

17 lektioner i TELEKOMMUNIKATION

Per Wallander

PERANT

Per Wallander
17 lektioner i TELEKOMMUNIKATION
Första upplagan
ISBN 91-86296-10-8

Innehållet i denna bok är skyddat enligt Lagen om upphovsrätt, 1960:729, och får inte reproduceras eller spridas i någon form utan skriftligt medgivande av Per Wallander, som är copyrightinnehavare. Förbudet gäller hela verket såväl som delar av verket och inkluderar tryckning, kopiering, överföring till annat medium etc.

Copyright © Per Wallander 2001

Tryckt i Sverige / Printed in Sweden 2001
Fälth & Hässler, Smedjebacken

Förlag

PERANT

Per Wallander Antenn AB
Strandängsstigen 2, S-125 52 ÄLVSJÖ, SWEDEN
e-mail: per@perant.se
<http://www.perant.se>

INNEHÅLL

1 — Sinusspänning	9
1.1 Sinusformade spänningar och strömmar	10
1.2 Fasvinkel är alltid en tidsförskjutning	11
1.3 Lägga ihop, dra ifrån och multiplicera	14
1.4 Derivera och integrera	16
1.5 $j\omega$ -metoden	18
1.6 Vad innebär "stationär" växelström?	19
1.7 Om man ändå vill lösa differentialekvationen	20
1.8 Decibel	22
2 — Fourierserier och spektrum	25
2.1 Fourierserier	26
2.2 Signalens spektrum	28
2.3 Medeleffekt under en period	29
2.4 Effektspektrum	30
2.5 Två sätt att alstra fyrkantvågen	31
2.6 Hur fungerar Fourierserietvecklingen?	33
2.7 Ett matematiskt "filter"	34
2.8 När $u(t)$ är periodisk	35
2.9 Jämn eller udda funktion?	36
2.10 Fourierserietveckling i praktiken	37
2.11 Korta pulser	38
2.12 Komplexa Fourier-serier, Fourier-transformen	41
2.13 Några vanliga funktioner	42
2.14 Vad är egentligen ett spektrum?	44
3 — Brus	45
3.1 Termiskt brus	46
3.2 Många generatorer	47
3.3 Motstånd brusar	48
3.4 Gaussiskt brus	50
3.5 Brus i förstärkare	52
3.6 Brusförhållandet NR och brusfaktorn $NF(F)$	54
3.7 Bestämning av förstärkarens brusfaktor	55
3.8 Flera förstärkare	56
3.9 Signal/brus-förhållandet före detektorn	58
3.10 Brusfri resistans	62

4 — Amplitudmodulering – AM	63
4.1 Basband och linjekodning	64
4.2 Modulering	65
4.3 Amplitudmodulering i tidsplanet	68
4.4 Amplitudmodulerade signalens spektrum	70
4.5 Verkliga informationssignaler	72
4.6 AM-signalens demodulering	74
4.7 Verklig AM-modulering	76
4.8 Effekt i AM-signalen	77
5 — Amplitudmodulering – VSB, DSB, SSB, QAM	79
5.1 Stympat sidband, VSB (Vestigial Sideband)	80
5.2 Dubbla sidband, DSB (Double Sideband)	82
5.3 DSB-signalens demodulering	86
5.4 Enkelt sidband, SSB (Single Sideband)	88
5.5 QAM (Quad Amplitude Modulation)	90
5.6 Signal/brusförhållandet efter detektorn	92
5.7 FM-stereo — exempel på DSB	93
5.8 Analog FDM-telefoni — exempel på SSB	96
6 — Superheterodynmodulering	97
6.1 AM-mottagaren	98
6.2 Decibel och förstärkning	104
6.3 Signal/brusförhållandet före detektorn	106
7 — Frekvensmodulering och fasmodulering	109
7.1 Argumentmodulering	110
7.2 Frekvensmodulering i tidsplanet	112
7.3 Frekvensmodulering och matematiken	114
7.4 Frekvensmodulerade signalens spektrum	116
7.5 Frekvensmodulerade signalens bandbredd	118
7.6 Varför så många spektralkomponenter?	120
7.7 FM-signalens visardiagram	122
7.8 Om FM-signalen matas genom ett BP-filter?	123
8 — FM-demodulering och faslåst loop	125
8.1 FM-signalens demodulering	126
8.2 FM-signalen och brus	128
8.3 Bruströskel vid FM-demodulering	130
8.4 Preemphasis och deemphasis	131
8.5 Faslåst slinga (loop)	132
8.6 Faslåst loop som FM-demodulator	134
8.7 Faslåst loop som lokaloscillator	135
9 — Datakommunikation	139
9.1 Protokoll, gränssnitt (Interface) och OSI	140
9.2 A/D-omvandling och samplingsteoremet	142
9.3 PCM-kodning (Puls Code Modulation)	144
9.4 PCM-kodning i verkligheten	146
9.5 Symbolhastighet och bandbredd (Hartley)	148
9.6 Antalet olika symboler ger bithastigheten	151
9.7 Bithastighet och brus (Shannon)	152

10 — Fysiska lagret	153
10.1 Fasta telenätet	154
10.2 Linjekodning	158
10.3 Nyckling (Keying) och signalvektorer	159
10.4 ASK (Amplitude Skift Keying)	160
10.5 Hur hög mottagen effekt P_m behövs?	162
10.6 Antipodal 2ASK och BPSK	163
10.7 Icke-koherent 2FSK (Frequency Shift Keying)	165
10.8 Koherent 2FSK, MSK (Minimum Shift Keying)	166
10.9 QAM (Quad Amplitude Modulation)	168
10.10 PSK (Phase Shift Keying)	169
11 — Länklagret	171
11.1 Kontrollsumma, CRC-beräkning	172
11.2 Felrättning, Hammingkodning	176
11.3 V.24, hårdvaruhandskakning	178
11.4 SDLC/HDLC, mjukvaruhandskakning	180
11.5 Kretskopplat nät	185
11.6 Paketförmedlande nät	186
11.7 Cellförmedlande nät, ATM	188
12 — Ledningar	189
12.1 Radiostrålning	190
12.2 Vad händer när det "flyter" ström?	191
12.3 Signalens hastighet på ledningen	192
12.4 Ledningens karakteristiska impedans Z_0	193
12.5 Framåtgående effekt och anpassning	195
12.6 Reflektad effekt	196
12.7 Reflexionskoefficient	197
12.8 "Masken"	198
12.9 Går det ström?	202
13 — Smithdiagrammet, anpassning	203
13.1 Våglängd	204
13.2 Stående vågor	205
13.3 Stående vågförhållande	206
13.4 Att mäta stående vågförhållandet	208
13.5 Smithdiagrammet	212
13.6 Impedansbestämning med slitsad ledning	216
13.7 Anpassning med kvartsvågstransformator	218
13.8 Vad händer vid anpassning?	221
13.9 Anpassning med kortsluten stubbe	222
13.10 Öppen stubbe eller kondensator	224
13.11 Anpassning i verkligheten	226
14 — Antenner	227
14.1 Radiostrålning från en elektron (laddning)	228
14.2 Hertz experiment och resonans	230
14.3 Dipolantenn	234
14.4 Polarisering	236
14.5 Olika dipolantenn	237
14.6 Mätning av strålningsdiagram	239

Innehåll

17 lektioner i TELEKOMMUNIKATION

14.7	Riktantenner	240
14.8	Direktivitet D	242
14.9	Antennvinst (gain) G	244
14.10	Effektiv mottagningsarea A_e	245
14.11	Mottagarantenner	246
15	— Radiokommunikation	247
15.1	Radiosändaren	248
15.2	Radiomottagaren	249
15.3	EIRP	250
15.4	ERP och V/m	251
15.5	Länkbudget	252
15.6	Brus	256
16	— Optokommunikation	259
16.1	Tre typer av optofibrer	260
16.2	Våglängd eller frekvens?	262
16.3	Optofibers dämpning	263
16.4	Kromatisk dispersion	264
16.5	Optoförbindelsens bandbredd	266
16.6	Våglängdsmultiplexering, WDM (Wavelength Division Multiplexing)	268
17	— En lek med bilder	269
17.1	Kondensatorn	270
17.2	Kondensatorn som transmissionsledning	271
17.3	Plattledningens "uppladdning"	272
17.4	Kondensatorns uppladdning	273
17.5	Att parallellkoppla två kondensatorer	274
17.6	Spolen som ledning	275
17.7	Koppling mellan lindningsvarven	276
17.8	E och D , H och B	277
17.9	Maxwells ekvationer	278
17.10	Varifrån kommer magnetiska fältstyrkan?	283
17.11	Varifrån kommer elektriska fältstyrkan?	284
17.12	Vad är en elektriskt laddad partikel?	285
17.13	Vad händer när det åskar?	286
17.14	Vad är en magnet för något?	287
17.15	Varför attraheras partiklar med olika laddning?	288
17.16	Varför attraheras neutrala partiklar (tyngdkraft)?	289
17.17	Hur snabbt är ljushastigheten?	290
17.18	Laddning i rörelse, magnetfält och massa	291
17.19	Radiostrålning och gravitationsstrålning	294
17.20	Batteriet	295
BILAGA	— Formelsamling	297
INDEX		309

FÖRORD

För många år sedan, i tv-programmet tekniskt magasin, intervjuade Erik Bergsten den gamle Chalmersprofessorn Olof Rydbeck, grundaren av Råö rymdobservatorium, som sade så här:

— "Jag har arbetat hela livet med elektriska fält och magnetfält. Men vad är egentligen ett elektriskt fält? Vad är ett magnetfält?"

Den starkaste drivfjädern att bli lärare i telekommunikation kan knappast vara att man tycker att det är roligt att undervisa. Jag tror snarare att man blir lärare i telekommunikation därför att man själv inte begriper. Finns det någon effektivare metod att själv komma underfund med ett ämne, än att stå inför åhörare och berätta?

17 lektioner i TELEKOMMUNIKATION innehåller 16 avsnitt som ingår i "telekommunikation grundkurs" vid KTH:s högskoleingenjörsutbildning i Haninge. Kapitel 17 innehåller det som jag lärt mig genom att stå framför er studenter och prata.

Om jag inleder en lektion med att ställa upp matematiska uttryck, då får jag genast frågan "vad håller vi egentligen på med?" Dessa studenter behöver bilden.

Om jag börjar med bilderna så tycker andra studenter att jag är virrig och ostrukturerad.

Det är min förhoppning att kapitel 17 kan tjäna som inledning till en matematisk behandling av elektrodynamiken. För min egen del så tycker jag i alla fall att jag vet vad ett magnetfält är. Och ett elektriskt fält med för den delen. Det som syns när vi har framåtgående och reflekterade signaler.

Älvsjö i augusti 2001



PS.: Vad är en signal för något?

1 — Sinusspänning

Det viktiga i detta kapitel är att förstå vad en sinusformad växelspanning $\hat{u} \cos 2\pi ft$ är för något.

När den sinusformade spänningen får en fasvinkel $\hat{u} \cos(2\pi ft - \varphi)$ så innebär detta att spänningen förskjuts på tidsaxeln. Positiv fasvinkel innebär att den sinusformade spänningen kommer lite tidigare, negativ fasvinkel att spänningen kommer lite senare.

Sinus och cosinus är samma sak, bara 90° fasskillnad (tidsförskjutning). Ofta utgår vi från cosinus eftersom multiplikation av två cosinusspänningar ger cosinusspänningar som svar. Vid sinus får man blandat d.v.s. spänningar som ligger förskjutna i förhållande till varandra, något som komplicerar beräkningarna.

Kapitlet avslutas med decibel, viktigt, eftersom vi räknar på effektförstärkning och effektdämpning, inte spänningsförstärkning.

1.1 Sinusformade spänningar och strömmar

Sinus och cosinus uppfanns för att de behövdes vid navigering efter stjärnorna. Vi använder sinus- och cosinusfunktionerna för att t.ex. beskriva läget för punkten P i ett rätvinkligt koordinatsystem. Cosinus för vinkeln ger x-koordinaten, sinus för vinkeln ger y-koordinaten.

Sinusformen

Om vi ritar upp y-koordinaten som funktion av vinkeln i ett rätvinkligt koordinatsystem, vinkeln längs x-axeln, så får vi "sinuskurvan". Om vi gör på samma sätt med x-koordinaten så får vi cosinuskurvan. Kurvorna är identiska till formen. Enda skillnaden består i att sinuskurvan är förskjuten 90 grader längs vinkelaxeln.

Sinusformad spänning

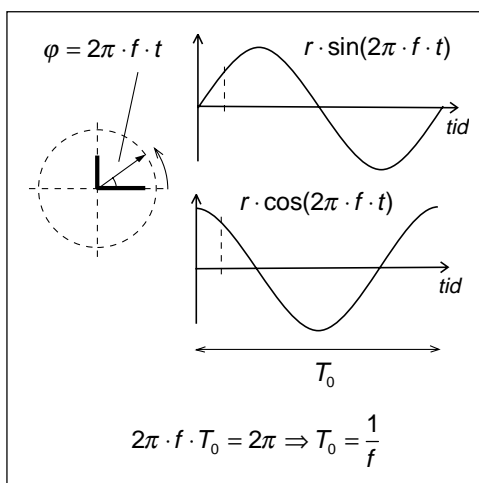
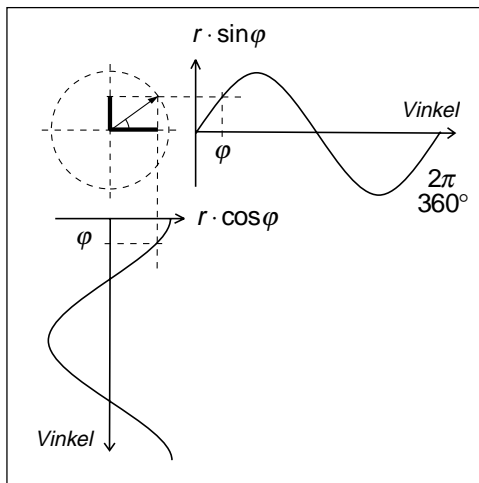
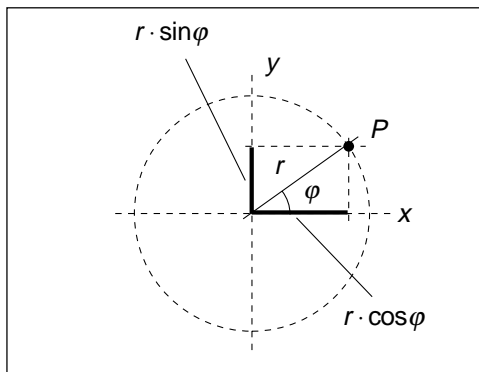
Växelspänningsgeneratoren har en spole som roterar i ett magnetfält, eller en magnet, ett magnetfält, som roterar i en spole. Då induceras en växelspänning i spolen. Om magnetfältet är utformat på lämpligt sätt så kommer spänningen att ändra sig i tiden enligt sinuskurvan.

Sinusfunktionens argument växer med tiden t , och efter tiden T_0 har växelspänningen genomlöp en period. Om vi känner växelspänningens periodtid T_0 så kan vi beräkna frekvensen, d.v.s. antalet perioder per sekund.

Enheten för frekvens var tidigare p/s (perioder per sekund, eng. c/s, cycles per second), men ändrades 1958 i samband med det geofysiska året till Hz för att hedra Heinrich Hertz, den förste som konstruerade en sändare och mottagare för radiovågor.

Sinusspänning eller cosinusspänning?

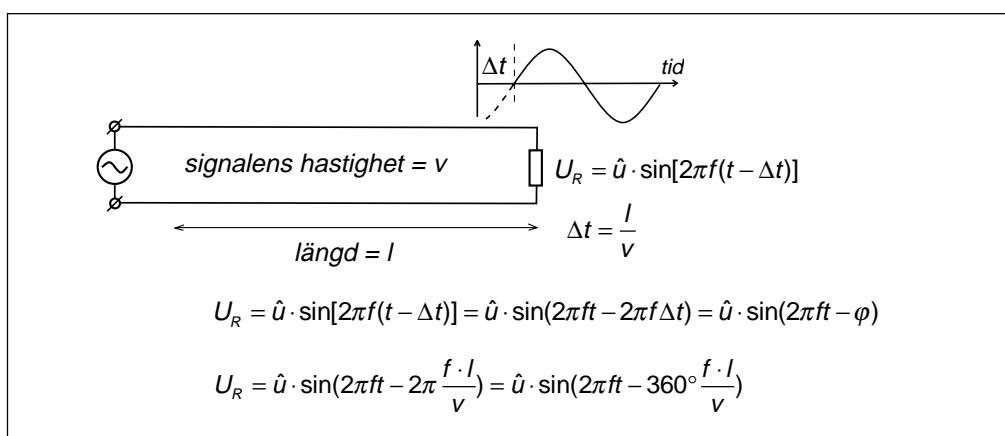
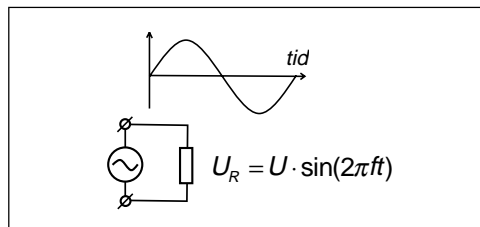
Lämnar generatoren en sinusspänning eller en cosinusspänning? Det beror på vid vilken tidpunkt vi sätter $t = 0$. Om vi utgår från tidpunkten när växelspänningen är noll, med positiv derivata, så får vi en sinusspänning. Om vi utgår från maximalt positiv så får vi en cosinusspänning.



1.2 Fasvinkel är alltid en tidsförskjutning

Bilden till höger visar ett motstånd som anslutits till en växelspänningsgenerator. Spänningen över motståndet blir identisk med generatorspänningen.

Vad blir spänningen över motståndet om vi har en förlustfri ledning mellan generatoren och motståndet?



Signalen går längs ledningen med ljushastigheten, som ju är väldigt snabbt, men inte oändligt snabbt. Väckelspänningen kommer fram lite senare, tiden Δt senare.

Utgå från det matematiska uttrycket för spänningen över motståndet. Fysikaliskt kommer signalen fram lite senare.

Nu kan vi multiplicera ihop och få den ursprungliga sinusspänningen, men förskjuten en negativ fasvinkel, som egentligen skall uttryckas i radianer, men som lättare kan förstås om den uttrycks i grader.

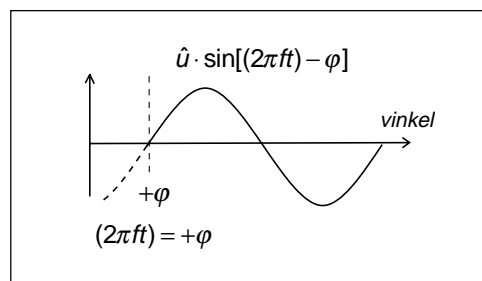
Att signalen kommer fram lite senare innebär att väckelspänningen har fått en fasvinkel, en negativ fasvinkel.

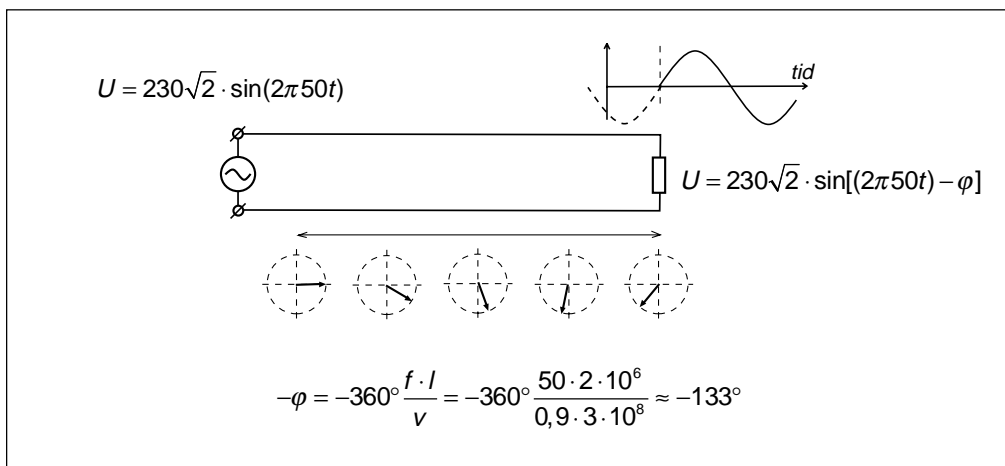
Denna väckelspänning som fördröjts och fått en negativ fasvinkel kommer, om den ritas i ett diagram där x-axeln graderats i vinkel, att gå genom noll vid motsvarande positiva fasvinkel. En förskjutning av sinusspän-

ningen i positiv riktning ger alltså motsvarande fasvinkel i det matematiska uttrycket, men med minustecken.

Positiv – tidigare, negativ – senare

En positiv fasvinkel i det matematiska uttrycket innebär alltså att signalen kommer tidigare (hur nu detta skall gå till), och en negativ fasvinkel att signalen kommer senare, är fördröjd.





Kraftverk i Norrland, lampa i Skåne

Vi tänker oss att nätspänningen alstras i Norrland och att lampan ansluts i Skåne, ett avstånd på 200 mil.

Effekten färdas på kraftledningarna med ljushastigheten på ledningen, som är aningen långsammare än ljushastigheten i vakuum, vi antar 90 % av 300 000 km/s. I detta fall motsvarar fördröjningen en fasvinkel på ungefär -133° .

GSM 1800

Men om generatoren är sändaren i ficktelefonen med en frekvens av ungefär 1800 MHz, och ledningen fram till antennen är 5 cm lång och effekten går med 80 % av ljushastigheten i vakuum, vilken fasvinkel ger detta?

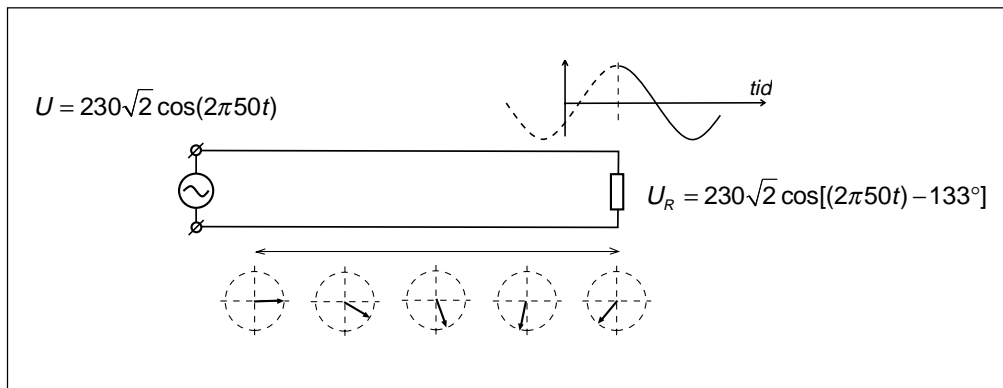
Även i detta fall får vi en fasvinkel på -135° . Vid hög frekvens behövs inte så långa ledningar innan fasvridningen blir betydande. Det är bland annat detta vi sysslar med i ämnet telekommunikation.

Visardiagram

I bilden ovan finns även spänningen inritad som en spänningsvisare i visardiagrammet. Detta är den riktning som spänningsvisaren får när man sätter $t = 0$.

Spänningens momentana amplitud vid $t = 0$ får man genom att ta visarens projektion på y-axeln (sinus-axeln), eftersom generatoren lämnade sinusspänning, d.v.s. generatorens momentana spänning är noll vid $t = 0$.

$$\begin{aligned}
 -\varphi &= -360^\circ \frac{f \cdot l}{v} = \\
 &= -360^\circ \frac{1800 \cdot 10^6 \cdot 0,05}{0,8 \cdot 3 \cdot 10^8} = -135^\circ
 \end{aligned}$$



Vad händer om generatorn lämnar cosinus-spänning?

- De matematiska uttrycken innehåller cosinusuttryck.
- Bilden på spänningen blir cosinus som förskjuts -133° .
- Visardiagrammen är oförändrade, men i detta fall avläses den momentana spänningen vid $t = 0$ på x-axeln (cosinus-axeln).

Var försiktig med att blanda sinus och cosinus i dina beräkningar

Om generatorn lämnar cosinus-spänning och vi vill ha spänningen över motståndet som sinusuttryck så måste fasvinkeln korrigeras för omvandlingen från cosinus till sinus.

För den som använder trigonometrin t.ex. för navigering så är det praktiskt med både sinus och cosinus, eftersom sinus för en vinkel är y-riktningen, medan cosinus för vinkeln är x-riktningen. Men i elläran använder vi inte sinus och cosinus för tvådimensionella problem, utan bara för att beskriva hur sinusspänningen ändras i tiden, hur amplituden ändras, i en dimension.

I avsnitten om modulering så använder vi cosinus, helt enkelt därför att om man multiplicerar två cosinustermer så blir även resultatet cosinustermer. Om vi multiplicerar sinustermer så blir resultatet cosinus-termer och det blir genast mycket besvärligare.

$$\begin{aligned} U_R &= 230\sqrt{2} \cos[(2\pi 50t) - 133^\circ] = \\ &= 230\sqrt{2} \sin[(2\pi 50t) - 133^\circ + 90^\circ] = \\ &= 230\sqrt{2} \sin[(2\pi 50t) - 43^\circ] \end{aligned}$$

$$\begin{aligned} \cos \alpha \cdot \cos \beta &= \frac{1}{2} \cos(\alpha - \beta) + \frac{1}{2} \cos(\alpha + \beta) \\ \sin \alpha \cdot \sin \beta &= \frac{1}{2} \cos(\alpha - \beta) - \frac{1}{2} \cos(\alpha + \beta) \end{aligned}$$

1.3 Lägga ihop, dra ifrån och multiplicera

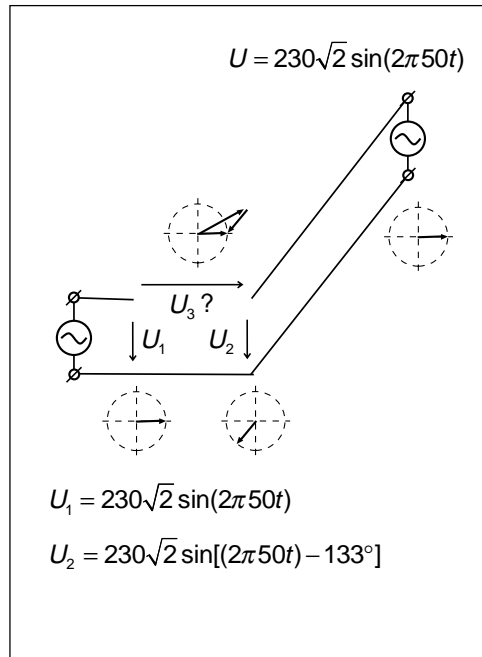
Kraftverk i Norrland och kraftverk i Skåne skall samköras

De svenska kraftverken är kopplade till ett gemensamt elnät. Samtliga kraftverk matar ut effekt på ett sammanhängande ledningsnät.

Vad händer om både kraftverket i Norrland och kraftverket i Skåne lämnar sinusspänning med fasvinkeln noll? Vi tänker oss att kraftverket i Skåne just startat och skall kopplas in på ledningsnätet.

Från Norrland kommer sinusspänning med fasvinkeln -133° . Från kraftverket i Skåne kommer sinusspänning med fasvinkel 0° . Mellan de två ledarna som skall sammankopplas finns en sinusspänning med effektivvärdet 421 volt och fasvinkeln $+23,5^\circ$. Detta resultat kan vi räkna ut med hjälp av den trigonometriska formelsamlingen, eller lösa grafiskt med visaridiagram.

Vad händer om vi kopplar ihop? Det säger PANG!



$$\begin{aligned}
 U_1 - U_2 = U_3 &= 230\sqrt{2} \{ \sin[(2\pi 50t) - \sin[(2\pi 50t) - 133^\circ]] \} = \\
 &= 230\sqrt{2} \left\{ 2 \sin \frac{133^\circ}{2} \cdot \cos \left[(2\pi 50t) - \frac{133^\circ}{2} \right] \right\} = \\
 &= 230\sqrt{2} \left\{ 2 \sin \frac{133^\circ}{2} \right\} \left\{ \sin \left[(2\pi 50t) - \frac{133^\circ}{2} + 90^\circ \right] \right\} = \\
 &= 230\sqrt{2} \{ 1,83 \} \{ \sin[(2\pi 50t) + 23,5^\circ] \}
 \end{aligned}$$

Trigonometriska formelsamlingen är viktig

I exemplet ovan tar vi skillnaden mellan två sinusspänningar med samma frekvens. Då är det lätt att lösa uppgiften med visaridiagram. Men om sinusspänningarna har olika frekvens, och vi inte bara lägger ihop eller drar ifrån utan multiplicerar spänningarna med varandra, så är det inte lika lätt att se resultatet i bilder. Därför är den trigonometriska formelsamlingen viktig.

Trigonometriska formler

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

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

(3) $\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta$

(4) $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$

(5) $\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$

(6) $\tan(\alpha + \beta) = \frac{\tan \alpha + \tan \beta}{1 - \tan \alpha \tan \beta}$

(7) $\tan(\alpha - \beta) = \frac{\tan \alpha - \tan \beta}{1 + \tan \alpha \tan \beta}$

(8) $\cot(\alpha + \beta) = \frac{\cot \beta \cot \alpha - 1}{\cot \beta + \cot \alpha}$

(9) $\cot(\alpha - \beta) = \frac{\cot \beta \cot \alpha + 1}{\cot \beta - \cot \alpha}$

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

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

(12) $\sin^2 \alpha = \frac{1 - \cos 2\alpha}{2}$; (13) $\cos^2 \alpha = \frac{1 + \cos 2\alpha}{2}$

(14) $\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$

(15) $\sin \alpha - \sin \beta = 2 \sin \frac{\alpha - \beta}{2} \cos \frac{\alpha + \beta}{2}$

(16) $\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$

(17) $\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}$

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

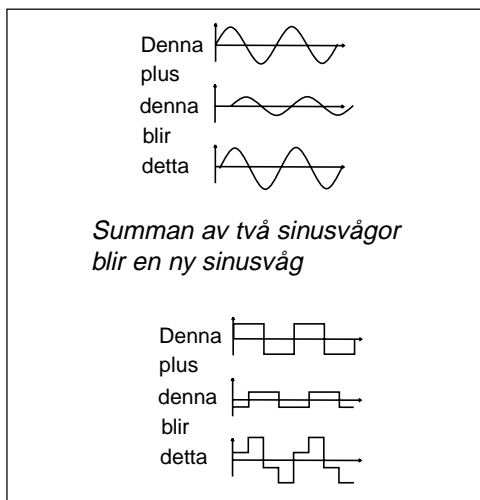
(19) $\cos \alpha \cos \beta = \frac{1}{2} \cos(\alpha - \beta) + \frac{1}{2} \cos(\alpha + \beta)$

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

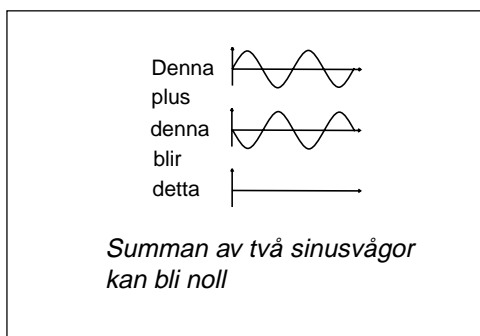
(21) $\cos \alpha = \frac{e^{j\alpha} + e^{-j\alpha}}{2}$; (22) $\sin \alpha = \frac{e^{j\alpha} - e^{-j\alpha}}{2j}$

1.4 Derivera och integrera

Sinusspänningen är den enda växelspanning där man kan lägga ihop spänningar med samma frekvens men olika amplitud och fasvinkel, olika fördröjning, och där summan fortfarande är sinusformad. Ingen annan form på växelspanningen har dessa egenskaper.



När man adderar två sinusvågor med samma amplitud och periodtid där den ena sinusvågen är fördröjd en halv period, så blir summan noll. Detta är vårt främsta problem vid radiokommunikation (flervägsutbredning), men utnyttjas även när vi medvetet vill stoppa vissa frekvenser, t.ex. i filter.

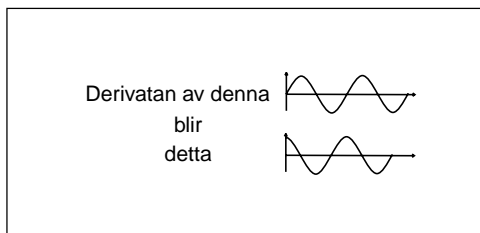


Derivatnan

Ytterligare en egenskap hos sinusvågen är att även derivatan är sinusformad, och då är naturligtvis även alla högre ordningars derivator också sinusformade.

$$\frac{d}{dt} \sin \omega t = \omega \cos \omega t$$

$$\frac{d}{dt} \cos \omega t = -\omega \sin \omega t$$

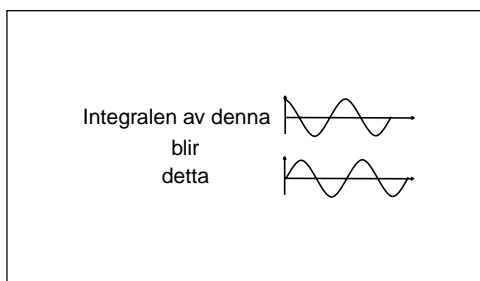


Integralen

Även integralen är sinusformad.

$$\int \sin \omega t dt = -\frac{\cos \omega t}{\omega}$$

$$\int \cos \omega t dt = \frac{\sin \omega t}{\omega}$$



1 — Sinusspänning

En spänningskälla $u(t)$ matar ett nät bestående av motstånd, spole och kondensator. Beräkna strömmen i kretsen.

Ett problem av denna typ leder till en ekvation som innehåller både derivator och integraler. För att beräkna strömmen $i(t)$ så deriverar vi en gång och får en inhomogen linjär differentialekvation av andra ordningen med konstanta koefficienter.

$$\text{Du vet väl att: } \omega = 2\pi f$$

Om $u(t)$ är sinusformad

Men om $u(t)$ är sinusformad så måste även strömmen vara sinusformad.

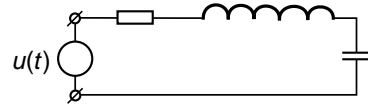
Då kan vi gissa, d.v.s. ansätta en sinusformad ström med okänd amplitud och fasvinkel men med samma frekvens som $u(t)$.

Nu kan vi beräkna derivatan av denna okända ström ...

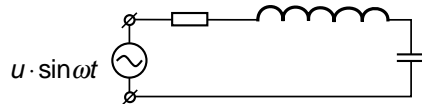
... och vi kan beräkna integralen.

Genom att stoppa in dessa uträknade värden på strömmens derivata och integral så får vi en vanlig trigonometrisk ekvation med enbart sinustermer.

Detta fungerar eftersom även derivatan och integralen har samma sinusform som den ursprungliga strömmen. Det är bara amplituden och fasvinkeln som är annorlunda.



$$R \cdot i(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int_{-\infty}^t i(t) \cdot dt = u(t)$$



$$R \cdot i(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int_{-\infty}^t i(t) \cdot dt = \hat{u} \cdot \sin \omega t$$

$$i(t) = \hat{i} \cdot \sin(\omega t + \varphi)$$

$$\begin{aligned} \frac{di(t)}{dt} &= \hat{i} \omega \cos(\omega t + \varphi) = \\ &= \hat{i} \omega \sin(\omega t + \varphi + \frac{\pi}{2}) \end{aligned}$$

$$\begin{aligned} \int_0^t i(t) \cdot dt &= \frac{-\hat{i} \cdot \cos(\omega t + \varphi)}{\omega} \\ &= \frac{\hat{i} \cdot \sin(\omega t + \varphi + \frac{\pi}{2})}{\omega} \end{aligned}$$

1.5 $j\omega$ -metoden

Nu skall vi utnyttja ett matematiskt knep:

$$e^{j\omega t} = \cos \omega t + j \sin \omega t$$

Med hjälp av detta kan vi skriva generatorspänningen som:

$$\begin{aligned} \hat{u} \cdot \sin \omega t &= \text{Im}[\hat{u} \cdot \cos \omega t + j \cdot \hat{u} \cdot \sin \omega t] = \\ &= \text{Im}[Ue^{j\omega t}] \end{aligned}$$

och vi kan ansätta strömmen:

$$i(t) = \hat{i} \sin(\omega t + \varphi) = \text{Im}[I \cdot e^{j\omega t}]$$

På samma sätt blir spänningen över spolen:

$$\begin{aligned} L \frac{d(I \cdot e^{j\omega t})}{dt} &= L \cdot I \frac{d(e^{j\omega t})}{dt} = \\ &= j\omega L \cdot I \cdot e^{j\omega t} \end{aligned}$$

och spänningen över kondensatorn:

$$\begin{aligned} \frac{1}{C} \int_0^t I \cdot e^{j\omega t} dt &= \frac{1}{C} \cdot I \cdot \int_0^t e^{j\omega t} dt \\ &= \frac{1}{j\omega C} \cdot I \cdot e^{j\omega t} \end{aligned}$$

Genom att i den ursprungliga ekvationen införa de komplexa uttrycken för spänning och ström, ...

$$R \cdot i(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int i(t) \cdot dt = \hat{u} \cdot \sin \omega t$$

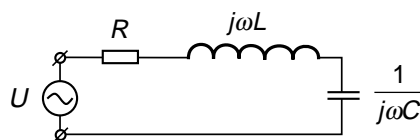
... komplettera induktansen med det som blir resultatet när strömmen deriveras, och komplettera kondensatorn med det som blir resultatet när strömmen integreras, ...

$$R \cdot I \cdot e^{j\omega t} + j\omega L \cdot I \cdot e^{j\omega t} + \frac{1}{j\omega C} \cdot I \cdot e^{j\omega t} = Ue^{j\omega t}$$

... förkorta bort e-funktionen, ...

$$R \cdot I + j\omega L \cdot I + \frac{1}{j\omega C} \cdot I = U$$

... så kan vi räkna som om det vore likström.



När vi beräknat den komplexa strömmen så multiplicerar vi med e-funktionen, ...

$$I = \frac{U}{R + j\omega L + \frac{1}{j\omega C}}$$

... varefter vi tar imaginärdelen av uttrycket om $u(t)$ var en sinusgenerator, ...

... eller tar realdelen av uttrycket, om $u(t)$ var en cosinusgenerator.

$$i(t) = \text{Im} \left[\frac{U}{R + j\omega L + \frac{1}{j\omega C}} \cdot e^{j\omega t} \right] = \hat{i} \sin(\omega t + \varphi)$$

Vi har fått en symbolisk lösningsmetod för differentialekvationen.

1.6 Vad innebär "stationär" växelström?

Den symboliska lösningsmetoden för differentialekvationen som vi utnyttjar i $j\omega$ -metoden fungerar bara vid stationär växelström. Vad innebär stationär?

Det stationära fallet

Innebörden av det stationära fallet är att växelspanningsgeneratoren aldrig får ändras. Visst ändras spänningen sinusformat, men sinusspänningens toppvärde, frekvens och fasvinkel får aldrig ändras. Det är detta som är det stationära fallet.

Men vi måste ju starta generatoren?

Om växelspanningsgeneratoren aldrig får ändra sin amplitud så måste den ha varit påslagen ända sedan tidernas begynnelse. I verkligheten räcker det att vänta en kort stund, tills insvängningsförloppet avklingat, innan vi har förhållanden som kan liknas vid det stationära fallet.

Superposition

Så länge som den elektriska kretsen består av linjära komponenter som motstånd, spolar, kondensatorer, transistorer m.m. som arbetar i det linjära området, så gäller superposition. Detta innebär att vi kan ha flera generatorer med var sin stationära växelspanning. Vi kan även ha likspänning och likström t.ex. för de transistorer som finns i kretsen.

För att beräkna totala strömmen så kortsluter vi alla generatorer och tar bort alla strömkällor utom en och beräknar strömbidraget från den källan. Sedan fortsätter vi med nästa. Börja med likströmmen. Fortsätt med växelspanningarna. Använd $j\omega$ -metoden på varje växelspanning för sig, och totala strömmen blir summan av alla delströmmar.

Laplace-transformen

I nästa avsnitt skall vi se att vilken komplicerad kurvform som helst kan åstadkommas genom att summera ett antal, kanske oändligt antal stationära växelspanningar, ett spektrum. Men vi behöver inte använda $j\omega$ -metoden på varje växelspanning för sig, utan

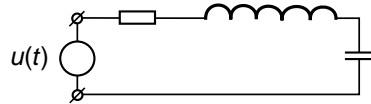
det har utvecklats symboliska metoder att utföra beräkningarna på alla spektralkomponenter samtidigt. En sådan symbolisk metod är Laplace-transformen.

Genom att beskriva komplicerade kurvformer som ett spektrum av stationära sinusspänningar, t.ex. spänningsändringen när generatoren startas, så klarar vi att beräkna strömmen i kretsen även under själva insvängningsförloppet, utan att behöva ge oss på själva differentialekvationen. Vi löser differentialekvationen med en symbolisk metod.

1.7 Om man ändå vill lösa differentialekvationen

Att lösa differentialekvationen

Bestäm strömmen $i(t)$, när vi lägger spänningen $u(t)$ över kretsen.



$$R \cdot i + L \frac{di}{dt} + \frac{1}{C} \int_{-\infty}^t i dt = u(t)$$

Eftersom vi inte vill ha både derivator och integraler i samma ekvation deriverar vi och stugar om:

Detta blir en *inhomogen linjär differentialekvation av andra ordningen med konstanta koefficienter*. Den klassiska lösningsmetoden är att söka dels efter en partikulärlösning $i_p(t)$ (en godtycklig speciell lösning) till den inhomogena differentialekvationen, ...

$$R \frac{di}{dt} + L \frac{d^2i}{dt^2} + \frac{1}{C} i = \frac{du(t)}{dt}$$

$$\frac{d^2i}{dt^2} + \frac{R}{L} \frac{di}{dt} + \frac{1}{LC} i = \frac{1}{L} \frac{du(t)}{dt}$$

Inhomogena differentialekvationen

... dels efter en allmän lösning $i_h(t)$ (innehållande ett antal integrationskonstanter, vars värden bestäms av begynnelsevillkoren) till den homogena differentialekvationen.

$$\frac{d^2i}{dt^2} + \frac{R}{L} \frac{di}{dt} + \frac{1}{LC} i = 0$$

Homogena differentialekvationen

Den fullständiga lösningen är summan av de båda lösningarna.

$$i(t) = i_p(t) + i_h(t)$$

1 — Sinusspänning

Vad innebär partikulärlösning och allmän lösning i verkligheten?**Partikulärlösningen**

Om vi förändrar spänningen över kretsen, en likspänning eller en växelspänning, händer två saker:

- Vi kommer att gå från ett **stationärt tillstånd** vid den ena spänningsnivån till ett annat stationärt tillstånd vid den andra spänningsnivån. Strömmens värde vid de stationära tillstånden kan beräknas så som du lärt dig vid stationär likspänning, eller med $j\omega$ -metoden vid stationär växelspänning.

Den ström $i_s(t)$ som vi får vid beräkning av stationära tillståndet är en **partikulärlösning** till den inhomogena differentialekvationen.

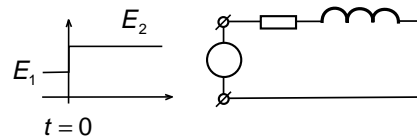
Allmänna lösningen

- Övergången från det ena stationära tillståndet till det nya stationära tillståndet ger under kort tid upphov till ett **övergångsförlopp** (transient förlopp, in- eller utsvängningsförlopp). Tiden kan vara så kort att övergångsförloppet försummas. Det är vad du hittills har gjort. Eller också är övergångsförloppet mycket viktigt, t.ex. i reglertekniken.

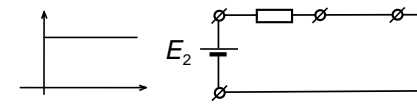
En allmän lösning till den homogena differentialekvationen utgöres av **övergångsförloppet**. De i lösningen ingående integrationskonstanterna bestäms av att strömmen genom en induktans och spänningen över en kondensator ej kan ändra sig språngvis. Såväl det induktiva spänningsfallet $L di/dt$ som kondensatorströmmen $C du/dt$ måste ha ändligt värde.

Vi väljer alltså den partikulärlösning $i_p(t)$ som är lika med stationära tillståndet $i_s(t)$. Då kommer den allmänna lösningen $i_h(t)$, när vi tar hänsyn till begynnelsevillkoren, att bli identisk med övergångsförloppet $i_o(t)$.

$$i(t) = i_p(t) + i_h(t) = i_s(t) + i_o(t)$$

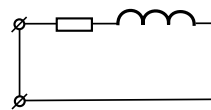


Beräkna strömmen när spänningen stiger vid $t = 0$



$t \gg 0 \Rightarrow i_s = \frac{E_2}{R}$ Strömmen vid $t \gg 0$

Partikulärlösning



$$L \frac{di}{dt} + Ri = 0$$

$$i_o = A \cdot e^{-Rt/L}$$

Allmän lösning

$$i(t) = i_s + i_o = \frac{E_2}{R} + A \cdot e^{-Rt/L}$$

Randvillkor:

$$\begin{cases} i(t) = \frac{E_1}{R} & t = 0 \\ i(t) = \frac{E_2}{R} & t \gg 0 \end{cases}$$

Detta ger:

$$i(t) = \frac{1}{R} (E_2 + (E_1 - E_2) \cdot e^{-Rt/L})$$

1.8 Decibel

Elektroingenjören räknar med dB av två orsaker:

- Hur mycket svagare blir signalen när den passerat kabeln? Är det spänningen som blivit svagare eller menar vi effekten? Genom att uttrycka kabeldämpningen i dB så behöver vi inte bekymra oss. Detta hanteras automatiskt vid omvandlingen till dB.
- När en signal passerar genom förstärkare och över ledningar så förstärks och dämpas signalen. Vi måste multiplicera med förstärkningen och dividera med dämpningen. Om förstärkningar och dämpningar räknas om till motsvarande värden i dB så övergår multiplikation och division till addition och subtraktion, något som lättare kan åskådliggöras grafiskt i ett länkbudgetdiagram.

Vilken slags dämpning avses?

Vi matar in 8 volt på kabeln och får ut 4 volt. Spänningen dämpas 0,5 ggr.

Vad händer med effekten? In i kabeln matas 8 volt och kabeln har karakteristiska impedansen 50 Ω. Vi matar in 1,28 W.

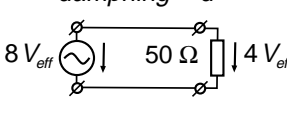
Hur mycket får vi ut? 4 volt över 50 Ω är 0,32 W. Effekten har dämpats 0,25 ggr. Vi måste alltså hålla reda på om vi menar spänningsdämpning eller effektdämpning (eller strömdämpning)!

Vad blir dämpningen i dB?

När dämpningen d omräknas till dB så blir siffervärdet i dB detsamma i båda fallen. Kabeldämpningen är helt enkelt 6 dB, eller -6 dB.

En förstärkare förstärker +6 dB om den förstärker, och förstärker -6 dB om den dämpar. I analogi med detta förstärker en kabel -6 dB, men eftersom kabeln aldrig förstärker, så säger man aldrig att kabelförstärkningen är -6 dB, utan att kabeldämpningen är 6 dB, eller -6 dB, vilket skall uppfattas som samma sak.

$dämpning = d$



$P_{ut} = P_{in} \cdot d$

$$d = \frac{P_{ut}}{P_{in}}$$

$$d [dB] = 10 \cdot \log d = 10 \cdot \log \frac{P_{ut}}{P_{in}} =$$

$$= 10 \cdot \log \left\{ \frac{U_{ut}^2 R_{in}}{R_{ut} U_{in}^2} \right\} =$$

$$= 10 \cdot \log \left\{ \frac{U_{ut}^2}{U_{in}^2} \right\} + 10 \cdot \log \left\{ \frac{R_{in}}{R_{ut}} \right\} =$$

$$= 20 \cdot \log \left\{ \frac{U_{ut}}{U_{in}} \right\} + 10 \cdot \log \left\{ \frac{R_{in}}{R_{ut}} \right\}$$

$$d [dB] = 10 \cdot \log \frac{P_{ut}}{P_{in}} = 10 \cdot \log \frac{0,32}{1,28} = -6 \text{ dB}$$

$$d [dB] = 20 \cdot \log \frac{U_{ut}}{U_{in}} + 10 \cdot \log \frac{R_{in}}{R_{ut}} =$$

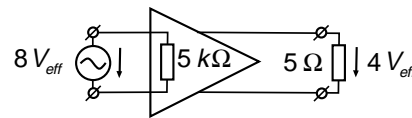
$$= 20 \cdot \log \frac{4}{8} + 10 \cdot \log \frac{50}{50} = -6 \text{ dB}$$

1 — Sinusspänning

Bilden visar en förstärkare. In matas 8 volt och ut kommer 4 volt. Har vi någon förstärkning eller fungerar "förstärkaren" som dämpsats?

Inspänningen ser en resistans på $5\text{ k}\Omega$, vilket betyder att den inmatade effekten är låg. På utgången levererar förstärkaren 4 volt till resistansen $5\ \Omega$, vilket ger hög effekt. Denna förstärkare har en förstärkning på $+24\text{ dB}$.

Det förekommer i datablad över förstärkare att dB-värdet för spänningsförhållandet redovisas. Den här förstärkaren har en spänningsförstärkning som är -6 dB . Men för att få förstärkarens verkliga förstärkning så måste man alltså även ta hänsyn till resistansförhållandet.



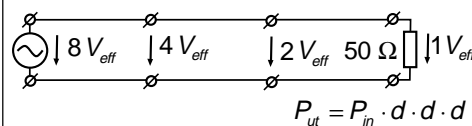
$$\begin{aligned} g[\text{dB}] &= 20 \cdot \log \frac{U_{ut}}{U_{in}} + 10 \cdot \log \frac{R_{in}}{R_{ut}} = \\ &= 20 \cdot \log \frac{4}{8} + 10 \cdot \log \frac{5000}{5} = \\ &= -6\text{ dB} + 30\text{ dB} = +24\text{ dB} \end{aligned}$$

Om vi tar en längre kabel?

I det tidigare exemplet beräknade vi kabeldämpningen till -6 dB . Vad händer med dämpningen om vi gör kabeln tre gånger så lång?

Den inmatade spänningen sjunker först till hälften och sedan till hälften av hälften, och sedan till ytterligare hälften av detta d.v.s. dämpningarna skall multipliceras.

Om kabeldämpningen för 1 m kabel är -6 dB , så blir dämpningen om kabeln är 3 m lika med -18 dB . Det är bara att multiplicera kabeldämpningen i dB/m med kabel-längden så får man totala kabeldämpningen i dB.



$$P_{ut} = P_{in} \cdot d \cdot d \cdot d$$

$$\begin{aligned} d_3[\text{dB}] &= 10 \cdot \log \frac{P_{ut}}{P_{in}} = 10 \cdot \log(d \cdot d \cdot d) = \\ &= 10 \cdot \log(d)^3 = 3 \cdot [10 \cdot \log d] = \\ &= 3 \cdot (-6\text{ dB}) = -18\text{ dB} \end{aligned}$$

Vad blir uteffekten?

Nu skall vi göra tvärt om. Vi känner den inmatade effekten och kabeldämpningen. Vad blir uteffekten?

Det är detta som gäller:

$$\frac{P_1}{P_2}[\text{dB}] = 10 \cdot \log \left\{ \frac{P_1}{P_2}[\text{ggr}] \right\}$$

$$\frac{P_1}{P_2}[\text{ggr}] = 10^{\frac{P_1[\text{dB}]}{10}}$$

$$d[\text{dB}] = 10 \cdot \log \frac{P_{ut}}{P_{in}}$$

$$\frac{d[\text{dB}]}{10} = \log \frac{P_{ut}}{P_{in}}$$

$$10^{\frac{d[\text{dB}]}{10}} = 10^{\log \frac{P_{ut}}{P_{in}}} = \frac{P_{ut}}{P_{in}}$$

$$P_{ut} = P_{in} \cdot 10^{\frac{d[\text{dB}]}{10}}$$

1 — Sinusspänning

Nu skall vi se på ett annat sätt att använda dB. Vi utgår från uttrycket för kabeldämpningen:

$$d[\text{dB}] = 10 \cdot \log \frac{P_1[\text{W}]}{P_2[\text{W}]}$$

Detta uttryck kan stuvras om med logaritmräkning så att vi får:

$$\begin{aligned} d[\text{dB}] &= 10 \cdot \log \frac{P_1[\text{W}]}{P_2[\text{W}]} = \\ &= 10 \cdot \log P_1[\text{W}] - 10 \cdot \log P_2[\text{W}] \end{aligned}$$

Men så kan man väl inte göra? Decibel skall ju vara ett effektförhållande?

Vi har ett effektförhållande. Förhållandet till referenseffekten 1 W.

$$10 \cdot \log P_2[\text{W}] = 10 \cdot \log P_1[\text{W}] - d[\text{dB}]$$

dBW

För att markera att referenseffekten är 1 W så anger vi enheten till dBW.

$$10 \cdot \log \frac{P_2[\text{W}]}{1\text{W}} = 10 \cdot \log \frac{P_1[\text{W}]}{1\text{W}} - d[\text{dB}]$$

$$P_2[\text{dBW}] = P_1[\text{dBW}] - d[\text{dB}]$$

dBm

Ett äldre sätt är att använda 1 mW som referenseffekt. Detta visar vi genom att ange enheten till dBm.

$$10 \cdot \log \frac{P_2[\text{mW}]}{1\text{mW}} = 10 \cdot \log \frac{P_1[\text{mW}]}{1\text{mW}} - d[\text{dB}]$$

$$P_2[\text{dBm}] = P_1[\text{dBm}] - d[\text{dB}]$$

Inom radioindustrin är det i dag vanligt att sändarens uteffekt anges i dBW. Däremot håller man oftast fast vid den äldre benämningen dBm när man skall ange mottagarkänslighet, d.v.s. den minsta effekt som skall matas in på mottagarens antenningång för att få fullgod mottagning.

Eftersom 1 W är 1000 ggr starkare än 1 mW så är enda skillnaden mellan dBm och dBW att man skall lägga till eller dra ifrån 30 dB när man går från det ena till det andra.

$$P[\text{dBW}] = P[\text{dBm}] - 30\text{ dB}$$

$$P[\text{dBm}] = P[\text{dBW}] + 30\text{ dB}$$

Exempel: GSM ficktelefonens mottagarkänslighet är -104 dBm, vilket är detsamma som -134 dBW.

Lite huvudräkning

Utgå från värdet för 3 dB och 10 dB:

$$0\text{ dB} \Rightarrow 1\text{ ggr}$$

$$3\text{ dB} \Rightarrow 2\text{ ggr}$$

$$10\text{ dB} \Rightarrow 10\text{ ggr}$$

Då kan resten beräknas med huvudräkning.

$$6\text{ dB} = 3\text{ dB} + 3\text{ dB} \Rightarrow 2 \cdot 2 = 4\text{ ggr}$$

$$7\text{ dB} = 10\text{ dB} - 3\text{ dB}$$

$$4\text{ dB} = 10\text{ dB} - 6\text{ dB}$$

$$9\text{ dB} = 6\text{ dB} + 3\text{ dB}$$

$$8\text{ dB} = 4\text{ dB} + 4\text{ dB} \Rightarrow 2,5 \cdot 2,5 = 6,3\text{ ggr}$$

$$5\text{ dB} = 8\text{ dB} - 3\text{ dB} \Rightarrow 6,3 / 2 = 3,2\text{ ggr}$$

$$2\text{ dB} = 5\text{ dB} - 3\text{ dB} \Rightarrow 3,2 / 2 = 1,6\text{ ggr}$$

$$1\text{ dB} = 4\text{ dB} - 3\text{ dB} \Rightarrow 2,5 / 2 = 1,25\text{ ggr}$$

2 — Fourierserier och spektrum

Kapitlet börjar med att beskriva hur man gör Fourierserieutvecklingen:

- hur vi beräknar likspänningsdelen,
- hur vi beräknar cosinustermerna,
- hur vi beräknar sinustermerna,

för att därefter slå ihop cosinus och sinus och få en enda frekvenskomponent med amplitud och fasvinkel på respektive frekvens — signalens spektrum.

Därefter jämför vi signalen som tidsfunktion med signalen som spektrum, frekvensfunktion.

Vi räknar på fyrkantspänningen därför att den är enkel att räkna på. Vi räknar även på korta spänningspulser som är intressantare eftersom spektrat liknar det vi får vid digital modulering med "ettor" och "nollor".

2.1 Fourier-serier

Den franske matematikern Jean Baptiste Joseph Fourier (1768 – 1830) visade att vilken periodisk signal $u(t)$ som helst kan ses som summan av likspänning och ett antal cosinus- och sinussignaler, alla med olika men konstanta amplituder a_n och b_n , och frekvens f_n , på följande sätt:

$$u(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cdot \cos 2\pi n f_0 t + b_n \cdot \sin 2\pi n f_0 t)$$

Integrering skall ske över en hel period T_0 som kan väljas hur som helst. Integreringen skall starta vid någon tidpunkt (t) och pågå till ($t+T_0$).

$$f_0 = \frac{1}{T_0}$$

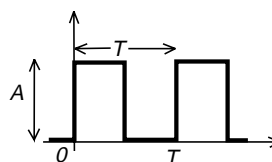
$$a_0 = \frac{1}{T_0} \int_t^{t+T_0} u(t) \cdot dt$$

$$a_n = \frac{2}{T_0} \int_t^{t+T_0} u(t) \cdot \cos(2\pi n f_0 t) \cdot dt$$

$$b_n = \frac{2}{T_0} \int_t^{t+T_0} u(t) \cdot \sin(2\pi n f_0 t) \cdot dt$$

Exempel: Fyrkantvågen

Fyrkantvågen beskriver vi matematiskt på följande sätt:



$$u(t) = \begin{cases} A & 0 \leq t \leq T/2 \\ 0 & T/2 \leq t \leq T \end{cases}$$

Så här gör vi Fourierserieuppdelningen:

Grundfrekvensen blir $1/T$:

$$f_0 = \frac{1}{T}$$

Beräkna likspänningskomponenten. Integrationsgränser kan väljas hur som helst så länge som vi integrerar över en period T :

$$a_0 = \frac{1}{T} \int_{-T/4}^{+3T/4} u(t) \cdot dt = \frac{1}{T} \int_0^{+T/2} A \cdot dt = \frac{A}{T} \left[\frac{T}{2} - 0 \right] = \frac{A}{2}$$

forts!

2 — Fourierserier och spektrum

forts!

Samtliga a_n -värden (utom likspänningskomponenten) blir noll. Denna signal innehåller alltså inga cosinustermer.

$$\begin{aligned} a_n &= \frac{2}{T} \int_0^T u(t) \cdot \cos 2\pi n f_0 t \cdot dt = \frac{2}{T} \int_0^{+T/2} A \cdot \cos 2\pi n f_0 t \cdot dt = \\ &= \frac{2A}{2\pi n f_0 \cdot T} \left[\sin \left(2\pi n \frac{1}{T} \frac{T}{2} \right) - 0 \right] = 0 \end{aligned}$$

Beräkna b_n -värdena:

$$\begin{aligned} b_n &= \frac{2}{T} \int_0^T u(t) \cdot \sin 2\pi n f_0 t \cdot dt = \frac{2}{T} \int_0^{+T/2} A \cdot \sin 2\pi n f_0 t \cdot dt = \\ &= \frac{2A}{2\pi n f_0 \cdot T} \left[-\cos \left(2\pi n \frac{1}{T} \frac{T}{2} \right) + 1 \right] = \frac{A}{\pi n} [1 - \cos \pi n] \end{aligned}$$

Nu kan vi skriva signalen som en oändlig serie.

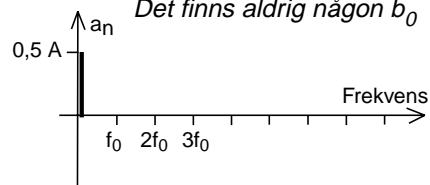
$$\begin{aligned} u(t) &= a_0 + \sum_{n=1}^{\infty} (a_n \cdot \cos 2\pi n f_0 t + b_n \cdot \sin 2\pi n f_0 t) = \\ &= \frac{A}{2} + \sum_{n=1}^{\infty} \left[0 \cdot \cos 2\pi n f_0 t + \frac{A}{\pi n} (1 - \cos \pi n) \sin 2\pi n f_0 t \right] = \\ &= \frac{A}{2} + \sum_{n=1}^{\infty} \frac{A(1 - \cos \pi n)}{\pi n} \sin 2\pi n f_0 t \end{aligned}$$

Vi ser att om $n = \text{udda}$ blir $(1 - \cos \pi n) = 2$ medan om $n = \text{jämn}$ blir $(1 - \cos \pi n) = 0$

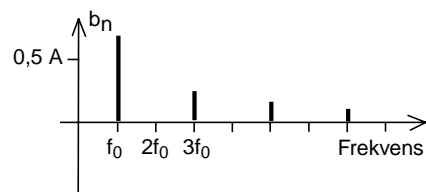
Vi får en oändlig serie som bara innehåller udda deltoner:

$$u(t) = \frac{A}{2} + \sum_{n=1}^{\infty} \frac{A(1 - \cos \pi n)}{\pi n} \sin 2\pi n f_0 t = \frac{A}{2} + \sum_{n=\text{udda}}^{\infty} \frac{2A}{\pi n} \sin 2\pi n f_0 t$$

OBS: Likspänningen ligger i a_0 .
Det finns aldrig någon b_0



Samtliga a_n -värden (utom likspänningskomponenten) blir noll. Signalen innehåller alltså inga cosinustermer.



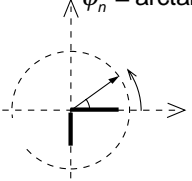
Signalen innehåller endast likspänningskomponenten och sinustermer, b_n -värden.

2.2 Signalens spektrum

När vi räknat ut alla a_n - och b_n -värden: $u(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos 2\pi n f_0 t + b_n \sin 2\pi n f_0 t) =$
 $= \frac{A}{2} + \sum_{n=udda} \frac{2A}{\pi n} \sin 2\pi n f_0 t$

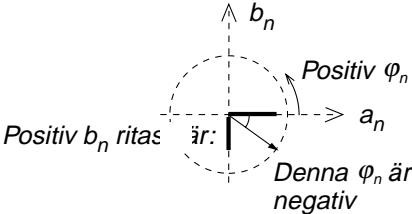
Utnyttja följande: $A_0 = a_0$
 $A_n = \sqrt{a_n^2 + b_n^2}$
 $\varphi_n = \arctan \frac{-b_n}{a_n}$

Observera att vinkeln ligger inom hela intervallet $0 - 360^\circ$. Håll reda på tecknen hos a_n och b_n eftersom arctg bara är definierad inom $\pm 90^\circ$. Rita enhetscirkel!



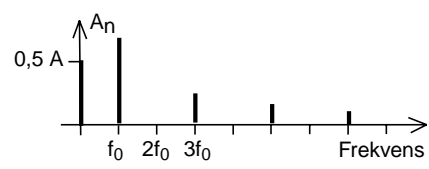
och skriv Fourier-serien för $u(t)$ på polär form:

$$u(t) = A_0 + A_1 \cos(2\pi f_0 t + \varphi_1) + A_2 \cos(2\pi 2f_0 t + \varphi_2) + \dots =$$

$$= A_0 + \sum_{n=1}^{\infty} A_n \cos(2\pi n f_0 t + \varphi_n) = \frac{A}{2} + \sum_{n=udda} \frac{2A}{\pi n} \cos(2\pi n f_0 t - 90^\circ)$$


Amplitudspektrum

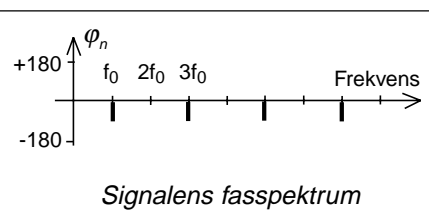
Genom att rita ett amplitud/frekvensdiagram över A_n -komponenterna så får vi signalens totala amplitud vid varje frekvens. Ett sådant diagram kallas signalens amplitudspektrum.



Signalens amplitudspektrum. A_n kan ej bli negativ. Tecknet ligger i fasvinkeln som i så fall är $+180^\circ$ eller -180° .

Fasspektrum

Men amplitudspektrat är inte en fullständig beskrivning av signalen. Vi måste även veta hur respektive frekvenskomponent ligger förskjuten i tiden, d.v.s. vi behöver varje frekvenskomponents fasvinkel, signalens fasspektrum.



Signalens spektrum

Signalens spektrum är alltså amplitudspektrat och fasspektrat. Men ofta är man bara intresserad av amplitudspektrat för att t.ex. få reda på vilken bandbredd en signal har. Därför menar man många gånger bara amplitudspektrat när man säger "signalens spektrum".

2.3 Medeleffekt under en period

Vad blir medeleffekten under en period om spänningen ses som summan av flera sinus-spänningar?

$$u(t) = u_0 + u_1 \sin 2\pi f_0 t + u_2 \sin 2\pi \cdot 2f_0 t$$

$$P_{medel} = \frac{1}{R \cdot T_0} \int_0^{T_0} u(t)^2 dt$$

Integrering (summering) av sinus eller cosinus över en eller flera hela perioder är alltid noll. Dessa termer stryker vi.

$$\begin{aligned} u(t)^2 = & u_0^2 + \cancel{u_0 u_1 \sin 2\pi f_0 t} + \cancel{u_0 u_2 \sin 2\pi \cdot 2f_0 t} + \\ & + \cancel{u_0 u_1 \sin 2\pi f_0 t} + u_1^2 \sin^2 2\pi f_0 t + u_1 u_2 \sin 2\pi f_0 t \cdot \sin 2\pi \cdot 2f_0 t + \\ & + \cancel{u_0 u_2 \sin 2\pi \cdot 2f_0 t} + u_1 u_2 \sin 2\pi f_0 t \cdot \sin 2\pi \cdot 2f_0 t + u_2^2 \sin^2 2\pi \cdot 2f_0 t \end{aligned}$$

Utnyttja att

$$\begin{aligned} \sin A \cdot \sin B &= \sin 2\pi f_0 t \cdot \sin 2\pi \cdot 2f_0 t = \\ &= \frac{1}{2} \cos(A - B) - \frac{1}{2} \cos(A + B) = \\ &= \frac{1}{2} \cos(-2\pi f_0 t) - \frac{1}{2} \cos(2\pi \cdot 3f_0 t) \end{aligned}$$

Då kan vi även stryka dessa termer eftersom de ger sinusformade signaler med frekvens som är multipel av grundfrekvensen.

$$\begin{aligned} u(t)^2 = & u_0^2 + u_1^2 \sin^2 2\pi f_0 t + \cancel{u_1 u_2 \sin 2\pi f_0 t \cdot \sin 2\pi \cdot 2f_0 t} + \\ & + \cancel{u_1 u_2 \sin 2\pi f_0 t \cdot \sin 2\pi \cdot 2f_0 t} + u_2^2 \sin^2 2\pi \cdot 2f_0 t \end{aligned}$$

$$\begin{aligned} P_{medel} &= \frac{1}{R \cdot T_0} \int_0^{T_0} u(t)^2 dt = \\ &= \frac{u_0^2}{R \cdot T_0} \int_0^{T_0} dt + \frac{u_1^2}{R \cdot T_0} \int_0^{T_0} \sin^2 2\pi f_0 t dt + \frac{u_2^2}{R \cdot T_0} \int_0^{T_0} \sin^2 2\pi \cdot 2f_0 t dt = \\ &= \frac{u_0^2}{R} + \frac{u_1^2}{R \cdot T_0} \int_0^{T_0} \frac{1}{2} (1 + \sin 2\pi \cdot 2f_0 t) dt + \frac{u_2^2}{R \cdot T_0} \int_0^{T_0} \frac{1}{2} (1 + \sin 2\pi \cdot 4f_0 t) dt = \\ &= \frac{u_0^2}{R} + \frac{u_1^2}{R \cdot 2} + \frac{u_2^2}{R \cdot 2} = \\ &= \frac{u_0^2}{R} + \frac{\left(\frac{u_1}{\sqrt{2}}\right)^2}{R} + \frac{\left(\frac{u_2}{\sqrt{2}}\right)^2}{R} \end{aligned}$$

Det som återstår är effektutvecklingen i likspänningskomponenten, samt effektutvecklingen i varje deltonskomponent för sig.

Varje delton, frekvens i amplitudspektrat, har sitt eget energiinnehåll. Effektutvecklingen under en period blir summan av den effekt som varje delton kan leverera.

Vi har beräknat medeleffekten under en hel period. Den verkliga effektutvecklingen följer naturligtvis variationerna hos $u(t)$ i kvadrat, d.v.s. det blir varmt "stötvis".

2.4 Effektspektrum

Signalens effektspektrum visar hur stor effekt som varje delton producerar i ett motstånd med värdet 1Ω .

Effektspektrum beräknas på följande sätt:

1. Gör Fourierserieutvecklingen, d.v.s. beräkna a_n - och b_n -komponenterna.

2. Omvandla till amplitudspektrum och fasspektrum. Fasspektrum behöver inte beräknas eftersom vi bara behöver amplitudspektrum för att ta fram effektspektrum.

3. Beräkna effektspektrum genom att kvadrera likspänningsvärdet, medan frekvenskomponenterna kvadreras och divideras med två (omvandling från toppvärde till effektivvärde).

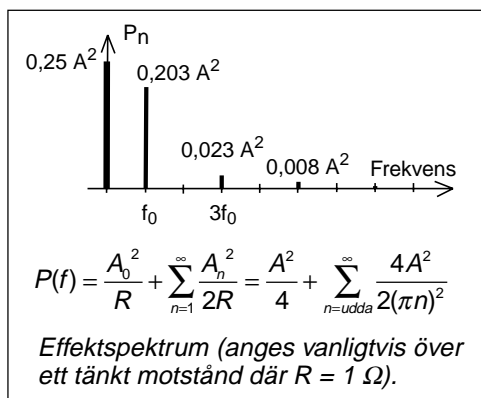
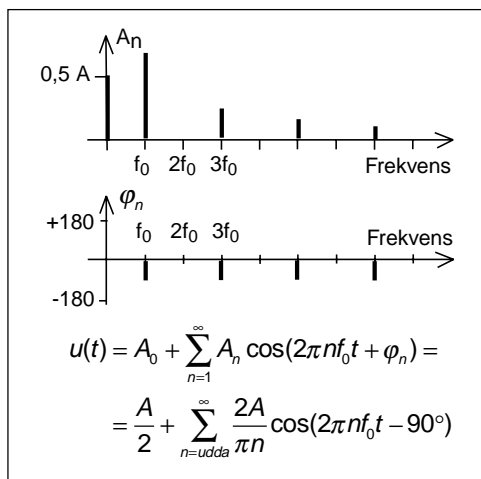
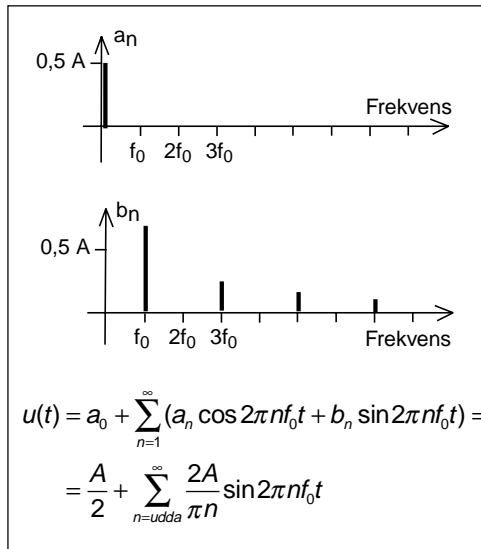
Vi kan rita ett effektspektrum över signalen $u(t)$ i motståndet R . Vi kvadrerar varje deltons effektivvärde och dividerar med resistansen. Ofta sätter man resistansen till det normerade värdet 1Ω .

På detta sätt kan vi enkelt se hur stor del av effekten i den totala signalen som finns vid olika frekvenser, och beräkna hur stor del av effekten som kommer fram om signalen passerar ett filter som inte släpper igenom alla deltoner.

Spektrumanalysator

En spektrumanalysator är en "mottagare" med oscilloskopskärm där effektspektrum ritas ut för inställt frekvensområde.

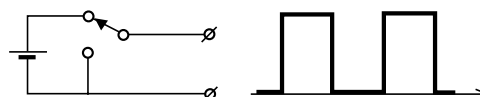
Spektrumanalysatorn har $R = 50 \Omega$ och effektaxeln är graderad i dBm.



2.5 Två sätt att alstra fyrkantvågen

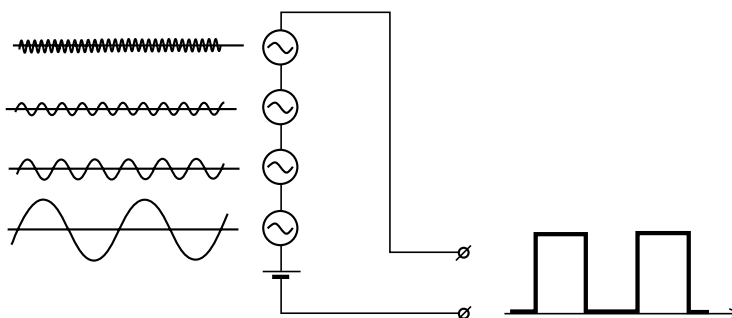
1. Signalen kan alstras i tidsdomänen

Fyrkantvågen kan alstras av en omkopplare som växlar i takt med fyrkantvågens periodtid.



2. Signalen kan alstras i frekvensdomänen

Man kan summera en likspänning och ett antal sinusvågor som är i gång hela tiden.



Syns ingen skillnad

Om vi tittar med ett oscilloskop på signalen så kan vi inte avgöra vilken av metoderna som använts för att skapa fyrkantvågen. Resultaten är identiska.

beskriva signalen i tidsplanet. För att avgöra om bandbredden räcker så måste vi få fram signalens amplitudspektrum.

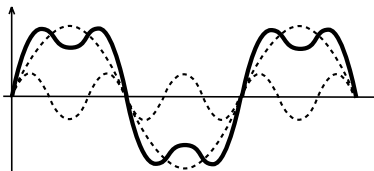
Skapa i tidsdomänen, men räkna i frekvensdomänen

Detta dubbla sätt att skapa en och samma signal kan vi utnyttja. Det är inget som hindrar att vi skapar signalen enligt metod 1, i tidsplanet, men räknar på signalen som om den skapats enligt metod 2, i frekvensplanet.

Antag att vi har en förstärkare som klipper signalen vid viss nivå och har begränsad bandbredd. För att avgöra om signalen kan förstärkas utan att klippas så måste vi

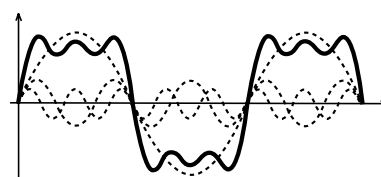
2 — Fourierserier och spektrum

Lägg ihop två sinustoner där den andra har tre gånger så hög frekvens och en tredjedel av amplituden. Man börjar kunna urskilja fyrkantvågen.



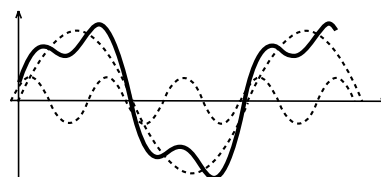
$$\sin 2\pi f_0 t + \frac{1}{3} \sin 2\pi 3f_0 t$$

Med tre sinustoner där den tredje har fem gånger så hög frekvens blir resultatet genast bättre.



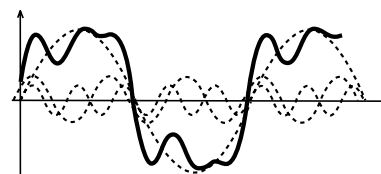
$$\sin 2\pi f_0 t + \frac{1}{3} \sin 2\pi 3f_0 t + \frac{1}{5} \sin 2\pi 5f_0 t$$

Sinustonerna måste ha inte bara rätt amplitud och frekvens, utan även rätt fasläge, komma i rätt tidsögonblick. I bilden kommer den tredje övertonen för tidigt, den som har tre gånger så hög frekvens. Trots att fasförskjutningen är minimal så blir vågformen hos summaspänningen kraftigt förändrad.



$$\sin 2\pi f_0 t + \frac{1}{3} \sin(2\pi 3f_0 t + 36^\circ)$$

Samma sak ser vi med tre sinus-spänningar.



$$\sin 2\pi f_0 t + \frac{1}{3} \sin(2\pi 3f_0 t + 36^\circ) + \frac{1}{5} \sin 2\pi 5f_0 t$$

På detta sätt kan man åstadkomma vilken periodisk vågform som helst bara genom att variera amplitud och fasläge hos de sinus-spänningar som läggs ihop.

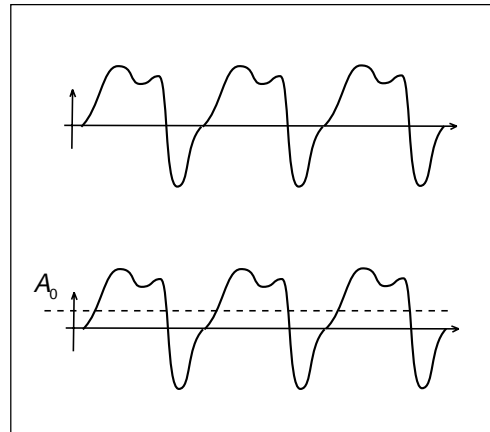
2.6 Hur fungerar Fourierserieutvecklingen?

Vi har en periodisk växelspanning $u(t)$, se figuren.

Likspänningskomponenten A_0 är enkelt, det är bara signalens medelvärde.

Men vilka frekvenser innehåller $u(t)$?

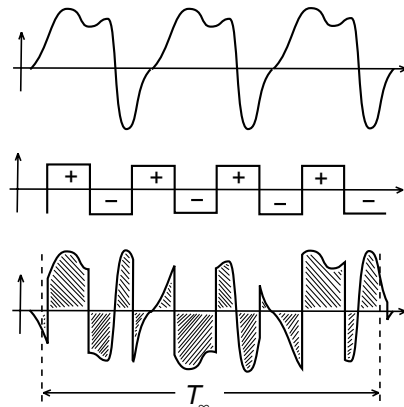
Tänk dig att du multiplicerar $u(t)$ med en fyrkantvåg som växlar mellan amplituderna $+1$ och -1 . Då kommer $u(t)$ att "byta sida" på de ställen där fyrkantvågen har amplituden -1 . Nu multiplicerar vi och integrerar, d.v.s. lägger ihop ytorna över x -axeln och drar ifrån ytorna under x -axeln. Dividera därefter med tiden, integrationsintervallet.



$$\frac{1}{T_\infty} \int_0^{T_\infty} u(t) \cdot [\text{fyrkant}] dt = ?$$

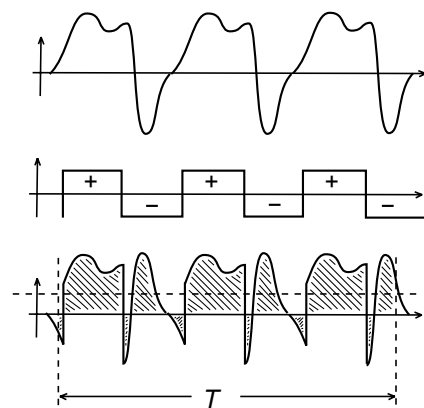
Multiplicera $u(t)$ med en fyrkantspänning ...

... lägg ihop ytorna ovanför x -axeln och dra ifrån ytorna under x -axeln samt dividera med tiden. Summan blir noll ...



... om inte fyrkantspänningen varierar i exakt samma takt som $u(t)$. Då får vi ett överskott på någon sida av x -axeln.

Om vi väljer integrationsintervallet T_∞ mycket långt, så kommer integralen att bli noll för alla periodtider hos fyrkantvågen utom de periodtider där fyrkantvågen varierar i takt med $u(t)$, har samma frekvens.



2.7 Ett matematiskt "filter"

Inte fyrkantvåg, men sinussignal

Det jag beskrev på föregående sida, att multiplicera med fyrkantvåg, fungerar inte i verkligheten. Men om vi byter fyrkantvågen mot en sinusformad signal så fungerar det. Orsaken till att det inte fungerar att multiplicera med fyrkantvåg är att fyrkantvågen själv är summan av flera sinusvågor.

Nu antar vi att $u(t)$ består av likspänning och olika sinusformade signaler. Vi multiplicerar med en sinusvåg med den frekvens vi vill undersöka. Vad händer när man multiplicerar ihop? Titta i trigonometriska formelsamlingen.

Alla termer blir nya sinussignaler som blir noll när man integrerar, UTOM när sinus-signalen multipliceras med sig själv, formel (12) och (13). Där får vi en konstant som blir kvar efter integreringen.

Eftersom konstanten är 0,5 så måste integralen multipliceras med 2 för att vi skall få den sökta sinussignalens amplitud U_x .

Rätt frekvens, men fasvinkel?

Om den sökta signalen har en av oss okänd fasvinkel så börjar vi med att undersöka hur mycket av signalen som är cosinussignal och hur mycket som är sinussignal. Därefter kan fasvinkeln beräknas. OBS att man får in ett minustecken framför sinusamplituden. Därför har vi minustecknet vid beräkning av fasvinkeln vid övergång till polär Fourierserie.

$$u(t) = U_0 \\ \Rightarrow U_0 \cdot \sin(2\pi f_x t) = U_0 \sin(2\pi f_x t)$$

$$u(t) = U_1 \sin(2\pi f_1 t) \\ \Rightarrow U_1 \sin(2\pi f_1 t) \cdot \sin(2\pi f_x t) = \\ = \frac{U_1}{2} \cos 2\pi(f_1 - f_x)t - \frac{U_1}{2} \cos 2\pi(f_1 + f_x)t$$

$$u(t) = U_x \cos(2\pi f_x t + \varphi_x) = \\ = U_x \cos \varphi_x \cdot \cos(2\pi f_x t) - \\ - U_x \sin \varphi_x \cdot \sin(2\pi f_x t) \\ \Rightarrow U_x \sin(2\pi f_x t) \cdot \sin(2\pi f_x t) = \\ = U_x \sin^2(2\pi f_x t) = \\ = \frac{U_x}{2} - \frac{U_x}{2} \cos 2\pi(2f_x)t$$

$$u(t) = U_x \cos(2\pi f_x t + \varphi_x) = \\ = U_x \cos \varphi_x \cdot \cos(2\pi f_x t) - U_x \sin \varphi_x \cdot \sin(2\pi f_x t)$$

2.8 När $u(t)$ är periodisk

Det finns bara en uppsättning sinusformade signaler som kan bilda $u(t)$. Signalen $u(t)$ har bara ett spektrum.

Frekvenser

Om $u(t)$ är periodisk med periodtiden T_0 , d.v.s. varje period är verkligen lik de övriga, då måste den uppsättning sinusformade signaler som bildar $u(t)$, dessa sinusformade signaler måste ha sina nollgenomgångar på exakt samma ställe inom varje period. Detta är bara möjligt om dessa sinusformade signaler har frekvenser som är multiplar av grundfrekvensen.

Vi behöver alltså bara beräkna Fourierintegralerna för frekvenser som är "grundton och övertoner" till periodtiden för $u(t)$.

Integrationstid

När $u(t)$ är periodisk, d.v.s. upprepar sig med periodtiden T_0 , så räcker det att integrera över en period. Ytterligare perioder innebär "mer yta", men integralen skall divideras med längre integrationstid. Resultatet blir det samma.

Deltoner eller övertoner?

Frekvenskomponenten som vi får om t.ex. $n = 3$, är detta tredje tonen, tredje deltonen, andra övertonen eller tredje övertonen?

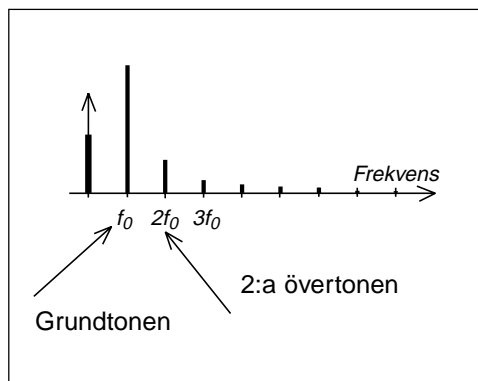
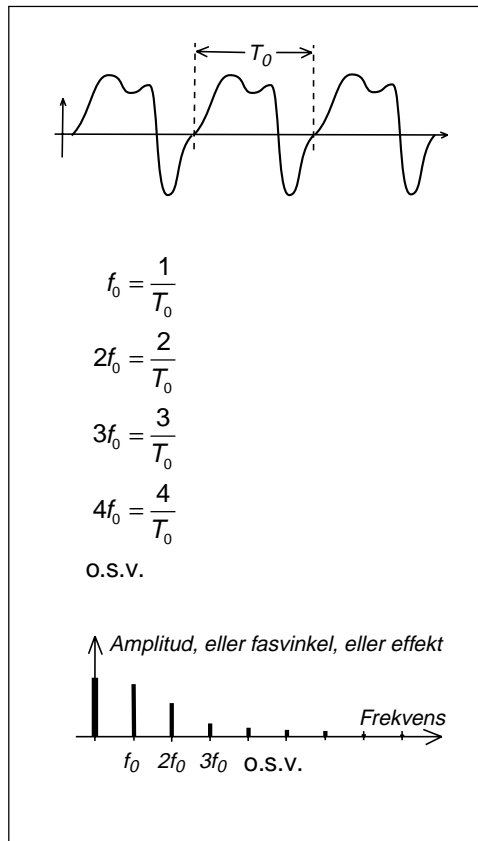
Harmonics

Den engelska benämningen är "harmonic". När $n = 2$ har vi "2:nd harmonic". Att kalla detta andra deltonen är logiskt. Detta språkbruk försöker man använda i undervisningen för att slippa missförstånd.

Men när du kommer ut bland radioingenjörer, då kallar alla den delton du får när $n = 2$ för andra övertonen.

Gammalt svenskt språkbruk är att allting är övertoner utom grundtonen som kallas grundton. När $n = 1$ har vi grundtonen, $n = 2$ är andra övertonen, $n = 3$ är tredje övertonen o.s.v.

Om du störs av att övertonerna tycks ha fel nummer, kalla dem då för deltoner.



Gör vad du vill, bara du inte kallar tredje övertonen för andra övertonen. Då kommer alla att missförstå dig. Deltonsnumret eller övertonsnumret skall vara lika med n .

2.9 Jämn eller udda funktion?

Vid beräkning av spektrat för fyrkantvågen blev samtliga cosinustermer noll. Om man kan se på funktionen i förväg att cosinustermerna eller sinustermerna blir noll, då slipper man lösa denna integral. Detta har lett fram till begreppen jämn respektive udda funktion.

Jämn funktion

En funktion $f(t)$ kallas jämn om $f(t)$ har perfekt symmetri runt y-axeln.

$$f(+t) = f(-t)$$

Tag ett värde på x-axeln, avläs y-värdet. Gå lika långt på negativa x-axeln. Du skall ha samma y-värde, med samma tecken.

Jämn funktion innehåller bara a_n -termer, cosinustermer och likspänning

Spektrat för den jämna funktionen innehåller bara a_n -termer, d.v.s. eventuell likspänning och cosinustermer. Om funktionen är jämn slipper du beräkna integralen för b_n -termerna.

Udda funktion

En funktion $f(t)$ kallas udda om $f(t)$ har perfekt symmetri runt både y-axeln och x-axeln.

$$f(+t) = -f(-t)$$

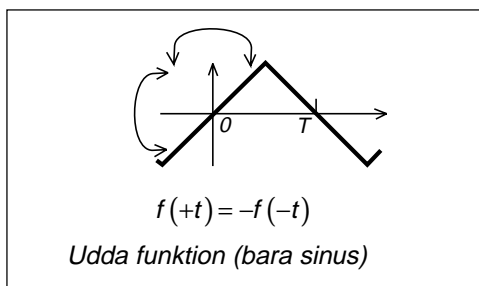
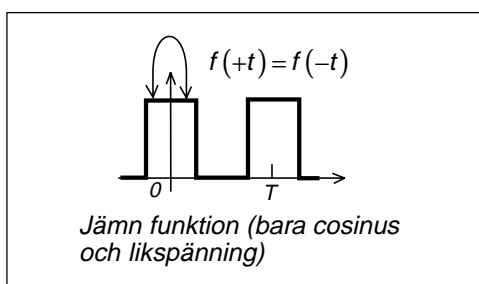
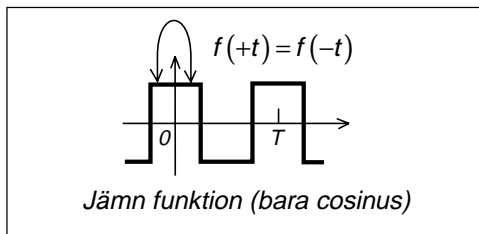
Tag ett värde på x-axeln, avläs y-värdet. Gå lika långt på negativa x-axeln. Du skall ha y-värdet, med motsatt tecken.

Udda funktion innehåller bara b_n -termer, sinustermer

Spektrat för den jämna funktionen innehåller bara b_n -termer, endast sinustermer. Om funktionen är jämn slipper du beräkna a_n -integralerna.

”Jämn funktion”, ”udda funktion”, eller ”varken eller”

Det finns inte bara udda eller jämna funktioner. Den största gruppen är varken jämn eller udda, den innehåller såväl a_n - som b_n -



termer, och då måste man lösa alla tre integralerna.

2.10 Fourierserietutveckling i praktiken

Amplitudspektrat påverkas inte om vi förskjuter signalen i "sidled"

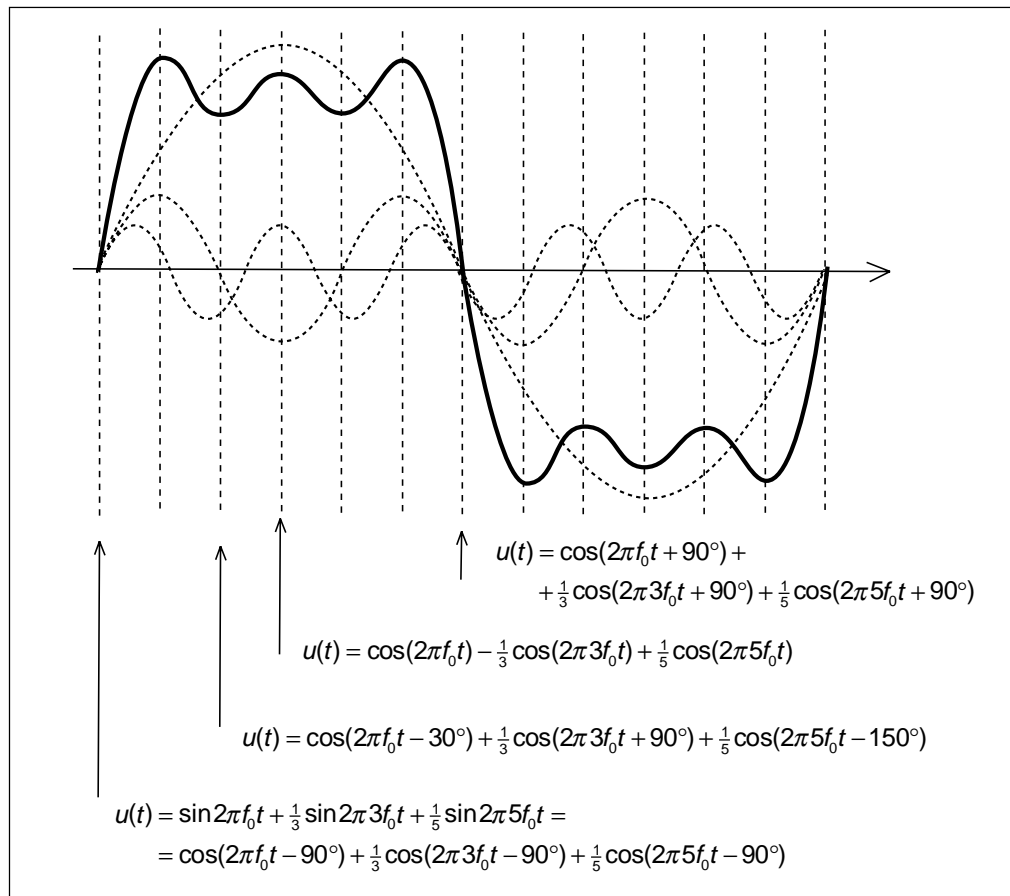
Bilden nedan visar hur fyrkantvägens sinus-toner matematiskt beskrivs på olika sätt beroende på var vi placerar $t = 0$.

Inget annat än fasvinklarna förändras när vi flyttar tidpunkten som vi kallar $t = 0$. Det är alltså bara fasspektrat som påverkas, inte amplitudspektrat.

förändras är likspänningskomponenten A_0 i spektrat.

Bara likspänningen påverkas om vi förskjuter signalen i "höjdled"

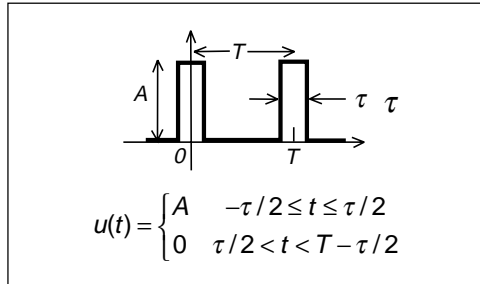
På motsvarande sätt förändras inte cosinustonernas amplitud eller fasvinkel om signalen förskjuts i "höjdled" genom att överlagras på en likspänning. Det enda som



2.11 Korta pulser

En signal som består av korta pulser är speciellt viktig vid telekommunikation eftersom digitala signaler ofta kan ses som summan av sådana signaler. Pulserna kan beskrivas så som i bilden.

Vi lägger $t = 0$ så att signalen blir jämn. Då slipper vi lösa integralen för b_n -termerna.



$$f_0 = \frac{1}{T}$$

$$a_0 = \frac{1}{T} \int_{-T/2}^{T/2} u(t) \cdot dt = \frac{1}{T} \int_{-\tau/2}^{\tau/2} A \cdot dt = \frac{A}{T} \left[\frac{\tau}{2} - \frac{-\tau}{2} \right] = \frac{A\tau}{T}$$

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} u(t) \cdot \cos 2\pi n f_0 t \cdot dt = \frac{2}{T} \int_{-\tau/2}^{\tau/2} A \cdot \cos 2\pi n f_0 t \cdot dt =$$

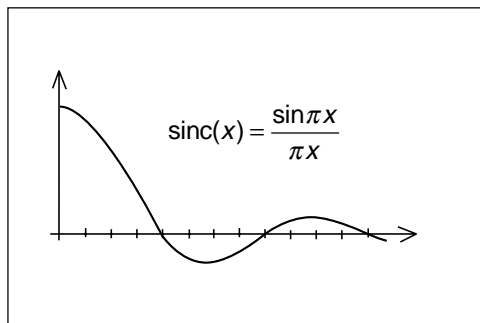
$$= \frac{2A}{2\pi n f_0 \cdot T} \left[\sin \left(2\pi n \frac{1}{T} \frac{\tau}{2} \right) - \sin \left(2\pi n \frac{1}{T} \frac{-\tau}{2} \right) \right] = \frac{2A}{\pi n} \sin \left(\pi n \frac{\tau}{T} \right) = 2A \frac{\tau}{T} \cdot \frac{\sin \left(n\pi \frac{\tau}{T} \right)}{n\pi \frac{\tau}{T}}$$

Den pulsförmade signalens spektrum får följande utseende:

$$u(t) = A \frac{\tau}{T} + \sum_{n=1}^{\infty} 2A \frac{\tau}{T} \cdot \frac{\sin(n\pi\tau/T)}{n\pi\tau/T} \cdot \cos 2\pi n f_0 t$$

Amplitudspektrat kommer att följa funktionen $\sin(\pi x)/\pi x$ som förekommer så ofta vid digital kommunikation att den fått ett eget namn.

$$\text{sinc}(n\tau/T) = \frac{\sin(n\pi\tau/T)}{n\pi\tau/T}$$



2 — Fourierserier och spektrum

Här ser du amplitudspektrat om periodtiden är fyra gånger pulsbredden. Då blir fjärde deltonen noll.

Om periodtiden är åtta gånger pulsbredden kommer åttonde deltonen att bli noll.

I detta fall har jag behållit pulsbredden och pulshöjden. Pulserna kommer mer sällan. Detta ger hälften så hög höjd på spektralkomponenterna. Hälften så hög höjd ger en fjärdedel av effekten. Men spektralkomponenterna kommer dubbelt så tätt. Två gånger en fjärdedel ger halva effekten. Detta stämmer. Om pulserna kommer hälften så tätt skall man få halva effektutvecklingen.

Vid datakommunikation kommer pulserna olika tätt beroende på datainnehållet. Detta ger varierande periodtid. Men så länge pulsbredden är densamma får spekrat noll vid exakt samma frekvens hela tiden oavsett datainnehåll. Vid digital rundradio, t.ex. DAB och DTV, så lägger man en annan signal vid denna frekvens och slipper att signalerna stör varandra.

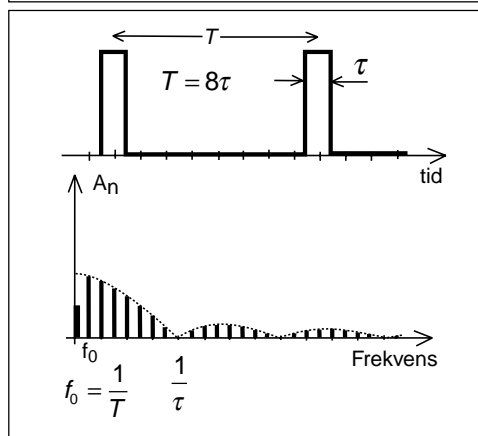
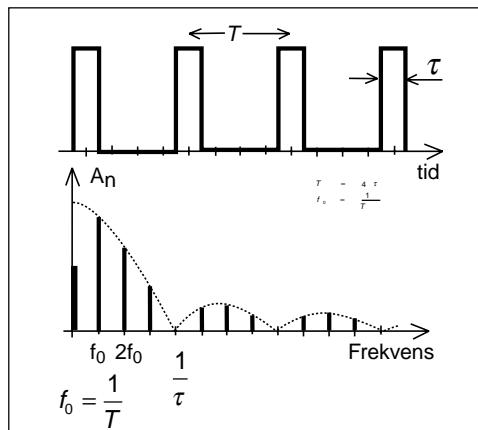
Den pulsformade signalen och effektutveckling

Den pulsformade signalen $u(t)$ som vi räknat på består av korta likspäningspulser med amplituden A . Mellan pulserna är amplituden noll, d.v.s. där finns ingen signal.

Om $u(t)$ matas in i ett motstånd så kommer värme att produceras under själva spänningspulsen. Däremellan sker ingen värmeutveckling.

När vi beräknar effektutvecklingen genom att se på signalens spektrum $U(f)$, då summerar vi effektutvecklingen som åstadkommes av likspänningskomponenten och samtliga deltoner. Dessa signaler finns hela tiden och producerar värme även mellan pulserna där amplituden för $u(t)$ är noll. Kan detta verkligen stämma?

Ja, det stämmer. Den medeleffekt vi får under en period genom att räkna på själva pulsen, $u(t)$, blir samma effekt som vi får genom att summera alla bidrag i signalens spektrum $U(f)$. Spektrat innehåller nämligen



ingen information om när saker och ting händer. Det finns inget i spekrat som talar om när pulsen har amplituden A respektive när pulsen har amplituden noll. För att få fram denna information, tidsinformationen, så måste vi summera samtliga spektralkomponenter i $U(f)$, för att få tidsfunktionen $u(t)$.

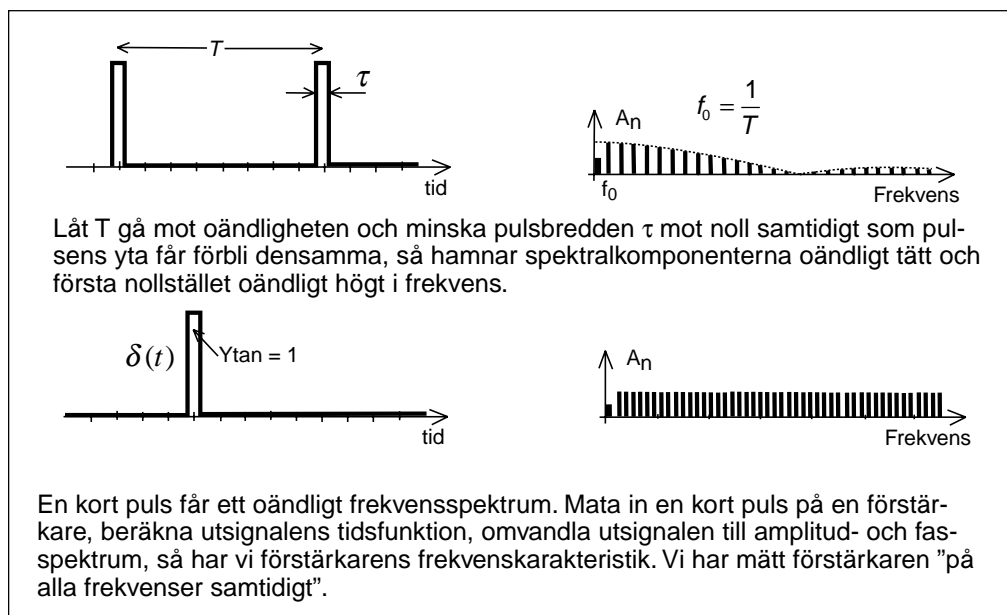
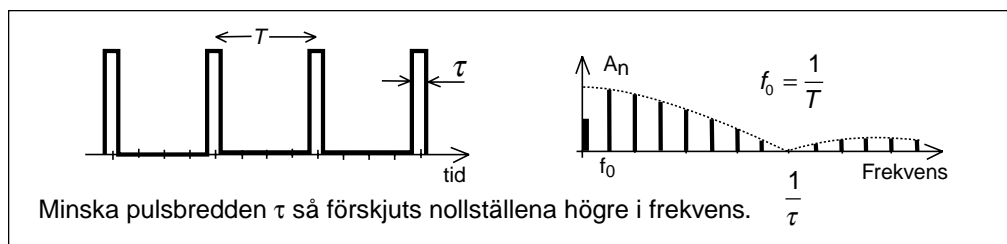
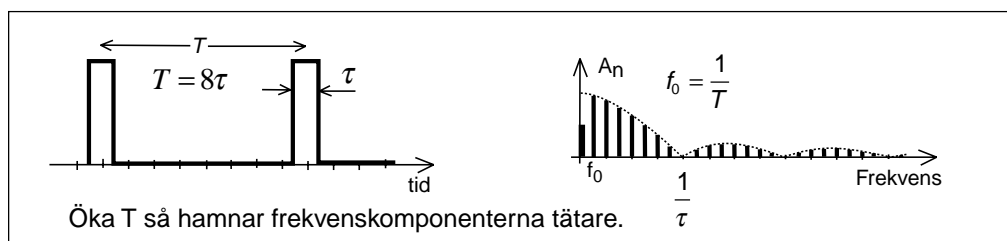
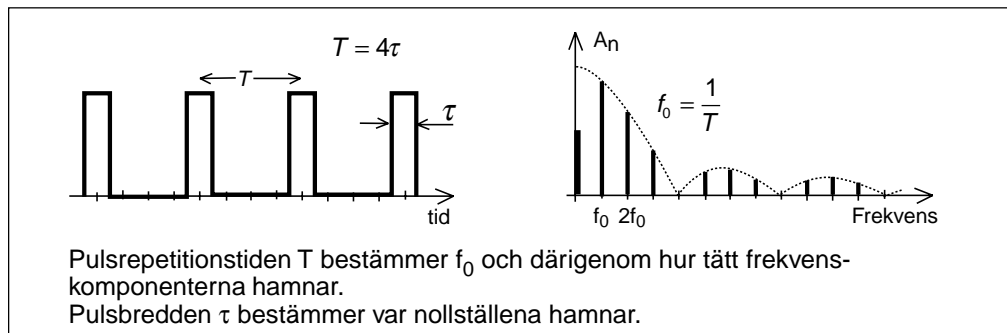
Det är detta som gör det så svårt

Vår hjärna tänker sekventiellt, först händer detta, sedan det där. Vi har lättast att hänga med när signalen beskrivs som funktion av tiden $u(t)$.

Vi kan beräkna spektrat. Vi kan rita spektrat. Vi kan med viss approximation se eller mäta spektrat med en spektrum-analysator. Men förstår vi? Och framför allt, förstår vi att vi måste lämna tidstänkesättet helt när vi ger oss in i frekvensplanet?

2 — Fourierserier och spektrum

17 lektioner i
TELEKOMMUNIKATION



2.12 Komplexa Fourier-serier, Fourier-transformen

Från matematiken vet du att $\cos\alpha = \frac{e^{j\alpha} + e^{-j\alpha}}{2}$

Använd detta så får vi: $u(t) = A_0 + \sum_{n=1}^{\infty} \frac{A_n}{2} (e^{j(2\pi n f_0 t + \varphi_n)} + e^{-j(2\pi n f_0 t + \varphi_n)}) =$

$$= A_0 + \sum_{n=1}^{\infty} \frac{A_n}{2} e^{j\varphi_n} e^{j2\pi n f_0 t} + \sum_{n=1}^{\infty} \frac{A_n}{2} e^{-j\varphi_n} e^{-j2\pi n f_0 t}$$

Inför: $\begin{cases} c_0 = A_0 \\ c_n = \frac{A_n}{2} e^{j\varphi_n} \\ c_{-n} = \frac{A_n}{2} e^{-j\varphi_n} \end{cases}$ så får vi:

$$u(t) = c_0 + \sum_{n=1}^{\infty} c_n e^{j2\pi n f_0 t} + \sum_{n=1}^{\infty} c_{-n} e^{-j2\pi n f_0 t} = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_0 t}$$

där $c_n = \frac{1}{T_0} \int_0^{T_0} u(t) \cdot e^{-j2\pi n f_0 t} \cdot dt$

Komplex Fourierserie

Ovan har jag visat hur $u(t)$ skrivs som komplex Fourierserie.

Övergången från cosinus till komplex e-funktion ger oss två e-funktioner, två oändliga serier, den ena med positiv e-funktion, den andra med negativ e-funktion.

Dessa båda serier sammanförs i en serie genom att införa negativ frekvens. Vi ser också att c_n -värdet bara har halva amplituden jämfört med trigonometriska serieutvecklingen. Den andra halvan av amplituden finns på negativa frekvensaxeln.

Den komplexa Fourierserien ger alltså ett spektrum som sträcker sig över såväl positiva som negativa frekvenser. Detta är matematik och har ingen koppling till verkligheten.

Om vi väljer $e^{-j2\pi n f_0 t}$ eller $e^{j2\pi n f_0 t}$ vid beräkning av c_n , d.v.s. plus- eller minustecken, spelar egentligen ingen roll fysikaliskt. Plus-tecken ger annat tecken på imaginärdelen hos c_n , det påverkar inte amplituden. Dessutom påverkas såväl positiv som negativ frekvens. När man går tillbaka från komplex Fourier-serie till tidsfunktion så får vi samma $u(t)$ oavsett hur vi definierar vårt "komplexa Fourier-rum".

Fouriertransformen

Om signalen $u(t)$ inte är periodisk, d.v.s. om T_0 går mot oändligheten, övergår de komplexa Fourier-serierna i Fouriers integralsats

som gäller även för ickeperiodiska signaler:

$$u(t) = \sum_{-\infty}^{\infty} c_n e^{j2\pi n f_0 t} \Rightarrow \int_{-\infty}^{\infty} U(f) e^{j2\pi f t} df$$

När T_0 går mot oändligheten så får man problem vid beräkning av c_n -värdena, de går mot noll. Därför multiplicerar man c_n med T_0 varefter T_0 får gå mot oändligheten. Detta ger ett ändligt värde ...

$$U(f) = \lim_{T_0 \rightarrow \infty} c_n \cdot T_0 = \int_{-\infty}^{\infty} u(t) e^{-j2\pi f t} \cdot dt$$

... om $u(t)$ är en funktion som är begränsad i tiden. Det finns därför signaler för vilka man inte kan skapa Fouriertransformen.

Laplacetransformen

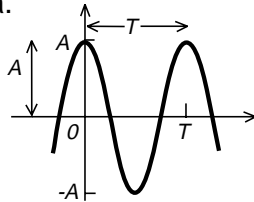
Om signalen inte är begränsad i tiden så kan vi dämpa signalen så att den avtar mot noll. En sådan dämpningsfunktion är e-funktionen. Vid beräkning av Laplacetransformen så dämpas signalen, vid omvandling från Laplacetransform till tidsfunktion förstärks signalen med motsvarande e-funktion.

$$U(s) = \int_0^{\infty} u(t) e^{-\sigma t} e^{-j2\pi f t} dt =$$

$$= \int_0^{\infty} u(t) e^{-(\sigma + j2\pi f)t} dt = \int_0^{\infty} u(t) e^{-st} dt$$

2.13 Några vanliga funktioner

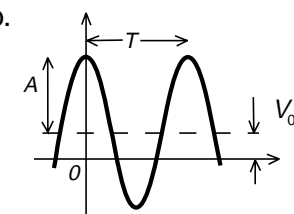
a.



$u(t) = A \cos 2\pi f_0 t$
 $f_0 = \frac{1}{T}$

Detta är den vanliga cosinussignalen där frekvensen är ett dividerat med periodtiden.

b.



$u(t) = V_0 + A \cos 2\pi f_0 t$

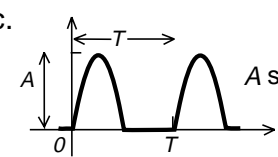
Om signalen ligger likspänningsförskjuten börjar man med att subtrahera likspänningen och beräkna spektrat för själva växelspänningen:

$$u(t) - V_0 = A \cos 2\pi f_0 t$$

Därefter lägger man tillbaka likspänningen:

$$u(t) - V_0 + V_0 = A \cos 2\pi f_0 t + V_0$$

c.



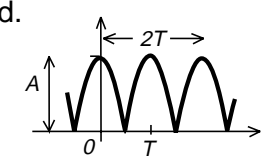
$A \sin 2\pi f_0 t$

Detta är halvågslikriktning där man behåller ena halvperioden av sinusspänningen. Vi har egentligen "en halv sinusspänning" plus "en halv halvågslikriktning" av sinusspänningen, se halvågslikriktning nedan.

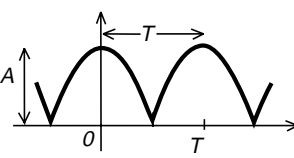
$$u(t) = \frac{A}{\pi} + \frac{A}{2} \sin 2\pi f_0 t - \frac{2A}{3\pi} \cos 2\pi(2f_0)t - \frac{2A}{15\pi} \cos 2\pi(4f_0)t - \dots$$

$$= \frac{A}{\pi} + \frac{A}{2} \sin 2\pi f_0 t + \sum_{n=2}^{\infty} \frac{A[1 + (-1)^n]}{\pi(1 - n^2)} \cos 2\pi(nf_0)t$$

d.



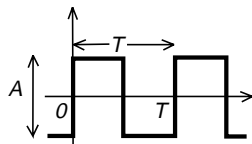
$u(t) = \frac{2A}{\pi} + \frac{4A}{3\pi} \cos 2\pi f_0 t - \frac{4A}{15\pi} \cos 2\pi(2f_0)t + \dots$

$$= \frac{2A}{\pi} + \sum_{n=1}^{\infty} \frac{4A(-1)^n}{\pi[1 - (2n)^2]} \cos 2\pi(nf_0)t$$


Detta är halvågslikriktning av cosinusspänningen. Observera att frekvensen plötsligt blivit en annan. Den halvågsliktade spänningen upprepar sig med halva den periodtid som cosinusspänningen hade. Halvågslikriktning av 50 Hz ger $f_0 = 100$ Hz.

2 — Fourierserier och spektrum

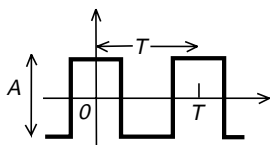
e.



$$u(t) = \frac{2A}{\pi} \sin 2\pi f_0 t + \frac{2A}{3\pi} \sin 2\pi(3f_0)t + \dots$$

$$= \sum_{n=\text{udda}} \frac{2A}{n\pi} \sin 2\pi(nf_0)t$$

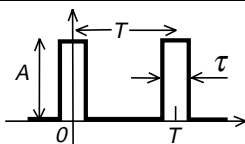
f.



$$u(t) = \frac{2A}{\pi} \cos 2\pi f_0 t - \frac{2A}{3\pi} \cos 2\pi(3f_0)t + \frac{2A}{5\pi} \cos 2\pi(5f_0)t - \dots$$

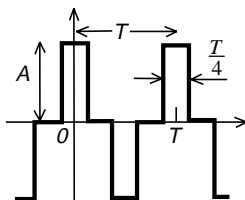
$$= \sum_{n=1}^{\infty} \left(A \frac{\sin n\pi/2}{n\pi/2} \right) \cos 2\pi(nf_0)t$$

g.



$$u(t) = A \frac{\tau}{T} + \sum_{n=1}^{\infty} \left(2A \frac{\tau}{T} \right) \left(\frac{\sin 2n\pi\tau/T}{n\pi\tau/T} \right) \cos 2\pi(nf_0)t$$

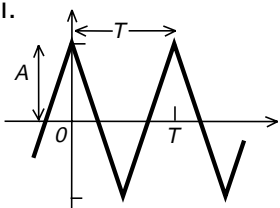
h.



$$u(t) = \sum_{n=\text{udda}} \left(A \frac{\sin n\pi/4}{n\pi/4} \right) \cos 2\pi(nf_0)t$$

Observera att uttrycket bara gäller när "pulsbredden" är periodtiden dividerat med fyra.

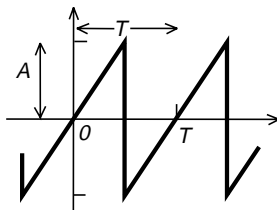
i.



$$u(t) = \frac{8A}{\pi^2} \cos 2\pi f_0 t + \frac{8A}{9\pi^2} \cos 2\pi(3f_0)t + \frac{8A}{25\pi^2} \cos 2\pi(5f_0)t + \dots$$

$$= \sum_{n=\text{udda}} \frac{8A}{(n\pi)^2} \cos 2\pi(nf_0)t$$

j.



$$u(t) = \frac{2A}{\pi} (\sin 2\pi f_0 t - \frac{1}{2} \sin 2\pi(2f_0)t + \frac{1}{3} \sin 2\pi(3f_0)t - \dots)$$

$$= \sum_{n=1}^{\infty} [(-1)^{n-1}] \left(\frac{2A}{n\pi} \right) \sin 2\pi(nf_0)t$$

2.14 Vad är egentligen ett spektrum?

Du har fått se att en signal som ändrar sig i tiden, så som vi är vana att tänka oss signaler, att en sådan signal kan åstadkommas genom att man summerar en eller flera sinusformade signaler som är stationära, inte ändrar sig.

Vi började med fyrkantvågen som alstras med ett batteri och en strömbrytare. Ibland finns signalen, ibland finns den inte.

Men spektrat av denna signal finns HELA TIDEN, även i mellanrummen då vår signal inte finns!

Jag kan filtrera ut en av spektralkomponenterna och se att spektralkomponenterna finns hela tiden.

Men vad händer efteråt, när jag kopplat bort batteriet? Fortsätter spektralkomponenterna? Stationär växelspanning skall ju aldrig ändras.

När signalen ej längre är periodisk blir perioden lång. Spektralkomponenterna hamnar så tätt, att jag inte kan mäta på endast en. Hur smalt filter jag än skaffar så släpper det igenom flera frekvenskomponenter som släcker ut varandra.

Men om signalen på detta sätt finns efteråt, då måste den även ha funnits före. Spektralkomponenterna måste ha funnits från början. Batteriet åstadkommer en förändring som vi uppfattar som fyrkantvågen.

I spektrat finns ingen tid. Där råder evighet. När skapades universum? Big Bang? I tidsplanet finns en början när allt small till, men spektrat d.v.s. energin måste ha funnits. I tidsplanet finns en början, i spektrat råder evigheten.

Fourier-serier

Uppdelning av en signal i Fourierkomponenter går att begripa. Fysikaliskt kan vi tänka oss signalens spektrum.

Komplexa Fourierserier

När man omvandlar den trigonometriska Fourierserien till komplex Fourierserie så får vi både ett positivt (positiva frekvenser) och ett negativt (negativa frekvenser) spektrum. Vi lämnar verkligheten, men summan av

amplituderna vid negativ och positiv frekvens ger oss fortfarande "rätt" amplitud.

Införandet av de komplexa talen underlättar vissa matematiska beräkningar.

Fouriertransformen

Komplex Fourierserie kan bara tillämpas på periodiska signaler, signaler som upprepar sig (i oändligheten). I Fouriertransformen låter man "periodtiden" gå mot oändligheten, man kan räkna på enstaka pulser.

Det spektrum som man nu får har ingen likhet med spektrat för komplex Fourierserie. Vid Fouriertransformen multiplicerar vi av matematiska skäl båda leden med "periodtiden", d.v.s. oändligheten vid beräkning av spektralkomponenterna, som nu ligger oändligt tätt.

Signalens energiinnehåll måste vara ändligt, annars fungerar inte Fourierintegralen.

Laplacetransformen

Vid Laplacetransformen multiplicerar vi signalen med en dämpningsfunktion $e^{-j\sigma}$ vid beräkning av "spektralkomponenterna", och förstärker signalen med $e^{j\sigma}$ när vi går tillbaka från spektrum till tidsfunktion. På så vis klarar vi att även behandla signaler som inte är begränsade.

Varför?

När man skall addera två signaler så kan det vara lättare att först beräkna signalernas spektrum och lägga ihop spektralkomponenterna och sedan gå tillbaka till tidsfunktionerna, än att direkt addera tidsfunktionerna. Så här gör man i signalprocessorer vid DAB (digital ljudradio) och DTV (digitala marknätet för tv).

3 — Brus

Ställ in tv:n på en tom kanal, en kanal där det inte pågår någon sändning. Vad ser du? "Myrkrig"! Detta är brus. Så ser brus ut. Troligtvis hör du inte bruset därför att tv:n automatiskt stänger av högtalaren så att du slipper bruset.

Brus är en signal där vi inte känner signalens tidsfunktion, utan bara dess amplitudspektrum.

Eftersom brus alstras av många oberoende generatorer så vet vi att tidsfunktionens amplitud ibland kan bli mycket hög, och att amplitudtäthetsfunktionen följer normalfördelningskurvan, Gausskurvan.

Vid telekommunikation, vid mottagaren, så har vi brus från källan. På ledningen kommer den önskade signalen men även brus, t.ex. en antennbrustemperatur T_{antenn} .

Vi har även brus som bildas i mottagarens olika förstärkarsteg. Mottagarbruset summeras och räknas om till en ekvivalent mottagarbrustemperatur T_{eq} på mottagarens ingång.

