

Modulatie en demodulatiemethoden

Roeland J. Dilz

15 mei 2007

Inhoudsopgave

1	Inleiding	3
2	Algemene wiskundige inleiding	4
2.1	De fourierreeks	4
2.2	De Fouriertransformatie	6
2.3	voorbeeld: Het RC filter	7
3	CW	10
3.1	algemene informatie	10
3.2	wiskundige representatie	10
3.3	modulatie	11
3.4	demodulatie	13
4	Amplitudemodulatie	15
4.1	algemene informatie	15
4.2	wiskundige representatie	16
4.3	modulatie	18
4.4	demodulatie	22
5	Enkelzijband	26
5.1	algemene informatie	26
5.2	wiskundige interpretatie	26
5.3	modulatiemethoden	27
5.4	demodulatie	35
6	FM	37
6.1	algemene informatie	37
6.2	wiskundige interpretatie	37
6.3	modulatie	43
6.4	demodulatie	43
7	Quadratuur Amplitude Modulatie	46
7.1	algemene informatie	46
7.2	wiskundige beschrijving	46
7.3	modulatie	49
7.4	demodulatie	49

1 Inleiding

In dit artikel behandel ik de meest gebruikelijke modulatiemethoden voor analoge signalen. Het is bedoeld voor mensen met enige ervaring met (HF-)electronica. Er staat redelijk wat wiskunde in. Maar het grootste deel valt met VWO wiskunde te snappen, als men de moeite neemt zich er in te verdiepen. Ik heb geprobeerd om er een redelijk aantal grafieken in te zetten om het wat begrijpelijker te maken. Ben je echter niet geïnteresseerd in de theorie, slechts in de praktische uitvoering, dan zijn er op het internet genoeg beter geschikte cursussen te vinden.

Voor vragen: Open een topic op www.circuitsonline.net, Alwaar de altijd vriendelijke forumgebruikers u maar al te graag zullen helpen! Ook voor suggesties om dit artikel te verbeteren kunnen daar opmerkingen geplaatst worden. enige wiskunde conventies waar ik me (tenzij anders vermeldt) aan zal houden:

symbool	betekenis
f_z	de zendfrequentie
f_a	de frequentie van een audiosignaal
T	de periodetijd van een signaal
$x(t)$	het HF signaal
$a(t)$	het audio signaal
$I(t)$	het I-signaal (in fase)
$Q(t)$	het Q-signaal (90 graden gedraaid)
$A(t)$	de amplitude van een signaal
$\hat{y}_{sin}(s)$	de fouriergetransformeerde van $y(t)$ naar s
$\hat{y}_{cos}(s)$	

2 Algemene wiskundige inleiding

2.1 De fourierreeks

In vele gevallen is het nodig om te kijken naar welke frequenties allemaal in een signaal aanwezig zijn. Denk alleen maar aan het antennesignaal, waar alle informatie van een hele hoop zenders in zit, maar waar je precies het goede stukje spectrum uit moet filteren. Daarnaast heb je vaak te maken met de vraag wat het frequentiespectrum van het een zender precies inhoudt. Hoe je dit zo goed mogelijk de lucht in krijgt, met zo min mogelijk storing op andere frequenties.

Maar wat is dat frequentiespectrum eigenlijk? Meneer Fourier en zijn nazaten hebben ons geleerd dat alle periodieke stuksgewijs continue functies zijn om te schrijven in een zogenaamde Fourierreeks. Niet periodieke stuksgewijs continue functies kun je omzetten door middel van een fouriertransformatie.

Wat stuksgewijs continu inhoud is onbelangrijk, alle elektrische signalen zijn het. Periodieke signalen zijn signalen die zich steeds herhalen, met een periode van T seconden. Voorbeelden zijn de slingerbeweging van een klok, de trilling van een pianosnaar enz.

Niet periodieke signalen zijn signalen die steeds veranderen, en waar geen steeds terugkomend element in zit. In de praktijk zijn alle elektrische signalen niet periodiek. Neem de vioolsnaar weer, eigenlijk is dat helemaal geen periodieke beweging, omdat de snaar langzaam steeds zachter gaat trillen. Toch is het meestal het handigst om met periodieke functies te werken, omdat je daar wiskundig een stuk meer mee kan.

