 1	$U_{nA1} = e_{nA1} A_1 A_2 \sqrt{B_F} = 0,39 \text{ mV}$	$e_{nRF} = \sqrt{4kTR_F} = 51 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_1	$U_{nR1} = e_{nR1} A_1 A_2 \sqrt{B_F} = 0,13 \text{ mV}$	$e_{nR1} = e_{nR3} = \sqrt{4kTR_1} = 4 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_2	$U_{nR2} = e_{nR2} A_2 \sqrt{B_F} = 0,01 \text{ mV}$	$e_{nR2} = e_{nR4} = \sqrt{4kTR_2} = 41 \text{ nV}/\sqrt{\text{Hz}}$
Modstand R_F	$U_{nRF} = e_{nRF} A_2 \sqrt{B_F} = 0,02 \text{ mV}$	$e_{nA1} = e_{nA2} = \text{Datablad} = 12 \text{ nV}/\sqrt{\text{Hz}}$
Forstærker 2	$U_{nA3} = e_{nA2} A_2 \sqrt{B_2} = 0,27 \text{ mV}$	hvor $A_1 = 1 + \frac{R_2}{R_1} = 101$
Modstand R_3	$U_{nR3} = e_{nR3} A_2 \sqrt{B_2} = 0,09 \text{ mV}$	$A_2 = 1 + \frac{R_4}{R_3} = 101$
Modstand R_4	$U_{nR4} = e_{nR4} \sqrt{B_2} = 0,01 \text{ mV}$	$B_1 = B_2 = \frac{GBP}{A} = 49,5 \text{ kHz} > B_F$
Total støj	$U_{nO} = \sqrt{\sum_{\text{Alle}} U_n^2} = 0,65 \text{ mV}$	$B_F = \frac{1}{2\pi R_F C_F} = 10 \text{ Hz}$
Signal/støj	$SNR = 20 \cdot \log_{10} \left(\frac{U_O}{U_{nO}} \right) = 64 \text{ dB}$	

¹⁰⁹Ved en meget lav grænsefrekvens er kredsløbet integrerende og kan fx omsætte acceleration til hastighed.

3.7.8 Operationsforstærkerens støjmodel

Operationsforstærkeren er kun tilnærmelsesvist beskrevet ved den termiske støj. Der er også haglstøj fra strømmen i indgangene som løber i det eksterne kredsløb. En mere generel model må derfor kombinere en spændings-støjtæthed e_n med en strøm-støjtæthed i_n for hver indgang. Som ved de øvrige modeller repræsenteres en OpAmp ved en ideel operationsforstærker med fejlene vist som eksterne kilder.



Figur 196 – Støjen i en forstærker repræsenteres af en spændingsgenerator i den ene indgang og en strømstøjgenerator for hver indgang.

Analysearbejdet er meget i stil med DC modellen for en operationsforstærker og det er muligt at benytte de samme kredsløbsligninger blot skal støjbidragene RMS adderes. Det er ikke muligt at kompensere for støjklidernes bidrag og der er ingen offsetstrøm i støjmodellen.

Operationsforstærkerens biasstrøm (side 168) er en forudsætning for at transistorerne i indgangen virker og strømmen skal leveres fra (eller optages af) det eksterne kredsløb. Der løber derfor en strøm i kredsløbet forårsaget af operationsforstærkeren. Strøm-støjtætheden i_n kan opfattes som en AC komponent der er overlejret denne DC strøm og strømmen omsættes af modstande til en spændings-støjtæthed der forstærkes på linje med den termiske støj fra modstandene.

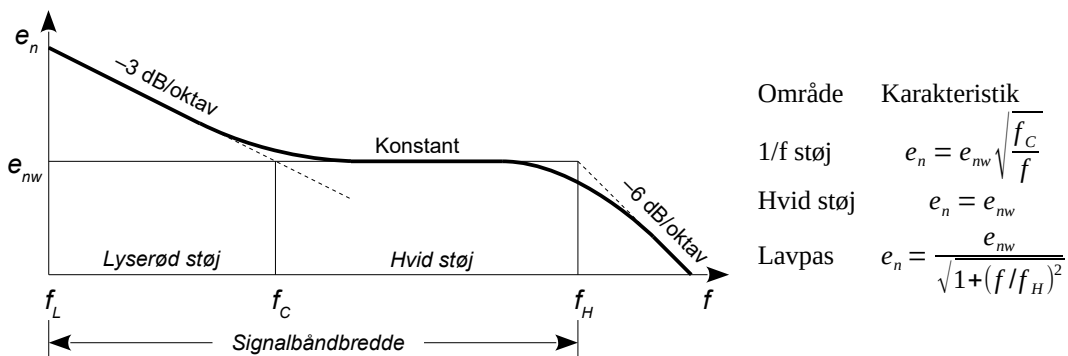
Støjkilderne har forskellig værdi for operationsforstærkere med enten bipolære transistorer eller MOSFET i indgangen.

- Ved bipolare transistorer er biasstrømmen relativt høj så bidraget fra i_n kan blive et problem hvis de eksterne modstande har høj værdi. Megen litteratur anbefaler derfor denne type OpAmp for signalkilder med lav indre modstand, hvilket som en tommelfingerregel bør være under 1 k Ω . Til gengæld kan e_n være meget lav og der kan bygges meget støjsvage opstillinger med en bipolar OpAmp hvis modstandsniveauet holdes under 100 Ω .
- Ved MOSFET er der næsten ingen biasstrøm så strøm-støjtætheden bliver lav og denne type OpAmp er derfor velegnet for en kildemodstand over 100 k Ω . Det er meget svært at bygge en støjsvag operationsforstærker med MOSFET så den høje termiske støj fra en sensor med høj indre modstand har mindre betydning da støjniveauet generelt er højt.

Operationsforstærkeren genererer ikke hvid støj for støjens RMS værdi er kraftigst ved lave frekvenser. Det kaldes for 1/f-støj og komplicerer beregningen af det samlede støjniveau, men det er almindeligvis tilstrækkeligt med en forenklet tilgang uden brug af avanceret matematik hvilket jeg viser på side 194. Det analytiske grundlag præsenteres herunder og der gives eksempler på databladets oplysninger på side 195.

3.7.9 Støjanalyse

Støjeffekten er ved lave frekvenser omvendt proportional med frekvensen så enheden bliver watt per hertz (W/Hz). Støjen kaldes for 1/f-støj på grund af afhængigheden af frekvensen og udtrykker støjen som en spændings-støjtæthed der aftager niveauet med kvadratroden af frekvensen. Det betyder at hældningen er -3 dB/oktav eller -10 dB/dekade og betegnes lyserød støj eller rosa støj (*pink noise*). Ved lave frekvenser er 1/f støjen dominerende og grænsen f_c (*corner frequency*) er opgivet i databladet for støjsvage operationsforstærkere.^{110,111}



Figur 197 – 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.

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.

For at kunne beskrive RMS værdien af støjspændingen skal forløbet integreres fra nul til uendelig. Det giver et problem for ved frekvensen nul er bidraget uendeligt. Løsningen på problemet er at måletiden må være endelig for resultatet skulle gerne være fundet inden man selv dør af alderdom. 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 termokobler eller en måling af meget svage strømme i en integreret kreds, men uanset applikationen vil den lave grænsefrekvens nødvendigvis være forskellig fra nul.

Resultatet af integrationen er vist i illustrationen som RMS værdien U_n af en spændings-støjtæthed 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 begrænses af operationsforstærkerens båndbrede. Udtrykket under kvadratroden kaldes for den ækvivalente støjbåndbrede.

For NE5532 er $e_{nw} = 5$ nV/√Hz og $f_c = 30$ Hz. RMS værdien af spændingsstøjen bliver $U_{nA} = 1,4$ μV ved grænsefrekvenserne $f_L = 1$ Hz og $f_H = 49,5$ kHz. Forstærkeren på side 188 har $A = 244$ så udgangens RMS støj bliver $U_{nO} = 0,34$ mV hvor den tidligere beregning gav 0,15 mV. Det er en ændring på 7 dB.

Det er oftest muligt at se bort fra 1/f-støjen der primært er af interesse ved måling af langsomt varierende signaler som måling af temperatur. Ved radiomodtagere benyttes JFET der har lav støj over cirka 50 kHz hvorved radioens aktive frekvensområde er i sikker afstand fra 1/f-støjen.

¹¹⁰ Mekanismen bag 1/f-støj er ukendt, men formodes at være relateret til omrokering af ladningskoncentrationer, strømme i overfladen af halvlederen og pludselige ændringer i krystalstrukturen fra dislokationer og diffusion af fremmedatomer.

¹¹¹ Ifølge engelske Wikipedia er der udført måling af 1/f-støj over flere uger som bekræfter at forløbet er som beskrevet.

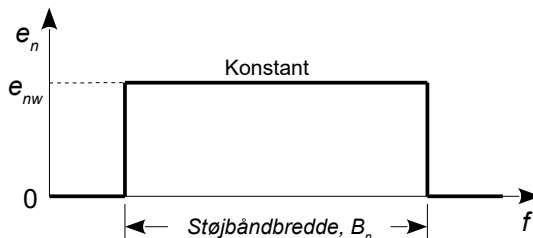
Herunder gennemføres integrationen over de tre frekvensbånd.

For at beregne en RMS værdi af støjspændingen U_n startes med at finde et udtryk for den samlede støjefekt P_n fra kredsløbets udgang. Effekttætheden beregnes fra p_n der beskrives ved en givet frekvens f . Støjefekten 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 er integration over alle frekvenser fra nul til uendelig.

$$P_n = \int_0^{\infty} p_n df = \int_0^{\infty} \frac{e_n^2}{R} df = \frac{U_n^2}{R} \Rightarrow U_n^2 = \int_0^{\infty} e_n^2 df \Rightarrow U_n = \sqrt{\int_0^{\infty} e_n^2 df}$$

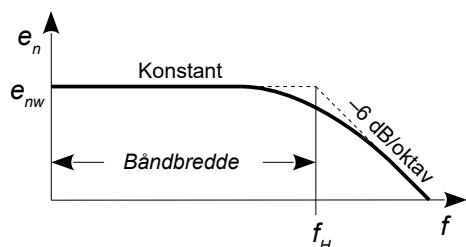
Effekttætheden p_n er givet ved kvadratet på spændingstætheden e_n delt med en modstand R og resultatet af integrationen kaldes for U_n som repræsenterer den RMS spænding der giver den samme effekt som støjen. Modstanden R er medtaget for at tydeliggøre relationen mellem effekt og spænding. Den antages at være den samme under hele beregningen idet måleopstillingen er invariant så den absolutte værdi af modstanden har ingen betydning.

I midterområdet hvor støjtætheden e_n er konstant vil integrationen løbe fra f_L til f_H . Værdien 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 side 184.



$$U_n = \sqrt{\int_{f_L}^{f_H} e_n^2 df} = \sqrt{e_n^2 \int_{f_L}^{f_H} df} = e_n \sqrt{f_H - f_L}$$

Båndbredden er ikke så skarp som vist i illustrationen for det såkaldte *brick-wall filter*, men det er en ganske god tilnærmelse som integration fra f_H og op viser. Støjspændingen bestemmes herunder for et første-ordens lavpasfilter der repræsenterer det høje område. Støjtætheden e_{nw} på indgangen af filtret antages konstant og den effektive værdi af spændingsstøjen viser at der kun kommer et bidrag på 2 dB ekstra fra området over -3 dB grænsefrekvensen.



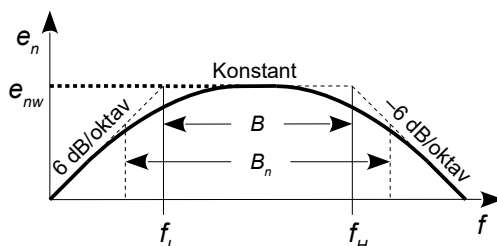
$$e_n = e_{nw} \left| \frac{\omega_H}{\omega_H + s} \right| = e_{nw} \frac{1}{\sqrt{1 + (f/f_H)^2}}$$

$$U_n = \sqrt{\int_0^{\infty} e_{nw}^2 \frac{1}{1 + (f/f_H)^2} df}$$

$$U_n = e_{nw} \sqrt{f_H \left[\arctan\left(\frac{f}{f_H}\right) \right]_0^{\infty}}$$

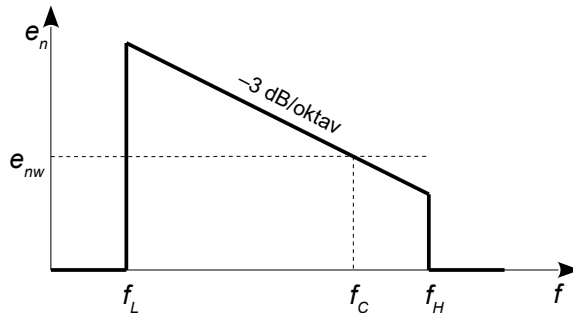
$$U_n = e_{nw} \sqrt{\frac{\pi}{2} f_H}$$

Der kan gennemføres en lignende integration for et båndpasfilter og resultatet bliver $\pi/2$ gange differensen mellem filtrets -3 dB grænsefrekvenser.



$B_n = 1,57 \cdot f_H$	Lavpasfilter -6 dB/oktav
$B_n = 1,11 \cdot f_H$	Lavpasfilter -12 dB/oktav
$B_n \approx f_H$	Lavpasfilter af højere orden
$B_n = 1,57 \cdot f_{BP}$	Båndpasfilter $f_{BP} = f_L = f_H$
$B_n = 1,57 \cdot (f_H - f_L)$	Båndpasfilter $f_H \gg f_L$

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 dB/oktav. Spændingsstøjens RMS værdi beregnes ved at integrere over området fra en lav grænse f_L til en høj grænse f_H . Resultatet er et udtryk baseret på en logaritme.



$$U_n = \sqrt{\int_{f_L}^{f_H} e_n^2 df} \Rightarrow U_n = e_{nw} \sqrt{f_C \int_{f_L}^{f_H} \frac{df}{f}}$$

$$e_n = e_{nw} \sqrt{\frac{f_C}{f}}$$

$$U_n = e_{nw} \sqrt{f_C \ln\left(\frac{f_H}{f_L}\right)}$$

Konklusionen på beregningerne kan samles til et fælles udtryk for en OpAmp der karakteriseres ved spændingsstøjtætheden e_{nw} og grænsen f_C mellem $1/f$ -støj og hvid støj. Forstærkningen er A og båndbredden er indstillet til f_L i den lave ende og f_H i den høje ende.

$$U_n = e_{nw} A \sqrt{f_C \ln\left(\frac{f_H}{f_L}\right) + \frac{\pi}{2}(f_H - f_L)}$$

Med $f_L = 1$ Hz, $f_C = 1$ kHz og $f_H = 10$ kHz samt $e_{nw} = 10$ nV/ $\sqrt{\text{Hz}}$ findes støjen til $U_n = 1$ μV .

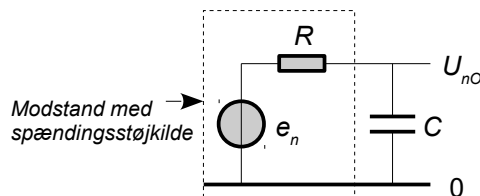
Den beregnede værdi stiger kun til $U_n = 1,3$ μV ved grænsefrekvensen $f_L = 0,001$ Hz der kræver en måletid på knapt en time for at give mening og forskellen er blot 2 dB.

Flertallet af anvendelser har et veldefineret frekvensområde der giver mulighed for at stille et kvalificeret krav til operationsforstærkerens f_C for overgangen fra $1/f$ støj til hvid støj.

Mellemfrekvensen i en AM radiomodtager har en båndbredde på 9 kHz centreret på 455 kHz og en FM radiomodtager har båndbredden 300 kHz centreret på 10,7 MHz. Analog videotransmission over et kabel benytter området fra 25 Hz til 7 MHz. En analog telefons omfang er fra 300 Hz til 3,5 kHz, måling af EKG benytter 0,1 Hz til 150 Hz og måling af vægt og temperatur er ofte i området 0,01 til 10 Hz.

3.7.10 Støjen fra et RC led

En modstand genererer en spændings-støjtæthed som er givet af modstandsværdien og med en kondensator i parallel vil støjen båndbegrænses af filtret dannet af modstand og kondensator. Den resulterende RMS værdi af støjens spænding er alene funktion af kondensatoren. Modstanden giver ganske vist mere støj fra sig når dens værdi stiger, men båndbredden aftager.



$$U_{nR} = e_{nR} \sqrt{B_n}$$

$$e_{nR} = \sqrt{4kTR}$$

$$B_n = \frac{\pi}{2} f_{-3\text{ dB}} = \frac{\pi}{2} \frac{1}{2\pi RC}$$

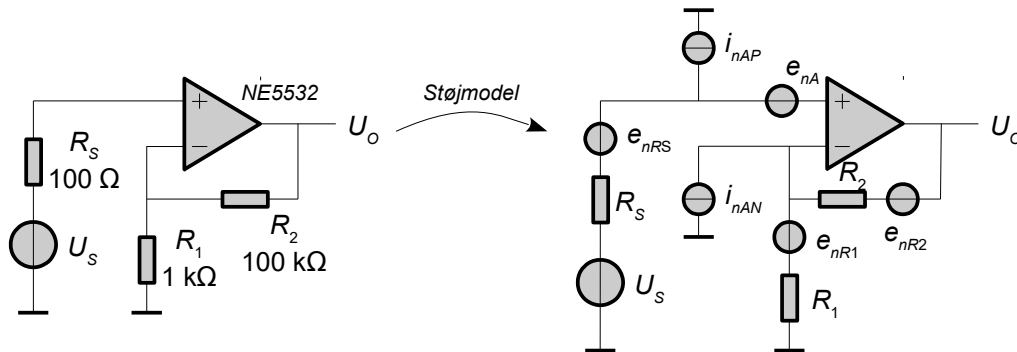
$$\Rightarrow U_{nO} = \sqrt{\frac{kT}{C}}$$

Figur 198 – Støjen fra en parallellforbindelsen af en modstand og en kondensator afhænger kun af kondensatoren. Selv om RMS værdien af støjspændingen er konstant ved varierende modstand så er støjens båndbredde givet af grænsefrekvensen defineret ved R og C .

Et RC led med en kondensator på 1 nF giver en RMS spænding på 2 μV indenfor båndbredden der med eksempelvis 100 k Ω og 1 nF er på $B_n = 2,5$ kHz.

3.7.11 Eksempel

Herunder illustreres brugen af modellen med et eksempel for en ikke-inverterende forstærker der er identisk med forstærkeren på side 188 og med forstærker 2 i kredsløbet på side 189. Der er to modstande i tilbagekoblingen som indstiller forstærkning og båndbredde (R_1 og R_2) og dertil en signalkilde med en indre modstand R_S . Beregningen af den resulterende støj på udgangen er blot 2 dB kraftigere end ved den simplere beregning.



Figur 199 – En ikke-inverterende forstærker kan med hensyn til støj opfattes som vist til højre hvor alle støjklender er tegnet ind og forstærkeren nu regnes for ideel.

Støjbidraget for NE5532 hentes fra databladet og modstandenes beregnes ved 300 K. Båndbredden beregnes ud fra GBP specifikationen og den indstillede forstærkning A . Den ækvivalente støjbandbredde er B_n . Den lave grænsefrekvens er sat til 1 Hz og forstærkeren korrigeres for $1/f$ -støj, men båndbredden er stort set ikke ændret fra B_n på grund af den lave værdi af f_c for NE5532. I den positive indgang løber strømstøjen i_n gennem R_S og forstærkes med A . I den negative indgang løber strømstrømmen udelukkende gennem R_2 .

OpAmp NE5532: $e_n = 5 \text{ nV}/\sqrt{\text{Hz}}$, $f_c = 30 \text{ Hz}$, $i_n = 0,7 \text{ pA}/\sqrt{\text{Hz}}$, $GBP = 10 \text{ MHz}$

Forstærkning: $A = 1 + \frac{R_2}{R_1} = 101,$

Båndbredde: $f_H = \frac{GBP}{A} = 100 \text{ kHz} \Rightarrow B_n = \frac{\pi}{2}(f_H - f_L) = 157 \text{ kHz}$
 $f_L = 1 \text{ Hz}$ (valgt)

Modstandsstøj: $U_{nRS} = A \sqrt{4kTR_S B_n} = 0,051 \text{ mV}$
 $U_{nR1} = A \sqrt{4kTR_1 B_n} = 0,163 \text{ mV}$
 $U_{nR2} = \sqrt{4kTR_2 B_n} = 0,016 \text{ mV}$

OpAmp spændingsstøj $U_{neA} = A e_{nA} \sqrt{f_c \ln\left(\frac{f_H}{f_L}\right) + B_n} = 0,200 \text{ mV}$

OpAmp strømstøj $U_{niRS} = A R_S i_n \sqrt{B_n} = 0,003 \text{ mV}$
 $U_{niR2} = R_2 i_n \sqrt{B_n} = 0,028 \text{ mV}$

Udgangens støj $U_{nO} = \sqrt{U_{nRS}^2 + U_{nR1}^2 + U_{nR2}^2 + U_{neA}^2 + U_{niRS}^2 + U_{niR2}^2} = 0,265 \text{ mV}$

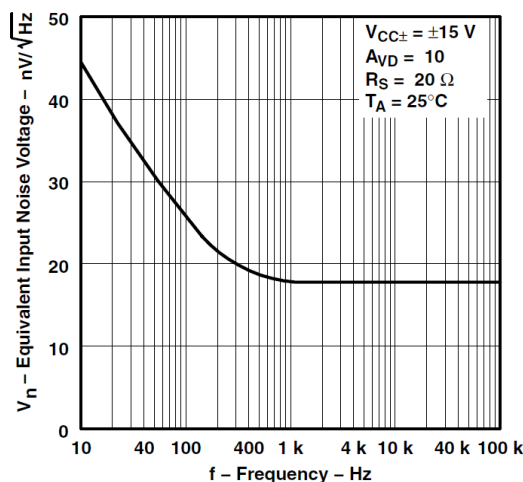
Signal/støj-forhold $SNR = 20 \cdot \log_{10}\left(\frac{U_O}{U_{nO}}\right) \text{ dB} = 71,5 \text{ dB}$

Ved beregningen for kredsløbet på side 188 var RMS værdien af støjen på 0,210 mV og den korrigerede værdi er kun 2 dB større som følge af den lidt større båndbredde i analysen.

Den primære støjkilde er spændingsstøjtætheden fra operationsforstærkeren med 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 betydende bidrag på grund af de lave modstandsværdier.

3.7.12 Databladet

Støjtætheden 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 spændings-støjtætheden som funktion af frekvensen og undertiden ligeledes en kurve for strøm-støjtætheden. De mere almindelige forstærkere har blot en talværdi. Det viste eksempel for TL071 repræsenterer den populære serie med TL072 og TL074 og grænsefrekvensen mellem de to støjområder ses at ligge ved $f_c = 300$ Hz.

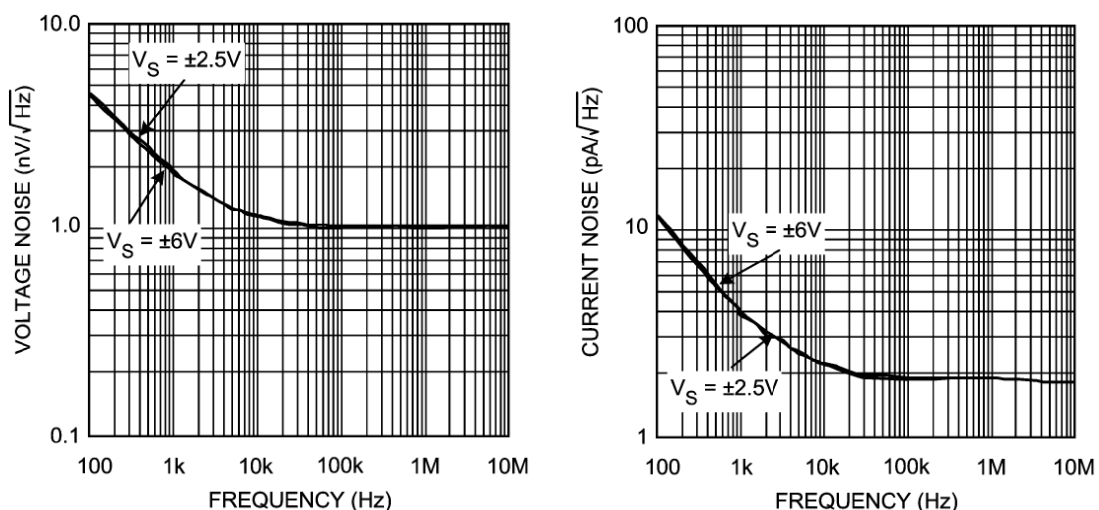


Forstærker	e_n nV/√Hz	i_n pA/√Hz	I_{CC} mA
LMH6624	1	2,3	11
AD797	0,9	2	8,2
NE5532	5	0,7	8
OP27	3,2	0,4	2,8
TL071	18	0,01	1,4
OP262	9,5	0,4	0,5
TLC271	25...68	—	0,02...2
OP293	65	0,05	0,015
ADA4051	95	0,1	0,01

Figur 200 – Støjtætheden for TL071-serien som funktion af frekvensen (venstre). Eksempler på specifikationerne for nogle udvalgte operationsforstærkere (højre). Kolonnen med strømforbrug viser at en støjsvag operationsforstærker har et højt strømforbrug.

LMH6624 er en bredbåndet forstærker (GBP på 1,5 GHz) som kræver en forstærkning på minimum 10 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 samler DC fejl og derefter subtraherer disse fejl til en specifikation på typisk ± 2 μ V.

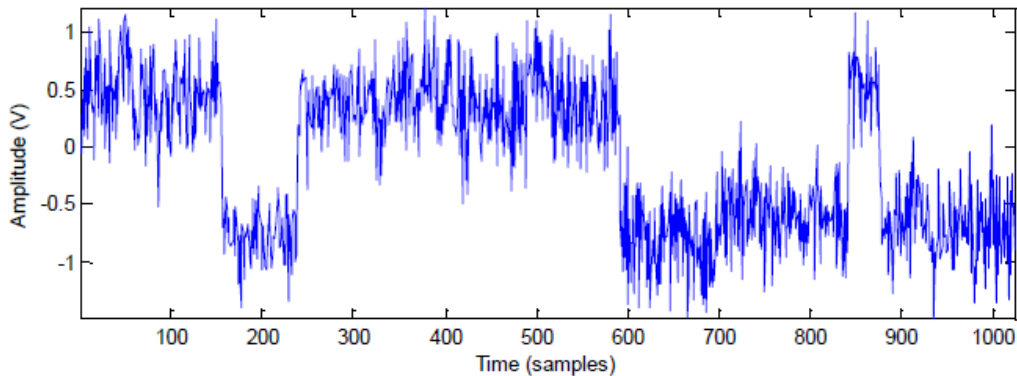
Støjsvage forstærkere kræver en høj værdi af strømmen i de interne trin som det vises i femte kapitel så der kan ikke laves støjsvage operationsforstærkere i *micro power* klassen.



Figur 201 – Støjtætheden (spænding og strøm) for LMH6624 og dobbeltudgaven LMH6626.

3.7.13 Popcorn støj

En noget speciel form for støj kaldes Popcorn støj (*popcorn noise*) som beskrivelse af lyden i en højttaler. Der gives flere forklaringer som opbygning og løsladelse af ladninger i halvlederen eller dens overflade, 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 pludselige ændringer kan være fra millisekunder til sekunder og amplituden kan i de mest ekstreme tilfælde være så høj som 100 μV .¹¹²



Figur 202 – Eksempel på popcorn støj i en OpAmp med en ukendt indstilling af forstærkningen.

Popcorn støjen er problematisk ved forstærkning af meget små signaler og ved lydbehandling hvor den høres som majs-korn der popper i en mikrobølgeovn. Den er specificeret for gode forstærkere i form af en garanteret spids-spids værdi af støjen fra 0,1 til 10 Hz. For den støjsvage AD797 er værdien garanteret indenfor 50 nV spids-spids for det angivne frekvensområde. Parameteren er verificeret som en del af testforløbet i produktionen og da den nedre grænsefrekvens fordrer en testtid på flere sekunder er det en kostbar test i et automatiseret produktionsforløb.

3.8 Sammenfatning

Modstande genererer støj på grund af elektronernes termisk forårsagede vibration og det giver en evigt varierende spænding mellem modstandens ledninger der betegnes spændings-støjtætheden med symbolet e_n og enheden $\text{V}/\sqrt{\text{Hz}}$. Strømmen bæres af et endeligt antal elektroner per tidsenhed og det giver en strøm-støjtæthed med symbolet i_n og enheden $\text{V}/\sqrt{\text{Hz}}$. Værdien er givet ved DC strømstyrken I_{DC} for dioder, bipolare transistorer (basis og kollektor) og felteffekt transistorer (gate). RMS værdien af støjen U_n ved udgangen fra et kredsløb kan beregnes fra forstærkningen A og båndbredden B_n for den pågældende støjkilde og det er tilstrækkeligt med -3 dB båndbredden i de fleste applikationer. Flere bidrag adderes kvadratisk hvilket også kaldes for RMS-addition.

Spændingsstøjtæthed:	$e_{nR} = \sqrt{4kTR}$	Boltzmanns konstant:	$k = 1,38 \cdot 10^{-23} \text{ J/K}$
Strømstøjtæthed:	$i_n = \sqrt{2q_0 I_{DC}}$	Elementarladningen:	$q_0 = 1,60 \cdot 10^{-19} \text{ C}$
RMS støjspænding:	$U_n = e_n A \sqrt{B_n}$	Omgivelsestemperatur:	$T \approx 300 \text{ K}$
Støjbåndbredde:	$B_n = f_H - f_L$	Forstærkning:	A
Addition af støj:	$U_{nTOT} = \sqrt{U_{n1}^2 + U_{n2}^2 + \dots}$	Strømstyrke:	I_{DC}
		Høj grænsefrekvens:	f_H
		Lav grænsefrekvens:	f_L

Supplerende information om støj i operationsforstærkere er tilgængelig fra internettet.¹¹³

¹¹² http://www.advsolned.com/example_popcorn_noise.html.

¹¹³ http://www.analog.com/static/imported-files/rarely_asked_questions/moreInfo_ra_q_opAmpNoise2.html.
http://en.wikipedia.org/wiki/Johnson%E2%80%93Nyquist_noise.



Din ekspert i
signalbehandlingsløsninger
og kommunikationsinterfaces

PERFORMANCE
MADE
SMARTER

TEMPERATUR | EX-BARRIERER | 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 styres af principper som enkelhed, fokus, mod og dygtighed og har fokus på at videreudvikle 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 styrker 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 indspark..."

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..."

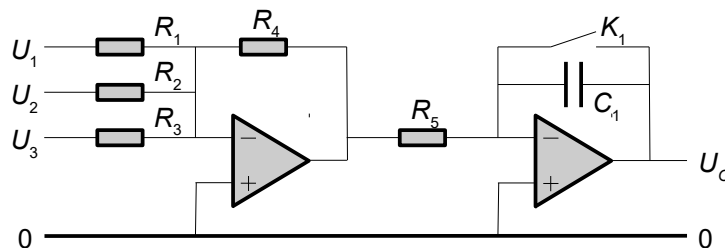
Kevin Grooters
Praktikant, Technology Department

Besøg <http://www.prelectronics.dk/career> og se, hvordan du kan blive en del af PERFORMANCE MADE SMARTER.

PR
electronics

3.9 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 integrerede forstærkere overhovedet blev udviklet var for anvendelse i analoge computere. Der var godt nok et voksende kommercielt marked, men det var videnskabelige og militære anvendelser, der på dette tidspunkt 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 203 – Simple 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 et 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, radarekniik, kemi og rumfart.

$$U_o = \frac{1}{T_1} \int_0^{\tau} (\alpha_1 U_1 + \alpha_2 U_2 + \alpha_3 U_3) dt, \quad T_1 = R_5 C_1, \quad \alpha_n = \frac{R_4}{R_n}, \quad -10 \text{ V} \leq U_o \leq 10 \text{ V}$$

3.9.1 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*.

$$U_{CMR \min} \leq (U_P, U_N) \leq U_{CMR \max}$$

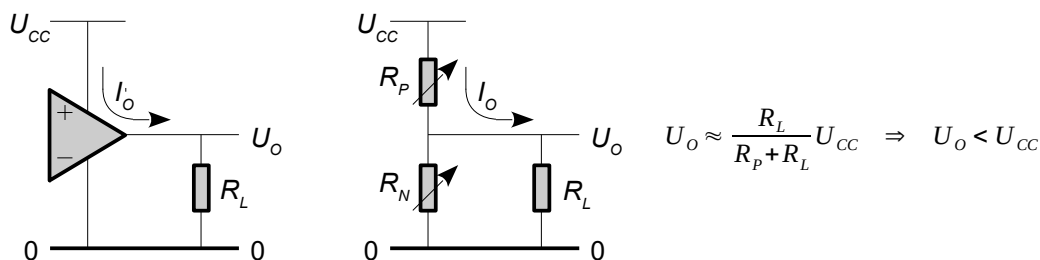
Området fra den laveste til den højeste spænding er en funktion af effektforsyningen så de to grænseværdier er relateret til de arbejdsbetingelser man byder forstærkeren. Ved ± 15 V var det i gamle dage nok at holde indgangene indenfor ± 12 V for at være på den sikre side, men ved de moderne kredsløb 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 benytter begge en P kanal JFET i indgangen og danner et eksempel på en undtagelse.

Går man udenfor databladets grænser kan der ske sære ting som f at udgangssignalets polaritet inverteres. En mere ubehagelig egenskab ved integrerede kredsløb er den destruktive *latch-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 har givet en del 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.

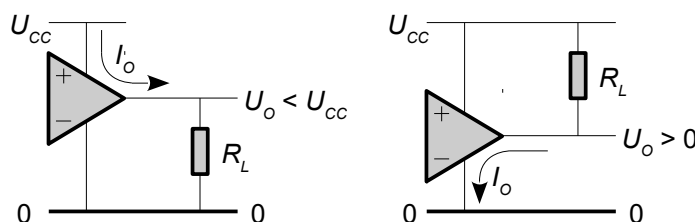
3.9.2 Udgangens spændingsområde

Operationsforstærkeren benytter transistorer til at stille udgangen på den ønskede spænding og samtidigt at levere den krævede udgangsstrøm. Udgangstrinnet er opbygget som *push pull* hvor der er to variable komponenter der herunder er vist som modstandene R_P og R_N . 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 204 – 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 bipolære transistorer er det muligt at nå 50 mV fra effektforsyningen hvis udgangen drives fra transistorens kollektor mens 600 mV er grænsen hvis basis-emitter er involveret. Med to transistorer i kaskade kan spændingsfaldet blive 1,2 V som det 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 hvilket udnyttes i CMOS forstærkere som fx TLC271. Uanset designet betyder det at man skal studere databladet grundigt hvis det er vigtigt at nå tæt på effektforsyningens grænser.



Figur 205 – Udgangen skal enten levere strøm til belastningen eller optage en strøm fra den. Der vil ligge et spændingsfald over udgangstrinnet så det er ikke altid muligt at drive udgangen helt op til effektforsyningen eller helt ned til nul.

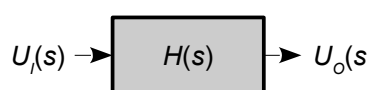
3.9.3 Udgangens strømområde

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.

Note 3

3.10 Overføringsfunktion

Et elektronisk system bearbejder meget ofte et oscillerende signal ved at ændre amplitude og fase som funktion af frekvensen af oscillationen. Udgangssignalet U_o er givet ved indgangssignalet U_i gange en overføringsfunktion H . Beregningen kan benytte alle typer af periodiske signaler for alle periodiske svingninger kan beskrives ved en Fourier serie af harmoniske svingninger der analytisk kan opfattes som individuelle harmoniske svingninger. Dertil kan benyttes ikke-periodiske signaler som opbygges med en blanding af kontinuerte funktioner, enhedsstep og Dirac delta pulser.



$$U_o = H U_i$$

$$U_o = |U_o| \exp(j\theta_o) \exp(j\omega t)$$

$$U_i = |U_i| \exp(j\theta_i) \exp(j\omega t)$$

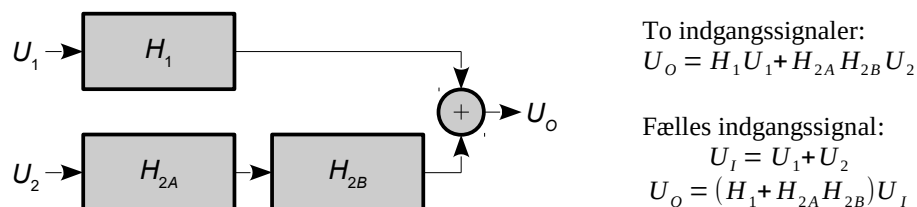
$$H = \frac{U_o}{U_i} = \left| \frac{U_o}{U_i} \right| \exp(j(\theta_o - \theta_i))$$

Figur 206 – Et systems overføringsfunktion beskriver hvordan udgangen reagerer på et signal påført indgangen. Alt er komplekse funktioner af s .

Selve oscillationen beskrives ved $\exp(j\omega)$ og kan forkortes væk som vist ovenfor da ind- og udgang svinger på den samme frekvens. Overføringsfunktionen kan måles i laboratoriet ved amplituden og fasen af U_i og U_o og det er nærliggende at definere tonegeneratorens signal som referencen hvor U_i defineres som 0 dB og θ_i som 0° . Overføringsfunktionen kan også findes analytisk ved brug af kredsløbsligninger baseret på Ohms og Kirchhoffs love som vist tidligere i dette kapitel.

3.10.1 Sammensatte kredsløb

Overføringsfunktionen kan umiddelbart benyttes 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 207 – Kredsløbet kan bestå af flere delkredse. Ved serieforbundne delkredsløb er den samlede funktion givet ved produktet af overføringsfunktionerne og ved parallelforbundne 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.

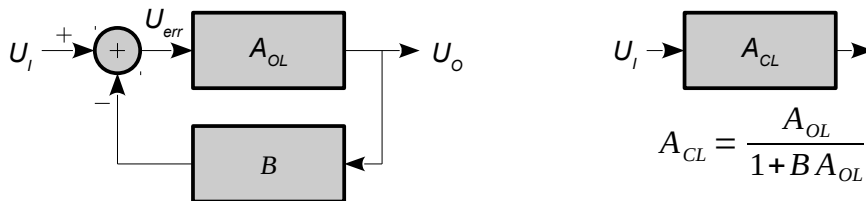
I eksemplet ledes det ene signal U_1 via overføringsfunktionen H_1 til udgangen U_o hvor det adderes med det andet signal U_2 via de serieforbundne overføringsfunktioner H_{2A} og H_{2B} . Hvis signalet på indgangen U_i er fælles for de to veje ($U_1 = U_2 = U_i$) vil den resulterende overføringsfunktion blive summen af de to funktionsudtryk¹¹⁴.

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 system af konstruktive og destruktive interferenser som øret opfatter som den karakteristiske lyd.

¹¹⁴ Det kan eksempelvis repræsentere det resulterende signal i lyttepositionen fra en højtaler 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} .

3.11 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. Tilbagekoblingen fører andelen B af udgangssignalet U_O tilbage til indgangen hvor det subtraheres fra indgangssignalet og det resulterende fejlsignal U_{err} forstærkes op med A_{OL} . Hvis A_{OL} er så høj at den kan betragtes som uendelig vil forstærkningen blive $A_{CL} = 1/B$.



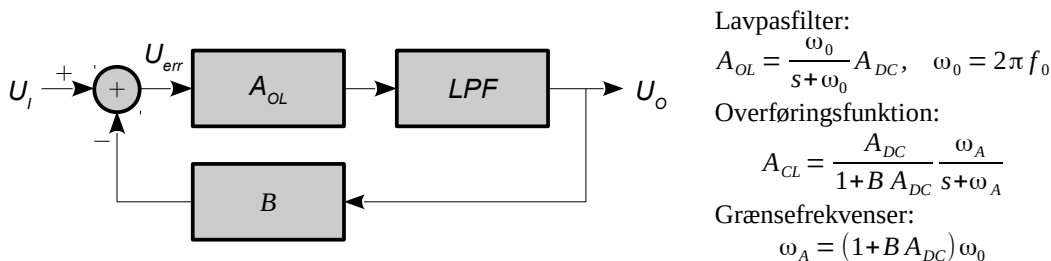
Figur 208 – Tilbagekobling af et kredsløb med overføringsfunktionen A_{OL} benytter brøkdelen B af udgangssignalet som fratrækkes indgangssignalet. Resultatet er en ny overføringsfunktion A_{CL} .

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} . Ledet $B A_{OL}$ i nævneren er sløjfeforstærkningen (*feed-back factor*) som normalt har en meget høj værdi.

$$\begin{aligned} U_O &= A_{OL} U_{err} \\ U_{err} &= U_I - B U_O \end{aligned} \Rightarrow U_O = A_{CL} U_I \quad \text{hvor} \quad A_{CL} = \frac{A_{OL}}{1 + B A_{OL}} \xrightarrow{A_{OL} \rightarrow \infty} \frac{1}{B}$$

3.11.1 Båndbredde

Operationsforstærkeren har en karakteristik der bedst beskrives som kombinationen af en DC forstærkning A_{DC} og et lavpasfilter med grænsefrekvensen f_0 (-3 dB) hvorover forstærkningen falder med -6 dB/oktav.



Med DC forstærkningen $A_{DC} = 10^5$ og tilbagekoblingen $B = 0,01$ bliver sløjfeforstærkningen $B A_{DC} = 10^3$ 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 $1/B = 100$.

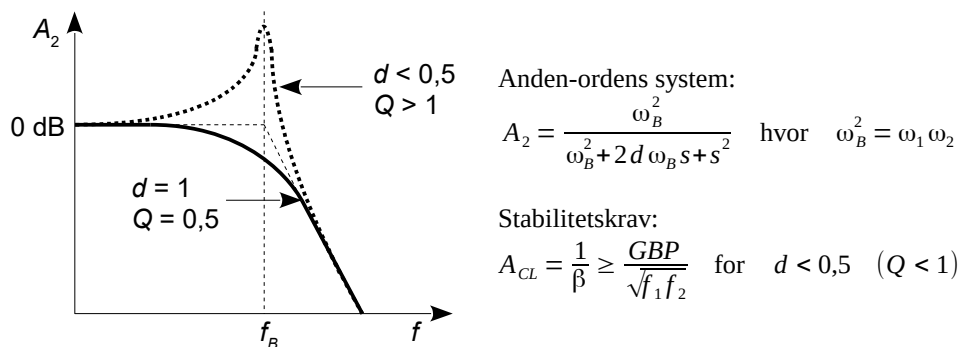
Selv om flertallet af tilbagekoblede systemer er mere komplekse vil mange kunne reduceres til denne simple model. Overføringsfunktionen A_{OL} beskrives ved en forstærkning A_{DC} ved DC og lave frekvenser samt et første-ordens lavpasfilter med polen f_0 . Det kan også være et elektro-mekanisk system hvor forstærkningen skyldes reguleringssystemet og motorens træghed er repræsenteret gennem et lavpas-filter. Filtret kan have flere poler hvorved tilbagekoblingen giver en udfordring af stabiliteten af systemet så det er ikke givet at man altid kan tvinge et system til at makke ret blot ved at øge tilbagekoblingen.

Tilbagekoblingen giver en forstærkning på $1/B$ og båndbredden forøges med faktor $1 + B A_{DC}$. Det væsentlige er her at åben-sløjfe forstærkningen A_{OL} er reduceret til en ikke-dominerende parameter så funktionen af kredsløbet i det væsentlige styres af B gennem eksterne komponenter der typisk er et par modstande hvor tolerancen kan specificeres.

3.11.2 Stabilitet

En operationsforstærker beskrives ved en dominerende pol, men der er ofte flere poler involveret i dens overføringsfunktion. De resterende poler ligger så højt oppe i frekvens at den dominerende pol dæmper den resulterende forstærkning til én eller derunder inden de høje poler får betydning og man kan normalt se bort fra dem. Derimod vil en operationsforstærker ofte udsættes for en eller flere ekstra poler på grund af kredsløbet og det kan påvirke stabiliteten.

Når operationsforstærkeren belastes af en kondensator på ind- eller udgang dannes der en ny pol som indgår i tilbagkoblingen og overføringsfunktionen bliver af anden orden. Det betyder at der er mulighed for en resulterende funktion hvor en høj frekvens forstærkes kraftigt og at der derfor er mulighed for oscillation på denne frekvens.



Figur 209 – Ustabilitet ved den tilbagekoblede forstærker vil vise sig som en resonans ved en given frekvens hvor systemets poler giver et lukket-sløjfe fasedrej nær ved 360°.

Stabiliteten af et tilbagekoblet system er funktion af tilbagkoblingen og kan lide under en for stor værdi af BA_{DC} . Det kan udtrykkes ved en dæmpningsfaktor d der bør være mindst én for stabilitet.

Overføringsfunktionen for det ikke-tilbagekoblede system A_{OL} beskrives ved to poler med ω_0 for den dominerende pol og ω_1 for en højtliggende pol. Dertil forstærkningen A_{DC} ved DC og lave frekvenser. Overføringsfunktionen for dette system bliver A_{CL} (*closed-loop amplification*) ved tilbagekobling med brøkdelen B af udgangssignalet.

$$A_{OL} = \frac{\omega_0 \omega_1}{(s + \omega_0)(s + \omega_1)} A_{DC}$$

$$A_{CL} = \frac{A_{OL}}{1 + BA_{OL}} = \frac{\frac{\omega_0 \omega_1}{(s + \omega_0)(s + \omega_1)} A_{DC}}{1 + B \frac{\omega_0 \omega_1}{(s + \omega_0)(s + \omega_1)} A_{DC}} = \frac{\omega_0 \omega_1 A_{DC}}{(s + \omega_0)(s + \omega_1) + B \omega_0 \omega_1 A_{DC}}$$

$$A_{CL} = \frac{\omega_0 \omega_1 A_{DC}}{s^2 + (\omega_0 + \omega_1)s + (1 + BA_{DC})\omega_0 \omega_1} = \frac{A_{DC}}{1 + BA_{DC}} \frac{(1 + \beta A_{DC})\omega_0 \omega_1}{s^2 + (\omega_0 + \omega_1)s + (1 + BA_{DC})\omega_0 \omega_1}$$

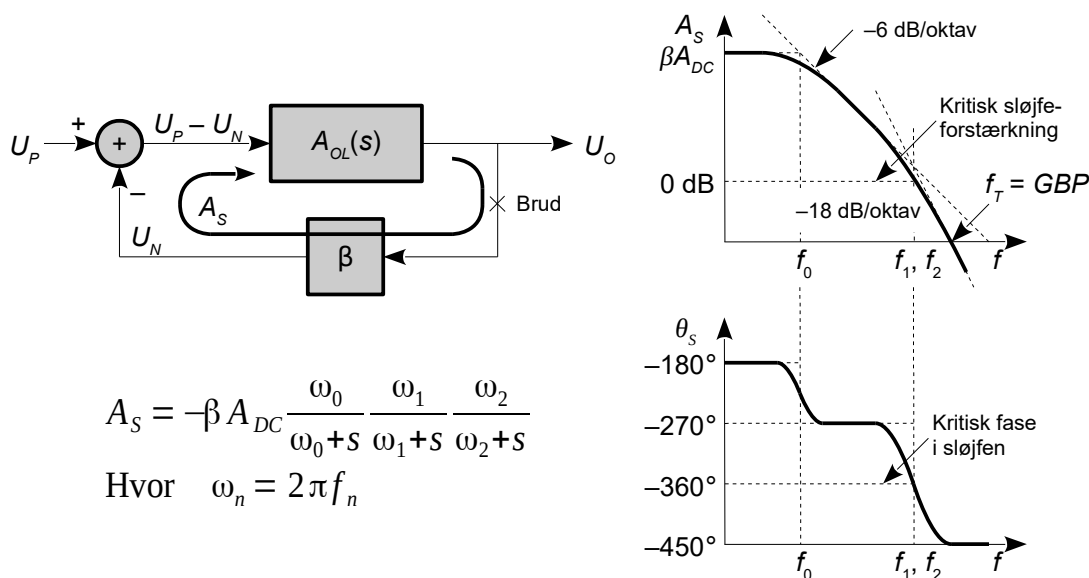
Den resulterende anden-ordens overføringsfunktion har større båndbredde ved multiplikationen af $\omega_0 \omega_1$ med faktor $1 + BA_{DC}$. Systemets stabilitet kan aflæses af dæmpningsfaktoren d .

$$\frac{(\omega_0 + \omega_1)s}{\omega_{SYS}^2} = \frac{2d\omega_{SYS}s}{(1 + BA_{DC})\omega_0 \omega_1} \Rightarrow d = \frac{\omega_0 + \omega_1}{2\omega_{SYS}} \approx \frac{\omega_1}{2\sqrt{(1 + BA_{DC})\omega_0 \omega_1}} = \frac{1}{2\sqrt{1 + BA_{DC}}} \sqrt{\frac{\omega_1}{\omega_0}}$$

For stabilitet er $d \geq 1$ for reelle rødder der opnås når den høje pol ligger mindst $4BA_{DC}$ over den dominerende pol. Det benyttes senere i dokumentet til at placere den dominerende pol lavt nok til at udgå problemer med stabilitet når den laveste af de høje poler er kendt.

$$BA_{OL} \gg 1 \Rightarrow d \approx \frac{1}{2\sqrt{BA_{DC}}} \sqrt{\frac{\omega_1}{\omega_0}} \Rightarrow \frac{\omega_1}{\omega_0} = 4BA_{DC}d^2 \xrightarrow{d=1} \frac{\omega_1}{\omega_0} = \frac{f_1}{f_0} = 4BA_{DC}$$

Forholdene beskrives gennem sløjfeforstærkningen A_S der er den tilbagekoblede andel β af udgangssignalet gange med åben-sløjfe forstærkningen A_{OL} . Sløjfeforstærkningen 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. Forstærkerens inversion beskrives ved -180° og dertil kommer -90° fra den dominerende pol. Fasedrejet er konstant på -270° for frekvenser under de høje poler ved f_1 og f_2 og begynder at stige ved frekvenser i nærheden af de høje poler. Ved en samlet fase på -360° skal sløjfeforstærkningen være reduceret til under én for at undgå selvsving så det beskrevne system har en fasemargin på 90° .



Figur 210 – Sløjfe-forstærkningen bruges til at beregne hvorvidt en OpAmp er stabil eller oscillerer. Sløjfen brydes og overføringsfunktionen opstilles fra den ene side af bruddet til den anden. Her er f_0 den dominerende pol og polerne f_1 og f_2 repræsenterer højtliggende poler eller eksterne poler.

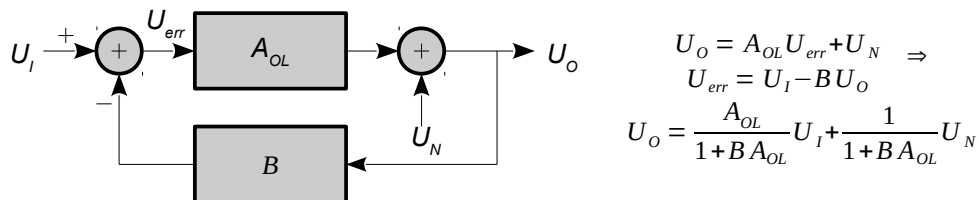
Hvis sløjfeforstærkningen bliver -1 vil overføringsfunktionen gå imod uendelig ved den aktuelle frekvens og operationsforstærkeren vil selvsvinge. Det kaldes Nyquists stabilitetskriterium og danner altså grænsen til oscillation. For en stabilt fungerende operationsforstærker må værdien af sløjfeforstærkningen ikke komme bare i nærheden af grænseværdien. Det beskrives almindeligvis ved en fasemargin der viser hvor meget reserve der er tilovers i faseren når forstærkningen er én.

Der kan opnås et marginalt stabilt system ved 45° fasemargin hvor der vil være en svag tendens til oscillation. Mange referencer benytter dette som stabilitetskriterium. Men jeg vil være forsigtig med at designe for skråsikkert efter dette. Der er altid nogle poler som man ikke har opdaget.¹¹⁵

¹¹⁵ 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.

3.11.3 Reduktion af uønskede signaler

Et tilbagekoblet kredsløb kan reducere problemer med at eksterne signaler kobler ind i kredsløbet. Det kan være 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 med temperaturen eller en ikke-linearitet i et internt kredsløb. En model af kredsløbet beskrives ved et uønsket signal U_N der kobles ind i signalvejen.

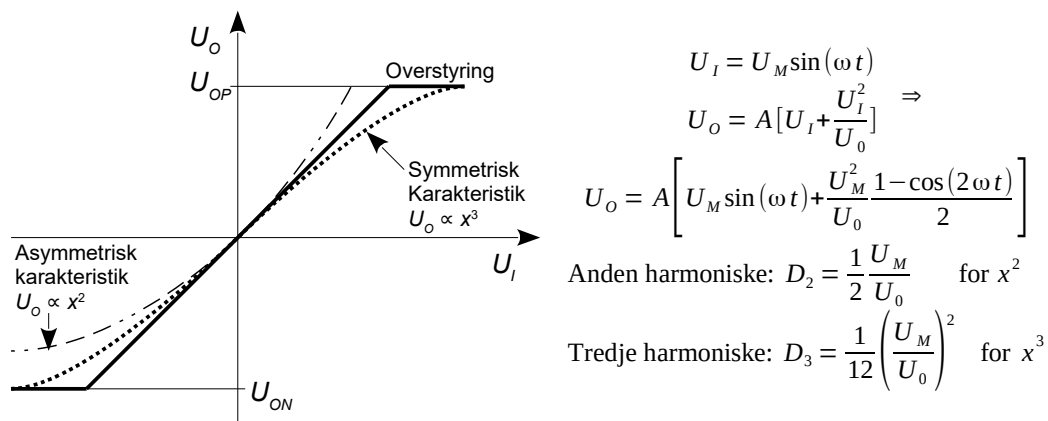


Figur 211 – Tilbagekobling af et kredsløb kan reducere uønskede signaler der kan repræsentere et eksternt støjsignal eller kredsløbets egen forvrængning.

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 $1 + B A_{OL}$ 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.

3.11.4 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 karakteristisk 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 svingning dannet af en sum af harmoniske svingninger. Det kaldes for harmonisk forvrængning.



Figur 212 – En forstærkers udgangssignal er sjældent en lineær funktion af indgangssignalet for oftest er relationen afhængig af spændingsniveauet hvilket giver en krum karakteristisk (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). Uledningen viser kun processen for anden-harmonisk forvrængning.

De simpleste former for forvrængning er de kvadratiske og kubiske relationer (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 dels 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 indgangssignalet¹¹⁶.

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 A_{OL} falder til nul ved et forsøg på at tvinge udgangen udenfor grænserne og så svigter selve idéen med at føre udgangens signal tilbage til indgangen.

Et lignende problem optræder ved den ubehagelige cross-over forvrængning i en effektforstærker hvor forstærkningen falder imod nul ved overgangen mellem svage positive og negative signaler på grund af udgangstrinnet. Det gælder primært for de såkaldte klasse B forstærkere. Da A_{OL} er nul kan forvrængningen ikke reduceres ved tilbagekobling.

3.11.5 Harmonisk forvrængning

Definitionen af total harmonisk forvrængning (*THD, total harmonic distortion*) er forholdet mellem effekten af forvrængningen og den totale effekt i signalet. En effekt kan skrives som summen af kvadratet på amplituden af de harmoniske der herunder kaldes h_1, h_2 etc.

$$THD = 100 \% \sqrt{\frac{h_2^2 + h_3^2 + \dots}{h_1^2 + h_2^2 + h_3^2 + \dots}} \quad \xrightarrow{\text{Aftagende niveau}} \quad \begin{aligned} D_2 &= 100 \% \frac{h_2}{h_1} \\ D_3 &= 100 \% \frac{h_3}{h_1} \end{aligned}$$

En savtaksvingning har en amplitude af de harmoniske på $h_n = 1/n$ så den harmoniske forvrængning bliver 70 % i forhold til grundtonen alene. En firkantsvingning har samme relation, men den mangler alle lige harmoniske så den giver en forvrængning på 42 %.

Forvrængning opfattes normalt som noget uønsket der skal undgås, men forvrængning kan være ønskværdig. Et eksempel er guitaristens *fuzz box* hvor to dioder begrænser signalets amplitude. En tones frekvens ændres ikke af overstyringen, men efterfølgende kredsløb kan få et mere veldefineret signal at arbejde med hvilket udnyttes i en FM radio til at undertrykke en uønsket amplitudemoduleret støj.

For den typiske forstærker er de harmoniske svage og det tillader en simplificering af formlen. Hvis det antages at de harmoniske falder hurtigt i niveau kan nævneren simplificeres til h_1 og hvis h_2 dertil er dominerende vil formlen kunne reduceres til udtryk D_2 der anvendes i kapitel fem til at beskrive den anden harmoniske forvrængning i en transistor. Tredje harmoniske forvrængning D_3 følger af en symmetrisk karakteristik som findes ved en operationsforstærkers differentialtrin.

Hørbarheden af harmonisk forvrængning er årsag til en livlig debat i litteraturen og på internettet. Der er videnskabeligt grundlag for at en *THD* under 0,25 % ikke kan høres af trænede lyttere hvis forvrængningen er domineret af lave harmoniske.

¹¹⁶ 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.

3.12 Opgaver

Analyse af DC forhold.

1 – Side 153 vises det matematiske udtryk for udgangens spænding U_O som funktion af differensen mellem de to indgange U_P og U_N . Brug det til at vise at bufferen side 154 vil have en forstærkning tæt på én. **Tip:** sæt $U_N = U_O$ og indsæt i ligningen. Løs ligningen for U_O som funktion af indgangssignalet U_P og vis at det giver følgende udtryk. Lad A_{OL} vokse over alle grænser hvilket matematisk kan udføres ved at dele i tæller og nævner med A_{OL} og derefter lade A_{OL} vokse over alle grænser.

$$U_O = \frac{A_{OL}}{1 + A_{OL}} U_P \xrightarrow{A_{OL} \rightarrow \infty} U_O \approx U_P$$

2 – Forstærkningen angiver forholdet mellem udgangens og indgangens amplitude. Vis at forstærkningen A bliver:

$$A = \frac{U_O}{U_P} \approx 1$$

3 – Åben-sløjfe forstærkningen A_{OL} har typisk en værdi på 300 000. Beregn forstærkningen. **Svar:** $A = 0,999\ 997$.

4 – For den ikke-inverterende forstærker på side 155 skal det vises at forstærkningen bliver som angivet. **Tip:** Sæt U_N til U_O gange med spændingsdelingen mellem R_1 og R_2 . Lad derefter A_{OL} vokse over alle grænser.

5 – Med åben-sløjfe forstærkningen $A_{OL} = 300\ 000$ skal forstærkningen beregnes med de angivne modstande. **Svar:** $A = 2,999\ 97$.

6 – Find operationsforstærkeren LM324 fra fx Texas Instruments på nettet og søg efter specifikationen af dens åben-sløjfe forstærkning, spændingsfejl og biasstrømme. Benyt minimum data for åben-sløjfe forstærkning og maksimum for de øvrige parametre for A-varianten **Svar:** $A_{OL} = 25\ 000$, $U_{OFS} = 3\ \text{mV}$, $I_{BIAS} = 100\ \text{nA}$ og $I_{OFS} = 30\ \text{nA}$.

7 – Side 155 vises en ikke-inverterende forstærker med en indstillet forstærkning på 3 gange. Hvis der benyttes en LM324 med en effektforsyning på $\pm 15\ \text{V}$ skal det vises at udgangssignalet ikke kan overstige $\pm 13\ \text{V}$ ved en sinus. **Tip:** Find data for *Output Voltage Swing* i databladet. Vis derefter at det maksimale indgangssignal bliver $\pm 4,3\ \text{V}$.

8 – Side Error: Reference source not found angives at et voltmeter med $10\ \text{M}\Omega$ indre modstand belaster dioden hårdt ved lav strømstyrke og at en operationsforstærker kan fungere som buffer. Giv en begrundelse for om LM324 vil kunne benyttes som buffer. **Svar:** LM324 giver ingen forbedring.

9 – Side Error: Reference source not found vises en forstærker for termokobler med AD8628 fra Analog Devices. Find databladet på nettet og giv en vurdering af hvor meget udgangen kan forventes at være forskudt på grund af offset voltage med de viste komponenter og hvor stor en fejlvisning giver det. **Svar:** Specifikationen $U_{OFS} = 10\ \mu\text{V}$ og en forstærkning på cirka 200 gange giver et DC potentiale på udgangen på cirka $\pm 0,2\ \text{mV}$ og fejlvisningen bliver på cirka $\pm 0,02^\circ\text{C}$. Talværdien afhænger af den valgte type termokobler.

10 – Forstærkeren for termokobler føres til en ADC med et indgangsområde på $0 \dots 5\ \text{V}$. Hvis termokobleren har den følsomhed som teksten angiver skal den højeste målbare temperatur T_{MAX} angives med 25°C som reference. Angiv også det nødvendige antal bit N for at måle med 1°C opløsning. **Svar:** $T_{MAX} = 525^\circ\text{C}$ og $N = 9$ bit.

11 – Side 156 er vist en inverterende forstærker. Signalet føres ind gennem modstand R_1 som belaster kilden. Vis at indgangsstrømmen ved $1\ \text{V}$ er på $100\ \mu\text{A}$.

Analyse af forstærkning og båndbredde.

12 – Mikrofonforstærkeren på side Error: Reference source not found forstærker AC med faktor A givet ved modstandene i tilbagekoblingen og forstærkningen skal beregnes. Kondensator C_1 sidder i serie med den forstærkningsbestemmende modstand R_2 og det giver den lave grænsefrekvens f_L når reaktansen C_1 er lig med modstanden. Bestem den lave grænsefrekvens. Find GBP parameteren ved at søge på nettet og beregn den høje grænsefrekvens f_H . **Svar:** $A = 201$, $f_L \approx 160\ \text{Hz}$, $GBP \approx 10\ \text{MHz}$ og $f_H \approx 50\ \text{kHz}$.

13 – Forstærkeren for en fotodiodes strøm side Error: Reference source not found fungerer som et lavpasfilter. Vis at overføringsfunktionen bliver som vist herunder. **Tip:** Antag at diodens strøm I_R løber gennem tilbagekoblingens impedans så udgangens spænding bliver $U_O = Z_{FB} I_R$ hvor Z_{FB} er impedansen af R_{FB} i parallel med C_{FB} . Vis også at grænsefrekvensen med de viste komponentværdier er $59\ \text{kHz}$.

$$U_O = \frac{\omega_{FB}}{\omega_{FB} + s} R_{FB} I_R \quad \text{hvor} \quad \omega_{FB} = \frac{1}{RC}$$

14 – Spændingsfølgeren på side 170 opfører sig som et lavpasfilter med grænsefrekvensen $f_0 = \omega_0/2\pi$. Brug en tabel over Laplace transformationen (fx fra andet kapitel) for at vise at stepresponsen bliver som angivet. Vis at udgangsspændingen er $1\ \%$ fra slutværdien ved $t_1 = 4,6\tau$. **Tip:** sæt $u_O(t_1) = 0,99\ U_M$ og løs ligningen for t_1 .

15 – Eftersis udtrykket for relationen mellem stigetid og grænsefrekvens på side 173. **Tip:** Benyt resultatet ovenfor til at vise at $t_{90} = 2,3025/\omega_0$ og $t_{10} = 0,1053/\omega_0$. Relationen bliver $t_R = 2,1972/\omega_0$ og ved indsættelse af udtrykket for ω_0 kan relationen eftervises.

16 – Slew rate specifikationen for en operationsforstærker angiver den maksimale hastighed i volt per sekund som udgangen kan flytte sig med. For signalet på side 173 skal det vises at $du_O/dt = 2\pi f_0 U_M \cos(2\pi f_0 t)$ og der skal gives en argumentation for at cosinus er sat til én.

Analyse af støj.

17 – En signalkilde med RMS værdien $U_S = 1\ \text{mV}$ har en indre modstand $R_S = 100\ \text{k}\Omega$ som forstærkes med $A = 100$ og båndbredden er $B = 20\ \text{kHz}$ (se side 187). Beregn støjætheden e_{nR} for modstanden, RMS værdierne på udgangen for signalet U_O og støjen U_{nO} samt det opnåede signal/støl-forhold SNR . **Svar:** $e_{nR} \approx 41\ \text{nV}/\sqrt{\text{Hz}}$, $U_O = 100\ \text{mV}$, $U_{nO} \approx 0,58\ \text{mV}$ og $SNR \approx 45\ \text{dB}$.

18 – Forstærkeren i den forrige opgave skal vælges så dens støjegenskaber matcher modstandens støjniveau. For at undgå for meget stigning i den samlede støj søges efter en specifikation for både spændingsstøjen e_{nA} og strømstøjen $i_{nA}R_S$ hver på halvdelen af e_{nR} . Bestem det samlede støjniveau e_{nTOT} ved forstærkerens indgang, det nye signal/støj forhold SNR og de specifikationer for e_{nA} og i_{nA} som forstærkeren skal have. Undersøg derefter om TL071 fra Texas Instruments kan imødekomme kravene.
Svar: $e_{nTOT} \approx 50 \text{ nV}/\sqrt{\text{Hz}}$, $SNR \approx 43 \text{ dB}$, $e_{nA} \approx 20 \text{ nV}/\sqrt{\text{Hz}}$, $i_{nA} \approx 0,2 \text{ pA}/\sqrt{\text{Hz}}$ og TL071 kan løfte opgaven.

19 – TL071 skal i opstillingen fra de foregående opgaver benyttes til at forstærke et signal i området 10 ... 20 Hz. Bestem en tilnærmet værdi af spændingsstøjtætheden e_{nA} for dette frekvensområde ud fra dens datablad og giv et bud på signal/støj forholdet for signalet $U_S = 1 \text{ mV}$ i dette frekvensområde. **Svar:** $e_{nA} \approx 41 \text{ nV}/\sqrt{\text{Hz}}$ og $SNR \approx 41 \text{ dB}$.

20 – En analog lydмиксерпульт har 64 kanaler og hver kanal

tænkes at referere til et 0 dB signalniveau på 1 V med et signal/støj-forhold på 100 dB. Beregn støjniveauet U_n for hver enkelt kanal samt det resulterende støjniveau U_{nTOT} hvis alle kanaler er indstillet ens og signalerne adderes.
Svar: $U_n \approx 10 \text{ } \mu\text{V}$ og $U_{nTOT} \approx 80 \text{ } \mu\text{V}$.

21 – En antennemodtager har modstandsværdien $50 \text{ } \Omega$ og båndbredden 10 kHz. Beregn støjspændingen U_{n1} . For at reducere støjen køles antennemodtageren ned til -270°C hvor værdien af den reducerede støjspænding U_{n2} skal estimeres. **Svar:** $U_{n1} \approx 90 \text{ nV}$ og $U_{n2} \approx 9 \text{ nV}$.

22 – Antennemodtageren fra forrige opgave har et for højt støjniveau. Den benyttes til at modtage digitale data med en *Bitrate* på 10 kbit/s. Hvad skal *Bitrate* (og dermed båndbredden) reduceres til for at signal/støj-forholdet kan forbedres med 10 dB? **Svar:** *Bitrate* = 1 kbit/s.

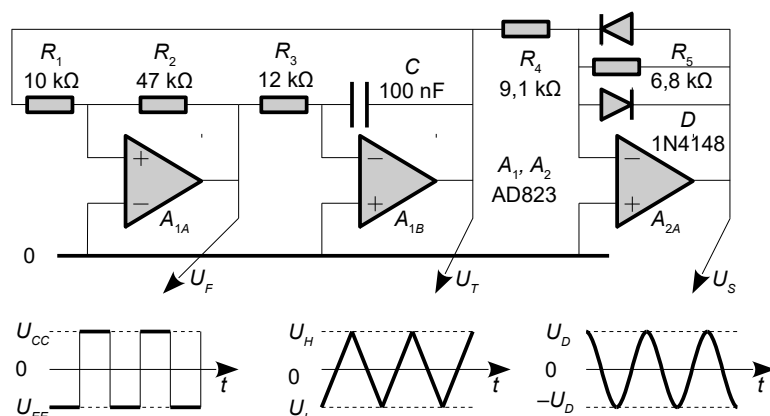
Til sammenligning har Rosetta satellitten en datahastighed på 2 kbit/s så det tager sin tid at sende billeder til jorden.

4 DESIGN

Dette kapitel viser hvordan man kan bygge oscillatorer, omsætte mellem forskellige fysiske størrelser og filtrere et signal. Dertil nogle råd om hvordan elektroniske kredsløb bør opbygges.

4.1 Oscillatorer

Herunder vises en tonegenerator for dannelse af firkant, savtak og sinus. Formlen for frekvensen af oscillationen gælder for operationsforstærkere af typen rail-to-rail hvor udgangen skifter helt ud til effektforsyningens grænser fx TLC274 og AD823. Med de viste komponenter (1 % tolerance) giver formelen 980 Hz og frekvensen blev målt til 960 Hz der er blot 2 % fra det beregnede.



Oscillatorfrekvens:

$$f_{osc} = \frac{R_2}{R_1} \frac{1}{4 R_3 C}$$

$$f_{osc} < \frac{SR}{20(U_{CC} - U_{EE})}$$

Spændingsamplitude:

$$U_H = -\frac{R_1}{R_2} U_{EE}$$

$$U_L = -\frac{R_1}{R_2} U_{CC}$$

Figur 213 – En oscillator bygget over en Schmitt-trigger der med de viste komponentværdier vil svinge på 1 kHz med ±5 V firkant, ±1 V triangel og ±0,6 V sinus ved ±5 V effektforsyning.

Forstærker A_{1A} er en Schmitt-trigger hvor udgangen enten er høj ($U_F = U_{CC}$) eller lav ($U_F = U_{EE}$). Tilstanden skifter som følge af niveauet fra U_T. For U_F lav skal U_T over U_H = 1,06 V ved ±5 V fra effektforsyningen for at U_F skifter høj. Tilsvarende skal indgangen under U_L = -1,06 V for at skifte U_F lav igen. Oscillatoren er selvstartende da udgangen fra A_{1A} altid vil være enten høj eller lav.

Forstærker A_{1B} er integrerende og med lavt niveau fra Schmitt-triggeren stiger udgangen U_T med en fast hastighed givet af strømmen gennem R₃ og kondensatoren C. Når niveauet kommer over den værdi hvor Schmitt-triggerens udgang skifter høj (U_H) ændres retningen så U_T integreres ned mod den lave grænse (U_L). Dette forløb fastsætter oscillatorens frekvens f_{OSC}.

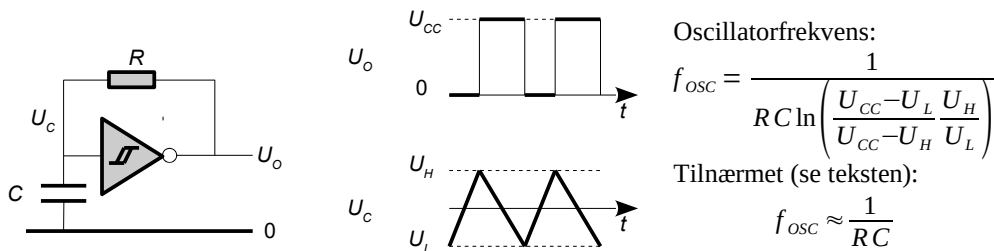
Forstærker A_{2A} benytter to dioder til at begrænse amplituden så den tilnærmes en sinus. Med de viste komponenter blev der målt en forvrængning på 2 % tredje harmoniske ved 1 kHz.

Schmitt-triggeren behøver tid for at skifte mellem yderstillingerne og det giver en øvre grænse for den frekvens oscillatoren kan arbejde ved. Det betegnes med *slew-rate* parameteren der angiver hvor hurtigt udgangen kan ændre sig. Den viste operationsforstærker har SR = 20 V/μs så et skift varer 0,5 μs ved ±5 V effektforsyning. En hel periode behøver 1 μs til skiftene så frekvensen skal holdes langt under 1 MHz. Hvis oscillationens periode er 10 μs (som svarer til 100 kHz) vil skiftet vare 10 % af tiden så amplituden af trekanten fra U_T bliver 10 % for høj og oscillatorens frekvens bliver 20 % under det beregnede hvilket passede med en måling.

En variation i effektforsyningen påvirkede frekvensen mindre end 1 % fra ±5 V ned til ±3 V, men kvaliteten af sinus-svingningen forringes væsentligt. Modstanden R₃ bør være fra 10 kΩ til 10 MΩ og kondensatoren C bør være plastfilm fra 1 nF til 1 μF så frekvensen bliver fra 0,1 Hz til 100 kHz. Som afkobling af U_{CC} og U_{EE} blev benyttet 1 μF tantal for hver forsyningslinje.

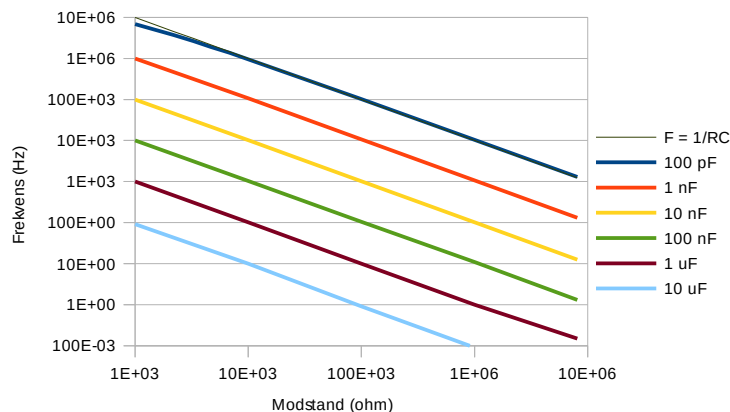
4.1.1 CMOS oscillator

Oscillatorer kan bygges med Schmitt inverter 74HC14 og Schmitt NAND 74HC132 der begge er bygget i CMOS teknologi. Komponenterne stiller udgangen lav når indgangen kommer over den høje tærskelværdi og den skifter først høj når indgangen kommer under den lave tærskelværdi. Med R og C dannes en oscillator som ramper op og ned mellem tærskelværdierne 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 214 – En CMOS inverter med hysteresis giver mulighed for en simpel oscillator der kan benyttes op til nogle få megahertz.

Formlen antager at tærskelværdierne er kendt, men specifikationerne for CMOS er meget usikre. Med opgivne data for 74HC132 beregnes en frekvens der er næsten det dobbelte af det målte så den tilnærmede formel er fundet eksperimentelt. Den ekstra indgang ved 74HC132 kan benyttes til start/stop af oscillatoren styret fra en mikroprocessor.



Figur 215 – Målt frekvens med 74HC132 ved 5 V. Der blev benyttet 1+2 for indgang og 3 for udgang.

