

Mikrovågsteknik:
Effektförstärkare och linjärisering

Krister Andreasson

Mikrovågsteknik: Effektförstärkare och Linjärisering

Copyright © 2004, 2009 Krister Andreasson

Tryckt hos: Författares Bokmaskin
Stockholm 2004

ISBN 91-7910-594-7

Mångfaldigandet av innehållet i denna bok, helt eller delvis, är enligt lagen om upphovsrätt den 30 december 1960 förbjudet utan medgivande av copyright innehavaren.

Förord

Utvecklingen inom mikrovågsområdet har lett till att det finns en mängd färdiga komponenter att köpa. Dessa komponenter är uppbyggda som modulenheter med anpassad in- och utgång. Mikrovågsarbetet har delats upp i dels komponentkonstruktion och dels systemkonstruktion.

Systemkonstruktören behöver en beskrivande förklaring av funktionen, för att på bästa sätt utnyttja komponenterna. Dessutom behövs en översikt över samtliga komponenter för att kunna bedöma alternativa systemlösningar. Det gäller ju att välja den kombination som ger den enklaste och billigaste slutprodukten.

Komponentkonstruktören bör ha en bred översikt över alla komponenter för att få nya idéer till kretslösningar. Det gäller ju att utnyttja varandras erfarenheter så mycket som möjligt. Idag använder man datorer för att dimensionera och optimera kretsarna. Det underlättar den matematiska hanteringen avsevärt, men det är fortfarande lika viktigt att välja lämpliga kretskopplingar.

Servicetekniker och testingenjörer behöver en beskrivande förklaring, av en stor mängd kretskopplingar, som inte är belastad med matematisk dimensionering. En tekniker behöver snabbt kunna sätta sig in i ett stort blockschema eller en speciell testuppkoppling.

Även försäljare och inköpare behöver en översiktlig förståelse för mikrovågsmarknaden. De måste kunna förstå teknikernas önskemål.

Det finns alltså idag många som arbetar med mikrovåg, som behöver en bred kunskap om mikrovågskomponenter, samt en förståelse för de speciella egenskaper som utnyttjas på mikrovåg. Målsättningen med boken är att ge den översikt och introduktion som behövs för att arbeta med avancerade effektförstärkare.

Innehåll

1. INLEDNING	1
Uteffekt	1
Verkningsgrad	2
Effektförbrukning	3
2. VAL AV TRANSISTOR	4
BJT	4
HBT	4
FET	5
E-FET	5
P HEMT	5
CMOS	5
LDMOS	6
SiC och GaN	6
Mobiltelefoner	7
Basstationer	7
TV-sändare	7
Digital rundradio	7
Radar	7
3. FÖRSTÄRKARNAS OLIKA KLASSER	9
Klass A	10
Klass B	13
Parallellresonans	14
Expanding av förstärkningen	15
Klass C	16
Klass AB	17
4. ÖVERTONER	18
Klass E	19
Klass F	21
Inverterad klass F	24
Formad insignal	25
5. DIMENSIONERING AV EFFEKTFÖRSTÄRKARE	29
Anpassning	29
Storsignal S ₂₂	29

Fysisk modell	30
Harmonisk Balans	30
mm-våg	30
Load-Pull	31
Aktiv Load Pull	33
Cripps metod	36
6. DOHERTY	38
7. FÖRSPÄNNING	41
Gate-ström	41
Olika uteffekter	43
Varierad uteffekt	44
Förskjutning av arbetspunkt	45
Viloström	45
Genombrottspänning	45
Adaptiv förspänning	46
Förspänning av bipolär transistor	47
Genombrott i bipolär transistor	48
Låg drivspänning	50
8. LÅG STRÖMFÖRBRUKNING	51
Arbetspunkt	52
Effektförstärkare	52
Spänningsdelning med source-följare	54
9. INTERMODULATION	56
Intercept Point	57
Transistorns olinjäriteter	58
Sweet Spot	61
Minneseffekt	62
Olika transistortyper	63
BJT	63
CMOS	63
Pulsdopad FET	63
LDMOS	63
HBT	64
Jämförelse	64
Användning av överton och blandfrekvens	65
Exempel på kretskoppling	66
10. PASSIV INTERMODULATION — PIM	68
Oxidering	69
Koaxialkabel	70
Antenn	70
Mätning av PIM	71

11. EFFEKTKOMBINERING I TRANSISTORER	76
Instabilitet i effektförstärkare	77
Transistorns impedans	78
Anpassning med cell-cluster	79
Isolerande effektdelare	80
Förstärkarmodul	81
12. KRETSAR FÖR EFFEKTKOMBINERING	82
Hybrider med binär delning	82
Hybrider med udda delning	83
Wilkinson	85
Wilkinson med flera utgångar	87
Wilkinson för högre effekt	88
Isolationsmotstånd i ring	89
Wilkinson på enkla kretskort	90
Kontinuerligt avtagande ledning	91
Uppdelning i serie	92
Extended resonance	94
Radialkombiner	95
Radialkombiner i microstrip	96
Radialkombiner i vågledare	97
Sektorkombinering	98
Sammanställning	99
13. EFFEKT FÖRSTÄRKARE FÖR KW	101
Modulförstärkare	101
Slutsteg med radialkombiner	102
Slutsteg med seriekombinering	103
Övervakning	104
Tillförlitlighet	104
14. MCPA	106
Kombinering med kaviteter	107
Kombinering med hybrider	108
Kombinering på låg effekt	109
Multiport effektförstärkare	110
IM-distorsion	112
IM-brus i förhållande till signal	113
Felvektor	114
Signal-stör-förhållandet	114
Grannkanalstörning	115
Topp-effekt	117
Reducering av crestfaktorn	119
15. LINJÄRISERING	121
16. FÖRDISTORSION	124

Fördistorsion med diod	127
Diod i serie	127
Diod i shunt	128
Motställda dioder	128
Bättre kontroll av distorsionen	130
Användning av både reflektion och transmission	131
Distorsion med signalen undertryckt	132
AM-PM reglering med IQ-krets	133
Fördistorsion med FET	134
Expansion vid låga spänningar	135
FET som variabel resistans	136
Linjärisering med både Δf och $2f_0$	137
Cascode	138
5:e gradens IMD	139
Separat generering av distorsion	140
Digital fördistorsion	142
17. ADAPTIV FÖRDISTORSION	144
Kompensering av temperaturdrift	144
Styrning med skillnaden	145
Kompensering av nivåskift	145
Detektering av distorsionen	146
Detektering av störning i grannkanal	147
Grannkanalfilter med signalen utfasad	148
Digital fördistorsion	149
Digital inställning av fördistorsion	150
18. ÅTERKOPPLING	153
Återkoppling av intermodulation	153
Återkoppling av övertoner	154
Återkoppling av skillnadsfrekvensen	154
Återkoppling av envelop	155
Kartesisk återkoppling	157
19. FRAMÅTKOPPLING	160
Balansering av fas och amplitud	161
Fördröjning	162
Riktkopplare	163
Automatisk balansering av signalen	165
Automatisk balansering av IM	166
Reglering med pilotsignal	167
Reglering från tabell	168
Utan reglering	168
Dubbel framåtkoppling	169
Verkningsgrad	171
Fördelar med framåtkoppling	172

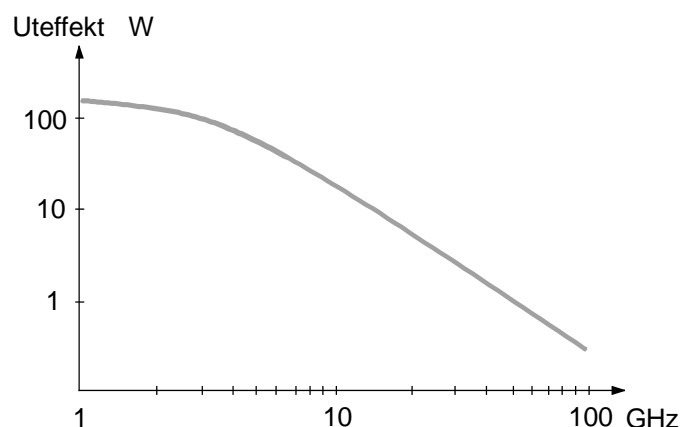
20. ENVELOP-ELIMINERING OCH ÅTERMODULERING	174
Återkoppling av envelop	175
Återkoppling av fas	176
Verkningsgrad	177
Distorsion från fördröjningen	177
Insignal på basbandet	178
Modulering av VCO	179
Bandbredd	179
Klass S kombinerat med Klass B	180
Fördistorsion av amplituden	181
Användning	181
21. LINC	183
Linjäriserad förstärkare	183
Förluster i summeringen	184
Linjäriserad sändare	185
LINC med fördistorsion	186
CALLUM	187
LINC med IQ-delning	188
Kombinering med Chireix	189
Jämförelse	191

Effekt förstärkare

1. Inledning

En effektförstärkare ska ha hög uteffekt, hög verkningsgrad och god linjäritet. De tre önskemålen går inte att uppnå samtidigt. Det blir en kompromiss beroende på applikationen. Många förstärkare kombineras i kretsar med låga förluster för att få högre uteffekt. Verkningsgraden förbättras med speciell hantering av övertonerna samt speciella kretsar till förspänningen. En förstärkare med högsta verkningsgrad har mycket dålig linjäritet. Det finns många sätt att förbättra linjäriteten.

Uteffekt



Figuren visar hur hög uteffekten kan vara på olika frekvensområden. Typiskt är att för högre frekvens ska halvledarnas dimensioner vara mindre. Det ger lägre genombrotts-spänning och lägre uteffekt. I batteridrivna handapparater används betydligt lägre effekt. Maximala effekten blir 1 eller 2 W vid 1-2 GHz. En förstärkare till radar innehåller många effektransistorer som kombineras till uteffekter på flera kW.

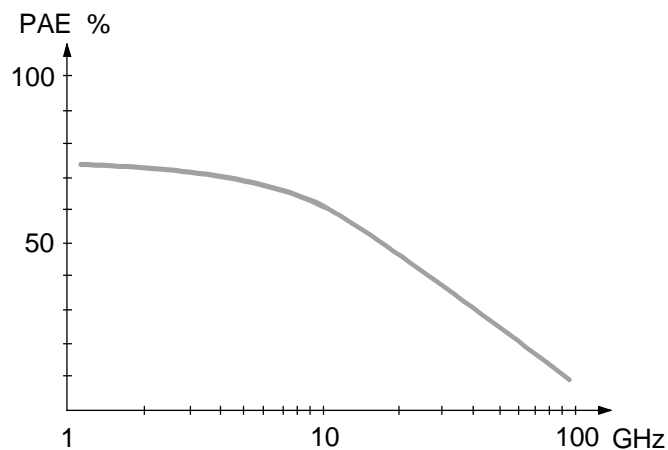
Verkningsgrad

$$\eta = \frac{P_{\text{ut}}}{P_{\text{DC}}}$$

η är verkningsgraden för drain (respektive kollektor), dvs. uteffekt i förhållande till DC-effekt.

$$\text{PAE} = \eta \cdot \left(1 - \frac{1}{G_p}\right)$$

PAE (Power Added Efficiency) är verkningsgraden för den totalt tillförda effekten, dvs uteffekt i förhållande till både DC-effekt och insignal. Om förstärkningen är mindre än 10 dB, får insignalens effekt en stor betydelse.



Figuren visar hur hög verkningsgraden kan vara för ett slutsteg med måttlig uteffekt. Förstärkare med mycket hög uteffekt har betydligt sämre verkningsgrad. Om det samtidigt ska vara god linjäritet blir ofta verkningsgraden bara några procent.

En radiolänk för telefoni eller TV innehåller en förstärkare på 10 - 25 W. Verkningsgraden för en TWT är 30 %, och en motsvarande halvledarförstärkare 20 %. Men för att få tillräcklig linjäritet minskas uteffekten (back-off). En TWT har vanligtvis en back-off på 6 - 8 dB. Det minskar verkningsgraden till 6 %. En halvledarförstärkare har bättre linjäritet och behöver bara 2 - 4 dB back-off. Det minskar verkningsgraden för halvledarförstärkaren till 10 %. Därigenom får halvledarförstärkaren bättre verkningsgrad än TWT.

Effektförbrukning

Om en 100 W klass A förstärkare behöver 1000 W från strömförsörjningen, så ska alltså förstärkaren kunna hantera 1000 W värmeutveckling. När en RF-signal kopplas till ingången, ger förstärkaren 100 W RF-effekt till belastningen. Transistorn behöver då bara absorbera 900 W. Den blir alltså inte lika varm under drift som utan insignal.

En 100 W klass AB förstärkare ger mycket mindre än 100 W effektförlust utan insignal. Med insignal kan effektförbrukningen öka till 500 W. Transistorn blir alltså inte lika varm som vid klass-A, och kan klara sig med enklare kylning.

En klass C förstärkare drar ingen ström, då det inte finns någon insignal. Effektförbrukningen blir alltså lägre än för klass AB.

En transistor klass A är konstruerad för att kunna ge ordentlig kylning. Det behövs en större transistor som ofta är uppdelad i flera deltransistorer, för att fördela värmeutvecklingen.

Om en klass C transistor ska användas i en klass A applikation, behöver den kunna hantera 4 gånger så hög effekt. En klass C transistor på 12 W ger endast 3 W i klass A.

Skillnaden mellan Klass C och klass AB är ca 3 gånger i effekt. Klass C transistorn på 12 W ger då 4 W i klass AB.

2. Val av transistor

BJT

Bipolära transistorer i kisel har använts på UHF och L-bandet (1-2 GHz). BJT har uppnått flera hundra watt pulseffekt på L-bandet. På S-bandet (2-4 GHz) går det att använda BJT, men prestanda blir sämre. Jämfört med LDMOS och FET kan BJT ge högre uteffekt. Men förstärkningen är bara ca 8 dB och den har högre distorsion.

HBT

I en HBT begränsas uteffekten av värmeutvecklingen istället för elektriskt genombrott. Det går därför att få 3 - 4 dB högre uteffekt för korta pulser (μ s). Men en nackdel med HBT är att den har ganska hög termisk resistans.

En HBT behöver endast en positiv spänningskälla. Det är speciellt fördelaktigt vid små batteridrivna apparater.

Den största olinjäriteten kommer från modulationen av bas-kollektor kapacitansen. Kapacitansens inverkan kan minskas med hjälp av en kortare kollektor och en lämplig dopningsprofil.

Vanligtvis används förstärkarkopplingar med gemensam emitter. Visserligen har HBT mycket högre genombrotts-spänning då den är kopplad med gemensam bas, men det blir stora problem med stabiliteten på de högre frekvenserna. Basen behöver ju förspänning och måste anslutas till jord via en kondensator. Kondensatorer har alltid en viss ströinduktans. Ovanför egenresonansen är kondensatorn induktiv. Stabiliteten för kopplingen med gemensam-bas är extremt känslig för induktans i basen.

Inimpedansen i en HBT bestäms främst av ströresistanserna i bas och emitter. Därför ändras inte inimpedansen av stora insignaler.

En HBT är en bipolär transistor, och kan därför ge termisk strömrusning. Ett extra motstånd i emittern eller basen motverkar strömrusningen. I basstationer, där det finns hög förstärkning, används främst emittermotstånd. Men motståndet ger ett spänningsfall som minskar tillgänglig spänning över transistoren. Därför används motstånd i basen till batteridrivna handapparater.

FET

Uteffekten begränsas av genombrottspänningen. Pulseffekten blir alltså inte speciellt högre än för kontinuerliga signaler.

Högre frekvenser kräver kortare gate-längd. Det för med sig att genombrottspänningen blir lägre. Uteffekten minskar alltså med kvadraten på frekvensen.

Det blir högre verkningsgrad för drain då strömmen genom transistorn minskar, dvs då man går från klass A till klass B. Men när strömmen genom en FET blir liten, blir också förstärkningen liten. Det betyder att den totala verkningsgraden PAE blir lägre. Därför används nästan alltid klass AB för effektförstärkare med FET. Det ger också mycket lägre distorsion än klass B.

E-FET

En E-FET har gate nersänkt så långt att avlänkningsområdet stänger av kanalen, utan att behöva negativ spänning på gate. Strömmen ökas med en positiv spänning på gate (Enhancement mode). Om transistorn är dimensionerad att ge halva maxströmmen vid $V_G = 0$ V, arbetar den i sitt linjära område (klass A) utan att behöva negativ förspänning.

Upp till 2 GHz kan en Pseudomorfsk E-HEMT ge 2 W med 61 % verkningsgrad. E-FET har fått bättre IM_3 än D-FET i klass AB på hög effekt.

P HEMT

En pseudomorfsk HEMT har högre verkningsgrad än både BJT och en del FET. Den kan användas upp till 10 W på mikrovåg. På mm-våg är det genomgående P HEMT som används. De ger några watt upp till 34 GHz och ett par hundra mW upp till 94 GHz.

CMOS

CMOS och BiCMOS har inte tillräckligt hög genombrottspänning för att ge mer än ett par hundra mW vid 2 GHz (0,5 W vid 1 GHz). Men för system med låg drivspänning och låg effekt, t.ex. Bluetooth, är den mycket lämplig. CMOS kan användas till förstärkare med drivspänning så låg som 0,9 till 1,65 V. Det betyder att det räcker med endast en battericell av NiCd. Med CMOS på kisel kan större del av systemet integreras monolitiskt jämfört med GaAs.

LDMOS

LDMOS är lämplig för effekter högre än 1 W, upp till ca 2,5 GHz. Förstärkaren med LDMOS försämras inte så mycket då lastens impedans varierar. Den kan klara VSWR=10 vid maximal uteffekt. Bipolära transistorer kan bara klara VSWR=3. En bipolar effektransistor behöver en skyddande ferritisolator på utgången. En MOSFET kan klara sig utan isolator.

LDMOS ger ca 12-13 dB förstärkning. Bipolära transistorer ger endast 9 -10 dB förstärkning. Högre förstärkning och verkningsgrad än andra effekt-transistorer betyder lägre systemkostnader på grund av färre förstärkarsteg och lägre effektförbrukning. En annan fördel är högre impedans, som alltså ger enklare anpassning.

LDMOS förspänd till klass AB ger bättre linjäritet än både FET och BJT. Men för att klara den linjäritet som WCDMA kräver, behövs ändå en minskning av signalen (back-off) på 10 - 13 dB.

LDMOS har mjuk limitering. Det kan vara fördelaktigt med mjuk limitering då topeffekten är mycket större än medeffekten, t.ex. vid flera signaler eller vid AM-modulering.

En nackdel med MOS transistorer är att spänningen som behövs för att transistorerna ska börja leda (Threshold Voltage), varierar mellan olika exemplar av transistorer. Dessutom kan V_{th} variera 5-20 % över transistorens livslängd. Största förändringen sker de första 5 timmarna. Transistorerna bör alltså vara för-åldrade (inbrända) innan förspänningen ställs in. Förspänningen kan dessutom vara automatiskt reglerad från medelströmmen eller förstärkningen. V_{th} är ca 2-3 V för en MOS transistor på hög effekt. System för mobiltelefoni ställer höga krav på sändarens brus i mottagarbandet. En MOSFET ger ca 5 dB lägre brus än en HBT.

SiC och GaN

Kiselkarbid SiC och galliumnitrid GaN är materialkombinationer med större bandgap än GaAs och Si. Det ger 5 – 6 gånger högre genombrotts-spänning.

SiC har mycket god termisk ledningsförmåga. Kisel leder värme 3 gånger bättre än GaAs och SiC så mycket som 10 gånger bättre än GaAs. GaN kan läggas som ett epitaxialskikt på SiC för att få god kylning.

En LDMOS på 25 W kan ha en strökapacitans på utgången på ca 25 pF. Motsvarande transistor i SiC har en strökapacitans på < 1 pF.

Mobiltelefoner

Till Amerikanska och Japanska mobiltelefoner (DAMPS, CDMA, PDC) används MESFET, P-HJFET eller HBT. Till Europeiska mobiltelefoner (GSM, DCS) används bipolära transistorer och LDMOS, samt GaAs MMIC på 1800 MHz.

Amerikanska CDMA och TDMA behöver en linjär förstärkare. Max uteffekt bör ligga 3 dB från mättad uteffekt (3 dB back-off). GSM-telefoner arbetar på konstant nivå mycket närmare mättnad.

Verkningsgraden kan bli ca 65 % för mättad uteffekt och ca 40 % för linjär förstärkning.

Basstationer

Basstationerna till GSM är uppdelade i olika effektklasser från 2,5 W till 320 W beroende på cellstorlek. GSM makroceller med uteffekter på 20 - 40 W har byggts med LDMOS på 50 - 80 W. Verkningsgraden för drain är ca 55 % vid 900 MHz och ca 43 % vid 2 GHz. Totala verkningsgraden för en effektförstärkare kan vara 15 - 30 %.

UMTS behöver flera hundra watt. Med en topp effekt som kan vara 18 dB högre än medeffekten, blir totala uteffekten ca 500 W vid 2 GHz. Den effekten kan uppnås med några effektkombinerade LDMOS.

TV-sändare

En LDMOS-förstärkare på 150 W CW (kontinuerlig effekt), kan täcka hela TV-bandet 470 - 860 MHz. Många kombinerade förstärkare kan tillsammans ge uteffekter på flera kW.

Digital rundradio

Digital rundradio (DAB) i Europa använder OFDM med 192 till 1536 bärvågor som var och en moduleras med olika bitströmmar. Internt matchad LDMOS i push-pull ger 135 W inom frekvensområdet 1,4 - 1,6 GHz.

Radar

En radar kan vara baserad på en mycket stabil syntesgenerator, för att uppnå god undertryckning av klutter. Signalen förstärks till pulseffekter på 1 - 30 kW eller mer.

Referenser

Böcker

Steve C Cripps, "RF power amplifiers for wireless communications", Artech House 1999 ISBN 0-89006-989-1

Nick Pothecary, "Feedforward linear power amplifiers", Artech House 1999 ISBN 0-1-58053-022-2

Peter B Kennington, "High linearity RF amplifier design", Artech House 2000 ISBN 1-58053-143-1

Diverse

C E Weitzel, "RF power devices for wireless communications", IEEE MTT-S 2002 pp285-288

Carsten Fallesen, "A CMOS power amplifier for GSM-1800 with 55 % PAE", IEEE MTT-S pp 911-914

Fazal Ali, "Design considerations for high efficiency GaAs HBT power amplifiers", EuMC 1994 pp 156-176

Nicolas Constantin, "GaAs FETs gate current behaviour and its effects on RF performance and reliability in SSPAs", IEEE Transactions on microwave theory and techniques, vol 43 no 12 dec 1995 pp 2918-2924

Hitoshi Hayashi, "Quasi linear amplification using self phase distortion compensation technique", IEEE Transactions on microwave theory and techniques, vol 43 no 11 nov 1995 pp 2557-2564

Adam L Martin, "A 33 GHz power amplifier based on an extended resonance technique", MTT-S 2000

Rolf Lohrmann, "A novel distributed multicell multistage amplifier structure", EuMC 2003 pp 379-382

3. Förstärkarnas olika klasser

Förstärkarna delas in i olika klasser beroende på hur arbetspunkten ställs in med transistorers förspänningar. Valet av arbetspunkt bestäms av kompromissen mellan linjäritet (distorsion) och verkningsgrad (effektförbrukning).

Klass	Arbetspunkt % av I_{\max}	Ledande del av perioden	Verkningsgrad % teoretiskt
A	50	360	50
AB	0 - 50	180 - 360	50 - 78
B	0	180	78
C	0	0 - 180	78 - 90

Dessutom finns det några klasser som inte är en beskrivning av arbetspunktens inställning.

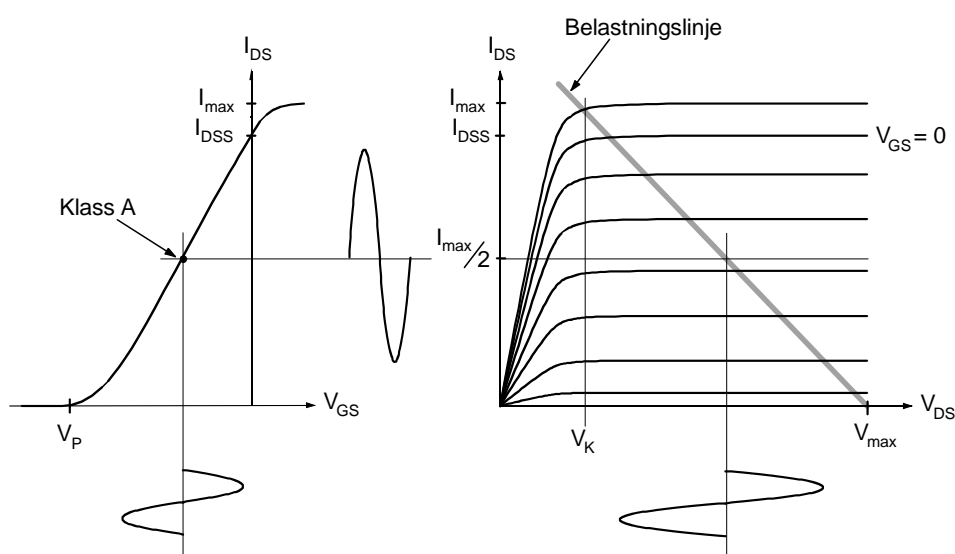
Klass D och E använder transistorerna som en switch. Klass D används enbart på lägre frekvenser. Det är vanligt med klass D i AM-sändare. Klass E har provats upp till ca 1 GHz.

Klass F och inverterad klass F utnyttjar övertonernas belastningsimpedans för att öka verkningsgraden. Inverterad klass F ger högsta verkningsgraden, men kräver också högre genombrottspänning.

Om övertonerna måste belastas med resonanskretsar, blir förstärkarna mycket smalbandiga. Med filter blir bandbredden högst 1,8 (dvs mindre än en oktav). Bredbandigare förstärkare behöver arbeta i klass A, som inte alstrar några starka övertoner. Klass B i push-pull kan också bli ganska bredbandiga eftersom de jämna övertonerna blir undertryckta.

Klass A

Arbetspunkten för klass A ligger mitt i transistorns linjära område. Det ger den lägsta distorsionen. Nackdelen är hög strömförbrukning, till och med då det inte finns någon RF-signal. Klass A ger alltså lägsta verkningsgrad. Dessutom behöver transistor vara speciellt konstruerad för att tåla den höga värmeutvecklingen.



De flesta datablad anger I_{DSS} , som är strömmen mellan drain och source då $V_{GS} = 0V$. Men Schottky-övergången till gate har ett avlänkningskikt även i vila. Det betyder att transistoren inte är helt öppen. Strömmen i transistoren kan ökas upp till 20 % genom att förspänna Schottky-dioden i framriktningen ca 0,5 V. I praktiken kan termiska problem och tillförlitlighet begränsa strömmen till I_{DSS} istället för maxströmmen I_{max} . Arbetspunkten ligger då på $I_{DSS}/2$.

V_P är spänningen då transistoren stryps. Området närmast V_P är ganska kvadratisk, för att sedan övergå till ett linjärt område.

Knäspänningen V_K är den drainspänning som behövs för att transistoren ska kunna ge den önskade strömmen. Drainspänningar lägre än knäspänningen är ett oanvändbart område, som enbart minskar verkningsgraden. Knäspänningen (ca 1,2 V) i en FET kan göras lägre genom att minska resistansen i source. Avståndet mellan gate och source bör också göras så litet som möjligt.

Högsta spänningen över transistorn V_{\max} begränsas av genombrott i transistorn. Genombrottet (lavineffekten) sker mellan drain och gate, eftersom gate är negativt och drain är positivt förspänd. Databladet anger genombrott-spänningen V_{DGB} . Högsta spänningen på drain blir då:

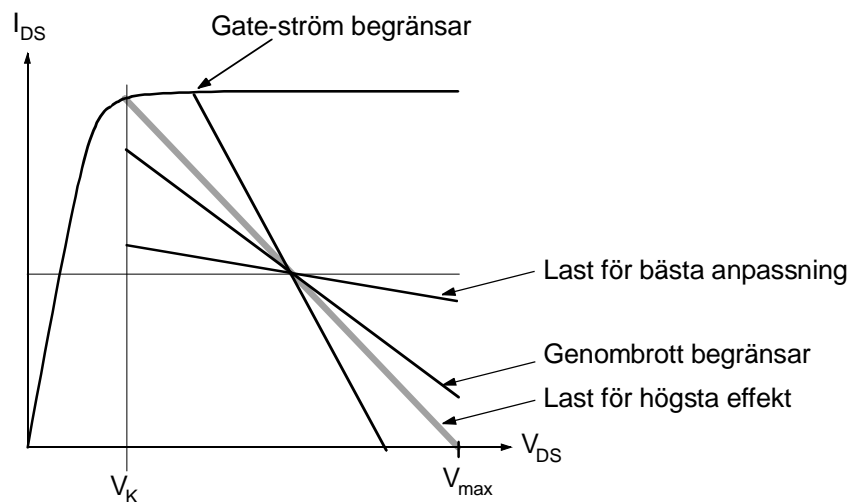
$$V_{\max} = V_{DGB} - V_P$$

Transistorn ger högsta uteffekt då variationerna i spänning sträcker sig från V_K till V_{\max} , och strömmen varierar upp till I_{\max} . Den optimala belastningsresistansen för högsta uteffekten ska alltså vara:

$$R_{\text{optP}} = \frac{V_{\max} - V_K}{I_{\max}}$$

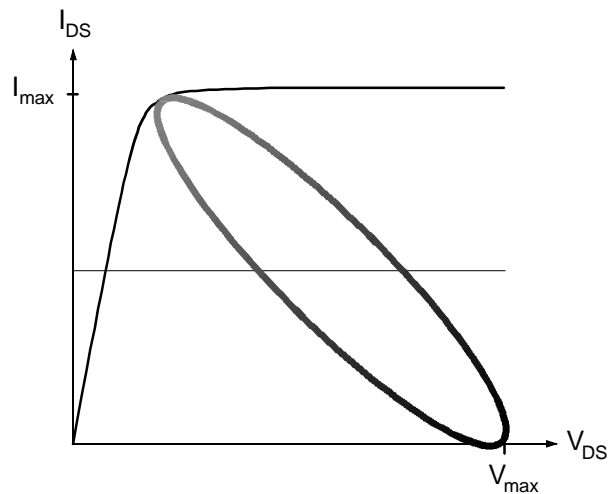
Den högsta RF-effekten blir då:

$$P_{\text{RFmax}} = \frac{1}{8} \cdot I_{\max} \cdot (V_{\max} - V_K)$$



Om lasten inte är optimerad för högsta uteffekt blir utstyrningen i ström respektive spänning mindre. Signalen klipps på grund av gateström respektive genombrott.

Belastningsresistansen för bästa anpassning är vanligen högre än för högsta uteffekt. Det betyder att en effektförstärkare alltid är missanpassad på utgången. Anpassningen får förbättras med en cirkulator eller hybridkoppling.

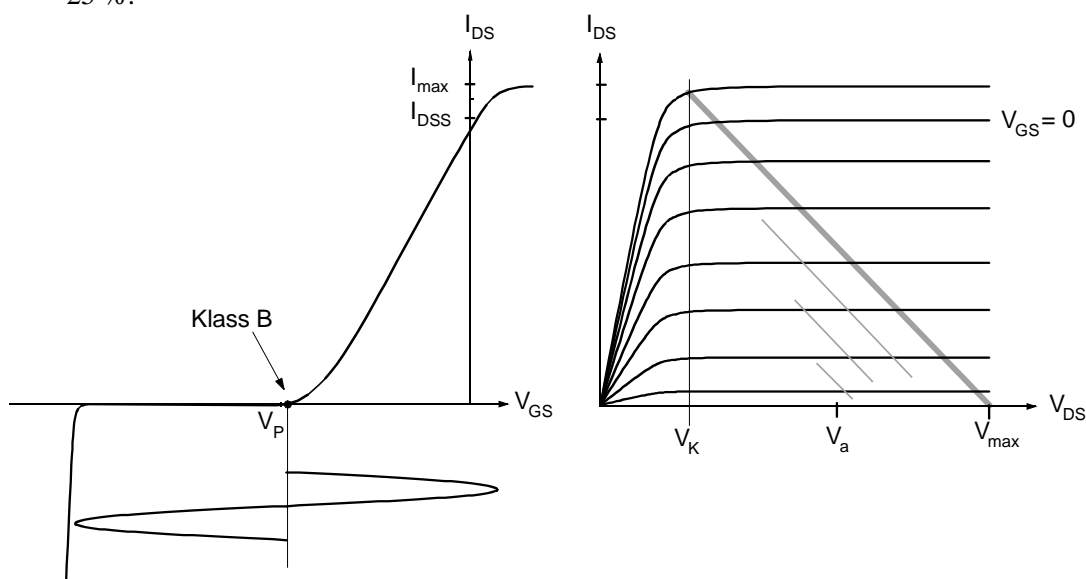


Transistorn har strökapacitans mellan drain och source, som laddas upp av signalen. Det gör att belastningslinjen istället blir en ellips. Men om man betraktar ströreaktanserna som en del av anpassningskretsarna och flyttar in gränssnittet, så blir det en resistiv linje för belastningen igen.

I praktiken kan belastningen vara ganska komplex med resonanser. Det gör att även belastningslinjen blir ganska komplex med slingor.

Klass B

Klass B är förspänd till strypning, V_P . Eftersom den inte har någon ström-förbrukning i vila blir verkningsgraden hög och kylproblemen måttliga. Vid lägre insignaler (Back-off) försämras inte verkningsgraden lika snabbt som för klass A. En klass B förstärkare får ganska hög verkningsgrad och konstant förstärkning över 10 dB dynamikområde. När amplituden minskas 10 dB i en klass A förstärkare, så har verkningsgraden minskats från 50 % till ca 5 %. Klass B har endast minskat till 25 %.



Transistorn leder enbart ström under den positiva halvperioden. För att få högsta uteffekt ska signalen vara så stor att den når fram till I_{max} , det vill säga förspänning av gate i framriktningen. Den negativa halvperioden blir då så stor som $2 \cdot V_P$. Spänningen på drain är då så stor att signalen precis når fram till genombrott.

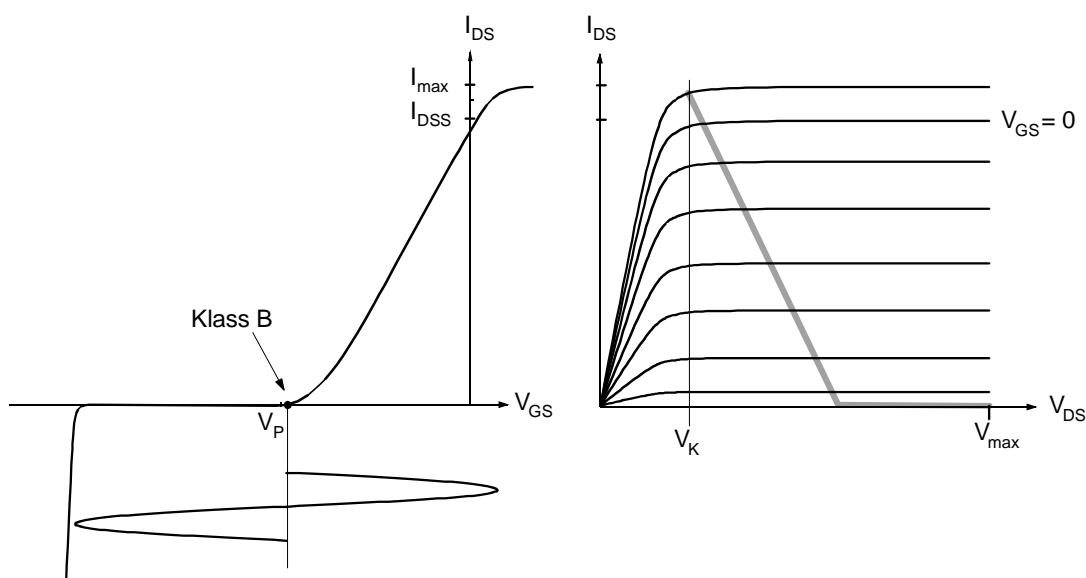
Jämfört med klass A ska alltså insignalen ha dubbelt så stor amplitud. Det betyder att klass B har teoretiskt 6 dB lägre förstärkning. I praktiken är försämringen bara 3-5 dB, efter det att anpassningen på ingången optimerats.

Drainspänningen V_{DS} är inställd till ungefär samma arbetspunkt som klass A. När insignalen ökar i styrka så ökar strömmens DC-medelvärde. Vid maximal utstyrning har transistorn ungefär samma variationsområde som för klass A. Belastningsresistansen är alltså densamma som för klass A.

Om förstärkaren inte är mättad, är uteffekten direkt proportionell mot insignalen. På så sätt kan den betraktas som en linjär förstärkare. Men eftersom strömmen består av endast ena halvperioden, innehåller utsignalen alltid övertoner. Andra övertonen ($2 \cdot f$) ligger bara 7,4 dB under grundtonen.

Parallellresonans

Utsignalen kan bli sinusformad igen med hjälp av en parallellresonanskrets tvärs över lasten. Eftersom strömkurvan inte innehåller några udda övertoner, behövs alltså inte någon kortslutning på udda övertoner. På mikrovåg kan därför en kvartvågsledning användas istället för resonanskrets.



Parallellresonanskretsen kortsluter övertonerna. Enbart strömmens grundton går till lasten. Spänningen över transistorn blir då sinusformad. Belastningslinjen ser lite annorlunda ut, men belastningsresistansen och uteffekten är densamma som förut. Lastresistansen är alltså inte lutningen på belastningslinjen.

Resonanskretsen på utgången förhindrar att övertonerna ser en resistiv belastning. Utan effektförluster på övertonerna blir det lite högre verkningsgrad. Nackdelen är att förstärkaren blir smalbandig. Med filter istället för en resonanskrets blir det lite större bandbredd.

Det är opraktiskt att använda en FET i klass B eftersom förstärkningen blir så liten i närheten av strypning. Dessutom blir genombrott-spänningen för låg, och läckströmmen vid strypning (pinch-off) blir för stor. Därför används inte FET i klass B eller C.

Vanligtvis används FET i klass A eller AB. Det innebär att det går DC-ström i vila. En batteridriven apparat behöver en särskild switch för att koppla bort den effektförbrukande förstärkaren, då det inte behövs någon RF-effekt.

Med bipolära transistorer går det däremot bra. En HBT kan användas i klass B eftersom transkonduktansen ökar exponentiellt då den börjar leda (turn on). Det ger liten knäspänning och hög verkningsgrad. En ganska tjock och svagt dopad kollektor ger hög genombrott-spänning mellan bas och kollektor. Det ger hög effektförstärkning.

Klass B ger också ett extremt lågt strömläckage för en HBT. Om insignalen är så liten som 0 dBm räcker den inte för att få transistorerna att börja leda. Förstärkaren förbrukar då ingen DC-effekt. Det behövs alltså ingen extra DC-switch för att stänga av den till viloläge.

Klass B används främst då två transistorer arbetar i push-pull.

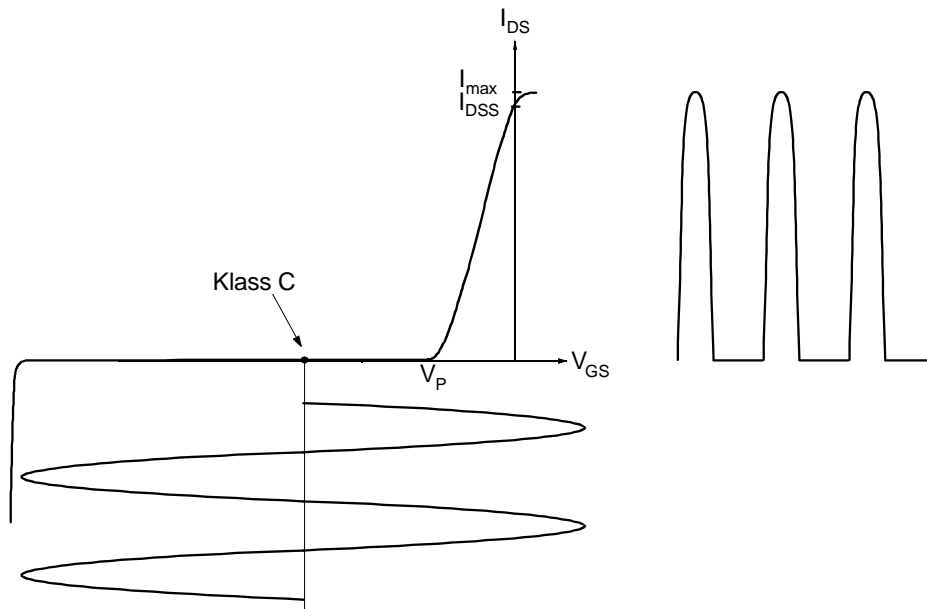
Expanding av förstärkningen

I en HBT klass B ökar bas- och kollektorströmmen då insignalen ökar. Det ger en förskjutning mot klass AB. Ökningen i ström ger högre förstärkning (gain expansion). Förstärkningsändringen ger signalen distorsion som försämrar grannkanalstörningen, ACLR.

Ett motstånd i serie med basens strömförsörjning kommer att motverka förskjutningen mot klass AB. En ökning av basströmmen ger ett spänningsfall över motståndet, som ger en minskning av strömmen.

Linjäriteten kan alltså förbättras i en HBT genom att förstärkningens expansion undertrycks. En FET har ingen gateström och ger alltså inget kompensande spänningsfall.

Klass C



Klass C innebär att transistorn är strypt under större delen av perioden. Endast de positiva topparna ger ström i transistorn. Insignalen ska vara så stor att strömmen helst når upp till I_{max} .

Strömmen består av korta pulser med låg DC-komponent. Det kan ge mycket hög verkningsgrad. Nackdelen är att de korta pulserna ger lägre amplitud på grundtonen. Till det behövs dessutom en ganska stor insignal. Förstärkningen är alltså betydligt lägre. Den önskade utsignalen filtreras fram ur strömpulserna med hjälp av en resonanskrets. Det gör klass C förstärkarna smalbandiga.

Klass C kan användas när det behövs extra hög verkningsgrad, till exempel vid satellitkommunikation. Satellitsystemen kräver däremot inte lika hög linjäritet som markburna system. Modulation med konstant amplitud gör att förstärkaren hela tiden precis uppnår mättnad. Det ger högsta verkningsgrad.

Uteffekten beror på hur stor drivspänningen är. Uteffekten kan alltså direkt regleras med drivspänningen.

Bipolära transistorer kan användas i klass A, B eller C. En FET för hög effekt har en tjock kanal, som behöver en stor negativ spänning för att avlänkas till strypling. Dessutom är genombrottspänningen ganska låg i en FET. Därför används FET transistorer nästan enbart i klass A eller klass AB på frekvenser över ca 6 GHz.

En HBT i klass C behöver ingen strömförsörjning till basen. RF-signalen är tillräckligt stor för att göra transistorn ledande vid de positiva topparna. Utan insignal är transistorn helt avstängd, den drar alltså ingen ström i vila. Kollektorspänningen kan lämnas på hela tiden.

Klass AB

Den låga förstärkningen i en FET nära strypning kan undvikas genom att förspänna till 10 - 20 % av I_{DSS} . Arbetspunkten blir en kompromiss mellan linjäritet och verkningsgrad.

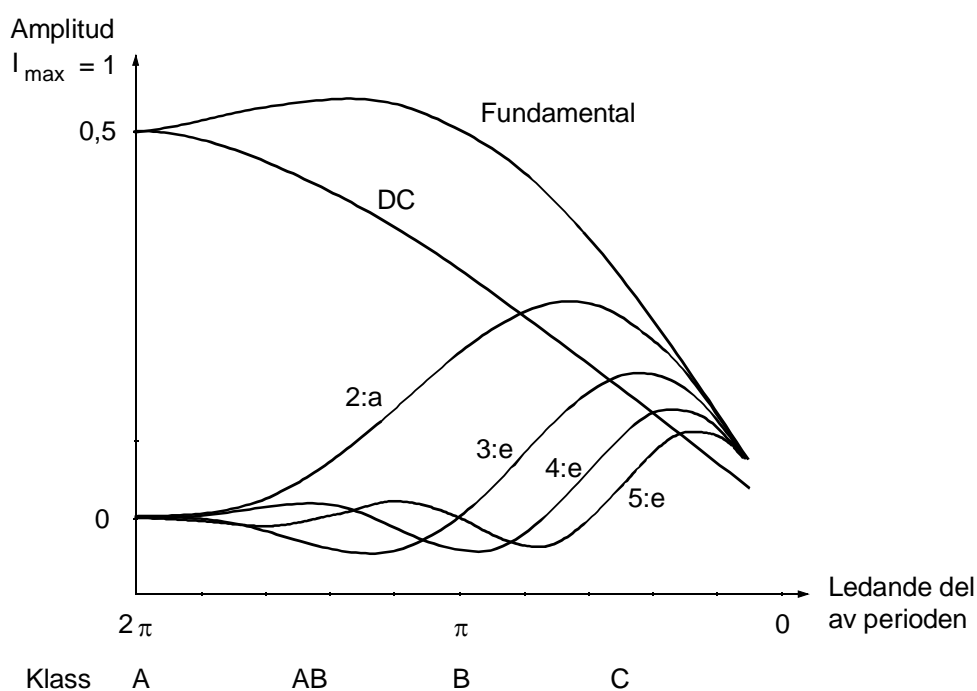
En klass A förstärkare kan få lägre distorsion genom att minska insignalen (back-off). En klass AB förstärkare har ganska konstant IMD över ett stort dynamikområde. Men så fort som signalen blir så stor att den börjar klippas försämras snabbt IMD. En IMD på -60 dBc tolererar i stort sett ingen klippning.

De flesta transistorerna är inte direkt optimerade att arbeta på så låga strömmar som i klass AB. Det går att få högre förstärkning nära strypning om transistorn är mycket svagt dopad närmast gate, och sen har en mycket starkt dopad väl definierad kanal för strömmen. Den dopningsprofilen kallas Spike-doped, Step-doped eller Delta-doped.

Klass AB har lägre värmeutveckling än klass A. Den behöver därför mindre kylning. Men det för också med sig att den inte kan hantera reflekterad effekt lika bra. Det kan behövas skyddskretsar som drar ner insignalen om belastningen ändras för mycket.

4. Övertoner

En klass A förstärkare för små signaler arbetar i sitt linjära område. Klass B förstärkaren har en ström i form av en halv sinuskurva. En likriktad sinussignal innehåller jämna övertoner. Strömmen i klass C förstärkaren består av rundade korta pulser. Den innehåller därför både udda och jämna övertoner.

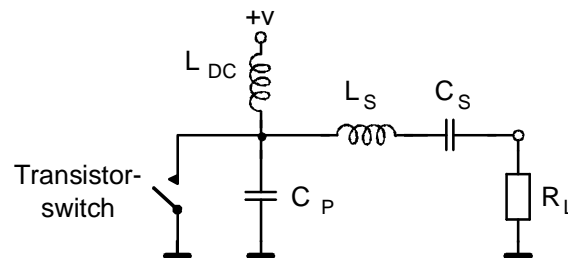


DC-komponenten minskar successivt då den ledande delen av perioden minskar. Den önskade fundamentala signalen, det vill säga grundtonen, är ungefär lika stor i klass A som i klass B. Lika stor signal och mindre förluster betyder högre verkningsgrad.

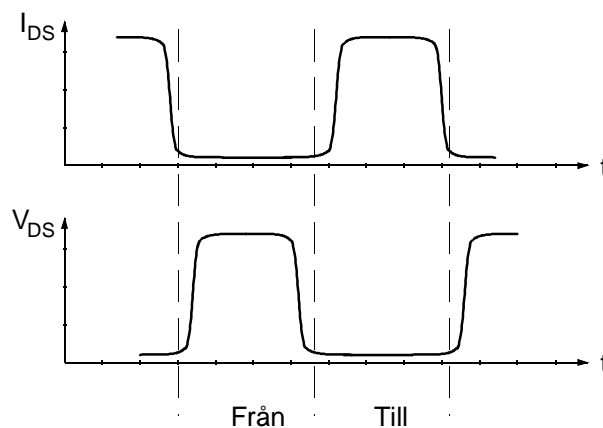
Klass AB innehåller dubbla frekvensen, det vill säga 2:a övertonen. Att den är positiv innebär att den ligger i fas med insignalen. Signalens vågform blir lite plattare på den negativa sidan och blir toppigare på positiva sidan. Resultatet är lägre medelström för samma topp-till-topp värde.

De udda övertonerna är noll för klass B. Men för klass AB är den 3:e övertonen betydande. Förspänningen kan justeras runt klass B för att kontrollera 3:e övertonen.

Klass E



Transistorn drivs som en switch. När switchen är till går det ström genom L_{DC} . När switchen slår från kommer den lagrade magnetiska energin att driva ström till C_P och utgångsnätet. Serieresonanskretsen gör att endast grundtonen ser en resistiv belastning.