3.1 Termiskt brus

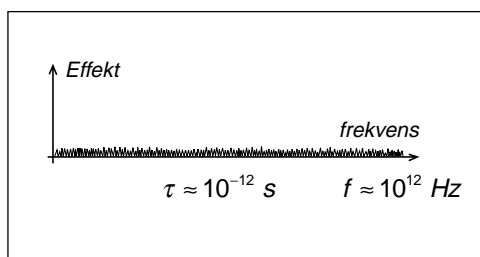
Vår föreställning om elektrisk ledningsförmåga är att atomerna hos elektriskt ledande material har löst bundna elektroner som kan hoppa från elektron till elektron och på det sättet transportera laddning.

Det flyter ström om man lägger på en spänning. Om det inte finns någon spänning så borde dessa fria elektroner stå still, och det gör de om ledaren kyls ner till absoluta nollpunkten, -273°C (0 Kelvin).

Om ledaren är varm så tillförs termisk energi till elektronerna som gör att de sporadiskt hoppar till en annan atom trots att vi inte lagt på någon spänning. Detta sporadiska hoppande sker i vilken riktning som helst och oftare ju varmare det är, proportionellt mot den termiska energin, temperaturen i Kelvin.

Elektroner som rör sig är elektrisk ström. Elektroner som hoppar ger upphov till strömpulser. Man har uppskattat pulsbredden till $\tau = 10^{-12}$ och av detta fått fram att spektrat bör ha i stort sett konstant amplitud upp till 10 – 100 GHz för att sedan avta upp mot 1000 GHz.

När elektronen hoppar så uppstår en lokal strömpuls där elektronen befinner sig. Denna strömpuls kan inte förorsaka något såvida inte elektronen hoppar i ett motstånd, för då får vi spänningsfall, en spänningspuls, och denna spänningspuls kan driva iväg ström på ledningen. Då avger motståndet bruseffekt.



3.2 Många generatorer

- Antag att vi har EN generator som lämnar 100 W i ett motstånd på 1 Ω . Spänningen över motståndet är en sinus-spänning med 10 volt effektivvärde, och högsta positiva eller negativa spänningstoppens är 14,14 volt.
- Vi byter ut generatorn mot 100 generatorer **på olika frekvenser** som var och en lämnar 1 W i motståndet på 1 Ω . Varje generator ger en sinus-spänning med 1 volt effektivvärde och 1,414 volt toppvärde över motståndet.
- Alla generatorerna tillsammans kommer att lämna 100 W till motståndet. Men vad blir toppspänningen?

Summaspänningen består av 100 spänningsvektorer som alla snurrar med olika hastighet (olika frekvenser) i visardiagrammet. Därför finns möjligheten att alla 100 spänningsvektorerna någon gång pekar åt samma håll, och toppspänningen blir 100 gånger 1,414 volt, d.v.s. 141,4 volt. Detta händer inte ofta, men det skulle kunna hända.

- Med hjälp av statistik kan man beräkna sannolikheten för att summaspänningen skall uppnå ett visst spänningsvärde.

Hur beskriver vi signalen från alla dessa 100 generatorer?

- Generatorerna lämnar 100 W effekt. På samma sätt med bruset. Vi beräknar effekten inom viss bandbredd.
- Spänningens amplitud kan beskrivas med statistik. På samma sätt med bruset.
- Men vi kan också säga att spänningen vars amplitud vi inte känner har **effektivvärdet 10 volt** eftersom denna spänning utvecklar 100 W i motståndet.

Mer om många generatorer

Det är inte bara brus som kan ge upphov till mycket höga toppspänningar trots att effekten är låg. Samma fenomen har man vid mobiltelefoni. Den ökande användningen av radiosändare, där effekten från varje enskild

radiosändare visserligen är låg, innebär samtidigt ökad risk för att det då och då kan uppträda mycket höga toppfältstyrkor.

Elektronikutrustning som blir störd av radiosändare påverkas oftast inte av sändarnas effekt utan av fältstyrkans toppvärde. Elektronikutrustning som inte är en radiomottagare innehåller inget filter för mottagning av viss frekvens, utan reagerar för radiosignaler inom ett brett frekvensområde, d.v.s. det är summafältstyrkan från många radiosändare på olika frekvenser som åstadkommer störningen.

Vid EMC-provning (elektromagnetisk kompatibilitet, förmågan att inte störas av elektromagnetiska fält) testas elektronikutrustning med en enda störsändare. Ett flygplan kan mycket väl testas med en eller två störsändare. Men vad händer om 300 ficktelefoner tappat kontakten med GSM-nätet och allihop upptäcker en ny basstation och börjar sända, anropa basstationen, när flygplanet håller på att landa? Sådana anrop utför ficktelefonen automatiskt för att tala om för GSM-nätet var den befinner sig. Enda möjligheten att förhindra detta är att ha ficktelefonerna avstängda.

300 ficktelefoner som sänder med 2 W kan störa lika kraftigt som en enda sändare som sänder med

$$(300)^2 \cdot 2 W = 180\,000 W$$

d.v.s. totala effekten multiplicerat med antalet sändare i kvadrat.

På samma sätt kan bruset från elektronerna ge upphov till toppspänningar som motsvarar summabrusets effektivvärde multiplicerat med roten ur antalet elektroner som brusar!

3.3 Motstånd brusar

Hur mycket bruseffekt kan ett motstånd lämna?

Brusets spektraltäthet

Motståndet kan maximalt lämna den effekt som frigörs när elektronerna hoppar, och den effekten kan inte vara större än tillförd termisk effekt.

Bruseffekten får ett effektspektrum som sträcker sig, med konstant amplitud (vitt brus), från låg frekvens, nästan 0 Hz, upp till nästan 100 GHz.

Men vi är inte intresserade av totala bruseffekten. Vi vill beräkna hur stor bruseffekt som tar sig igenom en förstärkare som har viss bandbredd B Hz. Därför har man beräknat summan av de spektralkomponenter i effektspektret som ryms inom en bandbredd av 1 Hz. Denna summa är oberoende av var vi placerar bandbredden 1 Hz inom frekvensområdet 0 Hz upp till 100 GHz (vitt brus).

Maximal bruseffekt lämnas vid anpassning

Motståndet lämnar ifrån sig maximal effekt till en lastresistans som har samma resistansvärde som det brusande motståndet.

Brusspänningens effektivvärde E_n

Det enda vi vet om brusspänningen är att den skall producera viss bruseffekt i lastresistansen. Vi kan därför räkna fram en ekvivalent spänning med konstant amplitud som ger samma effektutveckling. Detta är brusspänningens effektivvärde.

Nu beskriver vi det brusande motståndet som ett brusfritt motstånd i serie med en spänningskälla med en brusspänning vars effektivvärde är E_n .

Brusets spektraltäthet

Bruseffekt (termiskt brus) inom bandbredden 1 Hz:

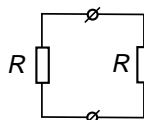
$$N_0 = k \cdot T \text{ [W / Hz]}$$

$$k = 1,38 \cdot 10^{-23} \text{ Boltzmanns konstant}$$

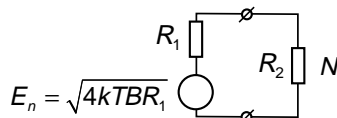
$$T = \text{temperaturen uttryckt i Kelvin (K)}$$

$$= 273 + \text{temperatur i grader Celcius}$$

Motstånd
som brusar



Brusfritt motstånd med
brusspänningskälla



$$N = \left(\frac{E_n}{R_1 + R_2} \right)^2 \cdot R_2$$

om $R_1 = R_2$

$$N = \frac{4kTBR}{4R^2} R = kTB$$

Bilden visar att även om vi kopplar ihop flera motstånd som brusar så kommer dessa motstånd att leverera bruseffekten kTB till en lastresistans, så länge som lastresistansen har samma värde som summan av brusmotstånden.

Observera att vi inte kan addera spänningarna. Vi måste beräkna effektutvecklingen från varje spänningskälla för sig.

Koppla ihop två motstånd

Vi kopplar ihop motståndet R_1 med ett annat motstånd R_2 . Då avger R_1 viss bruseffekt till R_2 .

Men självklart avger motståndet R_2 bruseffekt tillbaka till R_1 .

Det lustiga är att om motstånden har samma temperatur, $T_1 = T_2$, så avger det andra motståndet lika stor bruseffekt tillbaka till det första motståndet. Oavsett vilka resistansvärden motstånden har. Det råder jämvikt.

Om ett av motstånden är kallare?

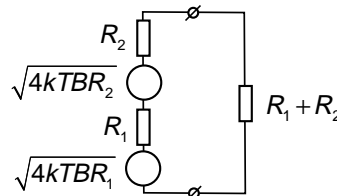
Kyl ner motståndet R_2 . Gör T_2 lägre än T_1 . Då brusar R_2 mindre. Då går mer effekt från R_1 till R_2 . Ända tills effekten värm upp det kallare motståndet till samma temperatur så att det blir jämvikt. Naturen strävar efter att utjämna olikheter, uppnå jämvikt.

Spole? Kondensator?

Men en spole? Eller kondensator? Vi får ju spänning om vi matar växelström genom en spole eller kondensator. Nej. Spolen och kondensatorn avger inget termiskt brus. Det är möjligt att spolen och kondensatorn brusar på annat sätt, se kapitel 17, men någon termisk bruseffekt kan de inte alstra.

Termiskt brus får vi bara från komponenter som själva kan omvandla ström till värme. De resistiva förlusterna i spolen och kondensatorn avger termiskt brus. Men förlustfria spolar och kondensatorer avger ingen termisk bruseffekt.

Motståndet tar alltså värme från omgivningen, eller producerar sin egen värme från ström som flyter igenom, och en del av denna värmeenergi omvandlas till växelspanning, brusspänning.

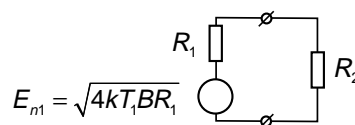


Bruseffekterna adderas
(inte spänningarna)

$$\text{Från } R_1: N_1 = \frac{4kTBR_1}{4(R_1 + R_2)^2} (R_1 + R_2)$$

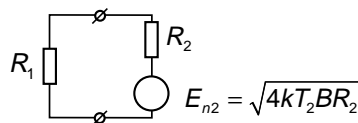
$$\text{Från } R_2: N_2 = \frac{4kTBR_2}{4(R_1 + R_2)^2} (R_1 + R_2)$$

$$N_1 + N_2 = kTB$$



$$E_{n1} = \sqrt{4kT_1BR_1}$$

$$P_2 [R_1 \Rightarrow R_2] = \left(\frac{E_{n1}}{R_1 + R_2} \right)^2 \cdot R_2 = \frac{4kT_1B(R_1 \cdot R_2)}{(R_1 + R_2)^2}$$



$$E_{n2} = \sqrt{4kT_2BR_2}$$

$$P_1 [R_2 \Rightarrow R_1] = \left(\frac{E_{n2}}{R_1 + R_2} \right)^2 \cdot R_1 = \frac{4kT_2B(R_2 \cdot R_1)}{(R_1 + R_2)^2}$$

3.4 Gaussiskt brus

Det termiska brusets spänning är summan av en stor mängd delspänningar alstrade av elektronernas värmerörelser. Dessa värmerörelser sker för oss helt oförutsägbart. Varje ledningselektron samlar på sig termisk energi tills det räcker för att hoppa. Då hoppar elektronen.

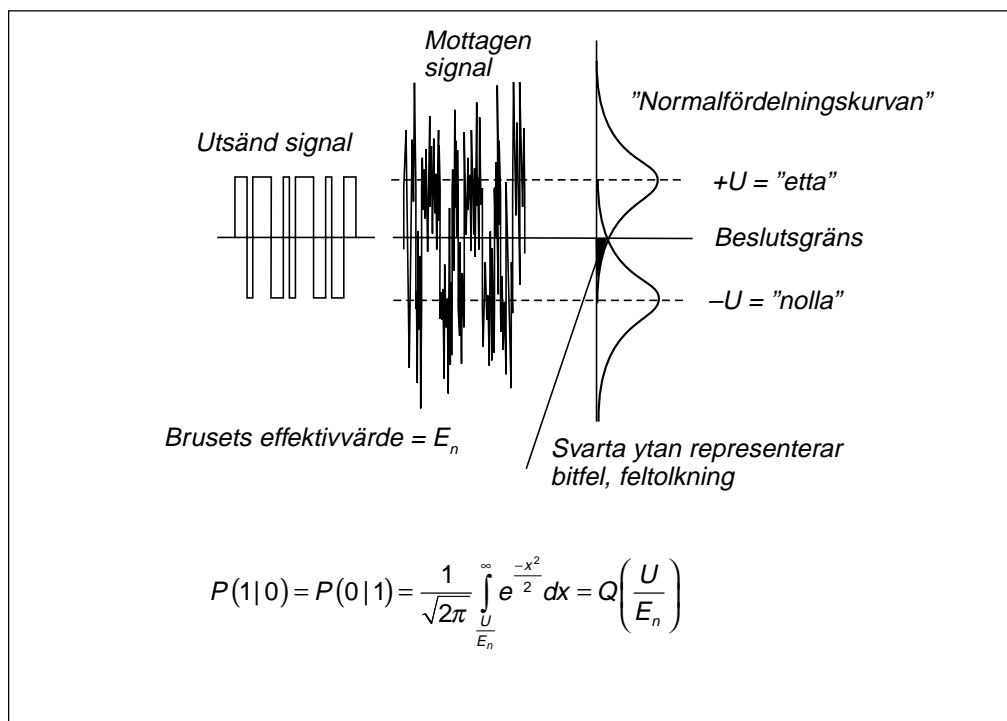
När elektronen hoppar längs ledaren så blir spänningsbidraget maximalt positivt, och när elektronen hoppar i andra riktningen så blir spänningsbidraget maximalt negativt. När elektronen hoppar "på tvären" i ledaren så får vi ett mindre spänningsbidrag. Brusspänningen är i varje ögonblick summan av alla dessa delspänningar. När vi summerar så kan vi råka ha extremt många elektroner som just då hoppar åt samma håll längs ledaren. Då blir spänningen hög. Eller också kan elektronerna hoppa kors och tvärs så att summaspänningen blir nästan noll.

Gaussiskt brus

Det enda vi kan säga om en spänning som varierar på detta sätt är att amplitudtätheten följer normalfördelningskurvan, Gauss-kurvan. Normalfördelat brus kallas ofta för Gaussiskt brus.

Bilden nedan visar utsänd signal, där ettor och nollor representeras av spänningarna $\pm U$. På överföringssträckan har brus adderats till signalen. När man skall avgöra om den mottagna symbolen motsvarar en etta eller en nolla, så gäller att positivt värde tolkas som etta, medan negativt värde tolkas som nolla.

När brusspänningen är så kraftig att amplituden överstiger beslutsgränsen (i riktning mot andra symbolen), då tolkar vi fel och får bitfel. Hur ofta detta sker representeras av den svarta ytan under Gauss-kurvan.



$\frac{U}{E_n}$	$Q\left(\frac{U}{E_n}\right)$	$\frac{U}{E_n}$	$Q\left(\frac{U}{E_n}\right)$	$\frac{U}{E_n}$	$Q\left(\frac{U}{E_n}\right)$
0,0	$5,00 \cdot 10^{-1}$	1,7	$4,46 \cdot 10^{-2}$	3,4	$3,37 \cdot 10^{-4}$
0,1	$4,60 \cdot 10^{-1}$	1,8	$3,59 \cdot 10^{-2}$	3,5	$2,33 \cdot 10^{-4}$
0,2	$4,21 \cdot 10^{-1}$	1,9	$2,87 \cdot 10^{-2}$	3,6	$1,59 \cdot 10^{-4}$
0,3	$3,82 \cdot 10^{-1}$	2,0	$2,28 \cdot 10^{-2}$	3,7	$1,08 \cdot 10^{-4}$
0,4	$3,45 \cdot 10^{-1}$	2,1	$1,79 \cdot 10^{-2}$	3,8	$7,23 \cdot 10^{-5}$
0,5	$3,09 \cdot 10^{-1}$	2,2	$1,39 \cdot 10^{-2}$	3,9	$4,81 \cdot 10^{-5}$
0,6	$2,74 \cdot 10^{-1}$	2,3	$1,07 \cdot 10^{-2}$	4,0	$3,17 \cdot 10^{-5}$
0,7	$2,42 \cdot 10^{-1}$	2,4	$8,20 \cdot 10^{-3}$	4,1	$2,07 \cdot 10^{-5}$
0,8	$2,12 \cdot 10^{-1}$	2,5	$6,21 \cdot 10^{-3}$	4,2	$1,33 \cdot 10^{-5}$
0,9	$1,84 \cdot 10^{-1}$	2,6	$4,66 \cdot 10^{-3}$	4,3	$8,54 \cdot 10^{-6}$
1,0	$1,59 \cdot 10^{-1}$	2,7	$3,47 \cdot 10^{-3}$	4,4	$5,41 \cdot 10^{-6}$
1,1	$1,36 \cdot 10^{-1}$	2,8	$2,56 \cdot 10^{-3}$	4,5	$3,40 \cdot 10^{-6}$
1,2	$1,15 \cdot 10^{-1}$	2,9	$1,87 \cdot 10^{-3}$	4,6	$2,11 \cdot 10^{-6}$
1,3	$9,68 \cdot 10^{-2}$	3,0	$1,35 \cdot 10^{-3}$	4,7	$1,30 \cdot 10^{-6}$
1,4	$8,08 \cdot 10^{-2}$	3,1	$9,68 \cdot 10^{-4}$	4,8	$7,93 \cdot 10^{-7}$
1,5	$6,68 \cdot 10^{-2}$	3,2	$6,87 \cdot 10^{-4}$	4,9	$4,79 \cdot 10^{-7}$
1,6	$5,48 \cdot 10^{-2}$	3,3	$4,83 \cdot 10^{-4}$	5,0	$2,87 \cdot 10^{-7}$

Den svarta ytan representerar bitfel när vi har ettor, men det har vi bara under halva tiden. Men vi får även bitfel när vi har nollor, och den ytan som inte är markerad i bilden, är lika stor som den svarta ytan.

Brusspänningen har effektivvärdet E_n , som svarar mot fördelningens standardavvikelse. När vi vet standardavvikelsen och amplituden så är det bara att lösa amplitud-sannolikhetsfunktionen, eller gå in i tabellverk och avläsa bitfelshalten.

Exempel:

- Om vi bara har brus, $U/E_n = 0$, så är det 50 % chans att avläsa rätt. Bitfelshalten blir 50 %.
- Om $U/E_n = 2,3$ så blir bitfelshalten ungefär 1 %.
- Om $U/E_n = 3,7$ så blir bitfelshalten ungefär 10^{-4} .
- Om $U/E_n = 5,0$ så blir bitfelshalten ungefär $2,87 \cdot 10^{-7}$.

Vi ser att även om signaleffekten är 25 ggr starkare än bruseffekten, så får vi fortfarande bitfel, även om bitfelen inte uppträder lika ofta.

$$\frac{U}{E_n} = 20 \cdot \log 2,3 = 7,2 \text{ dB}$$

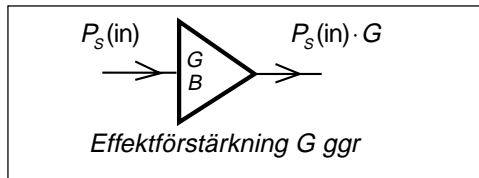
$$\frac{U}{E_n} = 20 \cdot \log 3,7 = 11,4 \text{ dB}$$

$$\frac{U}{E_n} = 20 \cdot \log 5,0 = 14 \text{ dB}$$

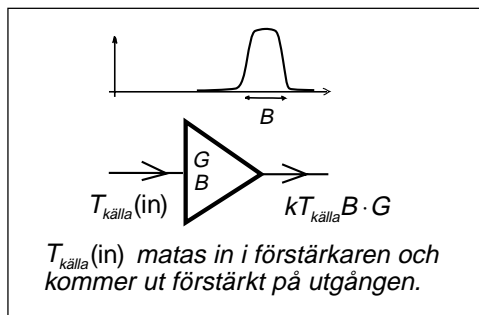
Detta är inte S/N, signal/brus-förhållandet, men någonting ditåt.

3.5 Brus i förstärkare

En förstärkare med effektförstärkningen G ggr och bandbredden B Hz matas av en signalkälla. Signalkällan levererar den önskade signalen med effekten P_S , och vi kan beräkna signaleffekten på förstärkarens utgång.



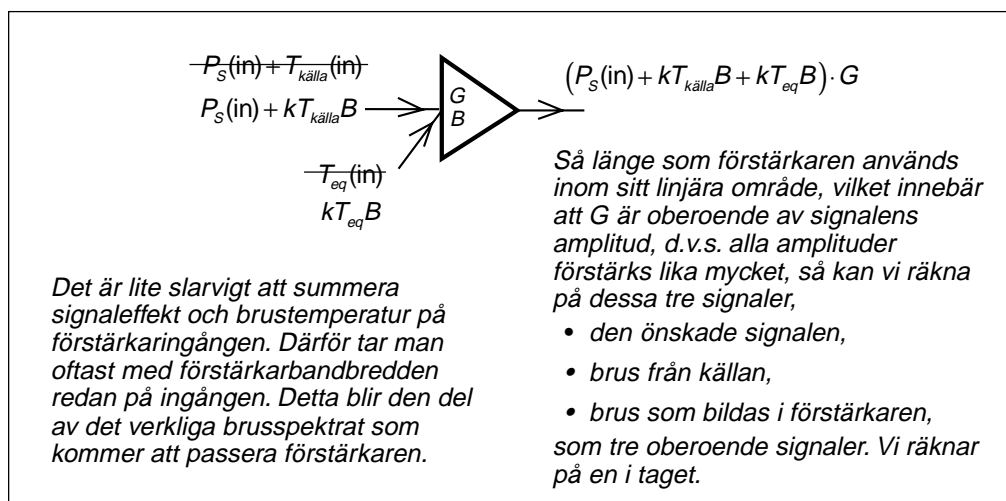
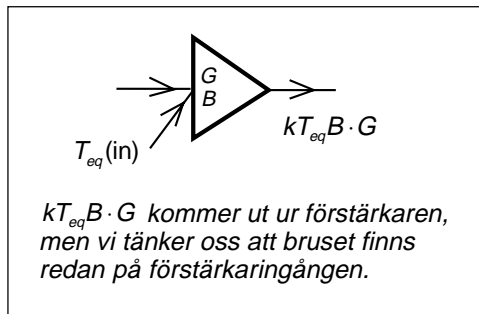
Signalkällan är inte brusfri. Den levererar även brus. Bruset har troligtvis ett brett effektspektrum. Eftersom vi inte vet något om bandbredden före förstärkaren nöjer vi oss med att ange att bruset från signalkällan har brustemperaturen $T_{källa}$. Detta brus går genom förstärkaren som har bandbredden B Hz och förstärkningen G ggr. Vi kan alltså beräkna bruseffekten på förstärkarens utgång.



Brus som bildas i förstärkaren

Men det räcker inte med detta. I förstärkaren bildas brus från motstånd och halvledare. Detta brus kommer ut på förstärkarens utgång.

För att underlätta beräkningarna så förlägger vi det brus som bildas i förstärkaren till en ekvivalent brus-källa på förstärkar-ingången med brustemperaturen T_{eq} och tänker oss att själva förstärkaren är brusfri.



Förstärkarens brustemperatur T_{eq}

Det brus som bildas i förstärkaren är en blandning av alla olika typer av brus, termiskt brus, hagelbrus o.s.v. men man räknar som om allt brus vore termiskt brus.

Ett sätt att ange förstärkarens brusegenskaper är därför att uppge förstärkarens ekvivalenta brustemperatur T_{eq} .

Att ange brustemperaturen är vanligt för mikrovågsförstärkare. Typiska värden på T_{eq} kan ligga i området 25 K – 100 K (mikrovågshuvuden för mottagning av satellit-tv) eller ända ner till 5 K för förstärkare som kyls med flytande helium. Sådana förstärkare används inom radioastronomin.

Hagelbrus (shot noise)

Det finns även andra former av brus i våra kommunikationssystem. Transistorer förstärker signalen genom att styra strömmen i en PN-övergång. Här sker laddningstransporten av "hål". Även om strömmen skall vara konstant så bildas hålen slumpvis. Därför kommer strömmen att gå "stötvis" genom transistorn. Brus adderas till den önskade signalen.

Brus från transistorer är den dominerande bruskillan i en radiomottagare. Det gäller därför att välja en transistor som brusar så lite som möjligt i tv-mottagarens ingångssteg, innan den önskade signalen förstärkts till sådan nivå att ett litet brustillskott från transistorerna inte märks.

Att välja transistor

Bruset ökar med likströmmen genom PN-övergången. Därför vill man ha så låg ström som möjligt genom transistorn i tv-mottagarens ingångssteg. Men om strömmen är låg så klarar inte transistorn att förstärka höga toppspänningar.

På mottagarens ingång finns inte något smalt filter. Det filter som väljer ut enbart den önskade signalen sitter längre in i mottagaren. Därför skall ingångstransistorn klara att förstärka flera signaler samtidigt utan att det bildas distorsion, intermodulationsprodukter (man ser en annan tv-kanal svagt i bakgrunden). För att klara detta bör likströmmen vara hög genom transistorn. I verkligheten måste man alltså kompromissa.

I brusgeneratorer som används för mätändamål så alstras bruset genom att man matar likström genom en zenerdiod i backriktningen, och därefter förstärker bruset i några transistorsteg. Detta ger en alldeles utmärkt brusgenerator.

"Man-made radio noise"

Varje gång ström startar eller stoppar finns risk att det bildas en radiostrålningspuls med frekvensspektrum som sträcker sig över ett mycket brett frekvensområde. Sådan störning bildas i naturen vid åskväder. Men de bildas även i datorer där ettor och nollor representeras av olika spänningsnivåer och därigenom även av olika strömnivåer.

Sådana störningar fångas upp av antenner och stör radiomottagaren. I vissa fall kan sådana störningar även fångas upp av kablar och störa ledningsbunden signalöverföring.

3.6 Brusförhållandet NR och brusfaktorn NF (F)

Det vanligaste sättet att ange förstärkarens brusegenskaper är inte med ekvivalenta brustemperaturen utan med brusfaktorn.

Förstärkarens brusförhållande NR

Uppmätning av förstärkarens brusegenskaper ger oss förstärkarens brusförhållande NR (Noise Ratio).

NR är förhållandet mellan den bruseffekt som kommer ut ur vår verkliga förstärkare, jämfört med den bruseffekt som skulle komma ut ur förstärkaren om den vore brusfri, när ingångsbruset har brustemperaturen $T = 290$ K.

Brustemperaturen 290 K (rumstemperatur, $+17^\circ$ C) betecknas T_0 .

Förstärkarens brusfaktor NF

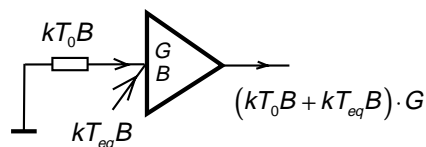
Man anger aldrig själva brusförhållandet för en förstärkare, utan brusförhållandet omvandlas till decibel.

Vad skall man kalla brusförhållandet när det omvandlats till dB? På engelska kallas detta "Noise Figure", NF . På svenska används ordet brusfaktor med beteckningen F men med sorten dB.

Om det står $F = 1,2$

Är detta NF dB eller är det brusförhållandet NR ggr? Det är F dB. En förstärkarens brusegenskaper anges alltid i dB.

I engelsk litteratur kallas brusförhållandet oftast för "noise factor". På svenska används ordet "brusfaktor" som betecknas med F .



$$F[\text{ggr}] = NR = \frac{\text{Brus ur brusig förstärkare}}{\text{Brus ur brusfri förstärkare}} = \frac{(kT_0B + kT_{\text{eq}}B) \cdot G}{(kT_0B) \cdot G} = 1 + \frac{kT_{\text{eq}}B}{kT_0B} = 1 + \frac{T_{\text{eq}}}{T_0}$$

OBS! $T_{\text{källa}} = T_0 = 290$ K

$$F[\text{dB}] = NF = 10 \cdot \log NR$$

$$NR = 10^{\frac{NF}{10}}$$

$$F[\text{dB}] = NF = 10 \cdot \log \left(1 + \frac{T_{\text{eq}}}{T_0} \right)$$

$$T_{\text{eq}} = T_0(NR - 1)$$

$$= T_0 \left(10^{\frac{NF}{10}} - 1 \right)$$

Brusegenskaperna för en transistor brukar anges i datablad (se ELFA-katalogen). Typiska värden på F är 0,8 – 1,5 dB för transistorer som skall användas i tillämpningar där lågt brus är viktigt. Motsvarande värden på brustemperatur blir 58,7 – 119,6 K.

$$T_{\text{eq}} = T_0 \cdot \left(10^{\frac{NF}{10}} - 1 \right) = 290 \cdot \left(10^{\frac{0,8}{10}} - 1 \right) = 58,66 \text{ K}$$

$$T_{\text{eq}} = T_0 \cdot \left(10^{\frac{NF}{10}} - 1 \right) = 290 \cdot \left(10^{\frac{1,5}{10}} - 1 \right) = 119,6 \text{ K}$$

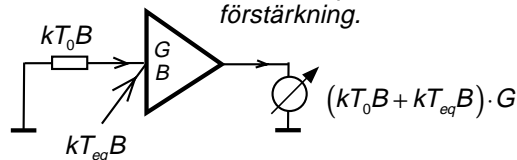
3.7 Bestämning av förstärkarens brusfaktor

En brusfaktormeter innehåller en brusgenerator som ansluts till förstärkaringången, och en "sant effektivvärdesmätande" voltmeter, som ansluts till förstärkarens utgång.

Brusgeneratorns impedans (resistans) måste vara anpassad till förstärkarens ingångsimpedans.

Vanliga voltmetrar mäter oftast växelspänningens toppvärde, medan skalan är graderad i effektivvärde för sinusformad signal (formfaktor 1,414). Vi behöver en voltmeter som klarar att mäta brusspänningens verkliga effektivvärde. Detta går med en likriktardiod som arbetar inom diodkurvas kvadratiske område.

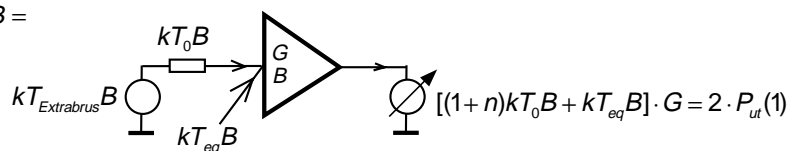
Först mäts uteffekten när brusgeneratorn är avstängd. Då producerar brusgeneratorn den bruseffekt som motsvarar $T = 290$ K.



Den bruseffekt man mäter på förstärkarens utgång blir brusets bruseffekt plus förstärkarens brus, multiplicerat med förstärkarens förstärkning.

Därefter startas brusgeneratorn och bruset ökas genom att vrida på en ratt som är kalibrerad i n , det extrabrus som alstras utöver brusets när brusgeneratorn är avstängd.

$$kT_0 B + kT_{\text{Extrabrus}} B = (1+n)kT_0 B$$



Bruset ökas tills förstärkarens uteffekt blir dubbelt så hög. Detta mäts med en sant effektivvärdesvisande voltmeter och spänningen skall öka 1,414 ggr.

Av dessa båda mätningar får vi $[(1+n)kT_0 B + kT_{\text{eq}} B] \cdot G = 2 \cdot (kT_0 B + kT_{\text{eq}} B) \cdot G$

$$(1+n)kT_0 B + kT_{\text{eq}} B = 2 \cdot kT_0 B + 2 \cdot kT_{\text{eq}} B$$

$$nkT_0 B = kT_0 B + kT_{\text{eq}} B$$

$$n = \frac{(kT_0 B + kT_{\text{eq}} B)}{kT_0 B} = 1 + \frac{T_{\text{eq}}}{T_0} = NR$$

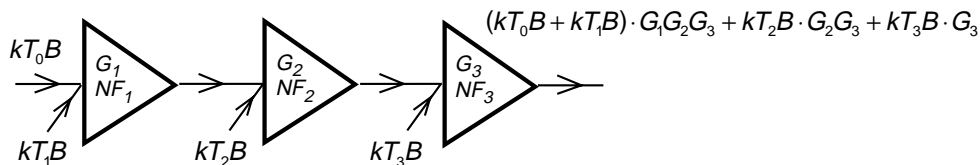
Den ratt som reglerar bruseffekten kan kalibreras i brusförhållandet NR eller ännu hellre direkt i dB, i brusfaktorn NF .

$$NF[\text{dB}] = 10 \cdot \log NR[\text{ggr}]$$

$$NR = 10^{\frac{NF}{10}}$$

3.8 Flera förstärkare

Vad händer med bruset om flera förstärkare kopplas i serie? Bruset från första steget förstärks i alla tre förstärkarna, bruset i andra steget förstärks i två förstärkare och bruset i tredje steget förstärks i ett steg.



Totala brusförhållandet

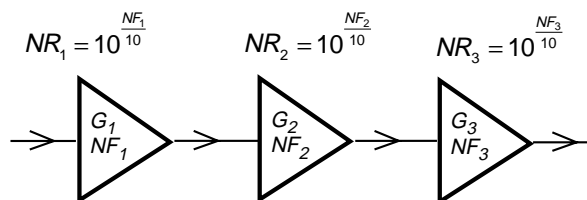
Vi kan ställa upp uttrycket för totala NR för dessa tre förstärkarsteg

$$NR_{tot} = \frac{kT_0B + kT_1B + \frac{kT_2B}{G_1} + \frac{kT_3B}{G_1 \cdot G_2}}{kT_0B}$$

$$= 1 + \frac{T_1}{T_0} + \frac{1 + \frac{T_2}{T_0} - 1}{G_1} + \frac{1 + \frac{T_3}{T_0} - 1}{G_1 \cdot G_2} =$$

$$= NR_1 + \frac{NR_2 - 1}{G_1} + \frac{NR_3 - 1}{G_1 \cdot G_2} =$$

$$= 10^{\frac{NF_1}{10}} + \frac{10^{\frac{NF_2}{10}} - 1}{G_1} + \frac{10^{\frac{NF_3}{10}} - 1}{G_1 \cdot G_2}$$



$$NR_{tot} = NR_1 + \frac{NR_2 - 1}{G_1} + \frac{NR_3 - 1}{G_1 \cdot G_2}$$

Formel för beräkning
av totala brusfaktorn

$$NF_{tot} = 10 \cdot \log \left\{ NR_1 + \frac{NR_2 - 1}{G_1} + \frac{NR_3 - 1}{G_1 \cdot G_2} \right\} =$$

$$= 10 \cdot \log \left\{ 10^{\frac{NF_1}{10}} + \frac{10^{\frac{NF_2}{10}} - 1}{G_1} + \frac{10^{\frac{NF_3}{10}} - 1}{G_1 \cdot G_2} \right\}$$

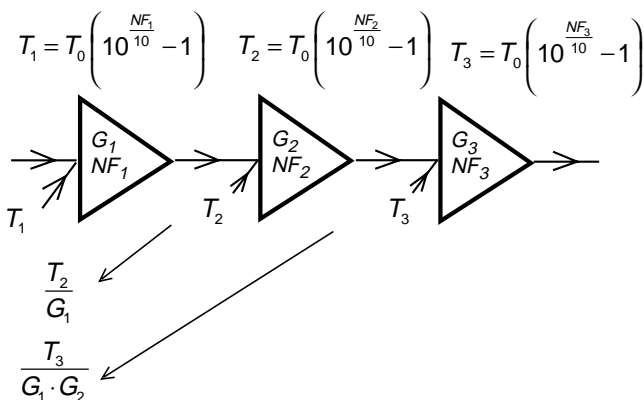
Så här enkelt blir det att beräkna förstärkarens totala brustemperatur när man känner brustemperaturen för varje ingående förstärkarsteg.

Brusbidrag från steg 1:

Brusbidrag från steg 2:

Brusbidrag från steg 3:

Totala brustemperaturen: $T_{eq} = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 \cdot G_2}$



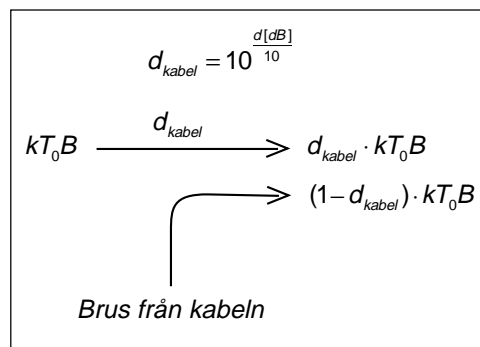
Brusbidrag från en dämpande kabel

En kabel som är förlustfri tillför inget brus till signalen. Men om kabeln dämpar så finns resistiva förluster, och dessa resistanser bidrar med termiskt brus.

1. Anslut ett motstånd på kabelns ingång, ett motstånd med temperaturen T_0 , vars värde är lika med kabelns karakteristiska impedans. Detta motstånd levererar bruseffekten kT_0B in i kabeln. Kabeln har en effektdämpning på d_{kabel} , varför den effekt som kommer ut ur kabeln är $d_{kabel} \cdot kT_0B$.

2. Men hur mycket effekt kommer ut ur kabeln? Både motståndet och kabelns dämpning är resistanser vid temperaturen T_0 . Därför måste bruseffekten ut ur kabeln vara kT_0B . Av detta kommer $d_{kabel} \cdot kT_0B$ från motståndet, medan $(1 - d_{kabel}) \cdot kT_0B$ har bildats i kabeln.

Bilden här bredvid visar hur kabeldämpningen inverkar vid radiomottagning.



Signal från antennen: P_{antenn}
Brus från antennen: T_{antenn}

Effektdämpning i kabeln: d_{kabel}

Kabelns temperatur: T_0

Signal efter kabeln: $P_{antenn} \cdot d_{kabel}$

Antennbrus efter kabeln: $T_{antenn} \cdot d_{kabel}$

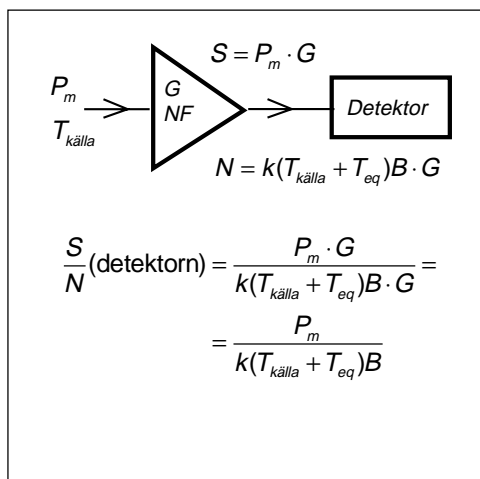
Kabelbrus efter kabeln: $(1 - d_{kabel}) \cdot T_0$

3.9 Signal/brus-förhållandet före detektorn

Mottagarkänslighet

Varför är man intresserad av att beräkna bruseffekten? Därför att nyttsignalen, den önskade signalen, måste vara ett visst antal gånger starkare än bruset för att man skall få fullgod mottagning. Det är bruset som bestämmer hur stark sändareffekt som behövs för att man skall kunna kommunicera.

Vi ser att signal/brus-förhållandet vid detektorn blir identiskt med signal/brus-förhållandet vid antenningången, så länge som vi adderar det brus som tillkommer på grund av mottagarens egen brusfaktor NF .



GSM-mottagarens detektor behöver $S/N = 9$ dB för fullgod mottagning. Radiokanalens bandbredd är 200 kHz, och mottagarens brusfaktor är i bästa fall $NF = 2$ dB. Vi antar att antennbruset motsvarar brustemperaturen 290 K.

$$P_m = 10^{\frac{9}{10}} \cdot 1,38 \cdot 10^{-23} (290 + 290[10^{\frac{2}{10}} - 1]) 200 \cdot 10^3 = 7,94 \cdot 1,38 \cdot 460 \cdot 2 \cdot 10^{-18} = 1,008 \cdot 10^{-14} \text{ W}$$

$$P_m[\text{dBW}] = 10 \cdot \log(1,008 \cdot 10^{-14}) = -140 \text{ dBW}$$

$$P_m[\text{dBm}] = 10 \cdot \log(1,008 \cdot 10^{-14} \cdot 1000) = -110 \text{ dBm}$$

Den maximala känsligheten hos en GSM-mottagare är alltså ungefär -110 dBm. Detta kan man uppnå för en basstationsmottagare.

I ficktelefonerna störs mottagningen av "man-made noise" från de digitala delarna i ficktelefonen, från mikroprocessorn m.m. Det är svårt att förhindra att sådant brus når ficktelefonens antenn, eftersom allt är så tätt ihopbyggt.

Dessutom får man viss dämpning i de filterkretsar som kopplar både mottagaren och sändaren till samma antenn.

Enligt GSM-specifikationen skall GSM-ficktelefonen ha en känslighet av -104 dBm på 900 MHz och -102 dBm på 1800 MHz.

Förenklat räknesätt är farligt!!!

Titta på uttrycket för S/N vid detektorn. Detta uttryck går att omvandla, så att S/N vid detektorn är lika med S/N in i mottagaren, d.v.s. det signal/brus-förhållande vi får från antennen, försämrat med mottagarens brus-förhållande. Men observera att detta bara gäller om $T_{källa} = T_0$.

Man kan inte räkna på detta sätt om källans brustemperatur avviker från 290 K, och det gör den oftast.

Räkna på signalen för sig ...

Vid beräkning av signal/brusförhållandet skall du alltid räkna på den önskade signalen P_m för sig. Vad händer med den önskade signalens effekt? Förstärkning? Dämpning?

... och brustemperaturerna för sig.

Vid beräkning av bruset, beräkna alltid brustemperaturtillskottet från systemets olika delar. Använd brusfaktorn för att få fram motsvarande brustemperatur. Därefter kan du kompensera för dämpningar och förstärkningar genom att flytta brustemperaturen till mottagaringången. Detta är en säker metod.

$$\begin{aligned} \frac{S}{N}(\text{detektorn}) &= \frac{P_m}{k(T_{källa} + T_{eq})B} = \\ &= \frac{P_m}{kT_{källa}B} \cdot \frac{1}{\left(1 + \frac{T_{eq}}{T_{källa}}\right)} = \frac{S}{N}(\text{in}) \cdot \frac{1}{\left(1 + \frac{T_{eq}}{T_{källa}}\right)} \end{aligned}$$

Om $T_{källa} = T_0$ så kan man räkna så här ...

$$\frac{S}{N}(\text{det.}) = \frac{S}{N}(\text{in}) \cdot \frac{1}{\left(1 + \frac{T_{eq}}{T_0}\right)} = \frac{S}{N}(\text{in}) \cdot \frac{1}{NR}$$

... och få följande enkla uttryck:

$$\frac{S}{N}(\text{det.})[\text{dB}] = \frac{S}{N}(\text{in})[\text{dB}] - NF[\text{dB}]$$

OBS! Gäller bara om $T_{källa} = T_0$ OBS!!! ...

Enda undantaget:

Det finns inga undantag. Även formeln för beräkning av totala brusfaktorn när man har tre förstärkarsteg går via brustemperatur, även om det inte syns.

Detta är brustemperaturerna återförda till ingångssteget:

$$\begin{aligned} NF_{tot} &= 10 \cdot \log \left\{ 10^{\frac{NF_1}{10}} + \frac{10^{\frac{NF_2}{10}} - 1}{G_1} + \frac{10^{\frac{NF_3}{10}} - 1}{G_1 \cdot G_2} \right\} = \\ &= 10 \cdot \log \left\{ NR_1 + \frac{NR_2 - 1}{G_1} + \frac{NR_3 - 1}{G_1 \cdot G_2} \right\} = \\ &= 10 \cdot \log \left\{ 1 + \frac{T_1}{T_0} + \frac{\frac{T_2}{G_1}}{T_0} + \frac{\frac{T_3}{G_1 \cdot G_2}}{T_0} \right\} = \\ &= 10 \cdot \log \left\{ 1 + \frac{T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 \cdot G_2}}{T_0} \right\} \end{aligned}$$

Ytterligare ett exempel:

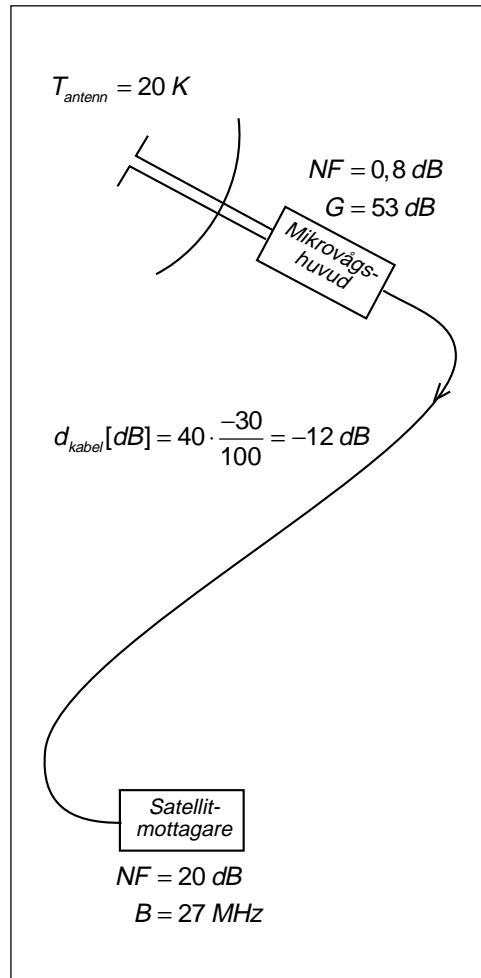
Vid mottagning av satellit-tv har man en parabolantenn på taket. Parabolantennen fångar upp radiosignalen från satelliten, men även brus från stjärnorna, och lite brus från jorden (sidolober). Antag att antennbruset uppgår till 20 K.

I mikrovågshuvudet sitter en förstärkare och frekvensomvandlare. Mikrovågshuvudet kan t.ex. ha brusfaktorn $NF = 0,8$ dB och förstärkningen $G = 53$ dB.

Från mikrovågshuvudet går en kabel ner till själva satellitmottagaren. Denna kabel kan vara 40 m lång och ha en dämpning av 30 dB per 100 m vid 2 GHz, den högsta frekvens i det frekvensband som används när tv-signalen transporteras till satellitmottagaren.

Satellitmottagaren har bandbredden 27 MHz och behöver 7 dB signal/brusförhållande för brusfri mottagning av den analoga tv-signalen (FM-modulerad vid sändning över satellit). Vi kan anta att satellitmottagaren har en brusfaktor på $NF = 20$ dB (den behöver inte vara bättre).

Beräkna mottagarkänsligheten, d.v.s. den lägsta effektnivån från satelliten som behöver fångas upp av parabolantennen för att man skall få fullgod tv-bild, d.v.s. en bild där brusprickarna uppträder så sällan att de inte uppfattas av ögat.



Satellitmottagaren behöver $S/N = 7$ dB och bandbredden är 27 MHz. Då kan vi direkt skriva ett uttryck för mottagarkänsligheten:

$$P_m = \frac{S}{N} \cdot k(T_{källa} + T_{eq})B = 10^{\frac{7}{10}} \cdot kT_{tot}B =$$

$$= 10^{\frac{7}{10}} \cdot 1,38 \cdot 10^{-23} \cdot T_{tot} \cdot 27 \cdot 10^6$$

Vad som saknas är totala brustemperaturen som får bidrag från antennen, mikrovågshuvudet, kabeln och satellitmottagaren, allt räknat vid mikrovågshuvudets antenningång.

$$T_{tot} = T_{antenn} + T_1 + T_{kabel} + T_{satellit}$$

Brusbidrag från antennen:

$$T_{\text{antenn}} = 20 \text{ K}$$

Brusbidrag från mikrovågshuvudet:

$$T_1 = 290(10^{\frac{NF}{10}} - 1) = 290(10^{\frac{0,8}{10}} - 1) = 58,6 \text{ K}$$

Kabeldämpningen är kabellängden multiplicerat med dämpning per meter i dB.

Nu kan vi räkna fram kabelbruset på utgången av kabeln, d.v.s. nere vid satellitmottagaren.

Kabeldämpning:

$$d_{\text{kabel}}[\text{dB}] = 40 \cdot \frac{-30}{100} = -12 \text{ dB}$$

Kabelbrus:

$$T_{\text{kabel}}(\text{ut}) = 290(1 - d_{\text{kabel}}) = 290(1 - 10^{\frac{-12}{10}}) = 290 \cdot 0,94 = 272 \text{ K}$$

Brusbidrag från kabeln återförd till antenningången:

$$T_{\text{kabel}} = \frac{T_{\text{kabel}}(\text{ut})}{G \cdot d_{\text{kabel}}} = \frac{T_{\text{kabel}}(\text{ut})}{10^{\frac{53}{10}} \cdot 10^{\frac{-12}{10}}} = \frac{272}{200000 \cdot 0,063} = 0,02 \text{ K}$$

Brusbidrag från satellitmottagaren återförd till antenningången:

$$T_{\text{satellit}} = \frac{290(10^{\frac{NF}{10}} - 1)}{G \cdot d_{\text{kabel}}} = \frac{290(10^{\frac{20}{10}} - 1)}{10^{\frac{53}{10}} \cdot 10^{\frac{-12}{10}}} = \frac{28710}{200000 \cdot 0,063} = 2,3 \text{ K}$$

Alltså får vi totala brustemperaturen beräknad vid mikrovågshuvudets antenningång ...

$$T_{\text{tot}} = T_{\text{antenn}} + T_1 + T_{\text{kabel}} + T_{\text{satellit}} = 20 + 58,6 + 0,02 + 2,3 = 80,92 \text{ K}$$

... och kan beräkna minsta effektnivån för fullgod tv-bild:

$$P_m = 10^{\frac{7}{10}} \cdot 1,38 \cdot 10^{-23} \cdot T_{\text{tot}} \cdot 27 \cdot 10^6 = 10^{\frac{7}{10}} \cdot 1,38 \cdot 10^{-23} \cdot 80,92 \cdot 27 \cdot 10^6 = 1,5 \cdot 10^{-13} \text{ W}$$

Vi omvandlar till dBm, det som vanligen används i datablad:

$$P_m[\text{dBm}] = 10 \cdot \log(1,5 \cdot 10^{-13} \cdot 1000) = -96 \text{ dBm}$$

3.10 Brusfri resistans

Alla fysiska motstånd, motståndet som komponent, avger termiskt brus. Dessutom finns motstånd som brusar mer än enbart det termiska bruset. Ett motstånd som består av hoppresat kolpulver kan, beroende på mikroskopiska oxidskikt mellan kolpartiklarna, göra att det behövs viss spänning för att bryta igenom oxidskikten. Strömmen går ryckigt. Det brusar. Därför används speciella lågbrusmotstånd i konstruktioner där man eftersträvar lågt brus. Men det termiska bruset slipper man inte.

Finns det resistanser som är brusfria? Nej, men det finns "resistans" som är brusfri.

När en antenn används som sändarantenn så matar sändaren ut spänning på antennkabeln, och vi får ström. Utgående från spänningen, strömmen och fasvinkeln så kan vi beräkna antennens impedans.

Den resistiva delen av antennimpedansen "absorberar" effekt, men inte så som ett motstånd skulle gjort, nämligen genom att omvandla effekten till värme, utan genom att omvandla effekten till radiovågor.

När samma antenn används som mottagarantenn så kommer den signal som fångas upp av antennen att uppträda som en spänningsgenerator i serie med antennimpedansen. Detta motstånd i antennimpedansen är brusfritt.

Men vad för slags brus kommer från en antenn?

Det brus som kommer från en mottagarantenn är brus som fångas in av antennen, brus som finns i form av radiostrålning och som tas emot av mottagarantennen.

Antennens strålningsdiagram

Varje antenn har ett strålningsdiagram som visar hur stor del av sändareffekten som strålar ut i olika riktningar. När antennen används som mottagarantenn så fångar den in signaler i motsvarande riktningar, i proportion till strålningsdiagrammet.

Vad finns för brusällor?

Varje "kropp" sänder ut brusstrålning som motsvarar kroppens temperatur. Jorden sänder ut brus med brustemperaturen 290 K.

En parabolantenn på en satellit som tar emot signal från jorden kommer samtidigt att ta emot brusstrålning från jorden och lämna ett antennbrus som i stort sett är 290 K.

Parabolantennerna på jorden som riktas mot satelliten fångar upp brus från stjärnor som befinner sig inom antennloben, brus från Vintergatan. Men ju lägre antennens elevationsvinkel är, desto större del av antennloben kommer att träffa jorden.

En parabolantenn som riktas rakt upp, elevation 90° , fångar upp bruseffekt motsvarande 3 K vid 12 GHz. Vid 10° elevation har bruset stigit till 20 K.

Men vid 100 MHz? Eller 10 MHz? Vad har vi för antennbrus?

Vid 10 MHz dominerar atmosfärbrus och vi kan räkna med $3 \cdot 10^5$ K på vintern. På sommarhalvåret stiger detta till $3 \cdot 10^6$ K.

Vid 100 MHz finns ett galaktiskt brus på 1500 K, men i tätbebyggt område finns man made noise och där är $3 \cdot 10^5$ K ett rimligare värde.

Strålningsresistansen brusar alltså inte, men ofta är antennbruset ändå betydligt högre än 290 K. Utom vid satellitmottagning, då antennbruset är avsevärt lägre än 290 K.

4 — Amplitudmodulering – AM

Att modulera betyder att anpassa något efter någonting.

Vad vi gör är att anpassa en signal (bärvåg) efter en annan signal (informationssignalen).

Den "anpassade" bärvågen (den modulerade signalen) överförs, varefter informationssignalen återskapas (detekteras, demoduleras).

I detta kapitel "anpassar" vi bärvågens amplitud efter informationssignalen.

Först studerar vi hur man bär sig åt.

Därefter tittar vi på den "anpassade" signalens tidsfunktion.

Sedan kommer vi till den "anpassade" signalens spektrum för att vi skall kunna avgöra hur stor plats som signalen tar på det som vi alla skall dela på, nämligen frekvensaxeln.

Därefter står demodulering på tur. Vi måste ju kunna återskapa något som åtminstone uppfattas som informationssignalen.

Som avslutning tittar vi på den modulerade signalens effektspektrum och konstaterar att merparten av sändareffekten går åt till att sända bärvågen, som inte innehåller någon information om informationssignalen.

4.1 Basband och linjekodning

Vid elektrisk informationsöverföring skall information, en informationssignal $m(t)$ överföras.

Informationssignalen

Informationen som vi vill överföra finns i form av en spänning. Det kan vara växelspänningen från en mikrofon, tal eller musik, eller växelspänningen från en tv-kamera, eller digital information i form av en spänning som växlar mellan två spänningsnivåer där ena nivån betyder etta och den andra nivån betyder nolla.

Dessa spänningar ändrar sig i tiden, de är funktioner i tiden, $m(t)$, men de upptar även ett visst frekvensområde, de har sina frekvensspektrum, $M(f)$.

Basband

Spektrumet för tal eller musik ligger inom 30 – 20 000 Hz, tv-signalen ligger inom 0 – 5,5 MHz, och spektrumet för växelspänningen som representerar den digitala signalen beror bl.a. på datahastigheten. Frekvensområdena som dessa spektrum upptar kallas basband.

Ibland överförs informationssignalen direkt eller linjekodad

Så sker i det analoga telefontätet, där mikrofonspänningen skickas på kopparledningen till telefonstationen. Ända in på 1950-talet telefonerade man över större delen av Europa genom att skicka informationssignaler, i stort sett ett samtal per trådpar.

Tv-signalen går som informationssignal från videon till tv:n om vi ansluter videon via SCART-kontakten.

Datasignaler i form av informationssignaler har vi på kabeln mellan datorn och modemmet.

ISDN-signalen på kopparkabeln mellan hemmet och telefonstationen är linjekodad. Spänningsnivåerna som betyder de olika databitarna har valts så att man skall klara så långa kablar som möjligt.

4.2 Modulering

Att modulera en bärvåg

En annan metod är att låta informations-signalen påverka en sinusspänning. Man låter informationssignalen modulera sinusspänningen, påverka sinusspänningen på sådant sätt att det i andra ändan är möjligt att återskapa informationssignalen. Sinusspänningen "bär" informationssignalen. Sinussignalen är bärvåg.

- En fördel med denna metod är att bärvågen kan läggas i ett helt annat frekvensområde. Vi kan "flytta" informations-signalens frekvensspektrum så att det hamnar runt bärvågen.
- Så här gör vi vid radiokommunikation. Vi måste skapa en signal vars frekvensspektrum ligger inom den "frekvenskanal" som avdelats för vår radiotrafik. Och det görs genom att modulera en bärvåg.
- Att flytta informationssignalen till ett helt annat frekvensområde genom att modulera en bärvåg började användas på 1950-talet för att överföra flera telefonsamtal på samma kopparledning, bärfrekvenssystem. Metoden utnyttjades ända in på 1990-talet, med upp till 10 800 samtidiga telefonsamtal på en ledning (koaxialkabel). När man övergick till digital teknik klarade man att överföra 565 Mbit/s på samma ledning, vilket innebar bara 7680 samtidiga telefonsamtal.
- En digital informationssignal har ett frekvensspektrum som sträcker sig ända från likström, 0 Hz. Men när telefonnätet använder bärfrekvensteknik begränsas frekvensområdet till 300 – 3400 Hz. Samma begränsning har införts vid digitaliseringen av telefonsamtal. Det går därför inte att överföra datasignaler med linjekodning i telefonnätet. Därför låter vi datorns informationssignal modulera en bärvåg som ligger ungefär mitt i det frekvensområde som överförs av telefonnätet, vid ungefär 1800 Hz. Denna utrustning, som på sändarsidan modulerar bärvågen och på mottagarsidan demodulerar signalen så att informations-signalen återskapas, kallas modem.

Eftersom den modulerade signalens spektrum ligger i samma frekvensområde som informationssignalen så kallas denna typ av modem för basbands-modem.

- När videon inte ansluts till tv:n via SCART-kontakten utan via antennkabeln, då alstras i videon en bärvåg i UHF-bandet, vid ca 600 MHz, som moduleras med tv-informationssignalen.

Tre olika sätt att modulera

Bärvågen kan moduleras på i princip tre olika sätt:

Detta är informationssignalen: $m(t)$

Detta är bärvågen: $E_c \cos 2\pi f_c t$

Man kan låta bärvågens amplitud ändras proportionellt mot informationssignalen, amplitudmodulering.

$$[E_c + k_a m(t)] \cos 2\pi f_c t$$

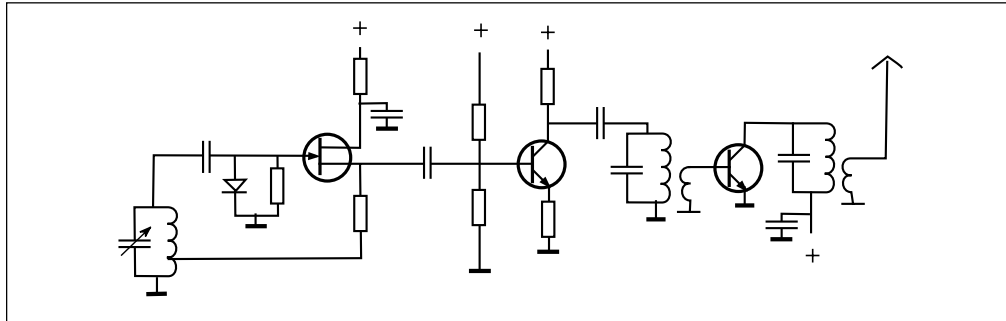
Man kan låta bärvågens frekvens ändras proportionellt mot informationssignalen, frekvensmodulering.

$$E_c \cos 2\pi [f_c + k_f m(t)] t$$

Man kan låta bärvågens fas ändras proportionellt mot informationssignalen, fasmodulering.

$$E_c \cos [2\pi f_c t + k_\phi m(t)]$$

I detta kapitel studerar vi amplitudmodulering, i kapitel 5 skall vi titta på några varianter av amplitudmodulering, i kapitel 7 kommer vi in på frekvens- och fasmodulering, och i kapitel 10 är vi framme vid överföring av digitala signaler, då man oftast utnyttjar moduleringsmetoder som kan ses som en kombination av alla tre moduleringsätten.



Vi skall se hur de olika moduleringsätten kan utföras i verkligheten. Först en omodulerad radiosändare. Radiosändaren består av tre steg.

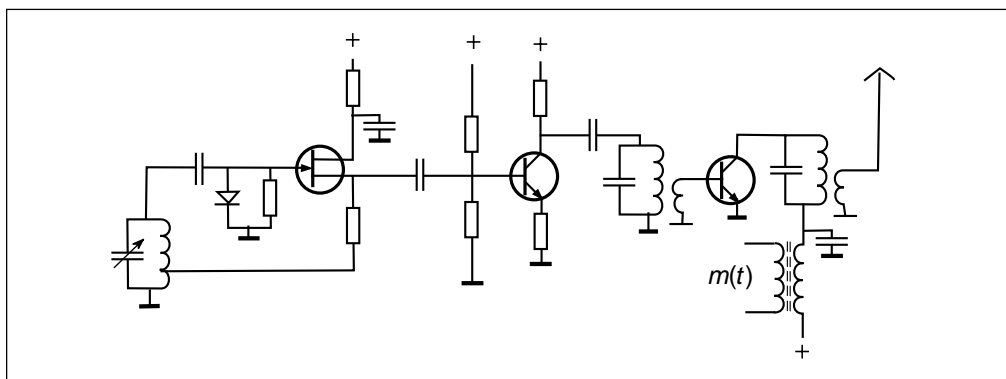
- Första steget är en oscillator där bärvågen alstras. Bär vågsfrekvensen bestäms av L och C i resonanskretsen.
- Det andra steget är en klass A-förstärkare, med bandpassfilter (resonanskrets) på utgången. Resonanskretsen är avstämmd till bärvågsfrekvensen.
- Det tredje steget är en förstärkare som arbetar i klass C. Transistorn ligger strypt utom under en del av drivsignalens positiva halvperiod, då transistorn leder för fullt. Klass C-steget fungerar nästan som en strömbrytare, switch. Kollektorspänningen kortsluts till jord i takt med bärvågsfrekvensen. Vi får fyrkantström genom resonanskretsen. Men resonanskretsen är avstämmd till bärvågsfrekvensen. Den släpper bara igenom fyrkantströmmens grundton, bärvågsfrekvensen, och spärrar övertoner.

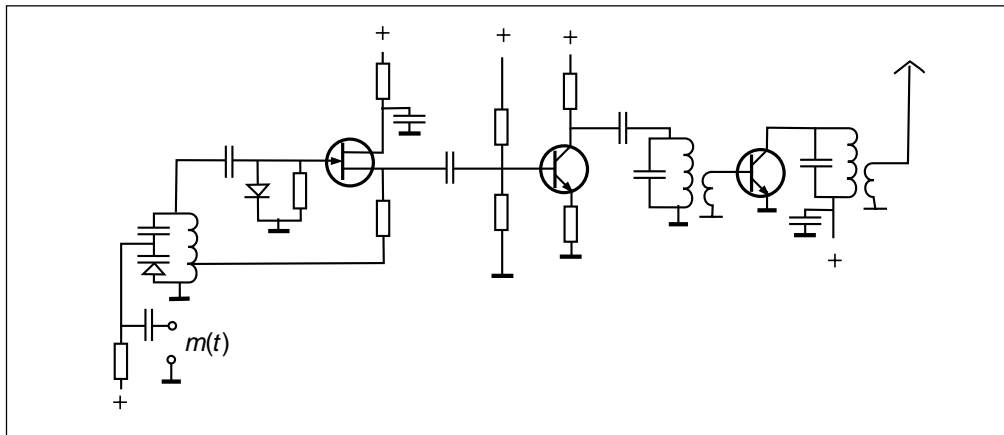
Resonanskretsen på utgången fungerar som bandpassfilter som släpper igenom frekvenser runt bärvågsfrekvensen, som matas till antennen.

Fördelen med klass C är hög verkningsgrad. Uteffekten är hög i förhållande till inmatad likströmseffekt (60 – 80% verkningsgrad). Nackdelen är att utsignalens amplitud i stort sett är konstant även om drivspänningen från oscillatorn skulle variera i amplitud. Klass C-förstärkaren fungerar därför som en begränsare, "limiter".

Amplitudmodulering

Vid amplitudmodulering skall utsignalens amplitud variera i takt med informations-signalen $m(t)$. Denna amplitudändring kan inte göras på oscillatorn eftersom klass C-steget lämnar en utsignal som bara beror på likspänningen som matas till kollektorn. Därför amplitudmodulerar vi genom att ändra denna likspänning i takt med $m(t)$. Detta åstadkommes genom att lägga en transformator i serie med likspänningen till klass C-steget.





Frekvensmodulering

Oscillatorns frekvens bestäms av L och C i resonanskretsen. Om vi ersätter kondensatorn C med en kapacitansdiod, en backspänd diod som fungerar som kondensator vars kapacitans beror på backspänningens storlek, så kan vi ändra oscillatorfrekvens genom att ändra backspänningen. Låt backspänningen bestå av en likspänning med överlagrad informationsspänning $m(t)$. Då kommer oscillatorfrekvensen att ändras i takt med $m(t)$. Detta är frekvensmodulering.

Fasmodulering

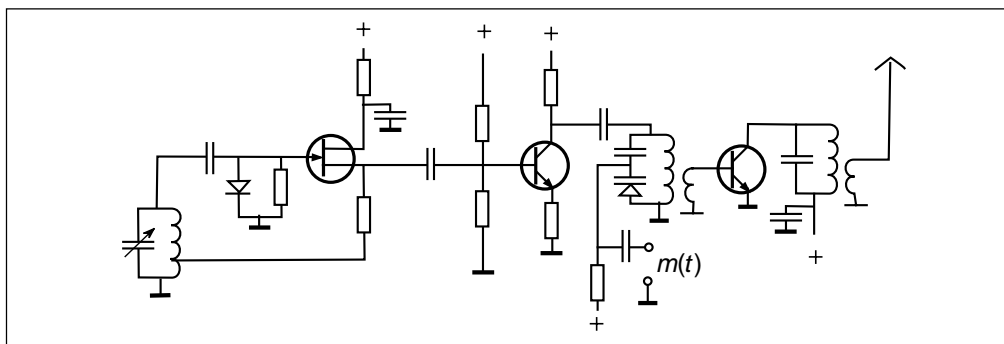
Återgå till den ursprungliga sändaren med konstant oscillatorfrekvens f_c . Mellan oscillatorsteget och effektsteget har vi en resonanskrets.

- Om resonanskretsen är avstämd till oscillatorfrekvensen f_c så passerar oscillatorspänningen opåverkad till effektsteget.

- Om resonanskretsens resonansfrekvens sänks så kommer oscillatorspänningen att dämpas. Men det gör ingenting eftersom effektsteget klipper. Utsignalens amplitud förblir oförändrad. Men det filter som bildas av resonanskretsen har inte bara en amplitud/frekvenskurva, det har även en fas/frekvenskurva. Även signalens fas ändras när resonanskretsens resonansfrekvens ändras.

Genom att utforma denna resonanskrets på samma sätt som oscillatorns resonanskrets vid frekvensmodulering och styra resonansfrekvensen med $m(t)$, så kommer utsignalens fas att variera i takt med $m(t)$. Detta är fasmodulering.

I verkligheten kan alla dessa modulationsätt åstadkommas på ett flertal olika sätt. Jag har bara visat ett exempel för att du skall förstå vad det är frågan om.

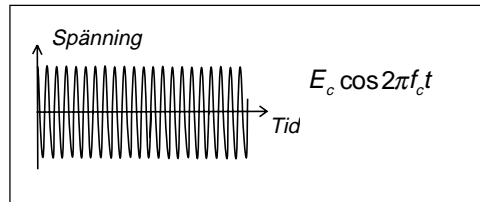


4.3 Amplitudmodulering i tidsplanet

Amplitudmodulering är som att koppla en dämpsats mellan bärvågsgeneratoren och sändarens utgång. Observera att sändaren kan mata en antenn om det är radio-kommunikation, eller en ledning om det är ledningsbunden kommunikation.

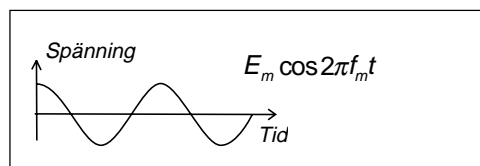
Bärvåg

Utgå från bärvågen, en sinusformad signal med amplituden E_c och frekvensen f_c , där "c" betecknar "carrier", bärvåg. Vi väljer att beskriva bärvågen med cosinusfunktionen eftersom beräkningarna blir något enklare än om man utgår från sinusfunktionen.



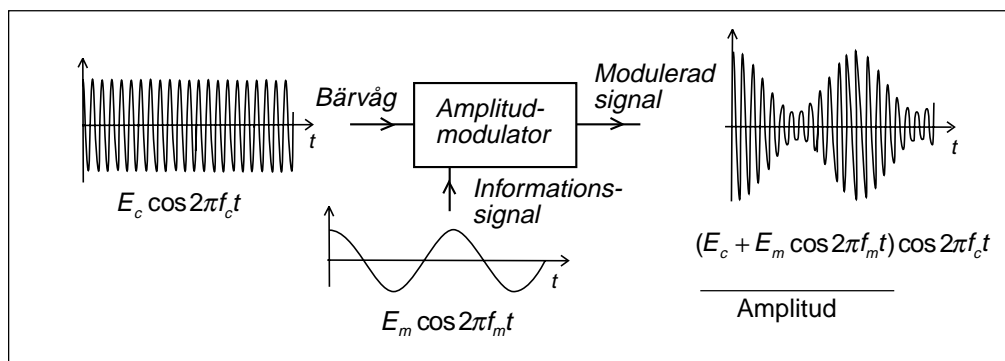
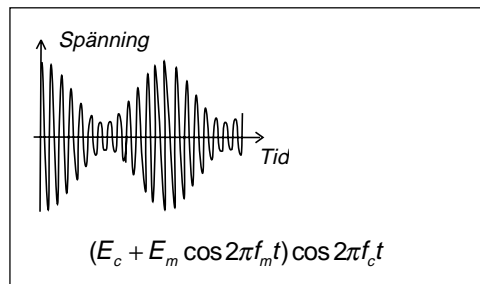
Informationssignal

Låt informationssignalen $m(t)$ vara en sinus-signal med amplituden E_m och frekvensen f_m . Även här väljer vi cosinusfunktionen. Prova själv med sinusfunktioner.



Modulerad signal

Hos den modulerade signalen kommer amplituden att ändras från $E_c + E_m$ till $E_c - E_m$ i takt med informationssignalen.



4 — Amplitudmodulering – AM

Övermodulering vid AM

Om E_m är större än E_c kommer $E_c - E_m$ att bli negativ, d.v.s. transistoren som moduleras får negativ kollektorspänning och stryps. Vi får ingen utsignal.

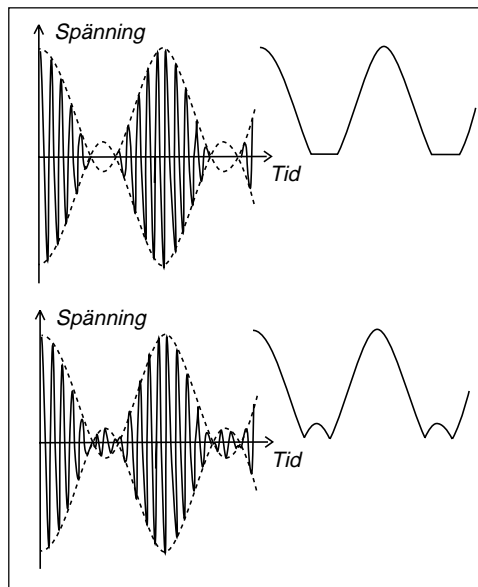
Andra typer av amplitudmodulatorer ger signal som "korsar tidsaxeln".

Oavsett vad som händer så blir den modulerade signalen distorderad. Amplituden följer inte informationssignalen. Det kallas övermodulering när E_m blir större än E_c vid amplitudmodulering. Övermodulering vid AM ger distorsion och måste undvikas.

Moduleringsgrad m_a

Vi kan omforma det matematiska uttrycket och införa moduleringsgraden m_a (modulationsindex).

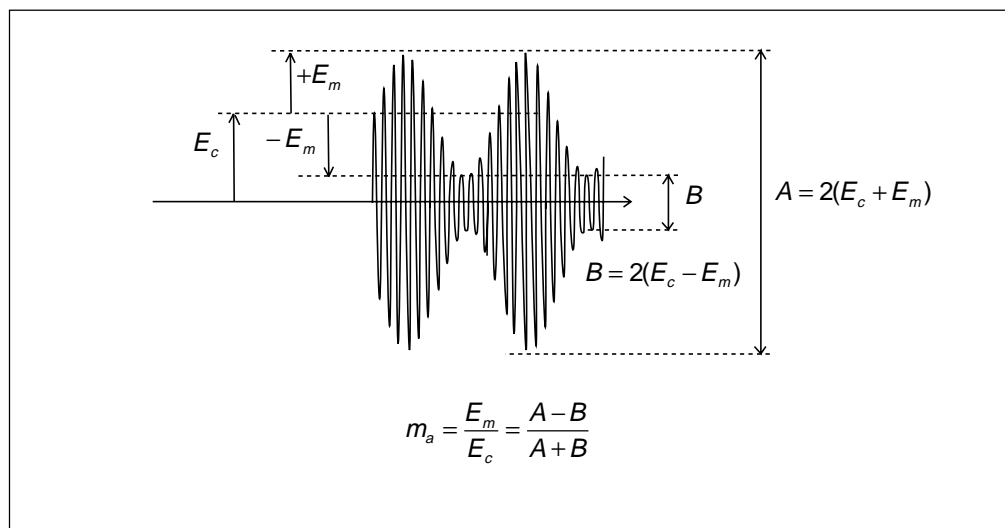
Vi får övermodulering om moduleringsgraden är större än 1.



$$E_c \left(1 + \frac{E_m}{E_c} \cos 2\pi f_m t \right) \cos 2\pi f_c t =$$

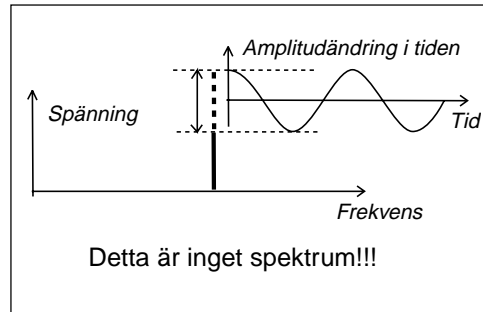
$$E_c (1 + m_a \cos 2\pi f_m t) \cos 2\pi f_c t$$

$$m_a = \frac{E_m}{E_c}$$



4.4 Amplitudmodulerade signalens spektrum

Den amplitudmodulerade signalens bärvåg ändrar sin amplitud sinusformat. Visst kan vi rita en frekvensaxel och visa att bärvågsamplituden ändras som funktion av tiden. Men detta är inget spektrum. Ett spektrum består av frekvenskomponenter som "aldrig" ändrar sin amplitud.



Hur beräknar vi AM-signalens spektrum? Genom att fourierserierutveckla det matematiska uttrycket för AM som funktion av tiden.

I detta fall behöver vi inte fourierserierutveckla. Det räcker att multiplicera ihop så får vi tre cosinustermer med amplituder som inte varierar i tiden. Exakt samma resultat får vi med fourierserierutveckling.

$$\begin{aligned}
 (E_c + E_m \cos 2\pi f_m t) \cos 2\pi f_c t = \\
 E_c \cos 2\pi f_c t + \frac{E_m}{2} \cos 2\pi(f_c + f_m)t + \frac{E_m}{2} \cos 2\pi(f_c - f_m)t = \\
 \frac{E_m}{2} \cos 2\pi(f_c - f_m)t + E_c \cos 2\pi f_c t + \frac{E_m}{2} \cos 2\pi(f_c + f_m)t
 \end{aligned}$$

undre sidband
bärvåg
övre sidband

Bärvåg, carrier

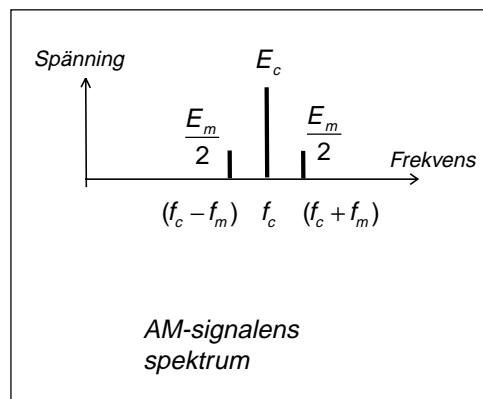
Vi får en frekvenskomponent med frekvensen f_c och amplituden E_c som är vår omodulerade bärvåg.

Undre sidbandet, LSB (lower sideband)

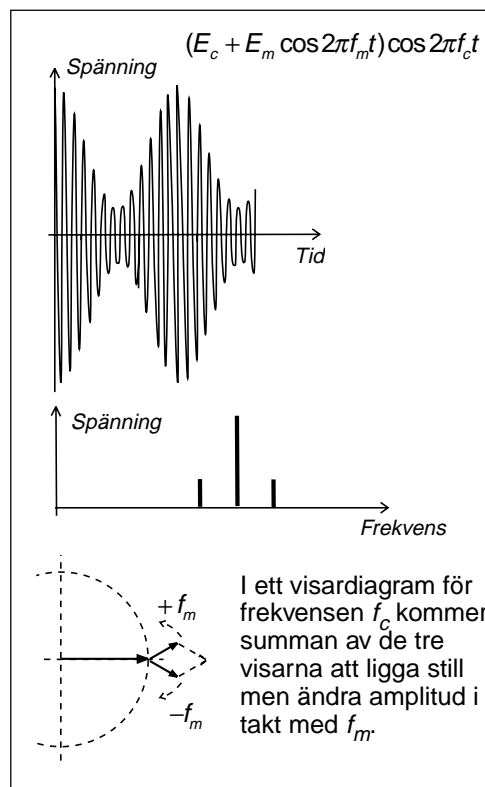
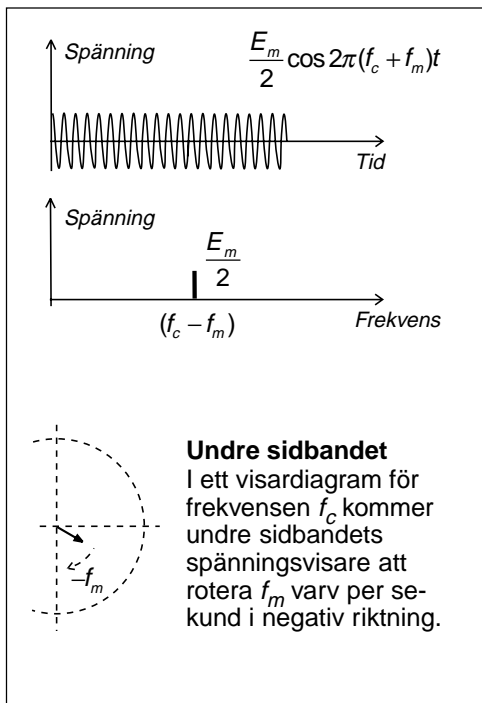
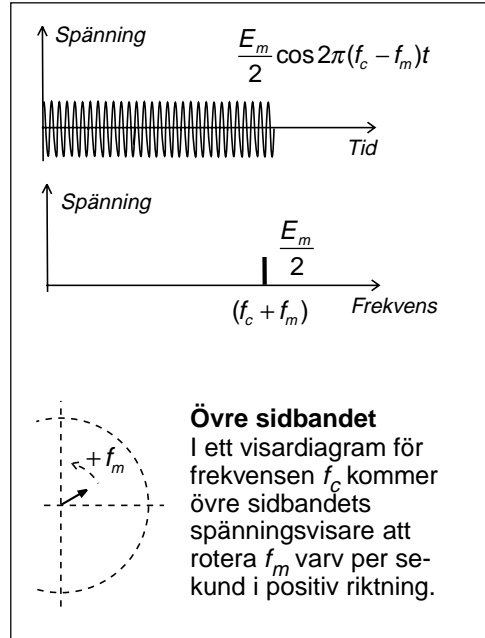
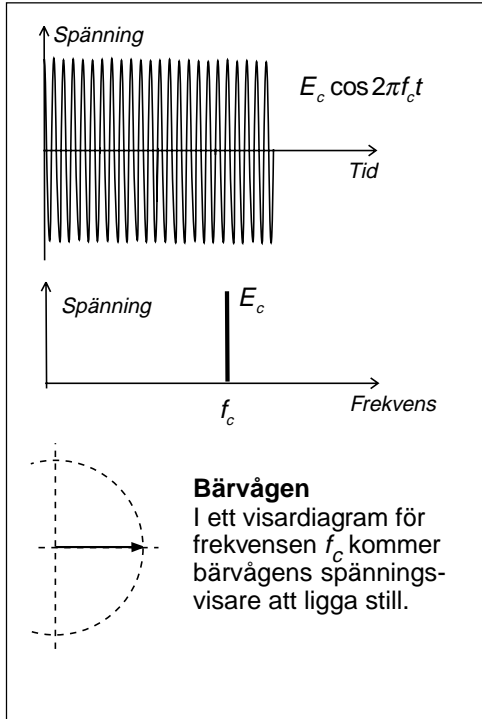
Vi får en frekvenskomponent med frekvensen $f_c - f_m$ och amplituden lika med halva informationsspänningen E_m .

Övre sidbandet, USB (upper sideband)

Vi får ytterligare en frekvenskomponent med frekvensen $f_c + f_m$ och amplituden lika med halva informationsspänningen E_m .



4 — Amplitudmodulering – AM



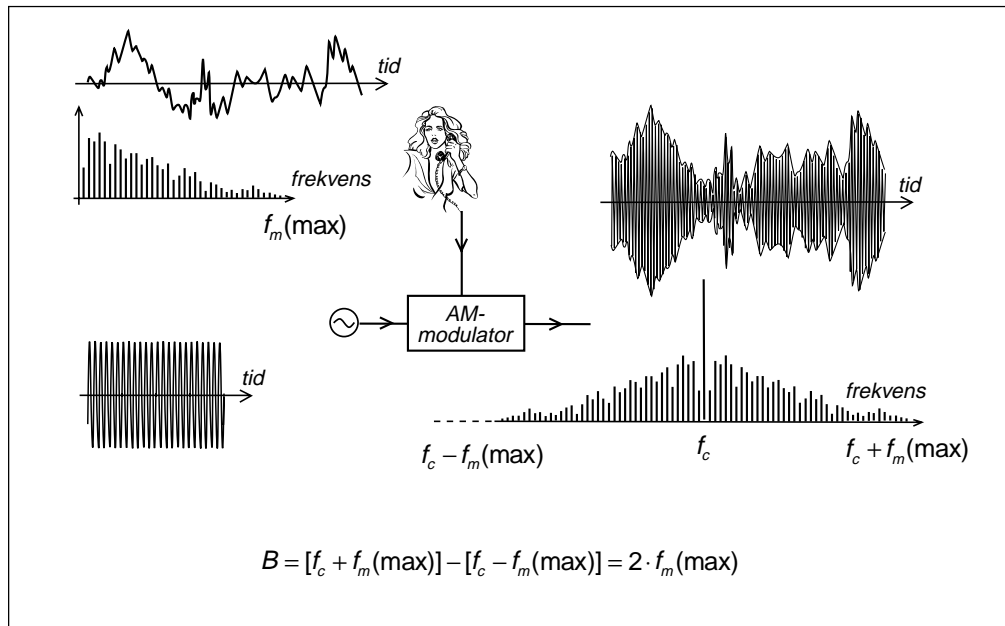
4.5 Verkliga informationssignaler

Varje frekvenskomponent ger summa- och skillnadsfrekvenser

Om $m(t)$ är en komplicerad tidsfunktion så fourieruppdelar vi $m(t)$, och utnyttjar ovanstående räknesätt för var och en av spektrumets frekvenskomponenter. Resultatet blir bärvågen, och frekvensspektrumet av $m(t)$ som övre sidband, och det spegelvända frekvensspektrumet av $m(t)$ som undre sidband.

AM-signalens bandbredd (B)

För att beräkna hur stor bandbredd som radiosignalens frekvensspektrum upptar, så behöver vi inte känna $m(t)$ i detalj. Det räcker att veta den högsta frekvensen i informationssignalens spektrum.



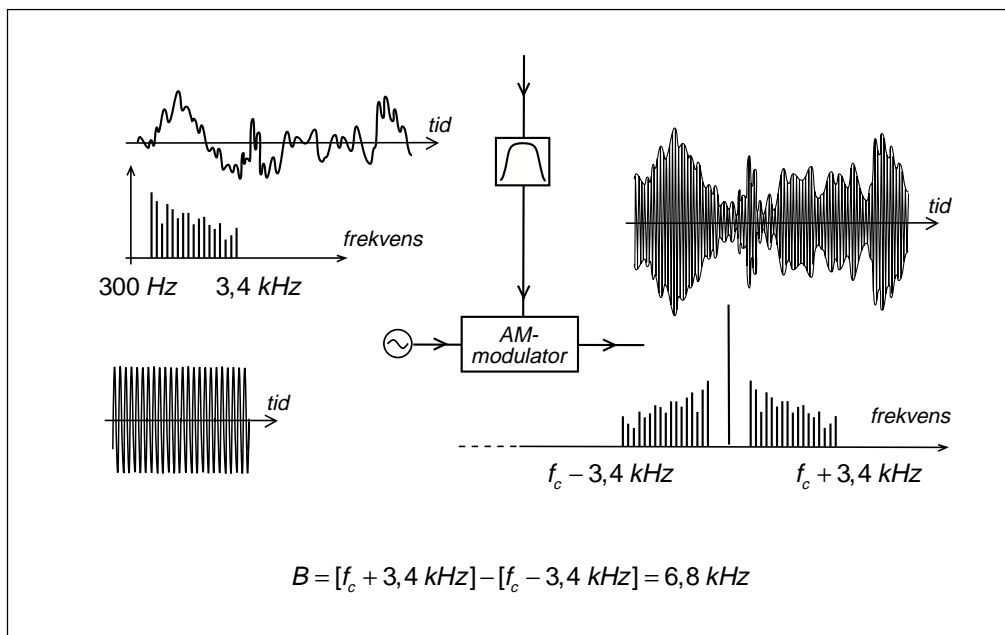
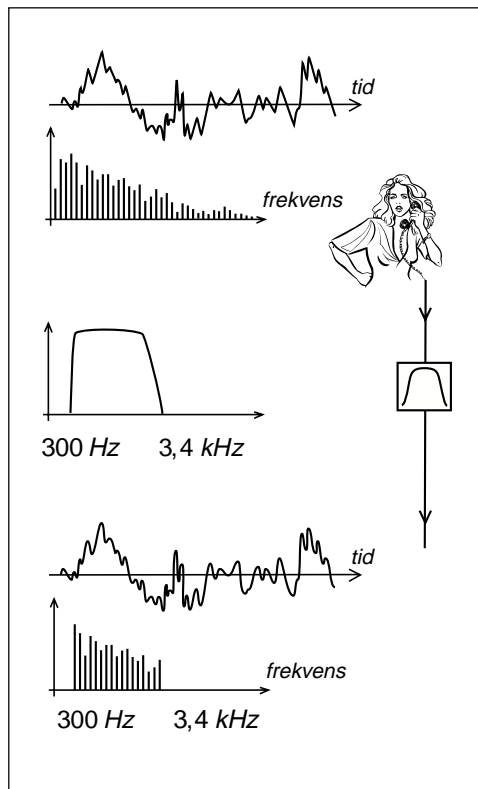
4 — Amplitudmodulering – AM

Informationssignalen filtreras för att begränsa bandbredden

Vid mobiltelefoni och kommunikationsradio ("telefonkvalitet") matas talsignalen genom ett filter som skär bort alla frekvenser över ca 3400 Hz. Om man amplitudmodulerar en sändare med sådant filtererat tal så får man en radiosignal med bandbredden 6,8 kHz.

Normalt filtrerar man även bort lågfrekvenser som ligger under 300 Hz. Dessa lågfrekventa frekvenskomponenter "stjäl" sändareffekt men talets uppfattbarhet försämras inte om man filtrerar bort dem. Observera att det "tomrum" på 600 Hz som bildas runt bärvågen inte kan användas till annat. Signalens bandbredd är 6,8 kHz oavsett om de lågfrekventa komponenterna filtreras bort eller ej.

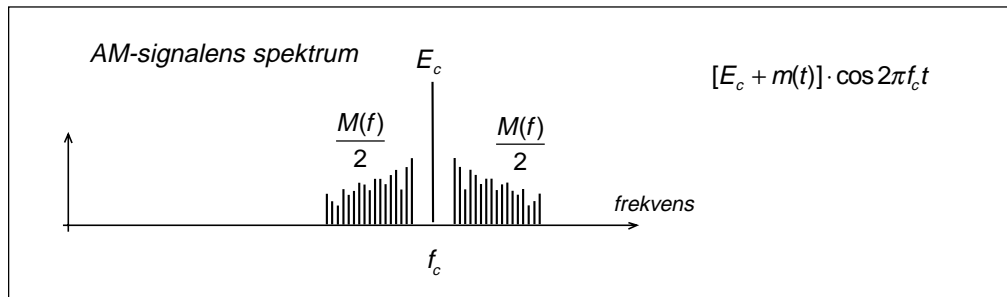
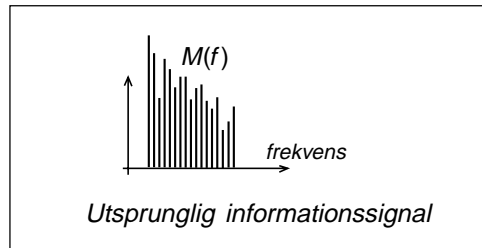
Rundradiosändarna på lång- och mellanvåg har tilldelats frekvenskanaler som är 9 kHz breda. Detta innebär att högsta modulationsfrekvens får vara 4,5 kHz. I verkligheten fuskar man och gör visst intrång på grannkanalerna så att högsta modulationsfrekvensen kanske hamnar uppåt 6 kHz. Detta är den ljudkvalitet man fick från en gammal 78 varvs grammofonskiva.



4.6 AM-signalens demodulering

Utgå från den ursprungliga informations-
signalen $m(t)$ som har spektrumet $M(f)$.

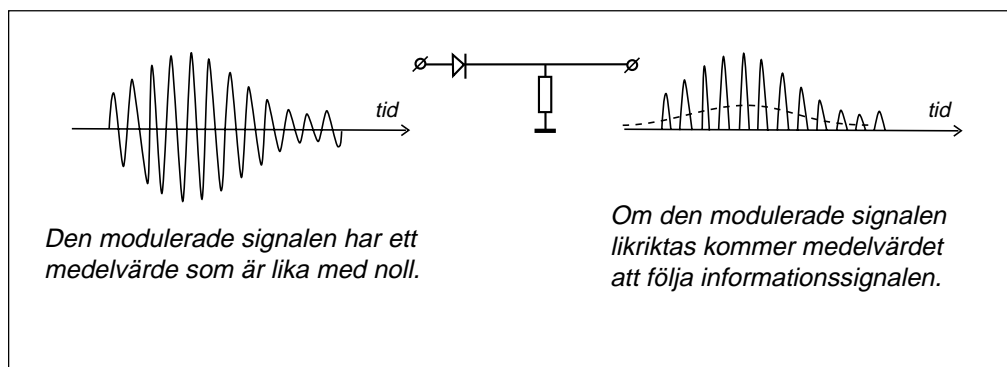
Skapa den modulerade AM-signalen som
får ett spektrum med bärvågen f_c och två
sidband bildade av summa- och skillnads-
frekvenserna.

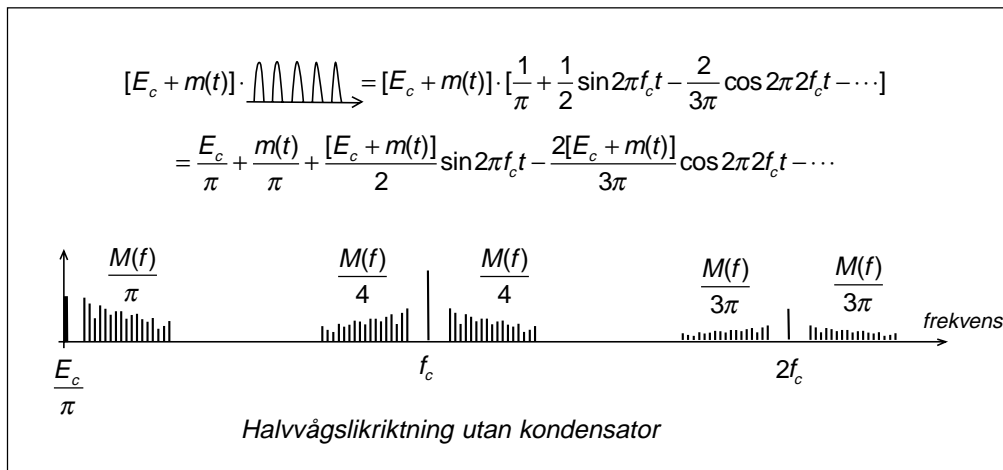


När AM-signalen tas emot av mottagaren,
så skall den ursprungliga informationssig-
nalen återskapas. Att återskapa informa-
tionssignalen kallas detektering och sker i
ett steg som kallas detektor eller
demodulator.

AM-detektorn

Enklaste sättet att återskapa den låg-
frekventa informationssignalen (LF-signa-
len) är att halvålslikrikta den högfrekventa
modulerade signalen (HF-signalen)



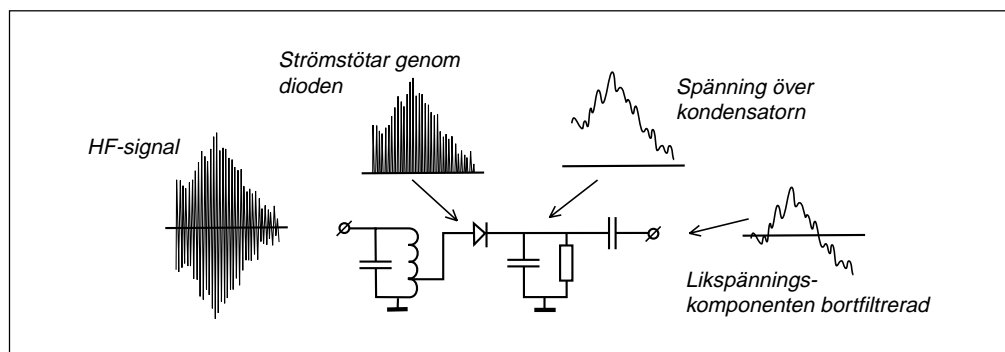
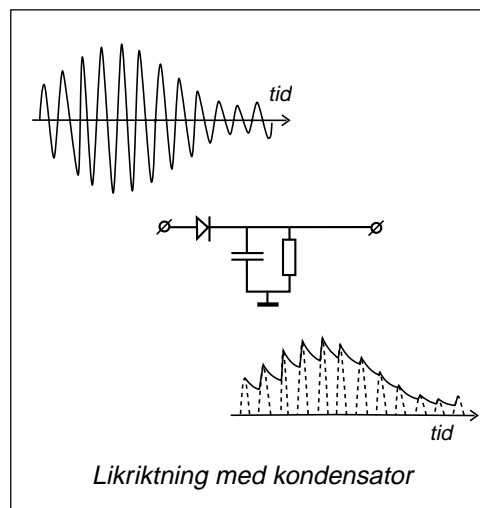


Matematiskt beskriver vi den likriktade signalen som amplituden multiplicerat med en halvvåglikriktad signal med amplituden=1 och periodtiden $1/f_c$. Denna kurvform finns i kapitel 2. Vi stoppar in Fourierserien med amplituden=1 och multiplicerar.

Resultatet blir ett spektrum där vi bl.a. finner den ursprungliga informations-signalen.

Envelopp-detektorn

Om man kopplar en kondensator parallellt med motståndet efter likriktardioden så kommer kondensatorn att laddas upp genom dioden (låg tidskonstant) men laddas ur genom motståndet (längre tidskonstant). Genom lämpligt val av tidskonstant så kan man återskapa en växelspanning som i stort sett följer "topparna" på HF-signalen, signalens "envelopp". Denna form av detektor kallas oftast envelopp-detektor.



4.7 Verklig AM-modulering

Ser ut som en strömbrytare

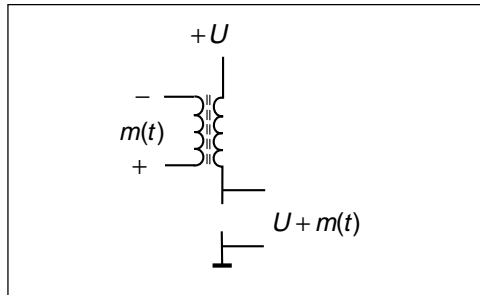
Den verkliga AM-sändaren uppför sig som en strömbrytare. Klass C-effektförstärkaren matas med en kollektorspänning som varierar i takt med $m(t)$. Basen styrs av bärvågsfrekvensen f_c , och transistorn är ömsom strypt (avbrott), ömsom kortsluten.

Matematiskt kommer kollektorspänningen att bestå av matningsspänningen multiplicerad med en fyrkantvåg med amplituden=1.

Fourierserieuppdelar fyrkantvågen och multiplicera så får vi

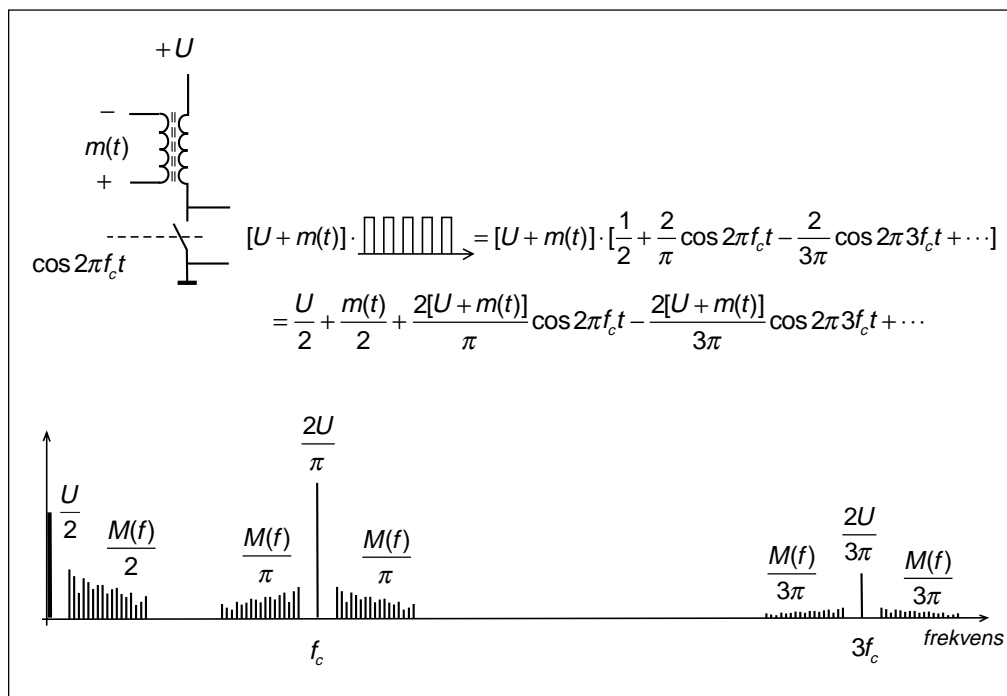
- en likspänningskomponent och halva ursprungliga informationssignalen,
- en AM-signal runt bärvågsfrekvensen f_c
- en AM-signal runt $3f_c$, runt $5f_c$ o.s.v.

Med en resonanskrets filtrerar vi ut AM-signalen runt f_c samtidigt som resonanskretsen fungerar som transformator till antenneresistansen.



AM-sändaren i Sölvesborg

I Sverige finns i dag endast en rundradio-sändare på långvågs- och mellanvågsbanden. Den ligger i Blekinge, vid Sölvesborg och täcker Östersjön och norra Europa. Sändaren är modern. AM-moduleringen åstadkommes inte så som beskrivits här, utan med konstant matningsspänning U , men fyrkantvågens pulsbredd varierar i takt med $m(t)$.



4.8 Effekt i AM-signalen

Antag att AM-sändaren ansluts till en belastning (antenn eller kabel) som ser ut som 50 ohm. Vad blir effekten som AM-sändaren levererar till denna belastning?

Att beräkna effektutvecklingen utgående från den modulerade spänningen är komplicerat. Betydligt enklare är att utgå från AM-signalens spektrum och beräkna den effekt som produceras av varje enskild spektralkomponent.

Effekt i bärvågen

Effekten i bärvågen är lättast att beräkna eftersom bärvågen alltid har amplituden E_c . I detta fall blir effekten

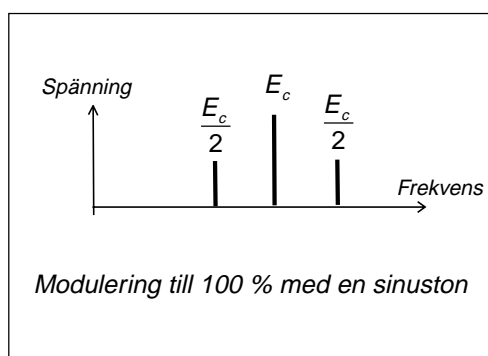
$$P_{\text{bärvåg}} = \frac{\left(\frac{E_c}{\sqrt{2}}\right)^2}{50} = \frac{E_c^2}{100}$$

Modulering med 1 sinuston

Om sändaren moduleras till 100 % med en sinuston så får varje sidband en amplitud som är lika med halva bärvågen.

$$P_{\text{sidband}} = \frac{\left(\frac{E_m}{2\sqrt{2}}\right)^2}{50} = \frac{E_m^2}{400} = \frac{E_c^2}{400} = \frac{P_{\text{bärvåg}}}{4}$$

Effekten i varje sidband blir 25 % av bärvågs-effekten.



Modulering med 2 lika starka sinustoner

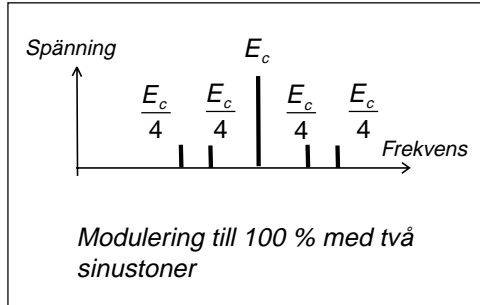
Om sändaren moduleras till 100 % med två lika starka sinustoner så får varje sidbandskomponent en amplitud som är en fjärdedel av bärvågsamplituden. Summan av dessa fyra sidbandskomponenter får max bli lika med E_c , annars blir sändaren övermodulerad.

$$P_{\text{sidband}} = 2 \cdot \left(\frac{E_c}{4\sqrt{2}} \right)^2 = 2 \cdot \frac{E_c^2}{1600} = \frac{P_{\text{bärvåg}}}{8}$$

I detta fall blir effekten i varje sidband bara 12,5 % av bärvågseffekten.

Modulering med verkligt spektrum

En verklig informationssignal har ett spektrum med många spektralkomponenter. Toppspänningen hos denna informationsignal får ej överstiga E_c om vi skall undvika övermodulering. Detta innebär att medeleffekten i varje sidband bara blir några procent av bärvågseffekten. Och det är i denna sidbandseffekt som själva informationen ligger.



5 — Amplitudmodulering – VSB, DSB, SSB, QAM

AM-signalen har två sidband som båda innehåller samma information om informations-signalen. Genom att filtrera bort merparten av det ena sidbandet så får signalen smalare bandbredd. Detta modulerings-sätt kallas VSB (Vestigial Sideband), stympar sidband, och används vid tv-sändningar.

Huvuddelen av effekten i en AM-signal ligger i bärvågen som inte innehåller någon information om informationssignalen (mer än vilken frekvens som är bärvågsfrekvensen). En AM-signal där man undertrycker bärvågen och bara sänder sidbanden kallas DSB (Double Sideband), dubbla sidband. Bärvågsfrekvensen kan återskapas genom att den ligger mitt emellan de identiska sidbanden.

Men det räcker faktiskt att bara sända ett sidband, SSB (Single Sideband). SSB upp-tar bara hälften så stor bandbredd som AM eller DSB, men ingen information om bärvågsfrekvensen. Därför används SSB bara för vissa informationssignaler, t.ex. tal.

Vid DSB så kan man sända två informations-signaler, två DSB-signaler, på samma frekvenskanal. Detta sändningssätt kallas QAM (Quad Amplitude Modulation) och är lika spektrumeffektivt som SSB.

5.1 Stympat sidband, VSB (Vestigial Sideband)

Vanlig tv är amplitudmodulerad

De analoga tv-sändarna amplitudmoduleras med bildsignalen. Hur stor bandbredd behövs för att överföra tv? Hur bred måste frekvenskanalen vara?

För att besvara den frågan behöver vi veta högsta frekvensen i informations-signalen (bildsignalen).

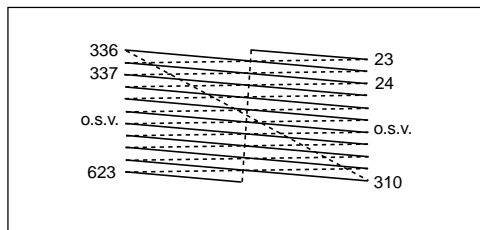
Tv-bilden är uppbyggd av 625 linjer med 25 bilder per sekund. Men för att få mindre flim-mer så sänder man 50 bilder per sekund, där varje bild innehåller varannan linje (interlaced scanning). Totalt skickas alltså 15 625 linjer per sekund. Därför har linjen 64 μ s på sig att gå över bildrutan från vänster till höger och tillbaka, som sker snabbare. Detta ger 52 μ s till själva linjen.

Tv-bilden består av 575 linjer i vertikallid. Det försvinner 25 linjer vid varje halv-bildväxling. Två av dessa linjer används för text-tv-datasignalen, 360 bit per linje (kan synas i överkanten på tv-bilden).

Klarar ögat att särskilja linjen som bild-element (pixel)? Nej, man räknar med en försämring motsvarande 70 %, d.v.s. 400 bildelement kan i bästa fall särskiljas, t.ex. som 200 svarta horisontella streck på tv botten.

Horisontellt vill vi kunna återge 4/3 så många bildelement. Vi vill kunna rita 270 vertikala svarta streck på vit botten. Varje linje måste amplitudmoduleras med 270 "toppar" och 270 "dalar", en sinusspänning med 270 perioder längs varje linje. Dessa 270 perioder skall sändas under 52 μ s. Detta blir 5,1 MHz. Därför har man valt 5 MHz som högsta frekvens i informations-signalen.

Titta på testbilden på tv. Strax under mitten ser du ett brett fält med vertikala streck. Till vänster är det glest mellan strecken. Sedan blir det gradvis tätare. I området runt den modulationsfrekvens som motsvarar 4,43 MHz flyter strecken ihop till ett färgmönster. Detta är störning från färg-moduleringen. Vid ännu högre frekvens flyter allt ihop och blir grått.



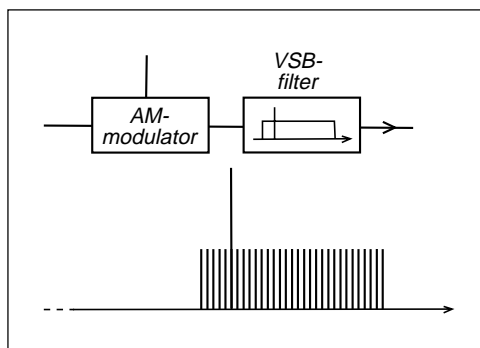
Tv-signalens spektrum

Den amplitudmodulerade tv-signalen har ett spektrum som sträcker sig 5 MHz åt båda håll runt bärvågen. Men informationen är identisk i de båda sidbanden. Det räcker att sända ett sidband.

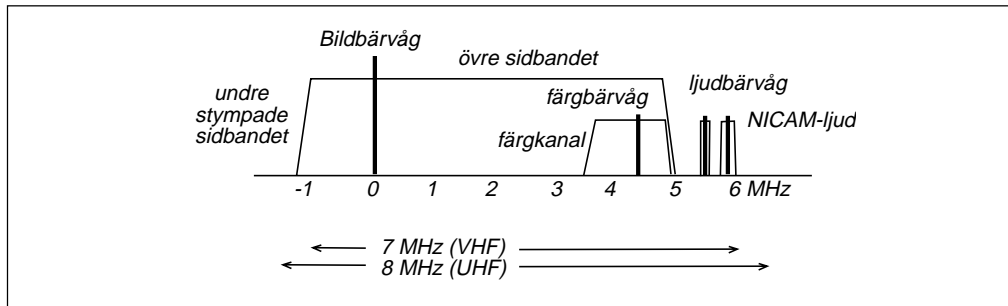
Stympat sidband, VSB

För att kunna sända tv-signalen på en smälare frekvenskanal så filtreras den amplitudmodulerade tv-signalen. Signalen matas genom ett filter som klipper bort större delen av det undre sidbandet.

Den frekvenskanal som används för tv är i Sverige 7 MHz bred på VHF-bandet (kanal 2 – 11) respektive 8 MHz bred på UHF-bandet (kanal 21 – 68). Förutom den amplitudmodulerade bildsignalen så sänder man tv-ljudet som ligger 5,5 MHz högre än bildbärvågen (FM-modulering) och NICAM stereoljud (digital modulering) som ligger 5,85 MHz högre än bildbärvågen.



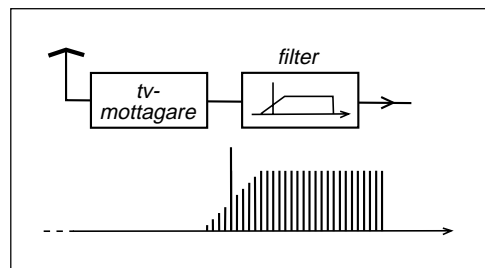
5 — Amplitudmodulering – VSB, DSM, SSB, QAM



I tv-mottagaren

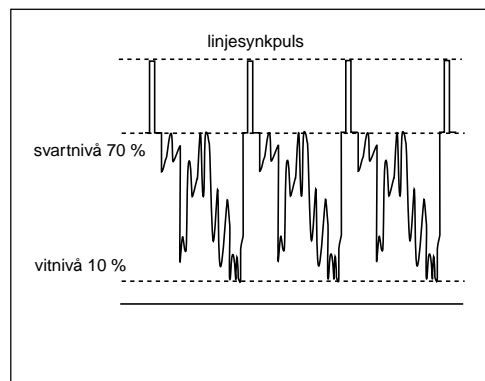
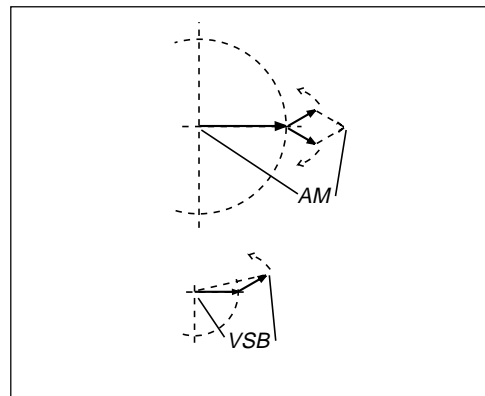
När den stympade tv-signalen kommer till tv-mottagaren så filtreras signalen ytterligare i ett filter som dämpar en del av det övre sidbandet, dämpar bärvågen till hälften, och dämpar det undre sidbandet. Varför?

Vid detekteringen kommer spektral-komponenterna i det övre och det undre sidbandet att adderas. Genom att skära på detta sätt kommer de lågfrekventa spektral-komponenterna att återfå rätt storlek. Att bärvågen dämpats till hälften gör inget eftersom vi bara har ett sidband.



Distorsion vid detekteringen

Summan av bärvågen och ett sidband ger inte bara amplitudändring, utan även fasändring, se visardiagram. Vid enveloppdetektering av en VSB-signal så kommer den detekterade signalen att bli för hög vid riktigt låga amplituder. Genom att lägga vitnivån vid 10 % av totala signalens amplitud (behövs för detekteringen av ljudet) så blir distorsionen vid VSB-detekteringen inte alltför störande i bilden (påverkar gråskalan).



5.2 Dubbla sidband, DSB (Double Sideband)

Balanserad blandare

Vi skall börja med en koppling som kallas "balanserad blandare" eller "multiplikator" eller "produkt-detektor", eller "ringmodulator". Den består av två transformatorer och fyra dioder. Blandaren har tre portar.

Anslut positiv spänning till port 2. Då kommer ström att flyta genom de yttre dioderna, medan de korslagda dioderna spär-rar.

Spänning som matas in på port 2 kommer inte ut genom port 1 eller 3, eftersom det blir motriktade strömmar i transformatorerna (balanseras bort). Det bildas inget magnetfält som kan ge spänning i port 1 eller 3.

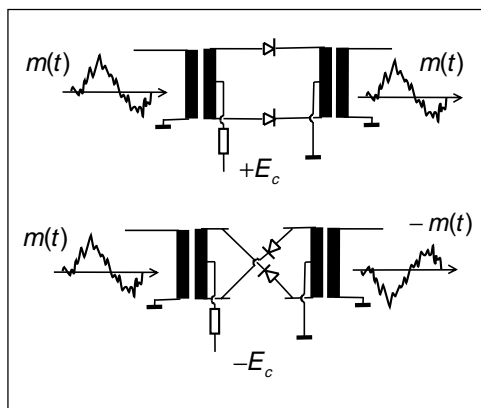
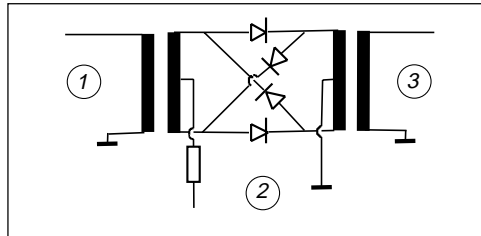
Anslut $m(t)$ till port 1

Anslut E_c till port 2. Om E_c är mycket större än $m(t)$ så kommer E_c att öppna och stänga dioderna. Dioderna fungerar som strömbrytare. Genom att ändra tecken på E_c så får även $m(t)$ motsatt tecken.

Låt E_c växla i takt med f_c

Om vi matar in bärvågen i port 2 och informationssignalen i port 1 så får vi en modulerad signal i port 3. Hur kommer denna signal att se ut?

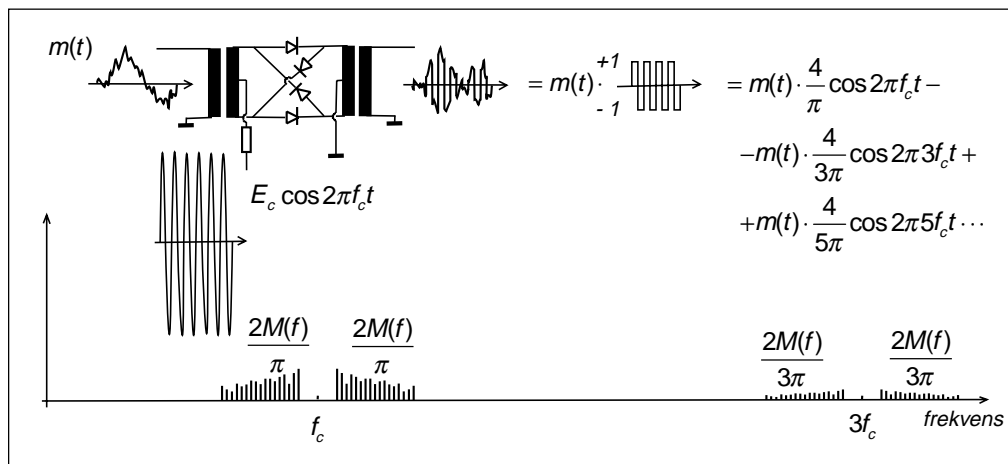
Signalen i port 3 kommer att bli $m(t)$ multiplicerad med en fyrkantvåg med amplituden ± 1 och periodtiden $1/f_c$.



Dubbla sidband, men ingen bärvåg!

Vad har vi fått? Dubbla sidband, men ingen bärvåg. Men behövs inte bärvågen? Jo, den behövs när den modulerade signalen skall skapas, men den behöver inte överföras. Det räcker att överföra sidbanden.

Vi får även sidband som ligger centrerade runt bärvågens udda övertoner.



5 — Amplitudmodulering – VSB, DSM, SSB, QAM

Vi kan styra dioderna från port 1

Det går lika bra att styra dioderna genom att lägga styrspänningen på port 1. Observera att styrspänningen E_c inte ger ström i transformatorn vid port 3, eftersom strömmen genom dioderna går åt motsatt håll. Däremot kommer $m(t)$ att ge ström antingen uppifrån och ner, eller nerifrån och upp, beroende på vilka dioder som leder.

Fördelen med att koppla på detta sätt är att nu kan $m(t)$ även innehålla likspänning.

Multiplikator

I fortsättningen ritas vi den balanserade blandaren som en fyrkant med ett kryss, och de tre ingångarna. Denna komponent kallas ibland multiplikator, eftersom den svaga signalen multipliceras med "takten" i den starka signalen, som styr dioderna.

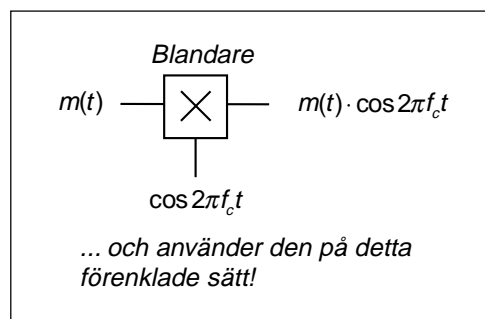
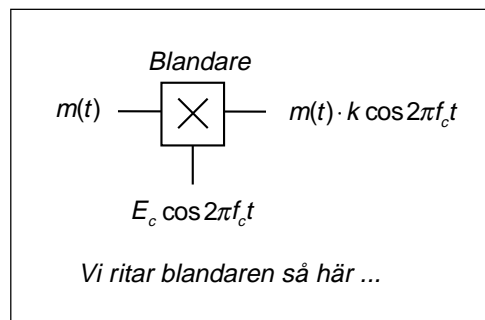
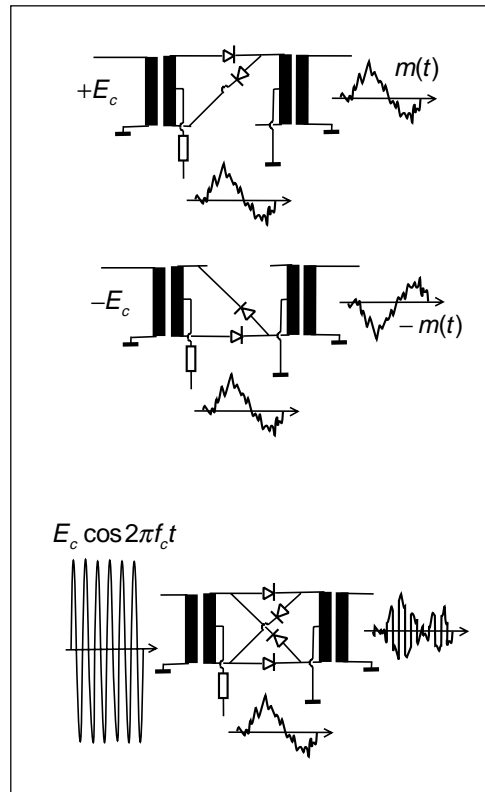
Produkt-detektor

Utsignalen blir produkten av de båda signalerna, dock utan att påverkas av bärvågens amplitud E_c . Villkoret är att E_c skall vara "tillräckligt" stor. Ett annat namn på den balanserade blandaren är därför "produkt-detektor".