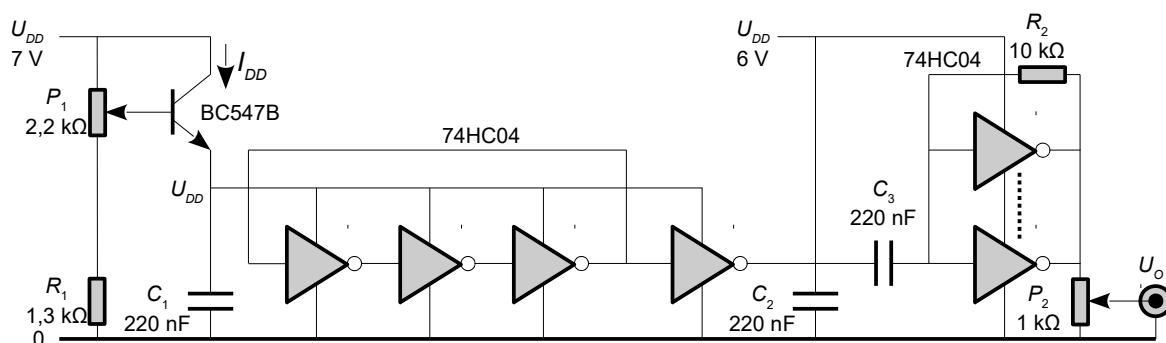
Oscillatoren kan benyttes indenfor det store frekvensområde fra 0,1 Hz til 10 MHz. Ved de høje frekvenser afviger det målte fra beregningen da den digitale kreds behøver 15 ns på at skifte mellem niveauerne så oscillatoren bruger 30 ns af hele perioden på at skifte og frekvensen bliver lavere end beregnet. Hvis skiftet varer 10 % af hele perioden bliver grænsen omkring 3 MHz så ved høje frekvenser kan formelen ikke forventes at passe.

Med en kondensator under 100 pF skal du være opmærksom på at der er en indgangskapacitet på den digitale komponent på cirka 3 pF og tilsvarende fra banerne på printkortet. Det virker som en forøgelse af kondensatorens værdi og det bliver mærkbart ved små værdier.

Ved lave frekvenser betyder lækstrømmen i kondensatoren at modstanden både skal levere strøm til opladning af kondensatoren og til vedligeholdelse af ladningen. Det begrænser den øvre værdi af modstanden der ikke bør komme over 1 MΩ ved brug af elektrolytiske kondensatorer.

4.1.2 Spændingsstyret oscillator

De foregående oscillatorer er begrænset af den tid den aktive komponent er om at skifte mellem lav og høj som følge af en ændring på indgangen (*propagation delay* t_d). Det er muligt at basere oscillationen på denne parameter for med tre invertorer i en lukket sløjfe vil udgangens niveau løbe igennem kæden og ende med at skifte til modsat stilling så kæden oscillerer på en frekvens der er givet af den samlede tidsforsinkelse. Ved CMOS er propagation delay en funktion af spændingen så oscillatorens frekvens kan styres ved at variere spændingen. For 74HC04 er $t_d = 45$ ns ved 2 V og aftager til 8 ns ved 6 V så et forsigtigt bud på frekvensen er fra 4 MHz til 20 MHz. En buffer giver konstant amplitude og med seks invertorer i parallel bliver udgangsmodstanden lav.

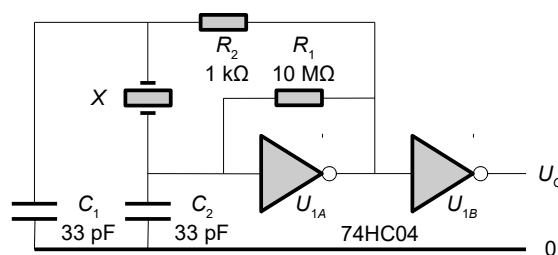


Figur 216 – En oscillator baseret på komponentens spændingsafhængige propagation delay.

Jeg har brugt opstillingen med 74AC04PC der dækkede fra 40 MHz ved 2 V til 100 MHz ved 6 V. Strømforbruget I_{DD} nåede op til 100 mA som gav 30°C temperaturstigning.

4.1.3 Krystaloscillator

Et krystal kan give en præcis indstilling af frekvensen og denne type oscillatorer benyttes flittigt ved mikroprocessorer. Herunder vises en såkaldt Pierce oscillator hvor modstand R_1 sikrer at inverter U_{1A} holdes i det aktive område mens kondensatorerne C_1 og C_2 danner resonans med krystallets selvinduktion. Oscillatoren svinger tæt på krystallets parallelresonans. For at undgå oscillation på en overtone kan modstand R_2 inkluderes for at danne et lavpasfilter med C_1 så forstærkningen reduceres ved høje frekvenser.^{117,118,119}



Figur 217 – En 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_{1B} er en buffer der sikrer at udgangen ved U_O kan drive et tilsluttet kredsløb uden at det influerer på selve oscillatorens funktion. Udgangssignalet er en firkant som er velegnet til at drive digitale kredsløb. Det kan være nødvendigt at justere på de viste komponentværdier for at få en stabilt virkende funktion ved den valgte frekvens.

¹¹⁷ https://en.wikipedia.org/wiki/Pierce_oscillator.

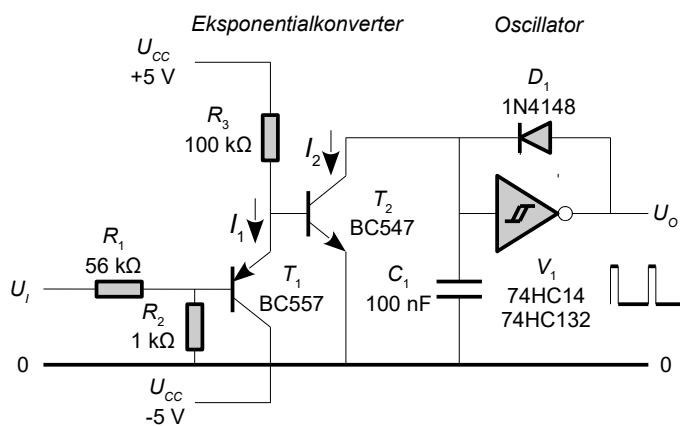
¹¹⁸ https://en.wikipedia.org/wiki/Crystal_oscillator.

¹¹⁹ <http://www.electronics-tutorials.ws/oscillator/crystal.html>.

4.1.4 Eksponentielt spændingsstyret oscillator

Den analoge musiksynthesizer blev introduceret omkring 1920 og blev brugt af en del klassiske komponister. Omkring 1970 slog den igennem ved jazz og rock. Et populært musikinstrument var Minimoog hvor tonegeneratorer og filter blev styret af en elektriske spænding fra et klavier med en eksponentiel karakteristisk på 1 oktav/volt så frekvensen steg til det dobbelt ved én volt stigning. En ændring på 1/12 volt svarer til et halvtone trin i den ligesvævende temperatur.^{120,121}

Herunder vises et eksempel på en eksponentialkonverter der driver en oscillator. Der er ikke en præcisionsoscillator, men eksemplet indeholder de basale dele. Analysen bygger på materiale fra femte kapitel. Se også litteraturen.¹²²



Eksponentialkonverter:

$$I_2 = I_{S2} \exp\left(\frac{U_{BE2}}{U_T}\right), \quad U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$$

$$U_{BE2} = \frac{R_2}{R_1 + R_2} U_I + U_{BE1}$$

$$U_{BE1} = U_T \ln\left(\frac{I_1}{I_{S1}}\right)$$

$$I_2 = I_1 \frac{I_{S2}}{I_{S1}} \exp\left(\frac{R_2}{R_1 + R_2} \frac{U_I}{U_T}\right)$$

Oscillator:

$$dq = C_1 dU = I_2 dt \Rightarrow dt = \frac{C_1 dU}{I_2}$$

$$f_{osc} = \frac{1}{dt} = \frac{I_2}{C_1 dU} = \frac{I_2}{C_1 (U_H - U_L)}$$

Figur 218 – En eksponentielt styret oscillator hvor en spændingsændring på indgangen på 1 V giver en fordobling af strømmen i kollektor. Udgangen kan drive en frekvensdeler som HEF4040.

Oscillatoren består af en Schmitt-trigger koblet som relaxations-oscillator hvilket 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 hvor udgangen går høj indtil spændingen er nået op til den høje tærskelværdi U_H . Frekvensen kan beregnes af kondensatorligningen ved C_1 for strømmen I_2 og differensen mellem de to tærskelværdier der er målt til $U_H - U_L = 1,15 \text{ V}$ for 74HC132.

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 vil strømmen stige eksponentielt med styrespændingen. Spændingsdelingen 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 $I_1 = 44 \mu\text{A}$ med diagrammets komponentværdier. Antages transistorerne ens på nær forskellen i polaritet bliver oscillatorens frekvens $f_{osc} \approx 380 \text{ Hz}$ for 74HC132 ved nul volt på indgangen. Frekvensen blev målt til 700 Hz ved $U_I = 0 \text{ V}$ så det kan konkluderes at transistorerne ikke er ens og at $I_{S2}/I_{S1} \approx 1,8$. Relationen på 1 oktav/volt gælder ved 25°C og kan temperaturkompenseres hvis R_2 har temperaturkoefficienten $0,33 \text{ }^\circ\text{C}^{-1}$ som fx opnås hvis den vikles af tynd kobbertråd. Ved mere end 3 V ind bliver frekvensen usikker da oscillatoren benytter 4 μs på opladning fra U_L til U_H og det giver 2 % fejl omkring 6 kHz. Det kan forbedres ved at benytte en mindre kapacitetsværdi. Ved større strøm afsættes en mærkbar effekt i T_2 som opvarmer den og det påvirker stabiliteten af frekvensen.

¹²⁰http://en.wikipedia.org/wiki/Analog_synthesizer.

¹²¹<http://www.torean.dk/artikel/Zink.pdf> for en introduktion til musikalsk temperatur (side 8 til 18).

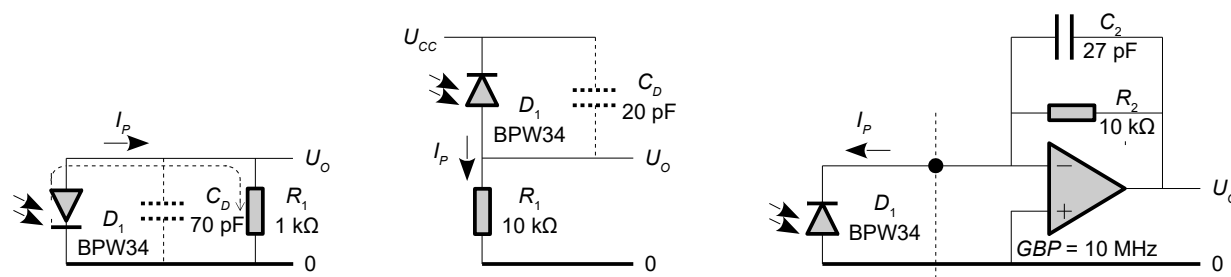
¹²²<http://yusynth.net/archives/WirelessWorld/Electronic-sound-Synthesizer.pdf>.

4.2 Omsætning af fysiske størrelser

Her følger eksempler på måling af lys, temperatur, acceleration og lyd.

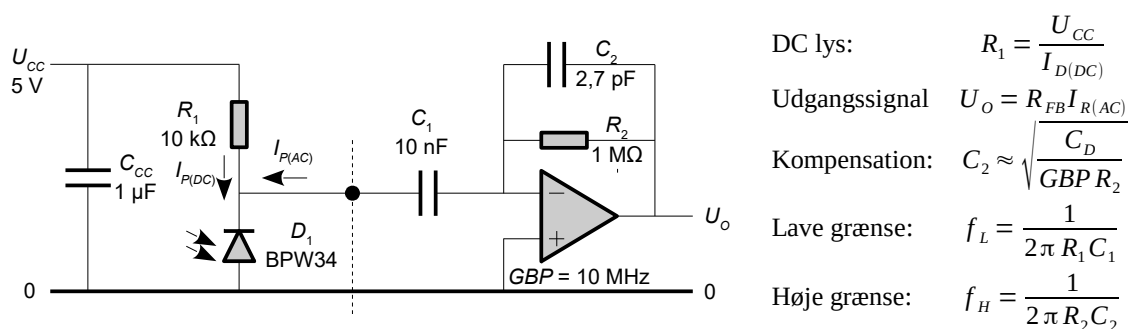
4.2.1 Lysmåler

En fotodiode genererer en strøm I_P som reaktion på det indfaldne lys hvis strøm af fotoner danner elektron-hulpar i diodens halvleder. Strømrretningen er imod pileretningen og forklaringen gives i femte kapitel. Strømmen kan måles med et amperemeter eller omsættes til en spænding ved en modstand for måling med et voltmeter. Spændingen over dioden skal holdes under 0,5 V for ellers vil dioden selv optage strømmen. Solpaneler benytter mange dioder i serie for at opnå høj spænding og mange i parallel for høj strømstyrke. De små fotodioder vi benytter i elektronikken giver blot nogle få milliampere ved direkte sollys.



Figur 219 – En fotodiode genererer strøm som reaktion på det indfaldne lys.

For at være praktisk anvendelig til lysmåling forspændes dioden af en ekstern spændingskilde så strømmen ikke løber gennem dioden når spændingen over modstanden stiger. Der er en kapacitet fra den spærrende diode som er vist stipleet i illustrationen. Den giver en tidskonstant sammen med modstanden og det påvirker båndbredden. Kapaciteten mindskes ved øget spænding over dioden i spærreretningen. En operationsforstærkers virtuelle nul leder strømmen gennem R_2 så udgangen kan drive en lav belastningsmodstand. Kondensator C_2 indskydes for at kompensere for diodens kapacitet C_D der danner en pol med R_2 og kan gøre operationsforstærkeren ustabil.

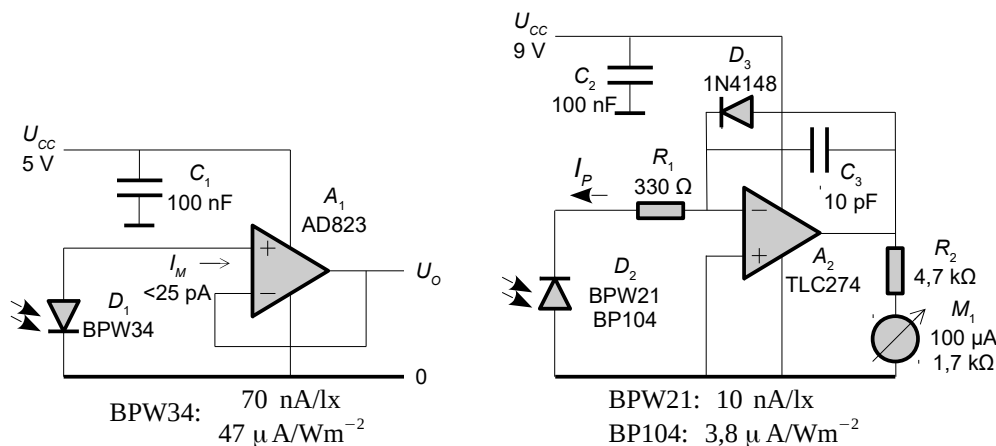


Figur 220 – En typisk opstilling hvor det absolutte lysniveau ignoreres og kun ændringer føres til forstærkeren.

Industrielle sensorer benytter pulset lys for at undertrykke baggrundsbelysningen og mindske den gensidige påvirkning mellem sensorer ved brug af individuelle frekvenser. Modstanden R_1 optager DC strømmen og kun de hurtige variationer overføres gennem C_1 der har så lille en værdi at de langsomme ændringer i omgivelsernes lysstyrke ikke overføres til udgangen. Forstærkningen af det oscillerende signal kan øges da R_2 ikke skal bære et DC signal. Støj fra effektforsyningen føres til operationsforstærkeren gennem R_1 og forstærkes med R_2/R_1 så effektforsyningen bør afkobles og C_{CC} er et minimum.

Diodens logaritmiske karakteristisk kan benyttes til at komprimere det optiske område fra svagt lys til stærkt lys så det kan vises på et måleinstrument uden områdeskift eller den logaritmiske relation kan benyttes til at efterligne øjets og hjernens opfattelse af lysændringer.¹²³

Opstillingen herunder til venstre lader fotostrømmen løbe gennem dioden selv og bufferen gør det muligt at måle spændingen uden at belaste dioden. Spændingen over dioden ændres med 60 mV per dekade ændring i lysstyrken fra cirka 0,7 V ved direkte solskin til 0,4 V ved netop synligt lys. Strømmen er her så lav at selv et voltmeter med 10 M Ω indre modstand ville give fejlmåling så en operationsforstærker med JFET eller MOSFET i indgangen er nødvendig. Fotodioden BPW34 er følsom for lys med en bølgelængde fra 1100 nm der er i det nær-infrarøde område ned til 450 nm ved 10 % af fuld følsomhed. Det dækker en stor del af bølgelængdeområdet fra en glødepære. Størst følsomhed ligger ved 900 nm så det synlige lys fra 800 nm (rød) til 400 nm (blå) vil let blive "overdøvet" af den infrarøde andel af spektret.



Figur 221 – Lysmåler med logaritmiseret visning. BPW34 dækker både det nær-infrarøde og det synlige område, BPW21 dækker det synlige område og BP104 dækker det nær-infrarøde område.

Fotodioderne kan leveres med filtre som indfarvet plast der blokerer for en del af spektret så der udelukkende måles infrarødt eller synligt lys. Kredsløbet til højre har jeg brugt i en batteridrevet lysmåler der kunne vise synligt lys fra 100 klx (stærkt solskin) ned til 1 lx (netop synligt lys) på et drejespoleinstrument uden behov for omskiftning af område. Fotodioden BPW21 er forsynet med et filter der fjerner det infrarøde lys så kun det synlige lys bliver registreret. Fotodioden BP104 er tilsvarende forsynet med et filter der fjerner det synlige lys så kun det infrarøde lys bliver målt. Logaritmeringen med 1N4148 giver en variation på 100 mV/dekade så det angivne lysområde dækker stort set hele instrumentets skala. Instrumentet blev "kalibreret" ved at injicere en kendt strømstyrke til den virtuelle nul og markere meterudslaget på skalaen.¹²⁴

4.2.2 pH-probe

Et kredsløb som det venstre kan benyttes ved en pH-probe der danner en spændingsforskel over to sider af en tynd membran af glas så den indre modstand er enorm. Spændingen varierer med cirka 60 mV per dekade ændring i koncentrationen af brintioner fra under -0,4 V ved stærke syrer til over +0,4 V ved stærke baser.

Den høje impedans af pH-proben gør opstillingen følsom for 50 Hz brum så der bør inkluderes et filter som dæmper fra cirka 5 Hz og opad. Et RC led på udgangen med 33 k Ω og 1 μ F keramisk vil nok være et minimum.¹²⁵

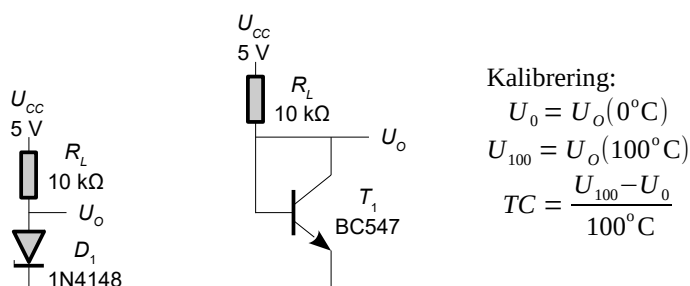
¹²³https://en.wikipedia.org/wiki/Stevens%27_power_law.

¹²⁴Min artikel "Optoelectronics by Design" publiceret i Electronics World and Wireless World, May 1994, side 365.

¹²⁵<https://www.sparkyswidgets.com/portfolio-item/ph-probe-interface/>.

4.2.3 Termometer

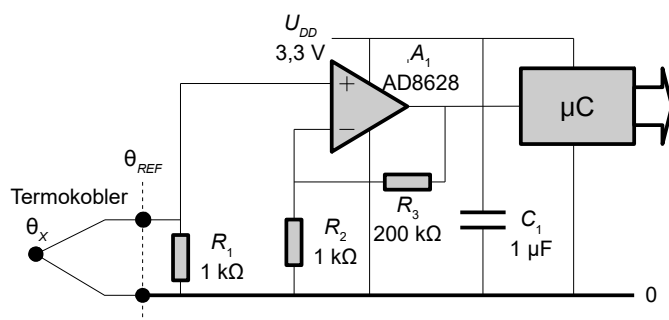
En siliciumdiode har et spændingsfald omkring 0,7 V der varierer med cirka $-2 \text{ mV}/^\circ\text{C}$ og det kan udnyttes som et termometer. Det er uden betydning om dioden er en almindelig diode, en lysdiode eller en diode-koblet transistor. Ved en lysdiode er spændingsfaldet dog omkring 1,8 V. For at være brugbart skal termometret kalibreres hvilket kan udføres at holde dioden i en blanding af is og vand for at finde spændingen U_0 ved 0°C og derefter i kogende vand for at finde U_{100} ved 100°C . Ud fra disse værdier kan man beregne den aktuelle værdi af temperaturkoefficienten TC .



Figur 222 – En diodes temperaturkoefficient kan udnyttes som termometer.

Spændingen U_o er en funktion af strømmen i dioden så det er vigtigt at strømmen holdes konstant, men det er ofte tilstrækkeligt med en modstand og en fast spændingskilde.

For måling af høje temperaturer benyttes en termokobler. Den består af to metaltråde der er i elektrisk kontakt. Ved en temperaturforskel mellem kontaktstedet (θ_x) og referencestedet (θ_{REF}) opstår der en spændingsforskel mellem metaltrådene. Forklaringen ligger i generation af frie elektroner der stiger med temperaturen og danner en spændingsforskel over lederen når enderne har hver sin temperatur. For at få en spændingsforskel mellem to ledere skal de være dannet af forskellige materialer som har forskellig grad af villighed for generation af frie elektroner.¹²⁶



$$U_o = (\theta_x - \theta_{REF}) S A$$

hvor:

- θ_x = Temperatur (fx 100°C)
- θ_{REF} = Reference (fx 25°C)
- S = Følsomhed (fx $50 \mu\text{V}/^\circ\text{C}$)
- $A = 1 + \frac{R_3}{R_2} = 201$

$U_o = 0,75 \text{ V}$ med eksemplets værdier

Figur 223 – En termokobler genererer en svag spænding, som skal forstærkes op for at kunne måles med en A/D konverter. Opstillingen måler positiv temperatur overfor referencepunktet.

En termokobler afgiver et meget svagt signal på $10 \dots 70 \mu\text{V}$ per grad celsius. Det er for lidt til selv en 16 bit A/D konverter så der er behov for en forstærkning der løfter signalet op til et niveau på fx $10 \text{ mV}/^\circ\text{C}$ og det indstilles af R_2 og R_3 . Modstanden R_1 holder indgangen på nul når termokobleren fjernes og værdien er ukritisk.¹²⁷

Forstærkeren AD8628 er chopper-stabiliseret så indgangens fejlspænding er på blot $1 \mu\text{V}$ og vil ikke influere på resultatet. Operationsforstærkeren kan arbejde med signaler nær ved nul og den trækker blot 1 mA fra effektforsyningen. Kondensatoren C_1 afkobler effektforsyningen for at opnå en stabil funktion og bør altid inkluderes uanset hvad man arbejder med.

¹²⁶ <https://en.wikipedia.org/wiki/Thermocouple>.

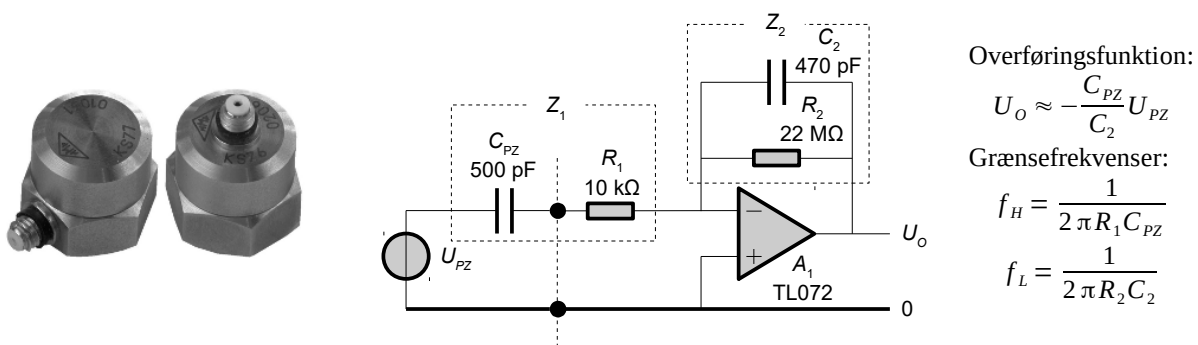
¹²⁷ En E-type termokobler består af NiCr og CuNi legeringer og har spændingsforskellen $68 \mu\text{V}/^\circ\text{C}$.

4.2.4 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 og benyttes i lightere til at danne en gnist, men som regel er signalets amplitude moderat. De piezo-elektriske krystaller blev brugt i den ældre grammofons pick-up hvor de omsatte rillens bevægelser til elektriske spændinger. De finder desuden anvendelse i visse højtalere 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 at 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 høj indgangsmodstand og den anden er at belaste krystallet så hårdt at det er strømmen der måles. Det sidste kaldes for en ladningsforstærker (*charge amplifier*) og skal eksemplificeres her.

Det piezo-elektriske krystal kobles til en virtuel-nul indgang hvilket effektivt kortslutter krystallet så strømmen bestemmes af krystallets spænding og impedansen af dets kapacitet. 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 224 – Et piezo-elektrisk accelerometer kan udformes så det er let at montere på en maskine for måling af vibration. Transduceren fungerer som en spændingsgenerator U_{PZ} i serie med en indre kapacitet C_{PZ} så ved belastning med en virtuel-nul indgang bliver signalet differentieret og det må derfor integreres for at genskabe transducerens spænding.

Et eksempel skal gives for interface til en sensor med 10 mV udgangsspænding for en acceleration på $g = 10 \text{ m/s}^2$ så følsomheden bliver $S = 1 \text{ mV/ms}^{-2}$. Forstærkningen i det viste kredsløb er én så følsomheden ændres ikke, men udgangsimpedansen bliver lav og kan forbindes til et efterfølgende dataopsamlingsystem.

Overføringsfunktionen beregnes ud fra impedansen Z_1 af C_{PZ} for det piezo-elektriske krystals kapacitet i serie med indgangsmodstanden R_1 samt af impedansen Z_2 for parallellforbindelsen af tilbagekoblingen ved R_2 og C_2 .

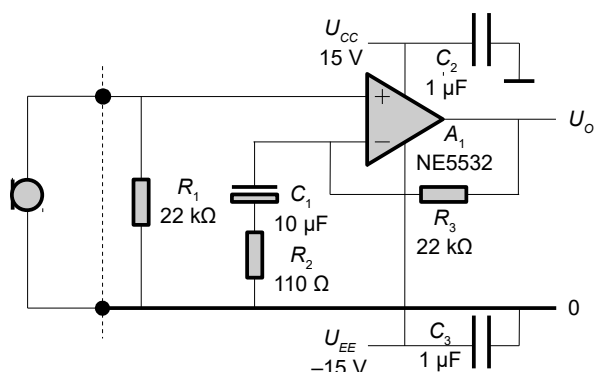
$$U_O = -\frac{Z_2}{Z_1} U_{PZ} \quad \text{hvor} \quad \begin{aligned} Z_2 &= \frac{R_2(1/sC_2)}{R_2 + 1/sC_2} \\ Z_1 &= R_1 + 1/sC_{PZ} \end{aligned}$$

Resultatet er et passivt båndpasfilter med grænsefrekvenserne $f_L = 15 \text{ Hz}$ og $f_H = 32 \text{ kHz}$. Indenfor dette område er udgangssignalet stort set uafhængigt af frekvensen og repræsenterer spændingen fra det piezo-elektriske krystal ganget med C_{PZ}/C_2 .

$$U_O = -\frac{sR_2C_{PZ}}{(1+sR_1C_{PZ})(1+sR_2C_2)} U_{PZ} = -\frac{C_{PZ}}{C_2} U_{PZ} \quad \text{for} \quad \frac{1}{2\pi R_2C_2} < f < \frac{1}{2\pi R_1C_{PZ}}$$

4.2.5 Lydmåling

En dynamisk mikrofon omsætter et akustisk signal til et tilsvarende elektrisk signal. Det akustiske signal får en let membran til at svinge i takt med lydsvingningerne og den bevæger en spole der er placeret i et magnetfelt. Det giver en elektrisk spænding over spolen efter Faradays lov. Niveaueet er ofte 5 mV ved 1 Pa og det er for lidt for at drive en effektforstærker eller en A/D konverter.¹²⁸



Indstilling af forstærkning A_p :

$$A_p = 1 + \frac{R_3}{R_2} \Rightarrow R_3 = (A_p - 1)R_2$$

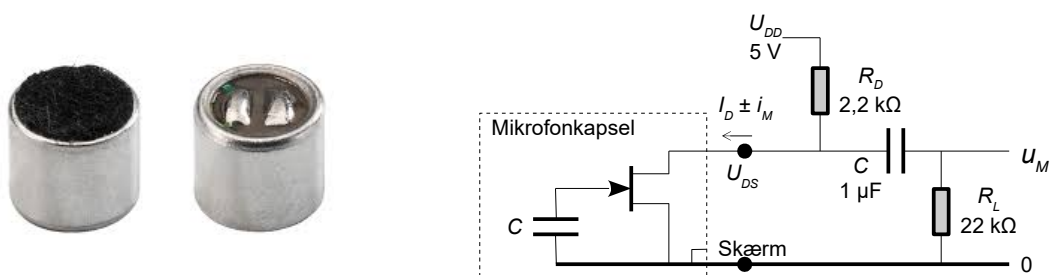
Grænsefrekvenser:

$$f_H = \frac{GBP}{A_p} \quad f_L = \frac{1}{2\pi R_2 C_1}$$

Figur 225 – En forstærker for dynamisk mikrofon.

Der er behov for at forstærke signalet og her forstærkes signalet 200 gange. Kondensator C_1 vil blokere for DC så forstærkerens offsetspænding ikke påvirker funktionen og den bestemmer den laveste frekvens der kan overføres til 13 Hz (–3 dB). Den højeste frekvens er givet af forstærkeren til 50 kHz (–3 dB) ved den valgte forstærkning. Indgangen holdes nær nul ved R_1 for at sikre et lavt potentiale på indgangen hvis mikrofonen udtages. Signalets amplitude kan måles med ensretteren side 223 hvor kondensator C_1 fx kan være 1 μF for at midle signalet til en DC spænding.

En populær mikrofontype er baseret på den kondensator der dannes af en tynd metalmembran og en bagvedliggende elektrode. Den tynde membran bevæges af lydfeltet og det ændrer værdien af kapaciteten. Hvis kondensatoren har en konstant ladning vil spændingen variere og det kan opnås med et dielektrikum hvor der er en "indfrosset" ladning hvilket kaldes for en elektret mikrofon. De sælges i store antal til lav pris. Lydkvaliteten er god, men egenstøj er ret høj.¹²⁹



Figur 226 – En elektret mikrofon har indbygget JFET for at mindske impedansen (Monacor).

Den lave værdi af kondensatoren på cirka 10 pF gør det upraktisk at føre signalet videre gennem et kabel så i stedet indbygges en JFET der omsætter spændingen til en strøm som bedre tåler at blive ført gennem et kabel. Følsomheden specificeres typisk til 5 mV/Pa hvor det forudsættes at mikrofonen belastes af 2 kΩ så det svarer til signalamplituden $i_M = 2,5 \mu\text{A/Pa}$. Transistoren behøver en DC strøm $I_D < 500 \mu\text{A}$ der leveres gennem R_D og udgangsspændingen U_M føres til en forstærker der kan være den samme som vist ovenfor.

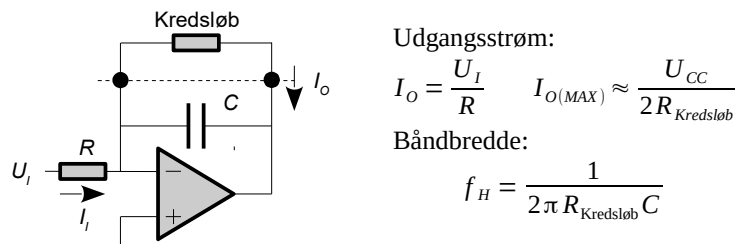
¹²⁸En mikrofons signalamplitude er produktet af lydtrykkets amplitude der måles i pascal og mikrofonens følsomhed på 5 mV/Pa som en typisk værdi. Et lydtryk på 1 Pa svarer til 94 dB SPL. Det niveau der findes umiddelbart udenfor ens mund når der tales højt, men uden at råbe er på 90 dB SPL. I en afstand på 1 m fra den talende ved en tilhørers øre er lydtrykket faldet til cirka 60 dB SPL.

¹²⁹<https://www.monacor.com/en-gb/monacor/products/installation-material/energy-measuring-and-diy/diy-mce-4000/>.

4.2.6 Strømkilder

Det er tit nødvendigt at styre en strøm som funktion af spændingen fra en mikroprocessor eller et en digital til analog konverter og her gives nogle eksempler på hvordan det kan klares.

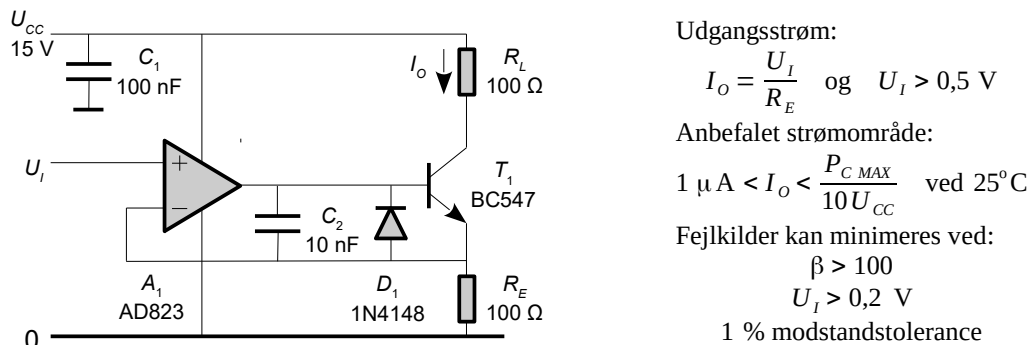
En operationsforstærker holder sin negative indgang nær ved den positive og med den positive på nul kan en modstand R i indgangen omsætte en spænding til en strøm der kan føres gennem et eksternt kredsløb. En ulempe ved kredsløbet er at det eksterne kredsløb skal "flyde" i forhold til operationsforstærkerens effektforsyning og nul. Kredsløbet indgår i tilbagekoblingen hvilket gør det svært at opnå en stabilt fungerende opstilling hvis kredsløbets impedans er kompleks.



Figur 227 – En operationsforstærker kan omsætte strøm til spænding.

For at fungere skal spændingen på udgangen holdes sikkert indenfor effektforsyningens grænser hvilket i formelen ovenfor repræsenteres ved $U_{CC}/2$. Det sætter en grænse for den impedans det eksterne kredsløb kan antage. Strømmen skal også være indenfor operationsforstærkerens rammer der som regel vil begrænse den til ± 10 mA. Kondensator C er påkrævet for stabilitet og med værdier i området 100 pF til 10 nF vil den høje grænsefrekvens være begrænset.

En DC strømkilde kan dannes med en transistor hvor en operationsforstærker holdes i balance ved at indstille spændingen over R_E som lig med indgangens U_I . Strømmen I_O er da givet af Ohms lov. Spændingen kan komme fra en mikroprocessors logiske udgang for en on/off styring af strømmen eller fra en D/A konverter for en programmerbar indstilling af strømkildens værdi.

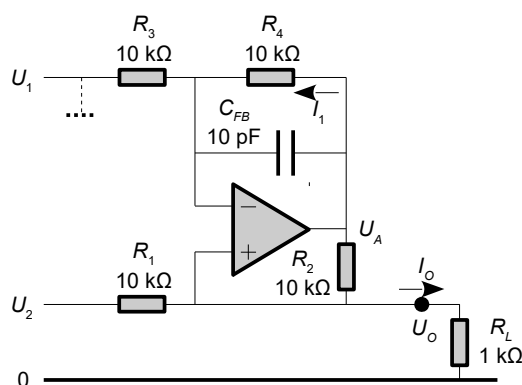


Figur 228 – En transistor kan styres af en operationsforstærker til en præcis DC strømkilde.

Kondensator C_1 er en afkobling af effektforsyningen. Kondensatoren C_2 modvirker selvsving og ved dobbelt effektforsyning (fx ± 15 V) vil dioden D_1 beskytte transistorens basis-emitter.

Man bør være konservativ med effektafsættelsen $P_C = U_{CE} I_C$ for at undgå opvarmning og her er der foreslået en tiendedel af det tilladte for transistoren. Spændingen over transistoren U_{CE} er sat lig med U_{CC} selv om den faktiske værdi er lavere. For BC547 er grænsen $P_{C MAX} = 0,3 \text{ W}$ ved 25°C og det giver $I_O = 2 \text{ mA}$ som maksimum. For BC139 er grænsen $P_{C MAX} = 1,25 \text{ W}$ ved 25°C som giver en maksimal udgangsstrøm på 10 mA og med køleplade stiger grænsen til 100 mA.

En Howland strømkilde kan drive strøm til (eller optage strøm fra) en belastning hvor den ene tilslutning er ført til nul. Der er mange referencer til kredsløbet på internettet og flere af dem er groft misforstået så jeg vil anbefale at du sætter sig grundigt ind i funktionen inden du benytter den. Modstandenes præcision bør være $\pm 0,1\%$ eller bedre og det kan opfyldes ved de små SIL-pakker med 4 ens værdier der leveres med $\pm 0,05\%$ indbyrdes tolerance.¹³⁰



Udgangsstrøm:

$$I_o = \frac{U_2 - U_1}{R_1} \quad \text{hvis} \quad \frac{R_1}{R_2} = \frac{R_3}{R_4}$$

Udgangsmodstand (parallel med R_L):

$$|R_o| \approx \frac{R_1}{4e_R} \quad \text{hvor} \quad \begin{aligned} e_R &= 0,01 \text{ for } \pm 1\% \\ e_R &= 0,001 \text{ for } \pm 0,1\% \end{aligned}$$

Udgangskapacitet (parallel med R_L):

$$C_o \approx \frac{1}{2\pi \text{GBP} R_2} \left(1 + \frac{R_4}{R_3} \right)^2$$

I praksis forventes: $C_o > 10 \text{ pF}$

Værdien stiger hvis C_{FB} benyttes.

Figur 229 – En strømgenerator der kan benyttes både ved AC og DC, men der stilles ret store krav til komponenternes præcision. Kondensator C_{FB} kan være nødvendig for stabilitet.

Funktionen ses ved at antage $U_1 = 0$ og $U_2 = 10 \text{ V}$ hvilket giver en udgangsstrøm på $I_o = 1 \text{ mA}$. Med belastningen $R_L = 1 \text{ k}\Omega$ bliver udgangsspændingen $U_o = R_L I_o = 1 \text{ V}$ og strømmen i R_1 bliver lig med $0,9 \text{ mA}$. Operationsforstærkeren indstiller sin udgang på $(1 + R_4/R_3)U_o = 2 \text{ V}$ så der løber en strøm i R_2 på $0,1 \text{ mA}$ der sammen med de $0,9 \text{ mA}$ fra R_1 giver $I_o = 1 \text{ mA}$.

Udgangsspændingen er typisk begrænset til $\pm 5 \text{ V}$ ved $\pm 15 \text{ V}$ effektforsyning hvis R_1 til R_4 er ens. Udgangsmodstanden R_o er funktion af modstandsvalget og deres tolerance. Den beregnede værdi for figurens værdier er $250 \text{ k}\Omega$ ved 1% tolerance, men det er ikke verificeret. Kondensatoren er nødvendig for stabilitet, den begrænser båndbredden og den vil givetvis kræve justering.¹³¹

Som det fremgår af analysen løber der en betydende strøm i indgangen så der skal være en lav udgangsmodstand fra det styrende kredsløb. Opstillingen benytter positiv tilbagekobling og kan derfor let gøres ustabil. Det anbefales at benytte et printkort med et fuldt dækkende stelplan og at afkoble hver gren af effektforsyningen med 100 nF keramiske kondensatorer og gerne suppleret af et par $100 \mu\text{F}$ aluminium elektrolytter.

¹³⁰Modstande i SIP finde hos www.bourns.com og www.vishay.com.

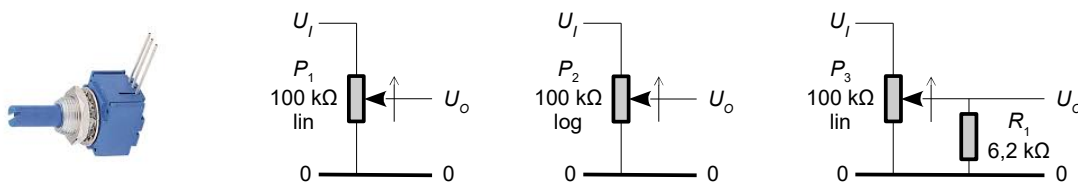
¹³¹Formlerne er fra <http://www.ti.com/lit/an/snoa474a/snoa474a.pdf>.

4.3 Korrektionskredsløb

Potentiometre er variable spændingsdelere der kan benyttes til at justere på forstærkningen af et signal eller indstillingen af et filterkredsløb. De fås både for manuel betjening ved drejning af en aksel eller for digital indstilling hvor MOSFET kontakter vælger delingsforholdet (*digipot*).

4.3.1 Styrkeregulering

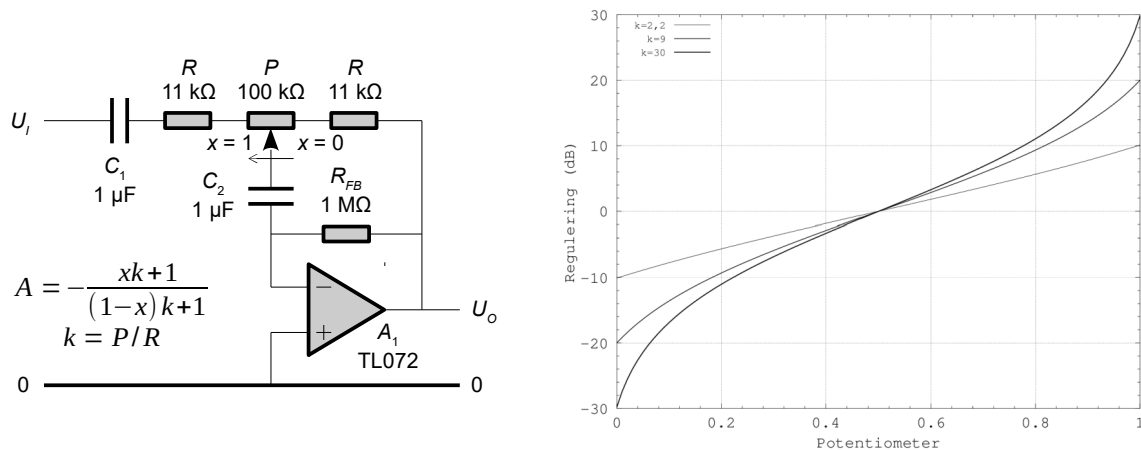
Et potentiometer kan benyttes som regulering af et signals amplitude fra 0 til 100 %. Ved et lineært potentiometer vil midterstillingen give 50 % af indgangssignalet. Illustrationen herunder viser en pil ved glideren for bevægelsesretningen som giver mere signal ved drejning mod uret.



Figur 230 – Styrkeregulering med potentiometer.

For styrkeregulering i et forstærkeranlæg er det lineære potentiometer ikke velegnet da der kun opnås en dæmpning på 6 dB i midterstillingen. Her foretrækkes logaritmiske potentiometre som vist med P_2 i illustrationen. Modstandsbanen er udført så signalet dæmpes cirka 30 dB i midten af reguleringsområdet. Det svarer til den subjektivt opfattede variation i amplituden i relation den den fysiske drejning på potentiometerets knap.

Det er muligt at opnå en pseudo-logaritmisk variation med et lineært potentiometer som P_3 viser. En modstand belaster potentiometeret så midterstillingen giver den ønskede dæmpning og med en modstand på 6,2 kΩ opnås 20 dB dæmpning i midterstillingen. Indgangsmodstanden aftager fra 100 kΩ ved 0 % signal til 5,8 kΩ ved 100 % signal så belastningen af kilden er ret hård.

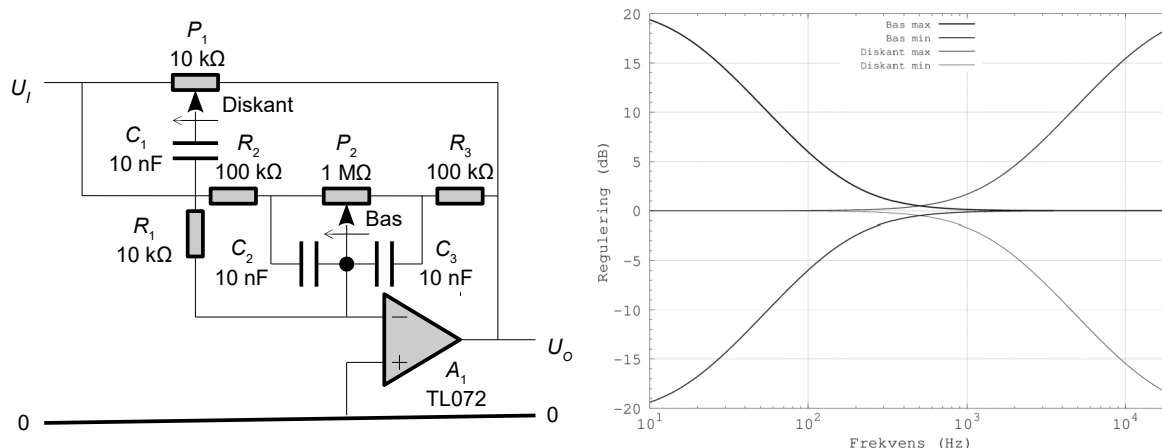


Figur 231 – En variabel forstærkning som funktion af et potentiometers drejning.

Potentiometeret kan også benyttes til at indstille forstærkningen med en operationsforstærker. Den inverterende forstærkers forstærkning er givet ved forholdet mellem to modstande og ved at bruge et potentiometer kan forstærkningen varieres kontinuerligt. Reguleringen er tilnærmet logaritmisk fra -20 dB til 20 dB hvilket giver en behagelig oplevelse af drejningen på potentiometret. C_1 og C_2 giver DC blokering for at undgå en knasende lyd når potentiometeret drejes og R_{FB} sørger for at operationsforstærkeren får sin biasstrøm.

4.3.2 Toneregulering

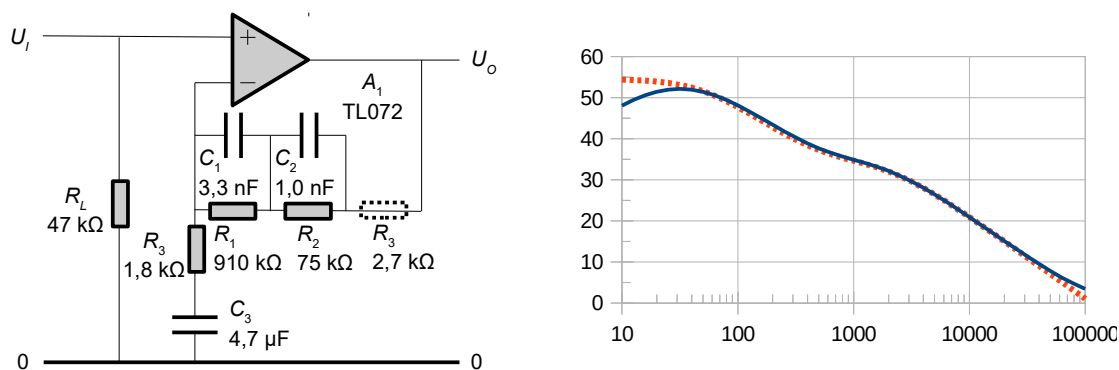
Bas- og diskantregulering opnås med en Baxandal tonekontrol der findes i mange udgaver. Ved høje frekvenser er P_2 kortsluttet af C_2 og C_3 så niveauet varieres ved at R_1 og C_1 lægges i parallel med R_2 for at øge niveauet eller R_3 for at sænke det. Ved lave frekvenser er P_1 ikke aktiv på grund af den høje impedans af C_1 så P_2 indstiller niveauet. Grænsefrekvensen for bassen er på 160 Hz som flyttes nedad ved at øge værdierne af $C_2 = C_3$ og $R_2 = R_3$. Grænsefrekvensen for diskanten indstilles ved R_1 og C_1 og til 1,6 kHz og flyttes opad ved at mindske C_1 .



Figur 232 – En klassisk toneregulering med separat potentiometer for bas og diskant.

4.3.3 RIAA forstærker

Afspilning af vinyl-plader med 33 eller 45 rpm benytter en elektro-dynamisk pick-up hvis nål følger et spiralformet spor i overfladen af pladen hvor lyden er nedfældet som små vibrationer omkring spiralen. For at få mest mulig musik ind på pladen dæmpes bassen mens diskanten er fremhævet for at øge afstanden fra pladens overfladeujævnhed. Ved afspilning skal processen inverteres og korrektionen kaldes RIAA (*Recording Industries Association of America*). Standarden fra omkring 1950 kan beskrives ved de tre tidskonstanter 75 μ s, 318 μ s og 3180 μ s.¹³²



Figur 233 – En RIAA forstærker med 35 dB forstærkning ved 1 kHz.

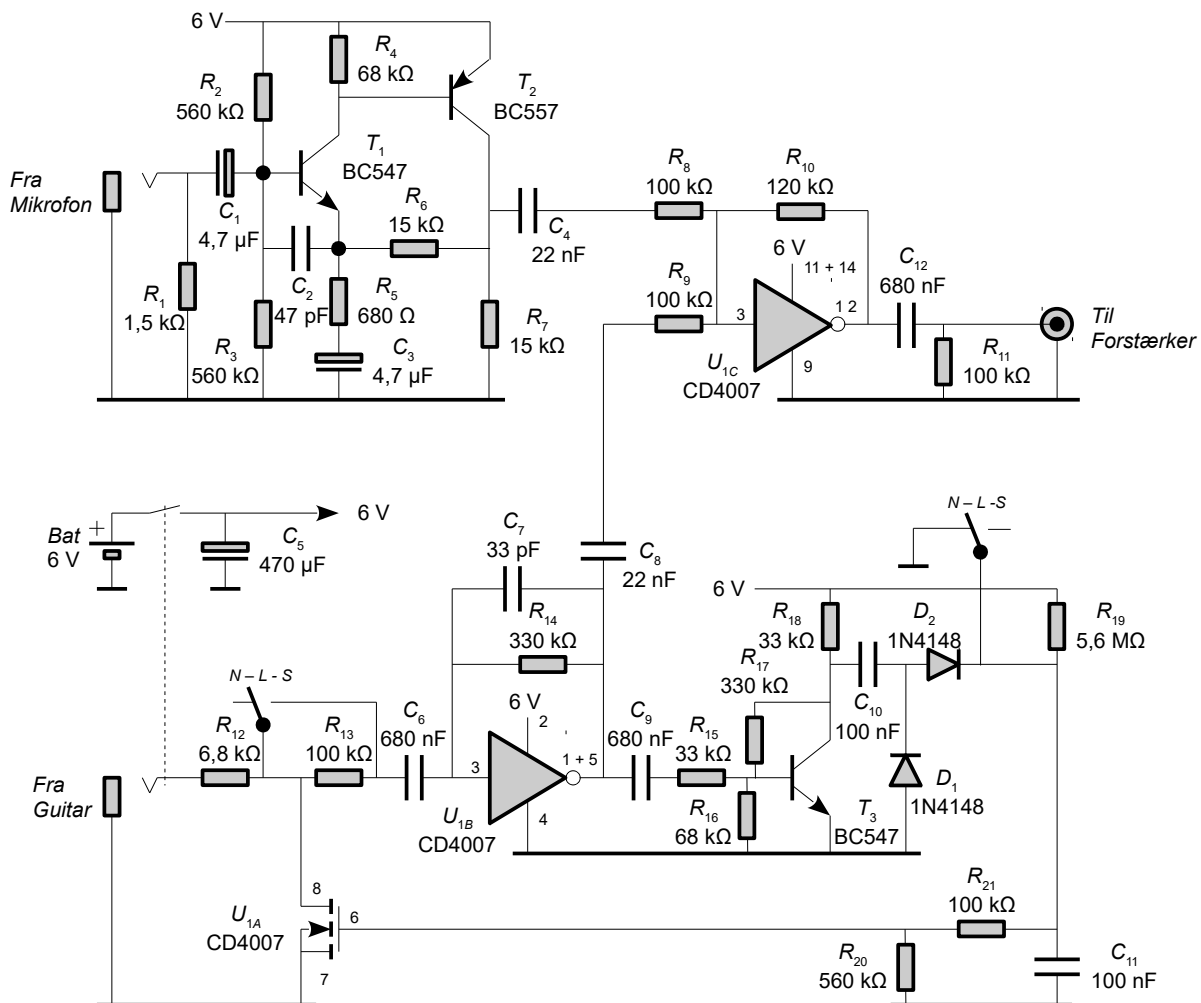
Specifikationen kan oversættes til en pol ved 50 Hz, et nulpunkt ved 500 Hz og endnu en pol ved 2120 Hz. R_L belaster pick-uppen så den arbejder optimalt. Opbygningen med en ikke-inverterende forstærker indfører endnu et nulpunkt. R_3 kan forbedre OpAmp stabiliteten og med 2,7 k Ω sænkes nulpunktet til 50 kHz. C_3 mindsker DC offset og giver en grænse ved 20 Hz (-3 dB). Begge dele diskuteres ihærdigt på nettet, men de er ikke del af standarden.¹³³

¹³²For en fin introduktion til emnet anbefales https://en.wikipedia.org/wiki/RIAA_equalization.

¹³³Dimensioneringen gennemgås detaljeret i http://andyc.diy-audio-engineering.org/phono_preamp.xhtml.

4.3.4 Guitarkompressor

En dynamisk mikrofon og en elektrisk guitar skulle mikses sammen og føres til en guitarforstærker for jazz-sammenspil med en sangerinde. Mikrofonen forstærkes 20 gange i T_1 og T_2 for at komme på niveau med guitaren. Guitaren forstærkes af U_{1B} der er en CMOS inverter og føres herfra til U_{1C} der samler mikrofonen og guitarens signaler og fører kombinationen til forstærkeren. Kredsen U_1 er en CD4007 der indeholder tre sæt P og N kanal MOSFET der enten kan bruges som to individuelle transistorer eller for analog forstærkning og begge muligheder udnyttes her.



Figur 234 – En dynamisk mikrofon mikses med en guitar hvor der er mulighed for at begrænse guitarens amplitude for en mere homogen lyd.

Guitarens signal kan påvirkes med en tre-stillings dobbelt omskifter af typen ON-OFF-ON. Den ene sektion indstiller forstærkningen ved modstandene R_{12} , R_{13} og R_{14} og den anden sektion aktiverer en N-kanal MOSFET som her er en spændingsstyret modstand der kan dæmpe guitaren. Signalet fra U_{1B} forstærkes i T_3 og ensrettes i D_1 og D_2 til en spænding der følger signalets amplitude. Modstandene R_{19} og R_{20} holder transistoren U_{1A} nær ved det punkt hvor den begynder at virke og ved stigende amplitude mindskes dens modstandsværdi. Resultatet er at guitarens signal dæmpes når niveauet kommer over en grænse der indstilles med guitarens styrkekontrol.

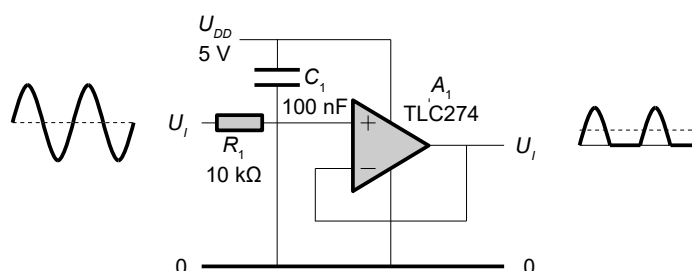
Ved stilling *N* forstærkes guitaren uden påvirkning af amplituden (*neutral*). Ved stilling *L* begrænses niveauet til en maksimal værdi (*limiter*). Ved stilling *S* øges forstærkningen så begrænseren holder amplituden mere konstant og guitaren får en syngende karakter (*sustain*).

4.4 Ikke-lineære kredsløb

Dioden kan lede strømmen i den ene retning, men ikke i den anden retning og udgangen for en operationsforstærker kan ikke føres udenfor effektforsyningen. Det kan udnyttes til ikke-lineære kredsløb.

4.4.1 Enkeltensretning

En del operationsforstærkere tillader at indgangssignalet går under den negative effektforsyning og at udgangen kan nå næsten helt ned til samme grænse. Det kan udnyttes til en enkelt form for enkeltensretning om vist her. Ved positive indgangssignal vil udgangen følge indgangssignalet og for negativt indgangssignal vil udgangen blot ligge stille nær ved 0 V.

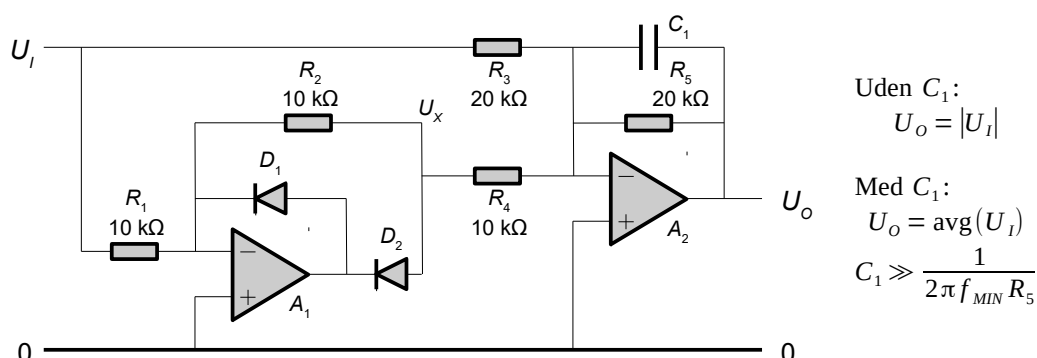


Figur 235 – Enkeltensretning af et analogt signal for brug ved signalbehandling.

De integrerede kredse er beskyttet imod elektrostatisk udladning gennem beskyttelsesdioder der holder indgangene indenfor effektforsyningens grænser. Det betyder at indgangen ikke må komme under $-0,3$ V som det fremgår af grænsedata for TLC274 (*Absolute Maximum Ratings*) og hvis det ikke overholdes skal strømmen begrænses til 5 mA. Det er grunden til modstanden R_1 der vil holde strømmen indenfor det lovlige ned til -5 V på indgangen.

4.4.2 Dobbeltensretning

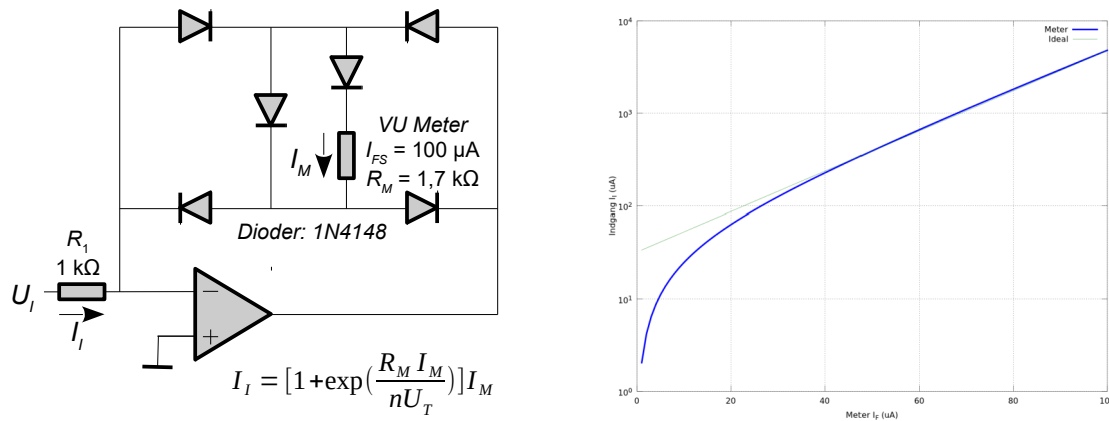
Spændingsfaldet over en diode er generende ved brug i analog signalbehandling og kan undgås ved at benytte en operationsforstærker for at danne en såkaldt "ideel diode". Det vises herunder hvor A_1 modvirker spændingsfaldet over dioden D_2 så U_x går negativ ved positivt signal og ligger på 0 V ved negativt fortegn. Forstærker A_2 samler U_i og U_x til absolutværdien af signalet så der opnås en dobbeltensretning. En kondensator i parallel med R_5 lavpasfilteres udgangssignalet. Operationsforstærkeren sætter en øvre grænse til et sted i kilohertz-området.



Figur 236 – 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.

4.4.3 Pseudo-logaritmisk måleinstrument

Indenfor lydverdenen er VU-metre populære, men deres dynamikområde er kun 20 dB. Den viste opstillingen kan derimod vise 50 dB på et instrument. En helbølgeensretter er bygget over en inverterende forstærker og fire dioder.^{134,135}



Figur 237 – Strømmen i instrumentet varierer omtrent logaritmisk med indgangsspændingen.

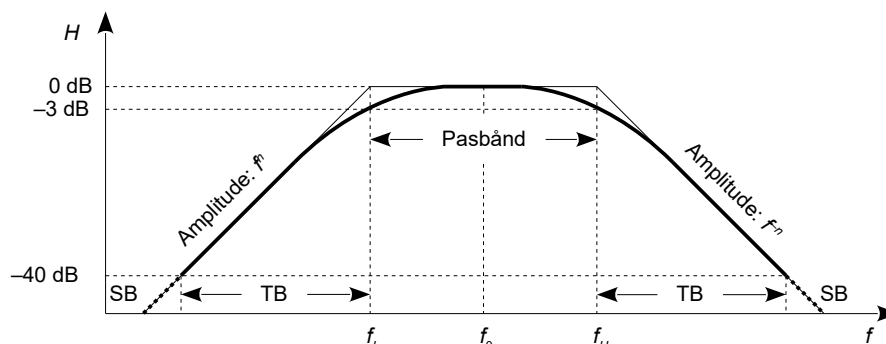
Kredsløbet er designet til at bruge 1N4148 ($n = 1,7$), men det vil fungere med de fleste dioder. Ved fuld udstyring på $100 \mu\text{A}$ ligger der 170 mV over R_M og strømmen i indgangen når 5 mA ved det kraftigste indgangssignal. Temperaturkoefficienten af drejespoleinstrumentets kobbertråd er på 0,33 %/K og kompenserer for temperaturkoefficienten af spændingen fra de to dioder i det logaritmerende kredsløb.

¹³⁴Det er et drejespoleinstrument med indbygget ensretter så instrumentet slår ud alt efter hvor kraftigt signalniveauet er. Området er fra +3 dB og ned til -20 dB og i den professionelle verden er der krav til instrumentets dæmpning (reaktionstid).

¹³⁵Se *Circuit Idea* i "Electronics World", July/August 1996, side 615.

4.5 Filtre

Filtre benyttes almindeligvis for at opnå en variation i amplituden som funktion af frekvensen med et eksempel illustreret herunder for et båndpasfilter. Der er dog også mulighed for at studere andre parametre som fasens variation med frekvensen eller stepresponsens udvikling over tid.



Figur 238 – Amplituden af et båndpasfilter vist i dobbelt-logaritmisk afbildning med et Bode plot (tynd streg) og den aktuelle amplitude (tyk streg). Grænsefrekvensen ved -3 dB er normalt markør for hvortil pasbåndet går. TB står for transmissionsbåndet hvor dæmpningen gradvist går imod asymptotens værdi. SB står for stopbåndet der arbitrært er vist ved -40 dB svarende til et niveau på højst 1 % af det oprindelige signal.

Det illustrerede filter tillader et signal at passere hvis frekvensen ligger i pasbåndet fra den lave grænsefrekvens f_L til den høje grænsefrekvens f_H hvor signalet er 3 dB under det maksimale niveau der i visse situationer kaldes centerfrekvensen f_0 . Under den lave grænsefrekvens og over den høje er signalet i transmissionsbåndet fra pas- til stopbånd. Niveaulet ved overgang til filtrets stopbånd har ikke en fast definition, men afhænger af formålet med filtret.

Filtre af typen lavpas, båndpas og højpas beskrives ved de grænsefrekvenser hvor stejlheden af amplitudens variation med frekvensen skifter. Det er almindeligt at skitsere et filters respons ved det såkaldte Bode plot hvor der tegnes rette linjer med en hældning på ± 6 dB/oktav eller vandret for filtre af første orden, ± 12 dB/oktav for filtre af anden orden og så videre. Det er vist ovenfor med tynd streg. Når denne asymptotiske karakteristik er tegnet kan man konstruere responsen med fri hånd og der rammes ofte ret præcist i forhold til den faktiske respons.

Det selektive filter afviger fra det beskrevne ved at kondensator og spole kan gå i resonans med hinanden og svinge på et signals frekvenskomponent. Amplitudens respons afviger fra et Bode plot og det beskrives almindeligvis ved den såkaldte Q-faktor der kort sagt angiver afvigelsen fra Bode plottet. Hvis et filter eksempelvis har en Q-værdi på 5 så vil amplituden løfte sig 5 gange (14 dB) over Bode plottets niveau og hvis Q-værdien er 0,5 vil amplituden være på 0,5 gange (-6 dB).

Filtre opdeles efter overføringsfunktionens orden der også beskrives ved antallet af poler. Det simplest mulige filter er af første orden (én pol) hvor mulighederne er højpas eller lavpas. Ved filtre af højere orden er ordenen givet ved antallet af reaktive komponenter (spoler og kondensatorer) og viser sig ved eksponenten til s i nævnerens polynomium. Overføringsfunktionen for filtre af anden orden har to poler som enten er reelle eller komplekst konjugerede. I første tilfælde er det muligt at splitte overføringsfunktionen op i to første-ordens blokke med hver sin individuelle grænsefrekvens og i andet tilfælde kan filtret beskrives ved en grænsefrekvens og den tilhørende godhed.

Filtre af højere orden kan altid opdeles i blokke af anden orden og for filtre af ulige orden tilmed en enkelt blok af første orden. Det er følgelig kun nødvendigt at behandle designet af første-ordens og anden-ordens filtre.

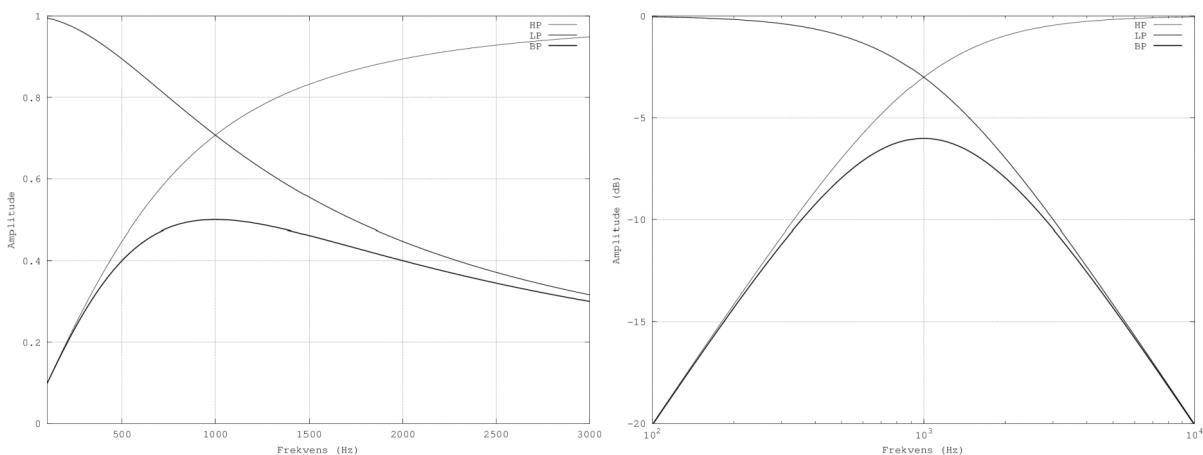
4.5.1 Filtre af første orden

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. Det kaldes også for halv-effekt grænsen. Stejlheden er moderat og svarer til en radios tonekontrol med bas- eller diskant drejet helt ned så der slipper en del signal igennem udenfor pasbåndet. Amplituden af et signal i en dekades afstand fra grænsefrekvensen er kun reduceret med 20 dB.¹³⁶

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	$H_{LP} = \frac{\omega_H}{\omega_H + s}$	$H_{LP} \underset{ s < \omega_H}{\approx} 1$	$H_{LP} \underset{ s > \omega_H}{\approx} \frac{\omega_H}{s}$

For højpas er amplituden proportional med frekvensen (6 dB/oktav) for frekvenser mindst en oktav under polen. For lavpas er amplituden omvendt proportional med frekvensen (-6 dB/oktav) fra en oktav over polen og op. Det forklarer hvorfor forløbet kan tilnærmes ved et Bode plot hvor filtret beskrives ved rette linjer med ± 6 dB/oktav eller en vandret linje.

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 tælleren 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.



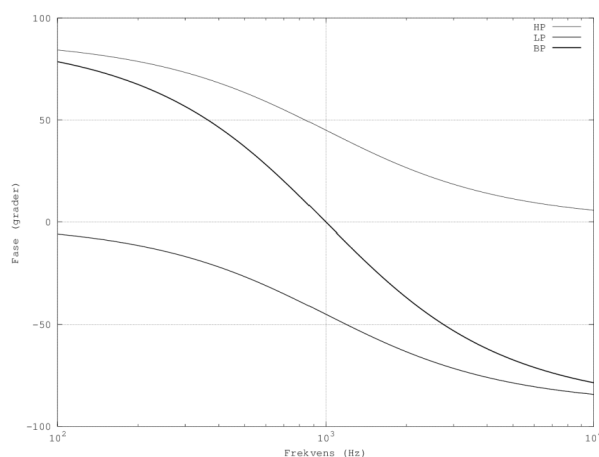
Figur 239 – Amplituden plottet for første-ordens filtre af typen højpas (HP) og lavpas (LP) samt for båndpas dannet af de to filtre i kæde (BP). Til venstre vises et plot med lineære akser og til højre et plot med logaritmiske akser. Grænsefrekvensen er begge 1 kHz hvor HP og LP er 3 dB nede.

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 leddene ω_L (eller ω_H) og det imaginære led $s = j\omega$, og amplituden af nævneren er givet ved RMS addition.

$$\begin{aligned}
 \text{Højpas: } A_{HP} &= \left| \frac{s}{\omega_L + s} \right| = \frac{\omega}{\sqrt{\omega_L^2 + \omega^2}} = \frac{1}{\sqrt{1 + (f/f_L)^2}} \\
 \text{Lavpas: } A_{LP} &= \left| \frac{\omega_H}{\omega_H + s} \right| = \frac{\omega_H}{\sqrt{\omega_H^2 + \omega^2}} = \frac{1}{\sqrt{1 + (f/f_H)^2}}
 \end{aligned}
 \quad \text{hvor} \quad
 \begin{aligned}
 \omega &= 2\pi f \\
 \omega_L &= 2\pi f_L \\
 \omega_H &= 2\pi f_H
 \end{aligned}$$

¹³⁶ Subjektivt svarer 10 dB reduktion til en halvering af lydniveauet, så ved -20 dB er det subjektive lydniveau kun reduceret til en fjerdedel. Filtret er i mange henseender for svagt til at kunne dæmpe et støjsignal tilstrækkeligt.

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 fassen forløbe blødt fra den ene asymptote til den anden som det vises i illustrationen herunder for højpas, lavpas og båndpas.



$$H_{HP} = \frac{s}{\omega_H + s}$$

$$H_{LP} = \frac{\omega_L}{\omega_L + s}$$

$$H_{BP} = \frac{s}{\omega_H + s} \frac{\omega_L}{\omega_L + s}$$

$$\theta = \arctan\left(\frac{\text{Im}(H)}{\text{Re}(H)}\right)$$

$$\theta_{HP} = \frac{\pi}{2} - \arctan\left(\frac{\omega}{\omega_H}\right)$$

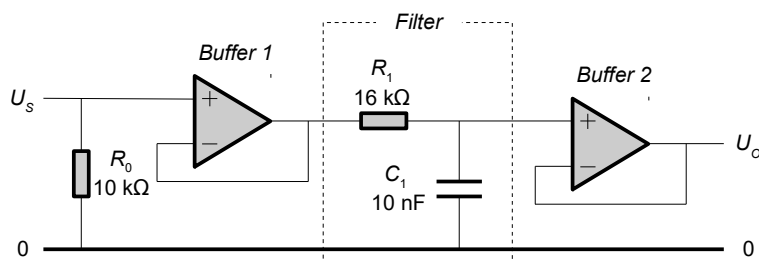
$$\theta_{LP} = -\arctan\left(\frac{\omega}{\omega_L}\right)$$

$$\theta_{BP} = \theta_{HP} + \theta_{LP}$$

Figur 240 – 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 fassen er drejet -45° .

4.5.2 Ikke-inverterende lavpasfilter

Et første-ordens lavpasfilter vises herunder med selve filtret i den stiplede boks. Der er indskudt en buffer både før og efter filtret for at gøre funktionen uafhængig af det eksterne kredsløb.



$$H = \frac{\omega_1}{\omega_1 + s}$$

$$\omega_1 = 2\pi f_1$$

$$f_1 = \frac{1}{2\pi R_1 C_1}$$

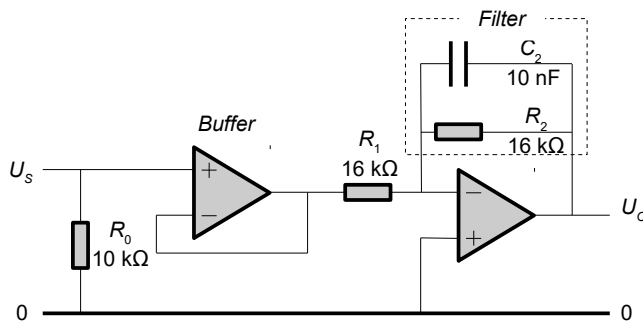
$$GBP > 100 f_1$$

Figur 241 – 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.

Grænsefrekvensen er givet af RC produktet så den lave udgangsmodstand fra Buffer 1 er garanteret for at modstanden R_1 ikke utilsigtet forøges af impedansen fra kredsløbet til venstre for U_s . Buffer 2 sikrer tilsvarende at udgangen fra filtret ikke belastes af det efterfølgende kredsløb.

4.5.3 Inverterende lavpasfilter med forstærkning

En populær realisation af første-ordens filtret vises herunder hvor en inverterende forstærker har båndbredden begrænset af en kondensator i parallel i tilbagelkoblingen. 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 . 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.



$$H = -\frac{R_2}{R_1} \frac{\omega_2}{\omega_2 + s}$$

$$\omega_2 = 2\pi f_2$$

$$f_2 = \frac{1}{2\pi R_2 C_2}$$

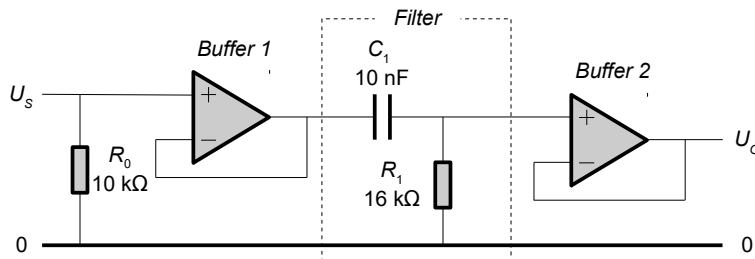
$$GBP > 100 \cdot \left(1 + \frac{R_2}{R_1}\right) f_1$$

Figur 242 – 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.

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 det vil almindeligvis forbedre forstærkerens stabilitet overfor utilsigtet kapacitet fra printkortets stelplan.

4.5.4 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 og 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.



$$H = \frac{s}{\omega_1 + s}$$

$$\omega_1 = 2\pi f_1$$

$$f_1 = \frac{1}{2\pi R_1 C_1}$$

$$GBP > 100 f_1$$

Figur 243 – 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.