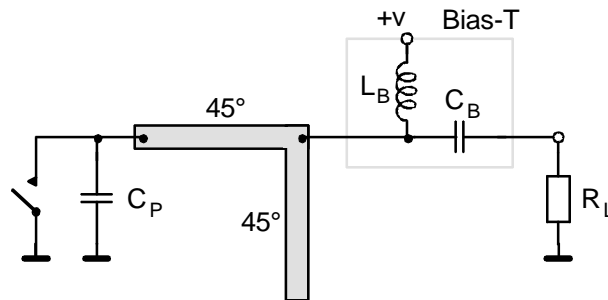


När switchen är till går det stor ström genom transistorn. Spänningen över transistorn är nästan noll, eftersom transistorn har mycket låg resistans när den är till. Med switchen i läge från går det ingen ström i transistorn. Istället är spänningen hög.

När spänningen är hög är strömmen noll och när strömmen är hög så är spänningen noll. Det ger hög verkningsgrad. Om det skulle finnas spänning och ström samtidigt blir det effektutveckling i transistorn. Det minskar verkningsgraden. På mikrovåg är omslagstiden en betydande del av perioden. Transistorn styrs därför ut så att spänningen blir fördröjd, samt går ner till noll innan strömmen ökar.

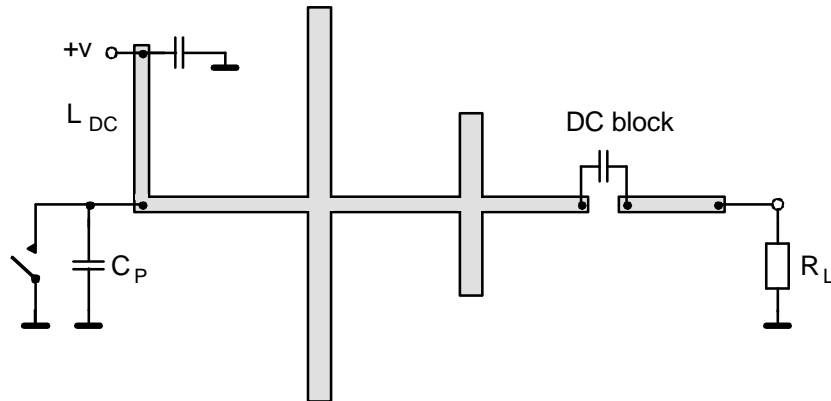
En fördel med klass E är att den fungerar med en betydande kapacitans på transistorns utgång. Transistorns strökapacitans kan inkluderas i den totala kapacitansen C_P .

Förstärkaren fungerar bara för signaler med konstant amplitud, till exempel GSM. Uteffekten kan i viss mån justeras med spänningskällan utan att verkningsgraden försämras. Spänningen över transistorn (V_{DS}) blir ca 3 gånger så stor som spänningskällan (+V).



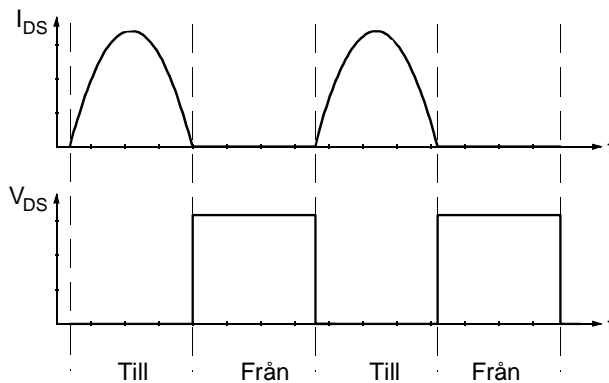
I praktiken räcker det med att 2:a övertonen ser en öppen avslutning för att det ska bli klass E. Ledningslängderna är 45° långa på grundtonen. De är alltså en kvarts våglängd på dubbla frekvensen. Kretsen ger alltså transistorn en öppen avslutning för 2:a övertonen. Ledningarnas impedans optimeras för grundtonen. Kretsen skulle också kunna innehålla en stubbe för 3:e övertonen.

De största förlusterna kommer från transistorn då den är i läget till. Kretsarna har högt Q-värde och behöver avstämmas mot komponenternas toleranser. De behöver också efterjusteras för att kompensera för komponenternas åldring.



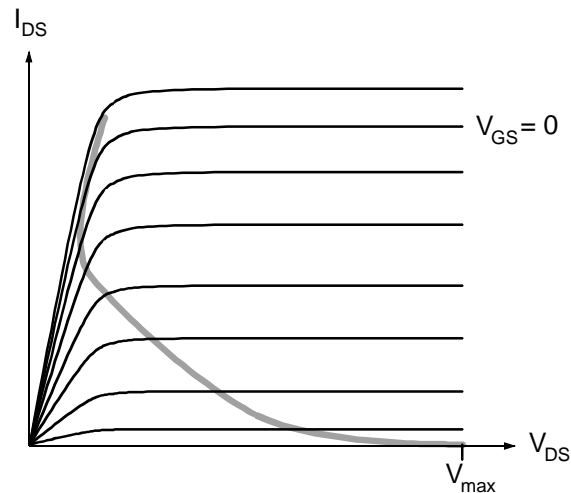
Kretsen kan innehålla flera stubbar för flera övertoner. Längderna är $\lambda/4$ för respektive frekvens. Den öppna stubben transformeras till ett avbrott vid transistorn. Impedanserna väljs så att grundtonen transformeras till 50Ω .

Klass F

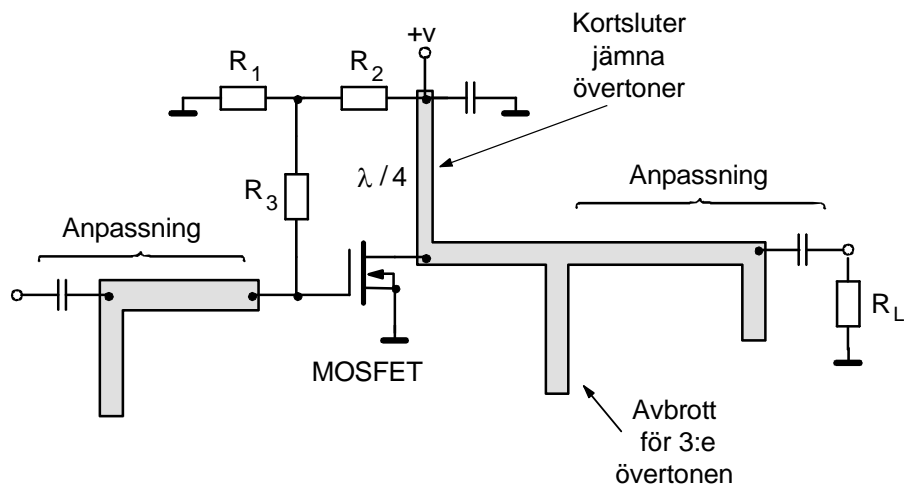


En klass B förstärkare har en ström i form av halv sinuskurva. För att få högsta verkningsgrad ska spänningen vara noll då det går ström, och max då strömmen är noll. Spänningens kurvform ska alltså vara en fyrkantvåg.

Klass F är den belastning som säkerställer dessa kurvformer. En halv vågs-likriktning innehåller enbart jämna övertoner, och en fyrkantvåg innehåller enbart udda övertoner. Belastningen ska alltså vara kortslutande för jämna övertoner och uppvisa ett höghmigt avbrott för udda övertoner.



Transistorn är förspänd till klass B, och har en belastningslinje som går in i området där strömmen också beror på gatespänningen. Under de omständigheterna kan spänningen på drain bli fyrkantig, om övertonerna är lämpligt avslutade. Den vertikala delen av belastningslinjen är alltså typiskt för klass F. Det är här som fyrkantvågen får sin tillplattade botten.



Kortslutning av de jämna övertonerna sker med kretsarna för transistorens förspänning. En kortsluten $\lambda/4$ ledning stör inte grundtonen eller de udda övertonerna. Däremot blir de jämna övertonerna kortslutna. Bandbredden blir större om stubbens impedans minskar. En bredare ledning ser alltså ut som en kortslutning över ett större frekvensområde.

En öppen stubbe som är $\lambda/4$ för 3:e övertonen ger en virtuell kortslutning. Denna kortslutning transformeras längs en ledning till ett avbrott vid transistorn.

Genom att på lämpligt sätt avsluta 2:a och 3:e övertonen kan verkningsgraden för drain bli 75 %, det vill säga ungefär som för klass B. Om man avslutar 3:e och alla jämna övertoner, kan verkningsgraden för drain komma upp till 88 %. I praktiken handlar det mer om att nå upp till 70 %. Men för att uppnå riktigt höga verkningsgrader ska även ingången vara reaktivt avslutad för 2:a övertonen.

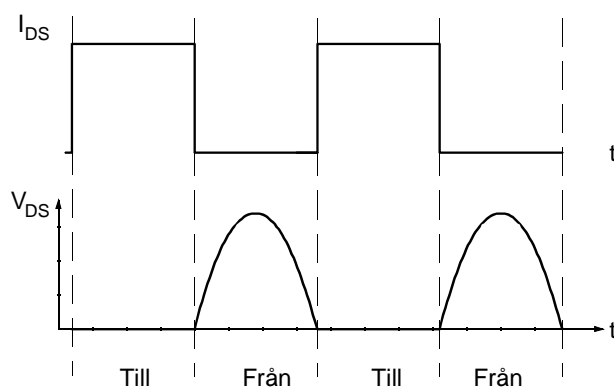
Anpassningen på in- och utgång består av T-kretsar med serieledning, öppen stubbe samt diskret kondensator.

Eftersom kretsen avslutar övertonerna reaktivt, är det bara den sinusformade grundtonen som når lasten på utgången. Övertonerna har alltså filtrerats bort från utgången. En annan fördel är att verkningsgraden blir högre då övertonerna inte ser någon resistiv last.

En P-FET har en kapacitans på gate som varierar kraftigt med spänningen på gate. Distorsion av den sinusformade signalen på gate påverkar vågformen på drain. Resultatet blir en försämring av verkningsgraden. Distorsionen kan elimineras genom att kortsluta alla övertoner. I praktiken räcker det med låg impedans på 2:a övertonen för att återfå det mesta av verkningsgraden. En vanlig MESFET har inte lika stora problem med kapacitansvariationer.

Inverterad klass F

Klass F ser kortslutning på 2:a övertonen och avbrott på 3:e övertonen. Inverterad klass F ser avbrott på 2:a och kortslutning på 3:e övertonen.



Övertonerna ändrar vågformen på ström och spänning. 3:e övertonen drar ner toppen på strömmen så att strömkurvan blir mer fyrkantig. 2:a övertonen adderas till spänningen så att spänningskurvan blir mer toppig.

Strömkurvan för klass F är toppig. Högre ström betyder mer förluster i transistorens ströresistanser. Inverterad klass F har lägre ström. Mindre förluster betyder högre verkningsgrad.

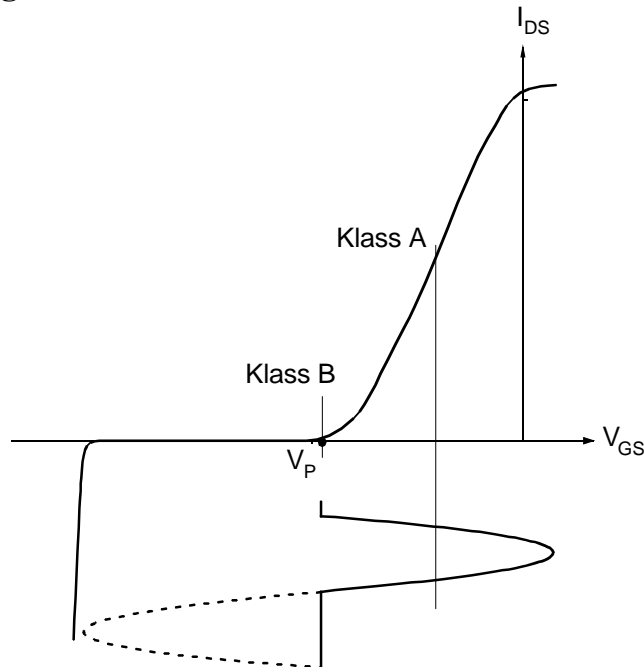
Spänningskurvan för inverterad klass F är mer toppig. Spänningen över transistorerna kan bli 3 - 4 gånger större än spänningen i arbetspunkten. Klass F som har fyrkantig spänning, uppnår endast dubbla spänningen. Inverterad klass F behöver alltså en transistor med högre genombrottspänning än klass F. Då batterispänningen är låg, ger inverterad F högre uteffekt än klass F.

Inverterad klass F har i praktiken uppnått över 80 % verkningsgrad för drain, och 67 % PAE.

Klass F har högre verkningsgrad än Inverterad klass F, om arbetspunkten ligger på mindre än 10 % I_{DSS} . Om transistorerna är förspänd till högre ström, klass AB eller klass A, får inverterad klass F högre verkningsgrad än klass F.

Förspänning till lägre ström än 10 % I_{DSS} för också med sig lägre förstärkning. Klass F får högsta PAE vid gainkompression mindre än 1 dB. Inverterad klass F får högsta PAE vid gainkompression större än 2 dB.

Formad insignal



En modifierad klass A förstärkare matas med en insignal i form av en halv sinuskurva. Med kortslutning för jämna övertoner och avbrott för 3:e övertonen får förstärkaren samma kurvform på drain som klass F. Det blir alltså samma verkningsgrad på drain som för klass F, med en förstärkning som för klass A. Eftersom det bara behövs hälften så stor insignal, blir förstärkningen 3 till 6 dB högre än klass F eller klass B. Kretsen får alltså några procent högre PAE än klass F.

Den negativa spänningen på gate är bara hälften så stor som för klass F. Spänningen på drain kan därför ökas i motsvarande grad, utan risk för genombrott. Det ger högre uteffekt.

Den önskade vågformen för spänningen alstras genom att addera dubbla frekvensen med lämplig amplitud och fas. Drivsteget före effektförstärkaren är av klass B med resistiv belastning. 2:a övertonen filtreras fram, fasvrids och adderas till grundtonen med lämplig amplitud (-7 dBc), så att spänningen blir formad som halva sinuskurvor.

Förstärkaren har beskrivits under benämningen hHCA (half sinusoidally Harmonic Control Amplifier).

Inverterad klass F kan på motsvarande sätt drivas från en förstärkare som ger en 3:e överton i lämplig fas och amplitud (-9 dBc). Spänningen från drivsteget är då fyrkantformad, och slutsteget är av klass A. Förstärkaren har kallats rHCA (rectangularly Harmonic Control Amplifier).

Drivningen kan alternativt ske från ett förstärkarsteg som klipper positiva och negativa topparna. Spänningen blir då trapetsiodformad. Förspänningen till slutsteget ställs in så att transistorn bara leder då spänningen är noll. Det eliminerar förlusterna i transistorn så att verkningsgraden (PAE) kan bli hög.

Sammanställning

Grundklasser

Klass	Arbetspunkt % av I_{\max}	Ledande del av perioden	
A	50	360	Hel sinus
AB	0 - 50	180 - 360	
B	0	180	Halv sinus
C	0	0 - 180	Toppen av sinus

Övertoner

Klass	Kortsluten	Öppen
E		2:a (alla)
F	2:a (4:e)	3:e (5:e)
Inv. F	3:e (5:e)	2:a (4:e)

Formad insignal

hHCA	Halv sinus + F
rHCA	Fyrkantig + inv. F

Referenser

Klasser

- Thomas B Mader, "The transmission-line high efficiency Class-E amplifier", IEEE Microwave and guided wave letters, vol 5 no 9 sep 1995 pp 290-292
- R Tayrani, "A broadband monolithic S-band class-E power amplifier". IEEE Radio frequency integrated circuits symposium 2002 pp 53-56
- Andrei Grebennikov, "High efficiency class-E monolithic HBT power amplifier for wireless applications", 6th European conference on wireless technology 2003 pp 313-316
- Akira Ohta, "Effekt of input marmonic termination on high efficiency HBT amplifiers for mobile communications", 32nd European microwave conference 2002 pp 749-752
- Andrey Grebennikov, "Circuit design technique for high efficiency class-F amplifiers" IEEE MTT-S 2000 pp 771-774
- T Heima, "A new practical harmonics tune for high efficiency power amplifier", 29th European microwave conference 1999 pp 271-274
- C J Wei, "Analysis and experimental waveform study on inverse class-F mode of microwave power FETs", IEEE MTT-S 2000 pp 525-528
- Paolo Colantonio, "A unified approach to high efficiency microwave power amplifier design", GAAS 1999 pp 272-275
- Bernhard Ingruber, "Rectangularly driven class-A harmonic control amplifier", IEEE transactions on microwave theory and techniques, vol 46 no 11 nov 1998 pp 1667-1672
- Bernhard Ingruber, "High efficiency harmonic control amplifier", ", IEEE transactions on microwave theory and techniques, vol 46 no 6 june 1998 pp 857-862
- Masahiro Maeda, "A high power and high efficiency amplifier with controlled second harmonic source impedance", IEEE MTT-S 1995 pp 579-582
- Pat Malloy, "Consider load tolerance in amplifiers for immunity susceptibility", Microwave & RF 2001 mars pp 111-117

5. Dimensionering av effektförstärkare

Anpassning

Dimensioneringen börjar med anpassningen på utgången.

Förstärkaren anpassas antingen för max förstärkning, högsta uteffekt, högsta verkningsgrad eller bästa linjäritet (IMD). Anpassning till maximal uteffekt ger 1 - 3 dB högre uteffekt än vid konjugatanpassning. Tyvärr blir förstärkningen samtidigt ca 1 dB lägre. Eftersom förstärkaren inte är konjugatanpassad, blir Return Loss (VSWR) för förstärkaren dålig.

Det kan vara lämpligt att välja anpassning för att få bästa verkningsgrad. Verkningsgraden kan nämligen endast förbättras med anpassningen. Uteffekten kan bli större med en transistor som har större periferi. 10 - 20 % större periferi kan vara ett billigt sätt att få 5 % högre verkningsgrad.

När anpassningen ändras varierar både uteffekt och förstärkning. Inställningen blir enklare om transistorn först drivs till mättnad och sen justeras för max uteffekt.

En förstärkare med fler kaskadkopplade transistorer brukar innehålla frekvenskompensering i kretsen mellan transistorerna. I en effektförstärkare behöver drivsteget optimal belastning för att ge tillräcklig effekt till slutsteget. Om kompenseringen för jämn förstärkning över bandet ligger på ingången, kan uteffekten fortfarande vara maximal. För att inte ingången ska bli missanpassad kan frekvenskompenseringen innehålla en resistans.

Storsignal S_{22}

Transistorns ingång S_{11} är ganska okänslig för signalnivån. Utimpedansen S_{22} varierar däremot starkt med effektnivån. En transistor för hög effekt behöver alltså mätas upp med stora signalnivåer. Ett enkelt och snabbt sätt att mäta storsignal S_{22} , är att reducera spänningen över transistorn V_{DS} , och sedan mäta småsignal S_{22} .

Nackdelen med S-parametrarna är att de är definierade för linjära komponenter med konstant belastning. En effektförstärkare drivs in i sitt olinjära område. S-parametrarna kan inte användas till att göra beräkningar på andra impedanser än den uppmätta. Värdet på storsignal S_{22} motsvarar inte den optimala belastningen för högsta effekt.

Fysisk modell

En fysisk modell av transistorn med dimensioner, dopningsprofil och förspänning, kan ge transistorns kurvor för ström/spänning, samt S-parametrarna för små signaler.

En fördel med en fysisk modell är att det går att se inverkan av temperaturens variation och processtoleranser.

Harmonisk Balans

Kretsen delas upp i linjära respektive olinjära delar. Transistorn är den olinjära delen. Den mäts upp till en olinjär ekvivalent krets. Beräkningarna görs sen i tidsdomänen. Det ger vågformen på ström och spänning. Resultatet transformeras till frekvensdomänen med diskret Fourier-transformation, för att sedan kopplas ihop med de linjära delarna.

De linjära delarna är anpassningskretsar, stabiliseringskretsar och förspänningskretsar. De kan enkelt beräknas i frekvensdomänen med befintliga CAD-program. Grundton och övertoner anpassas till sina respektive belastningar för att få hög verkningsgrad.

Program med Harmonisk Balans innehåller tio övertoner. Vid dimensionering av effektförstärkare räcker det med 2:a och 3:e övertonen. Högre övertoner kan i praktiken anses som kortslutna av transistorns strökapacitans på utgången.

Harmonisk Balans ger samma noggrannhet som vid olinjär CAD. Fördelen är att dimensioneringen blir mycket snabbare. Distribuerade komponenter har långa tidskonstanter som gör olinjär CAD (t.ex. Spice) mycket långsam.

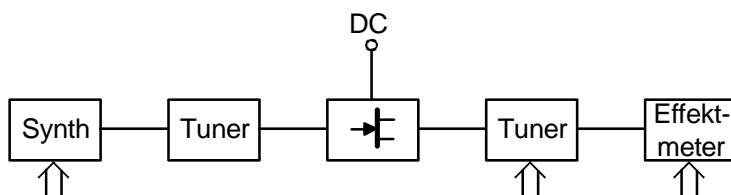
Med Harmonisk Balans kan man beräkna transistorns förstärkning för stora effekter, samt intermodulation, ACPR och verkningsgrad.

mm-våg

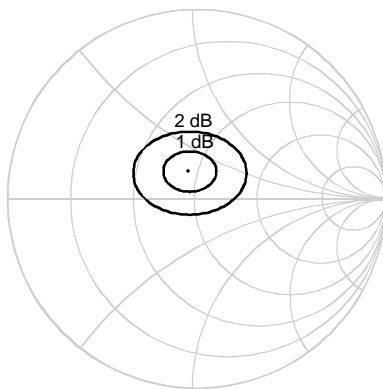
På mm-våg, där det inte finns lämpliga mätinstrument, kan transistorerna inte direkt mätas. Transistorns S-parametrar för små signaler mäts för frekvenser upp till ca 50 GHz. En ekvivalent linjär krets för transistorn anpassas till mätdata. Dessutom mäts transistorns I/V data (transistorkurvorna). Tillsammans ger det en komplett olinjär modell för att kunna beräkna uteffekten.

Load-Pull

En Tuner är en reaktiv krets som avstämmer transistorn till önskad anpassning. Load-Pull innebär att lastens impedans varieras, för att påverka transistorn till önskad prestanda.

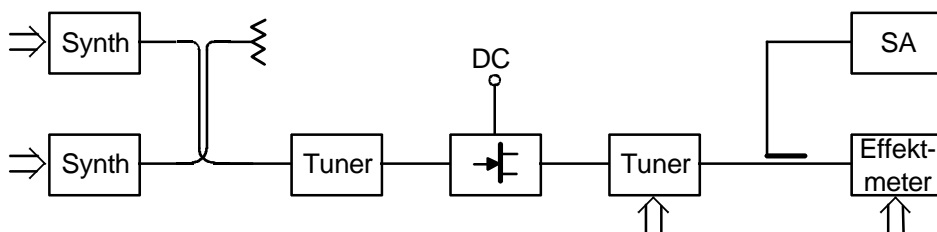


En Tuner på ingången ställer in transistorn till konjugatanpassning. På utgången sitter en programmerbar Tuner som ställer in olika belastningar till transistorn. Tunern är antingen uppmätt i förväg, eller också switchas Tunern till en nätverksanalysator för varje inställning. En effektmeter registrerar uteffekten för de olika inställningarna. Allt ska sedan upprepas på de olika frekvenserna inom bandet. Eftersom det blir mycket mätdata används en dator till att styra instrumenten och lagra all data. Därefter kan datorn beräkna den optimala belastningen för maximal uteffekt.

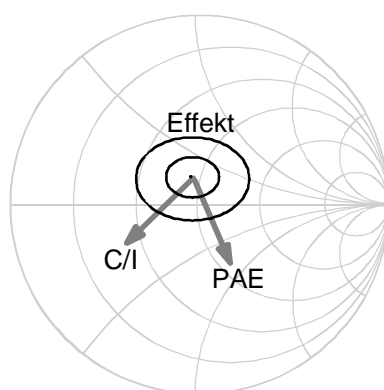


Datorn ritat upp konturerna för de belastningar som motsvarar 1 dB respektive 2 dB lägre uteffekt. Med en multimeter kan också effektförbrukningen från likspänningen mätas upp. Datorn beräknar verkningsgraden PAE, och i diagrammet kan man göra kompromisser mellan uteffekt och verkningsgrad.

Konturerna för konstant uteffekt respektive verkningsgrad är inte cirklar eftersom förstärkaren är olinjär. Tyvärr behövs en uppsättning konturer för varje frekvens. Det blir en stor mängd data att mäta upp.



Med två insignaler som skiljer sig några MHz, kan man på en spektrumanalysator avläsa distorsionen C/I. Konturer för konstant C/I kan sen ritas in i samma Smithdiagram som för effekt och verkningsgrad.



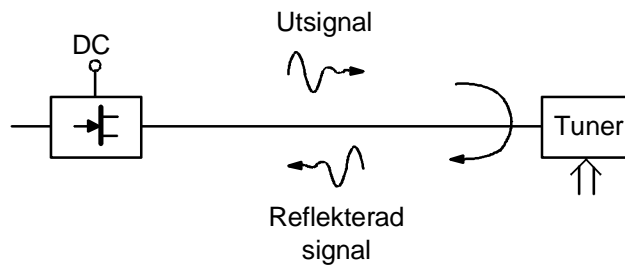
Högsta effekt, PAE respektive C/I sker vid olika anpassningar. Exemplet visar med pilar hur impedansen kan behöva förändras för att förbättra C/I och PAE. Med konturer för alla tre kan man lätt göra en lämplig kompromiss.

Hela uppmätningen kan sen göras för olika frekvenser, ineffekter och förspänningar. Det slutliga valet av lastimpedans kan sedan anpassas till 50Ω med ett linjärt analysprogram.

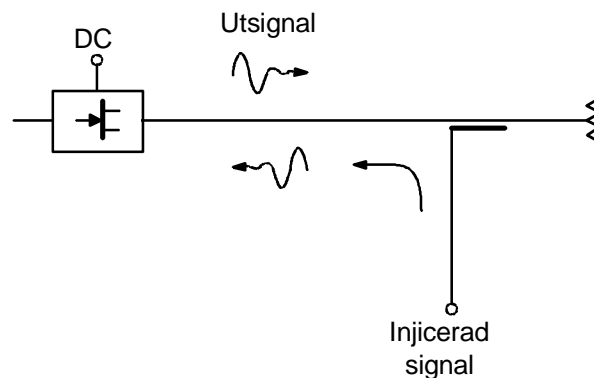
En nackdel med passiv Load-Pull är svårigheten att uppnå stora reflektioner. Tunern har förluster och mellan transistor och Tuner finns förluster som minskar reflektionen. 1 dB förluster minskar reflektionsfaktorn från 1 till 0,79.

En annan nackdel är att det bara är grundtonen som regleras. Övertonen blir okontrollerade. En effektförstärkare behöver rätt belastning på 2:a övertonen (och kanske 3:e) för att få högsta verkningsgrad.

Aktiv Load Pull

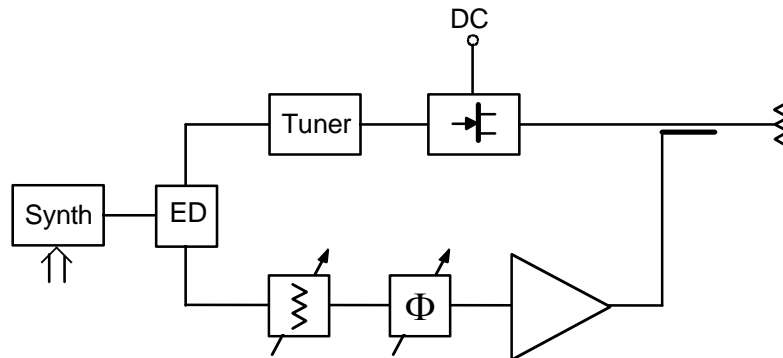


En tuner ställer in en viss impedans. Eftersom den inte är inställd till konjugat-anpassning sker en viss reflektion. Den reflekterade signalen är ett mått på hur mycket belastningen avviker från 50Ω .



Istället för den speciella avslutningen kan man använda en reflektionsfri avslutning. Med en injicerad signal ser det från transistorn ut som en missanpassning. En förutsättning är förstås att den injicerade signalen är koherent (har samma frekvens) som den förstärkta signalen.

Injiceringen sker på grundtonen, 2:a och 3:e övertonen ser 50Ω anpassning.

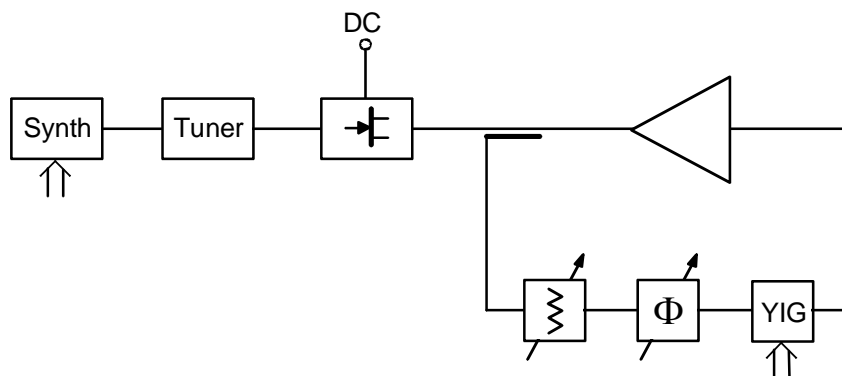


Den injicerade signalen hämtas i en parallell kanal från signalkällan (synthesizer). Genom att variera amplitud och fas på den injicerade signalen, kan vilka belastningar som helst simuleras.

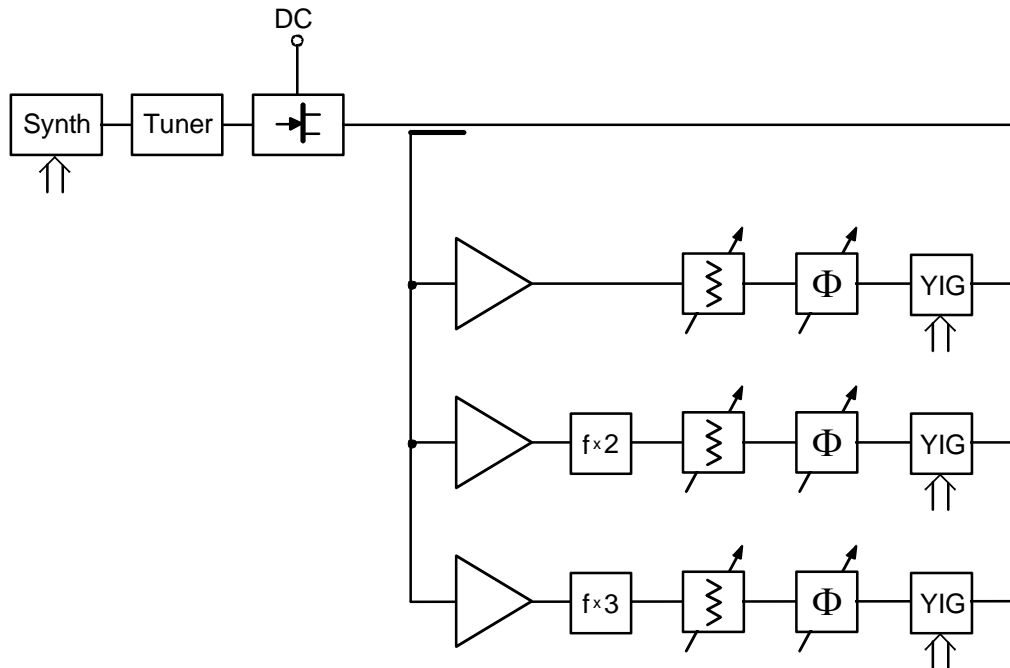
Det är lätt att simulera $\Gamma=1$, även om tillledningarna från transistorn har förluster. Signalnivån kompenseras lätt med den variabla dämpsatsen. Belastningen kan därför varieras över hela Smith-diagrammet.

En nackdel är att det är svårt att hålla belastningen konstant när effekten varieras. De två vägarna kan nämligen limitera på olika signalnivåer.

Kretsen med två parallella signalvägar kallas Takayama metoden.



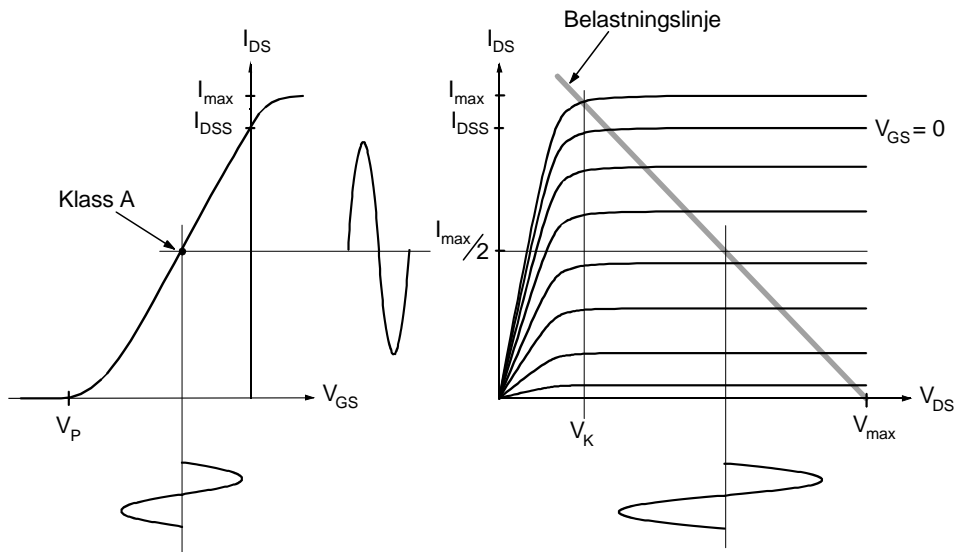
Ett annat alternativ är att alstra injiceringen från utsignalen, i en slinga (Loop) tillbaka till transistorn. Ett YIG-filter kan användas för att eliminera självsvängning i slingan.



Både Takayama och kretsen med slinga kan utvidgas till belastning av övertoner. Beroende på transistor och graden av limitering, blir verkningsgraden förbättrad 5 - 15 % genom att optimera fasen på övertonerna.

Cripps metod

Belastningen för högsta uteffekt beräknas från transistorns kurvdiagram.



Högsta strömmen är I_{DSS} eller I_{max} . Största spänningsvariationen är mellan V_K och V_{max} . För högsta effekt ska belastningsresistansen vara.

$$R_{optP} = \frac{V_{max} - V_K}{I_{max}}$$

Transistorns strökapacitans på utgången C_{DS} och dess tillledningsinduktans L_D mäts med små signalnivåer. Dessa reaktanser får ingå i anpassningskretsen, som nu kan dimensioneras med en linjär simulator.

Cripps metod är enkel och behöver varken data från Load-Pull eller olinjär simulator. Nackdelen är att det blir ca 1 dB lägre uteffekt med den enkla metoden.

Referenser

Design

Franco N Sechi, "Design procedure for high efficiency linear microwave power amplifier", IEEE Transactions on microwave theory and techniques, vol 28 no 11 nov 1980 pp 1157-1162

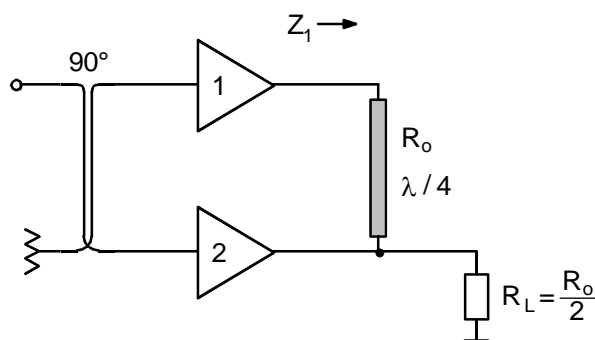
C Duvaud, "Optimization of trade-offs between efficiency and intermodulation in SSPAs based on experimental and theoretical considerations", IEEE MTT-S 1993 pp 285-288

Pierre Berini, "An experimental study of the effects of harmonic loading on microwave MESFET oscillators and amplifiers", IEEE Transactions on microwave theory and techniques, vol 42 no 6 jun 1994 pp 943-950

Steve C Cripps, "Old-fashioned remedies for GaAs FET power amplifier designers", IEEE MTT-S Newsletter summer 1991 pp 13-17

6. Doherty

Doherty är en kretskoppling som förbättrar verkningsgraden då signalerna har varierande amplitud.



Förstärkare 1 arbetar i klass AB eller klass B. Den kallas ibland för bärvågsförstärkare. Förstärkare 2 arbetar i klass C, och kallas ibland för toppamplitudförstärkare.

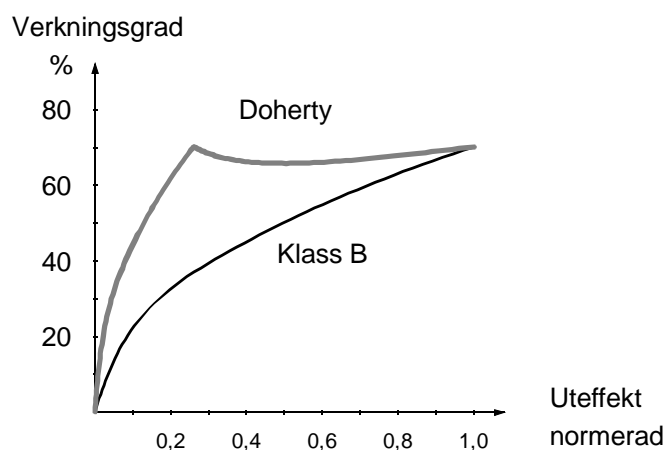
Vid små signaler är förstärkare 2 förspänd på gate så att den är avstängd (utgången höghögmig). Lasten R_L transformeras längs kvartvågsledningen, så att belastningsimpedansen för förstärkare 1 blir $2 \cdot R_o$. Hög impedans på belastningen ger hög verkningsgrad.

När insignalen är stor ger även förstärkare 2 en utsignal. Tillskottet minskar den lastimpedans som förstärkare 1 ser. Lastimpedansen förändras på samma sätt som vid aktiv "Load Pulling". Vid full utsignal från förstärkare 2 har lastimpedansen minskats till R_o . Förstärkare 1 är fortfarande driven till mättnad, men den lägre impedansen gör att den kan leverera mer ström till lasten.

Vid stora signaler kan förstärkare 1 ge dubbelt så hög effekt till lasten R_o . Jämfört med lasten $2 \cdot R_o$. Dessutom ger ju förstärkare 2 ett tillskott till uteffekten. Vid stora signaler ger de två förstärkarna lika mycket effekt till lasten.

När förstärkare 1 börjar komprimera så börjar förstärkare 2 att ge utsignal. Lastimpedansen Z_1 minskar successivt från $2R_o$ till R_o då insignalens amplitud ökar till max. Uteffekten står alltså i proportion till insignalen. Doherty kan därför användas till amplitudvarierande signaler.

En kommunikationssignal (t.ex. W-CDMA) kan ha en AM-variation så stor att toppeffekten är ca 10 dB större än medeleffekten. Det betyder att uteffekten är reducerad 10 dB under största delen av sändningstiden. Dessutom ska uteffekten regleras efter behov. Högsta uteffekten behövs endast för maximala räckvidden. En mobiltelefon i stadsmiljö har vanligtvis en medeleffekt ca 20 dB lägre än den maximala.



När förstärkare 1 uppnått mättnad och förstärkare 2 är avstängd är verkningsgraden maximal. Vid större insignal uppnås maximal verkningsgrad då båda förstärkarna är mättade. Verkningsgraden hålls alltså mycket hög under ett stort dynamikområde. En klass B förstärkare har betydligt sämre verkningsgrad vid reducerad uteffekt (Back-off).

Transformeringen till dubbla impedansen ger en extra topp i verkningsgraden vid 6 dB back-off. Med en transformering till 3 gånger så stor impedans, hamnar den extra toppen vid 7,24 dB back-off.

Eftersom Doherty innebär parallellkoppling av två transistorer kan det behövas isolationsmotstånd för att förbättra stabiliteten.

Referenser

Doherty

Charles F Campbell, "A fully integrated Ku-band Doherty amplifier MMIC", IEEE microwave and guided wave letters vol 9 no 3 mars 1999 pp 114-116

Sami Bousnina, "Analysis and experimental study of an L-band new topology Doherty amplifier", IEEE MTT-S 2001 pp 935-938

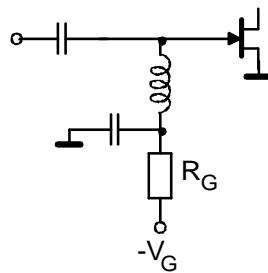
Youngoo Yang, "Experimental investigation on efficiency and linearity of microwave Doherty amplifier", IEEE MTT-S 2001 pp 1367-1370

Masaya Iwamoto, "An extended Doherty amplifier with high efficiency over a wide power range", IEEE Transactions on microwave theory and techniques, vol 49 no 12 dec 2001 pp 2472-2478

7. Förspänning

Gate-ström

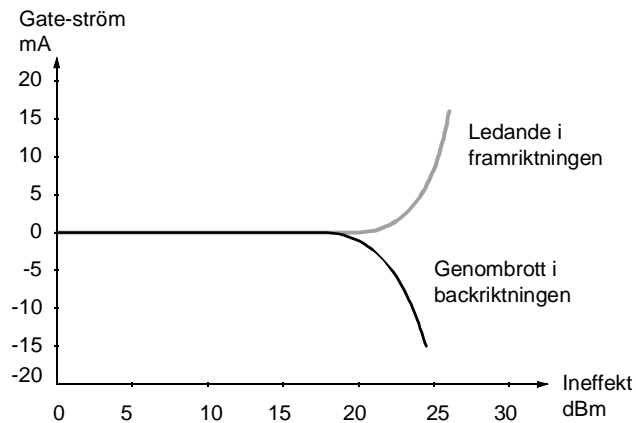
När signalen på gate blir tillräckligt stor, börjar det gå gateström. De positiva topparna på signalen blir så stora att Schottky-övergången på gate blir förspänd i framriktningen. Även de negativa topparna ger ström, då diodövergången blir förspänd till genombrott.



Ett motstånd kan placeras i serie med gate. Den likriktade strömmen ger då ett spenningsfall som förspänner transistoren bort från klippningen. En stor resistans ger en stabil DC-förspänning.

Nackdelen med en stor gate-resistans är att det kan uppstå ett termiskt problem. Utan insignal kan det gå en läckström ut från gate. Med en stor gate-resistans alstras ett stort spenningsfall. Transistorn förspänns mot större ström, som värmer upp transistoren. Uppvärmningen kan ge större läckage så att strömmen ökar allt mer (thermal runaway).

Det behövs alltså ett motstånd för att begränsa gate-strömmen. Men inte för stor resistans om det är risk för termiskt läckage. En kompromiss kan vara några hundra ohm (ca $400 / P_{\text{Sat}}$). Transistortillverkare ger rekommendationer hur stor R_G bör vara.



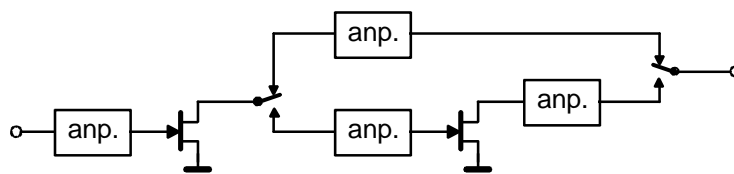
Om gate leder i framriktningen blir strömmen positiv. Om det är genombrott i backriktningen blir strömmen negativ. Ofta är transistor utstyrd så att både positiva och negativa topparna limiteras. Den sammanlagda strömmen kan då först gå lite negativt, för att sedan vända och bli positiv vid hårdare drivning av transistor.

När en transistor drivs in i 1 - 2 dB kompression för att få hög uteffekt och verkningsgrad, blir gate-strömmen normalt ca 1 - 2 mA per mm gate-periferi. Gate-strömmen kan till och med bli upp till 300 mA för stora transistorer.

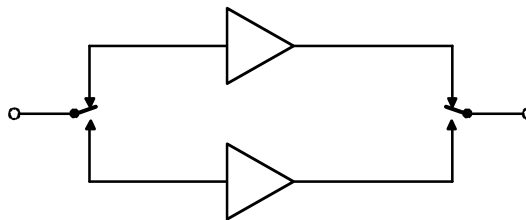
Spänningskällan behöver kunna ge både positiva och negativa strömmar. Eventuellt behövs ett parallellmotstånd över spänningskällan.

Olika uteffekter

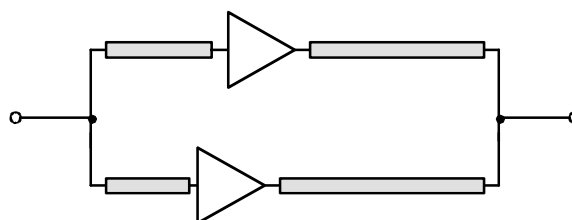
Ibland behövs en förstärkare som kan kopplas om mellan hög och medelhög uteffekt. Om en klass B förstärkare användas på 20 dB lägre uteffekt (back-off), blir verkningsgraden mindre än 5 %. En klass A förstärkare får, vid 20 dB reduktion, en verkningsgrad mindre än 1 %. Om verkningsgraden ska vara hög behövs olika slutsteg.



Ett alternativ är att koppla bort transistorn för högsta effekt, och mata antennen direkt från drivsteget.

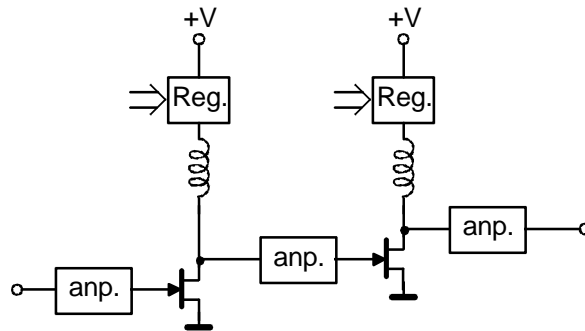


Ett annat sätt är att switcha mellan två olika parallella slutsteg. Det slutsteg som inte används ska då kopplas bort från strömförsörjningen.

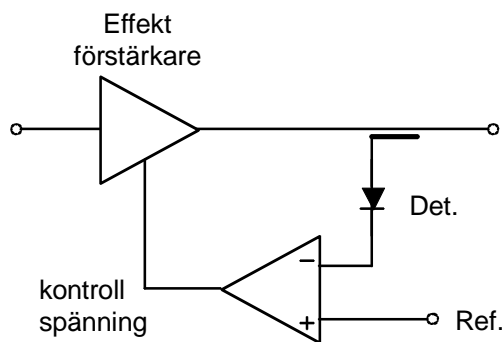


Själva omkopplarna kan elimineras, om den strömlösa (eller strypta) transistorns impedans transformeras längs en ledning till något mycket höghmigt vid knutpunkterna. Åtminstone behöver impedansen vara fyra gånger så hög, för att förlusterna i knutpunkten ska vara mindre än 1 dB.

Varierad uteffekt



Uteffekten är proportionell mot kvadraten av drivspänningen. Det kan utnyttjas till finjustering av uteffekten med hög verkningsgrad, åtminstone inom ett område på 10 - 20 dB.

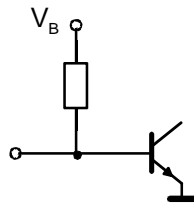


Via en riktkopplare detekteras uteffekten, och jämförs med en referens. Kontrollspänningen styr effektförstärkarens spänning på drain (kollektor). Uteffekten justeras genom att variera referensspänningen.

Eftersom det är en sluten slinga hålls uteffekten konstant även om det är variationer i temperatur eller insignalnivå.

Regleringen kan göras lagom snabb, så att amplituden kan starta och sluta mjukt under ett sändningsintervall (burst).

Förskjutning av arbetspunkt



Med ett motstånd mellan basen och dess spänningskälla kan både verkningsgrad och linjäritet förbättras. Vid stora signaler går det stor basström. Det ger ett spenningsfall över motståndet, och transistorn arbetar i klass B. Vid små signaler blir det mindre basström och mindre spenningsfall. Transistorn arbetar då i klass AB.

Viloström

När en batteridriven apparat inte används ska strömförbrukningen vara så liten som möjligt. Med en viloström mindre än $1 \mu\text{A}$ räcker batteriet länge.

BJT, HBT och E-FET kan lätt försättas i viloläge genom att jorda ingången.

En vanlig MES-FET behöver stänga av drainspänningen för att strömmen ska bli tillräckligt låg. En MOSFET kan användas som drain-switch. Samma MOSFET kan också användas för att variera uteffekten med drainspänningen.

Genombrottspänning

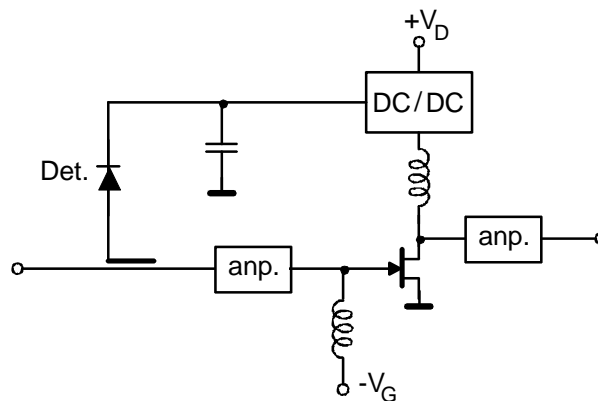
När en förstärkare drivs in i mättnad behövs en transistor som har en genombrottspänning som är 3 gånger så stor som arbetsspänningen.