Förenklat användningssätt

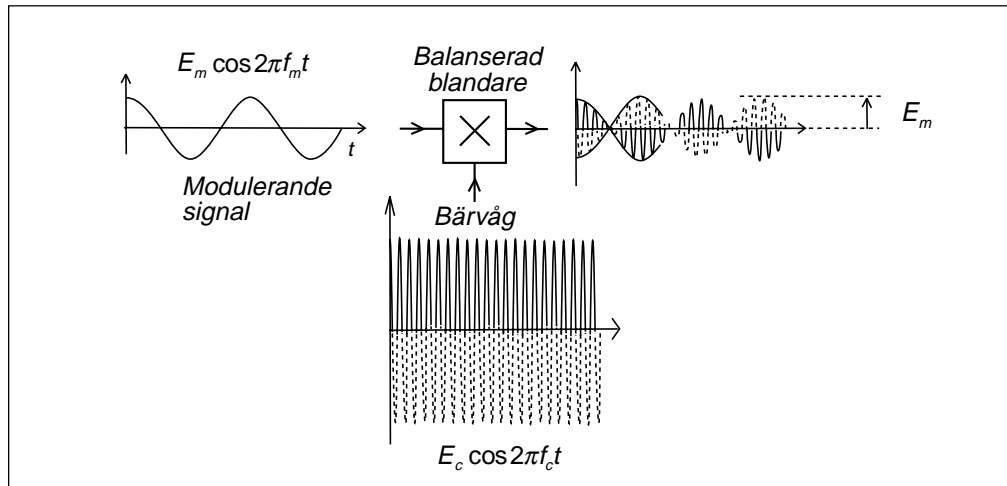
Vi kommer att använda oss av den balanserade blandaren så som visas i den undre bilden.

- Den signal som styr dioderna anges oftast utan amplitud, eftersom amplituden E_c är utan betydelse.
- Utsignalen anges som produkten av de båda signalerna. Egentligen skall det finnas en proportionalitetsfaktor k som bl.a. beror på transformatorernas varvtalsomsättningar, men vi slarvar. Utsignalens amplitud kan enkelt regleras i det efterföljande förstärkarsteget.
- Vi kommer även att få utsignal runt de udda övertonerna till bärvågen f_c , men dessa filtreras bort i efterföljande steg. Därför slarvar vi och påstår att utsignalen bara ligger runt själva bärvågen.



5 — Amplitudmodulering – VSB, DSB, SSB, QAM

17 lektioner i
TELEKOMMUNIKATION



DSB-modulering

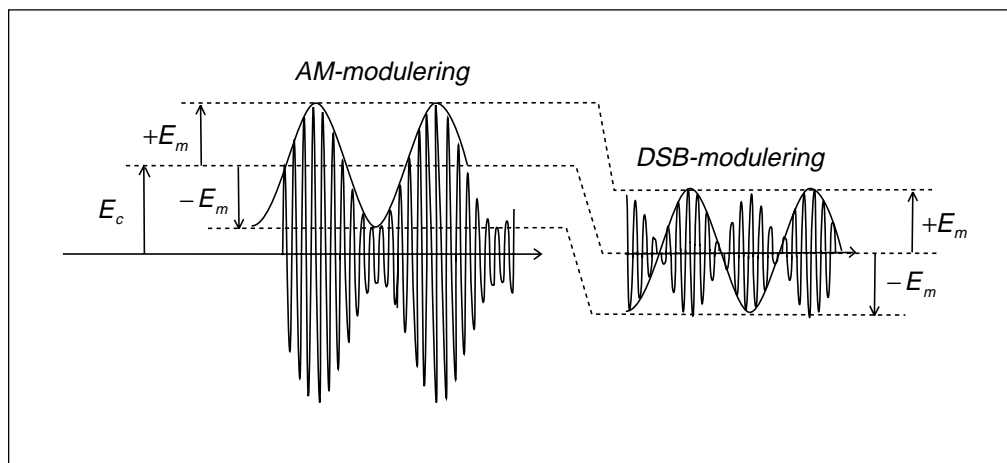
För att skapa en signal som består av två sidband runt en bärvåg som inte finns, så matar vi in informationssignalen på en balanserad blandare, och styr diodkopplingen med bärvågsfrekvensen.

På utgången får vi informations-signalen, med sin amplitud E_m , men signalen "pendlar" fram och tillbaka runt frekvens-axeln i takt med bärvågsfrekvensen f_c .

Jämför med AM

Vid AM så ändras bärvågens amplitud i takt med informationssignalen. Tänk dig att vi avlägsnar bärvågen så att bara ändringarna finns kvar. Då har vi DSB-signalen.

När vi skapar DSB-signalen så är bärvågs-amplituden E_c utan betydelse, så länge som den är "tillräckligt" stor, d.v.s. avsevärt större än E_m .



5 — Amplitudmodulering – VSB, DSM, SSB, QAM

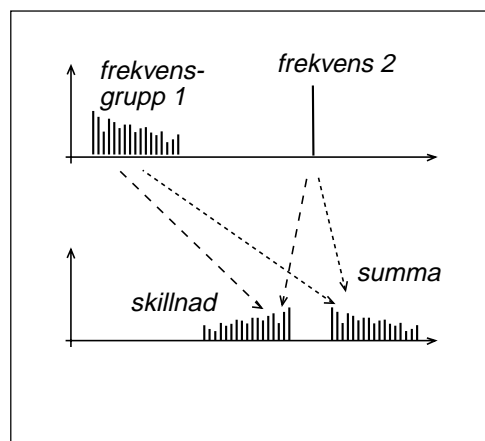
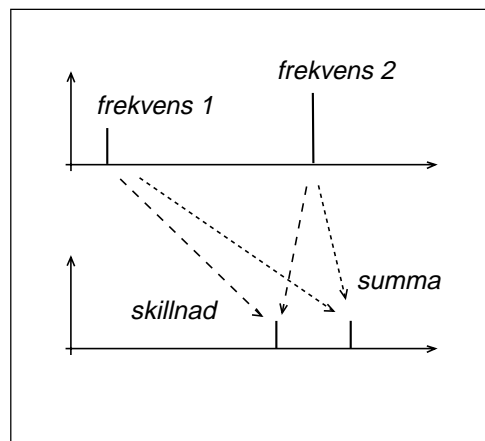
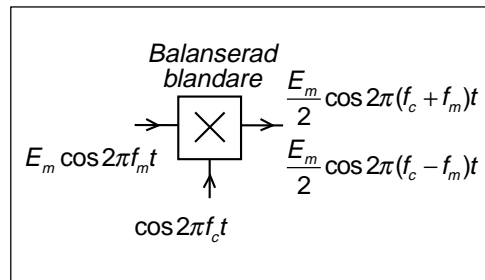
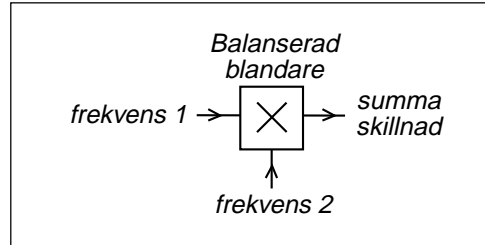
Balanserade blandaren skapar summa och skillnad

I den balanserade blandaren skapas signaler vars frekvenser är summafrekvensen och skillnadsfrekvensen av signalerna som matas in i port 1 och port 2. Man "blandar" signalerna och skapar summa och skillnad.

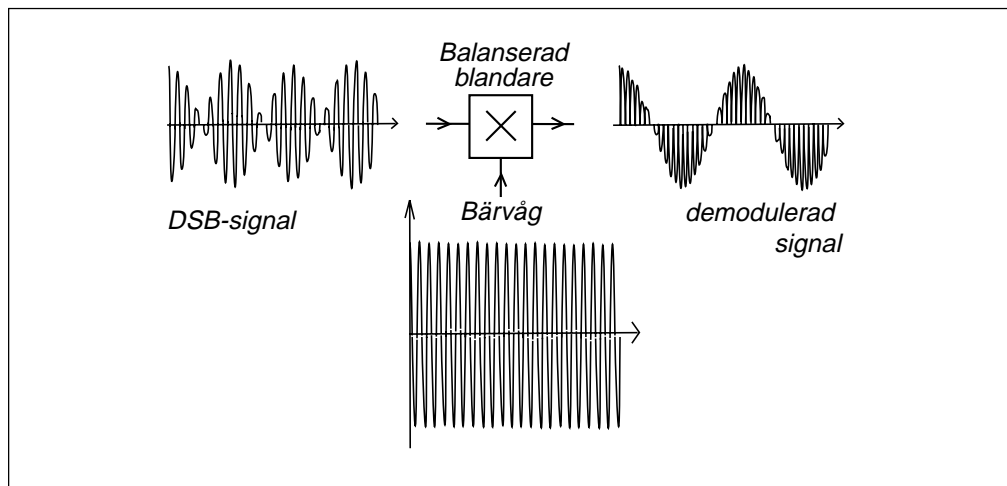
Informationssignal och bärvåg ger summan: övre sidbandet. Informationssignal och bärvåg ger skillnaden: undre sidbandet.

Modulering med verklig signal

Om informationssignalen inte är en sinus-signal utan en verklig signal, så beskriver vi den verkliga signalen med dess spektrum och tar summa och skillnad för varje enskild spektralkomponent.



5.3 DSB-signalens demodulering



Demodulering: att återskapa

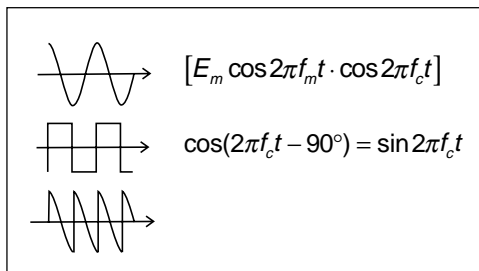
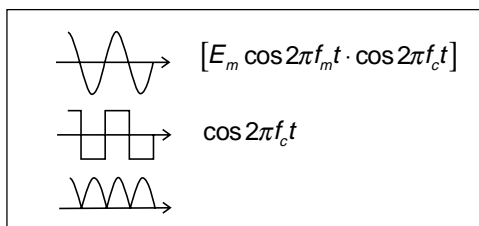
I mottagaren skall informationssignalen åter-skapas från DSB-signalen.

DSB-signalen skapades genom att informationssignalen "kastades" fram och tillbaka runt frekvensaxeln i takt med bärvågsfrekvensen. För att återskapa informationssignalen så gör vi samma sak en gång till. Vi kastar DSB-signalen runt frekvensaxeln i takt med bärvågsfrekvensen, så att DSB-signalen kastas tillbaka.

När DSB-signalen skapas så får vi egentligen "fyrkant"-pulser med övertonsspektrum runt alla udda övertoner till bärvågen f_c . Men dessa filtreras bort så att bara spektrumet runt grundtonen f_c sänds. Därför kommer den demodulerade DSB-signalen att bestå av sinuspulser liknande helvågslikriktning. Men likspänningskomponenten i denna "helvågslikriktade" signal följer exakt variationerna hos informationssignalen.

Viktigt att bärvågen ligger i rätt fas

I bilderna här bredvid har jag visat vad som händer om bärvågssignalen i mottagaren inte kopplar om dioderna exakt i rätt tidsögonblick. Om bärvågssignalen är fördröjd en kvarts period, 90° , så sker omkopplingen

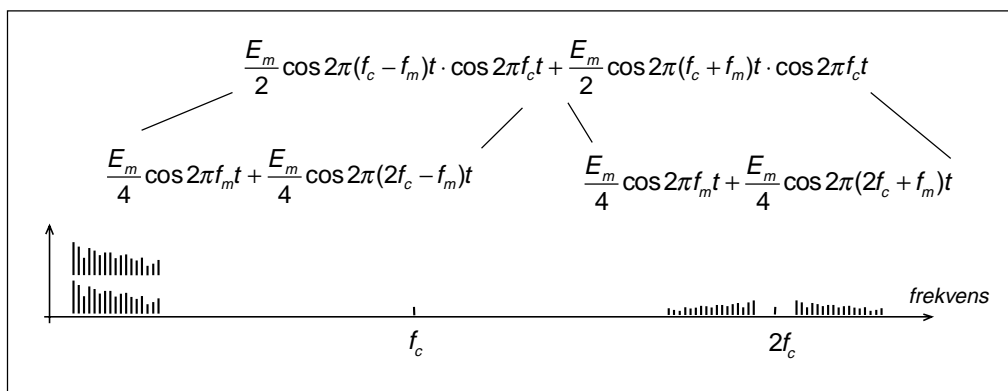
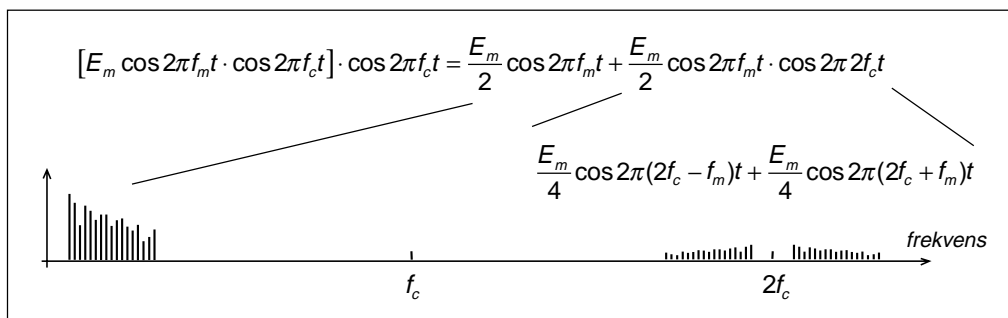
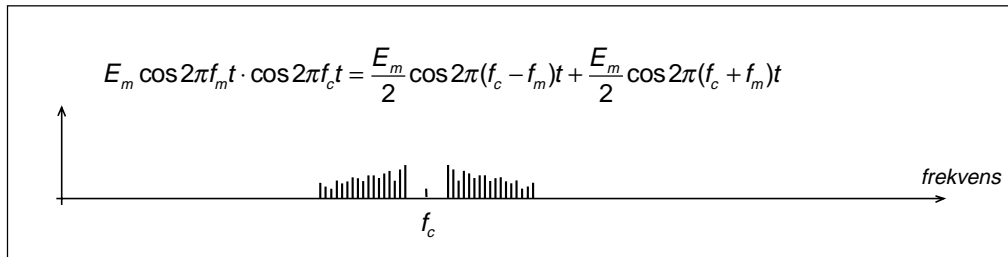


på sådant sätt att utsignalen blir en växelspanning utan likspänningskomponent, d.v.s. vi får ingen demodulerad signal.

Koherent detektor

DSB-demoduleringen behöver den ursprungliga bärvågen, inte bara med rätt frekvens f_c utan även fas. Den skall ligga rätt "i tid". En detektor med fasriktig bärvåg kallas "koherent detektor".

5 — Amplitudmodulering – VSB, DSM, SSB, QAM



Summa och skillnad

Än en gång skall vi se på DSB-signalen ur perspektivet "summa och skillnad".

I den översta bilden skapar vi DSB-signalen genom att multiplicera informations-signalen med bärvågen. Som resultat får vi summa och skillnad, övre och undre sidbanden.

I bild 2 detekterar vi genom att multiplicera DSB-signalen med bärvågen. Resultatet blir tre! signaler, summasignalen som hamnar vid dubbla bärvågsfrekvensen och skillnadsfrekvensen som hamnar som demodulerad signal. Varför tre?

I bild 3 ser vi tydligare. Den demodule-

rade signalen består av två signaler inom samma frekvensområde.

Om vi använder en detektor där bärvågen ligger rätt i fas så kommer skillnads-signalerna att bli identiska och sammansätta sig till dubbla amplituden på varje spektral-komponent.

Om bärvågen ligger 90° fel, så kommer frekvenskomponenterna i skillnads-signalerna att hamna 180° relativt varandra, d.v.s. summan av de båda skillnads-signalerna blir noll. Detta motsvaras av bilden på föregående sida, den bild där vi inte har någon likspänningkomponent.

5.4 Enkelt sidband, SSB (Single Sideband)

Det räcker med ett sidband

Vid amplitudmodulering, såväl vid AM som vid DSB, så innehåller de båda sidbanden identisk information. Det räcker att överföra ett sidband, antingen det övre eller det undre, för att signalen skall innehålla all nödvändig information om informationssignalen.

Nästan, men inte riktigt

Räcker det verkligen att överföra endast ett sidband? Nej, inte riktigt. För att man skall kunna återskapa en exakt kopia av informationssignalen så måste vi även veta bärvågens exakta frekvens. Vi skall se på detta.

Först skapar vi SSB-signalen

Att skapa en SSB-signal innebär att först skapa en DSB-signal och sedan mata den genom ett bandpassfilter som bara släpper igenom ett av sidbanden.

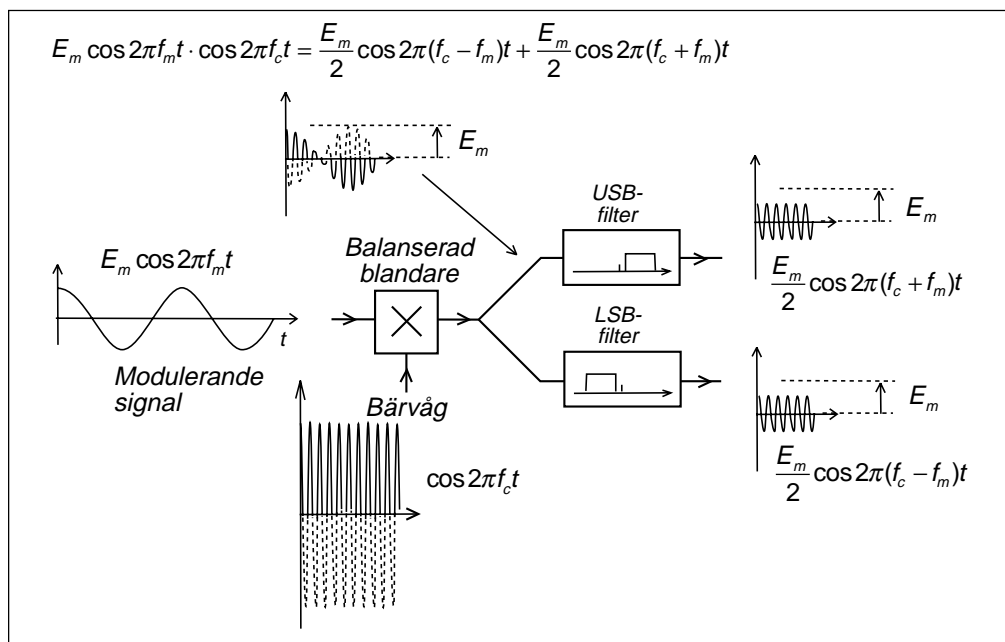
Övre sidbandet, USB (Upper Sideband)

Vi kan filtrera ut det övre sidbandet. Om informationssignalen består av en enda sinussignal, då blir även det övre sidbandet en enda sinussignal.

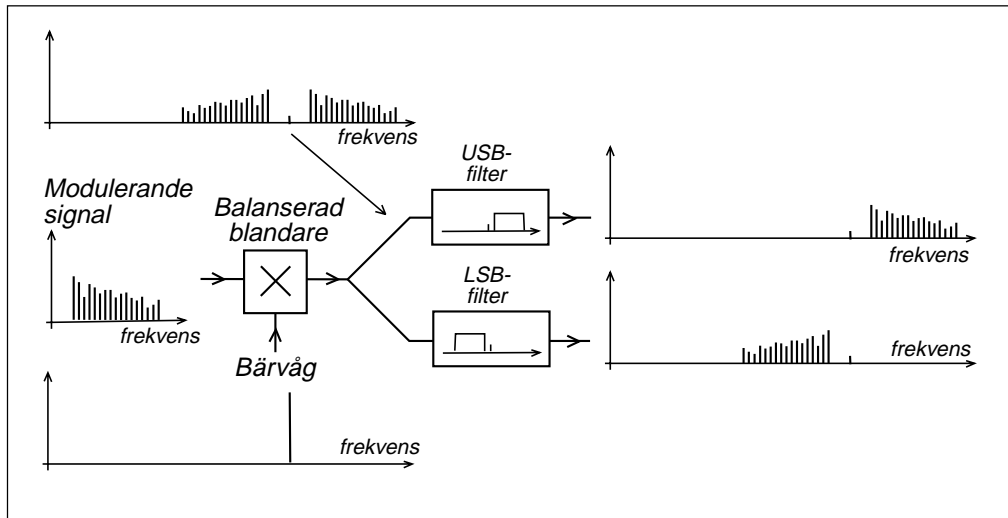
Undre sidbandet, LSB (Lower Sideband)

... eller vi kan filtrera ut det undre sidbandet. Vilket sidband man använder har egentligen ingen betydelse ur överföringssynpunkt.

Radiofrekvensbanden delas in i frekvenskanaler, där varje kanal skall rymma en radiosändare. Det kan vara fördel ur störningssynpunkt om radiosändarna i närliggande frekvenskanaler använder samma sidband, bl.a. beroende på att bärvågen aldrig kan avlägsnas helt, den kan bara undertryckas.



5 — Amplitudmodulering – VSB, DSM, SSB, QAM

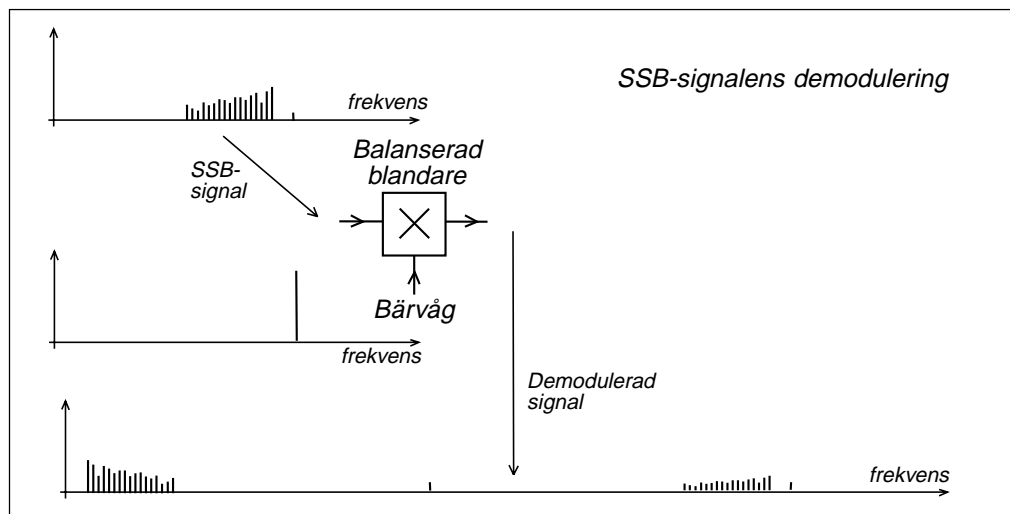


SSB-signalens demodulering

Demoduleringen sker på samma sätt som vid DSB. Vi multiplicerar med bärvågen och får summa och skillnad, varvid skillnadsfrekvenserna bildar informationssignalens spektrum.

Vid DSB bildades två demodulerade signaler som kunde hamna i motfas och släcka ut varandra. Därför var bärvågens fas viktig vid DSB-demodulering. Vid SSB bildas bara en demodulerad signal. Därför har fasläget hos demodulatorns "bärvåg" ingen betydelse.

Vid DSB så klarar man att återskapa bärvågen till rätt frekvens och fas därför att den skall hamna exakt mitt emellan sidbandens spektralkomponenter. Vid SSB finns ingen sådan information som kan utnyttjas. Därför ansånds SSB vanligtvis bara vid sådan överföring där ett litet frekvensfel kan tolereras, t.ex. talöverföring. Det räcker att den återskapade "bärvågen" ligger inom 20 Hz för att talet skall vara fullt förståeligt. Musik låter dock hemskt redan vid betydligt mindre frekvensfel.

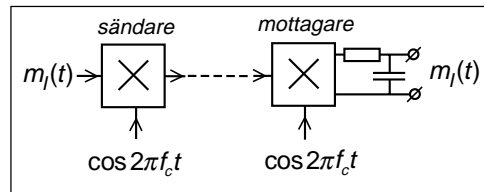


5 — Amplitudmodulering – VSB, DSB, SSB, QAM

17 lektioner i
TELEKOMMUNIKATION

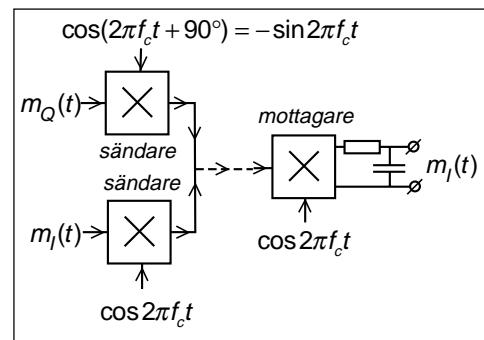
5.5 QAM (Quad Amplitude Modulation)

Skapa en DSB-sändare och en DSB-mottagare. Förbindelsen mellan sändaren och mottagaren kan vara radiovågor, men lika gärna en ledning. För enkelhets skull kopplar vi ihop sändaren och mottagaren med en lång ledning.



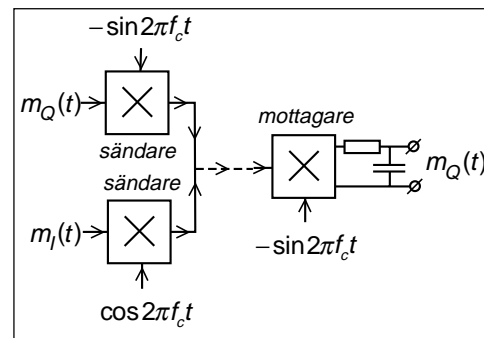
Modulera DSB-sändaren med $m_I(t)$. Ut från mottagaren, efter lågpasfiltret, får vi $m_I(t)$.

Skapa ytterligare en DSB-sändare som moduleras med $m_Q(t)$ och låt bärvågen ligga 90° fasförskjuten i förhållande till den första sändaren.



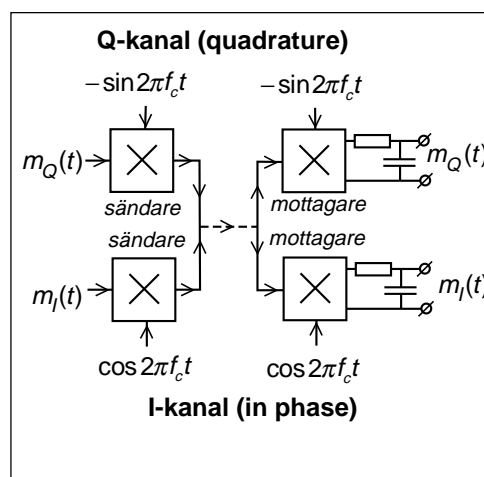
Vad får vi efter mottagaren? Samma som tidigare. Eftersom $m_Q(t)$ är modulerad på en bärvåg som ligger 90° fasförskjuten så kommer sidbanden från $m_Q(t)$ att ta ut varandra. Vi får fortfarande bara $m_I(t)$ efter vår mottagare.

Vi kan alltså använda ledningen för att transportera två signaler inom samma frekvenskanal centrerad runt f_c .



Med ytterligare en mottagare som åter-skapar bärvågen $-\sin 2\pi f_c t$ så detekterar vi $m_Q(t)$.

Vid SSB klarar vi att överföra signalen på halva bandbredden. Men vid DSB så kan mottagaren åter skapa bärvågen i rätt fasläge och då kan vi överföra två signaler på dubbla bandbredden och alltså få samma spektrumeffektivitet som vid SSB.



Metoden bygger på att vi har dubbla sidband. I den ena mottagaren hamnar de båda sidbanden i fas och adderas, i den andra mottagaren hamnar de båda sidbanden i motfas och subtraheras. Något liknande är ej möjligt vid SSB som ju bara har ett sidband.

I-kanal och Q-kanal

Man kallar de båda transmissionskanalerna för I-kanalen, som ligger i fas med den ursprungliga DSB-sändaren, och Q-kanalen, som ligger 90° fasförskjuten, i kvadratur.

5 — Amplitudmodulering – VSB, DSM, SSB, QAM

QAM-modulator

Eftersom vi amplitudmodulerar två signaler där den ena ligger i kvadratur, så kallas denna modulator bestående av dubbla DSB-modulatorer för QAM-modulator (Quad Amplitude Modulator).

QAM-modulator kan åstadkomma vilken fasvinkel som helst

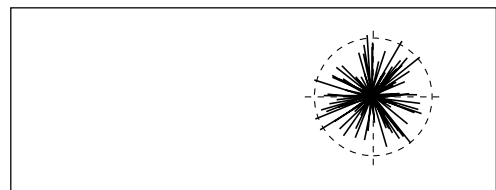
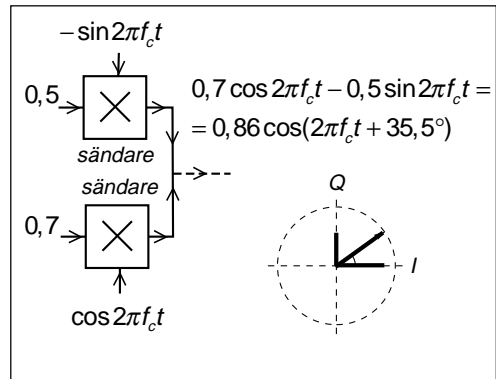
De båda DSB-signalerna på utgången av QAM-modulatorn kan ses som en signal med fasvinkel. Genom att välja lämpliga amplituder på I- och Q-signalerna så kan man åstadkomma vilken amplitud och fasvinkel som helst på bärvågen. Vi skall se att QAM-modulatorn även kan användas vid fas- och frekvensmodulering.

Brus

Brus kan ses som många spänningsvektorer som snurrar i visardiagrammet, med olika amplituder och rotationshastigheter. Summan av alla dessa spänningsvektorer är brusspänningen E_n .

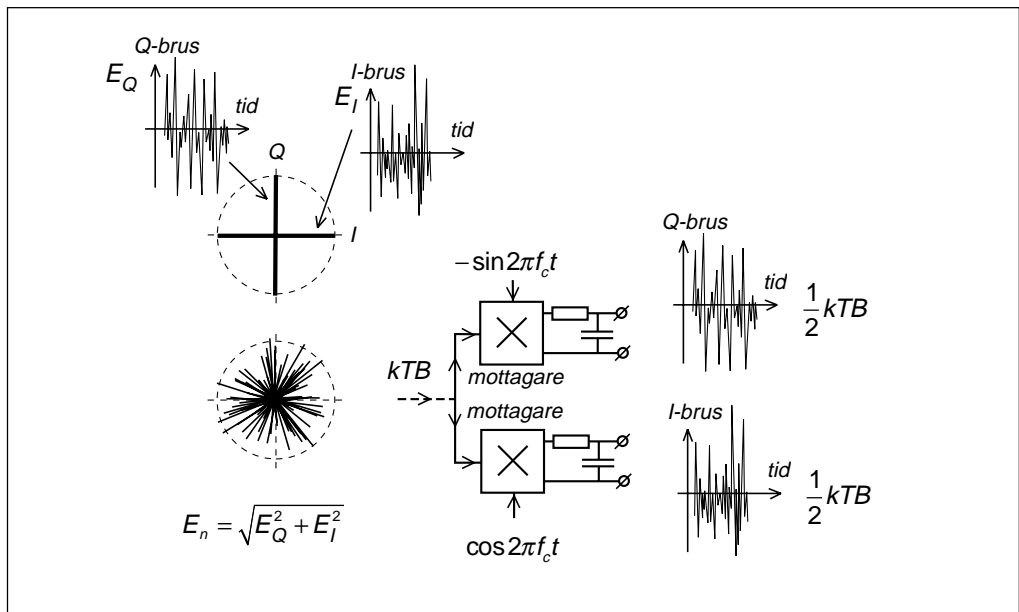
I-brus och Q-brus

Om man matar in brus på en QAM-mottagare så kommer mottagaren att dela upp brus signalen på en bruskomponent som lig-



ger i fas med I-kanalen, och en bruskomponent som ligger i fas med Q-kanalen.

Den brus effekt som kommer in till QAM-mottagaren är "brus före detektorn", d.v.s. det kTB som vi beräknat i kapitel 3. Denna brus effekt delas upp så att halva brus effekten påverkar I-kanalen, och andra halvan påverkar Q-kanalen.



5.6 Signal/brusförhållandet efter detektorn

Även brus

Fram till mottagarens detektor kommer inte bara den önskade signalen utan även brus.

AM utan modulering

Vi lyssnar på en AM-station utan modulering. Vad händer med bruset vid detekteringen?

Signalen till detektorn är bärvågen plus brusvektorn. Om bärvågen är stark så kommer I-bruset att påverka resultatens längd, medan Q-bruset bara ger fasmodulering. Det brus som hörs efter detektorn är alltså halva bruseffekten i varje sidband, multiplicerat med två sidband.

Om bärvågen är svag så kommer både I- och Q-bruset att påverka resultatens längd och vi får dubbelt så hög bruseffekt.

DSB

Vid DSB lägger man till bärvågen i mottagaren och kan göra bärvågen stark oavsett styrkan på DSB-signalen. Bruset håller alltid samma styrka som vid AM med stark bärvåg.

Vid svag DSB-signal kan det dock uppstå problem med att låsa bärvågen i rätt fasläge och demodulera själva signalen.

SSB

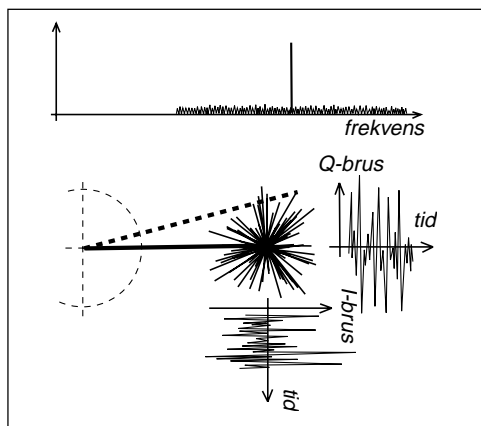
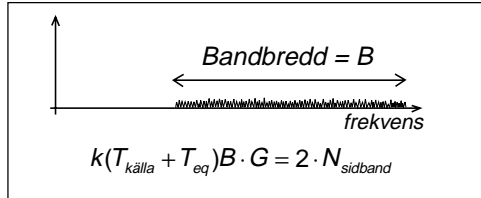
Vid SSB gör man som vid DSB, lägger till en stark bärvåg så att bara I-bruset detekteras. Dessutom finns ett filter i mottagaren som bara släpper fram brus i ett sidband till detektorn. Eftersom SSB inte behöver låsa bärvågen till rätt fasläge så är SSB överlägset bäst vid mycket svaga signalstyrkor.

Sidbandseffekten

Vid AM och DSB fördelas sändareffekten på de båda sidbanden. Vid detekteringen adderas sidbandskomponenterna till dubbla spänningen, vilket motsvarar 2 ggr så hög effekt som i de båda sidbanden.

Signal/brus-förhållandet

Signal/brus-förhållandet, om vi med signal menar totala effekten i sidbandet (SSB) eller totala effekten i de båda sidbanden (AM, DSB), blir identiska vid AM, DSB och SSB.



AM: stark bärvåg:

$$N_{AM} = 2 \cdot \frac{1}{2} N_{sidband} = N_{sidband}$$

AM: svag bärvåg:

$$N_{AM} = 2 \cdot N_{sidband}$$

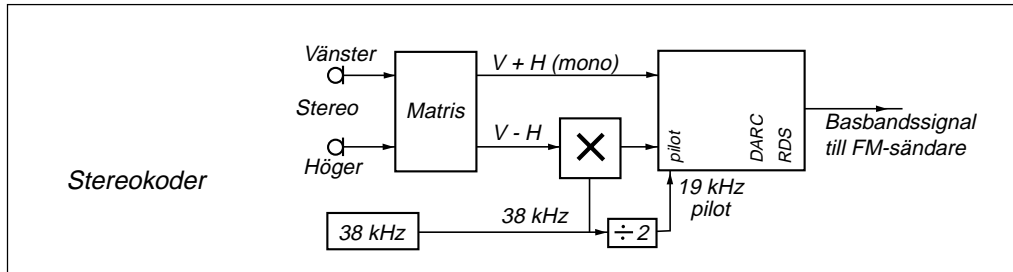
$$N_{DSB} = 2 \cdot \frac{1}{2} N_{sidband} = N_{sidband}$$

$$N_{SSB} = \frac{1}{2} N_{sidband}$$

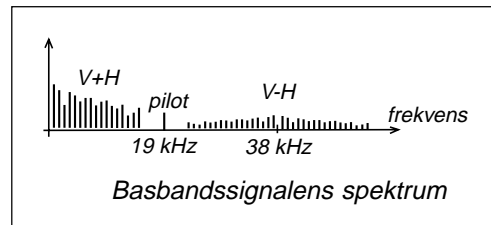
$$\left(\frac{S}{N}\right)_{AM,DSB} = \frac{2 \cdot (P_{sidband} + P_{sidband})}{N_{sidband}}$$

$$\left(\frac{S}{N}\right)_{SSB} = \frac{(P_{sidband})}{\frac{1}{2} N_{sidband}}$$

5.7 FM-stereo — exempel på DSB



Spänningarna från höger och vänster mikrofon kombineras i en matris så att man bildar summasignalen ($V + H$) som är monosignalen, samt skillnadssignalen ($V - H$). Man gör detta för att monomottagare skall kunna ta emot en monosignal, även om det är stereosändning.



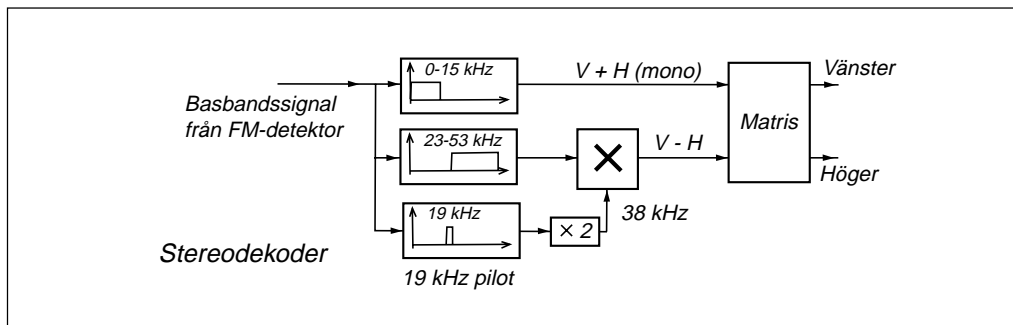
Monosignalen ($V + H$) täcker frekvensområdet 30 Hz – 15 kHz och skickas till FM-sändaren som den är.

Stereoinformationen ($V - H$) DSB-moduleras med bärvågen 38 kHz, så att den bildar två sidband inom frekvensområdet 23 kHz – 53 kHz. Dessa sidband matas till FM-sändaren. Eftersom stereoinformationen ligger över det hörbara frekvensområdet så kommer den inte att märkas i en monomottagare.

För att stereomottagaren skall kunna demodulera DSB-signalen så behöver mottagaren tillgång till 38 kHz bärvågen inte bara till rätt frekvens utan även med rätt fasläge. För att hjälpa stereomottagaren så matar man en liten signal som är halva

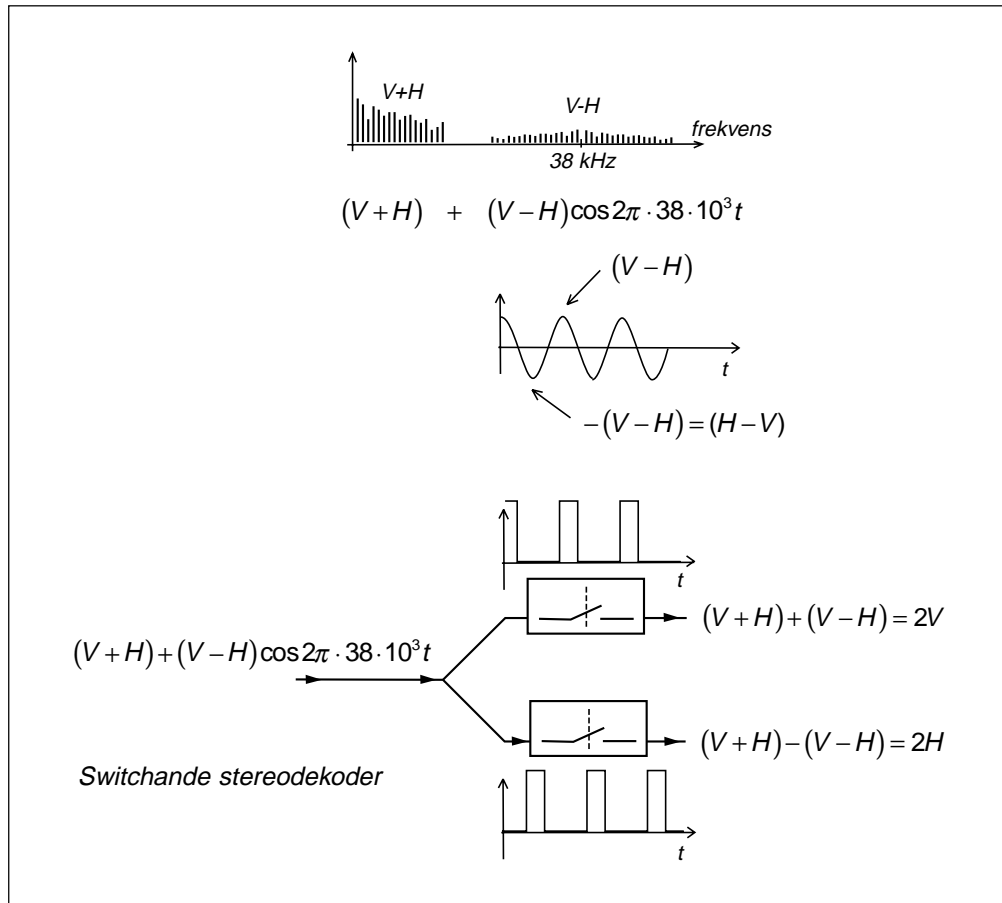
bärvågsfrekvensen, 19 kHz, till FM-sändaren. Denna hjälpsignal kallas "pilot", och stereosystemet kallas pilottonsystemet.

I stereomottagaren filtrerar man ut monosignalen, DSB-signalen och piloten. Piloten förstärks och multipliceras med 2 till 38 kHz och används för att demodulera DSB-signalen. När man väl har $V + H$ samt $V - H$ så får man vänster kanal genom att summera signalerna, samt höger genom att först fasvrider den ena signalen 180° och därefter summera.



5 — Amplitudmodulering – VSB, DSB, SSB, QAM

17 lektioner i
TELEKOMMUNIKATION



Switchande stereodekoder

Det enklaste sättet att i dag konstruera en stereodekoder är med digitalteknik.

Basbandssignalen som kommer från FM-detektorn består av $(V+H)$ och stereo-informationen $(V-H)$ som ömsom är positiv, ömsom negativ, i takt med 38 kHz.

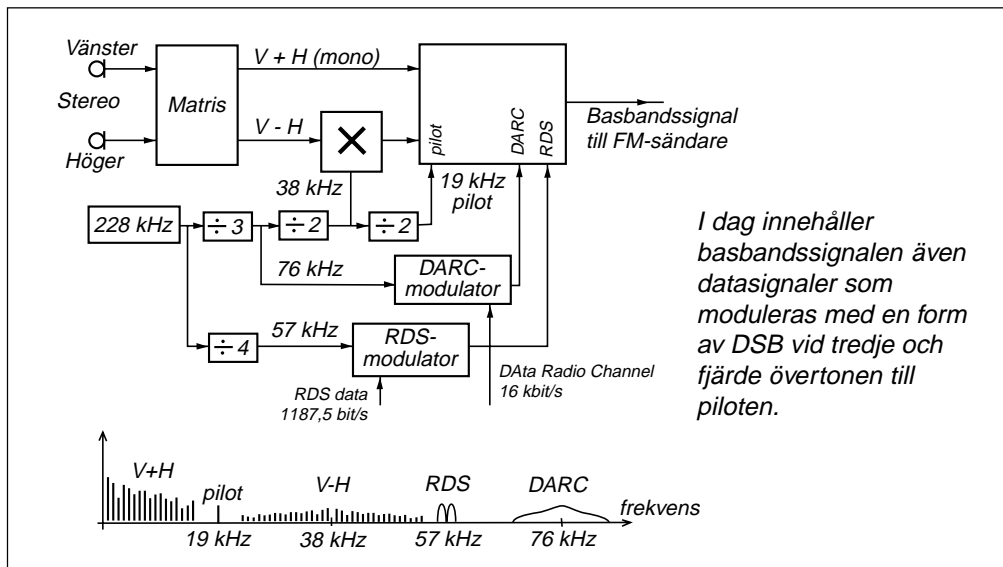
Genom att konstruera två strömbrytare och styra strömbrytarna så att den ena sluter när cosinusspänningen är positiv, och den andra strömbrytaren sluter när cosinus är negativ, så får vi på utgången av strömbrytarna spänningpulser vars medelvärde är vänster respektive höger stereokanal.

Fasfel ger sämre stereoseparation

Vad händer om strömbrytarna inte öppnar och sluter i exakt rätt tidsögonblick, d.v.s. om strömbrytarna har ett fasfel relativt den undertryckta 38 kHz bärvågen?

Det som händer vid fasfel är att man får lägre amplitud på "stereopulserna", något som resulterar i att vänster ljudkanal kommer att innehålla lite höger, och höger ljudkanal kommer att innehålla lite vänster. Vi får sämre stereoseparation. Men ljudkvaliteten som sådan kommer inte att påverkas. Vi får t.ex. ingen distorsion på själva ljudet.

5 — Amplitudmodulering – VSB, DSM, SSB, QAM

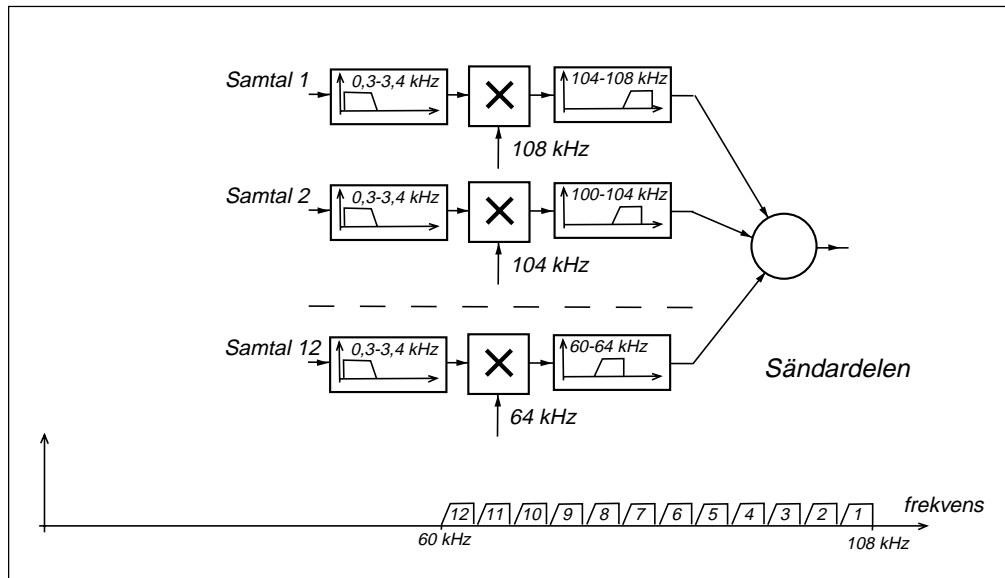


Dagens FM-sändare

Dagens FM-sändare används inte bara för stereoutsändning, utan på tredje övertonen till piloten, 57 kHz, ligger RDS (Radio Data System), en digital kanal med bithastigheten 1187,5 bit/s. Denna datakanal utnyttjades tidigare för RDS personsöksystemet (MBS). I dag finns programinformation på RDS-kanalen, samt vissa låghastighets datatjänster, t.ex. utsändning av korrektionssignaler för differentiell GPS positionsbestämning.

På fjärde övertonen till piloten, 76 kHz, ligger ytterligare en datakanal DARC (Data Radio Channel) modulerad med 16 kbit/s. Detta är helt och hållet en kommersiell tjänst.

5.8 Analog FDM-telefoni — exempel på SSB

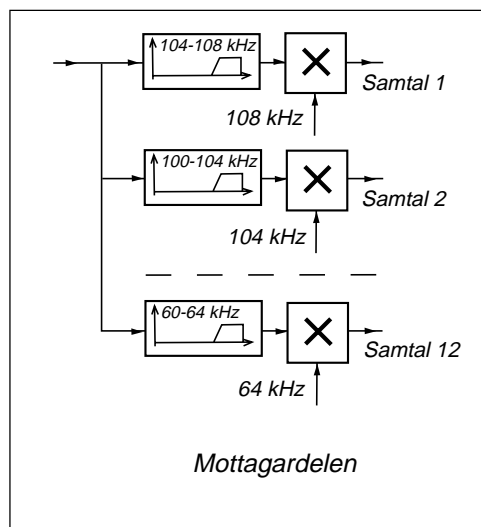


FDM (Frequency Division Multiplex)

Med FDM menas att flera signaler skickas på samma förbindelse, multiplexeras, genom att varje signal får sitt eget frekvensområde.

År 1949 infördes det första FDM-systemet i svenska telenätet, när man på en koaxialkabel mellan Stockholm och Norrköping överförde 600 samtidiga telefonsamtal med SSB-teknik. Koaxialkabeln var försedd med förstärkare på 9,6 km avstånd, som klarade att förstärka signaler upp till 2,6 MHz.

Tekniken utvecklades så att man 1977 på koaxialkabeln mellan Stockholm och Göteborg klarade att överföra 10 800 samtidiga telefonsamtal. Koaxialkabeln var försedd med förstärkare på 1,5 km avstånd, som klarade att förstärka signaler upp till 60 MHz.



6 — Superheterodynmodtageren

Så godt som alle radiomodtagerer endda sedan slutet av 1930-talet är uppbyggda enligt superheterodynprincippet.

Superheterodynprincippet innebär att man bygger en modtager för en fast frekvens, mellanfrekvensen, en modtager med den bandbredd och den typ av detektor som önskas.

Den antennisignal som man vill ta emot "blandas" i en blandare med en lokalt alstrad signal, så att skillnadsfrekvensen (även summasignalen förekommer) blir mellanfrekvensen.

Vi studerar hur detta fungerar i en rundradio-modtager för mellanvåg, i en privatradiomodtager för 27 MHz PR-bandet samt vid mottagning av satellit-tv med parabolantenn.

Avslutningsvis tittar vi på modtagarens förstärkning samt signal/brusförhållandet före detektorn.

6.1 AM-mottagaren

En radiomottagare som skall ta emot en AM-modulerad radiosignal innehåller en AM-detektor och en förstärkare, en LF-förstärkare som förstärker den lågfrekventa ljudsignalen till sådan styrka att den kan höras i en högtalare.

Vilka signaler kommer från antennen? Alla signaler på alla frekvenser som kan fångas upp av antennen. Det vi hör i högtalaren är moduleringen på den signal som har den starkaste bärvågen.

Men vi vill inte lyssna på den starkaste signalen, utan på en viss radiosändare på en viss frekvens. Därför behövs ett bandpassfilter som filtrerar ut den önskade signalen. Detta filter skall helst inte ha större bandbredd än absolut nödvändigt.

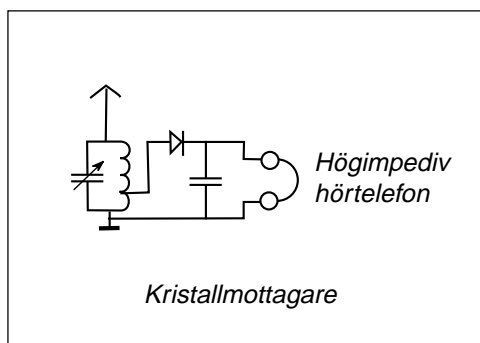
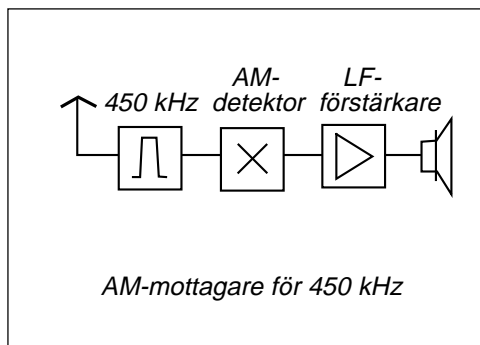
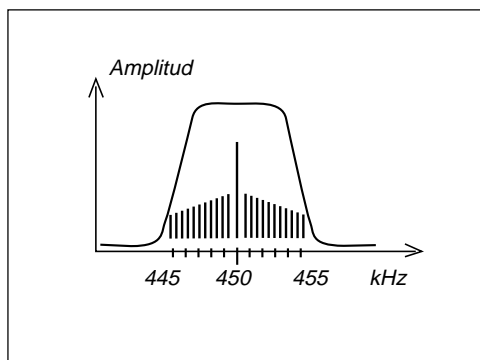
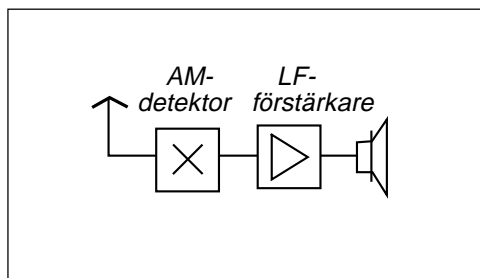
Vid den typ av AM-modulering som används för rundradiosändningar på långvåg (160 – 255 kHz) och mellanvåg (535 – 1605 kHz) så har varje radiosändare fått en frekvenslucka, frekvenskanal, som är 9 kHz bred. Vårt filter skall alltså helst inte vara bredare än 9 kHz.

När man skall tillverka enkla bandpassfilter så är det relativt okomplicerat att åstadkomma ett filter vars bandbredd är 2 % av mittfrekvensen. Ett enkelt bandpassfilter bestående av några resonanskretsar med mittfrekvensen 450 kHz får alltså en bandbredd, kanalbredd, på 9 kHz.

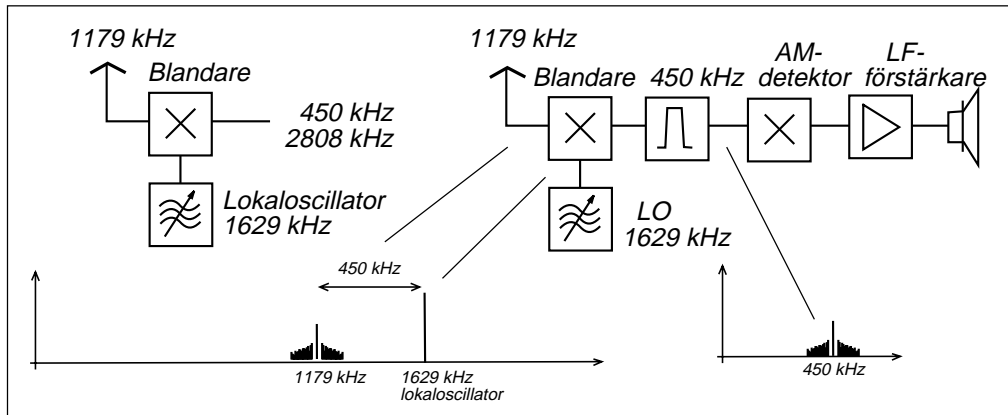
När vi kopplar in detta filter med bandbredden 9 kHz så kan vi alltså lyssna på radiosändningar på 450 kHz. Men vid denna frekvens ligger inga radiosändare!

Kristallmottagaren

Min barndoms radiomottagare för att lyssna på "Mannen i svart" (spökhistorier) från mellanvågssändaren i Göteborg bestod av en resonanskrets, alltså ett enkelt filter, en AM-detektor i form av en likriktardiod (kristalldiod), en kondensator som höjer LF-signalens nivå samt en höghörselhörtelefon (under huvudkudden).



6 — Superheterodyn-mottagaren

**Superheterodynprincipen**

Med hjälp av filtret, AM-detektorn och LF-förstärkaren så har vi en mottagare för 450 kHz. Hur bär vi oss åt för att t.ex. lyssna på mellanvågsstationen i Sölvesborg på 1179 kHz?

Vad gör vi? Vi "flyttar" 1179 kHz till 450 kHz i en blandare av samma typ som används när vi alstrar DSB-signaler.

Vad händer i blandaren? Det bildas summa och skillnad. Till blandarens två ingångar matar vi antenssignalen på ena porten och till andra porten en i mottagaren "lokal" alstrad växelspanning, alstrad i lokaloscillatorn, som vi ställer in på 1629 kHz. I blandaren alstras summa och skillnad mellan antenssignalens alla frekvenskomponenter och lokaloscillatorsignalen.

I detta fall kommer skillnaderna att utgöras av den önskade signalens spektrum centrerat runt 450 kHz. Denna signal kan passera filtret, AM-detekteras, förstärkas

och avlyssnas i högtalaren.

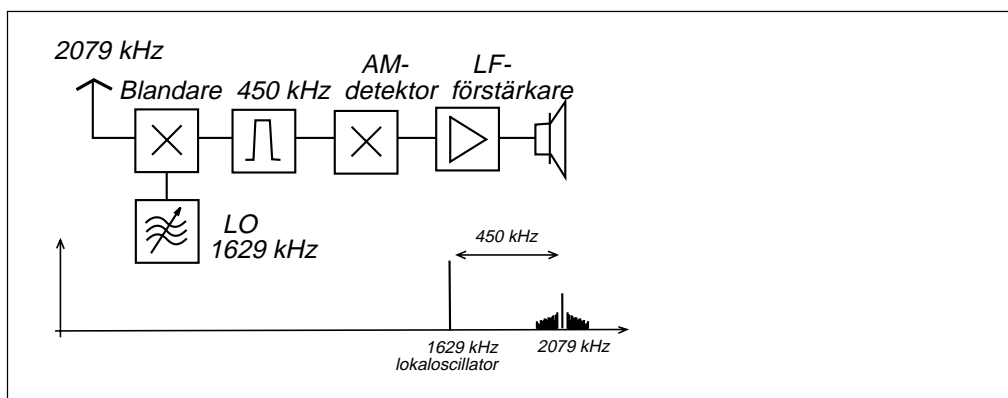
Att på detta sätt bilda skillnadsfrekvenser har fått namnet heterodyn-principen, och en radiomottagare som utnyttjar denna blandningsprincip kallas "superheterodyn-mottagare".

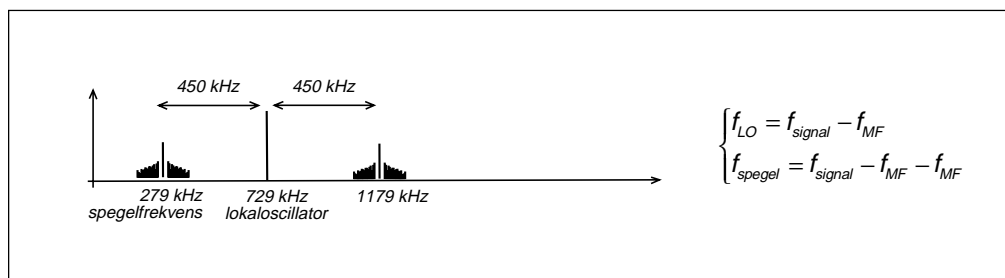
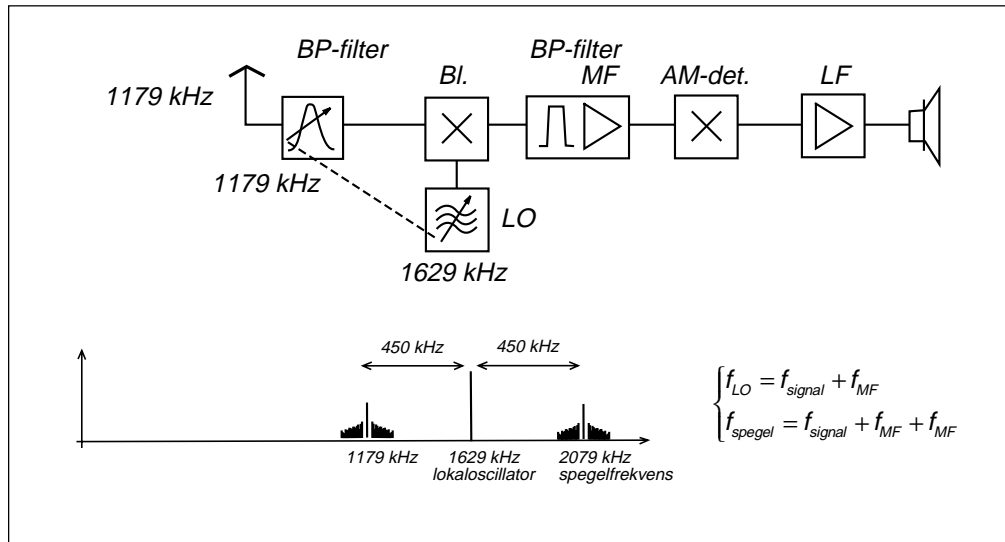
På 1930-talet fanns "raka mottagare", mottagare som hade ett varierbart filter på ingången, som en kristallmottagare med LF-förstärkare. Men vid slutet av 1930-talet kom superheterodyn-mottagaren och den är fortfarande helt dominerande.

Man börjar klara att bygga mottagare med koherent detektor för mottagning av DSB (datasignaler), som har huvuddelen av filtreringen efter detektorn, i form av lågpasfilter vid LF-förstärkaren.

Spegelfrekvensen

Det finns ytterligare en antenssignal som kan passera 450 kHz-filtret, nämligen 2079 kHz!





Spegelfrekvensen

Mellan antennen och blandarsteget måste finnas ett filter som dämpar spegelfrekvensen, ett bandpassfilter avstämt till den önskade radiosignalens frekvens. Kravet på detta filter är inte smal bandbredd utan att det dämpar tillräckligt mycket på spegelfrekvensen. Tillräckligt mycket kan i dessa sammanhang innebära en dämpning på ända upp till -100 dB. Starkast tänkbara signal på spegelfrekvensen skall inte störa svagast tänkbara önskad signal.

Lokaloscillatorfrekvensen

Skillnad innebär att lokaloscillatorfrekvensen minus den önskade antenssignalen, eller den önskade antenssignalen minus lokaloscillatorfrekvensen, skall vara mellanfrekvensen, i detta fall 450 kHz.

LO högre än den önskade signalen

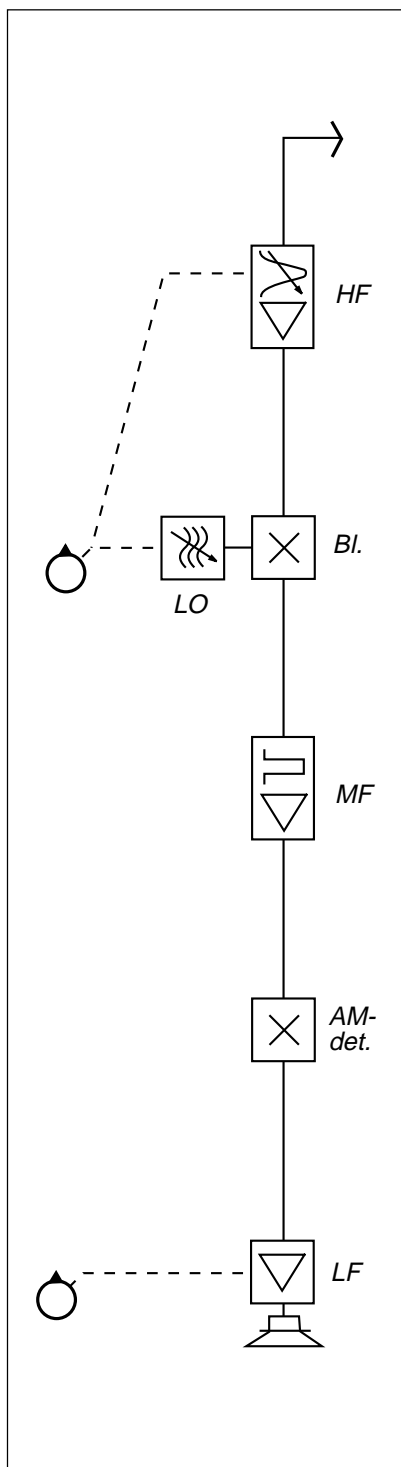
Vi kan alltså välja att lägga lokaloscillatorn högre i frekvens än den önskade antenssignalen. I detta fall kommer spegelfrekvensen att hamna ännu högre i frekvens.

Vanligast är att lokaloscillatorn ligger högre i frekvens.

LO lägre än den önskade signalen

Vi kan även välja att lägga lokaloscillatorn lägre i frekvens än den önskade antenssignalen. I detta fall kommer spegelfrekvensen att hamna ännu lägre i frekvens.

6 — Superheterodynmodtageren



Superheterodynmodtageren

Från antennen kommer antenssignalen. Antennen kan sitta högt belägen. Då har vi en antennkabel mellan antennen och radiomodtageren.

Det är även vanligt att antennen sitter på själva radiomodtageren eller t.o.m. är inbyggd i höljet. Även i dessa fall finns oftast en antennkabel, även om den är kort.

Högfrekvenssteget (HF)

Mellan antennen och blandaren finns filtret för dämpning av spegelfrekvensen. Avstämning till rätt frekvens sker med modtagerens frekvensrätt.

Det är svårt att göra ett blandarsteg med låg brusfaktor. För att förbättra modtagerens känslighet, minska egetbruset, så kan det finnas en transistorförstärkare (HF-steget) med låg brusfaktor.

Blandarsteget (Bl.) och lokaloscillatorn (LO)

Därefter kommer blandarsteget, där antenssignalen blandas med signalen från lokaloscillatorn. Lokaloscillatorns frekvens ändras med modtagerens frekvensrätt. Skillnadsfrekvensen ($HF - LO$ eller $LO - HF$) kallas mellanfrekvensen (MF).

Mellanfrekvenssteget (MF)

I MF-steget sitter det smala filter som ser till att du inte störs av frekvenser utanför frekvenskanalen. Dessutom är det denna bandbredd som begränsar bruset som når fram till detektorn.

Vid mottagning av AM är filtret 6 – 10 kHz brett och vid mottagning av SSB är filtret ca 3 kHz brett. I dessa fall är MF-frekvensen oftast ca 450 kHz.

Vid mottagning av FM-rundradio är filtret ungefär 200 kHz brett och MF-frekvensen 10,7 MHz.

Vid analog tv-mottagning är filtret 5 MHz brett centrerat vid 36,7 MHz. Tv-signalen är VSB utsänd med övre sidbandet, men nedblandad med en lokaloscillator som ligger högre i frekvens.

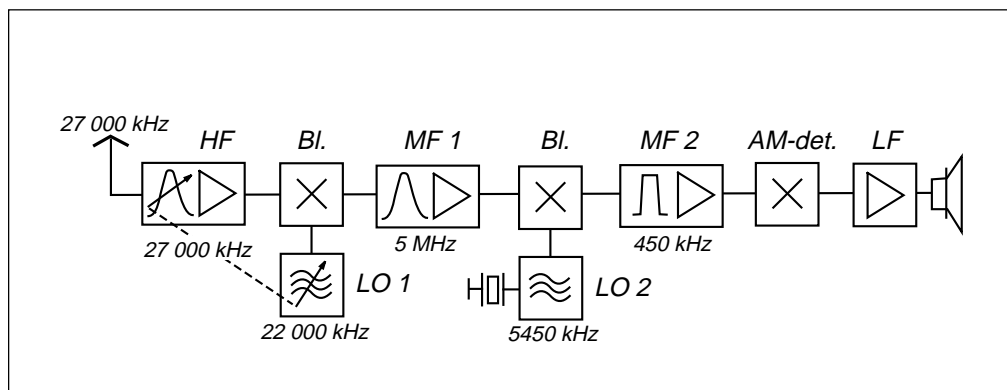
Även GSM-modtageren har ett filter med 200 kHz bandbredd centrerat vid 10,7 MHz.

Detektorn

Detektorn kan vara en AM-detektor, en SSB-detektor eller koherent detektor för DSB med bärvågs-generering (bärvågen skall ligga på MF-frekvensen), en FM-detektor eller någon form av datadetektor.

Lågfrekvensförstärkaren (LF)

I lågfrekvensförstärkaren förstärks informations-signalen till önskad nivå. Här sitter modtagerens volymkontroll.



Dubbelsuperheterodyn

HF-filtret som skall dämpa spegelfrekvensen har en bandbredd som kan vara 5 – 10 % av resonansfrekvensen. Om mellanfrekvensen är 450 kHz så kommer spegelfrekvensen att hamna 900 kHz från den önskade antennsignalen. Det är inte säkert att HF-filtret har tillräcklig dämpning 900 kHz från resonansfrekvensen om signalfrekvensen är hög. Då kan man blanda ner signalfrekvensen i flera steg, i en dubbelsuperheterodynmodtagare.

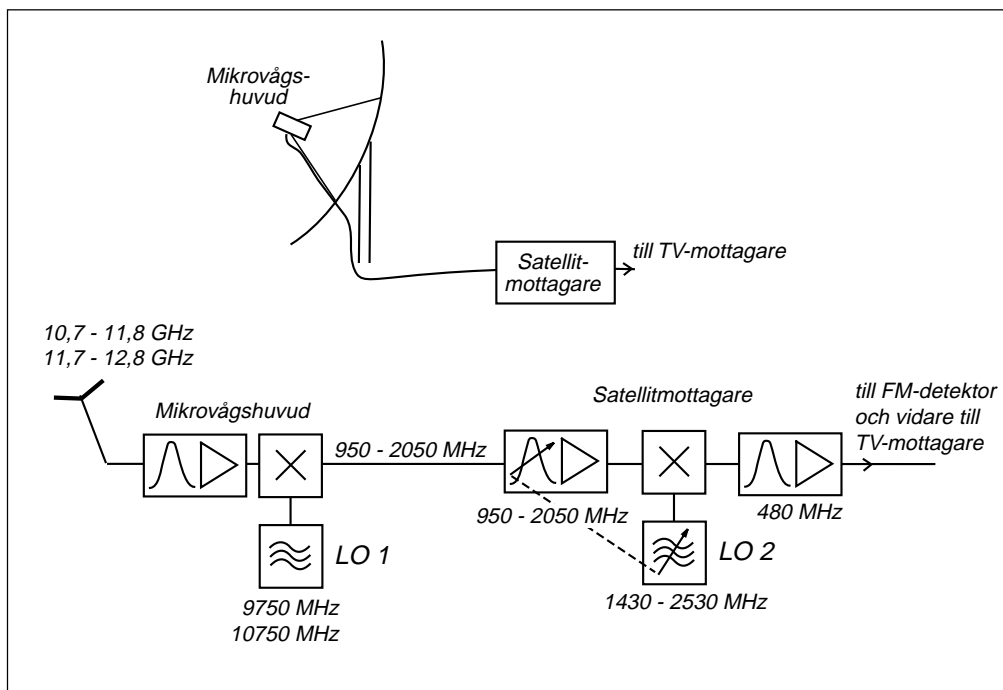
På bilden ovan syns en modtagare som skall ta emot 27 MHz. Genom att först blanda ner 27 MHz-signalen till 5 MHz, första mellanfrekvens (MF 1), så kommer spegelfrekvensen som skall dämpas av HF-filtret att hamna på 17 MHz, tillräckligt långt bort för att dämpas av ett enkelt HF-filter.

Därefter blandas 27 MHz-signalen som nu blivit en 5 MHz-signal med 5450 kHz från andra lokaloscillatorn, och vi får 450 kHz, andra mellanfrekvens (MF 2). I detta fall hamnar spegelfrekvensen på 5900 kHz, en signal som dämpas av MF 1-filtret, om den signalen skulle finnas på utgången av första blandaren.

Första lokaloscillatorn bestämmer modtagarens frekvens

I denna dubbelsuper varierar man första lokaloscillatorn, medan andra lokaloscillatorn ligger på en fast frekvens (5450 kHz).

6 — Superheterodyn-mottagaren

**Satellitmottagning — en dubbelsuper**

Vid satellitmottagning ligger signalfrekvensen i 11 GHz-bandet. Så hög frekvens dämpas kraftigt i koaxialkabel. Dämpningen i kabeln ökar med ökande frekvens. Därför har man placerat HF-förstärkaren och första blandaren med sin lokaloscillator i direkt anslutning till den mottagarantenn som sitter i fokus (mikrovågshuvudet) av parabolen.

Första lokaloscillatorn fast

I detta fall ligger första lokaloscillatorn på en fast frekvens. Detta innebär, om den fasta frekvensen är 9750 MHz, att alla frekvenser inom satellitbandet 10,7 – 11,8 GHz flyttas ner till bandet 950 – 2050 MHz. Detta frekvensband överförs på kabel ner till själva satellitmottagaren, som med en variabel lokaloscillator klarar att ta emot satellitkanalerna och omvandla dessa till 480 MHz mellanfrekvens med 27 MHz bandbredd.

Från satellitmottagaren kan man byta frekvens på LO 1 från 9750 MHz till 10 750 MHz genom att ändra matningsspänningen från 14 V till 16 V. På detta sätt ordnar satellitmottagaren "automatisk" omkoppling så att det går att ta emot alla kanaler inom satellitbandet 10,7 – 12,8 GHz.

Tv-signalen är frekvensmodulerad

När analog tv sänds i "marknätet" är bildsignalen amplitudmodulerad med VSB. Vid analog tv-sändningar via satellit är bildsignalen frekvensmodulerad med en bandbredd på 27 MHz. Tack vare att signalen sprids på en bredare frekvenskanal så klarar man sig med 7 dB signal/brusförhållande. Vid mottagning i analogt marknätet krävs 30 dB signal/brusförhållande för fullgod bildkvalitet.

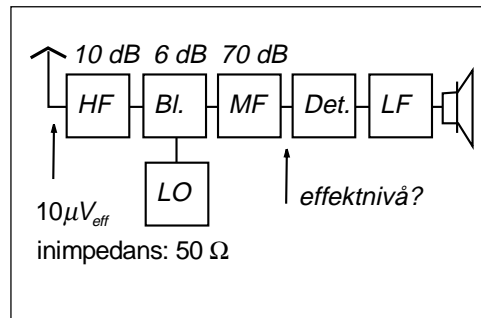
6.2 Decibel och förstärkning

Alltid effektförhållande

I denna bok används decibel ALLTID för att beteckna ett effektförhållande. Förstärkningen G i ett förstärkarsteg innebär alltid förhållandet mellan uteffekten och ineffekten.

Hur hög effekt når fram till detektorn?

Vi har en superheterodynmodulering där HF-steget förstärker 10 dB, blandarsteget förstärker 6 dB och MF-steget (flera transistorer) förstärker 70 dB. Om den önskade signalen har nivån $10 \mu V_{eff}$ på antenkabeln och HF-stegets inimpedans är 50Ω , hur hög effekt kommer då fram till detektorn?

**Det finns en "säker" metod:**

Beräkna insignalens effekt och multiplicera sedan med de olika stegens förstärkning.

$$P_{in} = \frac{(10 \cdot 10^{-6})^2}{50} = 2 \cdot 10^{-12} \text{ W}$$

$$P_{Detektor} = 2 \cdot 10^{-12} \cdot 10^{\frac{10}{10}} \cdot 10^{\frac{6}{10}} \cdot 10^{\frac{70}{10}} = 0,8 \text{ mW}$$

dBm och dBW

Decibel uttrycker ett effektförhållande, men kan även referera till en känd effektnivå. Det finns två referensnivåer:

- Om effekten anges i mW så är enheten dBm.
- Om effekten anges i W så är enheten dBW.

$$P_{Detektor} = 0,8 \text{ mW} \Rightarrow 10 \cdot \log(0,8 \text{ mW}) = -1 \text{ dBm}$$

$$P_{Detektor} = 0,0008 \text{ W} \Rightarrow 10 \cdot \log(0,0008 \text{ W}) = -31 \text{ dBW}$$

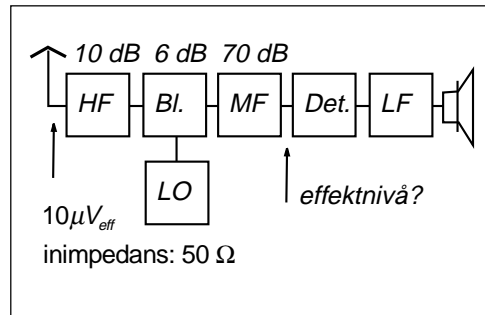
Enda skillnaden mellan dBm och dBW är att sifvervärdet för dBm är +30 dB högre.

6 — Superheterodynmodtagaren

En "elegant" metod

Om det finns en säker metod så borde det även finnas en elegant metod?

När man blir van att räkna med dB så är enklaste metoden att beräkna insignalens effekt i dBm och sedan addera förstärkningarna och subtrahera dämpningarna uttryckta i dB.



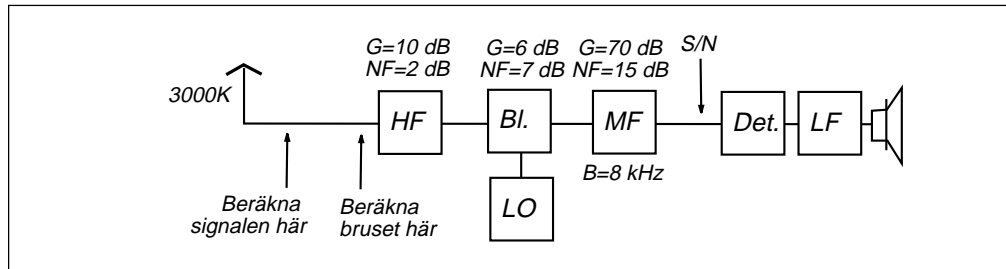
$$P_{in} [dBW] = 10 \cdot \log \frac{(10 \cdot 10^{-6})^2}{50} = -117 \text{ dBW}$$

$$P_{Detektor} [dBW] = -117 + 10 + 6 + 70 = -31 \text{ dBW}$$

Så här räknar jag:

$$\begin{aligned} P_{in} [dBW] &= 10 \cdot \log \frac{(10 \cdot 10^{-6})^2}{50} = \\ &= 10 \cdot \log(10 \cdot 10^{-6})^2 - 10 \cdot \log 50 = \\ &= -100 - 17 = \\ &= -117 \text{ dBW} \end{aligned}$$

6.3 Signal/brusforholdet før detektoren

**Vi har beräknat signalens nivå**

Signalens nivå vid mottagarens antenningång har redan beräknats, både som effekt i W och uttryckt i dBm och dBW.

$$P_{in} = 2 \cdot 10^{-12} \text{ W}$$

$$P_{in} = -117 \text{ dBW}$$

$$P_{in} = -117 + 30 = -87 \text{ dBm}$$

Beräkna mottagarens brustemperatur

Visst kan man utnyttja formeln för brusfaktor när man har flera seriekopplade förstärkare. Själv tycker jag att det är lika enkelt att beräkna brustemperaturbidraget från varje steg för sig (vilket är exakt vad formeln gör).

$$T_{antenn} = 3000 \text{ K}$$

$$T_{HF} = T_0(10^{\frac{2}{10}} - 1) = 169,6 \text{ K}$$

$$T_{Bl} = \frac{T_0(10^{\frac{6}{10}} - 1)}{10^{\frac{10}{10}}} = 86,5 \text{ K}$$

$$T_{MF} = \frac{T_0(10^{\frac{15}{10}} - 1)}{10^{\frac{10}{10}} \cdot 10^{\frac{6}{10}}} = 222 \text{ K}$$

Brustemperatur från antennen

I detta fall fångar antennen upp atmosfärbrus av sådan nivå att antennen ser ut som en brusälla med brustemperaturen 3000K.

$$N_{in} = k \cdot (T_{Antenn} + T_{HF} + T_{Bl} + T_{MF}) \cdot B =$$

$$= k \cdot (3000 + 169,6 + 86,5 + 222) \cdot B =$$

$$= 1,38 \cdot 10^{-23} \cdot 3478,1 \cdot 9000 =$$

$$= 0,43 \cdot 10^{-15} \text{ W}$$

Brustemperatur från HF-steget

HF-steget har brusfaktorn 2 dB och detta brus finns på "rätt ställe", på ingången till HF-steget.

$$N_{in} [\text{dBW}] = 10 \cdot \log(0,43 \cdot 10^{-15} \text{ W}) =$$

$$= -153,7 \text{ dBW} = -123,7 \text{ dBm}$$

Brustemperatur från blandarsteget

Blandarsteget har brusfaktorn 6 dB men detta brus måste dämpas med HF-stegets förstärkning för att hamna på ingången till HF-steget.

Brustemperatur från MF-steget

MF-steget har brusfaktorn 15 dB och detta brus måste dämpas med både blandarstegets och HF-stegets förstärkning för att hamna på ingången till HF-steget.

$$N_{in} = 0,43 \cdot 10^{-15} \text{ W}$$

$$N_{in} = -153,7 \text{ dBW}$$

$$N_{in} = -123,7 \text{ dBm}$$

Nu känner vi signaleffekten på ingången och bruseffekten på ingången. Då är det bara att beräkna signal/brusförhållandet.

Vi vill ha signal/brusförhållandet uttryckt i dB.

- Utgå från effekterna. Beräkna förhållandet i ggr och omvandla till dB.

$$\frac{S}{N} = \frac{P_{in}}{N_{in}} = 10 \cdot \log \frac{2 \cdot 10^{-12} \text{ W}}{0,43 \cdot 10^{-15} \text{ W}} = 36,68 \text{ dB}$$

- Utgå från effekterna uttryckta i dBW. Beräkna skillnaden.

$$\frac{S}{N} = P_{in} [\text{dBW}] - N_{in} [\text{dBW}] = -117 - (-153,7) = 36,7 \text{ dB}$$

- Utgå från effekterna i dBm. Beräkna skillnaden.

$$\frac{S}{N} = P_{in} [\text{dBm}] - N_{in} [\text{dBm}] = -87 - (-123,7) = 36,7 \text{ dB}$$

6 — Superheterodynmodtageren

17 lektioner i
TELEKOMMUNIKATION

7 — Frekvensmodulering och fasmodulering

Frekvensmodulering innebär att bärvågens frekvens ändras i takt med informations-signalen.

Fasmodulering innebär att bärvågens fasfördröjning ändras i takt med informations-signalen.

Om fasfördröjningen ändras (om spänningsvisaren ändrar fasvinkel) så uppstår en frekvensändring vid själva ändringen. Fasmodulering är därför samma sak som frekvensmodulering av informationssignalens ändringar, d.v.s. informationssignalens derivata.

Vi tittar på frekvensmodulering i tidsplanet samt hur FM-modulering hanteras matematiskt och kommer fram till FM-signalens spektrum, där vi behöver tabeller eller kurvor över Besselfunktionerna för att få fram spektralkomponenternas amplituder.

Ur spektrat får vi FM-signalens bandbredd, som även kan beräknas med Carsons regel.

Varför bildas så många spektralkomponenter? När vi förstår vad detta innebär fysikaliskt så studerar vi vad som händer när en bandbegränsad FM-signal matas genom en limiter.

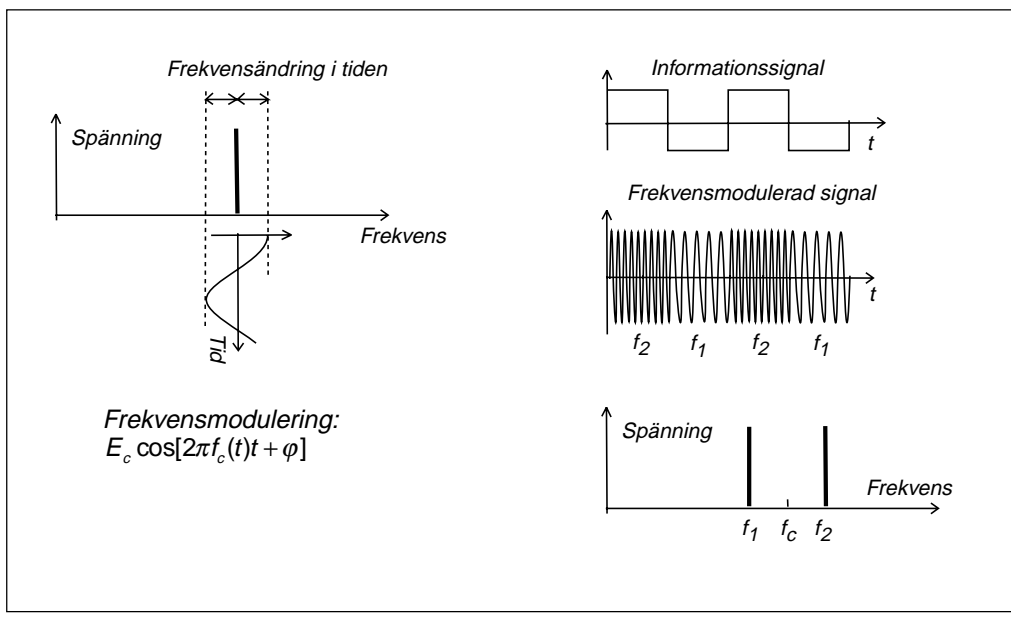
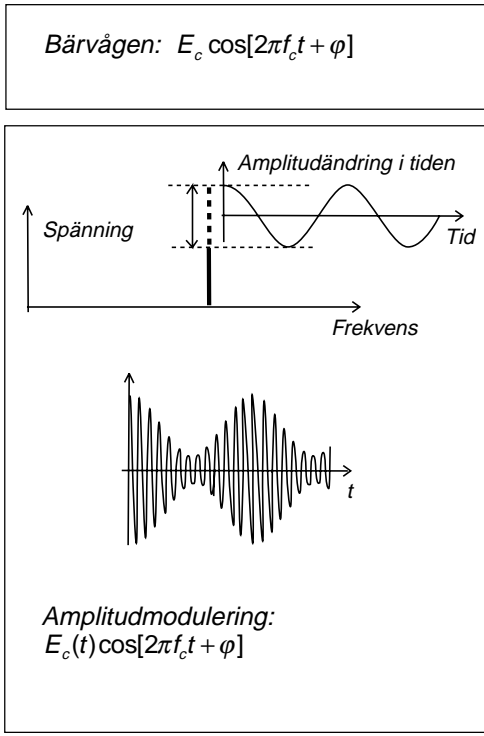
7 — Frekvensmodulering och fasmodulering

7.1 Argumentmodulering

Hittills har vi studerat olika former av amplitudmodulering, där vi påverkar bärvågens amplitud E_c

- Vid AM låter vi bärvågens amplitud variera i takt med informationssignalen.
- Vid VSB sänder vi bara "halva" AM-signalens spektrum.
- Vid DSB skapar vi en AM-signal men undertrycker bärvågen.
- Vid SSB sänder vi bara "halva" DSB-signalens spektrum.
- Vid QAM sänder vi två DSB-signaler 90° i fas.

I inget av fallen gör vi någon medveten förändring av bärvågens frekvens f_c eller fasvinkel φ . Men när vi filtrerar bort halva spektrat så som vi gör vid VSB och SSB så resulterar detta i att signalens spänningsvektor, förutom att variera i amplitud, dessutom får en frekvens/fasvariation.



7 — Frekvensmodulering och fasmodulering

Argumentmodulering

Argumentmodulering innebär att inte röra bärvågens amplitud E_c utan bara påverka bärvågens argument. Detta kan göras på två sätt:

- Frekvensmodulering, där vi ändrar frekvensen f_c i takt med informationen.
- Fasmodulering, där vi adderar en fasvinkel φ som är proportionell mot informationen.

Frekvensmodulering

Vid frekvensmodulering rör vi inte bärvågens amplitud utan låter dess frekvens f_c ändras proportionellt mot informations-signalen.

Om informationssignalen är en fyrkantvåg så kommer FM-signalen att växla mellan två olika frekvenser, se bilden.

Fasmodulering

Det andra sättet att argumentmodulera, att modulera fasvinkeln φ , innebär att bärvågs-signalens frekvens inte ändras, utan dess fasvinkel φ . Men fasvinkeln talar om när signalen kommer fram i tiden. Fasmodulering innebär alltså att bärvågen kommer fram lite tidigare (positiv fasvinkel) eller lite senare (negativ fasvinkel).

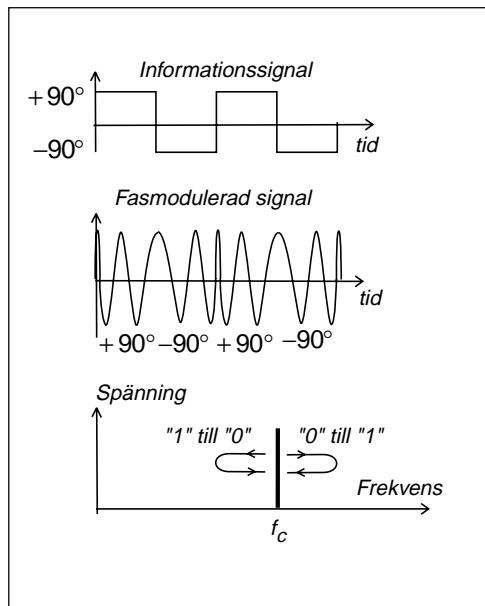
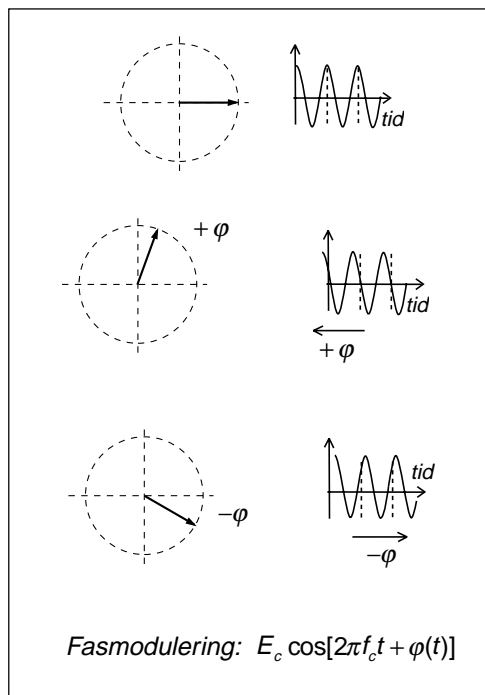
Bilden visar bärvågen som fasmoduleras. När fasvinkeln ändras från positiv fasvinkel till negativ fasvinkel så dras cosinus-kurvan isär. Visaren roterar momentant långsammare än f_c . Detta innebär en minskning av frekvensen.

När fasvinkeln ändras från negativ till positiv fasvinkel roterar visaren momentant snabbare och cosinuskurvan trycks ihop. Vi får en höjning av frekvensen.

Vid fasmodulering är ändringarna frekvensmodulerade

Vi ser att vid fasmodulering är informations-signalens ändringar frekvensmodulerade. Man kan därför derivera informations-signalen, d.v.s. mata informationssignalen genom ett högpasfilter och låta den styra en frekvensmodulator. Resultatet blir en fasmodulerad signal.

På samma sätt kan vi integrera informationssignalen genom att mata den genom



ett lågpasfilter varefter den får styra en fasmodulator. Resultatet blir en frekvensmodulerad signal.

7.2 Frekvensmodulering i tidsplanet

Resonanskretsen

Vi utgår från resonanskretsen. Resonansfrekvensen bestäms av spolens induktans och kondensatorns kapacitans.

Använd en variabel kondensator, t.ex. en vridkondensator. Vrid fram och tillbaka. Då ändras resonanskretsens resonansfrekvens i takt med "informationen", i takt med vridningen.

Låt resonanskretsen vara det frekvensbestämmande elementet i en oscillator, t.ex. en signalgenerator. Vrid fram och tillbaka på frekvensratten. Då blir utsignalen frekvensmodulerad.

Spänningsstyrd oscillator, VCO

Byt vridkondensatorn mot en spänningsstyrd kapacitansdiod (varicap). Då kan växelspanningens frekvens ändras med en spänning som styr kapacitansen i kapacitansdioden.

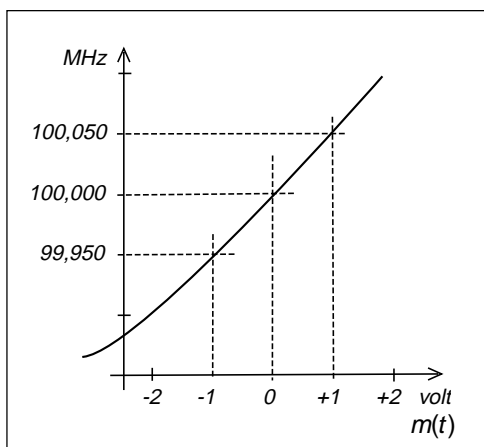
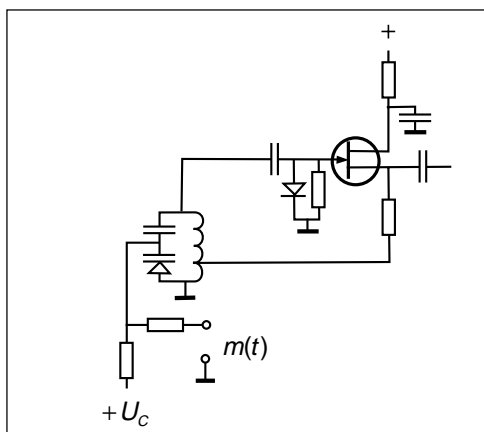
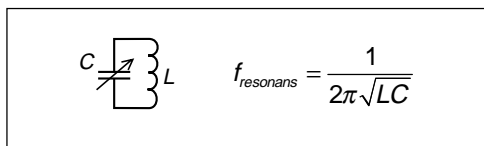
Kapacitansdioden måste ha backspänning samtidigt som den frekvensbestämmande spänningen skall kunna variera runt nollvärdet. Därför matas kapacitansdioden via en spänningsdelare där spänningen U_c har valts så att frekvensen blir 100 MHz när $m(t)$ har värdet 0 V.

En oscillator vars frekvens påverkas av en styrspanning kallas VCO (Voltage Controlled Oscillator).

Man kan rita ett diagram över hur frekvensen ändras med styrspanningen. När styrspanningen har värdet 0 V lämnar vår oscillator 100 MHz.

När styrspanningen ökar till +1 V så ökar backspänningen över kapacitansdioden och kapacitansen minskar. Frekvensen ökar till 100,050 MHz.

När styrspanningen minskar till -1 V minskar frekvensen till 99,950 MHz.



Vi kan beräkna k_f

Ur diagrammet kan vi beräkna frekvensändring per volt styrspanning:

$$k_f = \frac{100,050 - 99,950}{+1 - (-1)} = \frac{0,1}{2} = 50 \text{ kHz/V}$$

7 — Frekvensmodulering och fasmodulering

Nu tänker vi oss att $m(t)$ varierar från + 0,8 volt till - 0,8 volt. Detta innebär att den frekvensmodulerade bärsvågen kommer att svepa fram och tillbaka i frekvens, mellan yttervärdena 99,960 och 100,040 MHz.

Deviation

Maximala frekvensavvikelsen kallas FM-signalens deviation Δf .

$$\Delta f = k_f \cdot m(t)_{\max} = 0,050 \cdot 0,8 = 0,040 \text{ MHz}$$

INTE FM-signalens bandbredd

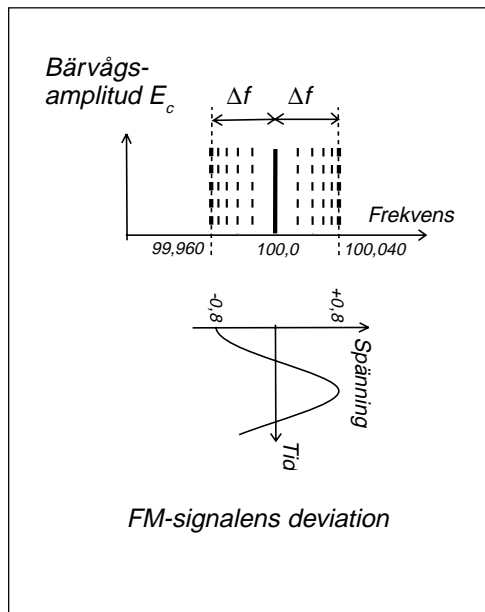
OBSERVERA att vi inte ritat FM-signalens spektrum. Bilden visar vad bärsvågen gör i tidsplanet.

På samma sätt som AM-signalen har en bärsvåg med konstant frekvens som varierar i amplitud, så har FM-signalen en bärsvåg med konstant amplitud som varierar i frekvens.

Bandbredden beror både på deviationen och informationssignalens frekvens

AM-signalens bandbredd påverkas av hur snabbt bärsvågen varierar i amplitud, och detta beror på informationssignalens frekvens.

På samma sätt får FM-signalen en bandbredd som överstiger $2\Delta f$. Hur mycket större bandbredden blir beror på hur snabbt FM-signalen varierar i frekvens, och detta beror på informationssignalens frekvens.



7.3 Frekvensmodulering och matematiken

Vid amplitudmodulering finns ett linjärt samband mellan informationssignalens spektrum och den amplitudmodulerade signalens spektrum. Därför är det relativt okomplicerat att ta sig från informationssignalens spektrum till den amplitudmodulerade signalens spektrum, antingen genom att betrakta en spektralkomponent i taget, så som vi gjort, eller med transformräkning via Fouriertransformer och faltning.

Vid argumentmodulering råder ett icke-linjärt förhållande mellan informationssignalens spektrum och den argumentmodulerade signalens spektrum. Därför finns inget enkelt sätt att beräkna den argumentmodulerade signalens spektrum vid valfri informationssignal $m(t)$. Vi håller oss därför till specialfallet att informationssignalen är en enkel sinussignal (1). Men även i detta fall blir beräkningarna komplicerade. Liksom vid amplitudmodulering så håller vi oss genomgående till cosinusfunktioner.

Bärvågen $s(t)$ har som vanligt amplituden E_c och växelspanningen beskrivs med cosinusfunktionen (2).

Vid fasmodulering låter vi fasvinkeln variera i takt med $m(t)$. Vi inför proportionalitetskonstanten k_p (phase modulation) och får $s_{PM}(t)$ som är den fasmodulerade signalen som funktion av tiden (3).

På motsvarande sätt är $s_{FM}(t)$ den frekvensmodulerade signalen som funktion av tiden (4).

Vid frekvensmodulering är frekvensen en funktion av tiden, som i sin tur är en funktion av tiden (multiplieras med t) som därefter utgör argument i en cosinusfunktion. Detta är svårt att hantera matematiskt. Därför gör vi på följande sätt:

Tidsderivatan av argumentet θ är ögonblicksvärdet av vinkelfrekvensen ω , som i sin tur är 2π multiplicerat med frekvensen (5).

Förläng uttrycket med dt och häng på

integraltecken och integrera, så får vi argumentet θ_{FM} för FM-modulering.

Nu har vi fått ett uttryck för den frekvensmodulerade signalen $s_{FM}(t)$ som egentligen är fasmodulering med informationssignalens integral (6).

Samtidigt inför vi en cosinusformad informationssignal och löser integralen (7). Vid integreringen får vi en integrationskonstant θ_k som vi inte behöver bry oss om. En konstant fasvinkel innebär bara att spänningsvisaren inte "började" snurra från riktningen 0° i visardiagrammet.

I uttryck (8) identifierar vi FM-signalens deviation d.v.s. den maximala frekvensavvikelsen hos frekvensens ögonblicksvärde, samt inför modulationsindex m_f för frekvensmodulering (8).

Dela isär argumentet, se trigonometriska formelsamlingen, så får vi uttryck (9).

I uttryck (9) finns två termer som kan skrivas som oändliga trigonometriska serier med amplituder som i sin tur är oändliga serier, vars summa är de s.k. Bessel-funktionerna av n :te ordningen med modulationsindex som argument (10).

Utnyttja detta så får vi (11) som är spektrum för en bärvåg med amplituden E_c som frekvensmoduleras av en cosinusfunktion med frekvensen f_m på sådant sätt att modulationsindex blir m_f . Vi har även att deviationen blir:

$$\Delta f = m_f \cdot f_m$$

$$1) \quad m(t) = E_m \cos 2\pi f_m t$$

$$2) \quad s(t) = E_c \cos(\theta)$$

$$3) \quad s_{PM}(t) = E_c \cos[2\pi f_c t + k_p \cdot m(t)]$$

$$4) \quad s_{FM}(t) = E_c \cos[2\pi(f_c + k_f \cdot m(t))t]$$

$$5) \quad \frac{d(\theta_{FM})}{dt} = \omega_i = 2\pi[f_c + k_f \cdot m(t)] = 2\pi f_c + 2\pi k_f \cdot m(t)$$

$$d(\theta_{FM}) = \omega_i dt = 2\pi f_c dt + 2\pi k_f \cdot m(t) dt$$

$$\int_{-\infty}^{\theta} d(\theta_{FM}) = \int_{-\infty}^t \omega_i dt = 2\pi f_c \int_{-\infty}^t dt + 2\pi k_f \int_{-\infty}^t m(t) dt$$

$$\theta_{FM} = \omega_i t = 2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(t) dt$$

$$6) \quad \begin{cases} s_{FM}(t) = E_c \cos \left[2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(t) dt \right] \\ m(t) = E_m \cos 2\pi f_m t \end{cases}$$

$$7) \quad \begin{aligned} s_{FM}(t) &= E_c \cos [2\pi[f_c + k_f \cdot E_m \cos(2\pi f_m t)]t] = \\ &= E_c \cos \left[2\pi f_c t + 2\pi k_f \int_{-\infty}^t E_m \cos(2\pi f_m t) dt \right] = \\ &= E_c \cos \left[2\pi f_c t + \frac{2\pi k_f \cdot E_m}{2\pi f_m} \sin 2\pi f_m t + \theta_k \right] \end{aligned}$$

$$8) \quad \begin{cases} \Delta f = k_f \cdot E_m \\ m_f = \frac{\Delta f}{f_m} \end{cases} \Rightarrow s_{FM}(t) = E_c \cos [2\pi f_c t + m_f \sin 2\pi f_m t]$$

$$9) \quad s_{FM}(t) = E_c \cdot \cos 2\pi f_c t \cdot \cos [m_f \sin 2\pi f_m t] - E_c \cdot \sin 2\pi f_c t \cdot \sin [m_f \sin 2\pi f_m t]$$

$$10) \quad \begin{cases} \cos [m_f \sin 2\pi f_m t] = J_0(m_f) + J_2(m_f) \cdot \cos 4\pi f_m t + J_4(m_f) \cdot \cos 8\pi f_m t + \dots \\ \sin [m_f \sin 2\pi f_m t] = J_1(m_f) \cdot \sin 2\pi f_m t + J_3(m_f) \cdot \cos 6\pi f_m t + \dots \end{cases}$$

$$11) \quad \begin{aligned} s_{FM}(t) &= E_c \cdot J_0(m_f) \cdot \cos 2\pi f_c t + \\ &+ E_c \cdot J_1(m_f) \cdot [\cos 2\pi(f_c + f_m)t - \cos 2\pi(f_c - f_m)t] + \\ &+ E_c \cdot J_2(m_f) \cdot [\cos 2\pi(f_c + 2f_m)t + \cos 2\pi(f_c - 2f_m)t] + \\ &+ E_c \cdot J_3(m_f) \cdot [\cos 2\pi(f_c + 3f_m)t - \cos 2\pi(f_c - 3f_m)t] + \\ &+ E_c \cdot J_4(m_f) \cdot [\cos 2\pi(f_c + 4f_m)t + \cos 2\pi(f_c - 4f_m)t] + \\ &+ E_c \cdot J_5(m_f) \cdot [\cos 2\pi(f_c + 5f_m)t - \cos 2\pi(f_c - 5f_m)t] + \dots \end{aligned}$$

7.4 Frekvensmodulerade signalens spektrum

Frekvensmodulerade signalens spektrum framgår av ekvation (11) på föregående sida. Låt oss titta på ett exempel.

Bärvågen $E_c = 30 \cos(2\pi 10^8 t)$ skall frekvensmoduleras. Man har funnit att en likspänning på 2 volt ger 12 kHz frekvensändring. Rita amplitudspektrum om modulationssignalen (informationssignalen) är $m(t) = 5 \cos(2\pi \cdot 10^4 t)$.

Börja med omodulerade bärvågens frekvens: $f_c = 100$ MHz
Detta är spektrats mittpunkt.

Bärvågens amplitud: $E_c = 30$ V (toppvärde).
 E_c behövs när frekvenskomponenternas amplituder skall beräknas.

Nästa steg är informationssignalens frekvens: $f_m = 10$ kHz
Detta är frekvensdelningen mellan spektralkomponenterna (på samma sätt som vid AM).

Nu har vi allt utom frekvenskomponenternas amplituder som är $E_c \cdot |J_n(m_f)|$.
Värdet på Besselfunktionerna hämtar vi ur diagrammet, men vi behöver veta var vi skall avläsa. Vi behöver m_f .

$$m_f = \frac{\Delta f}{f_m} = \frac{k_f \cdot |m(t)|_{\max}}{f_m}$$

Modulationsfrekvensen (informationssignalens frekvens) f_m känner vi. Men vi känner inte deviationen Δf .

Däremot känner vi informationssignalens maxvärde $|m(t)|_{\max} = 5$ volt (den spänning som ger maximal frekvensavvikelse, maximal deviation) och vi vet att 2 volt ger 12 kHz frekvensavvikelse, d.v.s.

$$k_f = \frac{12 \text{ kHz}}{2 \text{ V}} = 6 \text{ kHz/V}$$

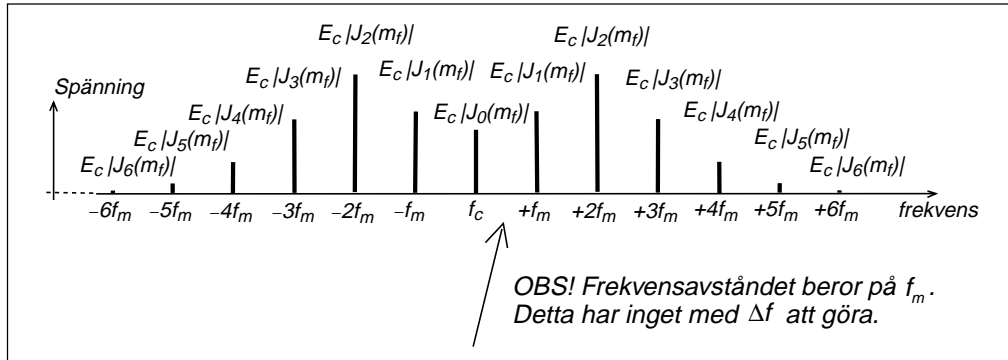
Ur detta får vi:

$$m_f = \frac{\Delta f}{f_m} = \frac{k_f \cdot |m(t)|_{\max}}{f_m} = \frac{6 \cdot 5}{10} = 3$$

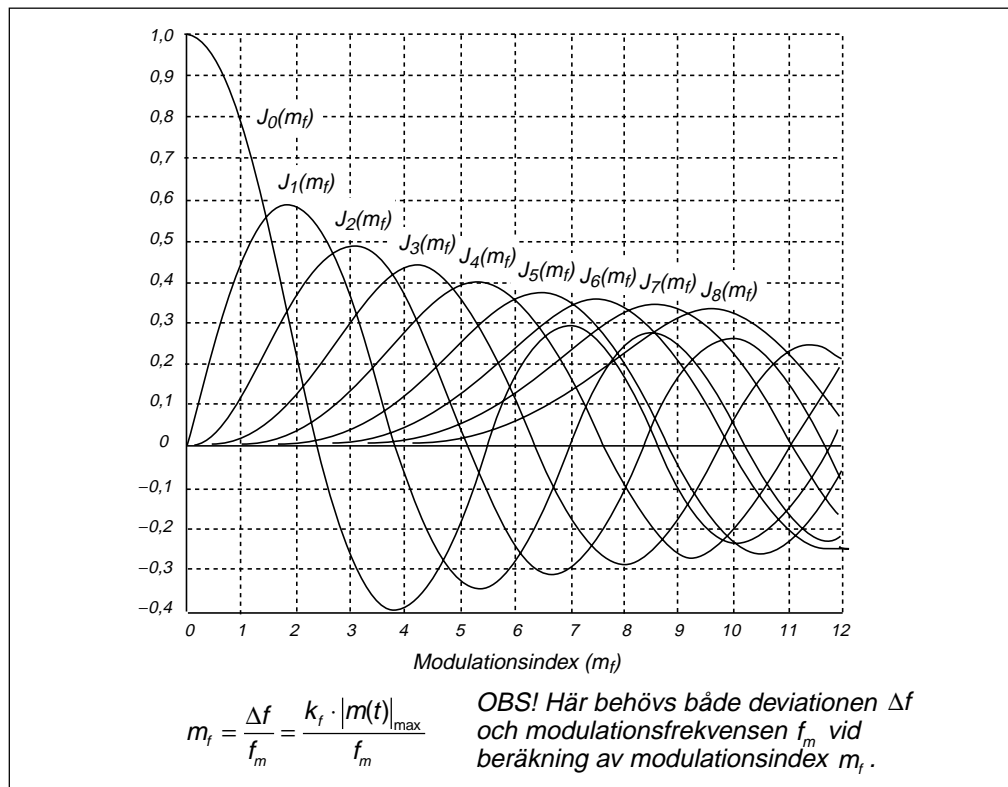
Nu avläser vi ur diagrammet och beräknar frekvenskomponenternas amplituder.

VIKTIGT: Vid AM fick vi halva amplituden på den övre frekvensen och halva på den undre frekvensen. Så är det inte vid FM. $E_c \cdot |J_n(m_f)|$ är amplituden på spektralkomponenten!

7 — Frekvensmodulering och fasmodulering



Frekvens MHz	Amplitud	Frekvens MHz	Amplitud
100,000	$E_c \cdot J_0(m_f) = 30 \cdot 0,26 = 7,8 \text{ V}$	99,990	$E_c \cdot J_1(m_f) = 30 \cdot 0,34 = 10,2 \text{ V}$
100,010	$E_c \cdot J_1(m_f) = 30 \cdot 0,34 = 10,2 \text{ V}$	99,980	$E_c \cdot J_2(m_f) = 30 \cdot 0,49 = 14,7 \text{ V}$
100,020	$E_c \cdot J_2(m_f) = 30 \cdot 0,49 = 14,7 \text{ V}$	99,970	$E_c \cdot J_3(m_f) = 30 \cdot 0,31 = 9,3 \text{ V}$
100,030	$E_c \cdot J_3(m_f) = 30 \cdot 0,31 = 9,3 \text{ V}$	99,960	$E_c \cdot J_4(m_f) = 30 \cdot 0,13 = 3,9 \text{ V}$
100,040	$E_c \cdot J_4(m_f) = 30 \cdot 0,13 = 3,9 \text{ V}$	99,950	$E_c \cdot J_5(m_f) = 30 \cdot 0,04 = 1,2 \text{ V}$
100,050	$E_c \cdot J_5(m_f) = 30 \cdot 0,04 = 1,2 \text{ V}$	99,940	$E_c \cdot J_6(m_f) = 30 \cdot 0,01 = 0,3 \text{ V}$
100,060	$E_c \cdot J_6(m_f) = 30 \cdot 0,01 = 0,3 \text{ V}$		



7.5 Frekvensmodulerade signalens bandbredd

Vi har beräknat amplitudspektrat för en signal som är frekvensmodulerad av en cosinussignal. Men det matematiska uttrycket är identiskt med en signal som är fasmodulerad av en sinussignal.

Av detta kan vi dra slutsatsen att vid modulering med sinusformad signal, så blir amplitudspektrat identiskt för frekvens- och fasmodulering, om modulationsindex väljs så att $m_f = m_p$.

Hur många spektralkomponenter?

Hur stor bandbredd behövs för att överföra en frekvensmodulerad signal? Det finns ju oändligt antal spektralkomponenter. Hur många skall man ta med?

En tumregel har blivit att man tar med så många att man inte förlorar mer än någon eller några procent av signalens effekt.

FM-signalens effekt

Hur beräknar vi FM-signalens effekt? Det enklaste är att se FM-signalen i tidsplanet. Där har vi en bärvåg med konstant amplitud som sveper Δf fram och Δf tillbaka. FM-signalens effekt är helt enkelt effekten i bärvågen.

$$s_{FM}(t) = E_c \cos[2\pi f_c t + m_f \sin 2\pi f_m t]$$

$$\begin{cases} s_{PM}(t) = E_c \cos[2\pi f_c t + k_p \cdot m(t)] \\ m(t) = E_m \sin 2\pi f_m t \end{cases}$$

$$s_{PM}(t) = E_c \cos[2\pi f_c t + k_p \cdot E_m \sin 2\pi f_m t] = E_c \cos[2\pi f_c t + m_p \sin 2\pi f_m t]$$

$$P_{FM} = \frac{\left(\frac{E_c}{\sqrt{2}}\right)^2}{R} = \frac{E_c^2}{2 \cdot R} = \frac{30^2}{2 \cdot R}$$

Summera effekten i spektralkomponenterna

När FM-modulering betraktas i frekvensplanet, som spektrum, så har bärvågen delat upp sig på spektralkomponenter. Men effekten i bärvågen skall vara densamma som summan av effekterna i spektralkomponenterna. Vi skall se om detta stämmer.

$$P_{FM} = \frac{\left(\frac{E_c \cdot J_0(m_f)}{\sqrt{2}}\right)^2}{R} + 2 \cdot \frac{\left(\frac{E_c \cdot J_1(m_f)}{\sqrt{2}}\right)^2}{R} + 2 \cdot \frac{\left(\frac{E_c \cdot J_2(m_f)}{\sqrt{2}}\right)^2}{R} + 2 \cdot \frac{\left(\frac{E_c \cdot J_3(m_f)}{\sqrt{2}}\right)^2}{R} + 2 \cdot \frac{\left(\frac{E_c \cdot J_4(m_f)}{\sqrt{2}}\right)^2}{R} + \dots = \frac{E_c^2}{2 \cdot R} (J_0^2(m_f) + 2 \cdot J_1^2(m_f) + 2 \cdot J_2^2(m_f) + 2 \cdot J_3^2(m_f) + 2 \cdot J_4^2(m_f) + \dots)$$

7 — Frekvensmodulering och fasmodulering

Andel av effekten i varje sidbandspar	Summa
$ J_0(m_f) ^2 = 0,260^2 = 0,0676$	0,0676
$+ 2 \cdot J_1(m_f) ^2 = 2 \cdot 0,339^2 = 0,2298$	0,2974
$+ 2 \cdot J_2(m_f) ^2 = 2 \cdot 0,486^2 = 0,4724$	0,7698
$+ 2 \cdot J_3(m_f) ^2 = 2 \cdot 0,309^2 = 0,1910$	0,9608
$+ 2 \cdot J_4(m_f) ^2 = 2 \cdot 0,132^2 = 0,0348$	0,9956
$+ 2 \cdot J_5(m_f) ^2 = 2 \cdot 0,043^2 = 0,0037$	0,9993
$+ 2 \cdot J_6(m_f) ^2 = 2 \cdot 0,011^2 = 0,00024$	0,9995

Vi ser i tabellen hur summan av sidbandsparen börjar närma sig 100 %. För att denna beräkning skall vara någorlunda rättvisande så krävs noggranna siffervärden på Bessel-funktionerna.

Om vi tar med mittkomponenten och fyra sidbandspar så får vi 99,6 % av signalens totala effekt. Då blir totala bandbredden $B = 8 \cdot f_m = 80$ kHz.

Hur många sidbandspar som måste tas med beror på modulationsindex m_f , som ju skiljer från fall till fall.

Carsons regel

Det finns ett förenklat sätt att beräkna erforderlig bandbredd som går ut på att man först behöver bandbredd för att få igenom deviationen, d.v.s. $2 \cdot \Delta f$. Men det räcker inte med detta. Då gör man som vid AM, man lägger till bandbredd motsvarande högsta modulationsfrekvensen på båda sidor.

Detta sätt att beräkna bandbredden har fått namnet Carsons regel. Det är inte en matematiskt härledd lag utan en tumregel. Vanligtvis får man med 98 – 99 % av signalens effekt inom denna bandbredd.

Carsons regel:

$$B = 2 \cdot (\Delta f + f_m) = 2 \cdot (30 + 10) = 80 \text{ kHz}$$

7 — Frekvensmodulering och fasmodulering

7.6 Varför så många spektralkomponenter?

Varför innehåller den FM-modulerade signalens spektrum oändligt antal sidbandspar när den AM-modulerade signalen bara innehåller ett sidbandspar när man modulerar med en sinusformad signal? För att besvara den frågan så måste vi titta på den FM-modulerade signalen som spänningsvisare.

Detta är AM-modulering. Vi har bärvågen och två sidband, där sidbandet med högre frekvens snurrar motsols och sidbandet med lägre frekvens snurrar medsols.

Resultanten ändrar endast sin amplitud.

$$s_{AM}(t) = E_c \cos 2\pi f_c t + \frac{E_m}{2} \cos 2\pi(f_c + f_m)t + \frac{E_m}{2} \cos 2\pi(f_c - f_m)t$$

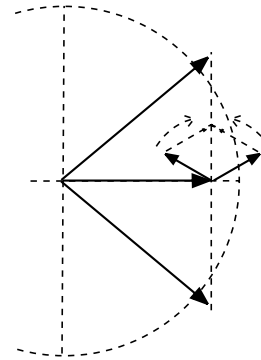
Genom att sätta minustecken framför sidbandet med lägre frekvens så kommer resultanten inte att ligga still utan ändra sin fasvinkel fram och tillbaka. Detta är FM-modulering.