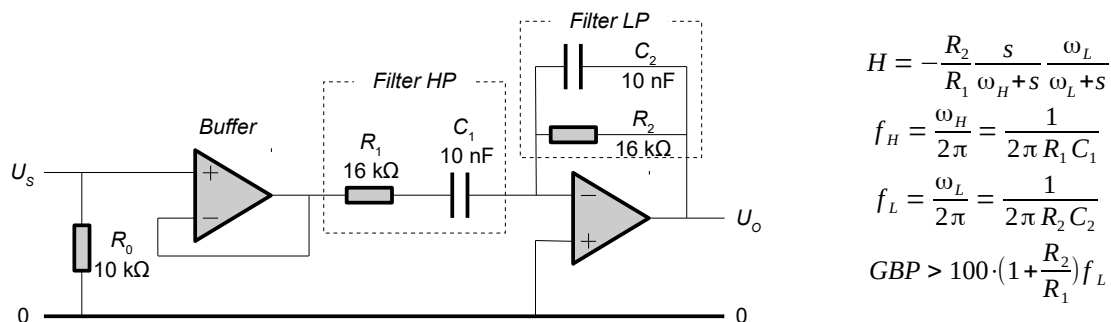
Med en grænsefrekvens på 10 Hz vil C_1 blive på 1 μF med den viste modstandsværdi. Til ukritiske formål vil en keramisk kondensator med X7R dielektrikum være mest attraktiv ud fra pladskrav og økonomi mens mere krævende formål vil have fordel af en plastfilm type. De er fri for mikrofoni og kan leveres uden dielektrisk absorption der kan være et problem ved krævende lydformål.^{137,138}

¹³⁷Kondensatorens dielektrikum er piezoelektrisk og genererer en spænding når den udsættes for chok og rystelser.

¹³⁸Dielektrisk 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.

4.5.5 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 244 – Et båndpasfilter med en inverterende forstærker. Bufferen kan være påkrævet for at opnå den ønskede forstærkning. Grænsefrekvensen er 1 kHz med de viste værdier.

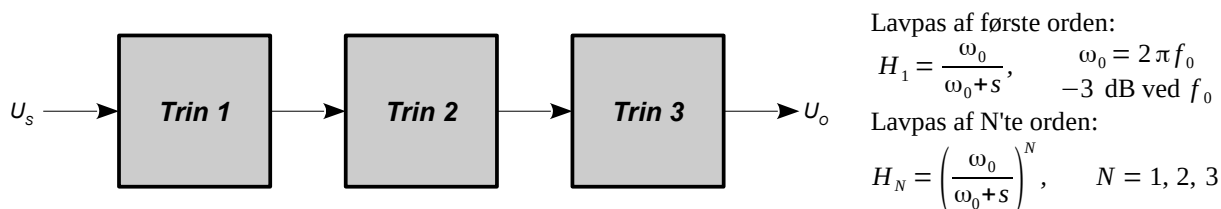
Forstærkningen i pasbåndet indstilles gennem modstandene R_1 og R_2 som dog ikke kan vælges helt frit idet begge modstande også varetager filtrering. For at undgå gensidig påvirkning mellem filtrene skal lavpasfiltrets grænsefrekvens skal ligge langt over højpasfiltrets ($f_H \gg f_L$).

Båndpasfilter	Lavfrekvens	Mellemfrekvens	Højfrekvens
$H = \frac{s}{\omega_L + s} \frac{\omega_H}{\omega_H + s}$	$H \approx \frac{s}{\omega_L}$	$H \approx 1$	$H \approx \frac{\omega_H}{s}$
$\omega_L \ll \omega_H$	+6 dB/oktav +90°	Pasbånd (0 dB) 0°	-6 dB/oktav -90°

Der er intet galt ved at vælge ens grænsefrekvens for filtrene ($f_H = f_L$) der blot giver en forstærkning i pasbåndet på -6 dB. Med højpasfiltrets grænsefrekvens højere end lavpasfiltret ($f_L > f_H$) bliver resultatet stadig et båndpasfilter, men dæmpningen i pasbåndet bliver endnu mere betydende.

4.5.6 Kaskade af filtre

Megen elektronik inkluderer mere end et enkelt lavpas- eller højpasfilter. Det er fx tilfældet når et signal skal igennem en kæde af flere operationsforstærkere. Hvis der eksempelvis er tre ens lavpasfiltre med $f_0 = 50$ kHz i kaskade vil de tilsammen give -9 dB ved grænsefrekvensen f_0 så niveauet er -3 dB allerede ved 25 kHz hvor hvert af filtrene blot dæmper 1 dB.



Figur 245 – Et større system kan indeholde mange filtre i kaskade så den resulterende respons kan blive degraderet af den successive begrænsning af frekvensområdet.

Der gælder tilsvarende for en kaskadekobling af forstærkere med DC blokering som beskrives ved et højtpasfilter. Med $f_0 = 15$ Hz for hvert af filtrene vil tre filtre i kaskade give -3 dB ved 30 Hz.

4.5.7 Filtre af anden orden

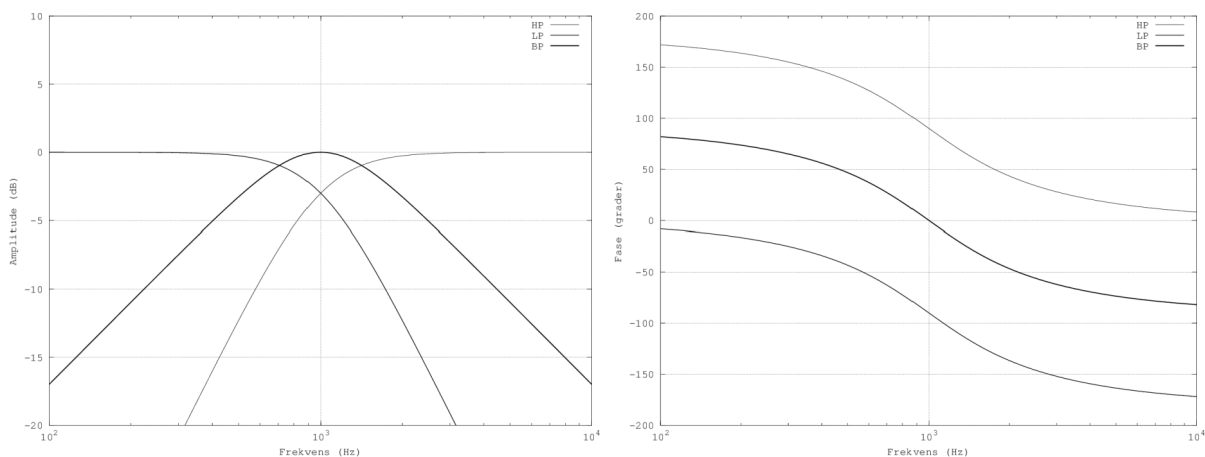
Et filter af anden orden har to poler og de er enten reelle eller komplekst konjugerede. I det første tilfælde kan filtret opfattes som bestående af to første-ordens filtre i kaskade som beskrevet på de foregående sider. I det andet tilfælde skabes et filter med mulighed for at påvirke responsen i omegnen 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.

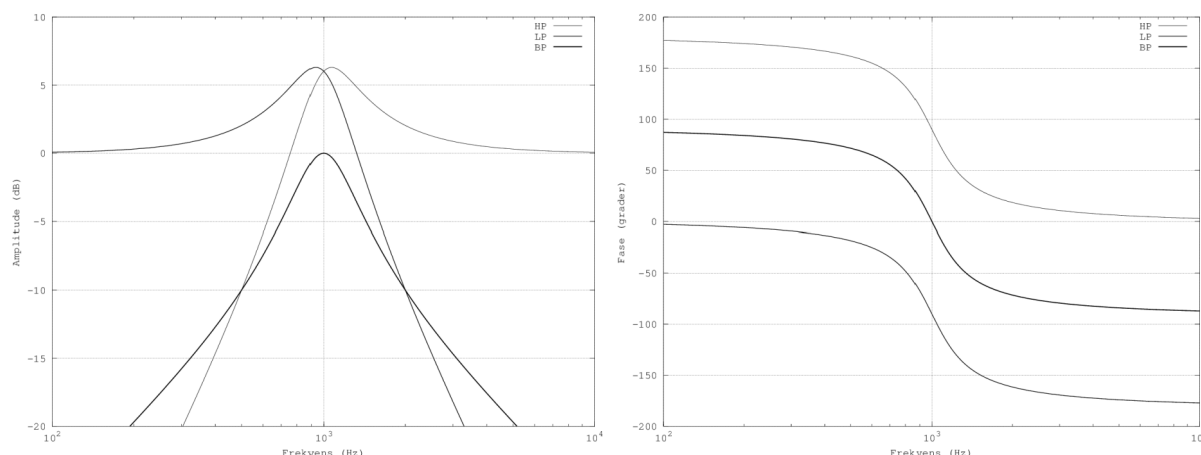
Den karakteristiske frekvens f_0 angiver filtrets asymptotiske grænsefrekvens som er den frekvens hvor asymptoterne for s gående mod nul og for s gående mod uendelig krydser hinanden. Det er den udæmpede resonansfrekvens f_0 . Frekvensen som filtret svinger på ved en påvirkning kaldes for den dæmpede 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 dæmpningen d eller ζ eller tilsvarende for filtrets godhed Q . Herunder vises responsen for $d = Q = 1/\sqrt{2} = 0,71$ der også betegnet som et maksimalt fladt filter (maximum flat filter) eller et Butterworth filter hvor niveauet er -3 dB ved grænsefrekvensen.



Figur 246 – 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.

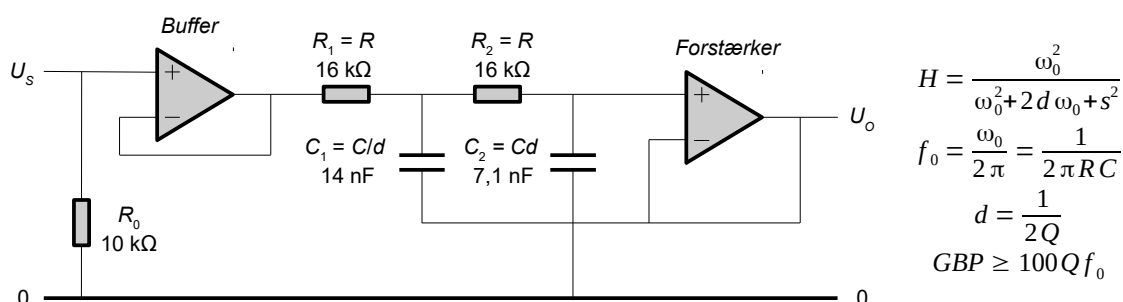
Responsen vises herunder for $d = 0,25$ ($Q = 2$). Niveauet er 6 dB ved $f_0 = 1$ kHz svarende til en forstærkning af signalet på to gange. Toppunktet ligger ved $0,97 f_0$ for lavpasfiltret og ved $1,03 f_0$ for højpasfiltret, og ved stigende godhed (aftagende dæmpning) rykker de to frekvenser hen imod f_0 . Fasen forløber stejletere i overgangsområdet end ved det første eksempel, så filtret holder fasen nærmere den asymptotiske værdi ved aftagende dæmpningen (stigende godhed).



Figur 247 – 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).

4.5.8 Ikke-inverterende lavpasfilter

Lavpasfiltret herunder er især kendetegnet 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 ofte mødt design benytter en dæmpning på 0,71 for at opnå den såkaldte Butterworth karakteristisk, der er vist herunder, og hvor amplituden er jævn med -3 dB ved grænsefrekvensen.



Figur 248 – 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 primært forholdet mellem kondensatorerne der indstiller godheden. Modstandene danner forholdet mellem den geometriske og aritmetiske middelværdi som normalt er tæt på én. Det kan 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 %.

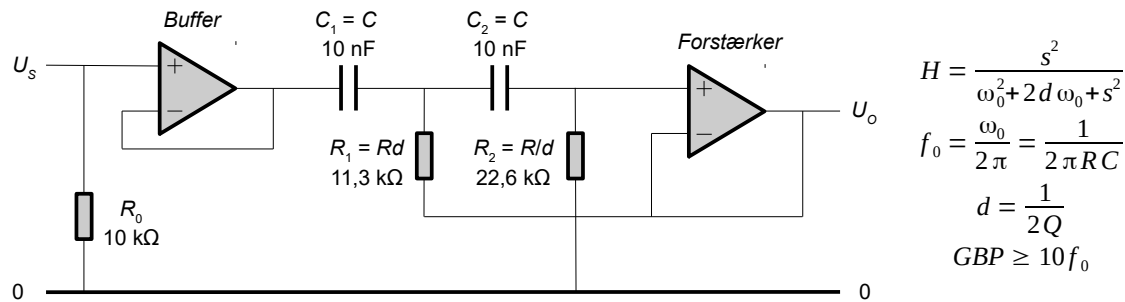
$$f_0 = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}} \quad \text{og} \quad Q = \frac{\sqrt{R_1 R_2}}{R_1 + R_2} \sqrt{\frac{C_1}{C_2}}$$

Den viste buffer i indgangen har samme funktion som ved de tidligere viste filtre nemlig 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.

4.5.9 Ikke-inverterende højpasfilter

Filtret kan ændres til højpas ved at ombytte modstande og kondensatorer. Der gælder samme krav til komponentværdierne som ved lavpasfiltret.

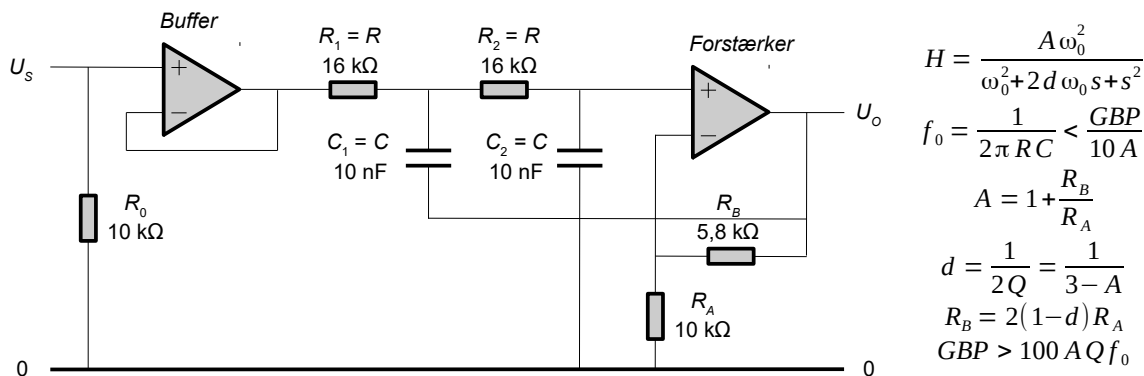
Et højpasfilter lader et signal passere når frekvensen er over grænsefrekvensen, lavere frekvenser dæmpes progressivt med faldende frekvens og DC blokeres. Operationsforstærkeren vil ved høje frekvenser virke som et første-ordens lavpasfilter på grund af den endelige båndbredde udtrykt ved operationsforstærkerens *GBP* specifikation. Det betyder i praksis at overføringsfunktionen for et højpasfilter bliver et asymmetrisk båndpasfilter.



Figur 249 – 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.

4.5.10 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.

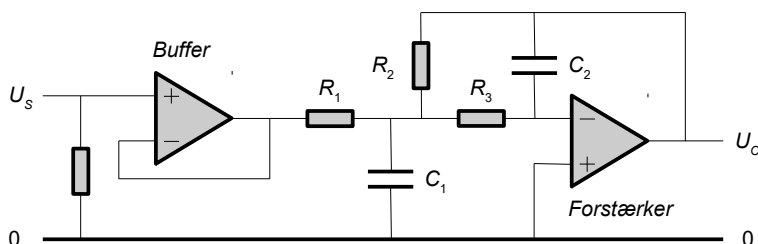


Figur 250 – Et anden-orden lavpasfilter med en buffer i indgangen og en variabel forstærkning A for tilpasning af dæmpningen. Grænsefrekvensen er 1 kHz og godheden er 0,71 med de viste værdier.

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 forsvares at kalde det for et anden-ordens filter.

4.5.11 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.



$$H = -\frac{R_2}{R_1} \frac{\omega_0^2}{s^2 + 2d\omega_0 s + \omega_0^2}$$

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{R_2 R_3 C_1 C_2}}$$

$$Q = (R_1 \parallel R_2 \parallel R_3) \sqrt{\frac{C_1}{R_2 R_3 C_2}}$$

$$GBP > 100 f_0 \sqrt{\frac{R_2 C_1}{R_3 C_2}}$$

For $R = R_1 = R_2 = R_3$:

$$f_0 = \frac{1}{2\pi R \sqrt{C_1 C_2}}$$

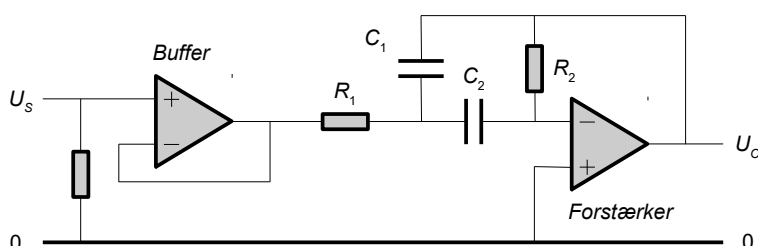
$$Q = \frac{2}{3} \sqrt{\frac{C_1}{C_2}}$$

Figur 251 – 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.

4.5.12 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 .



$$H = -A_0 \frac{2d\omega_0 s}{\omega_0^2 + 2d\omega_0 s + s^2}$$

$$A_0 = Q^2 \left(1 + \frac{C_2}{C_1}\right)$$

$$f_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}}$$

$$Q = \sqrt{\frac{R_2}{R_1} \frac{\sqrt{C_1 C_2}}{C_1 + C_2}}$$

$$GBP \geq 20 Q^2 f_0$$

For: $C = C_1 = C_2$:

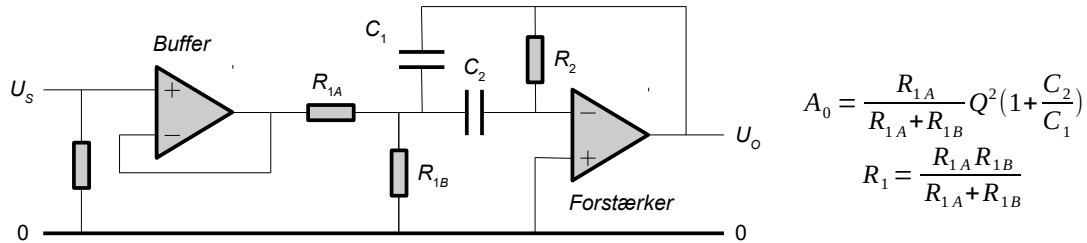
$$Q \approx \frac{1}{2} \sqrt{\frac{R_2}{R_1}}$$

$$f_0 = \frac{1}{4\pi Q R_1 C}$$

Figur 252 – 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 designs. 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 k Ω til 1 M Ω . 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.



Figur 253 – 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.

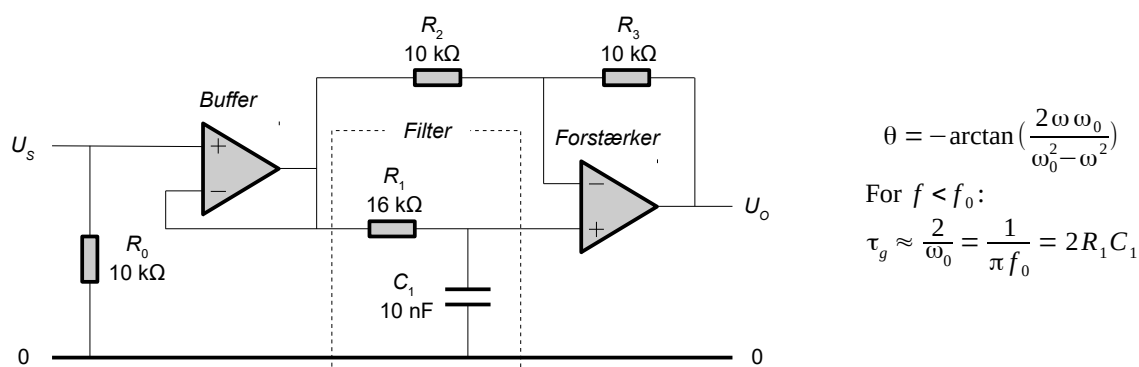
Den store værdi af forstærkningen ved resonans kan mindskes ved at dæmpe indgangssignalet. Det foretages helt enkelt ved en opdeling af R_1 i to modstande hvor parallelværdien benyttes som modstanden R_1 og forstærkningen reduceres ved spændingsdelingen mellem dem. Det kan indses ved at danne Thevenin-ækvivalentet for indgangssignalet U_s og de to modstande R_{1A} og R_{1B} . 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.

4.5.13 Fasedrejningsfilter

Et lidt specielt filter lader alle frekvenser passere uden dæmpning og påvirker kun fasen. Det kaldes på engelsk for et *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 frekvenser.¹³⁹

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.



Figur 254 – Et fasedrejningsfilter påvirker ikke amplituden. Fasen drejes 0° ved lave frekvenser, -90° ved grænsefrekvensen og -180° mod høje frekvenser. Grænsefrekvensen er 1 kHz med de viste komponentværdier hvor $\theta = -90^\circ$ og modstandene R_2 og R_3 skal være ens.

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 $f_0 = 1$ kHz er den på $\tau_g = 0,3$ ms), men det er alligevel nok til at finde praktisk anvendelse. Fasedrejningsfiltret ses anvendt i telefonsystemer til at korrigere for lydets 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 forpladen.¹⁴⁰

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.

¹³⁹Det er med vilje at en inverterende forstærker ikke betegnes med $+180^\circ$, 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.

¹⁴⁰Der kan foretages en "elektrisk flytning" af en højttalerenhed tilbage fra forpladen, så signalet bringes i fase med en anden højttalerenhed. Hvis korrektionen baseres på $f_0 = 10$ kHz er $\tau_g = 30$ μ s, og med lydets hastighed på $c = 340$ m/s vil det svare til en justering på $c\tau_g \approx 10$ mm. Hvis der er behov for en større værdi af tidsforsinkelsen kan flere fasedrejningsled seriekobles.

$$|H| = \left| \frac{\omega_0 - j\omega}{\omega_0 + j\omega} \right| = \frac{\sqrt{\omega_0^2 + \omega^2}}{\sqrt{\omega_0^2 + \omega^2}} = 1$$

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 tælleren igen går imod nul.

$$\theta = \arctan\left(\frac{\operatorname{Im}(H)}{\operatorname{Re}(H)}\right) = -\arctan\left(\frac{2\omega\omega_0}{\omega_0^2 - \omega^2}\right) = -\arctan\left(\frac{2x}{1-x^2}\right) \quad \text{hvor} \quad x = \frac{\omega}{\omega_0}$$

Taylor-serien for arctan vises herunder.

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots \Rightarrow \arctan(x) \approx x \quad \text{for} \quad |x| < 1$$

Brugt på udtrykket for fassen betyder det at ω^2 i nævneren kan ignoreres fordi frekvensen antages lavere end grænsefrekvensen og fassen 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° .

$$\theta = -\arctan\left(2\frac{\omega}{\omega_0}\right) \approx -2\frac{\omega}{\omega_0} \Rightarrow \theta \approx -2\frac{2\pi f}{\omega_0} = -2\pi(2RC)f = -2\pi\tau_g f \Rightarrow \tau_g = 2RC$$

I den teoretiske litteratur er tidsforsinkelsen givet af den afledte af fassen med hensyn til frekvensen og resultatet kaldes for gruppetidsforsinkelsen (*group delay*). Begrænses analysen til området under grænsefrekvensen hvor simplificationen er gyldig findes helt det samme udtryk som før.

$$\tau_g \stackrel{\text{def}}{=} -\frac{d\theta}{d\omega} \Rightarrow \tau_g \approx -\frac{1}{\omega_0} \frac{d\{-\arctan(2x)\}}{dx} = \frac{2}{\omega_0} = 2RC$$

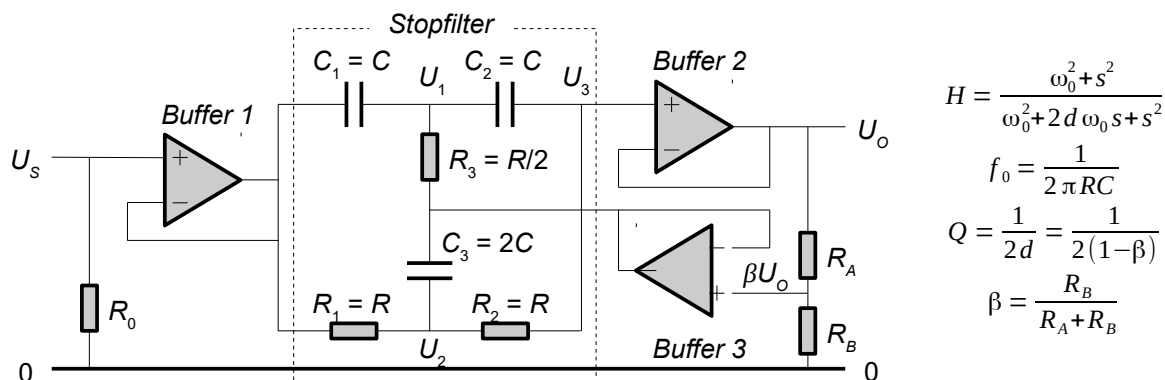
Med en sinussvingning på indgangen af filtret bliver resultatet en tidsforsinkelse τ_g der normalt kaldes for gruppeløbetiden (*group delay*).

$$u(t) = U_m \sin(2\pi t + \theta) \Rightarrow u(t) = U_m \sin(2\pi f(t - \tau_g))$$

$$\theta = -2\pi\tau_g f$$

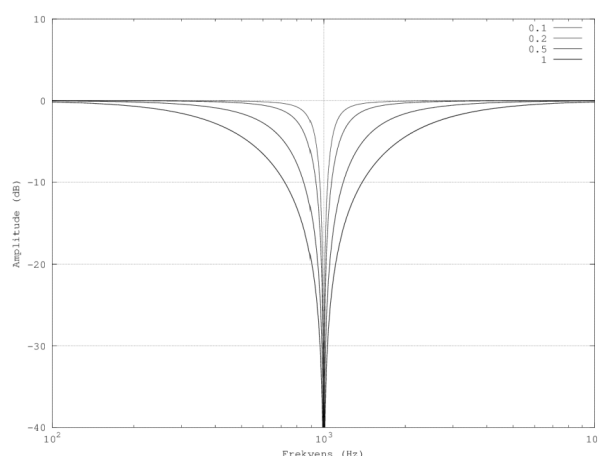
4.5.14 Stopfilter

Det er muligt at undertrykke et signal ved en given frekvens hvilket kan være attraktivt hvis der er en genevirkning fra 50 Hz brum i et måleapparat. Det kan også benyttes for måling af harmonisk forvrængning fra ved at undertrykke første harmoniske og kun lade forvrængningen passere. Det kaldes for et stopfilter (*notch filter*) og den mest populære form ses herunder. Der er to T-formede led så det kaldes et *Twin-T filter* i engelsk litteratur.



Figur 255 – 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.



$$|H_{BS}| \approx \frac{|f_0^2 - f^2|}{2df_0^2} \quad \text{hvor } f \approx f_0$$

$$\Delta f \approx 2d\alpha f_0 \quad \text{hvor } \alpha = 10^{-D/20\text{dB}}$$

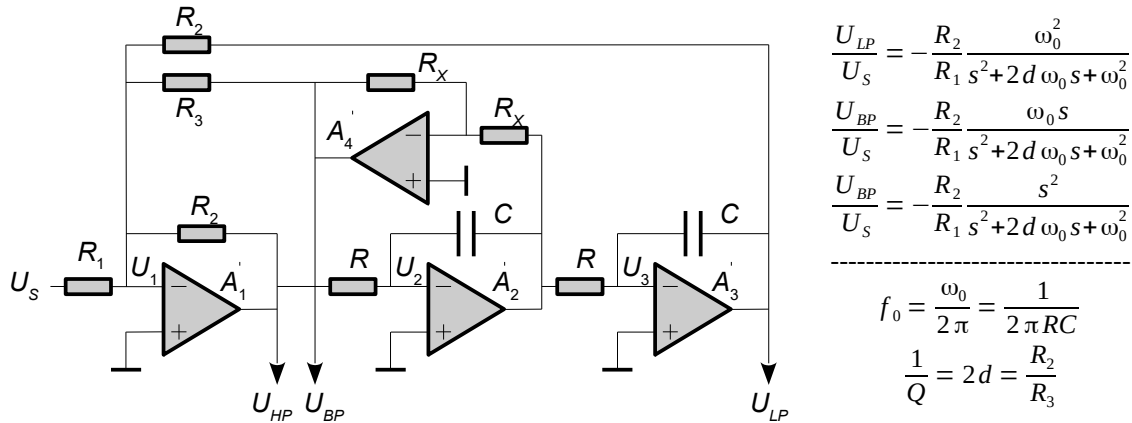
Figur 256 – Dæmpningen for et stopfilter ved 1000 Hz som funktion af dæmpningen fra $d = 0,1$ (smalt bånd) til $d = 1$ (bredt bånd med $\beta = 0$).

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). Det er 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.

4.5.15 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.



Figur 257 – 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 alle positive indgange på nul så problemer med common-mode vil ikke optræde. Dertil opnås et enklere udtryk for indstilling af filtrets godhed og alle udgange får samme polaritet i pasbåndet. 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 knudepunktsligningerne 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 inverttere signalet. Det udnyttes at de inverterende indgange er virtuel nul.

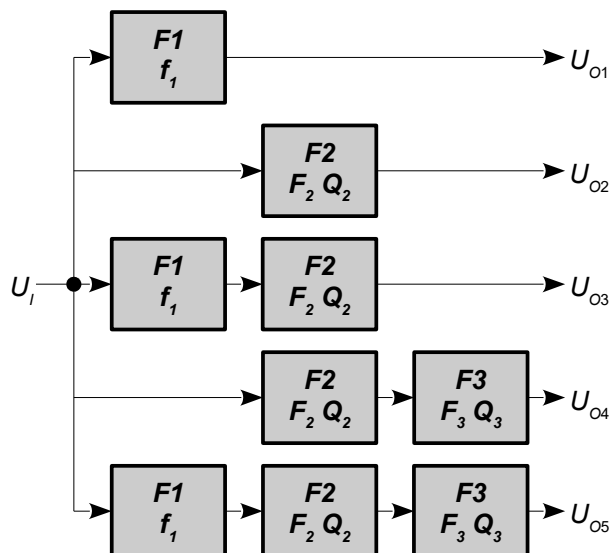
$$\begin{aligned} \frac{U_1 - U_S}{R_1} + \frac{U_1 - U_{HP}}{R_2} + \frac{U_1 - (-U_{BP})}{R_3} + \frac{U_1 - U_{LP}}{R_2} &= 0 & 0 &= \frac{U_S}{R_1} + \frac{U_{HP}}{R_2} - \frac{U_{BP}}{R_3} + \frac{U_{LP}}{R_2} \\ \frac{U_2 - U_{HP}}{R} + \frac{U_2 - U_{BP}}{1/sC} &= 0 & \Rightarrow & 0 = \frac{U_{HP}}{R} + sC U_{BP} \\ \frac{U_3 - U_{BP}}{R} + \frac{U_3 - U_{LP}}{1/sC} &= 0 & 0 &= \frac{U_{BP}}{R} + sC U_{LP} \end{aligned}$$

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.

$$\begin{aligned} U_{BP} &= -sRC U_{LP} \quad \text{og} \quad U_{HP} = -sRC U_{BP} = (sRC)^2 U_{LP} \\ U_S &= -\frac{R_1}{R_2} (sRC)^2 U_{LP} - \frac{R_1}{R_3} sRC U_{LP} - \frac{R_1}{R_2} U_{LP} \\ \frac{U_{LP}}{U_S} &= -\frac{1}{\frac{R_1}{R_2} (sRC)^2 + \frac{R_1}{R_3} sRC + \frac{R_1}{R_2}} = -\frac{R_2}{R_1} \frac{\frac{1}{(RC)^2}}{s^2 + \frac{R_2}{R_3} \frac{s}{RC} + \frac{1}{(RC)^2}} \\ \frac{U_{LP}}{U_S} &= -\frac{R_2}{R_1} \frac{\omega_0^2}{s^2 + 2d\omega_0 s + \omega_0^2} \quad \text{hvor} \quad \omega_0 = \frac{1}{RC} \\ & \quad 2d = \frac{1}{Q} = \frac{R_2}{R_3} \end{aligned}$$

4.5.16 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 der er ikke behov for at arbejde med polynomier af højere orden. Stejle filtre dannes som 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.



Figur 258 – 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.

Et femte-ordens lavpasfilter består af blokkene der leder til U_{O5} i illustrationen og herunder vises overføringsfunktionen. Filtret 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. Man kan som eksempel placere filtre med høj godhed sidst i kæden for at undgå overstyring ved højt signalniveau.

$$H_{LP5} = \frac{\omega_1}{\omega_1 + s} \frac{\omega_2^2}{\omega_2^2 + 2d_2\omega_2 + s^2} \frac{\omega_3^2}{\omega_3^2 + 2d_3\omega_3 + s^2} \quad \text{hvor} \quad d_n = \frac{1}{2Q_n}$$

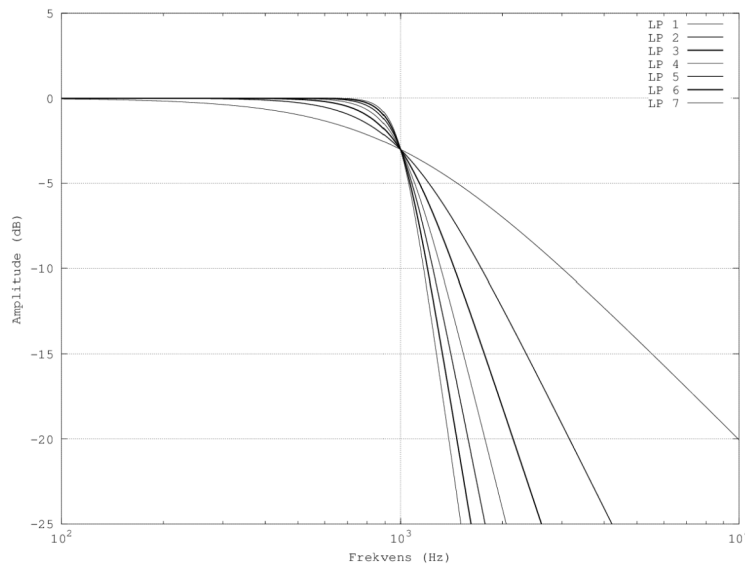
Overføringsfunktionen kan omformes til højpas som det vises lidt senere og dertil kommer en del valgmuligheder. For båndpasfiltre kan responsen være symmetrisk eller ej og der kan inkluderes nulpunkter ved at modificere tælleren. Emnet kan let vokse ud over alle grænser så jeg vil nøjes med at gennemgå nogle ganske få muligheder.

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 dB/oktav uanset de indgående parametres værdier.

$$H_{LP5} \xrightarrow{s \rightarrow 0} 1 \quad \text{og} \quad H_{LP5} \xrightarrow{s \rightarrow \infty} \frac{\omega_1 \omega_2^2 \omega_3^2}{s^5} \Rightarrow \omega_0^5 = \omega_1 \omega_2^2 \omega_3^2 \Rightarrow f_0 = \frac{1}{2\pi} \sqrt[5]{\omega_1 \omega_2^2 \omega_3^2}$$

4.5.17 Butterworth filter

Det mest populære af alle filtertyper 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 259 – 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 herunder 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.¹⁴¹

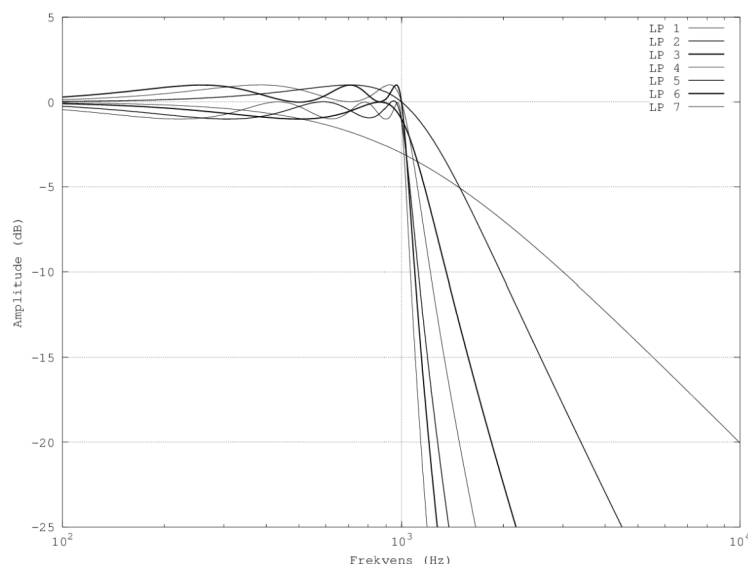
n	$F1$	$F2$ Anden orden		$F3$ Anden orden		$F4$ anden orden		A_2
	f_1/f_0	f_2/f_0	Q_2	f_3/f_0	Q_3	f_4/f_0	Q_4	
1	1,000							$-7,0$ dB
2		1,000	0,707					$-12,3$ dB
3	1,000	1,000	1,000					$-18,1$ dB
4		1,000	0,541	1,000	1,306			$-24,1$ dB
5	1,000	1,000	0,618	1,000	1,620			$-30,1$ dB
6		1,000	0,518	1,000	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 skrå asymptote (ordenen gange -6 dB/oktav) stort set er korrekt allerede ved første-ordens filtret og det er netop baggrunden for Bode plottets succes.

¹⁴¹ Sammensat af forskellige kilder på nettet og korrigeret for diverse skrivefejl. Illustrationerne benytter tabellernes talværdier.

4.5.18 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 260 – 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 filtre 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		A_2	ΔA_{B-C}
	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	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.

4.5.19 Højpasfilter

Overføringsfunktionen kan transformeres til højpas ved at substituere s/ω_1 med ω_1/s for såvel første-ordens blokken som for anden-ordens blokkene og 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.

$$H_{LP2} = \frac{\omega_2^2}{\omega_2^2 + \frac{1}{Q}\omega_2 s + s^2} = \frac{1}{1 + \frac{1}{Q}\frac{s}{\omega_2} + (\frac{s}{\omega_2})^2} \Rightarrow H_{HP2} = \frac{1}{1 + \frac{1}{Q}\frac{\omega_2}{s} + (\frac{\omega_2}{s})^2} = \frac{s^2}{\omega_2^2 + \frac{1}{Q}\omega_2 s + s^2}$$

4.5.20 Filterdesign

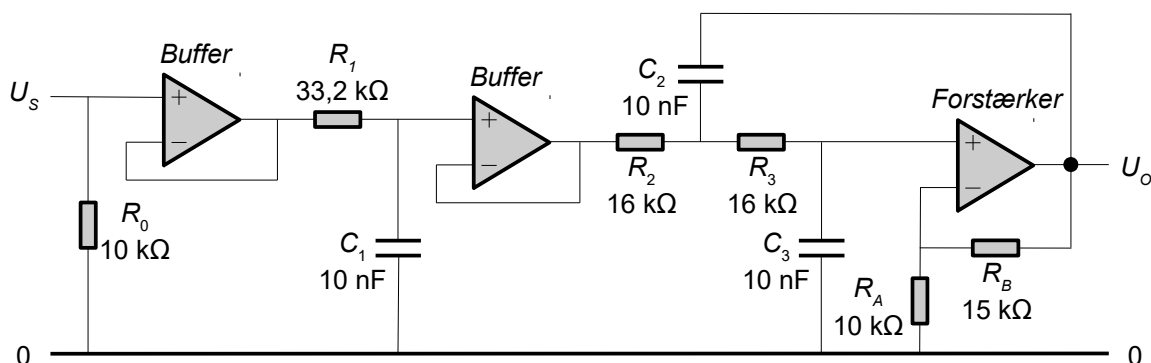
Ved design af filtre er grænsefrekvensen almindeligvis givet ud fra den påtænkte anvendelse, men den nødvendige filterorden og filtertype kan være svære at fastlægge og dertil kommer hvilken af de viste opbygninger der bør foretrækkes. I det følgende vil jeg gennemgå designet af et filter der skal dæmpe signaler over 1 kHz med mindst 20 dB i én oktavs afstand over grænsefrekvensen. Det er umiddelbart let at skalere filtret til en anden frekvens ved blot at ændre kondensatorerne.

4.5.21 Valg af filtertype og orden

Som udgangspunkt vil jeg anbefale Butterworth med mindre den stejlere 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. Det betyder at 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: $f_1 = 0,494 \cdot f_g = 494 \text{ Hz}$

Anden-ordens sektion F2: $f_2 = 0,997 \cdot f_g = 997 \text{ Hz}$ $Q_2 = 2,017$



Figur 261 – 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 $22,5 \text{ dB}$ 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. Det 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.

4.5.22 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 10 kΩ ved grænsefrekvensen og derefter vælge en kondensator så impedansen ligger i intervallet fra 1 kΩ til 100 kΩ hvilket bør sikre at operationsforstærkeren arbejder optimalt. Det giver her en kondensator på 15,9 nF og det vælges at benytte 10 nF i alle tre sektioner svarende til en impedans på cirka 16 kΩ.

$$\begin{aligned}
 X_C &= \frac{1}{2\pi f_0 C} \Rightarrow C = \frac{1}{2\pi f_0 X_C} \quad \text{hvor} \quad 1 \text{ k}\Omega \leq X_C \leq 100 \text{ k}\Omega \\
 R_1 &= \frac{1}{2\pi f_1 C} = 32,2 \text{ kohm} \quad \text{for} \quad \begin{matrix} C_1 = 10 \text{ nF} \\ f_1 = 494 \text{ Hz} \end{matrix} \\
 R_2 = R_3 &= \frac{1}{2\pi f_1 C} = 16,0 \text{ kohm} \quad \text{for} \quad \begin{matrix} C_2 = C_3 = 10 \text{ nF} \\ f_2 = 997 \text{ Hz} \end{matrix} \\
 R_B &= 2(1-d)R_A \\
 d = \frac{1}{2Q_1} &= 0,248 \Rightarrow R_B = 15,0 \text{ kohm} \quad \text{for} \quad R_A = 10,0 \text{ kohm}
 \end{aligned}$$

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 $R_A = 10 \text{ k}\Omega$.

4.5.23 Valg af operationsforstærkere

Den valgte operationsforstærker bør have en GBP specifikation på 100 gange grænsefrekvensen gange forstærkningen gange godheden med mindre man kan er klar til at acceptere en afvigelse fra den beregnede karakteristik. I det aktuelle tilfælde kræves en operationsforstærker med en båndbredde på $GBP > 300 \text{ kHz}$ og det tillader brug af næsten alt der kan købes.

$$\begin{aligned}
 GBP > 100 A Q f_0 \Rightarrow GBP > 303 \text{ kHz} \quad \text{for} \quad \begin{matrix} A = 1,5 \\ Q = 2,017 \\ f_0 = 1 \text{ kHz} \end{matrix}
 \end{aligned}$$

Dertil kommer krav til indgangsimpedans og -kapacitet for ikke at belaste de komponenter der definerer grænsefrekvensen og at udgangen 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 førnævnte optimale værdi omkring 10 kΩ.

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.

4.5.24 Passive filtre

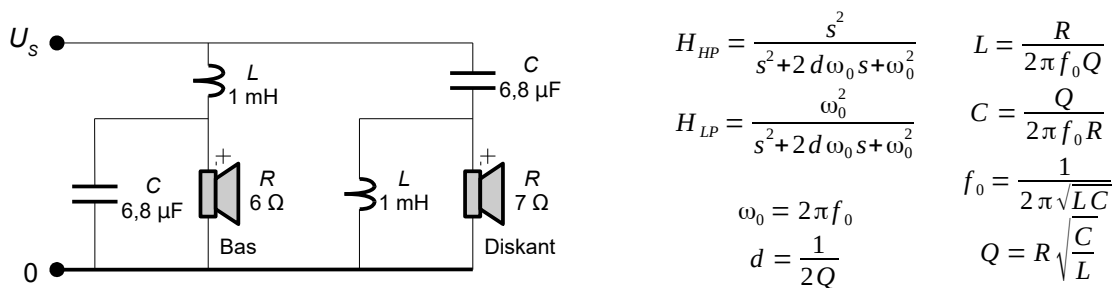
Filtre udelukkende med R , C og L kaldes for passive da der ikke er regeneration fra en aktiv komponent som operationsforstærkeren. Passive filtre er altid stabile og uden effektforstærkning. Udgangssignalets amplitude kan godt være højere end indgangssignalets amplitude, men det er kun i et smalt frekvensområde og der vil altid tabes effekt i et passivt filter.

I min optik er der to typer af passive filtre som kan have interesse ved moderne elektronik. Den ene type deler frekvensområdet mellem to højttalerenheder og den anden type benyttes ved højfrekvens hvor LC-filtret danner båndpasfiltre som udvælge et smalt frekvensområde for den videre signalbehandling. Dertil kommer SAW-filtre (*surface acoustical wave*) der minder om krystaller og kvartbølgefiltre, men de falder begge udenfor sigtet med bogen.

Der er skrevet metervis af bøger om passive filtre så de to følgende afsnit er blot en kort introduktion til emnet.

4.5.25 Delefilter

Den elektro-dynamiske højttaler kan ikke gengive hele det hørbare frekvensområde fra cirka 15 Hz til 15 kHz uden betydende kompromisser med hensyn til lydtryk og lydets spredning. Derfor deles frekvensområdet i to eller flere frekvensbånd som varetages af højttalere dedikerede til enten bas, mellemtone eller diskant. Herunder vises et klassisk to-vejs filter af anden orden.



Figur 262 – Et populært delefilter af anden orden. Komponentværdierne realiserer en delefrekvens på $f_0 = 1,9$ kHz og godheden er $Q = 0,5$ for bassen og $Q = 0,6$ for diskanten. De to enheder dæmpes med 6 dB for bassen og 4 dB for diskanten ved f_0 hvor fasen drejes 180° mellem enhederne.

Filtrene skal dimensioneres efter højttalerens DC modstand fra svingspolens kobberleder og den er aldrig 8Ω selv om megen litteratur tager udgangspunkt ved den nominelle værdi. Almindeligvis vil modstanden være omkring 80 % af den nominelle værdi så en 8Ω højttaler vil have en DC modstand omkring $6,4 \Omega$. Højttaleren har en egenresonans og en selvinduktion som vanskeliggør en korrekt dimensionering og ethvert akustisk design benytter derfor mange testforsøg.¹⁴²

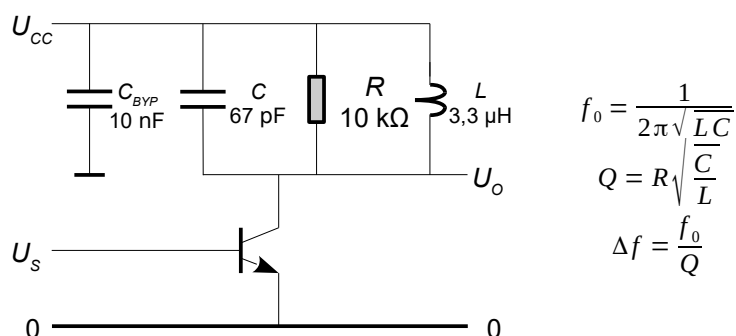
Overføringsfunktionen er af anden orden hvilket vil sige at højpasfiltret H_{HP} dæmper frekvenser under grænsefrekvensen f_0 med 12 dB/oktav eller 40 dB/dekade og lavpasfiltret H_{LP} dæmper på tilsvarende vis frekvenser over grænsefrekvensen.

Omkring grænsefrekvensen bestemmes forløbet af godheden Q der typisk vælges til $Q = 0,5$ for at sikre en konstant amplitude af udgangssignalet hvis faseforholdene ellers kan bringes til at passe. Ved grænsefrekvenser er H_{LP} 90° bagud og H_{HP} er 90° forud så der dannes destruktiv interferens ved delefrekvensen f_0 . Det løses i praksis ved enten at ombytte diskanthøjttalerens plus og minus eller ved at forskyde højttalerenhederne i forhold til hinanden svarende til en halv bølgelængde. Det er med 180 mm ved illustrationens 1,9 kHz ret meget, men der vil altid være en fasedrejning i højttalerenhederne som bør medtages i beregningen.

¹⁴²Se Elektroakustik fra www.torean.dk.

4.5.26 LC-filtre

En vigtig kobling benytter kondensator og spole i parallel. De danner resonans på den frekvens f_0 hvor impedansen stiger til maksimum. Brugt som kobling mellem to dele af et kredsløb kan det fungere som et båndpasfilter der kun tillader et smalt frekvensbånd at passere. Filtrets båndbredde Δf betegnes ved frekvensområdet hvor impedansen er indenfor -3 dB fra maksimum ved f_0 og det styres gennem resonansens godhed Q .



Figur 263 – Et LC-filter som kunne danne mellemfrekvensforstærker i en FM modtager.

Impedansens maksimum ved frekvensen f_0 er givet ved de samlede tab der skyldes modstanden i kondensator og spole samt især tabene fra det omgivende kredsløb. Tabene repræsenteres af R i illustrationen og er primært givet ved parallellforbindelsen af transistorens udgangsmodstand der er af størrelsesordenen $50\text{ k}\Omega$ og indgangsmodstanden på det efterfølgende kredsløb der normalt er så lav at spolen udformes som en transformator for at op-transformere belastningen. Derved tabes der noget af signalets forstærkning, men til gengæld opnås et smalt båndpas.

Med de viste komponenter er resonansens frekvens $f_0 = 10,7\text{ MHz}$ og godheden er $Q = 45$ for en belastning på $R = 10\text{ k}\Omega$ så båndbredden er $\Delta f = 230\text{ kHz}$. For at fungere i praksis i et kredsløb skal kondensatoren C inkludere udgangskapaciteten fra transistoren, spolens parasitiske kapacitet og kapaciteten fra det efterfølgende trin. De er alle af størrelsesordenen nogle få pico-farad så der vil måske komme til at stå 56 pF i diagrammet. Med nogen tilpasning kan det være et eksempel på en forstærker til en FM radios mellemfrekvens.

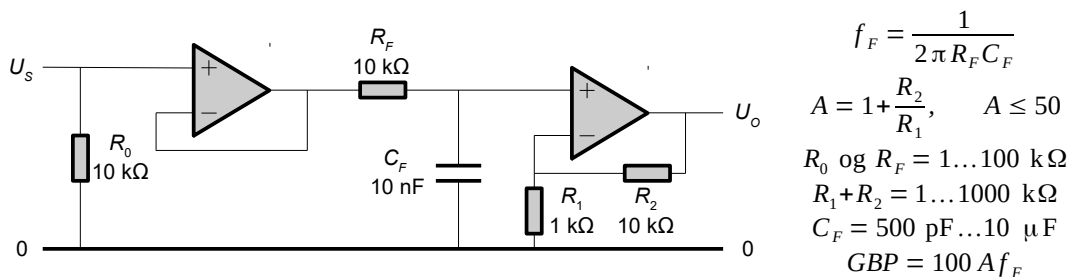
Kondensator C_{BY} danner afkobling af effektforsyningen og er nødvendig for at få kredsløbet til at virke. Den sikrer en lav impedans fra U_{CC} til nul og bør sidde tæt på kredsløbet. Værdien skal være stor nok til at have en lav impedans ved det aktuelle frekvensområde, men ikke så stor at den parasitiske selvinduktion giver problemer med selv-resonans (se første kapitel).

4.6 Designregler

Elektronik stiller nogle krav til et fornuftigt valg af komponenternes værdier, afkobling af effektforsyningen og den mekaniske opbygning på printet som ikke fremgår af de tidligere afsnit. Dertil kommer at komponenterne ikke er så ideelle som man kunne håbe på og det giver nogle begrænsninger når frekvensområdet udvides opad. Dette afsnit bringer nogle udpluk af forklaring, råd og vejledning der gælder generelt for alle typer af elektronik.

4.6.1 Komponentværdier

Valget af komponentværdier og krav til operationsforstærkeren kan virke baseret på tilfældigheder, men det er muligt at opstille nogle regler så et hensigtsmæssigt valg kan foretages.



Figur 264 – En operationsforstærker vil altid skulle drive en belastning. Det vejledende område for valget af komponenter begrundes i teksten.

Bufferen skal drive en strøm gennem R_F og ved frekvenser over grænsefrekvensen f_F for filtret er impedansen af C_F så lav at kun modstanden belaster bufferens udgang. En operationsforstærker kan typisk drive ± 10 mA til en belastning så ved et signal på ± 10 V vil $1 \text{ k}\Omega$ være et minimum. Modstanden skal levere den indgangsstrøm som forstærkeren behøver og det vil give et DC spændingsfald over den. For en strøm på 100 nA til operationsforstærkeren (*input bias current*) og et spændingsfald på $0,1 \text{ V}$ over modstanden bliver $100 \text{ k}\Omega$ maksimum.

Modstanden R_0 har det formål at sikre en DC forbindelse til nul og designes på samme måde.

For at opnå mindre end 1% påvirkning af filtrets funktion skal udgangsmodstanden fra bufferen være under $R_F/100 = 10 \text{ }\Omega$ og det kan normalt opnås. Indgangsmodstanden på forstærkeren skal tilsvarende være over $100 R_F = 10 \text{ M}\Omega$ og det er oftest tilfældet. Der opstår først problemer ved så høje frekvenser at operationsforstærkeren ikke har nok åben-sløjfe forstærkning til at garantere de krævede værdier. Forstærkerens *GBP* specifikation skal som en tommelfingerregel være mindst $100 A f_F$ hvor A er den indstillede forstærkning og f_F er den højeste frekvens af interesse.

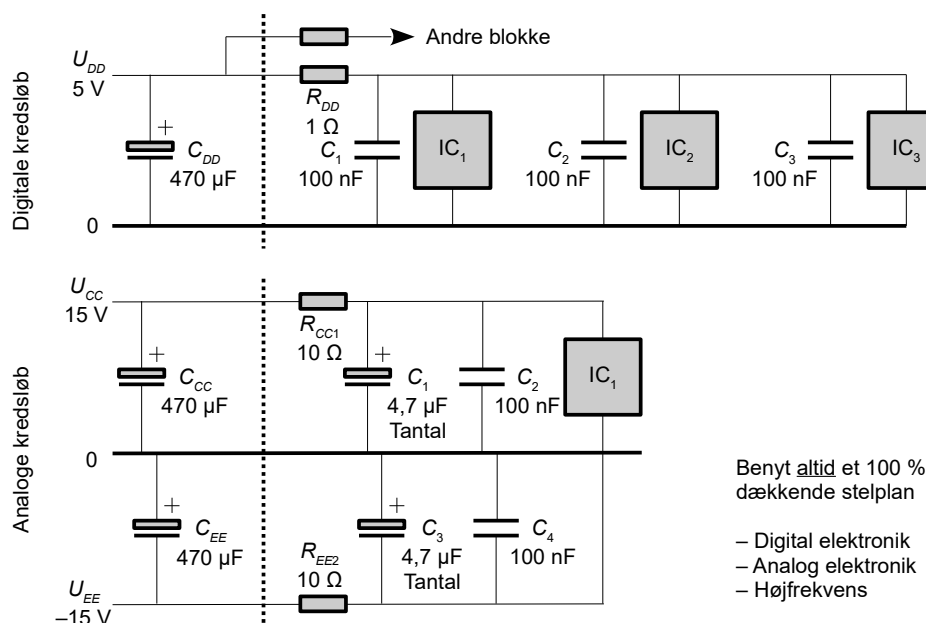
Operationsforstærkere har en fejlspejding (*input offset voltage*) der adderes til signalet som skal forstærkes. Med 5 mV som den typiske værdi og måske $0,5 \text{ V}$ acceptabel spænding på udgangen bør den indstillede forstærkning være under 50 gange da der her er to OpAmp i kaskade.

En tommelfingerregel siger at der må påregnes en kapacitet fra printkortets baner på cirka 1 pF per centimeter. Forstærkeren har dertil en indgangskapacitet på typisk 2 pF og de to kapaciteter ligger i parallel med C_F så den effektive værdi af kondensatoren øges med 5 pF eller deromkring. Det betyder at C_F ikke bør vælges mindre end 500 pF hvis fejlen skal holdes under 1% .

Der er sjældent problemer med kondensatorer hvis dielektrikummet er af keramik og plastfilm, men kapacitetsværdier over $10 \text{ }\mu\text{F}$ kan sjældent opnås uden kompromis med kvaliteten. Tilsvarende bør man generelt undgå at bruge elektrolytiske kondensatorer ved signalbearbejdning på grund af høj lækstrøm, harmonisk forvrængning, dårlig tolerance, ringe temperaturområde, kort levetid og en række af parasitiske komponenter som fx seriemodstand og serieinduktion.

4.6.2 Afkobling

Digitale kredse har kortvarige strømtræk når udgangen skifter og det kan ikke leveres gennem en kobberbane på printet på grund af dens serieinduktion. Det giver et behov for en kondensator tæt ved komponenten for at levere strømmen. Jeg anbefaler 100 nF for hver eneste digitale kreds. Ved større integrerede kredse anbefales 100 nF for hver eneste tilslutning til effektforsyningen og de skal monteres tæt på komponenten med det andet ben til stelplanet. En gruppe af digitale kredse kan isoleres fra andre kredsløb af ferritperler eller modstande på 1 Ω for at dæmpe svingning og mindske gensidig påvirkning. En modstand giver en mikroprocessor muligheden for at overvåge DC strømforbruget.



Figur 265 – Afkobling af elektroniske komponenter.

Analoge kredse har brug for afkobling af effektforsyningen for stabilitet og undertrykkelse af støj fra effektforsyningen, men der er sjældent et stort strømtræk. Som minimum anbefales 100 nF ved samtlige forbindelser til effektforsyningen hvilket typisk er to per komponent. Ved mere krævende anvendelse udbygges det som vist i illustrationen med 10 Ω og 4,7 μ F der danner et lavpasfilter som dæmper højfrekvente signaler. Databladets anbefalinger skal absolut følges.

Man bør som regel supplere den lokale afkobling af de digitale og analoge komponenter med en fælles afkobling af effektforsyningen med 470 μ F der sikrer mod kortvarige dyk i effektforsyningen fra relæer, lamper og andet der kan kobles ind og ud.

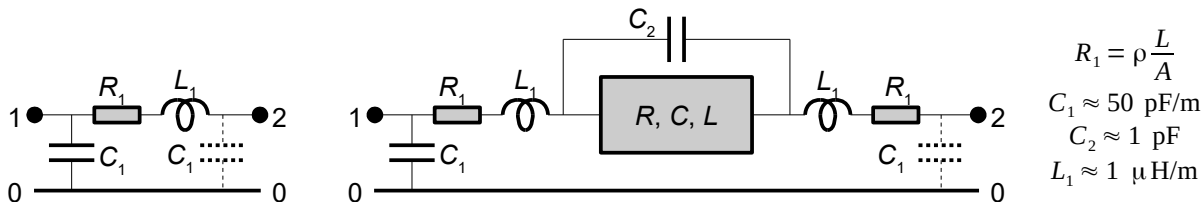
4.6.3 Stelplan

Erfaringen fra design af radiomodtagere og digitale kredsløb samt test efter EMC direktivet er en generel anbefaling om at benytte et gennemgående kobberlag på printkortet som 0 V (GND). Det kaldes for et stelplan (*ground plane*). Det kan lyde som luksus, men det giver en lav impedans ved høje frekvenser som både sikrer stabilitet i funktionen og immunitet overfor ekstern støj.

Det er muligt at benytte et stelplan allerede ved en prototype på "fumlebrædt" som leveres med et sækkende kobberplan i tilgift til loddeøer som via-huller. Et kort koster omkring 200 kr og kan spare en for mange timers fejlsøgning. Ved printudlægning anbefaler jeg ubetinget at reservere et af kobberlagene som et stelplan. Lad være med at bruge sammenkoblede kobberarealer (*area fills*) for at spare de ekstra lag for det betaler sig ikke.

4.6.4 Parasitiske komponenter

De passive komponenter betragtes ofte som om de realiserer definitionsligningerne, men ingen komponent er ideel og det kan være nødvendigt at tage hensyn til afvigelserne. Her vises modeller af både en leder og en generisk komponent som modstand, kondensator eller selvinduktion. Modellerne er relativt omfattende, men det er ikke nødvendigt at benytte alle dele ved en analyse. Man kan medtage det der er af betydning for den aktuelle applikation.



Figur 266 – Modeller af en lederbane og en passiv komponent sammen med tommelfingerregler for beregning af de parasitiske komponenters værdier.

En ledning eller en lederbane på et printkort er vist til venstre i illustrationen og vil repræsentere en modstandsværdi R_1 fra lederens materiale ρ , længden L i strømretningen og arealet A af dens tværsnit. En bane på et printkort vil have en modstand på cirka $2 \text{ } \Omega/\text{m}$ (se første kapitel) så 1 cm bane repræsenterer $0,02 \text{ } \Omega$ der ofte kan ignoreres.

Ledningen har en selvinduktion fra det magnetiske felt. En tommelfingerreglen for selvinduktionen er $1 \text{ } \mu\text{H/m}$ som skyldes at lederen danner en løkke med det omkringliggende kredsløb så en bane på 1 cm vil svare til 10 nH . Lederen kobler elektrisk med et stelplan og danner en kapacitet med en værdi på cirka 50 pF/m så en bane på 1 cm længde vil svare til 1 pF .

Opfattes en ledning som både kondensator og selvinduktion ses at blot 1 cm bane på et printkort vil danne resonans ved $1,6 \text{ GHz}$. Modellen holder næppe i dette frekvensområde, men det viser at man ikke kan benytte den simple antagelse om ideelle komponenter ved design af fx en Bluetooth komponent der arbejder på $2,5 \text{ GHz}$.

En modstand vil i et kredsløb have en kobling mellem de to tilledninger der let kan nå op på 1 pF . Impedansen af kondensatoren er $1,6 \text{ M}\Omega$ ved 100 kHz så en modstand på $1 \text{ M}\Omega$ kan ikke umiddelbart anvendes ved høje frekvenser.

En kondensator har en seriemodstand fra sin konstruktion kaldet den ækvivalente seriemodstand *ESR* (equivalent series resistance) som opgives i databladet. For et dielektrikum som keramik eller plastfilm er den under $0,1 \text{ } \Omega$ og vil kun sjældent genere. For aluminium elektrolytkondensatorer kan den nå op over $10 \text{ } \Omega$ ved de lave værdier i området under $10 \text{ } \mu\text{F}$, mens den falder med stigende værdi så man bør konsultere databladet. Ved en tantal elektrolytkondensator er den ækvivalente seriemodstand lav og de er populære for afkobling af effektforsyningen.

En kondensator har en serieinduktion fra tilledningen og banerne på printkortet. Ved 1 cm samlet længde udgør det 10 nH som vil gå i resonans med kondensatorens kapacitet. For en 100 nF kondensator anvendt for afkobling af en effektforsyning vil den gå i resonans ved 5 MHz og ved højere frekvenser er kondensatoren induktiv.

En selvinduktionsspole vil have en modstandsværdi fra den tråd den er viklet med og den beregnes af formelen for R_1 . Dertil en parallelkapacitet C_2 fra beviklingen der er omkring 100 pF for de større spoler så en spole på $100 \text{ } \mu\text{H}$ kan gå i resonans med sig selv på 2 MHz og ved højere frekvenser er spolen en kapacitet.

5 HALVLEDER

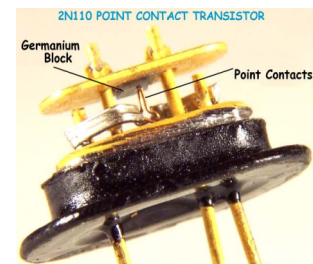
Ensretteren kan dannes af to ledende materialer der bringes i elektrisk kontakt og det var i brug inden radorøret blev opfundet selvom ingen vidste hvorfor det virkede. Den bipolare transistor BJT, felleffekt transistoren JFET og den dominerende MOSFET introduceres i dette kapitel. Moderne elektronik leveres som integrerede kredsløb, men der er stadig brug for diskret elektronik for effektforstærkere, radiokredsløb og energiforsyning. Her er fokus primært på analog elektronik og det er nødvendigt at starte med dioden.

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 272, felleffekt transistorerne JFET og MOSFET starter side 281 og den typiske opbygning af en operationsforstærker beskrives fra side 306 med nogle specialtyper fra side 321.

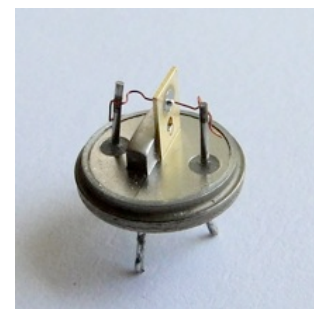
5.1 Historisk oversigt

Den bipolare transistor blev opdaget ved et tilfælde i 1947 hvor Bardeen og Brattain ved Bell Lab forsøgte at få Lilienfelds felleffekt transistor til at fungere med et krystal af germanium som basis. De brugte tynde metaltråde til at skaffe elektrisk kontakt med krystallet og vidste ikke at en kontakt mellem metal og halvleder udgør en diode. De fik kredsløbet til at give forstærkning og patenterede det og året efter fandt Shockley frem til den teoretiske forklaring.¹⁴³

1874	Blyulfid ensretter (Braun)
1897	Radiotelegrafi (Marconi)
1902	Radiotransmitter (Poulsen)
1907	Radorør (de Forest)
1926	Felleffekt transistor patent Al/Cu ₂ S (Lilienfeld)
1931	Halvlederteori (Wilson)
1935	MOSFET patent (Heil)
1939	MESFET teori (Shockley)
1940	PN halvlederen fremstilles (Schaff, Theuerer, Ohl)
1947	BJT ved Bell Labs (Bardeen, Brattain)
1948	BJT fremstillet ved diffusion (Shockley)
1953	JFET (Decay, Ross)
1958	Integreret kreds (Kilby)
1960	BJT i produktion (Fairchild)
1961	BIMOS (Sah)
1962	CMOS (Wanlass, San, Moore)
1962	MOSFET i produktion (RCA)
1963	Monolitisk operationsforstærker (Widlar, Fairchild)
1970	DRAM 1 kbit i PMOS (Intel)
1971	Mikroprocessor 4004 (Intel)
1972	Ur på chip
1974	Motorola 6800 kører på kun 5 V
1976	Mikroprocessor 1802 i CMOS (RCA)



Point-contact transistor (IBM).



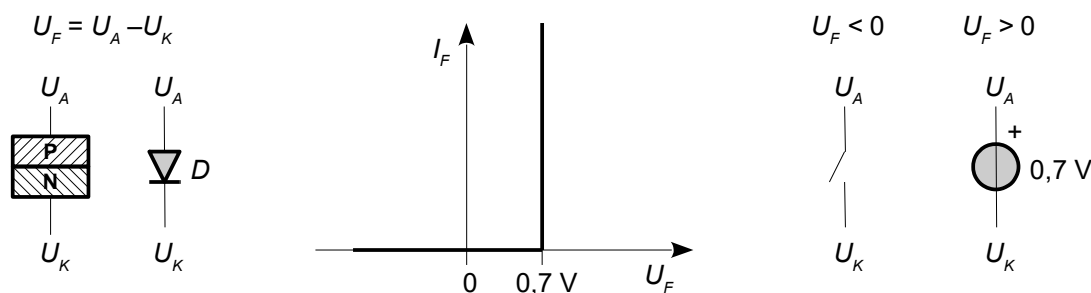
Junction transistor (IBM).

Figur 267 – Halvlederens udviklingsforløb fra diode til transistor og til integreret kredsløb.

¹⁴³https://en.wikipedia.org/wiki/History_of_the_transistor.

5.2 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 268 – Simple 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.

Spændingsfaldet over dioden i lederetningen udnyttes i mange kredsløb som en reference for det øvrige kredsløb. Det kaldes for et kunstigt nulpunkt. Andre kredsløb måler spændingsfaldet som et udtryk for temperaturen da der er en relation på cirka $-2 \text{ mV}/^\circ\text{C}$ for dioder uanset konstruktionen. Der afsættes en effekt i dioden når strømmen er stor og det er en af tabsmekanismerne i den populære DC-DC konverter som omhandles i det følgende kapitel.

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 1 nA som afhænger af materialets type, fysisk størrelse og temperatur. Det er diodens lækstrøm (*leakage current*) og den afhænger stort set ikke af spændingen over dioden, men 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 spærreretningen (*reverse*). Der gælder tilsvarende at spændingsfaldet normalt angives som positivt i lederetningen som U_F og i spærreretningen som U_R .

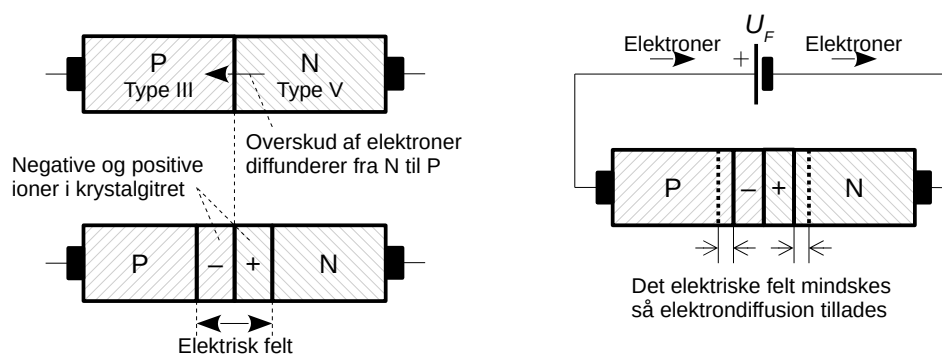
5.2.1 Diodens funktion

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 zonesmeltning. Før det blev selen og kobberoxyd brugt til ensretning af vekselstrøm og blyulfid til detektion af radiofoni. Dette afsnit giver en kortfattet introduktion til PN overgangens funktion og for en mere uddybende teori henvises til side 324.

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 naboatom 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 deltage i strømtransporten på linje med elektronen og et overskud af flytbare huller giver en P halvleder.

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 N laget. Den frie elektron vil i P halvlederen møde et type III atom der kan optage elektronen og danne en negativ ion i P laget. Diffusionen opbygger 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 og feltet. Området mellem N og P er tømt for ladningsbærere og kaldes for *depletion layer*.

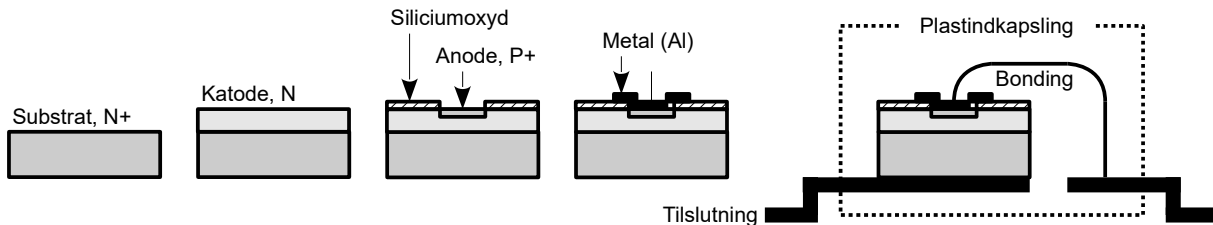


Figur 269 – Diodens N halvleder har et overskud af frie elektroner og P halvlederen har et overskud af flytbare huller. Diffusionen af elektroner danner et elektrisk felt som modvirker diffusionen. En ekstern spændingskilde kan ændre på det elektriske felt og derved styre strømmen i dioden.

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 spærrer.*

5.2.2 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 smeltede 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.



Figur 270 – Dioden fremstilles af smeltet silicium med kraftig dotering (N+) hvorpå et tyndt lag med svag dotering bygges (N). Ved en fotolitografisk proces dotes 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 dotes der flere gange og metallag benyttes til at danne interne forbindelser mellem komponenterne.

Først fremstilles et basismateriale fra et bad af smeltet silicium. Et krystal af rent silicium dyppes og trækkes langsomt op hvorved siliciumatomer udfældes på overfladen og opbygger et næsten perfekt krystal. Ved tilsætning af type III eller V under processen dotes materialet til en halvleder af type P eller N med henholdsvis underskud eller overskud af elektroner. Derefter opbygges et lag af siliciumoxyd på halvlederens overflade. Et lysfølsomt lag påføres og belyses gennem en maske og derefter kan dele af laget ætzes 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 kaldes for dotering og kan tilmed give fejl i krystalstrukturen. Ved dotering med basismaterialets forurening dannes et kraftigt doteret 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.

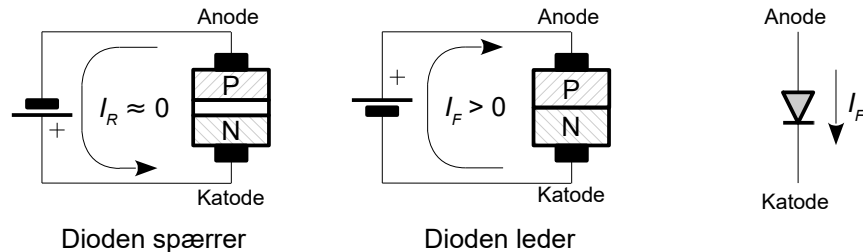
Selv om den doterede halvleder er en relativt god leder så er modstanden for høj til at man kan se bort fra den. Derfor blev processen forfinet med et *epitaxial layer*; en kraftigt doteret halvleder der udgør halvlederens basismateriale (*substrate*). Det kraftigt doterede lag (N+ i illustrationen) 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 halvlederens 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 glastrude 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 forureningen, men den høje energi som ionerne accelereres op til kan give betydelige skader på halvlederens krystal. Det formodes at være en af årsagerne til den såkaldte 1/f-støj i halvledere.

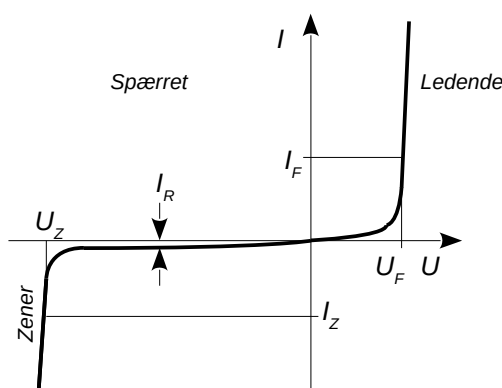
5.2.3 Diodens I-U karakteristik

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 omgivelsestemperaturen slår løs i halvlederen.



Figur 271 – Med dioden forspændt i spærreretningen øges bredden af PN overgangen og der løber ingen betydende strøm i dioden. Med dioden forspændt i lederetningen reduceres det elektriske felt over PN overgangen og diffusionsstrømmen får lov til at løbe. Ladningsbærerne i P laget er valensbåndets elektroner der hopper fra atom til atom. Det beskrives ofte som positive huller der bevæger sig i modsat retning af elektronerne. I N laget er det ledningsbåndets frie elektroner.

Diodens strøm varierer eksponentielt med spændingen over dioden og det gælder også for den bipolære transistor der introduceres side 272 og for en MOSFET i *sub-threshold* området. Det teoretiske fundament introduceres side 324.