Ik zal beginnen met het uitleggen wat een fourierreeks is, van daaruit is de overstap naar de fouriertransformatie vrij natuurlijk.

Het maken van een Fourierreeks is eigenlijk een methode om te bekijken hoeveel signaal van een bepaalde frequentie er aanwezig is in het totale signaal. Allereerst stel ik dat ik ieder T -periodiek signaal $a(t)$, kan schrijven als een som van sinussen en cosinussen op frequentie n/T met n een geheel getal. (dus de zogenaamde harmonischen, met een frequentie n maal de grond frequentie, met $n = 0, 1, 2, 3, \dots$) Best een sterke bewering. Maar stel nou dat er andere frequentiecomponenten in zitten, met frequentie $f_{anders} = 1/T_{anders}$. In dat geval zou de het signaal zich niet meer herhalen met periode T , maar met periode $T' = TT_{anders}$. En dus was het signaal al helemaal niet T -periodiek!. Alle frequentie componenten moeten dus wel een frequentie hebben van grondtoon $1/T$, of n maal die frequentie.

Dit verleidt mij ertoe $a(t)$ te schrijven als een som van sinussen en cosinussen:

$$a(t) = \sum_{n=0}^{n=\infty} c_n \sin \frac{2\pi n}{T} t + d_n \cos \frac{2\pi n}{T} t \quad (1)$$

Nu moeten alleen nog de coëfficienten c_n en d_n berekend worden. Dat klinkt eenvoudiger dan het is. Als het goed is is het volgende al bekend (met $m \neq n$):

$$\int_{t=0}^{t=T} \sin \frac{2\pi n t}{T} \sin \frac{2\pi m t}{T} dt = \frac{T}{2}$$
$$\int_{t=0}^{t=T} \sin \frac{2\pi n t}{T} \cos \frac{2\pi m t}{T} dt = 0$$

$$\int_{t=0}^{t=T} \sin \frac{2\pi nt}{T} \sin \frac{2\pi mt}{T} dt = 0$$

$$\int_{t=0}^{t=T} \sin \frac{2\pi nt}{T} \cos \frac{2\pi mt}{T} dt = 0$$

Dit blijkt de sleutel tot het probleem te zijn. De gewone functie $a(t)$ is bekend. Het is dus mogelijk om die functie te vermenigvuldigen met $\sin \frac{2\pi nt}{T}$ en daarna te integreren van $t = 0$ tot $t = T$. Invullen van (1), en daarna het laten wegvallen van bijna alle termen geeft de uitkomst¹:

$$\int_{t=0}^{t=T} \sin \frac{2\pi nt}{T} a(t) dt = \int_{t=0}^{t=T} \sin \frac{2\pi nt}{T} \sum_{m=0}^{m=\infty} c_m \sin \frac{2\pi mt}{T} + d_m \cos \frac{2\pi mt}{T} dt$$

$$= \int_{t=0}^{t=T} \sin \frac{2\pi nt}{T} c_n \sin \frac{2\pi nt}{T} dt = c_n \frac{T}{2}$$

Hetzelfde valt natuurlijk ook voor de cosinus te doen.

Dat brengt ons tot de conclusie:

$$\begin{cases} \frac{2}{T} \int_{t=0}^{t=T} \sin \frac{2\pi nt}{T} a(t) dt = c_n \\ \frac{2}{T} \int_{t=0}^{t=T} \cos \frac{2\pi nt}{T} a(t) dt = d_n \end{cases} \quad (2)$$

Het is dus mogelijk om voor iedere periodieke functie $a(t)$ de coëfficiënten c_n en d_n op te stellen met (2). c_n en d_n kunnen we maken om er even mooi naar te kijken, maar in sommige gevallen kan je er ook een leuke bewerking op doen. Na zo'n bewerking kan je het dan weer naar een echt signaal omzetten met (1).