Men vi har inte ren FM. Resultantens amplitud ändras när visaren sveper fram och tillbaka.

$$s_{FM}(t) = E_c \cdot \cos 2\pi f_c t + E_c \cdot J_1(m_f) \cdot [\cos 2\pi(f_c + f_m)t - \cos 2\pi(f_c - f_m)t]$$

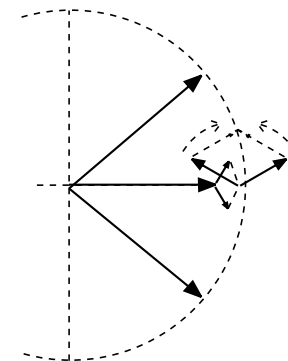
$$s_{FM}(t) = E_c \cdot J_0(m_f) \cdot \cos 2\pi f_c t + E_c \cdot J_1(m_f) \cdot [\cos 2\pi(f_c + f_m)t - \cos 2\pi(f_c - f_m)t]$$

Genom att minska amplituden på bärvågen så kommer resultanten närmare cirkeln.



$$s_{FM}(t) = E_c \cdot J_0(m_f) \cdot \cos 2\pi f_c t + E_c \cdot J_1(m_f) \cdot [\cos 2\pi(f_c + f_m)t - \cos 2\pi(f_c - f_m)t] + E_c \cdot J_2(m_f) \cdot [\cos 2\pi(f_c + 2f_m)t + \cos 2\pi(f_c - 2f_m)t]$$

Genom att lägga en amplitudmodulering på bärvågen med dubbla modulationsfrekvensen så kommer resultanten att ännu bättre följa cirkeln.

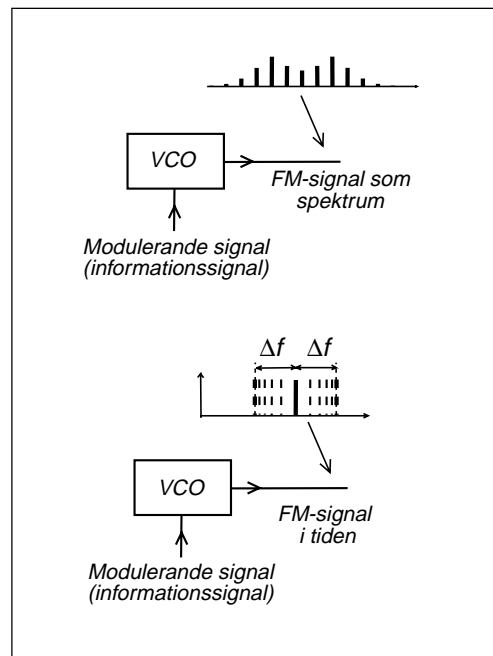


På detta sätt lägger man till fler och fler sidbandspar, där udda sidbandspar ger vinkelmodulering och jämna ger amplitudmodulering.

Två olika sätt att se på FM-signalen

Det skadar aldrig att påminna om att vi har två olika sätt att se på FM-signalen:

- Som spektrum. Oändligt antal spektralkomponenter som är stationära, d.v.s. inte ändrar sig.
- I tiden, där FM-signalen är en bärvåg som sveper $\pm\Delta f$ med konstant amplitud.



7.7 FM-signalens visardiagram

Visarens frekvens i visardiagrammet

En visare som sveper fram och tillbaka i visardiagrammet kommer att få högre frekvens än f_c när visaren roterar motsols, i riktning växande fasvinkel, men lägre frekvens än f_c när visaren roterar medsols, i riktning negativ fasvinkel. På detta sätt kan man även få en uppfattning om vad som händer med frekvensen (fasvinkelns derivata). Visardiagrammet visar ju egentligen bara amplitud och fasvinkel.

Fasvinkel vid PM

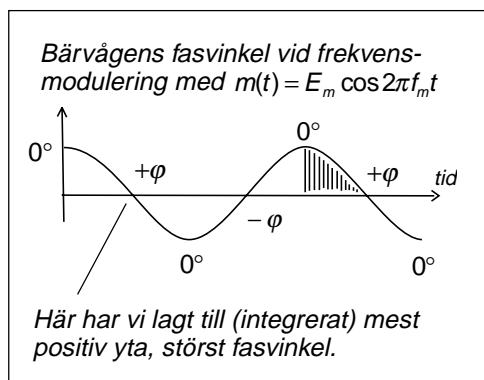
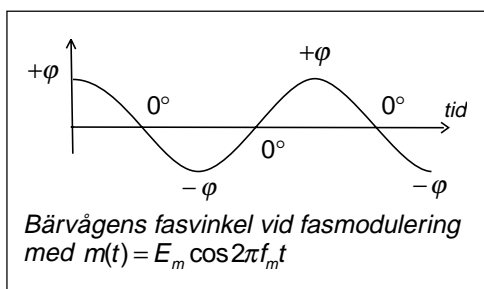
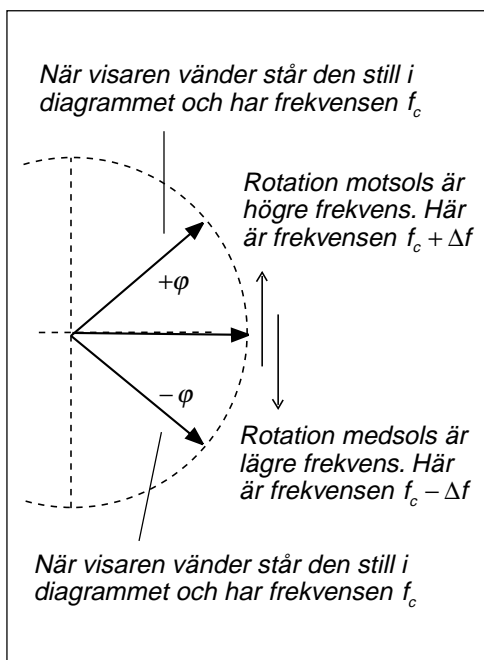
Nästa bild visar informationssignalen $m(t)$ där jag ritat in bärvågens fasvinkel vid fasmodulering, vid olika amplituder på informationssignalen. Vid fasmodulering är fasvinkeln direkt proportionell mot informationssignalens amplitud.

Fasvinkel vid FM

Bilden längst ner visar motsvarande diagram vid frekvensmodulering.

Vid frekvensmodulering är bärvågens frekvens direkt proportionell mot informationssignalens amplitud. Det innebär att frekvensen skall vara f_c vid informationssignalens nollgenomgångar. När har visaren frekvensen f_c ? Jo, när visaren står still i diagrammet, och det gör den när den vänder, i ytterlägena $\pm\varphi$.

Visaren har högst frekvens när den passerar $\varphi = 0^\circ$ i riktning mot $+\varphi$. Modulationssignalens toppar vid FM-modulering med sinusformad informationssignal kommer därför alltid att ligga vid fasvinkeln $\varphi = 0^\circ$. Efter en positiv topp skall följa en nollgenomgång med positiv fasvinkel, och fasvinkelns storlek vid nollgenomgången beror på föregående topps amplitud, eller snarare storleken av den positiva ytan som markerats i bilden, informationssignalens integral.



7.8 Om FM-signalen matas genom ett BP-filter?

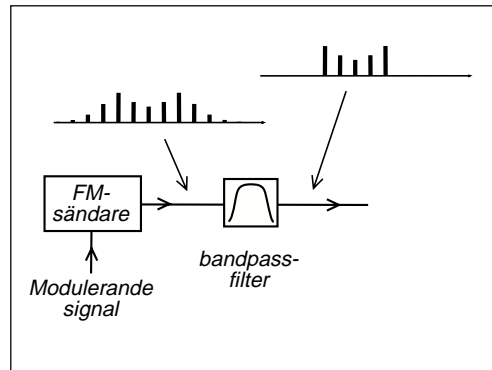
Om FM-signalen filtreras?

Vad händer om man matar FM-signalen genom ett filter, ett bandpass-filter som inte släpper igenom tillräckligt antal spektralkomponenter?

Bärvågen blir amplitudmodulerad

Vad gör spektralkomponenterna? De behövs för att summan av alla spänningsvektorer ska bli en resultatant som följer cirkeln, d.v.s. bilda en bärvåg som har konstant amplitud när den sveper $\pm\Delta f$ fram och tillbaka.

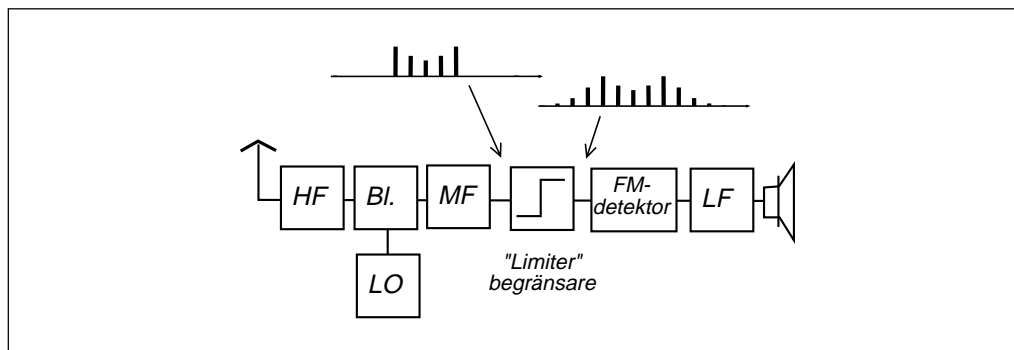
Om vi tar bort några sidbandspar så får vi distorsion, och distorsionens storlek blir summan av de spänningsvektorer som vi tog bort. Det blir alltså en liten amplitudmodulering av bärvågen, och en liten fasmodulering, med en modulationsfrekvens som ligger flera gånger högre än informationssignalens frekvens.



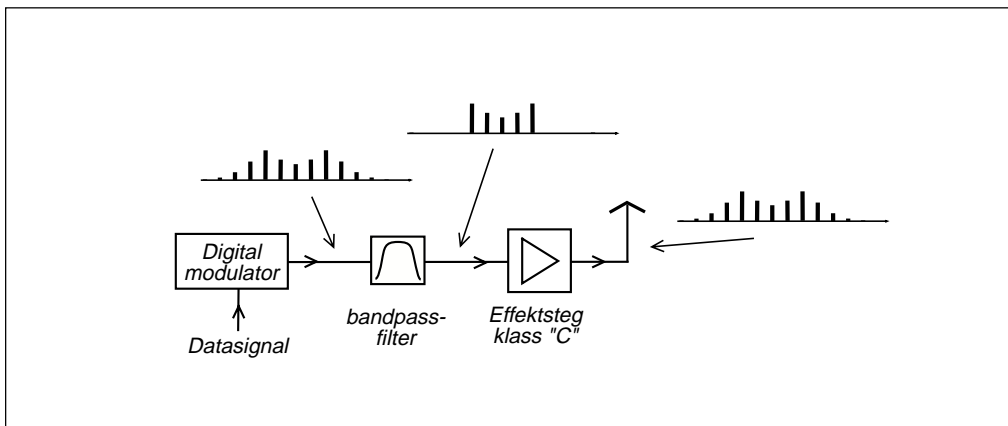
Hur gör vi i FM-mottagaren?

I FM-mottagaren, före FM-detektorn, så inför vi ett förstärkarsteg som "klipper" FM-signalens amplitud. Detta steg, en begränsare (limiter), tvingar bärvågen att hålla konstant amplitud oavsett vilken frekvens bärvågen har. Denna begränsare tvingar resultatanten att följa cirkeln. Vad blir följderna? De sidbandskomponenter som behövs för att resultatanten skall följa cirkeln kommer att bildas och adderas till spektrat.

Vi kan inte återskapa spektrat exakt som det såg ut, men vi blir av med merparten av distorsionen.



7 — Frekvensmodulering och fasmodulering



Filtrering av digital radiosignal

GSM-sändaren innehåller en digital modulator för att alstra den modulerade radiosignalen. Moduleringen är en form av fasmodulering.

För att få plats med så många radiokanaler som möjligt inom befintligt frekvensband så ville man göra radiokanalen så smal som möjligt. Därför filtreras GSM-signalen så att spektrat blir smalare. Distorsionen har ingen större betydelse eftersom det bara gäller att skilja mellan etta och nolla.

Den filtrerade radiosignalen matas till sändarslutsteget. För att få så hög verkningsgrad som möjligt, d.v.s. så stark radiosignal som möjligt i förhållande till inmatad likströmseffekt, så valde man en effektförstärkare som arbetar i klass C. Karakteristiskt för klass C är att transistorn arbetar som switch, den ger fyrkantspänning på utgången, som sedan filtreras till sinusspänning. Sinusspänningen får rätt frekvens, men alla amplitudvariationer hos inspänningen försvinner. Effektsteget arbetar som en "limiter".

Resultatet blev att hela den filtrering av spektrat som utförts före effektsteget gick förlorad. Effektsteget tvingade bärvågen att hålla konstant amplitud och då bildades de bortfiltrerade spektralkomponenterna.

GSM-sändarens effektsteg måste därför klara att förstärka de amplitudvariationer som finns hos den filtrerade signalen. Effektsteget kan inte arbeta i klass C utan måste vara linjärt i övre delen av amplitudkurvan.

8 — FM-demodulering och faslåst loop

Om frekvensen ändras i takt med informationssignalen?

Filter är något som finns på frekvensaxeln. Den ursprungliga FM-detektorn utnyttjade en filterkurva för att omvandla FM till amplitudvariationer, varefter signalen detekterades med en enveloppedetektor, som AM.

Kvadraturdetektorn och den pulsräknande detektorn reagerar direkt på signalens frekvens (hur tätt nollgenomgångarna kommer).

Ytterligare ett sätt är en faslåst slinga (loop). Man gör som sändaren, skapar en FM-signal som man "tvingar" att följa antennsignalen. På detta sätt blir styrsignalen identisk med informationssignalen.

Brus vid FM-demodulering ger krånglig matematik, men med bilder och visardiagram blir det tämligen enkelt.

FM ger högre signal/brusförhållande än AM, DSB och SSB vid samma sändareffekt, men det åtgår större bandbredd. Man slösar med bandbredd för att få brusfriare mottagning. Men detta kan inte utnyttjas för att detektera svaga signaler i bruset. Man får inte högre mottagarkänslighet. Tvärt om. Det finns en bruströskel. Under en viss signalnivå så drunknar FM-signalen i bruset (stor bandbredd ger hög bruseffekt) medan DSB och SSB fortfarande kan användas.

FM är alltså ännu bättre vid starka signaler, men värdelöst vid svaga signaler.

Efter FM-detektorn får man ett brus som ökar med frekvensen. Det "väser". Därför sänks diskanten (deemphasis) på FM-mottagaren.

Man kompenserar med motsvarande diskanthöjning (preemphasis) på sändaren för att informationssignalen skall förbli opåverkad.

Slutligen tittar vi på faslåst loop för att stabilisera mottagarens lokaloscillator.

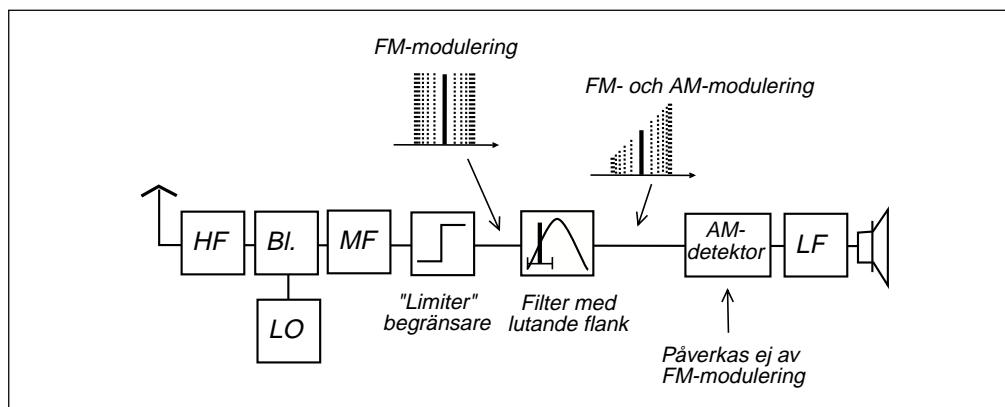
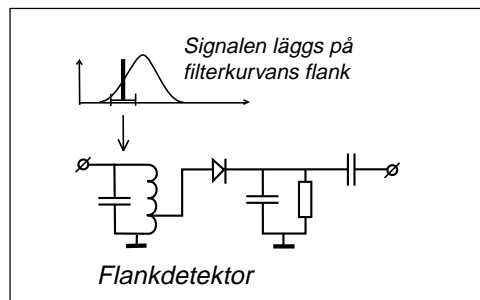
8 — FM-demodulering och faslåst loop

8.1 FM-signalens demodulering

Vi tar emot FM-signalen och den kommer ut ur MF-förstärkaren. Nästa steg är att åter skapa informationssignalen. Detta görs i FM-detektorn, eller FM-demodulern, eller FM-diskriminatorn, olika ord för i stort sett samma sak.

Utnyttja amplitudkurvan hos ett filter

Filterkurvan hos ett bandpass-filter har inte lodräta kanter, utan flankerna lutar mer eller mindre, beroende på hur många reso-



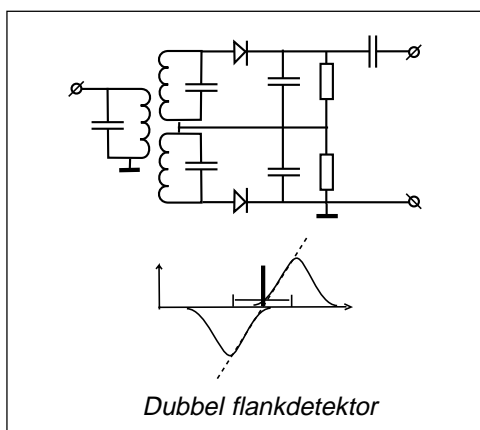
nanskreter som ingår i själva filtret.

Vi utnyttjar en enkel resonanskrets, men väljer filtrets resonansfrekvens så att FM-signalen ligger på ena flanken av filterkurvan. När FM-signalen går mot högre frekvens så släpper filtret igenom mer signal, när FM-signalen går mot lägre frekvens så kommer mindre del av signalen igenom.

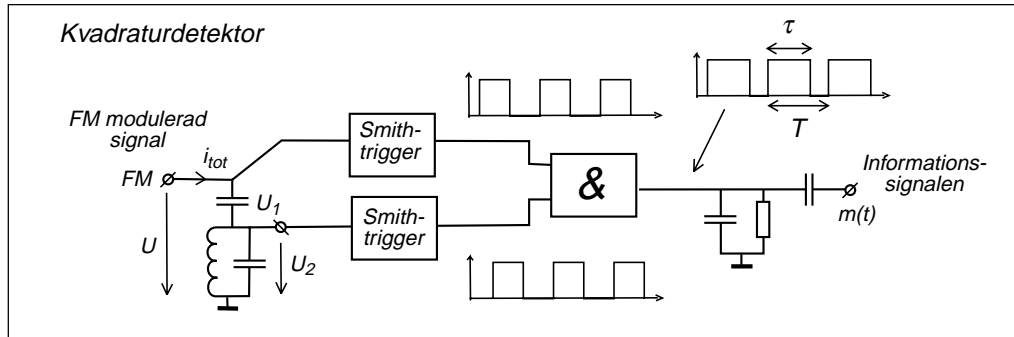
Vad har vi gjort? FM-signalens amplitud ändras beroende på frekvensen. FM-signalen blir AM-modulerad. Vi har fått en signal som är både FM- och AM-modulerad. Denna signal kan vi mata till en enveloppdetektor som är okänslig för signalens frekvensvariationer. Ut ur enveloppdetektorn kommer amplitudändringarna, informations-signalen.

Ett villkor för att detta skall fungera är att det inte redan finns amplitudvariationer på FM-signalen. Därför är det viktigt att mata FM-signalen genom en begränsare innan den matas till resonanskretsen.

Filterkurvans flank är inte helt linjär. Man



får distorsion. Genom att koppla två enveloppdetektorer med olika frekvens på resonanskretsarna, och lägga den andra dioden så att signalens negativa toppar detekteras, så får man en hyggligt rak detektorkarakteristik.



Utnyttja fasvridning i resonanskrets

FM-signalen kan delas upp på två signaler där den ena fasförskjuts (tidsfördröjs) med hjälp av en resonanskrets, som ger olika tidsfördröjning beroende på frekvensen.

Skapa fyrkantpulser av signalerna, fyrkantpulser som överlappar varandra, och summera i en digital "och"-krets (om någon ingång är hög så blir utgången hög).

Detta resulterar i en pulsspänning där pulsförhållandet τ/T (duty cycle) varierar med insignalens frekvens. Pulsspänningens medelvärde följer informationssignalen. Detta kallas kvadraturdetektor.

Pulsräknande detektor

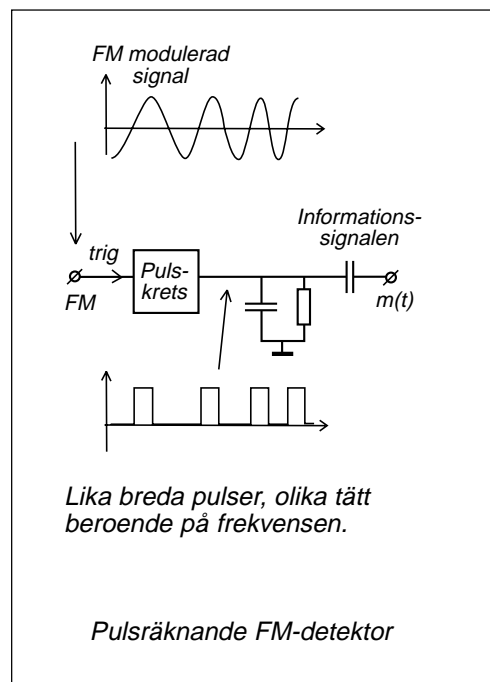
Om vi låter den FM-modulerade signalen trigga en pulsgenerator som lämnar pulser med konstant pulsbredd, så kommer vi att få lika breda pulser men olika tät beroende på den inmatade signalens frekvens. Medelvärdet av pulssignalen följer informationssignalen.

För att vi skall få så stora spänningsvariationer som möjligt måste frekvensändringen i förhållande till mittfrekvensen vara så stor som möjligt.

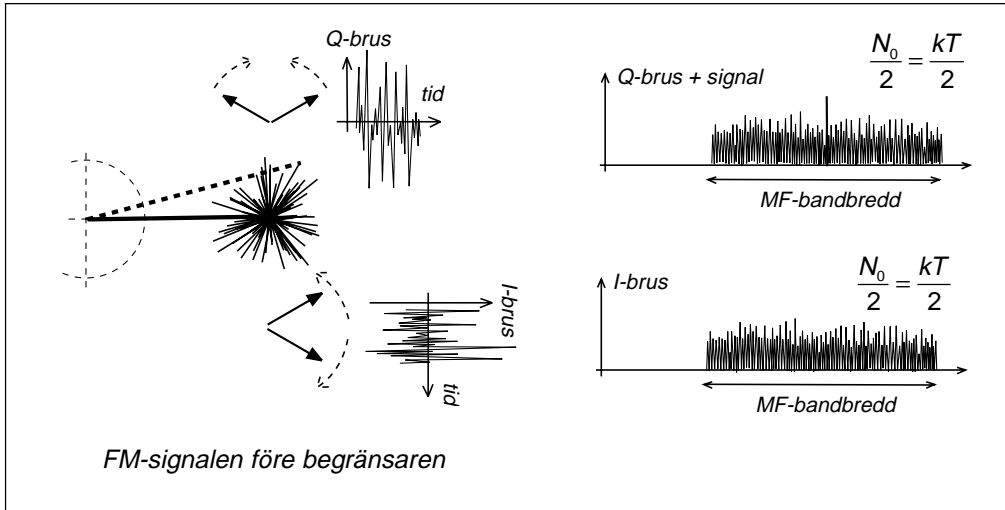
Mellanfrekvensen vid FM-mottagning är vanligtvis 10,7 MHz. Men en variation av $\Delta f = \pm 75$ kHz runt 10,7 MHz ger för små skillnader i pulstäthet.

Bättre är att blanda ner till 200 kHz och på den frekvensen ha $\Delta f = \pm 75$ kHz. Detta ger högre utspänning från den pulsräknande detektorn (detektorn som räknar hur tät pulserna kommer).

När man vill få in en komplett FM-mottagare i en integrerad krets passar den pulsräknande detektorn bra, eftersom man slipper resonanskretsar.



8.2 FM-signalen och brus



Före begränsaren

Före begränsaren finns signalen och brus. Totala bruseffekten är $N = kTB$, där B är MF-bandbredden.

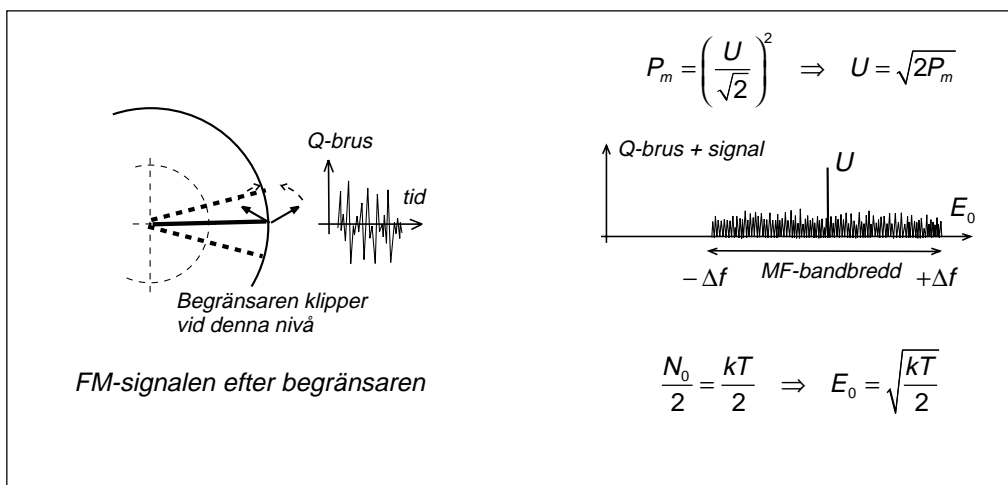
Bruset kan delas upp i två delar. I-bruset som ligger i fas med signalen och består av identiska sidbandskomponenter runt signalen. Q-bruset som ligger i kvadratur till signalen och har sidbandskomponenter som skiljer med ett minustecken.

Halva bruseffekten ligger som I-brus och andra halvan som Q-brus.

Efter begränsaren

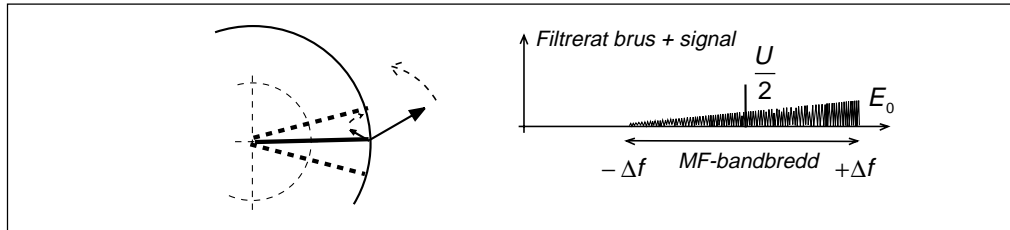
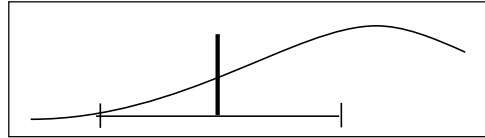
Begränsaren klipper amplituden på signal plus brus, "fyrkantar" signalen, men filtrerar ut grundtonen så att vi återfår sinusformen. Då försvinner amplitudvariationerna och vi blir av med I-bruset.

Om vi efter begränsaren förstärker till ursprunglig amplitud, så har vi alltså en signal med effekten P_m och toppspänningen U . Dessutom har vi halva bruseffekten som vi räknar om till brusspänning E_0 per Hz, effektivvärde.



8 — FM-demodulering och faslåst loop

Efter begränsaren har vi ett amplitudspektrum bestående av signal plus Q-brus. Detta amplitudspektrum matas genom flanken på ett filter så att bruset kommer att "luta".



Efter filtret

Efter filtret är inte längre alla sidbandspar i Q-bruset lika långa. När alla dessa sidbandspar snurrar åt olika håll så kommer de delar i varje par som är lika långa att ta ut varandra (kvarvarande Q-brus), medan resten är en visare per sidbandspar, ett sidband, som snurrar. Dessa visare ger visserligen fasvariation men framför allt amplitudvariation hos signalen.

AM-detektorn

Till enveloppdetektorn kommer själva signalen som fungerar som bärvåg, och sidband bestående av brus. Skillnaden mellan sidbanden är amplitudbrus. Detta amplitudspektrum för bruset finns efter AM-detektorn.

Beräkna bruset som vi hör

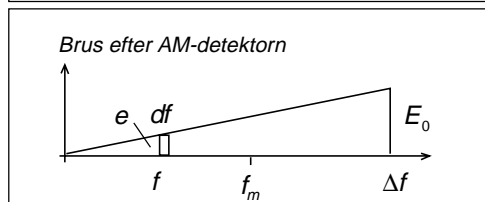
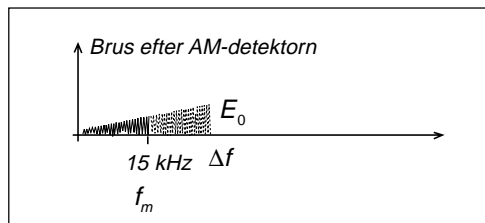
Först beräknar jag brusspänningen e vid valfri frekvens f . Därefter beräknar jag brus-effekten ΔN i det lilla delementet. Sedan summerar jag allt brus inom det hörbara frekvensområdet, d.v.s. upp till f_m .

Beräkna den detekterade signalen

Den detekterade signalen S_{ut} kommer att vara en sinusspanning med topp-till-toppvärdet U (100 % AM-modulering).

Signal/brus-förhållandet efter FM-detektorn

Nu är det bara att dividera och ersätta U och E_0 med de ursprungliga uttrycken.



$$\frac{e}{f} = \frac{E_0}{\Delta f} \Rightarrow e = \frac{E_0}{\Delta f} \cdot f$$

$$\Delta N = e^2 \cdot df = \left(\frac{E_0}{\Delta f} \right)^2 \cdot f^2 \cdot df$$

$$N_{ut} = \int_0^{f_m} e^2 \cdot df = \left(\frac{E_0}{\Delta f} \right)^2 \int_0^{f_m} f^2 df = \left(\frac{E_0}{\Delta f} \right)^2 \cdot \frac{f_m^3}{3}$$

$$\begin{aligned} \frac{S_{ut}}{N_{ut}} &= \frac{\left(\frac{U}{2\sqrt{2}} \right)^2}{\left(\frac{E_0}{\Delta f} \right)^2 \cdot \frac{f_m^3}{3}} = \frac{P_m \left(\frac{1}{2} \right)^2}{\frac{kT}{2} \cdot \frac{1}{\Delta f^2} \cdot \frac{f_m^3}{3}} \\ &= \frac{3}{2} \cdot \left(\frac{\Delta f}{f_m} \right)^2 \cdot \frac{P_m}{kTf_m} = \frac{3}{2} \cdot m_f^2 \cdot \frac{P_m}{kTf_m} \end{aligned}$$

8.3 Bruströskel vid FM-demodulering

Signal/brusförhållandet före FM-detektorn

Hela resonemanget kring FM-detektorn bygger på att signalen är mycket starkare än bruset så att amplitudbruset klipps bort av begränsaren. Vad händer om så inte är fallet?

Amplitudbruset klipps bort av begränsaren om brustopparna är svagare än själva signalens toppvärde. Detta kan sägas vara uppfyllt om signalens toppvärde är 5 ggr högre än brusets effektivvärde, se avsnittet om Gaussiskt brus. Vi behöver ett signal/brusförhållande på 12,5 ggr, ca 11 dB. Detta ger en bruströskel. Om den mottagna effekten P_m är svagare än 11 dB över kTB (där B är mellanfrekvensbandbredden) så dränks modulationssignalen i brus.

Om man vill jämföra två olika FM-system med olika modulationsindex m_f så måste man ta med i beräkningen att man behöver större bandbredd B när modulationsindex är högre. 11 dB över bruset i ett bredbandigt system är en kraftigare signal P_m än 11 dB över bruset i ett smalbandigt system. Därför normerar vi till informationssignalens bandbredd f_m , som ju även används vid DSB och SSB.

Bruströskel:

$$P_m = 12,5 \cdot kTB = 12,5 \cdot kTf_m [2(m_f + 1)]$$

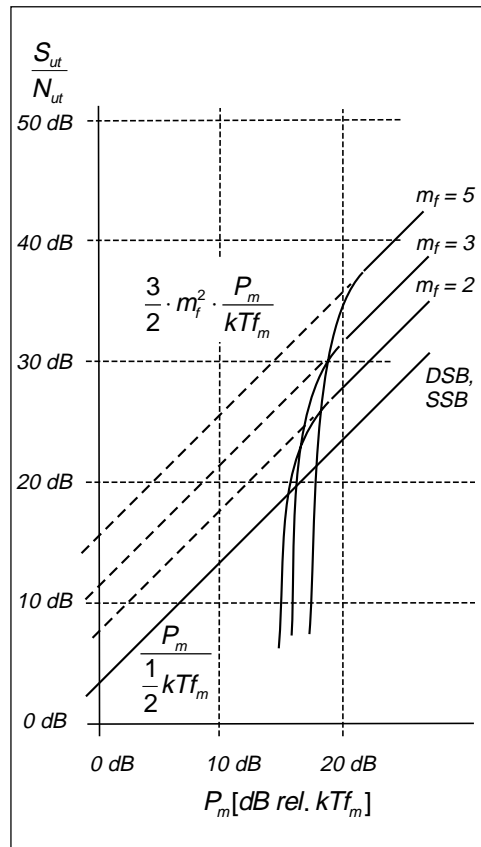
FM med modulationsindex $m_f = 5$ ger 12,7 dB högre signal/brusförhållande efter detektorn än DSB eller SSB, vid samma mottagen signaleffekt P_m , d.v.s. vid samma sändareffekt. Denna förbättring sker på bekostnad av bandbredden.

- SSB behöver bandbredden f_m ,
- DSB behöver $2 \cdot f_m$
- Ett FM-system med modulationsindex $m_f = 5$ behöver $B = f_m [2(m_f + 1)] = 12 \cdot f_m$.

$$\begin{cases} \text{signal} = P_m = \frac{1}{R} \left(\frac{\hat{U}}{\sqrt{2}} \right)^2 \\ \text{brus} = E_n^2 = \frac{1}{R} kTB \\ \frac{\hat{U}}{E_n} = 5 \end{cases}$$

$$\frac{P_m}{kTB} = \frac{\left(\frac{5E_n}{\sqrt{2}} \right)^2}{kTB} = \frac{25 \cdot kTB}{2 \cdot kTB} = 12,5$$

$$\frac{P_m}{kTB} [\text{dB}] = 10 \cdot \log 12,5 \approx 11 \text{ dB}$$



8.4 Preemphasis och deemphasis

Basbandsbruset ökar med frekvensen

Vi får brus efter FM-demodulatorn vars amplitudspektrum växer linjärt med frekvensen. Bruset ökar med +6 dB per oktav.

Diskantsänkande filter, deemphasis

Det är obehagligt att lyssna på brus som är starkare i diskanten, det väser. För att motverka väsandet hos bruset så kopplar man in ett diskantsänkande filter (deemphasis) som sänker med -6 dB/oktav från 3,2 kHz till 15 kHz. Men detta filter sänker även diskanten på den önskade signalen. Därför höjer man diskanten i sändaren (preemphasis) på motsvarande sätt med +6 dB/oktav.

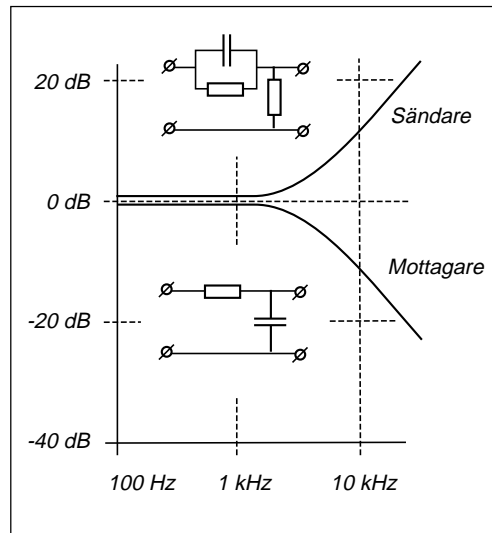
Fasmodulering över 3,2 kHz

Det diskanthöjande nätet är ett högpasfilter som deriverar modulationssignalen vid frekvenser över 3,2 kHz. Att derivera modulationssignalen och mata den till en FM-sändare ger fasmodulering som resultat.

Sedan tar vi emot denna fasmodulerade signal med en FM-detektor, men det diskantsänkande lågpasfiltret gör egentligen om FM-detektorn till en fasdetektor.

Deemphasis förbättrar S/N

Det diskantsänkande filtret minskar brusnivån med närmare -12 dB i mottagaren. Visserligen måste man öka modulationsindex på sändaren i motsvarande grad, men detta i ett frekvensområde där informations-



signalens spektrum innehåller låg effekt, på frekvenser över 3,2 kHz.

RDS och brus

RDS datakanalen ligger modulerad på en underbärvåg vid 57 kHz. Vid så hög modulationsfrekvens blir bruset mycket kraftigt efter FM-detektorn.

Tidigare fanns RDS personsökning, en tjänst som utnyttjade de starka FM-rundradiosändarna (60 kW ERP). Den har ersatts av Minicall (ca 50 W ERP) där sändaren endast moduleras med datasignalen. Systemen har jämförbar räckvidd.

$$N_{ut}(\text{minicall}) = \left(\frac{E_0}{\Delta f} \right)^2 \int_0^{3 \text{ kHz}} f^2 df = \left(\frac{E_0}{\Delta f} \right)^2 \cdot \frac{(3 \text{ kHz})^3}{3}$$

$$N_{ut}(\text{RDS}) = \left(\frac{E_0}{\Delta f} \right)^2 \int_{54 \text{ kHz}}^{60 \text{ kHz}} f^2 df = \left(\frac{E_0}{\Delta f} \right)^2 \cdot \left[\frac{(60 \text{ kHz})^3}{3} - \frac{(54 \text{ kHz})^3}{3} \right]$$

$$\frac{N_{ut}(\text{RDS})}{N_{ut}(\text{minicall})} = \frac{(60 \text{ kHz})^3 - (54 \text{ kHz})^3}{(3 \text{ kHz})^3} = \frac{216000 - 157464}{27} = 2168$$

$$\frac{N_{ut}(\text{RDS})}{N_{ut}(\text{minicall})} [\text{dB}] = 10 \cdot \log 2168 = 33,4 \text{ dB}$$

8 — FM-demodulering och faslåst loop

8.5 Faslåst slinga (loop)

Vi tar signalen från en generator, i detta fall en VCO, en spänningsstyrd oscillator där generatorfrekvensen påverkas av en styrspänning.

Fasdetektor

Signalen från VCO:n matar vi genom en strömbrytare som styrs av en symmetrisk fyrkantspänning (50 % stängd, 50 % öppen strömbrytare), med frekvensen f_{ref} .

Om VCO:n håller samma frekvens som f_{ref} , så kommer spänningen på andra sidan strömbrytaren att se olika ut beroende på fasläget, d.v.s. beroende på hur spänningarna ligger i tiden.

- Figur a) visar utspänningen när VCO:n ligger i fas med f_{ref} .
- Figur c) visar utspänningen när vi har en kvarts periods tidsförskjutning, -90° fasförskjutning mellan VCO-spänningen och f_{ref} .
- Figurerna b) och d) visar motsvarande när tidsförskjutningen är något mindre eller något större än -90° .

Spänningen efter strömbrytaren matas genom ett lågpasfilter (integrering) och ut får vi en likspänning som är noll om fasförskjutningen är -90° (och -270°), positiv om fasförskjutningen ligger mellan $+90^\circ$ och -90° , samt negativ om fasförskjutningen ligger mellan -90° och -270° .

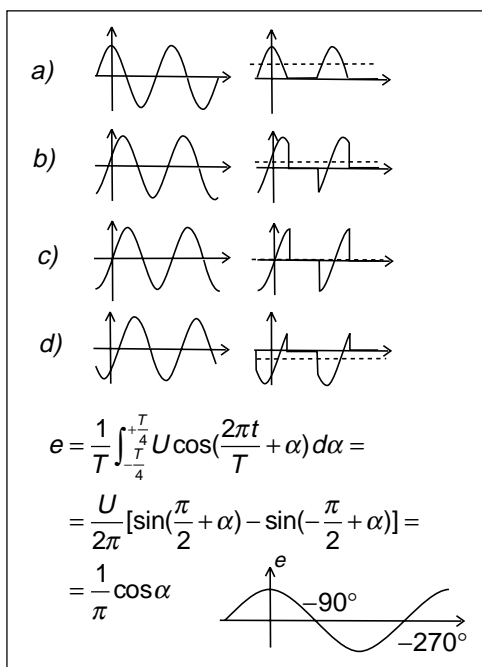
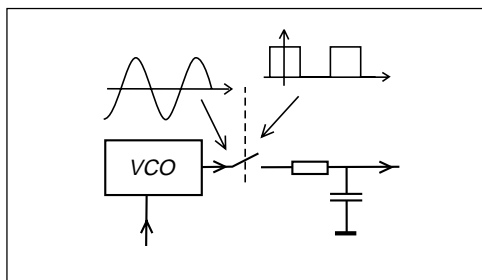
Strömbrytaren och lågpasfiltret fungerar som en fasdetektor, eftersom utspänningens medelvärde är proportionellt mot cosinus för fasvinkeln mellan VCO-spänningen och f_{ref} .

Faslåst slinga

Nu tar vi utspänningen e , förstärker den och matar in den som styrsignal på VCO:n.

Styrsignalen påverkar VCO-frekvensen, vilket i sin tur påverkar spänningen e som matas tillbaka som styrsignal. VCO-frekvensen och spänningen e kommer att pendla fram och tillbaka tills "slingan" hittar ett stabilt läge. Detta läge kan ligga antingen i närheten av -90° eller i närheten av -270° .

Om en ökning av styrspänningen i positiv

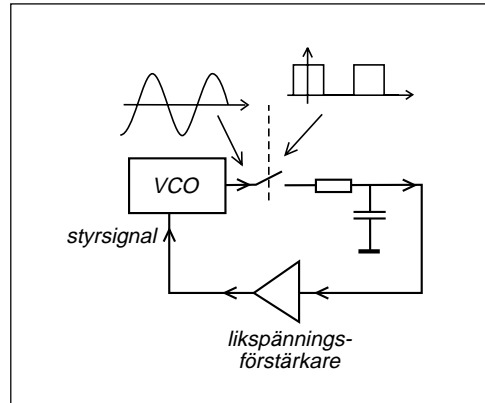


8 — FM-demodulering och faslåst loop

riktning innebär att VCO-frekvensen minskar, så kommer det stabila läget att ligga i närheten av -90° . Om ökning av styrspänningen innebär att VCO-frekvensen ökar så ligger det stabila värdet i närheten av -270° .

Vid det stabila läget kommer ett litet fel i styrsignal att driva tillbaka VCO:n i rätt fasläge. Vid det instabila läget kommer felet att öka så att VCO:n drivs bort och kommer att hitta det stabila läget.

Detta reglersystem är en faslåst slinga men kallas oftast faslåst loop efter engelskans "phase locked loop".



8 — FM-demodulering och faslåst loop

8.6 Faslåst loop som FM-demodulator

Den FM-modulerade signalen alstras i sändaren genom att informationssignalen $m(t)$ får styra en VCO.

I FM-mottagaren tar vi signalen från MF-steget och matar den till en fasdetektor. Fasdetektorns andra ingång matas från en lokal VCO i FM-mottagaren.

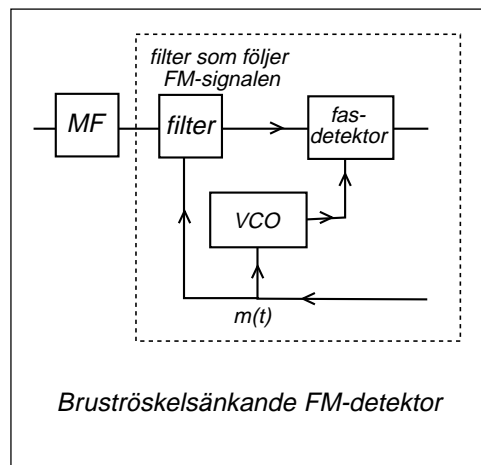
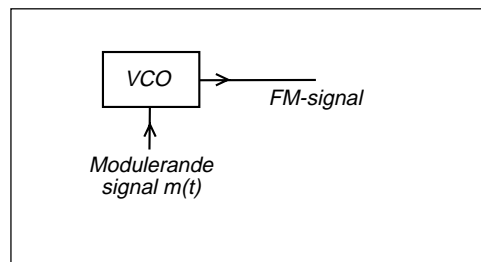
Utsignalen e från fasdetektorn matas via ett lågpasfilter och en förstärkare till den lokala VCO:n. Vad kommer att hända?

När den FM-modulerade signalen från MF-steget sveper i frekvens fram och tillbaka, den är ju FM-modulerad, så tvingas VCO:n att följa med. Den är faslåst till MF-signalen.

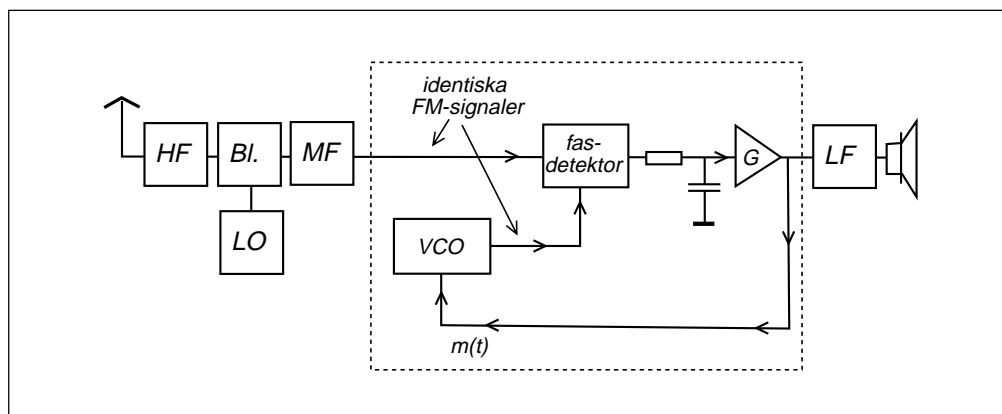
Men när VCO:n följer med i MF-signalens frekvensvariationer, vilken styrspänning måste då matas till den lokala VCO:n? Jo, $m(t)$. Den styrspänning som alstras i den faslåsta loopen är alltså identisk med den styrspänning $m(t)$ som matades till FM-sändaren.

Bruströskelsänkande FM-detektor

Bruströskeln bestäms av bruseffekten före FM-detektorn, alltså bandbredden på MF-filtret. MF-filtret måste egentligen vara så brett att det släpper igenom tillräckligt många spektralkomponenter, alltså $2 \cdot \Delta f$ och lite till. Men om man kan få filtret att flytta sig i frek-



vens i takt med bärvågens frekvensvariation, så räcker en bandbredd på $2 \cdot f_m$. Ett sådant filter kan styras på samma sätt som VCO:n, av $m(t)$.



8.7 Faslåst loop som lokaloscillator

Det är svårt att konstruera en noggrann och stabil oscillator

Det är svårt att konstruera en oscillator med noggrann och stabil frekvens, noggrann på så sätt att man vet vilken frekvens det är, och stabil innebärande att frekvensen inte ändrar sig t.ex. på grund av temperaturändringar. Speciellt gäller detta om frekvensen är hög.

Äldre tiders tv-mottagare hade en lokaloscillator uppbyggd kring en resonanskrets bestående av spole och spänningsstyrd kondensator (kapacitansdiod). Frekvensen ställdes in genom att man ändrade spänningen till kapacitansdioden med en potentiometer.

Frekvensen var inte särskilt noggrann. Det gick inte att först ställa in frekvensen och därefter slå på tv-mottagaren och tro att frekvensen var den rätta för den tv-kanal man ville titta på. Man måste först slå på tv:n och därefter skruva på frekvensratten, söka tills man hittade rätt tv-kanal. Man utnyttjade tv-sändaren för att ställa in rätt lokaloscillatorfrekvens.

När man ställt in rätt kanal och började titta så blev tv:n varm. Efter ett tag hade lokaloscillatorfrekvensen ändrat sig eftersom kapacitansen och induktansen i resonanskretsen ändrades med temperaturen. Inte mycket, men tillräckligt för att man skulle få t.ex. bildstörning i tv-ljudet. Då fick man gå fram och skruva på frekvensinställningen. Lokaloscillatorfrekvensen var inte tillräckligt stabil.

Radiosändaren utgör stabil referenssignal

Då infördes en enklare form av faslåsnings. Tv-signalen matades till en särskild FM-diskriminator med mycket lång tidskonstant på lågpasfilteret. Man fick en likspänning som var noll när tv:n var rätt inställd, men något positiv eller något negativ om tv:n låg för högt eller för lågt i frekvens. Denna likspänning matades tillbaka till lokaloscillatorns kapacitansdiod för att korrigera

frekvensdriften. Detta kallas AFC (Automatic Frequency Control). Oftast ställdes tv:n in med AFC bortkopplad, varefter AFC kopplades in för att motverka temperaturdrift.

Samma typ av AFC utnyttjas på FM-mottagare med analog inställning. Du märker hur stationen liksom hänger kvar när man ändrar frekvens.

Att låsa till radiosändaren är en enkel lösning, men den fungerar inte vid radiosystem där man inte kan leta efter sändaren utan måste hitta rätt frekvens med en gång.

Referensoscillatorn i radiomottagaren

Om vi inte kan utnyttja radiosändaren så behöver vi en egen referensoscillator i mottagaren.

Resonanskrets med spole och kondensator fungerar som en elektrisk pendel, se avsnittet om antenner. En stabilare resonanskrets får vi med en mekanisk resonanskrets. Det finns kvarts som är piezo-elektriskt. Det innebär att om en kvartskristall utsätts för mekaniskt tryck så bildas elektrisk spänning över kvartskristallen. Även det omvända fungerar, nämligen att om man lägger spänning över kvartskristallen så förändrar den sina dimensioner.

Placera en kvartskristall mellan två metallplattor och knuffa med en växelspanning så börjar kvartskristallen att vibrera, pendla som gungan i lekparken, på sin egen mekaniska resonansfrekvens.

Kvartskristallen används som frekvensbestämmande element i alla typer av referensoscillatorer, i elektriska klockor, som klockkrets i datorer, som referensoscillator i sändare och mottagare, men även som resonanskrets i MF-filter (kristallfilter).

SchemasyMBOL



styrkristall (X-tal)

8 — FM-demodulering och faslåst loop

Faslåst loop som lokaloscillator

Låt oss titta på ett exempel.

En FM-mottagare med 10,7 MHz mellanfrekvens skall användas för att ta emot stationer på FM-bandet (87,5 – 108 MHz). De frekvenser som skall kunna ställas in är 88,6 88,7 88,8 ... 107,8 107,9 MHz. Ingen FM-sändare får ligga på 87,5 eller 108 MHz eftersom sändarens hela spektrum skall ligga inom FM-bandet.

- Vilka frekvenser skall lokaloscillatorn kunna ställas in på?
- Beräkna referensoscillatorns frekvens och delningstalet N när man använder faslåst loop.

Uppgift a):

Lokaloscillatorn kan ligga högre i frekvens. Då får vi:

$$f_{\text{ant}} = 88,6 \text{ MHz} \quad f_{\text{LO}} = 88,6 + 10,7 = 99,3 \text{ MHz}$$

----- i steg om 100 kHz

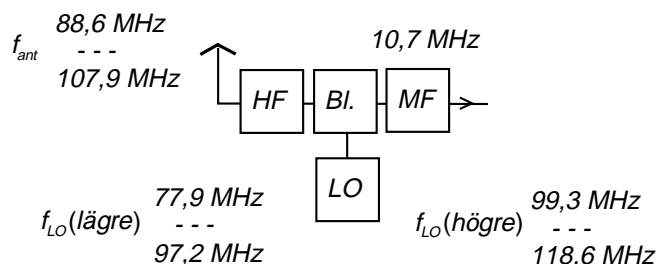
$$f_{\text{ant}} = 107,9 \text{ MHz} \quad f_{\text{LO}} = 107,9 + 10,7 = 118,6 \text{ MHz}$$

Lokaloscillatorn kan ligga lägre i frekvens. Då får vi:

$$f_{\text{ant}} = 88,6 \text{ MHz} \quad f_{\text{LO}} = 88,6 - 10,7 = 77,9 \text{ MHz}$$

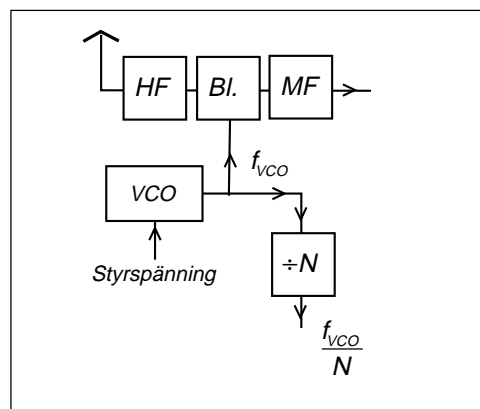
----- i steg om 100 kHz

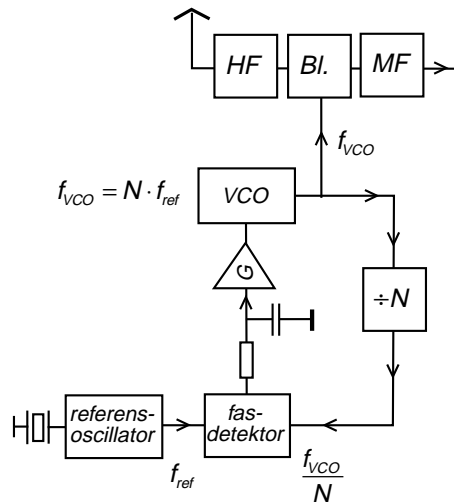
$$f_{\text{ant}} = 107,9 \text{ MHz} \quad f_{\text{LO}} = 107,9 - 10,7 = 97,2 \text{ MHz}$$



I uppgift b) skall vi använda en VCO som lokaloscillator samt en frekvensdelare.

Frekvensdelaren fungerar på så sätt att den frekvens som matas in, i detta fall f_{VCO} , görs om till spänningspulser varje gång växelspanningen passerar noll. Vi får två pulser per period. Sedan räknar vi pulser, och efter N pulser skiftar utspänningen från noll till plus, eller från plus till noll. Utspänningen blir en fyrkantvåg vars periodtid är N ggr insignalens, d.v.s. vi får en frekvens som är N ggr lägre. Delningstalet N kan bara vara heltal.





Loopen ställer in sig så att $\frac{f_{VCO}}{N} = f_{ref}$

Detta medför att $f_{VCO} = N \cdot f_{ref}$

Eftersom N bara kan vara heltal så väljer vi $f_{ref} = 100 \text{ kHz}$

Vi kan beräkna delningstalet $N = \frac{f_{VCO}}{f_{ref}}$

Uppgift b):

Referensoscillatorn kan ligga på 100 kHz eller på 100 kHz dividerat med ett heltal.

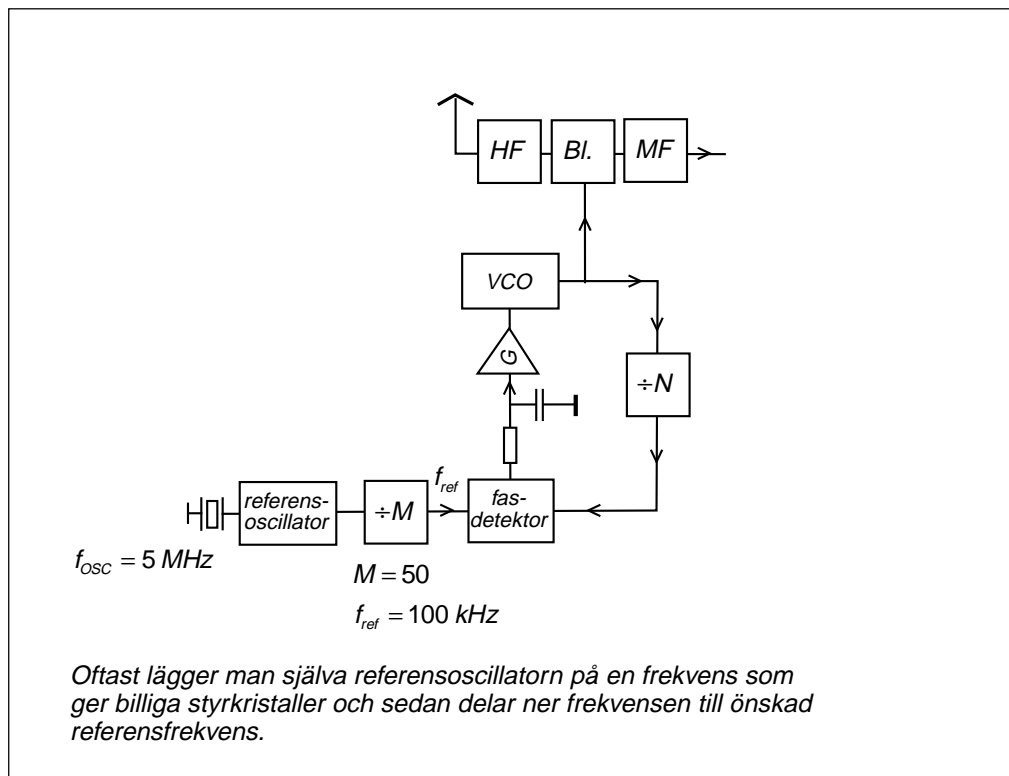
Om lokaloscillatorn skall ligga högre i frekvens får vi:

$f_{VCO} = 99,3 \text{ MHz}$ $N = 993$
i steg om 100 kHz *i steg om 1*
 $f_{VCO} = 108,6 \text{ MHz}$ $N = 1086$

Om lokaloscillatorn kan ligga lägre i frekvens får vi:

$f_{VCO} = 77,9 \text{ MHz}$ $N = 779$
i steg om 100 kHz *i steg om 1*
 $f_{VCO} = 97,2 \text{ MHz}$ $N = 972$

8 — FM-demodulering och faslåst loop



Faslåsta loops insvängningstid

När man ändrar delningstalet N i den faslåsta loopen så tar det viss tid innan slingan stabiliserat sig och låst VCO:n på den nya frekvensen. Vid specificeringen av GSM-systemet räknade man med 1 ms innan frekvensen stabiliserats.

GSM-ficktelefonen sänder dataskurar i tidluckor, en dataskur var åttonde tidlucka. I någon tidlucka däremellan tar ficktelefonen emot dataskurar från basstationen. Eftersom ficktelefonen inte sänder och tar emot samtidigt så kan samma VCO användas både för att alstra mottagarens lokaloscillator och att alstra sändarfrequensen, men det måste finnas 1 ms emellan för frekvensbyte. Detta har påverkat utformningen av GSM-systemets TDMA-ram (Time Division Multiple Access).

Finjustering av referensoscillatorn

Även en styrkristall kan ändra frekvens med temperaturen även om temperaturändringen är liten. Men man kan finjustera kristalloscillatorns frekvens med en variabel kondensator i serie eller parallellt över styrkristallen.

När GSM-ficktelefonen loggar in på ett GSM-nät så letar den efter en basstations-sändare som med jämna mellanrum sänder en speciell dataskur FCCH ("Frequency Correction Channel") som kan användas av ficktelefonen för att finjustera sin referensoscillator.

Ficktelefonens referensoscillator används inte bara för att generera sändar- och mottagarfrekvenserna utan även som data-klocka och tidsreferens.

9 — Datakommunikation

Först tittar vi på OSI-modellen för att förstå vad vi håller på med.

Många informationssignaler är från början analoga. De måste omvandlas till digital form. Vi tittar på samplingsteoremet och PCM-kodning av telefonsamtal och musik-cd, och inser att distorsionsfri återgivning av musik-cd är omöjligt.

Vid kommunikation så överförs symboler. Symbolhastigheten bestämmer hur stor bandbredd som behövs.

Genom att ha fler än två olika symboler att välja mellan så kan mer än en databit överföras på varje symbol.

Den övre gränsen för symbolhastigheten sätts av bandbredden.

Den övre gränsen för hur många olika utseenden man kan ha på symbolen sätts av bruset.

Shannons formel ger en övre gräns för bit-hastigheten när vi känner bandbredden och signal/brusförhållandet.

9.1 Protokoll, gränssnitt (Interface) och OSI

OSI-modellen

"Reference model for Open Systems Interconnection" är en referensmodell för digital telekommunikation framtagen av internationella standardiseringsorganisationen ISO.

OSI-modellen består av sju lager, där varje lager utför en bestämd uppgift.

Protokoll

Uppgiften som utförs av ett visst lager, t.ex. nätverkslagret, beskrivs i ett "protokoll". Det som utförts av ett lager tolkas av motsvarande lager i andra ändan av förbindelsen.

Flera av lagren behöver kunna "prata" med motsvarande lager i andra ändan. Sådana lager adderar information i ett huvud. Huvudet tolkas och avlägsnas på motsvarande nivå i mottagarändan.

Gränssnitt (Interface)

På sändarsidan tar varje lager emot data från överliggande lager, bearbetar enligt protokollet och skickar vidare till lagret under. Mellan varje lager finns ett gränssnitt där det beskrivs hur informationen skall vara ordnad för att nästa lager skall begripa.

Lager 1: Fysiska lagret

Fysiska lagrets uppgift är att ta emot databitar från datalänklagret och skapa "symbo-

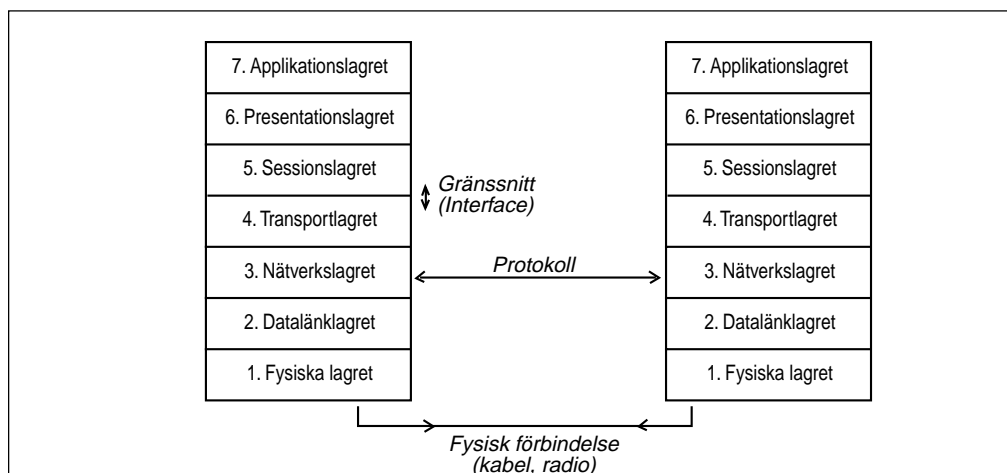
ler" som går att skicka på den fysiska förbindelsen.

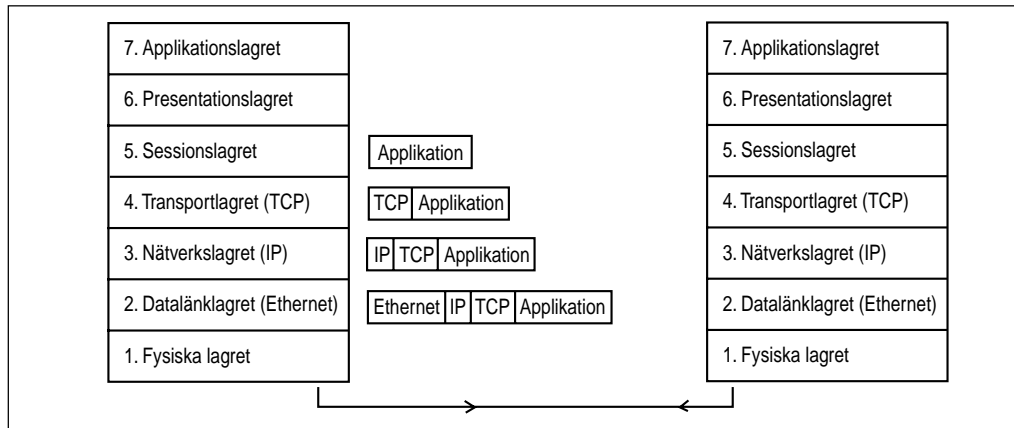
- Fysiska lagret kan skapa ljuspulser, symboler som skickas på optofibern.
- Fysiska lagret kan skapa olika spänningsnivåer, symboler som t.ex. skickas på kopparkabeln till abonnenten vid ISDN-abonnemang.
- Fysiska lagret kan förändra en bärvåg, modulera, skapa symboler som kan sändas på kabel (telefonmodem) eller skickas genom luften (radiomodem).

Gränssnittet mot fysiska förbindelsen beskriver hur symbolerna ser ut, t.ex. spänningsnivåer (V.28-gränssnittet) eller sättet att modulera (GSM:s radiogränssnitt).

Bitsynkronisering

Fysiska lagret kan lägga till extra databitar, extra symboler, utöver de databitar som kommer från datalänklagret. En orsak till att addera extra databitar är att ge mottagaren möjlighet att "bitsynkronisera", d.v.s. upptäcka exakt i vilket tidsögonblick symbolerna skall avläsas. En dataskur kan inledas med ett visst antal växlingar mellan två symboler, mellan "etta" och "nolla". Eller det kan, som i GSM, finnas en speciell "tränings-





sekvens” om 26 databitar som mottagaren utnyttjar för att hitta rätt tidsögonblick, för att som man säger ”bitsynkronisera”.

Faltningskodning

Att kontrollera om databitarna är rätt mottagna eller rätta eventuella bitfel är uppgifter för datalänklagret eller transportlagret. Men det finns bitfelsminimerande kodning som egentligen inte rättar bitfel utan bara minimerar bitfelen. Om jag rättar bitfel, då vet jag vilken databit som var feltolkad, vilken symbol som var fel mottagen. Den bitfelsminimerande kodningen kodar en hel datasekvens så att datasekvensen skall passa in i ett mönster (trellis) och mottagaren väljer det mönster som totalt sett ligger närmast den mottagna sekvensen. Mottagaren vet inte om denna sekvens är rätt, men den vet att utifrån det som mottagits så är detta den sekvens som uppvisar minst antal bitfel. Faltningskodning eller trelliskodning är en bitfellsänkande kodning som tillhör fysiska lagret.

Lager 2: Datalänklagret

Datalänklagret kopplar upp en förbindelse (datalänk) mellan två noder (datorer) på samma LAN (Local Area Network).

- Datalänklagret hanterar adresser och ser till att rätt datorer kommunicerar.
- Datalänklagret organiserar kommunikationen så att datorerna inte pratar i mun på varandra.

- Datalänklagret kan hantera ord- eller ramsynk, så att mottagaren klarar att identifiera åttabitsord eller dataramar i den mottagna bitströmmen. Alternativt kan detta ligga i fysiska lagret.

- Datalänklagret kan även kontrollera om det finns bitfel och begära omsändning från närmaste nod. Omsändning på nät-nivå hanteras av transportlagret.

Lager 3: Nätverkslagret

Nätverkslagret hanterar globala adresser, t.ex. IP-nummer, och sköter vägval via en eller flera routrar till någon bestämd dator på ett annat LAN (IP-protokollet).

Lager 4: Transportlagret

Transportlagret kan hantera uppdelning av långa datafiler i mindre paket, kontrollera om det finns bitfel samt begära omsändning av ursprungsdatorn (TCP-protokollet).

Lager 5: Sessionslagret

Exempel på sessionslager i Internet-världen är HTTP, FTP m.fl.

Lager 6 - 7:

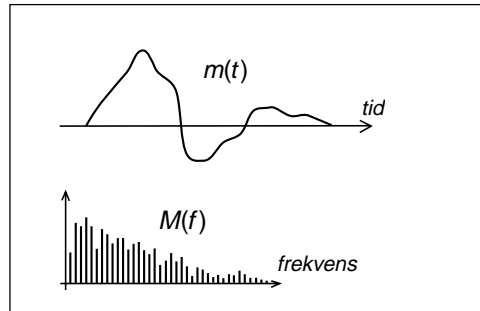
Gränserna mellan de övre lagren i protokollstacken kan vara diffusa. Verklighetens datakommunikation följer inte OSI-modellen strikt.

9.2 A/D-omvandling och samplingsteoremet

A/D-omvandling

Många informationssignaler är analoga t.ex. tal, och behöver överföras i digital form. Analog innebär att signalen kan anta vilka värden som helst inom två gränser. Analog beskrivning av signalen innebär att $m(t)$ ger oss signalens amplitud i varje punkt.

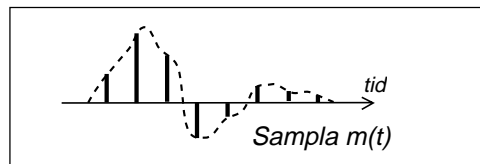
Den analoga signalen har tidsfunktionen $m(t)$ och spektrat $M(f)$.



Sampling

Vi mäter amplituden på $m(t)$ och tar ett "spänningsprov" (sampel) vid vissa bestämda tidpunkter. Vi samplar signalen.

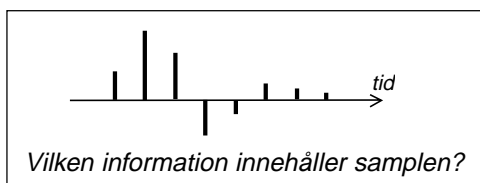
Samplingen är ett steg i att göra en digital beskrivning av $m(t)$. Spänningssampelen beskriver inte $m(t)$ i varje punkt utan bara i vissa punkter.



D/A-omvandling

Om vi överför dessa spänningssampel till en mottagare, vilken information har mottagaren då fått om den ursprungliga signalen?

För att ta reda på detta skall vi beräkna spektrat för spänningssampelen.

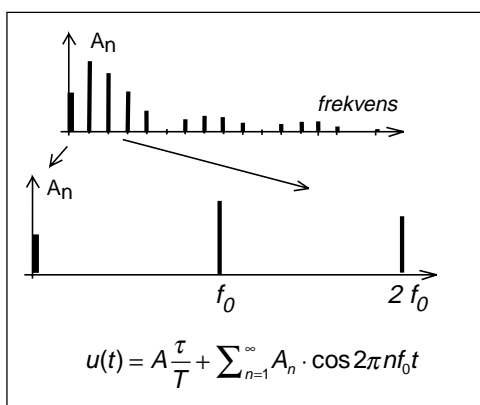


Spänningssampelen kan matematiskt beskrivas som $m(t)$ multiplicerat med en pulssignal bestående av mycket smala pulser.

Spänningssampelen kan beskrivas som den ursprungliga signalen $m(t)$ multiplicerat med smala pulser.

Spektrat för $m(t)$ känner vi, det är $M(f)$. Spektrat för oändligt smala spänningssampelen känner vi inte. Men spektrat för smala spänningssampelen har vi beräknat i avsnittet om Fourierserieutveckling.

De smala pulserna har ett spektrum bestående av en likspänningskomponent samt grundton och övertoner, där grundtonsfrekvensen f_0 är lika med samplingsfrekvensen.



Vi skall multiplicera funktionernas tidsuttryck. Detta motsvaras av att signalernas spektrer faltas. Men vi behöver inte lösa några faltningsintegraler. Faltningintegralen innebär att det bildas summa- och skillnadsfrekvenser. Det är samma sak som att se den pulsformade signalens frekvenskomponenter som "bärvågor". Resultatet blir DSB-modulering runt varje sådan bärvåg.

Dessutom får vi den ursprungliga signalens spektrum $M(f)$ multiplicerat med den pulsformade signalens likspänningskomponent A_0 .

Spektrat för spänningssampeln blir summan av alla dessa frekvenskomponenter.

Vikningsdistorsion (Aliasing distortion)

I figuren ser vi hur de olika spektrerna överlappar varandra. Det ser ut som om spektrat för $A_0 \cdot M(f)$ "viker" sig runt den frekvens som motsvarar halva samplingsfrekvensen. Detta kallas vikningsdistorsion (frequency folding).

Den signal som skall samplas får inte ha frekvenskomponenter med högre frekvens än halva samplingsfrekvensen. Annars uppträder vikningsdistorsion.

Samplingsteoremet

Samplingsteoremet formulerades av svensktätlingen Harry Nyquist år 1928 och innebär att:

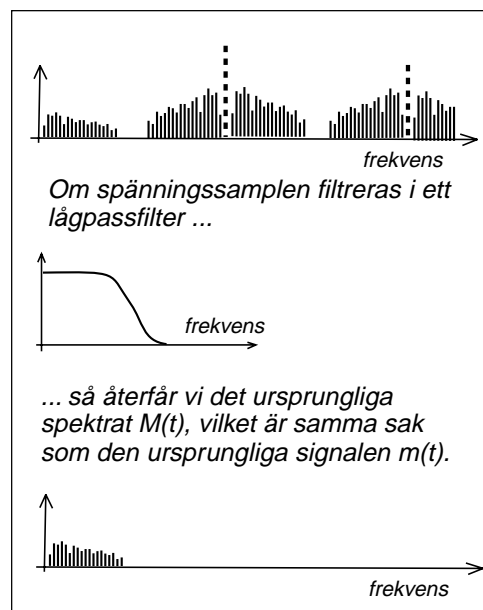
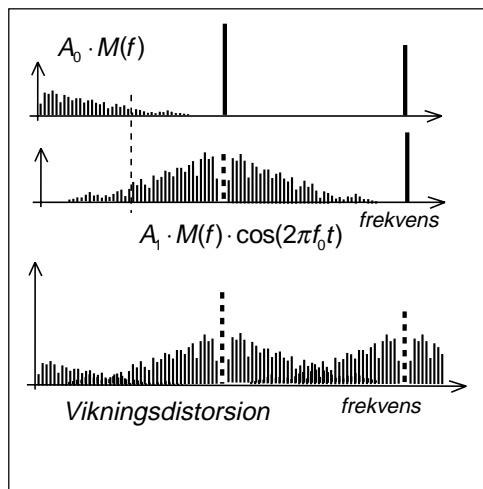
En signal $m(t)$ är entydigt bestämd av sina sampel, om samplingsfrekvensen är minst dubbelt så hög som den högsta frekvenskomponenten i signalens spektrum $M(f)$.

Det räcker med dessa spänningssampeln för att ha all information om signalen.

Digital signalbehandling

Sampling av analoga signaler blir allt viktigare. I takt med att digitaltekniken klarar att arbeta vid allt högre klockhastigheter så ersätts mycket av den analoga radiotekniken, filter, blandare m.m. av digitalteknik.

Radiosignalen A/D-omvandlas genom sampling. Därefter beräknas samplingspektrum med FFT, Fast Fourier Transform, en algoritm för snabb Fourierserietveckling



av korta tidselement. Filtrering sker genom att man bara tar med de frekvenskomponenter som man vill att filtret skall släppa igenom, blandning till andra frekvenser sker genom att subtrahera eller addera en fast frekvens till alla frekvenskomponenter, varefter man går tillbaka till tidsplanet med IFFT, Invers Fast Fourier Transform, och därefter skapar en analog signal av spänningssampeln i en D/A-omvandlare.

Det är detta man gör i digitala signalprocessorer.

9.3 PCM-kodning (Puls Code Modulation)

PCM-kodning av ljudsignaler används i många olika sammanhang. Vi skall titta på

- PCM-kodningen i fasta telenätet, 64 kbit/s-kodat tal.
- PCM-kodning för musik-cd.

1. Filtrering

Första steget är att filtrera informations-signalen så att den inte innehåller några frekvenskomponenter högre än halva samplingsfrekvensen. Detta är nödvändigt för att undvika vinkningsdistorsion.

- Filtret skär vid 3,5 kHz för telefoni.
- Filtret skär vid ca 20 kHz för musik-cd.

Efter filtreringen, när man avlägsnat de hög-frekventa frekvenskomponenterna, så blir signalens tidsfunktion lite mer "mjuk och avrundad".

2. Sampling

Nästa steg är att ta spänningssampel.

- 8000 sampel/s för telefoni.
- 44100 sampel/s för musik-cd.

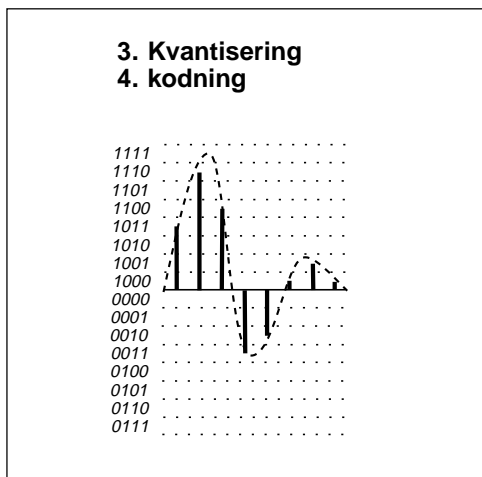
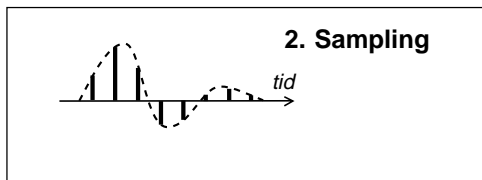
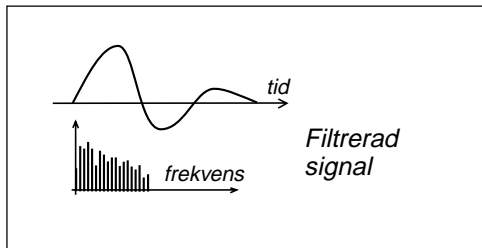
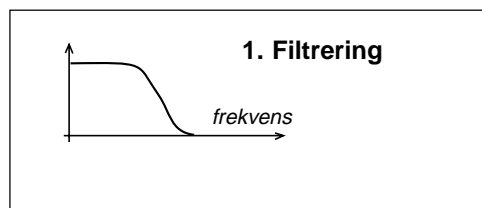
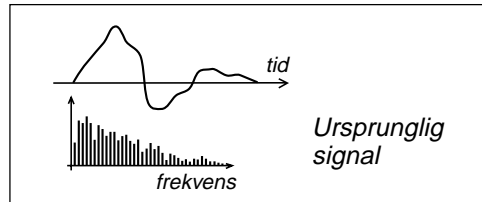
3. Kvantisering

Steket därefter har vi ännu inte behandlat. Spänningsnivån på ett sampel skall beskrivas med ett antal databitar. Men ett ändligt antal databitar kan inte beskriva vilken nivå som helst, utan bara vissa "trappsteg". Kvantiseringen innebär att bestämma vilket trappsteg som bäst beskriver det aktuella spänningssamplet.

4. Kodning

Kodning innebär att beskriva trappsteget med ett dataord.

- Vid telefoni används 8-bitsord. Detta ger 128 positiva och 128 negativa kvantiseringssteg (256 steg).
- Vid musik-cd används 16-bitsord vilket ger 32768 positiva respektive negativa kvantiseringssteg (65536 steg).



Att återskapa signalen

Vid telefoni skickas dataorden till mottagaren, och när man lyssnar på cd läses dataorden. Från dataorden skapas spänningssamplen.

Sedan återstår att mata signalen genom ett lågpasfilter som bara släpper igenom frekvenser lägre än halva samplingsfrekvensen, så har vi den ursprungliga signalen $m(t)$.

PCM-kodningens dynamik

Dynamiken är förhållandet mellan den starkaste signalens amplitud som omfattar alla kvantiseringssteg och den svagaste som omfattar ett enda kvantiseringssteg.

$$\text{dynamik}[dB] = 20 \cdot \log 2^n = n \cdot 6,02$$

Dynamiken vid musik-cd är 96 dB (16 bit).

Kvantiseringsbrus

Den återskapade signalens spänningssampel har inte exakt rätt nivå. De kan avvika med max ett halvt kvantiseringssteg. Denna avvikelse uppfattas som brus på den verkliga signalen. Det går att beräkna effektivvärdet av detta kvantiseringsbrus och få fram ett signal/kvantiseringsbrusförhållande:

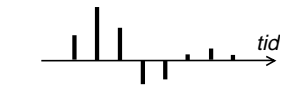
$$\frac{S}{Q}[dB] = 10 \cdot \log\left(\frac{3}{2} \cdot (2^n)^2\right) = 1,76 + n \cdot 6,02$$

Telefoni med A-lag-kompression

Vid PCM-kodning i telefoninätet så delar man inte in spänningssområdet i 256 lika stora kvantiseringssteg, utan stegen ligger tätare vid låg nivå och glesare vid hög nivå.

Vid hög nivå ligger kvantiseringsstegen så glest att man får kvantiseringsbrus motsvarande 6 bitars kodning. Detta ger ett kvantiseringsbrusförhållande på 37,9 dB. Vi skall senare se att det är detta brus som begränsar bithastigheten på våra 33,6 kbit/s-modem.

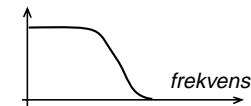
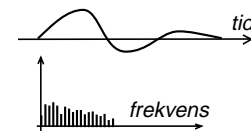
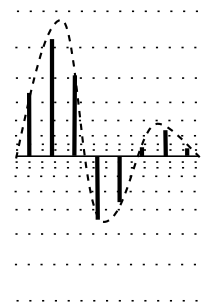
Vid låg nivå ligger kvantiseringsstegen så tätt att det motsvarar 12 bitars kodning av hela spänningssområdet. På detta sätt blir signal/kvantiseringsbrusförhållandet hyggligt även när man talar svagt i telefon.



Återskapa spänningssamplen



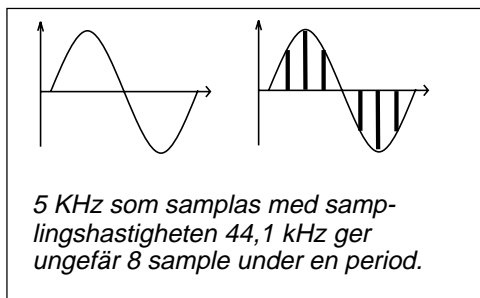
Spänningssamplen sett som spektrum matas genom LP-filtret ...

... och ut kommer den ursprungliga modulationssignalen $m(t)$.**A-lag och μ -lag**
Olinjär kvantisering **μ -lag används i USA och Japan**

I USA och Japan används en något avvikande kompressionskurva. Vid telefonsamtal mellan USA/Japan och övriga världen sker omräkning av kodorden så att mottagaren skall få rätt nivå på spänningssamplen.

9.4 PCM-kodning i verkligheten

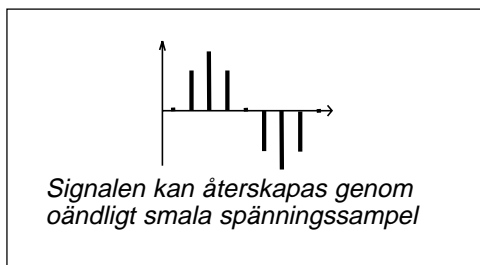
Vi har visat matematiskt, inte genom att räkna utan genom att visa vad matematiken gör, att om sampling sker med samplingshastighet som är minst dubbelt så hög som högsta frekvenskomponenten i signalen (Nyquist-villkoret), så har vi all information om signalen utan någon annan distorsion än de begränsningar som sätts av hur tätt kodningsstegen ligger. Med tillräckligt högt antal bitar i kodordet så blir det ingen distorsion. Hur är det i verkligheten?



D/A-omvandlingens spänningssampel

Hela teorin bygger på att signalen återskapas av spänningssampel som är oändligt smala. Men detta ger oändligt svag signal eftersom signalens amplitud är likspänningskomponenten av spänningssampelen.

Vi går åt andra hållet. Vi gör spänningssampelen lika breda som samplingsperioden. Detta är vanligt. Vi får hög utsignal. Men hur blir det med distorsionen?

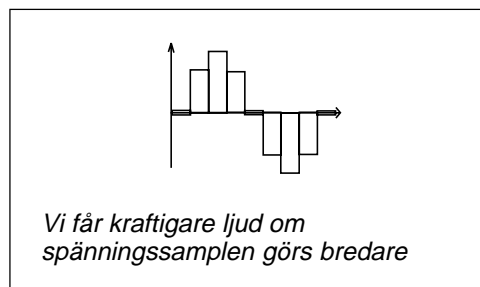


Hur låter 5 kHz på cd-skiva?

Vi utgår från 5 kHz. Detta är en hög frekvens men den förekommer som starka övertoner från vissa musikinstrument, t.ex. anslaget i en pianoton.

Först samplas signalen med 44,1 kHz samplingsfrekvens. Vi får ungefär 8 spänningssampel och vi antar att vi har så många bit i vårt kodord att det inte blir vare sig dynamikproblem eller kvantiseringsdistorsion.

Nu återskapar vi den ursprungliga signalen genom att bilda spänningssampel som är lika breda som samplingsperioden. Hur låter en sådan signal?



Skickas genom lågpasfilter

Den låter väl bra, tycker ni. Skicka signalen genom ett lågpasfilter så utjämnas "taggigheten" och signalen blir mjuk och rund. Förresten så har ju taggigheten en frekvens av 44,1 kHz. Det är ju ändå ingen som hör detta. Det kanske inte ens behövs något LP-filter.

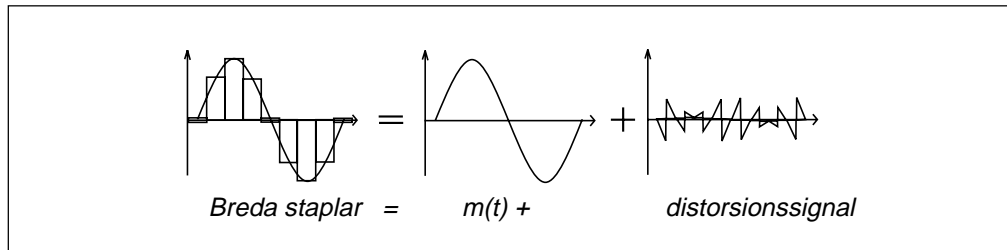
Tyvärr är det inte så enkelt. Vi kan inte bara säga att taggigheten försvinner om sig-

nalen LP-filtreras. Vi måste först beräkna spektrat för trappstegssignalen innan vi kan avgöra vad som kommer ut efter LP-filtret.

Hur beräknar vi spektrat?

Trappstegssignalen kan ses som summan av två signaler, dels sinussignalen på 5 kHz, dels skillnaden mellan sinussignalen och trappstegssignalen.

Skillnadssignalen är svår att beräkna matematiskt, men enkel att rita upp grafiskt. Om vi gör detta ser man att skillnadssignalen blir sågtandsformad. Dessutom blir amplituden på sågtänderna MYCKET HÖG. Vi har inget kvantiseringsbrus, men vi har brus från denna skillnadssignal som har amplituder som uppgår till 30 % av signalamplituden



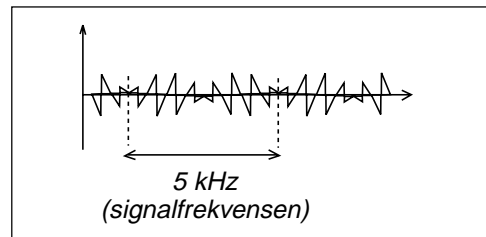
när signalfrekvensen är 10 % av samplingshastigheten.

Vad innehåller skillnadssignalen för frekvenskomponenter? Det är inte lätt att säga bara genom att titta på bilden, men en sak ser man klart och tydligt, nämligen att det finns en periodicitet i skillnadssignalen på 5 kHz. Ett lågpasfilter skär bort alla frekvenskomponenter över 20 kHz. Den del av distorsionssignalen som ligger på signalfrekvensen kommer inte att filtreras bort av lågpasfiltret.

Vi får alltså distorsion som är helt oberoende av kvantiseringen. Distorsionen ökar vid höga frekvenser och gör säkerligen att cd-ljud låter "kallt".

Kan inte distorsion undvikas?

Ett sätt att minska distorsionssignalens amplitud är att återskapa signalen med smala spänningssampel. Men då går spänningen ner till noll mellan varje spänningssampel och spektrat för den återskapade signalen kommer att innehålla DSB-signaler runt 44,1 kHz, 88,2 kHz o.s.v. Modulerad som DSB-signal finns inte bara den ursprungliga signalen som inte innehåller frekvenskomponenter över 20 kHz, utan även spektrat av den sågtandsformade distorsionssignalen



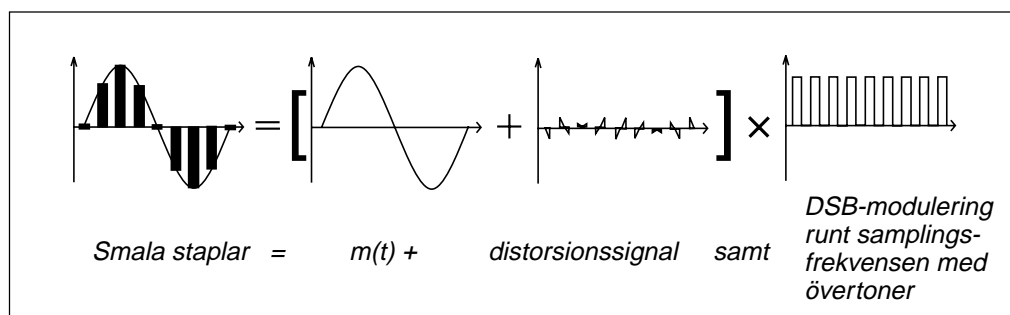
som ju inte är filtrerad. Höga frekvenskomponenter i distorsionssignalen hamnar inom det hörbara frekvensområdet, som vinkningsdistorsion från sidbanden runt 44,1 kHz, 88,2 kHz o.s.v.

Visst blir distorsionssignalen svagare om spänningssampelen görs smalare. Men även $m(t)$ blir svagare och distorsionen räknas i procent av $m(t)$.

Tätare spänningstaplar

Enda möjligheten att minska distorsionen vid hög frekvens är att se till att skillnaden i höjd mellan två närliggande spänningstaplar är så liten som möjligt.

Om skillnaden mellan två samplings-tidpunkter motsvarar fem kvantiseringsteg så bör denna spänningstapel delas upp i fem staplar som var och en ändrar sig ett kvantiseringsteg.



9.5 Symbolhastighet och bandbredd (Hartley)

Vi tänker oss serieporten på en pc. Ut på kabeln kommer "symboler" i form av spänningsnivåer. Ena symbolen betyder "1" medan andra symbolen betyder "0".

Symboler behöver bandbredd

Det är symbolerna som ger spänningsändringar på kabeln eller förändrar bärvågen, och det är ändringar som behöver bandbredd.

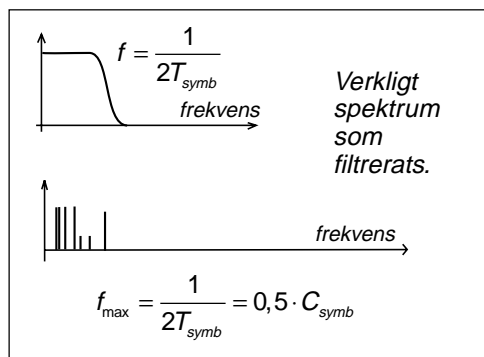
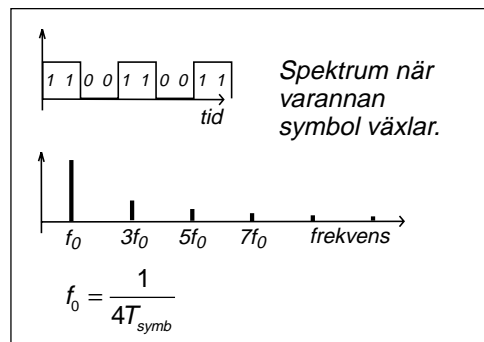
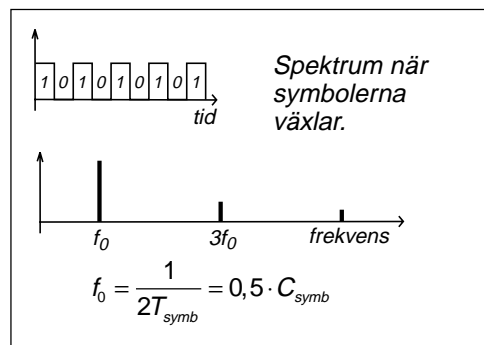
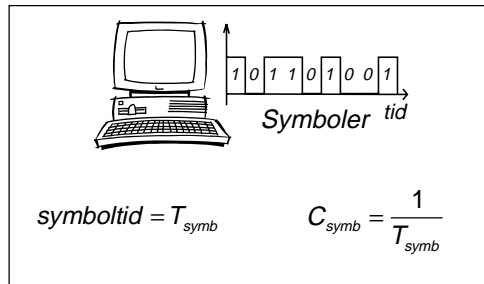
Vi skickar symboler på ledningen, ett visst antal symboler per sekund C_{symbol} . Denna signal har ett spektrum. Grundtonen i spektrat hamnar högst i frekvens när symbolerna hela tiden växlar. Då får vi en grundton i spektrat som hamnar på $f_0 = 0,5 \cdot C_{symbol}$.

Men vi kan även ha samma symbol två eller flera gånger i följd. Då hamnar grundtonen lägre i frekvens. En verklig datasignal där symbolerna kommer slumpvis får ett spektrum där man alltid får med åtminstone grundtonen i alla olika symbolkombinationer, om kanalen har en bandbredd på minst $B = 0,5 \cdot C_{symbol}$. Detta samband kallas ibland för Hartleys lag vid basbandskommunikation och säger att den maximala symbolhastigheten på viss bandbredd är lika med dubbla bandbredden.

Hartley (basband): $C_{symbol} = 2 \cdot B$

Symbolhastighet mäts i baud

Enheten för symbolhastighet är egentligen baud, men detta har så ofta blandats ihop med bithastighet bit/s att man numera i standardiseringsdokument från ETSI (GSM- och UMTS-standarderna) bara ser enheten utskrivna som "symbol/s" för att undvika missförstånd.



Radiokommunikation

Vid radiokommunikation måste man modulera en bärvåg. Vilken bandbredd behövs för att överföra basbandssignalen?

DSB

Vi kan amplitudmodulera i en balanserad blandare och få DSB. Som resultat får vi en bandbredd som är dubbelt så stor som modulationssignalens högsta frekvens. Vi får $B = C_{\text{symp}}$.

SSB

Vi kan minska på bandbreddsbehovet genom att bara sända ett sidband, använda SSB. Då får vi en bandbredd som blir lika med modulationssignalens högsta frekvens. Vi får $B = 0,5 \cdot C_{\text{symp}}$.

Bandbreddsbehov vid radiokommunikation

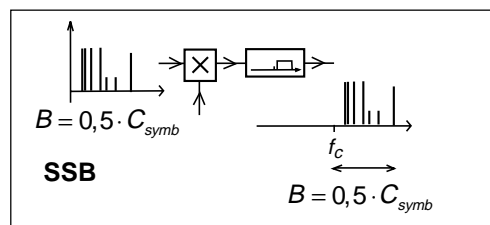
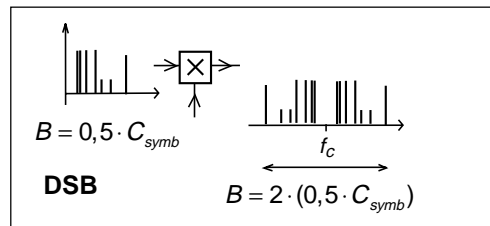
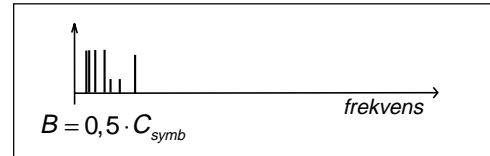
SSB ger lägsta bandbreddsbehovet. Vi klarar oss med $B = 0,5 \cdot C_{\text{symp}}$. Men den optimala mottagaren, den mottagare som har högst känslighet, en sådan mottagare använder koherent detektor (frekvensfel ger svävning som sänker detekterade signalen med -3 dB). Vid SSB finns ingen möjlighet att fastlåsa detektorn om man inte sänder bärvåg, vilket stjälar effekt från sidbandet.

I praktiken används DSB

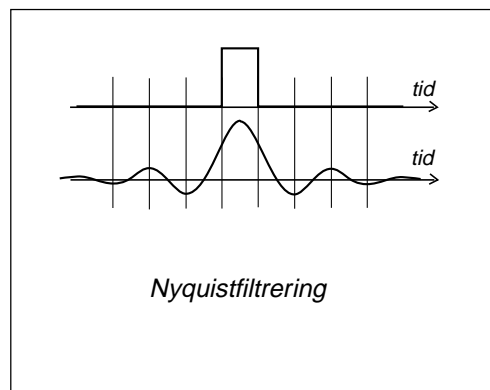
I praktiken används DSB, som visserligen kräver dubbla bandbredden, men fördelen med koherent detektor uppväger denna nackdel. Därför gäller vid radiokommunikation att den bandbredd som behövs är $B = C_{\text{symp}}$.

Hartley (basband): $C_{\text{symp}} = 2 \cdot B$

Hartley (modem- och radiokommunikation): $C_{\text{symp}} = B$

**Nyquistfiltrering**

Om en spänningspuls matas genom ett lågpasfilter så kommer pulsen att breddas och det uppstår ett "rippel" före och efter pulsen. Om lågpasfiltrets bandbredd är halva symbolhastigheten, d.v.s. Hartley (basband), så kommer nollställena i rippet att infalla när de andra symbolerna har maximum. Vi slipper "symbolinterferens", att symbolerna påverkar varandra. Filtrering för att undvika symbolinterferens brukar kallas Nyquistfiltrering, vilket i stort sett blir detsamma som Hartley för basband.



Från datorn kommer symboler med viss symbolhastighet C_{symp} .

Symbolernas spektrum

En bärvåg DSB-moduleras av symbolerna

$\cos 2\pi f_c t$

Modulerade signalens tidsfunktion och spektrum

frekvens

f_c

frekvens

DSB-signalen matas genom en transmissionskanal med för smal bandbredd.

$B < C_{\text{symp}}$

frekvens

frekvens

Inga spektralkomponenter passerar filtret. Därför finns ingen signal på utgången.

tid

DSB-signalen matas genom en transmissionskanal med tillräcklig bandbredd

$B \geq C_{\text{symp}}$

frekvens

frekvens

På utgången finns de spektralkomponenter som lyckas passera genom kanalen

Ett sidband lyckas passera filtret.

frekvens

frekvens

tid

Denna signal innehåller ingen information (ren sinus).

tid

tid

Signalen är förvrängd (sinusformad envelopp) jämfört med den utsända signalen, men informationen 10101 finns och kan tolkas med en koherent detektor.

9.6 Antalet olika symboler ger bithastigheten

I förra exemplet hade vi två utseenden på vår symbol. Ena utseendet representerade "1" och andra utseendet representerade "0".

Fyra olika utseenden på symbolen

Antag att vi har fyra olika utseenden på symbolen. Då kan dessa fyra utseenden symbolisera "11", "10", "01" och "00". Vi överför två databitar per symbol och får dubbelt så hög bithastighet som symbolhastighet.

$$C_{bit} = 2 \cdot C_{symb}$$

Med åtta olika utseenden på symbolen så överför vi tre bit/symbol, med 16 olika utseenden överför vi 4 bit/symbol, och med 1024 olika utseenden överför vi 10 bit/symbol.

Rent allmänt gäller:

$$C_{bit} = n \cdot C_{symb} = C_{symb} \cdot \log_2 M$$

$$n = \log_2 M$$

$$M = 2^n$$

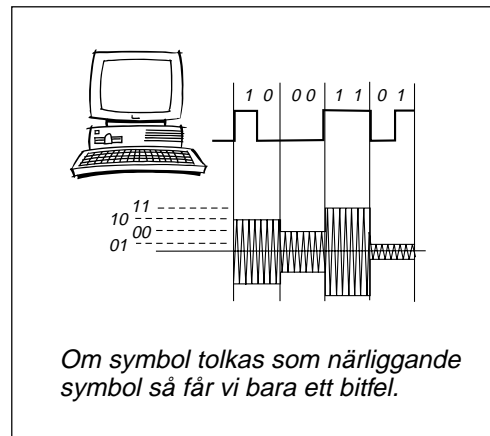
där

C_{bit} = bithastighet bit/s

C_{symb} = symbolhastighet symb/s, även Baud

n = bit/symbol

M = antal olika symboler



Exempel:

I fasta telenätet överförs frekvensområdet 300 – 3400 Hz. 33,6 kbit/s-modemet (V.34) använder en bärvåg på 1959 Hz som moduleras med 3429 symb/s. Detta ger något större bandbredd än 3,1 kHz, men filtren har sluttande flanker och man har prövat sig fram, gradvis höjt symbolhastigheten, och upptäckt att detta fungerar. Om detta inte fungerar

väljer modemmet en annan av nio bärvågor i intervallet 1600 – 2000 Hz och en annan symbolhastighet av sex i intervallet 2400 – 3429 symb/s. Man får dock lägre bithastighet än 33,6 kbit/s.

Symbolerna kan ha 1664 olika utseenden. Detta ger mer än 33,6 kbit/s, men de extra databitarna utnyttjas för bitfelsminimerande kodning (trelliskodning).

9.7 Bithastighet och brus (Shannon)

Vi kan öka bithastigheten utan att förändra symbolhastigheten, genom att ge symbolen fler utseenden. Men finns det ingen övre gräns?

Bruset begränsar

Hur ser signalen ut när den kommer fram? Det har adderats brus till signalen. Det som begränsar hur många olika utseenden symbolen kan ha är bruset på transmissionskanalen.

Shannons formel

Claude Shannon publicerade 1948 en uppsats som beskrev hur brus påverkar informationsöverföring. För en analog transmissionskanal som störs av Gaussiskt vitt brus, en AWGN-kanal (additive white Gaussian noise) gäller följande:

$$\begin{aligned} (C_{bit})_{\max} &= B \cdot \log_2 \left(1 + \frac{S}{N} \right) = \\ &= \frac{B \cdot \log \left(1 + \frac{S}{N} \right)}{\log 2} \end{aligned}$$

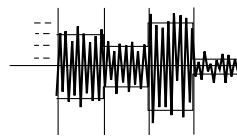
där

$$\begin{aligned} (C_{bit})_{\max} &= \text{maximal övre gräns [bit/s]} \\ B &= \text{kanalens bandbredd [Hz]} \\ \frac{S}{N} &= \text{signal/brusförhållande [ggr]} \end{aligned}$$

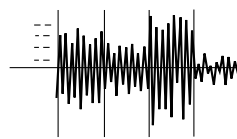
Exempel:

I analoga telenätet överförs frekvensområdet 300 – 3400 Hz, och signal/brusförhållandet, som begränsas av kvantiseringsbrus vid A/D-omvandlingen, är 37,9 dB. Beräkna den teoretiskt övre gränsen för bithastighet på en sådan kanal.

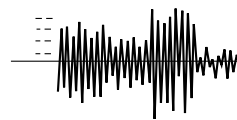
$$\begin{aligned} (C_{bit})_{\max} &= \frac{B \cdot \log \left(1 + \frac{S}{N} \right)}{\log 2} = \\ &= \frac{3100 \cdot \log(1 + 10^{\frac{37,9}{10}})}{\log 2} = \\ &= \frac{3100 \cdot \log 6167}{0,301} = 39 \text{ kbit/s} \end{aligned}$$



Signalen när den kommer fram.



Mottagaren skall inte bara avgöra vilken nivå som avses ...



Mottagaren vet inte ens när nivån skall mätas.

Shannon ger inga tips om hur denna bithastighet skall realiseras, utan värdet får ses som en teoretisk gräns som vi tror att vi inte kan överskrida.

10 — Fysiska lagret

Som inledning till modemkommunikation går vi först igenom hur fasta telenätet fungerar och kommer in på ISDN och ADSL.

Sedan är det dags för modemtekniken, MODulering och DEModulering.

Via signalvektorer och bitfelssannolikhet kommer vi fram till hur hög mottagen signal som behövs för viss bithastighet och upptäcker att den mottagna effekten, och därmed även sändareffekten, är direkt proportionell mot bithastigheten.

Efter att ha gått igenom amplitudskiftsmodulering (ASK), frekvensskiftsmodulering (FSK), fasskiftsmodulering (PSK) och QAM så upptäcker vi att skillnaden mellan de olika moduleringsätten egentligen är minimal.

10.1 Fasta telenätet

PSTN-nätet

PSTN (Public Switched Telephone Network) är det gamla analoga telefonnätet som bara är analogt mellan abonnenternas telefoner och telestationen. På telestationen finns en A/D-omvandlare som PCM-kodar talet. Det PCM-kodade talet (8 bit 8000 ggr/s) matas till växeln och vidare på det digitala rikslinjenätet som sammanbinder växlarna.

Analoga modem, V.34

När abonnenterna vill koppla upp sig till Internet så använder de "analoga modem". Med analogt modem menas ett modem avsett att anslutas till det analoga telefonnätet.

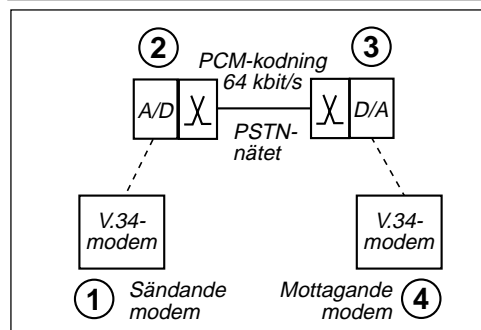
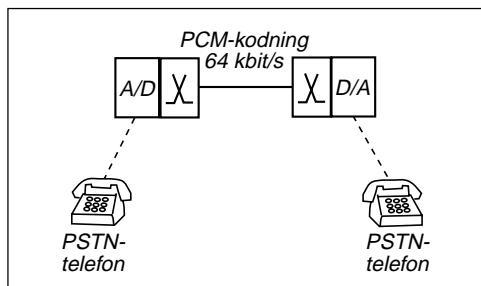
De analoga modemerna alstrar symboler (1) som skickas på kopparledningen till telestationen. Där A/D-omvandlas symbolen (2) genom PCM-kodning. Sedan skickas den digitala signalen hur långt som helst på det digitala nätet till mottagarens telestation, där man gör en digital/analogomvandling (3), d.v.s. alstrar spänningssamplarna som åter-skapar symbolen.

Vilka krav ställer man på den symbol som modemmet alstrar?

- Symbolen skall kunna transporteras på långa abonnentledningar, kopparkablar.
- Symbolen och symbolhastigheten skall ha ett spektrum som passar filtreringen och samplingen vid PCM-kodningen (2).
- Det mottagande modemmet (4) skall kunna se skillnad på så många olika symboler som möjligt så att vi kan skicka många bit per symbol.

Det mest effektiva med dagens teknik är att modulera en bärvåg. Man utnyttjar QAM, en form av dubbla DSB-modulatorer, som alstrar dubbla sidband.

Symbolhastigheten begränsas av bandpassfiltret vid (2), som släpper igenom 300 – 3400 Hz. Detta är emellertid ett ganska enkelt filter med sluttande flanker. Därför har man lagt bärvågen vid 1959 Hz och lyckats pressa symbolhastigheten till 3429



symp/s, vilket innebär att de intressanta delarna av sidbanden ligger inom 244,5 – 3673,5 Hz.

Antalet olika symboler bestäms av brusets på förbindelsen.

På kopparledningen från abonnent till telestation så adderas brus och störningar (överhörning från andra kopparledningar) till signalen, men inte mer än att signal/brusförhållandet uppgår till ca 60 dB.

Den kritiska punkten är PCM-kodningen vid (2). Avrundningsfelet vid kvantiseringen (kvantiseringsbruset) motsvarar ett signal/störförhållande på 37,9 dB. Det är detta avrundningsfel som begränsar hur många olika utseenden på symbolen som det mottagande modemmet kan särskilja.

När dataorden skickas på den digitala förbindelsen mellan telestationerna tillförs bara ett försumbart antal bitfel, och digital/analogomvandlingen, att återskapa spänningsstaplarna, ger inget större felbidrag. Sedan tillkommer brus på abonnentledningen till det mottagande modemmet, men detta brus ligger –60 dB under signalen.

ADSL

ADSL står för Asymmetric Digital Subscriber Line, och är en digitaliserad abonnentledning med högre bithastighet från telestationen mot abonnenten än från abonnenten mot telestationen. Det finns även utrustning med samma bithastighet i båda riktningarna, Symmetric DSL, SDSL. Därför använder man ibland samlingsnamnet xDSL.

Vid modemkommunikation via telenätet så måste det sändande modemmet alstra symboler vars spektrum ligger inom 300 – 3400 Hz, det frekvensområde som "släpps in" i telestationens A/D-omvandlare. Men själva abonnentledningen, kopparkabeln, kan transportera frekvenser från likspänning upp till den frekvens där dämpningen blir för hög. Detta har lett fram till ADSL-tekniken.

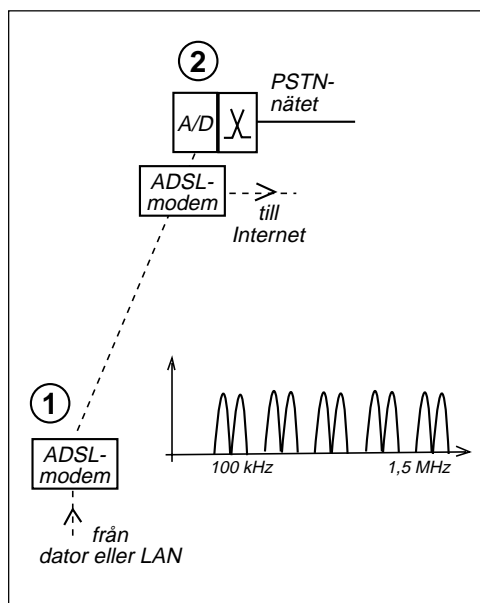
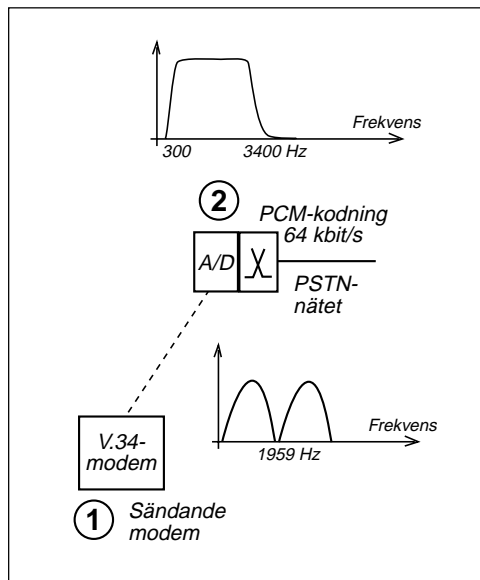
ADSL-modemet arbetar på samma sätt som V.34-modemet, med den skillnaden att bärvågen inte ligger vid 1959 Hz, utan uppåt 100 kHz. Och så lägger man ytterligare en bärvåg vid kanske 110 kHz, och en vid 120 kHz o.s.v. upp till så hög frekvens att signalen inte kommer fram på grund av dämpningen, som ökar med frekvensen. Varje bärvåg moduleras med t.ex. 4 ksymb/s, och vi kan ha många olika utseenden på symbolerna. Tillsammans ger detta bithastigheter på uppemot 6 – 8 Mbit/s vid kabellängder upp till ca 3 km. Vid kortare kabellängder, 500 m, klarar man 50 Mbit/s.

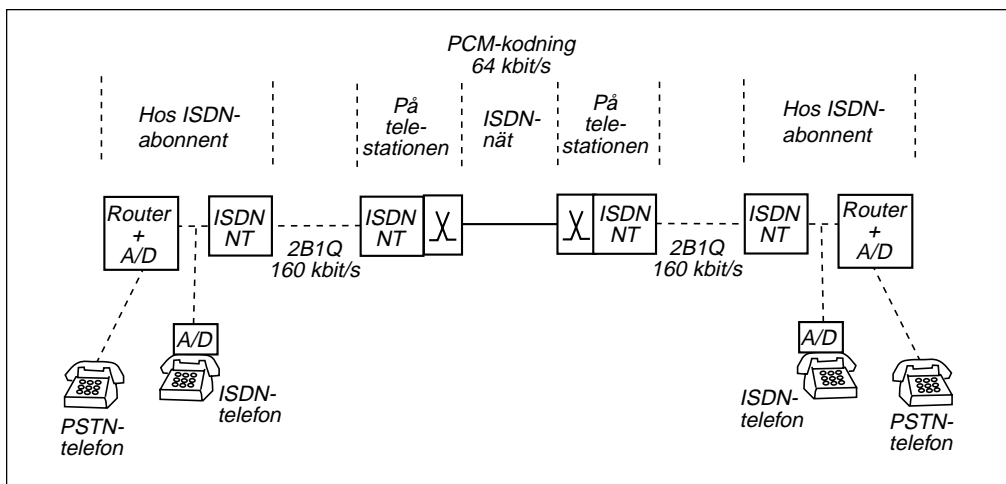
Alla siffervärden är ungefärliga. Varje tillverkare använder egna frekvenser och egen teknik. Överhörning mellan kopparpar i samma kabel kan göra att man inte kommer upp i dessa hastigheter. Nära telestationen ligger ofta 1000 kopparpar i samma kabel.

ADSL-signalen plockas ut från kopparledningen innan den når televäxels A/D-omvandlare. Eftersom ADSL bara utnyttjar frekvenser över det hörbara frekvensområdet så kan man samtidigt ha ett vanligt analogt abonnemang på samma kopparledning.

Multi Tone-teknik

Att utnyttja många bärvågor och på varje bärvåg ha en moduleringssteknik som liknar V.34-modemets kallas för "Multi Tone", många toner, många bärvågor.





ISDN-nätet

ISDN (Integrated Services Digital Network) är ett telenät där den PCM-kodade (64 kbit/s) signalen skickas ända hem till abonnenten. A/D-omvandlingen sker hemma hos abonnenten, i en ISDN-telefon eller, vilket är vanligast, i en ISDN-router som samtidigt används för att koppla en dator till det digitala ISDN-nätet.

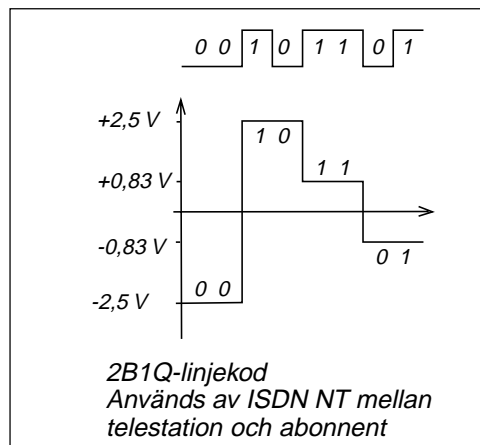
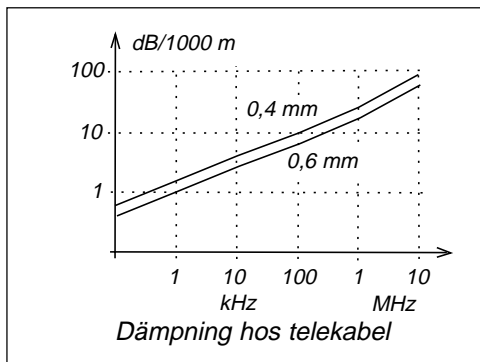
Linjekodning

När ISDN utvecklades, i början av 1980-talet, så fanns inte den modemteknik som utnyttjas vid ADSL. I stället fick man nöja sig med linjekodning, att överföra symboler som bestod av spänningspulser.

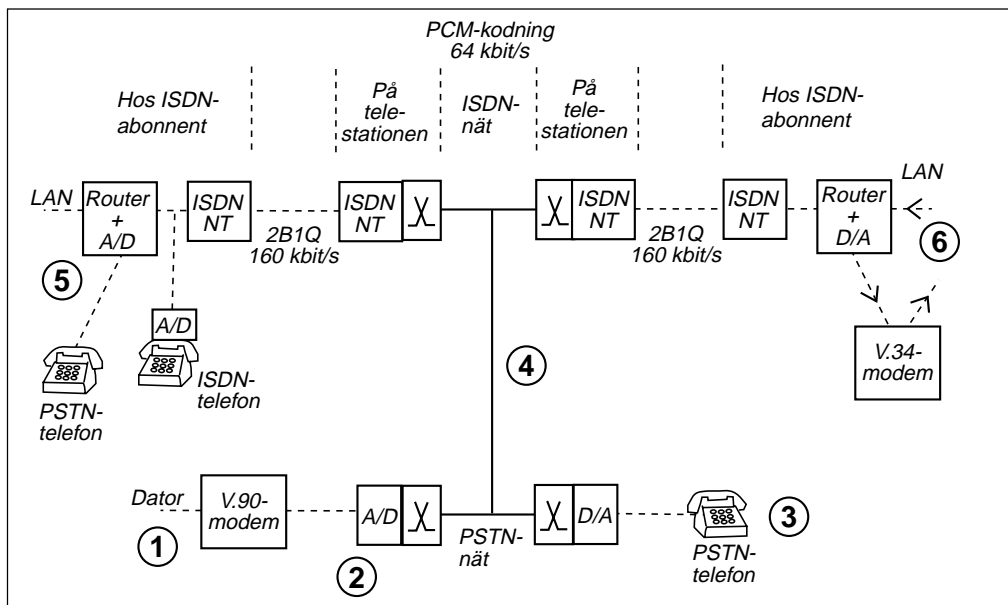
Vid ISDN skall 160 kbit/s skickas mellan telestationen och abonnentens nätterminal (NT). Hela ledningens frekvensområde fanns tillgängligt, med den begränsningen att hos ledningar ökar dämpningen med frekvensen. Den vanligaste kopparledningen i telenätet har ledare med 0,4 mm eller 0,5 mm diameter. Lägre symbolhastighet ger ett spektrum som ligger lägre i frekvens och inte påverkas lika mycket av kabeldämpningen.

I Sverige valde man en linjekod som kallas 2B1Q. Denna kod har fyra symboler. Symbolhastigheten blir 80 ksymb/s.

ISDN:s nätterminaler (NT) är 1980-talsteknik och kan utan vidare bytas mot ADSL-modem. Då skulle man utnyttja 160 kbit/s av



ADSL-kapaciteten för ISDN-abonnemanget och resten för fast uppkoppling mot Internet. Dessutom kan man ha ett vanligt analogt telefonabonnemang på samma kopparledning.



Hur man ringer i ISDN

En ISDN-abonnent vid (5) kan ringa till en PSTN-abonnent vid (3). A/D-omvandlingen sker hemma hos ISDN-abonnenten, anringen i en ISDN-telefon eller i den ISDN-router som används för Internetuppkoppling. Därefter skickas talet PCM-kodat 64 kbit/s via ISDN NT på kopparledningen till telestationen där den 2B1Q-linjekodade signalen tas emot och omvandlas till "vanlig" data-signal som matas till ISDN-stationen och vidare ut på ISDN-nätet.

Från ISDN-nätet kan man skicka telefonsamtal till PSTN-nätet. Där fortsätter det PCM-kodade talet till telestationen där digital/analogomvandling sker. Sedan matas spänningssamplen ut på kopparledningen och kan avlyssnas i PSTN-telefonen.