Lederetning:

$$I_F = I_S \left(\exp\left(\frac{U_F}{nU_T}\right) - 1 \right) - I_P$$

$$U_F \approx nU_T \ln\left(\frac{I_F}{I_S}\right) \quad \begin{matrix} I_F \gg I_S \\ I_P = 0 \end{matrix}$$

Temperaturspænding: $U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$

Absolut temperatur: $T = 300 \text{ K}$ ved 27°C

Idealitetsfaktor: $n = 1 \dots 2$

Afsat effekt: $P_D = U_F I_F$

Spærreretning:

$$I_R \approx I_{R0} 2^{(T-T_0)/\Delta T}, \quad \Delta T \approx 10 \text{ K}$$

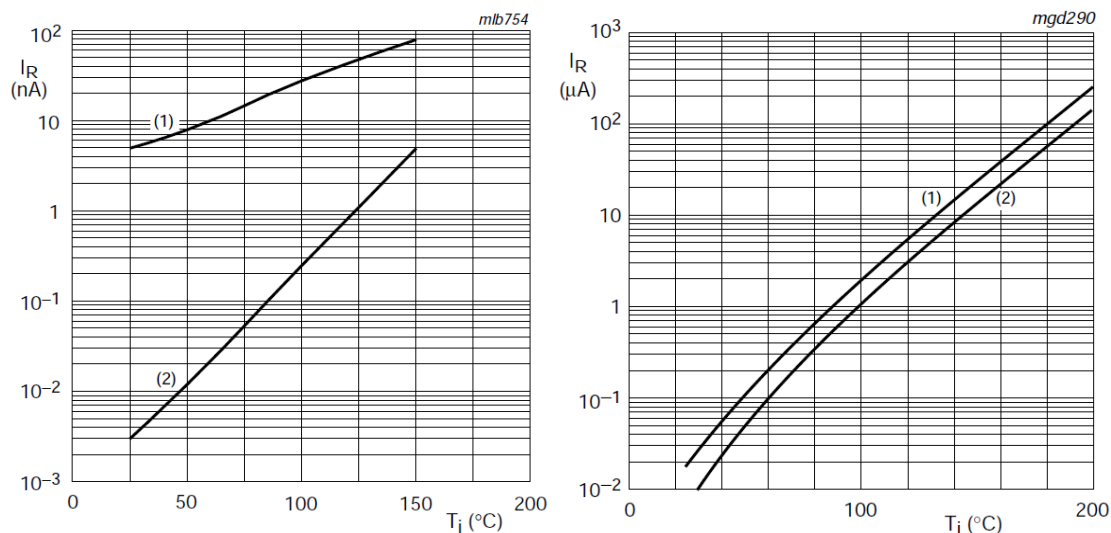
Figur 272 – I det *ledende* område varierer strømmen eksponentielt med diodens spænding. I det *spærrende* område er strømmen meget lille og varierer eksponentielt med temperaturen. I *zener* området vil diodens strøm vokse voldsomt når spændingen over dioden overskrider en grænse.

I ligningen er I_F strømmen i diodens lederetning 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 forklæres side 262.

I spærreretningen er strømmen I_R givet af flere bidrag bestående af mætningsstrømmen I_S for U_F negativ, de elektron-hulpar som temperaturen genererer i PN overgangen, strømmen i krystallets overflade og strømmen i diodens indkapsling hvis yderside er udsat for omgivelsernes forurening. Det kaldes under ét for diodens lækstrøm I_R der er stærkt afhængig af temperaturen.

5.2.4 Lækstrøm

Strømmen i spærreretningen 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 273 – 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.

5.2.5 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 324.

5.2.6 Idealitetsfaktor

Parameteren n kaldes for diodens idealitetsfaktor (*ideality factor*) der er et kompromis mellem de forskellige mekanismer der styrer strømmen i dioden. En del litteratur benytter bogstavet η (eta). Diffusionsstrømmen beskrives ved $n = 1$ mens elektron-hul rekombination i PN overgangen (*space recombination*) og strømmen i krystallets overflade beskrives ved $n = 2$. De to sidste dominerer ved lav strøm og forklarer hvorfor en transistors strømforstærkning aftager ved lav strømstyrke.

Værdien er omkring 1,7 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.

5.2.7 Fotostrøm

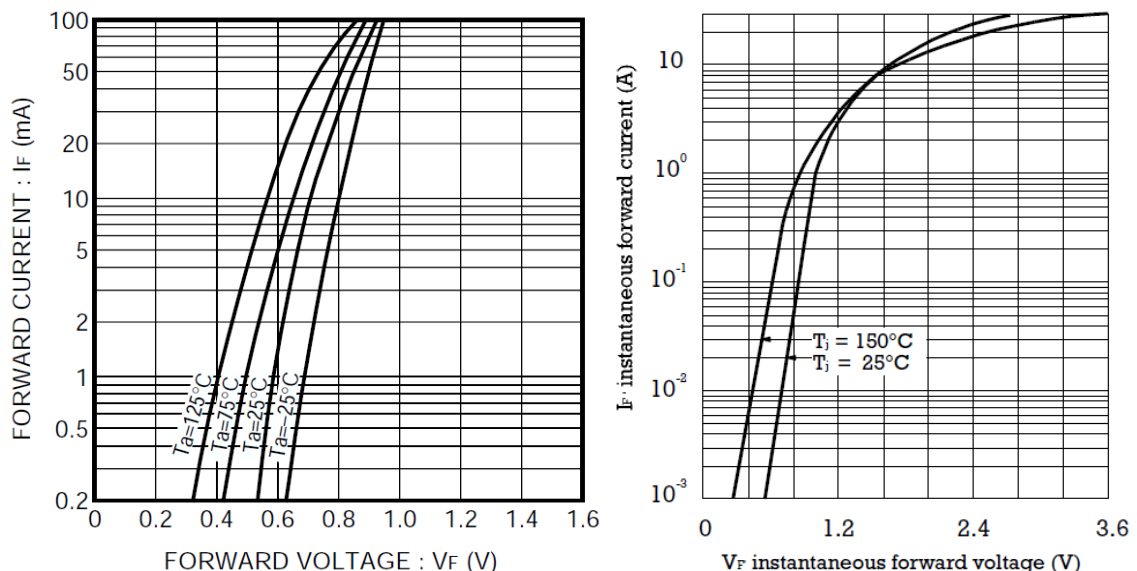
Parameteren I_p repræsenterer den strøm som absorption af en foton giver anledning til når den slår en elektron løs i PN overgangen. Elektronen tiltrækkes af den positive ladning fra ionerne i N halvlederen og føres til katoden som en overskydende elektron. Det er grunden til at en fotodiode har strømmen 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 269).

5.2.8 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 absolutte spændingsværdi varierer med teknologien som det vises side 256. Den logaritmiske egenskab er meget præcis og jeg skal senere vise eksempler baseret på den.

Herunder vises den publicerede karakteristik for to populære dioder. 1N4148 udmærker sig ved at dukke op alle vegne hvor der er brug for "en diode" og 1N4007 tåler en højere værdi af både strøm i lederretningen og spænding i spærreretningen før sammenbrud. Dioder for høj strømværdi har et stort areal ved PN overgangen og kan derfor opfattes som mange små dioder i parallel. De deles om strømmen og det er blandt andet derfor de "0,7 V i lederretningen" har bidt sig så godt fast.

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 omtrent logaritmisk op til 10 mA og 1N4001 til 1 A. Internettet er rig på data for andre dioder.¹⁴⁴



Figur 274 – 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 begrænsningen 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 betydende for dens værdi stiger med temperaturen.

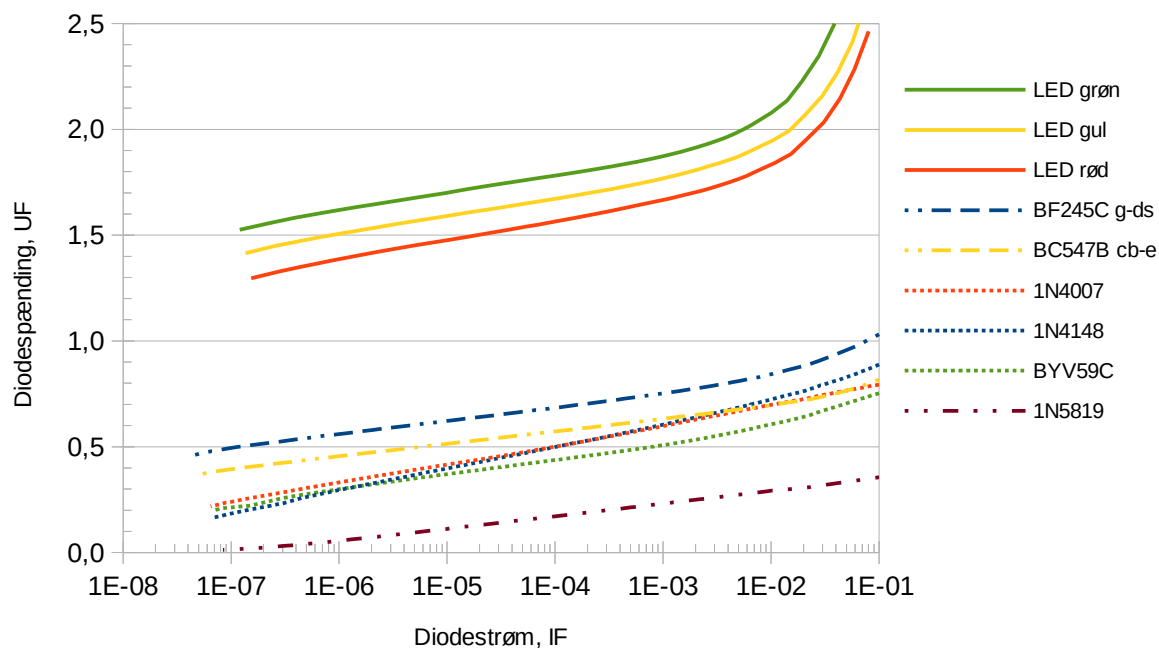
Der ses en relation til temperaturen ved den horisontale forskydning af kurverne. Spændingsfaldet aftager ved stigende temperatur for fastholdt strømværdi hvilket kan udtrykkes som større villighed til at lede strømmen ved stigende temperatur. For dioden 1N4148 indikerer kurven at spændingen aftager 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.

¹⁴⁴Katalogfirmaer: 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 ledespændingsfaldet er på cirka 2 V ved en strøm på 10 mA. Det ses at den grønne lysdiode har et højere spændingsfald end den røde fordi der kræves mere energi til at accelerere elektronerne op i fart. Materialet består af type III (aluminium, gallium og indium) og type V (fosfor og arsen) da en type III-V halvleder er bedre til at danne lys end silicium. Hældningen er 90 mV/dekade så idealitetsfaktoren er $n = 1,5$. Kurverne krummer opad ved en strøm over 1 mA på grund af modstanden i halvlederens materiale der beregnes til cirka $10\ \Omega$ ud fra de 0,1 V ekstra spændingsstigning ved 10 mA (se side 262).



Figur 275 – Spændingsfaldet over et antal dioder som funktion af strømmen.

Målingen inkluderer ikke dioder af germanium der ville ligge omtrent som Shottky-dioden. Det vil sige et lavt ledespændingsfald og en høj lækstrøm for en diode i spærreretningen.

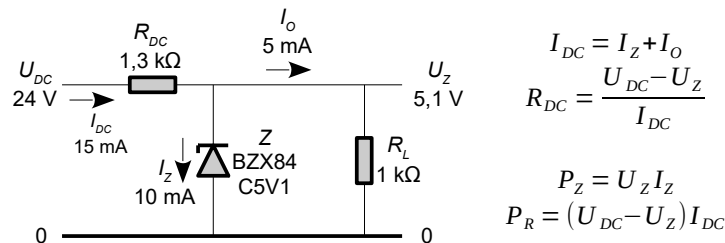
De følgende kurver er for halvledere af silicium hvor de to første er for diodekoblede transistorer. Øverst en JFET med drain forbundet til source ($n \approx 1,1$) og derefter en BJT med kollektor til basis ($n = 1,0$). Den beregnede værdi af halvlederens modstand var af størrelsesordenen $1\ \Omega$ 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 som det følger af ligningen side 253 når lederetningens strøm I_F fastholdes.

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 $1\ \Omega$ 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 for en strøm under 1 mA da idealitetsfaktoren er højere for dioderne.

Den sidste diode er 1N5819 ($n \approx 1,0$) som er en Shottky-diode og altså består af et metallag mod en siliciumhalvleder. Den har den laveste af de beregnede modstandsværdier på cirka $0,1\ \Omega$. Den lave værdi af ledespænding og modstand forklarer hvorfor Shottky-dioder er populære ved DC-DC konvertere hvor effekttabet er af primær betydning. Det lave spændingsfald vil også betyde at mætningsstrømmen bliver høj og den blev målt til 0,1 mA. Dioden er derfor ikke god til at spærre for strømmen.

5.2.9 Zenerdiode

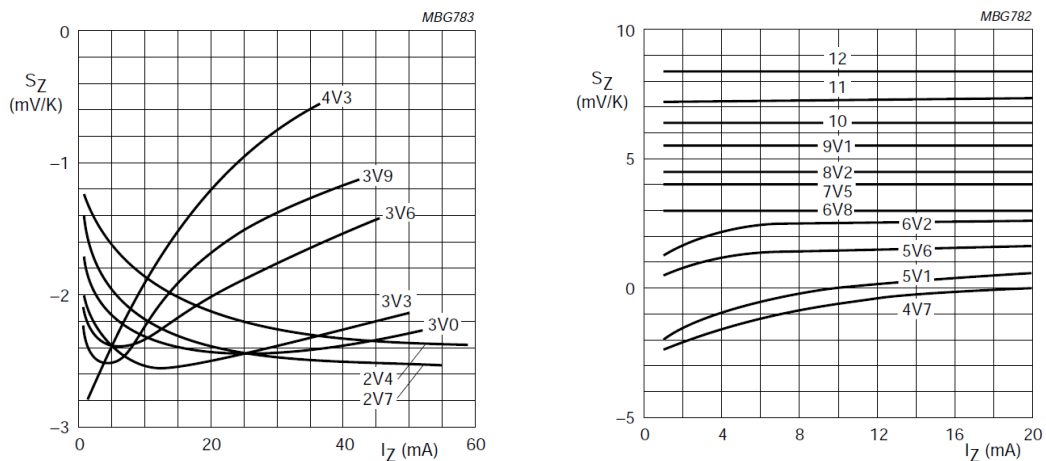
I spærreretningen er lækstrømmen lav indtil spændingen bliver så stor at PN overgangen bryder sammen og der løber en strøm I_Z i spærreretningen. Det er ikke destruktivt blot effektafsættelsen holdes vel under databladets specifikation på 0,25 W for en typisk småsignaldiode. Værdien af spændingen kan designs fra 3 V til flere kilovolt ved at styre halvlederens doteringsprofil.¹⁴⁵



Figur 276 – 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 $I_Z = 10$ mA for zenerdioden og dertil med $I_O = 5$ 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 ± 5 % af spændingsværdien U_Z ved en specificeret strømværdi, men de kan leveres med en tolerance på ± 2 %. Zenerdioden var en populær komponent indtil de integrerede kredse slog igennem med prisbillige alternativer som fx 7800-serien af spændingsregulatorer med en lignende tolerance. Spændingsreferencer leveres med en tolerance på eksempelvis $\pm 0,1$ % for LM4051A eller $\pm 0,02$ % for ADR4520.



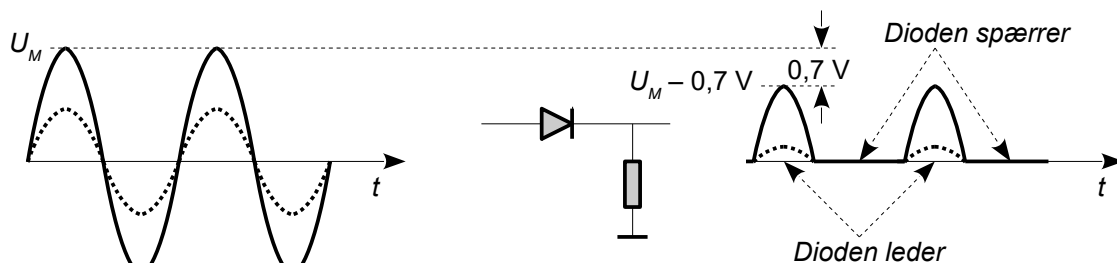
Figur 277 – Temperaturkoefficienten for dioden BZX84 serien er funktion af strømmen i dioden og skifter fortegn ved en spændingsværdi omkring 5 V.

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.

¹⁴⁵Der 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.

5.2.10 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.

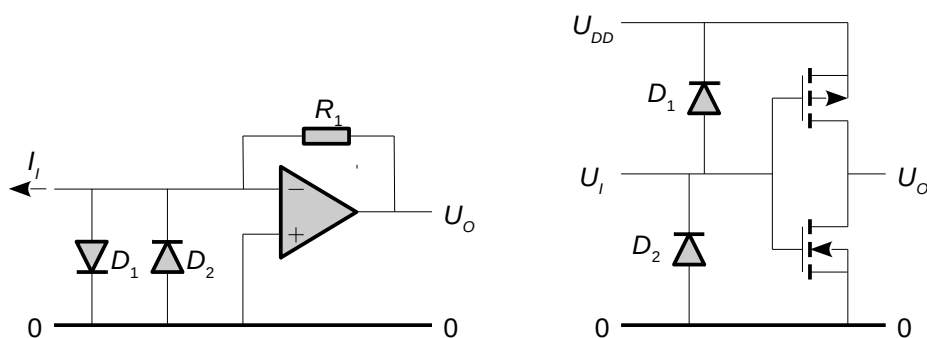


Figur 278 – Halvbølge ensretning af et vekselstrømsignal giver en pulserende jævnstrøm for der kan kun løbe strøm i diodens lederetning så udgangsspændingen kan ikke blive negativ.

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 bæreølge (*carrier*). Amplituden ændres i takt med det lavfrekvente signal og det resulterende signal sendes fra en antenne til modtageren der modtager den valgte bæreølge 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 bæreølgen mens TV-transmissionen fortsætter med AM indtil 2009.

5.2.11 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 antiparallelle dioder kan begrænse spændingen på indgangen til $\pm 0,7$ V. Diodernes mætningsstrøm kan påvirke funktionen og kapaciteten i dioden kan give ustabilitet.



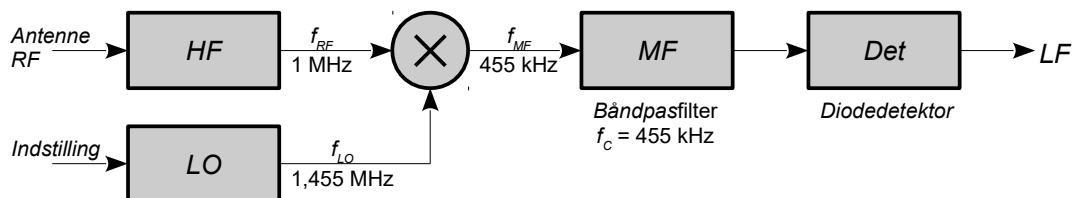
Figur 279 – Diodernes beskytter en følsom indgang imod overspænding (venstre), og hindrer en ødelæggelse af de højimpedante indgange på en CMOS inverter (højre).

De populære CMOS kredsløb 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).¹⁴⁶

¹⁴⁶Denne beskrivelse dækker hovedparten af CMOS4000 og 74HC-familierne og lavspændingsvarianter, alle mikroprocessorer og de programmerbare logiske komponenter, samt flertallet af analoge integrerede kredse.

5.2.12 Radiomodtager

En vigtig komponent i radiomodtagere er mikseren (*mixer*) der bruges til at omsætte bære-bølgen 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 gangetegn, 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.

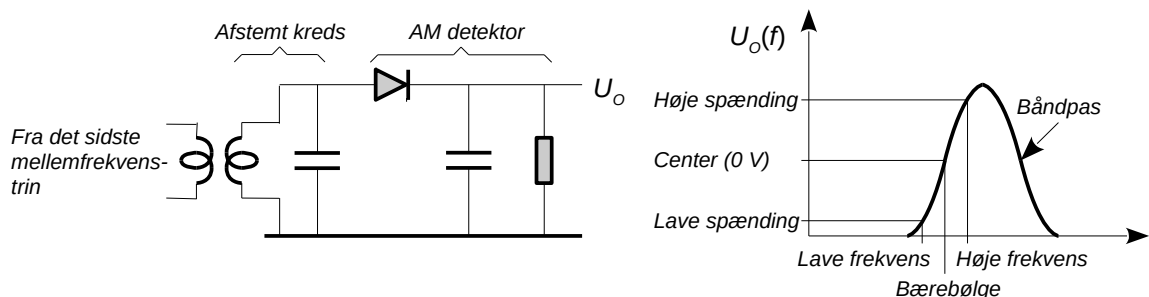


Figur 280 – En radiomodtager benytter dioden i en mikser for at omsætte et højfrekvenssignal til en lavere frekvens, der kan forstærkes op og efter detektion føres lavfrekvenssignalet til udgangen.

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 radiomodtager benytter en mellemfrekvens på 10,7 MHz og en satellitmodtager benytter flere lokaloscillatorer for trinvis nedsætning af frekvensen. Detektion 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.

En af mange metoder for at ændre en frekvensændring til en spændingsændring er at benytte den skrå 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 overstyrer og begrænses til en fast amplitude.

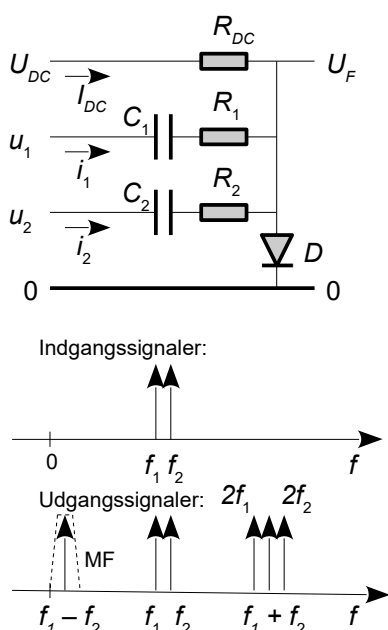


Figur 281 – I en FM modtager omsættes en frekvensændring i forhold til bære-bølgen til en analog spændingsændring som derefter kan omformes til et lavfrekvenssignal ved ensretning.

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 vendetangent i det skrå område og derved en mere symmetrisk kurve.

5.2.13 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 diode D til det ønskede arbejds punkt.



Analyse:

$$U_F = U_T \ln\left(\frac{I_{DC} + i_1 + i_2}{I_S}\right) \quad \text{hvor} \quad U_{DC} = U_T \ln\left(\frac{I_{DC}}{I_S}\right)$$

$$U_F = U_{DC} + U_T \ln\left(1 + \frac{i_1 + i_2}{I_{DC}}\right) \quad I_{DC} + i_1 + i_2 > I_S$$

Taylor approksimation:

$$U_F = U_{DC} + U_T \left[\frac{i_1 + i_2}{I_{DC}} - \frac{1}{2} \left(\frac{i_1 + i_2}{I_{DC}} \right)^2 + \dots \right] \Rightarrow$$

$$U_F = U_{DC} + u_F = U_{DC} + U_T \left[\frac{i_1 + i_2}{I_{DC}} - \frac{1}{2} \left(\frac{i_1}{I_{DC}} \right)^2 - \frac{1}{2} \left(\frac{i_2}{I_{DC}} \right)^2 - \frac{i_1 i_2}{I_{DC}^2} + \dots \right]$$

Harmoniske signaler og $|i_1| + |i_2| < I_{DC}$:

$$i_1 = I_1 \sin(2\pi f_1 t) \Rightarrow u_F \approx U_T \frac{i_1 i_2}{I_{DC}^2} = U_T \frac{I_1 I_2}{I_{DC}^2} \sin(2\pi f_1 t) \sin(2\pi f_2 t)$$

$$i_2 = I_2 \sin(2\pi f_2 t)$$

$$u_F \approx \frac{I_1 I_2}{2 I_{DC}^2} (\cos([2\pi(f_1 - f_2)]t) - \cos([2\pi(f_1 + f_2)]t))$$

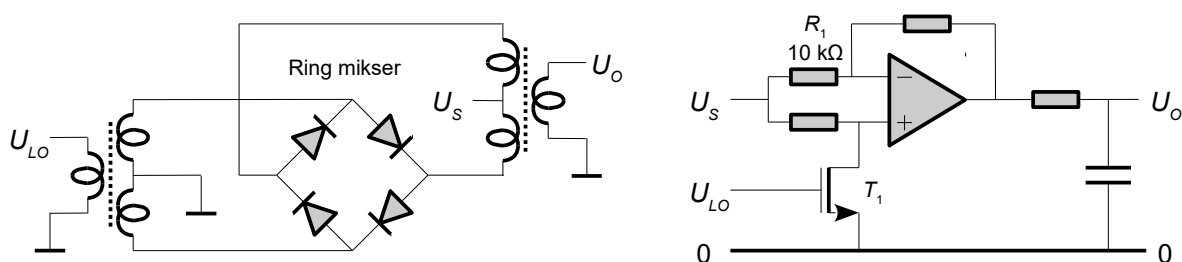
Efter filtrering i MF delen (kun differenssignalet benyttes):

$$u_F = U_T \frac{i_1 i_2}{2 I_{DC}^2}$$

Figur 282 – I en mikser benyttes diodens logaritmiske karakteristisk 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 indskudstab er stort. Som beregningen viser er tabet omkring 34 dB så det er nødvendigt at forstærke det højfrekvente signal ganske meget inden mikseren benyttes.

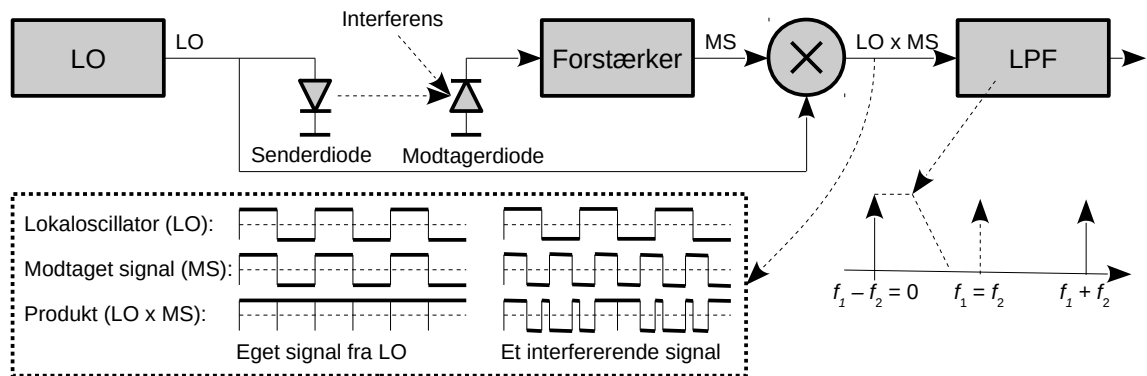


Figur 283 – En dobbeltbalanceret mikser for højfrekvente signaler ved radiomodtagelse (venstre) og en analog mikser med et lavpasfilter for brug i industriel elektronik (højre).

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. Lokaloscillatoren 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} .

5.2.14 Optiske sensorer

For brug i industrielle sensorer benyttes lokaloscillatoren LO til at styre en lysdiode der udsender lypulser. 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 sensors LO frekvenser og kan undertrykkes med et lavpasfilter. Detektion af signalet kan herefter foretages ved at spændingen fra lavpasfiltret sammenlignes med et DC niveau.

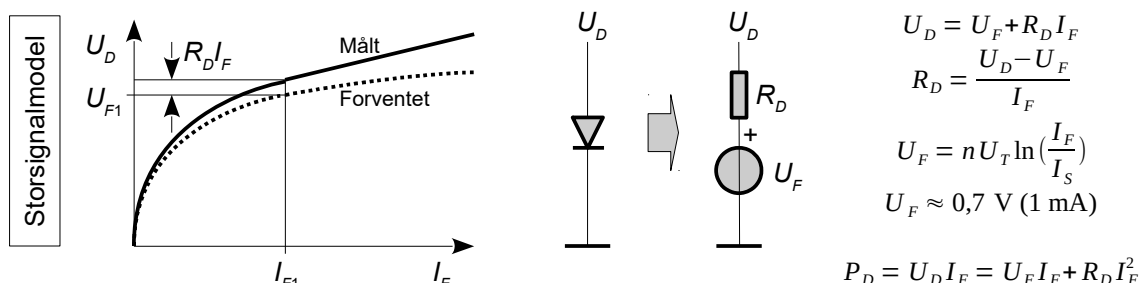


Figur 284 – Den balancerede mikser er meget anvendt for at opnå immunitet overfor forstyrrende interferens. Signalet fra lavpasfiltret LPF er en DC værdi hvis det modtagne signal er lig med lokaloscillatorens signal. Ved en forskel i frekvens vil differenssignalet fra mikseren svinge med frekvensforskellen mellem LO og den interfererende 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.

5.2.15 Storsignal – DC seriemodstand

Halvlederen har en DC modstandsværdi og ved måling på dioden ses det som et spændingsfald i lederetningen 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 effektforsynings ensretter og for dioden i en DC-DC konverter.



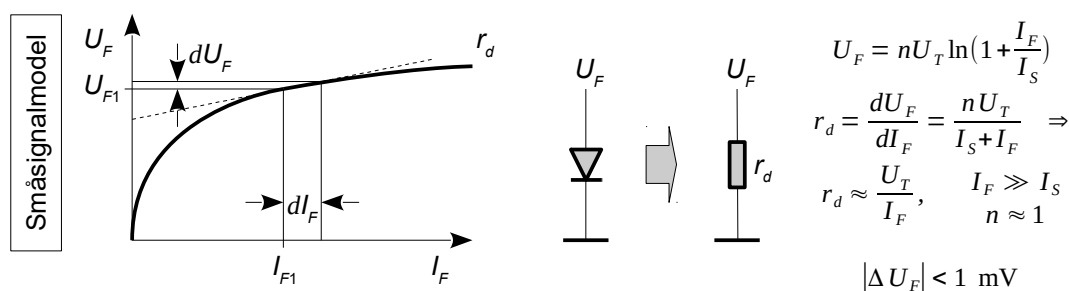
Figur 285 – Diodens seriemodstand skyldes ledningsevnen af halvlederens materiale og en model af dioden kan inkludere en modstand.

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 halvlederens 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.

5.2.16 Småsignal – Dynamisk modstand

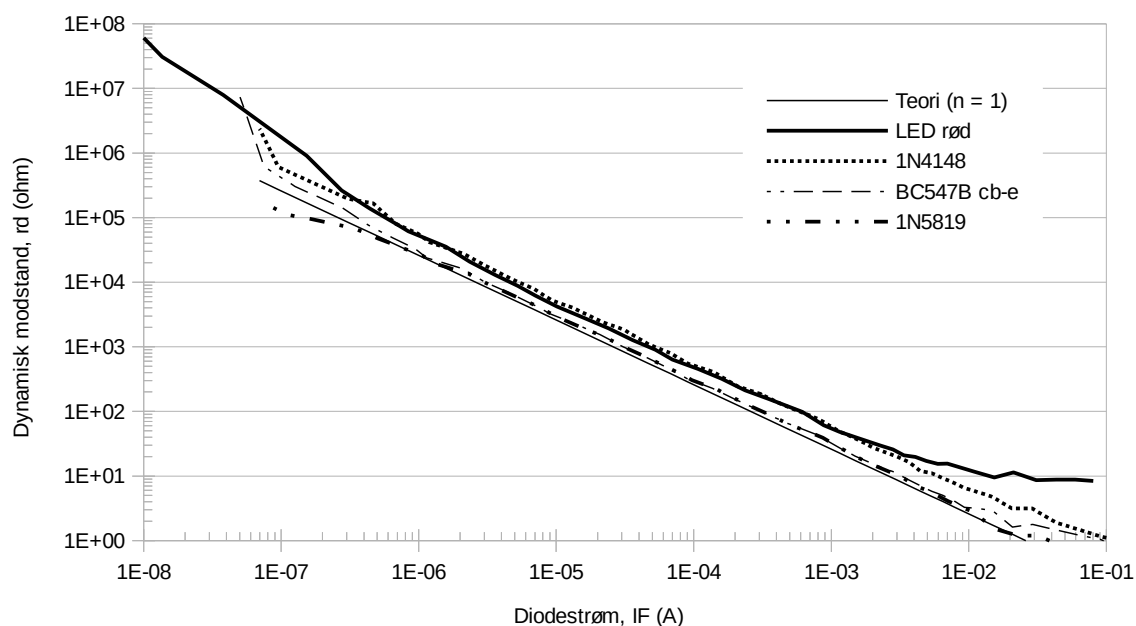
Ved lav strømstyrke hvor halvlederens modstand R_D ikke har en større betydning er spændingen over dioden givet af den logaritmiske funktion. Ved små variationer i strømmen kan dioden ses som en strømstyret modstand der kan varieres fra 10Ω til $300 \text{ k}\Omega$ ved en strømstyrke fra 3 mA til 100 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 263) hvor der i dag vil blive foretrukket anvendelse af et digitalt potentiometer eller en multipliserende DAC.



Figur 286 – Diodens dynamiske 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.

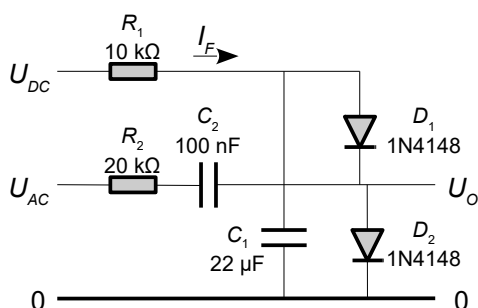


Figur 287 – Målte værdier af dynamisk modstand for en lysdiode (LED), en siliciumdiode (1N4148), en diode-koblet transistor (BC547B) og en Shottky diode (1N5819).

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. Bemærk, at trods de store forskelle i diodernes halvledermateriale (type III-V, silicium og metal-halvleder) er kurverne stort set ens.

5.2.17 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.



Spændingsdeling:

$$U_O = \frac{r_d/2}{r_d/2 + R_2} U_I \quad \rightarrow \quad U_O = \frac{r_d}{2R_2} U_I = \frac{nU_T}{2R_2 I_F} U_I$$

Parametre:

$$r_d = \frac{nU_T}{I_F} \quad \text{og} \quad I_F = \frac{U_{DC} - 2U_F}{R_1} \approx \frac{U_{DC} - 1,2 \text{ V}}{R_1}$$

Grænsefrekvens:

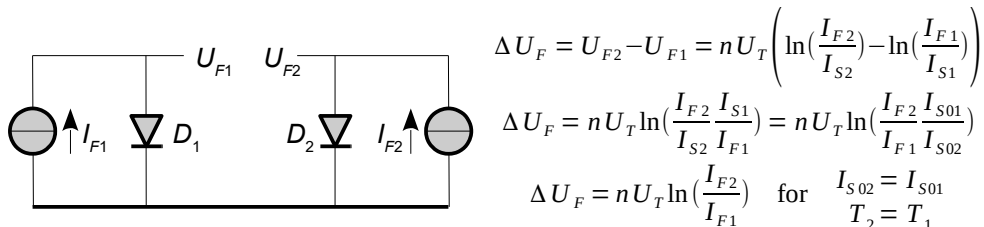
$$f_L \approx \frac{1}{2\pi R_2 C_2} \quad \text{og} \quad C_1 > \frac{1}{4\pi r_{d \min} f_L}$$

Figur 288 – 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 μA , 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.

5.2.18 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 strømstyrke. Det leder til en metode der bruges i de integrerede kredsløb for måling af temperatur internt på chippen og for opbygning af strømspejle (side 311).



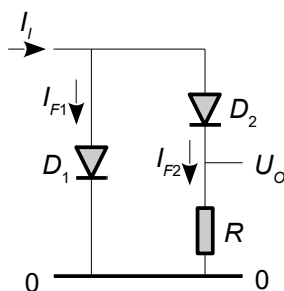
Figur 289 – Differensen mellem spændingen over to dioder er primært givet ved forholdet mellem de to strømstyrker og kun i ringe grad af de øvrige parametre.

Ved et forhold på 2 mellem strømstyrkerne 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 Schottky-dioder.

De to dioder arbejder ved hver sin strømstyrke 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 tilnærmelsesvist 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 strømstyrker og så går mætningsstrømmen ud.

5.2.19 Termometer

En enkel måde at trække to dioders spænding fra hinanden vises herunder hvor spændingen U_O er en funktion af den absolutte temperatur. Konstanten er $nk/q_0 = 86,2 \mu\text{V}/^\circ\text{C}$ for $n = 1$ så spændingen stiger 0,1 mV per grad celsius hvis logaritmen indstilles til 1,16. Det kræver at forholdet mellem strømstyrkerne er $e^{1,16} = 3,19$ så strømmene skal holdes på $I_{F1} = 3,19 I_{F2}$.



Ligninger:

$$U_{F1} = U_{F2} + U_O = 0$$

$$nU_T \ln\left(\frac{I_{F1}}{I_S}\right) = nU_T \ln\left(\frac{I_{F2}}{I_S}\right) + U_O \Rightarrow U_O = \left(\frac{nk}{q_0} \ln\left(\frac{I_{F1}}{I_{F2}}\right)\right) T$$

Design:

$$U_O = (0,1 \text{ mV/K}) \cdot T \quad \text{ved} \quad I_{F1} = 3,19 I_{F2}$$

$$R = \frac{U_O}{I_{F2}} = 3,19 \frac{U_O}{I_{F1}} \quad \text{og} \quad I_I = I_{F1} + I_{F2}$$

Figur 290 – 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.

5.2.20 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 ønsket.¹⁴⁷

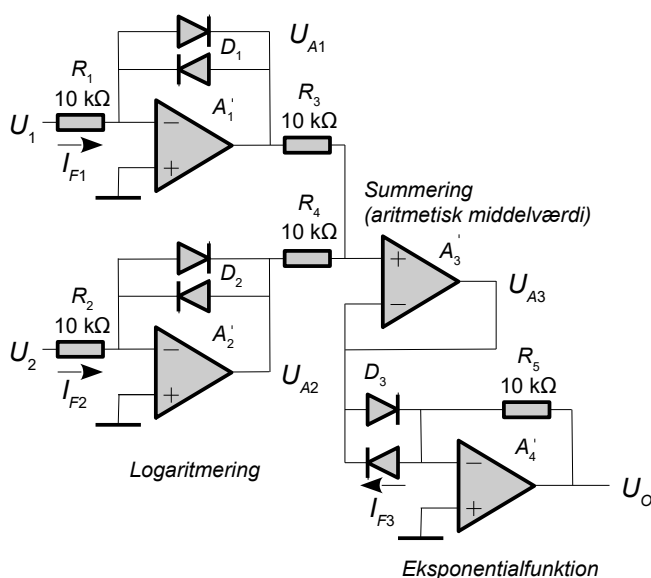
Herunder vises hvordan den geometriske middelværdi af to signaler kan dannes. Den aritmetiske middelværdi af to signaler er let at beregne for det kræver blot to modstande, men det er mere kompliceret at beregne den geometriske middelværdi der er kvadratroden af produktet af to signaler, men det kan 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.

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 negativ opnås division med den negative indgang og ved to negative indgangssignaler vil begge signaler være reciprokke.



$$U_{A1} = -U_T \ln\left(\frac{U_1}{I_S R_1}\right), \quad U_1 > 0$$

$$U_{A2} = -U_T \ln\left(\frac{U_2}{I_S R_2}\right), \quad U_2 > 0$$

$$U_{A3} = \frac{U_{A1} + U_{A2}}{2}$$

$$U_{A3} = -\frac{U_T}{2} \ln\left(\frac{U_1}{I_S R_1} \frac{U_2}{I_S R_2}\right)$$

$$U_{A3} = -U_T \ln \sqrt{\frac{U_1 U_2}{I_S^2 R_1 R_2}}$$

$$I_{F3} = I_S \exp\left(\frac{-U_{A3}}{U_T}\right)$$

$$I_{F3} = I_S \exp\left(\ln \sqrt{\frac{U_1 U_2}{I_S^2 R_1 R_2}}\right) = \sqrt{\frac{U_1 U_2}{R_1 R_2}}$$

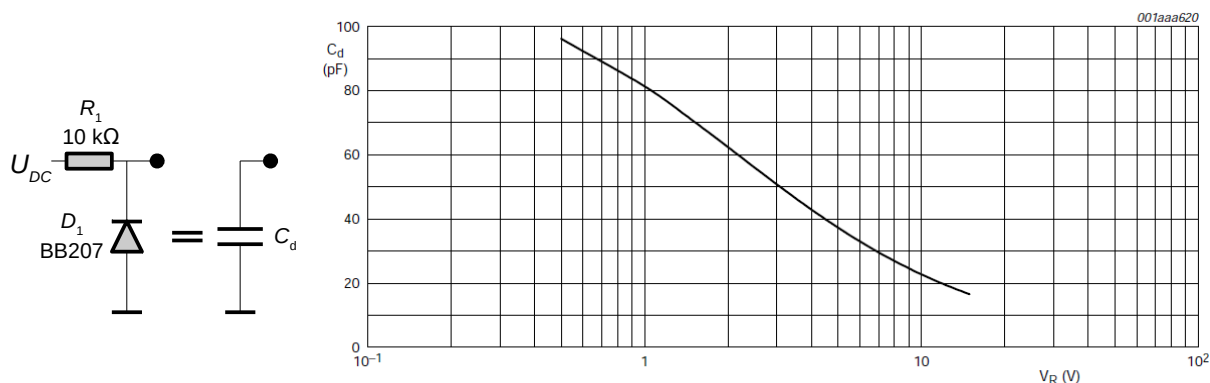
$$U_O = R_5 I_{F3} = R_5 \sqrt{\frac{U_1 U_2}{R_1 R_2}}$$

Figur 291 – 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.

¹⁴⁷ 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. Dioderne skal være meget nær ens så der blev benyttet en integreret kreds med det nødvendige antal dioder integreret på samme chip.

5.2.21 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 betydende og da tykkelsen af PN overgangen er funktion af spændingen over dioden kan kapaciteten ændres ved at ændre spændingen.



Figur 292 – 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.

5.3 Lysdiode

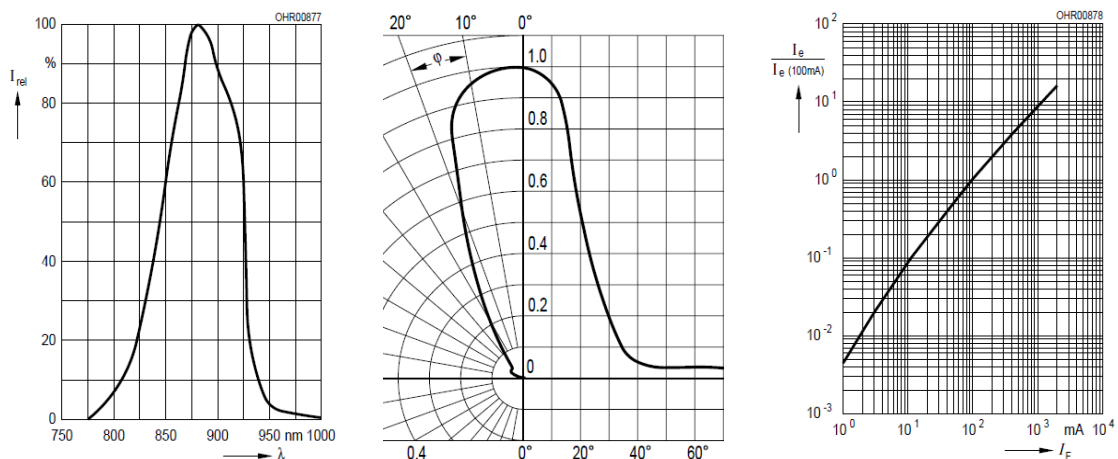
Når strømmen løber i halvlederen vil en elektron accelereres af spændingsfaldet og den kan få nok energi til at slå en elektron fra et af atomerne op til en højere bane omkring kernen. På et eller andet tidspunkt falder elektronen tilbage igen og den frigør da den oplagrede energi i form af en foton der alt efter halvledermaterialet vil lyse 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 (fotodiode).

5.3.1 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 lederetningen for elektronerne skal opnå en vis energi ved acceleration gennem spændingsfaldet, og de violette og ultraviolette lysdioder kræver en relativt høj energi.¹⁴⁸

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 maksimum. 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 μs . Dioden vil gå ON eller OFF indenfor 0,5 μs regnet mellem 10 og 90 % niveauerne.



Figur 293 – Den spektrale fordeling af lysudsendelsen (venstre), retningskarakteristikken (midte), og den relativt udsendte effekttæthed som funktion af strømmen (højre) for SFH485 (Osram).

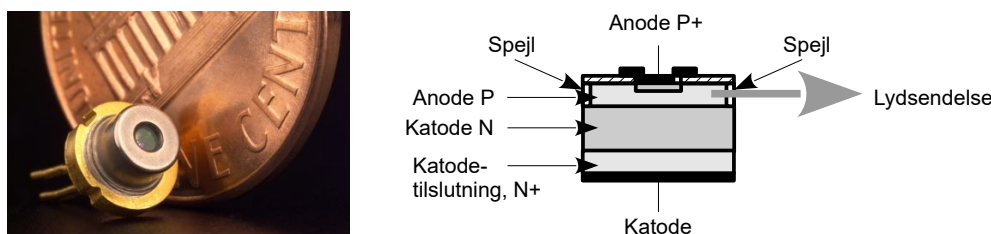
Strålingsintensiteten for SFH485 er 25 ... 160 mW/sr ved 100 mA hvor sr står for steradian¹⁴⁹. Det betyder at dioden belyser en flade med 25 ... 160 mW/m² i 1 m afstand. For omregning til anden afstand benyttes afstandsreglen der angiver en reduktion på kvadratet af forøgelsen af afstanden. Diodens strålingsintensitet 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 effekttætheden fra lysdioden er ikke farlig for øjet i modsætning til den koncentrerede lysstråle fra en laserdioden.

¹⁴⁸http://en.wikipedia.org/wiki/Light-emitting_diode.

¹⁴⁹En steradian er defineret som det vinkelrum, der afgrænser 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 .

5.3.2 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 rubin.¹⁵⁰



Figur 294 – 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 blå 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 multipla 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 takt.¹⁵¹

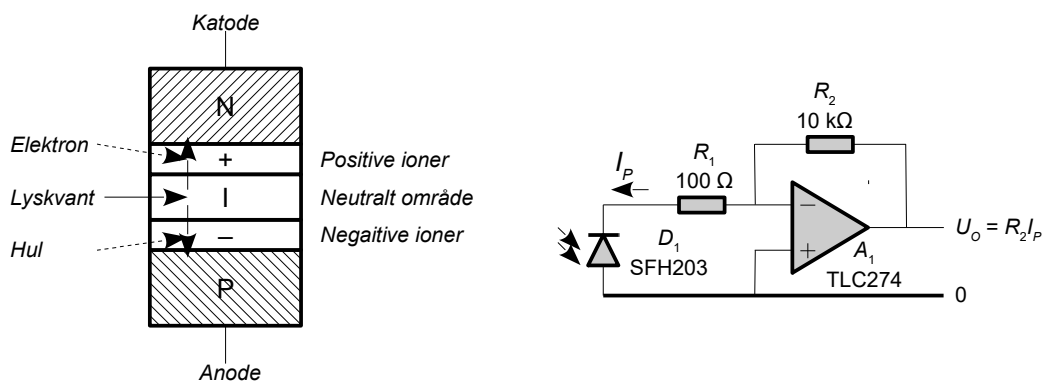
Laseren muliggør optiske systemer hvor lyset deles i to stråler ved endnu et delvist reflekterende spejl og de to stråler kan igen blandes efter at have passeret forskellige emner eller strækninger, og de vil nu interferere. Hvis den ene lysstråle er forsinket et helt antal bølgelængder i forhold til 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 omkring 1 nm.

¹⁵⁰<http://www.nbi.ku.dk/hhh/fysikkens/spinoff/laseren/>.

¹⁵¹http://en.wikipedia.org/wiki/Laser_diode.

5.3.3 Fotodiode

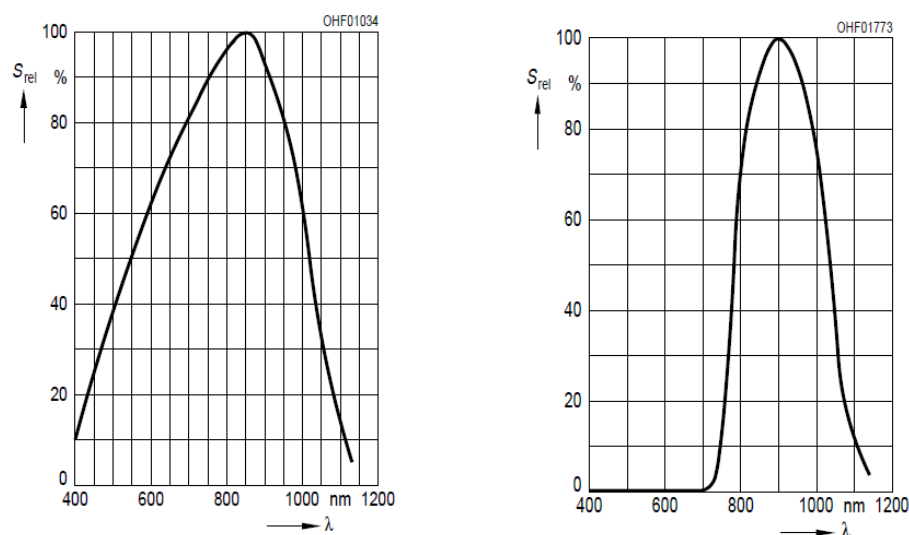
Alle halvledere er lysfølsomme og fotodioden er designet til at omsætte lys til elektrisk strøm. Anvendelsesområderne omfatter datakommunikation, fjernmåling af temperatur og omsætning af solens lys til elektrisk energi ved solpaneler. Kort beskrevet kan et lyskvant slå en elektron fri fra et af siliciumatomerne 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 imod diodens pileretning.



Figur 295 – 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 responsen for fotodioden SFH203.¹⁵²

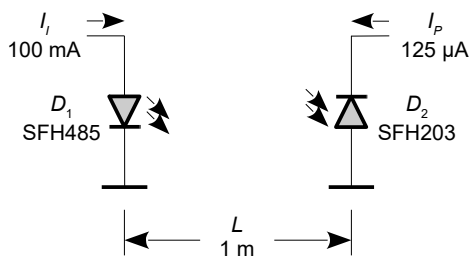


Figur 296 – Den spektrale følsomhed for SFH203 uden daglysfiltre (venstre) og SFH203F med et filter der dæmper den synlige del af spektret (højre). Dioden er mest følsom omkring 850 nm der ligger på grænsen mellem den røde farve og det infrarøde spektrum (Osram).

¹⁵²Ved 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 10^5 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 varmestråling i det infrarøde område.

Dioden SFH203 vil producere en strøm på $I_p = 50 \mu\text{A}$ ved energitætheden 10 W/m^2 og 950 nm . Den maksimale følsomhed er ved 850 nm hvor dioden kan forventes at give mest strøm så for et korrekt resultat skal spektret for IR-lysdioden ganges med den spektrale følsomhed for fotodioden og integreres over det fælles bølglængdeområde. Mindre kan dog gøre det hvis man kan leve med en betydelig usikkerhed for beregningen.

Den tidligere viste IR-lysdiode 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 belyningsniveauet 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.



IR-lysdiode SFH485:

$E_{IR} = 25 \text{ mW/sr}$ i akseretningen, 880 nm

$E_{IR} = 25 \text{ mW/m}^2$ i 1 m afstand, 880 nm

Fotodiode SFH203:

$I_p \approx 50 \mu\text{A}$ @ 10 mW/m^2 , 950 nm

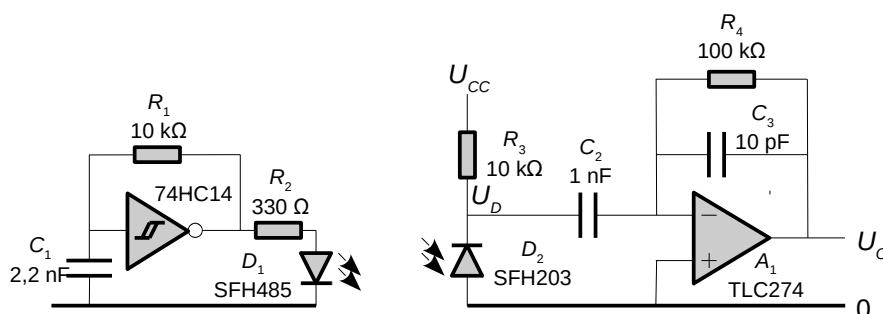
Figur 297 – Lyset fra den infrarøde senderdiode SFH485 modtages af fotodioden SFH203 i en afstand på 1 m .

Effekttætheden specificeres i steradianer så den reduceres med kvadratet på en ændring i afstanden. Det betyder at en fordobling af afstanden reducerer effekttætheden til en fjerdedel.

5.3.4 Lysdetektor

Et populært kredsløb indenfor den industrielle elektronik fjerner indvirkningen fra omgivelserne ved at benytte en pulset IR-lysdiode. Ideen er at omgivelsernes lys normalt varierer så langsomt at ændringerne kan undertrykkes med et højpasfilter som kun tillader passage af de korte pulser. Det omgivende lys kaldes også for DC lys.

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.



Spændinger:

$$U_D = U_{CC} - R_1 I_{P,DC}$$

$$U_O = R_2 I_{P,AC}$$

Filterkredsløb:

$$f_1 = \frac{1}{2\pi R_1 C_1} < f_{IR}$$

Lavpasled:

$$f_2 = \frac{1}{2\pi R_2 C_2} > f_{IR}$$

Figur 298 – Lysdetektor hvor de modtagne lyspulser 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 belyningsniveau 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 oscillatorens 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 oscillatorens frekvens. Dertil har C_3 som opgave at stabiliserer operationsforstærkeren på grund af fotodiodens indre kapacitet.

Ved et praktisk design bør man være opmærksom på at støj fra effektforsyningens U_{CC} kobles direkte til den inverterende indgang på operationsforstærkeren gennem R_3 og forstærkes med forholdet R_4/R_3 . Det er som regel nødvendigt at afkoble effektforsyningen for at holde højfrekvent støj nede i niveau. Indledningsvis bør man montere en kondensator på 47 ... 220 nF mellem U_{CC} og nul og hvis det ikke er nok kan man supplere med en 470 μ F elektrolytkondensator og måske også indskyde 10 Ω i serie med U_{CC} for at lavpasfiltrere støjen fra effektforsyningen.

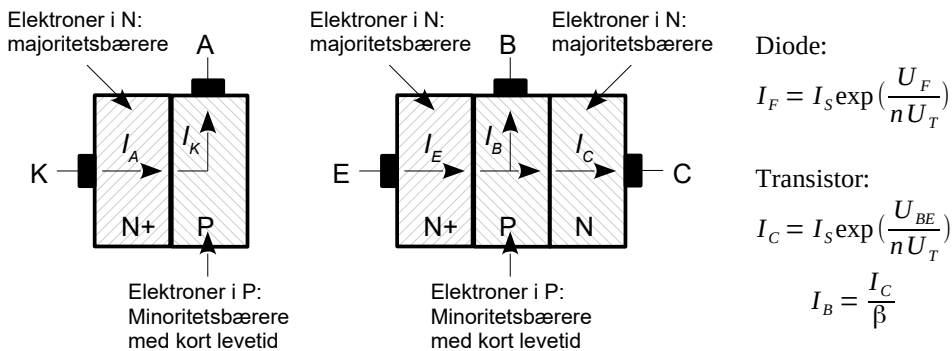
Ved industrielt brug er der ofte flere sensorer i brug samtidigt. Det kan for eksempel være udstyr for en vaskehal, en garagedør eller et transportbånd. For at adskille de pulsedede signaler fra de forskellige sensorer benyttes individuelle frekvenser for hver sensor og der kan diskrimineres mellem de enkelte sensors ved en mikser og et lavpasfilter som korrelerer det modtagne signal med oscillatoren der pulser IR-lysdioden (se side 261).

5.4 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 desuden i en JFET hvor den gennem et elektrisk felt i den spærreforspændte diode styrer strømmen i transistorens kanal. I en MOSFET beskriver diodens relation funktionen ved svag strømstyrke (*sub threshold*).

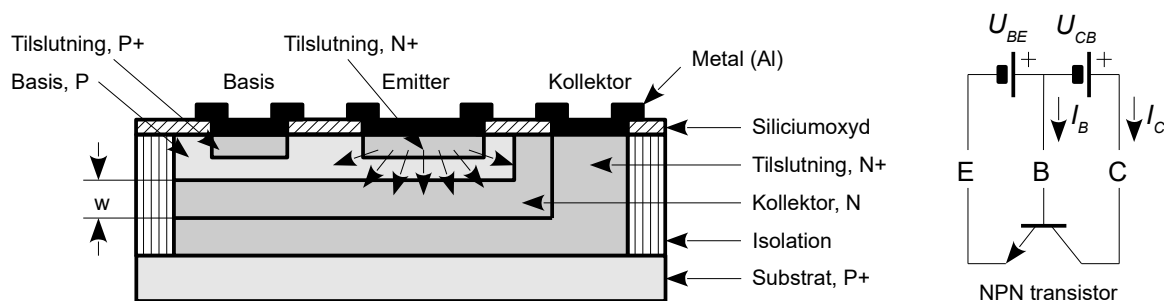
5.4.1 Den bipolære transistor

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 299 – 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.

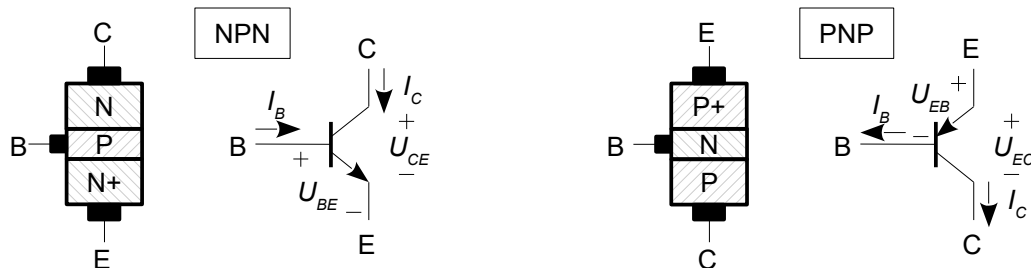
I en transistor udvides dioden i basis-emitter med endnu en N halvleder som danner kollektoren. Dioden opfører sig som en almindelig PN overgang hvor en ekstern spænding over basis-emitter styrer diffusionen af elektroner fra emitter til basis, men basislaget er væsentligt tyndere end indtrængningsdybden 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 300 – 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).

5.4.2 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 analoge integrerede kredse og diskret opbyggede effektforstærkere.



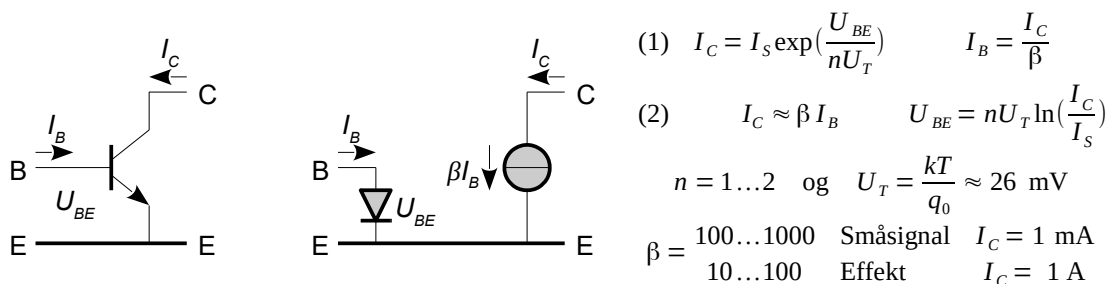
Figur 301 – To typer bipolære transistorer. Pilen i symbolet angiver strømretningen (modsat elektronernes bevægelsesretning) og viser samtidigt den elektrode der kaldes emitter.

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 princippet mindre egnet til højfrekvens for diffusionsstrømmen bæres af huller der har en op til fire gange lavere mobilitet end elektroner så ladningstransporten foregår lidt langsommere.

5.4.3 DC model

Der er to populære modeller for den bipolære transistor.

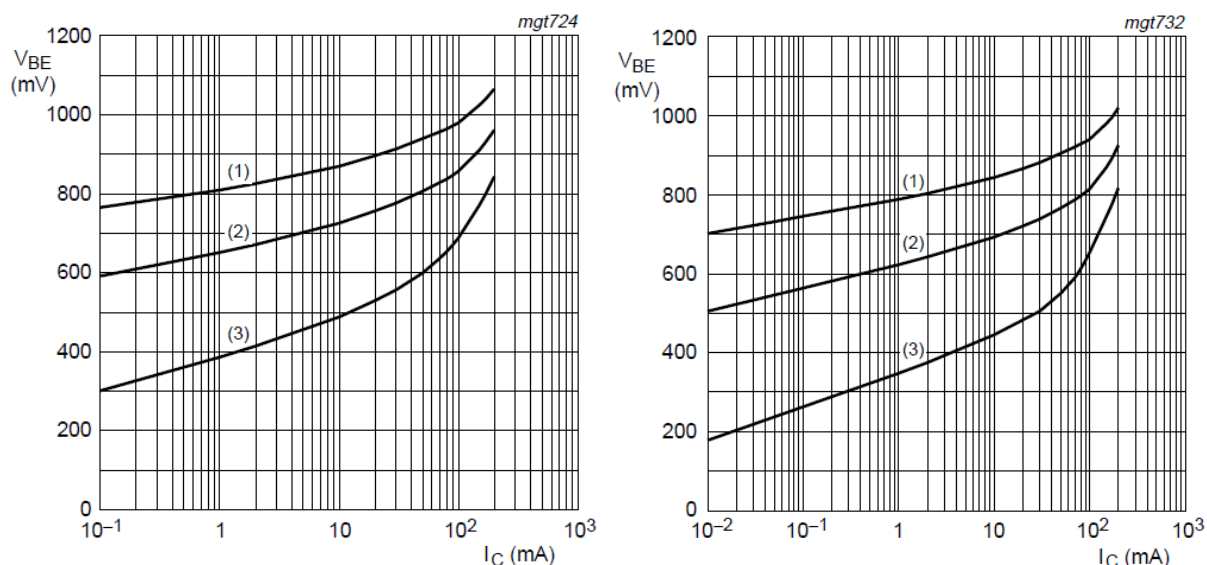
Model 1 har strømmen I_C i transistorens kollektor-emitter som funktion af den styrende spænding U_{BE} over dioden i basis-emitter og kun emitteren er den fælles terminal. Strømmen i basis I_B fra rekombinationen i det tynde basislag betyder at for hver elektron der rekombineres med et hul er der β gange så mange elektroner som når frem til kollektor så basisstrømmen er relativt lav.



Figur 302 – En model for den bipolære transistor med tilhørende matematisk fortolkning som en spændingsstyret strømkilde (1) eller en strømstyret strømkilde (2).

Model 2 har strømmen I_C i kollektor-emitter som funktion af den styrende strøm I_B i basis-emitter med transistorens strømforstærkning β som proportionaliteten mellem de to. Spændingen over basis-emitter er funktion af strømniveauet så tommelfingerreglen om 0,7 V over basis-emitter er kun en tilnærmelse. Modellen er velegnet til en DC analyse af et kredsløb som det skal vises på side 276.

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 303 – 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 elektrontætheden 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).

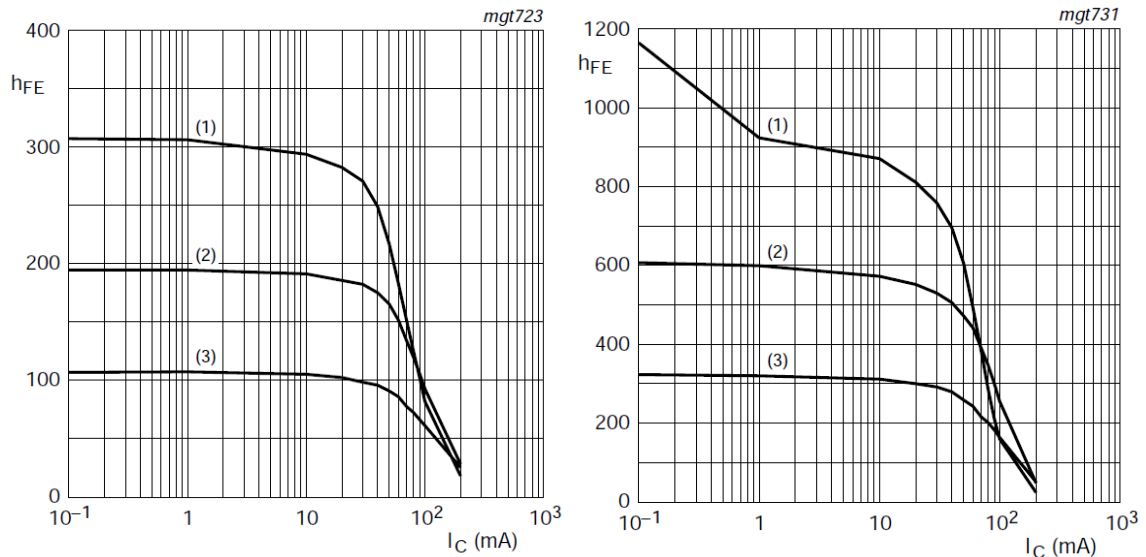
Mætningsstrømmen I_S (saturation current) afhænger af PN overgangens geometri og doteringsprofil og strømmen fordobles for hver 10°C temperaturstigning. Værdien er cirka 10 fA for BC547 ved 25°C , men det er ikke en værdi man kan finde i databladet. Mætningsstrømmen optræder som faktor til den eksponentielle 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 $-2,5 \text{ mV}/^{\circ}\text{C}$.

Transistorens 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 kredsløbet lukker ned ved så hård belastning at chip-temperaturen kommer over 150°C og bringer en transistor til at trække strøm ud af styringskredsløbet.

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.

Transistorens strømforstærkning 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 begrænses variationen til 1:2. Strømforstærkningen afhænger af arbejds punktet og temperaturen som illustrationen herunder viser. Det er en tilnærmelse at beskrive den som konstant, men det passer nogenlunde fra cirka 10 μA til 1 mA.

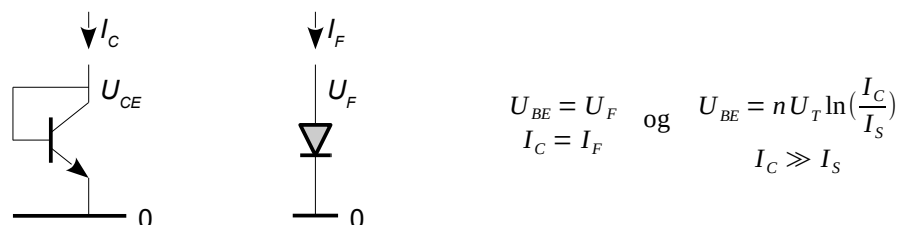


Figur 304 – Strømforstærkningen for BC547 og BC847 i A selektionen til venstre og C selektionen til højre. Tallene refererer til temperaturen med (1) for 150°C, (2) for 25°C og (3) for -55°C (NXP). Det viste knæk under 1 mA for BC547C er ikke kommenteret i databladet og har ingen fysisk årsag.

Temperaturkoefficienten af strømforstærkningen ses at være af størrelsesordenen 1 %/°C så den sædvanlige antagelse om en konstant værdi af strømforstærkningen holder ikke når temperaturen ændres og man kan være nødt til at tage hensyn til det.

5.4.4 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 311.

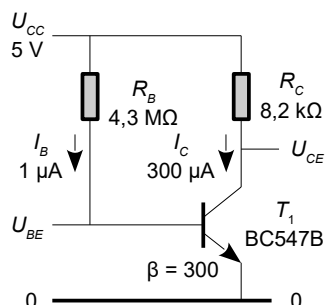


Figur 305 – En diode-koblet transistor opfører sig fuldstændig som en diode.

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 transistorens 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.

5.4.5 DC arbejds punkt

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.



Komponentdesign:

$$R_C = \frac{U_{CC} - U_{CE}}{I_C} \Rightarrow U_{CE} = \frac{U_{CC}}{2} \Rightarrow R_C \approx \frac{U_{CC}}{2I_C}$$

$$R_B = \frac{U_{CC} - U_{BE}}{I_B} \Rightarrow U_{BE} \approx 0 \Rightarrow R_B \approx \beta \frac{U_{CC}}{I_C}$$

DC arbejds punkt:

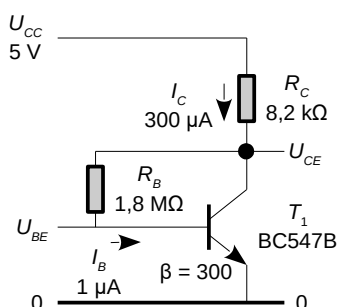
$$U_{CE} = U_{CC} - R_C I_C \rightarrow \begin{cases} U_{CC} & \text{for } \beta \rightarrow 0 \text{ Afbrudt} \\ 0 & \text{for } \beta \rightarrow \infty \text{ Mættet} \end{cases}$$

Figur 306 – 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 sjældent 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 $3,4 \text{ V}$ og 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 277).

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 arbejds punkt 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.



Komponentdesign:

$$R_C = \frac{U_{CC} - U_{CE}}{I_C} \Rightarrow U_{CE} = \frac{U_{CC}}{2} \Rightarrow R_C \approx \frac{U_{CC}}{2I_C} \Rightarrow R_C \approx \beta R_C$$

$$R_B = \frac{U_{CE} - U_{BE}}{I_B} \Rightarrow U_{BE} \approx 0 \Rightarrow R_B \approx \beta \frac{U_{CC}}{2R_B}$$

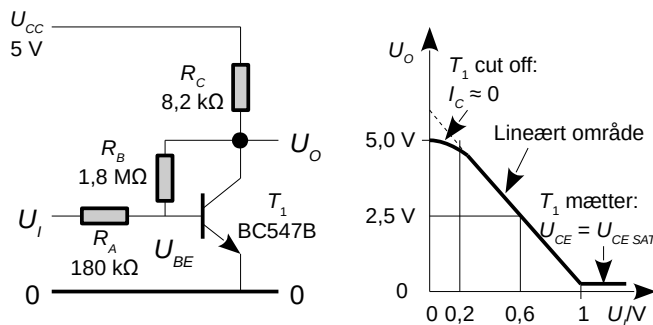
DC arbejds punkt:

$$U_{CE} = U_{CC} - R_C I_C \rightarrow \begin{cases} U_{CC} & \text{for } \beta \rightarrow 0 \text{ Afbrudt} \\ U_{BE} & \text{for } \beta \rightarrow \infty \text{ Lineær} \end{cases}$$

Figur 307 – En kobling hvor strømforstærkningen har mindre betydning for DC arbejds punktet som følge af tilbagekoblingen fra kollektor til basis. Forstærkeren er fælles emitter (*common emitter*) fordi emitter er fælles for ind- og udgang.

5.4.6 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.



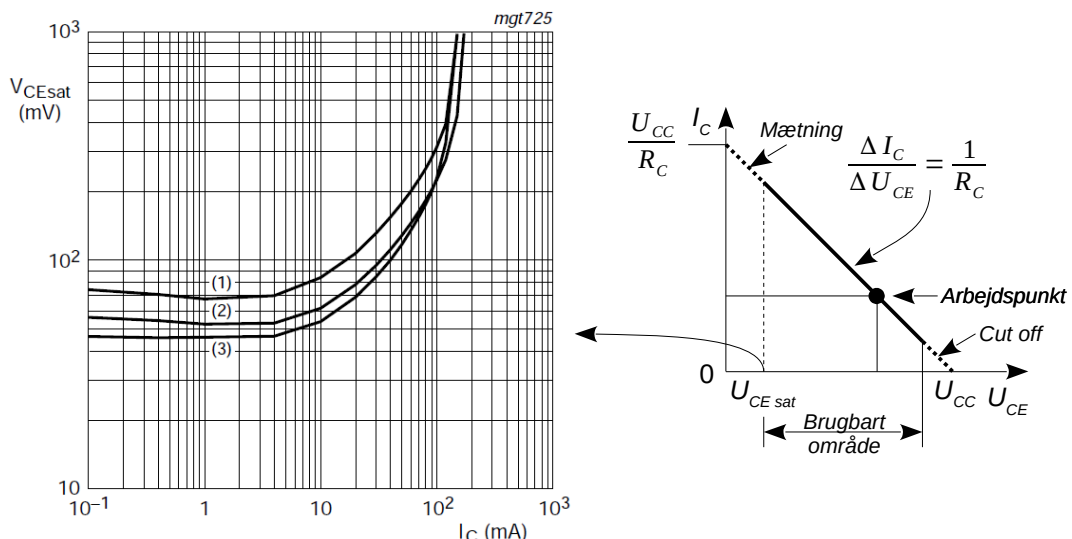
Figur 308 – En transistor kan forstærke et DC indgangssignal.

Udtrykket er ikke blandt de simpleste, men med komponenternes værdier indsat 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 formelen 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 286.

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 potentiale er tæt på emitterens.

5.4.7 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 309 – 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).

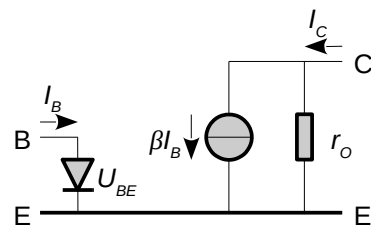
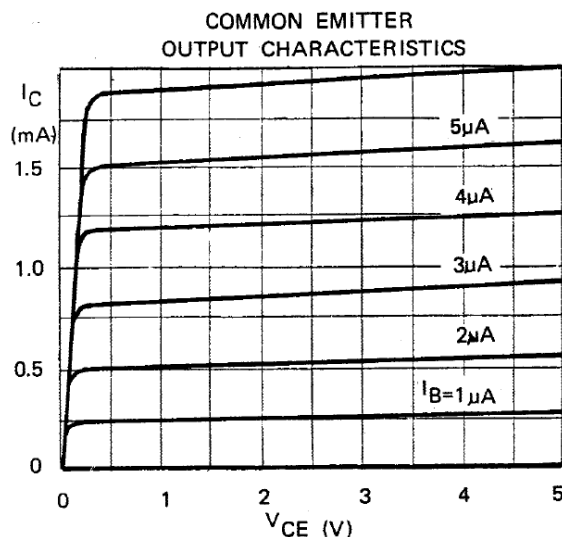
Kurven viser at transistorens aktive område ophører ved en kollektor-emitter spænding på 50 mV ved 1 mA og 25°C idet spændingen ikke kan blive lavere. For en strøm over 10 mA stiger den med strømmen og en mættet BC547 kan derfor beskrives ved en spændingskilde på 50 mV i serie med en modstand på 5 Ω . 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 langsomt.¹⁵³

Når transistoren ikke er mættet kan den arbejde med en spænding på kollektor der er lavere 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 med bipolære transistorer og en effektforsyning i nærheden af 1 V.

5.4.8 Udgangskaraktistik

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 250 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 310 – Udgangskaraktistikken for BC107 familien¹⁵⁴ viser strømmen i kollektor som funktion af spændingen over kollektor-emitter med strømmen i basis som parameter (Micro Electronics).