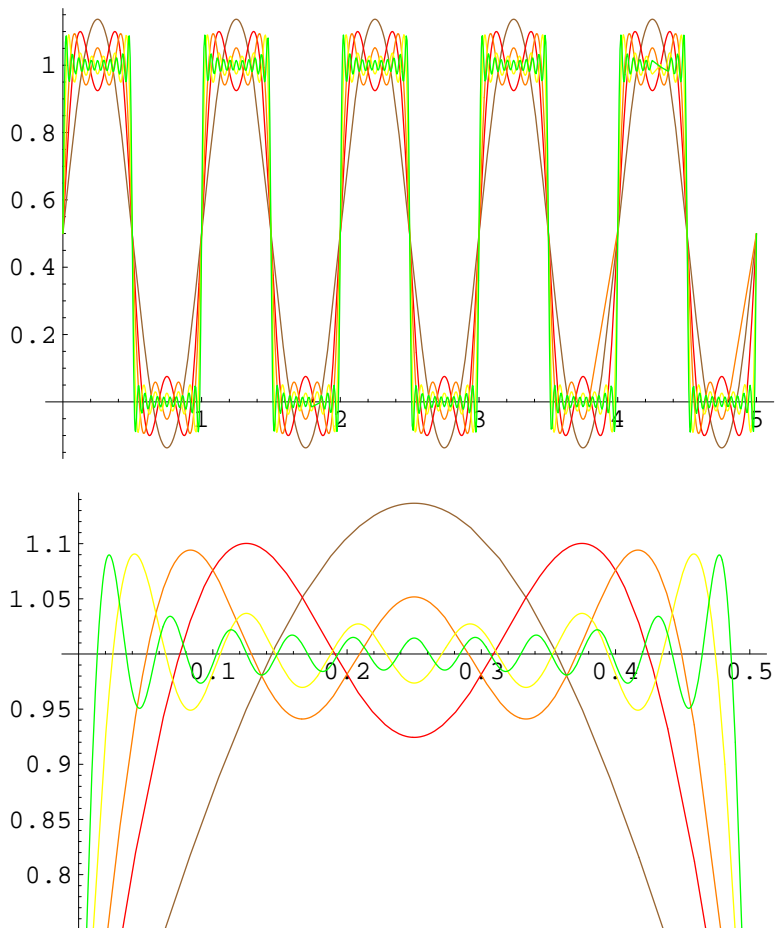
Als voorbeeld neem ik nu de $A(t)$ uit (figuur 3.2). De periode is $T = 1$. De integralen leveren op: $c_n = \frac{2}{\pi n}$ voor n oneven. $c_n = 0$ voor n even. $d_n = 0$ voor $n > 0$, en $d_0 = 0.5$. (zoals je misschien ziet: d_0 geeft de gemiddelde waarde van de $a(t)$. c_0 heeft verder geen betekenis.)

Nu is dit een oneindige reeks. Het is mogelijk om te kijken wat er gebeurt als we voor iedere t de hele reeks uit (1) te beschouwen. Maar dat is niet zo nuttig. Zoals je ziet, voor grote n wordt $c_n = \frac{2}{\pi n}$ steeds kleiner. Dus kijken we wat er gebeurt als we maar maar een paar termen van de reeks beschouwen. De plaatjes staan op de volgende bladzijde.

Het is te zien dat hoe meer termen er mee worden genomen, hoe beter het beeld. Bij een oscilloscoop kan je trouwens soms hetzelfde effect waarnemen, wanneer je blokgolven bekijkt met een frequentie in de buurt van de bandbreedte van de oscilloscoop. Omdat de harmonischen met een te hoge frequentie buiten de bandbreedte van de scoop vallen, zullen alleen de componenten met een lagere frequentie getoond worden. Dat levert dus een soortgelijk plaatje op.

In gevallen waarbij gewerkt wordt met signalen die plots verspringen lijkt het vaak alsof er altijd nog hele kleine 'ribbeltjes' blijven zitten op de plekken waar het signaal een sprong maakt. Het maakt dan niet veel uit of je 10 100 of 100 termen meeneemt, ze blijven er zitten, al worden ze hoogstens ietsje smaller. Die ribbels verwijnen echter alleen wanneer je ook echt de limiet berekend waarin het aantal termen in de som naar oneindig gaat. In het echt zijn zulke signalen ook niet mogelijk, en slechts een goede versimpeling van de werkelijkheid. Op vreemde plekken, zoals bij zo'n sprong, dan zie je dus ook dingen die in het echt niet zullen gebeuren omdat het signaal niet echt te produceren valt.