En linjär förstärkare med 3 - 6 dB lägre uteffekt (Back-off) behöver en transistor med en genombrottspänning som är dubbelt så stor som arbetsspänningen plus gatespänningen.

Vid genombrott går det ström genom transistorn. Eftersom det ger effektförluster i form av värme, blir verkningsgraden sämre. Dessutom försämras transistorns tillförlitlighet vid uppvärmning.

Adaptiv förspänning

Om den önskade signalen har AM-variationer kan strömförsörjningen regleras i motsvarande grad, för att få hög verkningsgrad. När insignalen är liten, så blir det liten RF-effekt ut. Om samtidigt drainspänningen minskas, så blir också förbrukningen av DC-effekt liten. Genom att justera DC-effekten efter behov blir verkningsgraden alltid hög.



Den detekterade insignalen styr en DC/DC omvandlare så att effekttransistorn får en drivspänning proportionell mot aktuell signalnivå. DC/DC omvandlaren är en switchad förstärkare med en pulsbredd, eller ett pulsintervall, som är proportionellt mot styrspanningen. Pulstågets DC-medelvärde filtreras fram med ett lågpasfilter. Med en switch-frekvens så hög som 10 MHz kan spolar och kondensatorer vara små.

Även spänningen på gate kan justeras för att få så hög verkningsgrad som möjligt. En klass B förstärkare har sämre transkonduktans för små signalnivåer. Om spänningen på gate görs mindre negativ så blir transkonduktansen högre. Det ger högre förstärkning och verkningsgrad. Även en klass AB förstärkare får bättre linjäritet om spänningen på gate justeras så att transkonduktansen hålls konstant.

Om en klass B förstärkare har 65 % verkningsgrad vid 1 W uteffekt, blir verkningsgraden ca 57 % vid 10 dB reducering (back-off) då både drain och gate justeras. Om enbart drain justeras blir verkningsgraden ca 49 %. Helt utan justering av strömförsörjningen ger en verkningsgrad på ca 18 %.

En klass A förstärkare kan ha adaptiv förspänning på enbart gate, för att reducera strömförbrukningen då signalnivån är liten. Vid 6 dB reducering (back-off) kan adaptiv förspänning på gate öka verkningsgraden från ca 12 % till ca 28 %. Vid 10 dB reducering blir verkningsgraden ca 3 %. Med en adaptiv justering av spänningen på gate eller drain, ökar verkningsgraden till ca 10 %. Med justering av både gate och drain blir verkningsgraden så stor som 24 %.

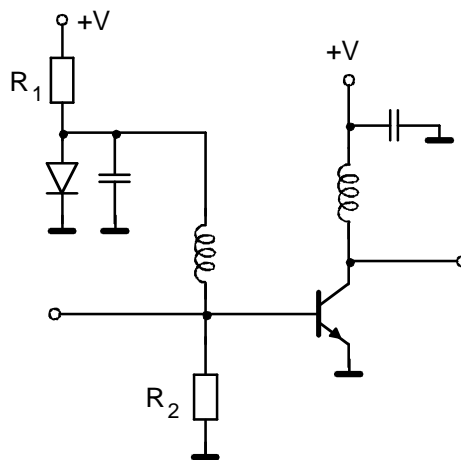
En klass AB förstärkare kan, vid 10 dB reduktion, ha en verkningsgrad på ca 15 % redan utan justering av strömförsörjningen.

Då signalen har stor bandbredd, till exempel W-CDMA, räcker inte switch-frekvensen i DC/DC omvandlaren. Kretsen kan därför inte reglera efter signalens topp effekt. Däremot kan den följa signalens medeleffekt, som varierar ganska långsamt med vågutbredningen. En krets som reglerar transistorens gate kan däremot justera för variationerna mellan medeleffekt och topp effekt.

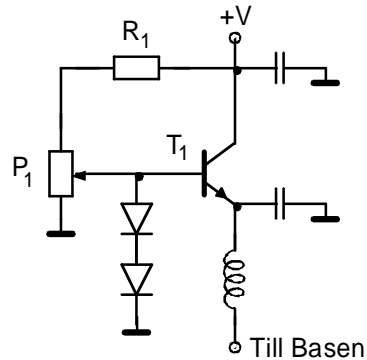
Förstärkaren anpassas för högsta uteffekten. När sedan uteffekten minskas med DC-förspänningen, så får transistorn andra S-parametrar och blir alltså missanpassad. Det gäller då att se till att den missanpassade förstärkaren fortfarande är stabil.

Förspänning av bipolär transistor

Vid låg effekt kan ett enkelt resistiv nät användas för att förspänna Basen. Vid högre RF-effekt behövs en stabil spänningskälla med låg impedans och förmåga att ge ganska stor ström till Basen.

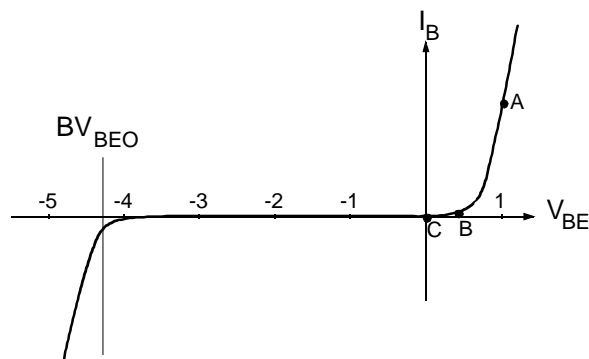


Dioden ger den önskade spänningen med låg drivimpedans. Eftersom det är hög effekt behövs det en ganska stor ström till basen. Dioden arbetar då på ännu högre ström, och behöver alltså vara dimensionerad för hög effekt. Dioden är monterad på transistorn eller dess kylfläns för att få temperaturkompensering. Helst ska diodens spänning som funktion av temperaturen vara densamma som för transistorn. R_2 ger Basen lite lägre spänning än dioden. R_1 ska ge ett mycket stort spenningsfall, och behöver alltså vara dimensionerad för hög effekt.

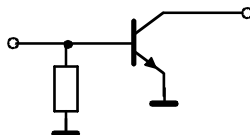


En effektt transistor kan användas istället för effektdiod och effektmotstånd. Genom att dioderna och motstånden arbetar på låg effekt blir verkningsgraden bättre. T1 är en transistor eller ett Darlingtonpar för audiodfrekvenser. Det åtgår två dioder. Den ena dioden för att kompensera spänningsfallet mellan bas och emitter på T1. Den andra för att ge rätt spänning till RF-transistorn. Det är alltså bara en diod som ska monteras på RF-transistorns kylfläns.

Genombrott i bipolär transistor

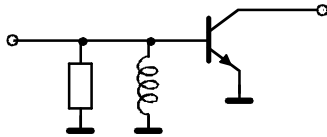


En klass C förstärkare behöver endast jordpotential på basen.



Om ett motstånd används för att ge jordpotential är det risk för genombrott. Insignalen likriktas av bas-emitter dioden, och likströmmen ger ett spänningsfall över motståndet. Transistorn alstrar alltså själv en negativ förspänning. En transistor på mikrovåg har en genombrott-spänning på endast några volt. Även om motståndet bara är på 5 - 10 Ω kan den likriktade spänningen tillsammans med RF-spänningen alstra genombrott.

Ett genombrott alstrar laddningsbärare som får sådan rörelseenergi att de kan ge defekter (traps) i kristallstrukturen. Minoritetsbärarna kan sporadiskt fastna i dessa fällor. Det ger lägre förstärkning. Speciellt vid låga strömmar blir h_{FE} lägre.



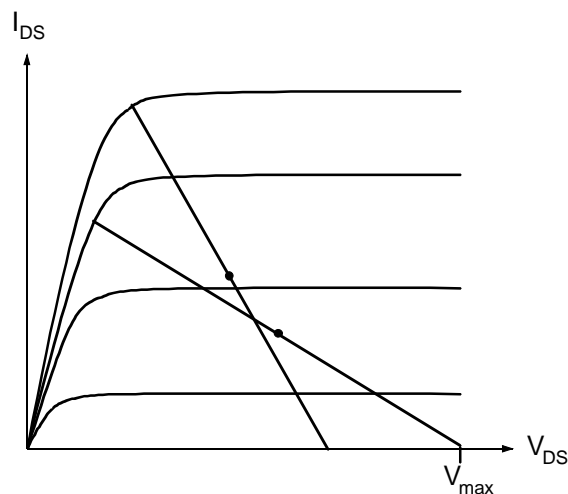
Ett bättre alternativ är att använda en spole för att säkerställa jordpotential. Ett motstånd kan användas för att ge spolen lägre Q-värde.

Låg drivspänning

Genom att minska antalet celler i batterierna, och arbeta på lägre spänning, reduceras volymen och vikten på handapparaterna väsentligt. En förstärkare som arbetar på 3,6 V eller mindre behöver bara en Li-ion cell.

Ett 3 V system innebär att effektförstärkaren ska kunna ge den önskade effekten även då batteriet har så dålig laddning att spänningen bara är 2,7 V. Dessutom ska den kunna klara en spänning på 4,5 till 5,5 V då batteriet håller på att laddas.

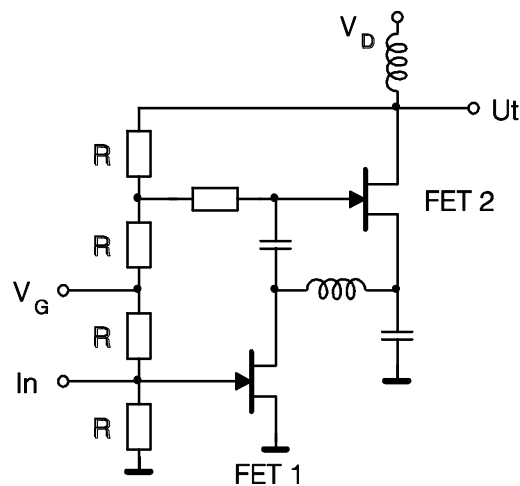
Genom att reducera gate-längden och optimera dopningen i epitaxialsiktet i en MESFET, är det möjligt att få en så låg knäspänning som 1 V och en genombrottspänning på 14 V.



När drivspänningen minskar blir också RF-spänningens amplitud mindre. För att få samma uteffekt behöver istället RF-strömmen öka, genom att resistansen i belastningen minskas. Större ström innebär en transistor med bredare gate. Dessutom innebär större ström att förlusterna, i transistor och övriga komponenter, blir mer påtagliga. Lägre drivspänning ger alltså sämre verkningsgrad.

8. Låg strömförbrukning

En FET-förstärkare för små signaler behöver en drainspänning på 3 - 5 V. Ett batteridrivet system på 12 V kommer alltså att slösa bort 75% av effekten i spänningsregulatorn.



Två kaskadkopplade transistorer kan DC-mässigt vara kaskodkopplade. Man utnyttjar då samma ström genom båda transistorerna. Det åtgår då bara hälften så mycket ström. RF-mässigt är det fortfarande två kaskadkopplade source-jordade förstärkarsteg. På mikrovåg är induktansen L_B höghögig och kondensatorerna är kortslutande. Med en spänningsdelare får man lika stora gate- och drainspänningar på respektive transistor.

Principen med att DC-mässigt koppla transistorerna i serie kan också användas för distribuerade förstärkare. De har vanligtvis ganska låg förstärkning i förhållande till strömförbrukningen. 4 transistorer ger bara 5-10 dB förstärkning. Med DC-koppling i serie blir strömförbrukningen bara en fjärdedel.

Arbetspunkt

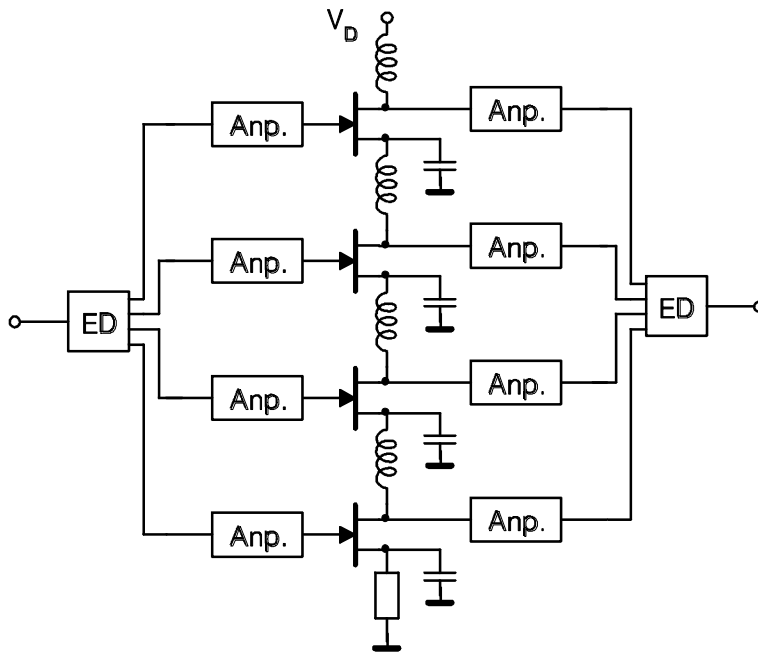
För att få låg effektförbrukning ska både ström och spänning hållas låga. Men den arbetspunkten ger lägre transkonduktans (förstärkning). Med lägre djup på kanalen får man högre förstärkning. Därför får man högre förstärkning med E-FET (Enhancement mode FET).

En nackdel med att förspänna till låg effektförbrukning är att cirkelarna för konstant gain (inritat i Smith-diagrammet) blir mindre. Det ställs då större krav på reaktanselementen i anpassningskretsen. Med större gatebredd blir gain-cirkelarna större, och det blir lättare att få anpassning.

En annan nackdel är att den lägre förstärkningen gör att det krävs fler steg. En FET har sin största förstärkning i förhållande till effektförbrukningen vid låga spänningar, i närheten av transistorens knäspänning (vanligtvis $< 1V$).

Effektförstärkare

En effektförstärkare behöver 6 - 10 V drivspänning. Om man istället kan arbeta på högre spänning kan systemets verkningsgrad bli högre. Det beror dels på högre verkningsgrad för en eventuell DC/DC omvandling från systemets högre arbetsspänning. Det beror också på de resistiva förlusterna i tilledningstrådarna. En fasstyrd antenn har väldigt många aktiva element med betydande förluster i strömförsörjningsnätet. Om man ökar DC-spänningen 4 gånger så minskar de resistiva förlusterna 16 gånger.



De individuella FET-transistorerna är kopplade DC-mässigt i serie och RF-mässigt i parallell. Varje FET har sin source RF-mässigt ansluten till jord via en kondensator. Spolarna som kopplar DC-strömmen mellan transistorerna är RF-mässigt höghögiga så att transistorerna signalmässigt blir isolerade från varandra.

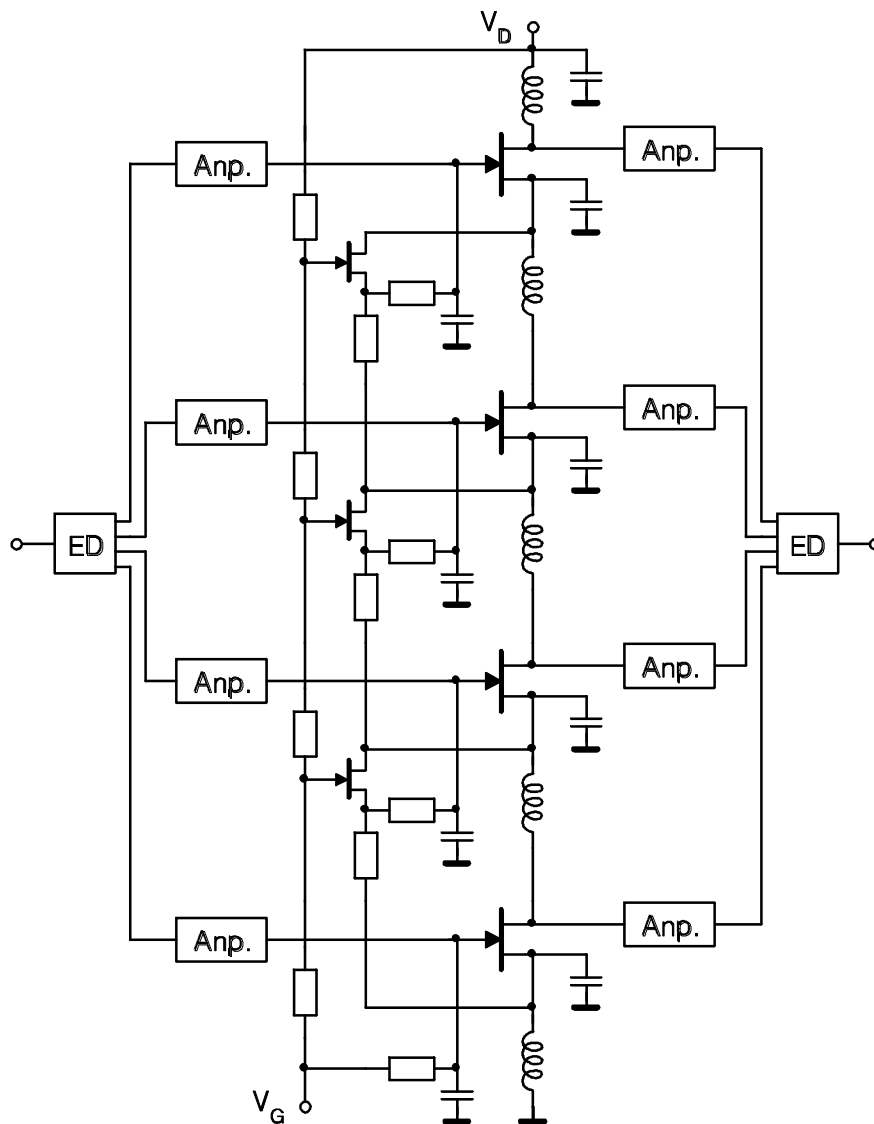
Istället för spolar kan man använda $\lambda/4$ ledningar. Det ger DC-förbindelse utan att RF-kretsarna lastas ner. Ledningarna kommer dessutom att kortsluta dubbla frekvensen. Det förbättrar verkningsgraden.

Exempelvis kan en antenn-array innehålla 1000 element. Varje element kan innehålla ett slutsteg som behöver 10 V och 1 A. DC-effekten matas via en kabel med resistansen 0,001 Ω . Totala strömmen 1000 A ger i kabeln en effektförlust på 1000 W. Om strömförsörjningen istället sker till grupper om 4 transistorer, minskas förlusterna till 62,5 W.

En nackdel med att koppla DC-mässigt i serie är att om transistorernas karakteristik varierar får man obalans i spänningsfördelningen. Det ger försämrade effektkombinering, eller kan rent av ge överslag mellan gate och drain.

Spänningsdelning med source-följare

När en klass-AB förstärkare når mättnadseffekten börjar den dra gateström. Denna ström går genom gatens spänningsdelare och gör att gatespänningen drivs mot klass-A. Det ger alltså försämrad verkningsgrad.



En FET används som sourceföljare för att gatespänningen ska ligga still även om det går gateström. Referensspänningen till sourceföljaren kommer från en spänningsdelare. Tack vare sourceföljaren kan spänningsdelaren vara höghög (dvs. lite effektförbrukning) utan att gateströmmen ändrar gate-spänningen. Mellan sourceföljare och RF-transistor finns ett RC-nät. Resistansen är mycket låg och används för att stabilisera förstärkaren för mycket låga frekvenser. Vid mycket låga frekvenser blir kondensatorn på RF-transistorns source mycket höghög. Transistorn ser istället nästa transistors utgångsresistans. Det ger negativ återkoppling och stabilitet ner till DC.

Referenser

Förspänning & styrning av effekten

Esko Järvinen, "Bias circuits for GaAs HBT power amplifiers", IEEE MTT-S 2001 pp 507-510

James Schellenberg, "High efficiency packaged Ka-band MMIC operating at 24 V", IEEE MTT-S 1998 pp 577-580

J Staudinger, "An overview of efficiency enhancements with application to linear handset power amplifier", IEEE Radio frequency integrated circuits symposium 2002 pp 45-48

Peter M Asbeck, "Synergistic design of DSP and power amplifiers for wireless communications", IEEE Transactions on microwave theory and techniques, vol 49 no 11 nov 2001 pp 2163-2168

M Ranjan, "Microwave power amplifiers with digitally controlled power supply voltage for high efficiency and high linearity", IEEE MTT-S 2000 pp 493-496

Cynthia Y Hang, "A new amplifier power combining scheme with optimum efficiency under variable outputs", IEEE MTT-S 2002 pp 913-916

Hiroshi Okazaki, "Efficient transmission power control scheme for Ku-band high power amplifiers in portable user earth terminals" IEEE Transactions on microwave theory and techniques, vol 49 no 6 june 2001 pp 1167-1172

Fredrick H Raab, "High efficiency linear amplification by dynamic load modulation", IEEE MTT-S 2003 pp 1717-1720

9. Intermodulation

Intermodulation är de icke önskade blandprodukter som alstras av en olinjäritet. Största problemet är den blandning med övertoner som ger IM-produkter inom det egna frekvensbandet och i grannkanalerna. Blandprodukterna inom egna bandet upplevs som IM-distorsion, och i grannkanalerna kallas de störsignaler.

En olinjäritet kan beskrivas med en Taylorserie

$$V_{ut} = a_1 V_i + a_2 V_i^2 + a_3 V_i^3 + a_4 V_i^4 + a_5 V_i^5 + \dots$$

Olinjäritetens kvadratiske del ger dubbla frekvensen

$$\begin{aligned} RF_{in} &= V \cdot \sin \omega_1 t \\ IM_2 &= V^2 \cdot \sin^2 \omega_1 t \\ &\quad \underbrace{\quad\quad\quad}_{\swarrow} \quad \underbrace{\quad\quad\quad}_{\searrow} \\ &\quad \text{lutningen 2} \quad \quad \quad 2 \cdot f \end{aligned}$$

Den kvadratiske funktionen av spänningen ger lutningen 2 i en log/log skala. Om det finns två insignaler ger kvadratiske delen signalernas summa- och skillnadsfrekvens. Olinjäriteten ger frekvensblandning som för en mixer.

Den kubiska termen ger vid två insignaler

$$(A+B)^3 = A^3 + B^3 + 3A^2B + 3AB^2$$

utöver 3:e övertonerna bildas alltså IM_3

$$\begin{aligned} &V^3 \cdot \sin(2\omega_1 - \omega_2)t \\ \text{och} & \\ &V^3 \cdot \sin(2\omega_2 - \omega_1)t \\ &\quad \underbrace{\quad\quad\quad}_{\swarrow} \quad \underbrace{\quad\quad\quad}_{\searrow} \\ &\quad \text{lutningen 3} \quad \quad \quad \text{IMD} \end{aligned}$$

Femte gradens term ger 5:e gradens IM-produkter.

$$IM_5 = V^5 \cdot \sin(3\omega_1 - 2\omega_2)t$$

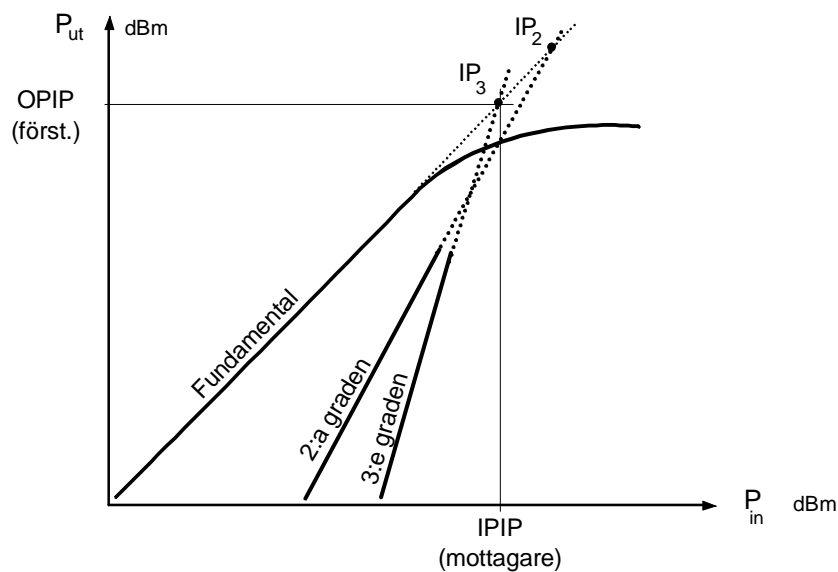
och

$$IM_5 = V^5 \cdot \sin(3\omega_2 - 2\omega_1)t$$

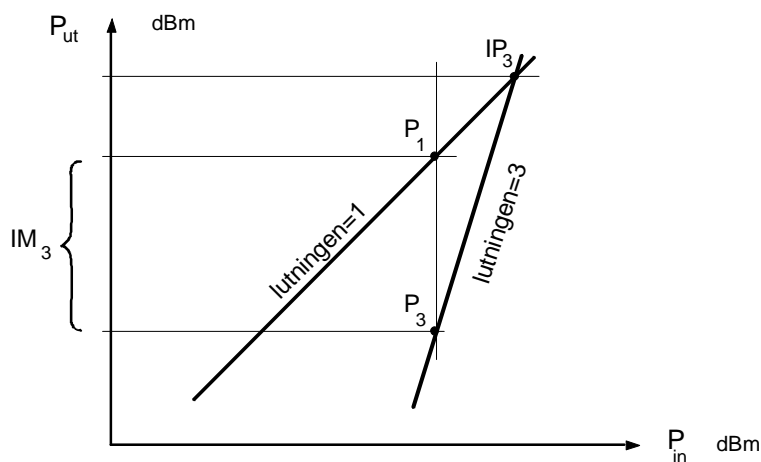
$$\underbrace{\quad}_{\substack{\updownarrow \\ \text{lutningen 5}}} \quad \underbrace{\quad}_{\substack{\updownarrow \\ \text{IMD}}}$$

Intercept Point

I en mottagare är det viktigt hur stora IM-produkterna är i förhållande till nyttosignalen. Men IM-undertryckningen får olika värden beroende på insignalens storlek. Ju större insignalen är desto större blir IM-produkterna. Istället används Intercept Point som ett mått på en förstärkares linjäritet.



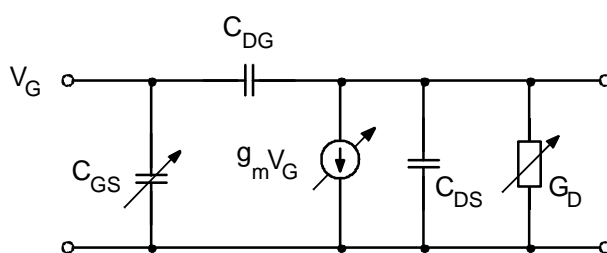
Naturligtvis kan inte förstärkaren uppnå dessa uteffekter. IM-produkterna mäts på låga signalnivåer, därefter förlängs linjerna fram till skärningen med nyttosignalens förlängning. Intercept Point för en förstärkare anges vanligtvis på dess utgång. För en ingångsförstärkare till en mottagare behöver Intercept Point räknas om till ingången, dvs. korrigeras med förstärkningen.



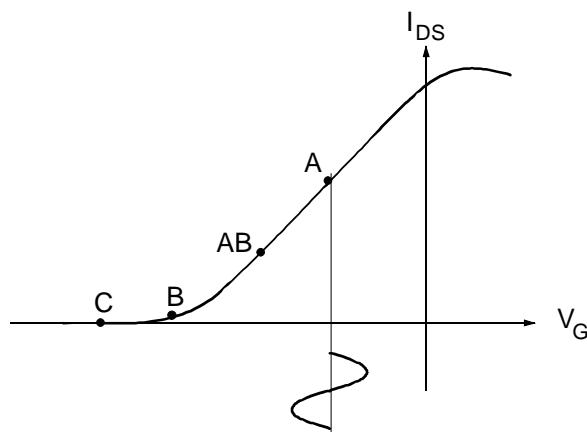
Om skärningspunkten (Intercept Point) är angiven, kan man lätt räkna ut hur stort IM-förhållandet är vid olika nivåer på signalen.

$$IM_3 = 2 \cdot (IP_3 - P_1) \quad \text{dBc}$$

Transistorns olinjäriteter



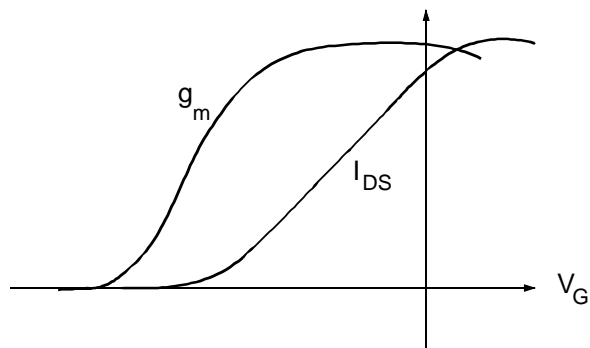
Transistorns kapacitans på ingången är olinjär som en varaktordiod. Konduktansen på utgången är också olinjär. Men den största olinjäriteten är drainströmmen som funktion av spänningen på gate, speciellt vid stora signalstyrkor.



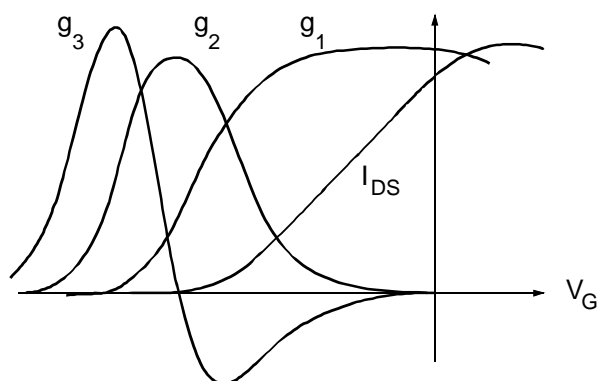
En klass A förstärkare är förspänd till halva maxströmmen. Signalen befinner sig hela tiden i transistorns linjära område. Vid små signalstyrkor kan det istället vara konduktansen på utgången som är den dominerande olinjäriteten. IMD kan då minskas med en högre spänning på drain. Signalvariationerna kommer då bort från de krökta transistorkurvorna för låga spänningar på drain. En klass B förstärkare leder bara ena halvperioden. Strömförbrukningen blir låg, och verkningsgraden blir hög. Men arbetspunkten ligger i den kraftiga olinjäriteten där transistorn stryps. Det ger mycket distorsion. Den största olinjäriteten kommer från transkonduktansen g_m . Kapacitansen på ingången är ganska liten och konstant, då gate har stor förspänning fram till strypning. Klass C förstärkaren leder mindre än en halv period. Endast de positiva topparna ger strömpulser. Det ger högre verkningsgrad och mer distorsion.

När insignalen ändras från arbetspunkten ΔV_G ändras strömmen ΔI_{DS} . Förhållandet mellan ström och spänning är en konduktans. Eftersom jämförelsen sker mellan in- och utgång kallas det istället transkonduktans.

$$g_m = \frac{\Delta I_{DS}}{\Delta V_G}$$



Transkonduktansen är alltså själva lutningen, dvs. strömkurvans derivata. Största olinjäriteten sker i närheten av strypning, där strömmen ändras exponentiellt. Här ändrar sig lutningen (transkonduktansen) mycket kraftigt. Transkonduktansens lutning (derivata) är alltså ett mått på olinjäriteten.



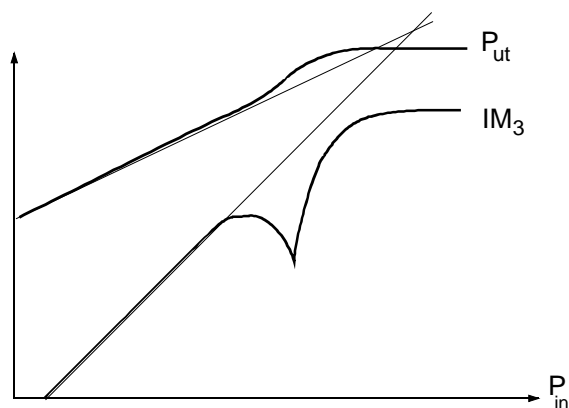
Då transistoren drivs in i kraftig olinjäritet (mättnad), ger andra och tredje derivatan en bra modell över distorsionen.

$$I_{DS} = g_1 V_G + g_2 V_G^2 + g_3 V_G^3$$

Utöver denna distorsion i amplitud (AM/AM), varierar också faser med insignalens amplitud (AM/PM). Amplitudens distorsion är den dominerande, fasdistorsionen ger en irriterande osymmetri i spektret.

En klass A förstärkare har en ganska konstant fas. Det är först vid 1 dB kompressionen som faser börjar ändras. I klass AB börjar faser dra iväg redan för låga signalnivåer, i amplitudens linjära område.

Sweet Spot



När en FET-transistor är förspänd i närheten av strypning (klass B), kan IM3 bli mycket låg för en viss signalnivå i närheten av 1 dB kompressionspunkten. Denna kombination av förspänning och signalnivå kallas för Sweet Spot. Här kan IM3 minska 20 dB. Minskningen beror på att samma frekvens kan alstras av både 3:e och 5:e gradens olinjäritet. Där g_2 är positiv och g_3 negativ tar de olika delarna ut varandra vid en viss signalnivå. Samtidigt sker ett tillskott på den önskade frekvensen från 3:e och 5:e gradens olinjäritet. Det ger en ökning av den önskade signalen, dvs. högre förstärkning.

Vid signalnivåer 20 -25 dB nedanför max signal domineras 3:e gradens termer av g_3 .

Positionen av Sweet Spot är ganska oberoende av övertonernas lastimpedanser. Det är främst förspänningen som ställer in placeringen av Sweet Spot. Storleken på undertryckningen bestäms av en kombination av förspänning och avslutningsimpedans för övertoner och skillnadsfrekvens (basbandet).

I vissa situationer kan alltså en klass B förstärkare användas istället för klass A, utan att försämra IMD.

Minneseffekt

Minneseffekt är att utsignalen inte bara beror på insignalen utan även tidigare insignal. Det beror på att energi har lagrats elektriskt, magnetiskt eller termiskt.

IM_3 mäts ofta upp med två generatorer på olika frekvenser. Skillnaden mellan frekvenserna är vanligen 2 - 20 MHz. Två lika stora insignaler på lite olika frekvenser ger ibland en addition i motfas och ibland en addition i fas. Utsignalens envelop varierar alltså mellan noll och dubbla amplituden. Variationen sker med skillnadsfrekvensen. Om skillnadsfrekvensen är mycket låg, lägre än 1 MHz, blir strömvariationerna så långsamma att transistorns temperatur varierar under perioden. Eftersom transistorns parametrar varierar med temperaturen alstras mer IM-distorsion.

Bipolära transistorer kan ha en basström som regleras av kollektorströmmen. Det blir en tidsfördröjning som ger en minneseffekt.

Ströreaktanserna i spolar och kondensatorer i förspänningskretsarna, samt jordningen av komponenterna, är viktiga parametrar för att minska minneseffekten.

Kondensatorerna för avkoppling behöver vara tillräckligt stora och ha högt Q-värde.

Distorsionens fas och amplitud varierar med modulationsfrekvensen när förstärkaren har minneseffekter. Det är speciellt besvärande, då förstärkaren ska linjäriseras genom att balansera bort distorsionen.

IM_3 alstras både över $(2f_2-f_1)$ och under $(2f_1-f_2)$ insignalerna. Minneseffekten gör att de båda sidbanden blir olika stora. Det blir alltså ett asymmetriskt spektra.

En klass A förstärkare har så hög ström, dvs. uppvärmning, att variationerna blir försumbara. Men en klass AB förstärkare som drivs till mättnad, får en klar försämring av IMD (4 - 6 dB) då skillnadsfrekvensen är låg (10 - 130 kHz). Men denna smalbandiga distorsion påverkar inte grannkanalstörningen nämnvärt, eftersom de flesta moderna system använder bredbandig modulation.

Olika transistortyper

FET

I en FET transistor är det I_{DS} som ger den största olinjäriteten vid stora signalstyrkor. Det är därför mycket viktigt med anpassningskretsen på utgången.

BJT

I en bipolär transistor är det emitterströmmen som ger den största olinjäriteten. Emitterströmmen bestäms av bas-emitter spänningen. Anpassningskretsen på ingången är därför viktigast då det gäller bipolära transistorer.

CMOS

Upp till några GHz är konduktansen på utgången den dominerande olinjäriteten. På Högre frekvenser är det istället transkonduktansen som ger den största olinjäriteten.

Pulsdopad FET

En pulsdopad FET har ett svagt dopat lager ovanpå ett starkt dopat lager. Med Gate på det svagt dopade lagret blir avlänkningskiktet mycket brett. Det för med sig att C_{GS} blir mer konstant. Pulsdopning ger också en mer konstant transkonduktans för olika spänningar på Gate. Vid små signaler blir då G_{DS} den dominerande olinjäriteten. IMD varierar därför med belastningens impedans.

LDMOS

LDMOS har lägre IM-distorsion än de andra teknologierna. Jämfört med en FET har LDMOS ett skarpare knä, där transistorn börjar leda. Det ger ytterligare en Sweet-spot i närheten av kompressionspunkten. En klass AB förstärkare får en linjäritet som en klass A förstärkare, fast med betydligt högre verkningsgrad.

HBT

Bas-Emitter övergångens olinjära kapacitans och resistans balanserar delvis ut varandra. Transkonduktansens olinjäritet kan vara betydande, men den dominerande olinjäriteten i en HBT är kapacitansen mellan bas och kollektor. Om transistorn har fullt avlänkad kollektor (punch-through collector), blir C_{BC} mycket liten och konstant. Det ger en mycket hög IP3.

Transkonduktansen kan linjäriseras med ett motstånd i emittern. Även ett motstånd i basen förbättrar transkonduktansens linjäritet i viss mån. IM3 kan också minskas med hjälp av en generatorimpedans som uppvisar en kortslutning för dubbla frekvensen och skillnadsfrekvensen.

Jämförelse

Både transkonduktansen g_m och konduktansen på utgången G_{DS} beror på dopningsprofilen i det aktiva skiktet i transistorn. En transistor ska naturligtvis ha en så hög IP3 som möjligt. Men när olika transistorer ska jämföras är det bättre att betrakta IP3 i förhållande till 1 dB kompressionen, eller IP3 i förhållande till DC-effekten.

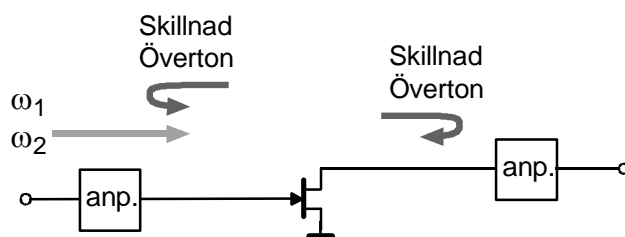
	IP3 dBm	IP3/1dB komp dB	IP3 / P_{DC} dB
Effekt FET	32	14	4
Dubbelt pulsdopad Pseudomorfic FET	33	17	6
Pulsdopad FET	43	24	50
HBT	35	21	44

Användning av överton och blandfrekvens

Insignaler	ω_1	ω_2		
IM ₂	$2\omega_1$	$2\omega_2$	$\omega_1 + \omega_2$	$\omega_1 - \omega_2$
IM ₃	$3\omega_1$	$3\omega_2$	$2\omega_1 - \omega_2$	$2\omega_2 - \omega_1$

IM₂ är inte direkt skadlig eftersom frekvenserna ligger så långt ifrån aktuellt band.

IM₃ innehåller delar som ligger inom bandet eller i grannkanalerna. Dessa IM₃ behöver undertryckas.



De IM₂ som kommer ut från transistorn kan ledas tillbaks in igen. De kommer då att blandas med insignalerna och alstra intermodulation på samma frekvenser som IM₃

$$\text{Skillnad} + \text{den ena} = \omega_1 - \omega_2 + \omega_1 = 2\omega_1 - \omega_2$$

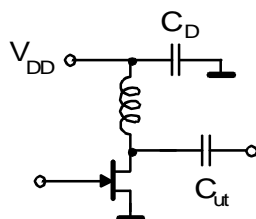
$$\text{Övertton} - \text{den andra} = 2\omega_1 - \omega_2$$

Skillnadsfrekvensen eller dubbla frekvensen kan alltså ge ett tillskott på den störande frekvensen. Genom att justera amplitud och fas kan störningarna på IM₃ balanseras bort. I praktiken kan det röra sig om 3 - 8 dB bättre IMD.

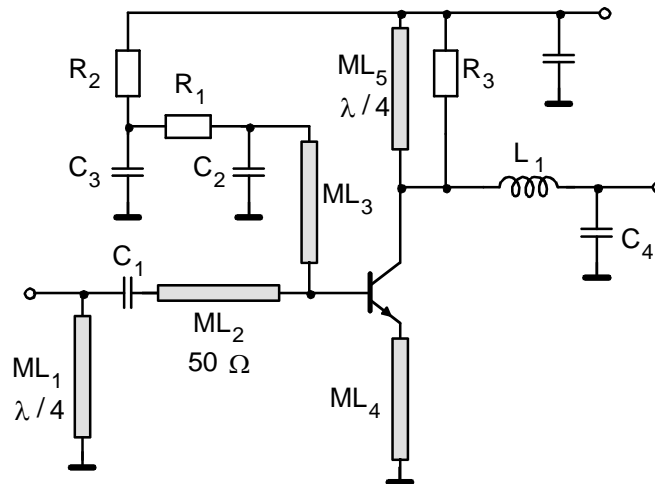
Anpassningskretsarna runt transistorn behöver alltså kunna ge rätt impedans på skillnadsfrekvensen och dubbla frekvensen, för att ge lägsta IMD. Även inom bandet kan anpassningen behöva korrigeras. Lägre generatorimpedans respektive högre lastimpedans kan ge lägre IM.

Kompensering med dubbla frekvensen fungerar bra vid två insignaler. Vid många insignaler blir inte alla blandprodukter undertryckta. Skillnadsfrekvensen ger bättre undertryckning vid flera insignaler. Kretskopplingen blir också enklare om det är fråga om låga frekvenser. Det kan däremot vara svårt att få undertryckning över tillräcklig bandbredd, för att klara alla signalavstånd.

Exempel på kretskoppling



Avkopplingskondensatorn C_D ska kortsluta mikrovågsfrekvenserna. Skillnadsfrekvensen ser däremot en betydande reaktans. IM_3 kan variera 10 dB beroende på kapacitansen C_D . Drainspänningen blir helt enkelt modulerad med skillnadsfrekvensen om C_D är för liten.



ML_1 och ML_5 är en kvarts våglängd långa och påverkar alltså inte det önskade bandet. Däremot blir dubbla frekvensen kortsluten ($\lambda/2$).

Transistorn strömförsörjs genom ML_5 . R_3 är ett stabiliseringsmotstånd.

ML_3 som är kortsluten med C_2 ger anpassning för önskat band. ML_2 och ML_3 ger lämplig impedans för dubbla frekvensen.

Skillnadsfrekvensen, 1 - 50 MHz, avslutas i förspänningskretsarna med R_1 . C_3 är kortslutande för skillnadsfrekvensen.

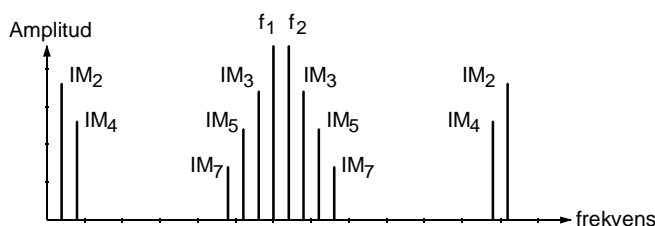
R_2 ger den önskade strömmen till basen. R_2 och C_3 filtrerar strömförsörjningen.

ML_4 tillsammans med ströinduktanserna i emittern ger den önskade återkopplingen för att få samtidig anpassning av både impedans och brus.

L_1 och C_4 ger anpassning till utgången.

10. Passiv intermodulation — PIM

När två eller fler sändare använder gemensam antenn och matarkabel, uppstår IM-produkter från de olinjäriteter som finns i de passiva komponenterna. Med 25 W uteffekt på respektive signal, kan blandprodukterna som hamnar i mottagarkanalerna vara så stora, att mottagningen störs eller blockeras. Om effekterna är riktigt stora, kan det till och med alstras störande PIM i de omgivande metallstrukturerna i närheten av antennen.



f_1 och f_2 är insignaler

$$IM_2 = f_1 \pm f_2$$

$$IM_3 = 2 \cdot f_1 \pm f_2$$

$$2 \cdot f_2 \pm f_1$$

$$IM_4 = 2 \cdot f_1 \pm 2 \cdot f_2$$

$$2 \cdot f_2 \pm 2 \cdot f_1$$

$$IM_5 = 3 \cdot f_1 \pm 2 \cdot f_2$$

$$3 \cdot f_2 \pm 2 \cdot f_1$$

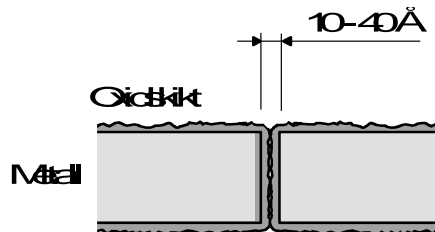
IM_3 kan störa grannkanalerna, som kanske ska ta emot en svag signal från en sändare långt borta. IM_2 kan bli problem i system som kombinerar två frekvensband som skiljer sig cirka en oktav, till exempel 900 MHz och 1800 MHz.

Det finns två typer av PIM. Dels olinjäriteter i själva materialet och dels olinjäriteter i kontakten mellan två metaller.

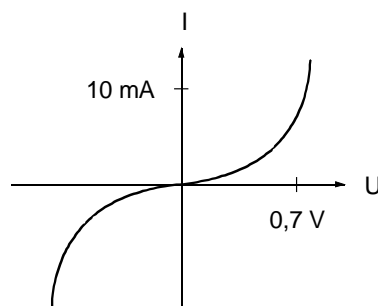
Till olinjära material räknas ferromagnetiska materialen såsom stål, rostfritt stål och nickel. Hysteresen är ett olinjärt fenomen. Komponenter som innehåller ferriter, till exempel cirkulatorer och fasskiftare, har alltså höga PIM-nivåer. Kolfiber är inte bra eftersom de har en olinjär resistans. Hermetisk tätning med Kovar bör också undvikas. Koppar och aluminium oxiderar lätt, och är därför inte heller så bra när det gäller PIM. Dessa material behöver pläteras med ett material som har låga PIM-nivåer.

Som bra material räknas silver, guld, vitmetall, berylliumkoppar och mässing. En förutsättning är förstås att de inte innehåller spår av ferromagnetiska material.

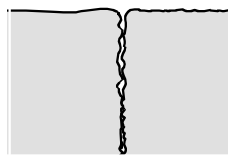
Oxidering



När två oxiderade metaller sammanfogas, blir det ett isolerande skikt mellan metallerna. Eftersom resistansen är hög kan det inte gå någon ström. Men skiktet är så tunt att det blir ett elektriskt genombrott med tunneleffekten. Elektronerna kan tunnla mellan metallerna redan för ganska små signalstyrkor.



Tunneleffekten är ett olinjärt fenomen, därför kommer flera samtidiga signaler att ge blandprodukter.



Även utan oxid blir det alltid små springor mellan två metaller, eftersom de inte kan vara perfekt plana med perfekt ytfinitet. De små potentialskillnaderna, som uppstår när strömmen tar sig fram, räcker för att tunnla över gapen. Högre kontakttryck deformerar oregelbundheterna i ytan så att det blir mer metallkontakt och mindre springor.



Skinneffekten gör att strömmen endast tränger in några μm i metallerna. Det är främst de ytor där strömmarna går som behöver god kontakt. Små sprickor och porer i metallen ger högre PIM-nivå. Även smuts och små partiklar på ytorna försämrar PIM. Innan komponenterna skruvas ihop behöver de alltså tvättas noggrant.

Om det är viktigt med låg PIM-nivå, bör den mekaniska strukturen vara så stor som möjligt, så att strömmarna fördelas över en stor yta. Det ger låg strömtäthet och stor kontaktyta. Alla skruvförbindelser behöver vara ordentligt åtdragna, så att metallerna ger god kontakt. Trimskruvar och rörliga delar ska inte vara placerade där de stora strömmarna går. Metallytorna ska vara så blanka som möjligt. I en gropig yta går strömmen växelvis genom metallen och växelvis genom oxiden och smutsen på ytan.