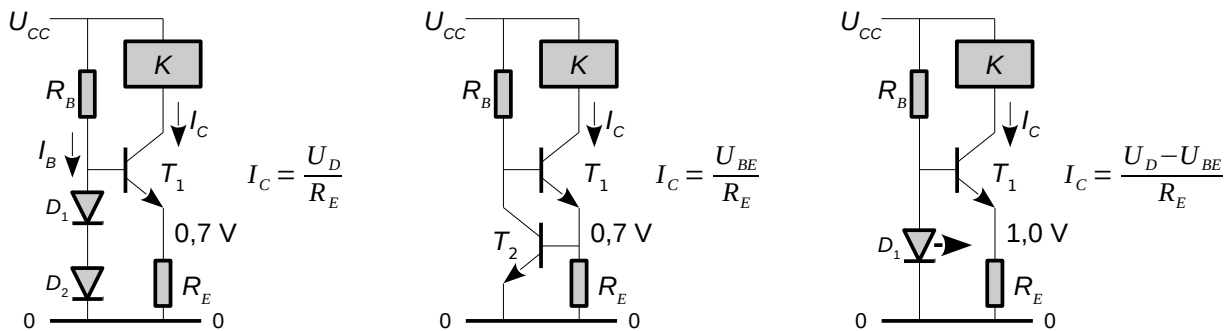
Når transistoren er aktiv vil strømmen i kollektor kun afhængig ganske lidt af spændingen over kollektor-emitter og det kan tolkes som om der var en modstand i parallel med transistorens udgang. I de fleste situationer er modstandens værdi så høj at man kan se bort fra den og blot antage at transistoren kan beskrives analytisk som en ideel strømgenerator.

¹⁵³Den ældre digitalteknik benyttede både afbrudt og mættet som de to binære tilstande med den omkostning at kredsløbene var langsomme 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 Schottky 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 transistorerne 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 bipolære. Hastigheden kan nå op i GHz området.

¹⁵⁴Databladet angives at passe for BC107-BC109, BC167-BC169, BC237-BC239 og BC317-BC319 og som fodnoten side 300 viser kan data forventes at passe med BC547-BC549.

5.4.9 Strømregulering

Udgangskaracteristikken viser at strømmen i transistorens kollektor er næsten uafhængig af den spænding der ligger over kollektor og emitter når spændingsfaldet er større end cirka 0,5 V. Det kan udnyttes til at bygge et meget anvendeligt kredsløb som sikrer at der løber en konstant strøm i en anden del af et kredsløb som herunder kaldes for K . Det kan være alt fra en modstand eller en lysdiode til et større elektronisk kredsløb.



Figur 311 – Populære typer af strømregulatorer.

Et lignende kredsløb kan naturligvis bygges med PNP transistorer hvor strømregulatoren sidder i "toppen" af kredsløbet og kredsløbet K sidder nær ved nul eller den negative forsyning.

Det første kredsløb benytter to siliciumdioder som spændingsreference. Som vist tidligere vil der ligge et spændingsfald over en diode på cirka 0,7 V ved et vist strømniveau så basis af transistor T_1 vil ligge 1,4 V over nul og transistorens emitter vil ligge 0,7 V over nul. Strømmen i transistoren er da fastsat af R_E og givet ved Ohms lov. For at virke som forventet skal strømmen i modstand R_B være tilstrækkelig til at forsyne transistoren med basisstrøm og dioderne med strøm. Diode D_1 vil have nogenlunde samme temperaturkoefficient som transistoren på cirka $-2 \text{ mV}/^\circ\text{C}$ så diode D_2 giver en resulterende temperaturkoefficient for regulatoren på $-0,3 \text{ }^\circ\text{C}$.

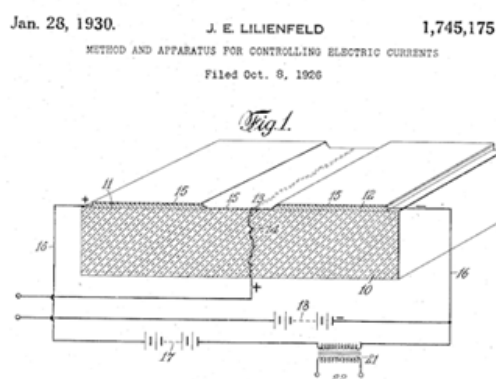
Det midterste kredsløb erstatter de to dioder med en transistor. Det er en overordentlig populær opstilling i specielt effektforstærkere samt i kredsløb hvor der skal indføres et maksimum for den strøm T_1 må trække. Det gælder både effektforstærkere og spændingsregulatorer som 7805. Her er det transistorens spændingsfald over basis-emitter der benyttes som reference.

Det tredje kredsløb benytter en lysdiode som reference. Spændingen ligger i området fra 1,7 V ved røde og grønne lysdioder til 3 V for blå og hvide lysdioder. Diodens lys kan bruges som en praktisk indikator for at kredsløbet er tændt. Det anføres ofte at opstillingen er stabil overfor en variation i temperaturen, men det er ikke rigtigt. Lysdiodens spænding varierer fra -3 til $-5 \text{ mV}/^\circ\text{C}$ hvor den aktuelle værdi må findes i lysdiodens datablad.

Ved brug af en strømregulator skal man være opmærksom på den effekt der afsættes i T_1 som beregnes ved strømmen i transistoren gange med spændingen over den. For en BC547 må der højst afsættes 0,3 W, men jeg vil ikke være tryk ved at gå over 0,1 W for at holde temperaturen på et rimeligt niveau. Ved 10 mA er spændingsfaldet over transistoren derfor begrænset til 10 V.

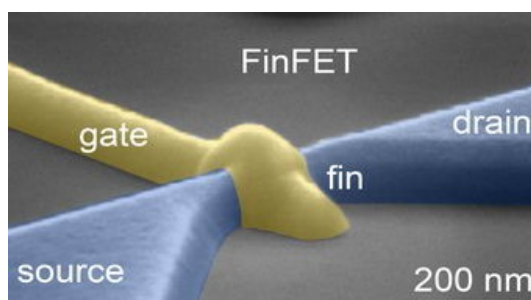
5.4.10 Felteffekt transistoren

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.^{155,156}



Figur 312 – 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 313 – I en FinFET er dimensionerne så små at kanalens modstandsværdi 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 patentet.¹⁵⁷

"[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 bipolære transistor. Den følgende gennemgang vil derfor fokusere på det der er specielt ved "fætteren".

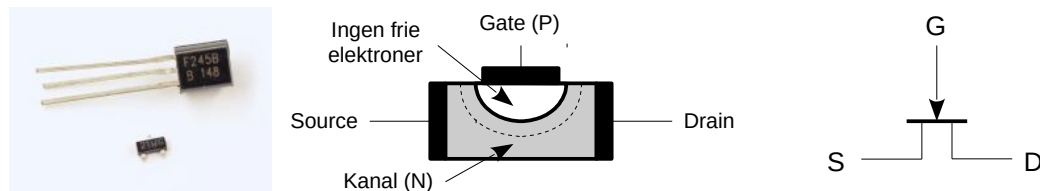
¹⁵⁵http://ethw.org/A_Very_Early_Conception_of_a_Solid_State_Device.

¹⁵⁶www.ecs.umass.edu/ece/ece344/MOSFET_talk.ppt.

¹⁵⁷<http://www.edn.com/electronics-blogs/edn-moments/4422371/Lilienfeld-patents-field-effect-transistor>.

5.4.11 JFET

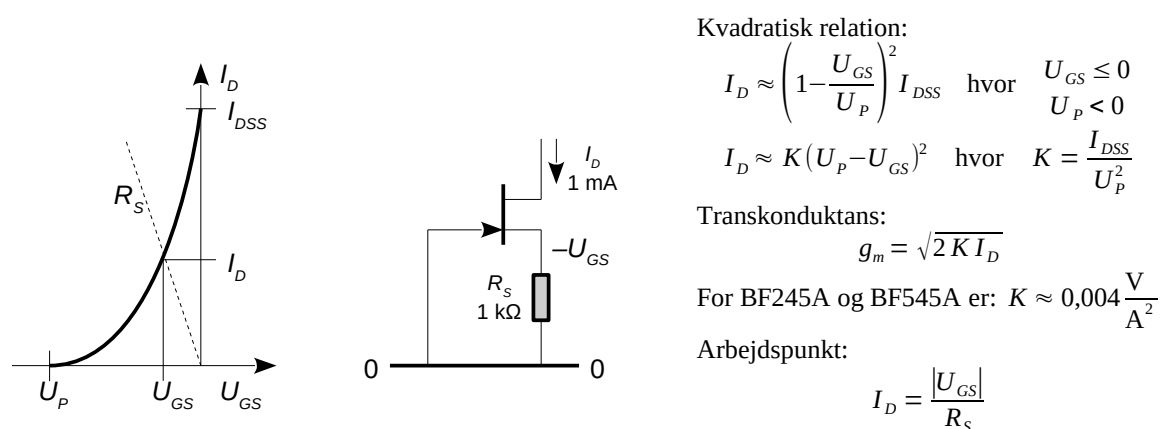
Den tidligste felfeffekt transistor var en JFET der står for *junction field-effect transistor*. Den har tre elektroder der benævnes *gate* for styreelektroden, *source* for strømkilden og *drain* for opsamling af elektronerne. Kanalen mellem source og drain er i mange tilfælde symmetrisk og de to elektroder kan da ombyttes. Hvis det ikke er tilfældet tegnes pilen nærmest ved source.



Figur 314 – En junction felfeffekt 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. Ved en N-kanal bæres strømmen af elektroner i ledningsbåndet og ved en P-kanal bæres strømmen af huller i valensbåndet.

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 forspændt betyder det at gate skal være negativ i forhold til source og drain for at skruer 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.



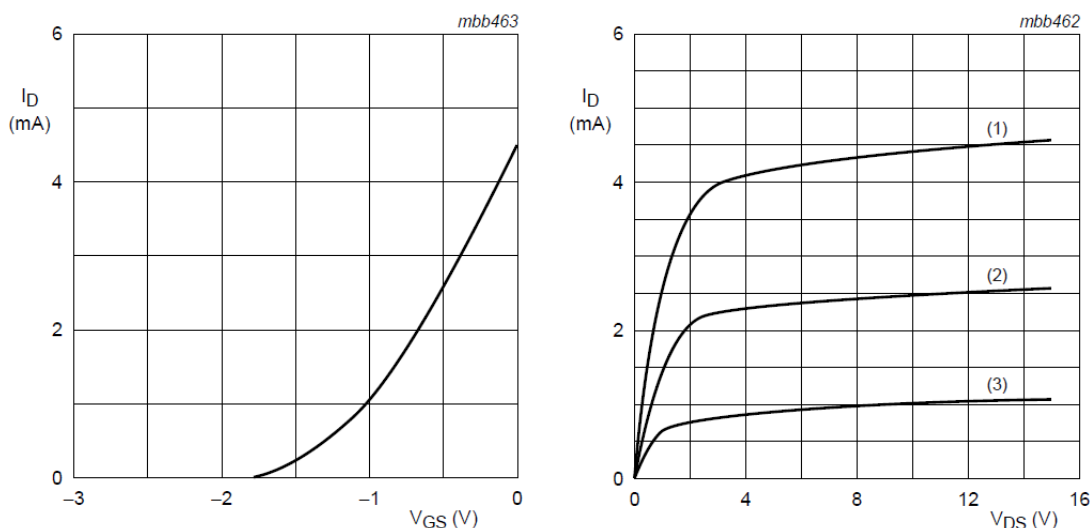
Figur 315 – En JFET anvendt som en strømkilde med strømmen indstillet af en modstand.

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 kanalen. Opstillingen kan benyttes som strømgenerator og den danner grundlag for opbygning af en AC forstærker som vist på side 301. Opstillingen er ækvivalent med en forstærker bygget med radorør (se side 336).

Strømmen styres af det elektriske felt over den spærrede diode så der en kapacitet mellem gate og kanal på omkring 5 pF for en småsignaltransistor. Dioden er fremstillet af silicium så den vil opføre sig som enhver anden siliciumdioder med cirka 0,6 V i lederetningen. Det er dog usædvanligt at benytte dioden i lederetningen 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 spærreretningen som ved BF245 og BF545 er under 10 pA.

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*) ved $U_{GS} = 0$.

Herunder vises karakteristikken for en JFET eksemplificeret ved BF545. Parameteren pinch-off U_P (hvor $I_D = 0$) varierer fra $-0,4$ til $-7,8$ V og I_{DSS} (I_D for $U_{GS} = 0$) varierer fra 2 til 25 mA. Transistoren leveres i tre selektioner hvor der er lidt bedre kontrol over parametrene. Her vises A-selektionen.



Figur 316 – Udgangskaraktistik for BF545A. (1) er $U_{GS} = 0$ V, (2) er $U_{GS} = -0,5$ V og (3) er $U_{GS} = -1$ V.

Udgangskaraktistikken består af strømmen i drain I_D som funktion af drain-source spændingen U_{DS} med gate-source spændingen U_{GS} som parameter. Ved nul volt over gate-source vokser strømmen i drain omtrent lineært med spændingen over drain og source indtil et par volt hvorover strømmen kun varierer lidt. En JFET kan derfor betragtes som en spændingsstyret modstand ved lav spænding over kanalen mens den fungerer som en strømregulator ved højere spænding.¹⁵⁸

Man kan få en idé om kanalens modstandsværdi ved lav værdi af U_{DS} fra karakteristikken ved beregning af hældningen $R_D = U_{DS}/I_D$. For $U_{GS} = 0$ er $R_D = (1 \text{ V})/(2,3 \text{ mA}) = 430 \Omega$ og ved $U_{GS} = -1$ V er $R_D = 1,5 \text{ k}\Omega$.

Felteffekt transistorens funktion er ikke baseret på rekombination af elektroner og huller i kanalen som ved den bipolære transistor så en JFET er typisk specificeret til anvendelse op til 1 GHz og den har desuden gode støjegenskaber som gør den velegnet til brug i radiokredsløb for VHF og UHF båndene.^{159,160,161,162}

I integrerede kredse benyttes en JFET af og til med kortsluttet gate og source som en strømkilde hvor strømstyrken beslutes på designtidspunktet. Transistoren er populær i indgangstrinnet på operationsforstærkere hvor den kan tilbyde særdeles lav indgangsstrøm (*input bias current*). Den almindelige JFET er dog langsomt på vej ud som følge af den succes der tilfalder den moderne MOSFET ved både analog og digital signalbehandling.¹⁶³

¹⁵⁸Jeg vil anbefale at U_{DS} holdes indenfor ± 20 mV hvis transistoren skal fungere som modstand i analog signalbehandling.

¹⁵⁹BF245 er specificeret til -3 dB ved 700 MHz.

¹⁶⁰BF545 er specificeret til en forstærkningsreduktion på 10 % ved 450 MHz.

¹⁶¹BF245 er specificeret til et støjtal på 1,4 dB op til 100 MHz.

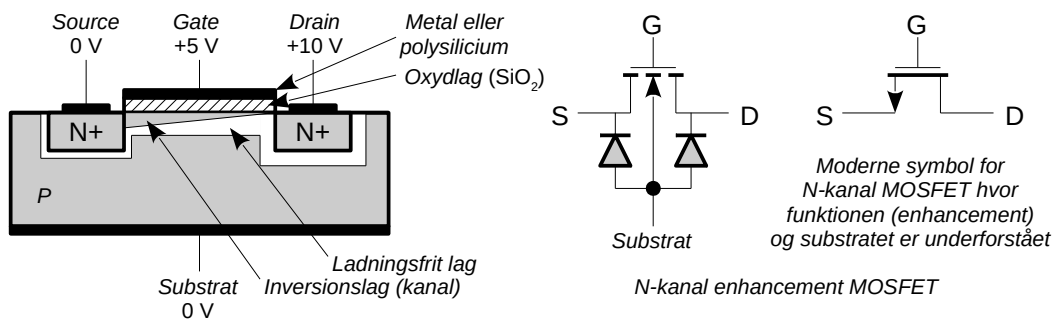
¹⁶²Se https://en.wikipedia.org/wiki/Radio_spectrum for en oversigt over frekvensbåndene.

¹⁶³BF245 (2N3819) fra omkring 1970 udgår i TO-92 på grund af vigende afsætning, men den produceres fremover som BF545 for overflademontage.

5.4.12 MOSFET

Den moderne udgave af feltereffekt 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 $\pm 20\text{ V}$. Elektroden 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 noget over 5 V mellem gate og source.



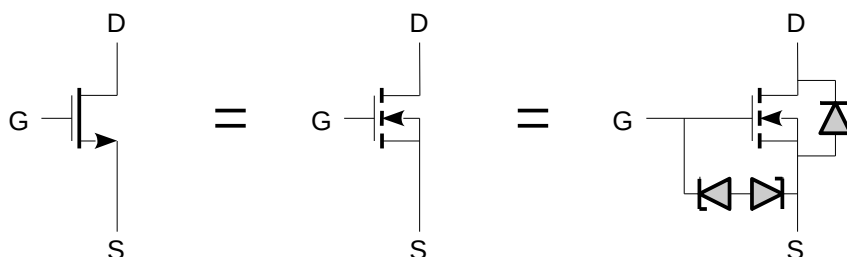
Figur 317 – En N-kanal enhancement MOSFET skaber kanalen fra drain til source ved en positiv spænding på gate i forhold til drain gennem forskydning af ladningen i et tyndt område lige under elektroden for gate.

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 spærende 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 nær 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 forbindes 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 318 – 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 svinge 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.

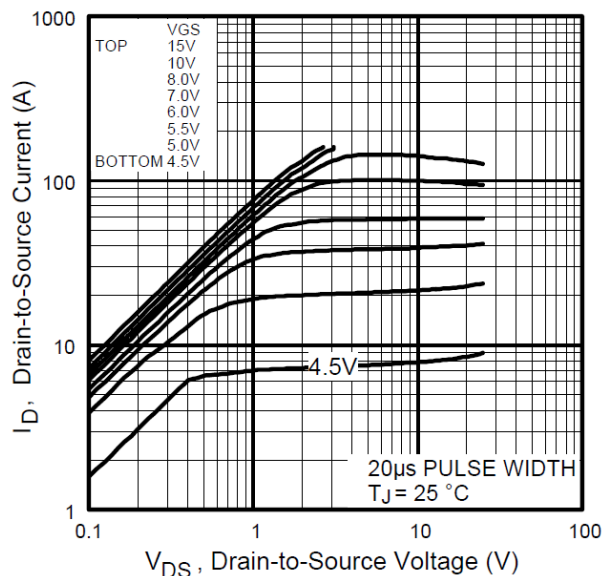
Der er en betydende kapacitet mellem gate og kanal som følge af at styringen benytter et elektrisk felt skabt mellem to adskilte ladninger. For en effekttransistor kan værdien let nå op på 1 nF og det stiller krav til driverkredsløbet hvis transistoren skal skifte hurtigt mellem ON og OFF.

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 transistorer.¹⁶⁴

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 effekttabet.

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 bipolære 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.

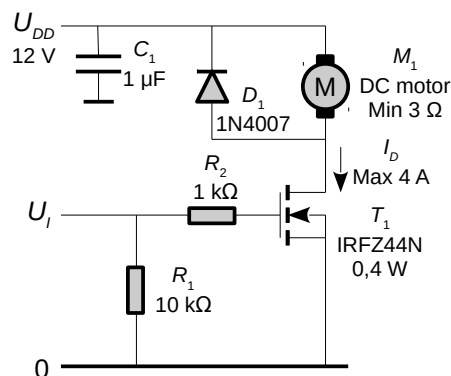
¹⁶⁴Se også: <http://en.wikipedia.org/wiki/MOSFET>.



Figur 319 – Udgangskaraktistikken for IRFZ44N specificeret til $U_{GS(th)} = 2 \dots 4 \text{ V}$ for $I_D = 0,25 \text{ mA}$.

Med $U_{GS} = 5 \text{ V}$ er strømmen I_D proportional med U_{DS} for en spænding indtil 0,5 V. Strømmen er $I_D = 15 \text{ A}$ ved $U_{DS} = 0,5 \text{ V}$ så modstandsværdien bliver $R_{DS} = U_{DS}/I_D = 0,03 \Omega$. Ved $U_{DS} > 2 \text{ V}$ er transkonduktansen på $g_m = \Delta I_D / \Delta U_{GS} = 34 \text{ S}$ 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 $g_m = 26 \text{ S}$ omkring 13 A hvor I_D ændres fra 7 til 20 A for U_{GS} fra 4,5 til 5 V.

En MOSFET anvendes ofte for styring af store belastninger som en mikroprocessor ikke kan klare direkte. En MOSFET kan kun opfattes som en ON/OFF kontakt hvis spændingen over kanalen er lav i ON stillingen. Karakteristikken for IRFZ44N viser at strømmen kan være 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 320 – MOSFET transistoren IRFZ44N som kontakt for en motor styret fra en mikroprocessor.

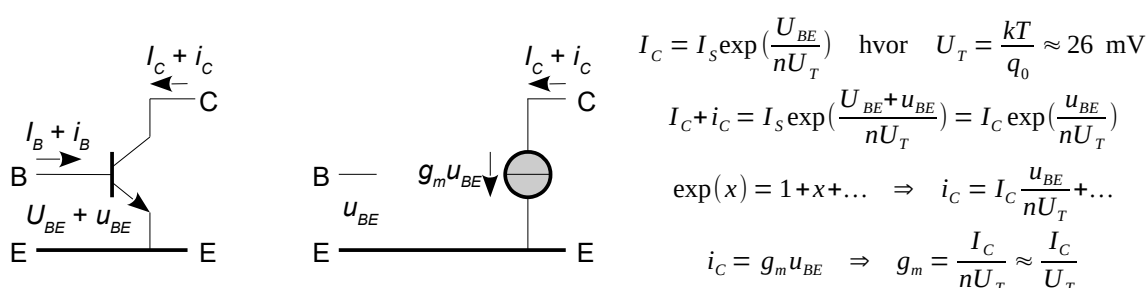
Dioden D_1 beskytter transistoren mod høje spændinger fra motorens kommutator. Dioden skal også benyttes hvis opstillingen skal drive et relæes spole eller en lampe gennem lange ledninger. Modstand R_1 afbryder for transistoren hvis der ikke er forbindelse til mikroprocessoren eller dens udgange er afbrudt som typisk er tilfældet i *power down* og lige efter tilslog 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.

5.5 Transistor ved AC

Den følgende fremstilling antager at transistoren er bragt til et fornuftigt DC arbejds punkt og vi er kun interesseret i hvordan AC strømme og spændinger varierer omkring arbejds punktet. Set over lang tid vil signalet blot svinge med en konstant middelværdi lig med arbejds punktet. Det vil også blive antaget at kredsløbet er lineært og det tillader superposition at blive brugt til at adskille DC og AC forholdene hvilket kaldes for en småsignalmodel (*small signal model*).¹⁶⁵

5.5.1 Simpel AC model

Karakteristikken for en transistor viser at strømmen i kollektor stort ikke afhænger af spændingen over kollektor-emitter (se side 278) og strømmen i transistorens kollektor kan derfor beskrives ved spændingen over basis-emitter efter eksponentialfunktionen fra side 273. For at blive brugbar som AC analyse beskrives den øjeblikkelige værdi af strøm og spænding ved summen af en DC værdi for arbejds punktet (I_B , I_C og U_{BE}) og en AC værdi der repræsenterer signalet (i_B , i_C og u_{BE}).

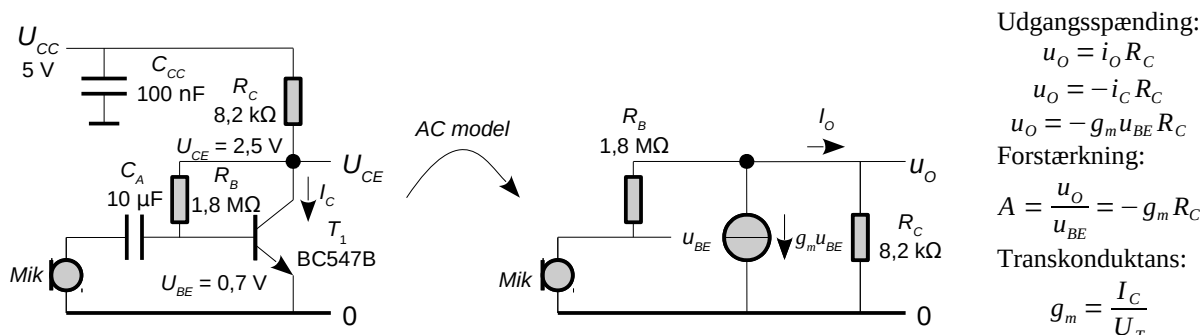


Figur 321 – En småsignalmodel af den bipolære transistor hvor den øjeblikkelige værdi af spænding og strøm opfattes som et DC arbejds punkt (U_{BE} og I_C) overlejret af en AC komponent (u_{BE} og i_C) der kan repræsentere et harmonisk signals amplitude eller et ikke-periodisk signal.

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.

5.5.2 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 322 – En mikrofon 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

¹⁶⁵Betegnelserne 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.

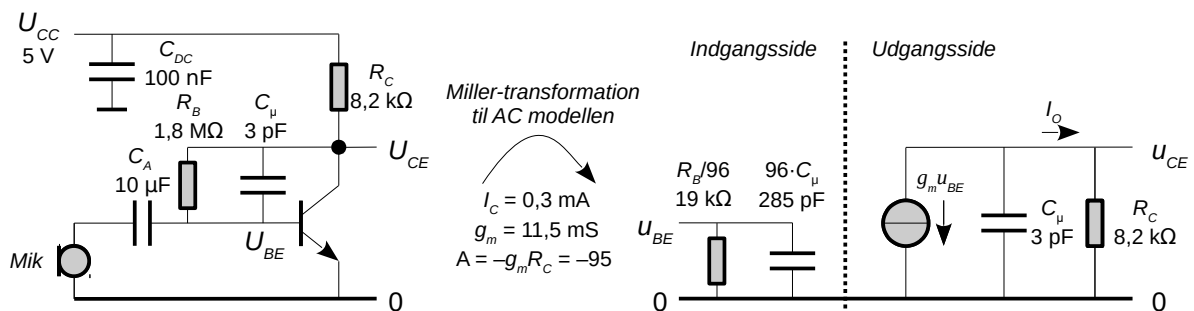
opnå maksimal amplitude på udgangen placeres DC værdien af kollektor-emitter ved halvdelen af effektforsyningens værdi ($U_{CE} = 2,5 \text{ V}$) og derved kan modstanden i kollektor bestemmes fra Ohms lov. For at indstille værdien til det ønskede designs modstanden i basis ud fra basisstrømmen på $I_B = I_C/\beta = 1 \mu\text{A}$ 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 $U_{BE} = 0,7 \text{ V}$. 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 arbejdsunktet 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 arbejdsunktet 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.

5.5.3 Miller transformation

Transistoren har en intern kapacitet C_μ i den spærrende diode i kollektor-basis og den begrænser den opnåelige båndbredde. 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 323 – Mikrofonforstærkeren med transistorens kapacitet C_μ i kollektor indtegnet i diagrammet.

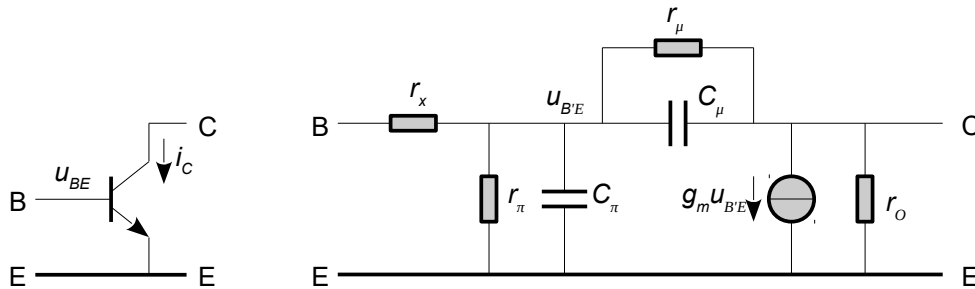
Strømmen i kollektor er $I_C = 0,3 \text{ mA}$ så transkonduktansen er $g_m = I_C/U_T = 11,5 \text{ mS}$ og med $R_C = 8,2 \text{ k}\Omega$ som belastning bliver forstærkningen $A = -g_m R_C = -95$.

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.

5.5.4 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 radorør selv om teknologierne er forskellige.



Figur 324 – 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 strømgeneratorens spændingsafhængighed.

$$\begin{array}{llll}
 r_\pi \approx \frac{\beta}{g_m} & C_\pi \approx \frac{g_m}{2\pi f_T} & r_x \approx 100 \, \Omega & g_m = \frac{I_C}{nU_T} & r_o = \frac{U_A}{I_C} = \frac{1}{h_{OE}} \\
 & & r_\mu \approx 10 \, \text{M}\Omega & & r_o \approx 35 \, \text{k}\Omega \\
 & & C_\mu \approx 1 \dots 5 \, \text{pF} & & U_A \approx 70 \, \text{V} \\
 \beta \approx 100 \dots 800 & f_T \approx \frac{(300 \, \text{MHz}) I_C}{10 \, \text{mA}} & U_T = \frac{kT}{q_0} \approx 26 \, \text{mV} & & h_{OE} \approx 30 \, \mu\text{S} \\
 & @ I_C < 10 \, \text{mA} & & & @ I_C = 2 \, \text{mA}
 \end{array}$$

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 arbejds punktet. For h-parametermodellen vises parametrene som tabelværdier med et sæt kurver for at omregne til et andet arbejds punkt end de 2 mA fra standarden. Den model introduceres på side 300. 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.

5.5.5 Basislagets modstand – r_x

Modstanden fra basis-terminalen og ind til den aktive del af basis kan være betydende 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 effekttransistor. 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 293.

5.5.6 Transkonduktans – g_m

AC strømmen i kollektor er givet ved AC spændingen over basis-emitter og proportionaliteten er transistorens transkonduktans (*transconductance*) med symbolet g_m hvor index kommer fra det engelske ord *mutual* stående for gensidig. Modellen har to porte (ind- og udgang) som forbindes af transkonduktansen. Strømmen i_c er en AC strøm ligesom spændingen u_{BE} er en AC spænding og de er oscillationer omkring DC arbejds punkterne. 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.

$$I_C = I_S \exp\left(\frac{U_{BE}}{nU_T}\right), \quad I_C \gg I_S, \quad U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$$

$n = 1 \dots 2$

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 arbejds punktet I_C mens andet led udelukkende afhænger af signalets amplitude u_{BE} .

$$I_C + i_c = I_S \exp\left(\frac{U_{BE} + u_{BE}}{nU_T}\right) = I_C \exp\left(\frac{u_{BE}}{nU_T}\right) \Rightarrow i_c = I_C \left[\exp\left(\frac{u_{BE}}{nU_T}\right) - 1 \right]$$

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 arbejds punktet bliver forskudt hvilket ikke normalt giver problemer og dertil et led med den dobbelte frekvens.

$$i_c(t) = I_C \left(\frac{1}{4} \left(\frac{U_M}{nU_T} \right)^2 + \frac{U_M}{nU_T} \sin(\omega t) + \frac{1}{4} \left(\frac{U_M}{nU_T} \right)^2 \cos(2\omega t) + \dots \right)$$

Leddene 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 leddet i forhold til amplituden af første-ordens leddet.

$$D_2 = \frac{U_M}{4nU_T} \Rightarrow D_2 \approx 1 \% \frac{U_M}{\text{mV}}$$

Som en tommelfingerregel betyder det at forvrængningen bliver 1 % ved en signalamplitude på 1 mV. Tilnærmelsen forudsætter $n = 1$ og $U_T = 25 \text{ mV}$. Forvrængningen kan regnes for hørbar¹⁶⁶ ved én procent så amplituden bør holdes under 1 mV.

¹⁶⁶Der 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.

5.5.7 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 strømforstærkningen $\beta = 300$ og arbejds punktet $I_C = 2 \text{ mA}$ er værdien $r_\pi \approx 4 \text{ k}\Omega$.

5.5.8 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- π modellen 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 potentiale. Den første af disse kapaciteter er funktion af strømmen og er normalt dominerende hvorimod den anden primært er givet ved fremstillingsprocessen.

$$C_\pi = C_{jBE} + C_{dBE} \xrightarrow{C_{jBE} \gg C_{dBE}} C_\pi \approx \frac{g_m}{2\pi f_T}$$

Transistorens transitfrekvens f_T angiver den frekvens hvor strømforstærkningen 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 omtrent konstant på cirka 150 pF under denne grænse.

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 transistfrekvensen.

$$C_{jBE} = \frac{dQ_{BE}}{dU_{BE}} = \frac{d(\tau_F I_C)}{dU_{BE}} = \tau_F g_m = \frac{g_m}{2\pi f_T}$$

5.5.9 Kollektorkapaciteten – C_μ

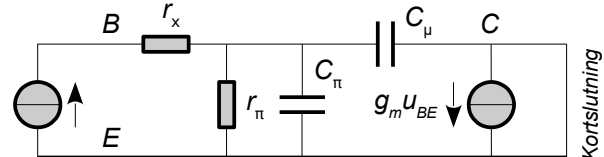
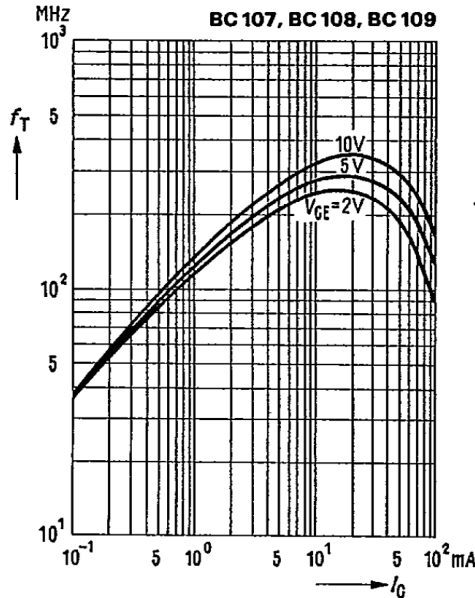
Normalt er dioden i kollektor-basis spærrende hvilket betyder at PN overgangen har to ladninger som adskilt af PN overgangen (*depletion layer*). Det er selve definitionen på en kondensator og den giver kapaciteten af kollektor-basis strækningen. Kapaciteten har fået betegnelsen C_μ og har en værdi på cirka 3 pF for typiske småsignaltransistorer.

Værdien er ikke konstant, men er en funktion af spændingen over kollektor-basis. Ved en forøget spændingsforskel 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 266.

Kondensatoren 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.

5.5.10 Transistrefrekvens – f_T

Transistorens funktion ved høje frekvenser defineres ved transistrefrekvensen (*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_μ har den samme værdi som indgangsmodstanden delt med DC værdien af strømforstærkningen β .



$$f = f_T \Rightarrow \frac{1}{\omega(C_\pi + C_\mu)} = \frac{r_\pi}{\beta} = \frac{1}{g_m} \Rightarrow C_\pi = \frac{g_m}{2\pi f_T} - C_\mu$$

$$C_\pi \approx \frac{g_m}{2\pi f_T}$$

Ved lav strøm er g_m og f_T proportional med I_C så C_π bliver konstant. For BC547 er værdien $C_\pi \approx 50$ pF.

Figur 325 – Transistrefrekvensen for typiske småsignaltransistorer som fx BC547 (Siemens).

Transistrefrekvensen 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 parallellforbindelsen 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 transistrefrekvensen f_T .

$$u_{BE} = \frac{i_B}{1/r_\pi + j\omega(C_\pi + C_\mu)} \Rightarrow \beta = \frac{i_C}{i_B} = \frac{1}{\frac{1}{g_m r_\pi} + \frac{j\omega(C_{dBE} + C_\mu)}{g_m} + j\omega\tau_F}$$

$$C_\pi = C_{dBE} + \tau_F g_m$$

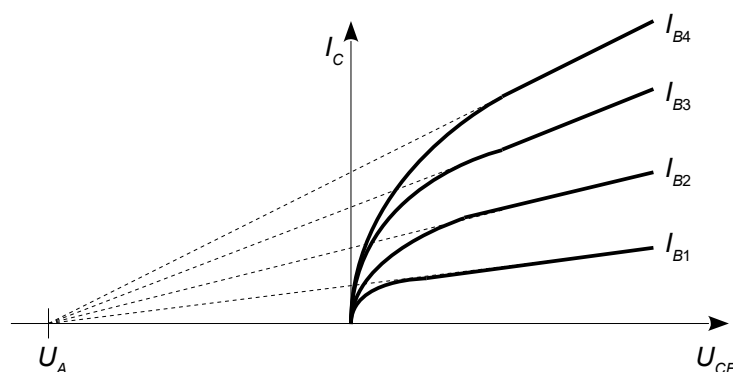
Ved DC er $\omega = 0$ så DC strømforstærkningen defineres som $\beta = g_m r_\pi$ og derved er $1/g_m r_\pi$ en lille værdi ligesom strømforstærkningen normalt er høj. Ved en tilstrækkelig høj frekvens bliver de to andre led større end $1/g_m r_\pi$ og reducerer strømforstærkningen. Transistrefrekvensen 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.

$$\frac{\omega_T(C_{dBE} + C_\mu)U_T}{I_C} + \omega_T\tau_F = 1 \Rightarrow f_T = \frac{1}{\frac{2\pi(C_{dBE} + C_\mu)U_T}{I_C} + 2\pi\tau_F} \Rightarrow f_T \approx \frac{I_C}{2\pi(C_{dBE} + C_\mu)U_T}$$

Transistrefrekvensen 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).

5.5.11 Udgangsmodstanden – r_o

Strømmen i transistorens 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. Udgangskaracteristikken 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 transistorens 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 326 – Udgangskaracteristikken for transistoren viser en relation mellem strømmen i kollektor og spændingen over kollektor-emitter. Hældningen i arbejds punktet 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 300) og fra SPICE simulatorer¹⁶⁷.

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 278.

For BC107, der er ækvivalent med BC547, er Early voltage fundet til $U_A = -96$ V så udgangsmodstanden bliver $r_o = 48$ k Ω ved 2 mA. De opgivne værdier for h-parameteren er $h_{oe} = 18$ μ S for BC107A stigende til $h_{oe} = 60$ μ S for BC107C. Udgangsmodstanden varierer tilsvarende fra $r_o = 56$ k Ω ned til 17 k Ω og dækker den beregnede værdi.

Det er normalt at ignorere udgangsmodstanden ved analytisk arbejde for en modstand i kollektor er normalt væsentligt mindre end transistorens 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 296.

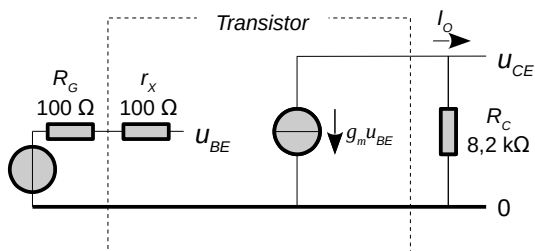
En strømgenerator benyttes i den integrerede kreds LM3900 hvor forstærkningen er 2800 gange.

For BC547 ved $I_C = 2$ mA er transkonduktansen $g_m = 77$ mS og med $r_o = 17 \dots 56$ k Ω 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.

¹⁶⁷Ved Multisim placeres transistoren på arbejdsfladen. Man dobbelt-klikker på komponenten og åbner for "Edit model" hvor SPICE parametrene er opført.

5.5.12 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 $g_m R_C$.



$$\begin{aligned}
 (1) \text{ Termisk støj: } & e_{nR} = \sqrt{4kT(R_G + r_x)} \\
 (2) \text{ Haglstøj fra } I_B: & i_{nB} = \sqrt{2q_0 I_B} = \sqrt{2q_0 \frac{I_C}{\beta}} \\
 & e_{nB} = (R_G + r_x) i_{nB} \\
 (3) \text{ Haglstøj fra } I_C: & i_{nC} = \sqrt{2q_0 I_C} \\
 & e_{nC} = \frac{R_C i_{nC}}{g_m R_C} = U_T \sqrt{\frac{2q_0}{I_C}} \\
 \text{Samlet indgangsstøj } & e_{nTOT} = \sqrt{e_{nR}^2 + e_{nB}^2 + e_{nC}^2}
 \end{aligned}$$

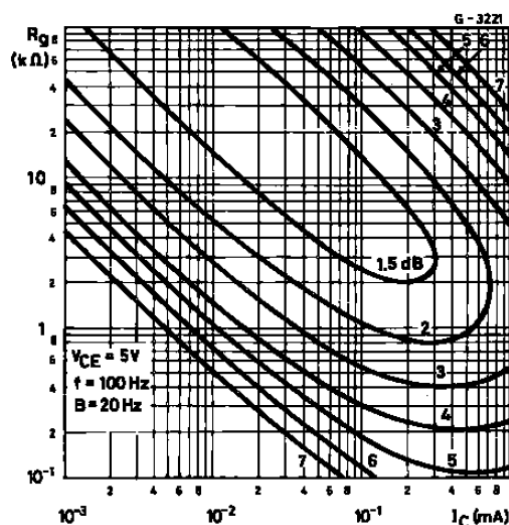
Figur 327 – Støjbidragene er signalkildens indre modstand R_G og transistorens modstand i basis r_x der begge er termiske støjklender og dertil haglstøjen fra strømmen i basis og i kollektor. Alle bidrag bliver henført til indgangen ved forstærkningen ved $g_m R_C$ der antager at udgangen er ubelastet.

For mikrofonforstærkeren fra side 286 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 $I_B = 1 \mu\text{A}$. Haglstøjen 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 udgangen er $80 \text{ nV}/\sqrt{\text{Hz}}$ der henføres til basis ved forstærkningen $g_m R_C = 95$ 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 $2 \text{ nV}/\sqrt{\text{Hz}}$ og er domineret af de to modstandes termiske støj. Til sammenligning er en ekstremt lav støjende operationsforstærker som fx AD797 på $0,9 \text{ nV}/\sqrt{\text{Hz}}$.

5.5.13 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 forringer 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.



Figur 328 – Støjtalet NF for BC109 (BC549) der er funktion af generatorens modstand og strømmen i kollektor (SGS Thomson).

Definition af støjtalet:

$$NF = 20 \cdot \log_{10} \left(\frac{SNR_G}{SNR_K} \right) \text{ dB} = 20 \cdot \log_{10} \left(\frac{U_G/U_{nRG}}{U_C/U_{n(RG+RX)}} \right) \text{ dB}$$

$$\Rightarrow \frac{NF}{\text{dB}} = 20 \cdot \log_{10} \left(\frac{U_{n(RG+RX)}}{U_{nRG}} \right) = 20 \cdot \log_{10} \left(\frac{e_{n(RG+RX)}}{e_{nRG}} \right)$$

$$\Rightarrow \frac{NF}{\text{dB}} = 20 \cdot \log_{10} \left(\frac{\sqrt{4kT(R_G + r_x)}}{\sqrt{4kT R_G}} \right) = 10 \cdot \log_{10} \left(1 + \frac{r_x}{R_G} \right)$$

$$\Rightarrow r_x = R_G (10^{\frac{NF}{10 \text{ dB}}} - 1)$$

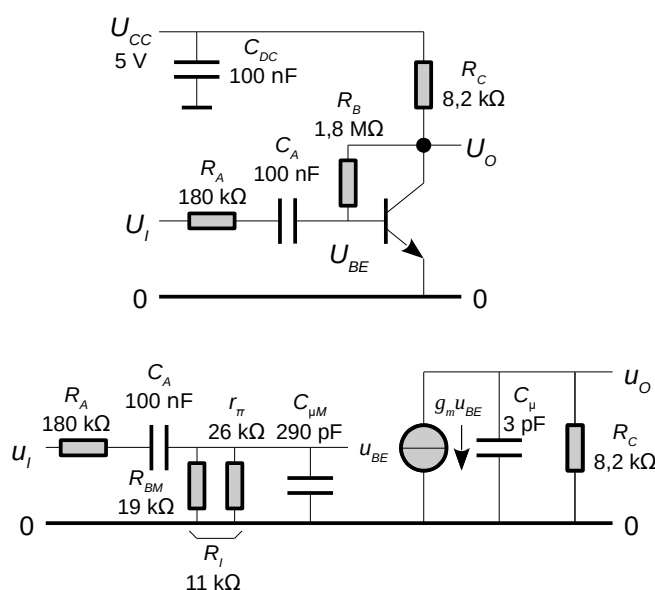
For $I_C < 500 \mu\text{A}$ og $R_G = 110 \Omega$ aflæses:

$$NF = 5 \text{ dB: } \Rightarrow r_x = 110 \Omega \cdot (10^{\frac{5 \text{ dB}}{10 \text{ dB}}} - 1) = 240 \Omega$$

5.5.14 Inverterende forstærker

Den inverterende forstærker kan opfattes som en operationsforstærker hvor signalkilden føres ind via R_A og med tilbagekobling via R_B . En forstærkning 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 277 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 $-5,8$.



DC design:

$$R_C = \frac{U_{CC} - U_O}{I_C}, \quad R_B = \beta \frac{U_O - U_{BE}}{I_C}$$

AC parametre:

$$g_m = \frac{I_C}{U_T}, \quad r_\pi = \frac{\beta}{g_m}$$

$$R_{BM} = \frac{R_B}{1 + g_m R_C}, \quad R_I = R_{BM} \parallel r_\pi$$

Forstærkning:

$$A = \frac{R_I}{R_A + R_I} A_{OL}$$

$$A_{OL} = -g_m R_C$$

Grænsefrekvenser:

$$f_L \approx \frac{1}{2\pi(R_A + R_I)C_A}$$

$$f_{H1} \approx \frac{1}{2\pi R_I C_{\mu M}}, \quad f_{H2} \approx \frac{1}{2\pi R_C C_\mu}$$

Figur 329 – Miller-transformering af tilbagekoblingen viser at i indgangsmodstanden og åben-sløjfe forstærkningen begge er for lave til at forstærkningen kan indstille alene ved R_A og R_B .

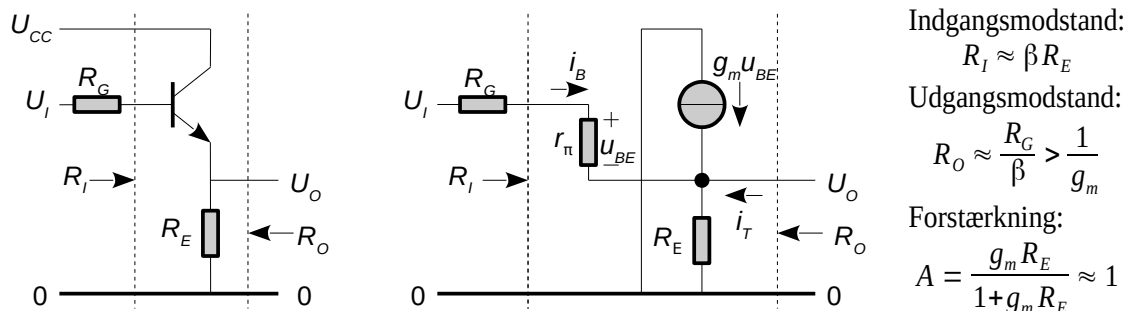
Med $U_{CC} = 5 \text{ V}$, $U_O = 2,5 \text{ V}$ og $I_C = 0,3 \text{ mA}$ bliver $R_C = 8,2 \text{ k}\Omega$ og for $\beta = 300$ er $R_B = 1,8 \text{ M}\Omega$. Transistorens transkonduktans beregnes til $g_m = 11,5 \text{ mS}$ så åben-sløjfe forstærkningen bliver $A_{OL} = -94$. Modstanden i basis 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 arbejds punktet 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 betydende og vil dæmpe forstærkningen for frekvenser over f_{H1} og f_{H2} . Den første skyldes indgangens lavpasfilter ved $C_{\mu M}$ og R_I . Parallelforbindingen med R_A ignoreres da den er meget større end R_I . En Miller-transformering af tilbagekoblingens kondensator giver $C_{\mu M} = 95 C_\mu = 290 \text{ pF}$ 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.

5.5.15 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 potentiale 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 330 – En emitterfølger har en høj indgangsmodstand og lav udgangsmodstand. 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 $\beta = 25$ giver en indgangsmodstand på 200 Ω, men med to emitterfølgere i kaskade (*Darlington pair*, se næste kapitel) bliver den over 20 kΩ.

Udgangsmodstanden er givet af signalkildens indre modstand divideret med strømforstærkningen og kommer sjældent over 100 Ω. Som minimum er udgangsmodstanden $1/g_m$ 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 301.

Emitterfølgerens spændingsforstærkning er lidt mindre end én idet der er en spændingsdeling mellem udgangsmodstanden R_O og emittermodstanden R_E .

Indgangsmodstanden på basis findes ved at beregne indgangsspændingen U_I som følge af strømmen i_B i basis og danne forholdet mellem de to.

$$U_I = r_\pi i_B + u_E \Rightarrow R_I = \frac{U_I}{i_B} = r_\pi + (1 + \beta) R_E \Rightarrow R_I \approx \beta R_E$$

$$u_E = R_E (i_B + i_C) = (1 + \beta) R_E i_B$$

Udgangsmodstanden 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.

$$U_O = R_E [i_B + i_C + i_T] \Rightarrow R_O = \frac{U_O}{i_T} = \frac{(R_G + r_\pi) R_E}{R_G + r_\pi + (1 + \beta) R_E} \approx \frac{r_\pi}{1 + \beta} \Rightarrow R_O \approx \frac{1}{g_m}$$

$$i_B = -\frac{U_O}{R_G + r_\pi}$$

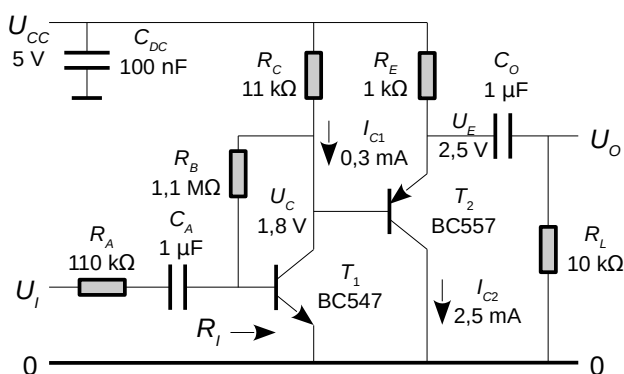
Forstærkningen bestemmes ved Kirchhoff på knudepunktet ved emitteren. Det sidste trin i udledningen herunder antager at strømforstærkningen kan regnes for stor.

$$\frac{U_O - U_I}{R_G + r_\pi} + \frac{U_O}{R_E} - g_m u_{BE} = 0 \Rightarrow A = \frac{U_O}{U_I} = \frac{1 + g_m r_\pi}{1 + g_m r_\pi + \frac{R_G + r_\pi}{R_E}} = \frac{1}{1 + \frac{R_G}{(1 + \beta) R_E} + \frac{r_\pi}{(1 + \beta) R_E}} \approx \frac{g_m R_E}{1 + g_m R_E}$$

$$u_{BE} = -\frac{r_\pi}{R_G + r_\pi} (U_O - U_I)$$

5.5.16 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 sænket og det øgede spændingsfald over R_C hæver åben-sløjfe forstærkningen i T_1 .



Designprocedure:

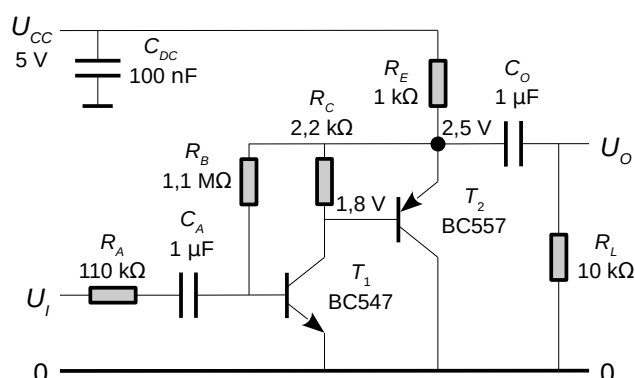
1. Vælg forsyningen U_{CC} .
2. Vælg strømmene I_{C1} og I_{C2} .
3. Aflæs β_1 og β_2 fra datablad.
4. $R_E = U_{CC}/2I_{C2}$.
5. $U_C = U_{CC}/2 - U_{BE2}$.
6. $R_C = (U_{CC} - U_C)/I_{C1}$.
7. $R_B = \beta_1(U_C - U_{BE1})/I_{C1}$.
8. $g_{m1} = I_{C1}/U_T$ og $g_{m2} = I_{C2}/U_T$.
9. $A_{OL} = g_{m1}R_C$.
10. $r_{\pi1} = \beta_1/g_{m1}$ og $r_{\pi2} = \beta_2/g_{m2}$.
11. $R_i = r_{\pi1} \parallel R_B/A_{OL}$.
12. $R_O = (R_C + r_{\pi2})/A_{OL}$.

Figur 331 – 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 $R_O = R_C/\beta_2 + 1/g_{m2} = 47 \text{ }\Omega$.

5.5.17 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:

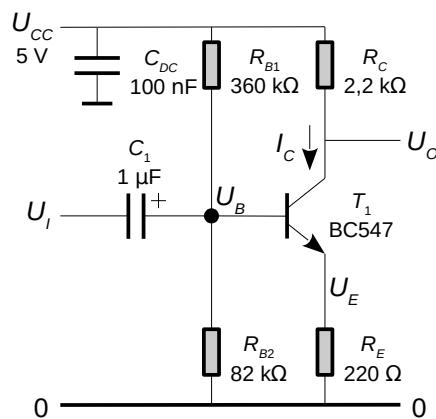
$$A_{OL} = g_{m1} R_{Ceff}$$

Figur 332 – 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 $-9,1$.

5.5.18 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, mindsker 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 opstilling fra omkring 1960 hvor transistoren tog over fra elektronrøret, men den 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:

$$R_O = r_o \parallel R_C \approx R_C$$

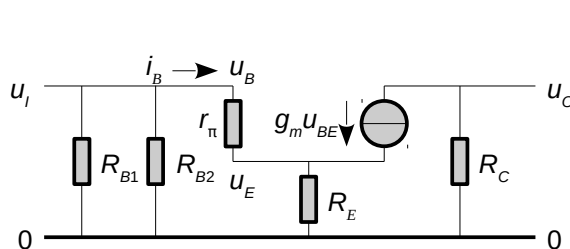
Figur 333 – En inverterende forstærker med tilbagekobling fra modstanden i emitter.

De viste komponenter giver $U_B = 0,93 \text{ V}$ så $U_E = 0,23 \text{ V}$ hvorved $I_C = 1 \text{ mA}$, $g_m = 38,5 \text{ mS}$ og $r_\pi = 7,8 \text{ k}\Omega$ for $\beta = 300$. Forstærkningen er 8,9 gange og indgangsmodstanden bliver $35 \text{ k}\Omega$.

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 genere 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 betydende forskel på β og $1 + \beta$.



$$u_I = r_\pi i_B + R_E (i_B + g_m r_\pi i_B) = (r_\pi + (1 + \beta) R_E) i_B \approx (r_\pi + \beta R_E) i_B$$

$$R_I = R_{B1} \parallel \frac{u_I}{i_B} \approx R_{B1} \parallel (r_\pi + \beta R_E)$$

$$R_B = R_{B1} \parallel R_{B2}$$

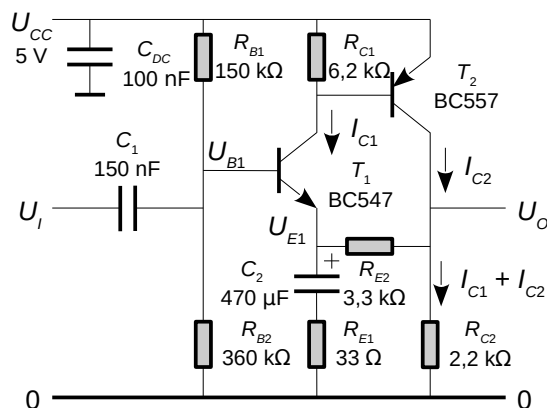
$$u_O = -g_m u_{BE} R_C \approx -\beta i_B R_C$$

$$A = \frac{u_O}{u_I} \approx -\frac{\beta R_C}{r_\pi + \beta R_E} = -\frac{g_m R_C}{1 + g_m R_E}$$

Figur 334 – Analysen af forstærkeren tager udgangspunkt i strøm og spænding ved basis.

5.5.19 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.



DC dimensionering:

$$U_{B1} = U_{BE1} + R_{E1} I_{C1} + U_O \quad U_O = \frac{U_{CC}}{2}$$

$$R_{B1} = 10 \frac{U_{CC} - U_{B1}}{I_{C1}} \quad R_{B2} = 10 \frac{U_{B1}}{I_{C1}}$$

$$R_{C1} = \frac{U_{BE2}}{I_{C1}} \quad R_{C2} = \frac{U_O}{I_{C1} + I_{C2}}$$

$$R_{E1} \approx \frac{R_{E2}}{1+A} \quad R_{E2} = \frac{U_{B1} - U_{BE1} - U_O}{I_{C1}}$$

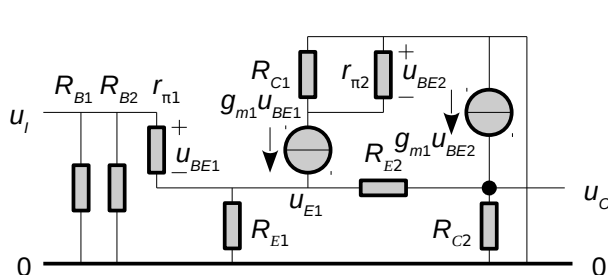
AC dimensionering:

$$C_1 > \frac{1}{2\pi f_L R_{B1} \parallel R_{B2}} \quad C_2 > \frac{1}{2\pi f_L R_{E1}}$$

Figur 335 – 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 $I_{C1} = 100$ μ A for at få en lav værdi af strømmen i basis på første transistor og $I_{C2} = 1$ 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} designses 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.



$$g_{m1} = \frac{I_{C1}}{U_T} \quad g_{m2} = \frac{I_{C2}}{U_T}$$

$$r_{\pi1} = \frac{\beta_1}{g_{m1}} \quad r_{\pi2} = \frac{\beta_2}{g_{m2}}$$

$$R_{L1} = R_{C1} \parallel r_{\pi2} \quad R_{L2} = R_{E2} \parallel R_{C2}$$

$$A_1 = -g_{m1} R_{L1} \quad A_2 = -g_{m2} R_{L2}$$

$$R_I \approx R_{B1} \parallel R_{B2} \quad A_{OL} = A_1 A_2$$

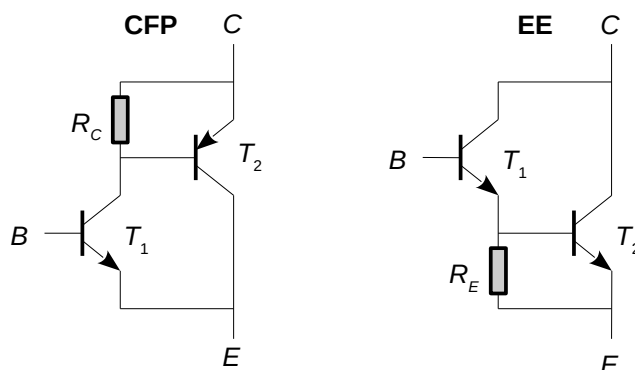
$$A_{CL} = \frac{A_{OL}}{1 + B A_{OL}} \quad B = \frac{R_{E1}}{R_{E1} + R_{E2}}$$

Figur 336 – Analyse af den ikke-inverterende forstærker.

Med $g_{m1} = 3,8$ mS og $g_{m2} = 38$ mS bliver $r_{\pi1} = 65$ k Ω og $r_{\pi2} = 13$ k Ω , som i parallel med R_{C1} giver en belastning af første transistor på $R_{L1} = 4,2$ k Ω . 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 $R_{L2} = 1,3$ k Ω 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 $A_{CL} = 89$.

5.5.20 Forstærket emitterfølger (CFP)

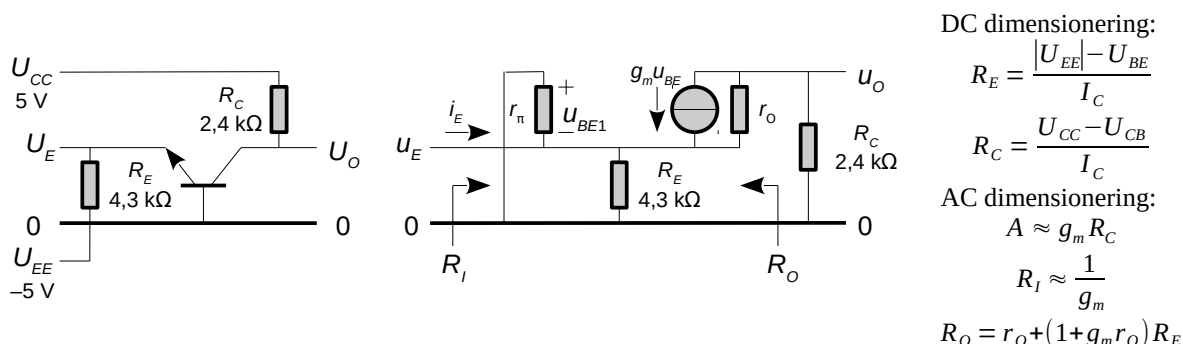
En forstærket emitterfølger med komplementære transistorer (CFP, *complementary feedback pair*) kan benyttes til at drive en stor strøm med en lav indgangsstrøm. Kredsløbet minder om meget om den dobbelte emitterfølger (EE, *Darlington*), men der er kun en enkelt basis-emitter strækning fra indgang til 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 337 – En forstærket emitterfølger (CFP) og en dobbelt emitterfølger (EE).

5.5.21 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 338 – 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 .

$$g_m u_{BE} + \frac{u_O - u_E}{r_o} + \frac{u_O}{R_C} = 0 \Rightarrow (g_m + \frac{1}{r_o}) u_E = (\frac{1}{r_o} + \frac{1}{R_C}) u_O \Rightarrow A = \frac{u_O}{u_E} = \frac{g_m + 1/r_o}{1/r_o + 1/R_C} \approx g_m R_C$$

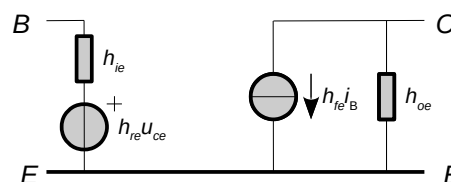
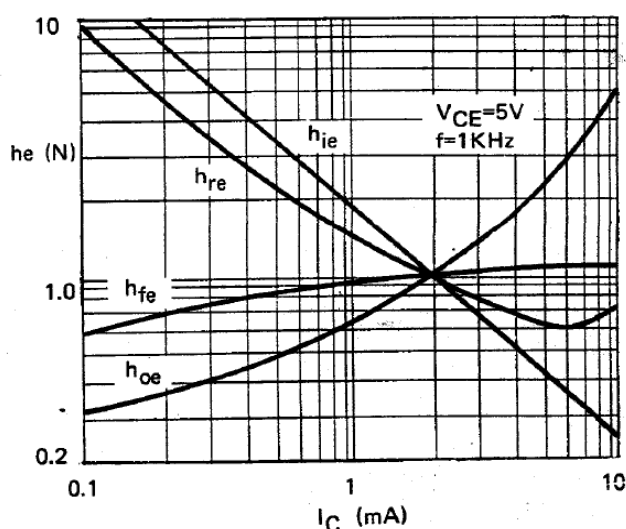
Indgangsmodstanden R_I findes ved at opstille Kirchhoffs ligning for knuden ved emitter. Den kan normalt tilnærmes ved U_T/I_C der er identisk med diodens dynamiske modstand på $R_I = 26 \Omega$ for en strøm i kollektor på $I_C = 1 \text{ mA}$.

$$-i_E + \frac{u_E}{r_\pi} + \frac{u_E}{R_E} - g_m u_{BE} = 0 \Rightarrow (\frac{1}{r_\pi} + \frac{1}{R_E} + g_m) u_E = i_E \Rightarrow R_I = \frac{u_E}{i_E} = \frac{1}{1/r_\pi + 1/R_E + g_m} \approx \frac{1}{g_m}$$

På samme måde bestemmes udgangsmodstanden R_O .

5.5.22 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.



$$r_{\pi} \approx h_{ie}, \quad r_{\mu} \approx \frac{r_{\pi}}{h_{re}}, \quad \beta \approx h_{fe}, \quad r_o \approx \frac{1}{h_{oe}}$$

$I_C = 2 \text{ mA}$	BC107A	BC107B	BC107C
$h_{ie} =$	2,7 k Ω	4,5 k Ω	8,7 k Ω
$h_{re} =$	$1,5 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	$3 \cdot 10^{-4}$
$h_{fe} =$	190	330	580
$h_{oe} =$	18 μS	30 μS	60 μS

Figur 339 – 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 overflademontage.¹⁶⁸

Kurverne krydser ved 2 mA der er arbejds punktet 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.

Udgangsmodstanden r_o kan estimeres fra h-parameter modellens værdier hvor de viste data giver $r_o = 17 \dots 56 \text{ k}\Omega$ ved $I_C = 2 \text{ mA}$ (se også side 278) stigende til $76 \dots 250 \text{ k}\Omega$ ved $I_C = 0,1 \text{ mA}$.

$$r_o = \frac{U_A}{I_C} = \frac{1}{h_{oe}} \Rightarrow U_A \approx \frac{I_C}{h_{oe}}$$

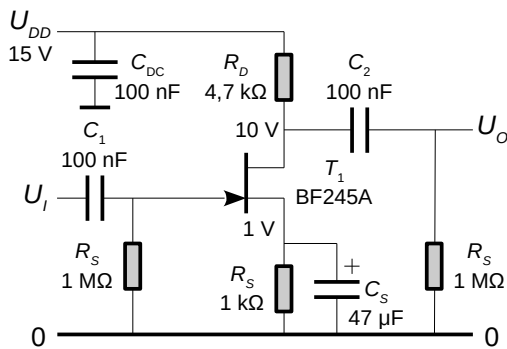
Værdien specificeres i hybrid-pi modellen ved transistorens Early-voltage værdi U_A der ganske vist ikke opgives i databladet, men kan findes i SPICE analysatorens modelbibliotek. Multisim oplyser parameteren til 80 V der svarer til $40 \text{ k}\Omega$ ved $I_C = 2 \text{ mA}$. Værdien passer med beregningen side 278, men jeg har set publicerede kurver der overhovedet ikke passer med beregninger. Med andre ord er udgangsmodstanden en usikker parameter.

Som en brugbar tommelfingerregel kan man antage $r_o = 50 \text{ k}\Omega$ ved et lavt strømniveau.

¹⁶⁸Ifølge <http://en.wikipedia.org/wiki/BC548> er BC108 den oprindelige betegnelse 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.

5.5.23 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:

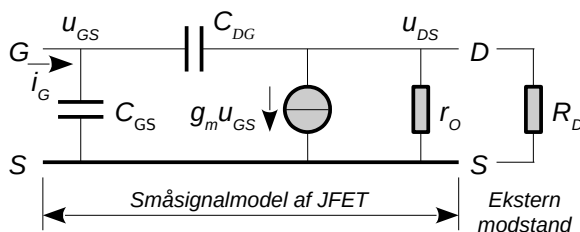
1. Vælg U_{DD} høj nok for strømregulering.
2. Vælg strømmen i drain, I_D .
3. Aflæs gate-source spændingen, U_{GS} .
4. Beregn $R_S = U_{GS}/I_D$.
5. Vælg udgangsspændingen U_D .
6. Beregn $R_D = (U_{DD} - U_D)/I_D$.
7. Beregn g_m fra formelen på næste side.
8. Beregn $A = g_m R_D / (1 + g_m R_S)$, 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 340 – 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 .

5.5.24 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:

$$A = -\frac{g_m r_o}{1 + s C_{DG} r_o}, \quad f_H = \frac{1}{2\pi C_{DG} r_o}$$

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 341 – 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 princippet 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 arbejds punktet I_D som en følge af en tilvækst i U_{GS} hvilket svarer til differentiation. Her benyttes kædere g len: $dy/dx = (dy/dz)(dz/dx)$. For at nå frem til udtrykket skal det indses at U_P er negativ.

$$g_m = \frac{dI_D}{dU_{GS}} = \frac{d}{dU_{GS}} \left(1 - \frac{U_{GS}}{U_P} \right)^2 I_{DSS} = 2 I_{DSS} \frac{d(1 - U_{GS}/U_P)^2}{d(1 - U_{GS}/U_P)} \frac{d(1 - U_{GS}/U_P)}{dU_{GS}} \frac{1}{U_P} = -2 \frac{I_{DSS}}{U_P} \left(1 - \frac{U_{GS}}{U_P} \right)$$

Udgangsmodstanden er givet ved hældningen af karakteristikken i arbejds punktet. For BF245A findes af databladet en strømvariation på $\Delta I_D = 0,1$ mA omkring arbejds punktet $I_D = 2,4$ mA ved ændringen ΔU_{DS} fra 5 V til 10 V så udgangsmodstanden bliver $r_O \approx 50$ k Ω .

$$r_O = \frac{dU_{DS}}{dI_D} \approx \frac{\Delta U_{DS}}{\Delta I_D} = \frac{U_{DS2} - U_{DS1}}{I_{D2} - I_{D1}}$$

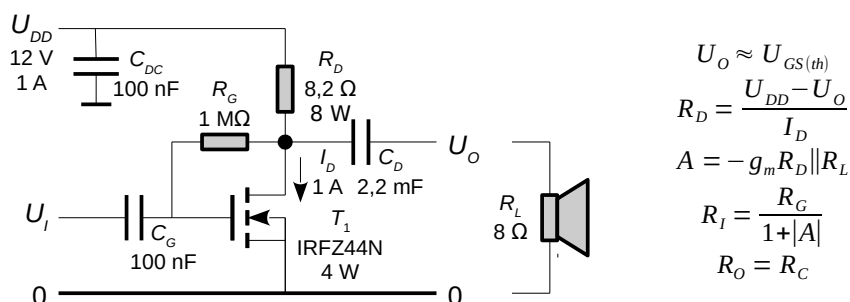
Med en spændingskilde på gate og Kirchhoff på drain kan overføringsfunktionen bestemmes.

$$\frac{U_{DS} - U_{GS}}{1/sC_{DG}} + g_m U_{GS} + \frac{U_{DS}}{r_O} = 0 \Rightarrow \frac{U_{DS}}{U_{GS}} = \frac{sC_{DG} - g_m r_O}{1 + sC_{DG} r_O} \approx -\frac{g_m r_O}{1 + sC_{DG} r_O} \quad \text{for } f \ll \frac{g_m r_O}{2\pi C_{DG}}$$

Forstærkningen er $g_m r_O = 100$ for $g_m = 2$ mS og $r_O = 50$ k Ω så værdien af leddet sC_{DG} er uden betydning. Grænsefrekvensen ved nævnerens polynomium er 3 MHz med $C_{DG} = 1$ pF.

5.5.25 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 betydende spændingsfald over R_G . DC arbejds punktets 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 arbejds punktet.



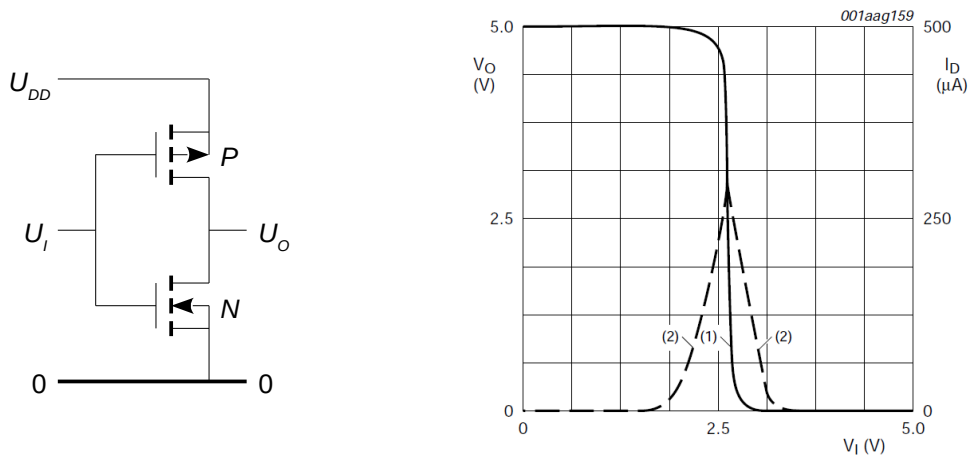
Figur 342 – En MOSFET effektforstærker der kan levere indtil 1 W til højttaleren, men da der ikke er en tilbagekobling er forvrængningen givet af transistorens udgangskaraktistik.

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 arbejds punktet 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 transkonduktans¹⁶⁹ på 6 S ved dette arbejds punkt så spændingsforstærkningen estimeres til 25 gange.

¹⁶⁹ Kurverne på side 285 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.

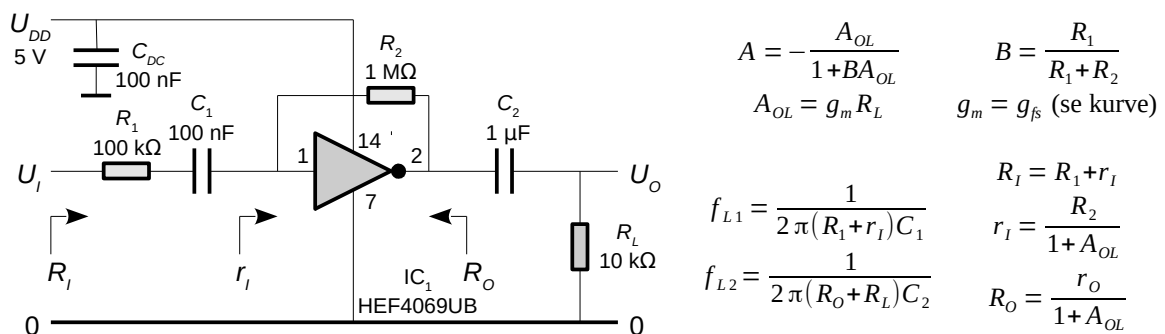
5.5.26 CMOS forstærker

To MOSFET, henholdsvis en P kanal og en N kanal, danner et komplementært par der er velegnet 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 343 – 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 kan bruges til analoge formål. De kaldes for *unbuffered* og arbejds punktet indstilles ved at holde ind- og udgang på det 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. 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.

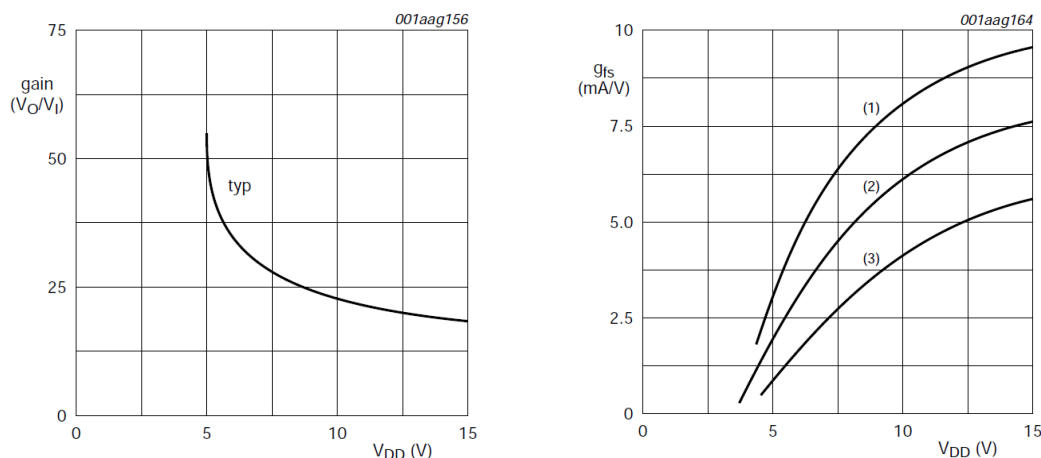


Figur 344 – En CMOS forstærker benytter et DC balancepunkt nær ved den halve effektforsyning.

Med de viste komponentværdier og $U_{DD} = 5$ V er $g_m = 2$ 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 $r_I \approx 50$ kΩ bliver indgangsmodstanden på $R_I \approx 150$ kΩ. Udgangsmodstanden er $r_O \approx 3,5$ kΩ. Grænsefrekvenserne bliver $f_{L1} \approx 10$ Hz og $f_{L2} \approx 12$ 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.