V.90-modem

Du har ett PSTN-abonnemang (1) och skaffar ett V.90-modem (56 kbit/s) för uppkoppling mot Internet. När din dator skall skicka symboler till din Internetoperatör ISP, Internet Service Provider (6), så måste symbolerna kunna PCM-kodas i din telestation (2). Därför är det effektivaste sättet att skicka symboler enligt V.34, d.v.s. 33,6 kbit/s. Dessa symboler skickas PCM-kodade till ISDN-nätet (4) och vidare till din ISP, som utnyttjar

en D/A-omvandlare för att göra analog signal av PCM-kodningen och mata den till ett V.34-modem som används för inkommande trafik.

Från din ISP skickas datasignalen i form av 8 bitsord, d.v.s. 56 kbit/s är kodad till 64 kbit/s med bitfelsminimerande kodning. ISDN-nätet tror att dessa 8 bitsord är ett PCM-kodat telefonsamtal och skickar dem via (4) till PSTN-nätet och telestationen. I telestationen (2) görs en D/A-omvandling, d.v.s. av varje 8 bitsord skapas en spänningsstapel som matas ut på kopparledningen till V.90-modemet.

I V.90-modemet finns en avancerad signalprocessor som klarar att nästan identifiera höjden på spänningsstaplarna, och med hjälp av den bitfelsminimerande kodningen återskapa de 7 bit som är kodade i de 8 bitsord som ligger bakom varje spänningsstapel.

Telestationens D/A-omvandlare (2) utnyttjas alltså för att göra en linjekodning med 8000 symb/s, 8 bit/symb (256 olika symboler). Detta går att tolka i V.90-modemet tack vare signalprocessorn. Att göra något liknande i andra riktningen fungerar inte ty PCM-kodaren i telestationen utgör en alltför osäker mottagare för "spänningsstaplar".

10.2 Linjekodning

Vi har sett två exempel på linjekodning, d.v.s. en teknik där symbolen består av en viss spänningsnivå i visst tidsögonblick.

ISDN:s NT

ISDN:s nätterminal utnyttjar enklast möjliga teknik. Databitarna kommer från länklagret i form av 8 bitsord (64 kbit/s) för de två telelinjerna (B1 och B2), och en 16 kbit/s bitström för signaleringskanalen (D).

Fysiska lagret sorterar in bitströmmarna i en ram bestående av 240 bit under 1,5 ms. Ramen inleds med 18 bit synkroniserings-signal som mottagaren måste känna igen för att veta var ramen börjar. Därefter följer 2 bit från signaleringskanalen samt ett 8 bitsord från vardera kretskopplade kanalen, något som upprepas 12 gånger. Ramen avslutas med 6 bit drift- och underhållsinformation, som talar om för ISDN-nätet att abonnentens NT "lever".

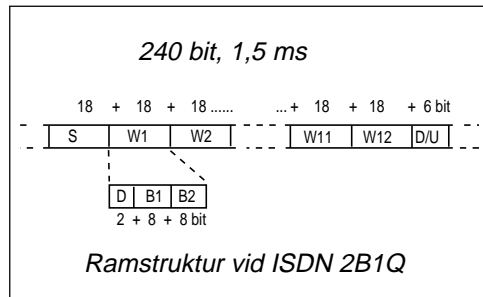
Protokollet innehåller ingen som helst teknik för att undvika bitfel. Hela kommunikationen bygger på att mottagaren måste klara att uppfatta symbolerna utan symbolfel.

V.90

Även V.90-modemet utnyttjar linjekodning i riktningen från telestationen ut mot abonnenten. Symbolerna består av de spänningsstaplar som skapas av digital/analog-omvandlaren. Men skillnaden i spänning mellan närliggande kvantiseringssteg är så liten att det är lätt att mottagaren tolkar fel. Dessutom, i de amerikanska och japanska PSTN-näten förekommer att man "lånar" en bit ur var 6:e 8 bitsord (var 48:e bit) för att skapa en signaleringskanal. Vi kan alltså få symbolfel därför att mottagaren tolkar fel, och vi kan få enstaka bitfel därför att tele-nätet "lånat" var 48:e bit.

Bitfelsminimerande faltningskodning

För att hantera detta så inför man en bitfelsminimerande faltningskodning för de databitar som bestämmer närliggande samlingsnivåer.



Spektrumformning

Från D/A-omvandlaren kommer 8000 spännings-sampel per sekund. Detta innebär att vi kan få starka frekvenskomponenter vid 4 kHz om två höga närliggande sampel har olika tecken (positiv respektive negativ spänning). Om mottagande modemet har svårt att tolka höjden på höga staplar som växlar tecken så snabbt, så utnyttjas "teckenbiten" i vissa 8 bitsord för spektrumformning. Man ger teckenbiten den polaritet som gör att signalens spektrum inte får alltför kraftiga högfrekventa komponenter.

Anpassning till kopparledningen

V.90-modemet arbetar med en ram bestående av 6 st 8 bitsord, totalt 48 bit som sänds på 0,75 ms. Dessa 48 bit är "framräknade" (bl.a. faltningskodade) av allt från 21 till 42 bit länkskiktsdata.

Vid 42 bit användardata blir bithastigheten 56 kbit/s. Men andelen länkskiktsdata kan minskas i steg om 1 bit ner till 21 bit på 0,75 ms. Då blir bithastigheten 28 kbit/s.

Vad vinner man på att fördela 21 bit under 48 utsända bit. Jo, kraftigare felminimerande kodning.

V.90-modemet arbetar hela tiden med 8 ksymb/s symbolhastighet. Man minskar inte heller på antalet symboler, t.ex. skulle man kunna använda bara varannat kvantiseringssteg. Nej, man fortsätter med hög bithastighet även om signalen är brusig och låter bitfelsminimerande kodning korrigera för den osäkerhet som bruset ger. Modemet väljer själv den högsta möjliga bithastigheten från 28 kbit/s till 56 kbit/s i steg om 1,333 kbit/s (21 steg).

10.3 Nyckling (Keying) och signalvektorer

Keying

Det finns ett ord som återkommer i alla förkortningar som gäller modulering, och det är "Keying". I radions barndom skickades telegram med morsetelegrafi, och sändaren "nycklades" (startades och stoppades, Keying) med en telegrafnyckel (Key).

Av modulering har vi behandlat amplitudmodulering, frekvensmodulering, fasmodulering och kvadraturmodulering. Alla dessa former återfinns vid digital modulering.

- ASK, Amplitude Shift Keying
- FSK, Frequency Shift Keying
- PSK, Phase Shift Keying
- QAM, Quad Amplitude Modulation

Signalvektorer

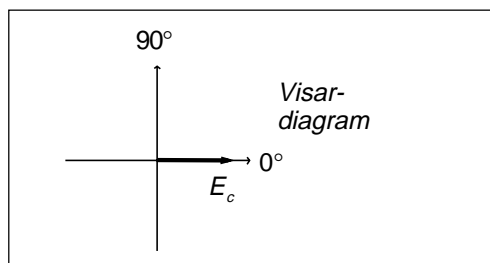
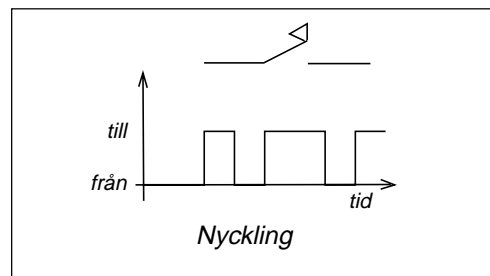
Vi är vana att rita $E_c \cos 2\pi f_c t$ i ett visar-diagram. Om denna signal datamoduleras med symboler som har symboltiden T_{symp} så kan vi beräkna energin i en symbol E_{symp} .

Vi ser även att roten ur E_{symp} är proportionell mot E_c . Roten ur E_{symp} kallas komplex signalvektor och kan ritas i ett vektordiagram.

Ordet vektor kan vara missvisande ur matematisk synpunkt. Signalvektorer adderas och subtraheras som vanliga vektorer, och vi kan multiplicera med skalär (förstärkning eller dämpning), men vi kan inte multiplicera två olika vektorer. Om vi däremot multiplicerar vektorn med sig själv så får vi energin under en symboltid.

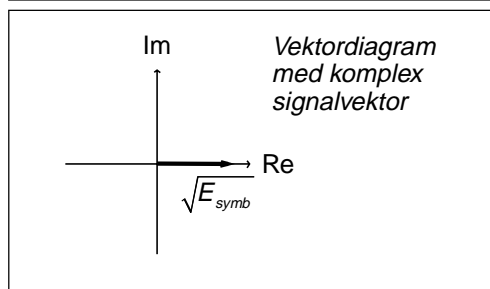
Egentligen E_s och E_b

Det normala är att skriva E_s , inte E_{symp} , och E_b , inte E_{bit} . Jag är övertydlig eftersom det är så lätt att blanda ihop symboler och bit.



$$E_{\text{symp}} = P_m \cdot T_{\text{symp}} = \left(\frac{E_c}{\sqrt{2R}} \right)^2 \cdot T_{\text{symp}}$$

$$\sqrt{E_{\text{symp}}} = E_c \sqrt{\frac{T_{\text{symp}}}{2R}}$$

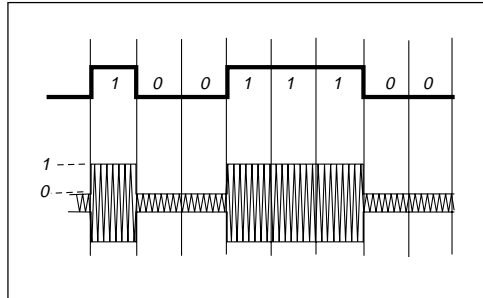


10.4 ASK (Amplitude Shift Keying)

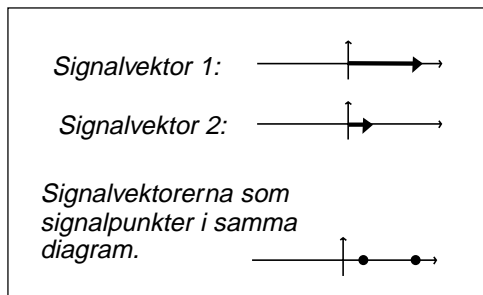
Vi kan ändra amplituden på bärvågen i takt med databitarna. Två symboler ger 2ASK.

Signalpunkter

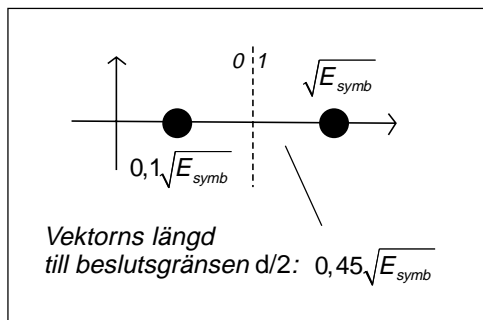
Nästa bild visar de två symbolerna som signalvektorer. Om man vill rita alla signalvektorer i samma diagram så blir diagrammet grötigt. Därför anger man signalvektorns spets med en signalpunkt. På så vis kan alla signalvektorer visas i samma diagram.



Den mottagna signalen är någon av signalvektorerna plus brus. Om bruset är så kraftigt att den mottagna signalen hamnar på andra sidan beslutsgränsen så får man symbolfel.



I kapitel 3 beskrev jag normalfördelat brus och Q-funktionen. Vad skall förhållandet vara mellan beslutsgränsen U och brusets effektivvärde E_n för att brustopparna skall överskrida beslutsgränsen under viss del av tiden? Som exempel räknar vi på bitfels-sannolikheten $p_{bit} = 10^{-4}$, där förhållandet blir ungefär 3,7.



Jag sätter E_{symb} till energin i den maximala symbolen. Avståndet mellan de två signalpunkterna kallar vi d . Avståndet till beslutsgränsen är alltså $d/2$.

$$d = 0,9 \cdot \sqrt{E_{symb}}$$

$$d/2 = 0,45 \cdot \sqrt{E_{symb}}$$

$$p_{bit} = p_{symb} = Q\left(\frac{U}{E_n}\right) = Q\left(\frac{d/2}{\sqrt{N_0 B \cdot T_{symb}}}\right) =$$

$$= Q\left(\frac{0,45 \sqrt{E_{symb}}}{\sqrt{N_0 B \cdot T_{symb}}}\right) = Q\left(\sqrt{\frac{0,2025 \cdot E_{bit}}{N_0}}\right)$$

Nyquistfiltrering

Vi räknar med optimal filtrering, rektangulär Nyquistkarakteristik, som visserligen inte är realiserbar, men för brusberäkningar blir resultatet ungefär som med ett verkligt filter (samma brusbandbredd, samma mottagen brusenergi). Detta ger $B = 1/T_{symb}$.

$$3,7 = \sqrt{\frac{0,2025 \cdot E_{bit}}{N_0}}$$

$$\frac{E_{bit}}{N_0} = \frac{E_b}{N_0} = 67,6 \Rightarrow 18,3 \text{ dB}$$

Ur detta kan vi räkna fram att denna detektor kräver $E_b / N_0 = 18,3 \text{ dB}$ för $p_{bit} = 10^{-4}$.

4ASK

Vi kan dela upp amplitudområdet på fyra olika symboler, vilket ger 4ASK.

I detta fall kommer två av symbolerna att kunna få symbolfel åt två håll. Därför blir det en enda svansyta under normalfördelningskurvan för de båda yttersta symbolerna, men två svansytor för de mellersta symbolerna som får symbolfel dubbelt så ofta. Medelsymbolfelen blir sex svansytor fördelade på fyra symboler.

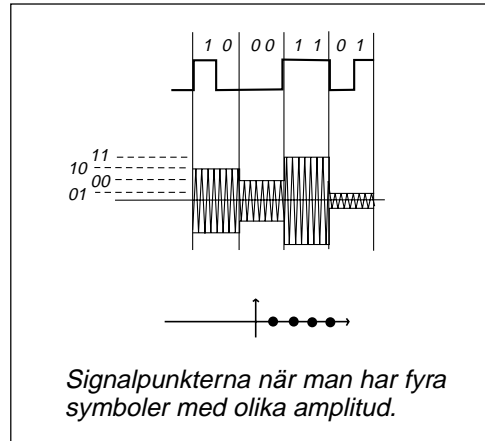
Bitfelsminimerande kodning

Om en symbol tolkas fel så att man tar den intilliggande i stället? Vi ser att jag fördelat databitarna på symbolerna så att intilliggande symbol bara ger ett bitfel av två överförda bit. En felaktig symbol på 5000 symboler motsvarar alltså ett bitfel på 10 000 bit.

0,5 dB vinst (10 % lägre sändareffekt)

Denna detektor kräver mer än fyra gånger högre E_b/N_0 för samma bitfelshalt jämfört med den tidigare detektorn med två symboler.

Den bitfelsminimerande kodningen ger en vinst på 0,5 dB.



$$\begin{aligned}
 p_{\text{symp}} &= \frac{6}{4} Q\left(\frac{U}{E_n}\right) = \frac{6}{4} Q\left(\frac{d/2}{\sqrt{N_0 B \cdot T_{\text{symp}}}}\right) = \\
 &= \frac{6}{4} Q\left(\frac{0,15\sqrt{E_{\text{symp}}}}{\sqrt{N_0 B \cdot T_{\text{symp}}}}\right) = \\
 &= \frac{6}{4} Q\left(\sqrt{\frac{0,0225 \cdot 2E_{\text{bit}}}{N_0}}\right) \\
 p_{\text{symp}} &= \frac{6}{4} Q\left(\sqrt{\frac{0,045 \cdot E_{\text{bit}}}{N_0}}\right)
 \end{aligned}$$

$$p_{\text{bit}} = 10^{-4} \Rightarrow p_{\text{symp}} = 10^{-4}$$

$$\begin{aligned}
 10^{-4} &= \frac{6}{4} Q\left(\frac{U}{E_n}\right) \\
 Q\left(\frac{U}{E_n}\right) &= 0,67 \cdot 10^{-4} \\
 \frac{U}{E_n} &\approx 3,85 = \sqrt{\frac{0,045 \cdot E_{\text{bit}}}{N_0}} \\
 \frac{E_b}{N_0} &= 329 \Rightarrow 25,2 \text{ dB}
 \end{aligned}$$

Utan bitfelsminimerande kodning

$$p_{\text{bit}} = 10^{-4} \Rightarrow p_{\text{symp}} = 2 \cdot 10^{-4}$$

$$\begin{aligned}
 2 \cdot 10^{-4} &= \frac{6}{4} Q\left(\frac{U}{E_n}\right) \\
 Q\left(\frac{U}{E_n}\right) &= 1,33 \cdot 10^{-4} \\
 \frac{U}{E_n} &\approx 3,65 = \sqrt{\frac{0,045 \cdot E_{\text{bit}}}{N_0}} \\
 \frac{E_b}{N_0} &= 296 \Rightarrow 24,7 \text{ dB}
 \end{aligned}$$

Med bitfelsminimerande kodning

10.5 Hur hög mottagen effekt P_m behövs?

Signal/brusförhållandet kan uttryckas som funktion av E_{symp} och C_{symp} . Bandbredden väljs som optimal Nyquistfiltrering, och bruset uttrycks som brus/Hz.

$$\begin{cases} \frac{S}{N} = \frac{P_m}{kTB} = \frac{E_{\text{symp}} \cdot C_{\text{symp}}}{kTB} \\ B = C_{\text{symp}} \\ kT = N_0 \end{cases}$$

$$\frac{P_m}{kTB} = \frac{E_{\text{symp}} \cdot C_{\text{symp}}}{kTB} = \frac{E_{\text{symp}}}{N_0} \cdot \frac{C_{\text{symp}}}{B} = \frac{n \cdot E_{\text{bit}}}{N_0}$$

$$P_m = n \cdot \frac{E_{\text{bit}}}{N_0} \cdot kTB = n \cdot \frac{E_{\text{bit}}}{N_0} \cdot kT \cdot C_{\text{symp}} = n \cdot \frac{E_{\text{bit}}}{N_0} \cdot kT \cdot \frac{C_{\text{bit}}}{n}$$

$$\therefore P_m = \frac{E_b}{N_0} \cdot kT \cdot C_{\text{bit}}$$

Vi ser att den effekt P_m som mottagaren behöver för att vi skall uppnå detektorkänsligheten E_b/N_0 endast är en funktion av detektorkänsligheten E_b/N_0 , antennens och mottagarens totala brustemperatur T och bithastigheten C_{bit} .

Medeleffekt

Om vi är intresserade av sändarens medeleffekt så kan vi antingen utgå från E_b/N_0 beräknad på en "medelbit", eller också kan vi räkna om till medeleffekt utgående från framräknad topp effekt P_m .

Det spelar ingen roll hur man gör så länge man vet vad man gör. Vad betyder egentligen E_b/N_0 om du hämtar ett värde ur en tabell?

Exemplet med de två ASK-modulationssätten visar att med fyra symboler så behöver mottagaren fyra gånger så hög mottagen effekt (och då behövs fyra gånger så hög sändareffekt) för samma bithastighet, men man behöver bara halva bandbredden (halva C_{symp}).

Priset för att minska bandbredden till hälften är alltså fyra gånger så hög sändareffekt.

Jag har samma räckvidd i båda fallen, men med fyra gånger så hög sändareffekt så "skräpar jag ner" större yta innan samma frekvenskanal kan återanvändas av någon annan.

Jag vann halva frekvensutrymmet men skräpade ner mer än dubbelt så stor yta. Alltså ingen vinst totalt sett.

Vid mobiltelefoni är det viktigare med lågt E_b/N_0 än smal bandbredd när man vill ha in så många samtliga samtal per ytenhet som möjligt på begränsat frekvensutrymme.

10.6 Antipodal 2ASK och BPSK

Störst detektorkänslighet får man när avståndet d är maximalt i signaldiagrammet, och med två signalpunkter så skall dessa ligga rakt genom origo, på diametralt motsatt plats, på antipoderna.

2ASK

Denna modulering kan kallas 2ASK, och då blir symbolerna lika, fast med minustecken.

Symbol 1: $E_c \cos 2\pi f_c t$

Symbol 2: $-E_c \cos 2\pi f_c t$

BPSK

Men fasmodulering med två faslägen, BPSK (Binary Phase Shift Keying) får samma signalpunkter. Då blir symbolerna:

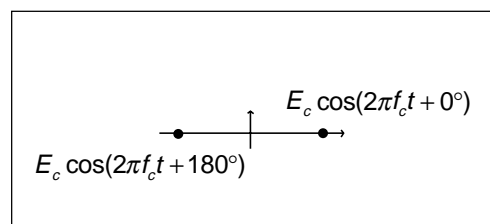
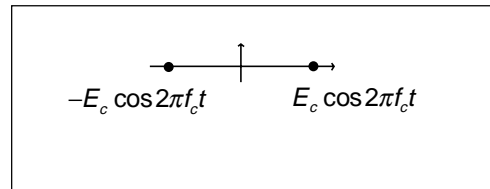
Symbol 1: $E_c \cos(2\pi f_c t + 0^\circ)$

Symbol 2: $E_c \cos(2\pi f_c t + 180^\circ)$

2ASK och BPSK har samma E_b/N_0

Beräkning av detektorkänsligheten följer samma mönster som förut. Avståndet d är avståndet mellan punkterna, och $d/2$ blir lika med symbolenergin, som är lika med bitenergin eftersom vi har en bit per symbol.

Det som är nytt är bruset. Det krävs en koherent detektor för att detektera denna signal. På samma sätt som vid DSB så kommer bara I- eller Q-bruset att detekteras och vi får bara halva bruseffekten efter detektorn.



$$p_{bit} = p_{symp} = Q\left(\frac{U}{E_n}\right) = Q\left(\frac{d/2}{\sqrt{\frac{1}{2}N_0B \cdot T_{symp}}}\right) =$$

$$= Q\left(\frac{\sqrt{E_{symp}}}{\sqrt{\frac{1}{2}N_0B \cdot T_{symp}}}\right) = Q\left(\sqrt{\frac{2 \cdot E_{bit}}{N_0}}\right)$$

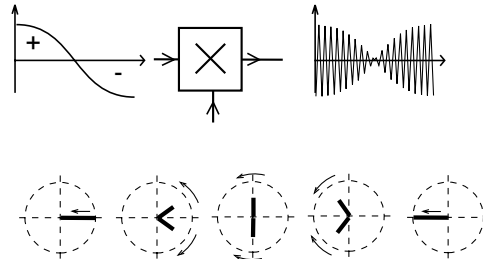
$$3,7 = \sqrt{\frac{2 \cdot E_{bit}}{N_0}}$$

$$\frac{E_{bit}}{N_0} = \frac{E_b}{N_0} = 6,8 \Rightarrow 8,3 \text{ dB}$$

2ASK

Modulationsformen 2ASK åstadkommes i en vanlig DSB-modulator. Utsignalen blir proportionell mot modulationssignalens amplitud och tecken.

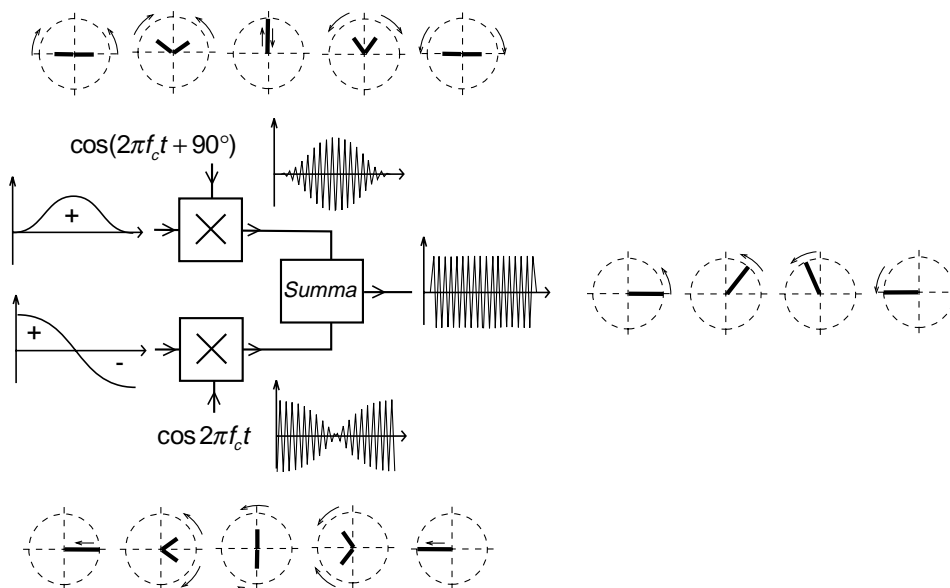
Förflyttning från ena signalpunkten till den andra innebär att bärvågen skall avta till noll och sedan växa med negativ amplitud. Förändring av amplitud är amplitudmodulering och kan åskådliggöras genom att dela signalen i två visare som vrider sig åt motsatt håll.



2ASK har lägre medeleffekt än BPSK

BPSK

BPSK innebär att bärvågen inte skall ändra amplitud. Förflyttning från ena signalpunkten till den andra skall ske längs cirkelns radie. Detta kan åstadkommas genom att man summerar två DSB-modulerade signaler i en QAM-modulator. Vi ser hur summan av visarna blir en enda visare med konstant amplitud, som vrider sig.



10.7 Icke-koherent 2FSK (Frequency Shift Keying)

Orthogonal mottagning

Om jag amplitudmodulerar en sändare (till/från) på frekvensen f_1 så kommer spektrat att gå genom noll på frekvensen $f_1 + C_{\text{symp}}$. Om jag lägger ett filter och en enveloppdetektor på frekvensen $f_1 + C_{\text{symp}}$ så får jag ingen utsignal. Spektralkomponenterna ger ingen amplitudvariation, bara fasvariation (olika tecken på spektralkomponenterna på ömse sidor om frekvensen $f_1 + C_{\text{symp}}$).

Frekvensskift

Om jag skapar två sändare, den ena på f_1 och den andra på f_2 och växlar mellan sändarna, ena symbolen sänds på f_1 och andra symbolen sänds på f_2 , så har jag skapat en orthogonal modulering, därför att ena sändaren ger utslag bara i en av mottagarna, inte i den andra mottagaren. Detta förhållande uppnås om frekvensskiftet är ett helt multiplicerat med C_{symp} .

Icke-koherent ortogonal 2FSK

Detta moduleringsätt kan alltså detekteras av en mottagare som inte återskapar någon bärvåg, en icke-koherent mottagare. Nackdelen med en sådan mottagare är att den detekterar både I- och Q-bruset. En sådan

mottagare får 3 dB högre E_b/N_0 än en koherent mottagare.

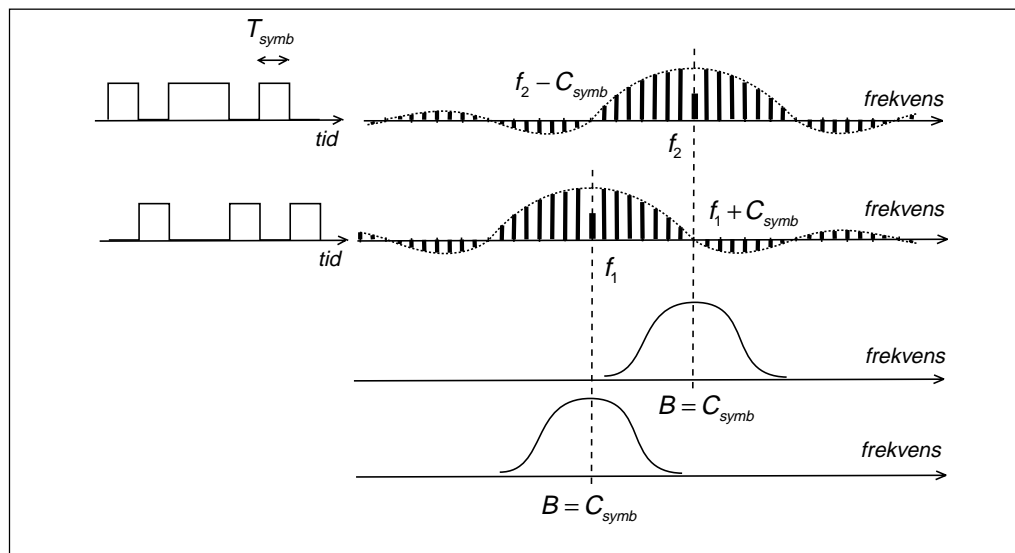
OFDM

Moduleringsättet är ortogonalt eftersom ett filter centrerat runt en av frekvenserna bara kommer att reagera för just den frekvensen.

Moduleringsättet som används vid digital-tv (DTV) och digital ljudradio (DAB) i marknätet, OFDM (Ortogonal Frequency Division Multiplex), innebär att man utnyttjar många bärvågor (1500 vid DAB, 8000 vid DTV) och frekvensskillnaden mellan bärvågorna är vald till C_{symp} .

Minst bandbredd om man sveper

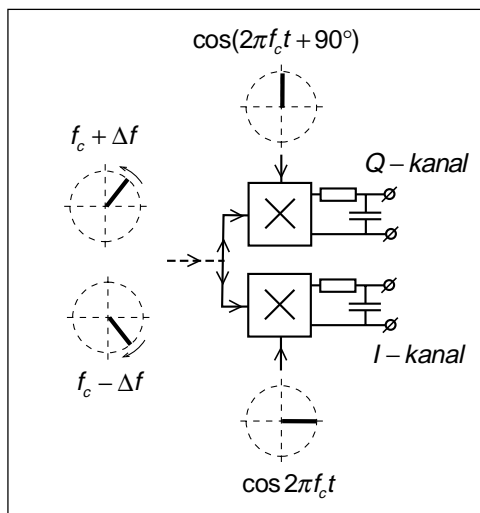
Man kan växla mellan två sändare, mellan två oscillatorer, vilket innebär att den ena stängs av och den andra startas (amplitudmodulering), men bättre är att ha en oscillator som FM-moduleras så att den "glider" $C_{\text{symp}}/2$ nedåt i frekvens på ena symbolen och glider tillbaka och upp $C_{\text{symp}}/2$ i frekvens på den andra symbolen. Då "hänger signalerna ihop", CPFSK (Continuous Phase FSK). Som alltid ger ett mjukt glidande, inga häftiga rörelser, det smalaste spektrat.



10.8 Koherent 2FSK, MSK (Minimum Shift Keying)

Koherent detektor

Vid FSK varierar frekvensen $\pm\Delta f$ runt f_c beroende på vilken symbol som sänds. När frekvensen är $f_c + \Delta f$ så kommer signalens fas relativt f_c att hela tiden öka. Om en sådan signal tas emot med en koherent detektor som återskapar bärvågen f_c , så kommer vi att få signal på I-kanalen som blir maximalt positiv varje gång fasen är 0° , och maximalt negativ när fasen är 180° , men ingen utsignal när fasen är 90° eller 270° . Då får vi i stället maximal positiv eller negativ signal på Q-kanalen.



Ortogonal modulering för koherent detektor

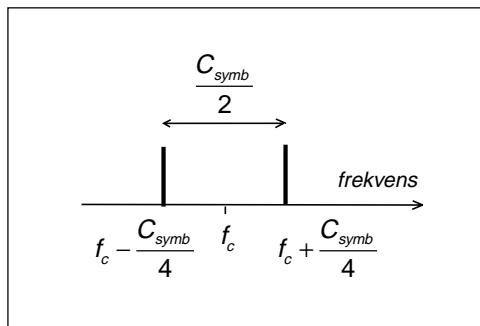
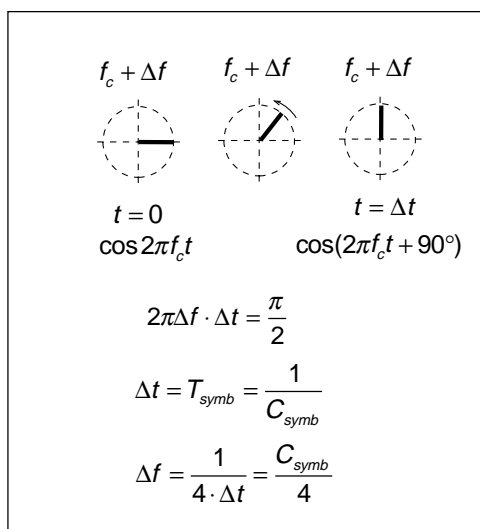
Om den ena symbolen avläses i det tidsögonblick när bärvågens fas snurrar förbi 0° eller 180° , medan andra symbolen avläses i det tidsögonblick när fasen snurrar förbi 90° eller 270° , så har vi ortogonal detektering, bara signal på den ena av utgångarna.

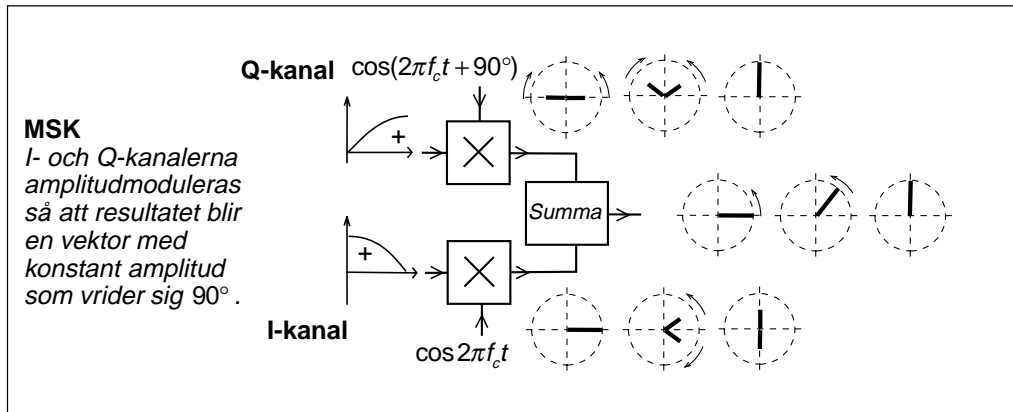
Hur lång skall symbolen vara?

Om vi sänder FSK så skall symbolen vara så lång, så att vi med frekvensskillnaden Δf hinner snurra 90° på en symboltid. Detta innebär att frekvensskillnaden Δf skall vara en fjärdedel av symbolhastigheten C_{symp} .

MSK (Minimum Shift Keying)

Om vi alltså använder en koherent detektor så kommer de båda symbolerna att vara ortogonala, om skillnaden mellan de båda frekvenserna är $C_{\text{symp}}/2$, hälften av frekvensskiftet vid icke-koherent FSK. Detta är det minsta frekvensskift som går att detektera ortogonalt. Därför kallas denna moduleringsform Minimum Shift Keying (MSK).



**GMSK (Gaussiskt filtrerad MSK)**

GSM-moduleringen är MSK, där signalerna som matas till I- och Q-modulatorerna är "avrundade" i filter med Gaussisk amplitudkurva. Detta ger mjuka ändringar hos de modulerande signalerna så att bandbredden inte blir större än nödvändigt.

Vid GSM är $C_{\text{symp}} = C_{\text{bit}} \approx 270 \text{ kbit/s}$. GSM-signalen ser ut att hoppa mellan två frekvenser. Den ena ligger på $f_c + 67 \text{ kHz}$, den andra ligger på $f_c - 67 \text{ kHz}$, d.v.s. en FSK-modulerad signal.

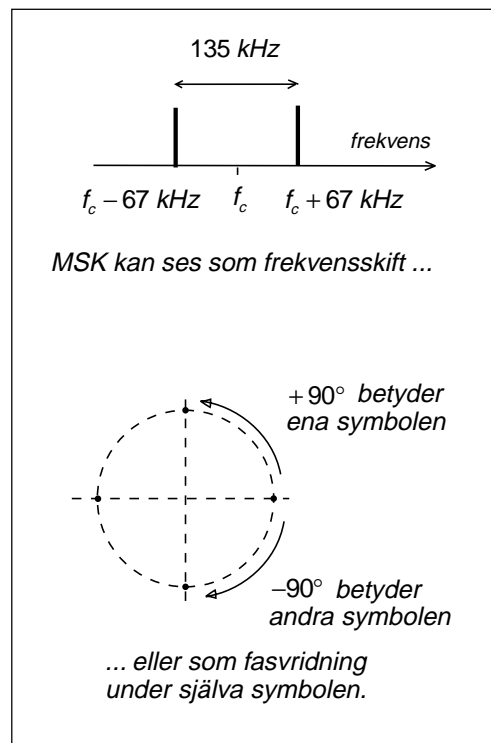
GSM-signalen alstras i en QAM-modulator. Detta är enda sättet att kunna garantera att fasvridningen blir exakt 90° under en symbol.

GSM-mottagaren känner av signalens fasläge i vissa bestämda tidsögonblick. Under ena symbolen vrider sig fasvinkeln $+90^\circ$, under andra symbolen vrider sig fasvinkeln -90° .

Att GSM alstras genom QAM-modulering innebär att man kan betrakta signalen som summan av två DSB-signaler som var och en moduleras med halva symbolhastigheten, ca 135 ksymb/s . Vi har alltså två sådana signaler på samma bandbredd. Det är detta som gör att GSM-signalen får plats i en 200 kHz bred frekvenskanal, trots att symbolhastigheten tycks vara 270 ksymb/s .

GFSK

Den modulationsform som utnyttjas av DECT, Bluetooth m.fl. kallas GFSK. Detta är GMSK som alstrats inte med QAM-



dulator utan genom att FM-modulera en oscillator. Detta är en förenklad modulationsform där man inte kräver exakt 90° fasvridning under en symbol. Man kan använda en enklare sändare och som mottagare någon enkel form av FM-detektor. Egentligen är detta ickekoherent FSK med för litet frekvensskift för att medge orthogonal detektering.

10.9 QAM (Quad Amplitude Modulation)

Med en QAM-modulator kan man åstadkomma vilken signalvektor som helst. QAM är därför ett modulationssätt där man kan sprida ut signalpunkterna optimalt i signaldiagrammet för att få så långt avstånd d till nästa signalpunkt som möjligt.

Telefonmodem 9,6 kbit/s (V.32)

Som exempel skall vi se på det gamla telefonmodemet för 9,6 kbit/s. Man använder symbolhastigheten 2,4 ksymb/s på telefonförbindelsen, överför 4 bit/symbol och behöver alltså 16 signalpunkter.

Signalpunkterna ligger tätare i mitten, vid låga amplituder än vid höga amplituder. Det beror på att bruset som adderas på telefonförbindelsen till övervägande del är kvantiseringsbrus, och eftersom kvantiseringsstegen (A-lag) ligger tätare vid låga amplituder så är kvantiseringsbruset lägre och därmed mindre störande vid låga amplituder.

Synkronisering

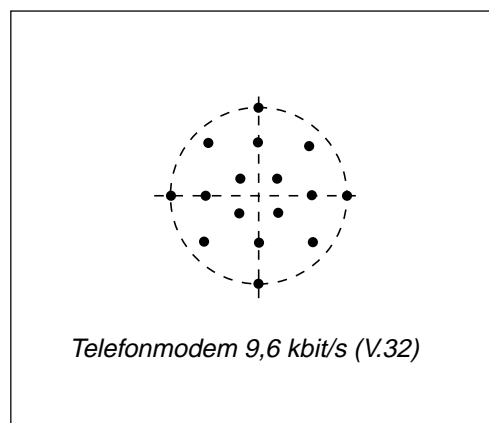
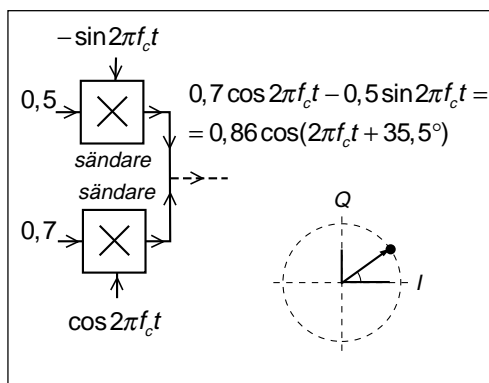
Signaldiagrammet ser ut på samma sätt om man vrider det ett kvarts eller ett halvt varv. Det är viktigt att sändande och mottagande modem är överens om vilken riktning som skall motsvara 0° innan kommunikationen börjar (synkronisering). Det är bl.a. sådan synkronisering som modemerna håller på med vid uppkoppling av en modemförbindelse.

Aven ett kort avbrott kan innebära att modemerna tappar synkroniseringen och måste börja om.

Vid mobiltelefoni, t.ex. GSM, så identifieras inte symbolen av punktens läge i diagrammet, utan av storleken på vridningen och vridningsriktningen. Vrids fasen åt ena eller andra hållet? Därför är GSM inte beroende av synkronisering av signaldiagrammen.

33,6 kbit/s-modem (V.34)

Signaldiagrammet för V.34-modemen ser ut på samma sätt som för V.32-modemet ovan, med skillnaden att vi inte har 16 signalpunkter utan 1664 punkter jämnt fördelade innanför cirkeln. Jag avstår därför från att rita någon bild.



10.10 PSK (Phase Shift Keying)

Fasskiftmodulering i betydelsen att signalen har ett visst fasläge under en symbol och sedan växlar till ett annat fasläge, något sådant existerar inte. En sådan modulering skulle innebära att signalen måste "hoppa" i sidled (på tidsaxeln) vid växling till nästa symbol. Man kan visserligen tänka sig att snabbt stänga av efter en symbol och starta vid nästa, men detta ger amplitudmodulering med fyrkantpulser och då bildas ett oändligt brett spektrum.

Vid växling från ett fasläge till ett annat, om man inte stänger av signalen, så måste sinuskurvan antingen "tryckas ihop" om nästa symbol skall ha positiv fas, komma tidigare, eller också måste sinuskurvan "dras isär" om nästa symbol skall ha negativ fas, komma senare. Trycka ihop innebär högre frekvens, dra isär innebär lägre frekvens.

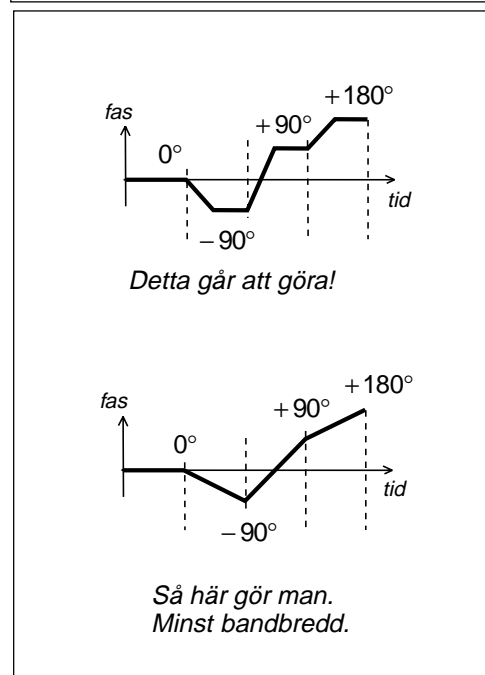
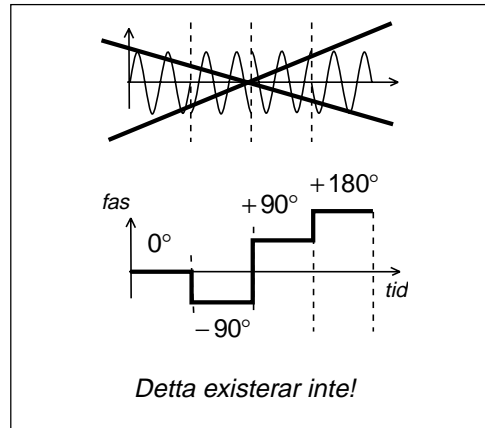
Om fasväxlingarna skall gå snabbt så krävs stor frekvensskillnad under växlingen. Frekvensskillnaden blir proportionell mot faskurvans derivata. Därför försöker man dra ut på tiden och låta växlingen från ett fasläge till nästa ta så lång tid som möjligt. Man utnyttjar hela symboltiden.

Egentligen FSK

Det som blivit kvar av fasskiftmoduleringen är egentligen frekvensskiftmodulering, eftersom man ändrar frekvens vid växling till nästa symbol beroende på hur fasen skall ändras. Sedan avläser man signalens fasläge vid slutet av symbolen.

PSK skall ha konstant bärvåg

Fasskiftmodulering innebär att signalen skall ha konstant amplitud, d.v.s. signalvisaren skall vrida sig med konstant längd mellan faspunkterna. Men bandbredden på signalen blir mindre om signalvisarens spets går raka vägen till nästa signalpunkt.



EDGE använder 8PSK

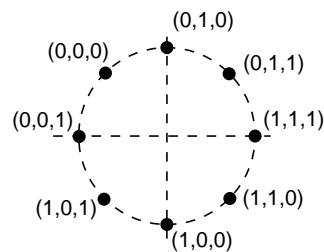
En vidareutveckling av GSM har fått benämningen EDGE (Enhanced Data rates for GSM and TDMA/136 Evolution).

För att öka bithastigheten med oförändrad symbolhastighet (270 ksymb/s) så har man infört åtta signalpunkter i signaldiagrammet. Punkterna ligger utspridda längs diagrammets cirkel.

Vi ser att signalpunkterna ligger utspridda längs cirkeln. Åtta punkter. Därför kallas detta 8PSK. Men signalen följer inte cirkeln mellan symbolerna, utan går i stort sett kortaste vägen. I verkligheten är detta alltså QAM-modulering.

Gray-kodning

Varje signalpunkt representerar tre databitar. Man har valt bitkombinationer så att skillnaden mellan två närliggande punkter bara är en enda bit. Om man får symbolfel och tolkar signalen till närliggande symbol så ger detta inte symbolfel med tre bitfel, utan bara ett enda bitfel. Två av databitarna är fortfarande rätt mottagna. På detta sätt håller man nere bitfelshalten till ett minimum.



8PSK med Gray-kodning

(I verkligheten vrids signaldiagrammet $67,5^\circ$ mellan varje symbol för att undvika att signalvisarens spets går genom origo från en punkt till en annan, men vi bortser från detta.)

11 — Länklagret

Efter att vi tittat på kontrollsummor, d.v.s. CRC-beräkning, och enkel felrättning med Hammingkodning, så kommer vi till länklagrets egentliga uppgifter.

Vi börjar med V.24, hårdvaruhandskakning, så som länklagret hanteras mellan ett telefonmodem och pc:n.

Därefter tittar vi på SDLC, mjukvaruhandskakning, så som länklagret hanteras mellan två modem, eller mellan två datorer anslutna till samma LAN (Local Area Network).

Som avslutning följer lite allmänt om kretskopplade, paketförmedlande och cellförmedlande nät, samt skillnaden mellan ett förbindelseorienterat och ett förbindelseöst paketnät.

11.1 Kontrollsumma, CRC-beräkning

Något sådant som bitfelsfri kommunikation existerar inte. Hur stark mottagen signal man än har, hur bra signal/störförhållandet än är, så uppträder bitfel. Mer eller mindre ofta.

Kontrollsumma

Hur kontrollerar man om det finns bitfel? Utan att behöva sända hela meddelandet igen? Man räknar fram en kontrollsumma vars värde påverkas om man fått bitfel.

Personnumrets kontrollsiffra

Personnumrets sista siffra är en kontrollsiffra. Den är framräknad på följande sätt:

- Siffrorna för år, månad, dag och nummer multipliceras växelvis med 2 och 1.
- Lägg samman siffrorna i produkterna. OBS! 12 räknas som 1 + 2.
- Entalssiffran i summan dras från talet 10. Detta blir kontrollsiffran.

Genom att multiplicera växelvis med 2 och 1 samt även ta med kontrollsiffran så skall summan bli helt tiotal om personnumret stämmer.

Varför multiplicera med växelvis 2 och 1? Jo, detta gör man för att kunna upptäcka om två siffror är omkastade.

CRC eller FCS

En liknande metod används vid datalagring och datakommunikation. Kontrollsumman kallas CRC (Cyclic Redundancy Check) efter metoden att räkna fram kontrollsumman, eller FCS (Frame Check Sequence) eftersom kontrollsumman kontrollerar ett datablock (Frame), eller BCS (Block Check Sequence).

ARQ (Automatic Repeat Request)

Vid datalagring (diskett m.m.) och datakommunikation, vid överföring av datafiler m.m. så finns kravet att alla databitar måste vara rätt. Men helt bitfelsfri läsning eller kommunikation existerar inte. Hur gör man? Jo, man räknar fram en kontrollsumma, på samma sätt som vid personnumret. Om kontrollsumman inte stämmer så gör man

$$\begin{array}{r}
 \overline{6\ 4}\ \overline{0\ 8}\ \overline{2\ 3}\ \overline{3\ 2\ 3}\ x \\
 \underline{2\ 1\ 2\ 1\ 2\ 1\ 2\ 1\ 2} \\
 12\ 4\ 0\ 8\ 4\ 3\ 6\ 2\ 6 \\
 \\
 1+2+4+0+8+4+3+6+2+6 = 36 \\
 x = 10 - 6 = 4
 \end{array}$$

på exakt samma sätt som vid personnumret:
Man frågar om!

Databitarna som skall skyddas med kontrollsumma delas in i block. Dessa block kan vara stora om sannolikheten för bitfel är liten (ibland ända upp till 8000 oktetter), eller små, kanske bara några oktetter, om risken för bitfel är hög.

Blocket betraktar vi som ett stort tal. Detta tal divideras med ett så kallat "generatorpolynom". Som resultat får vi kvoten (heltalsdelen) och en eventuell rest. Denna rest (CRC) adderas till blocket på sådant sätt att motsvarande division av blocket plus CRC går jämnt ut. Vi får bara heltalsdel.

Hela räkneoperationen utförs som "Modulo-2"-division. Jag nöjer mig med att visa hur man gör.

I mottagaren

Mottagaren utför samma modulo-2-division men med CRC i stället för nollor. Om divisionen går jämnt ut, d.v.s. om vi bara får nollor som rest, så är sannolikheten stor att samtliga databitar är rätt mottagna. Det finns dock en mikroskopisk chans att flera databitar är fel på sådant sätt att divisionen fortfarande går jämnt ut.

Om divisionen inte går jämnt ut så finns bitfel. Vi kan emellertid inte säga vilken eller vilka databitar som är fel. Vi kan heller inte säga hur många databitar som är fel. Bara att det finns bitfel.

Att räkna med X

Generatorpolynomen brukar inte presenteras som en sträng av ettor och nollor, utan med hjälp av X upphöjt till de positioner, från 0 till n-1, där det finns ettor. Det generatorpolynom som vi använt, 1010, skrivs som $X^3 + X$ eftersom det finns ettor i position tre och ett, men nollor i position två och noll. Ibland ser man enbart X, ibland X^1 . Däremot så skriver man inte X^0 utan 1.

Även blocket kan skrivas med X:

$$X^4 + X^2 + X^1 + 1$$

Om vi lägger till tre nollor så skiftas ettornas positioner och vi får:

$$X^7 + X^5 + X^4 + X^3$$

Detta kan vi använda i vår division.

Vissa tycker att det är enklare att räkna så här. De flesta rekommenderar dock den förra metoden.

1010	$ \begin{array}{r} 10011 \\ \hline 10111+110 \\ 1010 \downarrow \\ \hline 00011 \\ 0000 \downarrow \\ \hline 0011 \ 1 \\ 000 \ 0 \downarrow \\ \hline 011 \ 11 \\ 10 \ 10 \downarrow \\ \hline 01 \ 010 \\ 1 \ 010 \\ \hline 0 \ 000 \end{array} $
	Endast nollor innebär rätt mottaget ↑

Bild 5:
Mottagaren utför samma modulo-2-division men med CRC i stället för nollor. Då går divisionen jämnt ut om alla databitar är rätt mottagna.

Block: $10111 = X^4 + X^2 + X^1 + 1$
Generatorpolynom: $1010 = X^3 + X^1$
$ \begin{array}{r} X^4 + X^1 + 1 \\ \hline X^3 + X^1 \left \begin{array}{l} X^7 + X^5 + X^4 + X^3 \\ X^7 + X^5 \end{array} \right. \\ \hline X^4 \quad \quad \quad \downarrow + X^2 \\ X^4 \quad \quad \quad \downarrow + X^2 \\ \hline X^3 + X^2 \\ X^3 \quad + X^1 \\ \hline X^2 + X^1 = 110 \end{array} $
Sänt ord: $X^7 + X^5 + X^4 + X^3 + X^2 + X^1 = 10111 \ 110$

Generatorpolynom i verkligheten

CRC används vid lagring på diskett och hårddisk. En sektor på 512 oktetter skyddas av 16 bit CRC. Visst har du sett "CRC error" när diskettenheten inte klarat att läsa rätt från disketten?

Vid kommunikation med V.34-modem så utnyttjas ett omfrågningsprotokoll V.42 som begär omsändning vid bitfel. V.42 kan välja mellan två typer av CRC, 16 bit eller 32 bit.

Talkodaren i GSM skyddar ett antal extremt viktiga bitar med CRC. Om dessa databitar är fel mottagna så kastar GSM-telefonen 20 millisekunder tal och upprepar det tidigare 20 ms-blocket.

I GSM:s ursprungliga talkodare (FR, Full Rate) skyddas 50 bit med en CRC på 3 bit. Generatorpolynomet består alltså av fyra bit: 1011.

Den moderna GSM-talkodaren EFR (Enhanced Full Rate) skyddar 65 bit med en CRC på 8 bit.

Blocken kan vara hur långa som helst

Tekniken vid CRC-beräkning, att dividera och spara resten, innebär att blocket kan vara hur långt som helst.

Eftersom man begär omsändning av hela blocket om CRC inte stämmer, så får blocket inte vara större än att majoriteten av blocken kan överföras bitfelsfritt.

Vid TCP-protokollet (Internetkommunikation) skyddas ett block på 1500 oktetter med CRC. Om det blir CRC error så begär TCP-protokollet omsändning från värddatorn.

När Internetförbindelsen bärs av ett ATM-nät så delas blocken ner i ATM-celler med 47 okteters last i varje. Om en ATM-förbindelse överbelastas, d.v.s. om abonnenterna försöker mata in ATM-celler tätare än vad ATM-förbindelsen kan svälja, så kastar ATM vissa celler. Det kommer att saknas någon grupp om 47 oktetter i ett TCP-block på 1500 oktetter. TCP begär omsändning. Om det då fortfarande saknas 47 oktetter så fortsätter TCP att begära omsändning. Förbindelsen står och stampar, trots att majoriteten av oktetterna är rätt mottagna. Detta undviks om CRC beräknas över mindre block.

V.42 med 16 bit CRC:

$$X^{16} + X^{12} + X^5 + 1$$

V.42 med 32 bit CRC:

$$X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + \\ + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1$$

GSM:s äldre talkodare (FR),
3 bit CRC:

$$X^3 + X + 1$$

GSM:s nyare talkodare (EFR),
8 bit CRC:

$$X^8 + X^4 + X^3 + X^2 + 1$$

Vid mobiltelefoni

Vid radiokommunikation med rörliga terminaler, mobiltelefoni, så är risken för bitfel mycket hög. Men tal och bild kan ofta tolkas rätt även om det är enstaka bitfel. GSM kastar 20 ms tal då och då när det blir för många bitfel, men vi uppfattar ändå vad som sägs.

När man skall skicka datafiler över en GSM-förbindelse så är kravet att inte en enda databit får saknas och att alla databitar skall vara rätt mottagna. För att klara detta så behövs CRC kontrollsumma som omfattar samtliga databitar, och en mottagare som begär omsändning om det blir CRC-error.

I GSM kallas kontrollsumman för BCS (Block Check Sequence), och varje block består av de databitar som skall överföras på 20 ms. Vid 14,4 kbit/s blir det 288 bit.

På 20 ms skickar GSM ett radioblock på 456 bit. Skillnaden mellan 288 bit och 456 bit utnyttjas för CRC och bitfelsminimerande kodning. Om det blir för många omfrågningar så är det effektivare att övergå till 9,6 kbit/s med 192 bit. Detta ger fler bitar till bitfelsminimerande kodning, bitfelen minskar och man får färre omfrågningar.

11.2 Felrättning, Hammingkodning

CRC klarar att identifiera om det finns bitfel. Om man dessutom kan identifiera vilka databitar som är fel så skulle man kunna rätta dessa bitfel. En felaktig etta måste ju vara en nolla. Då slipper man begära omsändning.

Blockkoder

Det finns blockkoder som liknar CRC, som kan användas för att identifiera vilka bitar som är fel. En sådan blockkod används på FM-radions RDS-kanal.

RDS-kanalen (Radio Data System) är en datakanal med symbolhastigheten 1187,5 symb/s (multipel av 19 kHz stereopiloten). Databitarna är indelade i block där ett block om 16 bit skyddas av 10 bit. Denna kodning är så effektiv att av dessa 26 databitar kan 5 vara felaktiga, och de kan ligga i "skur", d.v.s. efter varandra. Om fler än 5 databitar är felaktiga så fungerar skyddsblocket som vanlig CRC. Vi får reda på att det finns bitfel som inte kan rättas.

Ett kort avbrott vid radiokommunikation ger oftast skurfel, eftersom man missar flera databitar i följd. Felrättande kodning är nödvändig. FM-mottagaren har ingen kontakt med FM-sändaren och kan inte begära omsändning.

Hammingkodning

Vi skall inte fördjupa oss i blockkoderna utan titta på ett exempel, Hammingkodning. Ett dataord om 7 bit kompletteras med 4 bit på sådant sätt att vi klarar att rätta ett bitfel (enkelfel). Antag att det dataord som skall skyddas är 001 0110.

Databitarna placeras i 11 positioner, där position 8, 4, 2 och 1 är reserverade för kontrollbitarna.

Kontrollbitarna beräknas genom modulo-2-addition av positionsnumren i binär form för de positioner som har "ettor". Modulo-2-additionen innebär att vi kompletterar under strecket så att vi får jämnt antal ettor i varje kolumn (vertikalt).

Pos: 11 10 9 8 7 6 5 4 3 2 1
 $D^7+D^6+D^5+X^4+D^4+D^3+D^2+X^3+D^1+X^2+X^1$
 Data: 0 0 1 0 1 1 0

Bild 1:

De sju databitarna placeras i rätt position.

9 = 1 0 0 1	11 = 1 0 1 1
6 = 0 1 1 0	10 = 1 0 1 0
5 = 0 1 0 1	9 = 1 0 0 1
= 1 0 1 0	8 = 1 0 0 0
$X^4 X^3 X^2 X^1$	7 = 0 1 1 1
	6 = 0 1 1 0
	5 = 0 1 0 1
	4 = 0 1 0 0
	3 = 0 0 1 1
	2 = 0 0 1 0
	1 = 0 0 0 1

Bild 2:

Kontrollbitarna beräknas som modulo-2-addition av de positioner som har "ettor".

Pos: 11 10 9 8 7 6 5 4 3 2 1
 $D^7+D^6+D^5+X^4+D^4+D^3+D^2+X^3+D^1+X^2+X^1$
 Data: 0 0 1 0 1 1 0
 Kontroll: 1 0 1 0
 Sänd sekvens: 001 1011 0010

Bild 3:

Den kompletta sekvensen om elva databitar.

Kontrollbitarna stoppas in på sina positioner.

I bild 4 kontrollräknar mottagaren och upptäcker att det inte finns något bitfel (Hammingkodningen skyddar bara mot ett enda bitfel).

I bild 5 syns motsvarande kontroll när det uppstått ett bitfel. Mottagaren klarar att identifiera vilken position som är felaktig.

Kodning i verkligheten

Blockkodning kan liknas vid ekvationssystem med många obekanta. Antag att vi skall överföra 40 databitar och adderar 10 databitar som bitfelsskydd. De första 40 databitarna kan vara originalbitarna, medan de resterande 10 kan ingå i ekvationssystem där man på olika sätt beräknar ett resultat utgående från alla 40 bitarna. Vi får något som skulle kunna liknas vid 40 obekanta, men fler än 40 ekvationssystem. Ur detta kan man lösa varje obekant på flera olika sätt och på detta vis sortera ut felaktiga lösningar.

Blockkoderna bygger på avancerad matematisk teori. Nackdelen är att avkodningen blir komplex vid stora block. Det blir stora ekvationssystem. Man kan därför inte utnyttja "mjuk" avkodning.

Ytterligare en nackdel är att visserligen är blockkoderna starka på att rätta bitfel upp till en viss gräns, men om bitfelen överstiger gränsen så blir resultatet totalt fel. Avkodningen spårar ur totalt.

Faltningskoder

Bakom faltningskoderna finns ingen matematisk teori som kan utnyttjas för att finna den optimala kodningen, utan man provar sig fram i datorsimuleringar.

Faltningskoderna liknar det som händer vid radiokommunikation. Man får reflexer i omgivningarna så att samma information kommer fram flera gånger, med olika tidsfördröjning. Som skuggbilder på tv.

För oss är skuggbilden störande, men om man låter ett bildbehandlingsprogram analysera bilden och leta efter mönster, upprepningar, så får man faktiskt en säkrare mottagning med reflexer. Samma information har kommit flera gånger.

Faltningsavkodningen kan enkelt kombineras med "mjuk" avkodning. Säg att en

Mottagen sekvens: 001 1011 0010

Pos: 11 10 9 8 7 6 5 4 3 2 1
0 0 1 1 0 1 1 0 0 1 0

9 = 1 0 0 1

8 = 1 0 0 0

6 = 0 1 1 0

5 = 0 1 0 1

2 = 0 0 1 0

= 0 0 0 0 Rätt mottaget

Bild 4:

Mottagaren kontrollerar den mottagna sekvensen när allt är rätt mottaget.

Mottagen sekvens: 001 101~~1~~ 0010

Pos: 11 10 9 8 7 6 5 4 3 2 1
0 0 1 1 0 1 ~~1~~ 0 0 1 0
0

9 = 1 0 0 1

8 = 1 0 0 0

6 = 0 1 1 0

2 = 0 0 1 0

= 0 1 0 1 Fel i position 5

Bild 5:

Mottagaren kontrollerar den mottagna sekvensen när det finns ett bitfel.

etta skall representeras av +1 V, medan en nolla är -1 V. Om detektorn indikerar +0,1 V så tolkas detta som en etta om man har "hård" avkodning, medan vid mjuk avkodning så är detta en mycket osäker etta. Den ligger nära gränsen till nolla. På detta sätt kan man sortera ut de databitar som man kan lita på, och vilka som är osäkra. Detta förbättrar avkodningen enormt och är en viktig del i GSM-mottagarens funktion.

11.3 V.24, hårdvaruhandskakning

Hittills har vi talat om kontrollsummor och felrättning eller bitfelsminimering. Nu skall vi titta på två olika länklager. Vi börjar med V.24-protokollet som används mellan en pc och ett modem.

Serieporten på pc, RS-232

Vid kommunikation mellan en pc (eller terminal) och ett modem så följer fysiska skiktet V.28-protokollet medan länkskiktet följer V.24-protokollet. I USA sammanfattas detta under RS-232. Vi skall inte fördjupa oss i V.28-protokollets spänningsnivåer, bit-synkronisering m.m. utan går till V.24-protokollet för att få exempel på uppgifter som hanteras av länkskiktet.

DTE (Data Terminal Equipment)

Pc:n eller terminalen kallas DTE.

DCE (Data Communication Equipment)

Modemet som hanterar kommunikationen mot yttrevärlden kallas DCE.

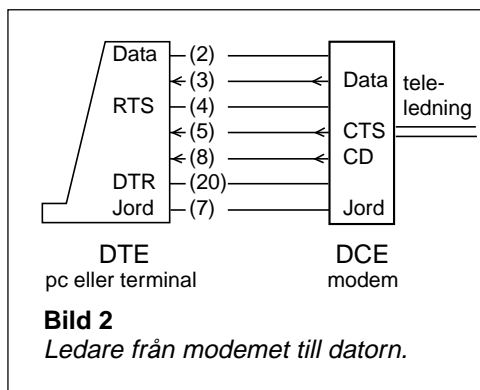
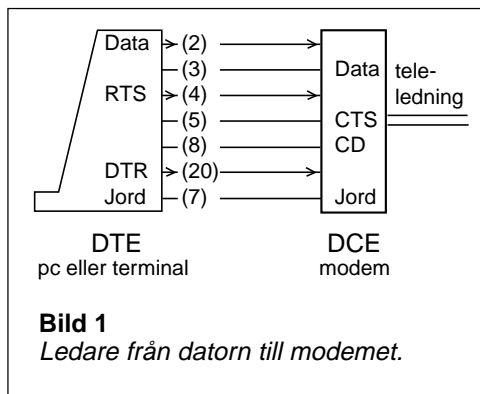
25-polig eller 9-polig kontakt

Pc:ns serieport har en 9-polig kontakt och eventuellt även den äldre 25-poliga kontakten. Siffror inom parantes anger stift i den 25-poliga kontakten.

Sex ledningar är intressanta

Kabeln mellan modemmet och pc:n innehåller minst sex ledare plus jordledning. Ledarna har följande beteckningar och utnyttjas till att transportera följande information:

- (2) Transport av data från DTE till DCE som skall vidare ut på teleledningen.
- (3) Transport av data som kommer från teleledningen till DCE och skall vidare till DTE.
- (20) DTR (Data Terminal Ready). Pc:n eller terminalen visar att den är beredd att ta emot data när DTR är hög.
- (4) RTS (Request To Send). Pc:n eller terminalen begär av modemmet att få sända när RTS är hög.



(5) CTS (Clear To Send). Modemet kontrollerar om "linjen är klar" (teleledningen och mottagaren där bortom) och ger pc:n eller terminalen klartecken att sända genom att lägga CTS hög.

(8) CD (Carrier Detect). Modemet upptäcker signal på teleledningen och beordrar pc:n eller terminalen att lyssna på linje (3) genom att lägga CD hög.

(7) Jord

Om två DTE skall kommunicera

Om man kopplar ihop två persondatorer via serieportarna så måste ledarna korsas. (2) skall till (3), (3) skall till (2) o.s.v. Modemkabeln har däremot inga korsade ledare. Välj rätt kabel!

Kommunikationsprogrammet startas

När terminalen slås på eller man startar kommunikationsprogrammet på pc:n, så talar terminalen om för modemmet att den kan ta emot data på (3) genom att lägga DTR hög.

Modemet bestämmer

Modemet är den som bestämmer. Varför? Jo, om det kommer data på teleledningen så finns ingen möjlighet att lagra data i modemmet tills terminalen är beredd att ta emot. Modemet saknar buffertminne.

När det kommer data på teleledningen (modulerad bärvåg, "carrier detect"), så beordrar modemmet terminalen/pc:n att lyssna på (3) genom att lägga CD hög.

CD aktiverar avbrottsfunktionen (IRQ) i pc:n, så att processorn avbryter det den håller på med och lyssnar efter data som kommer in via pinne (3).

När terminalen/pc:n vill sända så ligger dessa data i minnet. Terminalen kan vänta tills den får klartecken.

Terminalen/pc:n har ingen möjlighet att kontrollera att teleledningen är klar och att det finns en mottagare i andra ändan som är beredd att ta emot. Detta kan bara modemmet göra. Därför frågar terminalen/pc:n hur det ser ut på linjen genom att lägga RTS hög.

Modemet kontrollerar först att den mottagande enheten i andra ändan av teleledningen är beredd, och ger därefter klartecken genom att lägga CTS hög.

När terminalen/pc:n sänt sina data och är färdig så meddelar den modemmet att den är klar genom att lägga RTS låg, d.v.s. den drar tillbaka sin begäran att få sända. Modemet bekräftar genom att dra in tillståndet (lägger CTS låg).

Hårdvaruhandskakning

Modemet för befälet. Kommunikationen styrs med fyra ledningar, med hårdvara (DTR, RTS, CTS, CD), så att inte data går förlorad därför att enheterna talar i munnen på varandra eller inte är beredda. Detta är hårdvaruhandskakning, och det är länklagret som skapar ordning och reda i kommunikationen.

11.4 SDLC/HDLC, mjukvaruhandskakning

Mjukvaruhandskakning

Vad finns i andra ändan av teleledningen? Jo, ett modem. Hur hanteras kommunikationen? Det finns ju bara två trådar?

Kommandona som styr kommunikationen skickas som dataord inskjutna i dataflödet. Kommandona går som dataord, mjukvara. Detta kallas mjukvaruhandskakning.

SDLC (Synchronous Data Link Control)

I mitten av 1970-talet behövde IBM ett protokoll för att ansluta terminaler till sina stor datorer. Det blev SDLC (Synchronous Data Link Control), det första länklagerprotokollet för synkron bitorienterad kommunikation.

Vid SDLC är en dator "primary" medan de övriga är "secondary". All kommunikation hanteras av primary som ställer frågor. Ingen kommunikation kan initieras av secondary.

HDLC (High-level Data Link Control)

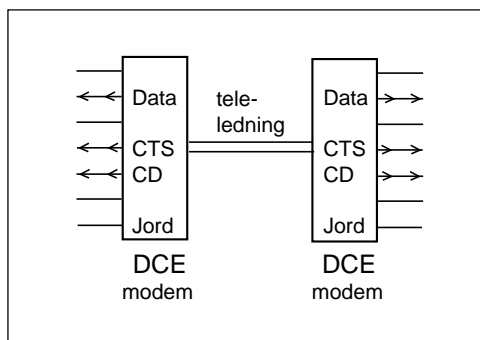
Men SDLC var ett IBM-specifikt protokoll. IBM lämnade SDLC till flera standardiseringsorgan för att försöka skapa en standard runt SDLC. Internationella standardiseringsorganisationen ISO utvidgade SDLC och skapade HDLC.

SDLC är i stort sett HDLC NRM (Normal Response Mode). Vidare införde man en mode där kommunikation kan initieras av secondary (HDLC ARM), samt en mode där primary kan växla mellan datorerna med full duplex (data skickas i båda riktningar samtidigt), HDLC ABM.

LAP, LAPB

Internationella Teleunionen ITU-T (tidigare CCITT) utgick från HDLC och skapade LAP (Link Access Procedure) från SDLC och LAPB (Link Access Procedure, Balanced) från HDLC ABM.

Delar av dessa protokoll finns i LAPD som är signaleringskanalen i ett ISDN-abonnemang, samt i SS7 som är telenätets signaleringsnät. När man skickar SMS till en GSM-telefon så överförs SMS-meddelandet i en SS7-ram ända ut till GSM-telefonen



(flaggor och bitstuffing avlägsnas i mobilväxeln).

Hur hanteras kommunikationen mellan dagens telefonmodem? Fysiska lagret följer V.34 eller V.90. Länklagret följer V.42 som är utvecklat ur LAPB.

IEEE 802.2

Organisationen IEEE modifierade HDLC och skapade IEEE 802.2 som även kallas LLC (Logical Link Control). Detta protokoll ingår i Ethernet, den vanligaste standarden för lokala nät, LAN.

Vi ser att de tankar om hur kommunikation mellan datorer skall hanteras, som utformades av IBM på 1970-talet, dessa tankar lever fortfarande. För att förstå vad som händer på länklagret så tittar vi närmare på SDLC.

En "primary"

SDLC bygger på att en nod utses till primary. Det naturliga för IBM på 1970-talet var att en stordator var primary.

En eller flera "secondary"

Till stordatorn ansluts en eller flera terminaler. En sådan terminal kan befinna sig långt bort och vara ansluten via en uppringd eller fast förbindelse. Dessa persondatorer eller terminaler kallas secondary.

Vi skall studera kommunikationen mellan primary och en secondary.

Synkron kommunikation

Går vi in och tittar på en synkron förbindelse så ser vi ettor och nollor rusa förbi (bild 1). Varken början eller slut.

Det behövs en "flagga"

Alla databitar som skall skickas placeras i "ramar" (frame). När börjar en ram? När slutar ramen?

Vi måste kunna identifiera början och slutet av en ram. Det gör vi genom att utnyttja en viss kombination av ettor och nollor som "flagga". Man har valt 0111 1110 (bild 2).

Bit-stuffing

Men kombinationen 0111 1110 förekommer bland de databitar som skall skickas. Inte bara som 8-bitsordet 0111 1110 utan även i kombinationer av två 8-bitsord, t.ex. 0101 0011 + 1111 0010. Varje grupp av sex ettor omgivna av nollor uppfattas som flagga.

Varför håller man sig inte till hela 8-bitsord? Därför att vid kommunikationen kan det bli "bitslip". Om sändarens och mottagarens klockor inte går exakt i takt kan mottagaren missa en bit, eller råka läsa dubbelt på en bit. Då förskjuts allt en databit. Men även vid bitslip skall mottagaren kunna identifiera när ramarna börjar och slutar.

För att undvika att kombinationen sex ettor omgivna av nollor förekommer i meddelandet så stoppar sändaren in en extra nolla efter fem ettor. Detta kallas bitstuffing.

Första steget för mottagaren är att identifiera flaggorna. Nästa steg är att avlägsna de extra bitstuffing-nollorna. Observera att sändaren stoppat in en extra nolla även om

```
... 000011011000100110111110
1000000111111001111110110
0000110011010001111000000
0000101000000111110010101
1011111000110000100010101
1011011000001001100110000
0011101111010011111001111
110011111011000001001101
11110100000...
```

Bild 1:

Mottagaren ser bara en massa ettor och nollor.

```
... 000011011000100110111110
10000001111111011111110110
0000110011010001111000000
0000101000000111110010101
1011111000110000100010101
1011011000001001100110000
0011101111010011111001111
11011111011000001001101
11110100000...
```

Bild 2:

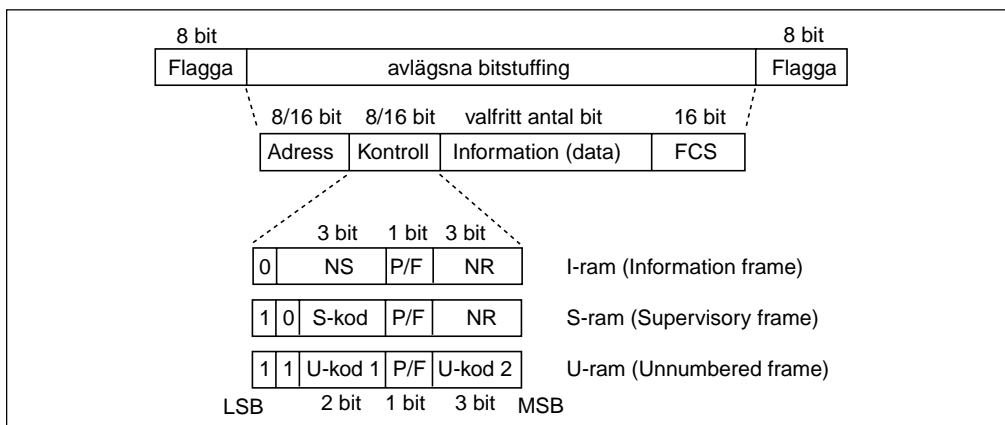
Mottagaren lerat efter flaggorna som utgör gräns för ramarna.

```
... 000011011000100110111110
100000 [ ] [ ] 110
0000110011010001111000000
0000101000000111110010101
1011111000110000100010101
1011011000001001100110000
00111011110100111110 [ ]
[ ] [ ] 11000001001101
11110100000...
```

Bild 3:

Mottagaren avlägsnar bitstuffing-nollorna.

det kommer en nolla efter fem ettor. Mottagaren kan ju inte veta om nollan efter fem ettor är en datanolla eller en bitstuffingnolla. Mottagaren stryker alltid en nolla efter fem ettor.



När mottagaren avlägsnat flaggorna och därefter strukit den nolla som stoppats in efter fem ettor, så återstår ramen som består av tre eller fyra fält.

Adress

Det finns bara ett adressfält, men adressfältet används på olika sätt beroende på vem som sänder.

- När primary sänder så innehåller adressfältet adressen till den secondary som skall ha ramen. Det finns bara en primary. Därför behövs ingen avsändare.

Om primary inte vet adressen till secondary så kan primary utnyttja "broadcast-adressen" FF (1111 1111) i en XID-ram. Då svarar secondary med sin egen adress i en XID-ram.

- När secondary sänder så måste paketet till primary. Och det finns bara en primary. Därför behövs inget adressfält som talar om vart ramen skall. Däremot behövs avsändaren. När secondary sänder så innehåller adressfältet avsändarens adress.

FCS (Frame Check Sequence)

Sista fältet är enkelt. Det består av 16 bit CRC som utnyttjas för att upptäcka om det finns bitfel. FCS beräknas på adress-, kontroll- och eventuellt datafält.

U-ram

Det finns tre olika typer av ramar. U-ramen är en ren kommandoram som används för

att koppla upp respektive koppla ner förbindelsen. Kontrolloktetten innehåller bara information om typ av U-ram, samt P/F-biten som används för att lämna över ordet till den andre.

U-ramen består bara av adress-, kontroll- och FCS-fält.

I-ram

Den information som skall utväxlas mellan datorerna skickas i I-ramar. Förutom adressfält, kontrollfält och FCS-fält så innehåller I-ramen ett informationsfält som kan bestå av valfritt antal databitar. Det finns inget krav på att informationen skall bestå av 8-bitsord, utan det går med vilket antal databitar som helst.

Kontrollfältet innehåller två räknare, den ena (NS) används av sändaren för att numrera de I-ramar som sänds.

S-ram

S-ramen är en kommandoram som används mitt i en kommunikation. Den består av adress-, kontroll- och FCS-fält. I kontrolloktetten finns en räknare (NR) som används för att meddela vilken I-ram man vill att sändaren skall sända nästa gång. Om man tar emot I-ram 5 och skriver NR = 6 så betyder detta att man har accepterat I-ram 5. Om man skriver NR = 5 så har man inte fått I-ram 5, vilket innebär att sändaren måste repetera.

I-ramen innehåller båda räknarna, d.v.s. man numrerar sina egna och bekräftar vad man tagit emot i samma kontrollfält.

U-ram (Unnumbered frame)						
1	1	U-kod 1	P/F	U-kod 2	HEX	
		1 1		1 0 1	BF	XID (Exchange information)
		0 0		1 0 0	93	SNRM (Set normal response mode) - primary
		1 1		0 1 1	73	UA (Unnumbered acknowledge) - secondary
		1 1		0 1 0	53	DISC/RD (Disconnect/Request disconnect)

XID

Vi skall titta på några U-ramar och börjar med XID (Exchange information). Den information som utväxlas är adressen.

Om primary inte vet adressen till secondary, då skickar primaryn broadcast-adressen FF (1111 1111) i en XID-ram. Secondary svarar med sin egen adress (i detta fall b2) i en XID-ram.

	Adr	Typ	NS	P/F	NR
Primary	FF	XID		1	
Secondary	b2	XID		1	

SNRM

SNRM (Set normal response mode) skickas av primary för att visa att primary startar förbindelsen. Samtidigt nollställer primary sina räknare (NS) = (NR) = 0.

	Adr	Typ	NS	P/F	NR
P	b2	SNRM		1	
S	b2	UA		1	

UA

UA (Unnumbered acknowledge) är svaret från secondary på en SNRM-ram. Även secondary nollställer sina räknare (NS) = (NR) = 0.

	Adr	Typ	NS	P/F	NR	
P	b2	I	0	0	0	data
	b2	I	1	1	0	data
S	b2	RR		1	2	
P	b2	RR		1	0	
S	b2	RR		1	2	
P	b2	RR		1	0	
S	b2	RR		1	2	
P	b2	RR		1	0	
S	b2	RR		1	2	
P	b2	RR		1	0	

P/F

P/F (Poll/Final) är en bit som används "för att lämna över ordet". Om P/F = 1 så är detta sista ramen i detta sändningspass. "Jag går över och lyssnar efter ramar från dig."

P/F = 0 innebär alltså att "jag är inte färdig. Det följer ytterligare en ram".

och läser. Ingen information finns att skicka på förbindelsen. Men då kan inte den ena datorn hålla på ordet. Tänk om den andra datorn får något att säga. Alltså måste datorerna lämna över ordet till varandra, ända tills någon yttrar sig. RR fram och tillbaka är att hålla liv i förbindelsen.

RR

RR (Receive Ready) betyder "jag lyssnar", d.v.s. samma sak som att lämna över ordet. Antag att du hämtat en Web-sida. Du sitter

S-ram (Supervisory frame)					
1	0	S-kod	P/F	NR	
		0 0			RR (Receive ready)
		1 0			RNR (Receive not ready)
		0 1			REJ (Reject, repetera från ram "NR")
		1 1			SREJ (Selective reject, repetera ram "NR")

NS-räknaren (Now Sending)

Informationsramarna numreras. I-ramarna från primary numreras från 0 till 7 (3 bit), varefter räknaren börjar om från 0.

Även I-ramarna från secondary numreras, från 0 till 7. Dessa båda NS-räknare löper helt oberoende av varandra. Var och en numrerar sina I-ramar.

NS visar alltså numret på I-ramen, den I-ram som sänds nu.

NR-räknaren (Next to receive)

Den andra räknaren visar vilken I-ram som man förväntar sig från den andre "nästa gång".

Bild 1 visar hur secondary förväntar sig I-ram 0, får I-ram 0 och 1, för att nästa gång förvänta sig I-ram 2.

När sätter secondary NR till 2? Jo, först när I-ram 0 och 1 är rätt mottagna, med korrekt FCS. Att fråga efter ram 2 innebär att godkänna alla ramar som föregår ram nr 2.

Max sju I-ramar i följd

Bild 2 visar hur secondary förväntar sig I-ram 2, hur primary sänder I-ramarna 2, 3, 4, 5, 6, 7 och 0, men inte fler? Varför?

Om primary även sänt nästa I-ram, en åttonde I-ram, så skulle den haft nummer 1. Vad svarar secondary? Med NR = 2. Men detta är vad secondary frågade efter innan de åtta ramarna sändes. Detta är samma sak som att inte ha fått några ramar alls.

Efter sju I-ramar så måste man lämna över för att NR-räknaren skall vara en säker bekräftelse på att allt är rätt mottaget.

Det finns en variant av SDLC-protokollet som har 16 bit i kontrollfältet. Då får man plats med 7-bitsräknare och NS och NR löper från 0 till 127.

Räknarna löper oberoende av varandra

Bild 3 visar hur secondary sänder sin första I-ram, med NS = 0, samt hur primary bekräftar med att nästa gång fråga efter I-ram nummer 1.

	Adr	Typ	NS	P/F	NR	
S	b2	RR	1	0	0	
P	b2	I	0	0	0	data
	b2	I	1	1	0	data
S	b2	RR	1	1	2	

Bild 1

	Adr	Typ	NS	P/F	NR	
S	b2	RR	1	2	2	
P	b2	I	2	0	0	data
	b2	I	3	0	0	data
	b2	I	4	0	0	data
	b2	I	5	0	0	data
	b2	I	6	0	0	data
	b2	I	7	0	0	data
	b2	I	0	1	0	data
S	b2	I	0	1	1	data
P	b2	RR	1	1	1	

Bild 2

	Adr	Typ	NS	P/F	NR	
S	b2	RR	1	2	2	
P	b2	I	2	0	0	data
	b2	I	3	0	0	data
	b2	I	4	0	0	data
	b2	I	5	0	0	data
	b2	I	6	0	0	data
	b2	I	7	0	0	data
	b2	I	0	1	0	data
S	b2	I	0	1	1	data
P	b2	RR	1	1	1	

Bild 3

Ordning och reda

Länklagret skapar ordning, ser till att datorerna inte pratar i mun på varandra.

Länklagret säkerställer säker kommunikation genom att inte fråga efter nya I-ramar förrän den pågående gruppen av I-ramar blivit rätt mottagna.

11.5 Kretskopplat nät

Nätverkslagret och transportlagret

Vi skall lämna länklagret och titta på kompletta nät.

Länklagret hanterar kommunikationen mellan två datorer, noder. Om man skall koppla upp en förbindelse i ett nät så måste man passera flera noder. Nätverkslagret håller reda på adressering och vägval i nätet och transportlagret garanterar säker kommunikation över flera noder. Men fortfarande finns ett länklager mellan de enskilda noderna.

Telenätet

När de första telefonerna installerades på 1880-talet så telefonerade man faktiskt över en enda koppartråd. Man utnyttjade jorden (jord) som den andra ledaren. Men detta gav störningar och kort räckvidd, max 50 km.

Ganska snart övergick man till tvåtrådsförbindelser. Ett telefonsamtal behövde ett eget trådpar, en egen elektrisk krets (circuit) mellan telefonerna. Telefonväxeln, först manuell men ganska snart automatisk (elektromekaniska växlar kom på 1920-talet), denna telefonväxel kopplade ihop trådparen mellan de abonnenter som ville ha en krets uppkopplad. Detta blev ett "kretskopplat" nät (circuit switched network).

FDM (Frequency Division Multiplex)

Varje telefonsamtal behövde ett eget kopparpar. Men på 1950-talet kom "radioteknik på kabel". Med SSB-modulering flyttades basbandet till någon högre frekvens. På så vis kunde flera simultana telefonsamtal utväxlas via samma koaxialkabel, se kapitel 5. Varje samtal fick sin egen del av frekvensområdet. Men varje samtal hade fortfarande en egen "krets", även om många "kretsar" delade på samma fysiska ledning på viss del av sträckan.

TDM (Time Division Multiplex)

På 1970-talet klarade man att PCM-koda telefonsamtalen. På 125 μ s fick man ett sampel, 8 bit. En kabel som överför 2048 kbit/s (2 Mbit/s) matar iväg 256 bit på 125 μ s. Detta räcker till 32 st 8 bitsord, 32 simultana tele-

fonsamtal. Varje samtal får sin del av tiden i en 125 μ s-ram. Varje samtal får sin egen tidlucka.

Detta är ett tidluckebaserat system, TDM-system. I verkligheten utnyttjas inte alla 32 tidluckorna till telefonsamtal. En tidlucka används till synkronisering m.m. Denna tidlucka innehåller ett speciellt 8 bitsord som alltid återkommer (egentligen varannan gång) som en flagga så att mottagarändan skall veta när 125 ms-ramen börjar. Mottagaren får ju bara en enda lång ström av ettor och nollor. Ur detta skall mottagaren plocka fram 8-bitsord, sampel, och rätt ord för rätt samtal. Mottagaren behöver alltså "bitsynk" för att avläsa databitarna och "ramsynk" för att gruppera bitarna i 8-bitsord och för att veta vilket 8-bitsord som hör till vilket samtal.

SDH (Synkron Digital Hierarki)

På en 155 Mbit/s-förbindelse (STM-1) får man plats med 63 st 2 Mbit/s-ramar under 125 μ s. Viss del av kapaciteten används för synkronisering, övervakning och styrning.

På en 622 Mbit/s-förbindelse (STM-4) får man in 4 st 155 Mbit/s-förbindelser.

På en 2,5 Gbit/s-förbindelse får man in 4 st 622 Mbit/s-förbindelser. Det är denna bithastighet som används på optokablar.

Även på optokabeln finns en indelning i 125 ms-ramar. I varje sådan ram finns 8-bitsord från $30 \cdot 63 \cdot 4 \cdot 4 = 30\,240$ telefonsamtal om hela kapaciteten utnyttjas för telefoni. I verkligheten blandar man. Några 2 Mbit/s-grupper utnyttjas för telefoni, andra för Frame Relay. Några 155 Mbit/s-förbindelser utnyttjas för ATM som i sin tur "bär" IP-trafik, eller också kör man IP-ramar direkt på STM-1.

11.6 Paketförmedlande nät

Visst skickas data som paket i de kretskopplade näten, paket som är 8 bit stora. Men paketen innehåller ingen adress. Det är tiden, när i tiden som tidluckan ligger, som identifierar paketet. Och om abonnenten inte har något att sända så är paketet tomt. Ingen annan kan utnyttja tidluckan. Abonnenten hyr tidluckan och får betala oavsett om han har något att sända.

Paket med adress

I de paketförmedlande näten skickas data när som helst i tiden. Varje paket innehåller en identifierare så att mottagarändan begriper vem som skall ha paketet.

Utnyttjar tomrum i kommunikationen

SDLC-förbindelsen skickar RR fram och tillbaka när änddatorerna inte har något att säga varandra. Men då finns möjlighet att skicka I-ramar med informationsfält som skall till någon annan abonnent.

Paketförmedlande nät är ett effektivt sätt att utnyttja förbindelsens kapacitet till verklig information. Nackdelen är att om abonnenterna i ett visst ögonblick vill sända fler paket än vad förbindelsen klarar att transportera, så kan det bli väntetid (upplevs som långsam förbindelse) eller man kan förlora paket, något som måste hanteras med begäran om omsändning från transportlagret (paket som inte får plats hanteras inte av länklagret).

Förbindelseorienterat paketnät

Vissa paketnät, t.ex. X.25-nät och Frame relay-nät, ställer upp en förbindelse mellan abonnenterna. Först skickas ett paket med fullständig adress till B-abbonenten. De växlar (routrar) som passeras noterar att i fortsättningen skall paketet med en viss förenklad adress skickas i viss riktning. Därefter startar trafiken.

Karakteristiskt för de förbindelseorienterade näten är att alla paket i en uppkoppling, t.ex. filöverföring, skickas samma väg genom nätet.

- X.25 är ett protokoll som utnyttjar länklagrets felkontroll i varje router. Om bitfel upptäcks så begärs omsändning från föregående router. Varje router måste alltså spara "I-ramarna" och vänta på bekräftelse på rätt mottaget. Detta gör kommunikationen långsam (max 64 kbit/s) men garanterat säker även på brusiga kopparpar med modem.
- På optofiber med låg bitfelshalt behövs inte individuell felkontroll på varje delsträcka. Frame Relay är X.25 utan felkontroll på länklagernivå. Här överlåter man felkontroll och omfrågning till transportlagret.

En variant av X.25 används i signaleringsnätet, det datanät som styr uppkopplingen av televäxlarna i det kretskopplade telenätet och mobiltelenäten. SMS transporteras på signaleringsnätet.

Tidigare kopplades kabeln, t.ex. optofibers ändutrustning, direkt till AXE-växeln. När man fick behov av dataförbindelser på dessa kablar så behövdes ett protokoll som AXE-växeln kunde hantera. Det blev Frame Relay som ju har stora likheter med signaleringsnätet, televäxlarnas eget datanät. Många sträckor i IP-nätet (Internet) går med Frame Relay som bärare.

Förbindelseöst paketnät

I förbindelseorienterade paketnät skickas först ett "uppkopplingspaket", varefter man använder en kort "etikett" som adress. Fördelen är att adressfältet bara består av ett fåtal oktetter (8-bitsord). I ett förbindelseöst paketnät finns den fullständiga adressen i varje paket. Nackdelen är att en mindre del av själva paketet kan utnyttjas för nyttodata.

I det förbindelselösa nätet sker vägval för varje enskilt paket.

Broadcastadress

För att en nod (router) skall kunna "tala" med en annan nod så måste den andra nodens adress vara känd. Därför har varje nod en routingtabell, adresslista, som visar vilken adress som skall väljas beroende på önskad slutdestination. I ett X.25-nät liksom vid Frame Relay så måste man skapa och mata in dessa routingtabeller i varje nod. På samma sätt är det i ett kretskopplat nät. Uppgifter om vart olika rikssamtal skall skickas måste matas in.

IP-protokollet innehåller en broadcast-adress som IP-routerna utnyttjar för att automatiskt bygga upp sina routingtabeller. Varje router frågar alla som han kan kontakta: "Vad är din adress och vilka andra routrar kan du nå?"

IP-protokollet

IP-protokollet är utvecklat av amerikanska försvaret som behövde ett datanät som klarar att stå emot fientliga anfall. Om någon router bombas så skall de kvarvarande routerna automatiskt bygga upp nya routingtabeller. Om hälften av IP-paketen i en filöverföring kommit fram och det blir avbrott så skall resten av paketen söka sig andra vägar.

Vid en jordbävning i Kalifornien gick det inte att ringa hem. Teleledningarna hade slitits av. Men det gick att skicka e-mail. IP-paketen letade sig via Japan, Asien och Europa till Kalifornien utan att någon manuell omkoppling behövde göras.

11.7 Cellförmedlande nät, ATM

IP-paketet innehåller ett huvud som består av 160 eller 192 bit (IPv4). Till detta kommer huvudet på det paket som bärs av IP. Det kan vara TCP -huvudet som består av 256 bit, eller UDP-huvudet som består av 64 bit.

Telefoni innebär att man har behov av att skicka 8 bit var 125 μ s. Att förse 8 bit med ett huvud på 160 + 64 bit är oekonomiskt. 97 % av förbindelsekapaciteten åtgår till att transportera huvuden, adressinformation.

Om man samlar ihop flera 8-bitsord för att skicka i klump så uppstår fördröjning. Man måste vänta tills paketet är fullt innan det kan sändas. Därför vill man att ett paketorienterat nät för telefoni skall ha så små paket som möjligt. Men då måste även huvudet vara litet.

Detta ledde fram till ATM, där alla paket är 53 oktetter, varav 5 oktetter till huvudet och 48 oktetter nyttolast. Vid telefoni rymmer nyttolasten 47 sampel, vilket ger en fördröjning på ca 6 ms, något som kan accepteras.

Alla ATM-paket är lika stora. Därför kallas dessa paket för "celler".

Förbindelseorienterat nät

ATM-huvudet innehåller inte fullständig adress utan bara en etikett (label).

När man skall ställa upp en ATM-förbindelse så skickas först ett antal uppkopplingsceller som meddelar respektive nod vart paket med en viss etikett skall skickas. Att ställa upp en väg genom nätet innan trafiken kan starta kallas ett förbindelseorienterat nät.

Andra förbindelseorienterade paketen är X.25 och Frame Relay. IP-nätet däremot är förbindelseöst. Varje paket har fullständig adress, och i varje nod görs vägval för varje paket.

En ATM-router ger liten fördröjning. Att avläsa etiketten går snabbt. En IP-router däremot måste göra en mer avancerad adressanalys, något som tar tid. Detta har lett fram till teknik där routern lär sig känna igen IP-huvuden och klarar att routa efterföljande paket till samma adress nästan lika snabbt som vid ATM. Vi ser hur fördelar och nackdelar mellan olika tekniska lösningar gradvis utjämnas.

12 — Ledningar

Vi börjar med några ord om radiostrålning, som behövs för att förklara signalens hastighet på ledningen och vad som ligger bakom ledningens karakteristiska impedans, samt ge en fysikalisk bild av vad en "laddning" är för något.

Därefter kommer vi till framåtgående effekt och anpassning, så som ledningar bör användas.

Tyvärr har vi ofta även reflekterad effekt på vår ledning. Vi inför begreppet reflektionskoefficient.

Därefter är det hög tid för "masken", en spänningspuls som kryper ut på ledningen och reflekteras. Hur stor del reflekteras? Kryper masken tillbaka på ovansidan eller på undersidan?

Vi avslutar med några filosofiska frågor kring vad som egentligen händer på en öppen ledning.

12.1 Radiostrålning

James Clerk Maxwell

Den skotske fysikern Maxwell föreslog 1864 inför vetenskapsakademien i London att Ampères lag borde kompletteras med en term som beror på förändringar i det elektriska fältet, det elektriska fältets derivata.

Maxwell såg hur Faradays lag beskriver vad som händer om vi ändrar magnetfältet (generatorprincipen) och tyckte att det även borde finnas ett samband som beskriver motsvarande för det elektriska fältet.

Radiostrålning

Om Maxwells antagande var riktigt så skulle detta leda till att en elektriskt laddad partikel som ändrar sin hastighet kommer att förlora energi, energi som strålar ut i form av ett elektriskt och magnetiskt fält, elektromagnetiska vågor.

Av matematiska beräkningar utgående från kompletteringen av Ampères lag fann Maxwell att denna elektromagnetiska energi borde utbreda sig med ljushastigheten, och föreslog att ljus skulle kunna vara sådan elektromagnetisk energi. Maxwell föreslog dessutom att det borde kunna finnas andra former av sådana elektromagnetiska vågor.

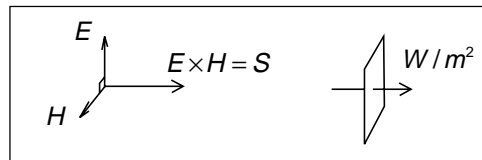
Heinrich Herz

År 1887, åtta år efter Maxwells död, lyckades Hertz utföra experiment som visade förekomsten av radiovågor, samt att radiovågorna bär sig åt på samma sätt som ljus. De kunde riktas med en parabolisk reflektor, reflekteras av en plåt på samma sätt som ljus i en spegel, brytas i ett dielektriskt prisma på samma sätt som ljus bryts i ett glasprisma.

Förutom att utföra experimenten så lyckades Hertz beskriva experimenten matematiskt utgående från de samband som vi kallar Maxwells ekvationer, trots att det enda som Maxwell gjorde, förutom att visa hur ekvationerna kan användas, var att komplettera Ampères lag med en term.

Poyntings vektor

När en elektriskt laddad partikel ändrar sin hastighet, eller rörelseriktning, så uppstår



elektriskt fält E V/m och magnetiskt fält H A/m som ligger vinkelrätt mot varandra.

Kryssprodukten $E \times H = S$ W/m² är effekttäthet, effekt per ytenhet, elektromagnetisk strålning, som utbreder sig med ljushastigheten i mediet.

Kryssprodukten får vi genom att vrida E mot H närmaste vägen. Då ligger S enligt skruvregeln.

Ljushastigheten

Ljushastigheten i mediet beror på permeabiliteten och kapacitiviteten i mediet.

$$\begin{aligned} c &= \frac{1}{\sqrt{\mu \cdot \varepsilon}} = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \cdot \frac{1}{\sqrt{\mu_r \varepsilon_r}} \approx \\ &\approx \frac{1}{\sqrt{(4\pi \cdot 10^{-7}) \left(\frac{10^{-9}}{36\pi}\right)}} \cdot \frac{1}{\sqrt{\mu_r \varepsilon_r}} = \\ &= \frac{3 \cdot 10^8}{\sqrt{\mu_r \varepsilon_r}} = \frac{c_0}{\sqrt{\mu_r \varepsilon_r}} \text{ m/s} \end{aligned}$$

Förhållandet E/H

Den elektromagnetiska vågen kan inte ha vilka värden som helst på E och H , utan det råder alltid ett bestämt förhållande, som även kan uttryckas med permeabiliteten och kapacitiviteten.

$$\begin{aligned} Z_{\text{vågimpedans}} &= \sqrt{\frac{\mu}{\varepsilon}} \approx \sqrt{\frac{4\pi \cdot 10^{-7}}{\frac{10^{-9}}{36\pi}}} \cdot \sqrt{\frac{\mu_r}{\varepsilon_r}} = \\ &= 120 \cdot \pi \cdot \sqrt{\frac{\mu_r}{\varepsilon_r}} \approx 377 \cdot \sqrt{\frac{\mu_r}{\varepsilon_r}} \Omega \end{aligned}$$

12.2 Vad händer när det "flyter" ström?

Ett batteri har överskott på elektroner vid minuspolen. Där finns fler elektroner än positiva joner. På motsvarande sätt råder underskott på elektroner vid pluspolen.

När batteriets poler ansluts till ledningens ledare

Ledarens ledare består av kopparjoner som sitter fast och fria rörliga elektroner. Elektronerna knuffar in sig så att de är jämnt fördelade längs ledarna. Elektronerna vill komma så långt från varandra som möjligt (lika laddning stöter bort).

Vid batteriets pluspol saknas elektroner. Då tar den yttersta elektronen på ledaren chansen att hoppa dit, så att han kommer längre från nästa elektron på ledaren. Vi får ett tomrum på ledaren, ett område som saknar en elektron. Då tar nästa elektron chansen att hoppa dit. Så hoppar nästa, och nästa o.s.v. På så vis förflyttas ett område längs ledaren, ett område som saknar en elektron. Och detta område som saknar en elektron har överskott på positiv laddning. Därför kallar vi detta "tomrum" för (en) positiv laddning.

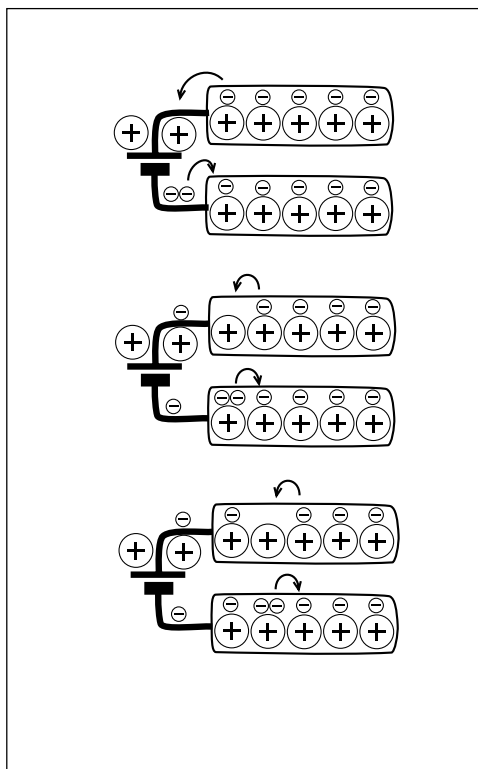
Hur fort rör sig laddningen?

Varje elektron har massa och kan inte röra sig särskilt fort. Det behövs oändlig energi för att få upp elektroner i ljushastigheten. Men själva laddningen, tomrummet, fortplantar sig med ljushastigheten (ljushastigheten i mediet om ledaren är omgiven av t.ex. plast. Denna hastighet är något lägre än ljushastigheten i vakuum).

Tomrummets hastighet beror inte på elektronernas individuella hastighet, utan på "timingen" mellan elektronerna. Elektronerna "tjuvstartar".

Negativ laddning

På samma sätt vid minuspolen. Där är så fullt med elektroner att en elektron tar chansen att hoppa över på ledaren trots att där redan finns en elektron. Men detta är bättre än att vara kvar i trängseln vid minuspolen. Då hoppar ledarens elektron ett steg till höger. Och nästa elektron hoppar, och nästa,



och nästa, i försök att komma bort från den elektron som kommer från vänster.

På ledaren som anslutits till minuspolen ser vi hur ett område med två elektroner, en elektronförtätning, rusar iväg åt höger. Varje elektron tar ett steg åt höger, i tur och ordning, på sådant sätt att elektronförtätningen tycks rusa iväg med ljushastigheten. Och elektronförtätningen, där vi har överskott på negativ laddning, kallar vi (en) negativ laddning.

(En) elektrisk laddning är inget man kan ta på, utan resultatet av elektronernas rörelser. Detta resultat av elektronernas rörelser är intressant vid studium av antenner och ledningar. Därför kommer vi fortsättningsvis inte att tala om elektroner, utan om laddningar som rör sig med ljushastigheten.

12.3 Signalens hastighet på ledningen

När man lägger spänning över ledningens ledare så kommer det alltså att flyta ström in i ledningen. På ledaren ansluten till batteriets pluspol så får vi positiv ström in i ledaren, och strömmen går från vänster till höger.

På den andra ledaren, ledaren ansluten till minuspolen, så får vi negativ ström och detta ritas vi med en pil med spetsen mot batteriet. Men åt vilket håll går denna negativa ström? Den går från vänster till höger på samma sätt som den positiva strömmen.

Att se ström som något som går in i ena ledaren och sedan kommer tillbaka i andra ledaren är en fungerande modell bara när man inte behöver ta hänsyn till att det tar tid för signalen att utbreda sig.

Radiostrålning mellan ledarna

När vi lägger spänning mellan ledarna så ger spänningen upphov till ett elektriskt fält i utrymmet mellan ledarna.

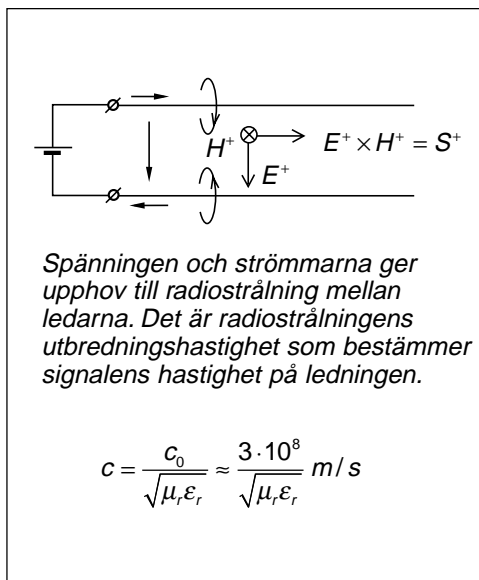
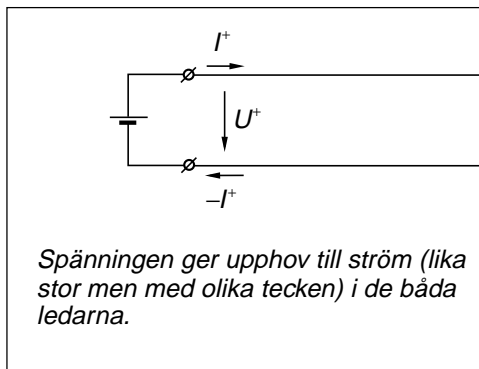
Men spänningen ger även upphov till ström i ledarna, och strömmen ger magnetfält riktat enligt skruvregeln runt ledarna (skruvregeln tillämpad på "pilarna").

Mellan ledarna kommer det elektriska fältet att vara riktat uppifrån och ner, medan magnetfälten är riktade in i papperet. Dessa båda visare representerar effekttäthet som utbreder sig i utrymmet mellan ledarna, och denna effekttäthet, signalen, utbreder sig med ljushastigheten i det medium som finns mellan ledarna, t.ex. någon plastisolering.

Man kan alltså säga att det är radiostrålningen mellan ledarna som knuffar elektronerna så att det ser ut som om positiv och negativ laddning rusar med ljushastigheten i ledarna.

Vad är signalen? Laddningar eller radiostrålning?

Är det radiostrålningen mellan ledarna som är själva signalen, medan laddningarnas (elektronernas) rörelser är en följd av radiostrålningen? I vissa sammanhang är detta faktiskt det lämpligaste synsättet.



Vi kommer i fortsättningen att anse att signalen är laddningarna som rör sig i ledarna, med ljushastigheten. Radiostrålningen mellan ledarna ser vi som en följd av laddningarnas rörelser. Orsaken till att vi väljer detta synsätt är att det är enklare att räkna på spänning och ström som håller sig till ledare, än att räkna på elektriska och magnetiska fält som breder ut sig i hela rymden.

12.4 Ledningens karakteristiska impedans Z_0

När vi ansluter ledningen till batteriet så kommer batterispänningen att utbreda sig "framåt" på ledningen med ljushastigheten. För att visa att detta är spänning som går från källan, generatoren eller batteriet, mot lasten som vi alltid förlägger till höger i bilden, så kallar vi spänningen U^+ , spänning som rör sig i positiv riktning.

Denna framåtgående spänning ger upphov till framåtgående ström. Hur stor blir strömmen?

- "Det vet vi inte eftersom vi inte kan se vad som sitter i ändan på kabeln".

Helt fel! Spänningen vet heller inte vad som finns i ändan på kabeln. Spänningen har ännu inte varit där och tittat efter.

Det är något helt annat som bestämmer strömmens storlek. Om jag säger att signalen går som radiostrålning. Vågimpedansen i mediet mellan ledarna är förhållandet mellan elektriska och magnetiska fältstyrkorna. Detta leder till att

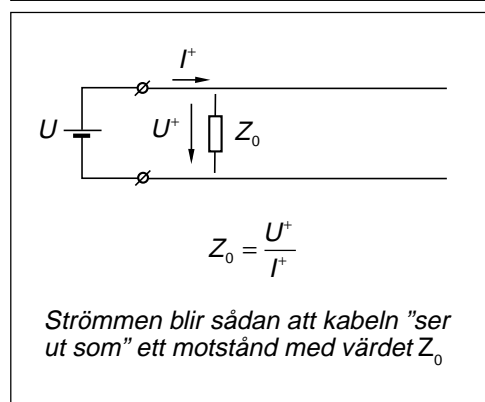
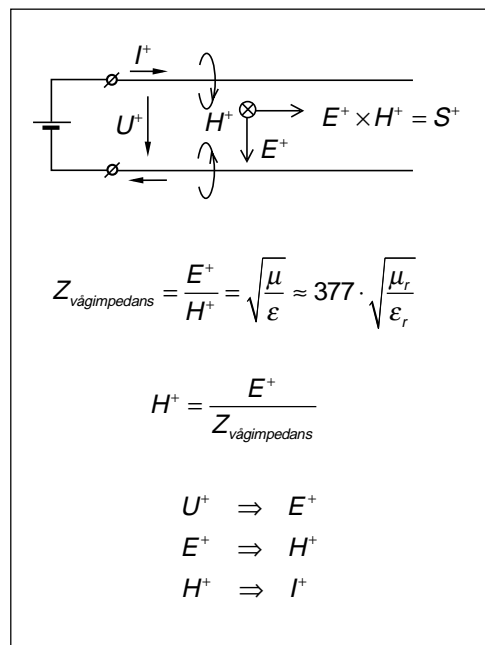
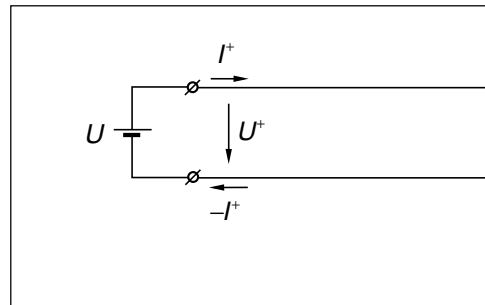
- Spänningen bestämmer den elektriska fältstyrkan.
- Elektriska fältstyrkan tillsammans med vågimpedansen i mediet mellan ledarna bestämmer storleken på den magnetiska fältstyrkan.
- Magnetiska fältstyrkan bestämmer storleken på strömmen.

Varje ledning har sin karakteristiska impedans Z_0

Detta i sin tur leder till att ledningens fysiska dimensioner tillsammans med kapacitetstalet bestämmer hur hög ström man får, om man lägger på en viss spänning.

Varje ledning ser ut som ett motstånd Z_0 . Den ström som flyter in i ledningen i det ögonblick vi ansluter batteriet bestäms av detta motstånd Z_0 , som kallas ledningens karakteristiska impedans.

Z_0 är rent resistiv ända ner till likström om ledningen är förlustfri. Om ledningen har förluster kommer Z_0 att förändras vid låg frekvens (< 10 kHz). I denna kurs studerar vi endast förlustfria ledningar.

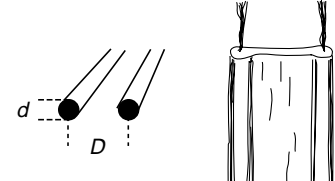


Parallellledningen

Karakteristiska impedansen Z_0 för två parallella ledare bestäms av avståndet mellan ledarna och ledarnas tjocklek.

Långt avstånd mellan ledarna ger hög karakteristisk impedans. Detta ger låg ström för viss signaleffekt. Låg ström innebär låga förluster, eftersom förlusterna är uppvärmningen av ledarna, ledarnas resistans multiplicerat med strömmen i kvadrat. Det var så här gamla telefonledningar såg ut, två "blanktrådar" upphängda i porslinsisolatorer på telefonstolpar. Det finns fortfarande telefonabonnenter i fjälltrakterna som har 10 mil till närmaste telefonstation. På dessa avstånd är blanktrådsledningar (4 mm kopparbronstråd) enda möjligheten om man vill undvika förstärkare längs ledningen.

När ledarna läggs i en kabel hamnar ledarna (0,5 mm koppar) närmare varandra och Z_0 minskar, dämpningen ökar. Längsta möjliga avstånd minskar till ca 5 km.

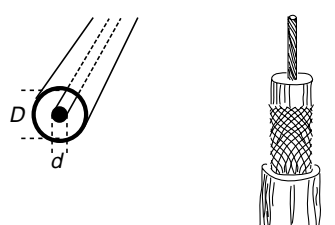


$$Z_0 = \frac{1}{\pi} \sqrt{\frac{\mu}{\epsilon}} \cdot \ln \frac{2D}{d} \approx \frac{276}{\sqrt{\epsilon_r}} \cdot \log \frac{2D}{d}$$

Koaxialledningen

Koaxialledningen är den enda ledningstyp som kan göras helt radiostrålningsfri, även vid hög frekvens. Därför används i dag alltid koaxialledning vid radiofrekvens, bortsett från vissa tillämpningar över 1 GHz där man använder vågledare.

Tunn mittledare ger högt Z_0 , som ger lägre ström. Men samtidigt ökar resistansen i mittledaren (den blir tunnare). Därför finns ett optimum.

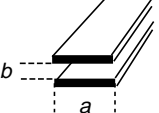


$$Z_0 = \frac{1}{2\pi} \sqrt{\frac{\mu}{\epsilon}} \cdot \ln \frac{D}{d} \approx \frac{138}{\sqrt{\epsilon_r}} \cdot \log \frac{D}{d}$$

$$Z_0 \text{ för lägsta dämpning: } Z_0 = \frac{77}{\sqrt{\epsilon_r}}$$

Solid polyetylenplast som dielektrikum har $\epsilon_r = 2,25$. För sådan kabel får man lägsta dämpning om $Z_0 = 51 \Omega$. Standard för radiotillämpningar har blivit 50Ω .

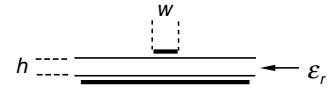
För skumplast (mest luft) kommer vi att få $\epsilon_r \approx 1$. Då blir lägsta dämpning vid 77Ω . Standard för kabel-tv har blivit 75Ω .



$$Z_0 = \frac{b}{a} \sqrt{\frac{\mu}{\epsilon}} \approx \frac{377}{\sqrt{\epsilon_r}} \cdot \frac{b}{a}$$

Plattledningen och "stripline"

Tryckt ledningsdragning är ett specialfall av plattledningen. Vid stripline består de två ledarna av den tryckta ledaren och ett "jordplan" på baksidan eller mitt i laminatet. Vid radiotillämpningar anpassas ledarens bredd så att man får önskat Z_0 .

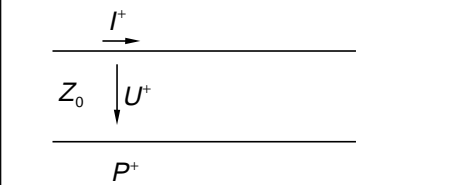


$$Z_0 \approx \frac{h}{w} \sqrt{\frac{\mu}{\epsilon}} \approx \frac{377}{\sqrt{\epsilon_r}} \cdot \frac{h}{w}$$

12.5 Framåtgående effekt och anpassning

Framåtgående effekt

Effekten, d.v.s. spänningen och strömmen, rusar in i ledningen med ljushastigheten (beror på isoleringens ϵ_r mellan och omkring ledarna). Hela vägen är förhållandet spänning/ström hos den effekt som rusar fram längs ledningen lika med karakteristiska impedansen Z_0 .



$$\begin{cases} Z_0 = \frac{U^+}{I^+} \\ P^+ = U^+ \cdot I^+ \end{cases}$$

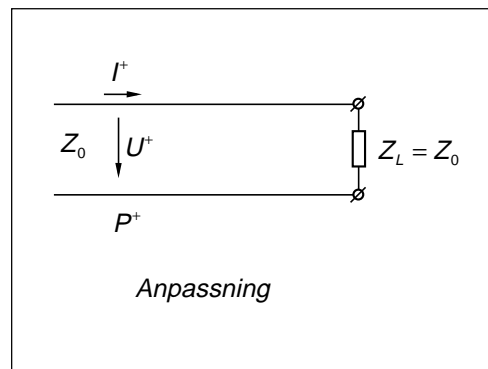
$$P^+ = U^+ \cdot I^+ = \frac{(U^+)^2}{Z_0} = Z_0 \cdot (I^+)^2$$

Anpassning

När ledningen avslutas med ett motstånd Z_L som har samma värde som ledningens karakteristiska impedans Z_0 så fortsätter signalen in i Z_L utan att märka någon skillnad.

- När I^+ går genom Z_L bildas en spänning som är lika med U^+ .
- När U^+ läggs över Z_L bildas ström som är lika med I^+ .

Randvillkoren är uppfyllda. Effektutvecklingen i Z_L kommer att bli lika med den framåtgående effekten. All effekt absorberas. Det blir ingen effekt över som kan reflekteras. Man säger att ledningen är anpassad.



Så här vill vi använda ledningar. All effekt som går framåt skall fortsätta in i belastningen, som kan vara en antenn eller radiomottagare eller kabel-tv-förstärkare. Även datakablar, t.ex. SCSI-flatkabel eller Ethernet (10BASE2) måste avslutas med $Z_L = Z_0$.

12.6 Reflekerad effekt

Om ledningen är öppen

Här kommer den framåtgående effekten till ledningens ände och där finns inget. Ledningen är avklippt. Här kan inte någon effekt förbrukas. Enda möjligheten är att effekten vänder och rusar tillbaka. Vi får reflekerad effekt, med reflekerad spänning och reflekerad ström. Vilka värden får spänningen och strömmen?

Om $Z_L = \infty$ så måste totala strömmen vara noll vid ledarens ände. Detta leder till att den reflekerade strömmen I^- blir lika stor som I^+ , men med motsatt riktning.

Den reflekerade signalen

Vi har räknat på den framåtgående signalen. Positiv riktning för den framåtgående signalen är från vänster till höger.

Nu skall vi räkna på den reflekerade signalen. Positiv riktning för den reflekerade signalen är från höger till vänster.

Strömmen I^- går i positiv riktning för den reflekerade signalen och ger alltså upphov till en positiv reflekerad spänning U^- .

Även för den reflekerade signalen är förhållandet mellan spänning/ström lika med Z_0 .

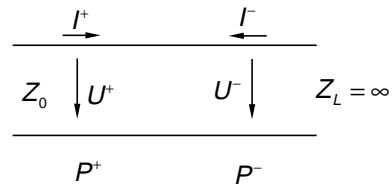
Räkna på varje signal för sig

Börja med att räkna på den framåtgående signalen. Det är enkelt.

Lika enkelt är det att räkna på enbart den reflekerade signalen.

Totala spänningen och strömmen

Först därefter lägger vi ihop signalerna för att beräkna den totala signalen. Att försöka klara två signaler i samma formel, signaler som rör sig åt olika håll, är snudd på omöjligt. Räkna på varje riktning för sig.

**Den framåtgående signalen:**

(positiv riktning från vänster till höger)

$$\begin{cases} Z_0 = \frac{U^+}{I^+} \\ P^+ = U^+ \cdot I^+ \end{cases}$$

Den reflekerade signalen:

(positiv riktning från höger till vänster)

$$\begin{cases} Z_0 = \frac{U^-}{I^-} \\ P^- = U^- \cdot I^- \end{cases}$$

Totala signalen:

(positiv riktning från vänster till höger)

Det du mäter med en voltmeter:

$$U_{\text{tot}} = U^+ + U^-$$

Det du mäter med en amperemeter:

$$I_{\text{tot}} = I^+ - I^-$$

Effekten som hamnar i lasten:

$$P_{\text{last}} = P^+ - P^-$$

OBS!

Observera särskilt att förhållandet mellan totala spänningen och totala strömmen kan bli vad som helst:

$$\frac{U^+ + U^-}{I^+ - I^-} \neq Z_0$$

12.7 Reflexionskoefficient

Även om det är enkelt att beräkna de framåtgående och reflekterade signalerna var för sig, så skall vi ändå härleda det uttryck som kopplar ihop den framåtgående och reflekterade signalen.