¹Goed moment om een kop koffie te pakken en te kijken of je het goed snapt



Figuur 1: Verschillende benaderingen van de blokgolf uit figuur1, bruin is de slechtste benadering, en groen de beste.

2.2 De Fouriertransformatie

Dan nu de fouriertransformatie. Stel we hebben een audiosignaal $a(t)$ van een nummer dat op repeat staat in de CD-speler. Dan is dit signaal wel degelijk periodiek. De periodetijd T is dan natuurlijk gelijk aan de lengte van het nummer. Het is dus mogelijk om er een fourierreeks van te maken. Gaan we echter kijken naar het resultaat, dan wordt duidelijk dat in de termen van (1) $\sin \frac{2\pi n}{T}$ de frequentiestapjes wel heel erg klein worden (ze worden $\frac{2\pi}{T}$). Er zullen dus ontzettend veel integralen uitgerekend moeten worden, en je zal ontzettend lange rijen c_n en d_n krijgen. Het is niet meer praktisch om dat in een reeksnotatie te houden. Het is veel handiger om een functie te zoeken met als variabele de frequentie. Dit noemen we de fouriergetransformeerde van $a(t)$. En het wordt op de volgende manier genoteerd c_n is vervangen door $\hat{a}_{sin}(s)$; en d_n is vervangen door $\hat{a}_{cos}(s)$. De variabele s is de frequentie, net zoals dat n eerst recht

evenredig was met de frequentie.:

$$a(t) = \int_{s=0}^{s=\infty} \hat{a}_{sin}(s) \sin ts + \hat{a}_{cos}(s) \cos ts \, ds \quad (3)$$

Waarbij $\hat{a}_{sin}(s)$ en $\hat{a}_{cos}(s)$ gegeven zijn door:

$$\begin{cases} \hat{a}_{sin}(s) = \int_{-\infty}^{\infty} a(t) \sin 2\pi st \, dt \\ \hat{a}_{cos}(s) = \int_{-\infty}^{\infty} a(t) \cos 2\pi st \, dt \end{cases} \quad (4)$$

Zoals je misschien al is opgevallen wordt een fouriergetransformeerde van een functie meestal aangegeven met een dakje. (in de meeste literatuur wordt de fouriergetransformeerde gedefinieerd met complexe getallen wat het mogelijk maakt om alles in 1 functie te schuiven. Om het simpel te houden gebruik ik een losse functie voor de sinussen, en een losse voor de cosinussen. Dit is echter wel een stuk minder elegant qua notatie. Qua rekenwerk is het echter eenvoudiger, omdat we de complexe getallen toch niet veel zullen gebruiken)

Een belangrijk punt om op te letten bij het werken met fouriergetransformeerden is dat het geen zin heeft om de fouriergetransformeerde van een oneindig lang signaal te nemen. Daar komt in de regel namelijk oneindig uit. Je kan het alleen doen bij 'absoluut integreerbare' functies². Meestal zal je de integraal echter niet van - oneindig tot oneindig laten lopen, maar bijvoorbeeld van 0 tot 10 periodes. (dat is hetzelfde als te zeggen dat het hele signaal slechts 10 periodes duurt) Dan kom je onder dat probleem uit. De enige fout die je dan krijgt op je berekende resultaten is dat de pieken in het signaal niet meer oneindig smal en oneindig sterk zijn, maar redelijk smal, en redelijk sterk. Duidelijk genoeg in ieder geval om conclusies te trekken, en het is mogelijk om het in plaatjes te zetten. En daar was het me eigenlijk om te doen.