Småsignalmodellen minder om en JFET. Transkonduktansen er vist for *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 specificeret 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 og risiko for sammenbrud (*latch up*) der er et problem ved en forsyningsspænding over 18 V.



Figur 345 – 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 k Ω 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.

$$R_o = \frac{r_o}{1+B|A_{OL}|}, \quad B = \frac{R_1}{R_1+R_2}$$

Støjstrømtætheden som følge af haglstøjen er givet af arbejdsunktets $I_D = 0,28 \text{ mA}$. Den giver et spændingsfald over udgangsmodstanden r_o og med to transistorer i kredsløbet forøges RMS værdien af støjspændingen med kvadratroden af to. Det giver spændingsstøjtætheden u_n og støjen kan henføres til indgangen som spændingsstøjtætheden e_n .

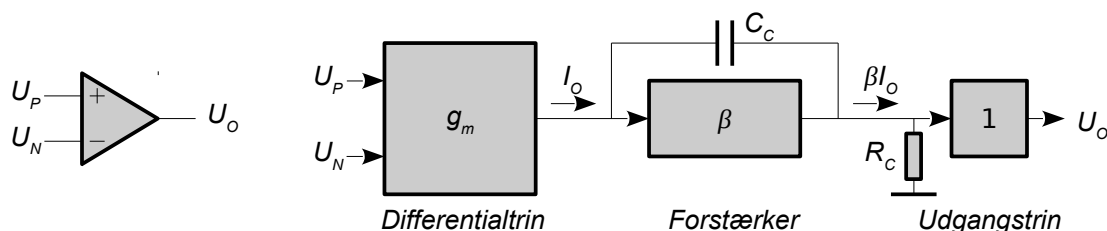
$$i_n = \sqrt{2q_0 I_D} = 9,5 \text{ pA}/\sqrt{\text{Hz}} \Rightarrow e_n = \frac{u_n}{A_{OL}} \approx 16,7 \text{ nV}/\sqrt{\text{Hz}}$$

$$u_n = \sqrt{2} r_o i_n \approx 334 \text{ nV}/\sqrt{\text{Hz}}$$

Note 4.

5.6 Operationsforstærker

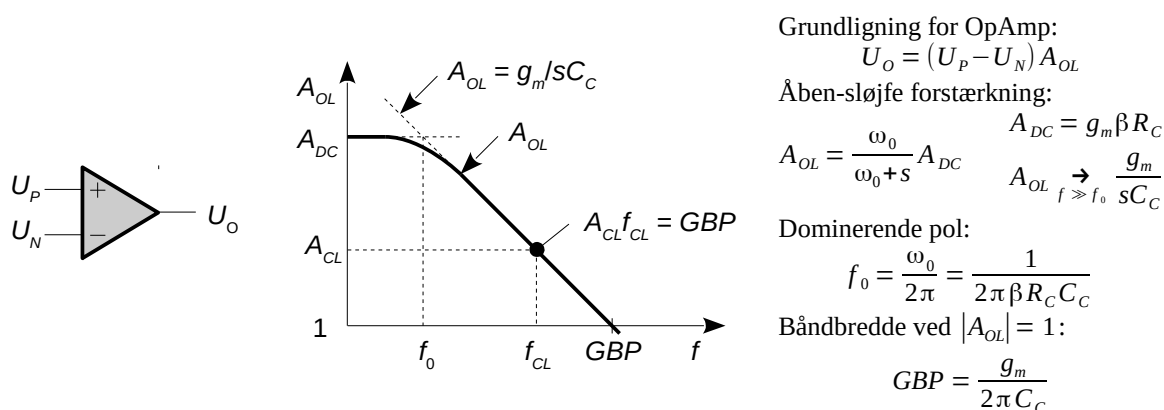
En OpAmp består af tre blokke. En spændingsdifferens ved indgangen omformes til en strøm ved første blok og forstærkes op i den midterste blok for at blive omsat til en spænding af en modstand. En buffer med en spændingsforstærkning på én driver en strøm til belastningen. Kondensatoren begrænser båndbredden og gør forstærkeren stabil, men er også med til at gøre den mere lineær.



Figur 346 – 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 321.

En OpAmp er kendetegnet 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.



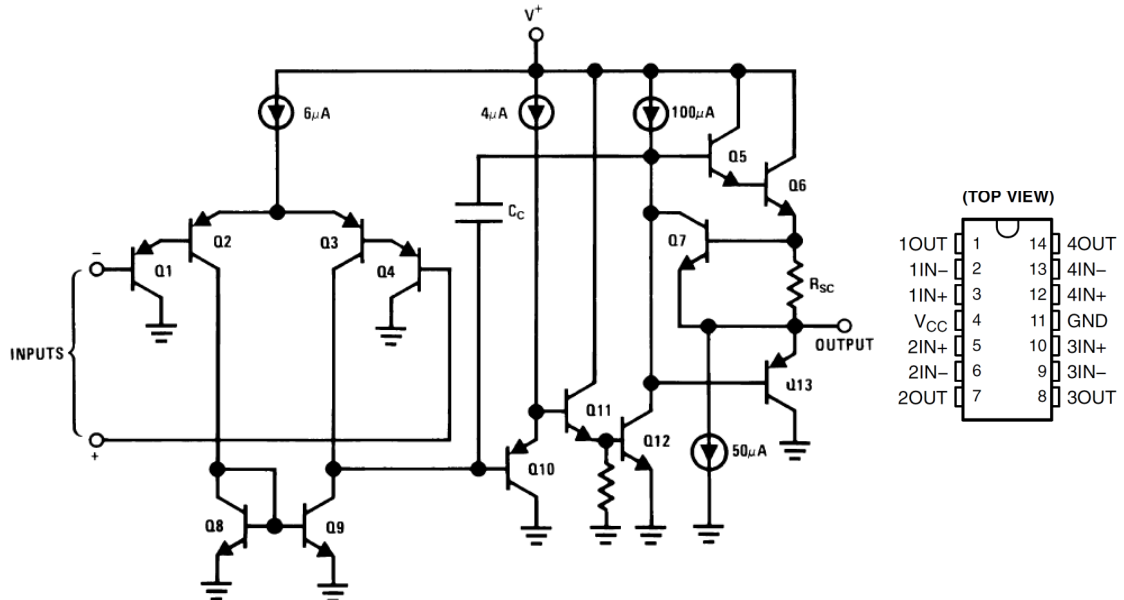
Figur 347 – Åben-sløjfe forstærkningen svarer til et lavpasfilter hvor grænsefrekvensen ofte ligger under den laveste frekvens for den pågældende applikation så forstærkeren er integrerende.

For frekvenser over den dominerende pol er forstærkningen omvendt proportionel med frekvensen og et specielt kendetegn ved en operationsforstærker er derfor at forstærkning gange båndbredde 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 transitfrekvens 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.

5.6.1 LM324

Herunder ses diagrammet for en af de mest populære operationsforstærkere nogen sinde. 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 operationsforstærkerer 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 348 – Den populære LM324 med de væsentligste detaljer vist (National Semiconductor).

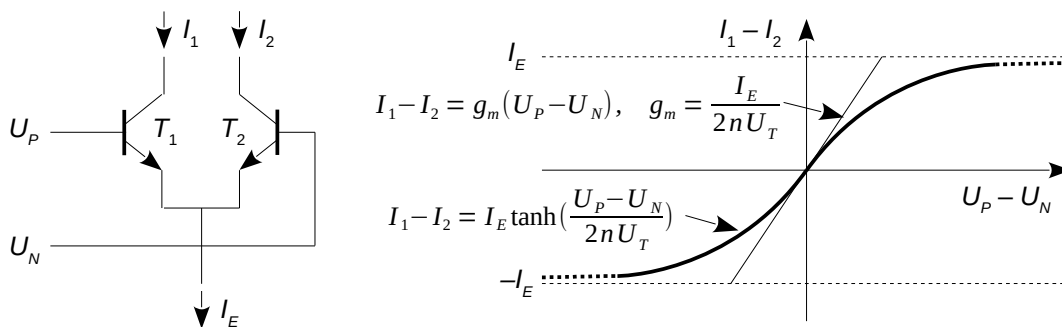
Indgangstrinnet består af differentialforstærkeren Q1 til Q4 med to PNP transistorer i hver gren der tillader at indgangen fungerer selv om transistorernes baser er under nul. Databladet tillader en indgangsspænding ned til $-0,3$ V. Indgangsstrømmen er 45 nA og varierer ikke med temperaturen så værdien 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 306 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 $50\mu A$ og Q13 er afbrudt da basis er positiv i forhold til emitter. Udgangstrinnet er derfor en dobbelt emitterfølger for en belastning på højest $\pm 50\mu A$. 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 $\pm 0,5$ V/ μs 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.

5.6.2 Differentialtrin

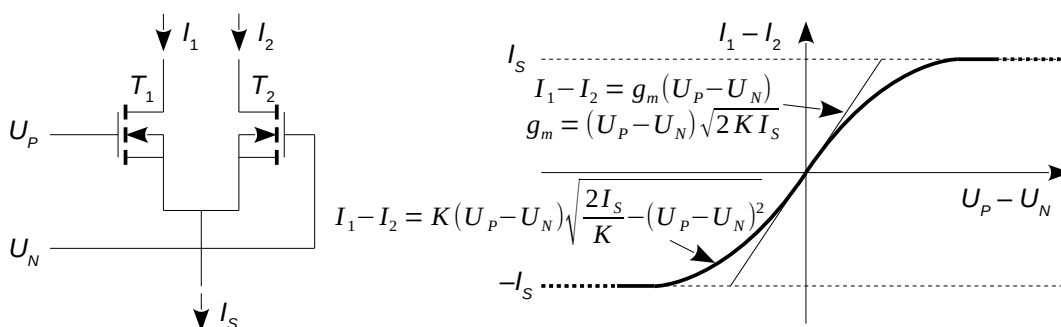
En operationsforstærker har to indgange med spændingen U_P på den positive indgang og U_N på den negative indgang. På udgangssiden løber der strømmen I_1 i transistor T_1 og I_2 i transistor T_2 . Differensen mellem de to strømme er funktion af differensen mellem de to indgangsspændinger og ved en svag 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 μV og det er så lidt at en kredsløbsanalyse normalt blot antager at $U_N = U_P$.



Figur 349 – 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 312.

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 350 – Et differentialtrin md 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 $\pm I_S$.

5.6.3 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 arbejds punktet 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 på $I_1 = I_2 = I_E/2$. Det er den balance som en operationsforstærker vil have som arbejds punkt på grund af tilbagelkoblingen. Med en differens mellem de to baser som følge af et eksternt signal vil arbejds punktet 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 bipolære transistor og det vil i det følgende blive forudsat at transistorerne er så tilpas ens at mætningsstrømmen 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 arbejds punktet $I_E/2$ gange med et udtryk for den forskydning strømmen er udsat for.

$$I_1 = I_S \exp\left(\frac{U_{BE1}}{nU_T}\right) = I_S \exp\left(\frac{U_{BE} + \frac{U_P - U_N}{2}}{nU_T}\right) = I_S \exp\left(\frac{U_{BE}}{nU_T}\right) \exp\left(\frac{U_P - U_N}{2nU_T}\right) = \frac{I_E}{2} \exp\left(\frac{U_P - U_N}{2nU_T}\right)$$

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 mætningsstrøm I_S ikke indgår. For at lede frem til en småsignalanalyse benyttes en Taylor rækkeudvikling af den hyperbolske 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 .

$$U_P - U_N = U_M \sin(\omega t) \Rightarrow \frac{I_1 - I_2}{I_E} = \frac{U_M}{2nU_T} \sin(\omega t) - \frac{1}{3} \left[\frac{3}{4} \frac{U_M \sin(\omega t)}{2nU_T} - \frac{1}{4} \left(\frac{U_M}{2nU_T} \right)^3 \sin(3\omega t) \right]$$

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.

$$D_3 = \frac{1}{12} \left(\frac{U_M}{2nU_T} \right)^2 \approx \frac{1}{48} \left(\frac{U_M}{U_T} \right)^2 \Rightarrow D_3 = 1\% \quad \text{ved} \quad U_M = U_T \sqrt{0,01 \cdot 48} = 18 \text{ mV}$$

5.6.4 Overføringsfunktionen for felt-effekt transistorer

S-kurven for en MOSFET ligner kurven for en bipolar 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 281 og 285).

$$I_1 = K(U_{GS1} - U_{GS(th)})^2 \Rightarrow (U_{GS1} - U_{GS(th)})^2 = \frac{I_1}{K} \Rightarrow U_{GS1} = \sqrt{\frac{I_1}{K}} + U_{GS(th)}$$

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 omroteres, 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 leddet med strømmen I_S i det tidligere udtryk så kun differensen ved udgangsstrømmen bliver tilbage.

$$(I_1 - I_2)^2 = 2K(U_P - U_N)^2 I_S - K^2(U_P - U_N)^4$$

Kvadratroden 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 kvadratroden.

$$I_1 - I_2 = K(U_P - U_N) \sqrt{\frac{2I_S}{K} - \frac{K(U_P - U_N)^2}{2I_S}}$$

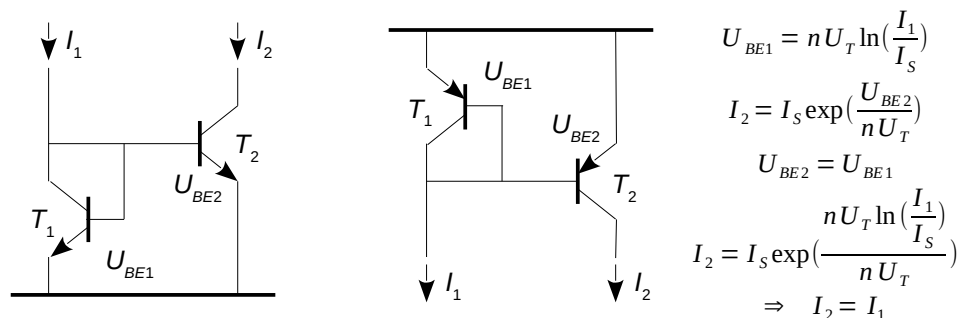
Leddene under kvadratroden er tæt på nul, idet de to transistorer deler strømmen I_S ligeligt og ved division med $U_P - U_N$ dannes et led der udtrykker ændringen i strømmen i drain som følge af en ændring i spændingen mellem indgangene og det giver transkonduktansen.

$$I_1 - I_2 \approx K(U_P - U_N) \sqrt{\frac{2I_S}{K}} \Rightarrow g_m = \frac{dI_D}{dU_{GS}} \approx \frac{\Delta I_D}{\Delta U_{GS}} = \frac{I_1 - I_2}{U_P - U_N} = \sqrt{2KI_S}$$

5.6.5 Strømspejl

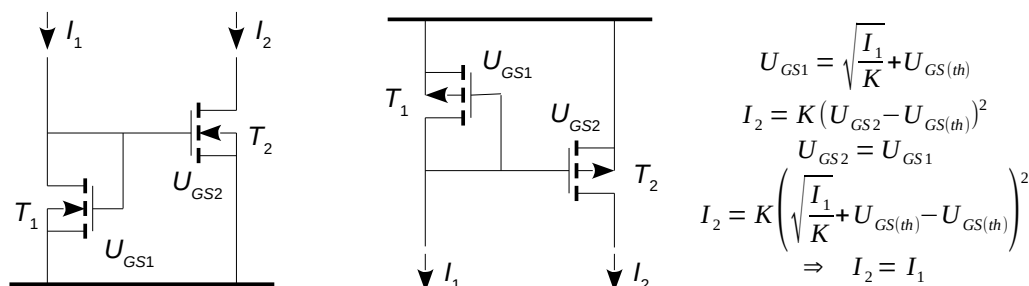
Differensen mellem de to strømme fra differentialtrinnet dannes af et strømspejl (*current mirror*) der herunder vises i sin mest enkle udførelse. Det er et populært trin i integrerede kredse hvor det benyttes både i operationsforstærkerens differentialtrin og i de øvrige kredsløb. Det ses også i kredsløb 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 351 – Et strømspejl udnytter transistorernes ens karakteristisk 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 nærmets 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 352 – 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 estimatet. 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.

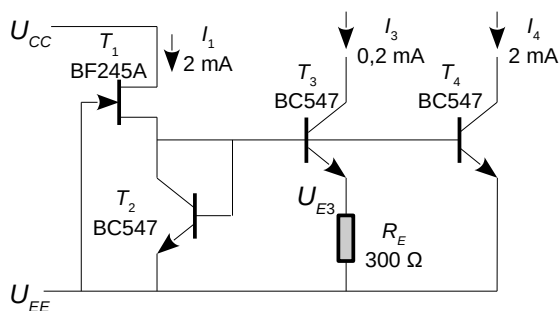
Ved BJT type PNP og P kanal MOSFET er ladbæreren huller hvis mobilitet er 2 til 4 gange lavere end elektronerne så de har en lavere grænsefrekvens. Den lavere mobilitet betyder dog blot at polen flytter ned i retning af 10 MHz og er derfor stadigvæk relativt højt oppe i frekvens.

5.6.6 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 operationsforstærker¹⁷⁰.

Stort set alle operationsforstærkere indeholder et kredsløb som holder de interne strømme relativt konstant¹⁷¹. En JFET er særdeles 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.



$$U_{E3} = n U_T \ln\left(\frac{I_1}{I_3}\right)$$

$$R_E = \frac{U_{E3}}{I_3}$$

Figur 353 – 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 $I_3 = 0,2$ 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.

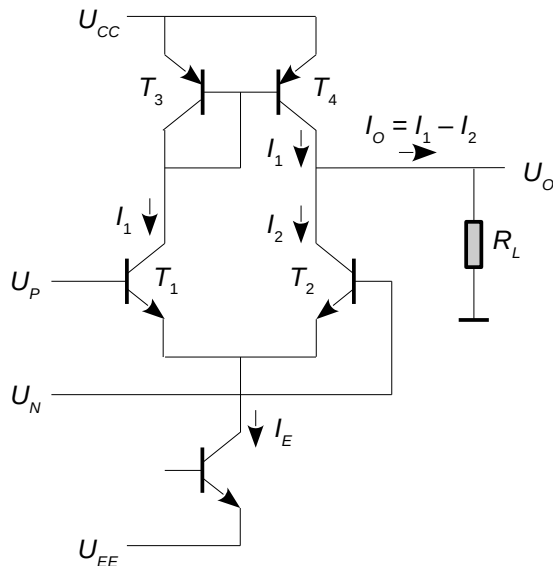
En ulempe ved opstillingen er at spændingen U_E har en positiv temperaturkoefficient på 0,35 %/K ved 60 mV som der ikke kompenseres for så for temperaturområdet -20°C til 85°C vil fejlen nå op i nærheden af $\pm 20\%$ i forhold til et design baseret på 30°C . En anden ulempe er at der ikke kan justeres 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.

¹⁷⁰CMRR (common-mode rejection ratio) angiver styrken af udgangssignalet fra operationsforstærkeren ved påvirkning af de to indgange med en harmonisk svingning. 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.

¹⁷¹En undtagelse er LM386 low-voltage power amplifier der er tiltænkt ukritiske formål som en forstærker for højttalere og motorer.

5.6.7 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:

$$A_{OL} = g_m R_L \Rightarrow A_{OL} = \frac{R_L}{2U_T} I_E$$

Forstærkningen er én når:

$$A_{OL} = 1 \Rightarrow g_m = \frac{1}{R_L}$$

Ligninger:

$$I_O = g_m (U_P - U_N)$$

$$U_O = R_L I_O$$

$$g_m = \frac{I_E}{2U_T}$$

$$U_T = \frac{kT}{q_0} \approx 26 \text{ mV}$$

Figur 354 – 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 strømgeneratoren 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 simple 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 åbensløjfe 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 opgave¹⁷².

Differentialtrinnet anvendes bredt indenfor analog signalbehandling hvor Gilbert cellen (*Gilbert cell*) skal nævnes¹⁷³. Den anvendes som fasedetektor i fase-låste oscillatorer og ved radiotransmission som en synkron detektor (mikser)¹⁷⁴.

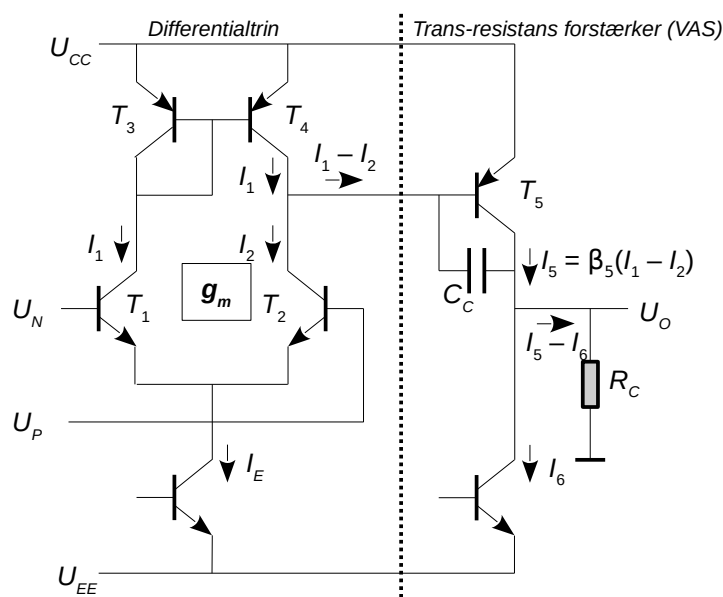
¹⁷²Se også min hjemmeside <http://www.torean.dk/artikel/publikationer.html> hvor du skal hente dokumentet "Gain Control".

¹⁷³Den blev opfundet af Howard Jones i 1963.

¹⁷⁴https://en.wikipedia.org/wiki/Gilbert_cell.

5.6.8 Trin 2 – Forstærker (VAS)

Vi er nu fremme ved operationsforstærkerens trin 2 der modtager strømmen fra differentialtrinet, 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:

$$A_{DC} = g_m \beta_5 R_C$$

AC forstærkning:

$$A_{OL} = \frac{g_m}{sC_C}$$

Gain-bandwidth product: $|A_{OL}| = 1$

$$GBP = \frac{g_m}{2\pi C_C}$$

Dominerende pol:

$$f_0 = \frac{GBP}{A_{DC}} = \frac{1}{2\pi\beta_5 R_C C_C}$$

Slew rate:

$$SR = \frac{I_E}{C_C}$$

Figur 355 – Kombinationen af et differentialtrin og et strømspejl giver en opstilling hvor indgangens differentielle spænding omsættes til en strøm på udgangen. En femte transistor kan give yderligere forstærkning af strømmen inden den føres til en belastningsmodstand og ender som en spænding.

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 endnu højere modstandsværdi bliver forstærkningen dog begrænset af transistorens indre modstand.

For at opnå stor DC forstærkning kan strømmen I_E i differentialtrinnets 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.

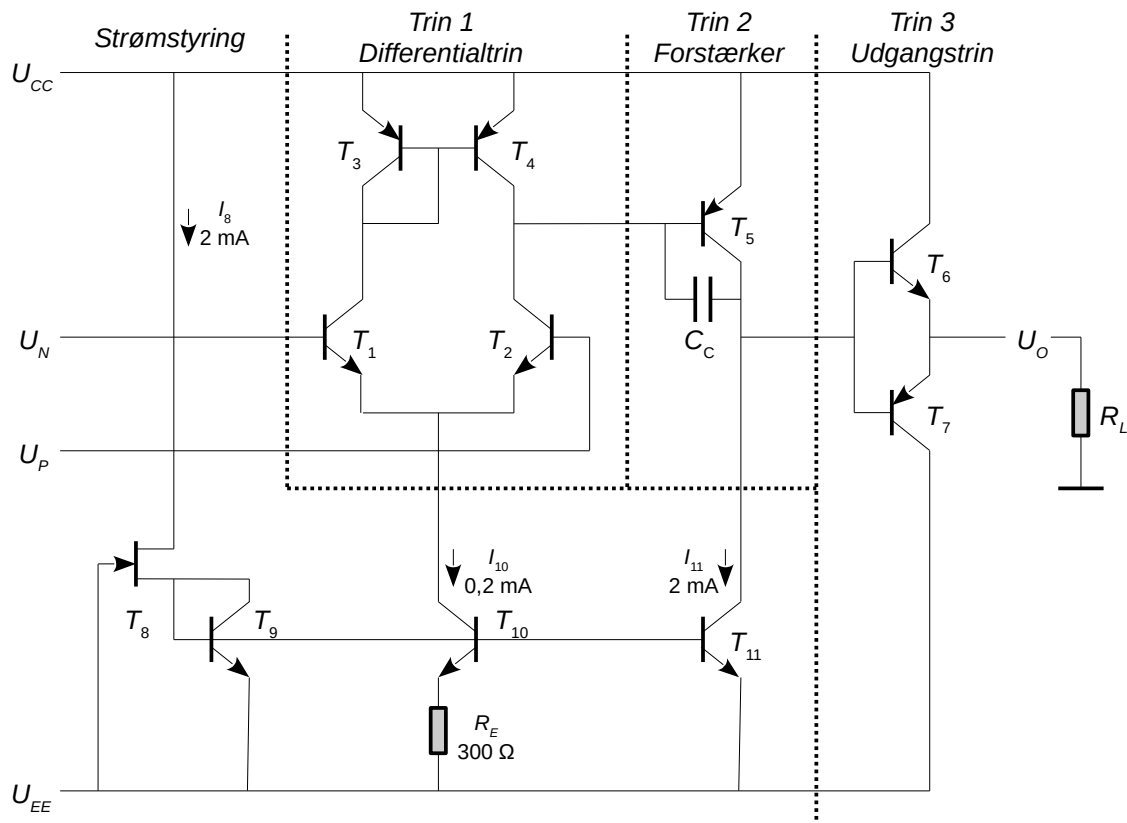
Basisstrømmen til T_5 kan kun løbe én vej og det kræver en ubalance i differentialtrinet ($I_2 > I_1$) for at optage basisstrømmen fra T_5 og trinnet må derfor køre en smule skævt, men det har ikke større betydning i praksis da emitterstrømmen I_E er meget større end basisstrømmen. Med MOSFET som forstærkende 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. For at flytte udgangens spænding skal kondensator C_C op- eller aflades og her er det vigtigt at $I_1 - I_2$ kan løbe begge veje. Strømmen kan maksimalt være $\pm I_E$ så der er en grænse for hvor hurtigt udgangsspændingen kan varieres og det beskrives ved slew rate specifikationen der uddybes på side 318.

5.6.9 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 den¹⁷⁵.

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 effektforstærkere.



Figur 356 – 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 nettet¹⁷⁶. Operationsforstærkere baseret på en kraftig udvidelse af det viste tre-trins kredsløb markedsføres som moduler for kommercielt salg¹⁷⁷.

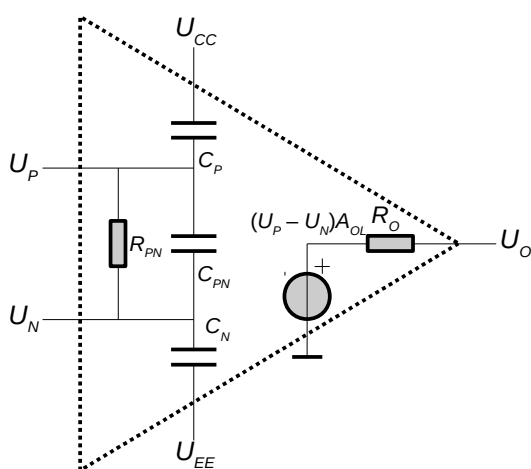
¹⁷⁵ http://www.ijetae.com/files/Volume3Issue12/IJETA_1213_17.pdf.

¹⁷⁶ http://www.sg-acoustics.ch/analogue_audio/discrete_opamps/.

¹⁷⁷ <http://www.johnhardyco.com/pdf/990.pdf>.

5.6.10 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 357 – 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 potentialet 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.

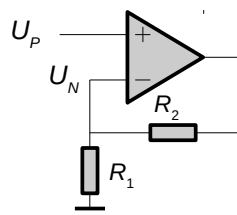
Udgangsmodstanden R_O er givet ved modstanden i punktet ved kollektor eller drain af T_5 divideret med værdien af strømforstærkningen for den aktive udgangstransistor (β_6 eller β_7). Modstanden r_O for T_5 og T_{11} kan beregnes fra h-parametermodellens h_{oe} eller transistorens Early Voltage som har 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.

5.6.11 Stabilitet

Det er ikke tilstrækkeligt at beskrive en reel operationsforstærker ved den dominerende pol alene for der er altid en eller flere højtliggende poler. Her modelleres operationsforstærkeren ved to poler. Ved tilbagekobling resulterer det i en anden-ordens overføringsfunktion hvor dæmpningen kan benyttes for at udlede et designkrav.



$$A_{OL} = \frac{\omega_0}{\omega_0 + s} \frac{\omega_1}{\omega_1 + s} A_{DC} \quad A_{CL} = \frac{A_{DC}}{1 + B A_{DC}} \frac{\omega_{CL}^2}{\omega_{CL}^2 + 2d \omega_{CL} s + s^2}$$

$$A_{CL} \xrightarrow{f < f_{CL}} \frac{A_{OL}}{1 + B A_{OL}} \xrightarrow{A_{OL} \rightarrow \infty} \frac{1}{B} \quad \omega_{CL} = 2\pi f_{CL} = \sqrt{(1 + B A_{DC}) \omega_0 \omega_1}$$

$$B = \frac{R_1}{R_1 + R_2} \quad d \approx \frac{1}{2\sqrt{(1 + B A_{DC}) \frac{\omega_0}{\omega_1}}}$$

$$f_0 < \frac{f_1}{4(1 + B A_{DC})} \xrightarrow{B A_{DC} > 1} f_0 < \frac{f_1 A_{CL}}{4 A_{DC}} \xrightarrow{A_{CL} = 1} f_0 < \frac{f_1}{4 A_{DC}}$$

Figur 358 – En tilbagekoblet operationsforstærker skal have en dominerende pol for stabilitet.

Operationsforstærkeren modelleres ved to poler hvor ω_0 er den dominerende pol og ω_1 er en højtliggende pol. Overføringsfunktionen for den tilbagekoblede forstærker 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 $(1 + B A_{DC})\omega_0 \omega_1$ der identificeres som kvadratet på systemets grænsefrekvens ved $\omega_{CL} = 2\pi f_{CL}$ ved at sammenligne med prototypen på et anden-ordens filter. Faktor $1 + B A_{DC}$ indføres i tælleren og giver et udtryk for DC forstærkningen A_{DC} . Koefficienten til s omskrives ved at gange og dele med ω_{CL} og resultatet giver et udtryk for dæmpningen.

$$A_{CL} = \frac{A_{DC}}{1 + B A_{DC}} \frac{(1 + B A_{DC})\omega_0 \omega_1}{(1 + B A_{DC})\omega_0 \omega_1 + \frac{\omega_0 + \omega_1}{\omega_{CL}} \omega_{CL} s + s^2} \Rightarrow$$

$$A_{CL} = \frac{A_{DC}}{1 + B A_{DC}} \frac{\omega_{CL}^2}{\omega_{CL}^2 + 2d \omega_{CL} s + s^2} \quad \text{hvor} \quad \omega_{CL} = 2\pi f_{CL} = \sqrt{(1 + B A_{DC})\omega_0 \omega_1}$$

$$d = \frac{\omega_0 + \omega_1}{2\omega_{CL}}$$

Det udnyttes her at den dominerende pol er meget lavere end den høje pol så tælleren kan reduceres til kun at bestå af den høje pol. Det antages desuden at $B A_{DC}$ er meget større end én at $1 + B A_{DC}$ kan reduceres til $B A_{DC}$. Begge antagelser er meget realistiske.

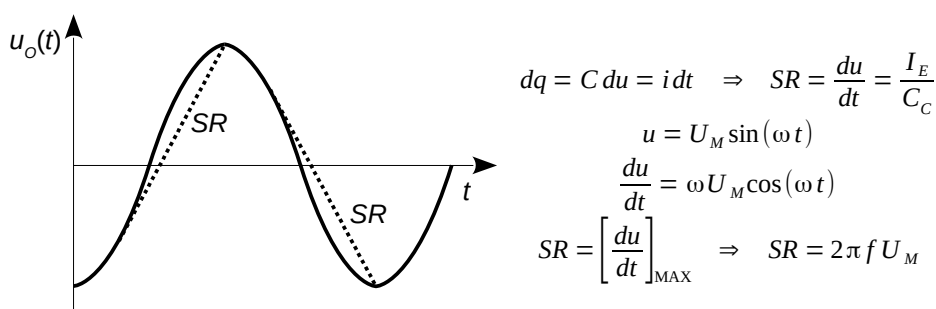
$$d = \frac{\omega_0 + \omega_1}{2\sqrt{(1 + B A_{DC})\omega_0 \omega_1}} \approx \frac{\omega_1}{2\sqrt{B A_{DC} \omega_0 \omega_1}} = \frac{1}{2\sqrt{B A_{DC} \frac{\omega_0}{\omega_1}}}$$

For at opnå stabilitet skal dæmpningen mindst være én for to reelle poler. Ved en svag grad af underdæmpning ($d = 0,5$) vil stepresponsen være svagt oscillerende.

$$d \approx \frac{1}{2\sqrt{B A_{DC} \frac{\omega_0}{\omega_1}}} \Rightarrow d^2 = \frac{\omega_1}{4 B A_{DC} \omega_0} \Rightarrow f_0 = \frac{f_1}{4 d^2 B A_{DC}} \Rightarrow \begin{aligned} f_0 &= \frac{f_1}{B A_{DC}} & d &= 0,5 \\ f_0 &= \frac{f_1}{4 B A_{DC}} & d &= 1 \end{aligned}$$

5.6.12 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 højst 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 359 – 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 kondensatorligningen 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 højeste 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 potentiale 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.

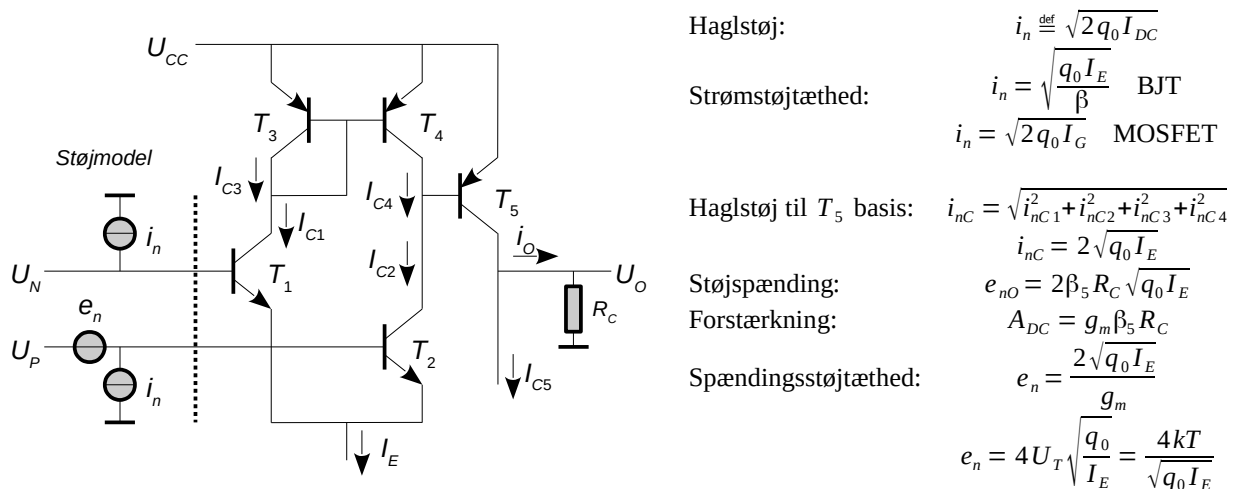
Denne introduktion til slew rate skyldes de debatter der fra tid til anden raser på internettet. Der fremsættes ofte ret fantastiske påstande om hvor stor en slew rate en forstærker behøver og det forsøges så skudt ned i debatspalterne med reference til hvor meget der er nødvendigt for at kunne gengive 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.

5.6.13 Egenstøj

Egenstøjen fra en operationsforstærker skyldes især haglstøj (*shot noise*). Strømmen bæres af elektroner 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 bipolære transistorer. Med cirka 100 Ω for småsignaltransistor er spændingsstøjtætheden e_n af størrelsesordenen 1 nV/ $\sqrt{\text{Hz}}$ og er kun betydende 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 støjkloder i både BJT og MOSFET, men resultatet er ret tæt på det der kan måles.



Figur 360 – Den ækvivalente spændingsstøjtæthed for et differentialtrin med bipolære transistorer kan beregnes alene ud fra arbejds punktet og antallet af aktive halvledere.

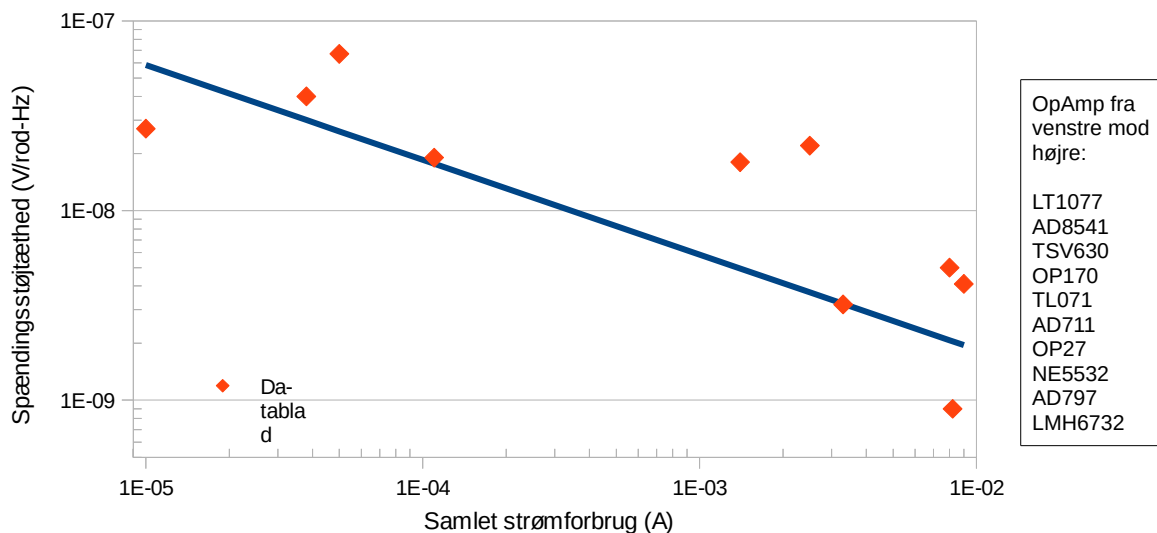
Strømstøjen i_n angiver strømmens fluktuation omkring den middelværdi I_{DC} der udgør strømmen i basis for den bipolære transistor. Basisstrømmen er $I_C/\beta = I_E/2\beta$, og med $I_E = 1$ mA beregnes en strømstøjtæthed på $i_n = 0,7$ pA/ $\sqrt{\text{Hz}}$ for en strømforstærkning på $\beta = 300$. 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 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} . Støjen henføres til indgangen som en ækvivalent støjkilde e_n ved at dividere med forstærkningen A_{DC} . For $I_E = 1$ mA bliver spændingsstøjtætheden $e_n = 1,3$ 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 $e_n = 18$ nV/ $\sqrt{\text{Hz}}$ og $i_n = 0,01$ pA/ $\sqrt{\text{Hz}}$, OP27 har $e_n = 3,2$ nV/ $\sqrt{\text{Hz}}$ og $i_n = 0,4$ pA/ $\sqrt{\text{Hz}}$ og den støjsvage AD797 har $e_n = 0,9$ nV/ $\sqrt{\text{Hz}}$ og $i_n = 2$ 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 μA til 9 mA. Det er antaget at der løber 3 % af det samlede strømforbrug i indgangstrinnet ved plotning af den blå linje som en værdi der gav rimelig visuel overensstemmelse med data.



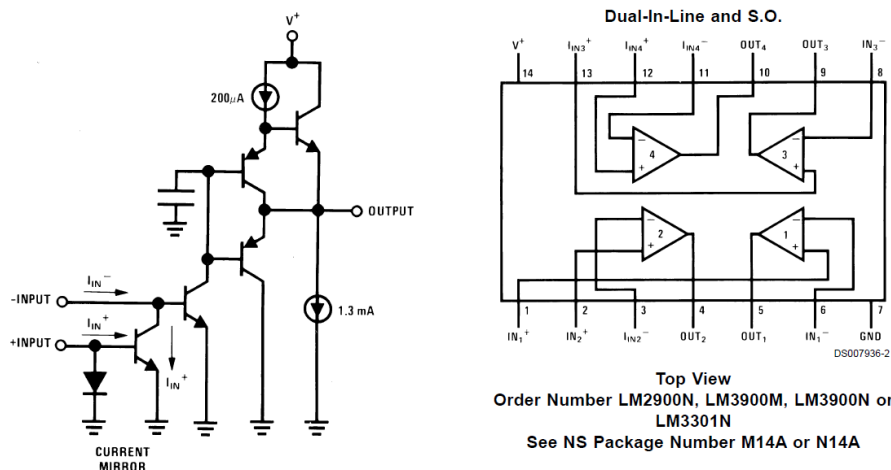
Figur 361 – Den estimerede spændingsstøjtæthed e_n (blå linje) sammenlignet med opgivne data (orange punkter) som funktion af det typiske strømforbrug for hele komponenten uanset antallet af operationsforstærkere der vises i parentes udfor den benyttede komponent.

De to operationsforstærkere der ligger under den blå linje har en strøm i indgangstrinnet på mere end 3 % af det samlede strømforbrug. For LT1077 er forklaringen at det samlede strømforbrug er på 10 μA 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 strømforbrug til hvert trin. De operationsforstærkere der ligger over den blå linje har en lavere strøm i indgangstrinnet end de benyttede 3 % af det samlede strømforbrug.

321

5.7.2 Transresistansforstærker

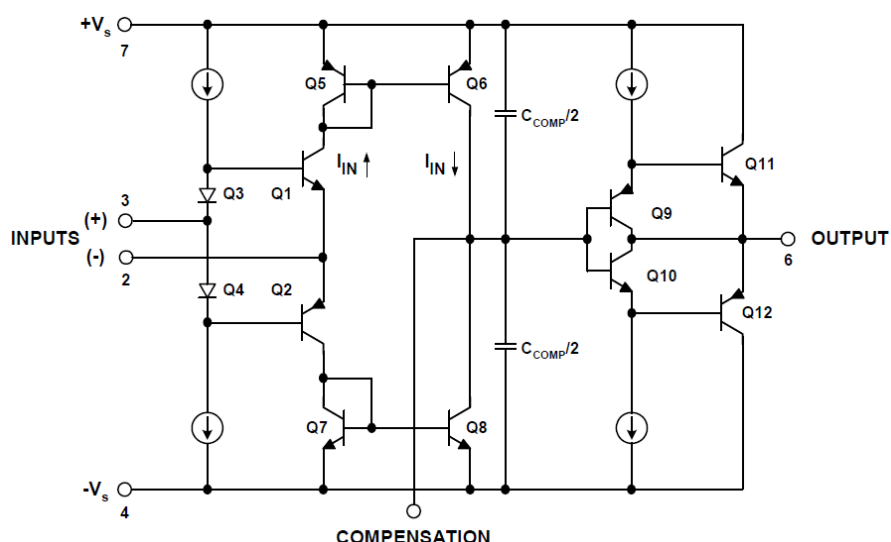
Et strømspejl på indgangen af en BJT giver transistoren en basisstrøm bestående af differensen mellem to indgangsstrømme og en buffer omsætter derefter kollektorens strøm til en spænding. Det kaldes for en *Norton amplifier* og den var overordentlig populær i en periode omkring 1980. Det er en alsidig 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 eksempler.¹⁷⁸



Figur 363 – En Norton-forstærker omsætter en strømdifferens til en spænding.

5.7.3 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 benyttes.¹⁷⁹



Figur 364 – I en current-feedback amplifier (som den nu udgåede AD846) afgør strømmen i den inverterende indgang hvor hurtigt operationsforstærkeren kan flytte udgangen (Analog Devices).

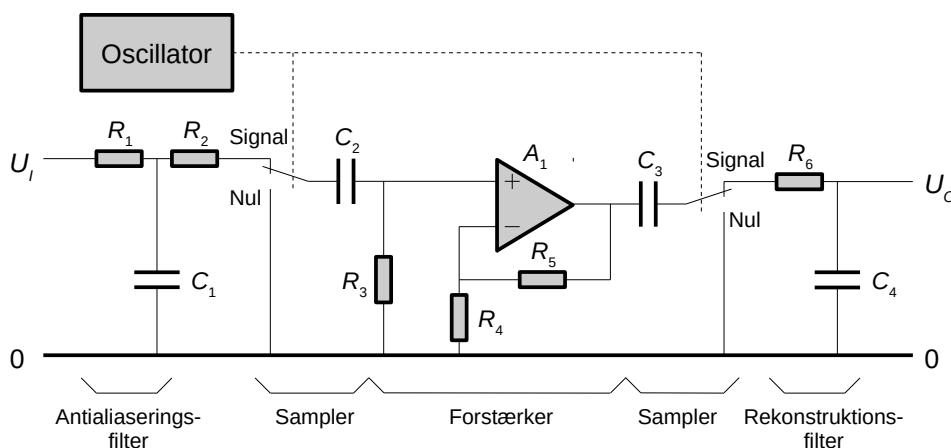
¹⁷⁸<http://www.ti.com/lit/ds/symlink/lm3900-n.pdf>.

¹⁷⁹<http://www.analog.com/static/imported-files/tutorials/MT-034.pdf>.

5.7.4 Chopperstabiliseret forstærker

Enhver forstærker har en DC forskydning af signalet på grund af indgangstrinnets fejlspænding 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 termokobler 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 fejlspænding måles af en mikroprocessor eller et potentiometer drejes manuelt indtil udgangen står på nul. Ud over besværet 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 fejlspænding 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 365 – 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 ovenfor. En moderne *chopper stabilized amplifier* vil benytte en langt mere avanceret teknik, men det viste trin giver en god introduktion til idéen.^{180,181}

- 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 , fejlspændingen 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 samlet 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.

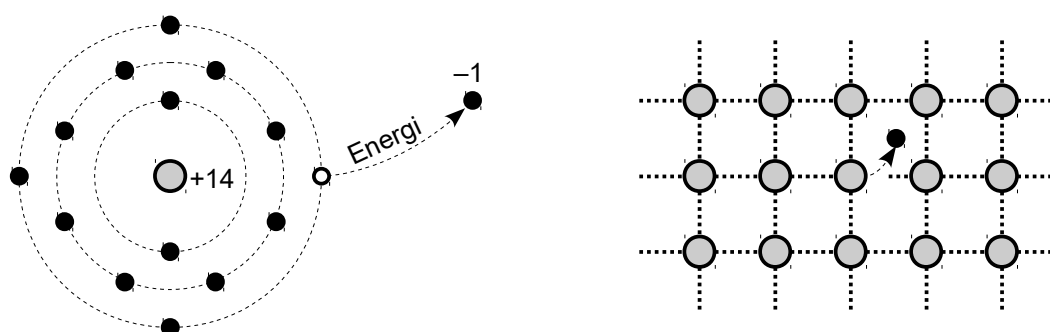
¹⁸⁰ <http://www.analog.com/media/en/training-seminars/tutorials/MT-055.pdf>.

¹⁸¹ http://www.analog.com/media/en/technical-documentation/data-sheets/ADA4051-1_4051-2.pdf.

5.8 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 udadtil neutralt. Elektronerne tiltrækkes af kernens positive ladning og holdes derved i en balance mellem en elektrostatiske 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 amorge 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 366 – 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 4 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.

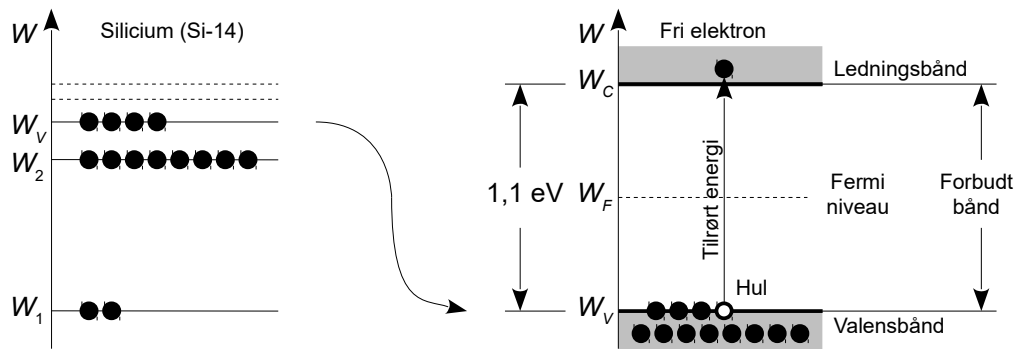
For et atom i lang afstand fra andre atomer er elektronens bane om kernen veldefineret og svarer til specifikke energiniveauer. For atomer arrangeret i et krystalgitter vil de yderste skaller derimod påvirkes af de andre atomers nærhed og i stedet for en skarp energiopdeling komprimeres deres skaller 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 anslået. 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 anslået atom. Ved et stort antal anslåede atomer vil tætheden af fotoner aftage eksponentielt efter at den eksterne påvirkning er 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 eV 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.¹⁸²

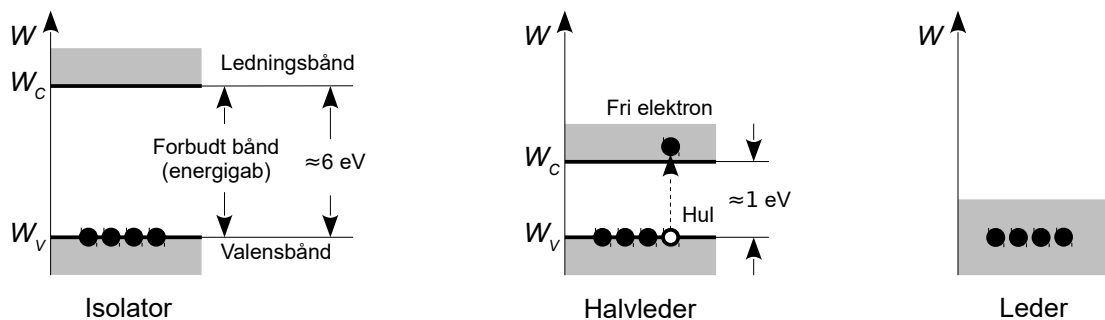
¹⁸²En elektron modtager energien 1 eV hvis den accelereres af et spændingsfald på 1 V. Elektronen har ladningen $q_0 = 1,6 \cdot 10^{-19}$ C så en energi på 1 eV svarer til $1,6 \cdot 10^{-19}$ J.

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 367 – Energimodel af grundstoffet 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 fri 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.



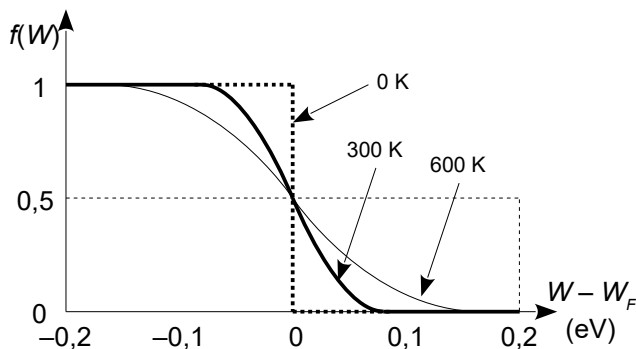
Figur 368 – Materialer kan opdeles i isolatorer, halvledere og ledere ud fra værdien af den energi der kræves for at elektroner kan passere det forbudte bånd og springe op til ledningsbåndet.

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.

5.8.1 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 specificeret 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:

$$n = \int_{W_C}^{\infty} \rho(W) f(W) dW$$

$$n = N_C \exp\left(-\frac{W_C - W_F}{kT}\right)$$

$$p = N_V \exp\left(-\frac{W_F - W_V}{kT}\right)$$

$$np = n_i^2 = N_C N_V \exp\left(-\frac{W_C - W_V}{kT}\right)$$

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 369 – 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 ligger midt mellem valens- og ledningsbåndene for den rene halvleder 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).

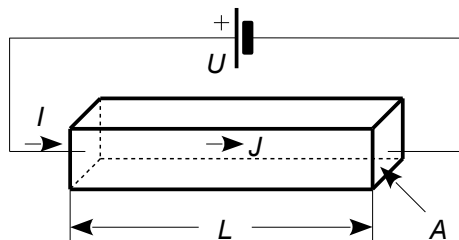
Den energi der mindst kræves for at slå en elektron fri at atomet og op til ledningsbåndet er givet ved differensen mellem energiniveauerne for ledningsbåndets W_C og valensbåndet W_V . For silicium kræves en energi på $W_G = 1,125$ eV og for germanium er det 0,663 eV. Der er derfor næsten ingen frie 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 slynget op i ledningsbåndet. Når en elektron forlader atomet efterlades en tom plads med en positiv ladning hvilket kaldes for et hul. Hvis pladsen derefter optages af en elektron fra et naboatom 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 ladningsbærere 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 .

5.8.2 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 mobiliteten μ_p og μ_n der udtrykker hvor villigt henholdsvis huller og elektroner flytter sig i et elektrisk felt.¹⁸³



$$\begin{aligned}
 J_p &= q_0 p \mu_p E & \Rightarrow & J = J_p + J_n = \sigma E \\
 J_n &= q_0 n \mu_n E & \sigma &= q_0 p \mu_p + q_0 n \mu_n \\
 E &= U/L & \Rightarrow & I = \sigma \frac{U}{L} A \Rightarrow U = \frac{L}{\sigma A} I \\
 I &= JA
 \end{aligned}$$

$$\begin{aligned}
 R &= \frac{L}{\sigma A} = \rho \frac{L}{A} & \mu_p &\approx 0,048 \text{ m}^2/\text{Vs} \\
 \rho &= \frac{1}{q_0(p\mu_p + n\mu_n)} & \mu_n &\approx 0,136 \text{ m}^2/\text{Vs}
 \end{aligned}$$

Figur 370 – 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.

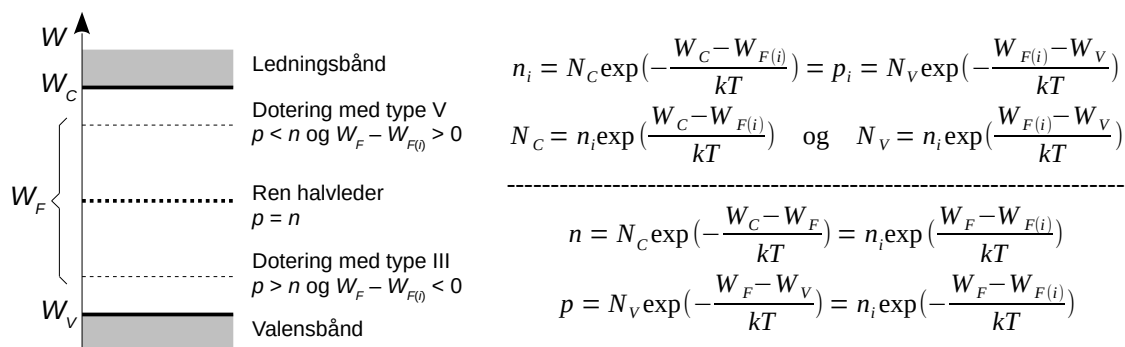
5.8.3 Den dotede halvleder

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.

¹⁸³Mobiliteten 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 371 – Ved den rene halvleder (intrinsic semiconductor) er Fermi niveauet W_F placeret midt i det forbudte bånd mellem valens- og ledningsbåndene ($W_F = W_{F(0)}$) 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 $W_F - W_{F(0)}$.

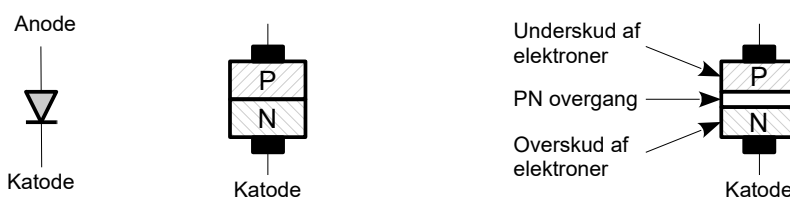
Med en forskydning af Fermi-niveauet på $W_F - W_{F(0)} = 0,1$ eV vil koncentrationen af elektroner øges fra det naturlige niveau på $n_i \approx 1 \cdot 10^{16} \text{ m}^{-3}$ med faktoren $\exp(3,86) \approx 50$ til $n \approx 50 \cdot 10^{16} \text{ m}^{-3}$. Tilsvarende aftager antallet af huller med faktoren $\exp(-3,86) \approx 1/50$ så elektronkoncentrationen dominerer. Den specifikke modstand bliver $\rho \approx 100 \text{ } \Omega\text{m}$ så modstandsværdien af en terning med 1 mm sidelængde vil aftage fra de tidligere beregnede 3,4 M Ω hvor elektroner og huller begge varetog transport af strømmen til cirka 100 k Ω 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 mangle 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 en halvleder udadtil 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 .

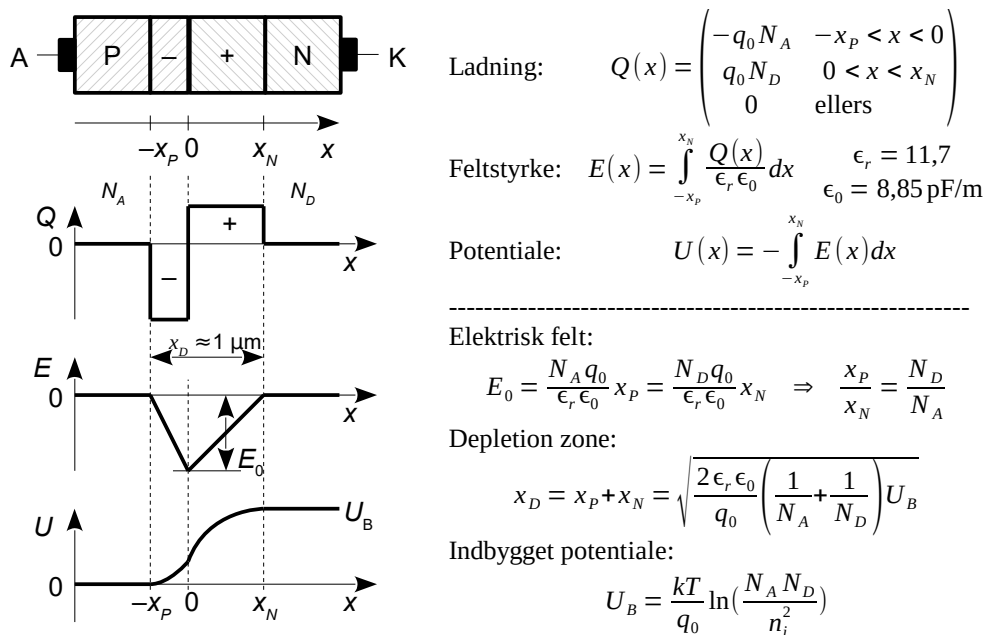
5.8.4 PN overgang

En diode består af to stykker silicium i elektrisk kontakt med hinanden. Det ene stykke er doteret med et antal acceptor-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 372 – 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 modsatrettede kræfter på elektronerne. Området ved PN overgangen er nu stort set tømt for frie ladningsbærere og det kaldes derfor *depletion layer* på engelsk.

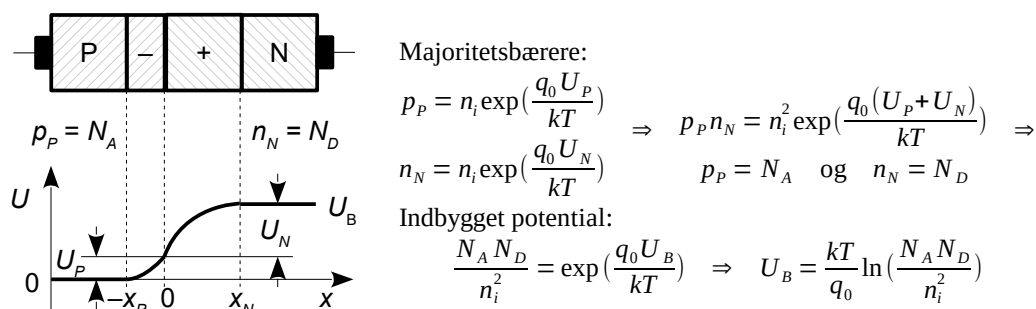


Figur 373 – Diodens interne forhold kan udledes fra ladningsfordelingen i de to halvledermaterialer og her antages en konstant doteringsprofil hvor N_A og N_D er uafhængige af positionen i materialet. Den elektriske feltstyrke $E(x)$ langs med x -aksen beregnes fra Gauss lov ved at integrere ladningen $Q(x)$ over x . Feltstyrken er nul i første og sidste del af halvlederen for halvlederen er elektrisk neutral uanset om den er ren eller doteret. Med en konstant koncentration af acceptor 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 acceptor- og donorkoncentrationer samt det indbyggede potentiale og udledes ikke her. [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 $x_D = 1,2 \text{ } \mu\text{m}$. Da $x_P/x_N = 0,5$ er det neutrale lags længder $x_P = 0,4 \text{ } \mu\text{m}$ og $x_N = 0,8 \text{ } \mu\text{m}$ og feltstyrkens maksimum bliver $E_0 = 10 \text{ V/m}$.

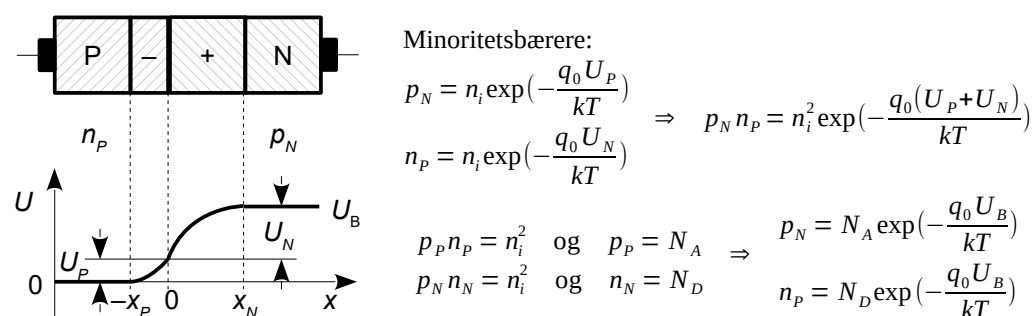
Længderne af områderne x_P og x_N er funktion af koncentrationen ved N_A og N_D og da der er lige mange positive og negative fastsiddende ioner på de to sider af PN overgangen bliver længderne funktion af koncentrationerne. Illustrationen viser den dobbelte koncentration af acceptor så x_P er halvt så lang som x_N . Det er muligt at foretage en individuel beregning af længderne, men jeg har valgt kun at gengive formelen for PN overgangens længde x_D og den interesserede 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 potentiale U_B (*build-in potential*) og som det vises i det følgende afsnit er det fundamentalt for diodens funktion.



Figur 374 – Det indbyggede potentiale 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 acceptor 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 hvorimod elektroner 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 elektroner n_N majoritetsbærere i N halvlederen og hullet p_N er minoritetsbærer.



Figur 375 – Koncentrationen af minoritetsbærere beregnes fra det indbyggede potentiale 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.

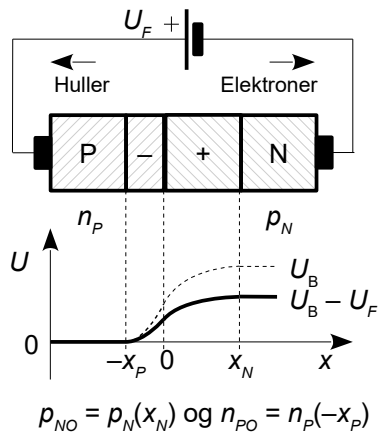
5.8.5 Strømmen i halvlederen

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 potentiale 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:

$$p_{NO} = N_A \exp\left(-\frac{q_0 U_B}{kT}\right) = \frac{n_i^2}{N_D} \quad \text{da} \quad \frac{N_A N_D}{n_i^2} = \exp\left(\frac{q_0 U_B}{kT}\right)$$

$$n_{PO} = N_D \exp\left(-\frac{q_0 U_B}{kT}\right) = \frac{n_i^2}{N_A}$$

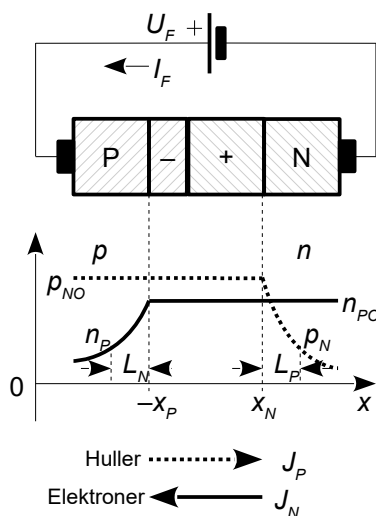
Diffusionsstrøm:

$$p_N^D = p_N - p_{NO} = \frac{n_i^2}{N_D} [\exp\left(\frac{q_0 U_F}{kT}\right) - 1]$$

$$n_P^D = n_P - n_{PO} = \frac{n_i^2}{N_A} [\exp\left(\frac{q_0 U_F}{kT}\right) - 1]$$

Figur 376 – Koncentrationen af de termisk genererede minoritetsbærere hæves af den eksterne spændingskilde fra et niveau givet af det indbyggede potentiale og det kan igen udtrykkes ved den rene halvleders n_i samt doteringen med N_A og N_D der netop skaber det indbyggede potentiale. 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:

$$p_N^D = \frac{n_i^2}{N_D} [\exp\left(\frac{q_0 U_F}{kT}\right) - 1] \exp\left(-\frac{x_N - x}{L_p}\right), \quad x > x_N$$

$$n_P^D = \frac{n_i^2}{N_A} [\exp\left(\frac{q_0 U_F}{kT}\right) - 1] \exp\left(-\frac{x_P + x}{L_n}\right), \quad x < -x_P$$

Strømtæthed:

$$J_p = \left(-q_0 D_p \frac{dp_N^D}{dx}\right)_{x=-x_P} = \frac{q_0 D_p n_i^2}{L_n N_A} [\exp\left(\frac{q_0 U_F}{kT}\right) - 1]$$

$$J_n = \left(q_0 D_n \frac{dn_P^D}{dx}\right)_{x=x_N} = \frac{q_0 D_n n_i^2}{L_p N_D} [\exp\left(\frac{q_0 U_F}{kT}\right) - 1]$$

Samlet strømtæthed:

$$J_F = J_p(x_N) + J_n(-x_P) \Rightarrow J_F = J_s [\exp\left(\frac{q_0 U_F}{kT}\right) - 1], \quad J_s = q_0 \left(\frac{D_p n_i^2}{L_p N_D} + \frac{D_n n_i^2}{L_n N_A} \right)$$

Figur 377 – Diffusionsstrømmens koncentration af minoritetsbærere aftager eksponentielt udenfor PN overgangen og det beskrives ved diffusionslængden L for hver af de to områder og fortegnet fjernes som resultat af differentiationen.

Den samlede strømtæthed J i halvlederen er summen af tæthederne for elektroner og huller. For at beregne strømtætheden 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 strømtætheden 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
$I_F = I_S [\exp(\frac{q_0 U_F}{kT}) - 1]$	$U_B = \frac{kT}{q_0} \ln(\frac{N_A N_D}{n_i^2})$
$I_S = q_0 A \left(\frac{D_P n_i^2}{L_P N_D} + \frac{D_N n_i^2}{L_N N_A} \right)$	$D_P = \mu_P kT / q_0$ $\mu_P \approx 0,048 \text{ m}^2/\text{Vs}$ $D_N = \mu_N kT / q_0$ $\mu_N \approx 0,136 \text{ m}^2/\text{Vs}$ $N_C \approx 2,84 \cdot 10^{25} \text{ m}^{-3}$ $N_V \approx 3,08 \cdot 10^{25} \text{ m}^{-3}$ $n_i \approx 1,07 \cdot 10^{16} \text{ m}^{-3}$ $\tau \approx 1 \text{ ns} \dots 1 \text{ ms}$
$n_i = \sqrt{N_C N_V} \exp(-\frac{W_G}{2kT})$	$L_P = \sqrt{D_P \tau_P}$ $L_N = \sqrt{D_N \tau_N}$
$W_G = W_C - W_V \approx 1,8 \cdot 10^{-19} \text{ J}$	

Figur 378 – Strømmen I_F i diodens lederetning som funktion af den eksterne spænding U_F , med parametrene mætningsstrømmen I_S , elektronens ladning q_0 , Boltzmanns konstant k og den absolutte 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 . Mobiliteten 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).¹⁸⁴

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 $\tau = 1 \text{ } \mu\text{s}$ er $L_P = 35 \text{ } \mu\text{m}$ og $L_N = 59 \text{ } \mu\text{m}$. 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 $1 \text{ } \mu\text{m}$ hvor arealet er $A = 10^{-6} \text{ m}^2$. Det indbyggede potentiale bliver på $U_B = 0,61 \text{ V}$.

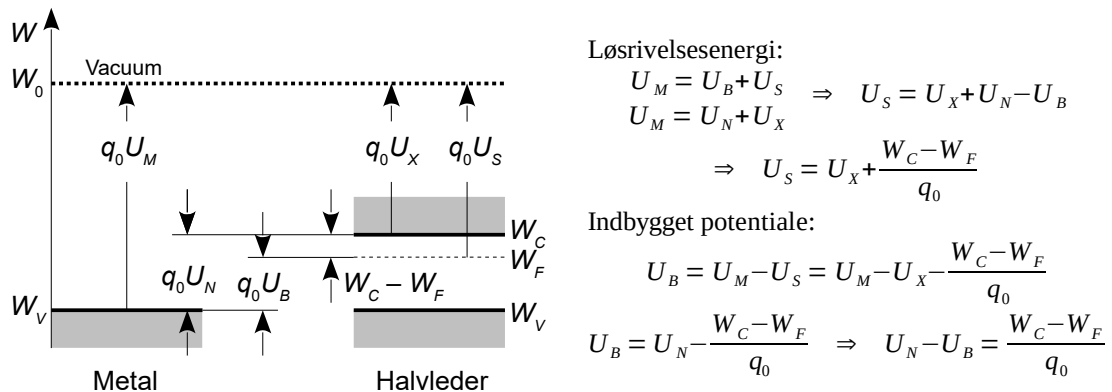
5.8.6 Shottky-diode

Det er ikke kun PN overgangen der kan fungere som diode for en metal-halvleder overgang giver også en ensrettende virkning. I modsætning til diodens PN overgang hvor ladningsbæreren går fra at være majoritetsbærende til minoritetsbærende så benytter Shottky-dioden kun én type af ladningsbærere som er majoritetsbærere (elektroner eller huller, men ikke dem begge). Fraværet af rekombination gør den hurtigere end en diode baseret på en PN overgang.¹⁸⁵

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 halvledere. Der er dog en anden mekanisme som giver en modsat rettet strøm af elektroner. Metallens 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.

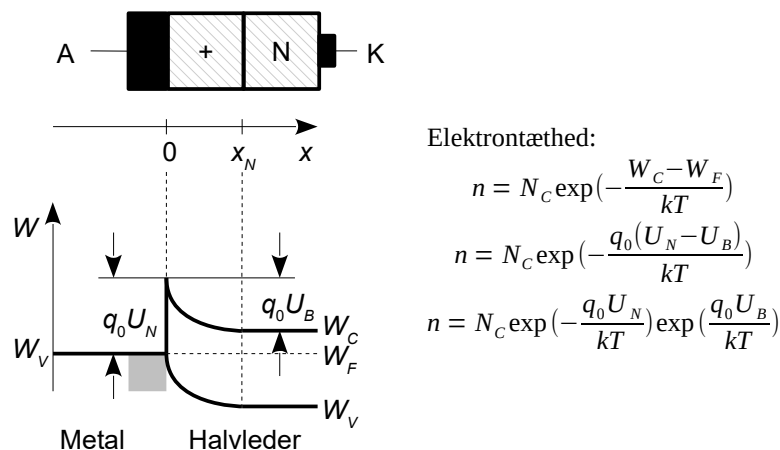
¹⁸⁴Data er fra referencen [CC] side 41, 45, 75, 89, 97, 101, 105, 155 og 156.

¹⁸⁵De første halvledere blev opbygget af krystallet blysulfid som en metalnål blev presset ned imod hvorved man eksperimentelt havde fundet frem til at komponenten kunne ensrette et vekselstrøm. Det var en vigtig komponent i de første radiomodtagere inden udviklingen radorøret muliggjorde en mere stabil funktion af ensretteren.



Figur 379 – De aktuelle energiniveauer omfatter de potentielle energier $q_0 U_M$ for metallet og $q_0 U_X$ for halvlederen hvilket efterlader en differens $q_0 U_N$ 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 potentiale 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 328).

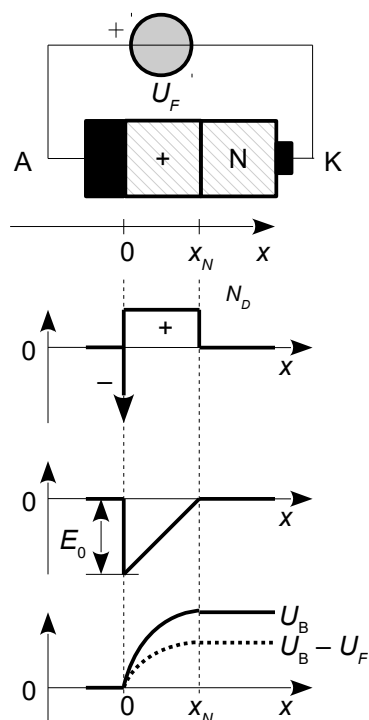


Figur 380 – Elektrontætheden bestemmes af to materialeparametre N_C og U_N samt det indbyggede potentiale U_B og for en halvleder i termisk ligevægt er Fermi-niveauet konstant.

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 potentiale 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 $C N_C$ kan skrives som $A_R T^2$, hvor A_R er en konstant ved de givne materialer (*Richardson constant*) og T er den absolute temperatur. Relationen er identisk med formelen for strømmen i en PN halvleder hvor de eneste forskelle er mætningsstrømmen I_s og det lavere spændingsfald i lederetningen U_F .



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:

$$I_F = A J_F$$

$$I_F = I_S \left[\exp\left(\frac{q_0 U_F}{kT}\right) - 1\right]$$

$$I_S = A A_R T^2 \exp\left(-\frac{q_0 U_B}{kT}\right)$$

$$A_R \approx \begin{pmatrix} 130 \text{ A/K}^2 \text{ cm}^2 & \text{P} \\ 250 \text{ A/K}^2 \text{ cm}^2 & \text{N} \end{pmatrix} \text{ Si ved } 20^\circ \text{C}$$

Figur 381 – 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 lederetningen skal fastholdes ved en lavere værdi af spændingen over dioden U_F så må mætningsstrømmen stige. Det er præcis hvad der sker ved at erstatte en almindelig diode med en Schottky-diode. Det lavere spændingsfald i lederetningen resulterer i en højere værdi af strømmen i spærreretningen som det vises herunder ved et taleksempel.