Koaxialkabel

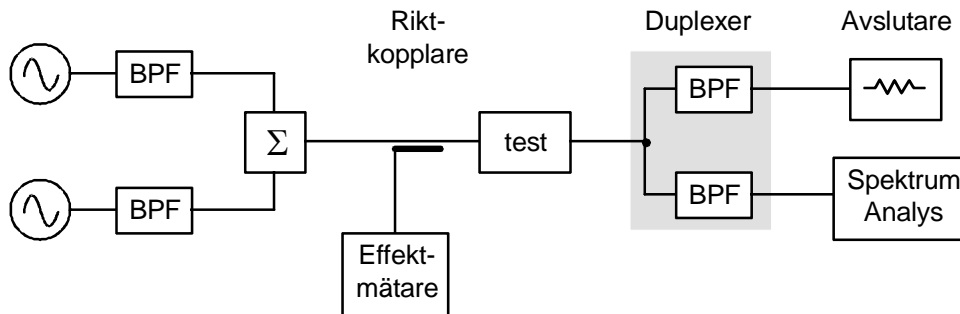
Strömmen går längs med koaxialkabeln, men en skärmfläta har trådar som lutar 45° . Det ger många metallövergångar mellan trådarna, det vill säga hög PIM-nivå. Helst ska man använda koaxialkabel med både innerledare och ytterledare i homogen koppar (semi rigid). Om man måste använda flexibel koaxialkabel ska skärmen helst bestå av tenn- eller silverpläterad koppartråd. Dessutom ska den vara fast monterad, speciellt om innerledaren består av tvinnade trådar. PIM-nivån blir högre när kabeln böjs fram och tillbaks, så att trådarna rör sig i förhållande till varandra. När kabeln är i rörelse kan PIM bli 15 dB sämre. En riktigt dålig kabel försämras 50 dB vid rörelse.

Koaxialkontakten har också en betydande PIM-nivå, speciellt om det är en klämd kontakt. Det blir betydligt bättre om mittstiftet löds fast. Skärmen kan lödas eller skruvas fast. Det kan skilja 15 - 50 dB beroende på monteringen då kabeln böjs fram och tillbaks.

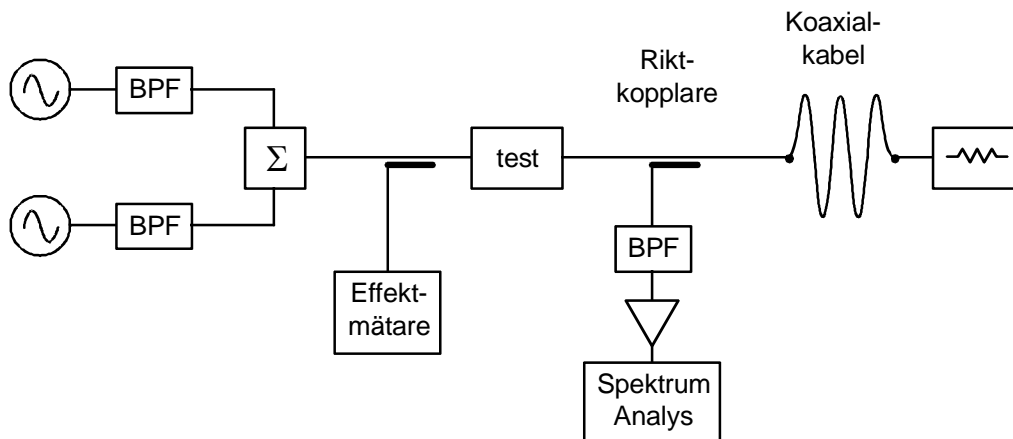
Antenn

En antenn ska helst inte bestå av olika metallbitar som monteras ihop. En reflektor av ett trådnät ger många kontaktpunkter, dvs hög PIM. Trådarna kan behöva lödas ihop för att få låg PIM. Ett ännu bättre alternativ är en stansad plåt. Roterskarvar och rörliga tillredningar som ibland används till styrda antenner, ger en klar försämring av PIM.

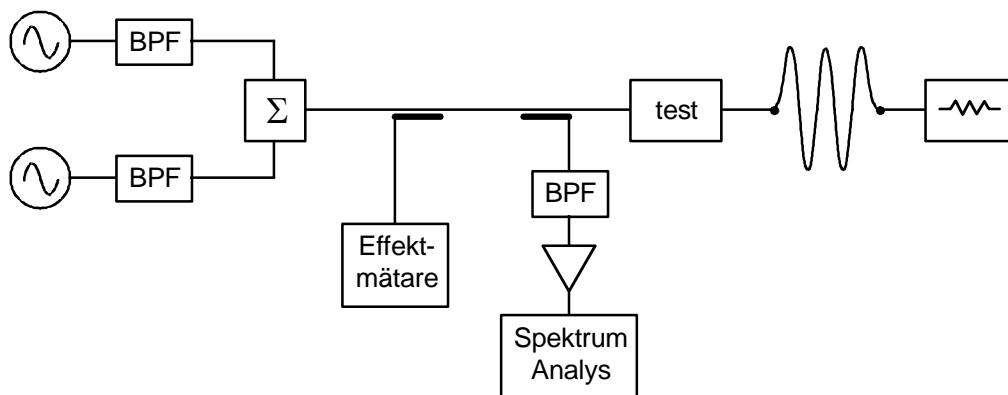
Mätning av PIM



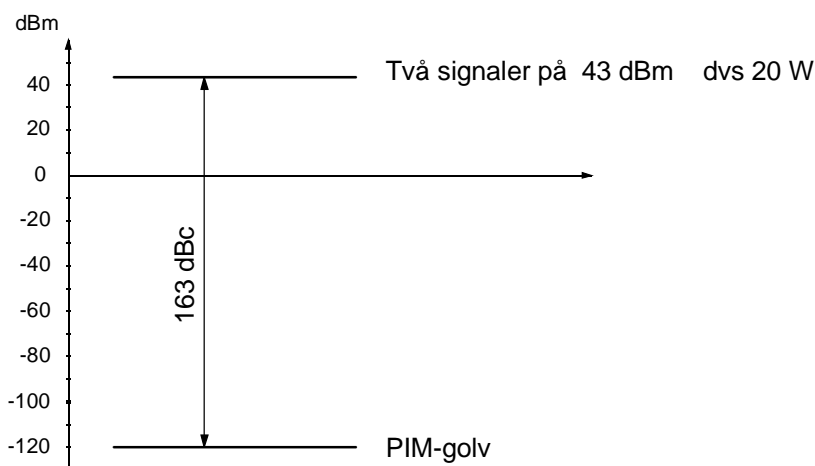
Eftersom intermodulationen i passiva komponenter är liten behövs två signalkällor på hög effekt. Signalkällorna behöver vara ordentligt isolerade från varandra för att inte själva alstra IM. Efter testobjektet ska de svaga blandprodukterna mätas med en spektrumanalysator. De starka uteffekterna får inte ta sig fram till spektrumanalysatorn, för då alstras kraftig intermodulation där. Testobjektets PIM och de höga uteffekterna separeras med en duplexer. Den PIM som alstras i lasten kan inte ta sig tillbaks genom duplexern till spektrumanalysatorn.



Mätsignal och uteffekter kan alternativt separeras med en riktkopplare. Men då behövs en last som inte själv alstrar PIM. En resistiv last har ganska hög PIM. En koaxialkabel kan ha mycket låg PIM. Med tillräckligt lång kabel blir de höga effekterna dämpade 20 dB. Den lägre signalnivån alstrar inte så mycket PIM i lasten, och det som alstras dämpas 20 dB på tillbakavägen. Till mobilbanden har använts 100 m RG214.



Om det är reflekterad PIM som ska mätas, används en riktkopplare för att separera reflektionen. Även om separationen gjordes med ett duplexfilter behövs fortfarande en precisionslast med koaxialkabel.



Vid mätning på mobilsystem används två insignaler som har 20 W per styck framme vid mätobjektet. PIM-nivån i dBc anges i förhållande till de två signaleffekterna 20 W. Effektmeteren visar den sammanlagda uteffekten, det vill säga 40 W vid mätobjektet. Spektrumanalysatorns känslighetsgräns (PIM-golvet) begränsas av PIM-alstringen i mätuppkopplingen. PIM-golvet kan ligga vid ca 160 - 175 dBc.

Signalen $2 \cdot f_1 - f_2$ minskar 1 dB då f_2 minskar 1 dB. Men om f_1 minskar 1 dB blir det 2 dB minskning, eftersom det är dubbla frekvensen som blandas. Om båda signalerna minskar 1 dB så blir det tillsammans 3 dB. Det betyder att om de två signalerna istället är 25 W (dvs 44 dBm) blir PIM 3 dB sämre. När PIM anges ska det alltså klart framgå vid vilken nivå det gäller.

Typ N adapter med nickellegering kan ge ca 120 dBc PIM. En typ 7/16 adapter som är försilvrad har en PIM-nivå lägre än 160 dBc. Cirkulatorer och fasskiftare har så dåligt PIM som 30 - 60 dBc. 1 m halvstyv (semi rigid) koaxialkabel kan ligga på 165 dBc. Är det istället en ½ tums kabel (jumper cable) med 7/16 kontakter kan den vara bättre än 175 dBc.

Referenser

Intermodulation

Min-Gun Kim, "An FET level linearization method using a predistortion branch FET", IEEE Microwave and guided wave letters, vol 9 no 6 june 1999 pp 233-235

John R Gajadharsing, "Low distortion RF LDMOS power transistor for wireless communications base station applications", IEEE MTT-S 2003 pp 1563-1566

Christian Fager, "Intermodulation distortion behaviour on LDMOS transistor amplifiers", IEEE MTT-S 2002 pp 131-134

Michael Jon Bailey, "Intermodulation distortion in pseudomorphic HEMTs and an extension of the classical theory", IEEE Transactions on microwave theory and techniques, vol 48 no 1 jan 2000 pp 104-110

Youngoo Yang, "Behavioral modeling of high power amplifiers based on measured two-tone transfer characteristics", Microwave Journal 2000 dec pp 90-104

H Yamada, "The effect of source impedance on linearity in InGaP/GaAs power HBTs", IEEE MTT-S 1996 pp 555-558

Nuno Borges de Carvalho, "A comprehensive explanation of distortion sideband asymmetries", IEEE Transactions on microwave theory and techniques, vol 50 no 9 sep 2002 pp 2090-2100

Colin S Aitchison, "Improvement of third order intermodulation product of RF and microwave amplifiers by injection", IEEE Transactions on microwave theory and techniques, vol 49 no 6 june 2001 pp 1148-1153

John F Sevic, "A novel envelope termination load-pull method for ACPR optimization of RF/microwave power amplifiers", IEEE MTT-S 1998 pp 723-726

Jose Carlos Pedro, "A novel set-up for co-channel distortion ratio evaluation", IEEE MTT-S 2000

P L Lui, "Passive intermodulation interference in communication systems", Electronic & Communication Engineering Journal, 1990 june pp 109-117

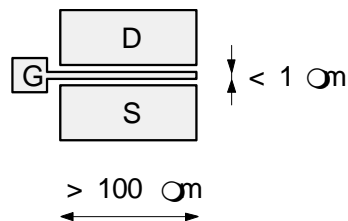
A D Woode, "Problems of passive intermodulation & multipactor in space borne microwave payloads", M&RF conference 1995 pp 162-166

J R Sanford, "Passive intermodulation considerations in antenna design", Antenn 1994 (Sweden) pp 157-160

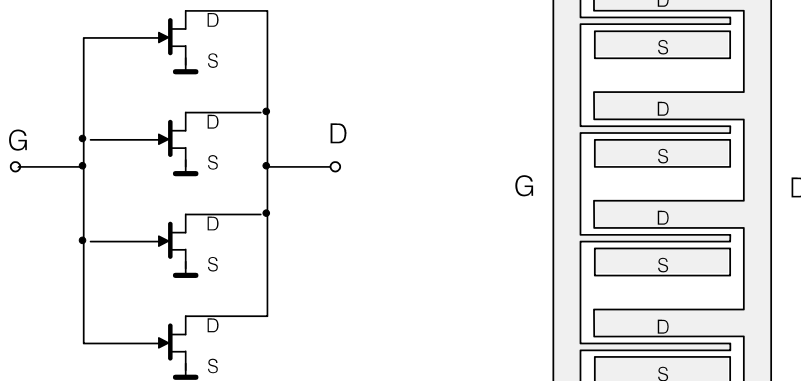
Eric Weibel, "Tests evaluate the influence of junctions on PIM", Microwaves & RF 1998 aug pp70-80

11. Effektkombinering i transistorer

En FET har på mikrovåg mycket små avstånd (gatelängd). Spänningen kan inte höjas utan risk för genombrott. För att få högre effekt måste istället strömmen ökas. En effekt-FET behöver vara mycket bred för att tåla stora strömmar.

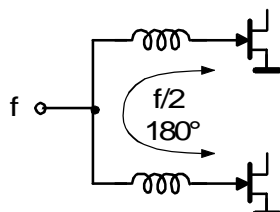


Gateledningen är då mycket lång och smal. Signalen kommer att dämpas och fasvridas då den går längs ledningen. Det minskar förstärkningen. I praktiken består därför effekttransistorn av ett antal små deltransistorer (celler) som är parallellkopplade. Gateledningen är utformad som ett antal fingrar med gemensam anslutningspunkt.



Gate och drain får enkla anslutningar. Tyvärr får man parasitreaktanser vid anslutningarna till source. För att få max uteffekt ska signalen delas lika mellan transistorcellerna. Ingen deltransistor får mättas före de andra. Dessutom måste delsignalerna ha samma fasgång. Den transversella bredden på systemet bör vara $< 1/16$ våglängd för att addera delsignalerna i fas.

Instabilitet i effektförstärkare



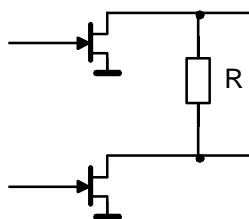
För att uppnå högre uteffekt parallellkopplas två eller flera transistorer. En effektt transistor består också av ett antal deltransistorer som är parallell-kopplade. Den backförspända gate-kapacitansen fungerar som en varaktordiod.

Om avståndet mellan två transistorer är 180° för halva infrekvensen, kan det uppstå parametrisk oscillering. Insignalen fungerar då som pumposcillator, och varaktordioderna uppvisar negativ resistans för halva frekvensen. Oscilleringen byggs upp från bruset, precis som för en oscillator.

Självsvängningen beror alltså på insignalen. Vid små signaler kan den vara stabil, men börjar självsvänga för stora signaler. Både grundton och 3:e överton kan vara starka, dvs både $f/2$ och $3f/2$.

Eftersom det är ett typiskt storsignalfenomen, kan det inte förutsägas med en traditionell linjär stabilitetsanalys.

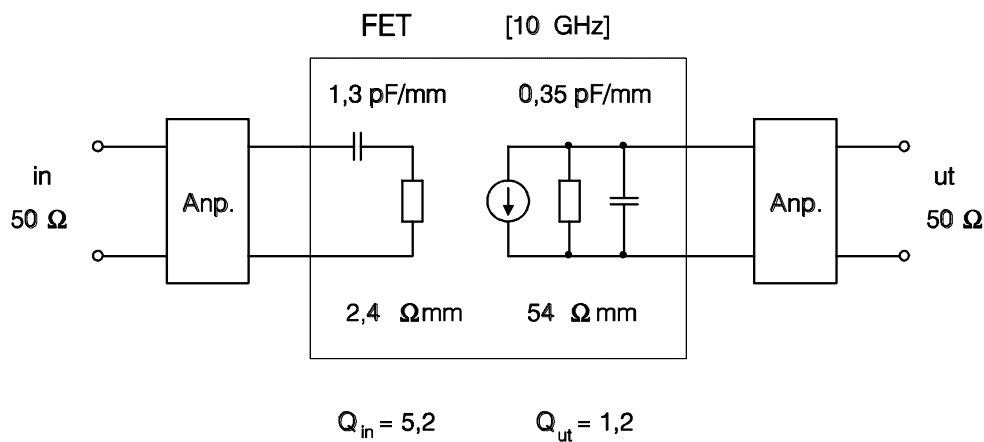
Självsvängningen kallas push-push mode eller udda mode, på grund av att de två transistorerna matas i motfas.



Ett motstånd mellan transistorernas drain kommer att dämpa ut svängningen som ligger i motfas, utan att påverka signalen på den önskade frekvensen.

Transistorns impedans

Genom att koppla ihop ett stort antal transistorceller kan man uppnå gatebredder på 5 - 10 mm. Det ger en uteffekt på över 1 W. 0,3 W/mm vid 1 dB kompressionen och 0,5 W/mm vid max uteffekt.



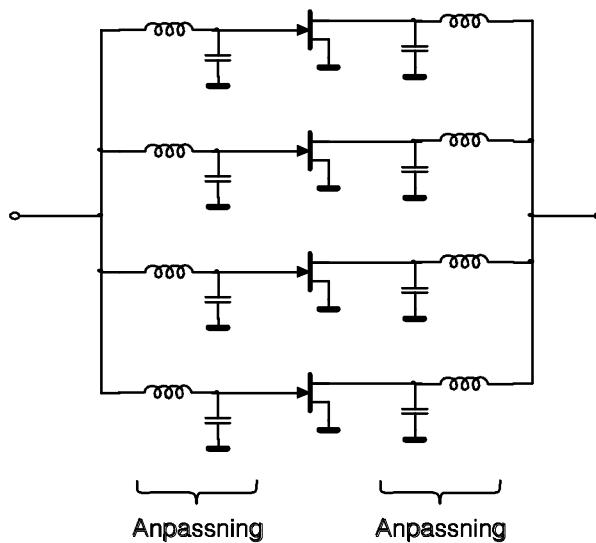
Det största problemet med effektransistorer är impedansen. Många parallellkopplade deltransistorer ger en mycket liten resistans och stor kapacitans. Denna låga impedans (1 - 2 Ω) ska sen transformeras till 50 Ω. Det är en mycket stor impedanstransformation som är svår att åstadkomma, speciellt över stora bandbredder.

Dessutom är Q-värdet på ingången ganska stort för att vara en förstärkare. Det begränsar bandbredden ytterligare.

Eftersom en FET för hög effekt har mycket låg impedans, blir även ledningarna för anpassning lågohmiga. Det är därför lämpligt att använda tunna substrat. Den mycket låga impedansnivån ger tyvärr stora förluster i kretselementen.

Anpassning med cell-cluster

Istället för att bunta ihop alla celler till en jättetransistor kan man dela upp cellerna i ett antal grupper (cluster).



Man får då ett antal transistorer som var och en har måttlig impedansnivå. De kan då lätt anpassas till en hög impedans över en stor bandbredd.

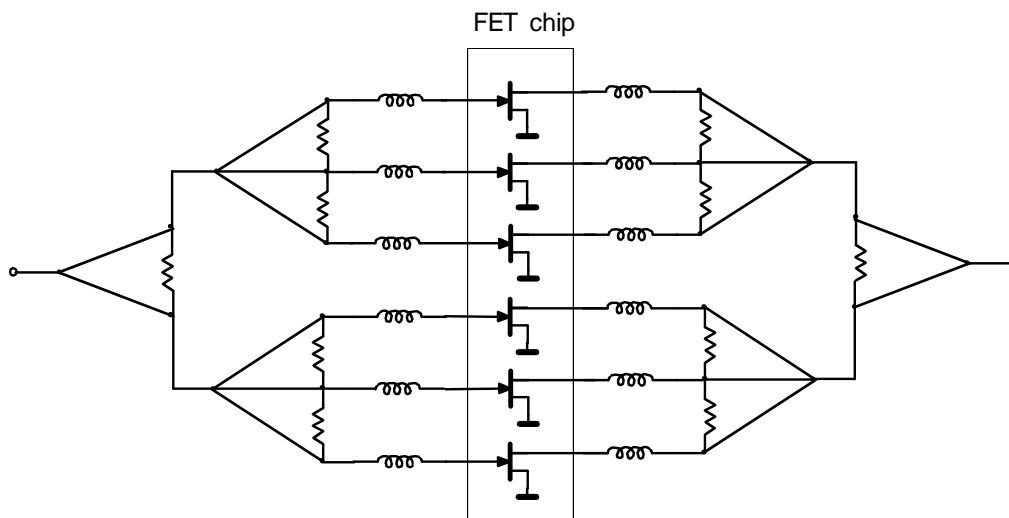
Även sammankopplingspunkterna på in- och utgång får en transformerbar impedans, eftersom det är höga impedanser som parallellkopplas.

Anpassningskretsarna blir ganska stora. De olika transistorerna kan därför ligga en bit åtskilda, utan att det blir några fasproblem. Det för också med sig att kylningen blir bättre, de värmer inte varandra lika mycket.

En nackdel med parallellkoppling är att det kan uppstå självsvängning mellan två transistorer, dvs push-pull moder.

Isolerande effektdelare

För att förhindra självsvängning i push-pull, använder man en effektdelare med isolerade utgångar.



Figuren visar en variant av Wilkinson i två steg, med sammanlagt 6 utgångar. Motstånden dämpar eventuella push-pull moder. Transistorerna blir isolerade från varandra. Effektdelarna transformerar samtidigt impedansen till 50Ω .

6 celler ger tillsammans 3,6 mm gatebredd. Med en sådan transistor kan man få en effekt på flera watt.

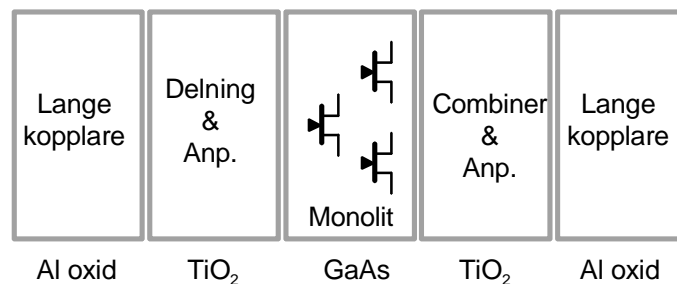
Hela transistorn, med effektdelare och anpassningskretsar, kan med fördel tillverkas monolitiskt. Med intern matchning kan man välja ett mycket tunt GaAs substrat. $50 \mu\text{m}$ tjockt substrat kan ge ledningsimpedanser så låga som $5 - 6 \Omega$. Tyvärr ökar förlusterna ju tunnare substrat man använder. Med monolitisk tillverkning blir kretsen mycket kompakt. Kopplingen mellan ledningarna måste då tas med i beräkningarna.

Förstärkarmodul

Nackdelen med monolitisk tillverkning är att kretsarna för anpassning och kombinerung får ganska stora förluster. En effektförstärkare behöver ha mycket låga förluster för att inte tappa uteffekt och förstärkning.

En förstärkare byggd med separata transistorer på kretskort får andra problem. Kapslar och bondtrådar ingår då i anpassningen. Det ger mindre bandbredd och kritisk montering av bondtrådarna.

En lämplig kompromiss, då det gäller effektförstärkare, är en monolitisk krets med delvis anpassning som kombineras med yttre kretsar. Eftersom monoliten kan innehålla fler transistorer ger det dessutom högre förstärkning. Vanligtvis innehåller monoliten ett eller två förstärkarsteg. Utgångssteget kan bestå av flera parallella förstärkare för att få hög effekt. Monoliten innehåller anpassningskretsarna mellan stegen och delvis anpassning på ingången. Övrig anpassning och effekttuppdelning sker på ett yttre kretskort. Den kombinationen gör att monteringen av bondtrådarna till monoliten inte blir så kritiska. Monoliten innehåller också så mycket som möjligt av kretsarna till förspänning och stabilisering. Kombinerung och anpassning på utgången sker på ett yttre kretskort, eftersom det är så viktigt med låga förluster. Ibland kan det finnas en kompensering av transistorns kapacitans på monolitens utgång.



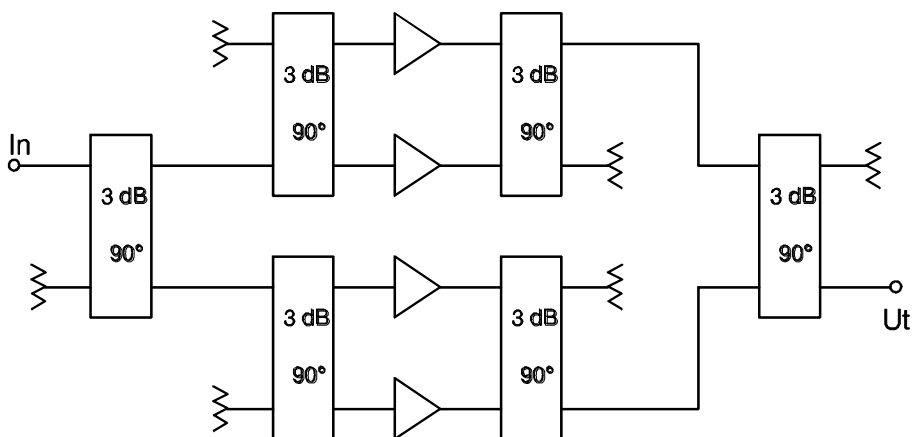
Med Lange-kopplare blir modulen anpassad på in- och utgång. De tillverkas ofta på aluminiumsubstrat. De andra kretsarna kan tillverkas på substrat av titandioxid, med högt dielektricitetsstal, för att få små dimensioner. Förstärkarmodulen kan alternativt byggas med substrat av kvarts för att få riktigt låga förluster. De olika delarna tillverkas och testas var för sig, för att sedan monteras på en gemensam bärare.

12. Kretsar för effektkombinering

En transistor har begränsad uteffekt. Genom att sammankoppla flera transistorförstärkare kan man uppnå högre effekter.

Kretsarna för kombinerings i en effekttransistor ger förluster. Ju högre effekt desto sämre verkningsgrad. Det kan bli bättre verkningsgrad med ett större antal mindre transistorer, och använda en yttre krets för kombinerings. Det ger dessutom större bandbredd och lägre fasbrus.

Hybrider med binär delning

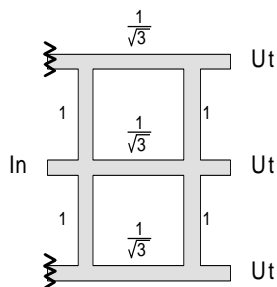


Med två förstärkare mellan två 90° hybrider får man en balanserad förstärkarmodul. Två sådana moduler kan på samma sätt kopplas mellan två hybrider. På så sätt får man alltså 4 gånger så hög uteffekt. Alternativt kan man få en förstärkare bredbandigare genom att använda flera transistorer för lägre effekt.

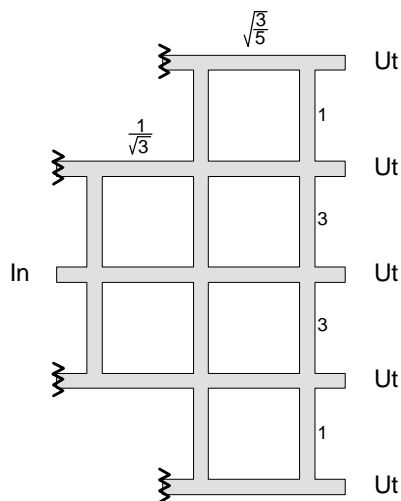
Med 90° hybrider sammansätts reflektionerna i avslutningsmotstånden. Förstärkaren får alltså ett ganska lågt VSWR på in- och utgång. Nackdelen är hybridernas förluster som ger minskad uteffekt och lägre gain.

Uteffekten kan ökas ytterligare med dubblerad krets mellan ytterligare två hybrider. De 8 transistorerna har då ett fördelningsnät med hybrider i tre steg. Det är lätt att öka antalet transistorer med binär delning, men det blir allt högre effektförluster ju fler hybrider som ska passeras.

Hybrider med udda delning

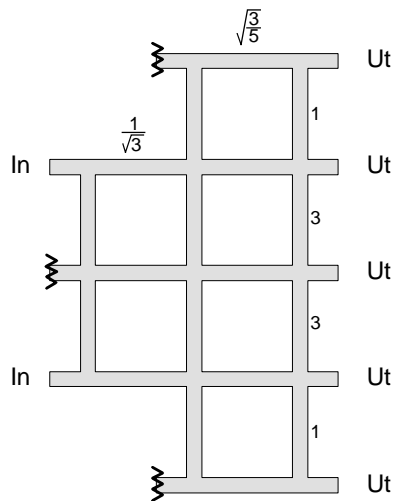


Två sammankopplade hybrider kan ge tre utgångar. Isolationen mellan utgångarna är ca 12 dB. Siffrorna på ledningarna representerar den normerade impedansen.



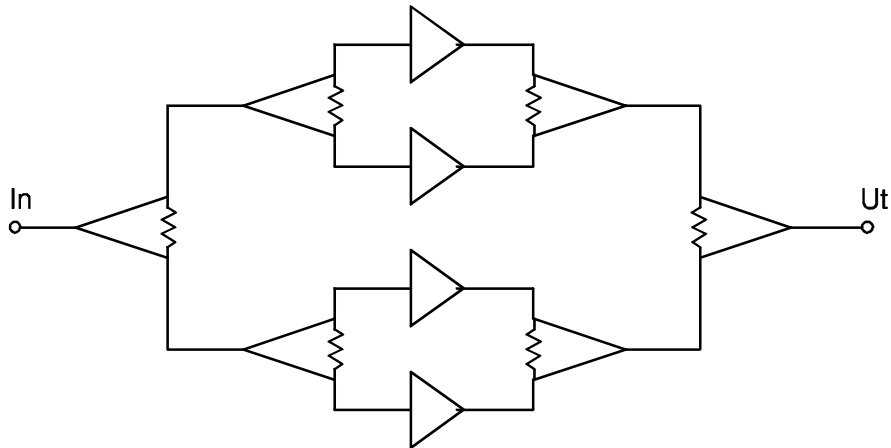
Med ytterligare 4 hybrider blir det 5 utgångar. Isolationen är ca 13 dB. Den normerade impedansen i de vertikala ledningarna är 1 om inget annat anges.

Uppdelningen kan byggas vidare med 6 hybrider i ett tredje steg.
 Det ger 7 utgångar.



Alternativet med 5 utgångar kan även matas från två ingångar.

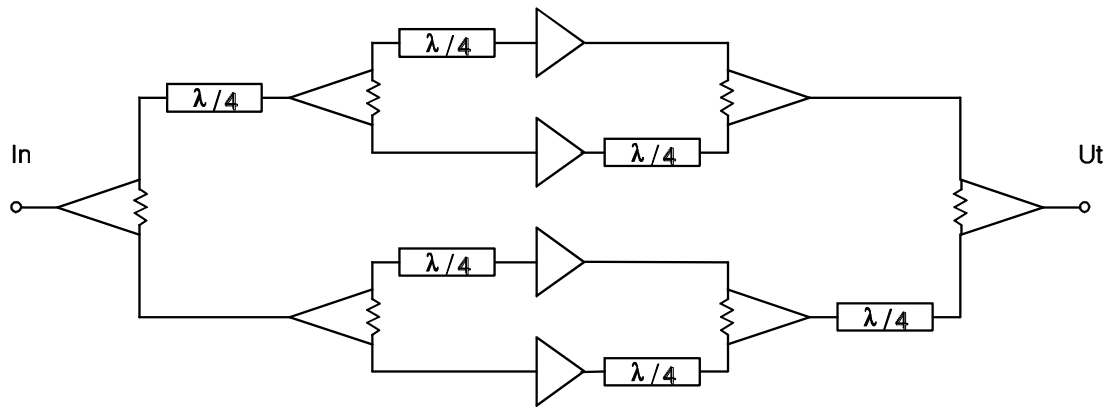
Wilkinson



En Wilkinson delare/combiner har små förluster med mycket litet rippel. Dessutom är amplitudbalansen mellan portarna mycket god. Kvartvågs-längderna kan samtidigt användas till impedanstransformation. Den stora nackdelen med Wilkinson är att utgångarna ligger i fas. Det betyder att två lika reflektioner sammansätts i ingången, dvs ett dåligt VSWR.

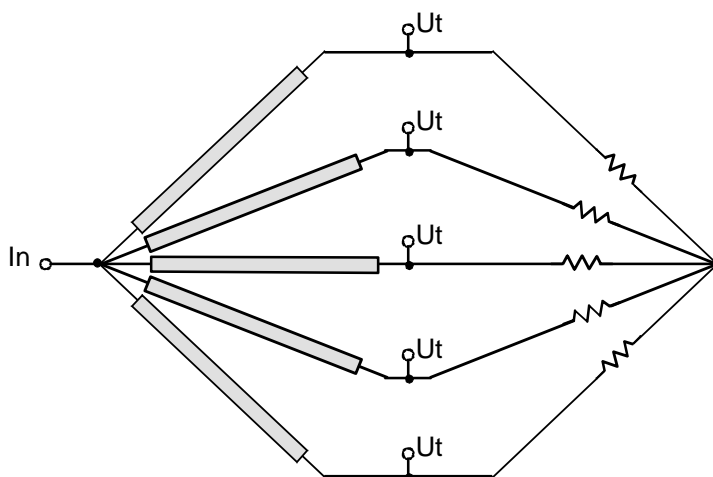
För att förbättra VSWR kan man använda 90° hybrider på in och utgång, men med Wilkinson för övrigt. Det är en kombination som ger både bra VSWR och ganska låga förluster.

Man kan också använda cirkulator (isolator) för att förbättra VSWR. Det har den fördelen att effektförstärkarna blir skyddade från en eventuellt stor reflektion från belastningen, t.ex. kortsluten eller öppen ledning.



Om bandbredden inte är så stor kan man använda en Wilkinson effektdelare samt en ledning på 90° för att uppnå motsvarande låga VSWR som för hybriderna. Den kallas ibland "offset Wilkinson". Man bör emellertid tänka på att den reflekterade effekten från transistorerna ska kunna absorberas i det inre motståndet i Wilkinson. En hybrid med yttre avslutningsmotstånd kan tåla högre effekter.

Wilkinson med flera utgångar

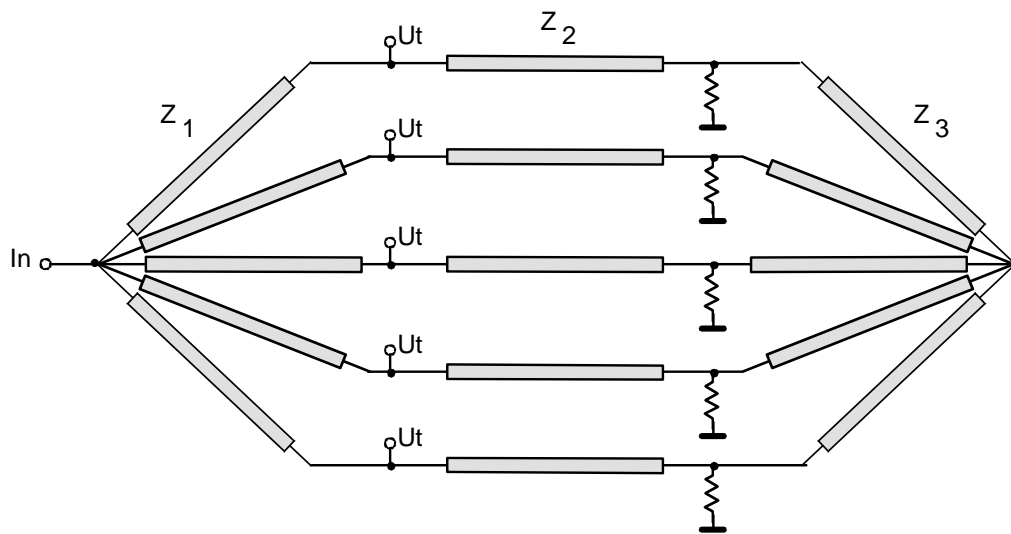


Wilkinson effektdelare kan byggas ut med flera portar. Alla utgångar behöver kopplas ihop resistivt för att få bra isolation mellan portarna. Motstånden är stjärnformigt hopkopplade till en gemensam punkt som inte är jordad. Strökapacitanserna mot jord försämrar prestanda. Strökapacitanserna kan kompenseras med en shuntande induktans från stjärnans mitt till jord. Förlusterna blir betydligt lägre, även bandbredd och isolation förbättras.

När N stycken portar parallellkopplas blir impedansen Z_0/N . Kvartvågsledningarna ska därför ha impedansen $Z_0 \cdot \sqrt{N}$ för att få anpassning.

Effektdelaren fördelar signalen till flera effektförstärkare. En likadan krets används sen för att kombinera effekterna till en gemensam utgång. Med ogynnsamma faslägen kan all effekt gå till motstånden. Tyvärr får motstånden dålig kylning eftersom den inte är anslutna till jord.

Wilkinson för högre effekt

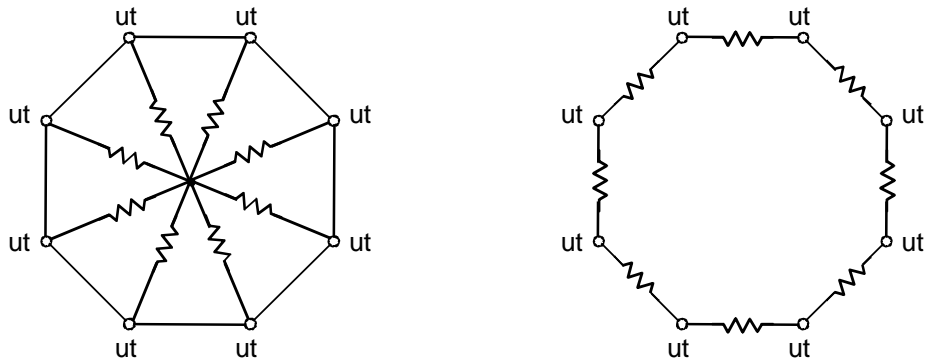


Motstånden kan monteras som shuntande avslutare om de placeras efter en transmissionsledning. Det ger god kylning så att kretsen kan kombinera högre effekter. Både motståndet och ledningen Z_2 har impedansen Z_0 (vanligen 50Ω).

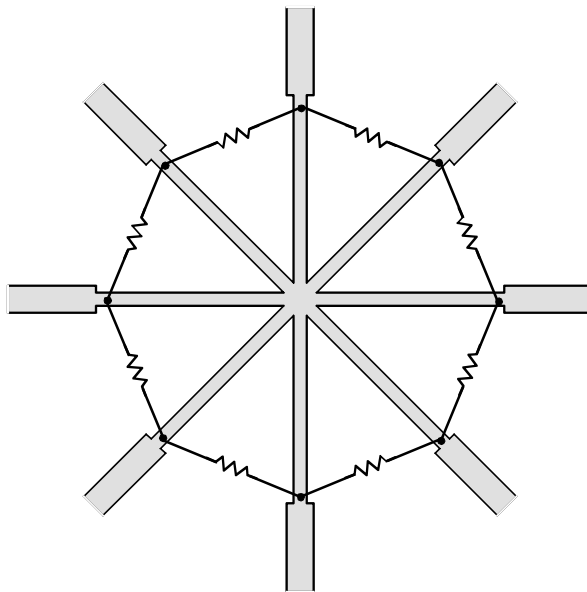
Ledningarna som kopplats ihop i en stjärna har impedansen $Z_3 = Z_0 / \sqrt{N}$.

Nackdelen med längre ledningar är att bandbredden blir mindre. Den blir användbar inom 20 % eller kanske 30 % bandbredd. Kretsen kallas ibland Gisel Combiner.

Isolationsmotstånd i ring



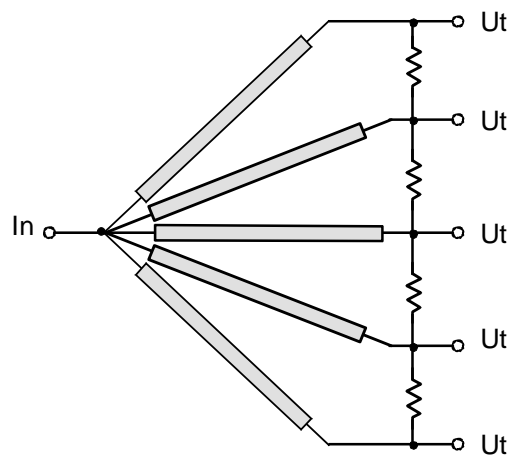
Istället för att koppla motstånden i stjärna, kan de kopplas i ring (stjärna/delta transformering).



Det blir en symmetrisk krets som alltså har mycket god balans i fas och amplitud. Wilkinson matas i centrum, så det krävs tyvärr fortfarande en tredimensionell struktur.

Wilkinson på enkla kretskort

En nackdel med Wilkinson är att den inte går att utföra i plana kretskort. Antingen behöver portarna anslutas vinkelrätt, eller också behövs ett flerlagerkort.

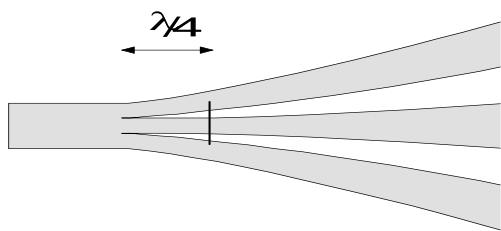


Genom att ta bort ett av motstånden får man en mycket enklare tvådimensionell krets. Nackdelen är att isolationen varierar kraftigt mellan de olika portarna. Ledningarnas impedanser och resistanserna kan istället optimeras för att få samma isolation mellan portarna. Tyvärr blir då anpassningen försämrade.

Kretsen kallas ibland ”Fork combiner” på grund av den gaffelformade uppdelningen.

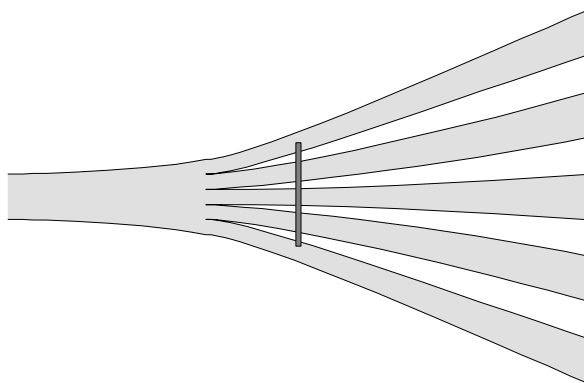
Kontinuerligt avtagande ledning

Wilkinson använder kvartvågsledningar för att transformera impedansen. Ett annat alternativ är att använda kontinuerligt avtagande ledning (Tapered Line). Den innehåller inga diskontinuiteter som Wilkinson, utan har en mjuk övergång mellan de olika impedanserna.



Den avtagande ledningen är betydligt längre än Wilkinson, ofta $1 - 2 \lambda$. Den får en viss undre gränsfrekvens beroende på ledningslängden. Effektdelaren har därför en högpasskaraktär, och kan alltså bli mycket bredbandig.

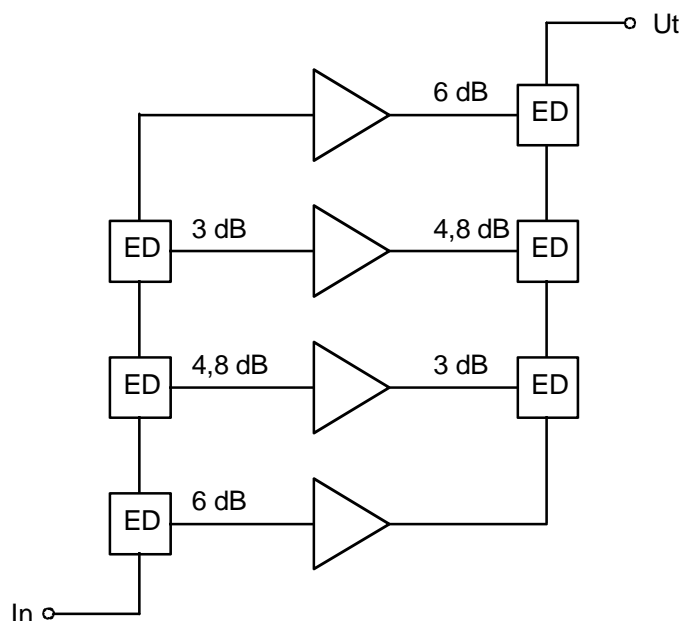
Isolationsmotstånden placeras $\lambda/4$ från delningspunkten. Om det ska vara en stor bandbredd används isolationsmotstånd på flera avstånd från delningspunkten, ungefär som en Wilkinson med transformering i flera steg. Motstånden kan vara i form av chipmotstånd eller en resistiv tunnfilm.



Om signalen ska delas upp till många utgångar behövs en mycket stor transformering av impedansen. Då kan det vara lämpligt att börja med en transformering före delningspunkten. Impedansen står ju i proportion till ledningsbredden. I delningspunkten ska den inkommande ledningen ha samma bredd som den sammanlagda bredden av alla utgångsledningar. Genom att dela upp en bredare ledning behöver det inte bli så smala (höghmiga) ledningar.

Uppdelning i serie

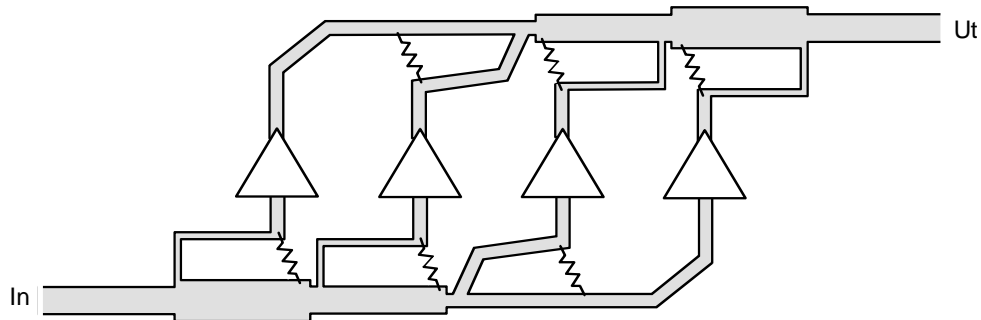
Seriedelning innebär att signalen tappas av på flera ställen längs en signalväg. Kretsen kallas ofta Traveling-Wave delare (respektive kombinerare).



Om det är fyra förstärkare, delas först insignalen upp så att första förstärkaren får $1/4$ av effekten. De övriga tre fjärdedelarna delas sen upp, så att andra förstärkaren får $1/3$ av den effekten. Resterande två fjärdedelar delas sen lika mellan tredje och fjärde förstärkaren. Resultatet är att alla förstärkarna har fått lika stor effekt.

På utgångarna sammansätts effekterna med en likadan krets. En förutsättning är att utgångskretsen är vänd så att signalerna adderas i fas.

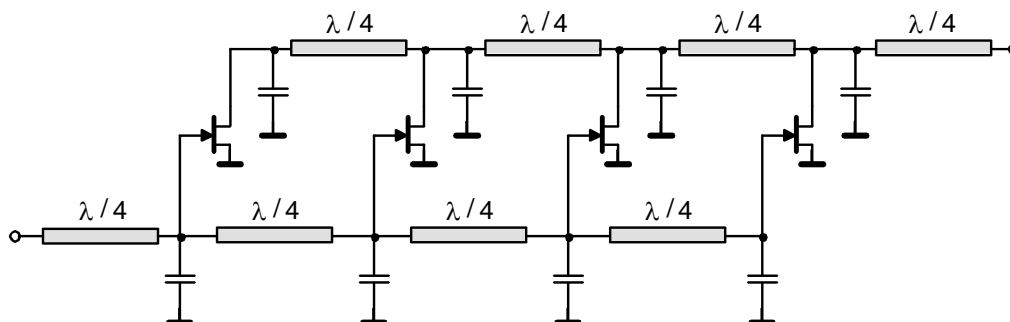
Om förstärkarna är balanserade med hybrider, kan även effektdelningen ske med hybrider. Allt läggs då på samma kretskort. Med kopplade ledningar är det lätt att uppnå den svaga koppling som behövs, om det är uppdelning till många förstärkare.



En Wilkinson effektdelare kan få olika effektdelning genom att välja olika impedanser på ledningarna. Med en kvarts våglängd avstånd mellan förstärkarna hamnar eventuella reflektioner i isolationsmotstånden. Kretsen får alltså bra anpassning på ingången.

Seriekombinering ger en kompakt krets med låga förluster, som även kan användas på mm-våg. Uppdelningen kan ske med valfritt antal utgångar. Förstärkarna kan placeras på valfritt avstånd, eftersom det är lika långa ledningar på in- och utgång. Förstärkarna sprids ut för att få plats och för att få god kylning. En nackdel är att signalen måste passera flera effektdelare. Det ger sämre verkningsgrad för kombineringsen, vilket begränsar antalet förstärkare.

Extended resonance



En FET har en strökapacitans C_G samt en konduktans G_G på ingången. Transistorns utgång har motsvarande C_D och G_D .

På en kvarts våglängd lång ledning inverteras impedansen. Den inverterade reaktansen kompenserar intilliggande transistors reaktans. Impedansen blir därmed helt resistiv. Transistorstegen konjugatanpassar varandra.

Den sammanlagda konduktansen behöver därefter transformeras till 50Ω med en $\lambda/4$ transformator.

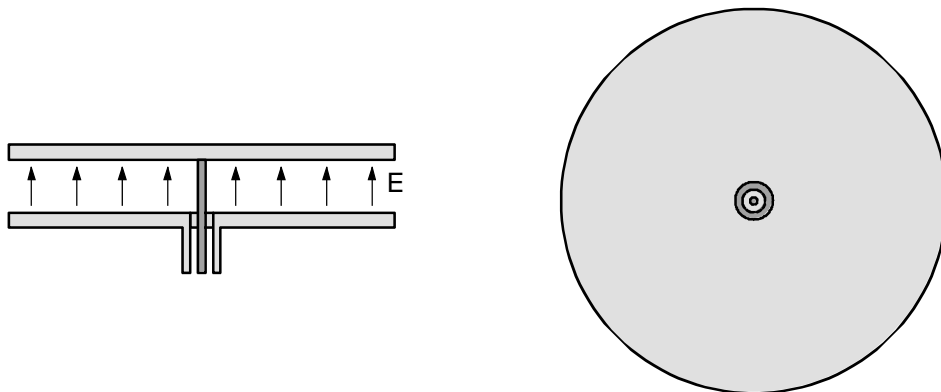
Yttre kapacitanser har adderats för att justera faskången mellan sektionerna. Kapacitanserna kan vara i form av öppna stubbar i microstrip. Ledningarnas impedanser är valda så att alla transistorer får lika stor signalamplitud.