Willen we weten wat de wat de intensiteit I_n of $I(s)$ van het signaal voor een bepaalde harmonische n , of frequentie s is, dan berekenen we die met pythagoras op de volgende wijze:

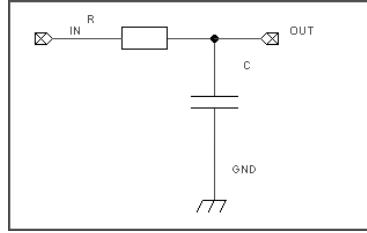
$$\begin{aligned} I_n &= \sqrt{c_n^2 + d_n^2} \\ I(s) &= \sqrt{\hat{a}_{sin}^2(s) + \hat{a}_{cos}^2(s)} \end{aligned} \quad (5)$$

2.3 voorbeeld: Het RC filter

Om een idee te geven hoe er nou mee gewerkt kan worden, zullen we het RC-filter beschouwen. Het enige wat we hoeven te weten is hoe zo'n ding reageert op een losse sinusspanningen van frequentie f . Dat is voldoende om daaruit het gedrag voor alle andere signaalvormen af te leiden. We zullen het uiteindelijk toepassen op de blokgolf van hierboven. Stel we hebben een RC-filter, zoals hierboven. Ik ben alleen genteresseerd in wat er gebeurt met een sinusspanning, met frequentie f ; dus ik stel $V_{in} = \sin ft$. De stroom door de weerstand is:

$$I = \frac{V_{in} - V_{out}}{R} = C \frac{dV_{out}}{dt}$$

²een functie waar de integraal van de absolute waarde, van -oneindig tot oneindig bestaat. Ik ga hier niet behandelen in welke gevallen dat kan, want dit artikel was tenslotte bedoeld voor radio-genteresseerden.



Figuur 2: RC-filter

Hieruit destilleren we dan de differentiaalvergelijking:

$$\frac{dV_{out}}{dt} = \frac{\sin ft - V_{out}}{RC}$$

In principe valt dit heel eenvoudig op te lossen door direct met de fourierge-transformeerde van \hat{V}_{out} te werken³. Het kan echter ook nog betrekkelijk simpel door het toepassen van enkele trucjes, en dat scheelt weer een hele berg theorie. Het kan in principe ook op de manier waarop je dergelijke problemen normaal oplost, maar de trucjes maken het wel veel eenvoudiger.

De belangrijkste truc om dit op te lossen is bedenken wat voor vorm V_{out} zou kunnen hebben. En dat is eigenlijk vrij simpel te zien; het moet wel een sinus of cosinus met dezelfde frequentie worden. Die oplossing stel ik dus voor: $V_{out} = A \sin ft + B \cos ft$. Nu moet dit ingevuld worden in de differentiaalvergelijking:

$$\frac{dV_{out}}{dt} = Af \cos ft - Bf \sin ft = \frac{\sin ft - A \sin ft - B \cos ft}{RC}$$

Dit stelt ons in staat om A en B te berekenen. Want aan beide kanten moeten gelijke getallen bij de sinus en cosinusstermen staan.

$$\begin{aligned} Af &= \frac{-b}{RC} \\ Bf &= \frac{A}{RC} - \frac{1}{RC} \\ A &= \frac{B}{fRC} \\ B &= \frac{A}{fRC} - \frac{1}{fRC} = \frac{B}{f^2 R^2 C^2} - \frac{1}{fRC} \end{aligned}$$

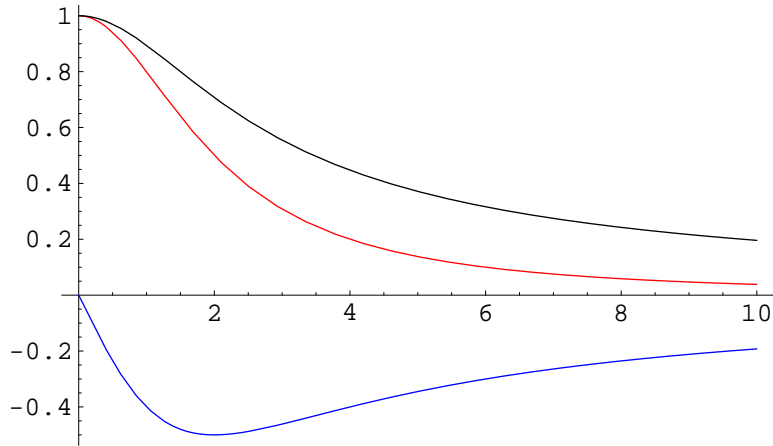
En hieruit kunnen we de oplossing bepalen, namelijk:

$$\begin{aligned} A &= \frac{1}{1+f^2 R^2 C^2} \\ B &= \frac{-fRC}{1+f^2 R^2 C^2} \end{aligned} \quad (6)$$