Reflexionskoefficienten

Vi definierar reflexionskoefficienten som förhållandet mellan reflekterade och framåtgående spänningarna:

$$\rho = \frac{U^-}{U^+}$$

Spänning och ström vid lastimpedansen

Vid Z_L måste följande samband gälla mellan totala spänningen över Z_L och totala strömmen genom Z_L :

$$Z_L = \frac{U^+ + U^-}{I^+ - I^-}$$

Efter diverse räkningar kommer vi fram till följande uttryck:

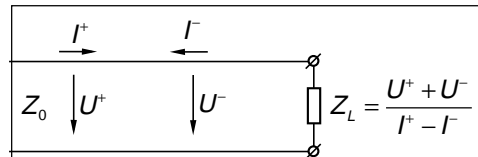
$$\rho = \frac{Z_L - Z_0}{Z_L + Z_0}$$

som vi skall använda för att beräkna reflekterade spänningen när vi känner den framåtgående spänningen:

$$U^- = \rho \cdot U^+$$

Reflexionskoefficienten är komplex

Z_L kan vara vilken impedans som helst. Detta innebär att reflexionskoefficienten får både amplitud och fasvinkel.



$$I^+ = \frac{U^+}{Z_0} \quad I^- = \frac{U^-}{Z_0}$$

Vi definierar reflexionskoefficienten:

$$\rho = \frac{U^-}{U^+}$$

utgår från sambandet som skall gälla över lastimpedansen

$$\frac{U^+ + U^-}{I^+ - I^-} = Z_L$$

Utnyttjar de vanliga sambanden mellan spänning och ström, och får:

$$U^+ + U^- = Z_L(I^+ - I^-) = \frac{Z_L}{Z_0}(U^+ - U^-)$$

$$Z_0 + \rho \cdot Z_0 = Z_L - \rho \cdot Z_L$$

$$\rho \cdot (Z_L + Z_0) = Z_L - Z_0$$

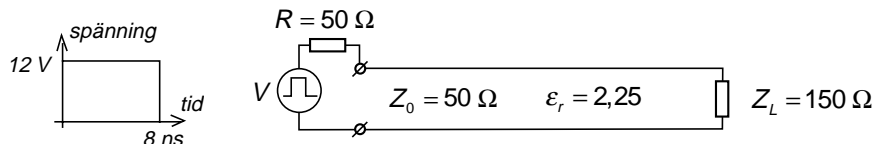
vilket ger:

$$\rho = \frac{U^-}{U^+} = \frac{Z_L - Z_0}{Z_L + Z_0}$$

12.8 "Masken"

Exempel 12.1

En pulsgenerator skickar ut en 8 ns lång puls på en 3 m lång ledning med $Z_0 = 50 \Omega$ och $\epsilon_r = 2,25$. Rita pulsens utseende och läge på ledningen efter 10 ns, 20 ns och 35 ns.



När pulsen stiger till 12 V så hamnar spänningen över motståndet och ledningen. Ledningen uppför sig som ett motstånd. Vi får spänningsdelning.

Generatoren är ansluten till ledningen via ett motstånd. Vi får spänningsdelning och beräknar nivån på den framåtgående spänningen.

$$U^+ = \frac{V \cdot Z_0}{R + Z_0} = \frac{12 \cdot 50}{50 + 50} = 6 \text{ V}$$

Pulsen kryper ut på ledningen med ljushastigheten på ledningen. Vi kan beräkna hur lång pulsen blir.

Pulsen är 8 ns lång. Pulsens utsträckning längs ledningen blir (maskens längd):

$$\Delta x = c \cdot \Delta t = \frac{c_0}{\sqrt{\epsilon_r}} \cdot \Delta t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 8 \cdot 10^{-9} = 1,6 \text{ m}$$

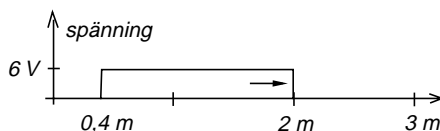
Det är denna puls som är masken, en mask som är 1,6 m lång.

Hur långt har maskens huvud hunnit efter 10 ns? Svar 2 m.

Hur hög är masken? Svar 6 V.

Efter 10 ns har pulsens framkant hunnit (har masken hunnit krypa):

$$x(10\text{ns}) = \frac{c_0}{\sqrt{\epsilon_r}} \cdot t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 10 \cdot 10^{-9} = 2 \text{ m}$$



När pulsen når ledningens ända reflekteras pulsen. Hur stor del av pulsen reflekteras (vad är reflexionskoefficienten vid lasten)?

$$\rho_L = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{150 - 50}{150 + 50} = 0,5$$

$$U^- = \rho \cdot U^+ = 0,5 \cdot 6 = 3 \text{ V}$$

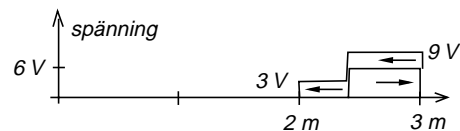
När maskens huvud når lastresistansen kommer bara en del av masken att reflekteras. Vi räknar ut att hälften, 3 V, reflekteras.

OBS att när masken vänder så kryper den tillbaka ovanpå sig själv. Spänningen blir summaspänningen, summan av framåtgående och reflekterade spänningarna.

Efter 20 ns har maskens huvud hunnit 4 m. Det är 3 m fram och 1 m tillbaka.

Efter 20 ns har pulsens framkant hunnit (har masken hunnit krypa):

$$x(20\text{ns}) = \frac{c_0}{\sqrt{\epsilon_r}} \cdot t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 20 \cdot 10^{-9} = 4 \text{ m}$$



När maskens huvud kommer tillbaka till generatorändan så träffar den resistansen vid generatoren. Vi beräknar reflexionskoefficienten i generatorändan och ser att ingen

spänning kommer att reflekteras. Vi har anpassning.

Reflexionskoefficient när pulsen når ledningens ända (generatorändan):

$$\rho_G = \frac{R - Z_0}{R + Z_0} = \frac{50 - 50}{50 + 50} = 0$$

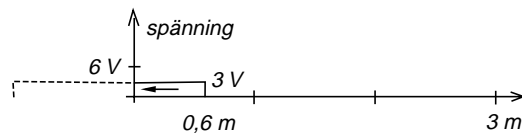
$$U^- = \rho \cdot U^+ = 0 \cdot 3 = 0 \text{ V}$$

Efter 35 ns har maskens huvud krupit 7 m, vilket är 3 m fram, 3 m tillbaka och sedan in i generatorändan eftersom inget reflekteras.

Vi ser maskens svans försvinna in i generatorändan.

Efter 35 ns har pulsens framkant hunnit (har masken hunnit krypa):

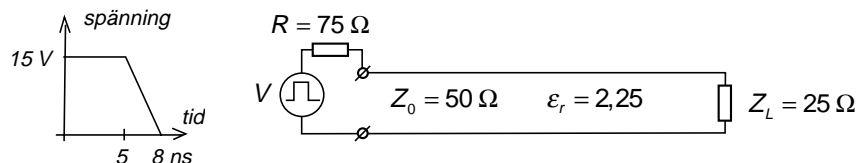
$$x(35\text{ns}) = \frac{c_0}{\sqrt{\epsilon_r}} \cdot t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 35 \cdot 10^{-9} = 7 \text{ m}$$



12 — Ledningar

Exempel 12.2

En pulsgenerator skickar ut en 8 ns lång puls på en 3 m lång ledning med $Z_0 = 50 \Omega$ och $\epsilon_r = 2,25$. Rita pulsens utseende och läge på ledningen efter 10 ns, 20 ns och 35 ns.



När pulsen stiger till 15 V så hamnar spänningen över motståndet och ledningen. Ledningen uppför sig som ett motstånd. Vi får spänningsdelning.

Generatoren är ansluten till ledningen via ett motstånd. Vi får spänningsdelning och beräknar nivån på den framåtgående spänningen.

$$U^+ = \frac{V \cdot Z_0}{R + Z_0} = \frac{15 \cdot 50}{75 + 50} = 6 \text{ V}$$

Pulsen kryper ut på ledningen med ljushastigheten på ledningen. Vi kan beräkna hur lång pulsen blir.

Pulsen är 8 ns lång. Pulsens utsträckning längs ledningen blir (maskens längd):

$$\Delta x = c \cdot \Delta t = \frac{c_0}{\sqrt{\epsilon_r}} \cdot \Delta t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 8 \cdot 10^{-9} = 1,6 \text{ m}$$

Det är denna puls som är masken, en mask som är 1,6 m lång.

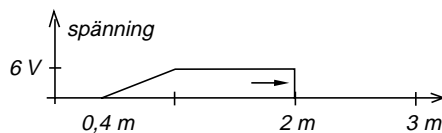
Hur långt har maskens huvud hunnit efter 10 ns? Svar 2 m.

Hur hög är masken? Svar 6 V.

Observera att den ändan av pulsen som kryper först är den del av pulsen som varit ute längst, d.v.s. ändan vid $t = 0$.

Efter 10 ns har pulsens framkant hunnit (har masken hunnit krypa):

$$x(10\text{ns}) = \frac{c_0}{\sqrt{\epsilon_r}} \cdot t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 10 \cdot 10^{-9} = 2 \text{ m}$$



När pulsen når ledningens ända reflekteras pulsen. Hur stor del av pulsen reflekteras (vad är reflexionskoefficienten vid lasten)?

$$\rho_L = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{25 - 50}{25 + 50} = -\frac{1}{3}$$

$$U^- = \rho \cdot U^+ = -\frac{1}{3} \cdot 6 = -2 \text{ V}$$

När maskens huvud når lastresistansen kommer bara en del av masken att reflekteras. Vi räknar ut att -2 V reflekteras.

OBS att när masken vänder med negativ reflexionskoefficient så kryper den tillbaka "på andra sidan av grenen". På samma

sätt som tidigare blir summaspänningen summan av framåtgående och reflekterade spänningarna.

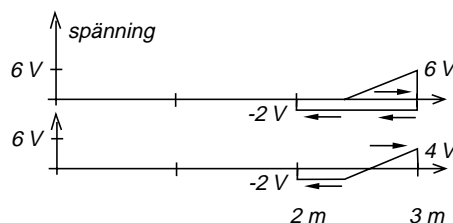
Efter 20 ns har maskens huvud hunnit 4 m . Det är 3 m fram och 1 m tillbaka.

Efter 20 ns har pulsens framkant hunnit (har masken hunnit krypa):

$$x(20\text{ns}) = \frac{c_0}{\sqrt{\epsilon_r}} \cdot t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 20 \cdot 10^{-9} = 4 \text{ m}$$

Det ser ut så här:

... eller så här om man summerar:



När maskens huvud kommer tillbaka till generatorändan så träffar den resistansen vid generatoren. Vi beräknar reflexionskoeff-

cienten i generatorändan. Den spänning som vänder adderas till framåtgående spänningen från generatoren (noll i detta fall).

Reflexionskoefficient när pulsen når ledningens ända (generatorändan):

$$\rho_G = \frac{R - Z_0}{R + Z_0} = \frac{75 - 50}{75 + 50} = 0,2$$

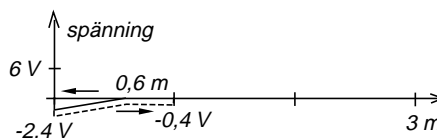
$$U^- = \rho \cdot U^+ = 0,2 \cdot (-2) = -0,4 \text{ V}$$

Efter 35 ns har maskens huvud krupit 7 m , vilket är 3 m fram, 3 m tillbaka och därefter 1 m ut på ledningen igen.

Positiv reflexionskoefficient betyder att masken inte byter sida.

Efter 20 ns har pulsens framkant hunnit (har masken hunnit krypa):

$$x(35\text{ns}) = \frac{c_0}{\sqrt{\epsilon_r}} \cdot t = \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot 35 \cdot 10^{-9} = 7 \text{ m}$$



12.9 Går det ström?

I de båda exemplen med "masken" så studerade vi en puls. Orsaken var helt enkelt den att om vi bara lägger på spänning, en mycket lång puls, och har reflexion i båda ändarna, så får vi multipelreflexer. Framåtgående spänningen blir en oändlig serie av framåtgående spänningar, och på samma sätt med de reflekterade spänningarna. För att slippa dessa serier så studerade vi en kort puls.

Ett annat sätt att slippa oändliga serier är att anpassa ledningen i ena ändan. Vi skall se på en ledning som är öppen i ena ändan men anpassad i generatorändan.

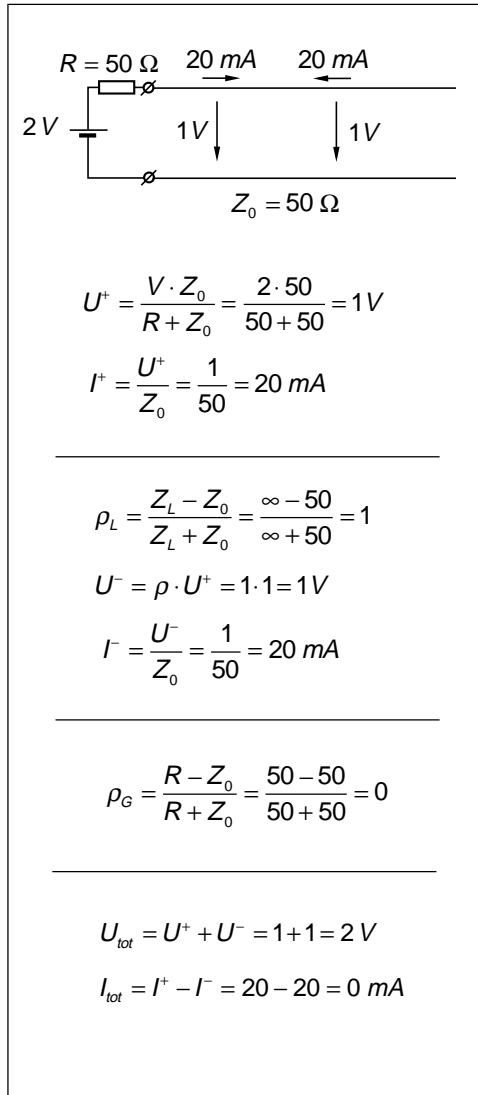
När vi ansluter batteriet så får vi spänningsdelning. Vi får 1 V framåtgående spänning, och detta ger 20 mA framåtgående ström.

I laständan får vi totalreflexion. Reflekterade spänningen blir 1 V och reflekterade strömmen blir 20 mA.

När den reflekterade spänningen kommer tillbaka till generatorändan så har summaspänningen på ledningen blivit 2 V, summaströmmen är 0 mA. Det går ingen ström genom motståndet. Vi får inget spänningsfall. Men detta tillstånd uppnås först efter den tid det tar för signalen att gå fram och tillbaka på ledningen.

Vi har ström på ledningen under detta första tidsögonblick. Frågan är, har vi även ström när spänningen kommit tillbaka? Har vi laddningar som går i båda riktningarna, vilket gör att elektronerna inte rör sig, så att vi inte får resistivt spänningsfall. Ingen värmeutveckling?

I första tidsögonblicket har vi framåtgående effekt på ledningen. Den framåtgående effekten reflekteras och kommer tillbaka. Sluttar batteriet leverera framåtgående effekt, eller fortsätter batteriet leverera framåtgående effekt men får tillbaka lika mycket, så att detta kan pågå i oändligheten? Ja, vem vet? Vi har inget instrument som kan mäta detta.



13 — Smithdiagrammet, anpassning

Vid växelspänning kommer den framåtgående spänningen och den reflekterade spänningen att få olika fasvinkel beroende på var på ledningen vi tittar. Hur lång tid har växelspänningen varit på väg? Och spänningarna rör sig i motsatt riktning.

Vi får punkter med maxspänning och punkter med minspänning, stående vågor, och definierar våglängd och ståendevågförhållande, samt ser på olika sätt att mäta ståendevågförhållande.

Från reflexionskoefficienten tar vi oss fram till Smithdiagrammet, som vi använder för att beräkna inimpedansen när lasten är ansluten med en ledning av viss längd.

Vi ser även hur en slitsad ledning kan användas för bestämning av impedans. En impedansmeter för hög frekvens.

Reflekterad effekt innebär att signalen går flera gånger fram och tillbaka, vilket kan ge problem i många tillämpningar.

Med kvartsvågstransformator anpassar vi en last med fel impedans, en vanlig metod i antennsystem med flera matade antenn-element.

Anpassning kan även ske genom att parallellkoppla ledningen med en ledningsstump (stubbe) eller kondensator, en vanlig metod i basstationssändare för mobiltelefoni.

13 — Smithdiagrammet, anpassning

13.1 Våglängd

Vi tar en ledning som är anpassad i båda ändar. På ledningen finns enbart framåtgående signal.

Ledningen matas av en växelspanningsgenerator. Framåtgående signalen beskrivs vi med sinusfunktionen (vi har använt cosinus vid så många andra tillfällen).

Ju längre bort från generatoren vi kommer, ju närmare lasten vi kommer, desto större blir tidsfördröjningen Δt .

Vi kan utveckla sinusuttryckets argument och byta ut frekvensen f mot periodtiden T . Då ser man att om tidsfördröjningen Δt är lika med en periodtid T så blir fasvinkel 2π , d.v.s. i alla punkter där tidsförskjutningen är hela multipler av periodtiden T , där har signalen samma fasläge.

Hur långt hinner signalen under en period?

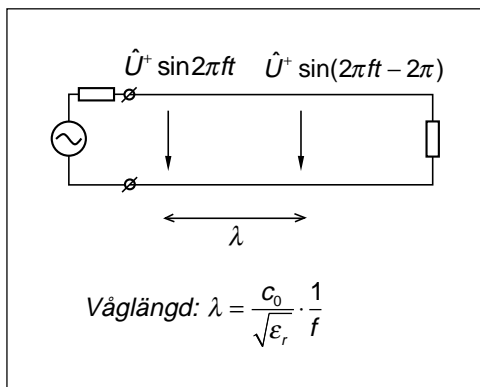
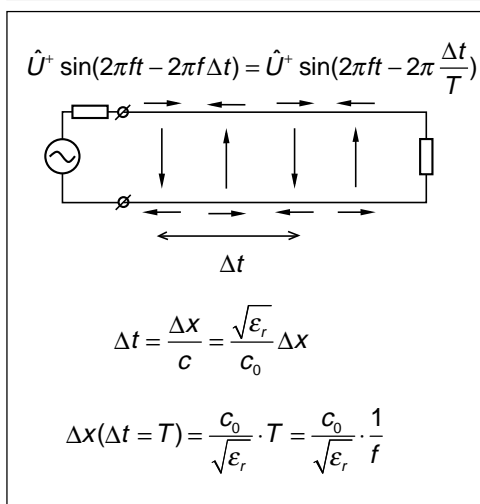
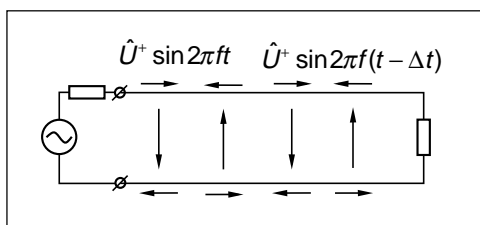
En sträcka som är speciellt viktig är avståndet mellan punkter där signalen har samma fasläge. Detta är samtidigt den sträcka som signalen hinner, med ljushastigheten, under en periodtid T . Denna sträcka är så betydelsefull att den fått ett eget namn, våglängd.

Om man ser den elektromagnetiska signalen som en vågrörelse så är våglängden avståndet mellan vågtopparna, avståndet mellan punkter som har samma fasläge.

Våglängden beräknas som signalens ljushastighet i det medium där den rör sig, multiplicerat med periodtiden. Periodtiden är samma som ett dividerat med frekvensen.

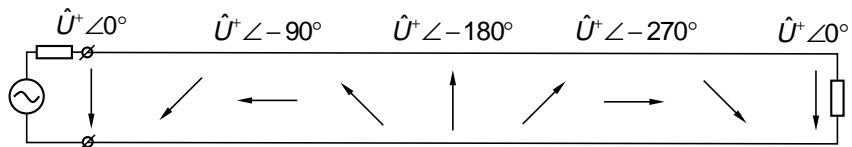
Grekiska bokstaven lambda (λ)

Våglängd betecknas med den grekiska bokstaven lambda (λ).

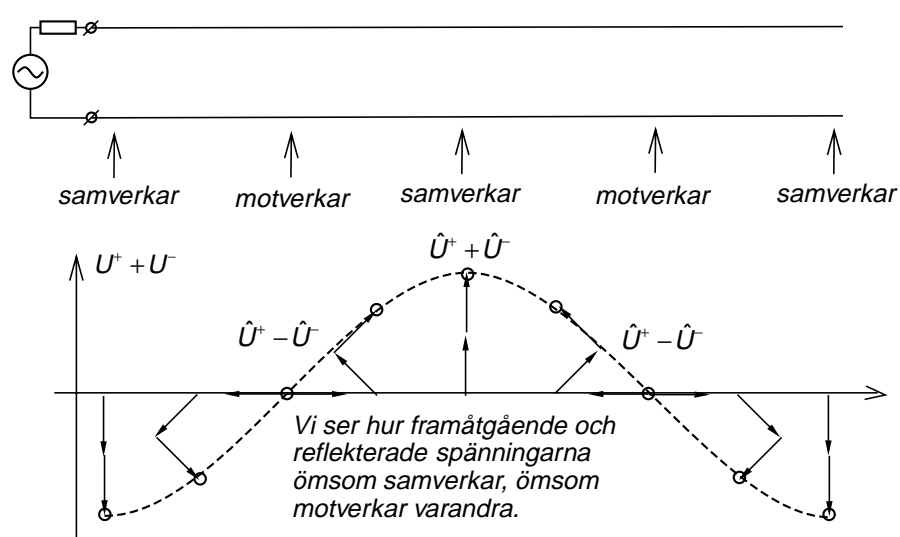
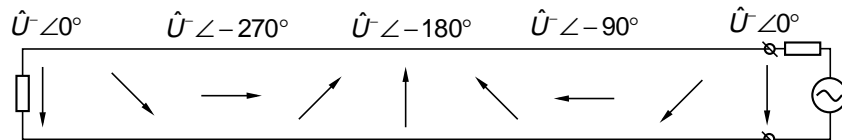


13.2 Stående vågor

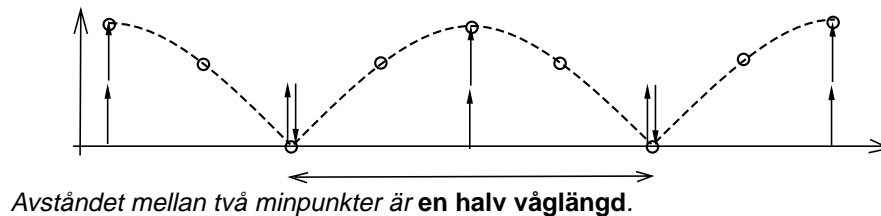
Här går signalen framåt på ledningen. Fasen vrids i negativ riktning.



Här är enbart en reflekterad signal. Fasen tycks vridas i positiv riktning om vi ser från vänster till höger.



$|U^+ + U^-|$ Summaspänningens amplitud bildar en stillastående våg, en stående våg.



13 — Smithdiagrammet, anpassning

13.3 Stående vågförhållande

En ledning med karakteristiska impedansen $Z_0 = 50 \Omega$ matas av ett batteri på 2 V i serie med en resistans på 50Ω . Belastningsimpedansen är $Z_L = 150 \Omega$.

Detta liknar "masken", se föregående kapitel. Vi får spänningsdelning. Framåtgående spänningen blir 1 V.

Reflexionskoefficienten vid lasten blir 0,5. Detta innebär att den reflekterade spänningen blir 0,5 V.

När den reflekterade signalen kommer till generatorändan ser den 50Ω , och den reflekterade signalen absorberas. Det blir ingen reflex tillbaka ut på ledningen.

Totala framåtgående spänningen blir alltså 1 V, och totala reflekterade spänningen blir 0,5 V. Vi har inga multipelreflexer.

Totala spänningen på ledningen, mätt med voltmeter, blir 1,5 V. I vilken punkt vi än mäter så är spänningen 1,5 V.

Byt batteriet mot en växelspanningsgenerator.

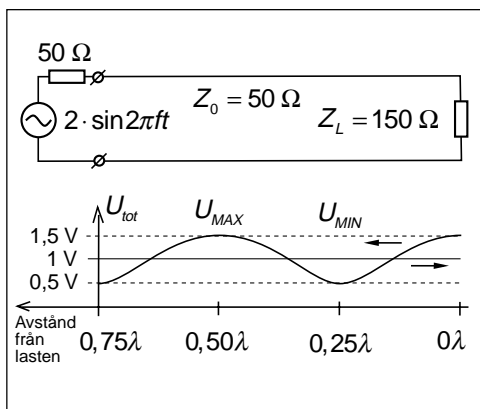
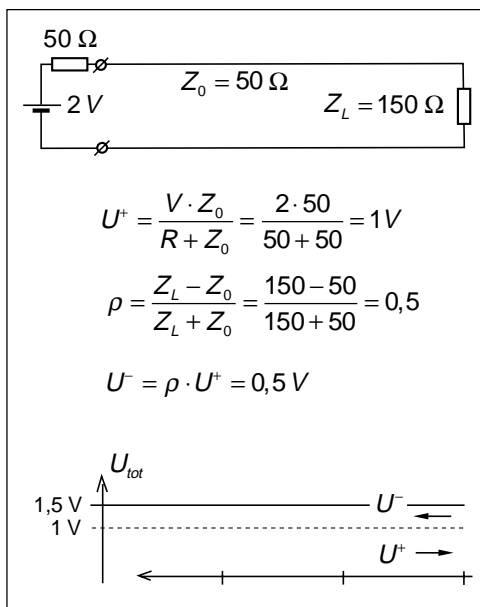
Nu byter vi ut batteriet mot en växelspanningsgenerator. Precis som tidigare får vi 1 V framåtgående spänning och 0,5 V reflekterad spänning. Detta är visserligen växelspanningar, men det finns ingen skillnad i beräkningen jämfört med likspänningsfallet, så länge som vi tittar på spänningarna var för sig.

Totala spänningen

Nu lägger vi ihop spänningarna för att få totala spänningen.

Vid lasten Z_L är reflexionskoefficienten 0,5. Detta innebär att reflekterade spänningen är 0,5 av den framåtgående. Summan blir 1,5 V.

Men nu flyttar vi oss bakåt från lasten. På avståndet $0,25\lambda$ från lasten så är gångvägsskillnaden (Δx) mellan de två spänningarna $0,5\lambda$, eftersom den reflekterade spänningen gått först $0,25\lambda$ till lasten och sedan $0,25\lambda$ tillbaka. En skillnad på $0,5\lambda$ innebär att fasvinkeln mellan spänningarna är π och framåtgående och reflekterade spänningarna hamnar i motfas.



Stående vågor

Summaspänningen, mätt med voltmeter, följer ett vågmönster, och avståndet mellan vågtopparna, eller mellan dalarna, är en halv våglängd.

Maxspänningen bildas av summan av framåtgående och reflekterade spänningarna. Spänningarna ligger i fas.

Minspänningen bildas av skillnaden mellan framåtgående och reflekterade spänningarna. Spänningarna ligger i motfas.

SVF (stående vågförhållande)

Maxspänningen dividerat med minspänningen kallas stående vågförhållandet, SVF (VSWR, Voltage Standing Wave Ratio).

Om vi inte har någon reflekterad spänning så blir SVF = 1. Det är detta som eftersträvas.

Om man mäter max- och minspänningarna så kan vi beräkna stående vågförhållandet SVF. Men vi kan även beräkna amplituden på den framåtgående spänningen och amplituden på den reflekterade spänningen.

När vi känner framåtgående och reflekterade spänningarna till beloppet, så kan vi beräkna reflexionskoefficientens belopp.

Vi ser även att reflexionskoefficientens belopp kan användas för att beräkna SVF. Faktum är att reflexionskoefficientens belopp och SVF är samma sak eftersom de utgår från samma siffror, bara olika beräkningar.

Ur framåtgående och reflekterade spänningarna kan vi även beräkna framåtgående och reflekterade effekterna, med hjälp av ledningens karakteristiska impedans.

Effekterna kan även användas för att beräkna reflexionskoefficientens belopp (eller SVF).

Reflexionskoefficient i dB

I antenndatablad är det fortfarande vanligast att t.ex skriva $SVF \leq 3$ för att ange maximala "missanpassningen", maximala avvikelserna från Z_0 . Men man övergår mer och mer till att ange reflexionskoefficienten i dB:

$$r = 20 \log |\rho| = 20 \log(0,5) = -6 \text{ dB}$$

$$\begin{cases} U_{MAX} = \hat{U}^+ + \hat{U}^- = 1,5 \text{ V} \\ U_{MIN} = \hat{U}^+ - \hat{U}^- = 0,5 \text{ V} \end{cases}$$

$$\begin{cases} SVF = \frac{U_{MAX}}{U_{MIN}} = \frac{\hat{U}^+ + \hat{U}^-}{\hat{U}^+ - \hat{U}^-} = \frac{1,5}{0,5} = 3 \\ \hat{U}^+ = \frac{U_{MAX} + U_{MIN}}{2} = \frac{1,5 + 0,5}{2} = 1 \text{ V} \\ \hat{U}^- = \frac{U_{MAX} - U_{MIN}}{2} = \frac{1,5 - 0,5}{2} = 0,5 \text{ V} \end{cases}$$

$$\begin{aligned} |\rho| &= \frac{\hat{U}^-}{\hat{U}^+} = \frac{U_{MAX} - U_{MIN}}{U_{MAX} + U_{MIN}} = \\ &= \frac{U_{MAX} - 1}{U_{MAX} + 1} = \frac{SVF - 1}{SVF + 1} = \frac{3 - 1}{3 + 1} = 0,5 \end{aligned}$$

$$\begin{aligned} SVF &= \frac{U_{MAX}}{U_{MIN}} = \frac{\hat{U}^+ + \hat{U}^-}{\hat{U}^+ - \hat{U}^-} = \\ &= \frac{1 + \hat{U}^-}{1 - \hat{U}^-} = \frac{1 + |\rho|}{1 - |\rho|} = \frac{1 + 0,5}{1 - 0,5} = \frac{1,5}{0,5} = 3 \end{aligned}$$

$$P^+ = \frac{\left(\frac{\hat{U}^+}{\sqrt{2}}\right)^2}{Z_0} = \frac{\left(\frac{1}{\sqrt{2}}\right)^2}{50} = 10 \text{ mW}$$

$$P^- = \frac{\left(\frac{\hat{U}^-}{\sqrt{2}}\right)^2}{Z_0} = \frac{\left(\frac{0,5}{\sqrt{2}}\right)^2}{50} = 2,5 \text{ mW}$$

$$\begin{aligned} \frac{P^-}{P^+} &= \left(\frac{\hat{U}^-}{\hat{U}^+}\right)^2 \Rightarrow \\ \Rightarrow |\rho| &= \sqrt{\frac{P^-}{P^+}} = \sqrt{\frac{2,5}{10}} = 0,5 \end{aligned}$$

13.4 Att mäta ståendevågförhållandet

Slitsad ledning

Mellan de två ledarna i en ledning finns ett elektriskt fält som är proportionellt mot totala spänningen mellan ledarna, d.v.s. summan av framåtgående och reflekterade spänningarna. Detta elektriska fält kan vi mäta, genom att föra in en liten "antenn". Denna korta ledningsstump gör inte kontakt med ledarna. Den befinner sig bara i det elektriska fältet. Om fältet är ett växelfält (från växelspänning) så knuffar och rycker fältet i ledningsstumpens elektroner. Det induceras en liten växelsignal på stumpen, och den signalen kan vi leda till en detektor och mäta. Detta fungerar alltså inte vid likspänning, bara vid växelspänning.

Bilden visar en "slitsad" koaxialledning. Slitsen är uppskuren i skärmen på sådant sätt att inga strömbanor skärs av (strömmen går längs ledaren, inte runt ledaren). Slitsens påverkan på ledningen kan därför försummas.

I slitsen skjuter vi in den lilla ledningsstumpen som förs längs slitsen. I varje punkt noterar vi ledningsstumpens läge samt spänningsnivån. Den slitsade koaxialledningen är luftfylld. Signalen går med c_0 . Vi kan mäta avståndet mellan minpunkter (maxima är bredare) och beräkna våglängden/frekvensen, och max och minpunkterna ger oss SVF. Vi skall även se att vi kan beräkna lastimpedansen.

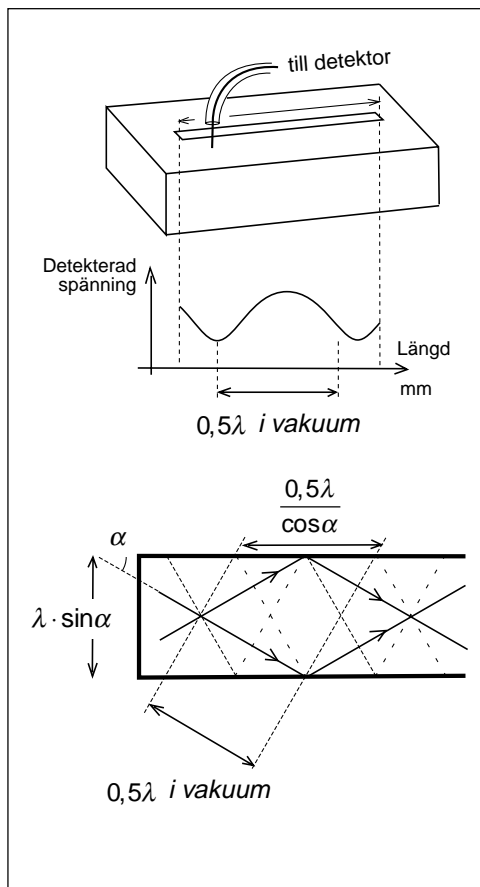
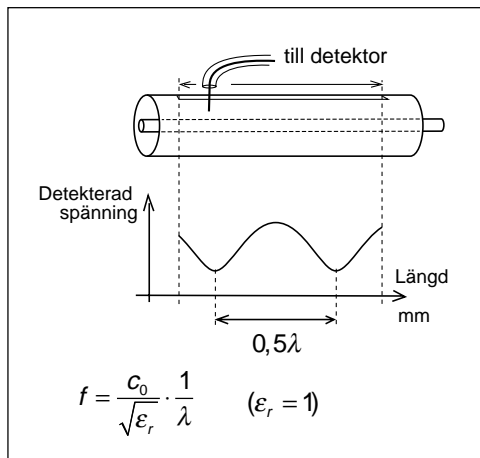
Vågledare

Samma mätning med en liten mätantenn kan utföras på en vågledare, och vi får max- och minvärden som svarar mot framåtgående och reflekterade signalerna i vågledaren, och kan beräkna SVF.

Avståndet är längre

När det gäller avståndet mellan minpunkterna så kommer detta att bli längre än en halv våglängd i luft. Det ser ut som om signalen går fortare än ljushastigheten i vakuum.

Vågledaren består av en "plattledning", övre och undre metallytorna, där plattledningens sidor förbundits med metall. Enda möjligheten för signal att gå i en så-



dan ledning är att signalen delar upp sig på två signaler som går i sicksack, reflekteras i väggarna. När signalen går med sådan vinkel att $\lambda \cdot \sin \alpha$ svarar mot bredden på vågledaren, då kommer summan av de två signalerna att ge ett totalt elektriskt fält som är noll vid plåtväggarna.

För den sammansatta signalen, de båda sicksack-signalerna tillsammans, definierar man en hastighet som kallas fashastighet. Fashastigheten bestämmer avståndet mellan vågtoppar vid SVF-mätning.

$$v_{fas} = \frac{c_0}{\sqrt{\epsilon_r}} \cdot \frac{1}{\cos \alpha}$$

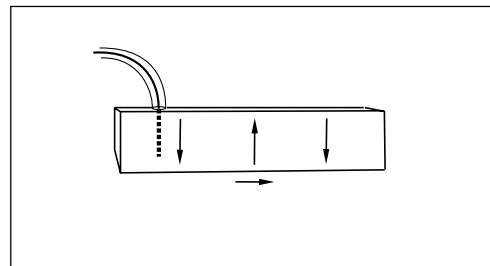
Samtidigt kommer signalen fram senare i borte ändan av vågledaren, eftersom signalerna går i sicksack. Tidsfördröjningen beräknas med hjälp av gruppshastigheten.

$$v_{grupp} = \frac{c_0}{\sqrt{\epsilon_r}} \cdot \cos \alpha$$

Begreppen fas- och gruppshastigheter existerar bara vid vågledare och är ett resultat av att man inte betraktar signalen som två signaler utan definierar en ny signaltyp som är summan av de båda sicksack-signalerna.

Övergången från vågledare till koaxialledning sker med en liten "antenn" som sticks ner i vågledaren.

Trattantennar som övergår i en vågledare är vanligt som matarantennar till parabolantennar.



13 — Smithdiagrammet, anpassning

Riktkopplaren

Om jag parallellt med en ledning lägger en liten ledningsstump så att ledningarna får ett gemensamt utrymme, så händer mycket förenklat följande:

När den framåtgående signalen rusar förbi ledningsstumpen så kommer en positiv laddning att "dra med sig" en negativ laddning. En negativ laddning är samma sak som en positiv laddning som rör sig i andra riktningen.

Innebörden av detta är att den framåtgående signalen gör en liten kopia av sig själv på ledningsstumpen, men kopian går i motsatt riktning.

Samma händer med den reflekterade signalen.

Nu har vi en ledningsstump med miniatyr-signaler som är kopior av förhållandet på huvudledningen. Om vi avslutar ledningsstumpen med ett motstånd som motsvarar karakteristiska impedansen på ledningsstumpen, så kommer ena kopian att hamna i motståndet medan andra kopian kan plockas ut och mätas.

Genom att sätta motståndet i andra ändan så plockar vi ut den andra kopian.

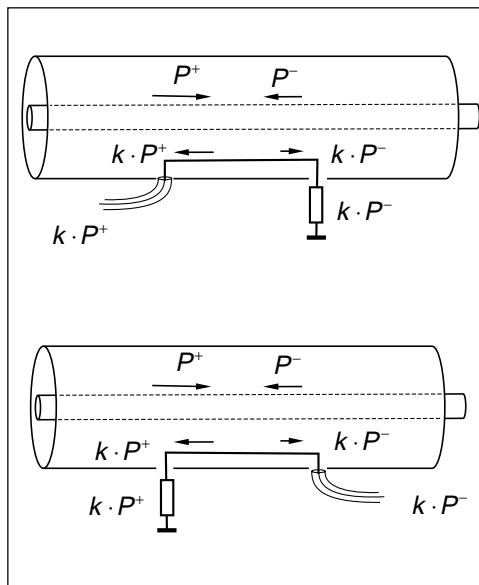
Vi har alltså en anordning som antingen plockar ut en liten kopia av den framåtgående signalen, eller av den reflekterade signalen, en riktkopplare, eftersom den är känslig för signalens riktning på ledaren.

Den vanligaste SVF-metern

Riktkopplaren med en parallell ledningsstump är i dag den vanligaste SVF-metern och används från 3 MHz och uppåt i frekvens. Det vanligaste är emellertid inte att kalibrera instrumentet i SVF utan direkt i framåtgående och reflekterad effekt.

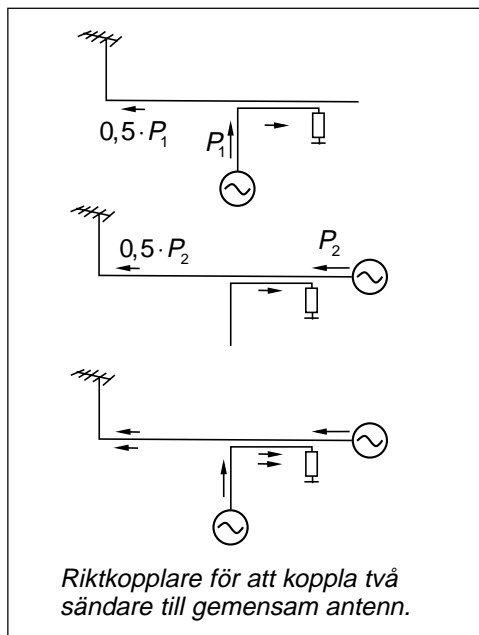
Effektmetern

Man avläser direkt den framåtgående effekten och den reflekterade effekten på ledningen. Hur mycket effekt hamnar i antennen? Skillnaden. Framåtgående minus reflekterad effekt.



Riktkopplare används även till annat

Om ledningsstumpen går långt in i den andra ledningen kan man få ut ända upp till halva signalen. En sådan riktkopplare kallas 3 dB-kopplare, eftersom man får ut halva effekten på ledningsstumpens port, och den andra halvan fortsätter rakt igenom.



13 — Smithdiagrammet, anpassning

Ytterligare en riktkopplare

Ytterligare en riktkopplare som används som effektmeter visas på bilden.

Den består av två delar, en kapacitiv spänningsdelare som ger oss en liten del av totala spänningen på ledaren.

$$k \cdot (U^+ + U^-)$$

Dessutom finns en strömtransformator, där en liten del av strömmen matas genom motstånd och ger en spänning som är proportionell mot totala strömmen på ledningen, som är proportionell mot skillnaden mellan framåtgående och reflekterade spänningarna.

$$R \cdot (I^+ - I^-) = k \cdot (U^+ - U^-)$$

Ena porten ger framåtgående signalen, andra porten ger reflekterade signalen.

Används i kabel-tv-uttaget

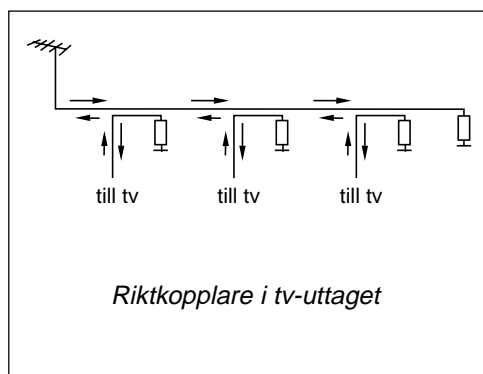
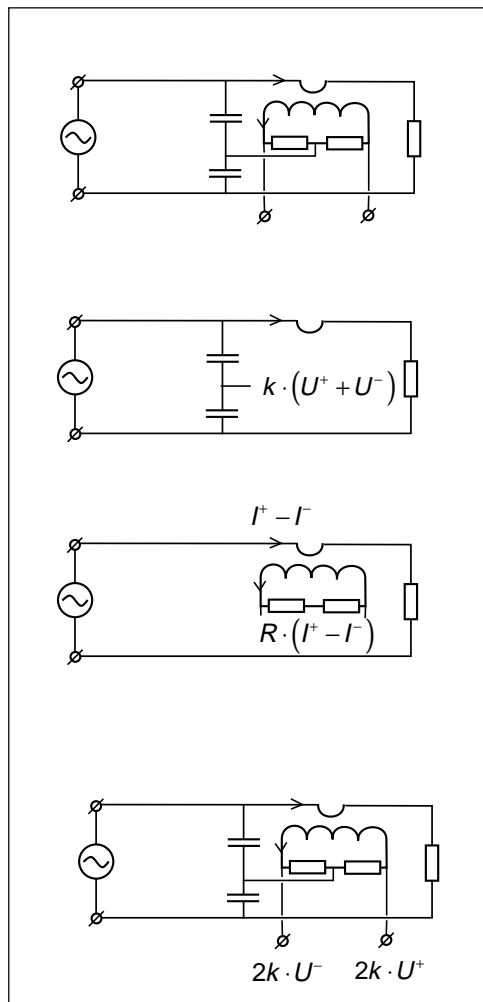
Riktkopplare av detta slag används dels i effektmätar från låg frekvens upp till några hundra MHz. Dessutom sitter anordningen (hälften) för att plocka ut den framåtgående signalen, i det vanliga kabel-tv-uttaget.

En tv-mottagare absorberar signalen på den kanal som den är inställd på för att ta emot. Men alla andra tv-signaler på alla andra frekvenser, de signalerna totalreflekteras och matas tillbaka ut på kabel-tv-nätet. Riktkopplaren gör att dessa reflexer bara finns på huvudmatarkabeln som reflekterade signaler, medan den enda signal som finns i framriktningen är den "rena" signalen från antennförstärkaren.

Nätverksanalysator

Om man mäter inte bara amplituderna på framåtgående och reflekterade spänningarna, utan även fasvinkeln mellan spänningarna, så kan man beräkna reflexionskoefficienten och, via karakteristiska impedansen, beräkna den verkliga impedansen.

Ett instrument som gör på detta sätt och som presenterar resultatet som en punkt på en oscilloskopskärm där Smithdiagrammet är inritat, kallas nätverksanalysator.



13 — Smithdiagrammet, anpassning

13.5 Smithdiagrammet

Vi utgår från en transmissionsledning som avslutas med en belastningsimpedans Z_L . Belastningsimpedansen kan vara reell eller komplex.

Över belastningsimpedansen har vi spänningar, den framåtgående spänningen och den reflekterade spänningen. Spänningarna har både amplitud och fasvinkel.

Vi kan beräkna reflexionskoefficienten ρ_L vid lastimpedansen, antingen genom att utgå från lastimpedansen Z_L och ledningens karakteristiska impedans Z_0 , eller genom att utgå från spänningarna U^- och U^+ . Även reflexionskoefficienten ρ_L är ett komplext tal, med amplitud och fasvinkel.

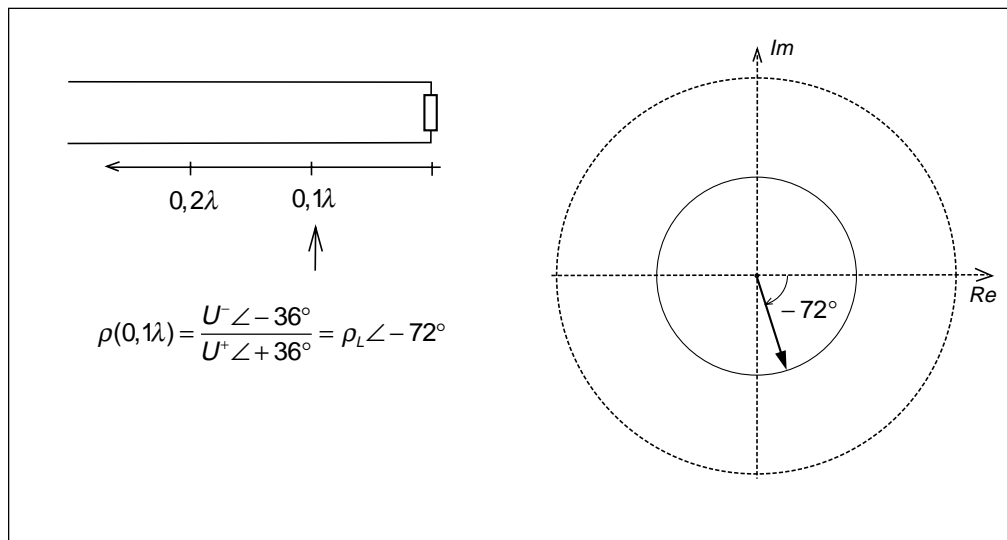
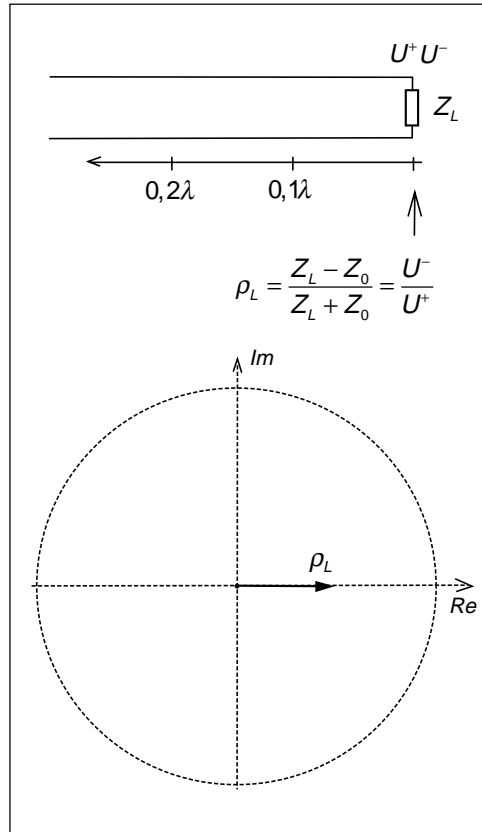
Rita reflexionskoefficienten i ett visardiagram

Vi kan rita ρ_L i visardiagrammet. I detta fall är fasvinkeln för ρ_L lika med noll.

Om vi flyttar oss längs ledningen?

Vad händer om vi flyttar oss $0,1\lambda$ mot generatorn? Framåtgående spänningen ökar sin fasvinkel. Spänningen kommer tidigare. $0,1\lambda$ svarar mot 36° . Den reflekterade spänningen har längre väg att gå. Den kommer senare, -36° senare.

Vi kan definiera en "tänkt" reflexions-



13 — Smithdiagrammet, anpassning

koefficient i den punkten, ett förhållande mellan U^- och U^+ , som är ρ_L som vridit sig -72° . Vi kan alltså gradera den yttre cirkeln i grader, och ett varv, 360° , svarar mot att man flyttar sig $0,5\lambda$ längs ledningen.

Enklare att gradera i våglängder

Men att hålla på och räkna om till grader (dubbla grader) är omständligt. Bättre är att införa ytterligare en cirkel graderad i våglängder, med gradering både för förflyttning mot generatoren (medurs), och gradering för förflyttning mot lasten (moturs).

Varför reflexionskoefficient?

Varför ha ett diagram för reflexionskoefficienten? Till varje reflexionskoefficient svarar ett visst bestämt värde på impedansen normerad till (dividerat med) ledningens karakteristiska impedans.

Impedansdiagram

Alltså graderar vi diagrammet i normerad impedans.

- Vi får en normerad resistansaxel från 0Ω till $\infty \Omega$ som löper från vänster till höger, med 1Ω mitt i diagrammet.
- Vi får en normerad positiv reaktansaxel som löper från $+j0 \Omega$ till $+j\infty \Omega$ längs den övre halvcirkeln, med $+j1 \Omega$ rakt upp.
- Vi får en normerad negativ reaktansaxel som löper från $-j0 \Omega$ till $-j\infty \Omega$ längs den undre halvcirkeln, med $-j1 \Omega$ rakt ner.

En ren resistans hamnar någonstans på resistansaxeln.

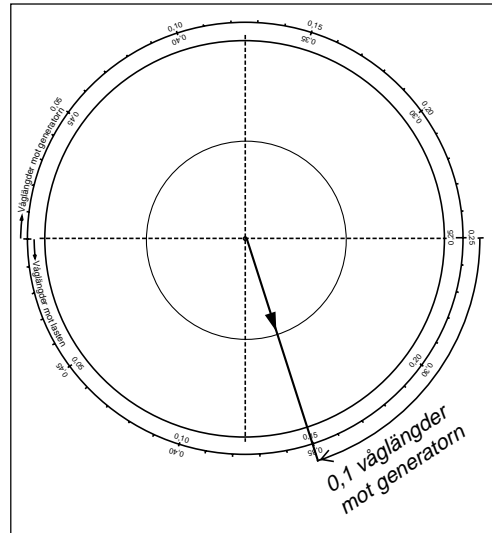
En resistans i serie med en spole hamnar i övre halvplanet.

En resistans i serie med en kondensator hamnar i undre halvplanet.

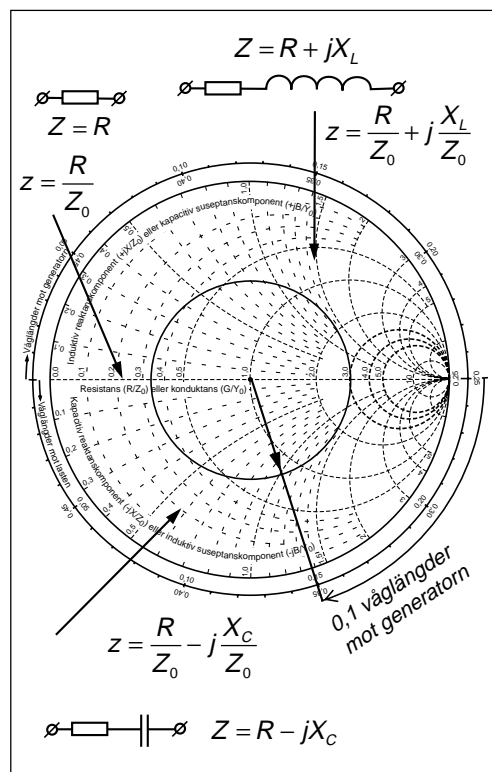
Smithdiagrammet

Detta impedansdiagram kallas Smithdiagrammet efter sin uppfinnare. Det publicerades första gången januari 1944 i tidskriften Electronics.

Smithdiagrammet går precis lika bra att använda för admittanser, se admittansgradering.



$$\rho = \frac{Z - Z_0}{Z + Z_0} \quad \frac{Z}{Z_0} = \frac{(1 + \rho)}{(1 - \rho)}$$



13 — Smithdiagrammet, anpassning

Uppgift:

Beräkna inimpedansen i en ledning med karakteristiska impedansen 50Ω och $\epsilon_r = 2,25$ som är 25 cm lång. Frekvensen är 80 MHz och ledningen avslutas med belastningsresistansen 150Ω .

1. Normera

Ledningen avslutas med $Z_L = 150 \Omega$. Beräkna z_L .

$$z_L = \frac{R}{Z_0} + j \frac{X}{Z_0} = \frac{150}{50} = 3 + j0$$

Belastningsimpedansen ligger vid värdet 3 på resistansaxeln.

2. Rita cirkel

Nästa steg är att rita cirkeln. Passarspetsen skall sitta mitt i diagrammet! Det är längs denna cirkel vi skall förflytta oss.

3. Drag streck

Drag ett streck från mittpunkten genom belastningsimpedansen så att vi kommer ut till cirkeln graderad i våglängder.

4. Förflytta dig längs ledningen

Nästa steg är att förflytta sig längs ledningen. Vart går vi? Mot lasten eller mot generatorm? I detta fall går vi mot generatorm, alltså medurs (läs i diagrammet så du inte snurrar åt fel håll).

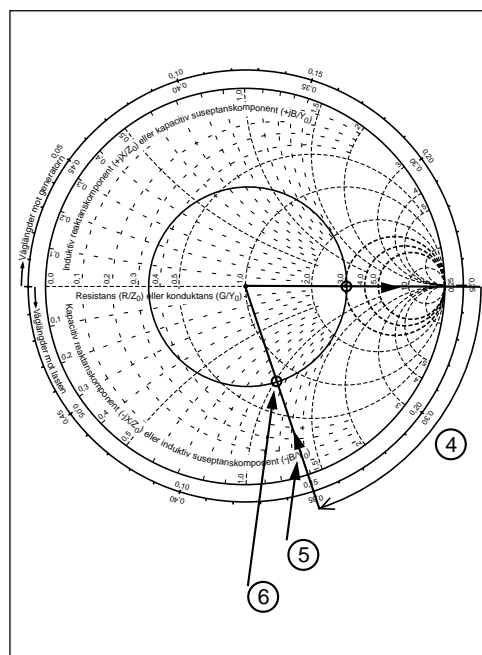
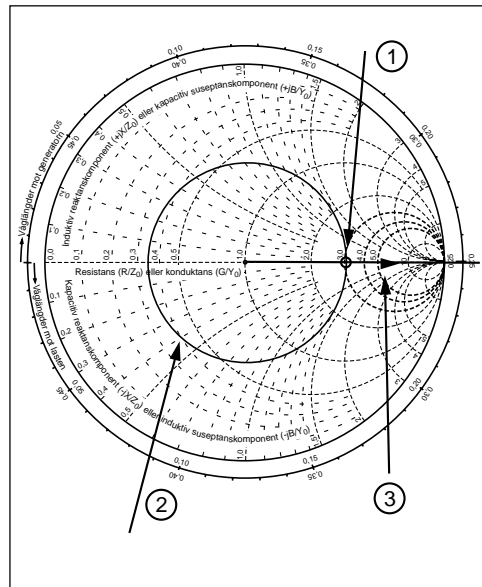
Hur långt skall vi förflytta oss? Hur många våglängder är 25 cm.

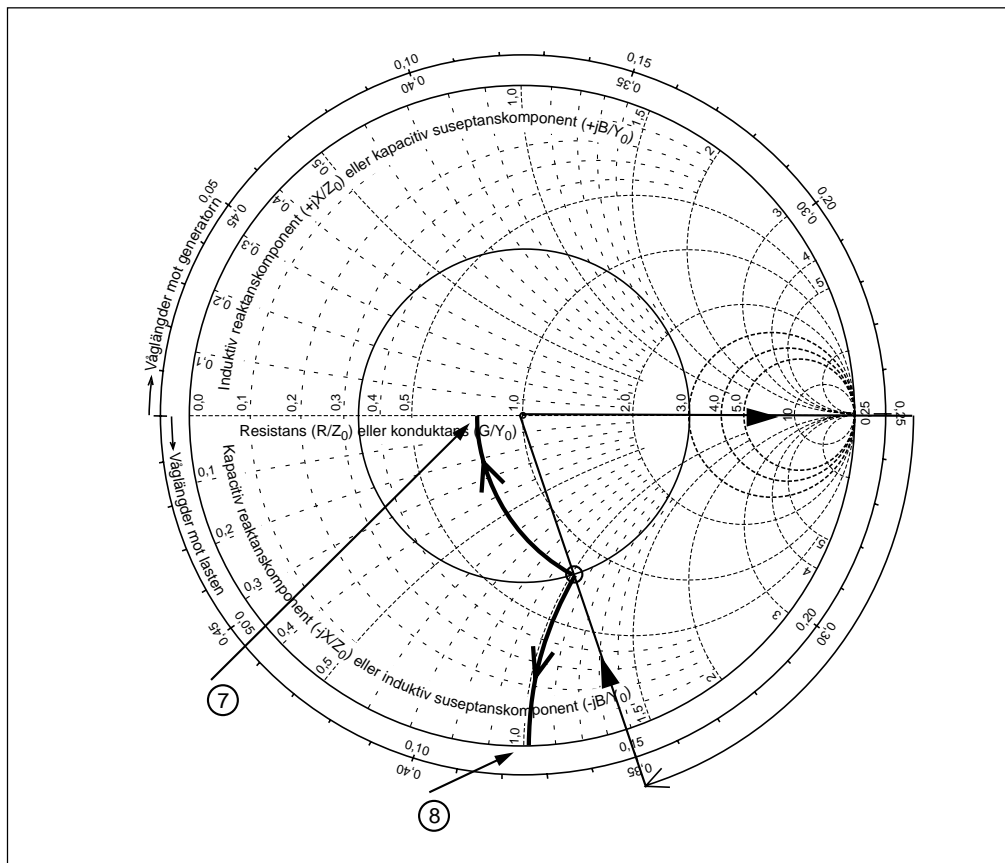
$$\begin{aligned} L(\lambda) &= \frac{25 \text{ cm}}{\lambda} = \frac{0,25}{\frac{c_0}{\sqrt{\epsilon_r}} \cdot \frac{1}{f}} = \\ &= \frac{0,25}{\frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot \frac{1}{80 \cdot 10^6}} = 0,1 \lambda \end{aligned}$$

5. Drag streck

Drag streck in mot diagrammets mittpunkt så att du får en skärningspunkt med cirkeln.

6. Avläs normerade impedansen





Avläs normerade impedansen på följande sätt:

7. Avläs normerade resistansen

Följ en hjälplinje in mot resistansaxeln. Se upp så att du inte följer din egen cirkel i stället för en hjälplinje.

Avläst värde: $r = 0,78$

8. Avläs normerade reaktansen

Följ hjälplinjerna ut mot yttercirkeln. Det finns siffermarkeringar på andra ställen i diagrammet. Använd inte dessa utan gå alltid ut mot yttercirkeln. Se upp så att du följer hjälplinjerna och inte ditt eget streck.

Avläst värde: $x = -j1,02$

Beräknad inimpedans:

$$\begin{aligned} Z_{IN} &= z_{IN} \cdot Z_0 = (r + jx) \cdot Z_0 = \\ &= (0,78 - j1,02) \cdot 50 = 39 - j51 \Omega \end{aligned}$$

Glöm inte att multiplicera med Z_0 . När vi går in i diagrammet normerar vi. När vi går ut "avnormerar" vi.

13 — Smithdiagrammet, anpassning

13.6 Impedansbestämning med slitsad ledning

Uppgift:

En okänd last ansluts till en slitsad ledning med $Z_0 = 50 \Omega$. Den slitsade ledningen har $\epsilon_r = 1,0$.

Man uppmäter minpunkter vid 102 mm och vid 418 mm när lasten inte är ansluten (öppen ledning).

När lasten ansluts blir maxspänningen 65 mV och minspänningen 45 mV. Spänningsmin uppträder vid 322 mm.

Beräkna generatorns frekvens och lastens impedans.

Beräkna våglängd och frekvens

Avståndet mellan två spänningsmin är $0,5\lambda$ (detta motsvarar en hel våglängd fram och tillbaka).

$$\lambda = 2 \cdot (0,418 - 0,102) = 0,632 \text{ m}$$

$$f = \frac{c_0}{\sqrt{\epsilon_r}} \cdot \frac{1}{\lambda} = \frac{3 \cdot 10^8}{\sqrt{1}} \cdot \frac{1}{0,632} = 474,68 \text{ MHz}$$

Avstånd till lasten

Vi måste veta var på ledningen lasten kommer att vara placerad. Om vi har minimum vid 418 mm när ledningen är öppen så måste ledningen ta slut $0,25\lambda$ bortom denna punkt (eller ytterligare halva våglängder d.v.s. hela varv i Smithdiagrammet).

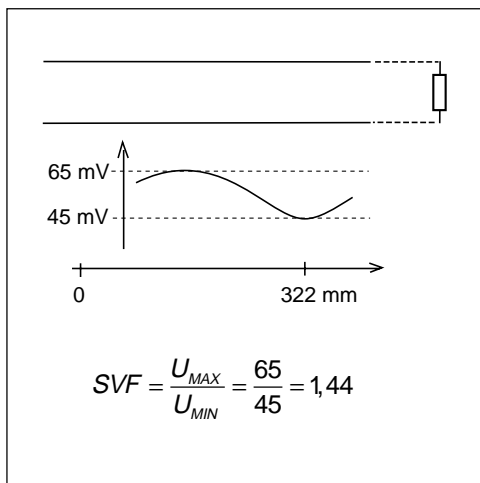
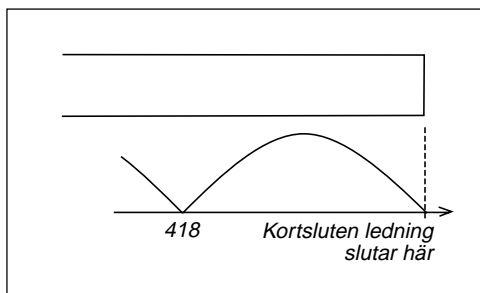
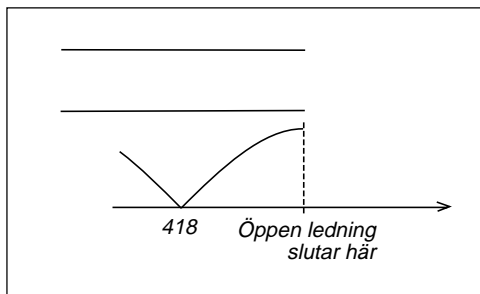
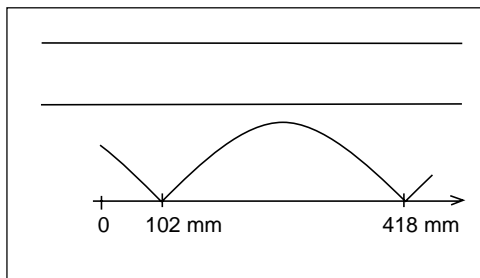
$$\text{Ledningen tar slut vid: } 418 + \frac{632}{4} = 576 \text{ mm}$$

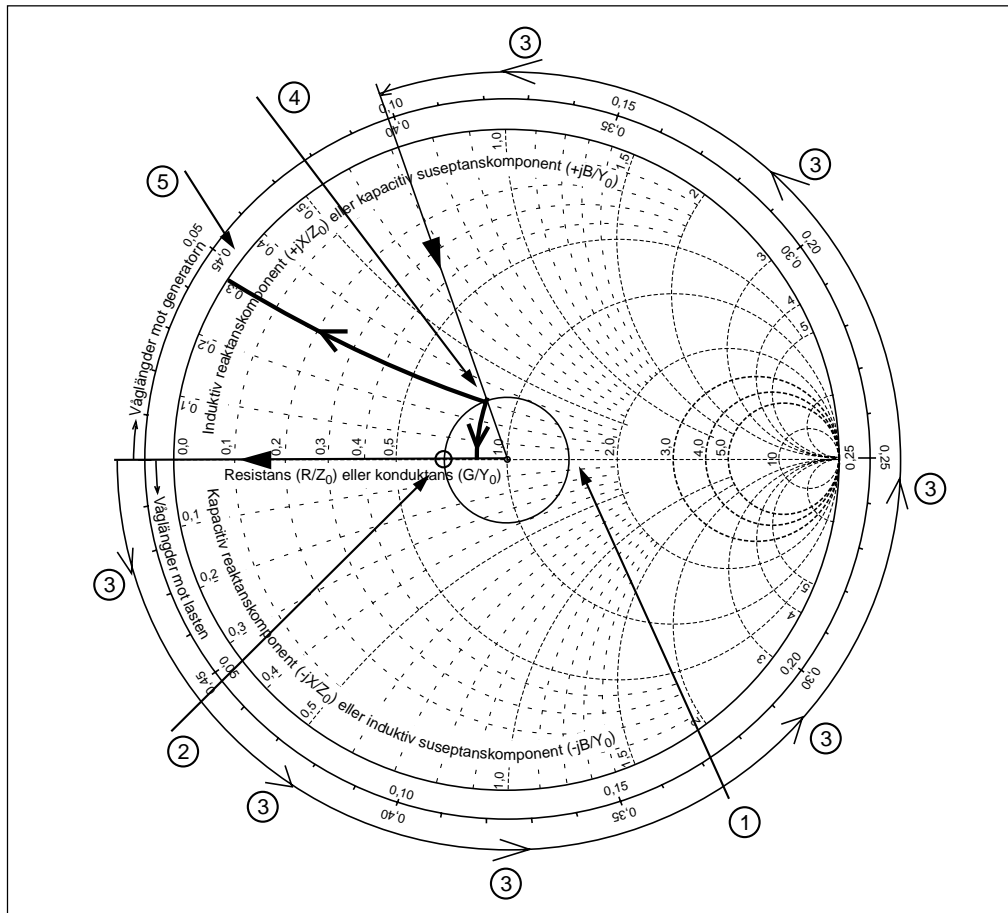
Om denna kalibrering utförts med kortsluten ledning så hade ledningen tagit slut vid 734 mm.

Anslut lasten

När lasten är ansluten har man uppmätt maxspänningen till 65 mV och minspänningen till 45 mV. Är detta effektivvärden eller toppvärden?

Vi vet inte, och vi bryr oss inte. Detta är inte de verkliga spänningarna på ledningen utan bara de små spänningar som bildas i mätsonden, den lilla mätantennen. Det enda vi skall göra med spänningarna är att dividera dem med varandra.





1. Rita cirkeln

Vi har SVF = 1,44. Detta representerar en cirkel som går genom $r = 1,44$. Rita cirkeln.

2. Var ligger spänningsminimum?

Spänningsminimum får vi där cirkeln skär resistansaxeln till vänster (-180° fasvinkel hos ρ). På motsvarande sätt ligger spänningsmaximum där cirkeln skär resistansaxeln på höger sida. Drag strecket!

3. Snurra mot lasten

Nu skall vi förflytta oss ett antal våglängder mot lasten. Tidigare har vi beräknat att ledningen slutar vid 576 mm. Vi befinner oss (minimum) vid 322 mm. Vi skall förflytta oss

$576 - 322 = 254$ mm. Detta skall divideras med våglängden.

$$L(\lambda) = \frac{0,576 - 0,322}{\lambda} = \frac{0,254}{0,632} = 0,402 \lambda$$

4. Drag streck och få skärningspunkten med cirkeln

5. Avläs

Nu avläser vi $r = 0,85$ och $x = +j0,30$ och beräknar

$$\begin{aligned} Z_L &= z_L \cdot Z_0 = (r + jx) \cdot Z_0 = \\ &= (0,85 + j0,30) \cdot 50 = 42,5 + j15 \Omega \end{aligned}$$

13.7 Anpassning med kvartsvågstransformator

Vi avslutar en ledning med resistansen R_1 . Ledningen görs $0,25\lambda$ lång. Vad ser vi in i öppningen på denna $0,25\lambda$ långa ledning?

Vi ser en ren resistans

Om vi ansluter en ledning som är $0,25\lambda$ lång så snurrar vi ett halvt varv i Smithdiagrammet.

Om det sitter en ren resistans i ena ändan, då ser vi även en ren resistans i andra ändan. Men den andra resistansen får annat värde, R_2 .

Transformator

Att omvandla från ett resistansvärde till ett annat resistansvärde är exakt vad transformatorn gör. Denna $0,25\lambda$ långa ledning kommer att uppföra sig som en transformator och kallas kvartsvågstransformator eftersom den är en kvarts våglängd lång.

Bara för resistanser

Kvartsvågstransformatorn kommer bara att uppföra sig som en transformator om den belastas med en ren resistans.

Transformationen beror på karakteristiska impedansen hos ledningen

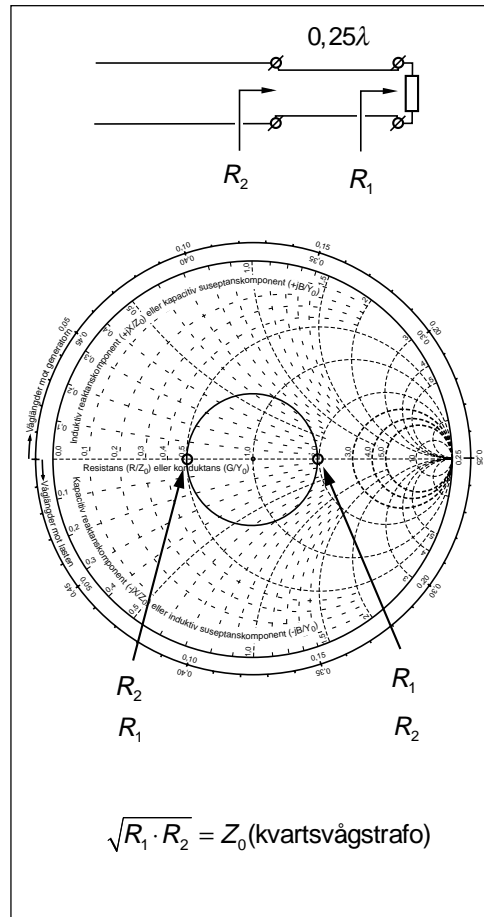
Antag att vi har $100\ \Omega$ och en kvartsvågstrafo med $Z_0 = 50\ \Omega$. Först normerar vi och får 2. Cirkeln går genom 2. På andra sidan går cirkeln genom 0,5. Alltså ser vi $25\ \Omega$ in i kvartsvågstrafon.

Nu byter vi till en kvartsvågstrafo med $Z_0 = 75\ \Omega$. Först normerar vi och får 1,33. Cirkeln går genom 1,33. På andra sidan avläser vi 0,75. Där ser vi alltså $56,25\ \Omega$.

Transformationsförhållandet beräknas på följande sätt:

Den resistans vi har i ena ändan multipliceras med det vi vill få i andra ändan, varefter vi drar roten ur. Detta ger kvartsvågsledningens karakteristiska impedans.

$$\sqrt{R_1 \cdot R_2} = Z_0 \text{ (kvartsvågstrafo)}$$



13 — Smithdiagrammet, anpassning

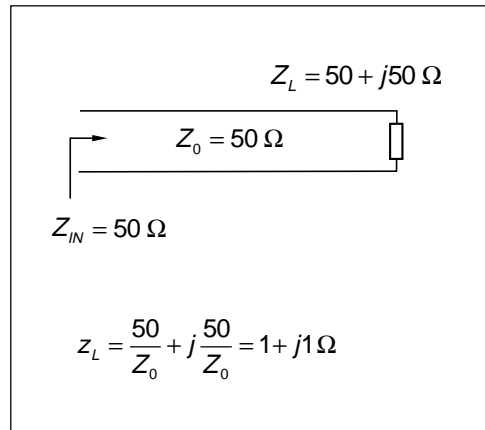
Uppgift:

En belastning på $Z_L = 50 + j50 \Omega$ ansluts till en ledning med $Z_0 = 50 \Omega$ och $\epsilon_r = 2,25$. Frekvensen är 500 MHz.

Anpassa lasten till ledningen med en kvartsvågstransformator.

- Vilken karakteristisk impedans skall kvartsvågsledningen ha?
- Var skall kvartsvågstransformatorn placeras?

1. Börja med att normera, pricka in lasten, och dra cirkeln.



2. Nästa steg är att förflytta sig mot generatoren, tills impedansen är rent resistiv. En sådan punkt har vi om vi flyttar oss $0,088\lambda$.

3. Avläs resistansvärdet och avnormera. Vi får 130Ω .

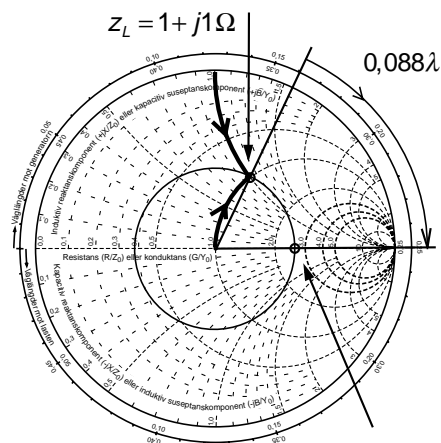
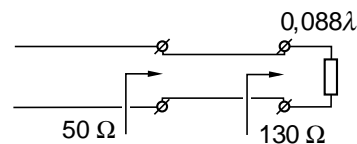
4. Nu vet vi vad vi vill ha i borte ändan. Vi vet också att vi vill ha 50Ω i andra ändan. Alltså kan vi beräkna karakteristiska impedansen för kvartsvågstransformatorn:

$$Z_0 (\text{kvartsvågstrafo}) = \sqrt{R_1 \cdot Z_0} = \sqrt{130 \cdot 50} = 80,6 \Omega$$

5. Det som återstår är att beräkna var kvartsvågstransformatorn skall kopplas in. Men vi vet redan att avståndet skall vara $0,088\lambda$.

Vi beräknar våglängden och multiplicerar:

$$L = 0,088 \cdot \frac{c_0}{\sqrt{\epsilon_r}} \cdot \frac{1}{f} = 0,088 \cdot \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot \frac{1}{500 \cdot 10^6} = 0,0352 \text{ m}$$



$$R_1 = 2,6 \cdot Z_0 = 130 \Omega$$

13 — Smithdiagrammet, anpassning

Det finns alltid två möjligheter

Men vi har ytterligare en möjlighet, nämligen att förflytta oss till den andra punkten där cirkeln skär resistiva axeln.

2. Vi kan även förflytta oss mot generatorn till den andra punkten där impedansen är rent resistiv. En sådan punkt har vi om vi flyttar oss $0,338\lambda$.

3. Avläs resistansvärdet och avnormera. Vi får $19,5 \Omega$.

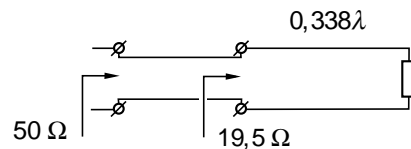
4. Nu vet vi vad vi vill ha i borte ändan. Vi vet också att vi vill ha 50Ω i andra ändan. Alltså kan vi beräkna karakteristiska impedansen för kvartsvågstransformatorn:

$$Z_0 (\text{kvartsvågstrafo}) = \sqrt{R_1 \cdot Z_0} = \sqrt{19,5 \cdot 50} = 31,2 \Omega$$

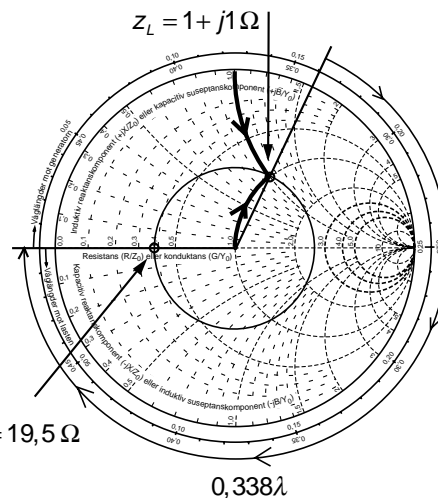
5. Det som återstår är att beräkna var kvartsvågstransformatorn skall kopplas in. Men vi vet redan att avståndet skall vara $0,338\lambda$.

Vi beräknar våglängden och multiplicerar:

$$L = 0,338 \cdot \frac{c_0}{\sqrt{\epsilon_r} \cdot f} = 0,338 \cdot \frac{3 \cdot 10^8}{\sqrt{2,25} \cdot 500 \cdot 10^6} = 0,1352 \text{ m}$$



$$Z_0 (\text{kvartsvågstrafo}) = \sqrt{R_1 \cdot Z_0} = \sqrt{19,5 \cdot 50} = 31,2 \Omega$$



$$R_1 = 0,39 \cdot Z_0 = 19,5 \Omega$$

Vad väljer man?

I verkligheten, vilken lösning väljer man?

Elektriskt kan det vara bäst att placera kvartsvågstransformatorn så nära lasten som möjligt. När man ändrar frekvens, när man avviker från den beräknade frekvensen, så flyttar sig punkten minst sträcka från resistiva axeln. Transformatorn blir bredbandigare.

Mekaniskt kan det vara lättast att konstruera en kvartsvågstransformator med låg karakteristisk impedans.

I verkligheten kan man ofta hitta lösningar där bägge villkoren är uppfyllda.

13.8 Vad händer vid anpassning?

Vi har en ledning som avslutas med en impedans som inte är lika med ledningens karakteristiska impedans. Vi får reflekterad effekt. Nu skall vi ta bort den reflekterade effekten genom att anpassa ledningen. Vart tar den reflekterade effekten vägen?

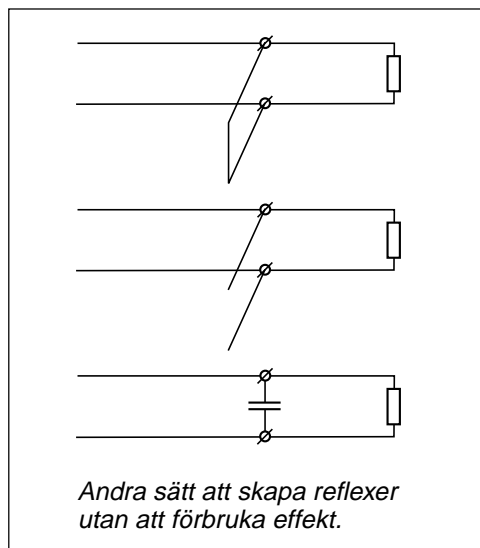
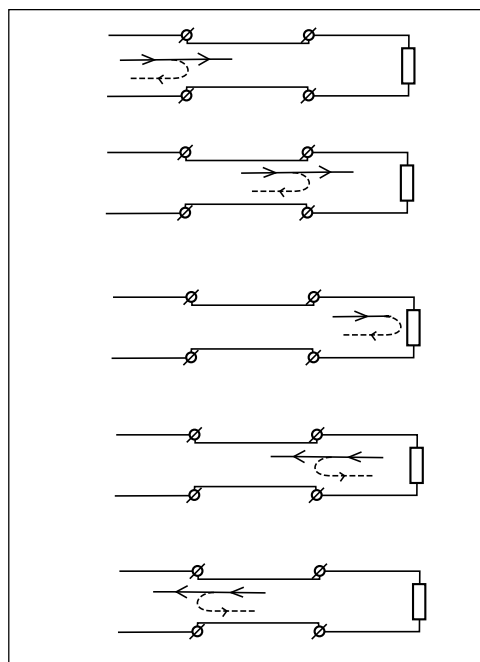
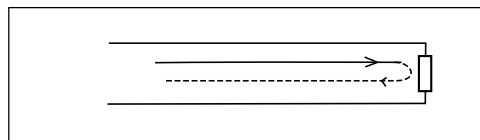
Anpassning med kvartsvågstransformator

När vi kopplar in en ledningsstump med annan karakteristisk impedans så händer följande:

- När den framåtgående signalen möter skarven mellan ledningarna så reflekteras en del av signalen.
- När den framåtgående signalen möter nästa skarv så reflekteras lika mycket fast med motsatt tecken. Denna andra reflex har gått en halv våglängd längre väg än den första reflexen. Därför byts tecknet så att dessa båda reflexer kommer att summeras. Om ledningen i stället varit en halv våglängd lång så skulle reflexerna tagit ut varandra. Detta sker t.ex. på antireflexbehandlade glasögon.
- Nu har vi skapat en reflekterad signal på vår ledning, en reflekterad signal som kommer från kvartsvågstransformatorn. Amplituden på reflexen bestäms av karakteristiska impedansen på kvartsvågsledningen.
- Den signal som går igenom kvartsvågstransformatorn träffar lasten, där en del av signalen reflekteras. Den del av denna reflex som tar sig igenom kvartsvågstransformatorn skall ha samma amplitud som reflexerna från kvartsvågstransformatorn. Sedan väljer vi avståndet mellan kvartsvågstransformatorn och lasten så att dessa reflexer hamnar i motfas och tar ut varandra.

Att skapa ytterligare reflexer

Anpassning innebär att skapa ytterligare reflexer så att summan av alla reflexer blir noll. Andra sätt att skapa reflexer är att parallellkoppla ledningen med en ledningsstump eller en kondensator.



13 — Smithdiagrammet, anpassning

17 lektioner i
TELEKOMMUNIKATION

13.9 Anpassning med kortsluten stubbe

Vi utför anpassningen, inte genom att titta på spänningarna som går fram och tillbaka utan genom att titta på impedanserna. Impedansen innehåller både den framåtgående och den reflekterade spänningen samtidigt (totala spänningen).

Parallellkoppling: Admittanser

Vid seriekoppling lägger vi ihop impedanser. Vid parallellkoppling lägger vi ihop admittanser. Vi skall räkna med admittanser. Men var lugn, det är enkelt när vi har Smithdiagrammet.

Tekniken går ut på att snurra i Smithdiagrammet tills vi korsar den cirkel där admittansen har konduktans ("resistiv" del) som är 1, normerad admittans. Men vi har susceptans. Susceptansen tar vi bort genom att parallellkoppla med ett objekt som har samma susceptans med motsatt tecken.

Exempel

En belastning på $Z_L = 100 + j25 \Omega$ ansluts till en ledning med $Z_0 = 50 \Omega$ och $\epsilon_r = 2,25$. Frekvensen är 500 MHz.

Lasten skall anpassas med en kortsluten stubbe, tillverkad av samma ledning.

Beräkna inkopplingspunkten A samt stubbens längd B (fysisk längd, inte våglängder).

1. Normera och rita cirkel

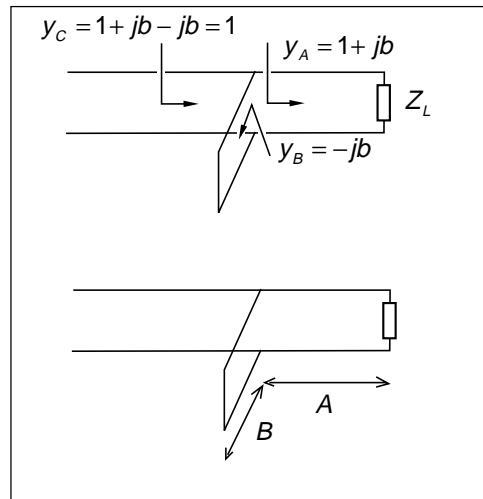
Börja på vanligt sätt. Beräkna normerade impedansen, pricka in och rita cirkeln.

2. Övergå till admittans

Nu övergår vi till admittanser. Lastens admittans ligger "på andra sidan" mittpunkten, i punkt 2. Vi avläser inte utan går ut till våglängdsgraderingen.

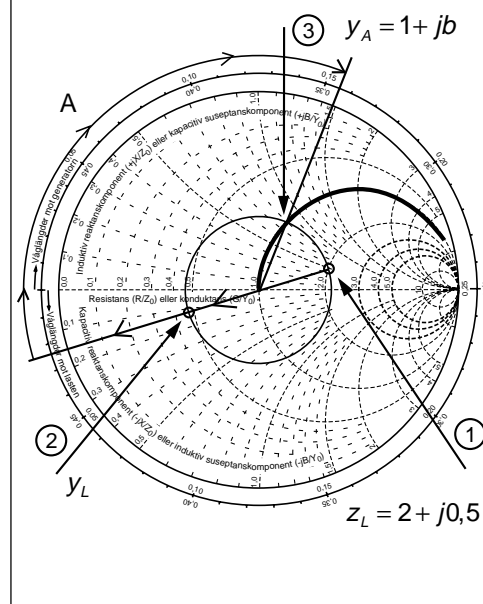
3. Förflytta dig längs ledningen tills admittansens konduktansdel är 1, i övre halvplanet.

Det är viktigt med övre halvplanet när vi har kortsluten stubbe. Det går även med skärningspunkten i undre halvplanet, men stubben blir längre. Håll dig till en lösningsmetod.



$$Z_L = 100 + j25 \Omega$$

$$z_L = \frac{R}{Z_0} + j \frac{X}{Z_0} = \frac{100}{50} + j \frac{25}{50} = 2 + j0,5$$



13 — Smithdiagrammet, anpassning

Nu kan vi avläsa sträckan A, som blir 0,179 våglängder. Detta motsvarar 7,16 cm vid 500 MHz och den aktuella ledningen.

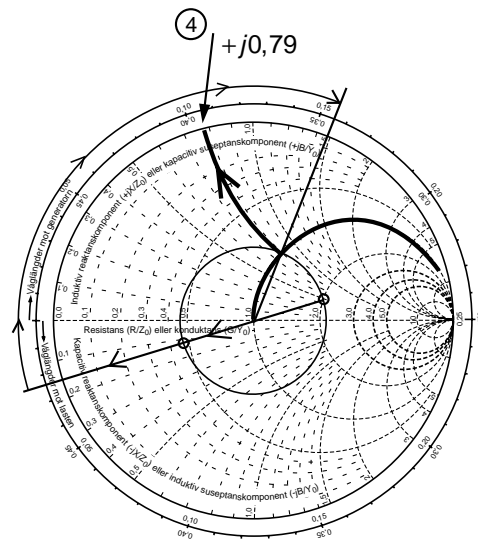
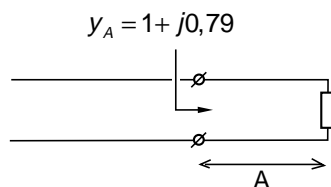
$$A = 0,024 + 0,155 = 0,179 \lambda$$

$$A(\lambda) = 0,179 \cdot \lambda =$$

$$= 0,179 \cdot \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot \frac{1}{500 \cdot 10^6} = 0,0716 \text{ m}$$

4. Admittansen mot lasten

Nu avläser vi normerade admittansen i punkt 3, som blir $y_A = 1 + j0,79$.

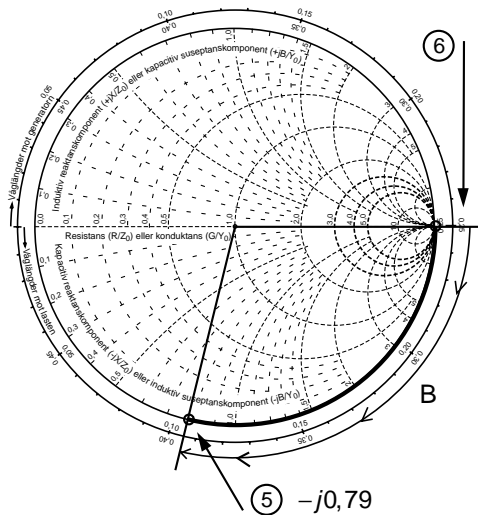
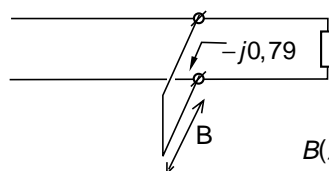


5. Skapa den motsatta susceptansen med stubben.

Stubbens borte ända är kortsluten. Då ligger den i punkt 6 (oändlig ledningsförmåga). Punkt 6 kan du även komma till genom att gå in på normerade impedansen 0 och omvandla till admittans (gå över till andra sidan).

Laständan på stubben är alltså kortsluten, punkt 6. Nu förflyttar vi oss mot generatoren tills vi ser $-j0,79$ i stubben.

Stubbens längd blir 0,144 våglängder, vilket motsvarar 5,76 cm.



$$B(\lambda) = 0,144 \cdot \lambda = 0,144 \cdot \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot \frac{1}{500 \cdot 10^6} = 0,0576 \text{ m}$$

13.10 Öppen stubbe eller kondensator

1. Normera och rita cirkel

Börja på vanligt sätt. Beräkna normerade impedansen, pricka in och rita cirkeln.

2. Övergå till admittans

Nu övergår vi till admittanser. Lastens admittans ligger "på andra sidan" mittpunkten, i punkt 2. Vi avläser inte utan går ut till våglängdsgraderingen.

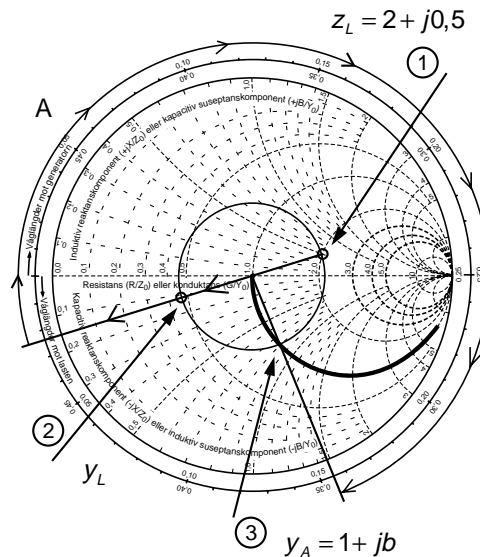
3. Förflytta dig längs ledningen tills admittansens konduktansdel är 1, i nedre halvplanet.

Det är viktigt med nedre halvplanet när vi har öppen stubbe eller kondensator.

Avläs längden A som blir 0,369 våglängder, vilket motsvarar 14,76 cm.

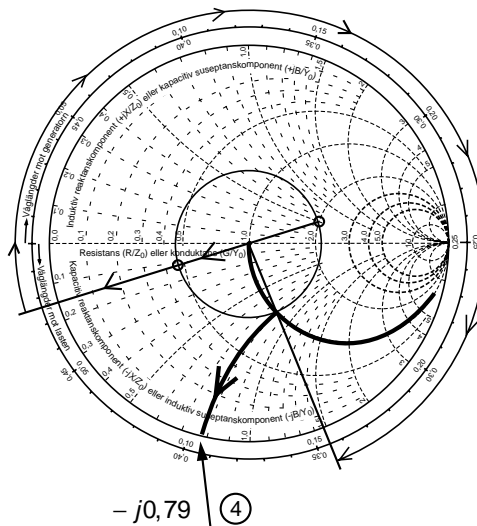
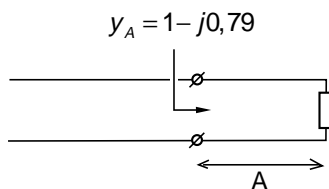
$$A = 0,024 + 0,25 + 0,095 = 0,369 \lambda$$

$$A(\lambda) = 0,369 \cdot \lambda = 0,369 \cdot \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot \frac{1}{500 \cdot 10^6} = 0,1476 \text{ m}$$



4. Admittansen mot lasten

Nu avläser vi normerade admittansen i punkt 3, som blir $y_A = 1 - j0,79$.

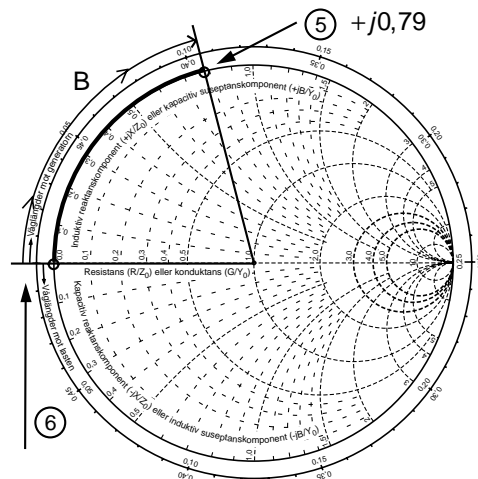
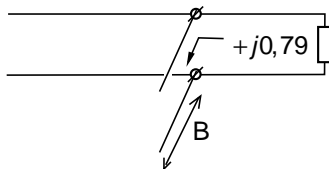


5. Skapa den motsatta susceptansen med stubben.

Stubbens bortre ända är öppen. Då ligger den i punkt 6 (noll ledningsförmåga). Punkt 6 kan du även komma till genom att gå in på normerade impedansen oändligheten och omvandla till admittans (gå över till andra sidan).

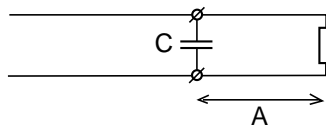
Laständan på stubben är alltså öppen, punkt 6. Nu förflyttar vi oss mot generatoren tills vi ser $+j0,79$ in i stubben.

Stubbens längd blir $0,106$ våglängder, vilket motsvarar $4,24$ cm.



$$B(\lambda) = 0,106 \cdot \lambda =$$

$$= 0,106 \cdot \frac{3 \cdot 10^8}{\sqrt{2,25}} \cdot \frac{1}{500 \cdot 10^6} = 0,0424 \text{ m}$$



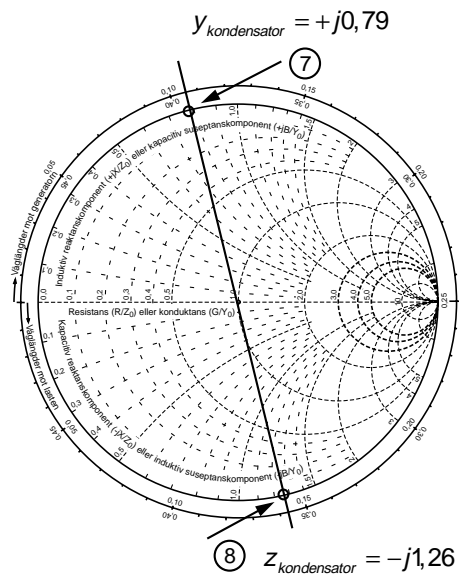
7. Anpassning med kondensator

Nu skall vi välja en kondensator som har normerade admittansen (susceptansen) $+j0,79$ vid denna frekvens. Men att räkna med admittanser känns osäkert. Därför övergår vi till impedans direkt i Smith-diagrammet. Impedansen (reaktansen) har vi på motsatta sidan, där vi avläser normerade reaktansen till $-j1,26$. Nu återstår bara att beräkna kapacitansen.

$$Z_{kond} = z_{kond} \cdot Z_0 = -j1,26 \cdot 50 = -j63 \Omega$$

$$X_C = \frac{1}{2\pi f \cdot C}$$

$$C = \frac{1}{2\pi f \cdot X_C} = \frac{1}{2\pi \cdot 500 \cdot 10^6 \cdot 63} = 5,05 \text{ pF}$$



13.11 Anpassning i verkligheten

Var, när och hur?

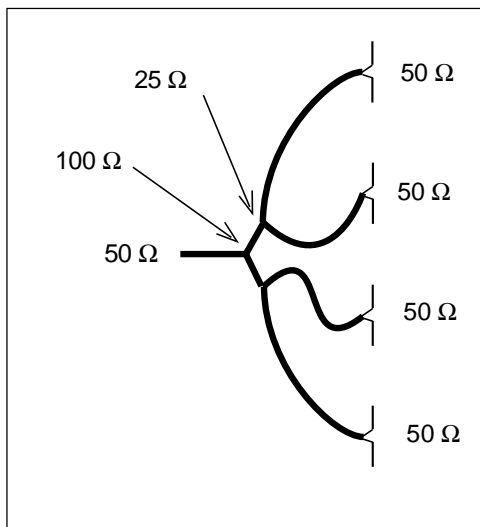
Anpassning av den typ som här beskrivits används i basstationer för mobiltelefoni.

Hopkoppling av antenner

Antennen skall vara anpassad till sändarkabeln. Detta utförs oftast direkt på antenn-elementet. Man ändrar längd på antenn-elementet så att antennens impedans hamnar så nära 50Ω som möjligt på det frekvensområde där den skall användas.

Men de flesta basstationsantenner består av flera antennelement. Man har alltså behov av att koppla ihop flera antennelement som var och ett ser ut som 50Ω . Detta utförs oftast med kvartsvågstransformatorer.

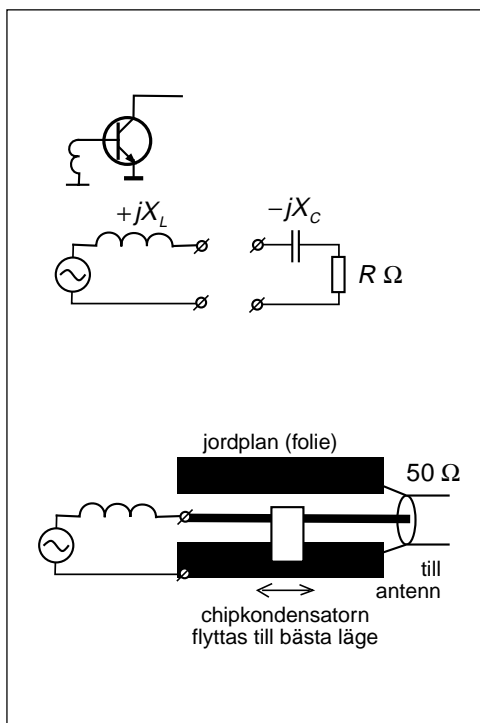
Bilden visar fyra 50Ω -antennor. Först parallellkopplas antennerna två och två. Detta ger 25Ω . Med en kvartsvågstransformator av 50Ω -kabel så transformeras detta till 100Ω , och efter parallellkoppling får vi 50Ω .



Sändartransistorn skall "se" en viss impedans

Sändartransistorn ser ut som en generator som i stort sett saknar inre resistans. Men transistorn kan ha induktans i ledningarna, eller en parallellkapacitans.

När antennkabeln kommer ner till sändaren och sändartransistorn tittar in i kabeln, då vill transistorn se en last som har motsatt reaktans så att reaktanserna tar ut varandra, och en resistans som ger önskad uteffekt.



Denna "impedansanpassning" av sändartransistorn utförs genom att löda fast en chip-kondensator mellan ledaren och jord på kretskortet. Genom att smälta lödtennet och förskjuta kondensatorn fram och tillbaka längs ledningen så finjusterar man främst den reaktiva delen, som kan skilja från transistor till transistor.

14 — Antenner

I vilka riktningar går strålningen från en elektron som strålar?

Resonansfenomenet är viktigt. Maxwell räknade på en förlustfri elektrisk pendel och upptäckte att den inte gungar i oändligheten. Vid varje pendling strålar lite energi ut som elektromagnetiska vågor.

Pendeln var även grundelementet vid Hertz försök med radiovågor, och än i dag är många antenntyper i resonans.

Halvvågsdipolen har visserligen samma längd som den resonanta halvvågsantennen, men halvvågsdipolen har sin längd för att minimera den reflekterade effekten på antennledningen.

Sedan följer polarisation, och varför antenner ser ut som de gör.

Genom att beskriva hur man mäter antennens strålningsdiagram så tar vi oss först igenom riktantenner, för att sedan hamna på antennparametrarna direktiviteten D , antennvinsten G och effektiva mottagningsarean A_e .

Som avslutning följer några ord om skillnader vid sändning respektive mottagning.

14.1 Radiostrålning från en elektron (laddning)

Elektroner som accelereras eller retarderas

Elektromagnetiska vågor, radiovågor, uppstår när elektriskt laddade partiklar, vanligtvis elektroner, ändrar hastighet eller rörelseriktning. En praktisk känsla för vad som händer får man av följande modell:

Stillastående elektron

Runt elektronen finns ett elektriskt fält, fältlinjer (som representerar energin i elektronens laddning). Dessa elektriska fältlinjer byggs upp med ljushastigheten, varefter de bara finns där, som långa rep.

Elektron som rör sig med konstant hastighet

Om elektronen rör sig med konstant hastighet (likström) så flyttas den punkt där fältlinjerna utgår. Fältlinjerna förändras för att stämma med elektronens nya läge. Förändringen sprider sig utåt med ljushastigheten. Detta medför att fältlinjer (repen) som egentligen utgår i sidled tycks luta bakåt. Men fältlinjerna är fortfarande räta. Ingen kan på avstånd avgöra om elektronen står still eller rör sig med konstant hastighet. Lutningen kan åskådliggöras genom att lägga till en liten fältkomponent i sidled.

Elektron som byter rörelseriktning

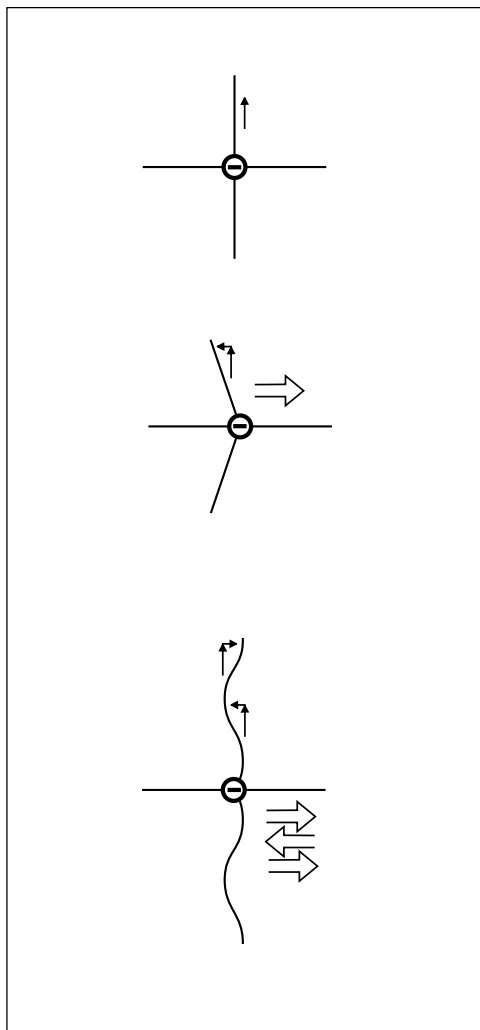
Om elektronen byter rörelseriktning skall fältlinjerna luta åt andra hållet. Fältlinjerna får en krök som rusar utåt med ljushastigheten.

Låt elektronen gunga fram och tillbaka (växelström). Då kommer de sidoriiktade fältlinjerna (repen) att "slingra" sig. Den lilla fältkomponenten "växlar" fram och tillbaka.

Elektriska fältet i radiostrålningen som bildas när elektronen ändrar rörelseriktning bär sig åt på samma sätt som den lilla fältkomponenten som knuffar fältlinjen i sidled.

Radiostrålning

När laddade partiklar ändrar hastighet eller rörelseriktning, d.v.s. accelereras eller retarderas, så bildas förändringar i de elektriska och magnetiska fälten som kan upptäckas

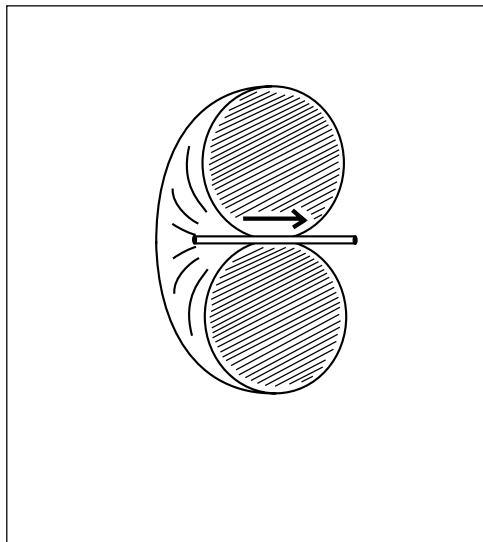


av en mottagarantenn. Sådana förändringar som vi kan ta emot i en antenn kallar vi radiostrålning.

Likström ger inte radiostrålning. Däremot får vi radiostrålning när likströmmen startar eller stoppar. I en dator representeras ettor och nollor av olika spänningsnivåer. Detta ger även olika nivå på strömmen. En dator kan sända ut kraftig störstrålning.

Strålningens riktningsdiagram

- Den sidoriiktade fältkomponenten (som åstadkommer "slingringen") är maximal i riktningar vinkelrätt mot elektronens rörelseriktning. I dessa riktningar har vi störst radiostrålning.
- Längs elektronens rörelseriktning är fältlinjerna alltid räta oavsett hur mycket elektronen gungar (repen slingrar sig inte). Det finns ingen sidoriiktad fältkomponent. I dessa båda riktningar får vi ingen radiostrålning.
- Storleken på den sidoriiktade fältkomponenten varierar med sinus för vinkeln mot elektronens rörelseriktning. Elektriska fältet i radiostrålningen bär sig åt på samma sätt.



Att förhindra strålning

Enda möjligheten att förhindra strålning är att se till att de elektriska och magnetiska fälten riktas på sådant sätt att de inte sammansätts till strålning.

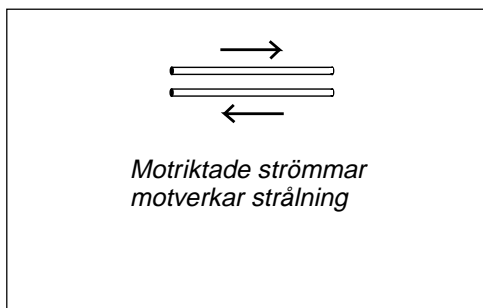
Motriktade strömmar

Ett sätt att undvika radiostrålning (i fel riktning) är att se till att alltid ha motriktade strömmar som är lika starka. Det är detta man gör i ledningen.

En ledning består av två ledare, och även om man transporterar växelström i ledningen, elektroner som skakar fram och tillbaka, så finns alltid i varje tvärsnitt lika många elektroner som skakar i båda riktningarna. Resultatet blir att radiostrålningen från elektronerna i ena ledaren motverkas av radiostrålningen från elektronerna i andra ledaren. Förutsättningen är dock att ledarna ligger tätt ihop.

Naturen vill slippa stråla

En enskild ledare som leder växelström kommer att stråla. Men naturen vill behålla sin



energi. Naturen vill slippa stråla om så är möjligt.

Om det finns metall, en plåt eller liknande, i närheten av ledaren som leder växelström, så kommer naturen att själv inducera den ström i plåten som behövs för att motverka strålningen från ledaren. Det är på detta sätt skärmning fungerar.

14.2 Hertz experiment och resonans

Elektrisk pendel

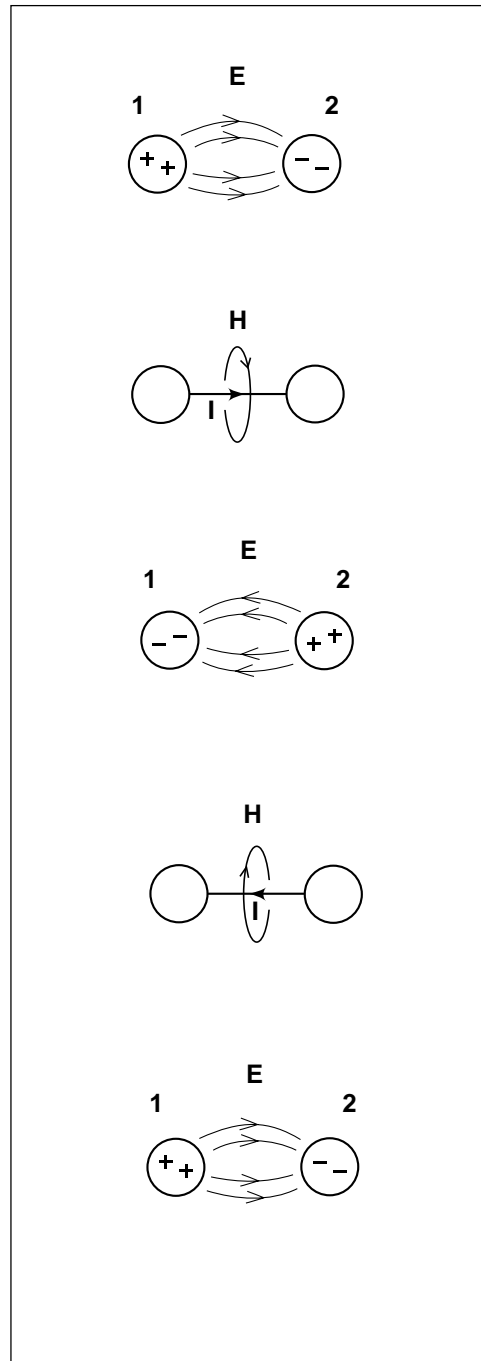
Utgå från två metallklot. Flytta över elektroner från klot 1 till klot 2. Detta kan du göra genom att ansluta ett batteri till kloten, pluspolen till klot 1 och minuspolen till klot 2 och därefter ta bort batteriet. Nu har klot 2 fått överskott på elektroner. Elektronerna vill tillbaka till klot 1. Elektronerna har lägesenergi (lika med energin i det elektriska fältet mellan kloten).

Sammanbind kloten med en metalltråd. Då rusar elektronerna över till klot 1. Men rusningen ger ström. Ström är rörelseenergi (energin i magnetfältet som bildas runt ledaren). Vad händer med rörelseenergin när alla elektroner kommit över?

När strömmen avtar minskar magnetfältet. Då frigörs rörelseenergi som "pressar" över ytterligare elektroner till klot 1. Nu får vi i stället alltför många elektroner på klot 1. Pendeln har svängt till andra läget.

På detta sätt kommer pendeln att svänga fram och tillbaka.

Vi har startat en elektrisk pendel. På ledaren mellan kloten rusar elektroner fram och tillbaka, växelström som ger radiostrålning.



Hertz experiment

Hertz tog två metallklot och ville flytta över elektroner från ena klotet till det andra. För att göra detta behövde han hög likspänning.

Den höga likspänningen fick han genom att ta en "tändspole", ett batteri och en brytare (brytarspets). När ström flyter genom spolen bildas magnetfält. När brytaren öppnas slutar strömmen att flyta, och magnetfältet "dör". Då induceras spänning i trådvarven. (Faraday, induktionslagen). Många trådvarv ger hög spänning.

Spänningen hamnar på kloten, plusspänning på ena och minusspänning på andra klotet. Spänningen hamnar även över "gnistgapet". När spänningen blivit så hög att det bildas en gnista över gnistgapet, då blir det elektrisk kontakt mellan kloten. Pendeln startar.

Elektrisk gnista är joniserad luft. Luftmolekylerna bryts sönder till joner och fria elektroner. De fria elektronerna i gnistan rör sig på samma sätt som elektronerna i metall. Gnistan är inget annat än elektriskt ledande luft.

Nu kan elektronerna rusa fram och tillbaka (växelström), och det bildas radiostrålning.

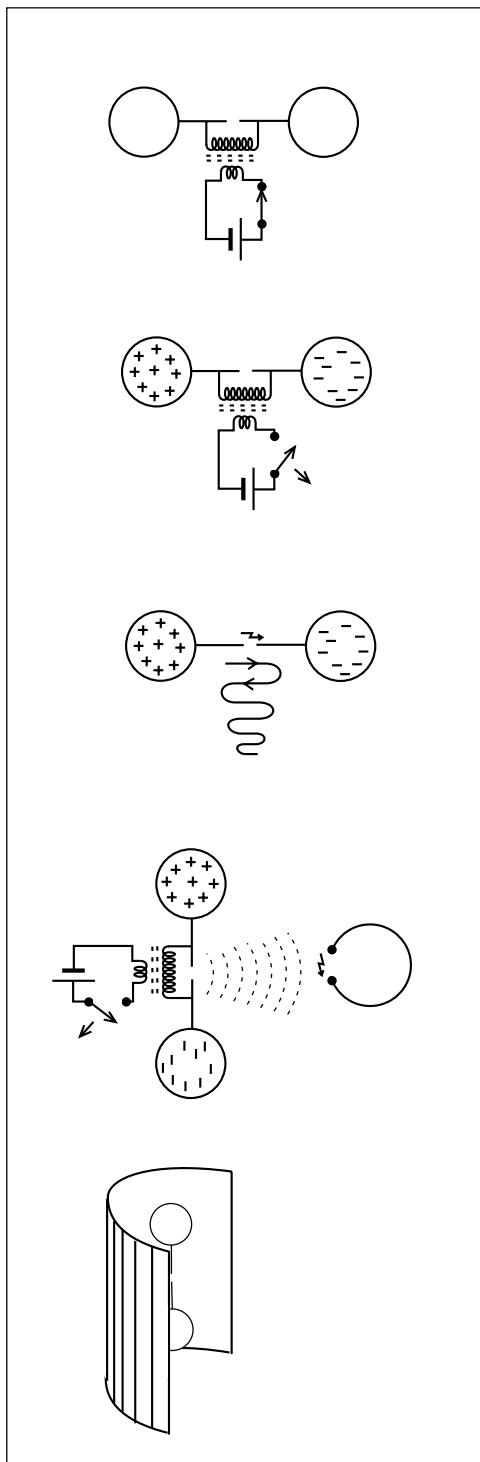
Hertz radiomottagare

Hertz mottagare var ytterligare en pendel som var hopvikt. När det kommer radiovågor och pendeln börjar pendla, då bildas ömsom ström i ledaren, ömsom spänning mellan kulorna. Hertz tittade efter gnistor i gapet mellan kulorna (det var mörkt i rummet).

Hertz parabol

Hertz upprepade experiment som var vanliga inom optiken. Maxwell hade påstått att radiovågorna skulle bära sig åt som ljuset.

- Hertz riktade radiovågorna med hjälp av en parabol och fick längre räckvidd.
- Hertz visade att radiovågorna reflekteras mot en metallplåt på samma sätt som ljus i en spegel.
- Hertz visade att radiovågorna böjs i ett prisma av paraffin, på samma sätt som ljus i ett glasprisma.



Hertz hade ingen radiosändare

Hertz flyttade över elektroner tills det blev överlag i gnistgapet. När det blev överlag började laddningar rusa fram och tillbaka med ljushastigheten mellan antennens ändrar.

När laddningarna rusade åt ena hållet gungar elektronerna åt ett håll. När laddningarna rusade tillbaka gungar elektronerna åt andra hållet. Elektroner som gungar fram och tillbaka är växelström.

Hertz hade ingen radiosändare. Växelströmmen som han behövde för att få radiostrålning, den bildades direkt på antennen.

Vilken frekvens fick Hertz?

När laddningarna rusat fram och tillbaka har växelströmmen genomlöp en period. Den tid det tar för laddningarna att rusa fram och tillbaka med ljushastigheten blev växelströmmens periodtid T . Frekvensen är antalet perioder per sekund.

Våglängd

Den sträcka man hinner med ljushastigheten c under periodtiden T har ett eget namn: våglängd. Var Hertz antenn en halv våglängd lång?

Halvvågsantenn

Hertz antenn var i halvvågsresonans. Den bar sig åt som en elektrisk pendel. Men metallkulorna medförde att antennen var förkortad, se nästa avsnitt.

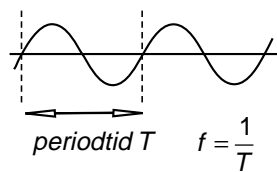
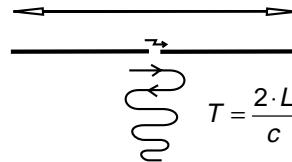
Dagens pendel sitter i radiosändaren

I dag låter vi inte växelströmmen bildas på antennen. Växelströmmen bildas i radiosändaren och matas ut på antennen. Men fortfarande alstras växelströmmen i en elektrisk pendel, antingen en resonanskrets med spole och kondensator, eller i en "hopvikt halvvågsantenn" som inte strålar (kvartsvågsresonator), eller i en mekanisk pendel (styrkristall, kvartskristall).

Kraftig radiostrålning — många elektroner

Kraftig radiostrålning innebär att sätta fart på många elektroner. Det har visat sig vara lättast för radiosändaren att sätta fart på elektronerna om elektronerna själva vill

avstånd mellan antennens ändrar L



$$L = \frac{c \cdot T}{2} = \frac{1}{2} \cdot \frac{c_0}{f \sqrt{\epsilon_r}} = \frac{1}{2} \cdot \frac{\lambda_0}{\sqrt{\epsilon_r}}$$

c är ljushastigheten i mediet runt antennen, som oftast är luft ($\epsilon_r \approx 1$). Men om antennen är plastisolerad så går viss del av fältet i plasten. Då är ljushastigheten aningen lägre än c_0 . Det finns tillfällen när antenner varit utsatta för så stark isbildning, flera cm tjock is, att antennen avsevärt ändrat sin resonansfrekvens.

gunga i den takt som sändaren "knuffar". Det är detta man uppnår på halvvågsantennen.

Två önskemål: "Många elektroner" — "strålning i rätt riktning"

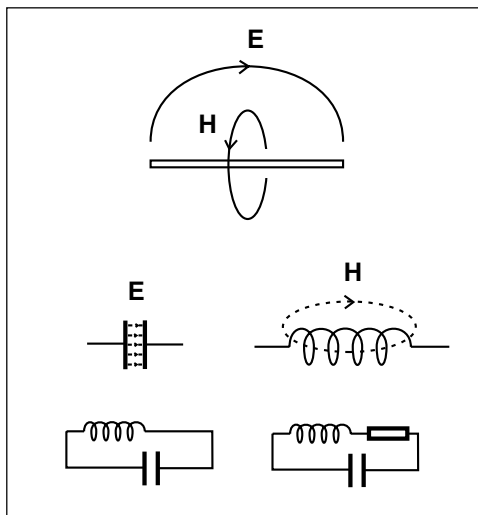
Halvvågsantennen är den bästa antennen när det gäller att underlätta för sändaren att sätta fart på så många elektroner som möjligt, d.v.s. att på enklast möjliga sätt åstadkomma radiostrålning.

Minst lika viktigt är att radiostrålningen går i rätt riktning. Men riktningsdiagrammet kan vi påverka, inte genom att förändra halvvågsantennen, utan genom att "störa" halvvågsantennen med metall i närheten, eller genom att kombinera flera halvvågsantenn till en riktantenn.

Som en resonanskrets

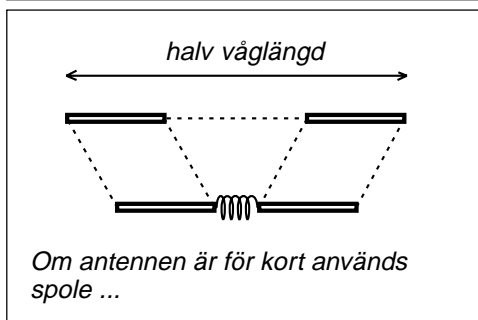
Energien på halv vågsantennen växlar mellan elektriskt fält och magnetfält. Elektriskt fält finns i en kondensator, magnetfält i en spole. Om man kopplar ihop kondensatorn och spolen så får man en resonanskrets där energin pendlar på samma sätt som på halv vågsantennen.