Kretsen ser ut som en distribuerad förstärkare. Men en distribuerad förstärkare har sina transistorer fördelade längs en transmissionsledning som är resistivt avslutad. Nackdelen med distribuerade förstärkare är att signalen dämpas längs ledningarna. Transistorerna blir därför inte utstyrda lika mycket. På utgången betyder dämpningen reducerad uteffekt. Dessutom beror uteffekten på transistorernas lastimpedans. I den distribuerade förstärkaren är enbart faskången optimerad, inte lastimpedanserna.

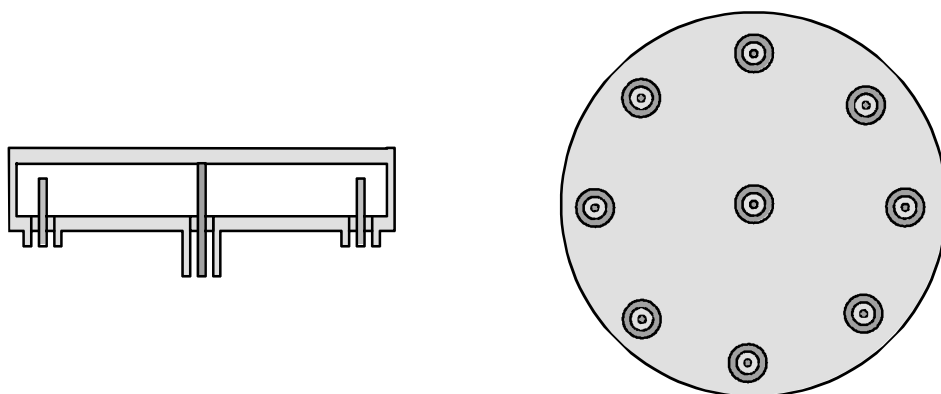
En extended resonance förstärkare har ingen resistiv avslutning på ledningarna. Kretsen får istället betraktas som ett filter, där sektionerna kopplas ihop med inverterare.

Den distribuerade förstärkaren är lämplig vid extremt stora bandbredder. Förstärkaren med extended resonance ger högre uteffekt, men för bandbredder mindre än 20 %.

Radialkombiner

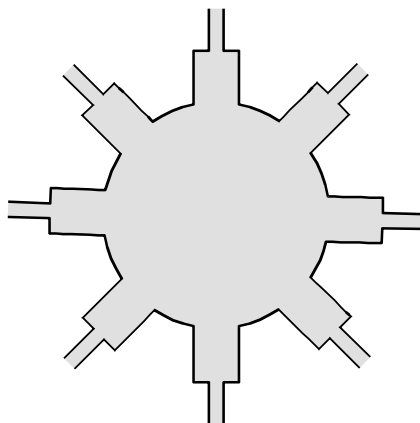


En radialkombiner består av två runda plattor. Signalen matas i centrum och går mellan plattorna radiellt ut mot periferin. Koaxialkontakten matar området mellan plattorna med en E-fält probe. Elektriska fältet är alltså orienterat vinkelrätt mellan plattorna. Magnetiska fältet går i slingor runt mittledaren och sträcker sig ut mot periferin. Signaltransporten sker alltså med en TEM-mode. Avståndet mellan plattorna är mindre än $\lambda/2$ för att inte vågledarmoder ska få plats. E-fält proben i centrum har mycket mindre diameter än plattorna.



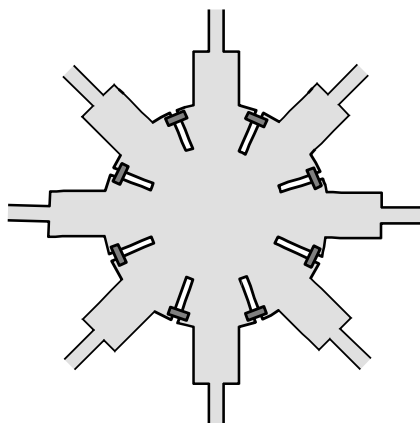
Eftersom fältet är lika runt periferin, kan ett valfritt antal portar kopplas med E-fält probar. Radialkombinern används om det är fler än 8 portar. Ju fler portar desto större diameter måste plattorna ha för att portarna ska få plats. Det har till och med gjorts kombinerings med över 100 portar.

Radialkombiner i microstrip



Fältet runt periferin kan kopplas ut till ett antal portar i microstrip. Fortfarande matas signalen i centrum med en E-fält probe vinkelrätt mellan plattorna. Men om det är ett tunt substrat blir impedansen mellan plattorna mycket låg. Koaxialingången kan då behöva en transformering med $\lambda/4$ ledning.

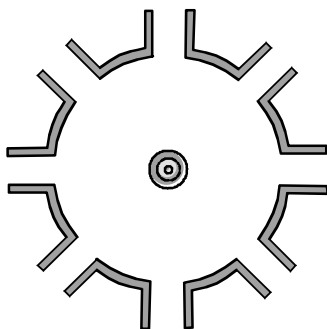
Den radiella transmissionsledningen ger en viss impedanstransformering. I microstrip kan dessutom utgångarna lätt förses med extra $\lambda/4$ transformering.



Högre moder kan exciteras av osymmetrier. Det försämrar både anpassningen, balansen i fas- och amplitud, isolationen och förlusterna. Den önskade TEM-moden alstrar strömmar radiellt utåt. Radiella $\lambda/4$ slitsar med resistans påverkar inte den önskade moden. De högre moderna ger strömmar på tvären, och dämpas därför resistivt. Den mer komplicerade kretsen blir dyrare och får i praktiken högre förluster i den önskade moden.

Radialkombiner används för att sammankoppla ett stort antal effektförstärkare. Det kan vara lämpligt att skydda en förstärkare med en ferritisolator (cirkulator) på utgången. I så fall kan man klara sig utan extra isolationsresistans i själva kombineringen.

Radialkombiner i vågledare



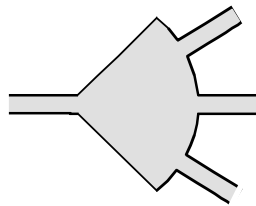
Portarna på periferin av den radiella ledningen kan ha formen av rektangulär vågledare. Det ger en mycket enkel krets som tål hög effekt. Porten i centrum är ofta en E-fält probe, men det går också att göra vågledarövergångar som alstrar det önskade cirkulära fältet.

Om anslutningen till vågledarna förses med bländare, kan den cylindriska kaviteten betraktas som en resonator. Det ger en mycket smalbandig kombineringskrets. Resonatorkombinering används främst till oscillatorer för hög effekt.

En vågledarstruktur blir ganska smalbandig. En ryggvågledare kan klara en bandbredd på över en oktav. Den radiella sektionen förses då med ryggar, som går ut mot periferin. Ryggarnas höjd kan varieras för att ge impedanstransformering. I periferin kan till och med ryggar vara så höga att det motsvarar 50Ω . I centrum har ryggar mycket låg höjd så att impedanserna blir höga. Det förenklar anpassningen eftersom alla utgångar i princip är parallellkopplade.

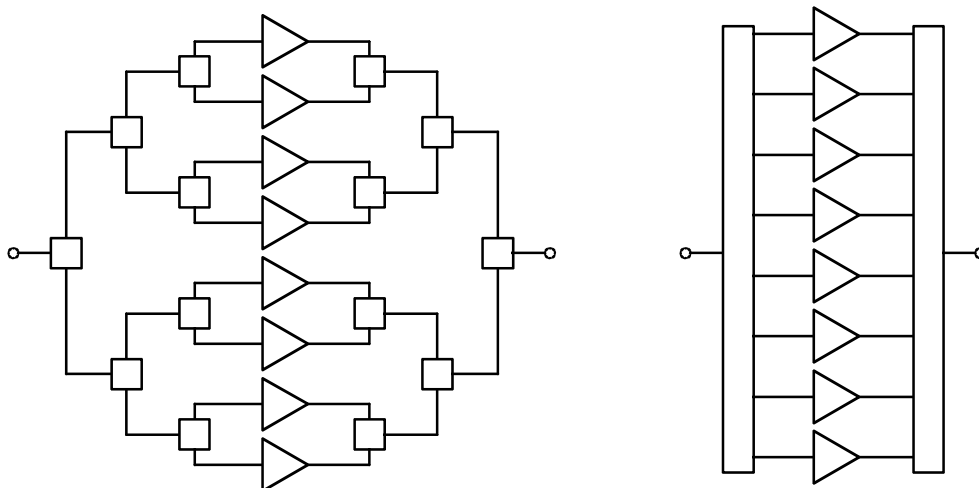
Sektorkombinering

En stor nackdel med radialkombinering är att porten i centrum ger en tredimensionell krets. En sektorkombinering är en förenklad variant, som blir helt plan och kan göras i microstrip.

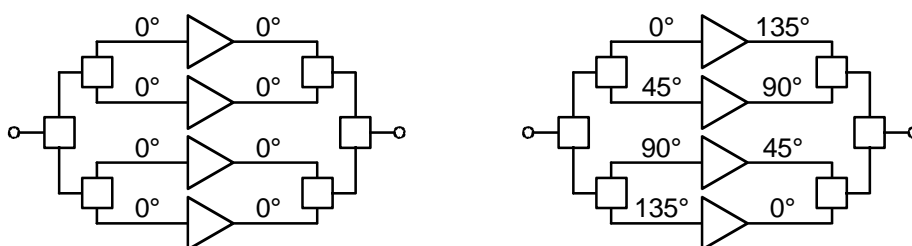


Signalstyrkan i de olika portarna beror på placeringen och bredden på stripledningarna.

Sammanställning



Binär uppdelning och kombinerings används då det är två eller fyra förstärkare som ska sammansättas. Om det är fler förstärkare, måste signalen passera fler passiva kretsar. Förlusterna ger minskning i uteffekt och förstärkning. Då åtta eller fler förstärkare ska kombineras används istället radialkombinern.



Olika kretskopplingar skiljer sig också i hur signalen delas upp i fas. En fördel med att dela upp i olika faslägen är att blandprodukten från dubbla frekvensen får dubbla fasen. IM-distorsionen som alstras av förstärkarna sammansätts alltså inte i fas på utgången.

Referenser

Effektkombinering

Yung-Jinn Chen, "A wide band multiport planar power divider design using matched sectorial components in radial arrangement", IEEE Transactions on microwave theory and techniques, vol 46 no 8 aug 1998 pp 1072-1078

Jae-Wook Rheem. "A microwave radial power combiner design", Microwave Journal 1997 nov pp 22-34

Gary Ferrell, "A high efficiency 10 watt HBT power amplifier assembly using combining techniques", IEEE MTT-S 1995 pp 327-330

J W Gipprich, "A power amplifier yields 10 watts over 8-14 GHz using GaAs MMICs in an LTCC serial combiner/divider network", IEEE MTT-S 1993 pp 1369-1372

D I Stone, "Q- and V-band planar combiners", IEEE MTT-S 1991 pp 1049-1052

Mohamed D Aabouzahra, "Multiport power divider-combiner circuits using circular-sector-shaped planar components", IEEE Transactions on microwave theory and techniques, vol 36 no 12 dec 1988 pp 1747-1751

R P Flam, "Radial power combiners for solid state power amplifiers", Military Microwave 1992 pp 23-28

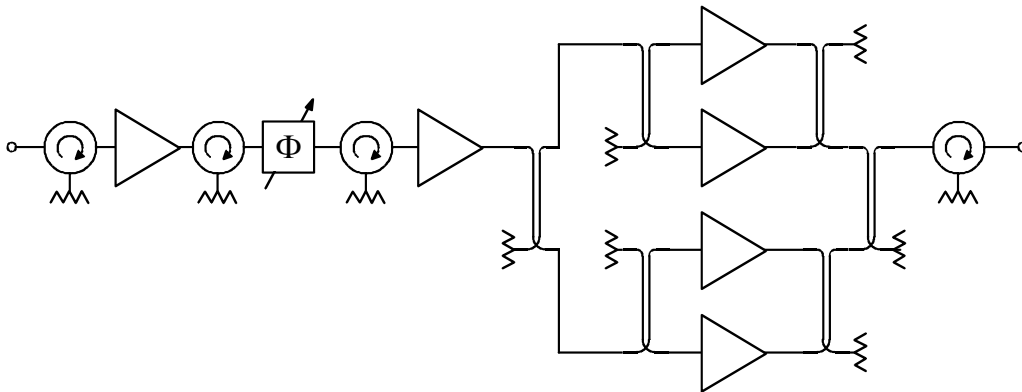
E Belohoubek, "30-way radial power combiner for miniature GaAs power amplifiers", IEEE MTT-S 1986 pp 515-518

Kenneth J Russell, "Microwave power combining techniques", IEEE Transactions on microwave theory and techniques, vol 27 no 5 may 1979 pp 472-478

13. Effektförstärkare för kW

För att uppnå kW effekter behövs det många separata förstärkarmoduler, som sammansätts med någon form av kombiner. Modulförstärkaren har en uteffekt på ca 200 - 500 W. Den kan också användas som drivförstärkare före slutsteget. Totala uteffekten är 1 - 30 kW pulseffekt ($<100 \mu\text{s}$). Det finns också förstärkare för kontinuerlig effekt på 1 kW.

Modulförstärkare



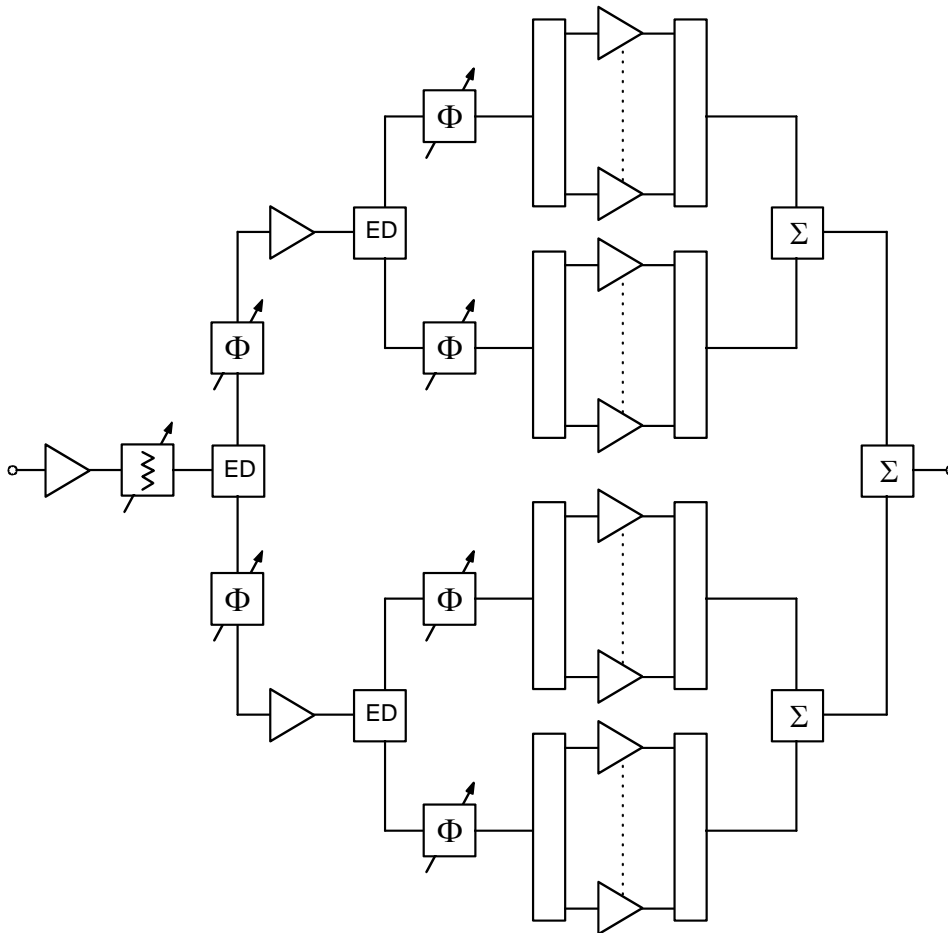
Modulförstärkaren innehåller några effektransistorer som kombineras med hybrider eller Wilkinson. Dessutom behövs några kaskadkopplade steg för att uppnå den önskade förstärkningen. Bipolära transistorer för pulser på S-bandet (2-4 GHz) kan ge 150 W pulseffekt med god verkningsgrad ($>40\%$).

När många modulförstärkare ska sammansättas, behöver samtliga signaler ha samma fasläge. Därför innehåller modulförstärkarna ofta en möjlighet att justera fasen.

På utgången sitter en cirkulator för att skydda transistorerna mot stora reflektioner. En antenn som kopplats loss, eller en kortsluten kabel, ger så stora stående-vågor att spänningen över transistorn kan bli fördubblad. Stor reflektion kan alltså försämra eller helt förstöra effektransistorerna.

På ingången sitter också en cirkulator. Den ska ge modulen god anpassning, även om dess transistorer är trasiga. Före modulförstärkarna sitter ett delningsnät (splitter). Om en modul ger stor reflektion kommer delningsnätet i obalans, så att ineffekten fördelas ogynnsamt mellan modulerna. Totala uteffekten minskar då med mer än den trasiga modulen. Modulen kan innehålla fler cirkulatorer för att få anpassning och fasstabilitet.

Slutsteg med radialkombiner

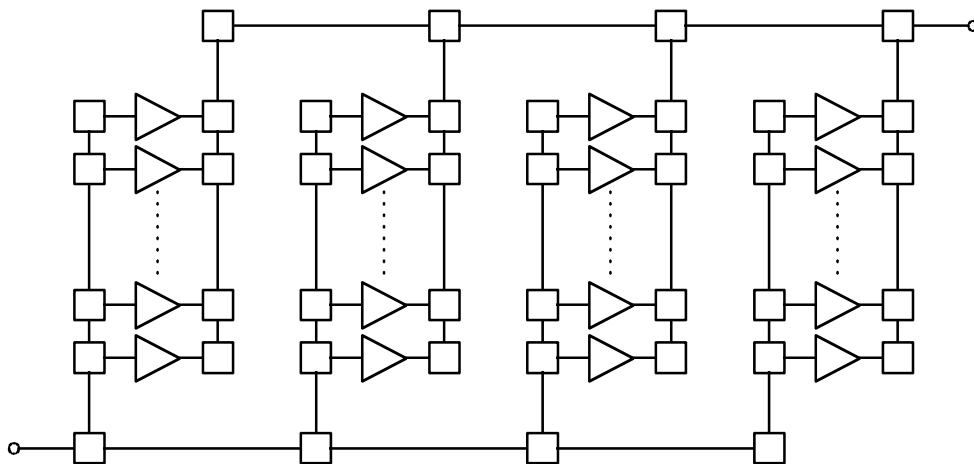


Slutsteget kan innehålla radialkombiner för att få låga förluster. I det här fallet är det en uppdelning till fyra separata grupper, som sen sammansätts i vågledare med mycket låga förluster. Kretsen för uppdelningen på ingångssidan kan vara byggd i strip-line med luftdielektrika.

Även slutsteget kan innehålla några intrimningar av fasen, för att sammansätta signalerna till maximal uteffekt.

Det finns ingen cirkulator för att skydda mot missanpassning i det här fallet. Varje modul har ju varsin cirkulator som skydd. Alternativet med en gemensam cirkulator på utgången av slutsteget, skulle behöva klara mycket högre effekt.

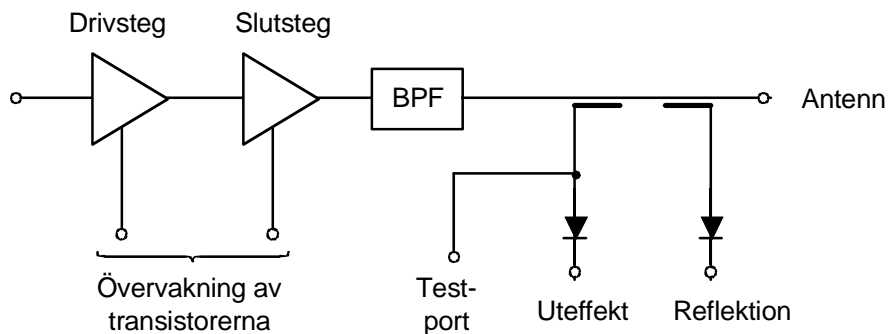
Slutsteg med seriekombinering



Seriekombinering (Trawling Wave) kan kombinera många modulförstärkare. I figuren är det ritat som ett antal förstärkare ovanför varandra. Därefter har fyra sådana seriekombinerade förstärkare sammansatts med seriekombinering.

På X-bandet (ca 10 GHz) har själva transistorerna inte så hög uteffekt. Med fyra grupper, som har 13 stycken seriekombinerade transistorer, har man uppnått 0,5 kW. Fyra sådana förstärkare har kombinerats till 2 kW kontinuerlig effekt.

Övervakning



Normalt förses effektförstärkaren med riktkopplare på utgången för att mäta uteffekten och den reflekterade signalen. De uppmätta signalerna kan vara kopplade till nivåreglering för att hålla konstant uteffekt, samt till automatisk avstängning då det blir fel på antenn. En testport används för att göra andra mätningar på den utsända signalen.

Samtliga effektransistorer kan övervakas med sensorer för temperatur och spänning. Systemet kan vara uppbyggt med extra modulförstärkare som automatiskt kopplas in om det blir något fel.

Tillförlitlighet

Effektförstärkarens tillförlitlighet kan bli högre, om transistorn inte arbetar på sin maximala effekt, utan på ca 10 % lägre kollektorspänning. Dessutom håller transistorn längre om den inte blir så het. Tio grader högre temperatur kan öka felfrekvensen (MTTF) med 60 %. Det vill säga livslängden halveras för varje 10 graders ökning av temperaturen. Dessutom minskar uteffekten med 0,2 dB.

Referenser

kW förstärkare

Michael Hanczor, "12 kW S-band solid state transmitter for modern radar systems", IEEE Transactions on microwave theory and techniques, vol 41 no 12 dec 1993 pp 2237-2242

H Ashoka, "An X-band 2 kW CW GaAs FET power amplifier for continuous wave illuminator application", IEEE MTT-S 1998 pp 1149-1152

J D Hay, "The exploratory development of a high power S-band solid state radar transmitter", IEEE International radar conference 1990 pp 135-140

14. MCPA

MCPA — Multi Carrier Power Amplifier

MCPA är benämningen på den effektförstärkare som används inom telekommunikation då flera kanaler (carriers) ska förstärkas samtidigt. Den kan också kallas flerkanalförstärkare. Med MCPA kan antalet signaler och avståndet mellan dem väljas fritt. Signalerna kan till och med ha olika slags modulation.

Satellitkommunikation använder främst TWT-förstärkare klass A, som arbetar på lägre effektnivå (back-off) för att vara tillräckligt linjära. På senare tid har de linjäriserats med fördistorsion för att kunna arbeta närmare maxeffekten. På så sätt kan verkningsgraden förbättras.

Kabel-TV behöver förstärkare för många simultana signaler. Förstärkaren behöver ha stor bandbredd och låg IMD.

Mobiltelefonsystem har större krav på dynamikområdet, eftersom en svag fjärran signal kan störas av en närliggande stark sändare. Effektförstärkaren behöver ha mycket god linjäritet, för att inte IMD (Inter-Modulation Distorsion) ska bli störande. IM-produkterna ska vara minst 60 dB lägre än bärvågen.

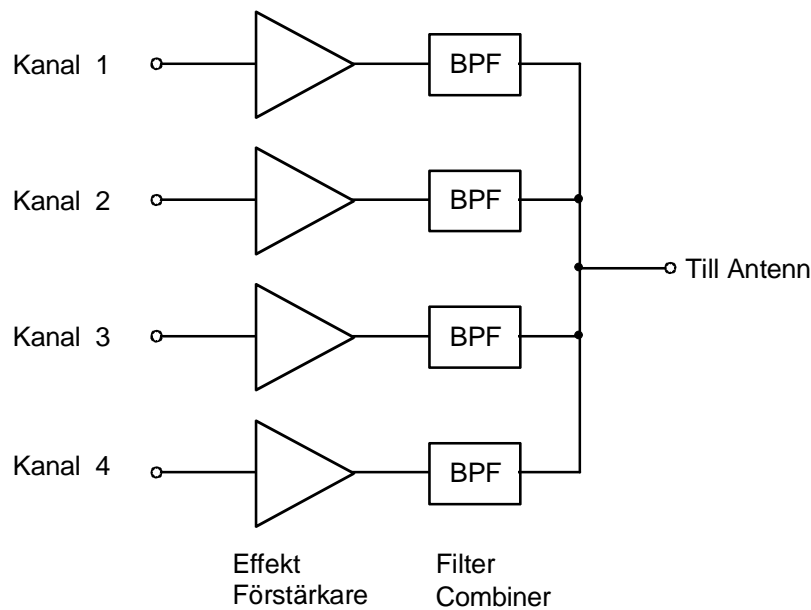
För att få så låg distorsion används reducerad uteffekt och linjäriseringskretsar. Det resulterar i att verkningsgraden blir bara 5 - 10 %.

Förstärkaren har en viss total effekt som fördelas till aktuella signaler. En 100 W förstärkare kan ge 100 W uteffekt om det bara är en signal. Två signaler får vardera 50 W, och om det är 5 signaler blir det 20 W per signal.

Ett system för mobiltelefon kan starta med ett fåtal basstationer på hög effekt. När sedan systemet byggs ut kan samma förstärkare användas för flera signaler. Vid utbyggnad med fler basstationer passar den lägre effekten.

Befintliga basstationer använder främst en förstärkare per kanal. När dynamisk frekvenstilldelning och frekvenshopp införs, blir det lämpligt att gå över till kombinerad hybrid. När antalet kanaler blir stort (t.ex. 8 kanaler per sektor) blir det billigare med MCPA. Dessutom blir det mer skrymmande med många separata effektförstärkare.

Kombinering med kaviteter

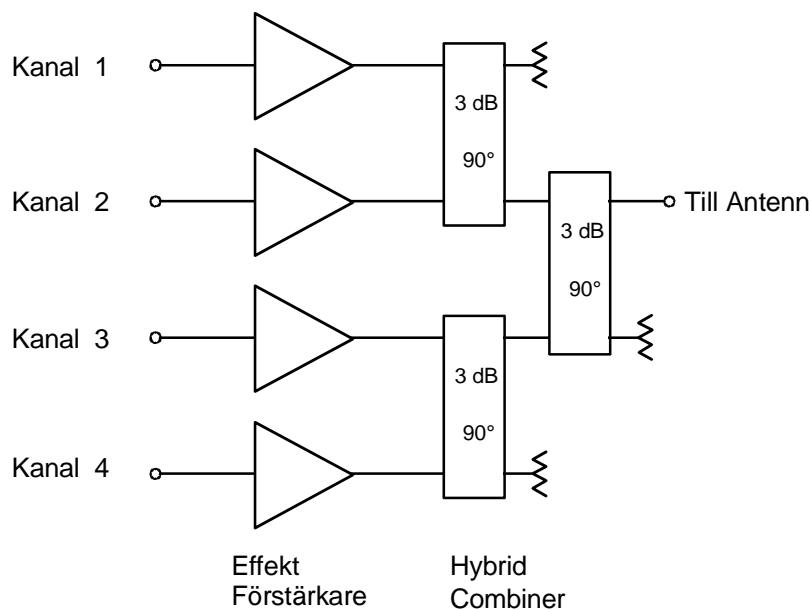


Befintliga system för mobiltelefon har använt filter till kombineringsen (cavity combiner). Med resonatorer i form av stora kaviteter blir det ganska låga förluster, ca 1,5 dB. Varje förstärkare behöver bara hantera en kanal. I GSM-systemet är dessutom amplituden konstant. Där kan alltså effektiva klass C förstärkare användas. Men varje förstärkare måste skyddas med dubbla cirkulatorer på utgångarna, för att inte grannkanalerna ska nå fram till slutsteget och skapa intermodulation.

Nackdelen med att kombinera med kavitet är att kanalerna blir låsta till sina frekvenser. Det finns visserligen kaviteter med automatisk avstämning, men det är en ganska långsam procedur.

En annan nackdel är att det behövs ett skyddsområde på en kanalbredd mellan de olika kanalerna.

Kombinering med hybrider



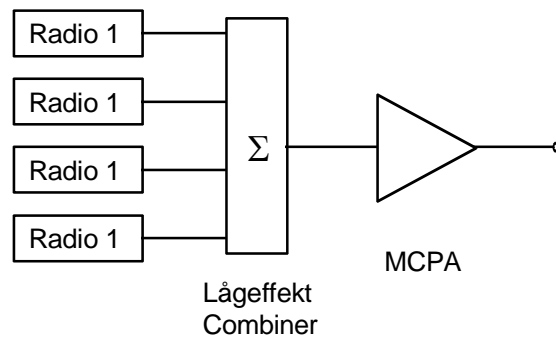
Två signaler kan kombineras i en hybrid. Med tre hybrider kan fyra kanaler kombineras. Fördelen med hybrider är att de kan göras bredbandiga. Kanalerna kan alltså ges valfri frekvens. Befintliga frekvenser kan dynamiskt utnyttjas i systemet (DCA - dynamic channel allocation). GSM-systemet kan använda långsamt frekvenshopp för att öka kapaciteten.

Eftersom kanalerna inte separeras med filter, behövs inte stora skyddsområden för filterflankerna. Med kanalerna tätt packade utnyttjas spektret effektivt.

Den stora nackdelen är att varje hybrid ger minst 3 dB extra förluster. Med fyra kanaler blir det alltså 6 dB extra förluster mellan förstärkare och antenn. Dessa förluster är särskilt besvärande eftersom de ligger efter effektförstärkaren. Mycket dyrbar effekt går förlorad innan den kommer till antennen.

En kompromiss kan vara att först fördela de udda kanalerna till en kavitetskombiner, och de jämna kanalerna till en annan kavitetskombiner. Därefter sammansätts de två grupperna med en hybridkombiner.

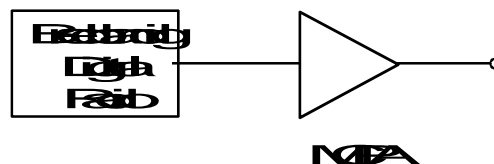
Kombinering på låg effekt



Om kretsen för kombinerings placeras före effektförstärkaren istället för efter, slipper man den stora förlusten i uteffekt. Men effektförstärkaren behöver nu kunna hantera flera signaler samtidigt, utan att det bildas störande blandprodukter. Flerkanalförstärkaren behöver ha mycket god linjäritet.

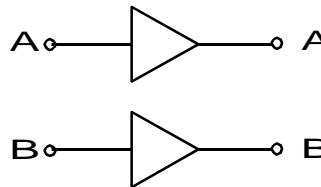
Eftersom förstärkaren är linjär, kan olika typer av modulation användas samtidigt. Frekvenser och antal kanaler kan variera dynamiskt.

En linjär effektförstärkare har ganska låg verkningsgrad, ca 5 - 10 %. Alternativet med kavitetskombinering ger också lika låg verkningsgrad, eftersom det behövs flera förstärkare. En extra fördel med MCPA är att distorsionen blir mycket lägre, både i egna bandet och i grannkanalerna.

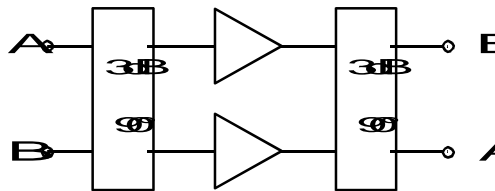


En linjär flerkanal förstärkare (MCPA) kan också användas i framtida system där största delen av radion ligger i en digital krets.

Multiport effektförstärkare



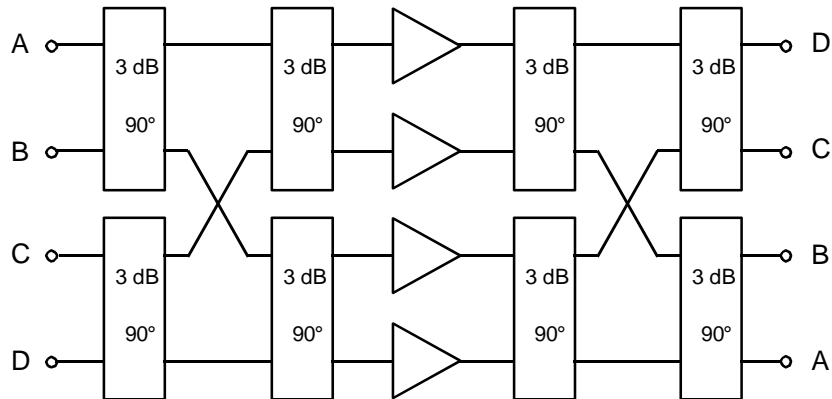
Två separata kanaler använder vanligtvis varsin effektförstärkare. En nackdel är att om ena förstärkaren går sönder, slutar den kanalen helt att fungera.



Om två förstärkare kopplas ihop med två hybrider kommer signalen att sammansättas diagonalt till utgången. Porten rakt fram blir isolerad, eftersom de två signalvägarna ligger i motfas. Två oberoende signaler kan därför samtidigt förstärkas till sina respektive utgångar.

Varje kanal använder alltså båda förstärkarna. Om en förstärkare går sönder minskar totala uteffekten för båda kanalerna. Däremot är det inte så illa att ena kanalen slutar fungera helt.

Ytterligare en fördel är att förstärkarna blir jämnt belastade, trots att ena kanalen har stor trafik samtidigt som den andra har mycket liten trafik.



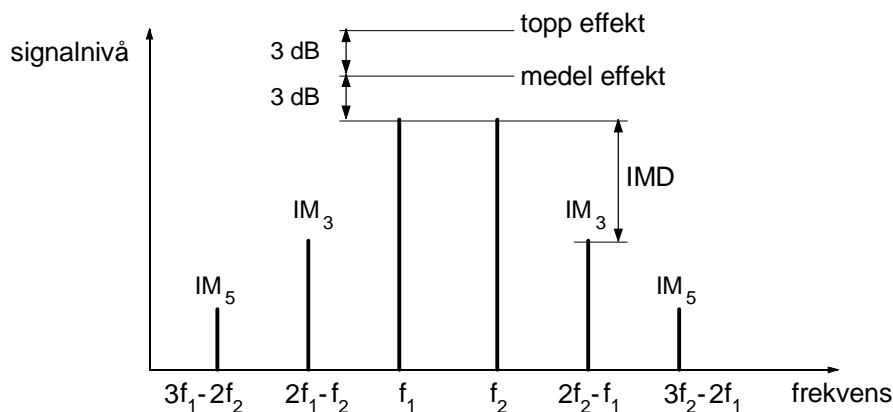
Systemet kan byggas ut med 4 eller 8 förstärkare. Det ger ännu bättre tillförlitlighet och jämnare fördelning av kanalernas kapacitet.

En nackdel är förlusterna i hybriderna som minskar den dyrbara uteffekten. Dessutom behöver förstärkarna vara tillräckligt linjära för att förstärka flera samtidiga signaler utan distorsion.

IM-distorsion

IMD – Inter Modulation Distorsion

En olinjär förstärkare ger distorsion i form av olika blandprodukter. De blandprodukter som hamnar inom det egna frekvensbandet och inom grannkanalen, går inte att filtrera bort utan måste undertryckas för att inte störa kommunikationen.



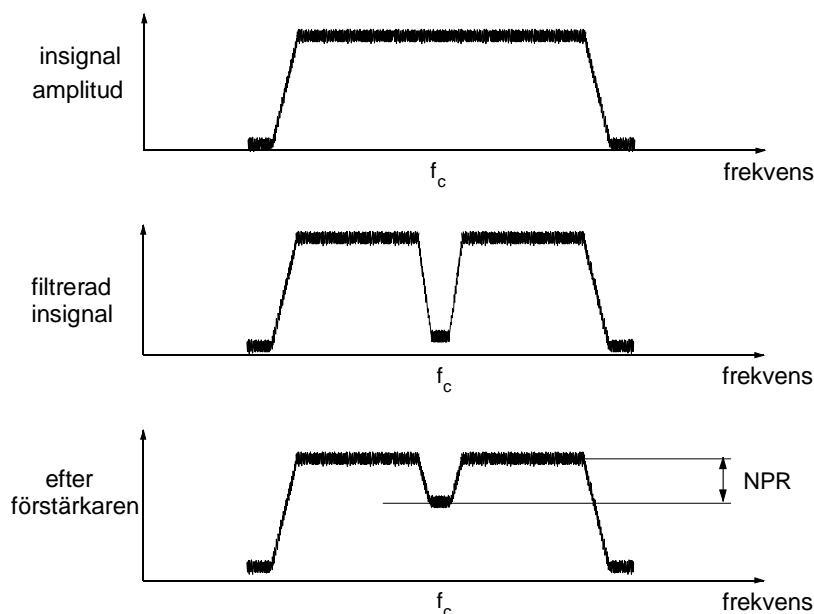
IMD kan mätas med två kontinuerliga signaler (CW-signaler). IMD definieras som förhållandet mellan den största IM-produkten och den största av insignalerna. Vid mätningen används vanligen två lika starka insignaler. IM-produkterna jämförs då med den ena av insignalerna. Båda insignalerna tillsammans ger dubbla ineffekten, dvs 3 dB högre signalstyrka. Toppeffekten är ytterligare 3 dB högre. Men det är alltså endast signalnivån för den ena insignalen som används som referens.

En klass A förstärkare har ca -30 till -35 dBc IM-distorsion. En linjäriserad förstärkare kan få en IMD på -60 till -70 dBc. De flesta förstärkare har IM3 som den största blandprodukten, men i en linjäriserad förstärkare kan det vara IM5 som är störst.

IM-brus i förhållande till signal

NPR – Noise Power Ratio

Med 10 insignaler bildas det totalt 450 stycken 3:e gradens blandprodukter, och 7050 stycken 5:e gradens blandprodukter. Vid flera jämnt fördelade signaler hamnar den största delen av 3:e gradens IM-produkter inom bandet. Eftersom insignalerna är modulerade blir också IM-produkterna modulerade. Spektret blir alltså mer bruslikt. NPR kan därför ersätta två-tons mätningarna för att få ett mått på distorsionen inom bandet.



Aktuellt band är fyllt av bruslika insignaler. Ett notch-filter tar bort en del av spektret mitt i bandet. Efter effektförstärkaren har det bortfiltrerade området istället fyllts med IM-distorsion.

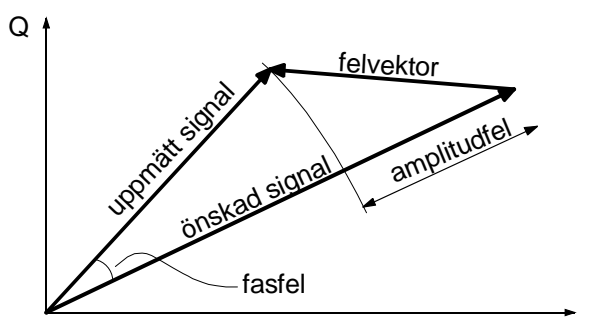
NPR är förhållandet mellan insignalernas medeleffekt och medeleffekten av IM-distorsionen i det filtrerade området.

Med filtrering mäts tyvärr inte den totala IM-effekten. Ett alternativ är att utsätta förstärkaren för hela bandets insignaler, och på utgången mäta IM-effekten med insignalerna undertryckta genom utfasning.

Felvektor

EVM – Error Vector Magnitude

GSM-systemet ger enbart variationer i fas. Det specificeras med fasfelet max 5° rms och 20° topp. EDGE och CDMA-system får dessutom variationer i amplitud. Det är då lämpligare att mäta felvektorn.



Uppmätt signal jämförs med den förväntade. Istället för att ange både fasfel och amplitudfel anges storleken på felvektorn. EMV är rms-värdet av felvektorns amplitud för olika tillstånd under kommunikationen.

EVM anges i % jämfört med den önskade signalen. EDGE får ha högst 9 % EVM över en burst. Toppvärdet av EVM under en symbol får vara högst 30 %. Specifikationen för 3G är EVM mindre än 17,5%.

EMV är en lämplig mätning, under förutsättning att mätsignalen har samma ISI (inter symbol interferens) som den förväntade referensen. Mätsignalen ska alltså ha samma filtrering av datapulserna som aktuellt system.

Signal-stör-förhållandet

IM-distorsionen ska vara så låg att den inte stör kommunikationen. Svaga signaler begränsas också av brusnivån. Systemet ska klara av ett visst signal-brus-förhållande (C/N — carrier to noise ratio). Signal-stör-förhållandet (C/I — carrier to interferens ratio) ska vara så stort att störningarna blir mycket mindre än bruset. Systemet begränsas av störningarna på samma sätt som bruset. Begränsningen är alltså den sammanlagda brus och distorsionen ($C/N+I$).

$C/I = C/N$ ger 3 dB försämring av systemets känslighet

$C/I = C/N + 6 \text{ dB}$ 1 dB

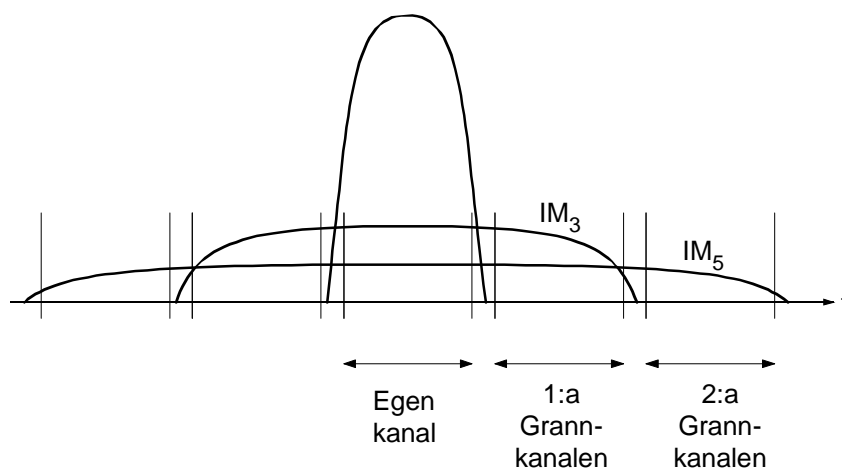
$C/I = C/N + 10 \text{ dB}$ 0,05 dB

En MCPA till ett GSM-system behöver vara extremt linjärt. Med 20 W uteffekt ska IMD klara -75 dBc. AMPS och IS-136 kräver att störningarna undertrycks minst -60 dBc.

Grannkanalstörning

ACPR - Adjacent Channel Power Ratio

Största delen av IMD hamnar i den egna kanalen, men lite effekt sprids också till grannkanalen. ACPR är ett mått på hur mycket effekt som hamnat inom grannkanalens bandbredd, i förhållande till effekten inom den önskade kanalens bandbredd.



Spektrumspridningen beror på förstärkarens olinjäritet. 3:e gradens olinjäritet ger ungefär tre gånger större bandbredd, och 5:e gradens olinjäritet ger fem gånger större bandbredd.

Både IM_3 och IM_5 hamnar inom 1:a grannkanalen. Totala störningen är vektorsumman, som alltså kan bli större eller mindre än IM_3 . 2:a grannkanalen innehåller endast IM_5 .

En W-CDMA signal med bandbredden 3,84 MHz har IM_3 inom 1,92 – 5,76 MHz från bärvågen. Ett system för flera samtidiga kanaler (multi-carrier), med 4 intilliggande W-CDMA kanaler, har en bredd på 18,84 MHz. IM_3 sträcker sig då upp till 28,26 MHz från centrum av de fyra signalerna. Det betyder att fler grannkanaler kan få problem med störningar (ACPR).

IMD inom bandet ökar med kvadraten på ineffekten. ACPR ökar med kuben på ineffekten. Dessutom kan grannkanalen lyssna på svaga signaler samtidigt som störningarna är stora. Så även om egna distorsionen inom bandet är större än grannkanalstörningen, blir det svårare att uppnå kraven för ACPR.

ACIR (adjacent channel interference ratio) är ett annat mått på mängden störningar från grannkanalen. Den beror på både sändarens olinjäritet och selektiviteten på mottagarens filter.

GSM	200 kHz offset	-30 dBc	3G	5 MHz offset	-45 dBc
	400 kHz	-60		10 MHz	-50 dBc
	600 kHz	-70			
IS95	750 kHz	-45	PCS	885 kHz offset	-45 dBc
	1.98 MHz	-60		1,25 MHz	-80

Topp-effekt

En signal som varierar i amplitud har en topp-effekt som är högre än medeleffekten (PAR – Peak to Average Ratio). Förhållandet mellan topp- och medeleffekten beror på typen av modulation och hur hårt signalen är filtrerad. Hårdare filtrering ger större överslängar i amplitud.

GSM EDGE	3,2 dB	
W CDMA	10,5	för 16 aktiva kanaler
	13,6	128 aktiva kanaler
	6	Handapparat på högsta datahastighet
$\pi/4$ QPSK	3,2 - 4,8	beroende på filterflankerna
16 QAM	5 - 6	beroende på filterflankerna
OFDM	7 - 13	

W CDMA innehåller olika spridningskoder för olika datahastigheter. De olika koderna skiljer sig i förhållandet mellan topp-effekt och medelvärde. För samma grannkanalstörning kan signalnivån behöva en justering på upp till 4 dB.

Kabel-TV system kan kombinera över 100 signaler. Förhållandet mellan topp- och medeleffekt blir ca 11 - 13 dB.

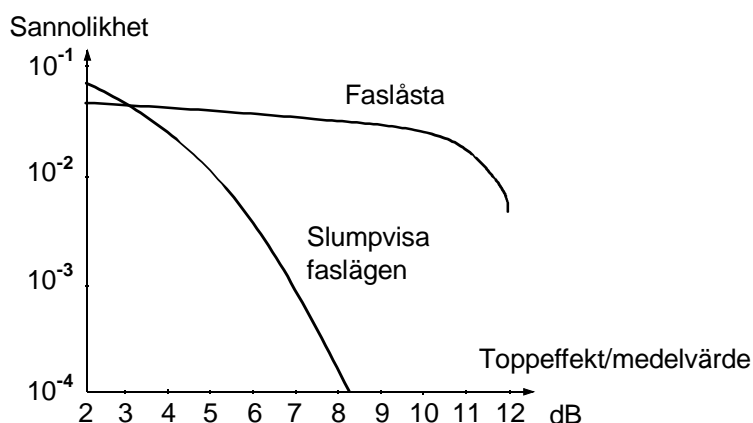
Om man räknar om till förhållande mellan amplituder används benämningen ”crest factor”. För en modulerad RF-signal används crestfaktorn för signalens envelop. Crestfaktorn för RF-signalen är alltså detsamma som för basbandet. En IQ-modulerad signal får en crestfaktor som är lägre än respektive kanal, eftersom de båda kanalerna inte går genom noll samtidigt. Men minskningen i crestfaktor är högst 3 dB.

Med två lika stora signaler blir den sammanlagda medeleffekten dubbelt så stor. Det motsvarar 3 dB högre effekt. Den största topp-effekten får man, vid de tillfällen då de två signalernas amplituder råkar adderas i fas. Dubbla spänningen motsvarar 6 dB högre topp-effekt. Varje dubbling av antalet signaler ökar medeleffekten med 3 dB och topp-effekten med 6 dB.

Med n stycken signaler där varje signal har effekten P , blir den totala medeleffekten nP och topp effekten (peak envelop power) n^2P . 16 stycken signaler på vardera 10 W, ger tillsammans en topp effekt på 2,56 kW. Förstärkaren och efterföljande kretsar behöver alltså klara den topp effekten. Om man istället utgår från förstärkarens maximala uteffekt, får varje signal en uteffekt som är $(20 \log_{10} n)$ dB lägre än förstärkarens maximala effekt. En förstärkare på 100 W ger alltså endast 0,4 W per signal, om det är 16 stycken.

Ju fler signaler desto lägre är sannolikheten att alla ska adderas i fas. Den största tiden har signalen ganska låg amplitud.

Signalens crestfaktor (PAR) specificeras med sannolikheten att överstiga en viss nivå (CCDF – Complimentary Cumulative Distribution Function).



Figuren visa sannolikhetsfördelningen för 16 omodulerade signaler (CW). Om alla signaler är låsta till samma referens, kan effekten öka 12 dB (dvs $10 \cdot \log 16$). Men om signalerna har slumpmässiga faslägen, blir det till och med mycket lägre sannolikhet att signalstyrkan överstiger medeleffekten med mer än 8 dB. Den effektiva distorsionen från 16 signaler, är så stor som man hade väntat sig från bara 6 signaler. När antalet signaler ökar kommer sannolikheten att alla ska adderas i fas att minska. Och när det uppstår en topp i effekten, så är den så kort att den kanske inte märks.