We weten nu dus wat er gebeurd met eender welke sinus die je op de ingang zet (en dus ook eender welke cosinus). Omdat dit een lineair systeem⁴ is, geldt dit ook voor de som van enkele sinussen en cosinussen. En daar zit de kneep, we kunnen het ingangssignaal door een fourierreeks, of transformatie omschrijven

³Voor de geïnteresseerden, koop een cursus fouriertheorie / signaalanalyse

⁴Een lineair systeem is een systeem waarbij alles met lineaire formules aan elkaar hangt. Alles met slechts spoelen, weerstanden en condensatoren is lineair. Zelfs met transistoren erbij is het tot op zekere hoogte lineair, maar dan moet je gaan uitkijken, want de vervorming is niet-lineair. Komen er diodes aan te pas, dan is het zeker NIET-lineair, en is deze methode ook niet verder bruikbaar.



Figuur 3: Rood is de sterkte van de sinus in V_{out} ,blauw van de cos, en zwart is de totale sterkte

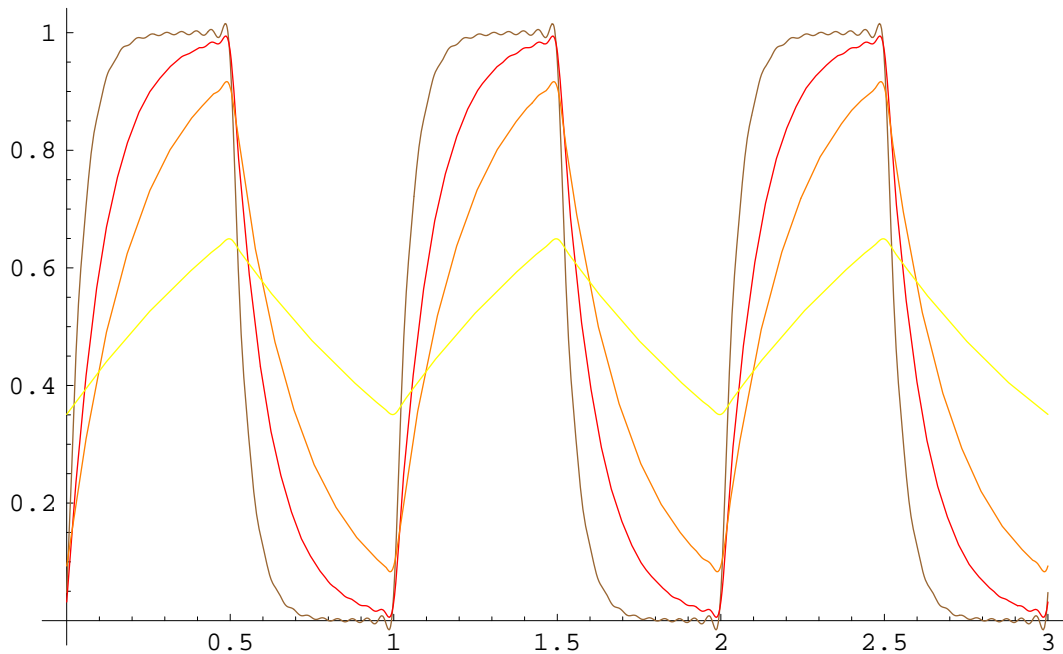
naar een som of integraal van sinussen en cosinussen, waarvan we het gedrag kennen. Iedere coëfficiënt in de fourierreeks van het ingangssignaal wordt door het RC filter volgens (6) omgezet. Voor de goede orde: frequentie f van iedere coëfficiënt is: $f = n/T$. En omdat we dit gaan invullen bij de blokgolf van een eindje terug vinden we:

$$c_{n,out} = \frac{c_{n,in}}{1 + \frac{n^2}{T^2} R^2 C^2} - \frac{-d_{n,in} \frac{n}{T} RC}{1 + \frac{n^2}{T^2} R^2 C^2} = \frac{2}{\pi n (1 + \frac{n^2}{T^2} R^2 C^2)}$$

$$d_{n,out} = \frac{c_{n,in} \frac{n}{T} RC}{1 + \frac{n^2}{T^2} R^2 C^2} + \frac{d_{n,in}}{1 + \frac{n^2}{T^2} R^2 C^2} = \frac{2 \frac{n}{T} RC}{\pi n (1 + \frac{n^2}{T^2} R^2 C^2)} + 0.5$$