Fast halv vågsantennen strålar. Komplettera resonanskretsen med ett motstånd (som vi kallar strålningsresistansen). Den effekt som blir till värme i motståndet motsvarar den effekt som strålar ut från antennen som radiovågor. Detta är halv vågsantennens ekvivalenta schema, en krets som ser ut på samma sätt som halv vågsantennen. OBS att komponentvärdena ändras om vi byter frekvens.

**Om antennen är för kort**

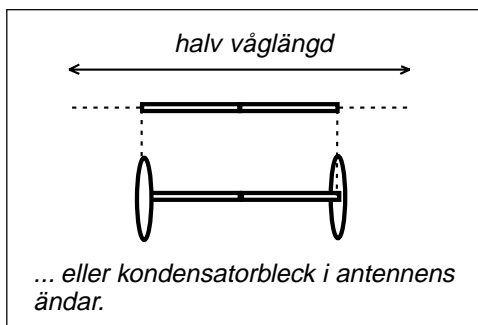
Om antennen mekaniskt är kortare än en halv våglängd, då kan man lägga till en liten del av det ekvivalenta schemat för att få den önskade resonansfrekvensen.

- Man kan komplettera med en del av spolen. För att göra nytta skall spolen placeras där antennen har sin magnetiska energi, i mitten.



- Eller man kan öka antennens elektriska energi genom att placera kondensatorbleck i antennens ändrar.

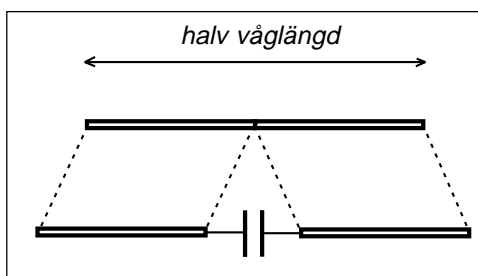
Hertz antenn med metallkuler i ändarna var förkortad på detta sätt.

**Om antennen är för lång**

Om antennen är för lång sätter man en spole mitt på antennen för att antennen skall se längre ut.

Vad gör man om antennen är för lång? Då sätter man en "negativ spole" mitt på antennen. Du som räknat med $j\omega$ -metoden vet att det enda som skiljer kondensatorn från en spole är ett minustecken. Kondensatorn är en negativ spole.

"Glasantennens" metallbleck som limmas på bilens glasruta är en sådan kondensator.



14.3 Dipolantenner

Hur matas halvvågsantennen?

Ett sätt kan vara med en liten transformator mitt på antennen. Strömknuffar från sändaren knuffar på antennens laddningar. Som att knuffa på gungan i lekparken.

Som två laddningar

Det som händer på halvvågsantennen kan liknas vid två laddningar, en positiv och en negativ, som släpps på antennens ändrar. Laddningarna kommer att rusa fram och tillbaka med ljushastigheten.

Först befinner sig laddningarna i ändarna. Då har vi spänning mellan antennens ändrar. Potentialen blir noll mitt på antennen.

När laddningarna kommit till halvvågsantennens mitt tar laddningarna ut varandra. Vi har ingen spänning, nollpotential. Däremot har vi ström. En positiv laddning åt ena hållet och en negativ åt andra hållet representerar båda ström åt samma håll. Max ström mitt på antennen.

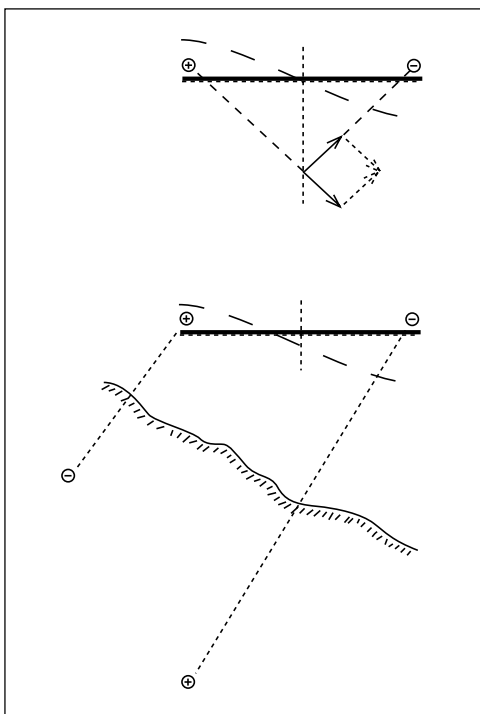
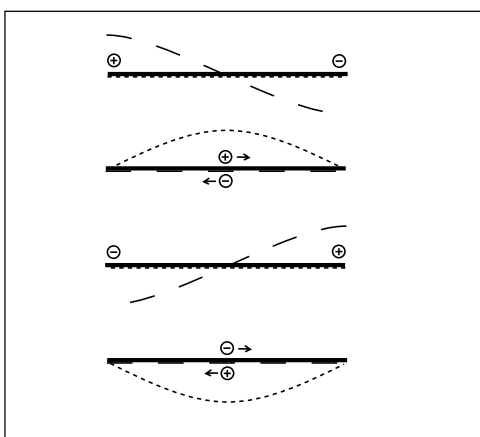
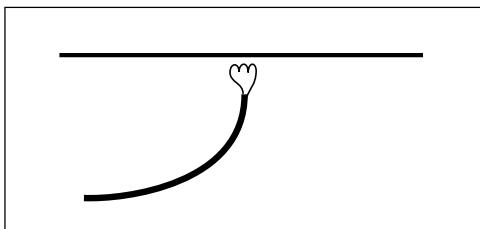
På detta sätt växlar det. Spänning mellan ändarna, maxström åt ena hållet, omkastad spänning mellan ändarna, maxström åt andra hållet.

Nollpotential i symmetriplanet

I ett plan mitt igenom halvvågsantennens mittpunkt kommer det bara att finnas elektriskt fält som är riktat vinkelrätt mot planet. Detta innebär att det inte finns något som vill sätta fart på elektroner om vi lägger metall i symmetriplanet. Här skall antennens matarledning dras för att vi inte skall få "antennström" (ström som strålar) på antennledningen. Bilden överst på sidan är alltså inte det optimala matningssättet.

Om antennen sitter snett över ett ledande plan så kommer laddningar i planet att attraheras av laddningarna på antennen. Då får vi inte längre nollpotential i antennens symmetriplan.

Matematisk analys av sådana här problem är komplicerade eftersom vi arbetar i tre dimensioner. En fysikaliskt korrekt bild med laddningar i kombination med praktiska försök och mätningar leder snabbast till ett ingenjörsmässigt resultat.



Dipolantennen

Om vi ansluter en generator till en öppen ledning så kommer generatören att mata ut laddningar på ledningen, ömsom positiva, ömsom negativa, laddningar som reflekteras på ledningen. Vi ser att den yttersta halva våglängden av ledningen ser ut på exakt samma sätt som halv vågsantennen. Men vi ser även att i varje tvärsnitt finns laddningar av olika polaritet som rör sig åt samma håll. Radiostrålningarna från ledarna tar ut varandra.

Två poler — dipol

Vik isär ledningens ledare. Då får vi två ledningsstumpar där det flyter laddningar utan att det finns motriktade laddningar. Dessutom kommer de båda benen att se identiska ut. Strömmen rör sig åt samma håll.

Inte bara att vi får strålning från varje ledare för sig, strålningen från de båda ledarna samverkar.

En antenn där man viker isär antennledningens ledare kallas dipolantenn. Bilden visar en dipolantenn som är en hel våglängd lång (helvågsdipol). Detta är (nästan) det optimala ut radiostrålningssynpunkt (optimum uppträder när varje antennben är $0,64\lambda$).

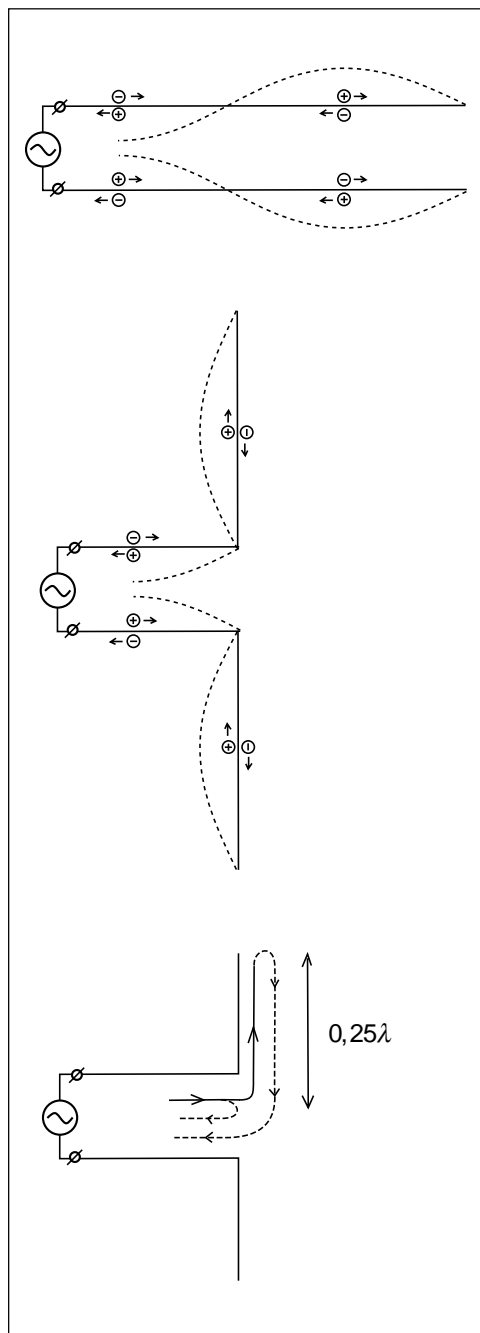
Helvågssantennen svår att mata

Helvågssantennen har visserligen goda strålningsegenskaper, men det är svårt för sändaren att få fart på elektronerna på antennbenen. Därför är det ytterst sällan man ser halv vågsantenner.

Halv vågsdipolen

Vad händer när signalen kommer fram till antennbenen? Ledningen övergår till något annat. En del av signalen reflekteras.

Resten av signalen fortsätter ut på antennbenen. När signalen kommer till antennbenens ändrar så reflekteras resten av signalen och går tillbaka, går tillbaka in på ledningen. På ledningen har vi alltså två reflekterade signaler, en signal från antennens matningspunkt och en annan från antennens ändrar. Om antennbenen görs $0,25\lambda$ långa så kommer de båda reflexerna att vara fasförskjutna $0,5\lambda$ och släcka ut varandra. Vi får ingen reflekterad effekt. Enda vägen för effekten är att bli radiostrålning (eller värme på grund av Ohmskt motstånd



i antennbenen). Detta är den vanligaste antenntypen, inte för att det är den bästa strålaren, utan för att den är lättast för sändaren att mata.

14.4 Polarisation

Horisontell polarisation

Man säger att radiovågen är horisontal-polariserad om elektriska fältstyrkan i radiovågen ligger horisontellt. Horisontell polarisation får man när elektronerna skakar i horisontell riktning. TV-sändningarna (marknätet) är normalt horisontal-polariserade.

Vertikal polarisation

Vertikal polarisation får man när elektronerna skakar i vertikal riktning.

Det är lättare att åstadkomma en rundstrålande antenn om den är vertikal-polariserad. Detta är orsaken till att man använder vertikal polarisation för kommunikationsradio och mobiltelefoni. Det vertikala antensprötet på bilen ger vertikal polarisation.

Helixantennen

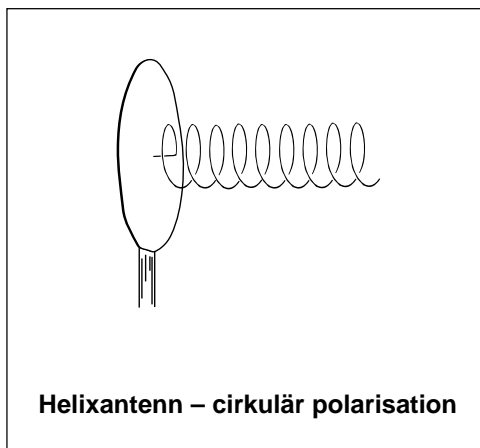
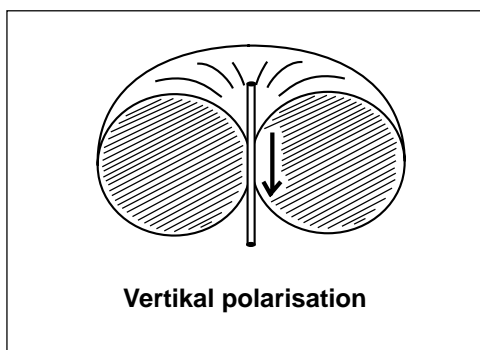
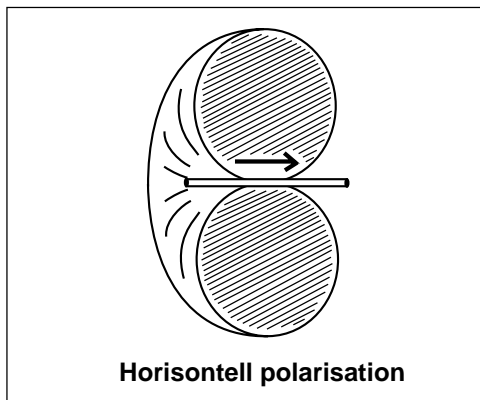
Helixantennen är en långtrådsantenn, tillverkad av en tråd som är många våglängder lång, där tråden lindats till en spole med omkretsen 1 våglängd. När elektronerna på ena sidan spolen är på väg uppåt kommer de att vara på väg uppåt även på andra sidan av spolen, en halv våglängd längre bort. På detta sätt samverkar alla elektroner, i spolens riktning.

Cirkulär polarisation

Strålningen från en helixantenn har cirkulär polarisation, som en dipolantenn som snurrar ett varv på periodtiden T . Den cirkulära polarisationen kan vara vänstervriden, eller högervriden, beroende på åt vilket håll spolen lindats.

Det intressanta med cirkulär polarisation är att när radiovågen reflekteras, i t.ex. marken eller i en husvägg, då ändras polarisationen så att den snurrar åt andra hållet. Vänstervriden blir högervriden efter reflex. Och högervriden kan inte tas emot av vänstervriden antenn. På detta sätt undviker man reflexerna.

Denna antenn kan ses på sportarenor. Fotografen bär på TV-kameran och efter springer teknikern med helixantennen på en pinne och försöker rikta antennen mot någon mottagarantenn.

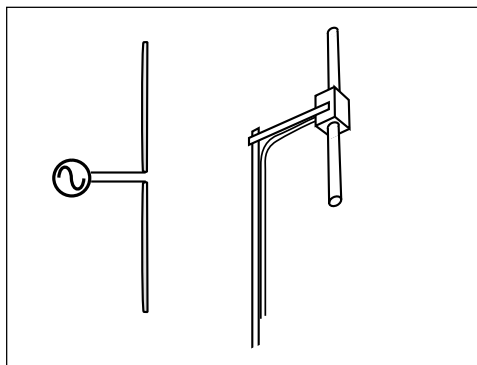


14.5 Olika dipolantenner

Halvvågsdipol (dipol)

Den absolut vanligaste antennen är en halv våglängd lång och matas i mitten. Det är så sändaren vill ha det. Radiomottagaren är inte lika kräsen utan klarar sig oftast hyggligt även om antennen inte är i halvvågsresonans.

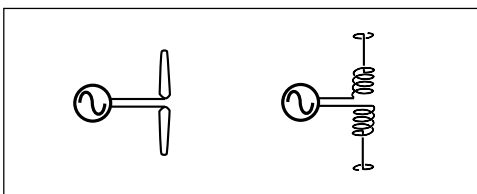
Bilden visar en vertikalpolariserad basstationsantenn, t.ex. monterad på ett vattentorn.



Förkortad dipol

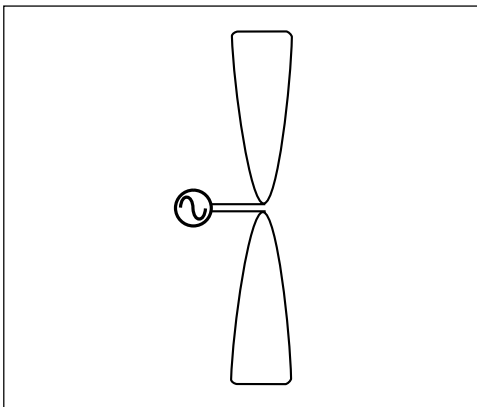
Om vi vill att antennen mekaniskt skall vara kortare än en halv våglängd, om längre antenn får inte plats, då ersätter vi den saknade längden med mittspole eller kondensatorbleck i ändarna.

Bilden visar en förkortad dipol. För att slippa ha alltför många varv i spolen har den även försetts med ändkapacitans. Detta gör antennen mer bredbandig än om enbart spole skulle använts.



Tjocka antennben ger bredbandig antenn

Vanligaste sättet att åstadkomma en antenn som sändaren klarar att mata över ett bredare frekvensområde, är att göra antennbenen tjocka. Tekniken används från pytt små mikrovågsantenner till jättestora kortvågsantenner.

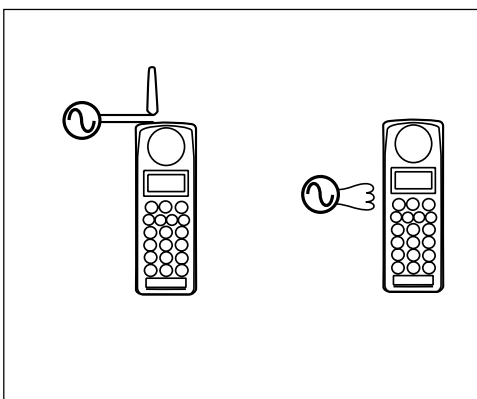


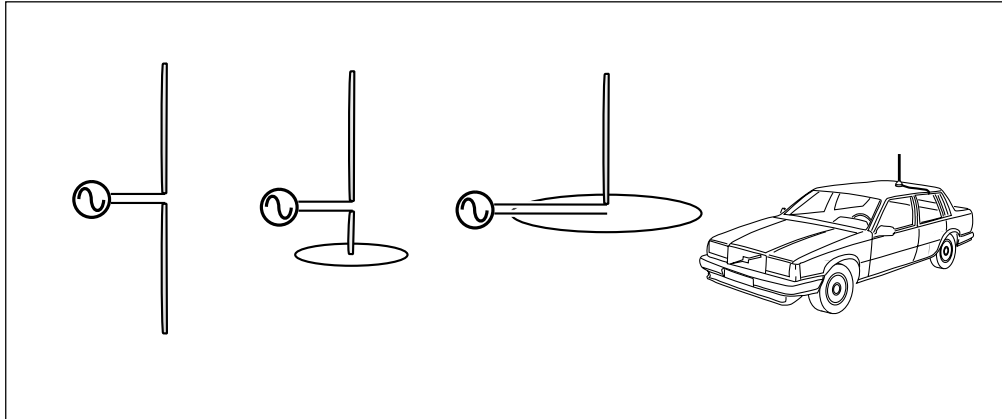
Ficktelefonen: Ett förkortat antennben och ett bredbandigt antennben

Ficktelefonen är exempel på en antenn där ena antennbenet är hälften av en förkortad halvvågsdipol, medan själva ficktelefonen utgör ett tjockt antennben, hälften av en bredbandig halvvågsdipol.

Det går alldeles utmärkt att kombinera två olika antennben på detta sätt. Man kan enkelt visa att det nästan enbart är det smalbandiga antennbenet som bestämmer antennens resonansfrekvens och bandbredd. Och tur är detta eftersom vi både håller i och lägger örat emot det bredbandiga antennbenet.

Vid "inbyggd" antenn är hela ficktelefonen en halvvågsantenn som matas via en "kopplingsanordning".



**Kvartsvågsantenn på jordplan**

Om ena antennbenet förkortas med kapacitansbleck och man gör kondensatorn så stor att den ersätter hela antennbenet, då får vi kvartsvågsantennen på jordplan. Bilplåten fungerar som ett stort kondensatorbleck som kraftigt förkortar ena antennsprötet.

14.6 Mätning av strålningsdiagram

Den antenn vars strålningsdiagram skall bestämmas (mätobjektet) placeras i en vridbar mast. På visst avstånd från masten finns en mottagarantenn och en mottagare, kalibrerad i mottagen signalstyrka. Mätobjektet roteras och signalstyrkans variation relativt signalstyrkan i maxriktningen, oftast omräknad i dB, ritas in i ett polärt diagram.

Men strålningsdiagrammet är tredimensionellt. Mätobjektet vrids 90° och man gör ännu en mätning. Två sådana projektioner ger oftast en tillräckligt god beskrivning av mätobjektets strålningsegenskaper.

E-plansdiagram

Vid mätningen roteras mätobjektet i horisontalplanet. Om mätobjektet placerats så att det sänder med horisontell polarisation, om elektriska fältet (E-vektorn) ligger i mätplanet, så kallas diagrammet E-plansdiagram.

H-plansdiagram

När mätobjektet vridits 90° så kommer det att sända med vertikal polarisation. Nu ligger magnetfältet (H-vektorn) i mätplanet. Därav namnet H-plansdiagram.

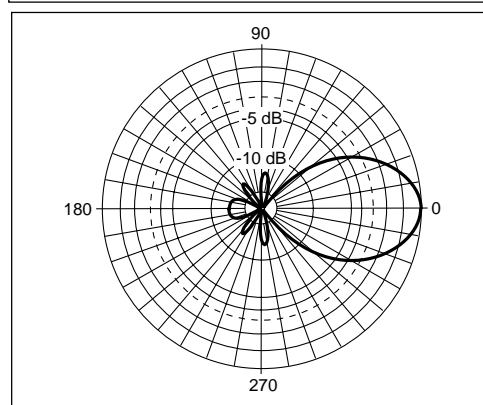
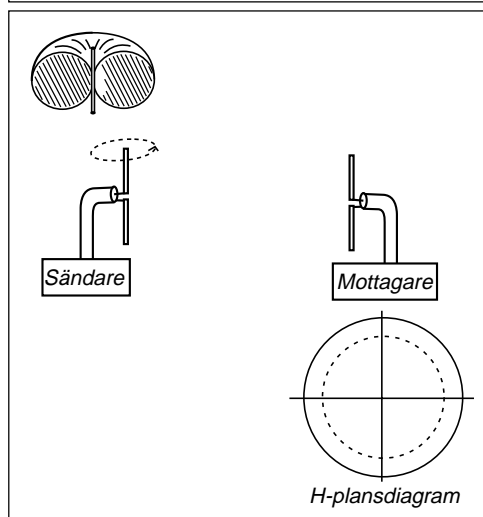
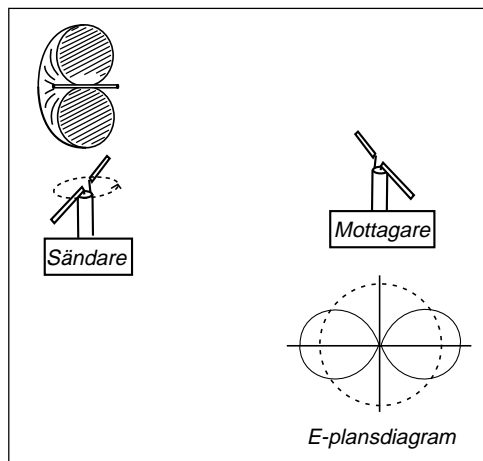
Lober

Den horisontella halvvågsdipolen har ett åtta-format strålningsdiagram. Halvvågsdipolen har två "lobber".

När halvvågsdipolen placeras vertikalt blir strålningsdiagrammet helt runt (i verkligheten får man viss inverkan av antennmasten). Här finns inga lobber, utan vi säger att antennen är rundstrålande.

Öppningsvinkel

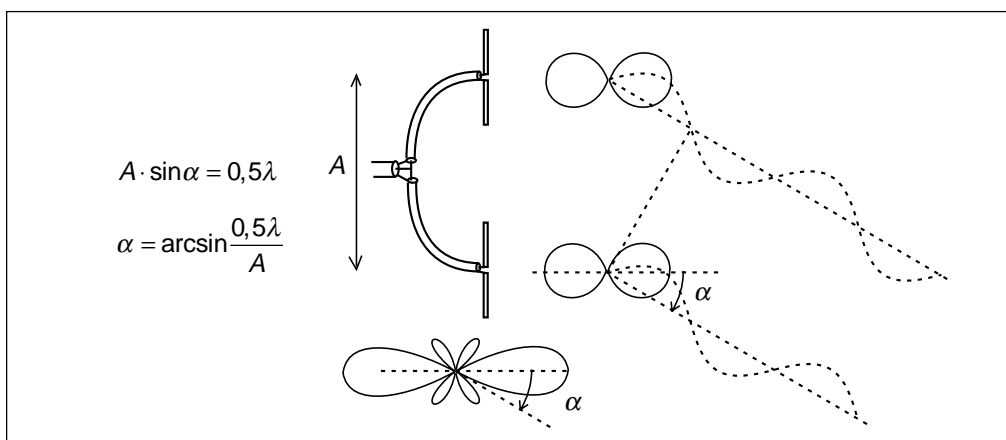
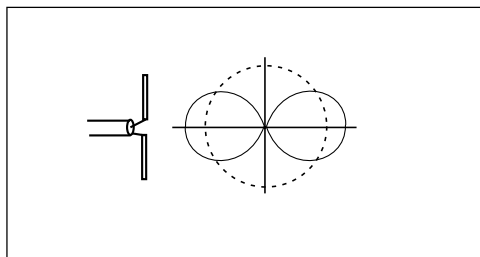
För riktantenner talar man om huvudlob och sidolober. Ofta vill man att huvudloben skall vara så smal som möjligt, medan sidoloberna skall vara svaga. Bilden visar något som skulle kunna vara en Yagi-antenn, en typisk tv-mottagningsantenn. I diagrammet ser vi att sidolober/backlobber är undertryckta -13 dB och att huvudlobens "öppningsvinkel" (vinkeln mellan -3 dB-punkterna) är 56° .



14.7 Riktantenner

Halvvågsdipolen har sitt åtta-formade strålningsdiagram. Men om vi vill ha någon annan form på diagrammet? Om vi vill ha kraftigare strålning i någon viss riktning?

Vi kan inte åstadkomma kraftigare strålning i någon viss riktning, vi kan bara hindra strålning från att gå i icke önskade riktningar och hoppas att strålningen i stället går i den riktning vi önskar.

**Stackade dipoler (kollinjär antenn)**

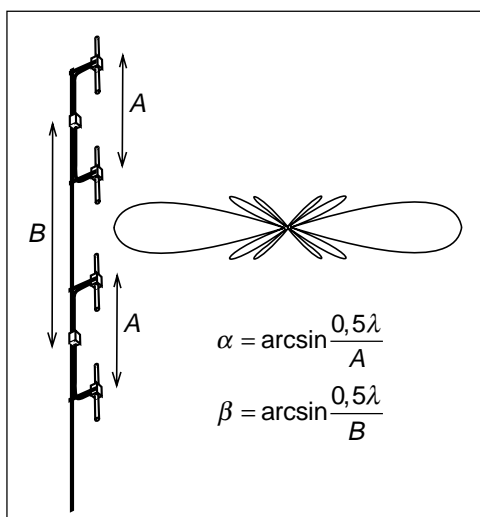
Placera två dipoler med sina strålningscentra på avståndet A . Då kommer signalerna från dessa båda antenner att hamna i de riktningar där ena signalen gått $0,5\lambda$ längre väg. I dessa riktningar går ingen signal. Den signal som skulle gått i dessa riktningar hamnar i stället i huvudloben (i detta fall).

Fyra dipoler

Med fyra dipoler får vi dels vinklar för varje grupp om två antenner, avstånd A , dels vinklar för avståndet mellan grupperna, avstånd B .

På detta sätt kan man öka antalet antenner i vertikal-led och få alltmer av effekten att gå i horisontal-led. Men loben blir mer och mer hoptryckt.

Förhållandena blir exakt desamma om dipolerna vrids till horisontell polarisation. UHF-sändarantennerna för tv består av 64



dipoler över varandra. Huvudloben är så hoptryckt att vinkeln mellan första nollställena är $3,5^\circ$. Man talar om "pannkakslob".

De stackade dipolerna trycker ihop strålningsloben, men antennen (antennsystemet) är fortfarande rundstrålande i horisontalplanet.

Verkliga riktantenner riktar strålningen i en riktning, eller i en sektor, i horisontalplanet.

Enkel riktantenn

Enklaste typen av riktantenn får man om två halvvågsdipoler placeras $0,25\lambda$ från varandra och matas med kablar där ena kabeln är $0,25\lambda$ längre.

- Åt höger tar den extra gångvägen i luften och den längre kabeln ut varandra. Signalerna samverkar.
- Åt vänster har ena signalen gått först $0,25\lambda$ längre väg i kabeln och sedan ytterligare $0,25\lambda$ längre väg i luften, totalt $0,5\lambda$. I denna riktning hamnar signalerna i motfas och tar ut varandra.

Tekniken bakom riktantenner är att utnyttja flera strålningskällor. I vissa riktningar samverkar signalerna, i andra riktningar motverkar signalerna varandra. För att åstadkomma detta krävs utrymme. Signalerna måste gå olika vägskillnad i olika riktningar. En riktantenn med viss antennvinst får normalt alltid en viss minsta storlek (i våglängder).

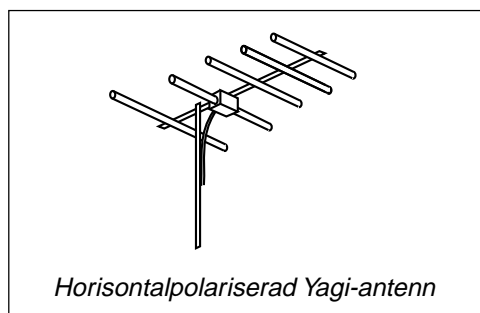
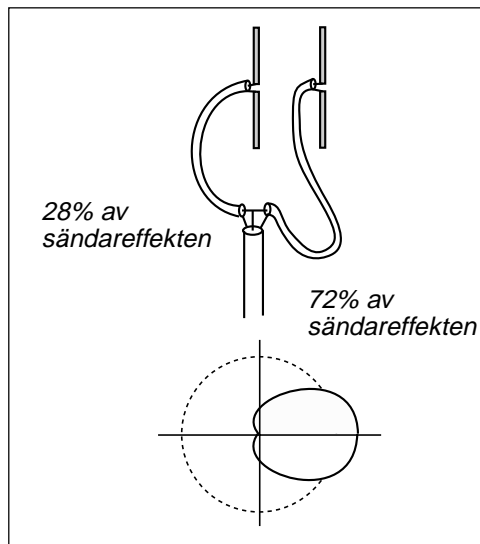
Yagi-antennen

Antennen ovan är enkel i teorin men svår att realisera. För att lika stark signal skall stråla ut från båda dipolerna måste det gå lika stor ström i de båda dipolerna.

Men dipolerna får inte bara signal via ledningarna. De är även mottagarantenner i förhållande till varandra (antennerna "kopplar" till varandra). I detta fall blir det så att på ena dipolen samverkar signalen från kabeln och det som kommer genom luften. På andra dipolen motverkar dessa båda signaler varandra. För att strömmarna skall bli lika stora måste sändareffekten delas olika.

Om ena dipolen får så stark signal genom luften att den inte behöver någon signal från sändaren? Då kan vi ta bort den kabeln.

Detta är Yagi-antennen. Man kombinerar flera halvvågsantenner, men bara en av



antennerna är ansluten med antennkabel till sändaren (mottagaren). De övriga är "anslutna" via strålning (koppling) genom luften, till det anslutna antennelementet.

Reflektorantenner

Dipolen kan placeras framför en plåt. Den signal som går mot plåten kommer att alstra ström i plåten. Denna ström ger radiostrålning. Vi får alltså strålning både från halvvågsdipolen och från plåten.

När avståndet mellan dipol och plåt är $0,25\lambda$ samverkar de båda signalerna i riktningen rakt ut från plåten genom dipolen.

Parabolantenner

När reflektorn blir mycket stor i våglängder kommer radiovågorna att bära sig åt som ljuset. Radiovågornas strålgång mot parabolens och reflexion till fokus följer optikens lagar.

14.8 Direktivitet D **Isotrop antenn**

Den isotropa antennen strålar lika kraftigt i alla riktningar i rymden (som en ballong). Både E-plans- och H-plansdiagrammet är helt cirkulärt. En sådan antenn kan inte konstrueras eftersom det skulle innebära att vi måste få elektronerna att röra sig i alla riktningar samtidigt. Med den isotropa antennen används som en tänkt referensantenn.

Signalen — effekttäthet

Radiosignalen består av elektrisk fältstyrka E V/m och magnetisk fältstyrka H A/m. Multiplicera dessa och vi får effekttäthet S W/m², effekt per ytenhet.

Lägg en sfär (ett klot) runt antennen och summera effekttätheten på sfärens inneryta, så får vi totalt utstrålad effekt. Om antennen har 100 % verkningsgrad så är detta lika med den effekt som sändaren matar in i antennen.

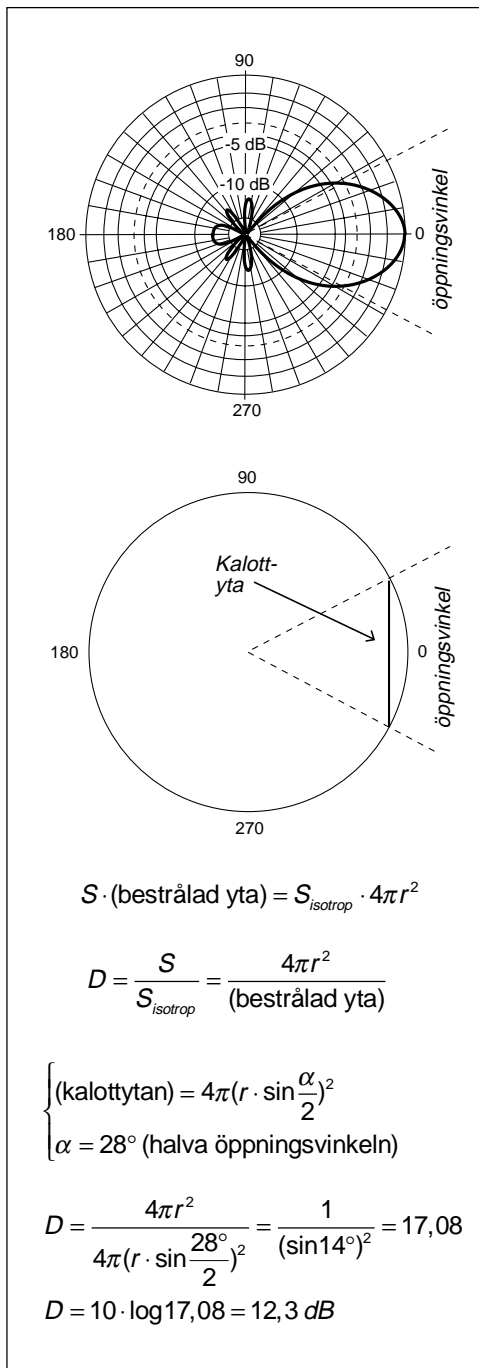
Direktivitet

Direktiviteten visar hur mycket kraftigare effekttätheten är i antennens huvudriktning, jämfört med medeleffekttätheten, d.v.s. den effekttäthet vi skulle fått från en isotrop antenn. Direktiviteten beräknar vi ur antennens strålningsdiagram.

Approximativt kan man säga att all effekt går i en kon som är lika med antennens öppningsvinkel, d.v.s. vinkeln mellan -3 dB-punkterna. Låt oss titta på det tidigare strålningsdiagrammet som hade 56° öppningsvinkel.

Effekttätheten från antennen multiplicerat med den yta som bestrålas är totalt utstrålad effekt. På samma sätt är effekttätheten från en isotrop antenn multiplicerat med sfärens hela yta lika med totalt utstrålad effekt. Ur detta kan vi beräkna antennens förmåga att rikta strålningen, antennens direktivitet.

Den bestrålade ytan blir en kalottyta. Med hjälp av formelsamling får vi att effekttätheten är 17 ggr högre, vilket blir 12,3 dB.



Basstationsantenn för mobiltelefoni

En basstationsantenn för mobiltelefoni, 900 MHz eller 1800 MHz, består oftast av ett antal strålande element i vertikalled så att strålningsdiagrammet trycks ihop. Vi tänker oss en sådan antenn med en vertikal öppningsvinkel på 10° . Ofta sänks loben (elektrisk tiltning) för att träffa marken. Vad blir antennens direktivitet om antennen är rundstrålande?

Vi beräknar ytan på sfären som bestrålas. Det blir ett band runt sfären (sfärens omkrets) som har höjden $2 \cdot r \cdot \sin 5^\circ$.

Sektorantenn

Hur hög blir direktiviteten om antennen bara strålar i en 90° bred sektor i horisontalplanet?

Från basstationen täcker man ofta tre celler, tre 120-graderssektorer, med var sin radiostation. Eftersom man bara kan ha ett visst antal samtidigt uppkopplade samtal i varje cell så måste man ha många celler om man har många abonnenter. Att täcka tre celler från samma mast är ett sätt att spara på antalet basstationsplatser.

En antenn som har en öppningsvinkel på 90° ger ofta optimal täckning av en 120-graderssektor.

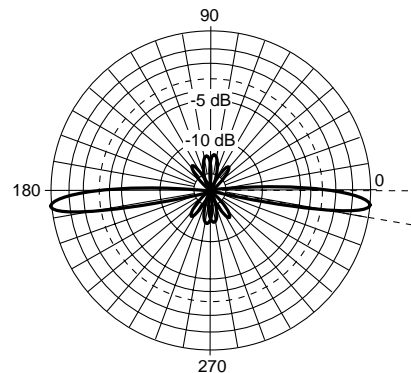
Den yta som belyses av denna antenn blir som i förra exemplet, med skillnaden att det bestrålade "bandet" bara sträcker sig ett kvarts varv runt sfären. Ytan blir en fjärdedel. Direktiviteten blir 4 ggr högre, vilket motsvarar 6 dB högre direktivitet.

Parabolantenn med 2 graders öppningsvinkel

På motsvarande sätt kan man beräkna direktiviteten för en parabolantenn för t.ex. satellit-tv, som har en öppningsvinkel på 2° . Här kan formeln med kalottytan användas:

$$D = \frac{4\pi r^2}{4\pi(r \cdot \sin \frac{1^\circ}{2})^2} = \frac{1}{(\sin 0,5^\circ)^2} = 13\,132$$

$$D = 10 \cdot \log 13\,132 = 41,2 \text{ dB}$$

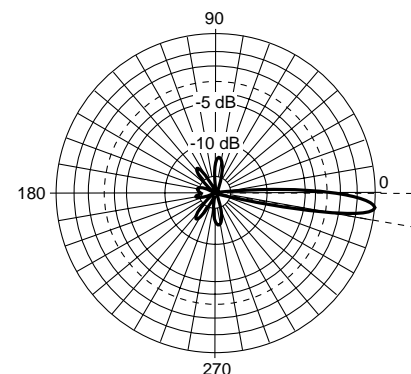


$$D = \frac{S}{S_{\text{isotrop}}} = \frac{4\pi r^2}{(\text{bestrålad yta})}$$

$$(\text{Bestrålad yta}) \approx (2\pi r) \cdot (2 \cdot r \cdot \sin 5^\circ)$$

$$D = \frac{4\pi r^2}{(2\pi r) \cdot (2 \cdot r \cdot \sin 5^\circ)} = \frac{1}{(\sin 5^\circ)} = 11,47$$

$$D = 10 \cdot \log 11,47 = 10,6 \text{ dB}$$



$$(\text{Bestrålad yta}) \approx \frac{90^\circ}{360^\circ} (2\pi r) \cdot (2 \cdot r \cdot \sin 5^\circ)$$

$$D = \frac{4\pi r^2}{\frac{90^\circ}{360^\circ} 2\pi r \cdot (2 \cdot r \cdot \sin 5^\circ)} = \frac{4}{\sin 5^\circ} = 45,9$$

$$D = 10 \cdot \log 45,9 = 16,6 \text{ dB}$$

14.9 Antennvinst (gain) G

Antennens direktivitet är en parameter som bara tar hänsyn till den effekt som strålar ut från antennen. Om antennen har 100 % verkningsgrad så är detta samma som den effekt som sändaren matar in i antennen. På något sätt måste vi alltså även ta hänsyn till antennens verkningsgrad. Detta gör vi genom att ange antennens antennvinst (gain).

$$G = \eta_L \cdot D$$

η_L = resistiva förluster (loss)

Mätning av antennvinst

På en antennmätplats, se "mätning av strålningsdiagram", monteras mätobjektet, vars huvudlob riktas mot mottagarantennen. Effekten från sändaren matas in i mätobjektet och mottagen signalstyrka antecknas.

Därefter byter man ut mätobjektet mot en kalibrerad referensantenn som kan vara en halv vågs dipol. Man kontrollerar att den effekt som sändaren matar in i referensantennen inte har förändrats, och avläser mottagen signalstyrka.

Förhållandet mellan signalstyrkorna, uttryckt i dB, är mätobjektets antennvinst relativt referensantennen. Till detta värde adderas referensantennens antennvinst relativt den förlustfria isotropa antennen.

Man gör tvärt om

Direktivitet, antennvinst, strålningsdiagram är oförändrade när antennen används som mottagarantenn. En viktig punkt vid noggrann mätning av antennvinst är att övertyga sig om att sändareffekten inte förändras när man skiftar mellan mätobjekt och referensantenn. Därför väljer man oftast att ha sändaren permanent ansluten till en antenn som skapar ett radiofält. Sedan mäter man mottagen signal från mätobjektet, respektive från referensantennen, och utför beräkningarna på samma sätt som ovan. Detta ger större mätnoggrannhet.

dB_i respektive dB_d

Antennvinst anges inte i dB, utan i dB_i. På detta sätt talar vi om att referensantennen är en förlustfri isotrop antenn.

Det är antennvinsten G som anges i antenndatabladet, inte direktiviteten D . Dessutom kan strålningsdiagrammet vara uppritat.

$$G = \eta_L \cdot D$$

Antennvinsten G kan antingen anges

- relativt en förlustfri halv vågs dipol [dB_d], eller
- relativt en förlustfri isotrop antenn [dB_i].

Man omvandlar mellan dessa två sätt genom att lägga till eller dra ifrån 2,15 dB.

$$G[\text{dB}_i] = G[\text{dB}_d] + 2,15 \text{ dB}$$

Vid själva mätningen så jämför man mot en referensantenn som inte är en isotrop antenn. På frekvenser under 1 GHz är det vanligt att referensantennen är en halv vågs dipol. Den antennvinst som man får vid jämförelse mot en förlustfri halv vågs dipol betecknas med dB_d, som visar att referensen var en dipol (egentligen halv vågs dipol om man skall vara petig).

Den förlustfria halv vågsdipolen har antennvinsten 2,15 dB_i. Om antennvinsten anges i dB_d så skall man addera 2,15 dB för att få antennvinsten uttryckt i dB_i.

14.10 Effektiv mottagningsarea A_e

Den signal som mottagarantennen befinner sig i är effekttäthet, effekt per ytenhet. För att beskriva hur mycket effekt antennen fångar upp så definierar vi antennens effektiva mottagningsarea, effektiva absorptionsarea A_e (Effective Aperture).

Mottagningsarea och antennvinst hör ihop

Antennens förmåga att fokusera effektutstrålningen i en viss riktning motsvaras av större mottagningsarea för signaler som kommer från den riktningen. Mottagningsarean är alltså inte konstant utan varierar med riktningen, helt proportionellt mot antennvinsten.

$$A_e = G \cdot \frac{\lambda_0^2}{4\pi}$$

Halvvågsantennens mottagningsarea

En förlustfri halvvågsdipol eller halvvågsantenn har en antennvinst som är 2,15 dBi. Översta bilden visar mottagningsarean för signaler som kommer in mot antennen i maxriktningen.

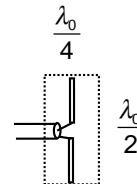
Parabolantennens mottagningsarea

Parabolantennen är intressant därför att den "fångar" upp den signal som kommer inom dess yta. Men när signalen skall tas emot av den antenn som sitter i parabolens fokus, så måste den antennen ha ett strålningsdiagram som täcker parabolytan, men inte mer. Detta kan inte realiseras. De yttre delarna av parabeln kan inte utnyttjas fullt ut. Förhållandet mellan "effektiva" ytan och fysiska ytan kallas aperturverkningsgrad $\eta_{apertur}$. Detta är ingen effektförlust utan bara en omräkningsfaktor mellan fysiska ytan och en tänkt yta utnyttjad till 100 %.

Men självklart finns även resistiva förluster (uppvärmning) η_L som måste med i beräkningen.

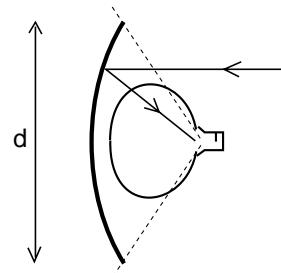
Ett realistiskt värde på $\eta_L \cdot \eta_{apertur}$ är ungefär 0,5.

Ur detta kan antennvinsten härledas som funktion av parabolens diameter. En parabol med 1 m diameter får antennvinsten 39 dBi vid 12 GHz.



$$\begin{aligned} A_e(\text{dipol}) &= G \cdot \frac{\lambda_0^2}{4\pi} = 10^{\frac{2,15}{10}} \cdot \frac{\lambda_0^2}{4\pi} = \\ &= 1,64 \cdot \frac{\lambda_0^2}{4\pi} \approx \frac{\lambda_0}{2} \cdot \frac{\lambda_0}{4} \end{aligned}$$

Halvvågsantennens mottagningsarea



$$Y(\text{fysisk yta}) = \frac{\pi d^2}{4}$$

$$A_e(\text{parabol}) \approx \eta_L \cdot \eta_{apertur} \cdot Y$$

$$A_e(\text{parabol}) \approx 0,5 \cdot \frac{\pi d^2}{4} \approx G \cdot \frac{\lambda_0^2}{4\pi}$$

$$G \approx 0,5 \cdot \left(\frac{\pi d}{\lambda_0} \right)^2$$

$$d = 1 \text{ m}, f = 12 \text{ GHz:}$$

$$G \approx 0,5 \cdot \left(\frac{\pi 1}{0,025} \right)^2 = 7896$$

$$G = 10 \cdot \log 7896 = 39 \text{ dBi}$$

14.11 Mottagarantenn

Jag har mest talat om sändarantenn. Vad skiljer mottagarantenn från sändarantenn?

Samma strålningsdiagram och antennvinst

Antennen får exakt samma strålningsdiagram oberoende av om den används för sändning eller för mottagning. Därav följer att den även har samma antennvinst i båda fallen.

Samma impedans

Antennens inre impedans, sådan den upplevs av sändaren, och av mottagaren, är exakt densamma i båda fallen.

Inte samma bandbredd

Sändaren är en spänningsgenerator med extremt låg inre impedans som ansluts till antennimpedansen. Detta ger en resonanskrets där Q -värdet, bandbredden, enbart bestäms av antennens strålnings- och förlustr resistans.

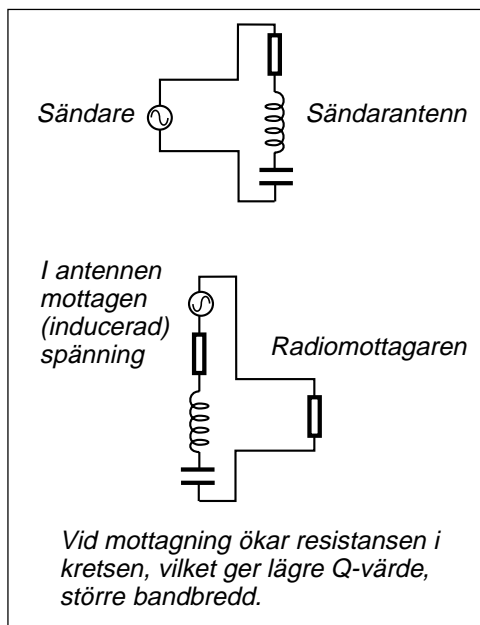
Radiomottagaren absorberar den mottagna effekten. Radiomottagaren ser ut som ett motstånd på ungefär 50Ω . Därför får vi vid mottagning en resonanskrets som belastas med högre resistans, vilket ger lägre Q -värde, större bandbredd.

Om antennens strålningsresistans avviker från det nominella, t.ex. 50Ω , så får vi en relativt kraftig förändring av den utstrålade effekten vid sändning.

Vid mottagning kommer effekten som går till radiomottagaren bara att ändras i ringa grad för samma förändring av antennens strålningsresistans.

OBS att detta resonemang är grovt förenklat eftersom transformation i antennkabeln tillkommer. Det förändrar dock inget i princip.

Varför fungerar Yagi-antennen bra som mottagarantenn över större delen av UHF-TV-bandet, medan den samtidigt är extremt smalbandig när den används som sändarantenn? Jo, vid mottagning accepterar vi större "missanpassning" därför att det inte



inverkar lika kraftigt på effektöverföringen till mottagaren.

En annan orsak är att vi vid TV-mottagning inte alltid är ute efter så kraftig mottagen signal. Vi kan acceptera att tappa signal på grund av missanpassning, därför att vi först och främst är ute efter ett smalt rikttningsdiagram för att undertrycka reflexer, slippa skuggbilder.

Verkningsgraden inte alltid så viktig vid mottagning

En sändarantenn bör ha hög verkningsgrad. Effekten skall bli radiovågor, inte värme i antennen (ohmskt motstånd i antennledaren).

Vid mottagning är signal/stör-förhållandet viktigt, inte signalens absolutvärde. Om störningarna kommer in via antennen medför låg verkningsgrad att signal och störning dämpas lika mycket. Signal/stör-förhållandet är oförändrat.

Exempel på en god mottagarantenn med extremt låg verkningsgrad är ferrit-antennen för mellanvägsmottagning (transistorradio för LF- och MF-bandet).

15 — Radiokommunikation

Sändareffekten sprids som ljuset. Att beräkna hur mycket ljus som infaller på en viss yta är inte speciellt komplicerat, så länge som det inte finns hinder i vägen.

Besvärligt blir det först när vi inte har fri sikt mellan ljuskällan och mottagningsarean.

Vid beskrivning av radiosystem förekommer begreppen EIRP respektive ERP.

Radiokommunikation består av en massa "vinster" och "förluster". Vi ställer upp en länkbudget.

Antennen fångar inte bara upp radiosignaler från radiosändare, utan även radiobrus från diverse brusällor.

15.1 Radiosändaren

Radiosändarens uppgift

Radiosändaren lämnar effekten P_s som matas in i antennen. Antennen skapar radiosignal som på avståndet r meter ger effekttätheten S W/m².

P_s = sändareffekt till sändarantennen

P_{ut} = utstrålad effekt från sändarantennen

G_s = sändarantennens antennvinst

D_s = sändarantennens direktivitet

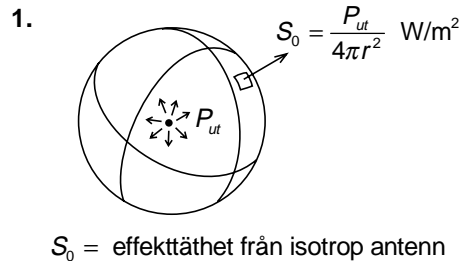
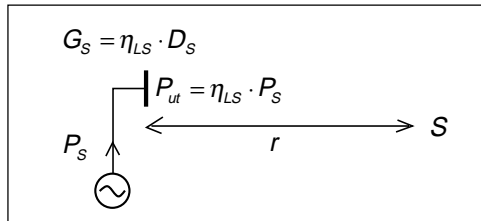
η_{LS} = sändarantennens verkningsgrad

r = avstånd till mottagarantennen

S = effekttäthet vid mottagarantennen

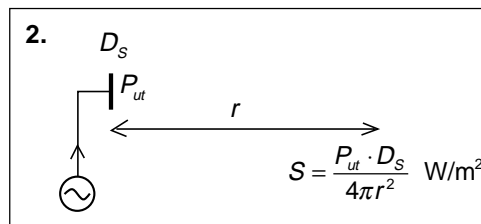
1. Om antennen vore isotrop

Om sändarantennen vore isotrop så skulle den utstrålade effekten P_{ut} sprida sig likformigt åt alla håll och fördela sig på ytan av en sfär med radien lika med r .



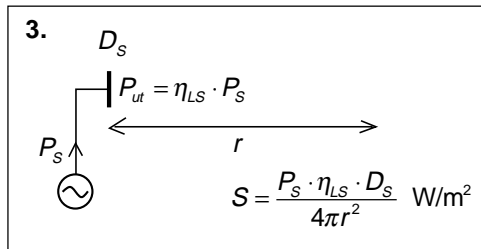
2. Sändarantennen har direktivitet

Sändarantennen sprider inte effekten likformigt utan en större del av den utstrålade effekten går i huvudloben. Hur mycket kraftigare denna signal är bestäms av antennens direktivitet, riktverkan. Här beräknar vi den verkliga effekttätheten vid mottagarantennen.



3. Vi känner inte den utstrålade effekten

Normalt är det sändareffekten P_s vi känner. Hur stor del av denna som strålar ut bestäms av antennens verkningsgrad η_{LS} .

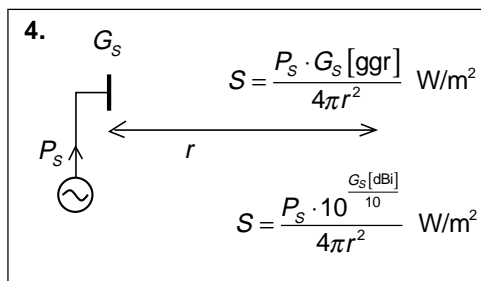


4. Men vi känner inte verkningsgraden?

Det gör inget. Vi känner oftast inte D_s heller utan det är $G_s = \eta_{LS} \cdot D_s$ vi känner.

Det är med dessa data vi utför beräkningar

Vi ser att alla uppgifter hänger ihop. Men de data vi känner vid beräkning av en radioförbindelse är vanligtvis sändarens uteffekt P_s och sändarantennens antennvinst G_s uttryckt i dBi samt naturligtvis avståndet r .



15.2 Radiomottagaren

Från radiosändarantennen kommer effekttäthet. Vår mottagarantenn har en viss mottagningsarea för signaler som kommer från den riktningen. Det är bara att beräkna den mottagna effekten.

$$P_m = S \cdot A_e$$

Normalt känner vi inte antennens mottagningsarea utan dess antennvinst.

$$A_e = G_m \cdot \frac{\lambda_0^2}{4\pi}$$

Detta ger:

$$P_m = S \cdot G_m \cdot \frac{\lambda_0^2}{4\pi}$$

Men antennvinst uttrycks i dBi:

$$P_m = S \cdot 10^{\frac{G_m[\text{dBi}]}{10}} \cdot \frac{\lambda_0^2}{4\pi} = S \cdot 10^{\frac{G_m[\text{dBi}]}{10}} \cdot \frac{c_0^2}{4\pi \cdot f^2}$$

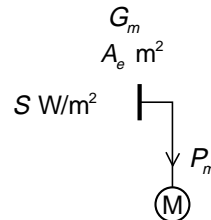
Exempel:

En basstation på 1800 MHz sänder med 5 W och har en sändarantenn med antennvinsten 7 dBi. Du befinner dig högt och fritt på 2 km avstånd och har fri sikt till basstationen. Din ficktelefon har antennvinsten -6 dBi (dålig verkningsgrad hos inbyggd antenn). Beräkna mottagen effekt uttryckt i dBm. Mottagarens känslighetsgräns är -100 dBm. Räcker signalen?

Först beräknar vi effekttätheten. Det är bara att sätta in siffrvärdena.

Därefter beräknar vi den mottagna effekten.

Sedan omvandlar vi effekten från W till mW genom att multiplicera med 1000, och detta omvandlar vi till dB.



P_m = effekt som matas till mottagaren
 G_m = mottagarantennens antennvinst
 A_e = antennens mottagningsarea
 S = effekttäthet från sändarantennen

$$S = \frac{P_s \cdot 10^{\frac{G_s[\text{dBi}]}{10}}}{4\pi r^2} = \frac{5 \cdot 10^{\frac{7[\text{dBi}]}{10}}}{4\pi 2000^2} = \frac{5 \cdot 5,01}{4\pi 2000^2} = 0,5 \cdot 10^{-6} \text{ W/m}^2$$

$$P_m = S \cdot 10^{\frac{G_m[\text{dBi}]}{10}} \cdot \frac{c_0^2}{4\pi \cdot f^2} = 0,5 \cdot 10^{-6} \cdot 10^{\frac{-6}{10}} \cdot \frac{9 \cdot 10^{16}}{4\pi \cdot 1,8^2 \cdot 10^{18}} = 0,5 \cdot 10^{-6} \cdot 0,25 \cdot 0,22 \cdot 10^{-2} = 2,75 \cdot 10^{-10} \text{ W}$$

$$P_m [\text{dBm}] = 10 \cdot \log(2,75 \cdot 10^{-10} \cdot 1000) = -65,6 \text{ dBm}$$

↑ 0 dBm
 + -20 dBm
 + -65 dBm
 + -100 dBm

-65,6 dBm är en starkare signal än -100 dBm. Signalstyrkan räcker.

15.3 EIRP

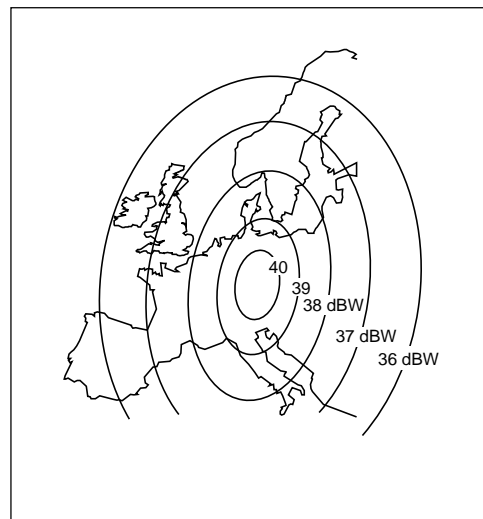
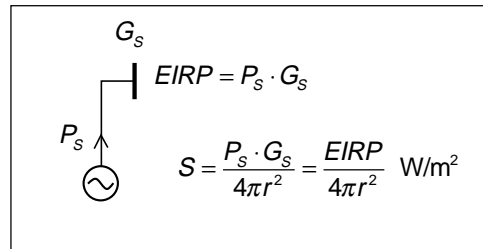
Effektivt isotropt utstrålad effekt EIRP

Vad karakteriserar sändarstationen? Jo, sändareffekten och antennvinsten. Därför multiplicerar man ofta ihop dessa båda och anger sändarstationens EIRP (Effective Isotropic Radiated Power), som alternativt till sändareffekt och antennvinst.

Bilden visar täckningskartan för en tv-satellit på 12 GHz. Satellitantennens huvudlob är riktad mitt i Europa. Ju längre från denna mittpunkt man befinner sig, desto längre ner på antennloben hamnar man, och antennvinsten sjunker.

På kartan finns markerat den EIRP som strålas ut mot olika områden. Mot stockholmsområdet får man 38 dBW, vilket är sändareffekten multiplicerat med antennvinsten i den riktningen. Vad blir effekttätheten om vi sätter avståndet till 40 000 km?

$$S = \frac{EIRP}{4\pi r^2} = \frac{10^{\frac{38}{10}}}{4\pi(4 \cdot 10^7)^2} = \frac{6310}{201 \cdot 10^{14}} = 3,14 \cdot 10^{-13} \text{ W/m}^2$$



15.4 ERP och V/m

I radions barndom när man höll till på långvåg, mellanvåg och kortvåg, så var det vanligast att beskriva radiosignalen med den elektriska fältstyrkan E . Eftersom förhållandet mellan E och H är 377Ω så kan vi enkelt omvandla effekttäthet till fältstyrka.

Halvvågsdipolen har antennvinsten 2,15 dBi. Om vi utnyttjar detta får vi ett uttryck för fältstyrkan från en halvvågsdipol:

$$E = \frac{7 \cdot \sqrt{P_s}}{r} \text{ V/m}$$

ERP (Effective Radiated Power)

För tv-rundradiosändare var det tidigare vanligt att man uppgav inte sändareffekten, utan sändareffekten multiplicerat med antennvinsten relativt dipol. På detta sätt fick man den sändareffekt som skulle matas till den halvvågsdipol för att ge samma signalstyrka som den verkliga sändaren.

UHF-tv-sändarna har en ERP på 1 MW (Effective Radiated Power, från dipol). Sändareffekten är 40 kW. Man har -3 dB dämpning i antennkabeln vilket ger 20 kW vid antennen. Själva antennen har 17 dBd antennvinst, d.v.s. 50 ggr relativt en halvvågsdipol.

Det är enkelt att omvandla från ERP till EIRP

EIRP-värdet är 2,15 dB högre, vilket motsvarar 1,64 ggr när man räknar med effekt.

$$ERP \cdot 1,64 = EIRP$$

Vi ser hur förvillande det blir med två olika referensantenner, dels den isotropa antennen, dels halvvågsdipolen. Som tur är håller man på att enas om att alltid referera till den isotropa antennen.

$$\begin{cases} \frac{E}{H} = 377 \Omega \Rightarrow H = \frac{E}{377} \\ S = E \times H \Rightarrow S = \frac{E^2}{377} \end{cases}$$

(Härledningen är lite "halvslarvig" eftersom E och H är vektorer som är vinkelräta mot varandra).

$$\begin{aligned} E(\text{från dipol}) &= \sqrt{377 \cdot S} = \sqrt{377 \cdot \frac{P_s \cdot G_s}{4\pi r^2}} = \\ &= \sqrt{\frac{377 \cdot P_s \cdot 10^{2,15}}{4\pi r^2}} = \\ &= \sqrt{\frac{49,2 \cdot P_s}{r^2}} = \\ &= \frac{7 \cdot \sqrt{P_s}}{r} \text{ V/m} \end{aligned}$$

15.5 Länkbudget

Lite decibel-räkning

Vi utgår från uttrycken för effekttäthet och mottagen effekt, samt löser ut mottagen effekt som funktion av sändareffekten, antennvinsterna, avståndet och frekvensen.

Sedan är det bara att räkna på. Tänk på att detta är mottagen effekt när det är fri sikt mellan antennerna.

$$\begin{cases} S = \frac{P_s \cdot G_s}{4\pi r^2} \\ P_m = S \cdot G_m \cdot \frac{\lambda_0^2}{4\pi} \end{cases}$$

$$P_m = \frac{P_s \cdot G_s}{4\pi r^2} \cdot G_m \cdot \frac{\lambda_0^2}{4\pi} = \frac{P_s \cdot G_s}{4\pi r^2} \cdot G_m \cdot \frac{c_0^2}{4\pi \cdot f^2}$$

$$\begin{aligned} P_m &= \frac{P_s \cdot G_s}{4\pi r^2} \cdot G_m \cdot \frac{\lambda_0^2}{4\pi} = \frac{P_s \cdot G_s}{4\pi r^2} \cdot G_m \cdot \frac{c_0^2}{4\pi \cdot f^2} = P_s \cdot G_s \cdot G_m \cdot \frac{c_0^2}{(4\pi)^2 r^2 \cdot f^2} = \\ &= P_s \cdot G_s \cdot G_m \cdot \frac{c_0^2}{(4\pi)^2 r_{km}^2 \cdot 10^6 \cdot f_{MHz}^2 \cdot 10^{12}} = P_s \cdot G_s \cdot G_m \cdot \frac{(3 \cdot 10^8)^2}{(4\pi)^2 \cdot 10^{18}} \cdot \frac{1}{f_{MHz}^2} \cdot \frac{1}{r_{km}^2} = \\ &= P_s \cdot G_s \cdot G_m \cdot \frac{(3 \cdot 10^8)^2}{(4\pi)^2 \cdot 10^{18}} \cdot \frac{1}{f_{MHz}^2} \cdot \frac{1}{r_{km}^2} = P_s \cdot G_s \cdot G_m \cdot \frac{1}{1755} \cdot \frac{1}{f_{MHz}^2} \cdot \frac{1}{r_{km}^2} \end{aligned}$$

Vi stoppar in avståndet i km och frekvensen i MHz och får detta uttryck:

$$P_m = P_s \cdot G_s \cdot G_m \cdot \frac{1}{1755} \cdot \frac{1}{f_{MHz}^2} \cdot \frac{1}{r_{km}^2}$$

Därefter omvandlar vi till decibel och får följande uttryck:

$$10 \cdot \log P_m = 10 \cdot \log \left[P_s \cdot G_s \cdot G_m \cdot \frac{1}{1755} \cdot \frac{1}{f_{MHz}^2} \cdot \frac{1}{r_{km}^2} \right]$$

$$10 \cdot \log P_m = 10 \cdot \log P_s + 10 \cdot \log G_s + 10 \cdot \log G_m - 10 \cdot \log 1755 - 10 \cdot \log f_{MHz}^2 - 10 \cdot \log r_{km}^2$$

Vi känner igen definitionerna på effekterna i dBW och antennvinsterna i dBi, det som står i databladen.

$$P_m [\text{dBW}] = P_s [\text{dBW}] + G_s [\text{dBi}] + G_m [\text{dBi}] - 32,44 - 20 \cdot \log f_{MHz} - 20 \cdot \log r_{km}$$

Det går precis lika bra att räkna i dBm, men det skall vara samma i både sändareffekten och mottagareffekten.

$$P_m [\text{dBm}] = P_s [\text{dBm}] + G_s [\text{dBi}] + G_m [\text{dBi}] - 32,44 - 20 \cdot \log f_{MHz} - 20 \cdot \log r_{km}$$

Vi prövar med det tidigare exemplet:

En basstation på 1800 MHz sänder med 5 W och har en sändarantenn med antennvinsten 7 dBi. Du befinner dig högt och fritt på 2 km avstånd och har fri sikt till basstationen. Din ficktelefon har antennvinsten -6 dBi (dålig verkningsgrad hos inbyggd antenn). Beräkna mottagen effekt uttryckt i dBm. Mottagarens känslighetsgräns är -100 dBm. Räcker signalen?

$$P_s [\text{dBm}] = 10 \cdot \log 5000 = 37 [\text{dBm}]$$

$$\begin{aligned} P_m [\text{dBm}] &= P_s [\text{dBm}] + G_s [\text{dBi}] + G_m [\text{dBi}] - 32,44 - 20 \cdot \log f_{\text{MHz}} - 20 \cdot \log r_{\text{km}} = \\ &= 10 \cdot \log 5000 + 7 - 6 - 32,44 - 20 \cdot \log 1800 - 20 \cdot \log 2 = \\ &= 37 + 7 - 6 - 32,44 - 65,1 - 6 = -65,54 [\text{dBm}] \end{aligned}$$

Länkbudget

Vi kan summera de olika bidragen, som i en "budget":

1. Sändareffekten

$$P_s [\text{dBm}] = 10 \cdot \log 5000 = 37 [\text{dBm}]$$

2. Sändarantennens antennvinst

$$G_s [\text{dBi}] = +7 \text{ dBi}$$

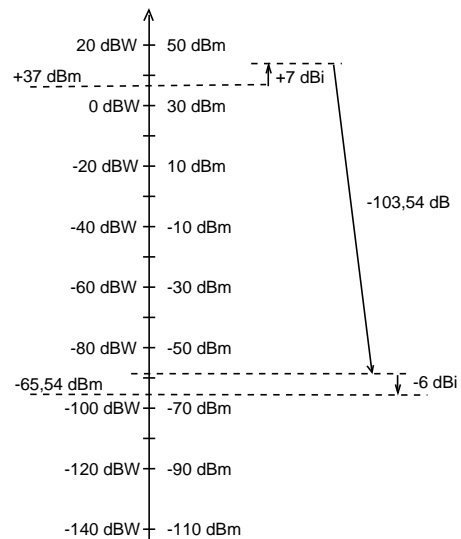
3. Dämpningen mellan isotropa antenner

$$\begin{aligned} L_s [\text{dB}] &= -32,44 - 20 \cdot \log f_{\text{MHz}} - 20 \cdot \log r_{\text{km}} = \\ &= -32,44 - 65,1 - 6 = -103,54 [\text{dB}] \end{aligned}$$

4. Mottagarantennens antennvinst

$$G_m [\text{dBi}] = -6 \text{ dBi}$$

Summera, så får vi P_m



Länkbudget:

1. Sändarens uteffekt: + ____ [dBW]
 2. Antennkabelns dämpning: - ____ [dB]
 3. Sändarantennens antennvinst: + ____ [dBi]
 4. Sträckdämpning mellan isotropa antenner: - ____ [dB]
 5. Dämpning i atmosfärgaser: - ____ [dB]
 6. Dämpning vid regn: - ____ [dB]
 7. Dämpning om sikten är skymd av träd: - ____ [dB]
 8. Dämpning vid reflexion mot husväggar: - ____ [dB]
 9. Fädningsmarginal (korta reflexer): - ____ [dB]
 10. Mottagarantennens antennvinst: + ____ [dBi]
 11. Antennkabelns dämpning: - ____ [dB]
- Summa effekt till radiomottagaren: ____ [dBW]

Verklig länkbudget

En verklig länkbudget kan bestå av många fler poster. Vi skall titta på posterna i länkbudgeten här ovan.

1. Sändarens uteffekt

Detta är ungefär lika med framåtgående effekt vid sändarutgången, om antennen har lågt SVF.

2. Antennkabelns dämpning

Antennen kan sitta högt i en mast. Då behövs lång antennkabel. Dämpningen beräknas genom att utgå från kabeltillverkarens värde på dämpning/meter, multiplicerat med kabellängden.

3. Sändarantennens antennvinst

Detta värde finns i antennens datablad. Det gäller bara att se upp så att den riktning där antennen har den uppgivna antennvinsten verkligen sammanfaller med den riktning där vi använder antennen.

4. Sträckdämpning mellan isotropa antenner

Fri sikt, "fritt rum", innebär inte bara fri sikt mellan antennerna, utan att en volym, första Fresnell-zonen, skall vara fri från störande objekt.

Verklig sändareffekt till antennen:

Framåtgående effekt vid sändaren multiplicerat med kabeldämpningen, minus reflekterad effekt vid sändaren dividerat med kabeldämpningen.

Sträckdämpning vid "fritt rum"

$$L_s [dB] = -32,44 - 20 \cdot \log f_{MHz} - 20 \cdot \log r_{km}$$

När det inte är fri sikt så utför man ofta någon typ av fältstyrkemätning för att utröna hur signalen utbreder sig. Diagrammet visar en sådan mätning där fritt-rum-dämpningen finns inritad som jämförelse.

Medan signalen vid fritt rum avtar med r^{-2} så kan man approximativt säga att i verklig miljö så avtar signalen med $r^{-3.5}$ eller r^{-4} , d.v.s. vi får

$$L_S [dB] = -32,44 - 20 \cdot \log f_{MHz} - 35 \cdot \log r_{km}$$

eller

$$L_S [dB] = -32,44 - 20 \cdot \log f_{MHz} - 40 \cdot \log r_{km}$$

5. Dämpning i atmosfärgaser

Vattenånga ger ett litet dämpningstillskott vid 23 GHz, och syreatomerna ger ett kraftigt dämpningstillskott vid 60 GHz.

6. Dämpning vid regn

Vid frekvenser över ca 10 GHz kan man få kraftiga dämpningstillskott vid häftiga regnväder.

7. Dämpning om sikten är skynd av träd

Vegetationsdämpning är ett stort problem vid 1800 MHz.

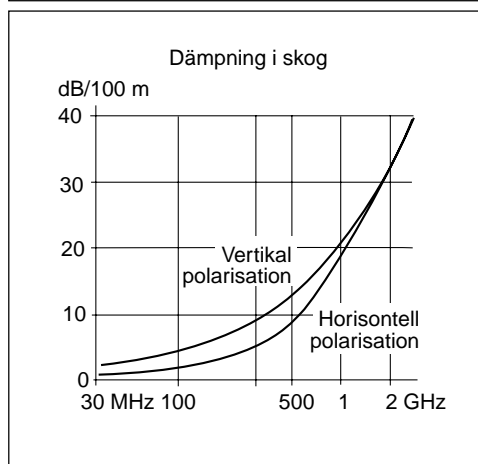
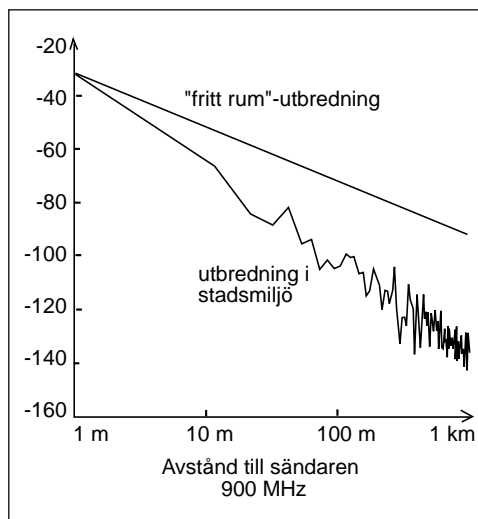
8. Dämpning vid reflexion mot husväggar

I stadsbebyggelse tar sig signalen fram genom reflexion mot husväggar. Hög frekvens (1800 MHz) reflekteras lättare än låg frekvens (450 MHz).

9. Fädningsmarginal

Vid flervägsutbredning hamnar signalerna ömsom i medfas och blir stark, ömsom i motfas och blir svag, som ståendevågor på en ledning. Även i verkligheten kan man få en halv våglängd mellan minpunkter. Detta kan hanteras om man sänder med 10 – 20 dB högre signal (fädningsmarginal) än vad som egentligen skulle behövts.

10 – 11. Se punkt 2 – 3.



15.6 Brus

Hittills har vi studerat storleken på den mottagna effekten P_m som härrör från den önskade sändaren. Men mottagarantennen tar inte bara emot den önskade signalen. Mottagarantennen fångar även upp icke önskade signaler och brus. Störningar från icke önskade sändare kan undvikas genom att välja en störningsfri frekvenskanal, men bruset kommer vi inte ifrån. Det finns överallt.

Om bruset kommer från alla riktningar?

Vi tänker oss att bruset strålar in mot antennen, lika starkt från alla riktningar. Vad innebär detta?

Om vi tänker oss en isotrop mottagarantenn. Den tar emot bruseffekt, och effekt-tätheten i respektive riktning multipliceras med mottagararean som ju beror på antennvinsten i den riktningen. Det totala mottagna antennbruset N_a består i detta fall av lika starka bidrag från alla riktningar.

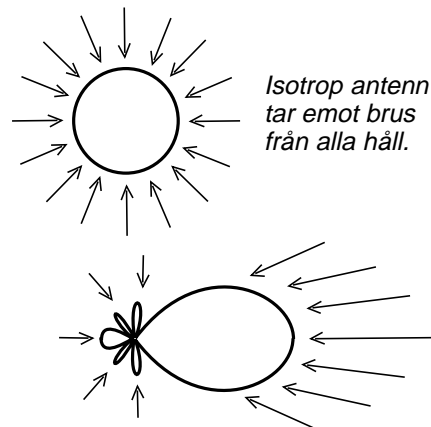
Nu byter vi antenn och väljer en riktantenn med hög antennvinst. Även denna antenn kommer att få ett mottaget antennbrus N_a med bidrag från alla riktningar. Men bidragen varierar med riktningen. Det brus som faller in i huvudloben ger ett stort bidrag till N_a , medan brus från andra riktningar bara ger ett litet bidrag.

Hur stort blir det totala bruset? Lustigt nog blir bruseffekten exakt lika stor i båda fallen. Det som skiljer är att den rundstrålände antennen tar emot brus från alla riktningar, medan riktantennen fångar upp mest brus i huvudloben men mindre från andra riktningar. Men totala bruseffekten blir densamma.

Självklart måste det bli på detta sätt. Det är ju så här sändareffekten sprids. Lika i alla riktningar, eller mer i en riktning. Men totala effekten är densamma.

Antennens verkningsgrad inverkar

Bruseffekten är inte exakt lika för alla antenner. Totala antennbruset påverkas av mottagarantennens verkningsgrad η_{Lm} . Detta leder till att signal/brusförhållandet blir proportionellt mot mottagarantennens



Riktantenn tar emot mer brus i huvudloben, men totala bruseffekten är densamma om antennerna har samma verkningsgrad.

$$P_m = S \cdot A_e = G_m \cdot \frac{\lambda_0^2}{4\pi} = S \cdot \eta_{Lm} \cdot D_m \cdot \frac{\lambda_0^2}{4\pi}$$

$$N_a = k \cdot T_a \cdot B \cdot \eta_{Lm}$$

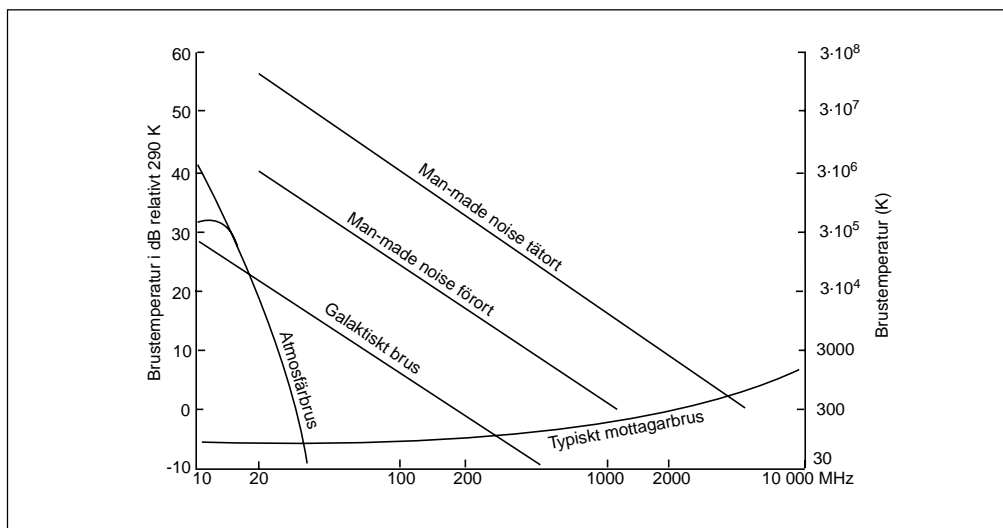
$$N = N_a + k \cdot T_{eq} \cdot B = k \cdot (T_a \cdot \eta_{Lm} + T_{eq}) \cdot B$$

$$\frac{S}{N} = \frac{P_m}{N} = \frac{S \cdot \eta_{Lm} \cdot D_m \cdot \frac{\lambda_0^2}{4\pi}}{k \cdot (T_a \cdot \eta_{Lm} + T_{eq}) \cdot B}$$

$$\frac{S}{N} = D_m \cdot \frac{S \cdot \eta_{Lm} \cdot \frac{\lambda_0^2}{4\pi}}{k \cdot (T_a \cdot \eta_{Lm} + T_{eq}) \cdot B} \approx$$

$$\approx D_m \cdot \frac{S \cdot \frac{\lambda_0^2}{4\pi}}{k \cdot T_a \cdot B}$$

direktivitet, om antennbruset är högt. Det gör alltså inget om mottagarantennen har låg verkningsgrad, bara den har hög direktivitet. Men detta gäller alltså bara om bruset är starkt och kommer från alla riktningar.



Antennbrustemperaturen T_a

Diagrammet här ovan visar på vänstra axeln brusnivån i dB relativt T_0 (290 K), och på högra axeln brustemperaturen i Kelvin. Som synes är bruset avsevärt högre än 290 K, åtminstone på frekvenser under 200 MHz.

Atmosfärbrus

Atmosfärbrus är benämningen på brusstörningar som uppstår vid blixurladdningar. Även om åskväder är vanligast under regnperioden vid ekvatorn, så sprider sig dessa blixstörningar runt hela jorden.

Galaktiskt brus

Som framgår av namnet härrör galaktiskt brus från andra källor än jorden och dess atmosfär. Största bruskillan är solen, men en hel del brus kommer även från stjärnor, främst från Vintergatan.

Man-made radio noise

Brus som alstras av elektrisk utrustning går under benämningen "man-made noise". Elektriska strömbrytare som startar eller stoppar ström ger en störningsspike. Överallt där det uppstår gnistor, där uppstår även radiostörningar. Häftig inbromsning av strömmen tillsammans med induktansen i ledningen ger upphov till hög spänning, som ger gnista. Men strömmens häftiga inbromsning ger även radiostrålning. Både radios-

törningen och gnistan beror alltså på strömmens häftiga inbromsning.

Inne i tätort är brusnivån oftast mycket hög, kanske mer än 10 dB över mottagarens egetbrus vid tv-nottagning. Men detta brus dämpas både av tv-antennens verkningsgrad och av dämpningen i antennkabeln. Det antennbrus som kommer fram till tv-mottagaren ligger därför oftast på samma nivå som tv-mottagarens egetbrus T_{eq} .

Brus vid satellitmottagning

När man riktar en parabolantenn upp mot himlen så kommer antennen att huvudsakligen fånga upp galaktiskt brus. Dock finns vid 23 GHz en brustopp som beror på dämpning i atmosfären (vattenånga).

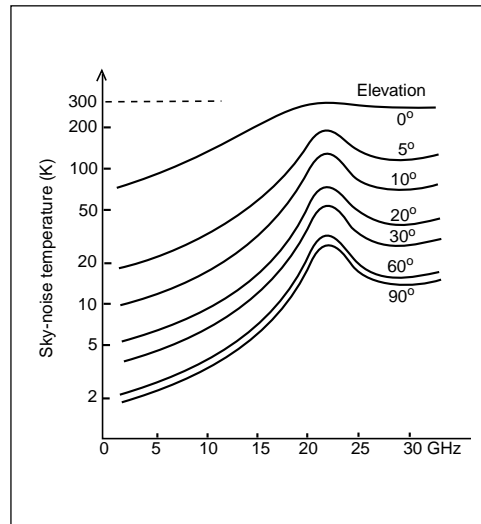
När signalen passerar en dämpningsfaktor så får man ett brustillskott vars storlek beror på dämpningen och dämparens temperatur.

Av bilden ser vi att bruset blir lägst när parabolen riktas rakt upp. Orsaken är att signalen går kortaste vägen genom atmosfären. Ju längre väg genom atmosfären, desto högre dämpning, och desto högre brustillskott.

Normalt tillkommer ytterligare brus som inte framgår av denna bild. Även en parabolantenn har sidolober och backlober även om de är svaga. Om dessa lobar riktas mot marken så kommer de att "se" 290 K och ge ett brustillskott.

Att rikta antennloben genom varma rökgaser från en skorsten ger också brustillskott.

Parabolantennen på satelliten riktar hela sin huvudlob mot jorden. Därför kommer satellitmottagaren att se en antenntemperatur på 290 K.



16 — Optokommunikation

Först tittar vi på tre olika typer av optofibrer, och dämpningen inom olika våglängdsområden.

Sedan kommer vi in på kromatisk dispersion och moddispersion, fenomen som begränsar symbolhastigheten på optofibern.

Som avslutning sägs några ord om WDM, våglängdsmultiplexering, den teknik som i dag införs på optofibrer som redan ligger nergrävda.

Optofibern är det mest fantastiska som hänt inom telekommunikation de senaste 100 åren. I början av 1980-talet var det ofta spänn när man kvällstid försökte ringa rikssamtal.

De långdistansförbindelser som fanns, koaxialkabel med förstärkare på 1,5 km avstånd, och radiolänkar där ytterligare frekvensband inte fanns att tillgå, var alla överbelastade. Och detta var innan Internet hade börjat.

Så kom optofibern. I stället för att gräva ner en koaxialkabel och installera förstärkare på 1,5 km avstånd, en koaxialkabel som kunde transportera 565 Mbit/s, så rullade man ut en kabel med 12, eller 24, eller 48 eller 96 optofibrer, 5 mil lång, och blåste på med en blinkande lampa, i hastigheter på 2,5 Gbit/s på varje fiber. I dag lyser vi med 32 lampor i olika färger på varje optofiber, och varje lampa blinkar i 2,5 Gbit/s.

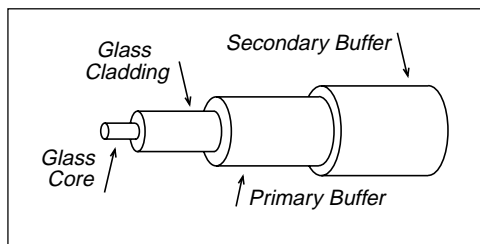
Utan optofibrer så skulle vi inte haft något Internet.

16.1 Tre typer av optofibrer

Optofibers uppbyggnad

Optofibers kärna (core) omges av en mantel (cladding) för att man skall få en väldefinierad övergång i brytningsindex från kärnan till omgivningen. Utanför manteln finns ett eller flera skyddande plastlager (buffer).

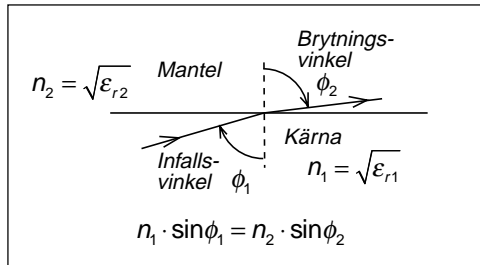
Kärnan och manteln är normalt av glas, men det finns även plastfiber för korta avstånd.



Snells lag

Snells lag från optiken visar hur ljusstrålen bär sig åt. Kärnan (core) har något högre brytningsindex ($n_1 = 1,50$) än omgivande mantel (cladding) ($n_2 = 1,485$). Därför blir brytningsvinkeln större än infallsvinkeln.

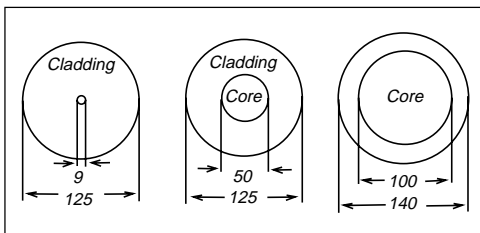
Vid någon viss infallsvinkel blir brytningsvinkeln 90 grader. För alla infallsvinklar som är större får vi totalreflexion. Det är dessa ljusstrålar som utbreder sig i fibern.



Kärnans diameter

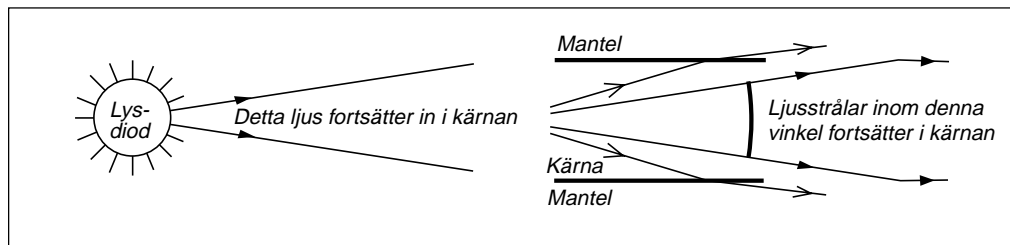
Kärnan kan ha stor diameter, 50 – 100 – 1000 μm , eller liten diameter, mindre än 10 μm .

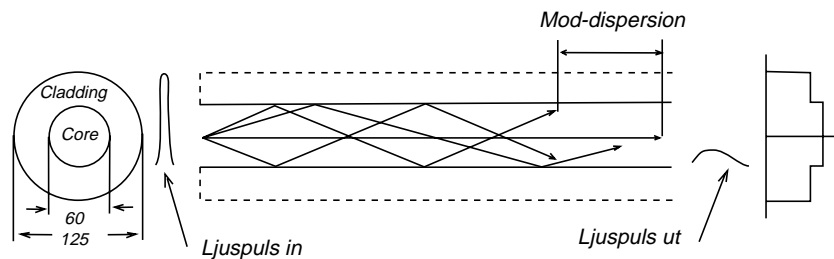
- Om kärnans diameter är stor är det lätt att få in ljuskällans ljus i fibern. Man klarar sig med billig ljuskälla, LED (Light Emitting Diode).
- Om kärnans diameter är liten går bara en mindre del av ljusenergin från en LED in i fibern. Fibrer med liten kärndiameter kräver dyra laserdioder som ljuskälla.
- Om kärnans diameter är liten fördelas hela ljusenergin på litet tvärsnitt. Detta ger hög effekttäthet i glasets som kan leda till olinearitet och intermodulation.



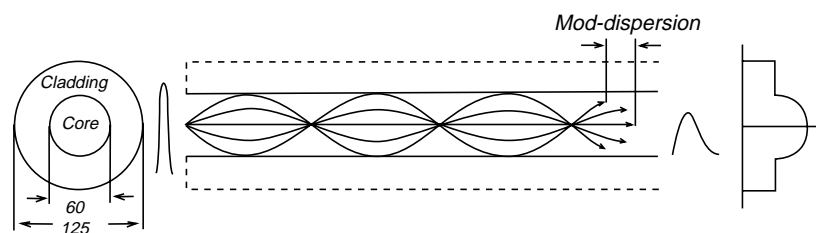
Mod-dispersion

Om kärnans diameter är stor kommer ljusstrålarna att kunna gå olika vägar, gå i olika moder, gå olika lång väg genom fibern. Ljusstrålarna kommer inte fram samtidigt. Ljuspulsen blir breddad på grund av att strålarna går olika väg, något som kallas moddispersion.

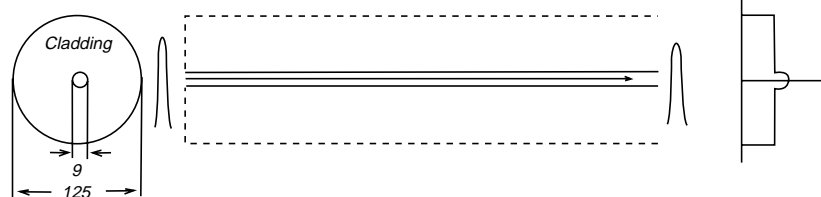


**Stegindexfiber**

Denna fiber har en skarp gräns mellan kärnan och manteln. Fibern karakteriseras av stor moddispersion, d.v.s. ljuspulsen blir kraftigt breddad även vid korta fiberlängder. Bandbredden p.g.a. moddispersion är 20 – 30 MHz-km.

**Gradientindexfiber**

Denna fiber har en flytande gräns mellan kärnan och manteln, vilket leder till att ljuspulserna böjs av mjukt i stället för en skarp reflexion. Fibern har betydligt mindre moddispersion trots samma kärndiameter. De bästa multimodfibrerna av gradientindextyp uppvisar en bandbredd p.g.a. moddispersion på 1 GHz-km.

**Singelmodfiber**

Om kärnan är extremt tunn kan ljusstrålarna bara gå en väg och vi slipper uppdelning på flera ljusstrålar, moddispersion, och får bara minimal breddning av ljuspulsen. Typisk bandbredd p.g.a. moddispersion är 100 GHz-km.

16.2 Våglängd eller frekvens?

Vid kommunikation på optokabel talar man om signalens våglängd, inte om dess frekvens. Men att använda våglängd är egentligen fel. Frekvensen är konstant, men våglängden förändras när ljuset kommer in i

glasfibern (ϵ_r). Överallt där vi i samband med optokommunikation ser begreppet våglängd så är det alltså ljusets (signalens) våglängd i vakuum som man menar.

Elektromagnetiska vågor

Frekvensområdet		Benämning	Frekvens	Våglängd
			30 Hz	10^7 m
ELF	Extremely Low Frequency		300 Hz	10^6 m
VLF	Very Low Frequency		3 kHz	10^5 m
LF	Low Frequency		30 kHz	10^4 m
MF	Medium Frequency		300 kHz	10^3 m
HF	High Frequency		3 MHz	100 m
VHF	Very High Frequency		30 MHz	10 m
UHF	Ultra High Frequency		300 MHz	1 m
SHF	Super High Frequency		3 GHz	10 cm
EHF	Extremely High Frequency		30 GHz	1 cm
THF	Tremendously High Frequency		300 GHz	1 mm
Infrarött	(värmestrålning)		3000 GHz	10^6 Å
			3×10^{13} Hz	10^5 Å
Optokommunikation	Synligt ljus		3×10^{14} Hz	10^4 Å
			3×10^{15} Hz	10^3 Å
Ultra-violett			3×10^{16} Hz	100 Å
	Mjuk röntgen		3×10^{17} Hz	10 Å
			3×10^{18} Hz	1 Å
	Hård röntgen		3×10^{19} Hz	0,1 Å
		Gamma-strålning	3×10^{20} Hz	

16.3 Optofibers dämpning

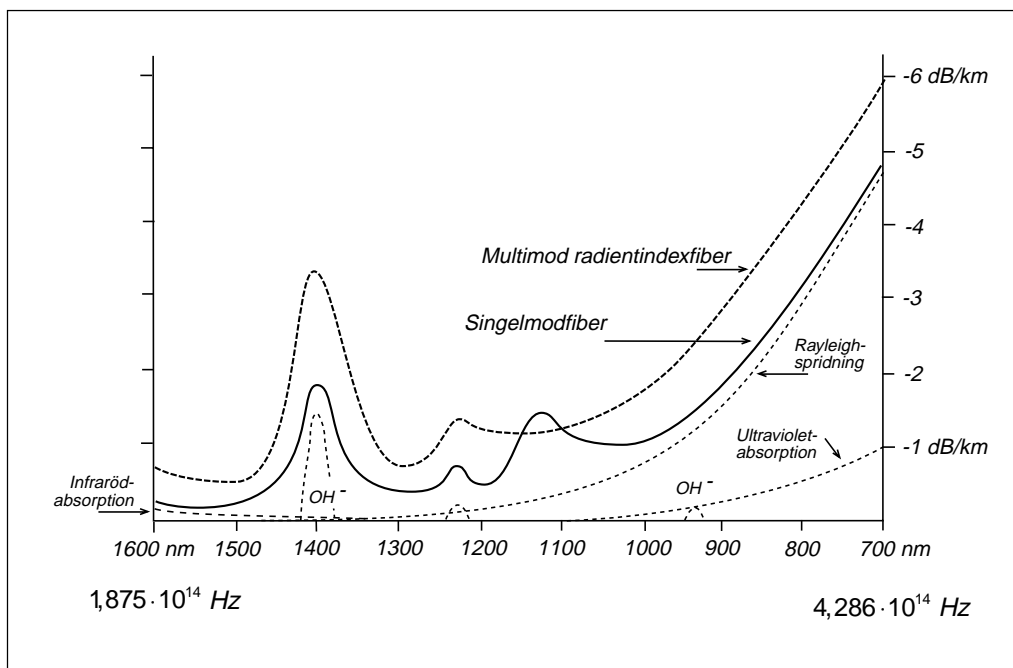
Dämpning

- Vibrationer i SiO_4 -molekyler stjälar effekt från ljusstrålen. Denna absorption kallas infrarödabsorption och avtar med ökande frekvens.
- Små variationer i brytningsindex hos glaset ger spridning, ljus som "sticker iväg" i andra riktningar. Denna spridning kallas Rayleigh-spridning och upplevs som dämpning av ljusstrålen. Rayleigh-spridningen ökar med frekvensen.
- Vibrationer i OH^- -joner (vatten) ger absorption vid $2,73 \mu\text{m}$ och övertonerna $1,38 \mu\text{m}$ och $0,94 \mu\text{m}$.

Tre "fönster"

Man talar om tre "fönster" vid optokommunikation.

- Första fönstret ligger vid 850 nm , inte för att denna våglängd ger låg dämpning, utan för att det var vid denna våglängd som man lyckades tillverka de första ljuskällorna och detektorerna.
- Andra fönstret ligger vid 1310 nm som har så låg dämpning, ca $0,5 \text{ dB/km}$, att man kunde börja använda optofiber i långdistansnätet.
- Tredje fönstret ligger vid $1500 - 1600 \text{ nm}$ där dämpningen är lägst. Man klarar i dag att tillverka optofibrer med en dämpning som inte är mycket högre än den teoretiska dämpningen, ca $0,2 - 0,23 \text{ dB/km}$.



16.4 Kromatisk dispersion

Allt beror på hastigheten

När vitt ljus går genom ett prisma så delas ljuset upp i spektrats alla färger. Varför?

All brytning av elektromagnetisk strålning bygger på att den infallande "strålen" inte är en enda stråle utan ett knippe strålar som utbreder sig som en plan våg, d.v.s. strålarna bildar en plan vågfront och strålningen utbreder sig vinkelrätt mot vågfronten.

När vågen infaller snett mot ett medium där ljushastigheten är en annan, så kommer under ett kort tidsögonblick Δt några strålar på ena kanten att gå fort, medan strålar på andra kanten går långsammare. Då ändrar vågfronten riktning, och utbredningen sker vinkelrätt mot denna nya vågfront.

Brytningsindex är inte konstant

Glaset brytningsindex varierar som funktion av frekvensen. Detta gör att olika färger går olika fort i glaset. Varje färg (frekvens/våglängd) har sitt eget värde på kapacitets-talet, ljushastigheten, brytningsindex, och varje färg får sin egen brytningsvinkel. Detta leder till att ljuset delas upp i regnbågens alla färger.

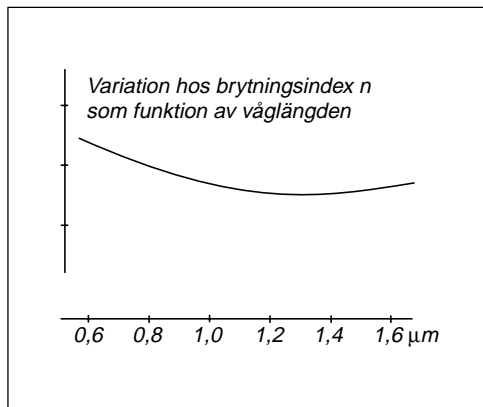
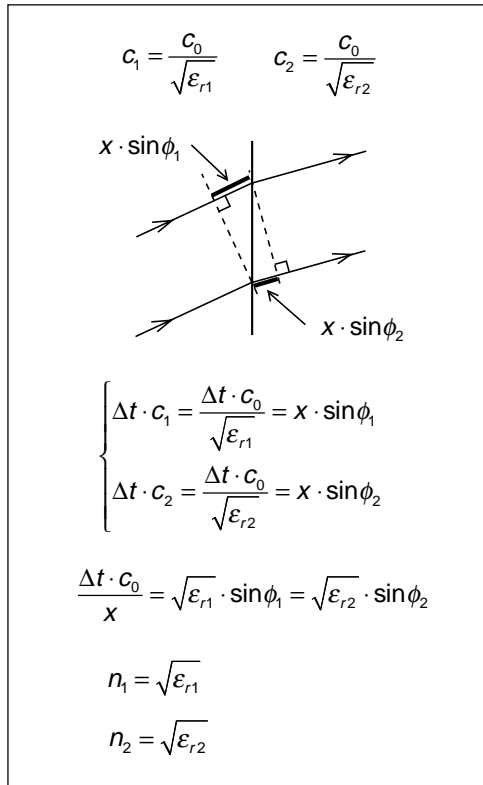
Diagrammet visar hur brytningsindex (och därmed ljushastigheten) varierar med våglängden i kiselglas. Derivatant är ett mått på "tidsspridningen" mellan olika spektralkomponenter från ljuskällan.

Materialdispersion

Hur påverkar detta optokommunikation? Jo, om ljuset från ljuskällan inte innehåller bara en frekvens utan ett spektrum av frekvenser, så går spektralkomponenterna olika fort i optofibern. De kommer inte fram samtidigt. Vi får en breddning av ljuspulsen, kallad materialdispersion, eftersom dispersionen beror på materialet i optofibern.

Vågledardispersion

En annan typ av dispersion som får till följd att olika våglängder går olika fort är vågledardispersion. Vågledardispersion drabbar bara singelmodfibrer. Visserligen går alla ljusstrålar som en enda stråle, men ändå



inte rakt fram. Strålarna kommer att studsas i väggarna, gå i sick-sack, men alla strålar med samma frekvens går på samma sätt.

Däremot kommer strålar av annan frekvens att sick-sacka på annat sätt. Detta betyder att olika frekvenser, olika våglängder, inte kommer fram samtidigt. Resultatet blir breddning av pulsen, vågledardispersion.

Kromatisk dispersion

Kromatisk dispersion, breddning av en ljuspuls om ljuspulsen består av flera våglängder, är summan av materialdispersionen och vågledardispersionen. Över 1300 nm motverkar dessa fenomen varandra.

Noll-dispersion

Hos glas varierar brytningsindex åt ett håll för låga frekvenser, men åt motsatt håll för höga frekvenser. Därför är materialdispersionen noll vid någon frekvens. Denna nollpunkt råkar infalla vid ungefär 1300 nm.

För singelmodfibrer blir den kromatiska dispersionen noll vid 1310 nm (förskjuts uppåt av vågledardispersionen). Därför är 1310 nm den bästa ljusvåglängden för överföring av hög symbolhastighet, om ljuskällan inte är helt "ren".

Ljuskällans spektralrenhet är viktig

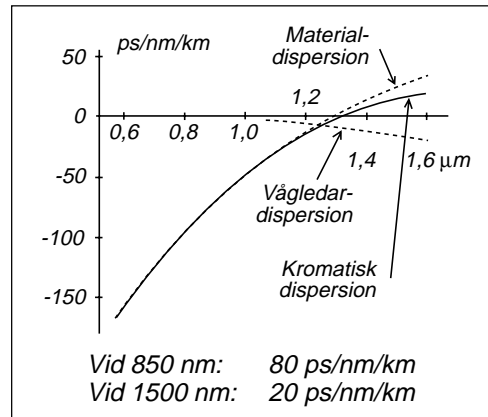
Enda möjligheten att undvika att ljuspulsen breddas om optofibern har kromatisk dispersion är att ljuskällan alstrar bara en enda färg, en enda linje i spektrat.

Men vi amplitudmodulerar ljuset, startar och stoppar ljuskällan, eller kopplar något emellan som stoppar ljuset i takt med moduleringen. Modulering ger breddning av ljusspektrat.

Dispersionsskiftad kabel

Helst skulle man vilja ha lägsta dämpning och noll i den kromatiska dispersionen vid samma frekvens. Detta har lett till utveckling av dispersionsskiftad kabel. Den materialdispersion som finns vid 1550 nm neutraliseras av ett lika stort dispersionsfel med motsatt derivata.

Tyvärr erhålls denna förbättring bara inom ett mycket smalt fönster runt 1550 nm. Utanför "fönstret" får man förvärrad dispersion. Dispersionsskiftad kabel kan därför inte



Exempel 1:

En ljuskälla (LED) har ett spektrum med bredden $\Delta\lambda = 50 \text{ nm}$ vid 850 nm. Hur stor blir pulsbreddningen τ_{krom} om optofibern är 2 km lång?

Ur diagrammet hämtas 80 ps/nm/km.

$$\tau_{krom} = 80 \cdot 10^{-12} \cdot 50 \cdot 2 = 8 \cdot 10^{-9}$$

Exempel 2:

En ljuskälla (laserdiod) har ett spektrum som består av en enda våglängd, 1500 nm. Ljuskällan moduleras med 10 Gsym/s. Vad blir spektrats bredd? Hur stor blir pulsbreddningen τ_{krom} om optofibern är 100 km lång?

Bandbredd = $\pm 5 \text{ GHz}$

1500 nm = $2 \cdot 10^{14} \text{ Hz}$

$$\Delta\lambda \approx 1500 - \frac{3 \cdot 10^8}{2 \cdot 10^{14} + 10 \cdot 10^9} = 0,075 \text{ nm}$$

Ur diagrammet hämtas 20 ps/nm/km.

$$\tau_{krom} = 20 \cdot 10^{-12} \cdot 0,075 \cdot 100 = 0,150 \cdot 10^{-9}$$

användas när man i dag uppgraderar med WDM-teknik, flera parallella optosystem inom 1500 – 1600 nm på samma fiber.

16.5 Optoförbindelsens bandbredd

Pulsen breddas

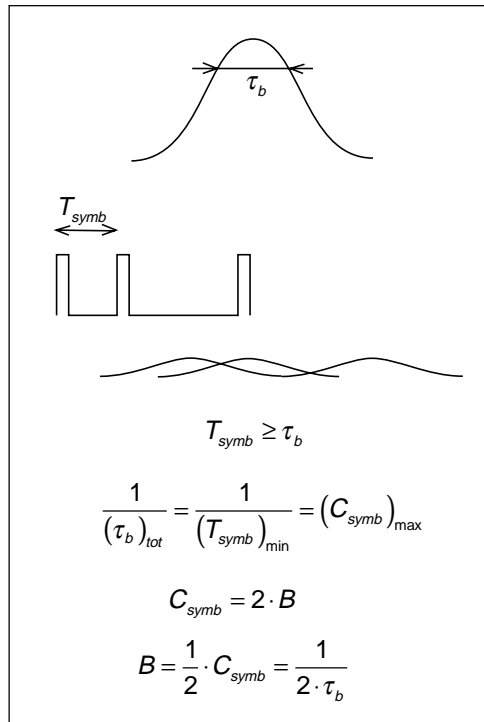
Om vi matar in korta ljuspulser på optofibern så är dessa breddade när de kommer fram. Pulsbredden vid -3 dB kallar vi τ_b . För att pulserna skall kunna urskiljas när de kommer fram krävs att symboltiden T_{symp} är större än τ_b , annars flyter pulserna ihop.

Optoförbindelsens bandbredd

Om vi har en högsta gräns för symbolhastigheten så ser optoförbindelsen ut som om den moduleras av en informationssignal som matats genom ett lågpasfilter med övre gränsfrekvens lika med halva symbolhastigheten, jfr. basbandskommunikation.

Bandbredd p.g.a. kromatisk dispersion

Utgående från spridningen p.g.a. kromatisk dispersion kan vi beräkna förbindelsens bandbredd med hänsyn tagen till kromatiska dispersionen.

**Exempel 1:**

$$\tau_{\text{krom}} = 8 \cdot 10^{-9} \Rightarrow$$

$$B_{\text{krom}} = \frac{1}{2 \cdot 8 \cdot 10^{-9}} = 62,5 \text{ MHz}$$

Exempel 2:

$$\tau_{\text{krom}} = 0,150 \cdot 10^{-9} \Rightarrow$$

$$B_{\text{krom}} = \frac{1}{2 \cdot 0,150 \cdot 10^{-9}} = 3,33 \text{ GHz}$$

Bandbredd p.g.a. moddispersion

I optofibers datablad anges pulsbreddningen på grund av moddispersionen direkt som bandbredd i Hz-km. Som typiska värden gäller att vanlig stegindexfiber har bandbredden 20 – 30 MHz-km, de bästa multimodfibrerna har en bandbredd på 1 GHz-km, medan singelmodfibrerna klarar 100 GHz-km.

Modkoppling

Moddispersionen är direkt proportionell mot fiberlängden. Men bara upp till ca 1 km fiberlängd. Små mikroskopiska ofullkomligheter i glaset gör att strålarna ändrar riktning. Sannolikheten för att en och samma stråle har gått maximalt i sicksack längs hela fibern är liten. Strålarna byter mod på vägen.

Vid fiberlängder som är längre än 1 km är det därför rimligare att räkna med roten ut avståndet.

Total bandbredd

Alla dessa olika fördröjningar som påverkar ljusstrålarna kan inte adderas direkt som de är. Den kromatiska dispersionen och moddispersionen är oberoende processer. Därför får man ett mer sannolikt resultat genom att addera fördröjningarna i kvadrat.

$$\tau_{tot}^2 = \tau_{krom}^2 + \tau_{mod}^2$$

Ur detta kan vi räkna fram totala bandbredden:

$$B_{tot} = \frac{1}{\sqrt{\frac{1}{B_{krom}^2} + \frac{1}{B_{mod}^2}}}$$

Exempel 1 och 3:

$$B_{tot} = \frac{1}{\sqrt{\frac{1}{62,5^2} + \frac{1}{50^2}}} = 39 \text{ MHz}$$

$$C_{symb} = C_{bit} = 78 \text{ Mbit / s}$$

Denna förbindelse klarar inte 100 Mbit/s Ethernet.

$$B_{mod} \approx \frac{B[\text{Hz} - \text{km}]}{L[\text{km}]} \quad L \leq 1 \text{ km}$$

$$B_{mod} \approx \frac{B[\text{Hz} - \text{km}]}{\sqrt{L[\text{km}]}} \quad L > 1 \text{ km}$$

Exempel 3:

En stegindexfiber har enligt datablad bandbredden 20 MHz-km. Beräkna bandbredden på grund av moddispersion om fibern är 400 m lång.

$$B_{mod} = \frac{20 \cdot 10^6 [\text{Hz} - \text{km}]}{0,4 [\text{km}]} = 50 \text{ MHz}$$

Exempel 4:

En singelmodfiber har enligt datablad bandbredden 100 GHz-km. Beräkna bandbredden på grund av moddispersion om fibern är 100 km lång.

$$B_{mod} = \frac{100 \cdot 10^9 [\text{Hz} - \text{km}]}{\sqrt{100 [\text{km}]}} = 10 \text{ GHz}$$

Exempel 2 och 4:

$$B_{tot} = \frac{1}{\sqrt{\frac{1}{3,33^2} + \frac{1}{10^2}}} = 3,16 \text{ GHz}$$

$$C_{symb} = C_{bit} = 6,32 \text{ Gbit / s}$$

Denna förbindelse klarar 2,5 Gbit/s (eller högre) i långdistansnätet.

16.6 Våglängdsmultiplexering, WDM (Wavelength Division Multiplexing)

850 nm

De första optosystemen för kommersiellt bruk låg vid 850 nm därför att man klarade att göra ljuskälla och detektor för denna våglängd. Minsta möjliga dämpning är relativt hög, i storleksordningen 2,4 – 3,5 dB/km.

Multimodkabel vid 850 nm används i dag i billig optoutrustning för korta avstånd som hopkoppling av LAN mellan olika våningsplan eller närbelägna byggnader.

1310 nm

Vid 1310 nm blev dämpningen så låg, ca 0,5 dB/km, att man kunde börja bygga långdistansnät. För att klara höga symbolhastigheter var man tvungen att välja singelmodfiber. Som ljuskälla krävdes en dyr laser för att få in tillräckligt med ljus i den tunna fiberkärnan. Största fördelen är att man slipper inverkan av kromatisk dispersion vid denna våglängd.

1550 nm

Absolut lägsta dämpning infaller vid 1550 nm. Tillverkningen av optofiber är i dag så högt utvecklad att man i stort sett uppnår den teoretiska dämpningen av ca 0,20 – 0,23 dB/km. Men däremot finns kromatisk dispersion vid denna våglängd som begränsar symbolhastigheten.

Många fibrer

När man på 1990-talet drog optokabel i långdistansnätet valde man kabel med många fibrer. Att gräva kan kosta 100 – 500 kr/m, optokabelns hölje kostar ca 50 kr/m medan själva glasfibern bara kostar ca 1 kr/m. Att lägga extra fibrer när man ändå höll på innebar ingen större investering. Därför finns i dag optokabel med allt från 12 fibrer, 24, 48, 96, 200, ja i stadsnätet förekommer t.o.m. 1000 fibrer i samma kabel.

WDM

Men ljuskällorna och optodetektorerna har utvecklats, liksom filtertekniken. I dag klarar man att mata in ljus från flera laserdioder på samma optofiber, laserdioder som alla

ligger på olika frekvenser (våglängd). I mottagarändan använder man bandpassfilter och filtrerar ut varje våglängd för sig och detekterar.

Tekniken kallas våglängdsmultiplexering, WDM (Wavelength Division Multiplexing), man sänder flera bärvågor (4 nm mellan varje bärvåg), och klarar att bygga system med 32 bärvågor, där varje bärvåg transporterar 2,5 Gbit/s, alltså 80 Gbit/s på en enda optofiber.

Med WDM-tekniken är det intressant att ha minimal dispersion i hela området 1500 – 1625 nm. Detta är inte möjligt med dispersionsskiftning.

Ytterligare ett problem när man har många bärvågor är att effekttätheten ökar i glaset. Det finns en övre gräns när glas-molekylerna reagerar för ljuset. Optofibern blir olinjär vilket leder till att det blir överhörning mellan de olika bärvågorna, intermodulation. Därför vill man öka fiberkärnans diameter. Men då börjar ljusstrålarna gå i flera moder vilket ger moddispersion.

Genom "smart" förändring av brytningsindex pågår utveckling där man försöker tvinga ljuset ut mot kanterna av kärnan. Man får "ett slags rör" där ljuset går i rörets väggar. På detta sätt har man lyckats öka effektiva tvärsnittsytan till nästan det dubbla utan att moddispersion uppstår.

Kanalisation

Det är lättast att hantera många olika förbindelser om varje förbindelse ligger i en egen optofiber. Med WDM-tekniken klarar man sig med färre fibrer men varje förbindelse kan få en egen våglängd.

Optokablar med många fibrer kan inte göras alltför långa. Fibrerna måste skarvas. Man får dämpning. I dag när man genom WDM-tekniken klarar sig med några få fibrer så gräver man ner kanalisation, plaströr. Sedan blåser man in själva glasfibrerna med tryckluft. Dessa fibrer kan vara mycket långa. Man klarar sig med färre skarvar.

17 — En lek med bilder

Många arbetsmoment i en ingenjörns liv kräver överblick, intuition och fantasi, färdigheter som ligger i den del av hjärnan där vi tolkar bilder, inte i den logisk-matematiska delen av hjärnan.

Många personer, liksom jag själv, har svårt att genomföra matematiska härledningar utan en bild av vad vi gör. Bilden kan vara ytterst förenklad, men den behövs som stöd "för att hålla reda på vad vi gör".

Om du fortsätter sina studier inom radio-kommunikationsområdet och går djupare in i elektrodynamiken, så kommer du att upptäcka en helt igenom matematisk värld som utgår från Maxwells ekvationer och därefter styrs av vektoralgebrans räkneregler. Detta är nödvändigt för att under kontrollerbara former kunna hålla reda på alla dessa fältstyrkor som bildas kors och tvärs i helt nya riktningar när man deriverar.

För att ändå ge ett stöd om du gärna vill se bilder, så bjuder jag på följande bildsvit där det mesta finns med.

För min del började det hela med "masken". Jag stod framme vid tavlan och berättade, och plötsligt undrade jag vad det egentligen är jag säger?

Så gott som alla bilder är nya, åtminstone för mig. När signalen strålar in i ledaren och ger värmeutveckling, sid 292, är hämtad från

Feynman: "Lectures on physics" del 2. Även uppgifterna om elektrostatiska fältet runt jordklotet kommer från samma källa.

En annan källa som innehåller allmän diskussion om energitransport på ledningar är Jordan, Balmain: "Electromagnetic waves and radiating systems".

Bilderna kan innehålla felaktigheter. Det kan finnas bilder som är ritade så att de misstolkas. Min egen tolkning är emellertid att jag inte har hittat något i bilderna som motsäger dagens matematiska beskrivning av elektrodynamiken. Däremot påstår jag att kondensatorns- och spolens uppladdningskurvor är trappstegsformade, men det har jag verifierat med oscilloskop.

Tänk nu bara på följande: Bilder ger idéer. Bilder spelar en framträdande roll vid tekniska framsteg. Men det är med matematiken som man kan visa att det finns en logisk väg utan luckor som når till samma mål.

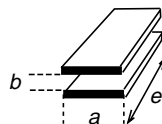
Einstein satt på solstrålen och åkte med. Men hur var det med Maxwell? Använde han bilder?

Om Maxwell berättas att han var snillet som blev mobbad i skolan av sina klasskamrater. En gång, när läraren fyllt tavlan tre gånger för att lösa och förklara ett problem, gick Maxwell fram och gav svaret med en enkel figur och några få rader. Efter det var han inte populär bland lärarna heller.

17.1 Kondensatorn

Plattkondensatorn

Utgå från en plattkondensator. Då är det enkelt att beräkna kapacitansen. Ytan $a \cdot e$ multiplicerat med kapacitiviteten ε dividerat med avståndet mellan plattorna b .

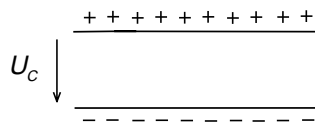
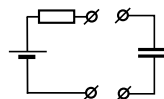


$$C = \varepsilon \cdot \frac{a \cdot e}{b} = \varepsilon_0 \varepsilon_r \cdot \frac{a \cdot e}{b}$$

Ladda upp kondensatorn

Anslut ett batteri till kondensatorn. Vi kan ha motstånd i serie. Det gör inget. Vänta med att ta bort batteriet tills kondensatorn är fulladdad.

På plattorna finns nu laddning som enkelt kan beräknas eftersom vi känner kapacitansen och spänningen.



$$U_C = \frac{Q}{C}$$

Beräkna energin i kondensatorn

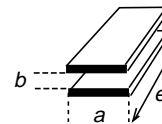
Nu skall vi beräkna den energi som lagrats i kondensatorn. Vi använder den vanliga formeln och ersätter kapacitansen med plattkondensatorns fysiska mått.

$$W_C = \frac{1}{2} \cdot C \cdot U_C^2 = \frac{1}{2} \cdot \varepsilon_0 \varepsilon_r \cdot \frac{a \cdot e}{b} \cdot U_C^2$$

17.2 Kondensatorn som transmissionsledning

Plattledningen

Bilden av plattkondensatorn liknar bilden av en plattledning. Vi beräknar plattledningens karakteristiska impedans.

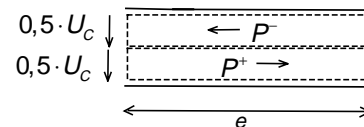
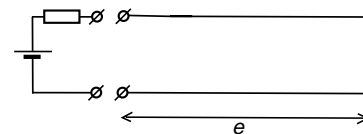


$$Z_0 = \frac{b}{a} \sqrt{\frac{\mu}{\epsilon}}$$

"Masken"

Nu ansluter vi batteriet för att "ladda upp kondensatorn". Vad har vi? "Masken" som vi studerade i kapitel 12.

Vi får framåtgående spänning som ger upphov till framåtgående effekt. Och vi har reflekterad spänning som ger upphov till reflekterad effekt.



$$P^+ = U^+ \cdot I^+ = \frac{(U^+)^2}{Z_0}$$

$$P^- = U^- \cdot I^- = \frac{(U^-)^2}{Z_0}$$

Totala energin i ledningen

Total energi är effekt multiplicerat med tid. Vilken tid skall vi använda? Jo, den tid det tar att gå från ena ändan till andra ändan, d.v.s. den tid det tar att gå sträckan e med ljushastigheten. Och detta gör vi för båda effekterna.