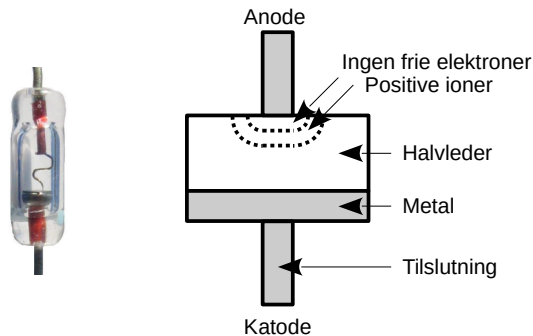
Hvis strømmen i diodens lederetning fastholdes på 1 mA så vil en siliciumdiode med et spændingsfald på 0,6 V have en mætningsstrøm på 100 fA og en Schottky-diode med et spændingsfald på 0,3 V vil have en mætningsstrøm på 10 nA. Det betyder at mætningsstrømmen øges med faktor 10^5 når spændingen over 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 lederetningen er mindre end spændingsfaldet for dioder af silicium så Schottky-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 spærre.¹⁸⁶

¹⁸⁶Mætningsstrømmen er $I_S \leq 1$ mA for Schottky-dioden 1N5819.

5.8.7 Point-contact diode

Siden 1874 har man kendt til dioder baseret på et krystal af blyulfid PbS (eng. *galena*) dog uden at vide hvordan det fungerede. Ferdinand Braun opdagede at strømmen kun kunne løbe i den ene retning hvis en metalledning blev presset ned mod overfladen af krystallet (*point contact diode*) og det blev brugt omkring 1900 i radiomodtagere for detektion af trådløs telegrafi.



Figur 382 – Diodens opbygning.

Krystallets atomer udstøder af og til en elektron på grund af omgivelsestemperaturens energi. Disse elektroner kan bevæge sig frit i krystallet og medvirker til at give det en ledningsevne. Det vil sige at det elektriske felt fra et eksternt batteri kan få elektronerne i bevægelse. Fra tid til anden falder en elektron på plads i et andet atom der mangler en elektron hvorved dens energi afgives. Denne konstante frigivelse og indfangning af elektroner opretholder et konstant niveau af frie elektroner og da der produceres flere elektroner ved stigende temperatur vil halvlederen være en bedre elektrisk leder ved forhøjet temperatur. Desuden kan lys slå elektroner fri.

Tilledningens elektroner har lav energi da et metal altid har frie elektroner. De energirige elektroner i halvlederen bevæger sig derfor mod metallederen hvor den overskydende energi kan afgives. De efterlade atomer i halvlederen mangler en elektron og det danner positive ioner. Der opbygges et elektrisk felt mellem metallederen og ionerne som bremser den fortsatte bevægelse af elektroner og der indstilles en balance hvor der ikke er elektroner i bevægelse.

Et eksternt batteri med plus til metallederen vil tiltrække de overskydende elektroner så det elektriske felt mindskes og elektronerne kan bevæge sig. De elektroner der forlader halvlederen erstattes af nye der leveres af batteriet. *Dioden leder den elektriske strøm.* Hvis batteriets poler ombyttes vil det forstærke det elektriske felt og blokerer for bevægelsen. *Dioden spærrer.* Der løber en svag strøm da der hele tiden genereres frie elektroner som tiltrækkes af batteriets positive pol, Det er diodens lækstrøm (*leakage current*).

Tilslutningen til diodens katode er også en diode, men dens areal er meget større end den tynde tråd så lækstrømmen bliver stor nok til at modsvare strømmen i diodens lederetning.

5.8.8 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 leder lige godt i begge retninger. En kraftig dotering af halvlederen sænker det indbyggede potentiale så der kun er et ganske ringe spændingsfald i lederetningen og mætningsstrømmen øges samtidigt så strømmen heller ikke bliver begrænset i spærreretningen. 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+).

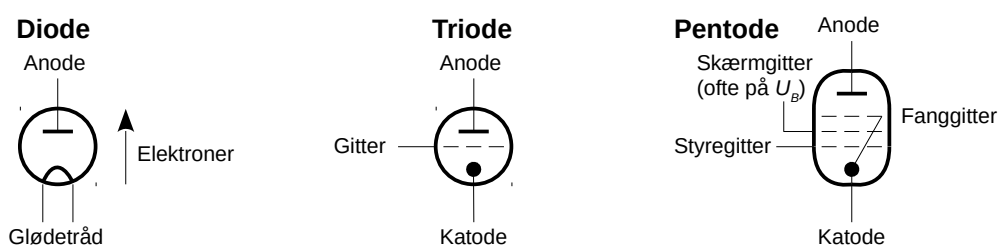
5.9 Elektronrøret

Mange vil nok mene at elektronrøret har udspillet sin rolle og at en omtale hører til ved museets udstillingsmontre, men elektronrøret bruges i dag ved radiotransmission og røntgenstråler (x-ray), ved mikrobølgeovne og langtrækkende radar samt videnskabelig brug. Transistoren egner sig kun til nogle få kilowatt og højere effekt kræver parallelkobling af flere trin. Jeg vil begrænse omtalen til analog signalbehandling ved lave effekter og der sigtes kun efter en introduktion.



Figur 383 – Elektronrør benytter indkapsling i en lufttom glaskolbe og katoden skal opvarmes for at udsende elektroner (Wikipedia).

Elektronrøret kaldes også for radorør eller blot rør og på engelsk *vacuum tube* eller *valve*. Det har 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). Det kræver at elektronrøret er pumpet luftomt så elektronerne ikke kolliderer med de meget tungere luftmolekyler. Med et negativt potentiale på anoden blev elektronerne frastødt så der ikke kunne løbe en strøm. Røret kunne derfor ensrette en vekselspænding og Fleming benyttede det i 1904 for en stabil detektion af radiotransmission med skibe.

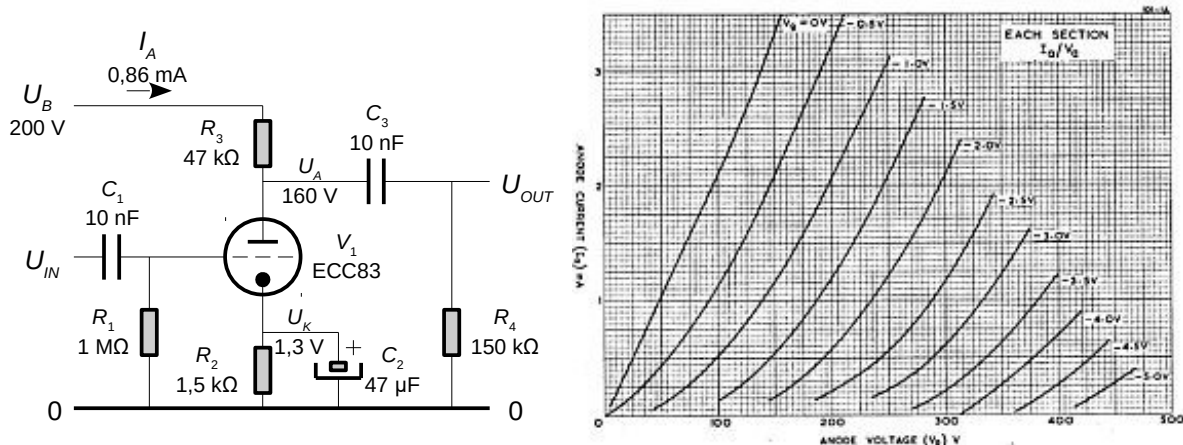


Figur 384 – Elektronrøret kan kun lede strømmen i én retning og er derfor egnet til ensretning. Med et gitter kan strømstyrken reguleres og flere gitre indføres for at forbedre funktionen. Flertallet af elektronrør har indirekte opvarmet katode så glødetråden er underforstået i diagrammet.

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 og det er for første gang muligt at forstærke et signal. 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.

5.9.1 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. Kondensatorerne blokerer for DC og C_2 holder katoden fast overfor AC signaler så der undgås tilbagekobling (se side 297).



Figur 385 – En forstærker med autobias (Philips 1970) og rørets udgangskaraktistik for den ene halvdel af det dobbelte rør ECC83 der kaldes 12AX7 i amerikansk litteratur (Mazda 1963).

Litteraturen benytter stejleheden $S = g_m$ og forstærkningen $\mu = g_m r_o$ hvor r_o er den indre modstand. Ved $U_A = 250$ V og $I_A = 1,2$ mA oplyses $g_m = 1,6$ mS, $r_o = 62,5$ kΩ og $\mu = 100$ for ECC83. Ifølge databladet giver den viste opstilling 34 ganges 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.

$$g_m = \frac{dI_A}{dU_G} \approx \frac{I_{A2} - I_{A1}}{U_{G2} - U_{G1}} \quad r_o = \frac{dU_A}{dI_A} \approx \frac{U_{A4} - U_{A3}}{I_{A4} - I_{A3}} \quad A = g_m R_L$$

$$R_L = r_o || R_3 || R_4$$

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.

Der er en kapacitet fra anode til styregitret og det begrænser funktionen ved høje frekvenser så et skærmgitter (*screen grid*) indbygges. Skærmgitret er forbundet til en positiv spænding så det vil opfange en del af elektronerne fra katoden og bærer en strøm. 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, men forbundet til katoden og det negative potentiale tvinger elektronerne tilbage til anoden. Elektronrør med fem elektroder kaldes for en pentode.

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 hævdes det at røret er uovertruffen indenfor lydgenjivelse.^{187,188}

187 Se fx <http://www.john-a-harper.com/tubes201/>.

188 Se fx <http://www.thevalvepage.com/index.shtml>.

5.10 Opgaver

Dioden.

1 – Vis at idealitetsfaktoren n kan findes fra diodens formel på side 253 ved at måle ledespændingen ved to forskellige strømværdier (U_{F1} og U_{F2} ved henholdsvis I_{F1} og I_{F2}). **Tip:** Benyt formelen 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.

$$n = \frac{U_{F1} - U_{F2}}{U_T \ln\left(\frac{I_{F1}}{I_{F2}}\right)}$$

2 – Bestem idealitetsfaktoren for dioden 1N4148 hvis data vises side 255. **Svar:** $n \approx 1,7$ ved $I_{F1} = 1$ mA og $I_{F2} = 2$ mA hvor der ved 25°C aflæses $U_{F1} = 0,585$ V og $U_{F2} = 0,615$ V. Bemærk at aflæsningen fra en kurve er meget usikker.

3 – Vis at mætningsstrømmen I_S kan udledes fra diodens formel på side 253 til udtrykket herunder når I_P ignoreres og $\exp(U_F/U_T)$ er så stor at "minus én" kan ignoreres. Brug kurven på side 255 til at finde værdien ved $I_F = 1$ mA, 25°C og $n = 1,7$. **Svar:** $I_S \approx 2$ nA.

$$I_S = \frac{I_F}{\exp\left(\frac{U_F}{n U_T}\right)}$$

4 – Vis at spændingen over en diode ændres med et fast bidrag ved en relativ strømförändring fra I_{F1} til I_{F2} som vist herunder.

$$\Delta U_F = U_{F2} - U_{F1} = n U_T \ln\left(\frac{I_{F2}}{I_{F1}}\right)$$

5 – For 1N4148 beregnes ændringen i spændingen ΔU_F over dioden når strømmen øges fra $I_{F1} = 1$ mA til $I_{F2} = 2$ mA samt fra $I_{F1} = 1$ mA til $I_{F2} = 10$ mA. **Svar:** $\Delta U_F = 30$ mV ved fordobling og $\Delta U_F = 100$ mV ved dekadet ændring.

6 – For kurven på side 254 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 .

7 – Beregn temperaturkoefficienten TC for 1N4148 på side 255 ved strømstyrkerne 1 mA og 10 mA og sammenlign data. **Svar:** $TC = -1,8$ mV/ $^\circ\text{C}$ og $-1,6$ mV/ $^\circ\text{C}$.

8 – For dioden 1N4148 beregnes DC modstanden R_D fra illustrationen side 262 ved brug af figuren side 255. Aflæs spændingsfaldet U_{F1} ved en strøm i lederetningen på 1 mA fra figuren på ved en temperatur på 25°C . Beregn den forventede spænding over dioden U_{F10} ved 10 mA ved brug af ΔU_F fra en tidligere opgave. Aflæs spændingen U_{D10} ved denne strøm og beregn DC modstanden R_D . **Svar:** $U_{F1} \approx 0,58$ V ved 1 mA, $U_{F10} \approx 0,68$ V ved 10 mA og $U_{D10} \approx 0,70$ V ved 10 mA så $R_D \approx 2$ Ω .

9 – Gentag beregningen ved 100 mA. **Svar:** $R_D \approx 3$ Ω .

10 – En zenerdiode af typen BZX84 på 6,2 V med data på side 257 forbindes i serie med en typisk siliciumdiode. Bestem det resulterende spændingsfald? **Svar:** 6,9 V.

Hvilken strøm skal benyttes for at opnå en forsvindende lille temperaturkoefficient? **Svar:** $I_Z = 3$ mA.

11 – For opstillingen til venstre på side 258 vil diodens strøm kunne genere funktionen. Der ønskes at måle en strøm på 1 μA indenfor ± 10 nA ved 25°C . Kan 1N4148 anvendes? **Svar:** Fra kurverne side 254 ses at dioden maksimalt har en lækstrøm på 5 nA og i lederetningen er strømmen under dette niveau når spændingen over dioden i lederetningen er under 70 mV så svaret er "Ja".

12 – Dioden 1N4148 benyttes i strømmåleren fra side 258 ved en temperatur på 125°C . Aflæs diodens maksimale lækstrøm og afgør om opstillingen kan måle indenfor den ønskede nøjagtighed. **Svar:** 10 μA giver svaret "Nej".

13 – For opstillingen til højre på side 258 ledes 20 mA til den digitale indgang. Overvej hvilken forsyningsgren 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 og beskyttelsen er illusorisk.

Mikser.

14 – For mikseren på side 259 vil en radiosender med frekvensen 1 MHz blive modtaget ved en lokalscillator på 1,455 MHz. Hvilken anden radiofrekvens vil også kunne modtages ved samme indstilling? **Svar:** 1,91 MHz.

15 – Der indsættes et båndpasfilter i HF blokken med en centerfrekvens på 1 MHz for at dæmpe den uønskede spejlfrekvens på 1,91 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 dB hvis filtret repræsenteres ved dets Bode-plot. Filtrets godhed er lig med forholdet mellem niveauet ved centerfrekvensen for filtret og dets Bode-plot ved den samme frekvens (0 dB). Der mangler 34,4 dB, så godheden skal som et estimat være $Q = 10^{34,4/20} = 52$.

16 – For mikseren på side 261 benyttes $f_{LO} = 10$ 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.

Analog signalbehandling.

17 – Ved styrkereguleringen side 263 er det maksimale AC niveau angivet til 200 mV ved en strøm på 10 μA . 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 er ikke opfyldt.

18 – Ved samme opstilling beregnes DC strømmen med en formel hvor diodens spændingsfald antages til 0,6 V.

Hvorfor benyttes ikke de sædvanlige 0,7 V? **Svar:** Der benyttes en strøm under 1 mA hvor diodens spænding i lederretningen er under 0,6 V, som derved er en bedre tilnærmelse end 0,7 V.

19 – Kapacitetsdioden BB207 side 266 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:** $L = 32 \text{ nH}$.

Lys- og fotodioder.

20 – En blå lysdiode (se side 267) 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 .

21 – Lysdioden udskiftes med en rød LED uden at ændre andet. Hvad bliver diodens strøm? **Svar:** Fra 6 til 13 mA.

22 – Brug oplysningerne fra side 267 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:** $B \approx 700 \text{ MHz}$ og 1,4 Mbit/s.

23 – Dioden SFH485 pulser 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.

24 – Fotodioden SFH203 side 269 anbringes på akse 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 μA ved 10 mW/m² og da den belyses med 630 mW/m² bliver $I_P \approx 3,15 \text{ mA}$ når SFH485 er tændt og 0 A når SFH485 er slukket.

25 – For lysdetektoren side 270 beregnes det kraftigste lys der kan accepteres når spændingen over dioden skal forblive i spærretilstanden. **Svar:** Den maksimale strøm i fotodioden SFH203 er 0,5 mA hvor katoden når 0 V. Følsomheden er 50 μA ved 10 mW/m² så belysningens maksimale intensitet bliver 100 mW/m² ved 850 nm.

26 – Lysdetektoren skal reagere på det pulsedede 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:** $I_P = 10 \mu\text{A}$, $E = 2 \text{ mW/m}^2$ 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.

Transistor DC.

27 – Fra illustrationen for BC547A på side 275 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 $TC = 0,5 \%/^\circ\text{C}$.

28 – Forstærkeren på side 276 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 μA så kollektorstrømmen vil 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 mætter

(saturation) og den analoge funktion ophører.

29 – Vis at spændingen U_{CE} på kollektor af forstærkeren på side 276 nederst kan udtrykkes som vist herunder ved brug af formlerne i figuren og beregn spændingen for en værdi på $\beta = 900$. **Svar:** $U_{CE} = 1,5 \text{ V}$.

$$I_B = \frac{U_{CE} - U_{BE}}{R_B} = \frac{I_C}{\beta} = \frac{U_{CC} - U_{CE}}{\beta R_C} \Rightarrow U_{CE} = \frac{\frac{U_{CC}}{\beta R_C} + \frac{U_{BE}}{R_B}}{\frac{1}{R_B} + \frac{1}{\beta R_C}}$$

30 – For forstærkeren side 277 skal indgangsspændingen U_i beregnes for en udgangsspænding på nul volt. **Svar:** Med $U_{CE} = 0$ skal kollektorens strøm være 600 μA så basis skal have strømmen 2 μA . Basismodstanden aftager cirka 0,4 μA så der skal flyde 2,4 μA i $R_A = 180 \text{ k}\Omega$ og det giver en indgangsspænding på 1,13 V.

31 – For forstærkeren side 277 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 $U_i = 0,7 \text{ V}$.

32 – For forstærkeren side 277 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.

33 – For forstærkeren side 277 kan udgangen U_{CE} ikke nå helt ned til 0 V. Ved brug af kurven side 277 beregnes den lavest mulige udgangsspænding. **Svar:** $U_{CE} > 55 \text{ mV}$.

34 – For JFET strømkilden med BF245A side 281 skal modstanden R_S beregnes for en strøm på $I_D = 0,5 \text{ mA}$. **Svar:** $U_{GS} = -1,3 \text{ V}$ og $R_S = 2,6 \text{ k}\Omega$.

35 – For samme strømkilde skal findes den laveste værdi af spændingen på drain. **Svar:** $U_D > 2 \text{ V}$.

36 – Motorstyringen side 285 skal kunne drive en DC motor med et strømforbrug på 5 A. Beregn den laveste værdi af spændingen på gate. **Svar:** $U_{GS} > 5,5 \text{ V}$.

Transistor AC.

37 – For AC modellen side 286 skal det vises at erstatning af U_{BE} med $U_{BE} + u_{be}$ giver det angivne udtryk for $I_C + i_c$. Erstat eksponentialfunktionen med Taylor seriens to første led (se andet kapitel) og vis at relationen bliver $i_c = g_m u_{be}$ hvor transkonduktansen er $g_m = I_C / U_T$.

38 – Det skal det vises at AC modellen forudsætter en signalamplitude på cirka 1 mV for at anden-ordens leddet er under 1 % af første-ordens leddet. **Tip:** Indsæt værdierne 1 mV, 2 mV, ... og beregn leddenes værdier.

39 – Med udgangspunkt i mikrofonforstærkeren side 287 skal det vises at grænsefrekvensen ved indgangen af T_S på side 314 er som følger, når transistoren fødes fra en høj modstandsværdi:

$$f_G = \frac{1}{2\pi\beta R_C C_u}$$

40 – For hybrid-pi modellen side 288 beregnes værdierne af parametrene for en BC547 ved $I_C = 0,3 \text{ mA}$. Oplyste værdier: $n = 1$ og $\beta = 300$. **Svar:** $g_m = 11,5 \text{ mS}$, $r_{\pi} = 26 \text{ k}\Omega$,

$f_T = 9 \text{ MHz}$, $C_\pi = 200 \text{ pF}$, $r_O = 230 \text{ k}\Omega$.

41 – For opstillingen side 297 gennemføres en beregning af DC arbejds punktet for $I_C = 1 \text{ mA}$, β på ∞ og $U_C = U_{CC}/2$. Derefter sættes $\beta = 300$ og ændringen i U_B vises at være på $0,23 \text{ V}$. Hvilken betydning har den ændring?

42 – For opstillingen side 298 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 299 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_I er som vist side 301. 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.

Operationsforstærker.

45 – Transkonduktansen g_m for differentialtrinnet side 310 udledes for MOSFET ved differenser hvor definitionen benytter differentiation i et arbejds punkt. Vis at det leder til det samme resultat for arbejds punktet hvor $U_P - U_N = 0$. Du får brug for differentiation af en sammensat funktion fg , hvor $(fg)' = f'g + fg'$, og kædereolen $df/dx = (df/dy)(dy/dx)$.

Halvlederteori.

46 – 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 .

47 – 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 formen for n_i^2 på side 326. 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 formen for elektronernes tæthed n .

48 – Beregn koncentrationen af ladningsbærere ved en temperatur på 100°C og vis at modstanden af en terning af silicium med sidelængden 1 mm bliver $R = 31 \text{ k}\Omega$.

49 – Den gennemsnitlige hastighed er $v_{th} = 10^5 \text{ m/s}$ for elektroner ved almindelig stuetemperatur. Vis at den gennemsnitlige energi (udtrykt som $0,5 m v_{th}^2$) er omkring $4,6 \cdot 10^{-21} \text{ J}$ og sammenlign med den energi, der kræves for at elektronen løsriver fra atomet.

50 – Udfør integrationen for det elektriske felt E på side 329 og vis at E_0 bliver som vist for x_p leddet. Begrund hvorfor x_n leddet skal give samme resultat og vis derefter at der gælder at $x_n/x_p = N_A/N_D$.

51 – Sæt $U_P = 0$ og $U_N = U_B$ i formlerne på side 329 og gennemfør beregning af det indbyggede potentiale U_B .

52 – Beregn diffusionskonstanten med tabeldata side 331 for N siden af en halvleder og derefter diffusionslængden, for $\tau_D = 1 \mu\text{s}$. **Svar:** $D_N = 1,2 \text{ m}^2/\text{s}$ og $L_N = 35 \mu\text{m}$.

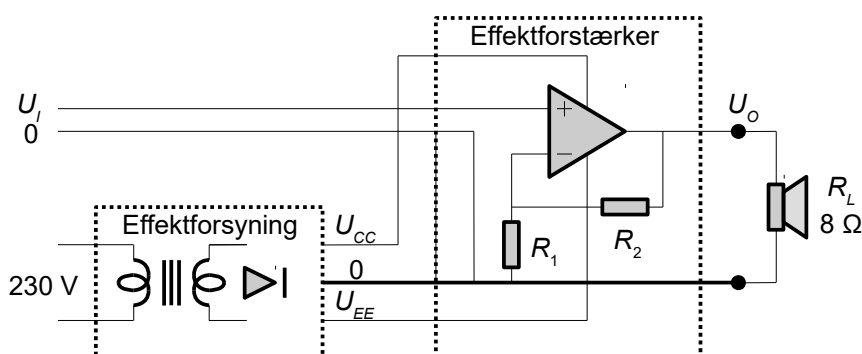
6 EFFEKTKREDSLØB

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 effektransistorer på en silicium chip for tabeffekten 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 indledes med effektforstærkere for lyd gengivelse 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 højere niveauer benyttes en parallelkobling af effektforstærkere med transformatorer for at kombinere udgangssignalerne, men det er stadig radiorøret, blandt andet i form af klystronen, der leverer den virkeligt 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 indtil 50 W i 8Ω . Der er ingen grundlæggende forskel fra effektforstærkere til højere effekter, men der kræves en del erfaring for at gennemføre designet.



Figur 386 – 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 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 363, derefter spændingsregulering ved lineære serieregulatorer på side 374 samt de energieffektive DC-DC konvertere der introduceres på side 378.

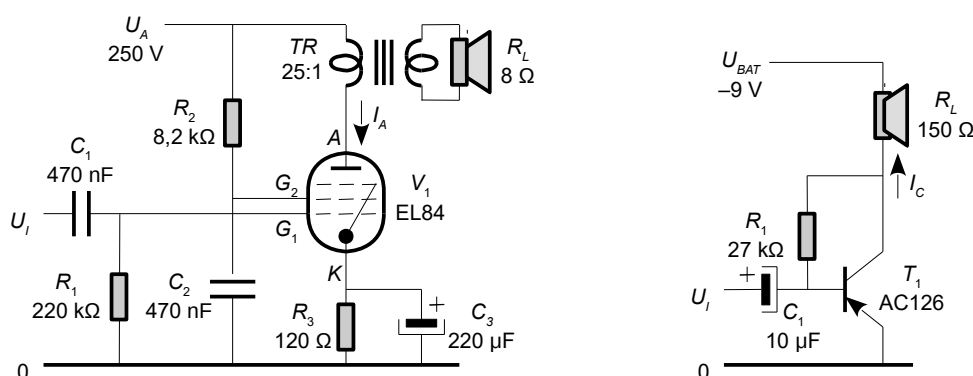
6.2 Klassifikation af udgangstrin

Udgangstrin opdeles i klasser der betegnes med bogstaver der beskriver de specifikke egenskaber ved kredsløbet. Dette kapitel sigter mod effektforstærkere for musikgengivelse over højttalere, men radiotransmission med radorør for var i brug allerede fra 1920 og det præger nomenklaturen.

De grundlæggende klasser er A, B, C og D som omhandles i det følgende. Dertil klasse E og F der udelukkende er for højfrekvens samt klasse G og H der er rettet mod lyd-gengivelse af høj kvalitet og højt lydniveau, men uden den eskalerende effektafsættelse. Litteraturen er rig på en lang række varianter som klasse S og T der begge er derivater af klasse D, *current dumping* der kombinerer klasse A og C samt klasse XD (*cross-over displacement*) der flytter et klasse B trins problematiske cross-over område 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 387 – 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 radorør og højttaleren var tilsluttet gennem en transformator for at undgå en DC strøm i svingspolen. I eksemplet arbejder pentoden EL84 ved forsyningens $U_A = 250\text{ V}$ og anoden A indstilles til strømmen $I_A = 48\text{ mA}$ ved modstanden R_3 så potentialet er $-5,5\text{ V}$ på gitteret G_1 overfor katoden K. Skærmgitteret 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 $5\text{ k}\Omega$ ved anoden og udgangseffekten når maksimum på $5,7\text{ W}$ når anodens strøm varierer med $\pm 100\%$. Virkningsgraden er 33 % og den harmoniske forvrængning oplyses i databladet til 10 %.

Det andet eksempel er fra elektronikbyggesættet Philips EE20 fra omkring 1965 hvor transistoren AC126 driver en $150\ \Omega$ højttaler uden transformator. Der løber en DC strøm i svingspolen hvilket afsætter effekt i den og membranen forskydes fra den neutrale position. En strømforstærkning på 100 får transistoren til at indstille sin strøm i kollektor på 20 mA så der optages 180 mW . Hvis strømmen svinger $\pm 20\text{ mA}$ om arbejds punktet bliver RMS værdien 14 mA og der vil afsættes en AC effekt på 30 mW i højttaleren så den kan ikke spille ret højt. Virkningsgraden på 17 % er lavere end for radorøret 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 karakteristisk og her er udgangseffekten under halvdelen af det maksimale. Bemærk at en PNP transistor arbejder med negativ spænding fra effektforsyningen.

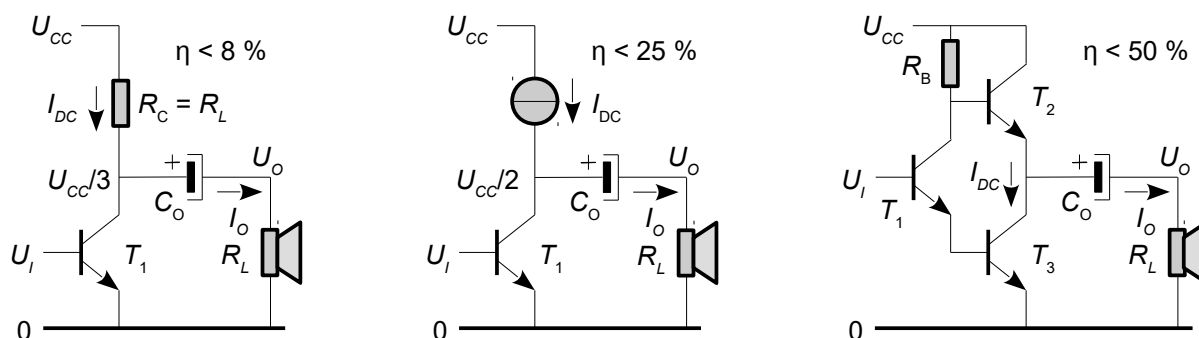
En klasse A effektforstærker har sine gode egenskaber fra det forhold at den aktive komponent primært arbejder ved små variationer omkring arbejds punktet og den harmoniske forvrængning er normalt lav af den simple grund at forvrængningen er proportional med signalniveauet. De to eksempler er derfor kun gode effektforstærkere ved moderat lytteniveau, men det er jo også det normale at en klasse A forstærker arbejder ved en procentdel af sin maksimale udgangseffekt.

Effektforbruget over en periode af signalet er konstant ved klasse A så effektforsyningen belastes konstant og 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 % så der afsættes altid en betydelig effekt i transistorerne. Den største udfordring ved klasse A er derfor at bringe effekten væk. Selv om det store effektforbrug burde gøre klassifikationen uinteressant så skal man aktivt foretage sig noget for ikke at bygge klasse A.

$$\text{Effektafsættelse i klasse A: } P_D = U_{CC} I_{DC}$$

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 I_{DC} .

Den første opstilling er en almindelig fælles-emitter forstærker hvor belastningen R_L tilsluttes via en kondensator så kun AC delen af udgangssignalet overføres. Strømvariationen i kollektor I_O 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 på I_{DC} hvilket hæver virkningsgraden til højst 25 %. Kondensatoren C_O kan undgås med en dobbelt effektforsyning og et differentialtrin i indgangen. Den sidste opstilling benytter tre transistorer hvor T_1 modulerer 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 388 – 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.

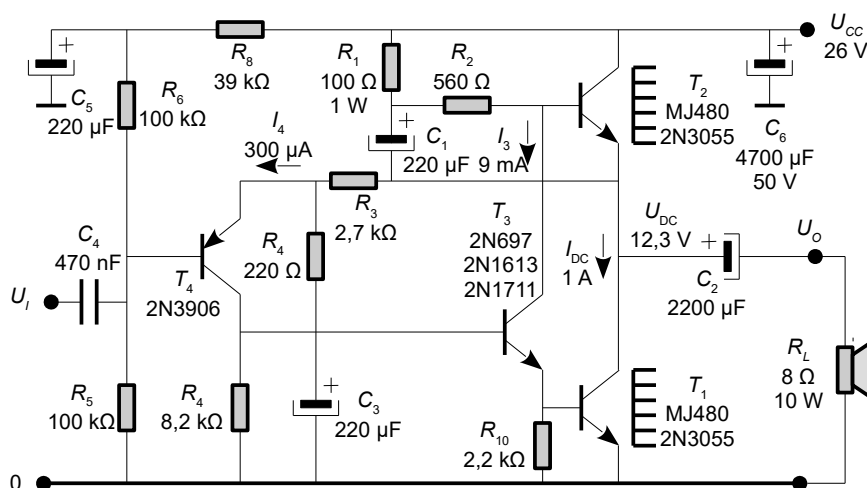
Udgangseffekten kan beregnes fra $P_O = U_{O\text{RMS}}^2 / R_L = I_{O\text{RMS}}^2 \cdot R_L$ og for en udgangseffekt på $P_O = 1\text{ W}$ bliver amplituden $U_O = \pm(2 \cdot P_O \cdot R_L)^{0.5} = \pm 4\text{ V}$ ved $R_L = 8\ \Omega$ der giver spændingsfaldet til $U_{CE} = 4\text{ V}$ i hvile. Strømdelingen mellem R_C og R_L giver amplituden af kollektorens strøm så der behøves $I_{DC} = 2 \cdot (2 \cdot P_O / R_L)^{0.5} = 1\text{ A}$ i kollektor i hvile. Effektforsyningen beregnes til $U_{CC} = R_C I_{DC} + U_{CE} = 12\text{ V}$ og dermed afgives effekten $P_{CC} = U_{CC} I_{DC} = 12\text{ W}$ fra batteriet så virkningsgraden bliver på $\eta = P_O / P_{CC} = 8\%$.

Effektforsyningen kan her reduceres til spids-spids værdien af udgangs svingning på $U_{CC} = (2 \cdot P_O \cdot R_L)^{0.5} = 8\text{ V}$ ved en udgangseffekt på $P_O = 1\text{ W}$ og $R_L = 8\ \Omega$. Udgangsstrømmen kan nå op på I_{DC} og DC arbejds punktet bliver $I_{DC} = (2 \cdot P_O / R_L)^{0.5} = 0,5\text{ A}$. Den optagne effekt reduceres til $P_{CC} = U_{CC} I_{DC} = 4\text{ W}$ 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 arbejds punktet, men udgangsstrømmen er givet ved differensen mellem T_2 og T_3 til værdien $\pm 2 \cdot I_{DC}$. DC arbejds punktet bliver $I_{DC} = 0,5 \cdot (2 \cdot P_O / R_L)^{0.5} = 0,25\text{ A}$ ved $P_O = 1\text{ W}$ og $R_L = 8\ \Omega$. Udgangsspændingens amplitude giver det samme krav til effektforsyningen som før på $U_{CC} = (2 \cdot P_O \cdot R_L)^{0.5} = 8\text{ V}$, men effektforsyningen afgiver nu kun en effekt på $P_{CC} = U_{CC} I_{DC} = 2\text{ W}$ så virkningsgraden er $\eta = P_O / P_{CC} = 50\%$.

Beregningen af virkningsgrad forudsætter at transistorernes kollektor-emitter kan nå ned til 0 V så det 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 angivet ved designformlerne og et eksempel skal gives.

Herunder vises en klasse A effektforstærker med *push pull* udgang der leverer 10 W til 8 Ω ved lav forvrængning. Kredsløbet er af samme type som vist til højre i forrige figur. Opstillingen blev bragt af John Linsley Hood i fagbladet Wireless World April 1969. Effektforsyningen var seriereguleret for at undertrykke netbrum, men den del medtages ikke i diagrammet.¹⁸⁹



Figur 389 – 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 radorørret 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. Effektransistorerne 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 arbejds punktet for effektransistorerne 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 effektransistorerne modtage 9 mA så strømforstærkningen skal være over 100 for det planlagte arbejds punkt. Strømforstærkningen er på mindst 30 og typisk 200 for MJ480 samt mindst 20 og typisk 70 for 2N3055 og det viser at denne type effektforstærker enten bør bruge udmålte transistorer eller at der bør være en justeringsmulighed.¹⁹⁰

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.

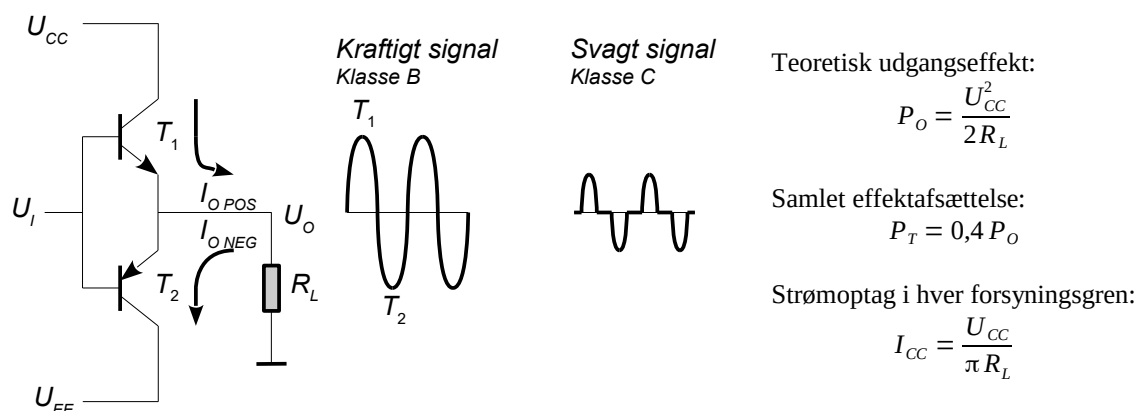
¹⁸⁹ Artiklen kan blandt andet findes på: <http://www.keith-snook.info/wireless-world-magazine/wireless-world-articles.html>.

¹⁹⁰ 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 5.

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 processeres 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 390 – 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 cirka 1960 anvendelse af transistorer til effektforstærkere der ikke i samme grad som radorøret tåler en stor effektafsættelse. Der afsættes maksimalt 40 % af den nominelle udgangseffekt i udgangen så en 100 W forstærker afsætter op til 20 W i hver udgangstransistor i modsætning til minimum 100 W i hver udgangstransistor ved klasse A.

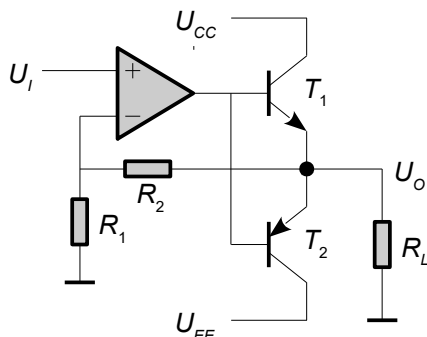
Der er sjældent behov for at afgive den fulde udgangseffekt over længere tid for et musikprogram består mest af svage signaler med få kraftige spidser og selv ved blot 10 dB til spidserne afsættes blot en tiendedel af fuld udgangseffekt lige inden forstærkeren overstyres. Fordelen ved klasse B er derfor at udgangstransistorerne kun opvarmes ganske lidt i sammenligning med klasse A.

Virkningsgraden kan teoretisk nå 78 % ved fuld udgangseffekt som det vises side 354. Litteraturen nævner det ofte som et vigtigt punkt, men det er efter min mening langt vigtigere med det lavere gennemsnitlige optag af effekt fra effektforsyningen.

Klasse B er i sin rene udformning en sjælden gæst indenfor lydengivelse 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 effektforstærkere med transistorer at det nedsættende blev betegnet med "transistorlyd" som en klar adskillelse fra de gode og velkendte effektforstærkere med radorør.

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* specifikation. Desuden fjerner det ikke cross-over forvrængningen. Problemet bliver reduceret, men det elimineres ikke.



$$\text{Indstillet forstærkning: } A = 1 + \frac{R_2}{R_1}$$

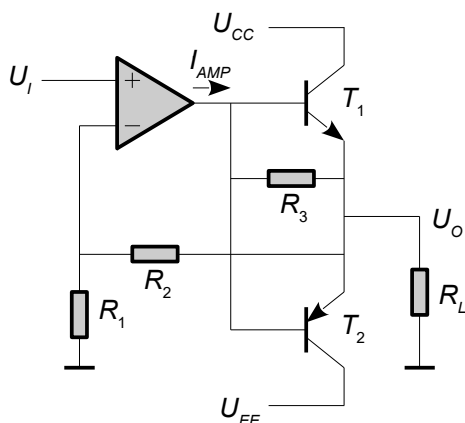
$$\text{OpAmp båndbredde: } GBP > \left(1 + \frac{R_2}{R_1}\right) \frac{2U_{BE}}{\Delta U_I} f_{MAX}$$

$$f_{MAX} < \frac{GBP}{\left(1 + \frac{R_2}{R_1}\right) \frac{2U_{BE}}{\Delta U_I}} = 650 \text{ Hz for } \begin{matrix} GBP = 10 \text{ MHz} \\ R_1 = 1 \text{ k}\Omega \\ R_2 = 10 \text{ k}\Omega \\ \Delta U_I = 1 \text{ mV} \end{matrix}$$

Figur 391 – 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.



Indstillet forstærkning:

$$A = 1 + \frac{R_2}{R_1}$$

Udgangsstrøm fra OpAmp:

$$I_{AMP} = \frac{U_{BE}}{R_3} + \frac{U_O}{\beta R_L}$$

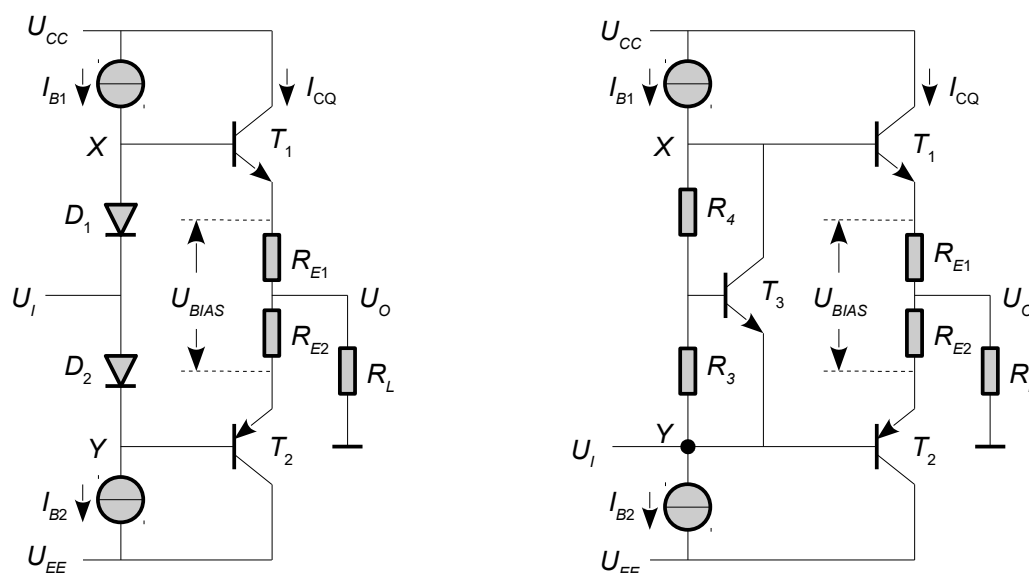
Figur 392 – Simpelt udgangstrin med korrektion for cross-over forvrængningen gennem brug af en operationsforstærker med tilbagekobling.

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 (langsomt 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 siden.¹⁹¹

¹⁹¹ <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 393 – To almindelige 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.

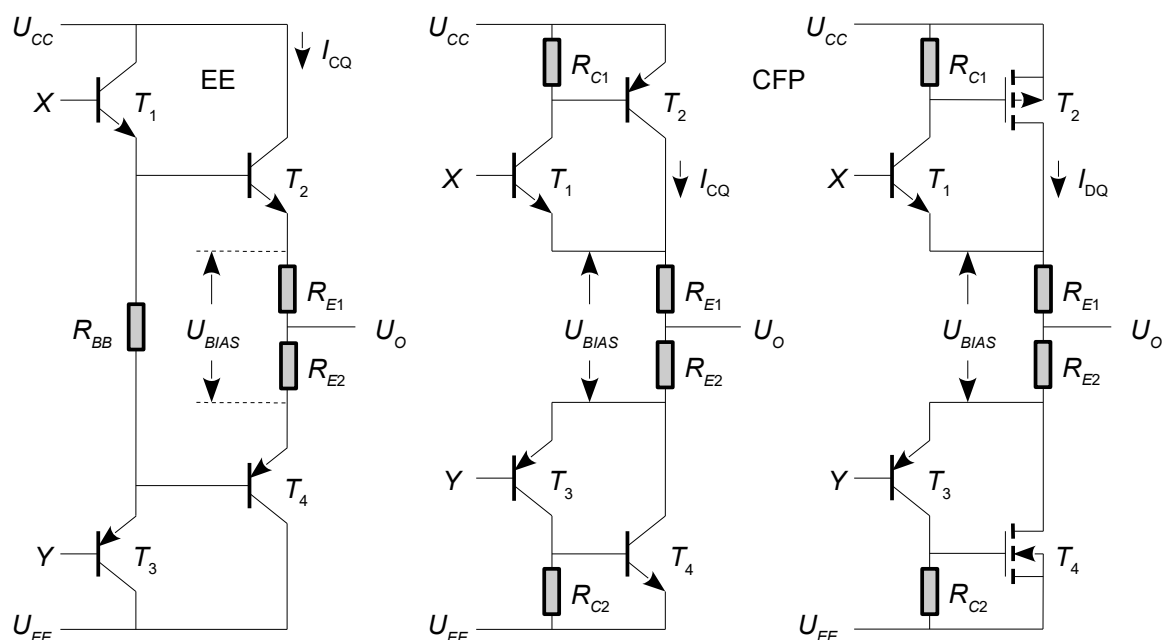
Ved det venstre kredsløb giver dioderne D_1 og D_2 en spænding på omkring 1,4 V der ideelt svarer til transistorernes basis-emitter strækninger plus spændingsfaldet over modstandene i emitter. Dioderne kompenserer for transistorernes temperaturkoefficient på cirka $-2 \text{ mV}/^\circ\text{C}$ hvis de er 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 spændingsdelingen 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 den hvilestrøm I_{CQ} der vil give 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} ændres med strømmen i transistorerne og stabiliserer hvilestrømmen I_{CQ} når temperaturen varierer. Strømgeneratorerne I_{B1} og I_{B2} giver strømmen til basis for 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 344.

DC strømiveauet I_{CQ} (*quiescent current*) er på 20 til 50 mA for typiske klasse AB effektforstærkere og kan måles som hvilespændingsfaldet U_{BIAS} over R_{E1} og R_{E2} . Der er ingen "gylden værdi" for det optimale arbejds punkt som må findes ved at måle den harmoniske forvrængning. Det indskyder desværre et testpunkt i produktionen, men det er svært at serieproducere klasse AB uden dette fordyrende testpunkt. Forvrængningen stiger 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 levere 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 394 – 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.

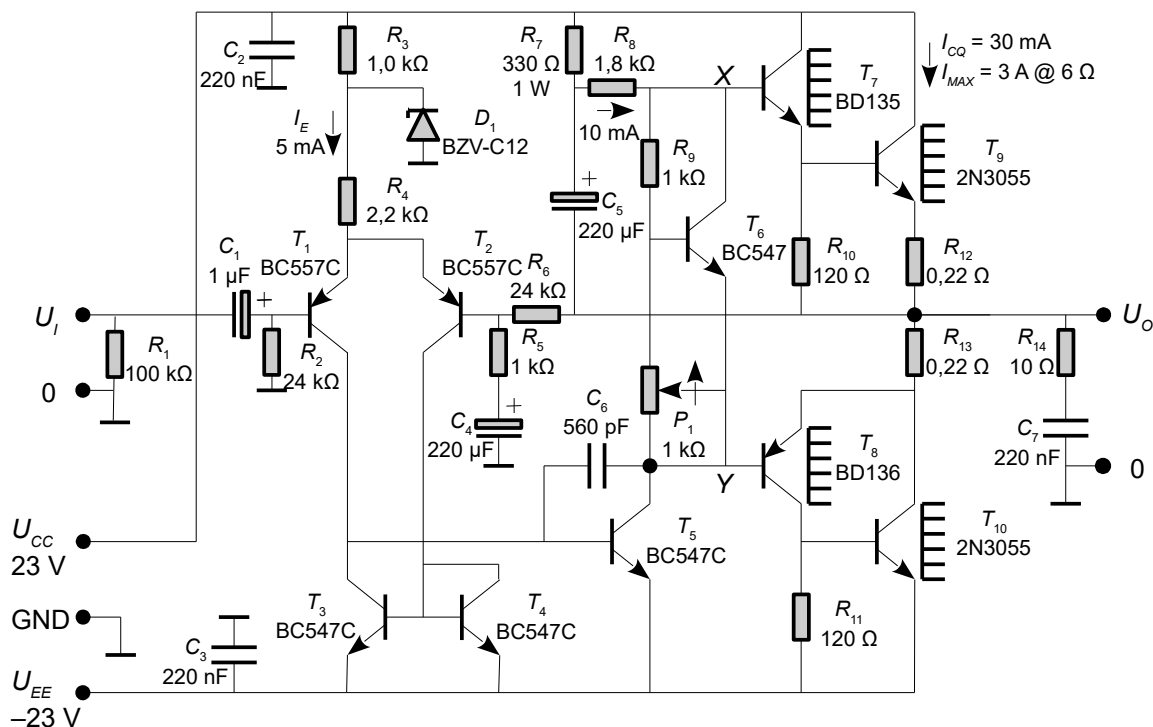
Til venstre vises to dobbelte emitterfølgere (EE) der populært kaldes for et *Darlington pair* hvor T_1 forstærker strømmen til basis af T_2 og tilsvarende for T_3 og T_4 . Der er fire basis-emitter strækninger fra X til Y, men arbejdstemperaturen varierer mellem transistorerne så det er svært at holde en konstant værdi af strømmen I_{CQ} . Enkelte effekttransistorer leveres med dioder på siliciumchippet for en mere præcis tilbagekobling af temperaturen. Modstanden R_{BB} forbedrer funktionen ved høje frekvenser ved at potentialet på basis kan trækkes forbi potentialet på emitter så rumladningen i basis-emitter dioden hurtigere fjernes.

I midten ses en to-trins forstærker 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 U_{BIAS} . Det er derfor lettere at holde I_{CQ} på et konstant niveau, 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 begrænser modstandenes værdier og 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 effekttab 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 Cdu/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 af afbryde for dem.

6.2.4 Designeksempel

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 effektransistorer hvilket var ret udbredt omkring 1980 hvor effektransistorer af typen PNP var ringere end NPN.



Figur 395 – Prototype på en klasse AB forstærker for 20 W i 8 Ω med quasi komplementær udgang. Der er ikke inkluderet beskyttelse mod kortslutning af udgangen.

Herunder beskrives et muligt DC designforløb.

- Effektforsyningen** planlægges med start ved udgangen. Ved $P_O = 20 \text{ W}$ i $R_L = 8 \Omega$ kræves en udgangsspænding på $U_{O\text{ RMS}} = (P_O R_L)^{0.5} = 12,6 \text{ V}$ med amplituden $U_O = 18 \text{ V}$. Der skal være plads til to basis-emitter strækninger for udgangstrinnet, noget lignende til den øvrige elektronik, ripple i effektforsyningen og 10 % på netspændingen så effektforsyningen bør placeres 5 til 7 V over kravet til U_O og det giver her værdien af U_{CC} og U_{EE} til $\pm 23 \text{ V}$.
- Strømniveauet** i udgangstrinnet estimeres. Med $R_L = 6 \Omega$ som belastning (se side 352) når strømmens spidsværdi $I_{MAX} = U_O / R_L = 3 \text{ A}$ (se side 354). Der afgives $P_O = U_O^2 / 2R_L = 27 \text{ W}$ og middelværdien af strømmen fra effektforsyningen bliver $I_{DC} = I_{MAX} / 3 = 1 \text{ A}$. Hvis den værdi holder så afgiver effektforsyningen 46 W og virkningsgraden bliver $\eta = 59 \%$.
- Effekttabet** i udgangstransistorerne T_9 og T_{10} bliver $P_T = 0,4 \cdot U_{CC}^2 / 2R_L = 18 \text{ W}$ samlet så der forventes et tab på 9 W i hver transistor (se side 354).
- Drivertrinnet** skal levere basisstrømmen til udgangstransistorerne og med $\beta_9 = \beta_{10} = 20$ for strømforstærkning i T_9 og T_{10} behøves $I_{B9} = I_{MAX} / \beta_9 = 150 \text{ mA}$. Effekttabet bliver på 1/20 af effekttabet udgangstransistorerne eller omkring 0,5 W i hver transistor. Man bør være meget konservativ og benytte en transistor der kan håndtere mindst 1 W da effekten vil variere som funktion af udgangsspændingen. Jeg ville bruge en på mindst 10 W uden et køleprofil og ellers må der planlægges køling af disse transistorer.
- Basisstrømmen** til drivertrinnet er $I_{B7} = I_{B9} / \beta_7 = 4 \text{ mA}$ ved $\beta_7 = \beta_8 = 40$. Strømmen leveres

gennem R_7 og R_8 der dog er designet til $I_{R7} = 10$ mA for at have et overskud. Der afsættes 0,23 W i transistor T_5 og spændingssvinget over den kan nå over 40 V så en BC547 bør genovervejes. Strømmen i R_8 holdes på en konstant værdi ved *boot strap* med C_5 som sikrer at spændingen over R_8 ikke varierer. Effekten i R_7 bliver på 0,5 W ved fuld udstyring så modstanden skal mindst kunne klare 1 W for at holde temperaturstigningen nede.

6. **Biaskredsløbet** behøver $U_{CE6} = 2,1$ V (fra T_7 , T_8 og T_9) som T_6 leverer ved at forstærke dens $U_{BE6} = 0,7$ V op til $U_{CE6} = (1 + R_9/P_1)U_{BE6}$. Potentiometret P_1 antages til 500 Ω .
7. **Indgangstrinnet** skal levere basisstrømmen til T_5 på $I_{B5} = I_{R7}/\beta_5 = 25$ μ A med $\beta_5 = 420$. Der skal vælges en højere strøm i den fælles emitter for at holde differentialtrinnets balance så nær ved fifty-fifty som muligt og strømmen vælges her til $I_E = 5$ mA. Det indstilles ved D_1 og R_4 hvor zenerdioden er indsat for at dæmpe ripplen fra effektforsyningen.

Med kendte DC arbejdsforhold kan AC analysen tage over.

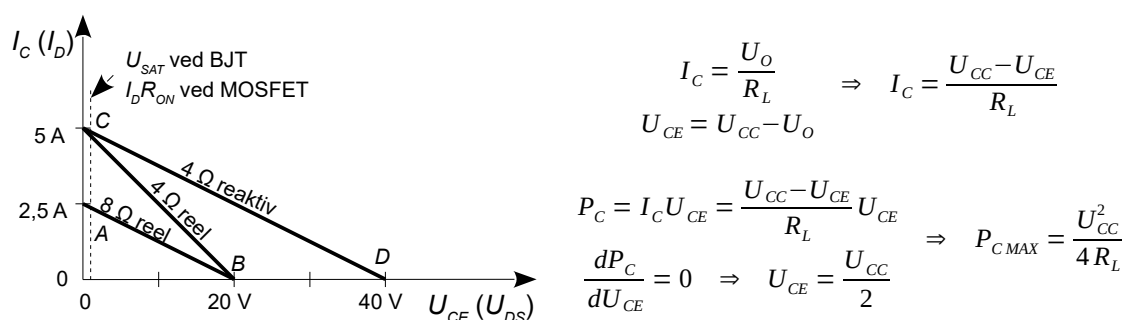
1. **Forstærkningen** indstilles så fuld udstyring på $U_{O_{RMS}} = 12,5$ V opnås ved $U_{I_{RMS}} = 0,5$ V på indgangen og indstilles af $A_{CL} = 1 + R_6/R_5 = 25$. Et valg af $R_5 = 1$ k Ω giver $R_6 = 24$ k Ω .
2. **Den lave grænsefrekvens** fastsættes af C_1 og R_2 til $f_L = 1/2\pi R_2 C_1 = 7$ Hz. Kondensator C_4 skal blot være "stor" så spændingsvariationen over den bliver "lille" i sammenligning med værdien over R_5 for at holde forvrængningen nede. Det skyldes at en aluminium elektrolyt kondensator ikke er en lineær komponent. En typisk værdi på 220 μ F giver 0,7 Hz.
3. **Transkonduktansen** i differentialtrinnet $T_1 \dots T_4$ beregnes til $g_m = I_E/2U_T = 0,1$ S.
4. **Åben-sløjfe forstærkningen** beregnes fra transkonduktansen, strømforstærkningen i T_5 og indgangsmodstanden på udgangstrinnet ved basis af T_7 og T_8 der giver et estimat på R_C i formen. Indgangsmodstanden for en emitterfølger er belastningen R_L gange med den samlede strømforstærkning til $R_C = \beta_7 \beta_9 R_L = 4,8$ k Ω . DC forstærkningen beregnes derefter til $A_{DC} = g_m \beta_5 R_C = 200\ 000$ ved en belastning på $R_L = 6$ Ω .
5. **Kompenseringen** skal sikre stabilitet ved at reducere forstærkningen til A_{CL} ved den høje pol der gives af R_C og C_μ for T_5 , T_7 og T_8 . Med $C_{\mu 5} = 2$ pF, $C_{\mu 7} = 10$ pF og $C_{\mu 8} = 18$ pF findes $C_\mu = 30$ pF så polen bliver $f_H = 1/2\pi R_C C_\mu = 1,1$ MHz og herfra beregnes kompenseringens kondensator til $C_C = g_m/2\pi f_H A_{CL} = 580$ pF.
6. **Slew rate** værdien bliver $SR = I_E/C_6 = 9$ V/ μ s der tilstrækkelig for musikformål som er på $SR = 2\pi f U_O = 2$ V/ μ s ved frekvensen 20 kHz.
7. **Båndbredden** er principielt på 1 MHz på grund af proceduren for kompensering, men det skal måles hvis parameteren er vigtig for anvendelsen.

Køleprofilen for udgangstransistorerne beregnes ud fra specifikationen af hvad temperaturen på transistorens chip må stige til. Det benævnes T_j for *junction temperature*.

1. **Chiptemperaturen** må for 2N3055 maksimalt være 185°C ved 9 W afsat effekt, men man bør aldrig gå til grænsen så her fastsættes værdien konservativt til $T_j = 100^\circ\text{C}$. Med 25°C i omgivelserne skal temperaturstigningen holdes nede på 75°C.
2. **Den termiske modstand** fra junction til omgivelser ved 9 W bliver 8,3°C/W med en køleprofil for hver transistor. Den termiske modstand fra junction til monteringsflade er 1,5°C/W ifølge databladet og der må påregnes en lignende værdi fra montagen så kravet bliver 5,3°C/W for køleprofilens termiske modstand til omgivelserne (K-værdien).

6.2.5 Arbejdslinje

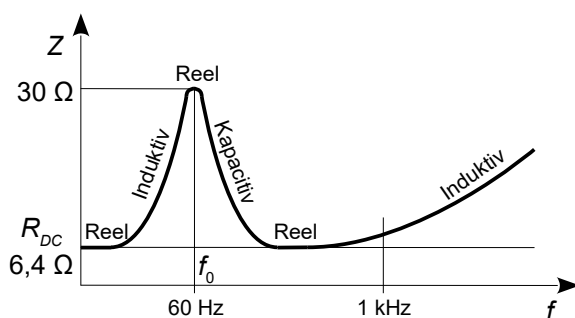
Belastningen af en effektforstærker regnes normalt for reel og der er da en lineær relation mellem udgangsspændingen U_O og strømmen I_O til belastningen og dermed I_C i udgangstransistoren. Det udtrykkes ved en arbejdslinje der viser sammenhængen mellem strømmen I_C og spændingen U_{CE} over transistorens 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 B ved $R_L = 4 \Omega$. Talværdierne ved akserne er for 20 W effektforstærkeren på side 350.



Figur 396 – Arbejdslinjerne for en effektforstærker med 8 Ω og 4 Ω belastning samt med en induktiv belastning (4 Ω reaktiv).

Transistoren kan ikke trække udgangen helt ud til effektforsyningen på grund af opbygningen med to transistorer og mætningsspændingsfaldet på $U_{SAT} \approx 1$ V for en bipolær transistor eller kanalens modstand $R_{ON} \approx 0,1 \Omega$ for en MOSFET, men det kan normalt ignoreres.

En højttaler er ikke alene beskrevet ved en modstand på 8 Ω . Højttalerens svingspole har en DC modstand R_{DC} på typisk 80 % af den nominelle værdi (det er 6,4 Ω , men jeg har målt lavere) så strømmen når let op på 3 A. Ved højttalerens mekaniske resonans vokser impedansen, den falder tilbage til DC værdien ved middel frekvens og stiger igen mod høje frekvenser hvor svingspolens selvinduktion kommer i spil. Forløbet varierer mellem at være kapacitiv, reel og induktiv med en variation i fasevinklen på $\pm 60^\circ$ som en reel mulighed.



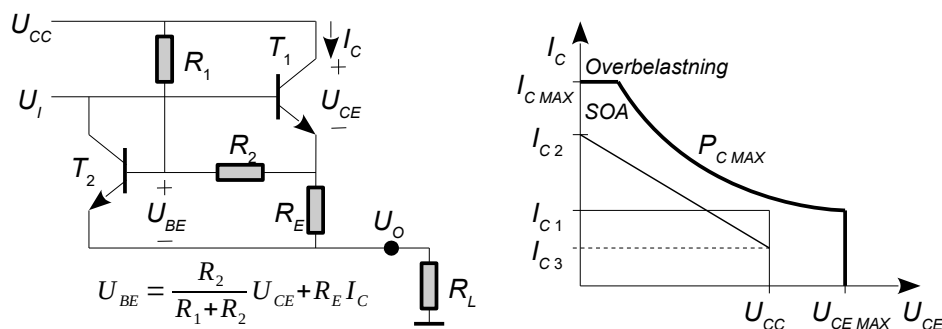
Figur 397 – Det typiske forløb af impedansen for en 8 Ω bashøjttaler i et lukket kabinet.

For at inkludere en margin overfor den faktiske belastning for en effektforstærker på 8 Ω kan der designes for 4 Ω belastning hvilket medfører at 20 W forstærkeren skal kunne klare en afsat effekt på 25 W der dog kun mødes kortvarigt ved et musiksignal. Den værste mulige reaktive værdi kan repræsenteres ved at designe efter den dobbelte værdi af spændingsfaldet for en reel belastning som vist ved arbejdslinjen fra C til D der er *worst case situationen* [DS-446].¹⁹²

¹⁹²Der 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.

6.2.6 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 udgangstranstoren så den afsatte effekt altid holdes indenfor det tilladelige. Det er transistorens *safe operating area* SOA. Det er ikke nok at beskytte transistoren med en smeltesikring for transistoren brænder af inden sikringen når at afbryde. En smeltesikring kan dog være en god investering for at undgå følgeskader i effektforsyningen [DS-443].



Figur 398 – 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 overlast af spænding, strøm og effekt. Transistoren kan klare en spænding $U_{CE MAX}$ over kollektor og emitter før 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 sammenbrud.¹⁹³

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 effektforstærkeren 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 (køleprofilen). 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 .

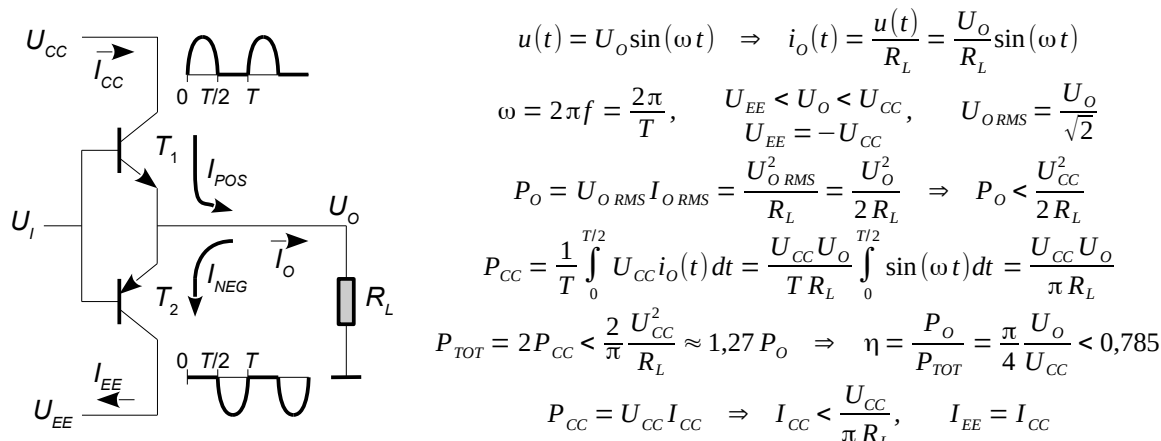
Ved en maksimal strøm i kollektor af T_2 på 100 mA og en strømforstærkning på 100 bliver strømmen 1 mA i basis. Spændingsfaldet over R_2 andrager da 0,1 V og kortslutningsstrømmen hæves med 15 %.

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 højeste mulige spænding på udgangen hvor $U_{CE} \approx 0$. Ved at inkludere R_1 kan strømmen reduceres ved kortsluttet udgang til $I_{C3} < P_{C MAX} / U_{CC}$.

¹⁹³ Ved 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.

6.2.7 Effektberegning

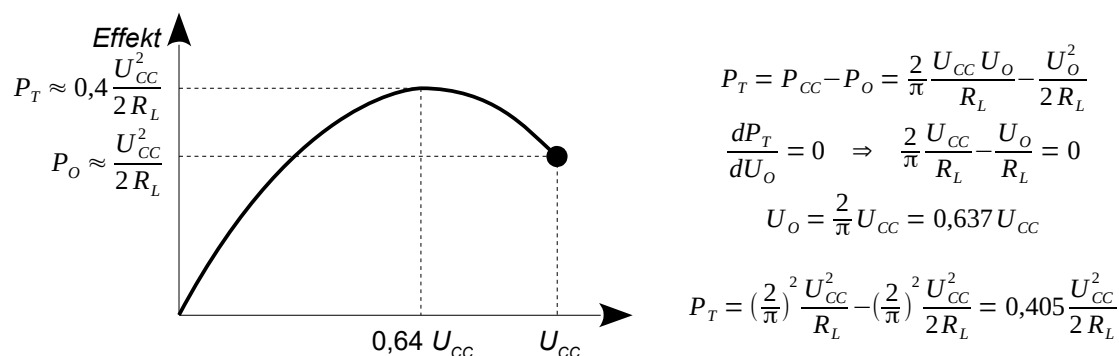
I dette afsnit beregnes den afsatte effekt i udgangstrinnet og kravet til effektforsyningen. Den maksimale udgangseffekt P_O opnås når udgangen svinger helt op til effektforsyningens U_{CC} hvilket kun er teoretisk muligt. Effektforsyningen skal da totalt levere P_{TOT} på knapt 30 % over P_O og det 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 399 – 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 $P_O = 20 \text{ W}$ til $R_L = 8 \Omega$ er $U_O = 18 \text{ V}$ og $I_O = 2,2 \text{ A}$ så der kræves $U_{CC} = -U_{EE} > 18 \text{ V}$ og $I_{CC} = I_{EE} > 0,7 \text{ A}$. Ved et design efter teoriens minimum på $U_{CC} = 18 \text{ V}$ skal effektforsyningen kunne levere $P_{TOT} = 25 \text{ W}$.

Den teoretiske virkningsgrad er 78 % for klasse B, men i praksis kan amplituden ikke nå helt op til effektforsyningen da transistorerne har et mætningsspændingsfald. Det er dog langt vigtigere at den afsatte effekt er en funktion af udgangseffekten. Det betyder at klasse B ikke afsætter en stor effekt ved lavt udgangsniveau.



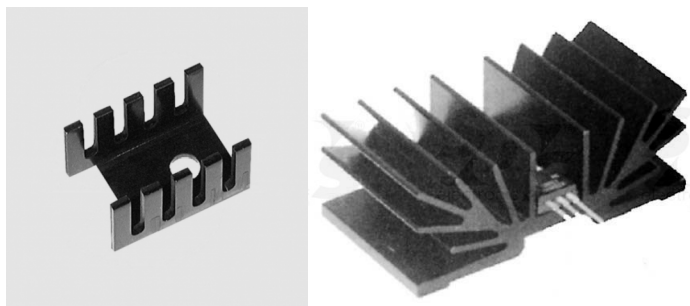
Figur 400 – Den afsatte effekt i udgangstransistorerne er tilsammen under 40 % af den maksimalt mulige udgangseffekt og værdien nås ved mindre end fuld udgangseffekt.

Med $U_O = U_{CC} = 20 \text{ V}$ og $R_L = 8 \Omega$ bliver den maksimale effektafsættelse i udgangstrinnet på $P_T = 8 \text{ W}$, 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 P_O ved den værst tænkelige amplitude af udgangssignalet. For en 100 W klasse B effektforstærker er $P_T = 40 \text{ W}$ så hver udgangstransistor skal optage 20 W .

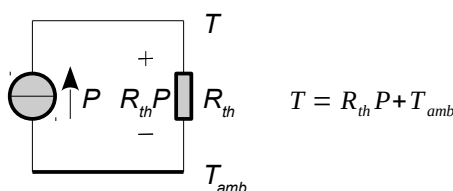
6.2.8 Termisk design

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 ekstruderet 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 401 – Køleprofiler (venstre) med målene 19,1 x 13,2 x 12,7 mm³ og termisk modstand 27 K/W samt (højre) 22,5 x 37,5 x 54 mm³ med termisk modstand 5,5 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 °C/W eller K/W der her betyder det samme.



Figur 402 – Beregning af temperaturstigning ved en effektstrøm P i en termisk modstand.

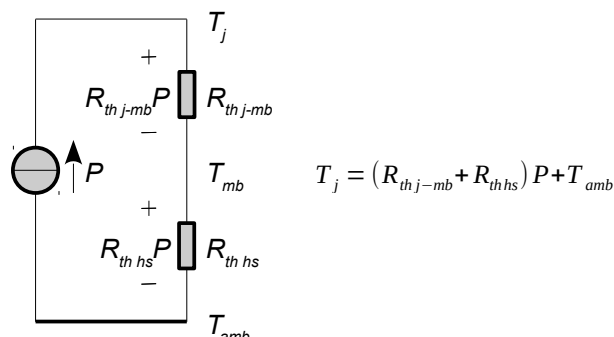
Med den store køleprofil på $R_{th} = 5,5$ K/W vil en afsat effekt på $P = 4$ W hæve temperaturen med 22°C og med en omgivelsestemperatur på $T_{amb} = 25$ °C bliver temperaturen ved monteringsstedet på $T = 47$ °C.

Den termiske modstand af et køleprofil er almindeligvis fra 1 til 50°C/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°C/W som den termiske modstand fra den interne chip til omgivelserne. En småsignaltransistor som BC547 har tilsvarende en værdi på 250°C/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°C/W fra PN overgangen (*junction*) til monteringsflade (*mounting base*). Værdien opgives i databladet for effektransistoren.

Hvis monteringsfladen kan holdes på 47°C vil en transistor med 4°C/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*.

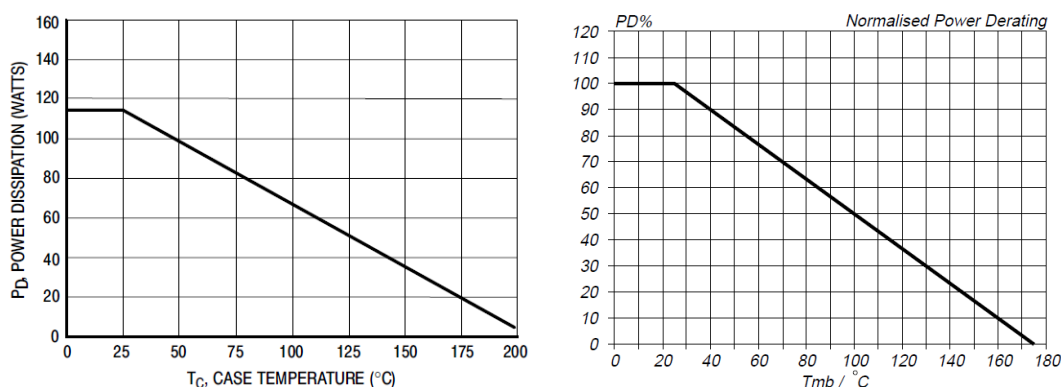


Figur 403 – Beregning af chippens temperaturstigning ved en transistor monteret på et køleprofil.

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 halvlederens 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.

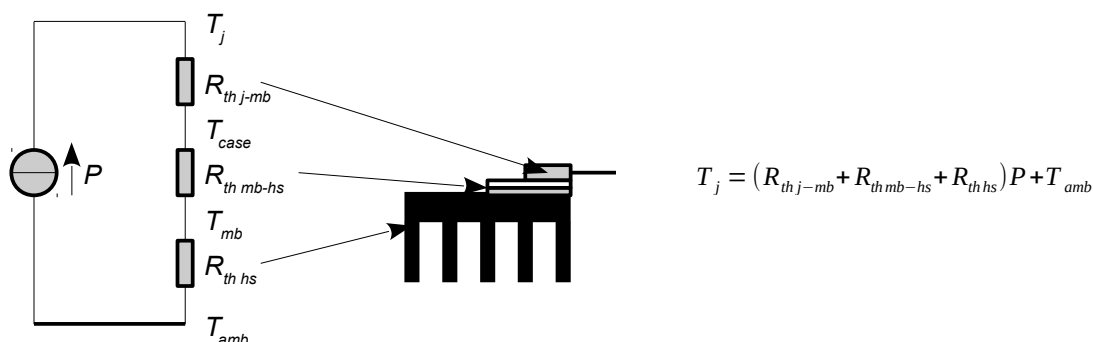
Herunder vises den maksimalt afsatte effekt som funktion af den temperatur køleprofilet giver ved monteringsstedet for to forskellige transistorer. Kurverne er typiske med en konstant effekt op til en temperatur på 25°C og med en værdi på 0 W ved en højere temperatur der for 2N3055 er noget over 200°C og for IRFZ44N er på 175°C .



Figur 404 – 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 temperaturforskel 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.

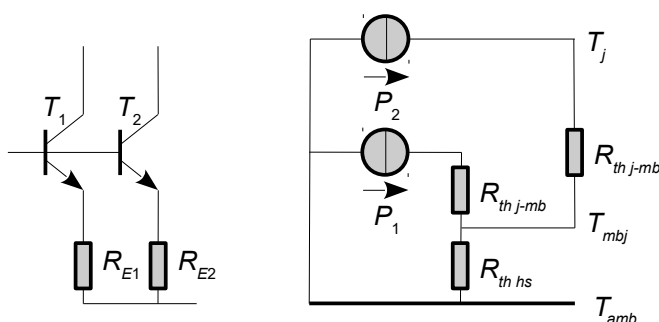


Figur 405 – Beregning af temperaturstigning ved en transistor monteret med isolationsskive.

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 vejforhold.

Hvis køleprofilet sidder udenpå kabinettet 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 kabinettet 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 reduceres.¹⁹⁴



Figur 406 – Beregning af temperaturstigning ved to transistorer i parallel.

Parallelkobling af effektt transistorer 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 mindsker 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.

¹⁹⁴Jeg har engang designet en brokølet 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.

6.2.9 Beregning af termisk modstand

Den afsatte effekt P i transistoren ledes ofte til køleprofilen 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 køleprofilen overalt har samme temperatur og som et beregningseksempel benyttes den store køleprofil fra side 355.

Varmen fra køleprofilen 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 køleprofilen opvarmes, stiger op og trækker kold luft til. Den termiske modstand af køleprofilen $R_{th HS}$ (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 konvektion. 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 $P = 1,3$ W + $3,0$ W = $4,3$ W (se eksemplerne nedenfor) beregnes $R_{th HS} = 5,8$ 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 køleprofilen betragtes som en kasse uden detaljerne ved køleribberne. Dertil T_K for køleprofilens temperatur i kelvin og T_{AMB} for omgivelsestemperaturen [AL-21 og Wikipedia].

$$P_R = e \sigma A_R (T_{HS}^4 - T_{AMB}^4) \quad \begin{array}{l} e_{\text{Maling}} \approx 0,9 \\ e_{\text{Poleret}} \approx 0,1 \end{array} \quad \sigma = 5,67 \cdot 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$$

For det store køleprofil er de ydre mål $22,5 \times 37,5 \times 54$ mm³ så $A_R = 0,0082$ m². Ved 50°C på køleprofilens overflade ($T_{HS} = 323$ K) og en omgivelsestemperatur på 25°C ($T_{HS} = 298$ K) afgives $P_R = 1,3$ W for en malet overflade der er antaget som repræsentativ for sort eloxering.