Let goed op hoe $d_{n,in}$ met een min-teken verschijnt bij $c_{n,out}$. Dit komt omdat $\pi/2$ graden verder in fase dan de cosinus een min-sinus geeft, en hij moet er dus van afgetrokken worden. Let er ook even op dat dit geldt wanneer n oneven of 0 is; alle andere termen van $c_{n,in}$ en $d_{n,in}$ zijn nul, en verschijnen dus niet aan de uitgang.

Dat levert dan inderdaad precies het gewenste resultaat, wat geplotted staat in de plaatjes die volgen.



Figuur 4: Dit zijn de resultaten van een benadering tot $n = 21$ van de gemaakte fourrierreeks. Bruin is bij de kleinste RC , en geel bij de grootste.

3 CW

3.1 algemene informatie

De eenvoudigste vorm van modulatie is veruit CW (continuous wave). Dit is de methode die meestal voor morse wordt gebruikt. CW wil zeggen dat er een sinusgolf wordt opgewekt die aan of uitgezet kan worden. Meer niet.

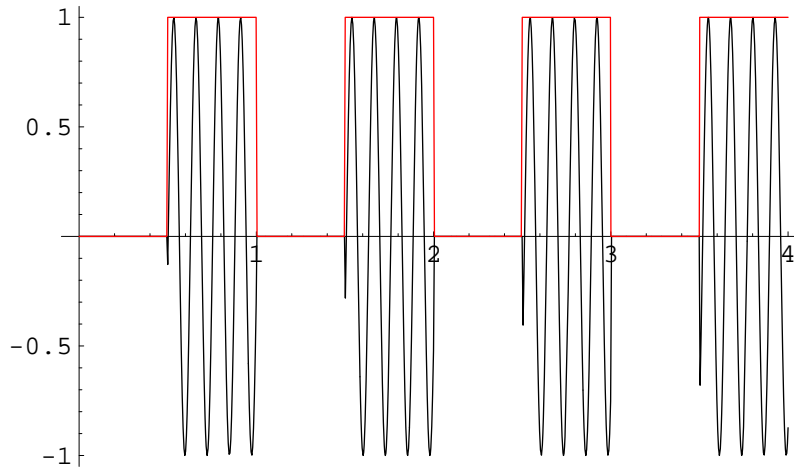
Omdat het de simpelste vorm van modulatie is, was dit ook de eerste eerste vorm van modulatie die werd gebruikt. Het is alleen geschikt om morsecode en dergelijken over te brengen. De bekende piepjes. Ondanks dat het zeer ouderwets is, heeft het toch wel voordelen. Het is namelijk wel een modulatiemethode die erg ver rijkt, vergeleken met de andere methoden. Onder zendamateurs wordt het nog steeds vrij veel gebruikt op de kortegolf. Al zal je het misschien niet altijd horen op een gewone radio, omdat het op een gewone AM-ontvanger niet goed hoorbaar is.

3.2 wiskundige representatie

Zoals gezegd: bij CW-modulatie is er slechts een sinusgolf die onderbroken wordt. Stel dat $A(t)$ 1 is indien er sinus moet zijn op tijdstip t , en 0 is indien

er geen signaal moet zijn (d.w.z. 1 bij piep, en 0 bij stilte). Dan komt in de simpelste benadering het signaal er dus als volgt uit te zien:

$$f(t) = A(t) \sin ft \quad (7)$$



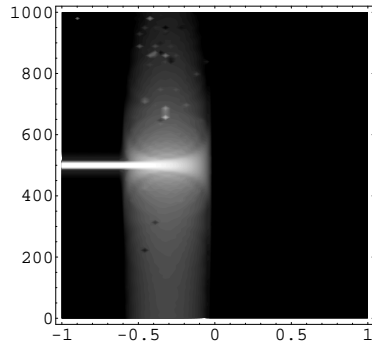
Figuur 5: Het CW signaal (zwart), en $A(t)$ rood. De zendfrequentie is VEEL lager gekozen dan realistisch.