Förhållandet mellan toppeffekt och medeleffekt (PAR) anges alltså för ett visst mått på sannolikheten. Vid till exempel 10^{-4} kommer 1 topp av 10000 att bli större, och klippas i förstärkaren. Distorsionens effekt blir då bara 1/10000 dvs 40 dB lägre. Om en förstärkare ger -20 dBc distorsion vid klippning, blir alltså distorsionen -60 dBc. Det är alltså inte enbart PAR som påverkar distorsionen utan även hur ofta som amplituden begränsas av kretsen.

En CDMA-signal för IS95 anges för en sannolikhet på 10^{-4} (dvs 0,01 %). Toppeffekten blir 9,7 dB över medeleffekten, då kanalen innehåller 9 koder. Som mest är toppeffekten 14 dB högre än medeleffekten. Upplänken, som använder OQPSK, har 0,01 % sannolikhet att uppnå 5 dB PAR.

Effektförstärkaren har själv en maximal signalnivå ut, och ger alltså en limitering. Men det är bättre att använda en särskild limiter på förstärkarens ingång. Med begränsad signalnivå på ingången, ställs mindre krav på slutstegets effekttålighet.

Reducering av crestfaktorn

Om man har kontroll över hela systemet, är det bra att försöka minska crestfaktorn. Det minskar kraven på effektförstärkarens linjäritet, och resulterar i högre verkningsgrad.

W-CDMA har vanligtvis högre crestfaktor för intilliggande orthogonala koder. Det är alltså bättre att kombinera koder som är utspridda. En signal med de fem kontrollkanalerna och åtta trafikkanaler kan få 3 dB lägre crestfaktor, genom att välja lämpliga koder.

Pilotdata är inte kanalkodat. Alla DPCH signaler har samma pilotsymboler. De sammansätts därför till hög effekt. Genom att tidsförskjuta bärvågens trafikkanaler, kan crestfaktorn minskas 2 dB.

Tidsförskjutningen kan också utnyttjas för att minska crestfaktorn då flera bärvågor sammansätts (multicarrier). 3GPP föreslår en fördröjning på 1/5 slot (133 μ s) för signaler på intilliggande bärvågor.

Limitering är ytterligare ett sätt att minska crestfaktorn. Det kan göras med en RF-krets eller på basbandet, med en matematisk funktion, innan pulsformningen. Nackdelen med limitering är att en del av signalen klipps bort. Det ger högre BER. Men om klippningen sker sällan, är det bättre att låta de felrättande koderna (FEC) hantera bitfelen.

Referenser

MCPA

Harald Pret, "Linearity considerations of W-CDMA front ends for UMTS", IEEE MTT-S 2000 pp 433-436

J S Kenney, "Design considerations for multicarrier CDMA base station power amplifiers", Microwave Journal 1999 feb pp 76-86

Charles J Meyer, "Test peak to average power of vector modulated signals", Microwave & RF 1993 jan pp 123-141

Markus Banerjee, "Generate wide dynamic range WCDMA / 3GPP signals", Microwave & RF 2003 feb pp 101-108

15. Linjärisering

En olinjär förstärkare alstrar blandprodukter mellan olika frekvenser. Dessa IM-produkter (Inter Modulation) hamnar dels i egna kanalen som IMD (IM-distorsion) samt i grannkanalerna som störning. Linjärisering innebär att IM-produkterna undertrycks.

En klass AB förstärkare med FET eller bipolär transistor har en IM-distorsion på ca 30 dBc. Om önskemålet är IMD på 70 dBc behövs ytterligare 40 dB undertryckning. En klass AB förstärkare med LDMOS har en lägre distorsion på 40 dBc och behöver därför bara 30 dB extra undertryckning.

En effektförstärkare i en handapparat har inte alls så hårda krav på IMD som en flerkanalförstärkare (MCPA). En enkel linjärisering kan användas för att få högre avkastning i produktionen (yield). Även en lågbrusförstärkare kan ha krav på IMD som kan förbättras med en enkel linjärisering.

Ett sätt att minska IMD är att minska signalstyrkan (back off). När insignalerna minskar 1 dB kommer grundtonerna att minska 1 dB, men IM3 minskar hela 3 dB. En förstärkare som har IM3 på -25 dBc, kan uppnå -55 dBc genom att minska signalstyrkan 10 dB. Nackdelen är just att uteffekten blir lägre. Eller om man vänder på det, istället för en förstärkare på 100 W behövs det en förstärkare på 1 kW.

En klass A förstärkare har låg verkningsgrad. Om den ska kunna ge stor uteffekt blir det mycket stor effektförbrukning. En klass AB eller klass B förstärkare har mycket bättre verkningsgrad, men den har dålig linjäritet. De olika kretskopplingarna som förbättrar linjäriteten är:

- Fördistorsion
- Adaptiv fördistorsion
- Återkoppling
- Kartesisk återkoppling
- Framåtkoppling
- EER
- LINC

En viktig användning av linjäriserade effektförstärkare är flerkanalförstärkare (MCPA). Det behövs mycket god linjäritet för att inte kanalerna ska störa varandra.

Fördistorsion

Fördistorsion är en ganska enkel krets som kan förbättra en redan befintlig förstärkare. Den kan till och med förbättra linjäriteten då förstärkaren är nära mättnad. Men vid mycket höga linjäriteter ($C/I > 50$ dB) behövs istället framåtkoppling (feed forward) eller adaptivt styrd fördistorsion.

En fördel med fördistorsion på RF eller MF är att hela bandbredden, för en förstärkare eller ett system, kan linjäriseras på en gång. Den är därför lämplig till bredbandiga flerkanal förstärkare i t.ex. satellitsystem eller basstationer till mobiltelefon. Fördistorsion kan användas till olika typer av signaler samtidigt, till exempel D-AMPS, IS-95 CDMA och WCDMA.

Tyvärr förbättrar inte fördistorsion verkningsgraden nämnvärt. Däremot används fördistorsion tillsammans med framåtkoppling. I den kombinationen kan fördistorsion ge en betydlig förbättring av verkningsgraden.

Framåtkoppling

Framåtkoppling ger mycket god linjäritet över stor bandbredd och dynamik, så att förstärkaren kan användas till flera samtidiga signaler. Nackdelarna är att det blir en ganska komplicerad krets, samt att komponenterna efter förstärkaren minskar uteffekten några dB. Framåtkoppling används till GSM-EDGE och WCDMA. Eftersom det är en stor kretskoppling blir det ganska låg verkningsgrad. Den är därför mest lämplig till basstationer.

Kartesisk återkoppling

Kartesisk återkoppling är en ganska enkel krets som ger mycket bra linjäritet, men är tyvärr ganska smalbandig. Insignalen kan med fördel vara en basbandsignal. Därigenom blir hela sändaren linjäriserad. Eftersom den är återkopplad kan den automatiskt kompensera för drift. Kartesisk återkoppling har använts i handapparaterna till AMPS.

LINC

LINC består av en enkel RF-krets, som använder två Klass C förstärkare för att tillsammans ge linjär förstärkning. Den passar smalbandiga sändare som har måttliga krav på distorsion. LINC är en intressant teknik under utveckling.

Referenser

Linjärisering

A D Vare, "Selection of transistor technology for linearised power amplifiers", ECWT 2002 pp 143-146

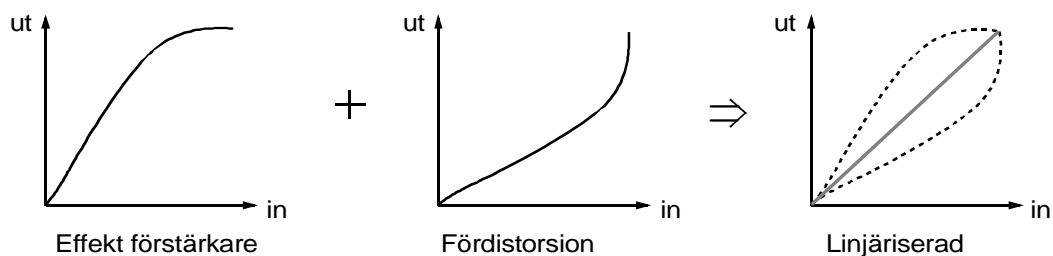
Allen Katz, "Linearizing high power amplifiers", EuMC 2001 WS12

Frederick H Raab, "Power amplifiers and transmitters for RF and microwave", IEEE Transactions on microwave theory and techniques, vol 50 no 3 mar 2002 pp 814-825

Peter B Kennington, "Multicarrier transceivers for software radio base stations", ECWT 2000 pp 324-327

16. Fördistorsion

En effektförstärkares olinjäritet kompenseras före slutsteget. På så sätt kan man använda komponenter på låg effekt, och kretsarna ger inte förluster på den dyrbara uteffekten.

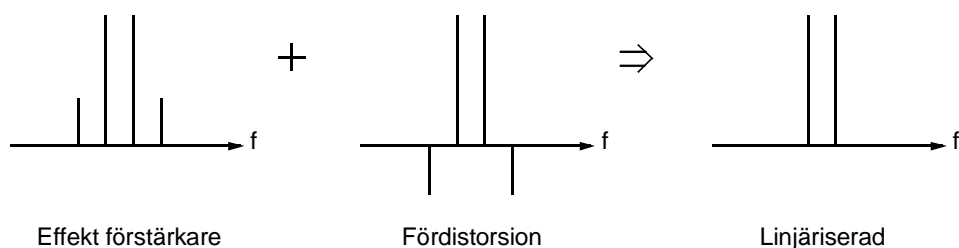


Förstärkaren uppnår en viss maxeffekt. När förstärkaren mäts får utsignalen fel amplitud, dvs. distorsion. En krets med fördistorsion ger en amplitud-karakteristik som går åt andra hållet. Tillsammans ger det en linjär karakteristik.

En TWT ger hög effekt och hög verkningsgrad på mikrovåg. Då effekten är flera hundra watt, blir den mindre och billigare än en motsvarande förstärkare med transistorer. Därför används den inom satellitkommunikation. Nackdelen med en TWT-förstärkare är att avrundningen sker inom ett stort område, den har alltså en mjuk limitering. För att få en någorlunda linjär förstärkare behöver uteffekten minskas (back-off) med 6 - 10 dB. Det ger IM-produkter på 22 - 30 dBc. Fördistorsion förbättrar störmarginalen med 5 - 13 dB. Alternativt kan man öka uteffekten med 6 dB om man använder fördistorsion.

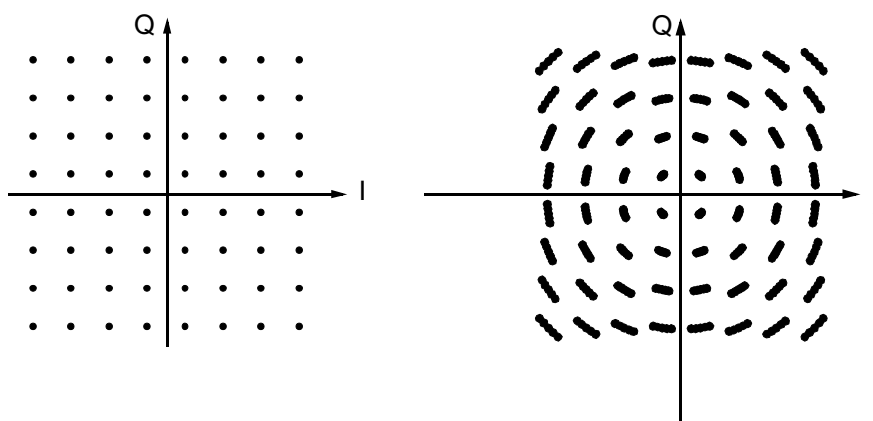
Basstationer till mobiltelefoni använder LD-MOS i slutstegen. De har också mjuk limitering och kan alltså förbättras med fördistorsion.

En nackdel med fördistorsion är att kretsen är optimerad för en viss signalnivå och en viss temperatur. Men med styrkretsar kan fördistorsionen justeras så att de störande blandprodukterna blir undertryckta även då omgivningen varierar. Med adaptiv återkoppling kan kretsen automatiskt justeras till bästa prestanda, även då effektförstärkaren åldras.

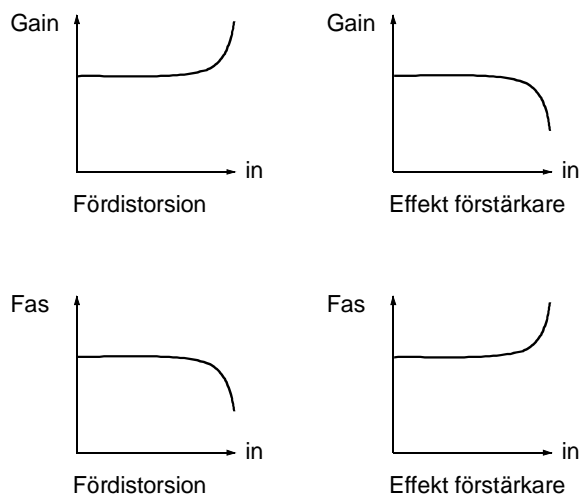


Alla olinjäriteter ger upphov till övertoner och IM-produkter. Det största problemet är 3:e gradens blandprodukter eftersom de kan hamna i egna bandet som distorsion eller i grannkanalen som störning. Kretsen med fördistorsion ger blandprodukter som är lika stora som effektförstärkarens men ligger i motfas. Störningarna blir alltså undertryckta genom utfasning.

De enklaste och vanligaste kretsarna undertrycker 3:e gradens IMD (Inter Modulations Distorsion) så att de blir lägre än 5:e gradens IMD. Det lönar sig inte att undertrycka mer än så mycket. Men det finns lite mer komplicerade kretsar som också ger en viss undertryckning av 5:e gradens IMD.



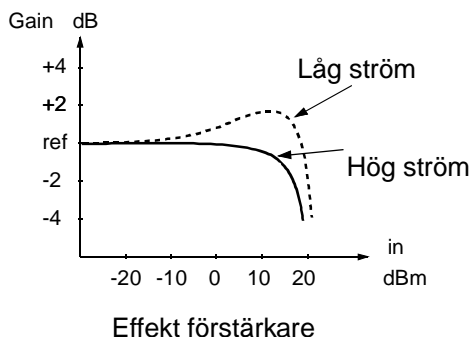
Det är inte bara amplituden som förvrängs då insignalen ökar (AM-AM), även fasen vrids av förstärkaren då den går in i limitering (AM-PM). Vid digital kommunikation (t.ex. QPSK, MSK eller QAM) ger fasvriddningen feltolkning med bitfel som följd. En TWT ger en fasvriddning på ca 40° . Kretsen med fördistorsion ger en motsvarande fasvriddning åt andra hållet. I praktiken minskas fasfelet från 40° till 5° .



En förstärkare som används nära sin 1 dB kompressionspunkt (exempelvis en handapparat) behöver en fördistorsion med expanderande karakteristik. 1 dB punkten kan då flyttas till lite högre uteffekt. Det kan utnyttjas till högre uteffekt eller lägre störning i grannkanalerna eller till högre verkningsgrad.

Förstärkaren i en basstation behöver ha mycket låg distorsion. Ofta arbetar förstärkaren långt från nivån för mättnad (stor back-off). I det fallet används en fördistorsion med kompression för att minska den totala distorsionen.

En FET-förstärkare ger ofta en fasvridning som ökar då ineffekten ökar. Fördistorsionen behöver då ha en negativ fas, dvs. en fas som minskar då ineffekten ökar. En bipolär förstärkare ger istället negativ fasvridning. Dess fördistorsion har då positiv fasvridning.

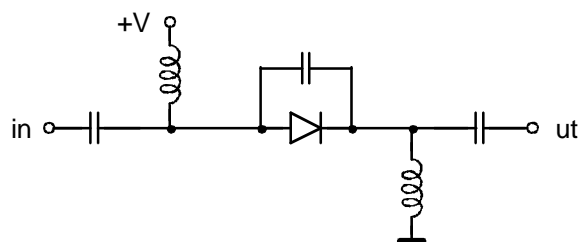


Förstärkare som arbetar på 2,5 % av maxströmmen eller mer får vid ökande signalstyrka en gradvis minskning av förstärkningen. Men om transistorn arbetar på låga strömmar, mindre än 1 % av maxströmmen, får förstärkaren först en ökning av förstärkningen innan den minskar vid mätning.

En förstärkare i en batteridrivna apparat (t.ex. mobiltelefon) kan använda dynamisk förspänning. Genom att minska strömmen då uteffekten ska vara låg, kommer den totala verkningsgraden att bli högre. I det fallet behöver fördistorsionen justeras för de olika olinjäriteterna vid de olika inställningarna av effektförstärkaren.

Fördistorsion med diod

Diod i serie

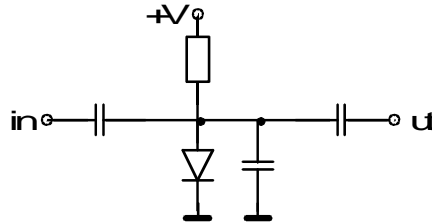


När insignalen ökar så ökar medelströmmen i dioden. Med större ström i framriktningen blir diodens impedans lägre. Det ger då lägre dämpning av signalen till utgången. Större dämpning för små signaler och mindre dämpning för stora signaler resulterar i en expanderande karakteristik.

Kondensatorn som shuntar dioden kommer att ge en fasvridning som varierar beroende på diodens impedans. Större insignal ger lägre diodimpedans, så att inverkan av kondensatorns fasvridning blir mindre. Med lämplig kapacitans kan slutstegets AM-PM kompenseras. Kretsen ger alltså positiv amplitud och negativ fas, då insignalen ökar.

Med denna enkla krets har störningarna i grannkanalen minskats cirka 5 dB. Modulationen var i det fallet $\pi/4$ QPSK på 1,9 GHz.

Diod i shunt

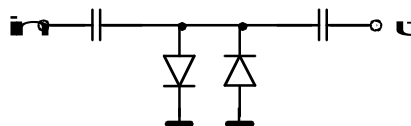


En shuntande diod ger normalt kompression av signalen eftersom dioden klipper signalen. Men med ett lämpligt motstånd i förspänningskretsen kan man få expansion istället.

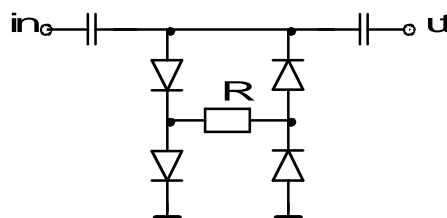
När insignalen ökar kommer den detekterade likströmmen att öka. Likströmmen ger ett större spänningsfall över motståndet. Det resulterar i motsvarande lägre spänning över dioden. Lägre förspänning ger högre impedans hos dioden, det vill säga mindre dämpning av signalen.

Signalens expansion och fasing kan justeras med spänningen $+V$.

Motställda dioder

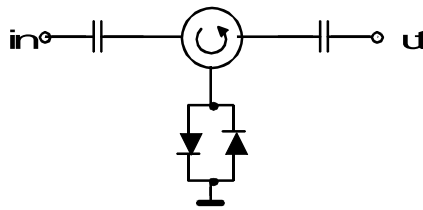


Med två motställda dioder får man en symmetrisk krets som undertrycker jämna övertoner och blandprodukter. Kretsen har alltså effektivt alstrat 3:e gradens blandprodukter, som sen ska kompensera slutstegets intermodulation. Eftersom kretsen är symmetrisk blir IM_2 undertryckt.

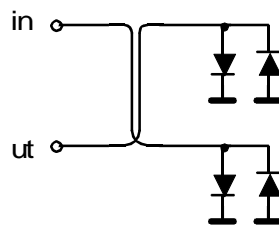


Det blir bättre balans, dvs. undertryckning av IM_2 , med ytterligare två dioder (eller två motstånd) samt ett överbyggande motstånd R. Förbättringen av IM_2 blir 10 - 25 dB.

Tyvärr är amplituden komprimerad istället för expanderad.

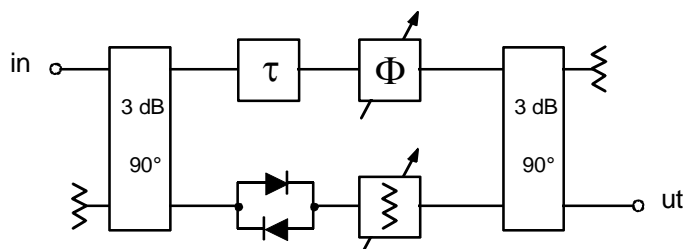


Om transmissionen är komprimerad så blir istället reflektionen expanderad. En cirkulator kan användas för att separera in- och utgång på en reflektionskomponent.



Istället för en cirkulator kan man använda en 3 dB hybrid. Men då behövs det två reflekterande diodkretsar istället.

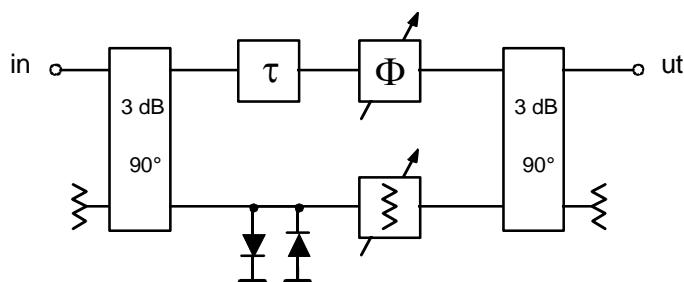
Bättre kontroll av distorsionen



Signalen delas upp i två olika vägar, den ena är expanderande och den andra är linjär. På utgången summeras signalen och blandprodukterna. Med en variabel dämpsats kan förhållandet mellan signal och störning lätt justeras.

Med den variabla fasskiftaren får vektoradditionen önskat fasläge. När insignalen varierar i amplitud kommer den sammanlagda utsignalen att vridas i fas. Fasskiftaren ställs in så att fasvridningen lagom kompenserar slutstegets AM-PM distorsion.

Dämpare och fasskiftare placeras i varsin signalväg. Eftersom de ger ungefär lika mycket fördröjning kan den extra fördröjningen τ vara mycket liten, eller kanske helt elimineras. Om dämpare och fasskiftare hade placerats direkt efter varandra, så skulle de ha påverkat varandra. Med komponenterna i varsin länk blir styrningen enklare.

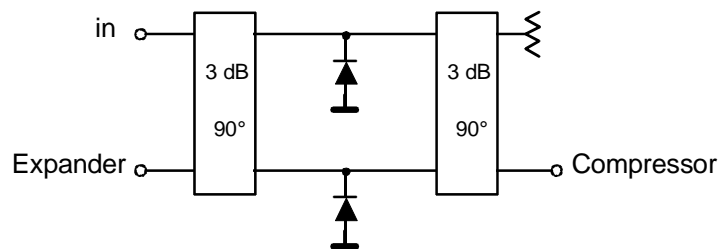


Shuntande dioder ger kompression, därför behövs en annan summering.

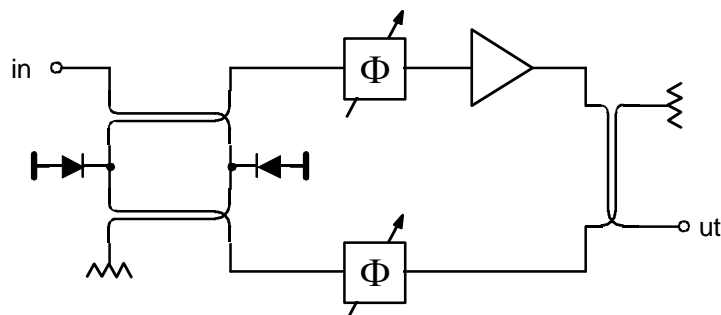
Vid små signaler adderas signalvägarna i motfas så att utsignalen blir liten. Då insignalen ökar blir signalerna inte lika stora längre. Det ger en stor signal ut. Slutresultatet är en expanderande krets.

Istället för två 90° hybrider kan man använda 180° på ena sidan och 0° på andra sidan. Det ger också addition i motfas.

Användning av både reflektion och transmission



Den signal som passerar shuntande dioder får komprimerande karaktär. Dioderna är ju limiterande. Den signal som inte passerar, reflekteras istället till den andra utgången. Den utsignalen blir alltså expanderande.



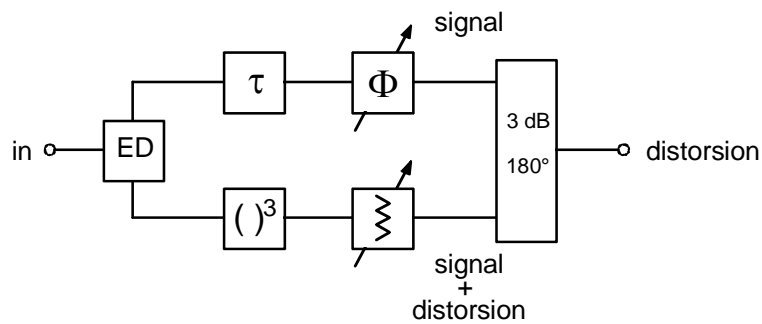
De två signalerna adderas sedan i en hybrid på utgången. Små signaler ger ingen reflektion vid dioderna och passerar den undre vägen till utgången. Stora signaler reflekteras till övre vägen mot utgången. Med en förstärkare i övre signalvägen blir det en övervägande expanderande karaktär.

Genom att addera signalerna i en 90° hybrid kan man få en stor fasvridning som en funktion av insignalens amplitud. Små signaler vrids 0° och stora signaler vrids 90° . Med fasskiftarna kan storleken på AM-PM ställas in. Önskas en fasvridning åt andra hållet, så kopplas signalen ut vid avslutaren istället.

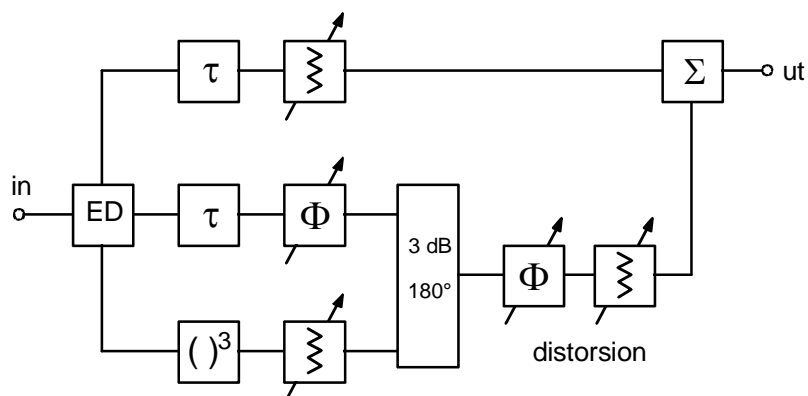
Hybriden på ingången ser två lika dioder. Reflektionerna är alltså undertryckta i ingångsporten. Anpassningen blir därför mycket bra över nästan en hel oktavs bandbredd.

Distorsjon med signalen undertryckt

Genom att separera distorsjonen från signalen blir det lättare att styra fördistorsjonen.



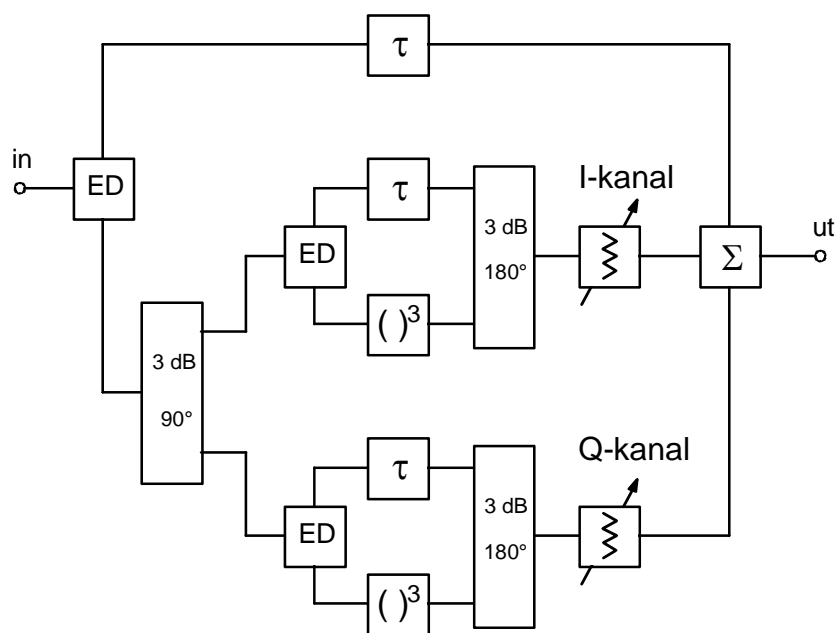
Signalen går en parallell väg och adderas i motfas i utgången. Signalen blir undertryckt så att det endast är distorsion i utgången.



Distorsjonen ställs in till lämplig amplitud och fas för att sedan adderas till signalen. När väl distorsionen har genererats, justeras den alltså för att passa aktuell slutförstärkare.

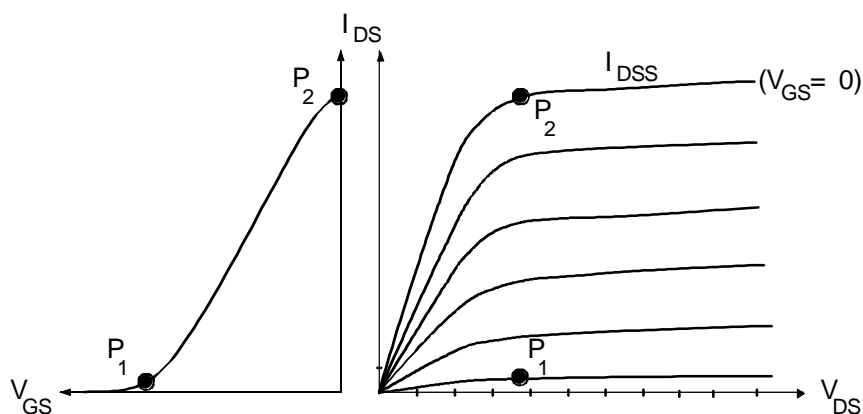
AM-PM reglering med IQ-krets

Diodkretsarna ger vanligtvis expanderings med konstant fas.



De två kretsarna som alstrar distorsion har matats med en 90° hybrid. Distorsionen från de två kretsarna skiljer sig alltså 90° i fas. Efter vektoraddition får man önskad fasvridning så att både AM-AM och AM-PM i efterföljande förstärkare kan kompenseras.

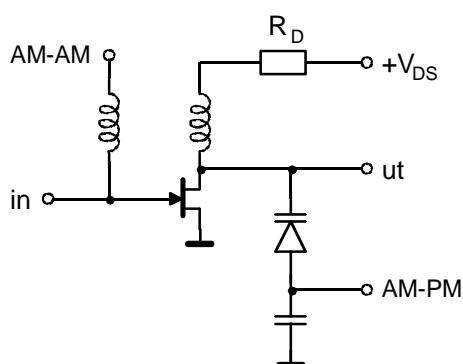
Fördistorsion med FET



En FET blir olinjär då den förspänns till strypning (pinch-off). Den har då en stor negativ spänning på gate, så att arbetspunkten ligger vid P_1 . Om insignalen ökar kommer också drainströmmen att öka. Arbetspunkten flyttas då in i ett område med högre förstärkning. Det ger alltså en expansion i amplitud. Vid punkten P_1 är transkonduktansen kraftigt olinjär. Det ger blandprodukter som kan utnyttjas till fördistorsion.

Vid punkt P_2 börjar strömmen bli mättad. Det ger en olinjäritet i konduktansen på utgången.

FET vid strypning och varaktor för AM-PM

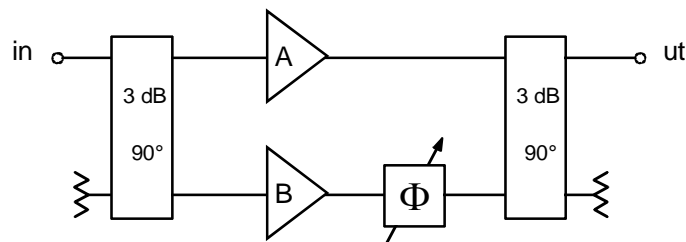


Spänningen på gate ställer in arbetspunkten till olinjäriteten vid strypning. När insignalen ökar kommer strömmen i transistorn att öka, och det ger högre förstärkning. Resultatet är en olinjäritet med ökande förstärkning. Genom att justera gatespänningen kan mängden AM-AM ställas in.

Även spänningen på drain kommer att variera och styra varaktordioden. Insignalens storlek kommer alltså att styra kapacitansen dvs fasvridningen. Det ger önskad AM-PM distorsion. Med en spänning på andra sidan varaktordioden kan dioden arbetspunkt ställas in, dvs mängden AM-PM.

Expansion vid låga spänningar

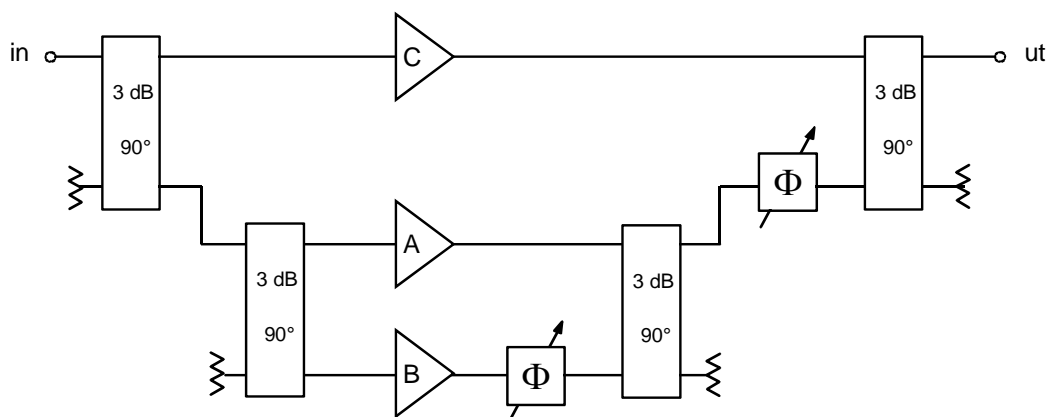
Kompressionen vid låga spänningar kan vändas till expansion med hjälp av en balanserad krets.



Transistorn A är förspänd till sitt linjära område. B är med låga spänningar i sitt olinjära område. Vid låga signaler balanserar de två förstärkarna ut varandra. Då insignalen är stor blir signalnivån i B mättad. Eftersom de förstärkta signalerna inte är lika stora balanserar de inte ut varandra och signalnivån i utgången blir stor. Resultatet är att utsignalen ökar mer än insignalen, dvs en expansion.

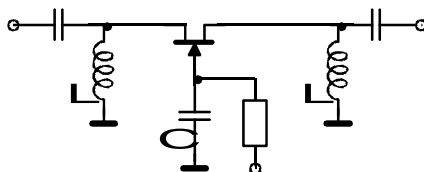
Storleken på intermodulationen bestäms av insignalens nivå och förspänningen på förstärkaren B. Fasen på IM-produkterna ställs in med fasskiftaren, som i sin enklaste form är en bit ledning.

Kretsen får ganska bra anpassning eftersom båda förstärkarna uppvisar samma impedans då gatespänningarna är lika.



Den undre delen med förstärkarna A och B har helt balanserat bort den linjära komponenten. Ut från den kretsen kommer endast IM-distorsionen. Dessa IM-produkter adderas, med lämpligt fasläge, till den linjära förstärkaren C.

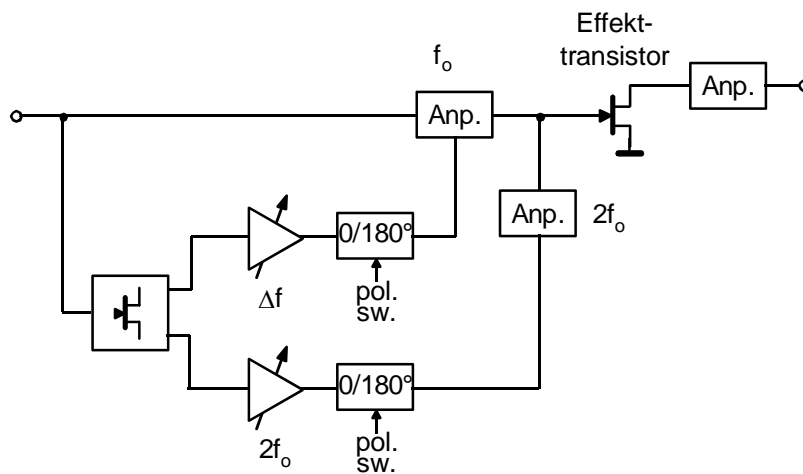
FET som variabel resistans



En HJ-FET används som variabel resistans. Kondensatorerna på in- och utgång används både som DC-block och till anpassning. Kretsens dämpning bestäms av transistorens resistans (r_{ds}). Resistansen bestäms dels av förspänningen på gate och dels av signalen. Om transistoren är förspänd nära strypning kommer kanalens resistans att minska då signalens effekt ökar. Det ger önskad expansion i amplitud.

Induktanserna används för att ge negativ fasvridning för högre insignal. Om L och C minskar blir avvikelser i fas större, utan att försämra expansionen i amplitud. Om L inte används blir fasavvikelsen istället positiv.

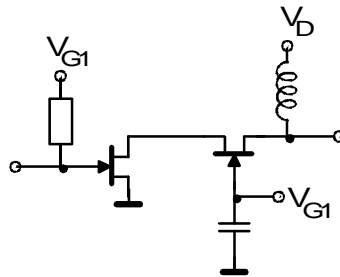
Linjärisering med både Δf och $2f_0$



En FET som är förspänd i närheten av strypning genererar fördistorsionen. Δf och $2f_0$ separeras med filter och förstärks. Därefter injiceras de på effekttransistorns ingång, där de blandas med insignalen f_0 . Med lämplig storlek på Δf och $2f_0$ blir förstärkarens distorsion kompenserad. En fördel med kretsen är att det inte behövs någon noggrann inställning av fasen. Det räcker med att välja rätt polaritet.

Kretsen har gett 27 dB undertryckning av IM_3 . Med injicering av enbart Δf blev undertryckningen bara 5,5 dB.

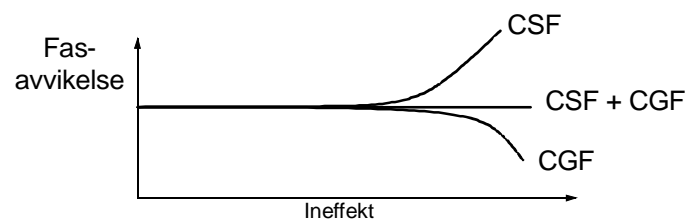
Cascode



Cascode är en kombination av gemensam Source FET (CSF) och gemensam Gate FET (CGF). Den kan alternativt bestå av två bipolära transistorer med gemensam emitter och gemensam bas.

Inimpedansen blir högre och alltså lättare att anpassa till 50Ω . Mellan transistorerna behövs inga anpassningskretsar. Det blir alltså en enkel och kompakt krets.

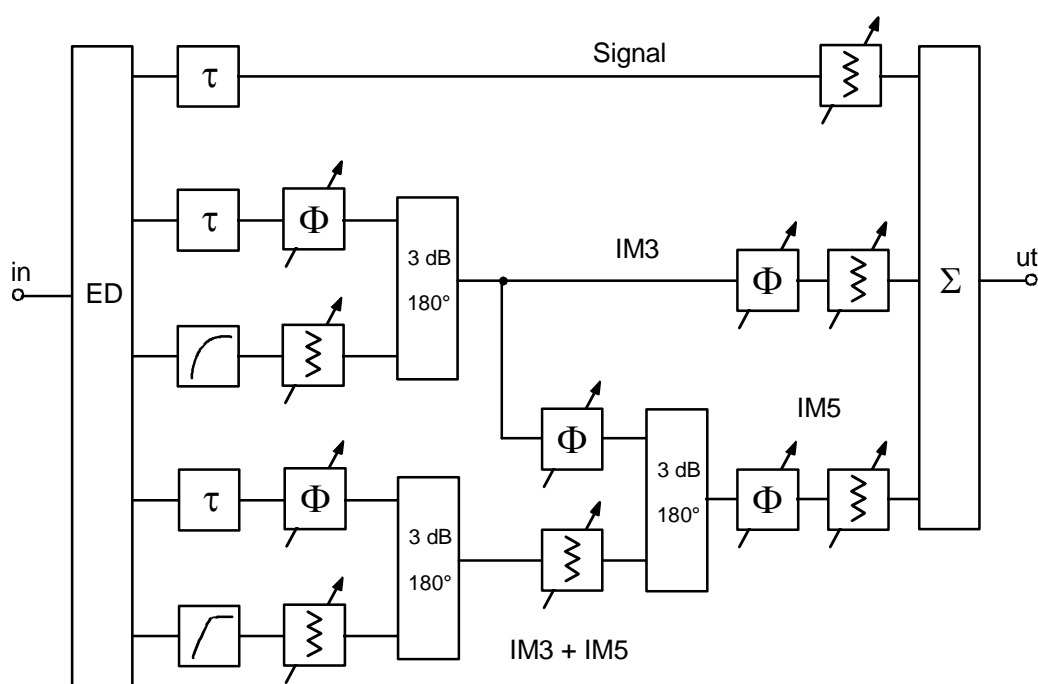
De två transistorerna har strömförsörjning i serie. Eftersom det är samma ström genom båda transistorerna, blir det lägre effektförbrukning än kaskadkoppling. Totala spänningen blir högre och strömmen lägre när de ligger i serie. Men vid riktigt låga batterispänningar är det inte lämpligt med cascode.



Vid höga effekter drivs förstärkaren till mättnad. En FET med gemensam Source får då en fas som ökar med insignalen. Gemensam Gate ger däremot en fasvridning åt andra hållet. Tillsammans kompenserar de då varandra, så att det blir en mycket liten fasavvikelse.

5:e gradens IMD

Fördistorsion med både 3:e och 5:e gradens IMD ger ännu bättre undertryckning av de icke önskade blandprodukterna. Följande krets kan enkelt justera amplitud och fas på både IM3 och IM5 oberoende av varandra.



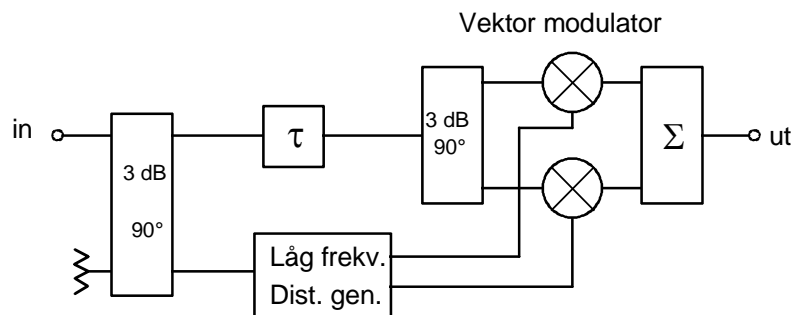
IM3 med undertryckt signal genereras som förut. Signalens undertryckning kan bli ca 30 - 35 dB. IM5 alstras med en ännu kraftigare olinjäritet. Signalen blir undertryckt, men det alstras också IM3 med större amplitud än IM5. Denna icke önskade IM3 balanseras bort ca 20 dB med hjälp av IM3 från kretsen ovanför.

Utsignalen innehåller alltså både IM3 och IM5 med väl kontrollerad amplitud och fas.

Separat generering av distorsion

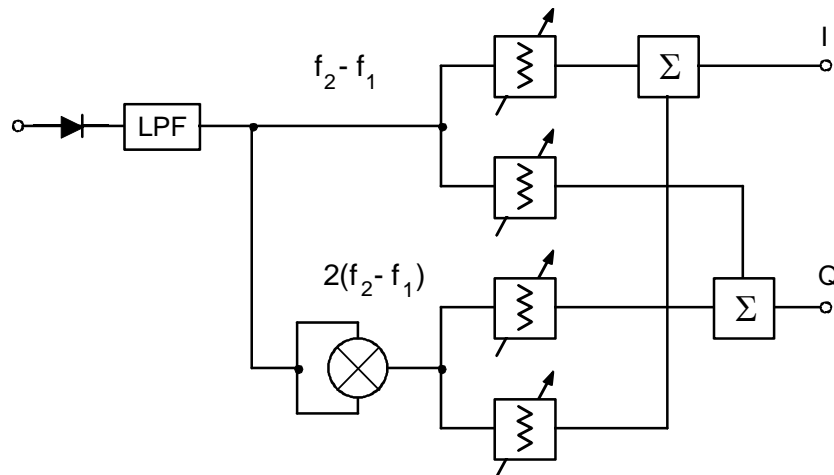
3:e gradens blandprodukter ($2f_2-f_1$, $2f_1-f_2$) kan alstras genom att blanda grundtonen (f_2 , f_1) med 2:a gradens blandprodukt (f_2-f_1).

5:e gradens blandprodukter ($3f_2-2f_1$, $3f_1-2f_2$) kan alstras genom att blanda grundtonen (f_2 , f_1) med 4:e gradens blandprodukt ($2f_2-2f_1$).



I en parallell kanal detekteras insignalens amplitud. Denna amplitud korrigerar fas och amplitud i en vektormodulator, så att AM-AM och AM-PM distorsionen i efterföljande effektförstärkare kompenseras.

Om det är flera signaler är det lättare att betrakta kretsen i frekvensdomänen. Distorsionens blandprodukter alstras på låga frekvenser. Därefter moduleras blandprodukterna på signalen i en vektormodulator (IQ-modulator).



Distorsionsgeneratoren har en diod på ingången, där insignalerna blandas med varandra. Alternativt kan man använda en FET-förstärkare med lagom olinjäritet. Skillnadsfrekvensen $f_2 - f_1$ filtreras fram. Det är den sökta 2:a gradens blandprodukt. 4:e gradens blandprodukt alstras genom att multiplicera 2:a gradens blandprodukt med sig själv.

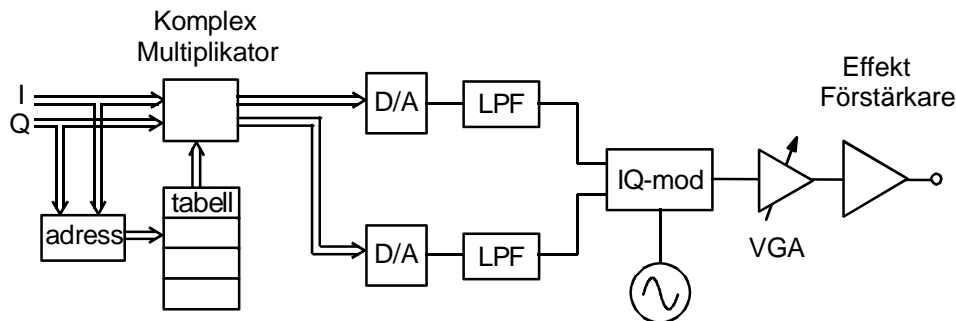
2:a och 4:e gradens blandprodukter adderas till I- respektive Q-kanalen. I efterföljande IQ-modulator alstras de önskade 3:e och 5:e gradens IMD efter blandning med insignalen.

Med de variabla dämpsatserna (eller variabla förstärkare) justeras nivåerna på blandprodukterna i I- och Q-kanalerna så att de lagom kompenserar distorsionen i slutförstärkaren.

Istället för att skapa en linjär och en kvadratisk funktion (work function) med vägda amplituder, kan de önskade nivåerna lagras i en digital tabell. Den detekterade insignalens amplitud AD-omvandlas och väljer lämpliga värden från tabellen.

Kretskopplingen blir enkel och kompakt eftersom det mesta görs på ganska låg frekvens. Med en slutförstärkare på 44 dBm har IM3 undertryckts 28 dB, och IM5 har undertryckts 17 dB.

Digital fördistorsion



Fördistorsion på RF eller MF är enkelt och billigt. Dessutom blir systemet linjäriserat över en stor bandbredd. Nackdelen är att linjäriteten blir begränsad till korrigering av 3:e gradens IMD, eller möjligen 3:e och 5:e gradens IMD.

Ett annat alternativ är att åstadkomma fördistorsion redan i basbandet. Till det behövs en PROM-tabell och en DSP. Men ett modernt kommunikationssystem använder ändå någon form av signalprocessning på basbandet. Koefficienterna till fördistorsionen lagras i en 2D-tabell till I- respektive Q-kanalen. Alternativt kan fördistorsionen lagras i en en-dimensionell tabell med komplexa tal, i polär eller kartesisk form. En fördel är att det går att korrigera för sådana olinjäriteter som de analoga kretsarna inte klarar av. En MCPA för W-CDMA kan kombinera digital fördistorsion med digital reducering av crestfaktorn. Det går då att få en verkningsgrad på 20 %.

Nackdelen är att det blir stor effektförbrukning och liten bandbredd. Bandbredden som DSP och efterföljande uppblandning behöver kunna hantera är dels signalens bandbredd plus de grannkanalerna där distorsionen behöver minskas. Det blir 5 – 7 gånger större bandbredd än basbandet.

Ett annat alternativ är att den digitalt alstrade fördistorsionen injiceras på effektförstärkaren via förspänningskretsarna. Det är inte lika effektivt som att korrigera den digitala signalen, men det behövs ingen extra bandbredd i DSP och uppblandning.