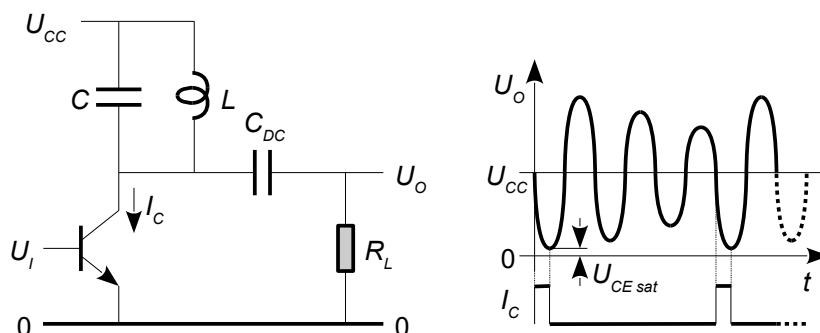
Konvektion betegnes med P_K for effekten overført til luften tæt på køleprofilen. Den afsatte effekt er givet ved det samlede areal A_K luften passerer langs med køleprofilens lodrette højde L . Værdierne indsættes i SI-enheder og effekten P_K fås i watt. Formlen er stærkt simplificeret [AL-16].

$$P_K \approx 1,2 \frac{A_K}{L^{0,25}} (T_K - T_{AMB})^{1,25}$$

For det store køleprofil er der 14 nogenlunde ens finner og dermed i alt 28 overflader og med et gæt på målene $20 \times 37,5$ mm² som gennemsnit for finnerne bliver arealet $A_R = 0,02$ m² og højden er $L = 37,5$ mm. Ved en overfladetemperatur på 50°C og 25°C ved omgivelserne bliver den afgivne effekt $P_K = 3,0$ W.

6.2.10 Klasse C

Dette princip anvendes primært indenfor radiotransmission hvor udgangssignalet er en højfrekvent svingning der enten føres til en antenne 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. Virkningsgraden er omkring 85 %. Er transistoren aktiv i 50 % af tiden er det en én-transistor klasse B, ved mere end 50 % af perioden er det klasse AB og når transistoren er aktiv i hele perioden arbejder den i klasse A [KBR-472].



Figur 407 – Et klasse C udgangstrin er aktivt i mindre end 50 % af perioden, og anvendes fx i radio-sendere, hvor en LC svingningskreds får tilført effekt i korte pulser for at opveje den afgivne effekt til en antenne. 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 $U_{CE sat}$ på cirka 0,5 V. 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 bære-bølge. 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.

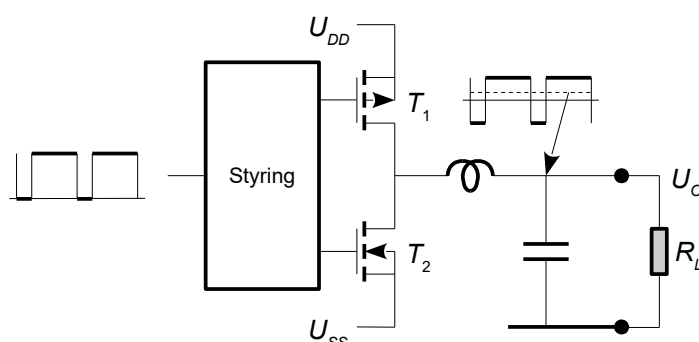
Klasse C kan anvendes til multiplikation af frekvensen fra et krystal. Svingningskredsen LC vil her afstemmes til anden eller tredje harmoniske af det pulserende signal så kredsen oscillerer på en højere frekvens end det indkommende signal.

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.

6.2.11 Klasse D

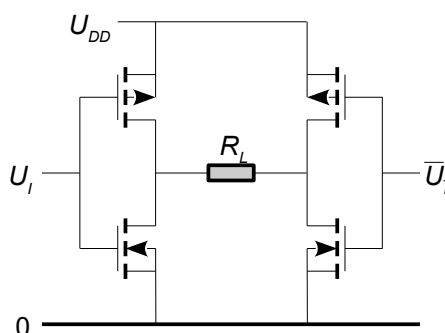
Princippet blev beskrevet af Baxandall i 1959 og Osborne i 1968 som en effektiv måde at generere en sinussvingning med høj effekt for radiotransmission og er senere udnyttet ved effektforstærkere til lyd gengivelse. Det er i teorien muligt at opnå 100 % virkningsgrad, men en praktisk opstilling vil typisk 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. Skiftfrekvensen er over 100 kHz for lyd gengivelse og gerne over 500 kHz hvis transistorerne er hurtige nok. Den integrerede elektronik i headsets og høreapparater kan have en lavere skiftfrekvens for at reducere effektforbruget ved skiftene hvor de interne kapaciteter i transistorerne skal op- og aflades.



Figur 408 – Et klasse D udgangstrin benytter transistoren som kontakt 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 og ved PA anlæg hvor udgiften til køling af elektronikken kan holdes nede selv om der afgives en effekt på flere kilowatt.



Figur 409 – 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 363) 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 lyd gengivelse og klasse D er givetvis fremtiden indenfor effektforstærkere.

6.2.12 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].

6.2.13 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].

6.2.14 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.

Note 6.

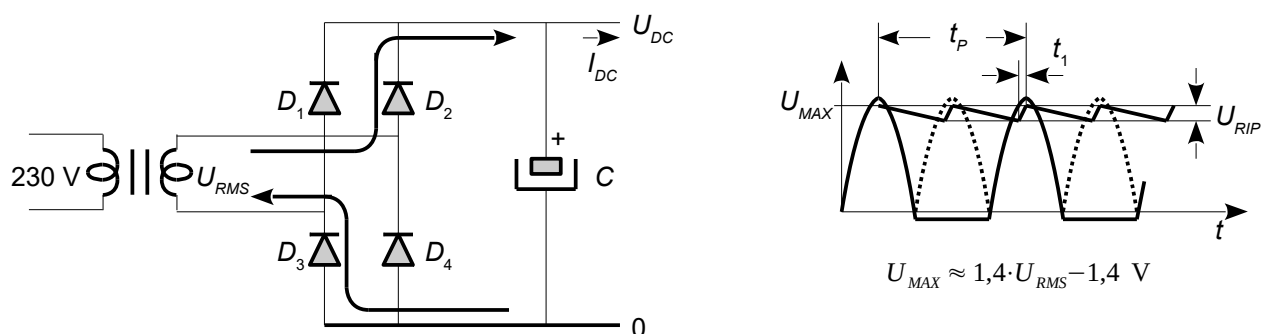
6.3 Effektforsyning

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 vekselspænding anvendes i stor udstrækning som effektforsyning og præsenteres 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 lydproduktion så omtalen af DC-DC konverteren vil kun berøre batteriforsynet 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.¹⁹⁵

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 udglattes.¹⁹⁶



Figur 410 – 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 lederetningen 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 spændingsfald.

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

¹⁹⁵Efter 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).

¹⁹⁶De 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.

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 $U_{MAX} - U_{RIP}$.

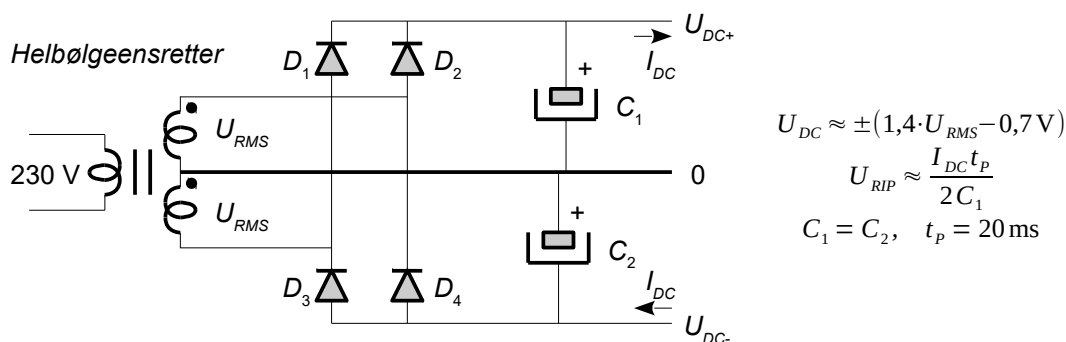
Kondensatorligningen kan benyttes til at beregne hvor meget spændingen falder i perioden med afladning. Det simpleste resultat nås hvis det antages at belastningsstrømmen er konstant over hele den analyserede periode hvilket i praksis er en glimrende approksimation.

$$\begin{aligned} dQ = C dU = i dt &\Rightarrow dU = \frac{i dt}{C} = \frac{I_{DC} t_P}{2C}, \quad t_P = 20 \text{ ms} \\ i = I_{DC} \end{aligned}$$

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 Dobbelt effektforsyning – helbølge

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 $\pm 15 \text{ V}$ til analog elektronik eller $\pm 25 \text{ V}$ til effektforstærkere og de kan dannes med en minimal ændring af kredsløbet.

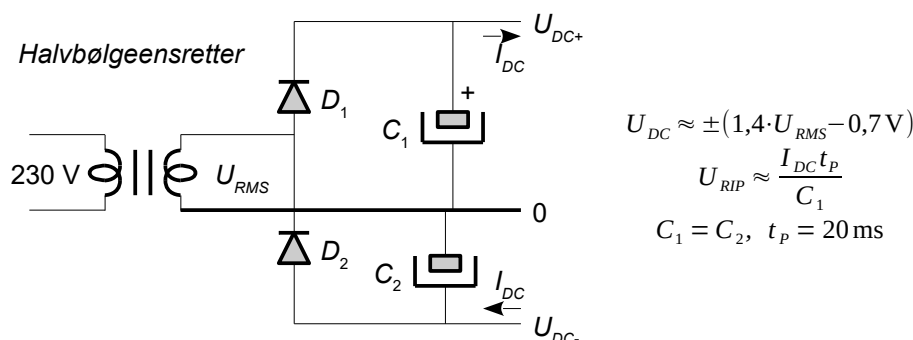


Figur 411 – En transformator skal ofte give to komplementære spændinger der som regel belastes lige hårdt og transformatoren leveres da med to ens sekundærviklinger.

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.

6.4.3 Dobbelt effektforsyning – halvbølge

Transformatorer med kun en enkelt vinding kan udmærket danne to komplementære spændinger som vist herunder. Der er kun en enkelt diode i hver udgang og kondensatorerne oplades kun ved 50 Hz så afladningstiden bliver dobbelt så lang. Ripplen bliver derfor dobbelt så stor sammenlignet med helbølgeensretteren ved den samme forbrugsstrøm.

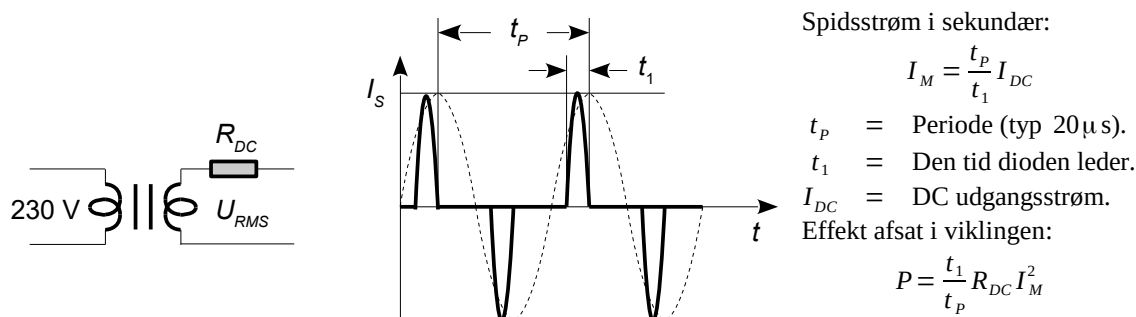


Figur 412 – En transformator med en enkelt vinding kan forsyne elektronik med både en positiv og en negativ forsyning, men de to forsyningslinjer bliver kun halvbølgeensrettet.

Der er som antydnet 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.

6.4.4 Effektforhold

Den afgivne effekt fra en effektforsyning kan ikke overstige den tilførte effekt minus tabene for ellers havde vi opfundet en evighedsmaskine.¹⁹⁷



Figur 413 – Strømmen i transformatorens sekundær løber kun i korte intervaller når dioderne er åbne, så strømmens topværdi bliver høj for at kunne levere forbrugsstrømmen.

Den pulserende strøm løber igennem transformatorens DC modstand R_{DC} fra beviklingen 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. På grund af de stærke harmoniske i strømmspidserne har EU begrænset denne type ensretter til effekter op til 75 W. Der er dog en lempelse hvis strømmspidserne kan holdes under et acceptabelt niveau.

¹⁹⁷ En 50 W transformator, der afgiver 25 V efter ensretning vil derfor højst 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.

Induktive komponenter til din konstruktion



DANTRAFO A/S

Fabriksvej 2, 7620 Lemvig

Tlf +45 76 28 29 30

mail dt@dantrafo.dk

www.dantrafo.dk



DANTRAFO Group

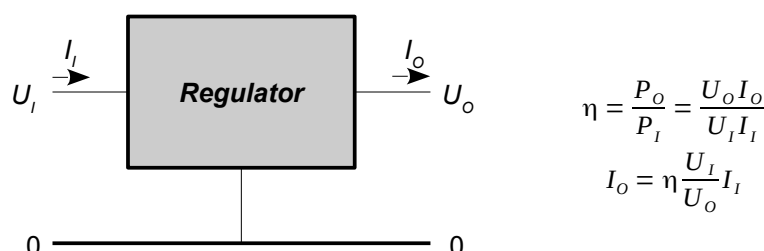
DANTRAFO A/S, Transformator-Teknik, Dantrafo Electronics Suzhou

6.5 Spændingsregulator

Ensretterens DC spænding er tilstrækkelig til effektforstærkere for lyd gengivelse 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 her skal gennemgås tre muligheder: en shuntregulator, en serieregulator og en DC-DC konverter. Listen 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 regulatorens virkningsgrad η (*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 414 – 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.

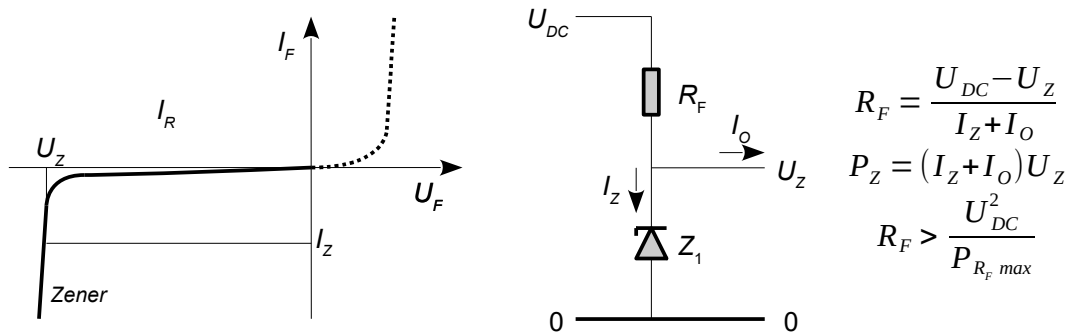
1. En shuntregulator har en meget ringe effektivitet da både udgangsstrøm og -spænding er lavere end indgangssidens ($U_o < U_i$ og $I_o < I_i$) så det er sjældent at virkningsgraden kan komme ret meget over 25 %.
2. Ved en serieregulator løber strømmen så godt som direkte til belastningen ($I_o \approx I_i$), 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 populær ved omsætning af store effekter 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 reguleret spænding 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 et integreret kredsløb.¹⁹⁸

En alternativ brug af zenerdioden er at beskytte imod overspænding, og der findes specialtyper som er udviklet til at optage energi ved transient overbelastning. De kaldes for transient absorbere og markedsføres som TranZorb og TranSil som omtales side 370.



Figur 415 – 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 levere 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 komponenterne.¹⁹⁹

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.

$$\Delta U_Z = \frac{R_Z}{R_F + R_Z} \Delta U_{DC}$$

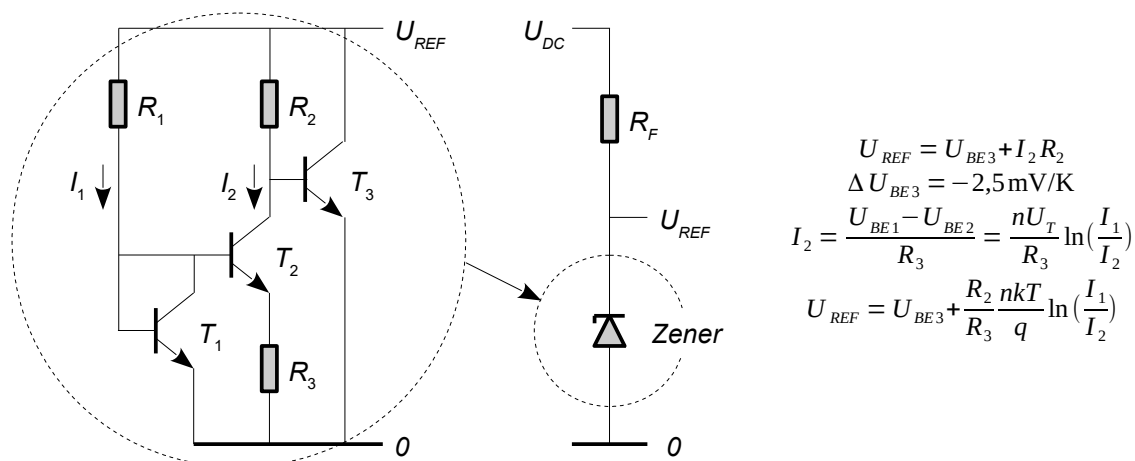
Med $R_F = 3 \text{ k}\Omega$ og $R_Z = 10 \Omega$ vil en ændring af effektforsyningens spænding med $\Delta U_{DC} = 1 \text{ V}$ give en variation af udgangsspændingen på $U_{Z \text{ RIP}} = 3,3 \text{ mV}$.

¹⁹⁸Integrerede 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.

¹⁹⁹Man 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 som spændingsreference på helt samme måde som en zenerdiode.



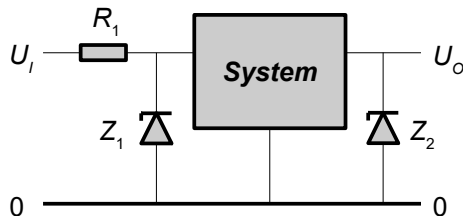
Figur 416 – Et kredsløb med nogenlunde samme funktion som en zenerdiode. Det kræver dog at transistorerne er meget nær ens og at de har samme temperatur.

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 $-2,5 \text{ mV/K}$. 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 cirka 20 mV ved en fordobling og det svarer til en indre modstand på 20Ω ved $I_Z = 1 \text{ mA}$.

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 seriemodstand eller af det ydre kredsløb selv.

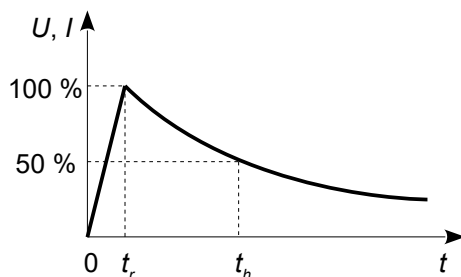


Kategori	Spænding	Modstand
	U_{SURGE}	R_{SURGE}
Kommerciel	500 V	42 Ω
Industri (signal)	500 V	2 Ω
Industri (DC linje)	1000 V	2 Ω

Figur 417 – Beskyttelse af ind- og udgang imod transient overspænding. Strømmen i zenerdioden kan begrænses af en modstand eller alene af det ydre kredsløb (testgeneratorens indre modstand). 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 spærreretningen 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 modstand 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.



Type	Stigetid t_r	Halveringstid t_h
ITU 10/1000	10 μ s	1000 μ s
EMC spænding	1,2 μ s	50 μ s
EMC strøm	8 μ s	20 μ s

Figur 418 – 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 transienter 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 forbrugeren. Det forventes at transienten kobler mellem parallelførte ledere i kabelbakkerne og har ledt til en række krav til signal- og forsyningslinjer 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.

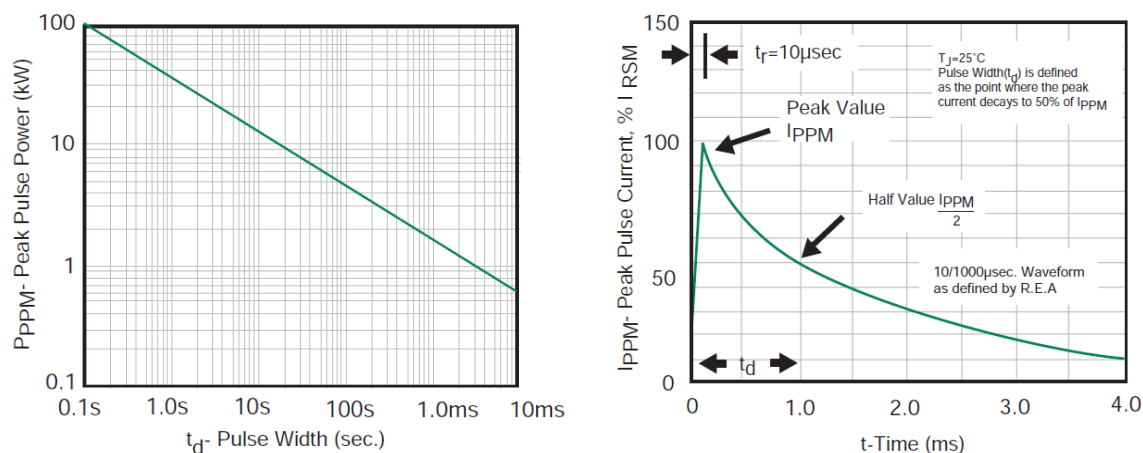
Lyntransienten er defineret som en kortvarig ikke-repeteret puls der stiger til maksimum i løbet af kort tid og derefter klinger eksponentielt ud. Pulsen ligner meget en transient fra telefonstandard ITU, men forløbet er meget hurtigere og der er en specifikation for både spænding- og strømpulser, men jeg vil kun komme ind på spændingspulsen i denne bog. Lyntransienten beskrives ved dens topværdi $U_{SURGE} = 500 \text{ V}$ 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 skrapere 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 hvorved 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 lignende og diodens zenerspænding kaldes for dens *clamp voltage*.²⁰⁰

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 varmestråling. 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 zenerdiodens evne til momentant at oplagre energi.



Figur 419 – 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).

EMC direktivets lyntransient er meget kortere end ITU standardens transient så energien bliver 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.

²⁰⁰ 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 μ s kan den optage 6 kW, men den afsatte energi aftager. Data er verificeret eksperimentelt ved måling på dioder med en *surge generator*.²⁰¹

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 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.

$$R_1 \geq \frac{U_{SURGE} U_C}{P_{PPM}} - R_{SURGE} \quad \Rightarrow_{R_1=0} \quad U_C \leq \frac{R_{SURGE} P_{PPM}}{U_{SURGE}}$$

Modstanden belastes hårdt så man må bruge en modstand specielt for absorbering af transient energi. Herunder vises data for to trådviklede modstande 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 μ s hvilket opfyldes af EMC direktivets lyntransient.²⁰²

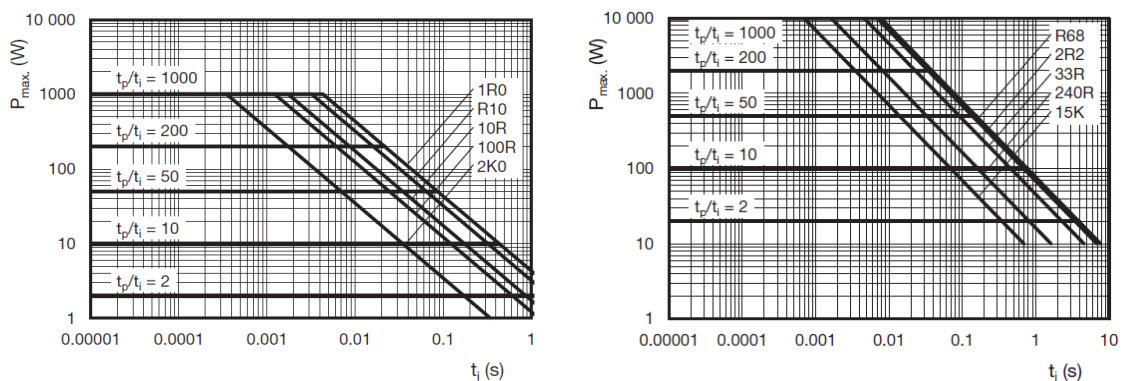
Pulsens effektive varighed findes ved at sætte arealet under en rektangulær puls med en varighed på t_p lig med arealet under den eksponentielt hændende transient hvor opstartens korte forløb er ignoreret da det kun drejer sig om at finde en enkel relation mellem de to pulser.

$$\int_0^{t_p} dt = \int_0^{\infty} \exp\left(-\ln(2) \frac{t}{t_h}\right) dt \Rightarrow t_p = \frac{1}{\ln(2)} t_h \Rightarrow t_p = 1,4 t_h$$

Resultatet er at en rektangulær puls skal regnes som $t_p = 70$ μ s 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.

²⁰¹ 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 homogent og angav at dioderne ved EMC direktivets lyntransient kunne optage 22 % af energien fra ITU transienten. Publiceret i EMC Avisen nummer 15, 1995.

²⁰² AC01 er 4,3 mm i diameter og længden er 11 mm. AC10 har diameteren 8 mm og længden 44 mm hvortil kommer tråddenderne. Modstanden dannes af en metaltråd der vikles omkring en keramisk kerne. Effekbelastning er opgivet ved 40°C og overfladens temperatur stiger med 100°C for AC01 og 320°C for AC10 ved fuld belastning (Vishay).



Figur 420 – 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

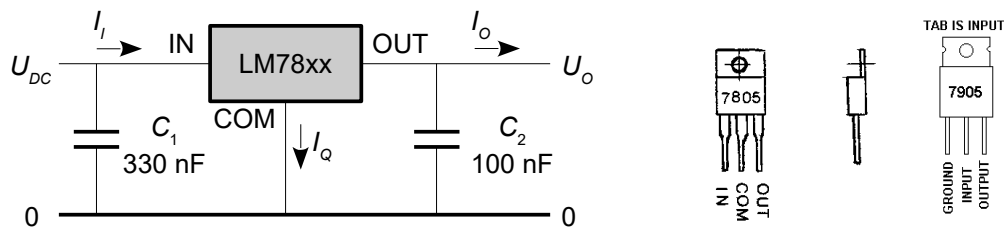
Følgende algoritme kan benyttes som vejledning for et design.

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 dioden. 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 formelen for R_1 . Hvis formelen angiver en negativ værdi skal der ikke benyttes en seriemodstand.
4. Beregn modstandens effektklasse fra formelen 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 og meget større end en 150 W zenerdiode så hvis der i stedet blev valgt en diode med $P_{PPM} = 400$ W vil seriemodstanden kunne undværes og der kan spares plads på printkortet.

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 tilnærmelsesvist proportional med udgangsstrømmen. Regulatorer af 7800-typen kan klare op til cirka 1,5 A og findes også for negative spændinger som LM7900. 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 selv.²⁰³



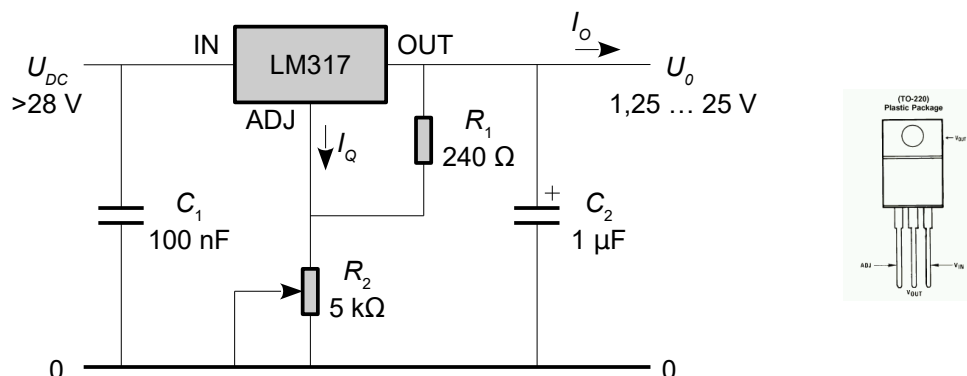
Figur 421 – Den typiske opstilling med en serieregulator nedsætter spændingen fra en ustabil effektforsyning til en stabil og kortslutningssikker forsyningslinje. De viste kondensatorer er krævet af hensyn til stabilitet og det anbefales kraftigt at studere databladet for den valgte komponent. Bemærk 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 som vist herunder hvor $U_{REF} = 5\text{ V}$. Strømmen I_Q kan nå 8 mA ved 78-typen hvilket besværliggør brugen.

$$U_O = \left(1 + \frac{R_2}{R_1}\right) U_{REF} + I_Q R_2 \quad \text{og} \quad U_{DC} > U_O + 2\text{ V}$$

Regulator LM317 har $U_{REF} = 1,25\text{ V}$ mellem udgang og reference og $I_Q \approx 50\text{ }\mu\text{A}$ gør den særdeles velegnet til at justere udgangsspændingen med to faste modstande. Forbrugsstrømmen fra det interne kredsløb føres nu til udgangen så strømmen i I_O må ikke komme under 10 mA.

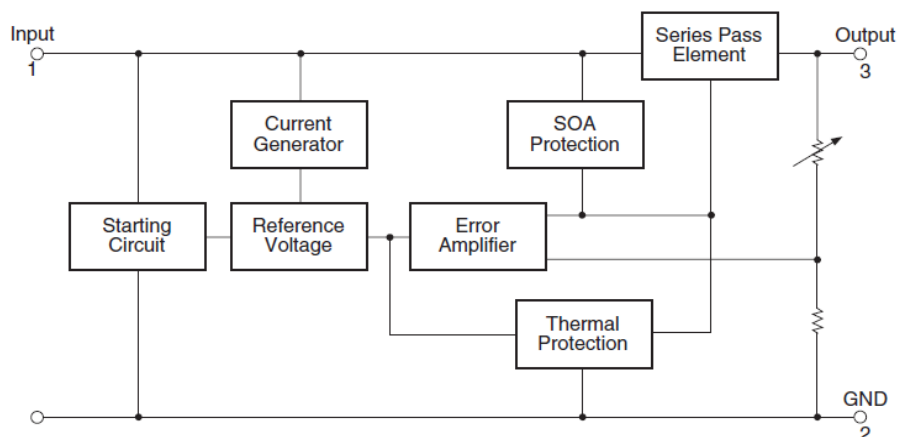


Figur 422 – Variabel udgangsspænding med LM317 (National Semiconductor).

²⁰³Der findes et rigtholdigt 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. Som et eksempel 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 forbrugeren løber igennem. Det er en NPN transistor med emitteren til udgangen, og den styres af en forstærker der sammenligner en neddelte version af udgangsspændingen med en reference der ofte er på 1,25 V. Ved 7800 er den dog opbygget anderledes som det vises senere. 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 423 – 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.

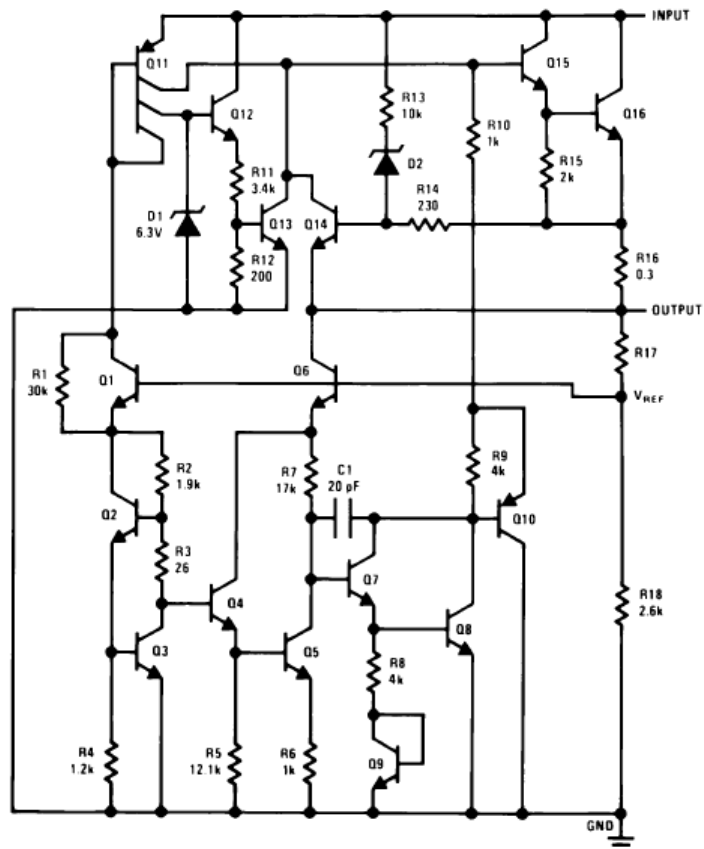
Serieelementet Q16 skal bære en ret stor strøm så den har Q15 som strømforstærker og basis af Q15 ligger cirka 1,4 V over den indstillede udgangsspænding hvilket er en væsentlig grund til at databladet foreskriver at der skal ligge mindst 2 V fra ind- til udgang.

Strømmen i Q16 giver et spændingsfald over R16 der med 0,3 Ω 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 mV/°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 fejlforstærkeren 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 fejlforstærkeren 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 424 – 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 μA 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 μA . Det kan nu konkluderes at spændingsfaldet over R7 bliver 1,7 V og derfor vil referencespændingen blive på 3,5 V på grund af basis-emitter spændingsfaldet ved Q6. Strømmen i Q3 kan nu beregnes til 850 μA 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 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 fejlforstærkeren. Det er dog kun spekulation.

Shuntregulatoren har et konstant effektforbrug på $U_{\text{DC}}I_o$ uanset om der trækkes strøm ud af den eller ej og det gør den kun egnet til simple småopstillinger. Serieregulatoren har et lavt internt

forbrug på 8 mA for 78-typerne og mindre for LM317 så effekttabet er stort set proportionelt med udgangsstrømmen I_O .

$$P_D = U_{DC} I_Q + (U_{DC} - U_O) I_O$$

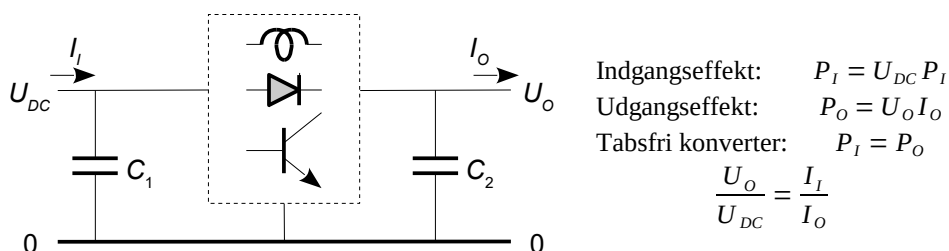
For LM7805 med $U_O = 5$ V og $I_Q < 8$ mA er effekttabet $P_D = 2,2$ W ved indgangsspændingen $U_{DC} = 25$ V og en udgangsstrøm på $I_O = 100$ mA. Af effektafsættelsen kommer de 0,2 W fra det faste effekttab.

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 .

6.8 DC-DC konverter

En serieregulator taber 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 oplagring i en kondensators elektriske felt. I modsætning til de lineære kredsløb i en serieregulator er diode og transistor her kontakter som styrer energitransporten.



Figur 425 – 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 bipolar eller felleffekt.

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 computerchippet i en PC. Der kan opnås en virkningsgrad over 95 % hvilket er umuligt ved serieregulatoren. Konverteren kan hæve eller sænke spændingen, den kan skabe isolation mellem ind- og udgang som for udstyr tilsluttet bynettet og den kan konstrueres for at levere flere forskellige udgangsspændinger.

Konverterens funktion baseres på at strømmen i en spole ikke kan ændres momentant. Strømmen vil vokse op når der lægges en spænding over spolen og den vil aftage hvis polariteten ændres.

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.

Opkonvertering (*boost converter*) benyttes eksempelvis ved effektforsyning fra en enkelt battericelle 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.²⁰⁴

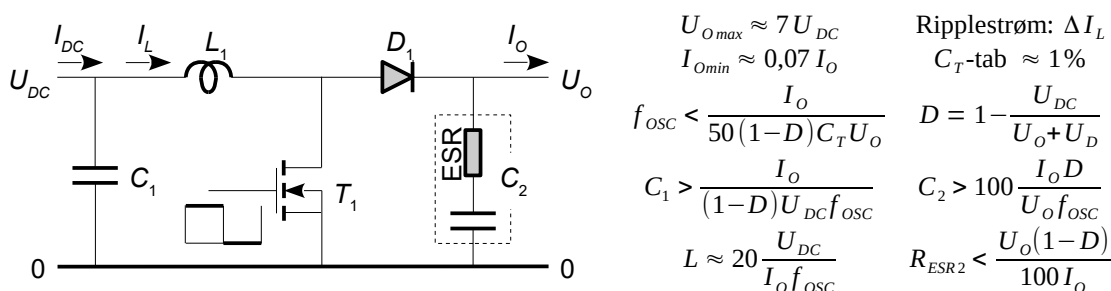
Nedkonvertering (*buck converter*) benyttes i headsets og høreapparater hvor levetiden af et Li-Ion batteri skal være så lang som muligt samt i en PC hvor effektforsyningens 12 V lokalt omsættes til processorens arbejdsspænding på cirka 1,7 V da en strømstyrke på 50 A ikke er velegnet til intern distribution gennem kabler, stik og printbaner.

Der er ikke den store forskel fra de beskrevne konvertere til konvertering mellem DC og AC som benyttes i motorstyring og effektoverførsel gennem undersøiske kabler eller omsætning mellem to AC niveauer med hver sin frekvens som benyttes i vindmøller og elforsyning generelt, men det ligger udenfor rammerne af denne bog hvor DC-DC konverteren blot skal introduceres.

²⁰⁴Cypress PSOC komponenter kan startes op fra en enkelt celle på blot 0,9 V ved en indbygget DC-DC konverter og konverteren kan følge med 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 startes. Med transistoren ON er spolen forbundet direkte over effektforsyningen U_{DC} så strømmen i spolen vokser lineært med andelen ΔI_L i tidsrummet t_{ON} . Med transistoren OFF vil strømmen fra spolen løbe gennem dioden til udgangen så kondensatoren oplades over tidsrummet t_{OFF} og samtidigt aftager strømmen i spolen med ΔI_L . Strømmen i spolen svinger med ΔI_L omkring middelværdien I_L og kondensator C_1 udjævner den oscillerende strøm så indgangsstrømmen I_{DC} bliver tilnærmelsesvis konstant.



Figur 426 – Udgangsspændingen U_O hæves over indgangsspændingen U_{DC} givet ved parameter D .

Svingningen i strøm omkring den nominelle værdi (*ripple*) har amplituden ΔI_L . Oscillatorens frekvens f_{OSC} vælges af designeren fra en maksimalt tilladelig værdi. Dioden bør være af typen Schottky for at holde spændingstabet 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.

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 regulering af udgangsspændingen behøves et kredsløb der måler spændingen og indstiller D og det vil normalt være den samme mikroprocessor som genererer PWM signalet.

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 384 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.

Designeksempel for konvertering fra $U_{DC} = 3,3 \text{ V}$ til $U_O = 5 \text{ V}$ ved en strøm på $I_O = 100 \text{ mA}$.

Mindste belastningsstrøm bliver 7 mA så udgangen bør ikke være ubelastet. Dioden vælges til en Schottky med et spændingsfald på $U_D = 0,3 \text{ V}$ og oscillatoren skal give $D = 0,38$. Transistoren, spolen og dioden antages at have en kapacitet på $C_T = 1 \text{ nF}$ så frekvensen er højest 640 kHz og her vælges $f_{OSC} = 100 \text{ kHz}$. De to kondensatorer bliver $C_1 > 490 \text{ nF}$ og $C_2 > 7,6 \text{ }\mu\text{F}$ hvor der vælges to ens kondensatorer på 10 μF . Den ækvivalente seriemodstand er $R_{ESR2} < 0,3 \text{ }\Omega$ for C_2 og det er til at opfylde. Spolen findes til $L = 6,6 \text{ mH}$ som afrundes til 5 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. Fasen slutter med en strøm I_2 i spolen som danner startbetingelsen når transistoren går OFF. Her aftager strømmen med et bidrag der tilsvarende kaldes for ΔI_{OFF} og ender ved startværdien I_1 for den første fase. Herefter vil hele forløbet blot gentages i det uendelige.

Transistor ON. Transistoren forbinder spolens højre side til nul så dioden spærrer 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 definitions ligningen $u_L = L_1 di_L/dt$.

$$-U_{DC} + u_L = 0 \Rightarrow L_1 \frac{di_L}{dt} = U_{DC} \Rightarrow di_L = \frac{U_{DC}}{L_1} dt$$

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 0,3 V for Shottky dioder (0,7 V for siliciumdioder). 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_O til belastningen vil strømmen i spolen svinge op og ned omkring en 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.

$$\Delta I_{ON} + \Delta I_{OFF} = 0$$

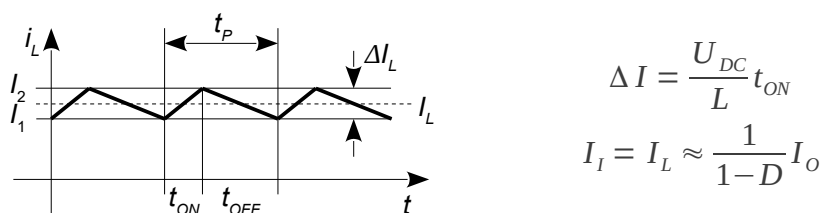
Resultatet er at udgangsspændingen U_o kan udtrykkes som funktion af indgangsspændingen U_{DC} og parameteren D . Bemærk at hverken spolens selvinduktion, de to kondensatorer eller den frekvens som oscillatoren arbejder på indgår i ligningen. Det betyder at man har stor frihed til at vælge disse parametre 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.

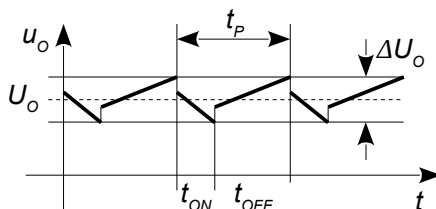
Det resulterende forløb i den stationære tilstand vises herunder for tre perioder ud af et forløb med et meget stort antal ens perioder. Den stiplede linje står for middelværdien af strømmen i spolen I_L og er lig med den strøm der optages fra effektforsyningen U_{DC} .



Figur 427 – Strømmen i spolen svinger i den stationære tilstand omkring en middelværdi I_L med en spids-spids værdi på ΔI der er styret af hvor lang tid transistoren er ON og OFF.

Ignorerer 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 seriemodstand ESR (se senere).

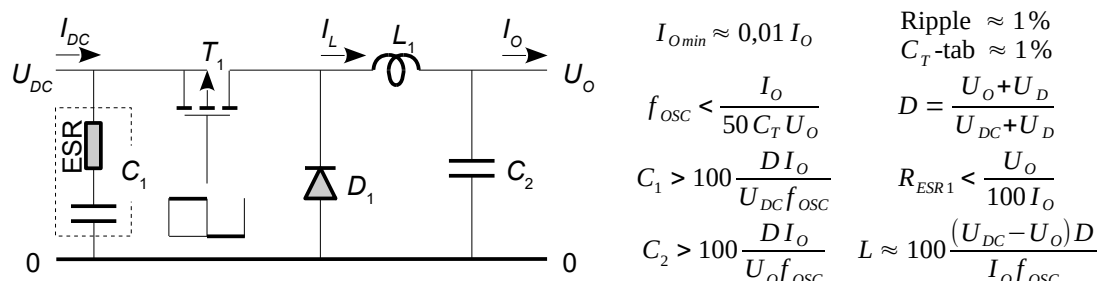


Figur 428 – Den resulterende spændingsprofil på udgangen af konverteren består af lineære segmenter fra op- og afladning af kondensatoren, samt spændingsspring fra den skiftende strømretning i kondensatorens ækvivalente modstand.

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

Omroking 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.

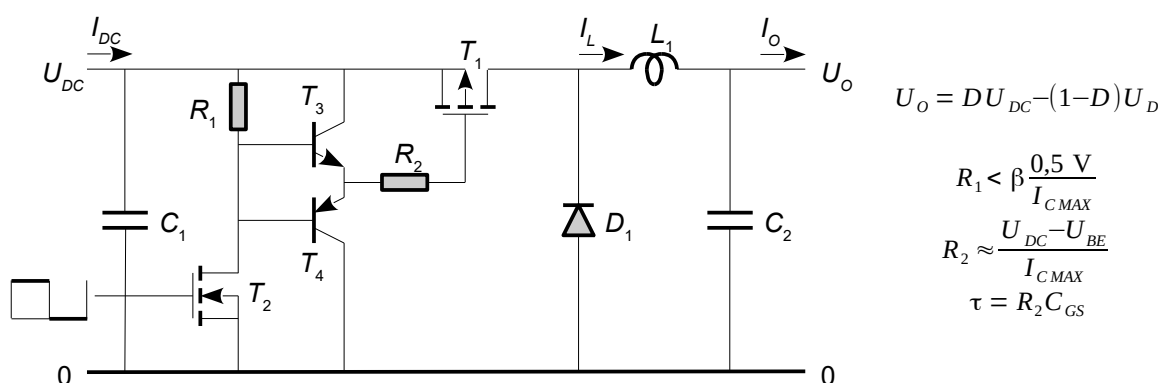


Figur 429 – Udgangsspændingen U_o sænkes til D gange indgangsspændingen U_{DC} og spolens strøm er lig med belastningsstrømmen.

Svingningen 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 levere 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 op mod 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. Et eksempel vises herunder hvor T_2 fungerer som inverter med R_1 og omsætter det logiske niveau til U_{DC} niveauet.



Figur 430 – Udgangsspændingen U_o sænkes til D gange indgangsspændingen U_{DC} og spolens strøm er lig med belastningsstrømmen.

Den indre kapacitet i en MOSFET er høj ($C_{GS} \approx 1 \text{ nF}$) og det forsinker omskiftningen af T_1 så der er indsat et klasse B effektrrin ved T_3 og T_4 der accelererer op- og afladning af kapaciteten. R_1 skal trække baserne af T_3 og T_4 til U_{DC} indenfor cirka 0,5 V. Strømmen i T_3 og T_4 begrænses af R_2 og med BC547 og BC557 hvor topværdien er $I_{C_{MAX}} = 100 \text{ mA}$ bliver $R_2 = 100 \Omega$ ved $U_{DC} = 10 \text{ V}$. Tidskonstanten er $\tau = 100 \text{ ns}$ og vil først give problemer ved en oscillator væsentligt over 100 kHz. Et alternativ er at benytte en IC der produceres netop for dette formål.

6.10.1 Analyse

Forløbet er det samme som før med kredsløbsligninger for hver af faserne. Med transistoren ON spærres 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 $L di_L/dt$ og på udgangen er spændingen U_O .

$$\begin{aligned} \text{ON:} \quad -U_{DC} + L \frac{di_L}{dt} + U_O &= 0 & \Rightarrow \quad \Delta I_{ON} &= \frac{U_{DC} - U_O}{L} t_{ON} \\ \text{OFF:} \quad U_D + L \frac{di_L}{dt} + U_O &= 0 & \Delta I_{OFF} &= -\frac{U_D + U_O}{L} t_{OFF} \end{aligned}$$

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.

$$U_O = D U_{DC} - (1 - D) U_D \Rightarrow D = \frac{U_O + U_D}{U_{DC} + U_D}$$

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 udgangsstrømmens mindst tilladelige værdi blive for høj så belastningen ikke kan 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 $U_{Xop} = U_O$ $U_{Xned} = U_{DC}$
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.

$$P_C < \frac{U_O I_O}{100} \Rightarrow \frac{1}{2} C_T U_X^2 f_{OSC} < \frac{U_O I_O}{100} \Rightarrow f_{OSC} < \frac{U_O I_O}{50 C_T U_X^2} \quad \text{hvor} \quad \begin{matrix} U_{Xop} = U_O \\ U_{Xned} = U_{DC} \end{matrix}$$

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	Oscillatorfrekvens
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	$\Delta I_{ON} = \frac{U_{DC}}{L_1} t_{ON}$	$L_1 = \frac{D U_{DC}}{\Delta I_L f_{OSC}}$
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 $U_{DC} = 5 \text{ V}$ til $U_O = 15 \text{ V}$ bliver $D = 0,68$. Med $\Delta I_O = 10 \text{ mA}$ og $f_{OSC} = 1,3 \text{ MHz}$ bliver spolen $L = 250 \text{ }\mu\text{H}$.

For ned-konvertering fra $U_{DC} = 10 \text{ V}$ til $U_O = 5 \text{ V}$ bliver $D = 0,5$. Med $\Delta I_O = 10 \text{ mA}$ og $f_{OSC} = 400 \text{ kHz}$ bliver spolen $L = 625 \text{ }\mu\text{H}$.

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 tilnærmelsesvist 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.

$$dq = C du = i dt \Rightarrow C > \frac{i dt}{du} = \frac{I_O t_{ON}}{U_{RIP}} = \frac{D I_O}{U_{RIP} f_{OSC}}$$

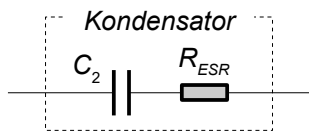
For en op-konverter med $I_O = 100 \text{ mA}$, $D = 0,68$ og $U_{RIP} = 10 \text{ mV}$ beregnes $C_2 > 5 \text{ }\mu\text{F}$ ved $f_{OSC} = 1,3 \text{ MHz}$.

For en ned-konverter med $I_O = 100 \text{ mA}$, $D = 0,5$ og $U_{RIP} = 10 \text{ mV}$ beregnes $C_1 > 12 \text{ }\mu\text{F}$ ved $f_{OSC} = 400 \text{ kHz}$.

6.11.4 Kondensatorens ESR

Kondensatoren har en ækvivalent seriemodstand ESR (*equivalent series resistor*) der optræder som en seriemodstand. 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 $0,1 \text{ }\Omega$ 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 431 – Den ækvivalente seriemodstand (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 seriemodstand 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 378. 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 $U_{DC} = 5\text{ V}$ til $U_O = 15\text{ V}$ med $U_{RIP} = 10\text{ mV}$ og $I_O = 100\text{ mA}$ kræves $R_{ESR} < 0,03\ \Omega$.

For ned-konvertering fra $U_{DC} = 10\text{ V}$ til $U_O = 5\text{ V}$ med $U_{RIP} = 10\text{ mV}$ og $I_O = 100\text{ mA}$ 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 $(U_O/U_{DC})I_O$ 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 kommercielt tilgængelige kredse for opbygning af DC-DC konvertere tilbyder mulighed for at arbejde i denne tilstand under betegnelser som *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. Det er ikke nødvendigt at opstille ligningerne påny for at aflure konverterens tab for det er oftest tilstrækkeligt med en idé om hvor store tabene er for at vide hvor man skal forbedre kredsløbet.

Type	Transistor ON	Transistor OFF	Modstandstab	Indre modstand
Op	$(R_L + R_T) I_L^2$	$(R_L + R_D) I_L^2$	$P_R = R_{T_{op}} I_O^2$	$R_{T_{op}} = \frac{R_L + D R_T + (1-D) R_D}{(1-D)^2}$
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 der omregnet fra strømmen i spolen I_L til strømmen i udgangen I_O og heraf kvadratet på $1 - D$.

Bemærk at der ikke medregnes tab fra spændingen over dioden eller kondensatorens ESR. Der er heller ikke medtaget tab fra de parasitiske kapaciteter. Værdien af R_T er derfor kun vejledende.

7 APPENDIKS

Her er samlet information af referencekarakter.

7.1 Konstanter

<i>Symbol</i>	<i>Værdi</i>	<i>Definition</i>	<i>Kommentar</i>
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

<i>Symbol</i>	<i>Enhed</i>	<i>Definition</i>	<i>Kommentar</i>
A	m^2		Areal
A	-		Forstærkning
α	-		Generel talværdi
B	T	$T = Wb/m^2$	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	Vm^{-1}		Elektrisk feltstyrke
e_n	V/ \sqrt{Hz}		Spændingsstøjtæthed
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

Appendiks

i_n	A/√Hz		Strømsstøj tæthed
J	A/m ²		Strømtæthed
L	H = Wb/A = Vs/A		Selvinduktion
L	dB	$20 \log_{10}(p_1/p_{REF})$	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	$C = As$	Elektrisk ladning
Q	-	$Q = 1/2d$	Godhedsfaktor
R	$\Omega = V/A$	$R = U/I$	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	$\Delta T = T_2 - T_1$	Temperaturdifferent, 1 K = 1 °C
t	m		Tråddykkelse
Δt	s	$\Delta t = t_2 - t_1$	Tidsinterval
θ	radian		Vinkel
U, u	V		Spændingsforskel
W	J	$W = \int p dt$	Energi (<i>work</i>) = 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.²⁰⁵

Et fælles enhedssystem sikrer at handel mellem to lande kan gennemføres uden stridigheder om hvordan de forskellige landes mål og vægt skal omsættes. Det økonomiske incitament ved den stigende handel førte til nedsættelse af metrologiske udvalg i de enkelte lande som varetager implementering og sporbarhed. Et stigende samarbejde mellem landene om tekniske landvindinger nødvendiggør fælles enheder for at undgå kostbare uheld på grund af trivielle misforståelser.²⁰⁶

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) defineres 1793 som en ti milliontedel af afstanden fra nordpol til ækvator. Masseenheden gram (g) defineres samtidigt som massen af vand ved 0°C med et volumen på en milliontedel af en kubikmeter. Den 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 enhedssystemet var der tilføjet strømstyrken ampere. Senere indføres stofenheden mol og lysenheden candela. Længdeenheden er siden redefineret som et antal bølglængder af en karakteristisk lyskilde og tidsenheden er nu defineret ved frekvensen af lyset fra en bestemt lyskilde. Enhederne for tid og afstand er derved kombineret i en fælles definition ud fra lysets hastighed. Der pågår arbejde med at redefinere de øvrige enheder ud fra basale fysiske konstanter.²⁰⁷

7.4.2 Enhed

Følgende grundenheder er defineret.

Enhed for ...	Symbol	Navn	Definition	Årstal
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 trippelpunkt er 273,16 K = 0,01 °C.	1954
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

²⁰⁵https://en.wikipedia.org/wiki/International_System_of_Units.

²⁰⁶https://en.wikipedia.org/wiki/Mars_Climate_Orbiter.

²⁰⁷<http://www.kvant.dk/upload/kv-2012-1/kv-2012-1-kv121-FBR-kg.pdf>.

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 $N = kg\ m\ s^{-2}$ så enheden for modstand bliver $\Omega = kg\ m\ s^{-2}\ A^{-2}$.

En enhed er et symbol og ikke en forkortelse så der skal ikke punktum efter et enhedssymbol med mindre det står sidst i en sætning.

7.4.3 Præfix

Et præfix er en skalering i spring af 1000 og dækker i øjeblikket området fra 10^{-18} til 10^{18} .

Præfikset skrives ligesom enheder med lodret skrift, det sættes altid foran enheden og SI systemet tillader ikke sammenstilling af enheder, så μm tillades ikke som erstatning for p (pico). Listen vist herunder viser ikke definitionen af centi, deci, deca og hecto da SI anbefaler at undgå disse.

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 et produkt af de to.

$$\text{Symbol} \quad x = \text{Talværdi} \quad \{x\} \cdot \text{Enheden} \quad [x]$$

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.

7.5 Omregning

Det er muligt at repræsentere de forskellige enheder ved grundenhederne 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 · 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 · 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

En ofte set fejl er at benytte en matematisk funktion på en fysisk størrelse uden at tage hensyn til enheden, Det giver ikke nogen mening hvis funktionen indvirkning på enheden ikke kan defineres. Et eksempel er ved brug af logaritmer. Hastigheden $V = 36 \text{ km/h}$ giver $\log(V) = 1,56$. Den samme hastighed kan udtrykkes som $V = 10 \text{ m/s}$ der giver $\log(V) = 1,00$. De to resultater burde være ens.

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 *Højttalere*, 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, Pretience 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.....15
0 dB båndbredde.....171

1

1/f-støj.....190f.
10/1000 puls.....370

3

3 dB grænsefrekvens.....226

4

4-tråds måling.....11
455 kHz.....259

A

Absolut temperatur.....184
Absolute Maximum Ratings.....353
Afkobling.....271
Aliasing.....126f.
AM.....258
Ampere timer.....48
Amperes lov.....51
Analog signalbehandling.....264
Analyse.....
 Afsat effekt.....13
 Elektrisk ladning.....43
 Ohms lov.....9
Anders And.....4
Annonce.....
 Dantrafo.....66, 366
 IDA.....2
 PR electronics.....197
Anode.....336
Antiparallelle dioder.....258
Arbejdslinje.....352
Arrhenius.....353
Atan2.....127

B

Balance.....153
Basisjordet.....299
Batteri.....48
Beskyttelsesdiode.....284
Beskyttelseskredsløb.....353
Biasstrøm.....168, 190
Bin.....127
Binær repræsentation.....126
Biot og Savarts lov.....57
Blyulfid.....251
BNC.....62
Bode plot.....225
Bonding.....252
Boost converter.....378f.
Boot strap.....296, 344, 351
Brick-wall filter.....192
Brinck elektronik.....4

Buck converter.....378, 382
Buffer.....154, 295
Butterworth.....231
Båndbredde.....91

C

CFP.....349
Chassis.....15
Chopper stabilised amplifier.....167
Chopperstabilisering.....323
Chorus.....201
cis(θ).....93
Clamp voltage.....371
CMOS.....303
 74AC04PC.....211
 74HC04.....211
 74HC132.....210
 74HC14.....210
 HEF4007.....284, 304
 HEF4069UB.....304
CMRR.....159f., 312
Cosinus.....86
Cross-over forvrængning.....206, 346
Current dumping.....347

D

Darlington pair.....295, 349
DC lys.....270
DC-DC konverter.....378
 Discontinuous mode.....386
 Nedkonvertering.....382
 Opkonvertering.....379
Decibel.....74
Depletion layer.....251, 329
Derating.....13
Designprocedure.....
 BJT forstærker.....296
 JFET forstærker.....301
Diameter.....86
Dielektrikum.....34
Dielektrisk absorption.....32, 228
Dielektrisk konstant.....46
Differentialkvotient.....96
Differentiation.....96
Diffusion.....251
Diffusionslængde.....332
Diode.....250, 328
 1N4007.....255
 1N4148.....254f., 263
 1N5819.....254, 263, 334
 AA119.....254
 BAS116.....254
 BB207.....266
 BP104.....214
 BPW21.....214
 BPW34.....214
 BYV59C.....256
 LED.....263

SFH203.....269
SFH485.....267
Diodemikser.....260
Diodemodstand.....262
Diodespænding.....262
Diskret Fourier transformation.....128
Dominerende pol.....170f.
Dotering.....327
Driftsstrøm.....327
Duty cycle.....379
Dynamikområde.....181
Dynamisk mikrofon.....217
Dynamisk modstand.....262

E

Early voltage.....292
ECL.....278
EE.....349
Effekt.....13
 Modstande.....13
 Udgangstrin klasse A.....343
 Udgangstrin klasse B.....354
Effektivværdi.....70
Eksponentialfunktion.....93
Elektret mikrofon.....217
Elektretmikrofon.....35
Elektrisk felt.....44
Elektrisk konstant.....34, 44
Elektrisk strøm.....43
Elektro-dynamisk højttaler.....68, 147
Elektromagnetisme.....50
Elektronrør.....336
ELF.....63
EMC direktivet.....370
Emitterfølger.....295
 CFP og EE.....299
Emittermodstand.....295
Energi.....13
Enhed.....389
 ampere.....43
 bel.....74
 coulomb.....43
 dBm, dBu, dBV.....75
 farad.....46
 henry.....52
 kelvin.....184
 kelvin.....184
 neper.....74
 Radianer per sekund.....86
 tesla.....50
 weber.....51
Enkeltensretning.....223
Epitaxial layer.....252
Eulers identitet.....106, 120
Europakort.....34

F

Fakultet.....103

Indeks

Faradays induktionslov.....	52
Fasedrejningsfilter.....	235
Fasehastighed.....	90
Fasemargin.....	177, 204
Feed back.....	202
Feed forward.....	361
Fejlspænding.....	166
Ferrit.....	65
Filter.....	225
Anden orden.....	230
Butterworth.....	230, 240
Chebychev.....	241
Design.....	242
Første orden.....	226
Højere orden.....	239
Højpas transformation.....	242
Passiv.....	244
Firetråds kobling.....	11
Fluorescens.....	324
Flux-vindingstal.....	52
FM.....	258
Formler.....	3
DC-DC konverter.....	382
Decibel.....	74
Diodespænding.....	255
Diodestrøm.....	253, 332
Egenstøj.....	319
Elektroner per rumfang.....	327
Energi.....	40
Energi (spole).....	40
Fermi-Dirac fordeling.....	326
Forstærker med BJT.....	296f.
Forvrængning.....	309
Godhed.....	41
Halvledermodstand.....	327
Hybrid-pi model.....	288
Indbygget potentiale.....	329
Kapacitet.....	46
Kondensatorligning.....	47
Magnetisk induktion.....	51
Minoritetsbærere.....	330
RMS algoritme.....	73
Selvinduktion.....	37
Brooks coil.....	56
Dobbeltleder.....	53
Enkelt vinding.....	59
Ledning.....	53, 62
Maxwell.....	59
Toroid med luftspalte.....	65
Wheeler (enkeltlag).....	55
Wheeler (flere lag).....	56
Slew rate.....	318
Strømfortrængning.....	63
Temperaturspænding.....	254
VAS.....	314
Forstærket diode.....	348
Forvrængning.....	205, 289
Fotodiode.....	213
Fotolitografi.....	252
Foton.....	324
Fourier.....	118
DFT.....	128
Transformation.....	122
FR4.....	34
Frekvens.....	86
Frekvensbånd.....	91
Frekvensområde.....	193
Fuzz box.....	206
Fysisk størrelse.....	78
Fælles basis.....	299
Fælles emitter.....	276
Fælles kollektor.....	295
G	
Gasudladning.....	371
GBP.....	171, 306
Geometrisk middelværdi.....	265
Germanium.....	251
Gibbs fænomen.....	119
Gitter.....	336
Glimmerskive.....	356
GND.....	15
Godhed.....	91, 230
Graetz kobling.....	363
Ground.....	15
Grundtone.....	119
Gruppetidsforsinkelsen.....	236
Grænseværdi.....	91
H	
h-parameter model.....	300
Haglstøj.....	184, 319
Harmonisk forvrængning.....	205
Cross over.....	346
Radiator.....	342
Harmonisk række.....	119
Harmonisk svingning.....	86
HDTV.....	118
Henry.....	36
Howland strømkilde.....	219
Hvid støj.....	181, 184
Hvirvelstrøm.....	65
Hybrid-pi model.....	288
Hybride kredsløb.....	199
Højrehåndsregel.....	50, 67
Højspændingskabel.....	63
Højtalere.....	68
I	
Idealitetsfaktor.....	254
BC547.....	274
IEC.....	15
IGBT.....	382
Impedans.....	91, 95
Indgangskapacitet.....	156
BJT.....	290
OpAmp.....	316
Indgangsmodstand.....	157
BJT.....	290
MOSFET.....	283
OpAmp.....	316
Indre modstand.....	19
Indtrængningsdybde.....	63, 272
Instrumentationsforstærker.....	160
Integration.....	99
Intel 4004.....	152
Isolator.....	10
ITU.....	370f.
J	
JFET.....	281
Josty Kit.....	4
K	
Kapacitet.....	34, 46
JFET gate.....	281
Skærmkabel.....	47
Tolederkabel.....	46
Katalogfirmaer.....	255
Katode.....	336
Kelvinkobling.....	11
Kirchhoffs love.....	16
Kirchhoffs strømlov.....	16
Kirchhoffs spændingslov.....	16
Klasse A.....	342
Klasse AB.....	348
Klasse B.....	346
Klasse C.....	359
Klasse D.....	361
Kobberoxyd.....	251
Kompleks Fourier serie.....	120
Komplekse tal.....	92
Komplementær transistor.....	273
Komplementært signal.....	159
Kondensator.....	30
Beregning.....	34
Elektrolyt.....	32, 363
Keramisk.....	31
Mikrofon.....	35
NP0.....	31
Parasitisk.....	33
Plastfilm.....	32
Printkort.....	34
Variabel.....	33
X7R.....	31
Kondensatorligningen.....	318
Kondensatormikrofon.....	35
7.1Konstanter.....	387
Boltzmann.....	326
Elementarladning.....	43
Elementarladningen.....	324
Permeabilitet.....	51
Permittivitet.....	46
Specifik modstand.....	10, 327
Vacuumpermittivitet.....	44
Kontron.....	4
Konvektion.....	358
Kraftfaktor.....	67, 147
Kredsløb.....	
78-type.....	376
Addition af signaler.....	156
Beskyttelsesdioder.....	258
Båndpasfilter.....	229, 233
Chebychev lavpasfilter.....	242
Differensforstærker.....	158f.
Differentialtrin.....	308
Differentiator.....	162
Effektforsyning.....	363
Elektretmikrofon.....	35
Fasedrejningsfilter.....	235
Forstærker med BJT.....	286, 296
Forstærker med BJT og buffer.....	296, 298

Indeks

Forstærker med CMOS.....	303	Li-ion batteri.....	19	AD846.....	322
Forstærker med JFET.....	301	Lige funktion.....	106, 119	AD8628.....	166f.
Forstærker med MOSFET.....	302	Ligesvævende temperatur.....	212	ADA4051.....	152, 166, 195
Fotodiode.....	213	Limes.....	96	CA3080.....	321
Geometrisk middelværdi.....	265	Linearitet.....	17, 77	L165V.....	152, 199
Halvbølge ensretning.....	258	Lokaloscillator.....	261	LF356.....	166, 168, 198
Helbølgeensretter.....	223	Luftspalte.....	65	LF357.....	204
Højpasfilter.....	161	Lyntransient.....	370	LM13700.....	321
Højpasfilter, anden orden.....	232	Lyserød støj.....	191	LM2902.....	307
Højpasfilter, første orden.....	228	Lysets hastighed.....	50	LM324.....	152, 166, 168, 198, 307
Instrumentationsforstærker.....	160	Lyttetræthed.....	346	LM358.....	307
Inverterende forstærker.....	156	Lækstrøm.....	250, 254	LM386.....	312
Klasse C udgangstrin.....	359	M		LM3886.....	152
Kontakt, BJT.....	27	Magnetfelt.....	50	LM3900.....	292, 322
Kontakt, MOSFET.....	285	Jorden.....	50	LM675T.....	152, 199
Lavpasfilter.....	163	Magnetisk induktion.....	50	LMH6624.....	195, 204
Lavpasfilter, anden orden.....	231	Centrum.....	58	NE5532.....	168, 188
Lavpasfilter, første orden.....	227	Kort leder.....	57	NE5534.....	152, 318
LM317 serieregulator.....	374	Magnetisk modstand.....	65	OP27.....	166, 168, 319
Lysdetektor.....	270	Majoritetsbærer.....	330	TDA2040.....	152
Lysmåler.....	214	Maksimalt fladt filter.....	230	TDA7294.....	152
Mikrofonforstærker.....	217, 286	MATLAB.....	129	TL071.....	152, 160, 168, 195, 198, 318
Mikroprocessor PWM.....	69	Mayer-Norton ækvivalent.....	19	TLC271.....	152, 195, 199, 313
Mikroprocessorer PWM.....	69	Mekanisk kraft.....	45	TLC274.....	168, 209
Mikser.....	260	Middellevetid.....	324	uA702.....	152
OpAmp.....	315	Middelværdi.....	69	uA741.....	166, 168, 173, 198
Oscillator, CMOS.....	210	Mikrofoni.....	228	Operationsforstærker.....	
Oscillator, EXP styret.....	212	Mikser.....		Udgangsmodstand.....	176
Oscillator, Krystal.....	211	Balanceret.....	260	Opfinder.....	
Oscillator, OpAmp.....	209	Diode.....	260	Alan Blumlein.....	152
Oscillator, spændingsstyret.....	211	Gilbert celle.....	313	Bardeen, John.....	23
Piezo-elektrisk accelerometer.....	216	Miller transformation.....	287	Brattain, Walter.....	23
PT100 temperaturføler.....	11	Minimoog.....	212	Ferdinand Braun.....	20
Radiomodtager.....	259	Minoritetsbærer.....	330	H. C. Ørsted.....	50
Spændingsforstærker.....	155	Mobilitet.....	273, 327	Heaviside, Oliver.....	65
Spændingsfølger.....	154	Modstand.....	10	John Bardeen.....	152
Spændingsstyring.....	263	0603.....	368	Julius Edgar Lilienfeld.....	152
State-variable filter.....	238	AC01.....	13, 372	Lilienfeld, Julius.....	28
Stopfilter.....	237	SFR16.....	368	Maxwell, J. C.....	50
Strain gauge (brokobling).....	12	Modulus.....	120	Oliver Heaviside.....	130
Strømkilde, AC.....	219	MOSFET.....	88	Schockley, William.....	23
Strømkilde, DC.....	218	Motor boating.....	297	Walter Brattain.....	152
Strømkilde, JFET.....	281	Mættet.....	153	Overtone.....	119
Strømspejl.....	311	Måling af.....		P	
Subtraktion af signaler.....	158	mekanisk deformation.....	12	Pasbånd.....	225
Termokobler.....	215	strøm.....	258	Peltier element.....	363
Termometer.....	264	N		Pentode.....	337, 342
Toneregulering.....	221	Norton amplifier.....	322	Periode.....	86
Udgangstrin.....	349	Norton ækvivalent.....	19	pH-probe.....	214
Variabel forstærkning.....	220	Nul.....	15	Phantom Power.....	35
VU-meter.....	224	Nyquist.....	72, 127	Phasing.....	201
Kvarts.....	251	Stabilitet.....	177	Philips EE20.....	4, 342
Kæderegel.....	98	Nyquists stabilitetskriterium.....	204	Pierce oscillator.....	211
L		O		Piezo-elektrisk krystal.....	216
Ladning.....	43	Octave.....	129	Planarteknologi.....	252
Ladningsforstærker.....	216	Ohms lov.....	9	Popcorn støj.....	196
Laplace.....	130	Omkreds.....	86	Positiv retning.....	15
Transformationspar.....	139	OpAmp.....	153, 306	Potentiale.....	45
Laser.....	268, 324	AD797.....	152, 168, 188, 196, 293, 319	Potentiel energi.....	40, 45
Latch-up.....	198	AD823.....	209	Printkort.....	10
LC led.....	105			Producenter.....	255
LC svingningskreds.....	359			Propagation delay.....	211
Ledningsbånd.....	324				

Indeks

Præfix.....390
PT100.....11
Push pull.....199, 343f.
PWM.....359f.

Q

Q.....91
QUAD 405.....347

R

Radiator.....42, 152, 336, 342
 12AX7.....337
 ECC83.....337
 EL84.....342
Radiotransmission.....258
Radius.....86
Regulator, 7800.....374
Regulator, LM317.....374
Regulering, DC-DC.....379
Regulering, zener.....368
Relativ permeabilitet.....64
Relativ permittivitet.....34
Reluktans.....65
Resistivitet.....10
Resonans.....39
Resonansfrekvens.....91, 148
Ripplespændingen.....364
RMS.....70
RMS addition.....185
Root mean square.....70
Rosa støj.....181, 191
RS422.....159
Rør.....336

S

Sampling.....126
Sampling frequency.....126
Samplingfrekvens.....127
Schmitt-trigger.....209
Seebeck element.....363
Sekundærbatteri.....48
Selen.....251
Selvinduktion.....37
 Kinetisk energi.....40
Sensivitetsanalyse.....105
Settling time.....170
Shannon.....72, 127
Shottky-diode.....332
SI enhedssystem.....389
Signal-støj forhold.....181
Silicium.....251
Siliconeskive.....356
Sinc.....122
Sinus.....86
Skineffekt.....63
Slew rate.....173, 314, 318
Slutværdisætningen.....141
Sløjfeforstærkning.....202
Sløjfeforstærkningen.....204
SMA.....62
SMD.....9
Småsignalmodel.....286, 304
 BJT.....288
 JFET.....301

SOA.....353
Sollys.....267
Solpanel.....269
Specifik ledningsevne.....327
Specifik modstand.....327
Specifik varmeledning.....358
Spejlfrekvens.....259
Spektrum.....127
Spændingsdeler.....16
Spændingsfølger.....154
Spændingsreference.....
 ADR4520.....257
 LM4051A.....257
 Widlar kredsløb.....369
Stagger-tuned filter.....259
Standardfordeling.....183
Startværdisætningen.....141
State-variable filter.....238
Stationær tilstand.....90, 94, 380
Stel.....15
Stelplan.....62, 247
Steradian.....267
Stevens lov.....74
Stigetid.....173
Stopbånd.....225
Strain gauge.....12
Strømforstærkning.....273, 275
Strømkilde.....218, 312
Strømretning.....9
Strømtæthed.....51
Strålingsintensitet.....267
Støjtal.....288, 293
Støjtæthed.....184
Subtraktor.....158
Superposition.....17
Surge Transient.....370
Svingningskreds.....39
Svingspole.....68
Switch-mode converter.....378

T

Taylor polynomium.....103
Temperaturkoefficient.....10
 1N4148.....255
 Bipolær transistor.....274
 Kondensator.....32
 Strømforstærkning.....275
 Transistor.....273
 Zenerdiode.....257
Termisk modstand.....355
Termisk støj.....184
Termoelement.....187
Termometer.....255
 PT100.....11
THD.....206
Thevenin ækvivalent.....19
Tid.....86
Tidskonstant.....48
Tilbagekobling.....202, 297
Tilbagekoblingsfaktor.....172
Tomgangsspænding.....19
Tommelfingerregel.....
 Banelængde.....62
 Effektætsættelse.....13

Fasemargin.....177
Gate sammenbrud.....283
GBP specifikation.....246
Harmonisk forvrængning.....289
Kapacitet.....246
Kapacitet til omgivelser.....248
Lav støj.....190
Parasitiske komponenter.....248
Sampling.....127
Selvinduktion af leder.....248
Strøm i basis.....27
Støj.....188
Udgangsmodstand, BJT.....300
Åben-sløjfe forstærkning.....165
Transformator.....66
TranSil.....368, 371
Transistor.....272
 2N3055.....344
 AC126.....342
 BC547.....355
 IRFZ44N.....88, 355
 MJ480.....344
 TO220.....355

Transistor.....
 2N3819.....282
 BC547.....273, 275, 277
 BC547B.....263
 BC549.....293
 BC847.....273, 275
 BF245.....281f.
 BF545.....281f.
 JFET auto-biasing.....281
 Maksimal forstærkning.....292
 Mætning.....277
Transitfrekvens.....171, 291, 306
Transkonduktans.....
 BJT.....286, 289, 309
 CMOS.....304
 Differentialtrin, BJT.....308
 ECC83.....337
 JFET.....301
 JFET og MOSFET.....310
 LM13700.....321
 MOSFET.....302
Transmissionsbånd.....225
TranZorb.....368, 371
Triode.....336
TTL.....278

U

Udgangsmodstand.....180
 BJT.....278
 CMOS.....304
 Hybrid-pi model.....292
 JFET.....302
Uendelig baffel.....147
Ulige funktion.....106, 119
Undervandsbåd.....63
USB.....159

V

Valens.....251
Valensbånd.....324
Valenselektron.....42, 324

Indeks

Varistor.....	371
Varmestråling.....	358
Vektorprodukt.....	67
Vinkel.....	86
Vinkelhastighed.....	86, 93
Virkningsgrad.....	367
Radiorør.....	342
Virtuelt nul.....	157
Volt.....	45
Volumehastighed.....	148

W

Wireless World.....	344
---------------------	-----

Z

Zener.....	257
1.5KE.....	371
BZX84.....	368
LM129.....	368
LM368.....	368
LT1021.....	368

MC1400.....	368
Widlar kredsløb.....	369
Zero padding.....	129

Æ

Ækvivalent seriemodstand.....	385
Ækvivalent støjbandbredde.....	191

Å

Åben-sløjfe forstærkning.....	165
-------------------------------	-----