Vad blir resultatet? Exakt detsamma som för plattkondensatorn.

$$\begin{cases} W = P \cdot t = P^+ \cdot \frac{e}{c} + P^- \cdot \frac{e}{c} \\ U^+ = U^- = 0,5 \cdot U_c \end{cases}$$

$$W = 2 \cdot \frac{(0,5 \cdot U_c)^2}{Z_0} \cdot \frac{e}{c} = 2 \cdot \frac{(0,5 \cdot U_c)^2}{\frac{b}{a} \sqrt{\frac{\mu}{\epsilon}}} \cdot \frac{e}{\frac{1}{\sqrt{\mu\epsilon}}} = \frac{1}{2} \cdot \epsilon_0 \epsilon_r \cdot \frac{a \cdot e}{b} \cdot (U_c)^2$$

17.3 Plattledningens ”uppladdning”

Nu ansluter vi plattledningen till en puls som stiger från 0 volt till U volt. Hur kommer spänningen att ändras, mätt i anslutningspunkten mellan generatoren och plattledningen?

Först spänningsdelning

När pulsen kommer sker spänningsdelning precis som för ”masken”. Denna spänning U_0 kryper ut på ledningen, totalreflekteras och kommer tillbaka efter den tid det tar att gå fram och tillbaka på ledningen.

Därefter reflexionskoefficient

När spänningen U_0 kommer tillbaka så träffar den resistansen som ligger i serie med generatoren. Vi kan beräkna reflexionskoefficienten.

Första trappsteget blir alltså lågt, det består bara av spänningsdelning, U_0 . Nästa trappsteg blir nästan dubbelt så högt. Sedan fortsätter det på samma sätt i oändligheten.

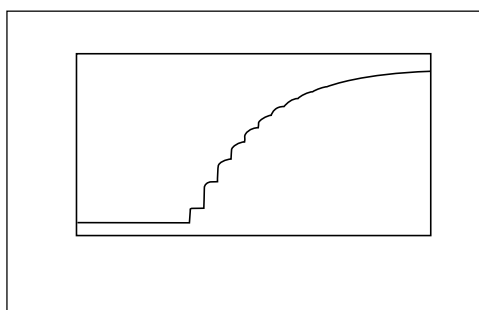
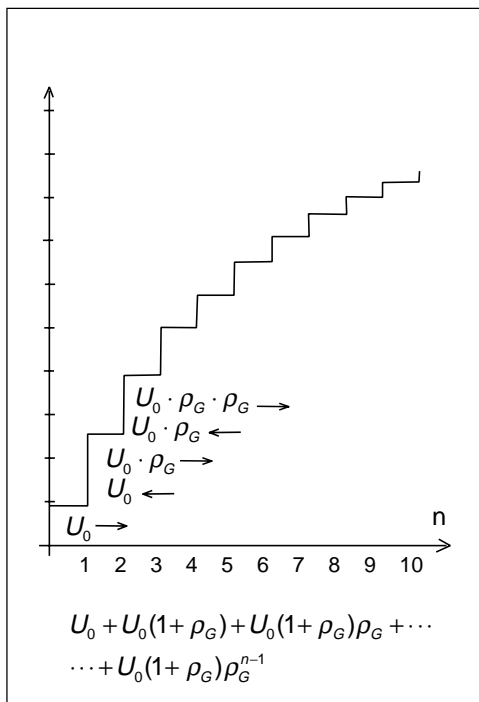
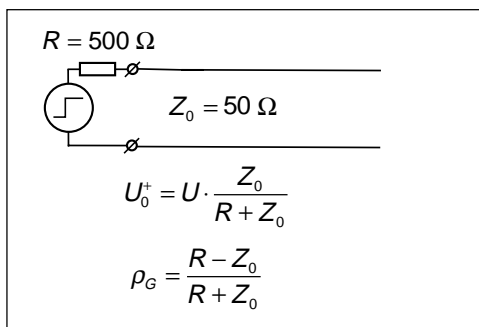
Serien konvergerar, och till slut är summan av alla framåtgående delspänningar lika med halva batterispänningen. På motsvarande sätt blir summan av alla reflekterade delspänningar andra halvan av batterispänningen.

Mätning med oscilloskop

Exakt det som här beskrivits har jag undersökt med oscilloskop. Resistansen R var 500Ω och ledningen var en 25 m lång 50Ω koaxialkabel.

Oscilloskopbildens har fotograferats och överförs till bild. Vi ser tydligt hur första trappsteget är ungefär hälften så högt som andra trappsteget.

Trappstegen blir mer och mer avrundade beroende på att pulsens högfrekventa komponenter dämpas mer än de lågfrekventa.



17.4 Kondensatorns uppladdning

Nu gör vi på samma sätt med kondensatorn. Vanlig ellära.

Problemet är e-funktionens exponent. Vi måste skriva uttrycket på ett sätt så att resultatet kan jämföras med föregående beräkning.

Jämför plattledningens kapacitans med dess karakteristiska impedans. Vi ser att kapacitansen kan skrivas som en funktion av ledningens karakteristiska impedans och den tid det tar för signalen att gå längs ledningens längd e .

Detta är ett generellt samband för ledningar. En 50Ω koaxialkabel där signalen går med $200\,000$ km/s har kapacitansen 100 pF/m.

Genom att stoppa in detta uttryck i e-funktionens exponent så kan vi uttrycka tiden t i antal gånger (n) som signalen rusat fram och tillbaka på ledningen ($2 \cdot \Delta t$).

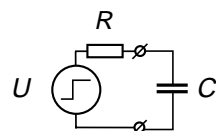
Nu kan vi rita upp den välkända kurvan över kondensatorns uppladdning och ser att kurvan följer trappstegskurvan som vi fick när vi betraktade kondensatorn som en ledning.

Ingen kondensator är väl 25 m lång?

Om kondensatorn inte är 25 m lång utan betydligt kortare så kommer trappstegen att bli mindre, och vi får bättre överensstämmelse med e-funktionen.

Elektrolytkondensatorn är exempel på en kondensator som kan vara mycket lång. Mekaniskt består kondensatorn av två metallfolier uppåt en meter långa som rullats ihop. Mellan folierna finns en elektrolyt med extremt hög kapacitivitet, som gör att signalen går mycket långsamt, folierna ser ut att vara långa.

Att betrakta kondensatorn som en behållare för laddningar är 1700-talssynsätt. När man utgående från detta synsätt integrerar fram spänningskurvan för kondensatorns uppladdning så är detta visserligen ett elegant exempel på integrering, men bara en grov beskrivning av verkligheten, som tycks vara trappstegsformad.

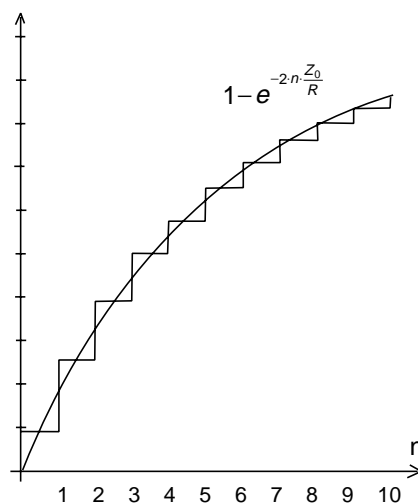


$$U_C = U(1 - e^{-t/RC})$$

$$\begin{cases} C = \epsilon \cdot \frac{a \cdot e}{b} \\ Z_0 = \frac{b}{a} \sqrt{\frac{\mu}{\epsilon}} \end{cases}$$

$$C = \epsilon \cdot \frac{e \sqrt{\frac{\mu}{\epsilon}}}{\frac{b}{a} \sqrt{\frac{\mu}{\epsilon}}} = \frac{e \sqrt{\mu \epsilon}}{Z_0} = \frac{e}{c} \cdot \frac{1}{Z_0} = \frac{\Delta t}{Z_0}$$

$$\frac{t}{RC} = \frac{t}{\Delta t} \cdot \frac{Z_0}{R} = 2 \cdot n \cdot \frac{Z_0}{R}$$



17.5 Att parallellkoppla två kondensatorer

Vi utgår från två identiska kondensatorer. Den ena laddas upp till spänningen U och får laddningen Q . Den andra kondensatorn har ingen spänning/laddning. Totala energin utgör energin i den uppladdade kondensatorn.

Nu parallellkopplar vi kondensatorerna. Det klassiska synsättet är att nu fördelar sig laddningarna på de båda kondensatorerna. Laddning kan inte försvinna.

Båda kondensatorerna får halva laddningen, vilket ger halva spänningen. Inget konstigt så här långt.

Nu beräknar vi totala energin som summan av energierna i de två kondensatorerna och upptäcker att totala energin nu bara utgör hälften av vad vi hade före parallellkopplingen. Vart har energin tagit vägen? Värme? Radiostrålning? Vi räknar på ett förlustfritt fall och kondensatorerna är utformade så att vi inte får radiostrålning.

Kondensatorerna som ledningar

Nu övergår vi till ledningar. Den ena ledningen laddas med framåtgående och reflekterad effekt så att totala spänningen blir U .

Vad händer när vi kopplar ihop ledningarna, d.v.s. parallellkopplar kondensatorerna? Jo, "masken" kryper in i nästa ledning. Den spänning vi mäter i sammankopplingspunkten är antingen spänningen när masken kryper till höger, eller spänningen när masken kryper till vänster. I båda fallen är detta halva spänningen.

Hur går det med energin? Totala energin är storleken på masken (effektens storlek) multiplicerat med maskens längd (egentligen den tid det tar för masken att krypa förbi). Inget av detta har förändrats. Totala energin är oförändrad.

I verkligheten

I verkligheten måste vi räkna med multipelreflexer i sammankopplingspunkten. Då leder detta till att vi till slut får en enda dubbelt

$$Q = U \cdot C \quad Q = 0$$

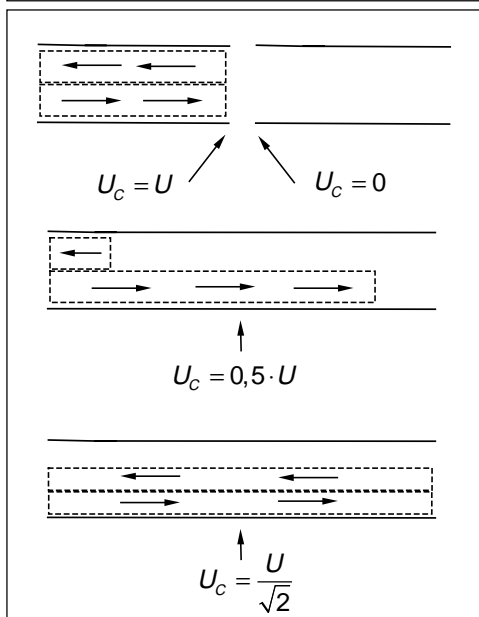
$$U_c = U \quad U_c = 0$$

$$W = \frac{1}{2} C \cdot U^2 + \frac{1}{2} C \cdot 0^2 = \frac{1}{2} C \cdot U^2$$

$$0,5 \cdot Q = 0,5 \cdot U \cdot C \quad 0,5 \cdot Q = 0,5 \cdot U \cdot C$$

$$U_c = 0,5 \cdot U \quad U_c = 0,5 \cdot U$$

$$W = \frac{1}{2} C \cdot \left(\frac{U}{2}\right)^2 + \frac{1}{2} C \cdot \left(\frac{U}{2}\right)^2 = \frac{1}{4} C \cdot U^2$$



så lång mask där totala spänningen blir $U/\sqrt{2}$. Men vi kommer aldrig att kunna mäta detta. Så snart som framåtgående och reflekterad mask i ett tvärsnitt inte är lika stora så flyter ström i kondensatorblecken. Vi får uppvärmning, resistiva förluster.

17.6 Spolen som ledning

Nu kortsluter vi plattledningen så att den bildar en spole på ett varv. Induktansen är lätt att beräkna. Även här ser vi att induktansen är proportionell mot karakteristiska impedansen och den tid det tar för signalen att gå sträckan e .

På vanligt sätt beräknar vi totala energin som är lagrad i spolens magnetfält.

Spolen som ledning

På den kortslutna ledningen har vi framåtgående signal och reflekterad signal på samma sätt som vid kondensatorn, med skillnaden att den reflekterade signalen fasvänds 180° .

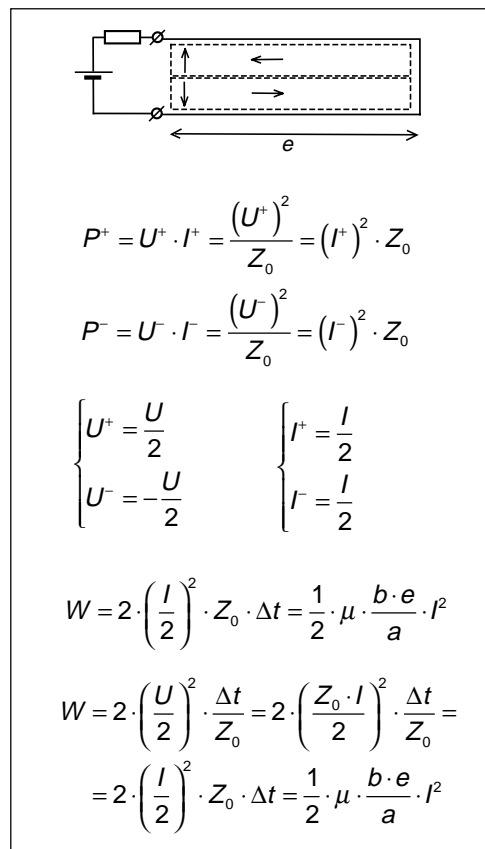
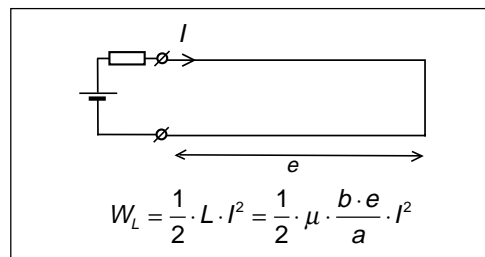
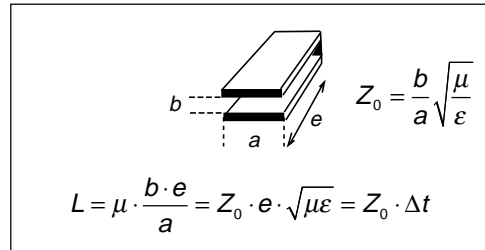
I framåtgående signalen har vi halva strömmen, och i den reflekterade signalen den andra halvan av strömmen. Totala strömmen blir summan av strömmarna.

I framåtgående signalen har vi halva spänningen och i den reflekterade signalen den andra halvan av spänningen, fast 180° fasviden. Summan blir noll. Vi ser ingen spänning. Men spänningarna finns där. Vi kan alltså lika gärna beräkna energin i spolens magnetfält genom att utgå från spänningarna.

$$\begin{cases} W = 2 \cdot \left(\frac{U}{2}\right)^2 \cdot \frac{\Delta t}{Z_0} \\ \frac{U}{I} = Z_0 \end{cases}$$

Även spolen laddas upp trappstegsformat

På samma sätt som att spänningen över kondensatorn stiger trappstegsformat, så kommer strömmen att växa trappstegsformat när man lägger spänning över spolen. Den vackra e -funktionen är en matematisk approximation.



17.7 Koppling mellan lindningsvarven

Vi utgår från samma spole som på föregående sida, men efter halva längden så vrider vi slingan ett halvt varv. Detta får till följd att den framåtgående signalen vrider sitt elektriska fält så att det blir riktat uppåt. På samma sätt får den reflekterade signalen ett elektriskt fält som är riktat åt motsatt håll.

Att vrida spolen på detta sätt förändrar inget. Vi har fortfarande samma totala energi i spolen.

Nu viker vi tillbaka yttre delen av spolen och får på detta sätt en spole med två lindningsvarv, men halva längden. Vad händer nu med energin i spolen?

Spänningarna adderas

Det som händer är att framåtgående signalen i varv 1 och reflekterade signalen i varv 2 har elektriska fält i samma riktning (nedåt). Dessutom rör sig signalerna åt samma håll. Då ligger även H-fälten i samma riktning så att de samverkar. Då skall signalerna adderas d.v.s. spänningen blir den dubbla.

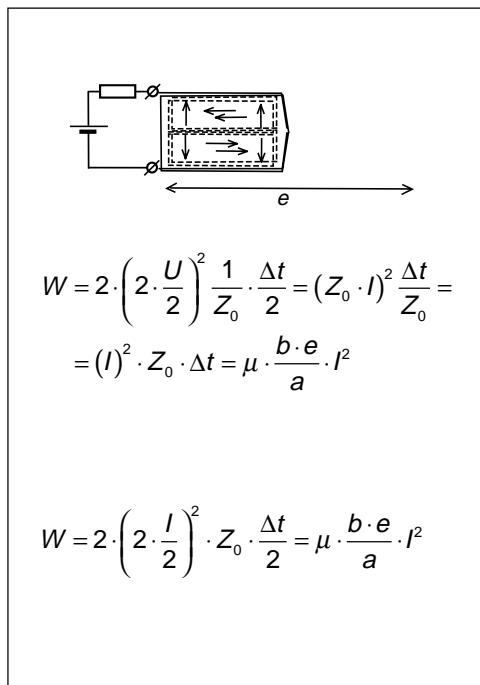
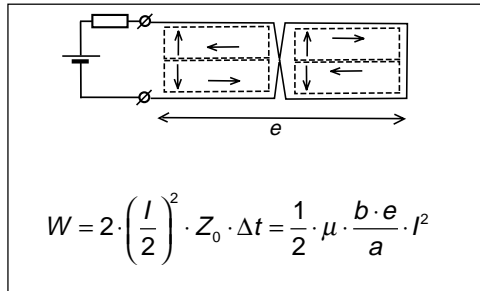
Samma sak händer med den reflekterade signalen i varv 1 och den framåtgående signalen i varv 2. Även de samverkar.

När vi nu beräknar totala energin i denna spole på två varv så upptäcker vi att energin är dubbelt så stor, trots att totala längden inte förändrats. Vi har bara "snott ihop" den ursprungliga spolen till två varv.

Självklart syns detta på spolens induktans som blivit dubbelt så stor. Vi säger att de två lindningsvarven "kopplar" till varandra.

Mycket viktigt

Detta exempel visar något principiellt mycket viktigt. Signaler som går åt olika håll behåller sin individualitet. Vi beräknar effekten i varje signal för sig. Men om signalerna går åt samma håll så skall E -fälten och H -fälten adderas, och av två lika starka signaler får vi en signal med fyra gånger så hög effekt. Mer om detta längre fram.

**Fyll utrymmet med järn**

Utrymmet i spolen kan fyllas med järn, transformatorplåt med högt μ . Transformatorplåten är visserligen elektriskt ledande, men det gör inget eftersom totala elektriska fältet från de båda signalerna är noll.

Transformatorplåt med högt μ medför att ljushastigheten sjunker. Signalerna går extremt långsamt. Det tar lång tid att gå från ena ändan till den andra. Δt blir stort, och spolen lagrar mycket energi.

17.8 E och D, H och B

E och H är "signalen"

E är elektriska fältstyrkan och H är elektriska fältstyrkan i signalen som har effekttätheten S . Förhållandet mellan E och H bestäms av mediet, mediets karakteristiska impedans.

E och H uppträder ALLTID parvis och vinkelrätt mot varandra. När vi säger att vi bara har elektrisk fältstyrka, ingen magnetisk fältstyrka, så innebär detta ALLTID att vi har två eller flera signaler som går åt olika håll, så att summan av deras elektriska fältstyrkor blir hög, medan summan av de magnetiska fältstyrkorna blir låg.

E och H är "tvådimensionell"

E och H finns i ett plan vinkelrätt mot signalens utbredningsriktning, bara i detta plan. Fältstyrkorna beskriver hur mycket effekt som tar sig igenom planet (per ytenhet). Men är det hög energi som går långsamt eller låg energi som går snabbt?

Om vi känner både E och H så kan vi via den karakteristiska impedansen ta oss fram till signalens hastighet. Men om vi bara känner en av fältstyrkorna, som i en kondensator?

E över en "yta"

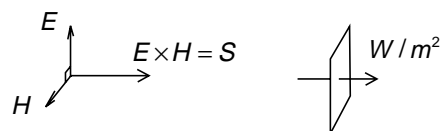
När vi beräknar totala laddningen Q på kondensatorblecket så räcker det inte att känna den elektriska fältstyrkan mellan kondensatorplattorna. Vi måste även på något sätt ta hänsyn till signalernas hastighet, d.v.s. hur tätt fältlinjerna hamnar.

D, elektrisk flödestäthet

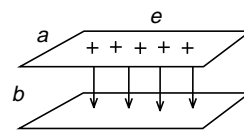
Den elektriska flödestätheten D beror på summan av de elektriska fältstyrkorna i signalerna, och ϵ_r . Signalernas hastighet påverkas ju av $\sqrt{\mu_r \epsilon_r}$, men vid beräkning av D tar vi alltså bara hänsyn till den ena delen i denna faktor.

B, magnetisk flödestäthet

På samma sätt när vi behöver summera totala antalet magnetiska fältlinjer genom en yta, då är det den magnetiska flödestätheten B som skall användas.



E och H är bara definierade i ett "plan" vinkelrätt mot signalen, tvådimensionellt.



$$\begin{cases} D = \epsilon E \\ Q = D \cdot a \cdot e \end{cases}$$

$$Q = C \cdot U_C = \epsilon_0 \epsilon_r \cdot \frac{a \cdot e}{b} \cdot U_C =$$

$$= \epsilon_0 \frac{U_C}{b} a \sqrt{\epsilon_r} \cdot e \sqrt{\epsilon_r} = \epsilon_0 E \cdot a \sqrt{\epsilon_r} \cdot e \sqrt{\epsilon_r}$$

Vi tycks ha en större "dielektrisk yta".



$$\begin{cases} B = \mu H \\ \Phi = B \cdot e \cdot b \end{cases}$$

$$\begin{aligned} \Phi &= D \cdot e \cdot b = \mu H \cdot e \cdot b = \\ &= \mu_0 H \cdot e \sqrt{\mu_r} \cdot b \sqrt{\mu_r} \end{aligned}$$

Hur stor "magnetisk yta" har spolen?

17.9 Maxwells ekvationer

Året är 1785. Människan har upptäckt elektrisk laddning. Hon kan ladda upp metallklot. Hon kan få laddning att röra sig i en metalltråd som ansluts till ett batteri.

Även magnetismen är välbekant i form av det jordmagnetiska fältet, permanentmagneter och kompassnålar.

Men ingen ser ännu något samband mellan elektricitet och magnetism. Detta betraktas som två helt olika fenomen.

Coulomb

Detta år (1785) upptäckte den franske fysikern C. de Coulomb (1736 – 1806) sambandet för dragningskraften F mellan elektriskt laddade klot. Utgående från detta härleder vi kraften på ett av kloten som befinner sig i det elektriska fältet från det andra klotet.

Genom att integrera den elektriska flödestätheten D runt en sluten yta a så får vi den totala laddningen som finns i den av ytan a inneslutna volymen v , ett samband som kallas Gauss lag.

Tidigare beskrevs detta med "divergensen" för D , d.v.s. skillnaden mellan elektriskt flöde in i volymen och ut ur volymen ger den inneslutna laddningen. "Fältlinjer slutar på laddningar". Numera utnyttjar man räkneregler för ∇ -operatoren vid motsvarande beräkningar.

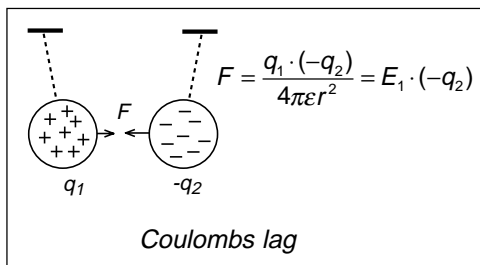
Ørsted

År 1819 upptäckte dansken H.C. Ørsted (1777 – 1851) att en kompassnål visar fel, om den placeras nära en metalledare där det flyter laddning. Laddningarna måste påverka, eller ge upphov till magnetfält.

Nu hade man funnit något som kopplade ihop två fenomen som hittills betraktats som oberoende av varandra, elektricitet och magnetism.

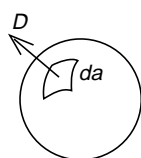
Biot och Savart

Året efter formulerar fransmännen J.B. Biot (1774 – 1862) och F. Savart (1791 – 1841) ett uttryck för magnetisk fältstyrka H från en strömförande ledare (Biot-Savart's lag).



$$\Rightarrow E = \frac{q_1}{4\pi\epsilon r^2}$$

$$\epsilon E \cdot 4\pi r^2 = q_1 \Rightarrow D \cdot 4\pi r^2 = q_1$$



$$\oint_a D \cdot da = \int_v \rho \cdot dv$$

$$\text{div } D = \rho$$

$$\nabla \cdot D = \rho$$

Gauss lag

F = kraften mellan laddningarna

q_1 och q_2

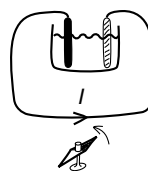
E = elektrisk fältstyrka från laddning q_1

D = elektrisk flödestäthet

r = avståndet mellan laddningarna

q_1 och q_2

a = ytan av den volym v som innesluter laddningstätheten ρ



Ørsted.

Ampère

Arbetet fortsattes av fransmannen A.M. Ampère (1775 – 1836), som summerade H -fältet vektoriellt i en slinga runt en strömförande ledare och visade att summan blir totala strömmen genom slingan. Jfr summering av E -fältet. Oberoende av vilken väg man väljer så blir vektoriella summan av E -fältet längs en sträcka lika med spänningen mellan ändpunkterna.

Vi ser hur totala E -fältet hos de framåtgående och reflekterade effekterna i slingan ger upphov till spänningen mellan ledarna, och hur H -fältet ger upphov till strömmen i ledarna.

Vidare måste vi veta hur fort effekterna rör sig för att kunna beräkna magnetiska flödestätheten B genom slingan, och totala flödet Φ genom slingan. Motsvarande gäller när vi skall beräkna den elektriska flödestätheten D och totala elektriska flödet.

Faraday

År 1831 demonstrerade engelsmannen M. Faraday (1791 - 1867) för sina elever, han var lärare, hur ström ger magnetfält, han upprepar Ørstedts försök. Men han visade även att magnetfältet från en permanentmagnet varken ger ström eller spänning.

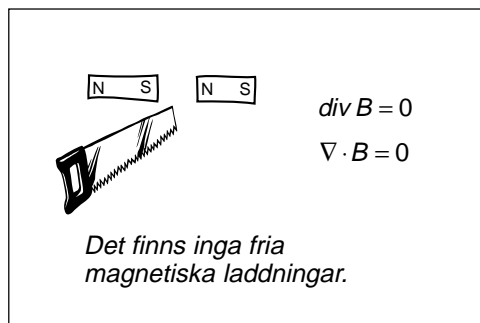
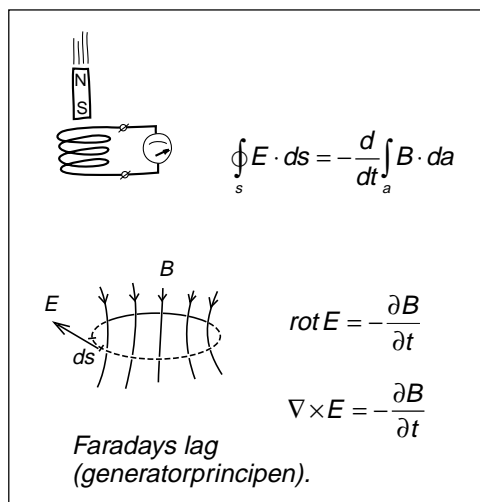
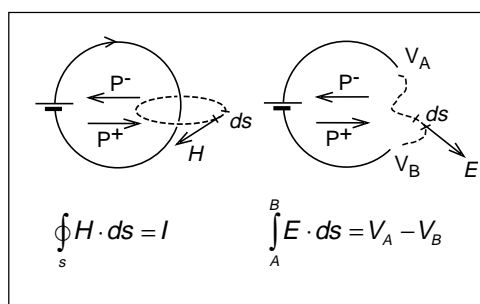
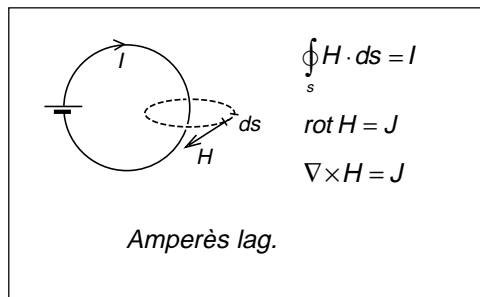
Faraday hade en spole och över spolen en voltmeter. Han håller permanentmagneten vid spolen för att visa att voltmeteren inte gör utslag.

Vad händer? Jo, han råkar tappa magneten. Och se! Nu får han utslag på voltmeteren. Historien är full av fumliga fysiker som upptäckt det mest fantastiska när de snubblat eller tappat diverse prylar.

Vad Faraday upptäckte var elektrisk induktion, att det alstras en elektrisk fältstyrka som ger spänning i ledaren, om magnetfältet runt ledaren ändras (generatorprincipen).

Gauss utvecklade matematiken

Under åren 1830 – 1840 sysslade den tyske matematikern J.C.F. Gauss (1777 – 1855, känd från statistiken) med elektromagnetism. Han utvecklade matematiken och gjorde det möjligt att matematiskt formulera att man inte kan sära på magnetens nord- och sydpoler.



Maxwell

Vid 1800-talets mitt fanns många som forskade i elektromagnetismen. En av dessa var skotten J.C. Maxwell (1831 – 1879) som utgick från de resultat som utarbetats av Coulomb, Ampère och Faraday. Låt oss titta på dessa ekvationer så som vi i dag är vana att se dem, skrivna med operatorkalkyl utvecklad av den självlärde engelsmannen O Heaviside (1850 – 1925):

- Ekvation 1 är en omskrivning av Coulombs dragningskraft mellan elektriska laddningar. Elektriska nettoflödet ut ur en sluten volym beror på nettoladdningen i volymen. Elektriska fältlinjer startar eller slutar på en laddning.
- Ekvation 2 är motsvarande formulering för det magnetiska fältet. Man kan inte sära på nord- och sydpoler, man kan inte få fristående nord- eller sydpoler i en sluten volym. Delar man en magnet bildas två nya magneter som har både nord- och sydpoler. Matematiskt beskrivs detta som att vi inte kan ha något nettomagnetflöde ut ur en sluten volym. Det går alltid lika många magnetiska fältlinjer in i volymen som ut ur volymen.
- Ekvation 3 är Faradays lag om induktion. Elektriska fältstyrkan och därmed även spänningen beror på ändringar i det magnetiska flödet. Kryssprodukten beskriver att fälten är riktade vinkelrätt mot varandra enligt "skruvregeln".
- Ekvation 4 är Ampères lag. Summan av magnetiska fältstyrkan runt en ledare ger strömmen i ledaren, och kryssprodukten beskriver att riktningen på magnetiska fältstyrkan och strömmen bestäms av skruvregeln.

$$\nabla \cdot D = \rho$$

$$\nabla \cdot B = 0$$

$$\nabla \times E = -\frac{\partial B}{\partial t}$$

$$\nabla \times H = J$$

Men derivatan av det elektriska flödet?

Maxwell tyckte att något fattades. I Faradays lag beskrivs vad som händer om det magnetiska flödet ändras. Varför finns ingen ekvation som beskriver vad som händer om det elektriska flödet ändras? För visst borde väl något hända?

I Faradays lag visas hur derivatan av det magnetiska flödet ger elektrisk fältstyrka. Varför skulle inte, med samma resonemang, derivatan av det elektriska flödet kunna ge magnetisk fältstyrka?

Om trådslungan i Ampères lag förses med en kondensator kommer ingen ström (inga laddningar) att kunna flyta mellan kondensatorplattorna. Men här finns ett elektriskt flöde. Skulle inte förändringar i detta elektriska flöde kunna påverka magnetiska fältstyrkan på samma sätt som strömmen?

Maxwell införde elektriska flödestätheten D (som kallades "förskjutningen"), betraktade derivatan av D som en förskjutningsström, och lät denna förskjutningsström ingå i Ampères lag.

Om man lägger integrationsslingan av H runt kondensatorn så får man förskjutningsströmmen som resultat. Om man lägger integrationsslingan av H runt själva ledaren så får man strömmen I som resultat.

Maxwell införde förskjutningsströmmen i Ampères lag helt på ovetenskapliga grunder. Han hade inga mätningar som stöd för sitt påstående. Det gick heller inte att härleda denna "ström" ur andra accepterade fysikaliska lagar. Han bara "tyckte" att detta fattades. Hans fysikaliska känsla sade honom att om naturen bär sig åt på ett visst sätt när magnetiska flödet ändras, varför skulle det inte kunna vara möjligt att naturen även bär sig åt på motsvarande sätt när det elektriska flödet ändras?

Strålar energi ut?

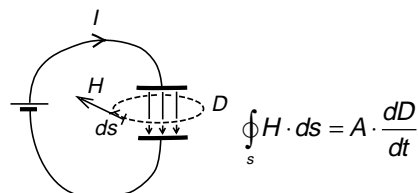
Men Maxwell nöjde sig inte med att bara komplettera Ampères lag. Han tillämpade ekvationerna på den elektriska pendeln och resultatet blev mycket märkligt.

En mekanisk pendel i vakuum som är upphängd utan friktion kommer att pendla i det oändliga, trodde man. Man har växling mellan läges- och rörelseenergi. Ingen energi kan försvinna. Pendlingen kommer

$$I = \frac{dq}{dt} = C \frac{dV}{dt}$$

$$C = \frac{A\epsilon}{b} \Rightarrow I = \frac{A\epsilon}{b} \cdot \frac{dV}{dt} = A\epsilon \cdot \frac{d}{dt} \frac{V}{b} = A\epsilon \cdot \frac{dE}{dt} = A \cdot \frac{dD}{dt}$$

"Förskjutningsströmmen"



Integrering runt kondensatorn ger förskjutningsströmmen.

$$\oint_s H \cdot ds = I + A \cdot \frac{dD}{dt}$$

$$\text{rot } H = J + \frac{\partial D}{\partial t}$$

$$\nabla \times H = J + \frac{\partial D}{\partial t}$$

Maxwells komplettering av Ampères lag.

aldrig att avstanna.

Men när Maxwell tillämpade ekvationerna på en elektrisk pendel i vakuum, med en ledare utan resistans, vad blev resultatet?

Han fick växling mellan lägesenergi och rörelseenergi, elektriskt flöde respektive magnetiskt flöde, precis som hos den mekaniska pendeln. Men dessutom försvann en liten del av energin, den strålade ut som en samtidig svängning av elektrisk och magnetisk fältstyrka. Inte antingen eller, utan en

samverkad svängning som ger energitransport. Maxwell kallade denna energi för elektromagnetiska vågor.

Energi försvann från pendeln! Pendlingen kommer enligt Maxwell att avklinga mot noll! Detta var något på den tiden mycket sensationellt.

Är ljuset elektromagnetiska vågor?

Maxwell grundade sin fysikaliska känsla på jämförelser. Han räknade ut att dessa elektromagnetiska vågor är energi som breder ut sig med ljushastigheten. Men ljus är ju energi som kommer från solen, eller från ljuskällan. Skulle inte ljus kunna vara elektromagnetiska vågor? Och säkerligen finns elektromagnetiska vågor av annat slag som man ännu inte kände till?

Många tvivlade

Maxwell beskrev sina teorier för Vetenskapsakademien i London år 1864. År 1873 publicerade han en bok i ämnet.

Vid Maxwells död 1879 tyckte den tyske professorn H.L.F. von Helmholtz (1821 – 1894) att nu fick det vara nog med svammel och ogrundade teorier. Maxwell hade splittat den akademiska världen i två läger, de som trodde och de som inte trodde. Och detta var farligt eftersom tro inte hör ihop med vetenskap. Därför förmådde Helmholtz Akademin i Berlin att utlysa ett pris till den som experimentellt kunde verifiera om Maxwell hade rätt eller fel. Helmholtz övertalade sin elev Hertz (1857 – 1894) att försöka vinna priset. Det lär finnas brev bevarade där Hertz gradvis, på ett försiktigt sätt, försöker klargöra för sin gamle professor att det nog ändå var Maxwell som hade rätt.

Energitransport i universum

På Maxwells tid kunde ingen förklara varför elektromagnetiska vågor bildas när energi pendlar fram och tillbaka på en elektrisk pendel. Senare verifierade man genom mätningar att elektromagnetiska vågor, radiovågor, bildas.

Vi kan räkna ut, med hjälp av matematik utvecklad långt före radiovågornas tid, att radiovågor bildas. Men varför bildas radiovågor?

Gravitationsvågor

Nu tror man (Einstein inspirerades av Maxwell) att det finns strålning även från en mekanisk pendel. Inte elektromagnetiska vågor, men väl gravitationsvågor.

Man har inte fått fram någon mätmetod för att påvisa gravitationsvågor. Visserligen skulle gravitationsvågorna kunna vara kraftiga, så kraftiga att väggarna i våra hus gungar. Men gravitationsvågorna påverkar även ljuset. Även ljusstrålarna "gungar". Därför kan vi inte se om väggarna rör sig.

Forskarna har letat i världsrymden och funnit en mekanisk pendel, två tunga himlakroppar, supernovor, som roterar runt varandra, och ändringen i rotationshastighet svarar mot den energiförlust man skulle få på grund av energitustrålning via gravitationsvågor, jfr Nobelpriset i fysik 1993.

17.10 Varifrån kommer magnetiska fältstyrkan?

När elektriska flödet ändras så bildas en magnetisk fältstyrka. Varifrån kommer magnetfältet?

Bild 1 visar en ledning som matas av ett batteri i serie med en resistans som är lika med ledningens karakteristiska impedans. Ledningen är avslutad med en lastresistans som även den har resistansen lika med ledningens karakteristiska impedans.

På ledningen finns en framåtgående signal från vänster till höger karakteriserad av elektrisk fältstyrka E och magnetisk fältstyrka H riktad in i papperet.

Bild 2 visar samma ledning matad av ett batteri från andra ändan. Vi har bara effekt som går från höger till vänster, karakteriserad av elektrisk fältstyrka E och magnetisk fältstyrka H riktad ut från papperet.

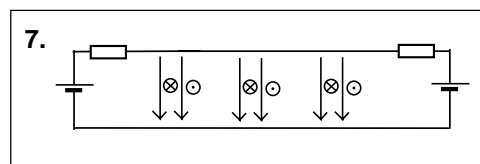
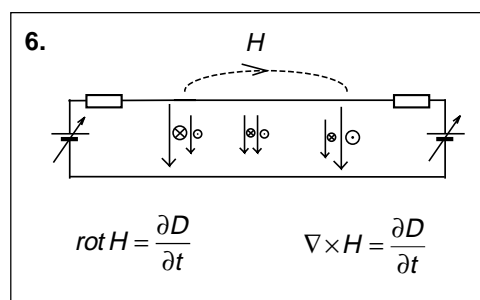
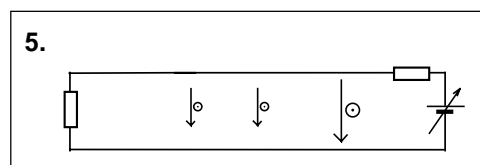
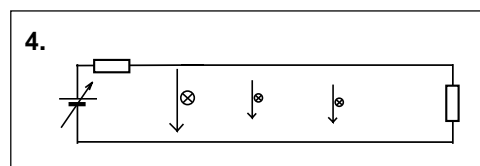
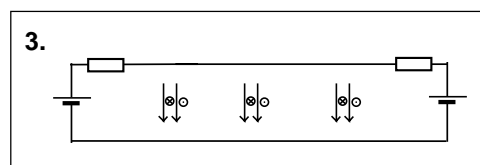
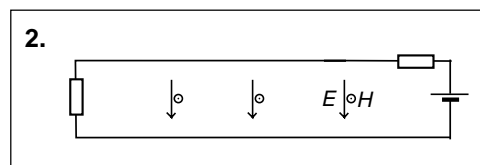
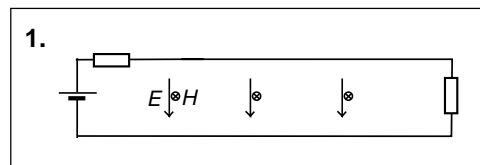
Bild 3 visar samma ledning med båda batterierna anslutna. Magnetfälten tar ut varandra. Vi ser bara elektrisk fältstyrka, elektrisk flödestäthet mellan ledarna.

I bild 4 ändrar vi den ena batterispänningen. Ändringen utbreder sig med ljushastigheten. I visst ögonblick har ändringen hunnit halvvägs ut på ledningen.

Bild 5 visar samma sak med det andra batteriet.

Bild 6 visar ledningen när båda batterierna ändras samtidigt. Vi ser hur det mitt på ledningen finns ett område där de magnetiska fältstyrkorna inte tar ut varandra. På höger sida får vi en resultant riktad ut från papperet, på vänster sida en resultant riktad in i papperet. Denna magnetiska fältstyrka roterar runt det område där elektriska flödet håller på att ändras ($\text{rot } H$).

Bild 7 visar ledningen efter ytterligare en tid, när förändringarna hunnit till andra ändan av ledningen.



17.11 Varifrån kommer elektriska fältstyrkan?

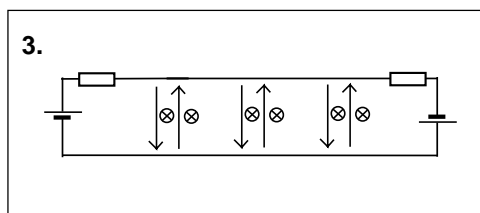
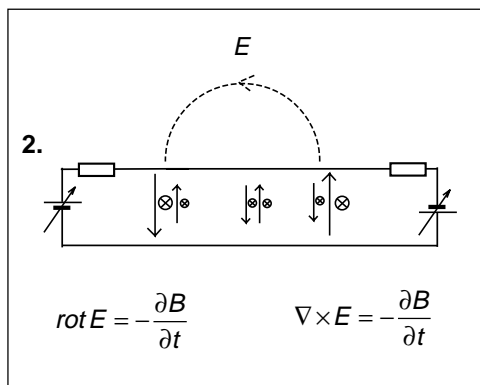
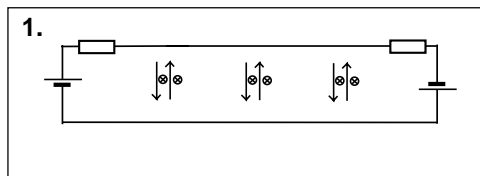
Vi skapar ett magnetiskt flöde på samma sätt som det elektriska flödet, genom att mata en ledning från två håll, men denna gång vänder vi det ena batteriet upp och ner.

De elektriska fältstyrkorna E tar ut varandra, medan magnetiska fältstyrkorna H samverkar in i papperet.

Nu höjer vi batterispänningarna samtidigt. Förändringen sprider sig ut på ledningen, och i ett visst ögonblick ser vi hur de elektriska fältstyrkorna inte tar ut varandra utan skapar en resultant som roterar ($\text{rot } E$) moturs runt det magnetiska flödet riktat in i papperet. Moturs rotation ger minustecknet.

Visst blir det elegant om vi kan acceptera att det inte finns något sådant som enbart elektrisk fältstyrka. Eller enbart magnetisk fältstyrka. Elektriska och magnetiska fältstyrkorna följs ALLTID åt, i form av effekt som förflyttar sig med ljushastigheten.

Men om det nu verkligen är på detta sätt, varför framgår inte detta av matematiken? Därför att så snart vi skriver upp ett matematiskt uttryck för elektriskt flöde i form av framåtgående och reflekterad effekt, så "försvinner" magnetiska fältstyrkorna. Termerna tar ut varandra. Det går alltså inte att räkna på detta sätt. Men man kan rita bilder!



17.12 Vad är en elektriskt laddad partikel?

Positiv laddning

Egentligen är det enkelt. Ta en uppladdad kondensator och dra ihop det kondensatorbleck som har positiv laddning till en punkt. Samtidigt låter vi det negativa kondensatorblecket expandera mot oändligheten. Där har vi en positiv laddning.

Direkt ur figuren inser vi riktigheten i Gauss lag.

Framåtgående och reflekterad signal

Hur går detta ihop med synsättet att kondensatorns laddning består av framåtgående och reflekterad effekt? Alldeles utmärkt. Ledningen har två ändrar där signalen reflekteras. Genom att "snurra" ihop ledningen så att de två ändarna möts så får vi två maskar som biter sig själva i svansen. Den ena snurrar medurs, den andra moturs.

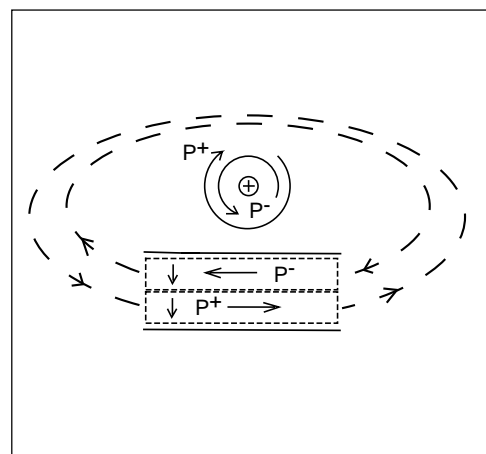
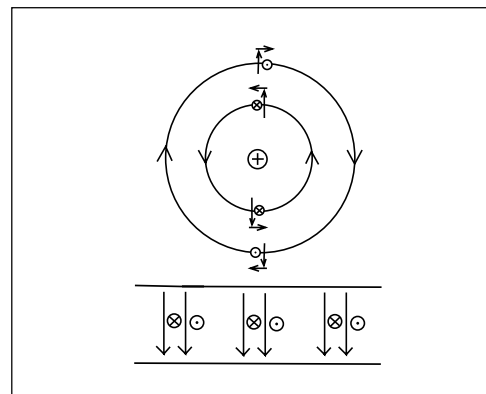
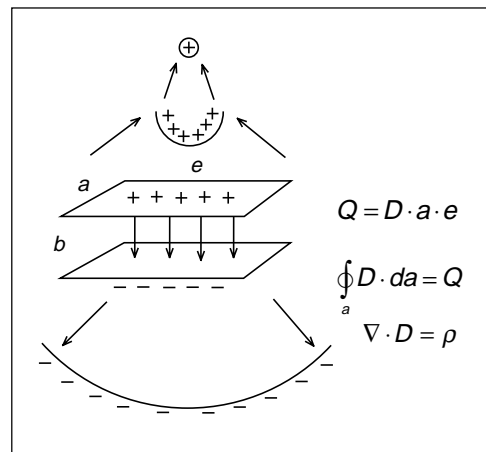
När de två signalerna snurrar runt den laddade partikeln, så kommer de elektriska fältstyrkorna att summeras radiellt ut från partikeln, medan de magnetiska fältstyrkorna ligger som motriktade koncentriska skal runt partikeln, tar ut varandra och inte syns.

Motriktade effekter som rör sig med ljushastigheten

Tänk dig dessa effekter som snurrar med ljushastigheten. Vad händer när vi kommer närmare partikeln? Effekterna hinner snurra fler varv per tidsenhet.

Ju längre in mot partikeln vi kommer, desto mindre blir omkretsen. Maskarna kryper upp på sig själva. När omkretsen minskat till hälften och varje delmask når två varv så ökar fältstyrkan till det dubbla. Men vad händer med energin? Jämför med fler varv i spolen. Den lagrade energin kommer att fyrdubblas. På detta sätt ökar energitätheten ju längre in mot partikeln vi kommer.

Men finns det någon partikel? Det kanske är detta som är partikeln. När energitätheten överskrider en viss gräns kanske vi uppfattar detta som en partikel?



17.13 Vad händer när det åskar?

Jordklotet är elektriskt laddat. Man har mätt en vertikal elektrisk fältstyrka på 100 V/m riktad mot marken, d.v.s. jordklotet är negativt laddat.

Runt jorden rusar framåtgående och reflekterad effekt, på samma sätt som när vi mätte på den slitsade ledningen. Skillnaden är att i den slitsade ledningen hade vi en växelspanningssignal, runt jorden har vi en likspänningssignal.

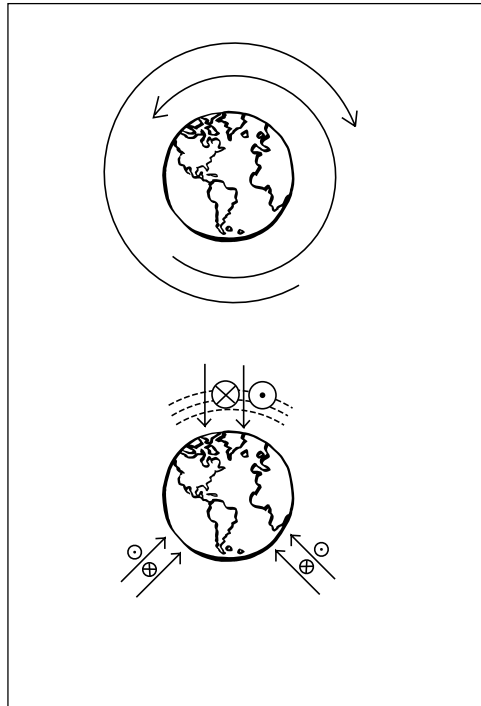
Om vi ansluter ett batteri till den slitsade ledningen så skapar vi en framåtgående och reflekterad likspänningssignal i ledningen. Men vi har ingen möjlighet att mäta på denna signal med en liten mätantenn eller med en riktkopplare. Mätanten- nen eller riktkopplaren fungerar bara för växelspanningssignaler.

Eftersom vi inte kan bygga någon sändarantenn för likspänningssignal så kan vi heller inte konstruera någon mottagarantenn. De likspänningssignaler som rusar runt jorden är helt enkelt svåra att mäta.

En möjlighet är att hissa upp en läckande plåthink med vatten. När plåthinken kommit upp i luften så har den fortfarande samma potential som nere vid marken. Men när vattendropparna läcker ut så tar de med sig laddning så att hinken, när vattnet är slut, har fått samma potential som omgivningen på den höjden. Sedan hissar man ner hinken och mäter spänningen mellan hinken och jord med en känslig elektrometer.

Regnväder innebär annan kapacitivitet.

I regnmolnet rusar signalerna genom ett dielektrikum. Detta innebär annat ϵ_r , annat förhållande mellan E och H . I detta fall skall E -fälten vara lika starka i och utanför regnmolnet. Det innebär att vi får kraftigare H -fält, högre effekttäthet S på signalen när den går i regnmolnet. Vad händer när det slutar regna? E -fältet stiger över en sträcka som motsvarar molnet. Vi får något som liknar kondensatorn som jag snodde ihop till en laddning. Samma sak händer med över-skottsenergin. Den snor ihop sig till en laddning, eller två laddningar, en plus- och en



minusladdning. Med lite fantasi så kan man nästan tänka sig att energin drar ihop sig till punktladdningar och kanske t.o.m. skapar ny materia.

17.14 Vad är en magnet för något?

Kortslut transmissionsledningen. Låt den framåtgående och reflekterade effekten fasvridas i ändarna och "studsas". De elektriska fältstyrkorna är motriktade, medan de magnetiska fältstyrkorna samverkar rakt igenom ledningen.

Magnetisk fältlinje kan inte "ta slut"

De magnetiska fältlinjerna slutar inte på någonting. De är slutna banor eller går från oändligheten till oändligheten. Det finns inget sätt på vilket vi kan dela ledningen så att vi får fältlinjen att sluta på något. Hur vi än lägger volymen över vars yta vi integrerar, så kommer varje fältlinje som går in i volymen också att komma ut på andra sidan. Detta beskriver vi på något av följande tre sätt:

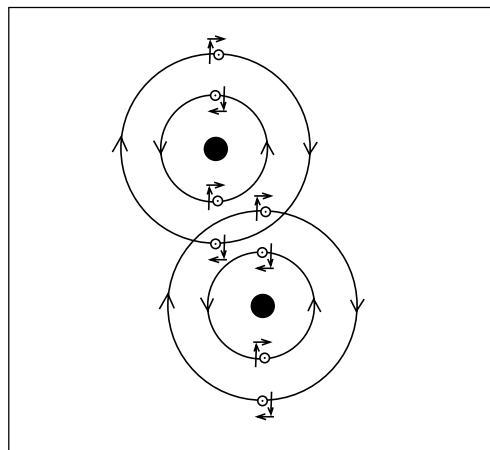
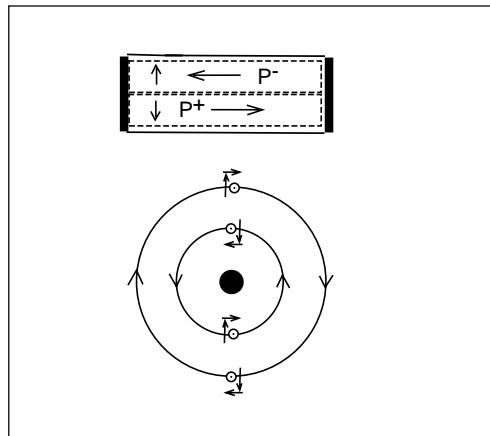
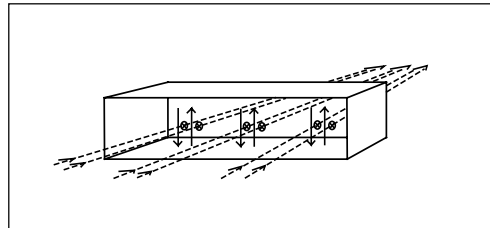
$$\oint_a \mathbf{B} \cdot d\mathbf{a} = 0$$

$$\operatorname{div} \mathbf{B} = 0$$

$$\nabla \cdot \mathbf{B} = 0$$

Det går att tänka sig en magnetisk partikel bestående av signaler som snurrar åt olika håll, på samma sätt som för den elektriskt laddade partikeln. Elektriska fältstyrkorna är överallt motriktade. Däremot kommer de magnetiska fältstyrkorna bara att samverka i en riktning, i bilden "ut från papperet". Den magnetiska partikeln är inte rotations-symmetrisk.

Om två magnetiska partiklar justerar in sig, sida vid sida som i den undre bilden, så kommer de magnetiska fältstyrkorna att samverka. Dessutom kommer signalerna som snurrar runt partiklarna att samverka i utrymmet mellan partiklarna, ett tillstånd där partiklarna kan dra nytta av varandras energier.



17.15 Varför attraheras partiklar med olika laddning?

Första bilden visar de cirkulerande signalerna runt en partikel som är positivt laddad. Jag har gjort en delförstoring som visar hur det ser ut på ena sidan av partikeln.

På visst avstånd är summan av de elektriska fältstyrkorna 10 V/m. På lite längre avstånd är de elektriska fältstyrkorna svagare, säg 8 V/m. Dessa fältstyrkor representerar effekten i signalerna som snurrar runt laddningen, och totala energin som är bunden till denna laddning antar vi helt rätt är proportionell mot kvadraten på dessa signaler.

$$W_1 = 10^2 + 8^2 = 164$$

Vi tar ytterligare en liten laddning med fältstyrkan 1 V/m. Den har ungefär energin 2.

$$W_{tot} = (10^2 + 8^2) + 2 = 166$$

För in en negativ laddning

I detta fall kommer fältstyrkorna att adderas närmast partikeln, medan fältstyrkorna subtraheras på andra sidan. Vi beräknar totala energin.

$$W_2 = (10 + 1)^2 + (8 - 1)^2 = 170$$

Vi ser att totala energin har ökat.

Addera en positiv laddning

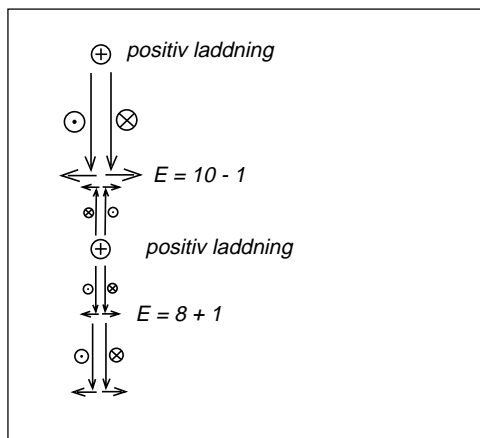
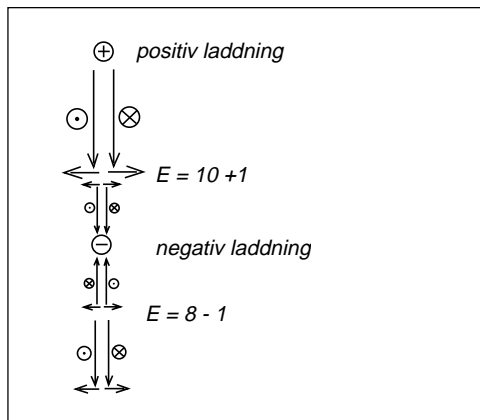
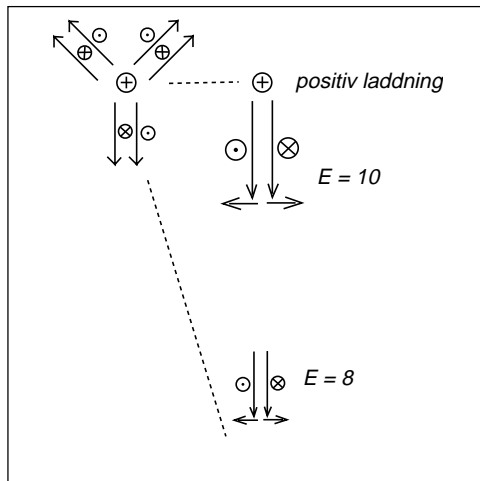
Nu gör vi samma sak med en positiv laddning och får svagare fältstyrka närmast partikeln men starkare fältstyrka längre ifrån.

$$W_3 = (10 - 1)^2 + (8 + 1)^2 = 162$$

Vi ser att totala energin blir mindre än i första exemplet.

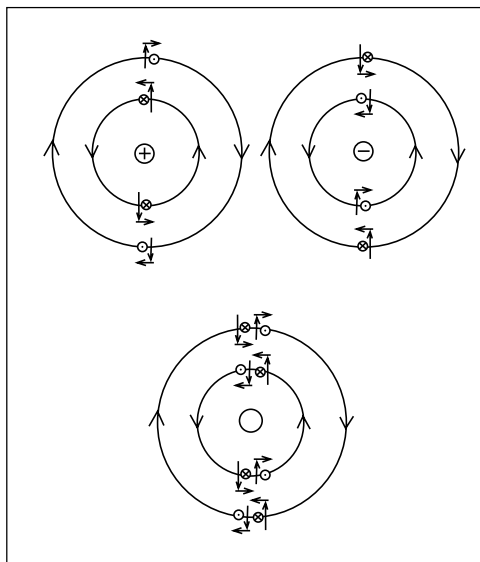
Naturen är "lat"

Naturen strävar efter att vara så energisnål som möjligt. I fallet med den positiva och negativa laddningen så behöver detta inte leda till att totala energin ökar till 170, utan det finns möjlighet för dessa partiklar att tillsammans ha energin 166 utan att de var och en behöver ha 164 + 2. De lever i symbios, drar nytta av varandra, och kan inte skiljas för då hamnar båda partiklarna under den kritiska energigränsen.



17.16 Varför attraheras neutrala partiklar (tyngdkraft)?

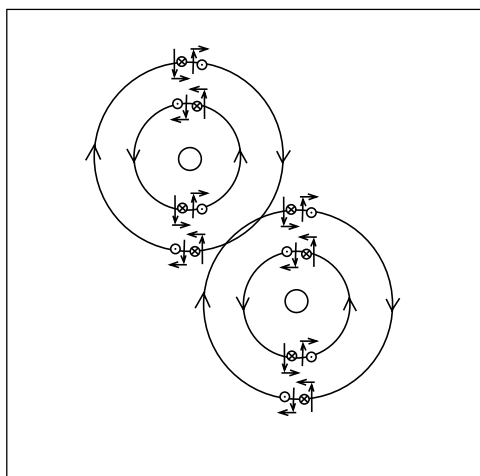
En neutral partikel består av en positiv partikel och en negativ partikel. Runt denna neutrala partikel finns fyra signaler, som två och två har motriktade magnetiska fältstyrkor, och i den andra kombinationen två och två som har motriktade elektriska fältstyrkor.



Tyngdkraft

När två sådana neutrala partiklar placeras i varandras närhet så ser man att det finns signaler som samverkar (samma riktning, samma polarisation). Detta innebär högre energitäthet i det gemensamma området.

Om man sår på partiklarna så kommer totala energimängden att minska, ett tillstånd som partiklarna försöker förhindra. De dras mot varandra.



17.17 Hur snabbt är ljushastigheten?

När ljus passerar en tung himlakropp så böjs strålgången. Detta förutsades av Einstein, och mätningar vid en solförmörkelse var första experimentet som visade att Einstein kunde ha rätt.

Om himlakroppen består av eller omges av signaler som snurrar så kan man tänka sig att dessa signaler kan "låna" energi ur andra signaler som passerar, och att ett sådant lån kan innebära att den lånade strålen under viss sträcka följer cirkelbanan, d.v.s. går längre väg. På så vis kommer vågfronten att krökas.

Men ljuset har ljusfrekvens. De signaler som snurrar är likspänning? Hur går det ihop?

Finns egentligen likspänning? Om vi haft "Big Bang" så har allt en början, allt är pulsformat, och pulser kan delas upp i sitt spektrum som för pulsers del innehåller i stort sett alla frekvenser.

Kapacitivitetstalet ϵ_r

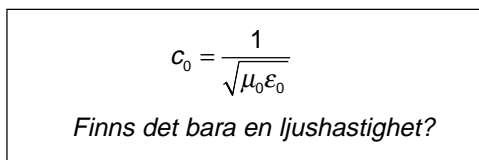
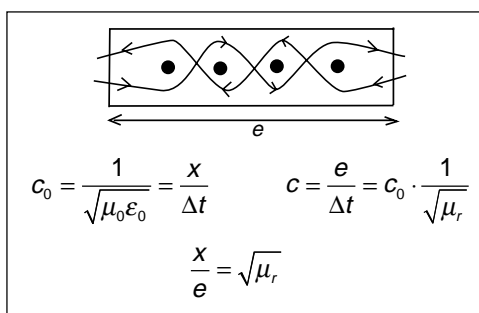
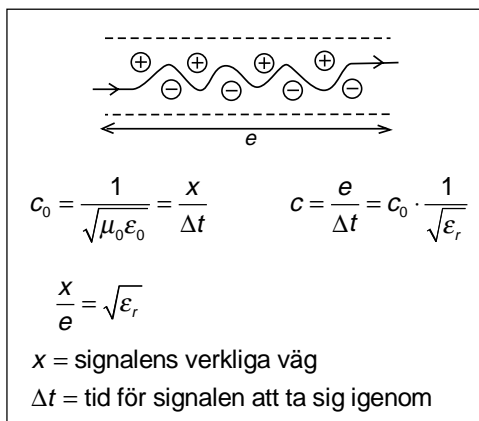
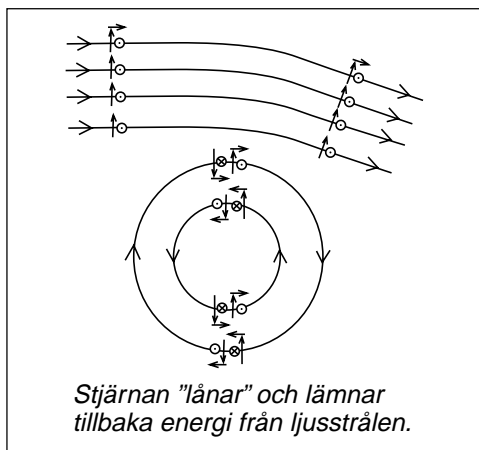
I ett dielektrikum kan man tänka sig att dielektrikats molekyler ställt in sig i ett så energikonserverat läge som möjligt. Positivt och negativt laddade partiklar utnyttjar varandra maximalt.

När man skickar signal genom detta dielektrikum hittar molekylerna andra sätt som är optimala genom att "låna" energin i signalen. Signalen kommer att åka slalom, och det tar längre tid att komma fram.

Detta upplever vi som om signalen går långsammare, trots att vi kanske bara har en enda ljushastighet. Exakt detsamma sker i vågledaren. Signalerna går med ljushastigheten i vakuum, men i sicksack.

Relativa permeabiliteten μ_r

Om materialet innehåller molekyler som är magnetiska så får signalerna (i elektriskt ledande material måste det vara två signaler) molekylerna att inordna sig så att magnetiska fältstyrkorna samverkar, samtidigt som signalerna åker slalom, tycks gå långsammare.



17.18 Laddning i rörelse, magnetfält och massa

Vi utgår från en positivt laddad partikel och de två signalerna som snurrar runt partikeln. E -fälten samverkar medan H -fälten är motriktade och inte märks.

Nu tittar vi på de signaler som snurrar i andra projektionen. Där ser vi spetsen på de elektriska fältstyrkorna, och de båda motriktade magnetiska fältstyrkorna.

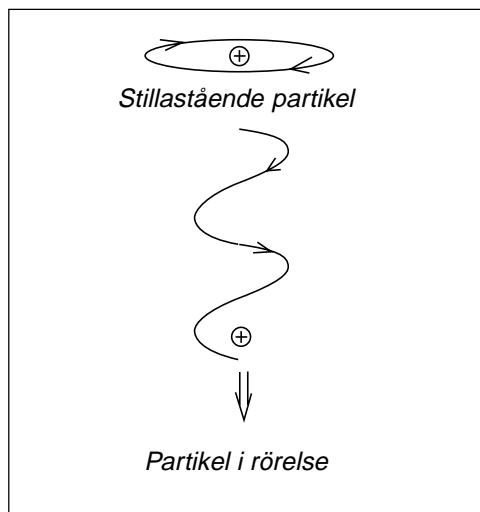
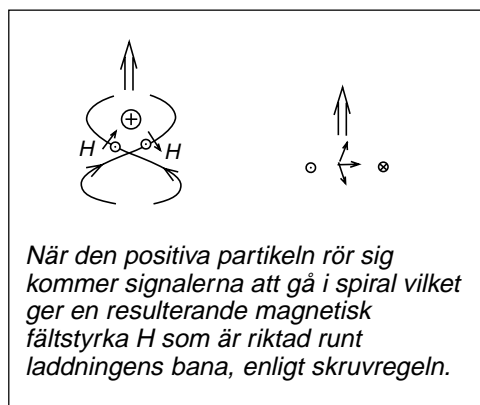
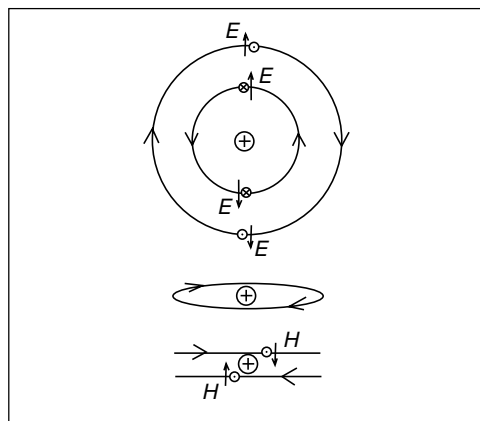
Partikeln rör sig med konstant hastighet

Nästa bild visar partikeln när den rör sig med konstant hastighet. Vad som då händer är att signalerna inte längre snurrar i cirklar utan i spiraler. De följer med när partikeln rör sig. H -fälten är riktade vinkelrätt mot signalernas rörelseriktning. H -fälten tar inte längre ut varandra utan vi får en resulterande magnetisk fältstyrka som snurrar runt partikelns rörelsebana med riktning enligt skruvregeln. Detta är det "magnetfält", egentligen den magnetiska fältstyrka, som bildas när ström flyter i en ledare.

Partikelns massa

När en partikel ökar sin hastighet så ökar inte rörelsemängden proportionellt mot hastigheten i kvadrat, utan betydligt mer, och går mot oändligheten när hastigheten närmar sig ljushastigheten. Tidigare sade man populärt att partikelns massa ökade när hastigheten ökade.

Titta på bilden här bredvid. De signaler som rusar runt partikeln i spiral när partikeln rör sig, de signalerna (som ju går med ljushastigheten) kommer att gå i spiral med allt mindre radie, ju fortare partikeln rör sig. Signalerna "kramas" ihop, elektriska fältstyrkorna ökar och partikeln "binder" mer energi. Partikeln kanske verkligen får större "massa".

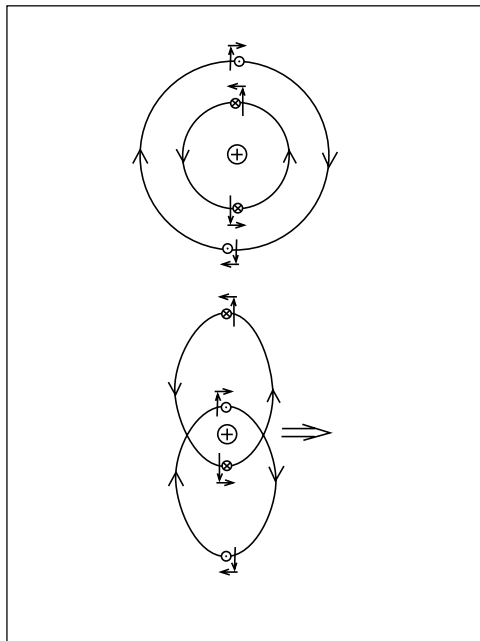


Även signalerna i andra planet

Nu tittar vi på de signaler som snurrar runt den positiva laddningen i samma plan som rörelsen.

Ge laddningen fart så att den rör sig åt höger. Vad händer med signalerna som snurrar? för att de skall hinna komma runt ett helt varv med ljushastigheten samtidigt som partikeln rör sig, så måste signalerna "gina", d.v.s. gå en kortare väg när de skall förbi i partikelns rörelseriktning, och ta en längre väg i andra riktningen för att inte hamna för långt bakom partikeln. Vi får en förskjutning av signalerna i sidled.

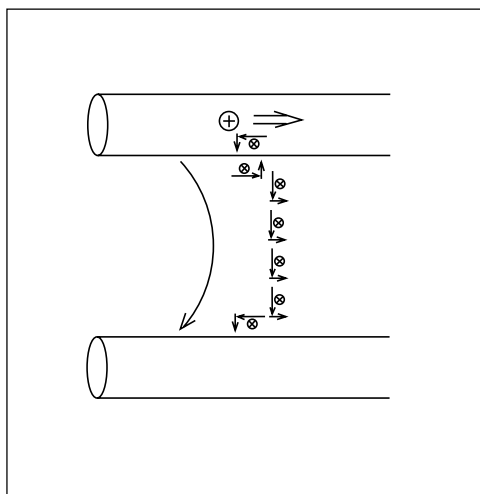
Denna förskjutning leder till att fältstyrkorna koncentreras när de går närmast partikeln. Här binds mest energi, och denna energi rör sig i samma riktning som partikeln. Vi ser även i denna bild den resulterande magnetiska fältstyrkan som snurrar enligt skruvregeln runt partikeln.

**Signal på ledningen**

Nu låter vi framåtgående signal gå i utrymme mellan ledarna i en ledning. Signalen utbreder sig som en sfärisk våg, där fältstyrkan kan delas upp i en plan vågfront och två små sidoriiktade komponenter som matar signal in i ledningens ledare.

Den fältstyrka som ligger parallellt med ledaren driver en positiv laddning åt höger så att det blir laddningsöverskott till höger. Då uppstår en motriktad elektrisk fältstyrka, och tillsammans med magnetiska fältstyrkan runt ledaren matar denna elektriska fältstyrka signal ut ur ledaren.

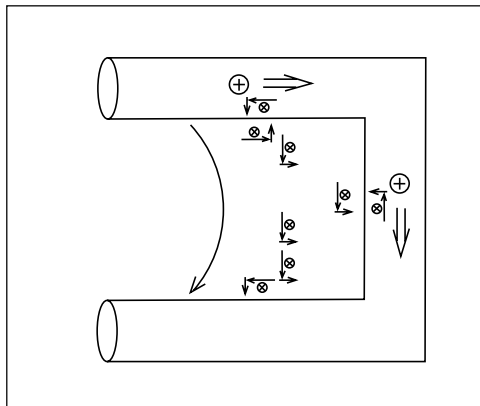
Om det finns resistans i ledaren kommer laddningarna inte att kunna ge tillräckligt hög motriktad fältstyrka. Då matas det in mer effekt i ledaren än vad som kommer tillbaka i form av elektromagnetisk signal. Det blir värmeutveckling. Ohmskt motstånd.



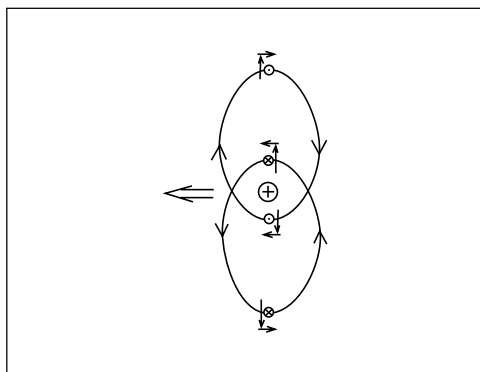
Kortsluten ledning

Nu kommer signalen fram till ledningens ända som är kortsluten.

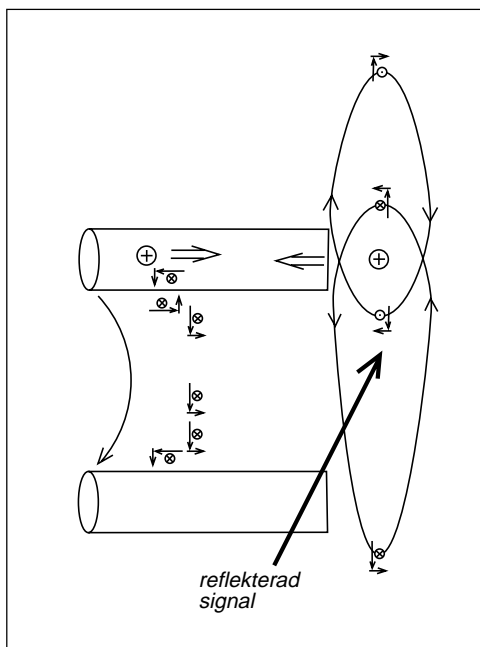
Vi får ett elektrisk fältstyrka som matar positiva laddningar nedåt, som ger laddningsöverskott i underkanten och en elektrisk fältstyrka riktat uppåt. Tillsammans med den magnetiska fältstyrkan bildar detta den reflekterade signalen på ledningen.

**Laddningsöverskott driver tillbaka laddningen**

När de positiva laddningarna kommer till ledningsändan så bildas efter ett litet tag överskott på elektrisk laddning, som ger för hög motriktad elektrisk fältstyrka, som driver tillbaka laddningarna. Signalerna förskjuts åt andra hållet.



Vi ser hur det bildas en signal som går i motsatt riktning, en reflekterad signal med "rätt" polarisation.



17.19 Radiostrålning och gravitationsstrålning

Elektromagnetiska vågor

Men det händer något mer när signalen når den öppna ledningen. Laddningarna bromsas in.

Titta på bilden. Den positiva partikeln accelereras åt höger. Då förskjuts signal-cirklarna. Den ena cirkeln förskjuts uppåt. Den signal som roterar i denna cirkel kommer att bli starkare, även långt bort, eftersom cirkelarna finns, svagare och svagare, ända ut till oändligheten.

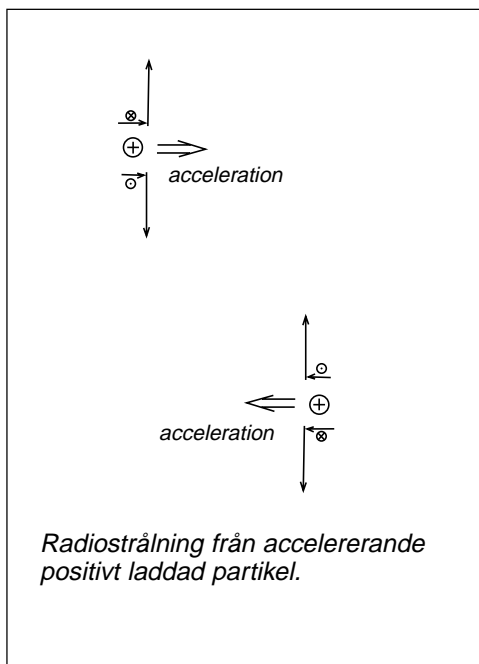
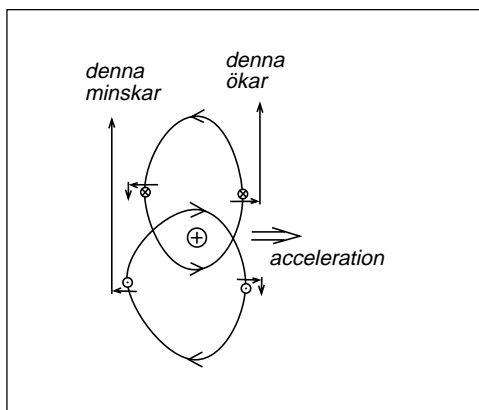
Den andra cirkeln förskjuts nedåt. Den signal som roterar i denna cirkel och strålar uppåt kommer att bli svagare. Men det är samma sak som att den motsatta polarisationen blir starkare. Förändringen i signalerna samverkar, och denna förändring kan vi upptäcka på långt avstånd. Det är dessa förändringar som vi kallar radiostrålning.

Elektriskt laddade partiklar som accelererar i ena eller andra riktningen (det inbegriper även retardation) kommer att åstadkomma "krusningar på vattenytan", elektriska och magnetiska fältstyrkor som sätter fart på elektriskt laddade partiklar, oavsett om partiklarna befinner sig i "metall" eller i ett plasma.

När likströmssignalen kommer till ledarens ända accelereras laddningen i motsatt riktning. Vi får radiostrålning. Men vi får även lika mycket radiostrålning med omvänd polaritet från den andra ledarens ända. Radiostrålningarna tar ut varandra.

Gravitationsvågor

När en neutral partikel ökar sin hastighet så finns fyra signaler som snurrar runt partikeln. Två av dessa signaler kommer att förskjutas åt vänster, två förskjuts åt höger. Vi får alltså en tyngdpunktsförskjutning i dessa båda energifält, som sprids utåt, som gravitationsvågor. Men dessa gravitationsvågor ger krusningar av båda polarisationsriktningarna. Det blir ingen resulterande elektrisk och/eller magnetisk fältstyrka, inget som kan detekteras i en radioantenn.



17.20 Batteriet

Hur beskriver vi batteriet?

Den vanliga beskrivningen av ett batteri är spänningskällan som har inre resistans. När batteriet belastas med en last vars resistans är lika med batteriets inre resistans så får vi ut halva batterispänningen.

Men detta är ju transmissionsledningen

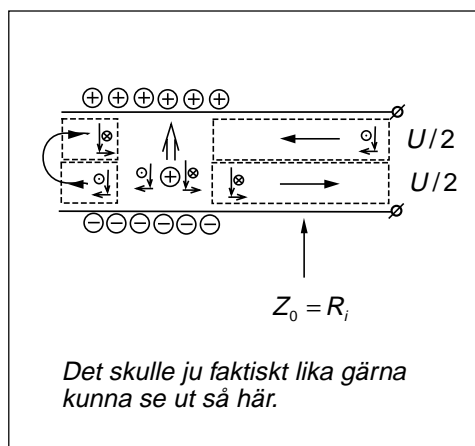
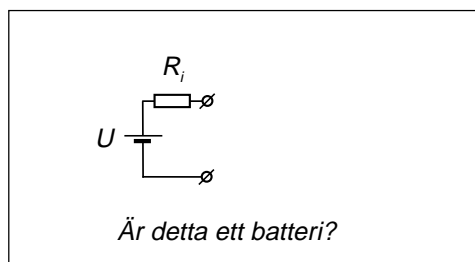
När transmissionsledningen belastas med en lastresistans vars värde är lika med ledningens karakteristiska impedans, då får vi ut halva spänningen.

Man kan tänka sig att i batteriet finns någon kemisk process som driver positivt laddade partiklar från minuspolen till pluspolen. Detta ger överskott på negativ laddning vid minuspolen och överskott på positiv laddning vid pluspolen. Vi får en elektrisk fältstyrka riktad från plus till minus i batteriet. Mot denna elektriska fältstyrka driver den kemiska processen positivt laddade partiklar. Det uppstår en roterande magnetisk fältstyrka, som tillsammans med den elektriska fältstyrkan kommer att driva elektromagnetisk effekt ut på ledningen, i båda riktningarna. Om inget finns anslutet så reflekteras signalerna.

Bara växelström kan "stråla"

Det intressanta med denna bild av batteriet är att vi åstadkommer elektromagnetisk likströmsstrålning. Men för att kunna åstadkomma denna likströmsstrålning så måste de positiva laddningarna ha någonstans att ta vägen. Vi behöver ledningen som för bort laddningarna så att den kemiska processen kan driva nya laddningar genom batteriet, så att det bildas mer "likströmsstrålning".

Om laddningarna däremot får gunga fram och tillbaka, som på antennen, då behöver vi inga ledningstrådar som för bort laddningarna.



Formelsamling

Formelsamlingen bygger på den formelsamling i telekommunikation som sedan tidigare får användas vid tentamen i telekommunikation vid ingenjörsskolan i Hanninge.

Det kan finnas bokstavsbeteckningar som avviker från de som jag använder i boken, men detta är den värld som väntar den färdiga ingenjören. Det finns ingen enhetlig standard för bokstavsbeteckningar när det gäller telekommunikation, utan man måste veta vad man gör.

0. Övrigt

Decibel

- 0.1 $P[dBm] = 10 \log P[mW]$ Effekt uttryckt i dBm.
- 0.2 $P[dBW] = 10 \log P[W]$ Effekt uttryckt i dBW.
- 0.3 $A_p [dB] = 10 \log \frac{P_{ut}}{P_{in}} = 10 \log P_{ut} - 10 \log P_{in}$ Dämpning, eller förstärkning.
- 0.4 $P[W] = 10^{\frac{P[dBW]}{10}}$ Omräkning från dBW till effekt.

Dataskommunikation

- 0.5 $n = \log_2 M = \frac{\log M}{\log 2}$ M = antal olika symboler, n = bit/symbol.
- 0.6 $C_{bit} [bit/s] = n \cdot C_{symb} [symb/s \text{ eller } baud]$
- 0.7 $C_{bit} [bit/s] = \frac{B \cdot \log \left(1 + \frac{S}{N} \right)}{\log 2}$ Shannons formel.
- 0.8 $C_{symb} [symb/s \text{ eller } baud] = 2 \cdot B [Hz]$ Max symbolhastighet vid linjekodning.
- 0.9 $C_{symb} [symb/s \text{ eller } baud] = B [Hz]$ Max symbolhastighet vid modem-kommunikation.

1. Ledningar

$$1.1 \quad Z_0 = \sqrt{\frac{r + j\omega l}{g + j\omega c}}$$

Karakteristiska impedansen beräknad ur ledningens primärkonstanter.

$$1.2 \quad Z_0 = \sqrt{\frac{l}{c}}$$

Karakteristiska impedansen vid förlustfri ledning.

$$1.3 \quad V = V^+ + V^-$$

Framåtgående och reflekterad spänning på ledningen.

$$1.4 \quad I = I^+ - I^- = \frac{V^+}{Z_0} - \frac{V^-}{Z_0}$$

Framåtgående och reflekterad ström. OBS strömmens riktning.

$$1.5 \quad \rho = \Gamma = \frac{V^-}{V^+} = \frac{Z_L - Z_0}{Z_L + Z_0}$$

$\rho = \Gamma =$ reflexionskoefficient.
 $Z_L =$ belastningsimpedans.

$$1.6 \quad \beta = \frac{\omega}{c}$$

Faskonstant (rad/m).

$$1.7 \quad \lambda = \frac{2\pi}{\beta} = \frac{c}{f} = \frac{c_0}{\sqrt{\epsilon_r \mu_r}} \cdot \frac{1}{f}$$

Våglängd.

$$1.8 \quad Z_m = Z_0 \frac{Z_L \cos \beta s + jZ_0 \sin \beta s}{Z_0 \cos \beta s + jZ_L \sin \beta s}$$

Inimpedansen för en s m lång ledning.

$$1.9 \quad SVF = \frac{V_{\max}}{V_{\min}} = \frac{|V^+| + |V^-|}{|V^+| - |V^-|} = \frac{1 + |\rho|}{1 - |\rho|}$$

Ståendevågförhållandet.

$$1.10 \quad Z_{\max} = \frac{V_{\max}}{I_{\min}} = Z_0 \cdot SVF$$

Impedans i ett spänningsmaximum.

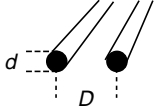
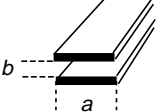
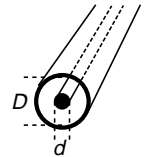
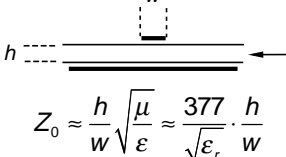
$$1.11 \quad Z_{\min} = \frac{V_{\min}}{I_{\max}} = \frac{Z_0}{SVF}$$

Impedans i ett spänningsminimum.

$$1.12 \quad Z = \sqrt{Z_0 \cdot R_L}$$

Kvartsvågstransformatorns karakteristiska impedans.

1.13 Karakteristisk impedans för några ledningar:

 $Z_0 = \frac{1}{\pi} \sqrt{\frac{\mu}{\epsilon}} \cdot \ln \frac{2D}{d} \approx \frac{276}{\sqrt{\epsilon_r}} \cdot \log \frac{2D}{d}$	 $Z_0 = \frac{b}{a} \sqrt{\frac{\mu}{\epsilon}} \approx \frac{377}{\sqrt{\epsilon_r}} \cdot \frac{b}{a}$
 $Z_0 = \frac{1}{2\pi} \sqrt{\frac{\mu}{\epsilon}} \cdot \ln \frac{D}{d} \approx \frac{138}{\sqrt{\epsilon_r}} \cdot \log \frac{D}{d}$	 $Z_0 \approx \frac{h}{w} \sqrt{\frac{\mu}{\epsilon}} \approx \frac{377}{\sqrt{\epsilon_r}} \cdot \frac{h}{w}$

2. TEM-vågor

Allmänt

Konstanter: $\mu_0 = 4 \cdot \pi \cdot 10^{-7} \text{ H/m}$ $\epsilon_0 \approx \frac{1}{36\pi} \cdot 10^{-9} \text{ F/m}$ $c_0 \approx 3 \cdot 10^8 \text{ m/s}$

- 2.1 $S = H_{\text{eff}} \cdot E_{\text{eff}} = H_{\text{eff}}^2 \cdot Z_{\text{TEM}} = \frac{E_{\text{eff}}^2}{Z_{\text{TEM}}}$ Strålningstäthet.
- 2.2 $Z_{\text{TEM}} = \sqrt{\frac{\mu}{\epsilon}}$ Vågimpedansen.
- $Z_{\text{TEM}} = \sqrt{\frac{\mu_0}{\epsilon_0}} \approx 120\pi \approx 377 \text{ } \Omega$ Vågimpedansen i vakuum.
- 2.3 $c = \frac{1}{\sqrt{\mu\epsilon}}$ Elektromagnetiska signalens hastighet.
- 2.4 $\delta = \frac{1}{\sqrt{\pi f \mu_0 \sigma}}$ Skindjup, σ = materialets konduktivitet.
- 2.5 $P = a \cdot b \cdot \frac{E_{\text{eff}}^2}{Z_{\text{TEM}}}$ Signalens effekt i en plattledning.

TE₁₀-mod i rektangulär vågledare

- 2.7 $R_{\text{TE}} = R_{\text{TEM}} \cdot \frac{\lambda_x}{\lambda_0} = \frac{R_{\text{TEM}}}{\sqrt{1 - \left[\frac{\lambda_0}{2a}\right]^2}}$ Vågimpedans.
- 2.8 $v_f = f \cdot \lambda_x = \frac{c}{\sqrt{1 - \left[\frac{\lambda_0}{2a}\right]^2}} = \frac{c}{\sqrt{1 - \left[\frac{f_g}{f}\right]^2}}$ Fashastighet.
- 2.9 $v_g = \frac{c^2}{v_f} = c \sqrt{1 - \left(\frac{\lambda_0}{2a}\right)^2}$ Grupphastighet.
- 2.10 $P = \frac{abE_{\text{eff}}^2}{2R_{\text{TE10}}}$ Effekt i vågledare.
- 2.11 $f_g = \frac{c}{\lambda_g} = \frac{c}{\sqrt{\epsilon_r}} \cdot \frac{1}{2a}$ Gränshastighet.

3. Antenner

$$3.1 \quad \eta_L = \frac{P_{ut}}{P_S} = \frac{R_s}{R_s + R_f}$$

P_{ut} = Utstrålad effekt, P_S = till antennen matad effekt, R_s = strålningsresistans, R_f = förlustresistans.

$$3.2 \quad S_0 = \frac{P_{ut}}{4\pi r^2}$$

S_0 = strålningsstäthet från isotrop antenn på avståndet r .

$$3.3 \quad D = \frac{S_{\max}}{S_0} = \frac{S_{\max}}{P_{ut}/4\pi r^2}$$

Antennens direktivitet. S_{\max} = strålningsstäthet i bästa riktningen.

$$3.4 \quad G = \frac{S_{\max}}{P_S/4\pi r^2} = \eta_L \cdot D$$

Antennvinst ("gain").

$$3.5 \quad A_e = G \cdot \frac{\lambda_0^2}{4\pi} = \frac{P_m}{S}$$

Absorptionsarea.

$$3.6 \quad EIRP = P_S \cdot G$$

Effektivt isotropt utstrålad effekt.

$$3.7 \quad L_s [dB] = -32,4 [dB] - 20 \log d_{km} - 20 \log f_{MHz} \quad \text{"Free space loss" mellan isotropa antenner.}$$

4. Brus

Konstanter: Boltzmanns konstant: $k = 1,38 \cdot 10^{-23} \text{ J}^\circ\text{C}$ Temperatur: $T_0 = 290 \text{ K}$

$$4.3 \quad N_0 = kT$$

Spektralfördelning, vitt brus (W/Hz).

$$4.4 \quad E_n = \sqrt{4kTB R}$$

Tillgänglig brusspänning (effektivvärde).

$$4.5 \quad N_{th} = kTB$$

Tillgänglig termisk bruseffekt.

$$4.6 \quad NR = \frac{S_{in}}{N_{in(290K)}} \bigg/ \frac{S_{ut}}{N_{ut}} = \frac{N_{ut}}{G \cdot N_{in(290K)}}$$

Brusförhållande (Noise Ratio).

$$4.7 \quad T_{eq} = (NR - 1) \cdot T_0$$

Ekvivalent brustemperatur.

$$4.8 \quad NR = NR_1 + \frac{NR_2 - 1}{G_1} + \dots + \frac{NR_n - 1}{G_1 G_2 \dots G_{n-1}}$$

NR, kaskadkopplade förstärkarsteg.

$$4.9 \quad T_{eq} = T_{eq1} + \frac{T_{eq2}}{G_1} + \dots + \frac{T_{eqn}}{G_1 G_2 \dots G_{n-1}}$$

T_{eq} , kaskadkopplade förstärkarsteg.

$$4.10 \quad NF [dB] = 10 \log NR$$

Brusfaktor i dB.

$$4.11 \quad N_{th 290K} [dBm] = -174 [dBm] + 10 \log B$$

Termiskt brus i dBm vid 290 K.

$$4.12 \quad N_{th 290K} [dBW] = -204 [dBW] + 10 \log B$$

Termiskt brus i dBW vid 290 K.

5. Modulering

Amplitudmodulering

_____ Omodulerad bärvågssignal _____

1. $E_c \cos \omega_c t$

_____ Modulerande signal (informationssignal) _____

1. $E_m \cos \omega_m t$

2. $m(t)$

_____ AM _____

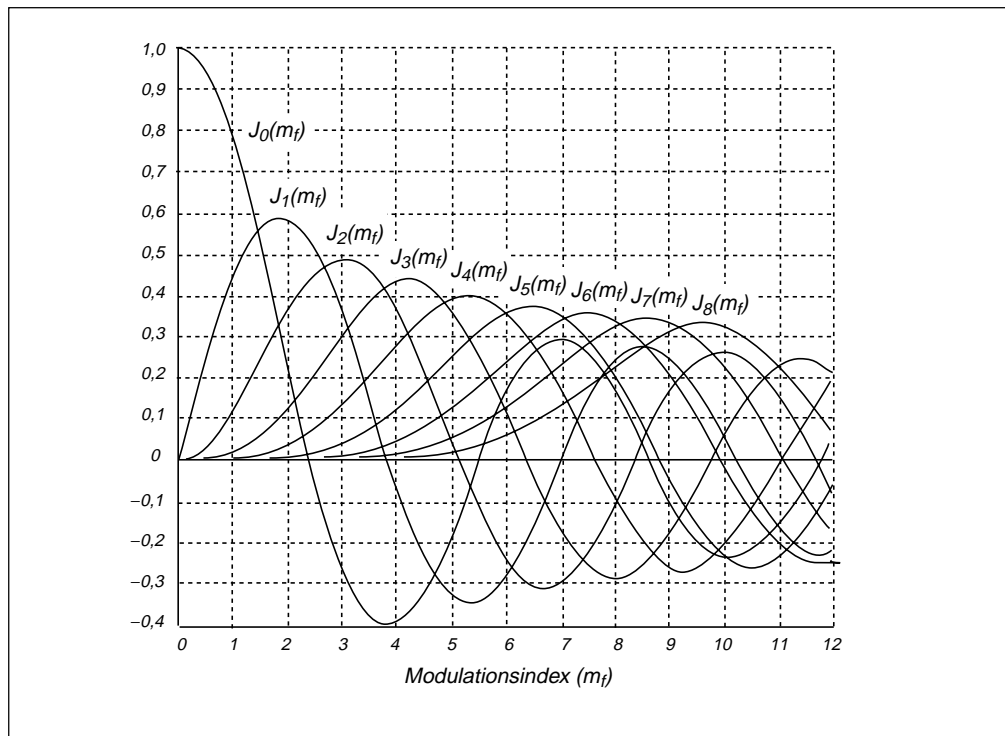
1. $E_c \left(1 + \frac{E_m}{E_c} \cos \omega_m t\right) \cos \omega_c t$ modulationsindex: $m_a = \frac{U_{\max} - U_{\min}}{U_{\max} + U_{\min}} = \frac{E_m}{E_c}$

2. $(E_c + m(t)) \cos \omega_c t$

_____ DSB utan bärvåg _____

1. $kE_c E_m \cos \omega_m t \cos \omega_c t$ k är en apparatkonstant (förstärkning)

2. $kE_c m(t) \cos \omega_c t$



Argumentmodulering

PM	FM
Omodulerad bärvågssignal	
$E_c \cos \omega_c t$	$E_c \cos \omega_c t$
Modulerande signal (informationssignal)	
1. $E_m \cos \omega_m t$	$E_m \cos \omega_m t$
2. $m(t)$	$m(t)$
Modulerad bärvågssignal	
1. $E_c \cos(\omega_c t + m_p \cos \omega_m t)$	$E_c \cos(\omega_c t + m_f \sin \omega_m t)$
2. $E_c \cos(\omega_c t + k_\theta m(t))$	$E_c \cos(\omega_c t + 2\pi \int_0^t k_f m(\tau) d\tau)$
Modulationsindex (= fasdeviation)	
$m_p = k_\theta E_m$	$m_f = \frac{k_f E_m}{f_m} = \frac{\Delta f_c}{f_m}$
Momentan fas	
1. $\omega_c t + m_p \cos \omega_m t$	$\omega_c t + m_f \sin \omega_m t$
2. $\omega_c t + k_\theta m(t)$	$\omega_c t + 2\pi \int_0^t k_f m(\tau) d\tau$
Momentan vinkelfrekvens	
1. $\omega_c + m_p \omega_m (-\sin \omega_m t)$	$\omega_c + m_f \omega_m \cos \omega_m t$
2. $\omega_c + k_\theta m(t)$	$\omega_c + 2\pi k_f m(t)$
Momentan frekvens	
1. $f_c + m_p f_m (-\sin \omega_m t)$	$f_c + m_f f_m \cos \omega_m t$
2. $f_c + \frac{k_\Phi}{2\pi} \cdot \frac{dm(t)}{dt}$	$f_c + k_f m(t)$
Bandbredd	
$B = 2(f_m + \Delta f)$ (Carsons regel)	

Trigonometriska formler

- (1) $\sin^2 \alpha + \cos^2 \alpha = 1$
- (2) $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$
- (3) $\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta$
- (4) $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$
- (5) $\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$
- (6) $\tan(\alpha + \beta) = \frac{\tan \alpha + \tan \beta}{1 - \tan \alpha \tan \beta}$
- (7) $\tan(\alpha - \beta) = \frac{\tan \alpha - \tan \beta}{1 + \tan \alpha \tan \beta}$
- (8) $\cot(\alpha + \beta) = \frac{\cot \beta \cot \alpha - 1}{\cot \beta + \cot \alpha}$
- (9) $\cot(\alpha - \beta) = \frac{\cot \beta \cot \alpha + 1}{\cot \beta - \cot \alpha}$
- (10) $\sin 2\alpha = 2 \sin \alpha \cos \alpha$
- (11) $\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha = 2 \cos^2 \alpha - 1 = 1 - 2 \sin^2 \alpha$
- (12) $\sin^2 \alpha = \frac{1 - \cos 2\alpha}{2}$; (13) $\cos^2 \alpha = \frac{1 + \cos 2\alpha}{2}$
- (14) $\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$
- (15) $\sin \alpha - \sin \beta = 2 \sin \frac{\alpha - \beta}{2} \cos \frac{\alpha + \beta}{2}$
- (16) $\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$
- (17) $\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}$
- (18) $\sin \alpha \cos \beta = \frac{1}{2} \sin(\alpha - \beta) + \frac{1}{2} \sin(\alpha + \beta)$
- (19) $\cos \alpha \cos \beta = \frac{1}{2} \cos(\alpha - \beta) + \frac{1}{2} \cos(\alpha + \beta)$
- (20) $\sin \alpha \sin \beta = \frac{1}{2} \cos(\alpha - \beta) - \frac{1}{2} \cos(\alpha + \beta)$
- (21) $\cos \alpha = \frac{e^{j\alpha} + e^{-j\alpha}}{2}$; (22) $\sin \alpha = \frac{e^{j\alpha} - e^{-j\alpha}}{2j}$

6. Optokommunikation

$$6.1 \quad B = \frac{1}{2} C_{\text{symp}} = \frac{1}{2 \cdot \sigma_b} \quad \text{Bandbredd p.g.a. pulsbreddning.}$$

$$6.2 \quad B_{\text{mod}} \approx \frac{B[\text{Hz} - \text{km}]}{L[\text{km}]} \quad L \leq 1 \text{ km} \quad \text{Bandbredd p.g.a. moddispersion.}$$

$$B_{\text{mod}} \approx \frac{B[\text{Hz} - \text{km}]}{\sqrt{L[\text{km}]}} \quad L > 1 \text{ km}$$

$$6.3 \quad B_{\text{tot}} = \frac{1}{\sqrt{\frac{1}{B_{\text{krom}}^2} + \frac{1}{B_{\text{mod}}^2}}} \quad \text{Total bandbredd.}$$

9. Fourierserier

Definition 1: (komplexa Fourierserier)

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad \text{där}$$

$$c_n = \frac{1}{T} \int_p x(t) \cdot e^{-jn\omega_0 t} dt$$

Definition 2: (trigonometriska Fourierserier)

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t) \quad \text{där}$$

$$a_0 = \frac{1}{T} \int_p x(t) dt \quad a_n = \frac{2}{T} \int_p x(t) \cdot \cos(n\omega_0 t) dt \quad b_n = \frac{2}{T} \int_p x(t) \cdot \sin(n\omega_0 t) dt$$

Definition 3: (alternativa Fourierserier)

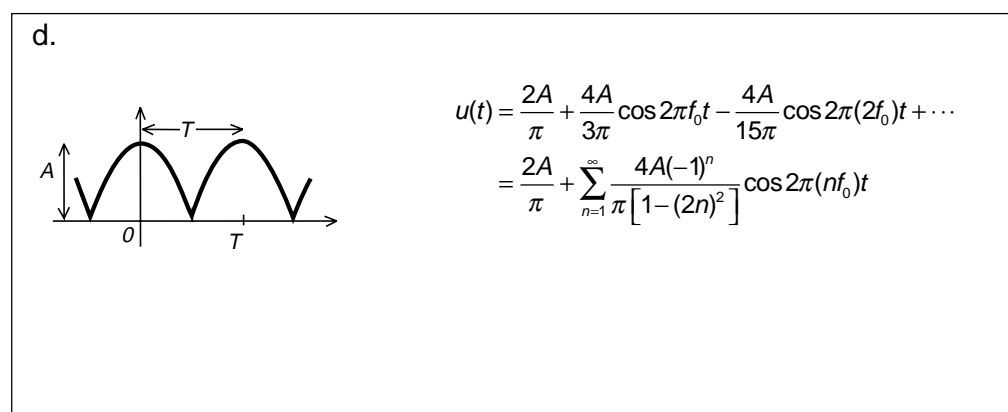
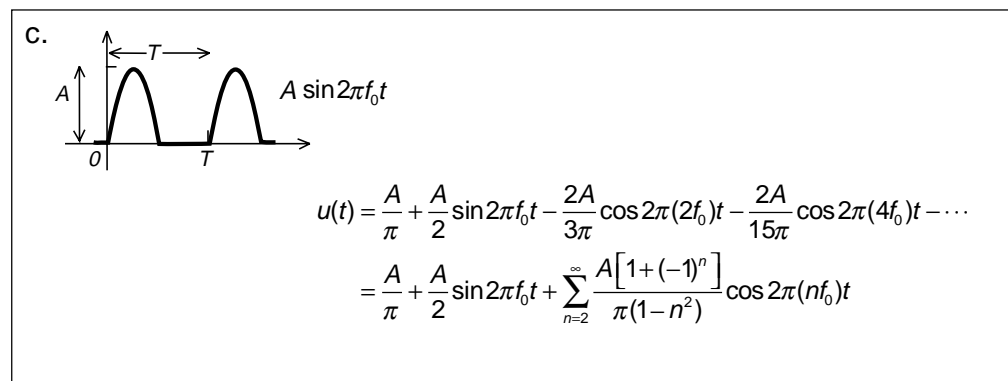
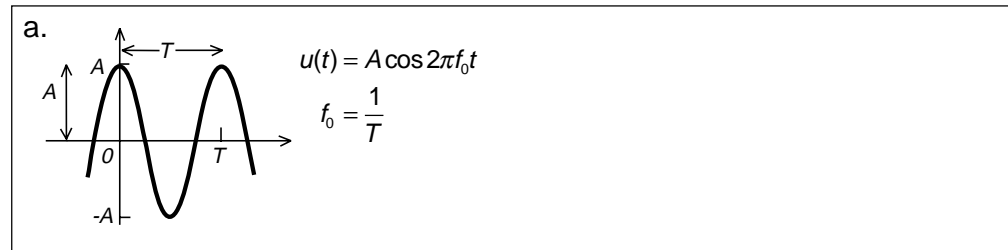
$$x(t) = A_0 + \sum_{n=1}^{\infty} A_n \cdot \cos(n\omega_0 t + \alpha_n) \quad \text{där}$$

$$A_0 = a_0 = \frac{1}{T} \int_p x(t) dt$$

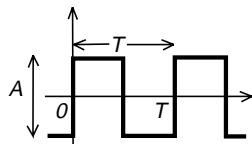
$$A_n = \sqrt{a_n^2 + b_n^2}$$

$$\tan(\alpha_n) = \frac{-b_n}{a_n} \quad \alpha_n = \arctan\left(\frac{-b_n}{a_n}\right) + [(\pi) \leftarrow \text{om } a_n < 0]$$

Några vanliga funktioner:



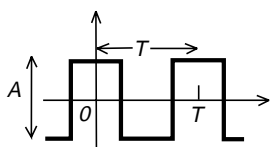
e.



$$u(t) = \frac{2A}{\pi} \sin 2\pi f_0 t + \frac{2A}{3\pi} \sin 2\pi(3f_0)t + \dots$$

$$= \sum_{n=\text{udda}}^{\infty} \frac{2A}{n\pi} \sin 2\pi(nf_0)t$$

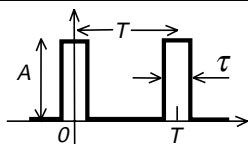
f.



$$u(t) = \frac{2A}{\pi} \cos 2\pi f_0 t - \frac{2A}{3\pi} \cos 2\pi(3f_0)t + \frac{2A}{5\pi} \cos 2\pi(5f_0)t - \dots$$

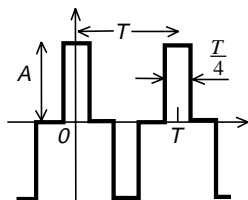
$$= \sum_{n=1}^{\infty} \left(A \frac{\sin n\pi/2}{n\pi/2} \right) \cos 2\pi(nf_0)t$$

g.



$$u(t) = A \frac{\tau}{T} + \sum_{n=1}^{\infty} \left(2A \frac{\tau}{T} \right) \left(\frac{\sin 2\pi\tau/T}{n\pi\tau/T} \right) \cos 2\pi(nf_0)t$$

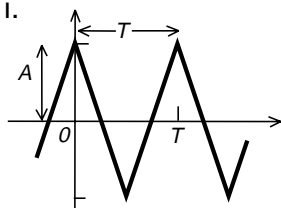
h.



$$u(t) = \sum_{n=\text{udda}}^{\infty} \left(A \frac{\sin n\pi/4}{n\pi/4} \right) \cos 2\pi(nf_0)t$$

Observera att uttrycket bara gäller när "pulsbredden" är periodtiden dividerat med fyra.

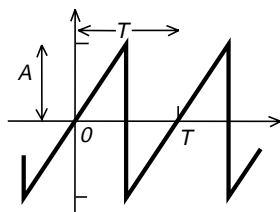
i.



$$u(t) = \frac{8A}{\pi^2} \cos 2\pi f_0 t + \frac{8A}{9\pi^2} \cos 2\pi(3f_0)t + \frac{8A}{25\pi^2} \cos 2\pi(5f_0)t + \dots$$

$$= \sum_{n=\text{udda}}^{\infty} \frac{8A}{(n\pi)^2} \cos 2\pi(nf_0)t$$

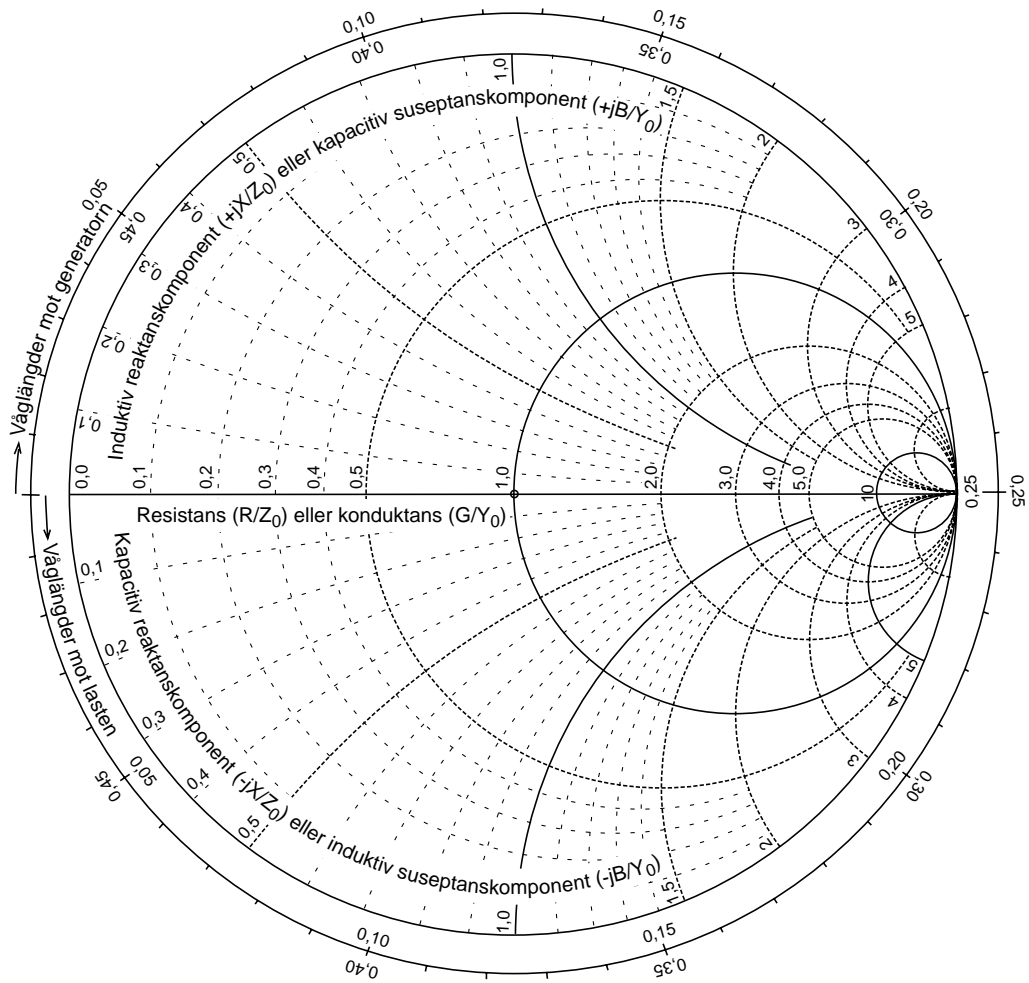
j.



$$u(t) = \frac{2A}{\pi} \left(\sin 2\pi f_0 t - \frac{1}{2} \sin 2\pi(2f_0)t + \frac{1}{3} \sin 2\pi(3f_0)t - \dots \right)$$

$$= \sum_{n=1}^{\infty} [(-1)^{n-1}] \left(\frac{2A}{n\pi} \right) \sin 2\pi(nf_0)t$$

Förenklat Smithdiagram



INDEX

- 2B1Q, 156
- 8PSK, 170

- A/D-omvandling, 142
- ADSL, 155
- A-lag, 145
- aliasing distorsion, 143
- AM-detektor, 74
- AM-mottagare, 98
- Ampère, 279
- amplitude skift keying (ASK), 160
- amplitudmodulering, 66
- amplitudspektrum, 28
- anpassning, 195, 218, 221
- antennbrustemperatur, 257
- antennvinst (G), 244
- antipodal 2ASK, 163
- argumentmodulering, 110
- ATM, 188
- atmosfärbrus, 257
- atmosfärgas, 255
- automatic repeat request (ARQ), 172

- balanserad blandare, 82
- bandbredd, 72, 266
- basband, 64
- basstationsantenn, 243
- baud, 148
- begränsare, 126, 128
- Besselfunktioner, 117
- bithastighet, 151
- bit-stuffing, 181
- bitsynkronisering, 140
- block check sequence (BCS), 172
- blockkoder, 176
- BPSK, 163
- broadcastadress, 182, 187
- brus, 91, 256
- brusfaktor, 54
- brusförhållande, 54
- brustemperatur, 53, 106
- bruströskel, 130
- bruströskelsänkande FM-detektor, 134

- brytningsindex, 264
- bärvåg, 70

- carrier, 70
- Carsons regel, 119
- cellförmedlande nät, 188
- cirkulär polarisation, 236
- Coulomb, 278
- cyclic redundancy check (CRC), 172

- D/A-omvandling, 142
- data communication equipment (DCE), 178
- data radio channel (DARC), 95
- data terminal equipment (DTE), 178
- datalänklagret, 141
- dBd, 244
- dBi, 244
- dBm, 24, 104
- dBW, 24, 104
- decibel, 22, 104, 252
- deemphasis, 131
- deltoner, 35
- deviation, 113
- dipolantenn, 234
- direktivitet (D), 242
- dispersions-skiftad fiber, 265
- double sideband (DSB), 82
- dsb-modulering, 84
- dubbelsuperheterodyn, 102
- dubbla sidband, 82
- dynamik, 145
- dämpning, 22, 263

- EDGE, 170
- effective aperture, 245
- effective isotropic radiated power (EIRP), 250
- effective radiated power (ERP), 251
- effekt, 77
- effektiv absorptionsarea, 245
- effektiv mottagningsarea, 245
- effektmeter, 210
- effektspektrum, 30
- effekttäthet, 242

- elektron, 228
enkelt sidband, 88
envelopp-detektor, 75
E-plansdiagram, 239
- F, 54
faltningkodning, 141, 176
Faraday, 279
fasdetektor, 132
faslåst loop, 132
faslåst slinga, 132
fasmodulering, 67, 111
fassetpektrum, 28
fasta telenätet, 154
fasvinkel, 11
FDM-telefoni, 96
felrättning, 176
flagga, 181
flankdetektor, 126
FM-stereo, 93
Fourier-serie, 26
Fouriertransform, 41, 44
frame check sequence (FCS), 172, 182
framåtgående effekt, 195
frekvens, 10
frekvensdomän, 31
frekvensmodulering, 67, 111
frequency division multiplex (FDM), 96, 185
fyrekantvåg, 26
fysiska lagret, 140
fädningsmarginal, 255
förbindelselöst, 186
förbindelseorienterat, 186
förkortad dipol, 237
- gain (G), 244
galaktiskt brus, 257
Gauss, 279
Gaussiskt brus, 50
Gaussiskt filtrerat MSK (GMSK), 167
generatorpolynom, 173
GFSK, 167
gradientindexfiber, 261
Gray-kodning, 170
gränssnitt, 140
- hagelbrus, 53
halvvågsantenn, 232
halvvågsdipol, 235, 237
Hammingkodning, 176
Hartley, 148
HDLC, 180
- helixantenn, 236
Hertz experiment, 231
Hertz, 190
horisontell polarisation, 236
H-plansdiagram, 239
hårdvaruhandskakning, 178
högfrekvenssteg (HF), 101
- I-brus, 91
icke-koherent 2FSK, 165
IEEE 802.2, 180
I-kanal, 90
impedansbestämning, 216
interface, 140
IP-protokollet, 187
I-ram, 182
ISDN-nätet, 156
isotrop antenn, 242
- j ω -metoden, 18
jämn funktion, 36
- kabel-tv, 211
karakteristisk impedans, 193
klass A-förstärkare, 66
klass C-förstärkare, 66
koaxialledning, 194
kodning, 144
koherent 2FSK, 166
koherent detektor, 86, 166
kollinjär antenn, 240
komplex Fourier-serie, 41, 44
kondensator, 224
kontrollsumma, 172
kortsluten stubbe, 222
kretskopplat nät, 185
kristall, 135
kristallmottagare, 98
kromatisk dispersion, 265
kvadraturdetektor, 127
kvantisering, 144
kvantiseringsbrus, 145
kvartskristall, 135
kvartsvågsantenn, 238
kvartsvågstransformator, 218
- laddning, 228
Laplacetransform, 41, 44
laserdiod, 260
light emitting diode (LED), 260
limiter, 126
linjekodning, 64, 156

-
- link access procedure (LAP), 180
 - ljushastighet, 190
 - lob, 239
 - logical link control (LLC), 180
 - lokalscillator, 99
 - lower sideband (LSB), 70, 88
 - lågfrekvenssteg (LF), 101
 - länkbudget, 253

 - man-made noise, 53, 257
 - mantel, 260
 - masken, 198
 - materialdispersion, 264
 - Maxwell, 190, 280
 - medeleffekt, 29
 - mellanfrekvenssteg (MF), 101
 - minimum skift keying (MSK), 166
 - mjuk avkodning, 176
 - moddispersion, 260
 - modkoppling, 267
 - modulationsindex, 69, 114
 - modulering, 65
 - moduleringsgrad, 69
 - modulo-2, 173
 - motståndsbuss, 48
 - mottagarantenn, 246
 - multi tone-teknik, 155
 - multiplikator, 82
 - mätning av antennvinst, 244

 - next to receive (NR), 184
 - noise factor (NF), 54
 - noise ratio (NR), 54
 - normalfördelat buss, 50
 - now sending (NS), 184
 - Nyquistfiltrering, 149, 160
 - nätverksanalysator, 211
 - nätverkslagret, 141

 - ODFM, 165
 - optofiber, 260
 - ortogonal mottagning, 165
 - oscillator, 66
 - OSI, 140

 - P/F, 183
 - paketförmedlande nät, 186
 - parabolantenn, 241, 243
 - parallellledning, 194
 - PCM-kodning, 144
 - periodisk signal, 35
 - periodtid, 10

 - phase shift keying (PSK), 169
 - pilot, 93
 - plattledning, 194
 - polarisation, 236
 - Poyntings vektor, 190
 - preemphasis, 131
 - primary, 181
 - produktdetektor, 82
 - protokoll, 140
 - PSTN-nätet, 154
 - puls code modulation, 144
 - pulsräknande detektor, 127

 - Q-brus, 91
 - Q-kanal, 90
 - quad amplitude modulation (QAM), 90

 - radio data system (RDS), 95
 - radiomottagare, 249
 - radiostrålning, 190
 - radiosändare, 248
 - reflekterad effekt, 196
 - reflektorantenn, 241
 - reflexionskoefficient, 197, 207
 - regn, 255
 - resonans, 230
 - riktantenn, 240
 - riktkopplare, 210
 - ringmodulator, 82
 - RR, 183
 - RS-232, 178

 - sampling, 142
 - samplingsteomet, 143
 - satellitmottagare, 103
 - SDLC, 180
 - secondary, 181
 - sektorantenn, 243
 - sessionslagret, 141
 - Shannons formel, 152
 - shot noise, 53
 - signal/bussförhållande, 58, 92, 106
 - signalvektor, 159
 - single sideband (SSB), 88
 - sinusformad spänning, 10
 - slitsad ledning, 208, 216
 - SNRM, 183
 - spegelfrekvens, 99, 100
 - spektrum, 28
 - spektrumanalysator, 30
 - spänningsstyrd oscillator (VCO), 112
 - S-ram, 182

-
- stackade dipoler, 240
 - stationär växelström, 19
 - stegindexfiber, 261
 - stereodekoder, 93
 - stereokoder, 93
 - stripline, 194
 - ström, 191
 - stubbe, 222
 - stympat sidband, 80
 - styrkristall, 135
 - stående våg, 205
 - stående vågförhållande (SVF), 207
 - superheterodynprincipen, 99
 - SVF-meter, 210
 - symbolhastighet, 148
 - synkron digital hierarki (SDH), 185
 - synkronisering, 168

 - telenätet, 185
 - termiskt brus, 46
 - tidsdomän, 31
 - time division multiplex (TDM), 185
 - transportlagret, 141
 - trigonometriska formler, 15
 - tråd, 255
 - tröskelsänkande FM-detektor, 134
 - tv-signal, 80

 - UA, 183
 - udda funktion, 36
 - undre sidband, 70, 88
 - upper sideband (USB), 70, 88
 - U-ram, 182

 - V.24, 178
 - V.32-modem, 168
 - V.34-modem, 154, 168
 - V.90-modem, 157
 - wavelength division multiplexing (WDM), 268
 - vertikal polarisation, 236
 - vestigial sideband (VSB), 80
 - vikningsdistorsion, 143
 - voltage controlled oscillator (VCO), 112
 - voltage standing wave ratio (VSWR), 207
 - vågledardispersion, 264
 - vågledare, 208
 - våglängd, 204
 - våglängdsmultiplexering, 268

 - XID, 183