Referenser

Fördistorsion

Kazuhisa Yamauchi, "A novel series diode linearizer for mobile radio power amplifiers", IEEE MTT-S 1996 pp 831-834

Kazuhisa Yamauchi, "An 18 GHz band MMIC linearizer using parallel diode with a bias feed resistance and parallel capacitor", IEEE MTT-S 2000 pp 1507-1510

Wei Huang, "Residual second order intermodulation suppression in third order distortion generator", IEEE MTT-S 1998 pp 737-740

Gary Hau, "A highly efficient linearized wideband CDMA handset power amplifier based on predistortion under various bias conditions", IEEE Transactions on microwave theory and techniques, vol 49 no 6 jun 2001 pp 1194-1200

Youngoo Yang, "New predistortion linearizer using low frequency even order intermodulation components", IEEE Transactions on microwave theory and techniques, vol 50 no 2 feb 2002 pp 446-452

Jaehyok Yi, "Analog predistortion linearizer for high power RF amplifiers", IEEE Transactions on microwave theory and techniques, vol 48 no 12 dec 2000 pp 2709-2713

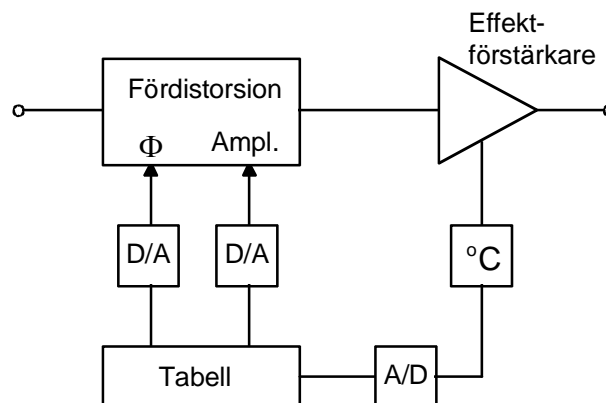
Sang Won Kim, "A novel predistorter using a balanced type IM3 generator", IEEE MTT-S 2003 pp 339-342

N Ceylan, "Mobile phone power amplifier linearity and efficiency enhancement using digital predistortion", EuMC 2003 pp 269-272

17. Adaptiv fördistorsion

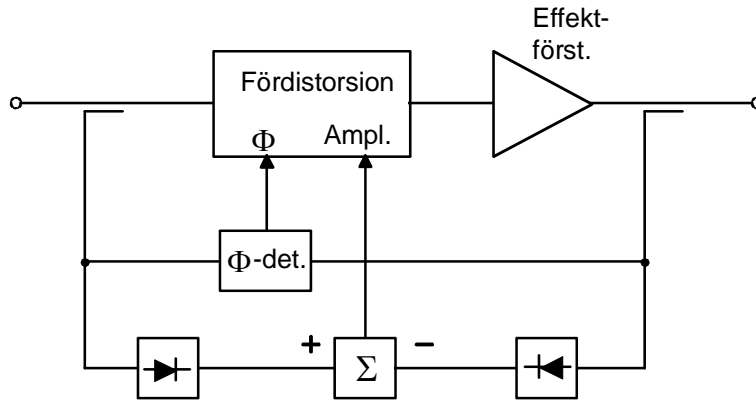
Linjärisering med fördistorsion kan lätt ge 10 dB undertryckning av IMD. Bättre undertryckning kräver noggrannare kompensering i fas och amplitud. Men vad som är rätt inställning varierar beroende på bl.a. temperatur och signalnivå. Om det behövs mer än 50 dB signal/störförhållande ($C/I > 50$ dB), blir inställningarna mycket kritiska. En förändring i fas på mindre än en grad, kan göra att förstärkaren inte håller specifikationen. För att kunna behålla IM-produkterna i motfas behövs en reglerkrets som styrs av omgivningen eller distorsionen. I de flesta fall är det ganska långsamma förändringar som behöver justeras adaptivt, därför kan reglerkretsarna vara långsamma.

Kompensering av temperaturdrift



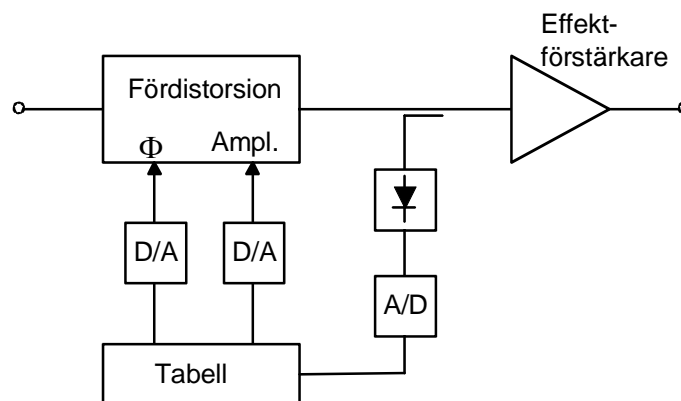
I en PROM-tabell lagras de korrigeringar som behövs för att få låg distorsion vid olika temperaturer. Effektförstärkarens temperatur mäts och beroende på temperaturen väljs en lämplig inställning av amplitud och fas i fördistorsionen.

Styrning med skillnaden



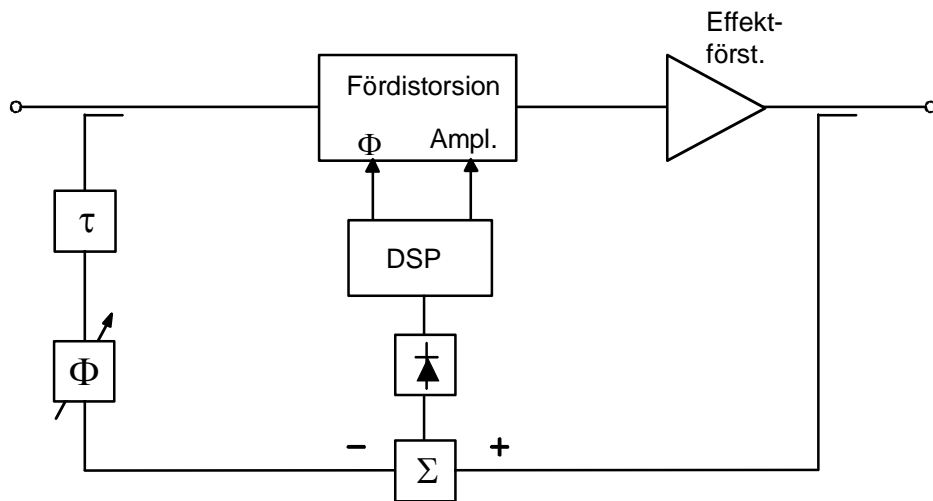
Förstärkarens utsignal jämförs med insignalen. Skillnaden i amplitud och fas används för att ställa in fördistorsionen. Med en snabb reglering kan signalens envelop korrigeras, men fördröjningen i kretsen begränsar då bandbredden till några MHz. Istället används en långsam reglering av fördistorsionen. I den här kretsen detekteras både distorsionens fas och amplitud. Men det räcker med att mäta amplituden för att minimera distorsionens storlek.

Kompensering av nivåskift



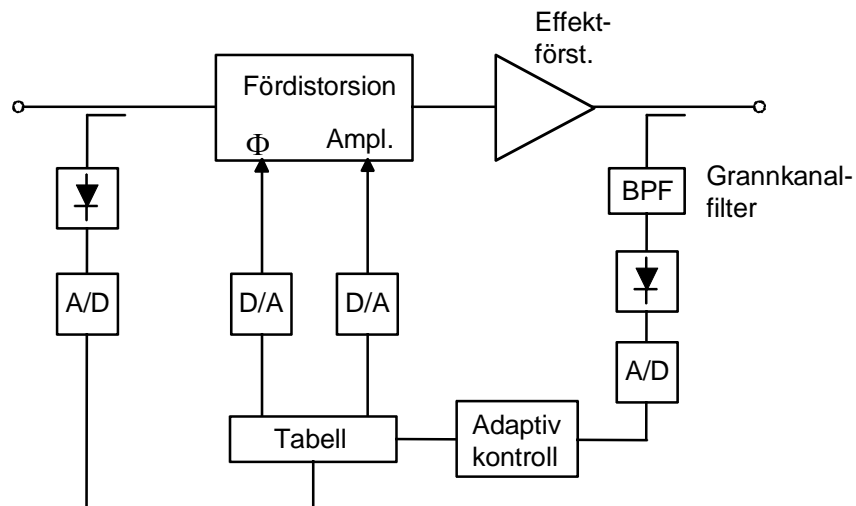
Effektförstärkaren behöver olika fördistorsion för olika signalstyrkor. De rätta inställningarna lagras i ett digitalt minne. Endast signalstyrkan på linjäriseringskretsens utgång behöver mätas. En digital styrkrets (mikrodator) ställer in linjäriseringens amplitud och fas, vid aktuell signalnivå.

Detektering av distorsionen



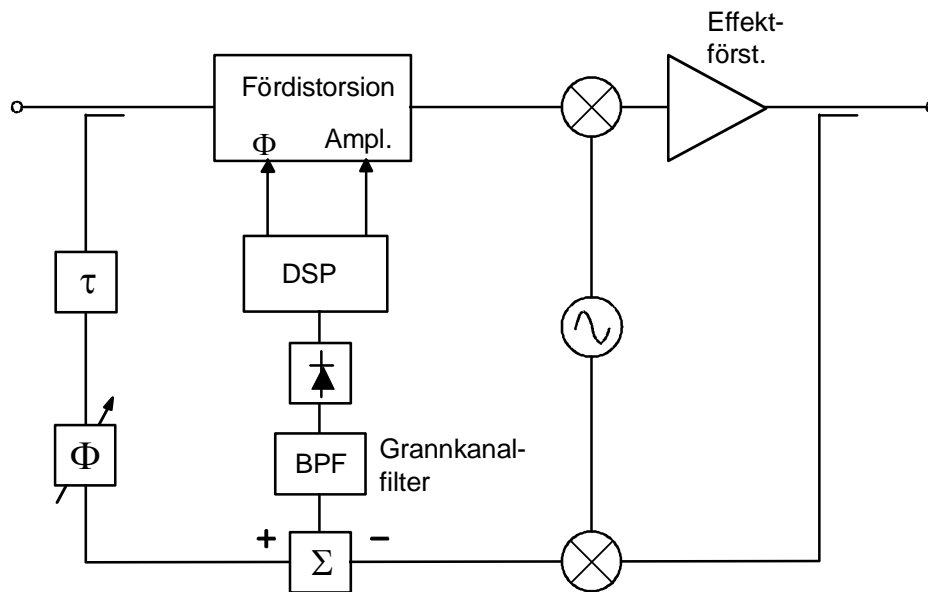
Skillnaden mellan utsignal och insignal är själva distorsionen. Eftersom själva signalen undertrycks ca 30 dB med motfaskopplingen, blir det ett ganska litet dynamikområde som detektorn behöver klara av. En digital signalprocessor mäter den detekterade distorsionen och ställer in lämplig fördistorsion.

Detektering av störning i grannkanal



I vissa situationer är det inte distorsionen i egna kanalen som är det största problemet. Det kan istället vara de IM-produkter som hamnar i grannkanalerna som behöver undertryckas. På utgången detekteras den totala effekten från IM-produkterna som hamnar i grannkanalerna, eller del av grannkanal. Fördistorsionen styrs till att ge så lite grannkanalstörning som möjligt efter effektförstärkaren. På så sätt kan störningarna i grannkanalerna undertryckas 10 - 20 dB.

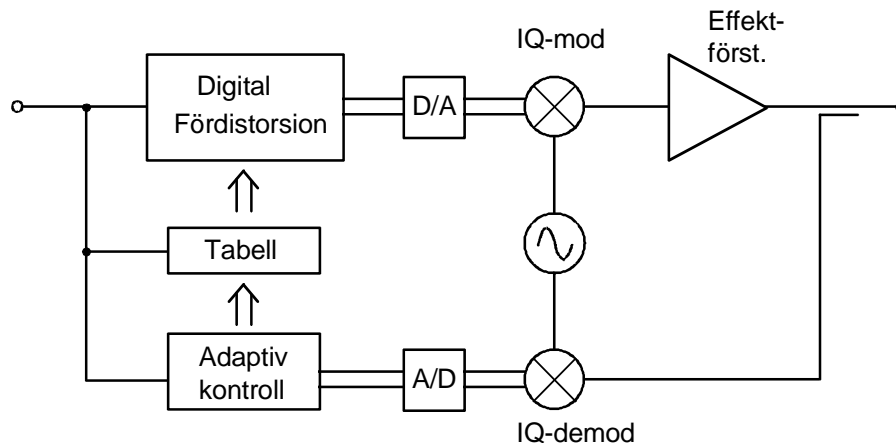
Grannkanalfilter med signalen utfasad



Fördistorsionen sker här på lägre MF-frekvens, därefter blandas signalen upp till RF-förstärkaren. Skillnaden mellan den nerblandade utsignalen och insignalen är alla IM-produkter. De störningar som ligger i grannkanalerna detekteras och A/D-omvandlas till en DSP. Signalprocessorn söker efter den inställning som ger minsta möjliga störningar i grannkanalen.

Genom att insignalen fhasas bort från utsignalen blir det bara distorsionen kvar. Det behövs då inte så branta flanker på grannkanalfiltret. Men den utfasningen kan också behöva en automatisk reglering för att inte försämrars vid åldring.

Digital fördistorsion



Med digital fördistorsion menas att alstringen av fördistorsionen sker med digitala kretsar på basbandet. Därefter moduleras signalen på önskad RF-frekvens. Utsignalen måste sedan demoduleras till basbandet igen. Hela digitala kretsen kan bestå av en DSP (digital signal processor).

Digital fördistorsion är speciellt lämpligt då insignalen är digital. De digitala kretsarna kan integreras med systemets övriga signalprocessning. I digital kommunikation sker kodning, filtrering av basbandssignalen och korrigering av modulatorfel med digitala kretsar på basbandet. Det är alltså inte en linjäriserad RF-förstärkare, utan får mer betraktas som en modul i systemet.

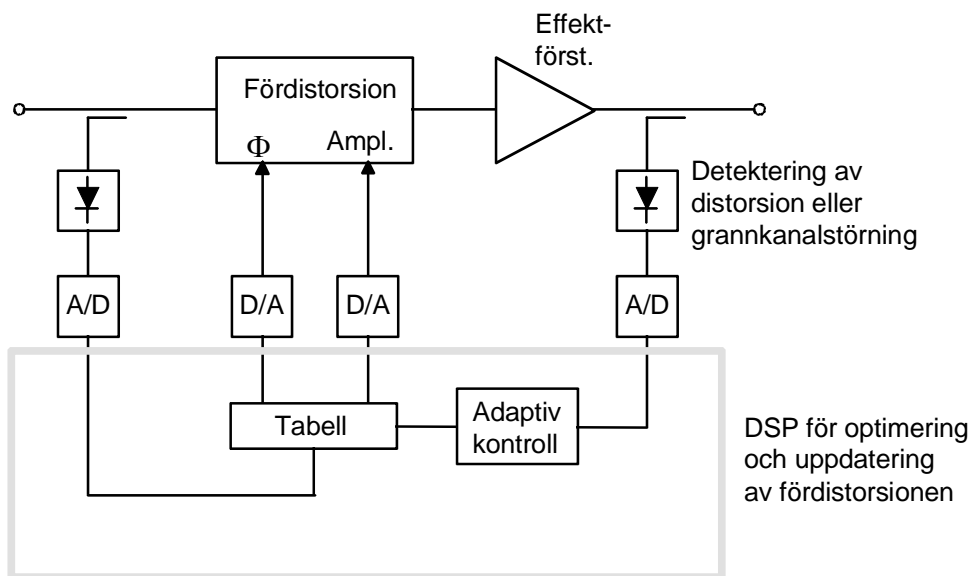
Inom digital kommunikation används ofta IQ-uppdelning vid modulationen. Det kan därför vara lämpligt att även göra fördistorsionen i kartesisk form (I, Q) istället för polär form (fas, amplitud).

LO-läckaget genom IQ-modulatorn sätter en gräns för hur mycket IM-produkterna kan minskas. Läckaget kan undertryckas med en parallellkoppling av ytterligare ett läckage, som justeras i motfas med hjälp av en fasskiftare och dämpsats. Det kan minska läckaget från 20 dBc till 40 dBc.

En nackdel är att digital fördistorsion kräver mycket processorkraft, dvs strömförbrukning. En standard DSP har en ordlängd på 16 bitar eller 32 bitar. Med en ASIC (application specific IC) kan ordlängden flexibelt väljas, för att få lägre effektförbrukningen eller större bandbredd.

Kretsarna för digital fördistorsion blir helt oberoende av frekvensen på RF-signalen. Däremot blir bandbredden mycket begränsad. Fördistorsion på RF ger den största bandbredden. Nerblandning till MF begränsas av bandbredden på MF-kretsarna. Vid digital fördistorsion på basbandet begränsas bandbredden av samplingshastigheten. Beroende på kretslösning behöver referenssignalen samplas 16 - 32 gånger per symbol. Det ger ett ganska smalbandigt system (30 kHz).

Digital inställning av fördistorsion



Insignalens amplitud väljer lämpligt värde från en tabell. Det värde som anses lämpligt blir kontinuerligt uppdaterat av processorn. Processorn justerar tabellen så att störsignalen minimeras. Tabellen kommer alltså att följa förändringarna i temperatur och frekvens samt effektförstärkarens arbetspunkt och åldring.

En diod som detekterar effekten ger en enkel krets men den konvergerar långsamt och blir känslig för brus. Med konvertering till en digital mottagare, både före och efter effektförstärkaren, kan distorsionen snabbt mätas upp och korrigeras.

Tabellen ska korrigera både amplitud och fas på signalen (AM/AM och AM/PM distorsion). Det blir alltså två tabeller (eller en 2-dimensionell tabell). Men det räcker med en 1-dimensionell tabell med komplexa tal, som styrs av distorsionens amplitud. Det är ju bara amplitudvariationer som ger distorsion.

En tabell, med förstärkningen som komplexa tal, kan kompensera valfri slutförstärkare. Tabellen innehåller 32 eller 64 komplexa tal, som kontinuerligt ska uppdateras. Genom att använda interpolation kan tabellen göras mindre. Istället åtgår mer processorkraft till själva interpolationen. Ett annat sätt är att lagra olinjäriteten som två ekvationer (polynom). Det ger en enklare krets, men kan inte anpassas till en valfri olinjäritet lika enkelt som med tabell. Fördelen med polynom är att de är enklare att lagra och uppdatera, eftersom de innehåller färre variabler. Det går att uppnå bra prestanda med 5:e gradens polynom.

Referenser

Adaptiv fördistorsion

Allen Katz, "Input adaptive linearizer system", IEEE MTT-S 2000 pp 1499-1502

Youngoo Yang, "Adaptive RF Cartesian predistorter on the low frequency even order IM terms", IEEE MTT-S 2001 pp 793-796

E G Jeckeln, "Adaptive baseband/RF predistorter for power amplifiers through instantaneous AM-AM and AM-PM characterization using digital receivers", IEEE MTT-S 2000 pp 489-492

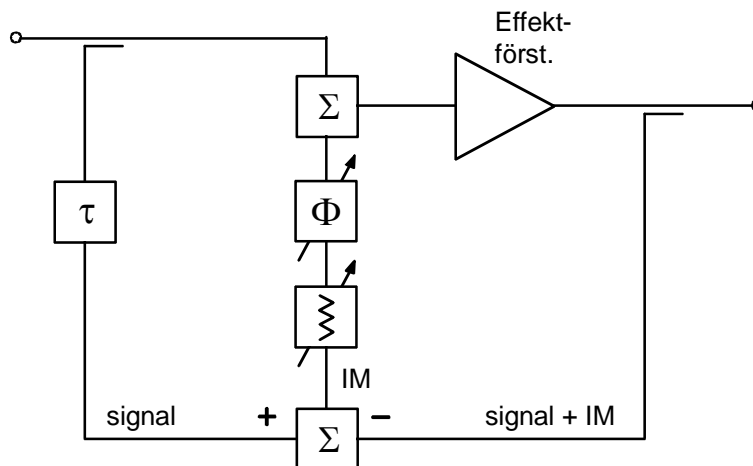
Michael Faulkner, "Adaptive linearization – experimental results", IEEE transactions on vehicular technology vol 43 no 2 may 1994 pp 323-332

Sante Andreoli, "Digital linearizer for RF amplifiers", IEEE transactions on broadcasting vol 43 no 1 mar 1997

18. Återkoppling

På lägre frekvenser används återkoppling av signalen, för att få högre linjäritet. Nackdelen med att återkoppla signalen är att förstärkningen blir lägre. På mikrovåg används inte återkoppling på grund av att en transistor har för låg förstärkning. Med fler transistorsteg blir förstärkningen visserligen större, men slingan blir så stor att kretsen lätt blir instabil och självsvänger. Förstärkaren måste vara stabil även då temperaturen ändras och transistorn åldras. Monolitförstärkare på låg effekt kanske kan fungera, men förstärkare på hög effekt kräver andra varianter av återkoppling.

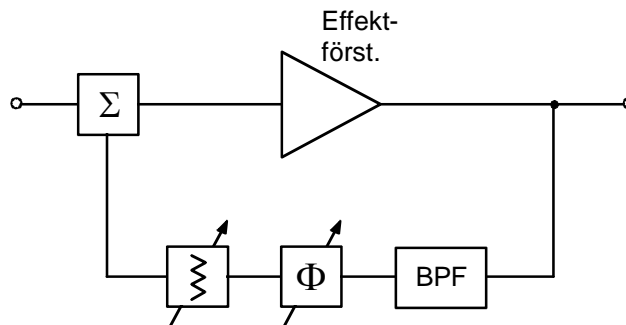
Återkoppling av intermodulation



Ut från förstärkaren kommer signal och IM-produkter. Signalen undertrycks (subtraheras) så att endast IM-produkterna blir återkopplade. Återkopplingen förbättrar alltså IMD utan att signalens förstärkning minskar.

För att signalen verkligen ska bli undertryckt i återkopplingen, behöver den justeras i amplitud och fas med en reglerkrets. Kretsen blir därför komplicerad, med endast måttlig undertryckning av IMD.

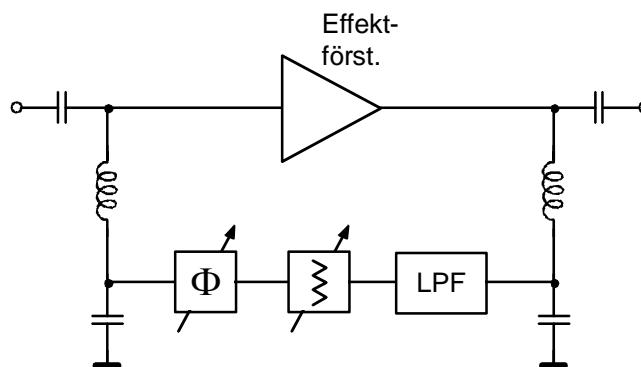
Återkoppling av övertoner



3:e gradens blandprodukter är av formen $2f_1 - f_2$. De övertoner som olinjäriteten också alstrar återkopplas till ingången. Där kommer övertonerna att blandas med insignalerna och alstra ytterligare 3:e gradens blandprodukter. Dessa extra blandprodukter justeras i fas och amplitud så att de ursprungliga blandprodukterna fasa bort.

I återkopplingen behövs alltså ett bandpass filter som släpper igenom övertonerna men inte nyttosignalerna. IM3 har på så sätt undertryckts 15 dB.

Återkoppling av skillnadsfrekvensen



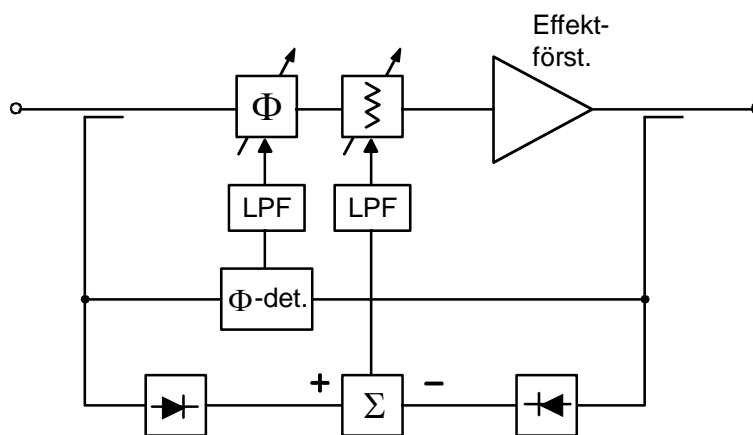
Skillnadsfrekvensen $f_2 - f_1$ är en 2:a gradens IM-produkt. Om denna skillnadsfrekvens återkopplas, kan den blandas med signalen med hjälp av förstärkarens olinjäritet. Resultatet blir $2f_2 - f_1$, dvs en tredje grads IMD. Med rätt injusterad fas och amplitud kan denna blandprodukt fasa bort förstärkarens ursprungliga IM_3 .

Återkopplingen sker alltså inte på mikrovåg, utan på lägre frekvenser. Men alla blandprodukter måste få rätt justering i amplitud och fas. Kretsen behöver alltså vara tillräckligt bredbandig för att klara alla skillnadsfrekvenser.

IM_3 har undertryckts 15 dB och IM_5 9 dB. Men förstärkarens distorsion varierar med temperatur och åldring. Återkopplingens fas och amplitud kan därför behöva en automatisk injustering.

Återkoppling av envelop

AM-AM och AM-PM distorsion innebär att utsignalens amplitud och fas varierar med insignalens amplitud. En krets som har konstant förstärkning och konstant fas innehåller ingen distorsion.

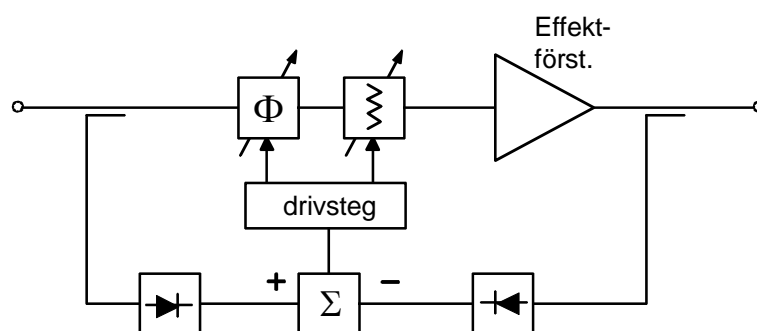


In- och utsignalens amplitud (envelop) detekteras. Skillnaden justerar en dämpningsfaktor. Om utsignalen inte ökar lika mycket som insignalen kommer dämpningsfaktorn att regleras till att ge högre utsignal. Resultatet är att utsignalen följer insignalens amplitud, förstärkningen är alltså konstant.

Fasdetektorn mäter fasskillnaden mellan ut- och ingång och justerar sedan fasskiftaren, så att fasen hålls konstant även då amplituden ändras.

Den största förbättringen av EVM kommer från justeringen av amplituden. Fasreglering behövs för att få EVM mindre än 1 %.

Återkopplingen ska reglera tillräckligt snabbt så att signalens envelop hinner justeras. Skillnadssignalens bandbredd står alltså i proportion till modulationens bandbredd på insignalen. Korrektionen i slingan behöver en bandbredd som är 5 gånger större än signalens bandbredd. En GSM-signal med 200 kHz bandbredd behöver alltså en slinga med 1 MHz bandbredd. Men den får inte vara större än nödvändigt eftersom återkoppling kan ge instabilitet och oscillering. Maximal bandbredd begränsas av fördröjningen i slingan. En monolitiskt tillverkad förstärkare kan alltså få stor bandbredd, medan en effektförstärkare uppbyggd med diskreta kretsar blir ganska smalbandig.



Det räcker att detektera amplitudens avvikelse och justera både fas och amplitud med samma reglerspänning. Både AM-AM och AM-PM kommer ju ursprungligen från samma amplitudvariation i insignalen.

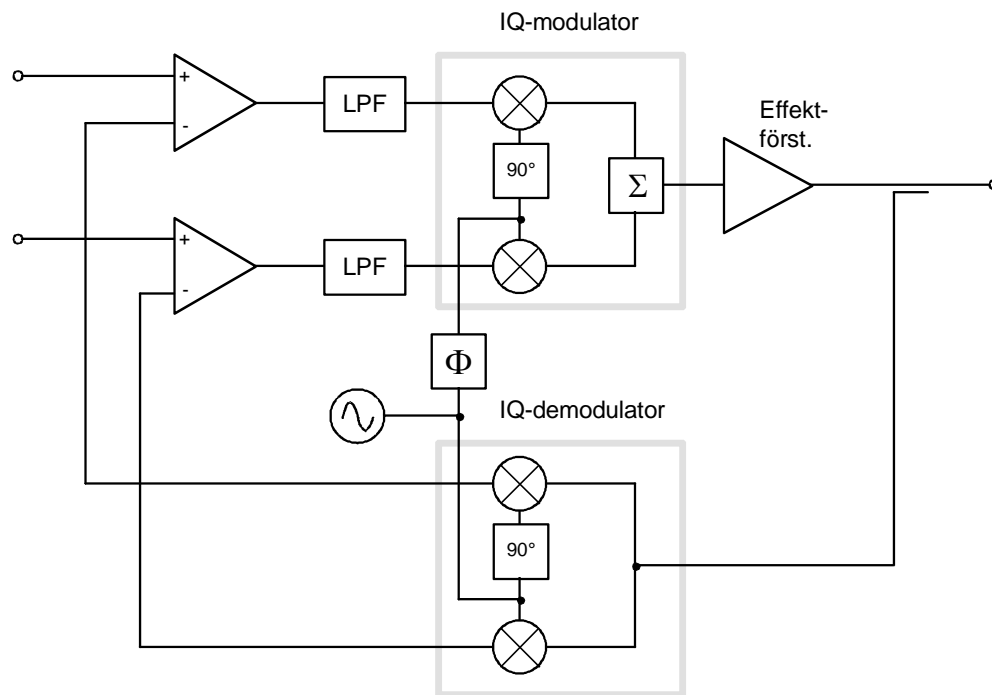
Linjärisering med återkopplad envelop kan minska IMD med 10 - 15 dB jämfört med en klass AB förstärkare. Förstärkningen hålls konstant och fasen ändras endast 4 grader, då insignalen varierar inom en dynamik på 10 dB.

Eftersom det är ett återkopplat system, så anpassar det sig efter förändringar i omgivningen. Linjäriseringen fungerar även om frekvensen eller temperaturen ändras, eller om transistorn åldras.

Maximal undertryckning bestäms av slingans förstärkning. Med 20 dB förstärkning i slingan kan intermodulationen, som faller inom slingans bandbredd, undertryckas med 20 dB. En förutsättning är förstås att slingan har mycket mindre distorsion än effektförstärkaren.

Kartesisk återkoppling

Återkoppling med styrning av fas och amplitud använder de polära koordinaterna. Kartesisk återkoppling använder de rätvinkliga kartesiska koordinaterna (I och Q). Vid QPSK-modulering kan återkopplingen med fördel ske ända till basbandet.

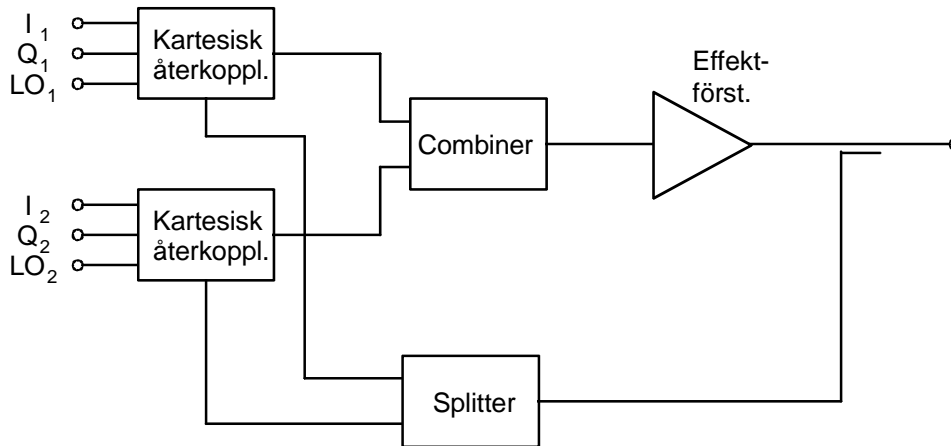


En IQ-demodulator blandar ner återkopplingen till basbandet. Det är på basbandet som insignalen jämförs med återkopplingen. Skillnaden, dvs. felspänningen, blandas upp med en IQ-modulator för att sedan förstärkas till hög effekt.

Modulator och demodulator använder samma oscillator. Fasskiftaren behöver justeras så att modulator och demodulator arbetar i fas. Om de inte arbetar i fas blir det överhörning mellan I- och Q-kanalen. Det försämrar stabiliteten i slingan och ger sämre undertryckning av IM-produkterna.

DC-offset från demodulator och differentiaförstärkare kommer att försämma balansen i modulatore. Det ger läckage av lokaloscillatorn, som hamnar mitt i det önskade bandet. En fördel med att först blanda ner till en mellanfrekvens, är att DC-offset från demodulatore är konstant eftersom frekvensen är konstant.

Bandbredden för linjäriseringen begränsas av fördröjningen i den återkopplade slingan. Bandbredden blir högst några MHz. Linjäriseringen behöver ha 5 - 10 gånger större bandbredd än själva kanalen för kommunikationen. Kanalen är vanligtvis i storleksordning 25 - 30 kHz.



Om systemet ska kunna hantera flera 30 kHz kanaler samtidigt, används flera kartesiska återkopplingar parallellt. Varje kartesisk modul innehåller både modulator och demodulator, samt differentialförstärkare och filter. Varje modul har sina ingångar för lokaloscillator och IQ-signaler. Kombineringen av RF-signalerna sker på låg effekt före förstärkaren.

Utöver effektförstärkaren blir också olinjäriteterna i modulator och drivförstärkare kompenserade. Hela sändaren är alltså linjäriserad. Däremot blir inte komponenterna i återkopplingen linjäriserade. Demodulatorens behov av extra låg distorsion för att inte dess IM-produkter ska överföras till effektförstärkaren.

Kartesisk återkoppling av en klass A eller klass AB förstärkare ger den bästa linjäriteten. Det ger undertryckning av störningarna i grannkanalen upp till 70 dB. En klass C förstärkare ger ett stort fasskift då transistorn börjar leda. Därför behövs en lägre slingförstärkning. Kartesisk återkoppling har använts till klass C förstärkare i handapparater till D-AMPS och PDC.

En fördel med återkoppling är att den linjäriserar förstärkaren trots att temperaturen varierar och transistorens förspänning ändras samt laständringar och komponenternas åldring.

En 900 MHz förstärkare för $\pi/4$ QPSK med 25 kHz bandbredd har fått en undertryckning av störningarna i grannkanalerna med 35 dB. Linjäriteten har i olika system förbättrats 25 - 45 dB vid olika bandbredder.

Referenser

Återkoppling

N Males-Ilic, "Amplifiers with improved IMD performance for multichannel wireless systems", EuMC 2002 pp 461-464

M R Moazzam, "A low third order intermodulation amplifier with harmonic feedback circuitry IEEE MTT-S 1996 pp 827-830

W J Jenkins, "Power amplifier linearization through low frequency feedback", ECWT 2000 pp 94-97

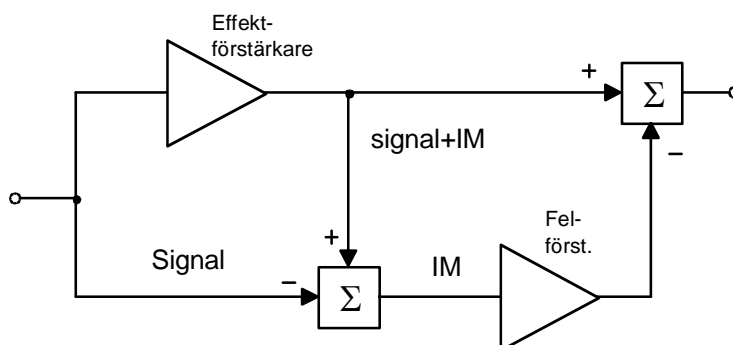
Hyun-Min Park, "A predistortion linearizer using envelope feedback technique with simplified carrier cancellation scheme for class A and class AB power amplifiers", Transactions on microwave theory and techniques, vol 48 no 6 june 2000 pp 898-903

Jean-Serge Cardinal, "A new adaptive double envelope feedback linearizer for solid state power amplifiers" Transactions on microwave theory and techniques, vol 43 no 7 july 1995 pp 1508-1515

Mats Johansson, "Linearization of RF power amplifiers using Cartesian feedback", Lund Institute of Technology, Applied Electronics, dec 1991

19. Framåtkoppling

Framåtkoppling (Feed Forward) används för att undertrycka IM-produkterna till 40 - 60 dB under nyttosignalen. Även 70 dB signal/stör förhållande är praktiskt användbart.



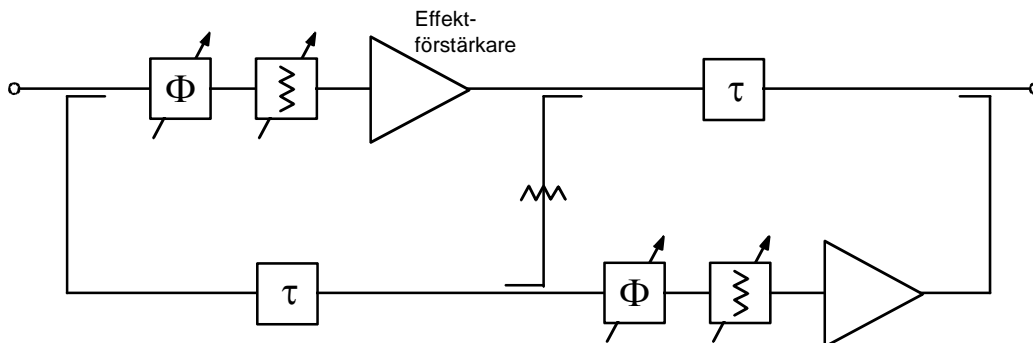
Kretsen för framåtkoppling består av två slingor. I den första slingan adderas den förstärkta signalen och insignalen i motfas. Om amplituderna ställs in lika så blir signalen helt undertryckt. IM-produkterna från effektförstärkaren blir däremot inte motfaskopplade eftersom de inte finns med i insignalen.

I den andra slingan adderas IM-produkterna till signalen från effektförstärkaren. Eftersom summeringen sker i motfas blir IM-produkterna undertryckta. Det enda som kommer ut är den önskade signalen på hög effekt.

Alla de störande signalerna som effektförstärkaren alstrar kommer att undertryckas i utgången, både distorsion och brus. Om förstärkaren består av både drivsteg och slutsteg blir den totala förstärkarkedjan linjäriserad. Även variationerna i strömförsörjningen, som moduleras på signalen, kommer att bli undertryckta.

Förstärkaren i andra slingan ska endast förstärka distorsionen från effektförstärkaren. Den här förstärkaren får inte alstra egna störningar, allt som alstras här hamnar i utgången. Felförstärkaren behöver alltså ha god linjäritet och lågt brus. Problemet med linjäritet och brus har alltså flyttats från effektförstärkaren till felförstärkaren. Men eftersom distorsionen har lägre nivå än signalen, kan felförstärkaren vara på ca 10 - 15 dB lägre effekt. Förstärkningen behöver däremot vara ganska stor, kanske upp till 60 dB. Blandprodukterna som alstras i effektförstärkaren hamnar dels inom bandet och dels i grannkanalerna. Bandbredden för andra slingan behöver därför vara tre (eller fem) gånger så stor som bredden på kanalen.

Balansering av fas och amplitud



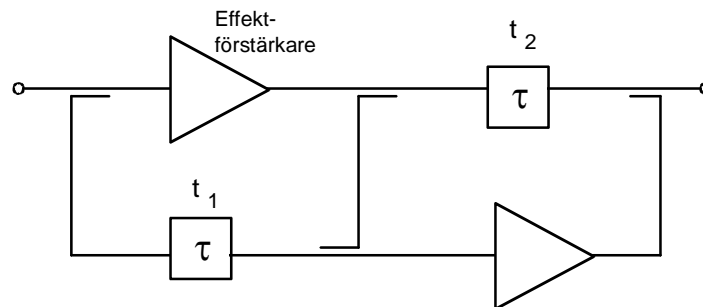
Uppdelning och summering sker med riktkopplare, hybrider eller effektdelare. För att undertrycka signalen i den första slingan behöver de två signalvägarna vara i motfas och ha samma amplitud. Det behövs en fastrimmer och justering av dämpningen för att få stor undertryckning av signalen. Med för dålig undertryckning av signalen, kommer efterföljande felförstärkare att bli överstyrd och alstrar IMD i utgången. En för stor amplitud av insignalen kommer också att försämma nyttsignalen i utgången, eftersom signalvägarna till utgången ligger i motfas.

I den undre signalvägen av första slingan finns en fördröjningsledning, för att kompensera den fördröjning som sker i effektförstärkaren. På så sätt blir signalvägarna motfaskopplade över en större bandbredd. Amplitud- och fas-trimmer kan placeras i den undre signalvägen istället. Men det är bättre i den övre signalvägen, för då blir dess distorsion undertryckt på samma sätt som för effektförstärkaren.

Slingan med IM-undertryckning behöver också motsvarande fördröjning samt amplitud- och fasjustering. Fasskiftare och dämpsats sitter här i den undre signalvägen, för att inte dämpa den höga uteffekten i onödan. Tyvärr måste fördröjningsledningen sitta i den övre vägen. En av de stora nackdelarna med framåtkoppling är just förlusterna i uteffekt på grund av komponenterna som kommer efter effektförstärkaren.

Det är dessutom lättare att hålla konstant förstärkning i systemet om amplitud- och fastrimmer sitter i serie med förstärkarna.

Fördröjning



I båda slingorna ska övre och undre gångväg adderas i motfas. Ju bättre fasinställningen är, desto bättre blir undertryckningen. Men de båda signalvägarna behöver också ha samma tidsfördröjning för att få en bredbandig undertryckning.

Fördröjningen i första slingan t_1 består ofta av en koaxialkabel. Nackdelen med fördröjningsledningen är att dess dämpning försämrar brusfaktorn, men i en effektförstärkare är inte bruset så stort problem.

Den andra slingans fördröjning t_2 är ett större problem. Förlusterna ligger här efter effektförstärkaren och minskar alltså uteffekten. Ett fördröjningsfilter med 6 - 12 resonatorer är ett alternativ med lägre förluster. Med lämpliga korskopplingar får filtret jämn faskång, ca $\pm 0,5^\circ$.

Filtret kan med fördel byggas ihop med intilliggande riktkopplare i samma mekaniska struktur. Filtret har vanligtvis 3 - 6,5 % bandbredd.

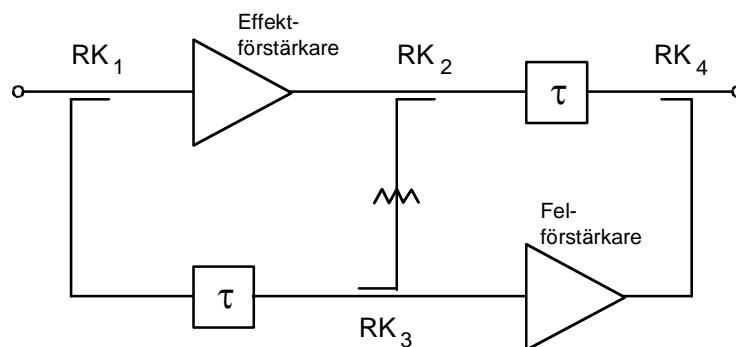
Naturligtvis är ett filter dyrare än en koaxialkabel, men koaxialkabeln har högre förluster. Ett filter kan ha 0,5 dB förluster, och en 0,250 tums halvstyv (semirigid) koaxialkabel har 1.22 dB förluster. Motsvarande högre uteffekt från effektförstärkaren ger ett ännu dyrare system. Fördröjningsfiltret blir alltså ett billigare alternativ. Dessutom blir filtret mycket mindre än motsvarande 0,250 tums kabelslinga.

Tidsfördröjningen behöver inte vara exakt lika i de båda signalvägarna. Det viktiga är att utfasningen blir tillräckligt stor över hela bandbredden.

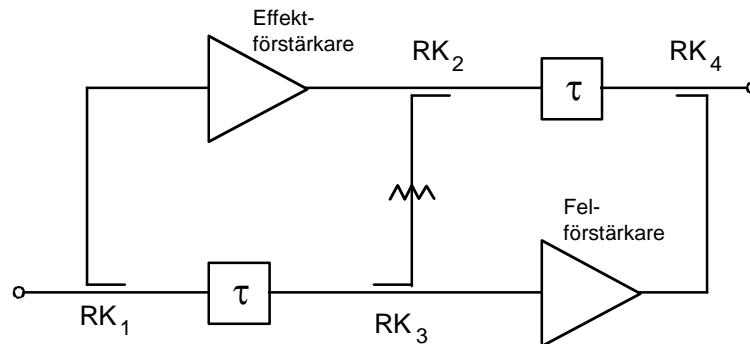
Kommunikation via satellit kan använda 200 MHz bandbredd på 10 GHz. Om fördröjningen är begränsad så att signalvägarna skiljer en period, blir undertryckningen 20 dB. Det är bättre än vad fördistorsion kan ge. Visserligen ger fördröjningen förluster i uteffekten, men fördistorsion kräver en motsvarande minskning i uteffekt (back-off) för att bli tillräckligt linjär.

Smalbandiga system, till exempel 5 kHz på 900 MHz, tolererar ännu större tidsskillnad. Med fyra perioders skillnad mellan signalvägarna är det möjligt att uppnå 80 dB undertryckning. Det är alltså möjligt att helt eliminera fördröjningen eftersom det troligtvis är fasgången i felförstärkaren som begränsar undertryckningen.

Riktkopplare



Effektförstärkarens IMD och brus blir undertryckt av felförstärkaren. Felförstärkarens bör ha låg IMD och lågt brus, eftersom dessa störningar hamnar i utgången. Systemets brusfaktor bestäms av felförstärkaren samt alla förluster framför den. Riktkopplaren RK_1 kommer med 10 dB koppling att försämra brusfaktorn med 10 dB.



RK_1 kopplas vanligtvis så att förlusterna till felförstärkaren blir låga. Med 10 dB koppling till effektförstärkaren blir det 0,5 dB kopplingsförluster till felförstärkaren. En försämring av brusfaktorn med 0,5 dB är ganska måttligt. Förlusterna till effektförstärkaren kompenseras med högre förstärkning där.

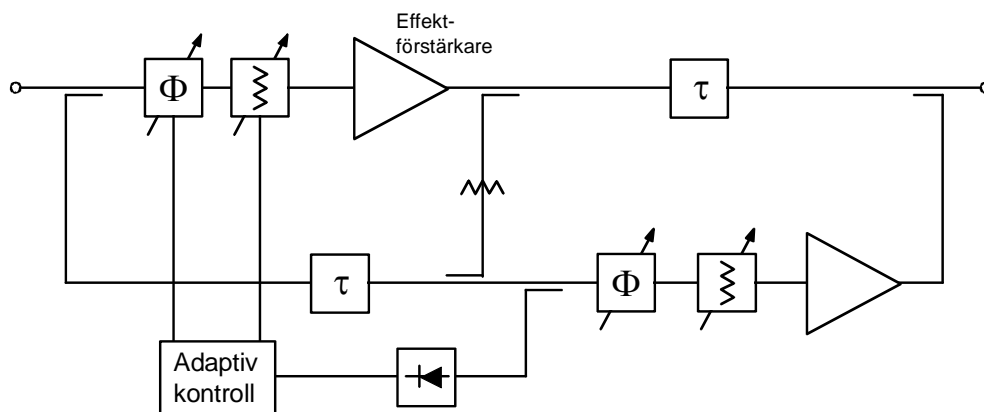
RK_2 ska ha så låga förluster som möjligt, för att inte dämpa uteffekten. Vanligtvis används 20 dB koppling för att få försumbara förluster. RK_3 och den eventuella dämpsatsen ska väljas så att effektförstärkarens signal fram till felförstärkaren blir lika stor som insignalen framme vid felförstärkaren. De ska ju balansera bort varandra.

RK_4 ska också ha låga förluster för att inte förlora uteffekt och verkningsgrad. Men en svag koppling kräver högre effekt från felförstärkaren, dvs. sämre verkningsgrad. Största verkningsgrad blir det med något mindre än 10 dB kopplingsfaktor. Om felförstärkaren är av klass A med dålig verkningsgrad ska kopplingsfaktorn vara neråt 6 dB.

Riktkopplarna måste också ha så hög direktivitet att signalen inte kan koppla tillbaks till respektive förstärkares ingång. Det ger då instabilitet och självsvängning. Uteffekten får inte heller koppla ner till felförstärkarens utgång och alstra ytterligare IMD.

Automatisk balansering av signalen

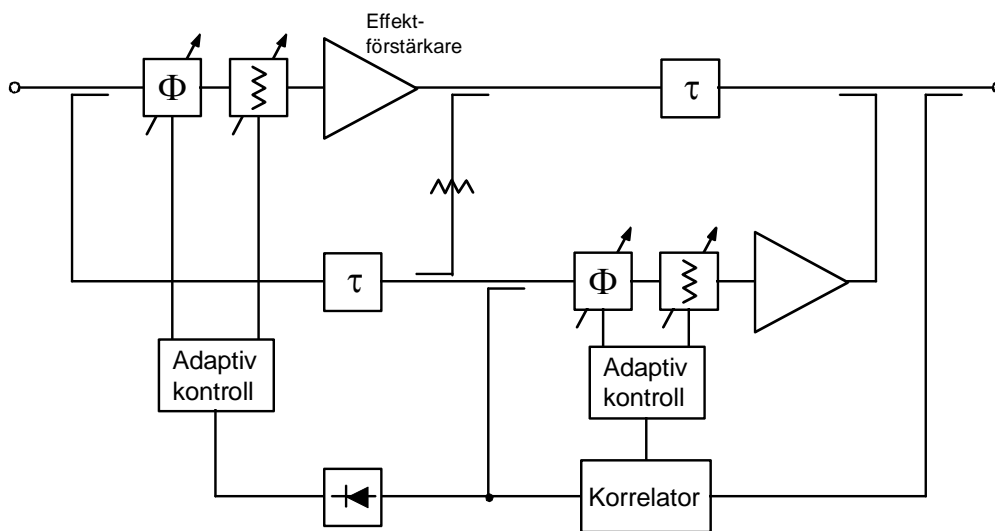
Åldring och temperaturvariationer i effektförstärkaren ger variation i fas och amplitud. Även signalnivå och DC-spänning påverkar linjäriseringen. För att få 25 dB undertryckning behöver amplitudfelet vara mindre än 0,5 dB och fasfelet mindre än 5° . En undertryckning på 40 dB behöver en noggrannhet på 0,1 dB och $0,1^\circ$. För att få en stor undertryckning behöver alltså fas och amplitud regelbundet efterjusteras.



En detektor mäter undertryckningen av signalen, och en kontrollkrets ställer in fas och amplitud till bästa undertryckning. Det räcker med att signalen blir i samma storleksordning som IM-produkterna. Detektorn mäter därför totala effekten, och kontrollkretsen optimerar fas och amplitud till att få lägsta nivå.

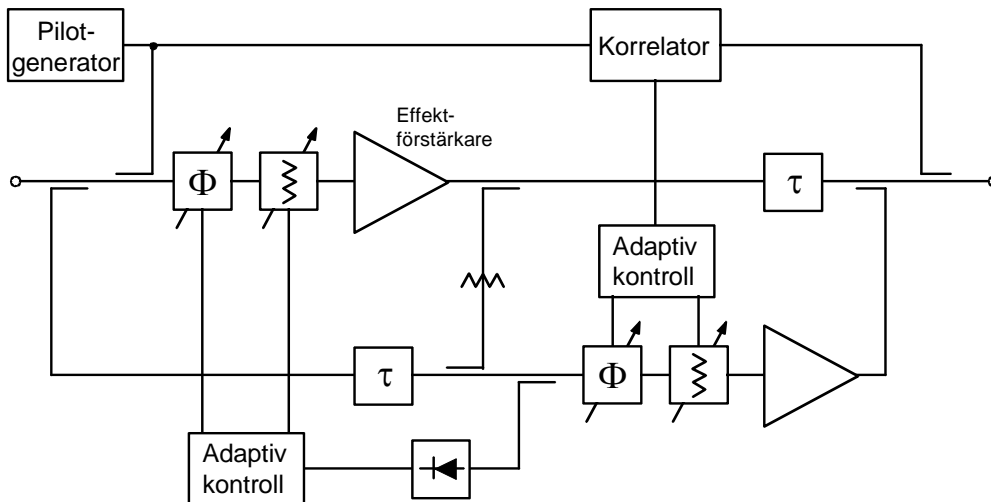
Automatisk balansering av IM

I slingan för IM-undertryckning ska distorsionen mätas, så att kontrollkretsen kan ställa in fas och amplitud till lägsta IMD. Problemet är att mäta en svag signal då det samtidigt finns en signal på mycket hög effekt. En mätning av totala effektnivån ger endast en grov indikation på hur bra amplitud och fas har ställts in.



Distorsionen i utgången kan mätas med en korrelator, med den extraherade felsignalen (distorsionen) som referens. En förutsättning är att nyttsignalen blir tillräckligt undertryckt, så att den detekterade signalen endast beror på distorsionen som finns i utgången. Reglerkretsen justerar adaptivt fas och amplitud så att distorsionen i utgången minimeras.

Reglering med pilotsignal

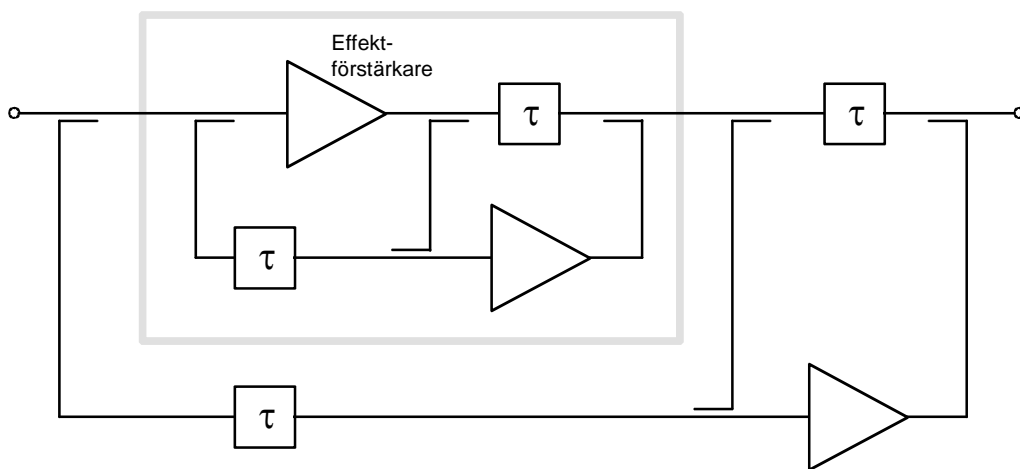


En liten testsignal injiceras på förstärkarens ingång. Denna pilotsignal kommer att undertryckas, precis som de andra störningarna från effektförstärkaren. Storleken på pilotsignalen är ett mått på IM-undertryckningen. På utgången mäts storleken av pilotsignalen med en smalbandig mottagare eller korrelator. Men den svaga testsignalen måste kunna mätas trots utsignalens mycket starkare effekt. Med riktkopplaren och en extra dämpsats minskas signalen så att inte mottagaren blir överstyrd. Det betyder att mottagaren måste avskärmas mycket noggrant för att inte de stora nivåskillnaderna i systemet (ca 120 dB) ska besvärmas av överhörning. För att få bästa prestanda över hela bandet bör pilotsignalen dessutom vara en bredbandig bruslik signal.

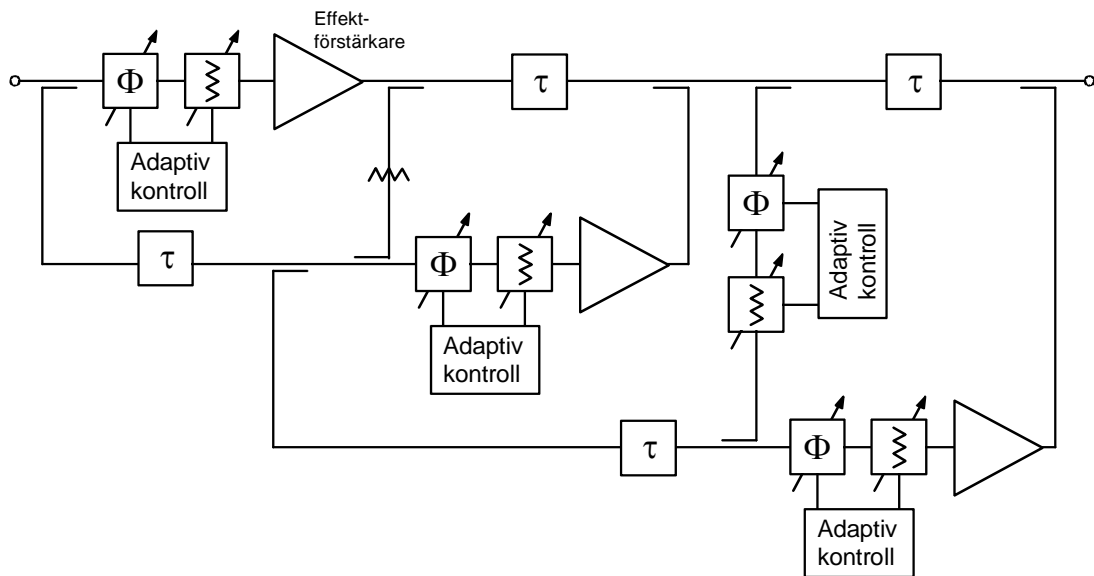
Dubbel framåtkoppling

Dual Loop , Nested Loop

Om det inte räcker med den linjäriserade förstärkaren, kan man göra ytterligare en linjärisering med en framåtkoppling till.



Den framåtkopplade kretsen har som effektförstärkare ett framåtkopplat slutsteg. IM-
undertryckningen blir den sammanlagda undertryckningen från båda stegen, ca 25 dB
per steg.



Det behövs mycket kretsar för att uppnå de riktigt stora undertryckningarna som -75 dBc. Men för en viss undertryckning går det att använda enklare kretsar. Det är lättare att uppnå 15 dB i två kretsar än 30 dB i en krets. Det stora antalet kretsar gör att verkningsgraden bara blir ca 5 %.

Verkningsgrad

En GSM-1800 basstation med 8 carrier kan ha 2,5 kW effektförbrukning, och varje operatör har tusentals basstationer. Eftersom förlusterna domineras av RF-kretsarna är det mycket viktigt att förbättra verkningsgraden på effektförstärkaren.

En klass AB effektförstärkare har ca 25 % verkningsgrad, och en linjär klass A förstärkare har ca 5 % verkningsgrad

En klass C förstärkare kan ha en verkningsgrad på 60 % och IM_3 på -15 dBc. Om både effektförstärkaren och felförstärkaren är av klass C, blir den maximala verkningsgraden 42 %. Med 1 dB förluster i fördröjningsledningen minskar verkningsgraden till 35 % och med 2 dB blir det 28 % verkningsgrad. Det blir alltså en ganska hög verkningsgrad tillsammans med en linjäritet som är jämförbar en klass A förstärkare.

Om felförstärkaren är av klass A blir det lägre verkningsgrad. I praktiken blir verkningsgraden för framåtkoppling mindre än 10 %. Fördelen med att kombinera en effektförstärkare klass C med en felförstärkare klass A är att både linjäritet och verkningsgrad blir mycket bättre än en självständig klass A effektförstärkare.

Om effektförstärkarens linjäritet har förbättrats med fördistorsion, blir IM -produkterna lägre. Felförstärkaren kan då arbeta på lägre effektnivå, och på så sätt få högre verkningsgrad, och riktkopplaren på utgången kan ha svag koppling det vill säga små förluster.

Fördelar med framåtkoppling

Framåtkoppling kan ge mycket god linjäritet över en stor bandbredd. Undertryckningen sker över ett stort dynamikområde. Andra metoder fungerar bara för små variationer i signalens amplitud. Framåtkoppling förbättrar också verkningsgraden betydligt

Framåtkoppling kan användas till valfri modulationstyp, även de med varierande amplitud. Den har använts till IS-95 och D-AMPS, men kan också användas till WCDMA. Flera olika signaler kan till och med förstärkas samtidigt. En förstärkare som kan hantera flera kanaler samtidigt kallas MCPA (Multi Carrier Power Amplifier).

Fördistorsion minskar endast IM-produkter av lägre grad, 3:e gradens IMD och eventuellt 5:e gradens. Framåtkoppling minskar all distorsion lika mycket.

Framåtkoppling kompenserar även minneseffekter, eftersom dessa variationer är inkluderade i felsignalen (IM-produkterna). Minneseffekterna innebär att förstärkning och färgång beror på tidigare signalamplitud. Det kan bero på förändringar i temperatur eller förspänningskretsar.

Referenser

Framåtkoppling

Barney Arntz, "Second order effects in feedforward amplifiers", Applied microwave & wireless, jan 2000 pp 66-75

P B Kennington, "High efficiency power amplifier linearization for mobile communications", Microwave & RF conference 1995 pp 24-28

P B Kennington, "Analysis of instability in feedforward loop", Electronic letters sep 1997 vol 33 no 20 pp 1669-1671

David Wills, "A control system for a feedforward amplifier", Microwave Journal april 1998 pp 22-34

Yougoo Yang, "Digital controlled adaptive feedforward amplifier for IMT-2000 band", IEEE MTT-S 2000 pp 1487-1490

Ahmad Khanifar, "Error signal reuse in a feedforward amplifier", IEE MTT-S 2002 pp 473-475

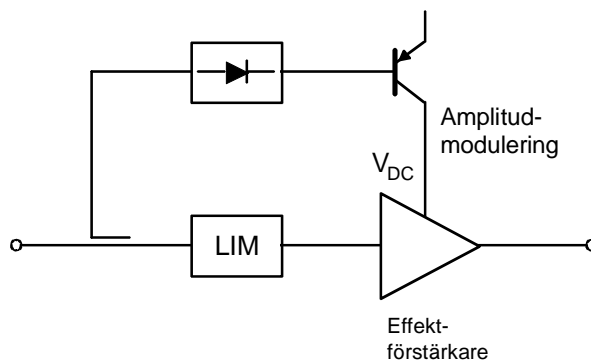
Ashok K Talwar, "Reduction of noise and distortion in amplifiers using adaptive cancellation", Transactions on microwave theory and techniques, vol 42 no 6 june 1994 pp 1086-1087

20. Envelop-eliminering och återmodulering

EER — Envelope Elimination and Restoration

Uppdelningen i amplitud och fas kallas polär modulation, till skillnad mot IQ-moduleringen som har en kartesisk uppdelning. Polär modulation, det vill säga EER, kallas också ibland Kahns teknik.

En klass C förstärkare har hög verkningsgrad. Om signalen innehåller variationer i amplitud behövs en linjär förstärkare, som tyvärr har låg verkningsgrad. EER undviker problemet genom att ta bort AM-variationerna och lägga till dem senare.

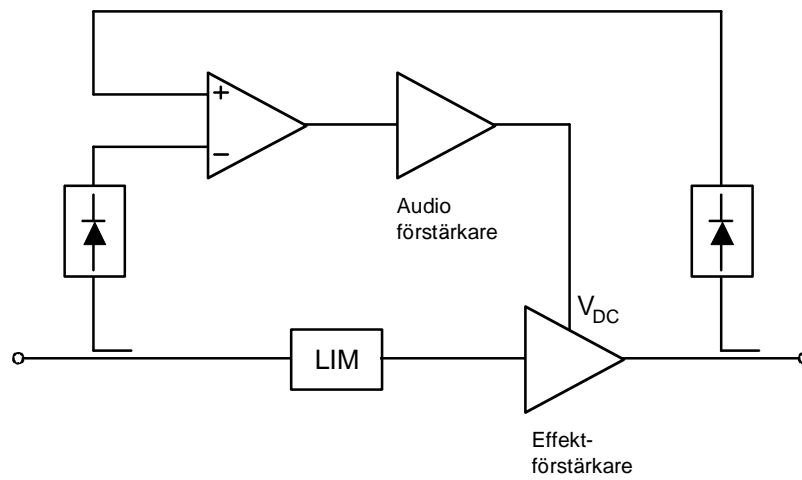


Signalens amplitud tas bort med en limiter. Istället detekteras amplitudvariationerna i en parallell signalväg. Efter limitern har signalen den önskade fasvariationen men med konstant amplitud, och kan alltså förstärkas med en klass C förstärkare. Den detekterade amplitudvariationen moduleras sedan på slutsteget genom att styra drivspänningen.

Om effektförstärkaren innehåller både drivsteg och slutsteg, kan även drivsteget amplitudmoduleras. Vid reducerad uteffekt (back-off) förbättras drivstegets verkningsgrad. Totala verkningsgraden över hela dynamikområdet förbättras därigenom markant. Alternativet med drivsteg på max effekt ger dessutom mer läckage till utgången då amplituden ska vara låg.

Amplitudinformationen behöver förstärkas med en audioförstärkare på hög effekt. Audioförstärkaren kan vara en pulsbreddsmodulerad switchad klass S förstärkare. RF-transistorns drivspänning styrs genom att variera pulsbredden. Ett annat alternativ är en delta-modulerad styrning för att få bättre linjäritet och bandbredd. Det fungerar bra upp till några tiotal kHz. Vid större bandbredd blir verkningsgraden betydligt försämrad.

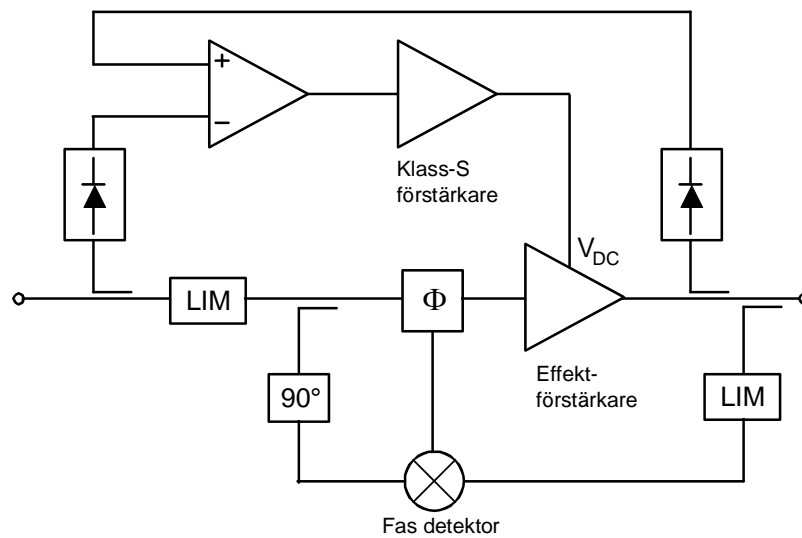
Återkoppling av envelop



En del av utsignalen detekteras och återkopplas. Utsignalens envelop jämförs med insignalens envelop, och det är skillnaden som styr matningsspänningen till effekttransistorn. Med återkopplingen säkerställs att utsignalens envelop följer insignalens envelop. Det ger alltså bättre linjäritet.

En förutsättning är att signalvägarna har lika fördröjning, så att amplitudvariationerna blir synkroniserade med fasvariationerna. I annat fall blir det distorsion och störningar i grannkanalerna. Då fördröjningen ökar blir störningarna i grannkanalerna allt större.

Återkoppling av fas



Utsignalens fas jämförs med insignalen i en fasdetektor. Felspänningen korrigerar en fasmodulator så att utsignalens fas följer insignalen. Störningarna i grannkanalerna (ACPR) blir betydligt lägre. Nackdelen är att kretskopplingen har blivit mer komplicerad.

Verkningsgrad

En effektförstärkare behöver en strömförsörjning på hög effekt. Den svaga detekterade signalen måste därför förstärkas med en effektförstärkare på basbandet (audio). RF-förstärkare har hög verkningsgrad eftersom det kan vara en klass C förstärkare som drivs till max effekt. Men även audioförstärkaren behöver ha god verkningsgrad, eftersom det är den totala verkningsgraden som är intressant.

En klass S förstärkare har mycket hög verkningsgrad, 90 % vid full effekt och 80 % vid 10 dB reduktion. Den totala verkningsgraden är en produkt av både RF-förstärkarens och audioförstärkarens verkningsgrader. Med klass C förstärkare för RF och klass S till styrningen blir den totala verkningsgraden ca 54 %.

En mycket viktig egenskap är att den håller hög verkningsgrad även då signalamplituden är ganska låg. RF-förstärkaren drivs ju alltid till mättnad, dvs klass C. Den har ca 60 % verkningsgrad även då signalnivån minskat 18 dB. Den totala verkningsgraden, då signalamplituden minskat 18 dB, är fortfarande ca 35 %. En klass B förstärkare skulle bara ha fått ca 6,5 % verkningsgrad. En klass E förstärkare ger 80 % verkningsgrad över dynamikområdet. Totala verkningsgraden blir då 64 %.

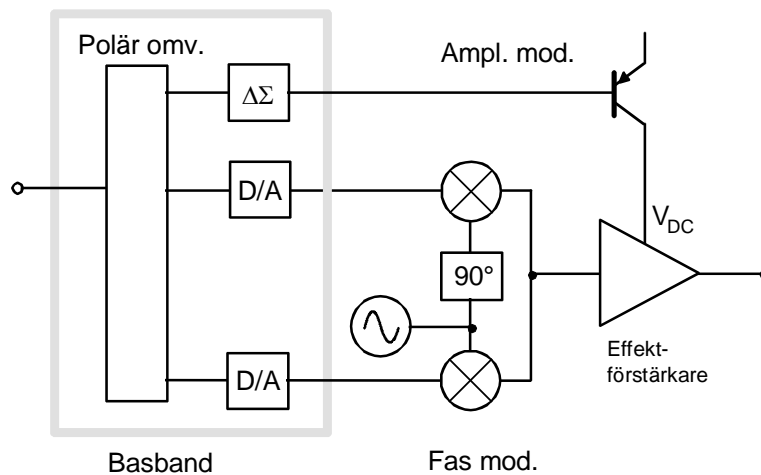
Drivspänningen, som har modulerad pulsbredd, behöver en switchfrekvens som är mycket högre än amplitudvariationernas bandbredd, för att kunna filtreras. Ju högre switchfrekvensen är desto sämre verkningsgrad får klass-S förstärkaren. Deltamodulering switchar bara då felsignalen ändrar polaritet. Det ger bättre verkningsgrad.

Distorsion från fördröjningen

IMD beror främst på modulatorns bandbredd och fördröjningen mellan amplitudmoduleringen och den limiterade fasinformationen. Efter klass-S förstärkaren måste signalen filtreras. Det ger en fördröjning på amplitudinformationen på 5 - 20 μ s. Bärvägen som innehåller fasinformationen bara är fördröjd 1 - 2 μ s. För att få C/I bättre än 50 dBc behöver klass-S förstärkaren ha 6 gånger större bandbredd än själva informationen. Skillnaden i fördröjning mellan amplitud och fas ska vara mindre än $0,03 / B_{RF}$.

En signal med 5 MHz bandbredd ska alltså ha sin amplitud och fas inom ca 5 ns. Det kan behövas en extra fördröjning i fas-kanalen. Återkoppling av envelopen kan minska inverkan av fasfelet.

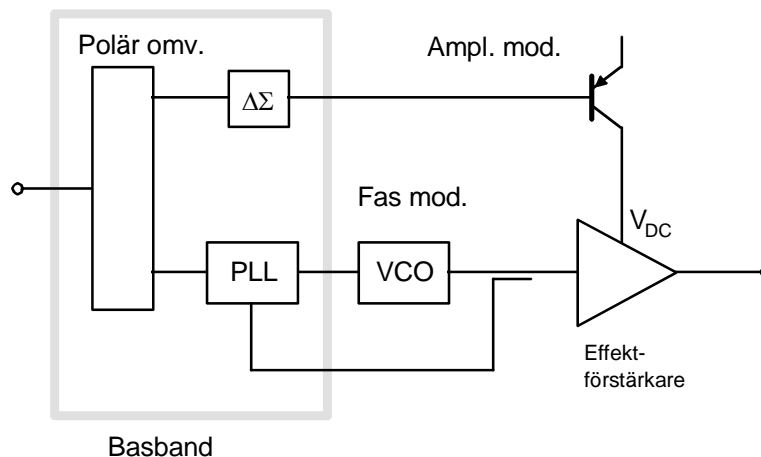
Insignal på basbandet



Om förstärkaren är en del av en sändare så finns redan amplitudinformatjonen och behöver alltså inte detekteras. Insignalen på basbandet delas upp i separat amplitud och fasinformation. Amplituden styr DC-omvandlaren. IQ-modulatorn modulerar fasen med konstant amplitud.

Fördröjningen av en samplad signal kan i första hand göras med samplingsintervallen. Noggrannare inställning kan interpoleras fram.

Modulering av VCO



Fasinformationen kan styra en faslåst slinga med en oscillator på den önskade frekvensen. Amplituden moduleras genom slutstegets strömförsörjning. En stor del av signalprocessen sker på basbandet. Det blir alltså en enkel kretskoppling för en linjäriserad sändare.

Alternativt kan oscillatoren arbeta på lägre frekvens, som sen blandas upp till den rätta frekvensen. På så sätt undviks att effektförstärkaren stör oscillatoren.

Eftersom oscillatoren kan ställas in på valfri centerfrekvens blir kretsen en synthesizer med inbyggd kanalväljare.

Bandbredd

En CDMA-signal är vanligtvis QPSK-modulerad och filtrerad. Förflyttningen mellan tillstånden kan då passera ganska nära origo. Det ger en minskning av amplituden och en variation i fas upp till 180°. Variationerna i amplitud och fas blir mycket snabba, jämfört med de filtrerade I- och Q-signalerna. Modulation med amplitud och fas (polär modulation) behöver en betydligt större bandbredd än IQ-modulatorens.

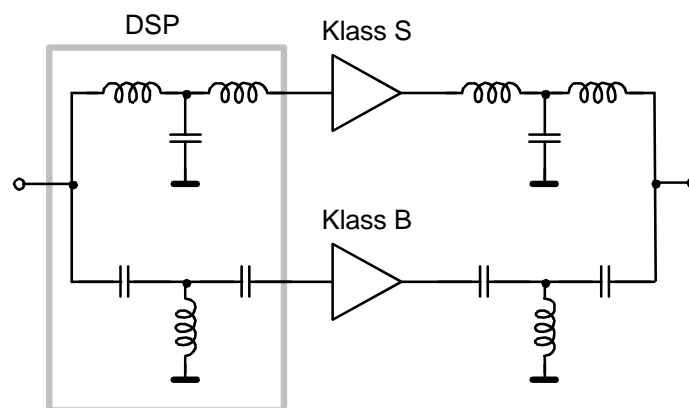
Bandbredden på amplitudvariationerna bör vara 4 - 8 gånger så stor som RF-bandbredden. Bandbredden kan förbättras genom att lyfta de högre frekvenserna som ligger ovanför modulatorens bandbredd (pre-emphasis). På basbandet adderas alltså ett

högpassfilter med en pol på 10 gånger RF-bandbredden. Den förbättringen i bandbredd kan ge 6 - 18 dB undertryckning av IM-produkterna.

PLL behöver också ha en mycket stor bandbredd på modulationen. Frekvensområdet för fasinformation kan också förbättras på motsvarande sätt som för amplituden.

Klass S kombinerat med Klass B

Klass S kan användas för att effektivt förstärka låga frekvenser. För de flesta vågformer ligger ca 80 % av effekten i DC-komponenten, och 99 % av effekten finns på frekvenser lägre än RF-bandbredden. Men de högre frekvenserna behövs också för att undertrycka störningarna i grannkanalerna.



En klass B förstärkare som användas för att förstärka de höga frekvenserna kopplas parallellt med klass S förstärkaren. De kombineras med en diplexer på utgången. På ingången separeras de två frekvensområdena med motsvarande diplexer i en DSP.

Ytterligare en förstärkare försämrar verkningsgraden. Men klass B förstärkaren kan arbeta på en låg effektnivå eftersom den bara behöver förstärka de korta topparna. Det blir alltså en mycket stor ökning av bandbredden men bara en liten försämring av verkningsgraden.

Fördistorsion av amplituden

När RF-amplituden är tillräckligt stor för att switcha effekttransistorn mellan strypning och mätnad, är det ett linjärt samband mellan utsignalens amplitud och DC-spänningen. Vid låga nivåer på insignalen kommer inte effekttransistorn att switchas ordentligt. Den lägre förstärkningen betyder att det inte längre är ett linjärt samband. Men den avvikelsen kan korrigeras med fördistorsion. Transistorns förstärkning mäts upp för olika nivåer på insignalen. Avvikelsen lagras på basbandet i en PROM-tabell, eller räknas om till ett polynom.

Distorsionen som uppstår i den polära modulatoren kan alltså korrigeras på två sätt. Den ena sättet är fördistorsion av amplituden respektive fasen. Det andra sättet är detektering och återkoppling av amplitud respektive fas. Fördistorsionen ger en enklare krets utan förluster på RF-sidan.

Användning

EER ger hög verkningsgrad även då signalens amplitud varierar, inte enbart för maxeffekten som hos andra linjäriseringar. Hur låga amplituder som kan användas begränsas av RF-transistorn.

Kretsen utvecklades för att ge linjär förstärkning för SSB signaler på HF. Men den har också använts till radio- och TV-sändare på hög effekt. Den kan användas antingen till linjärisering av en förstärkare eller till linjärisering av en hel sändare.

Om förstärkaren är förspänd till klass A blir det en krets med adaptiv förspänning (envelop tracking). Förstärkaren arbetar hela tiden i sitt linjära område oavsett insignalens amplitud. Adaptivt förspänd klass A ger den bästa linjäriteten. EER med klass C eller klass E används för att få den högsta verkningsgraden.

Referenser

EER

Frederick H Raab, "L-band transmitter using Khan EER technique", Transactions on microwave theory and techniques, vol 46 no 12 dec 1998 pp 2220-2224

Kang-Chun Peng, "A novel EER transmitter using two point delta-sigma modulation scheme for WLAN and 3G applications", IEEE MTT-S 2002 pp1651-1654

Manoja D Weiss, "Linearity of X-band class F power amplifiers in high efficiency transmitters", Transactions on microwave theory and techniques, vol 49 no 6 june 2001 pp 1174-1179

B Bakkaloglu, "Bandwidth extension technique for polar modulated RF transmitters", Electronics Letters 13th April 2006 Vol. 42 No 8

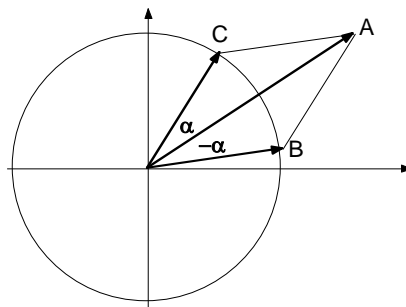
T. Nesimoglu, "Improved EER transmitters for WLAN", IEEE Radio and wireless symposium 17-19 januari 2006 pp 239-242

Frederick H Raab, "Split-band modulator for Kahn-technique transmitters", IEEE MTT-S 2004 pp 887-890

21. LINC

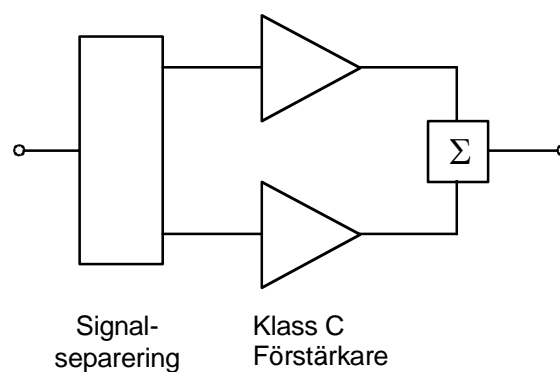
LINC — **L**inear amplification using **n**on-linear components

LINC använder olinjära förstärkare för att ge linjär förstärkning



En signal (A) med viss amplitud och fas, kan delas upp i två delvektorer med lika stora amplituder (B och C). Eller om man vänder på resonemanget, så kan två signaler med samma amplitud ge resultatanten (A). Det intressanta är att resultatanten kan få olika amplituder beroende på hur långt isär de två vektorerna ligger. Resultanten kan alltså bli modulerad i amplitud genom att variera fasläget på de två delsignalerna.

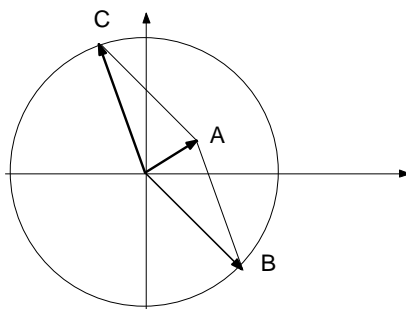
Linjäriserad förstärkare



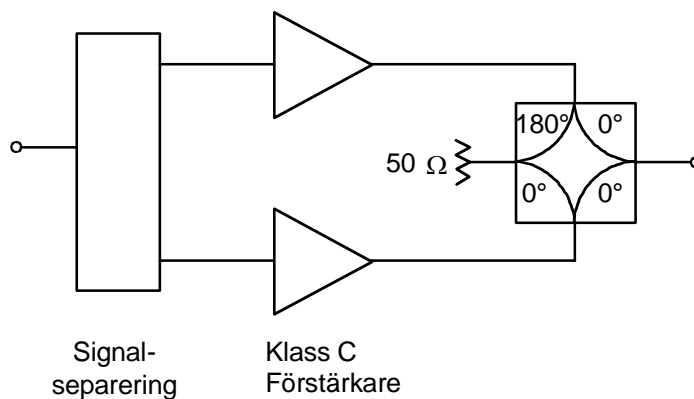
Insignalen delas först upp i två delsignaler. Förstärkningen kan nu ske med Klass C förstärkare, eftersom de två delsignalerna hela tiden har konstant amplitud. Därefter summeras de två delvektorerna till en utsignal som kan vara AM-modulerad.

Förluster i summeringen

Klass C förstärkarna har hög verkningsgrad. Men förlusterna i kretsen för summering försämrar verkningsgraden.



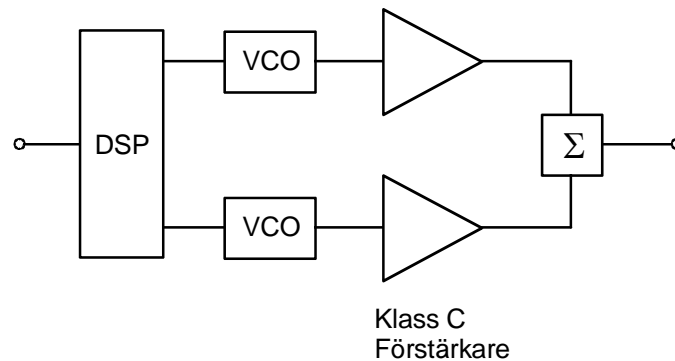
Då utsignalens amplitud (A) ska vara låg, orienteras de två delvektorerna (B och C) nästan i motfas. De två effektförstärkarna ger alltid hög uteffekt. När utsignalen ska vara liten, är den största delen av effekten att betrakta som förluster.



Hybriden summerar den önskade signalen till utgången. All övrig effekt hamnar i avslutningsmotståndet på skillnadporten. Det ger en klart försämrad verkningsgrad. Med en amplitudvariation som för 64QAM kan verkningsgraden bli så låg som 20 %.

Istället för att låta effekten gå förlorad som värme, kan skillnadskanalens effekt detekteras till en likspänning som återförs till strömförsörjningen. Om verkningsgraden är 60 % vid maxeffekten, blir verkningsgraden åtminstone 40 % då uteffekten minskas 10 dB.

Linjäriserad sändare



Uppdelningen av insignalen görs enklast med en digital signalprocessor på basbandet. Dessa signaler på basbandet överförs sedan till RF-signaler med varierande fas, med hjälp av spänningsstyrda oscillatorer (VCO).

Eftersom insignalen ligger på basbandet, är det egentligen en hel sändare som har blivit linjäriserad. De två signalerna styrs direkt från DSP. Det betyder att systemet snabbt kan hoppa mellan olika modulationstyper.

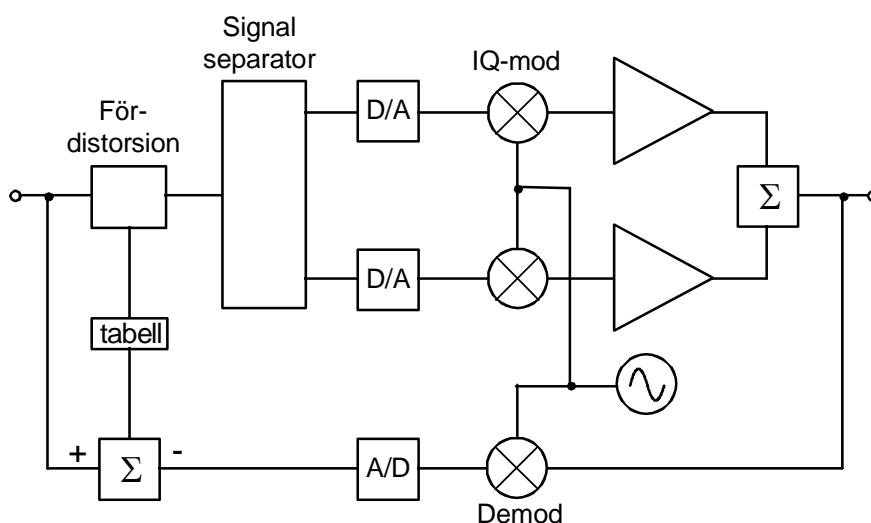
Nackdelen med DSP är den stora strömförsörjningen som minskar systemets verkningsgrad. En annan nackdel är den stora bandbredden som DSP behöver. Styrningen av VCO behöver minst 10 gånger större bandbredd än vad den önskade utsignalen ska ha.

LINC med fördistorsion

Det är ganska lätt att få en två-tons IMD på 20 dBc, men det går också att uppnå 40 - 60 dBc. En förutsättning är att de två signalvägarna är lika. Dessutom ska signalvägarna vara lika över hela frekvensområdet, och då temperatur och omgivning varierar. Det är svårigheterna med matchning mellan kanalerna som har begränsat användningen av LINC.

Då utsignalen ska ha låg amplitud är de två delsignalerna nästan i motfas. Det är i det läget som kanalernas balans är som viktigast för att få låg distorsion. Till skillnad från vanliga förstärkare får LINC mindre distorsion vid större signalstyrka.

Ett CDMA system kan behöva en fasmatchning på $0,2^\circ$ och amplitudmatchning på 0,5 dB mellan kanalerna.



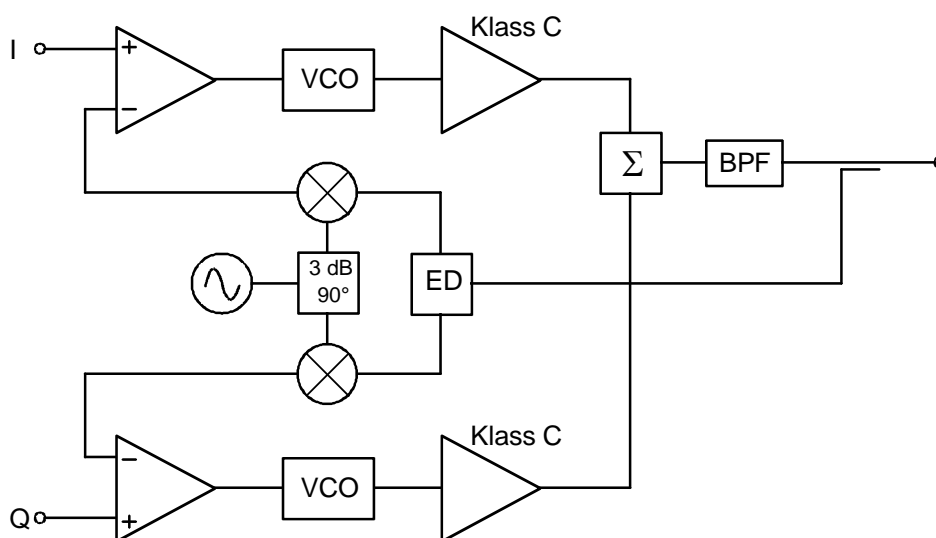
Om en DSP används så kan den dessutom utnyttjas för att ge fördistorsion. Utsignalen blandas ner till basbandet och jämförs med insignalen. Felsignalen väljer lämplig inställning av fördistorsionen från en PROM-tabell. Fördistorsionen består av en skalning av amplituden och rotation av fasen. Fas- och amplitudfel justeras alltså med hjälp av en tabell, som adaptivt kan uppdateras då och då.

Ett annat alternativ är kontinuerlig återkoppling i form av CALLUM.

CALLUM

CALLUM — Combined analog locked loop universal modulator

De två signalvägarna i LINC behöver ha en mycket god matchning i både amplitud och fas. CALLUM är en krets med återkoppling, för att de två signalvägarna ska regleras till god matchning.



Den återkopplade signalen demoduleras till basbandet med en IQ-demodulator. Insignalen på basbandet är också uppdelad i I- och Q-signaler. Skillnaden mellan insignal och återkoppling ger felfspänningar som styr de spänningsstyrda oscillatorerna (VCO).

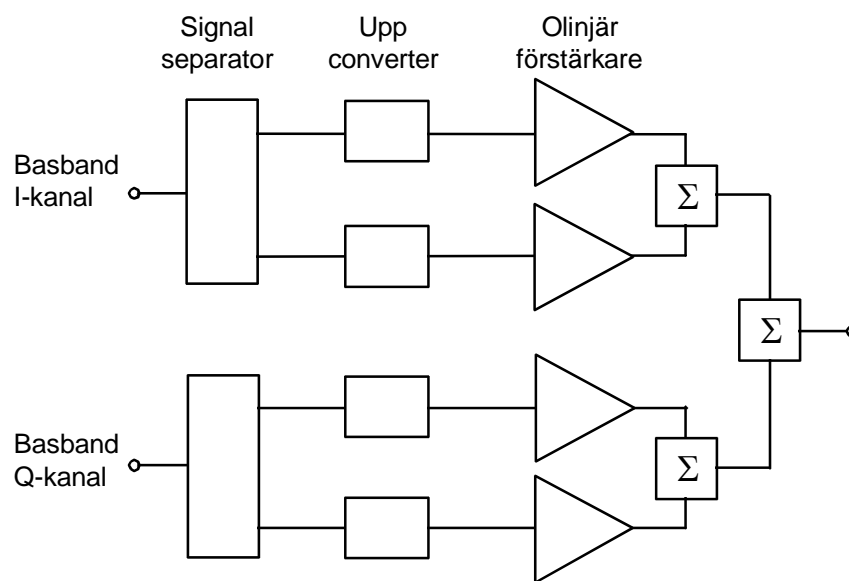
Respektive VCO kommer att låsa till den frekvens som demodulatorns oscillator är inställd till. Det är alltså demodulatorns oscillator som används som kanalväljare. Den kan därför vara utformad som en syntesgenerator.

Återkopplingen kommer alltså att justera VCO så att utsignalen blir en förstärkt kopia av insignalen. Resultatet är en linjär förstärkare, som genom återkopplingen behåller sin linjäritet även under temperaturvariationer och åldring.

En nackdel är att återkopplingen behöver en bandbredd som är mycket större än signalens bandbredd. Dessutom behöver återkopplingens bandbredd begränsas för att återkopplingen inte ska ge självsvängning. Resultatet är att signalen blir ganska smalbandig.

LINC med IQ-delning

En uppdelning av signalen i amplitud och fas kräver större bandbredd. Speciellt faser varierar mycket snabbt i närheten av origo. Om man istället använder IQ-uppdelning och sen adderar LINC i respektive kanal, krävs mindre bandbredd.



I-kanalen och Q-kanalen har varsin LINC-förstärkare. I utgången summeras I och Q och bildar den önskade utsignalen. Upp-convertern kan bestå av en fasmodulator, till exempel en vektormodulator (IQ-modulator), eller en VCO (synthesizer) på rätt frekvens som fasmoduleras.

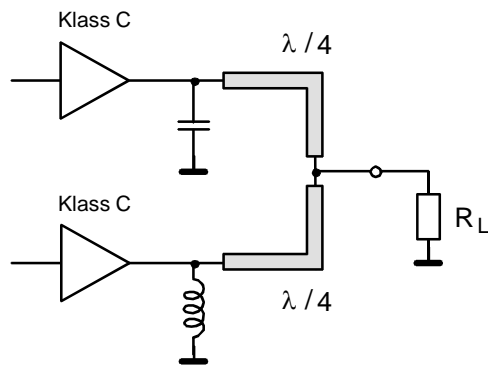
Förstärkarna arbetar med konstant amplitud på signalen. De kan alltså vara förspända till klass C för att få hög uteffekt och verkningsgrad. Alternativt kan de vara switchade förstärkare av klass E eller klass F.

Kretskopplingen blev dubbelt så stor, men fördelen är att den inte kräver så stor bandbredd. En annan fördel är att de två kanalerna med konstant amplitud inte behöver vara lika noggrant matchade.

Kombinering med Chireix

LINC sammansätter de två signalerna med en krets som är anpassad. Den delen av signalerna som inte sammansätts i utgången blir istället resistivt avslutade i kretsen. Det betyder en stor effektförlust då utsignalen ska vara liten.

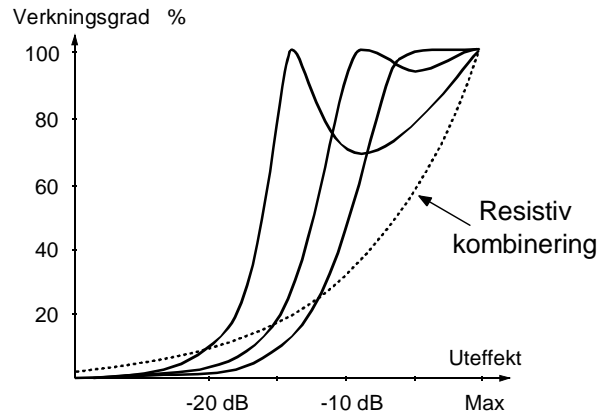
Chireix är en helt reaktiv kombineringskrets utan effektförluster.



Vid max amplitud ligger de två förstärkta signalerna i fas. Efter kvartvågsledningarna adderas de i fas i lasten. Vid inställning till lägre amplitud ut är de två signalerna fasvridna åt varsitt håll. Den del av signalerna som inte sammansätts i lasten, reflekteras tillbaka till förstärkarna. Den reaktiva delen av lasten som förstärkarna ser kan kompenseras med reaktans. Eftersom de två insignalerna har modulerats med fasvridning åt varsitt håll, behöver den ena förstärkaren kompenseras med en kapacitans och den andra med en induktans.

Reflektionen tillbaka från knutpunkten ser vid förstärkarna ut som en ökning av lastimpedansen då fasskillnaden ökar. Det för med sig att förstärkarna levererar mindre RF-effekt, och konsumerar mindre DC-effekt. Verkningsgraden är alltså hög trots att RF-amplituden är liten.

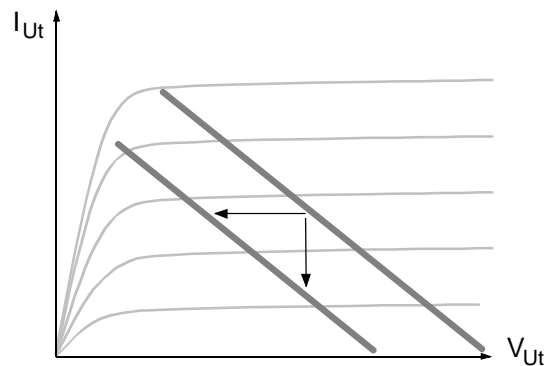
När förstärkarna ser en rent resistiv last är verkningsgraden hög. Lägre amplitud ger en last med reaktans, som kompenseras vid förstärkaren. Man kan valfritt välja vid vilken signalnivå som lasten ska bli resistiv.



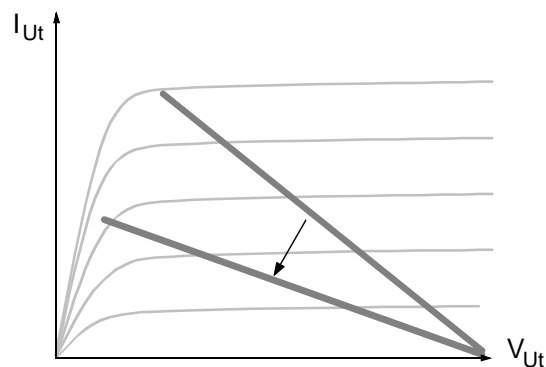
Kombineringsens verkningsgrad är maximal då uteffekten är max, samt vid den nivå som har kompenserats. Figuren visar tre kurvor för tre olika kompenserings. Den streckade kurvan visar en jämförelse med den isolerade resistiva kombineringsen, till exempel med en Wilkinson.

Jämförelse

LINC använder en isolerad anpassad kombinerings. Det ger en god linjäritet men med försämrade verkningsgrad. Chireix använder en reaktiv kombinerings med hög verkningsgrad. Nackdelen är att variationer i belastningens impedans ger distorsion. Den distorsionen kan kompenseras med en fördistorsion av fasen. Men en krets med fördistorsion behöver också en viss strömförsörjning som minskar verkningsgraden.



Adaptiv förspänning (Envelope Tracking) och EER minskar DC-spänningen och strömmen i proportion till önskat effektuttag.



Doherty och Chereix ökar lastens impedans för att på så sätt minska strömförbrukningen.

Referenser

LINC

Robert Langridge, "A power reuse technique for improved efficiency of outphasing microwave power amplifiers", Transactions on microwave theory and techniques, vol 47 no 8 aug 1999 pp 1467-1470

Paloma Garcia, "Nonlinear distortion cancellation in OFDM systems using adaptive LINC structure" PIMRC 2004 pp1506-1510

Ilkka Hakala, "A 2.14 GHz Chireix outphasing transmitter", Transactions on microwave theory and techniques, vol 53 no 6 June 2005 pp 2129-2137