Er is hier echter een probleem. Het lijkt heel netjes te zijn. Maar ga je de fouriergetransformeerde hiervan bekijken, dan zie je dat er toch nog wel wat op aan te merken is. Nou is dat op zich nog niet zo heel erg duidelijk te zien op het plaatje dat ik gemaakt heb, maar je moet er rekening mee houden dat in de praktijk een signaal dat bijvoorbeeld $60dB$ zwakker is dan de draaggolf toch nog erg hard is in de omgeving van de zender. En voor de burens die naar de radio proberen te luisteren is het erg vervelend. Iedere keer dat je je zender in en uitschakelt horen ze een klikje uit hun luidsprekers.

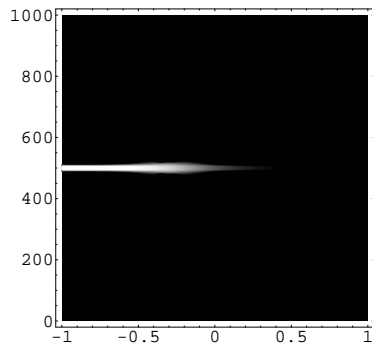
Gelukkig is het niet zo heel lastig dit te voorkomen. Het enige wat gedaan moet worden is de zender rustig aan en uitschakelen. (immers, iedere abrupte overgang veroorzaakt hoge frequentiecomponenten) Doordat er in de meeste gevallen toch genoeg LC-kringen in de zender zitten is het over het algemeen helemaal niet zo'n groot probleem om het rustig uit te schakelen. De filters in de buffers en eindtrap, en eventueel de antenne zelf blijven nog wel even resoneren, zodat het probleem zo goed als wegvalt. Maar ga dus niet een zender moduleren door de eindtrap kort te sluiten, tenzij je goed weet wat en hoe je het doet.

3.3 modulatie

Over de modulatie voor een CW zender valt eigenlijk niet zo heel veel te zeggen. Het is dan ook vrij eenvoudig. Over het algemeen zal bij een CW zender een oscillator gebruikt worden die continu aan staat. Er zal dan een buffertrap of een eindtrap in en uitgeschakeld worden. Het is NIET aan te raden om de oscillator aan en uit te schakelen. De reden hiervoor is dat de meeste oscillatoren tijdens



Figuur 6: Fouriergetransformeerde van een abrupt afgebroken draaggolf: Rechts zie je een draaggolf van 500Hz , die op $t = 0$ abrupt verbroken wordt. Je ziet duidelijk een enorme verbreding van het signaal. De paar vlekjes zijn afkomstig uit onvolkomenheden in mijn berekening.

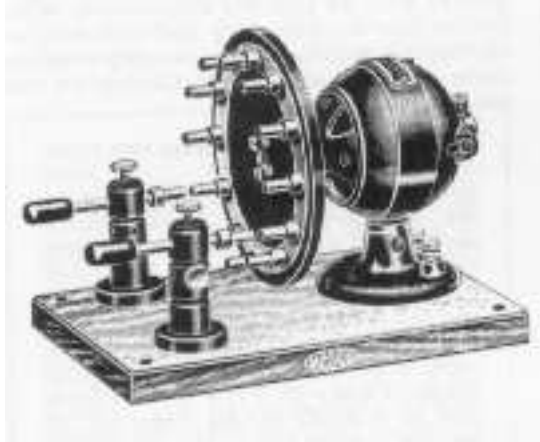


Figuur 7: Fouriergetransformeerde van een rustig afgebroken draaggolf. Dit geeft stukken minder storing. Hier is het effect extra dramatisch, omdat de draaggolf erg rustig werd afgebroken.

het opstarten niet stabiel zijn qua frequentie. Dit zorgt dus voor ongewenste storing op naburige frequenties. Met PLL-gestabiliseerde oscillatoren is het natuurlijk helemaal not-done om hem uit te schakelen.

De enige opvallende schakeling die ik ooit ben tegengekomen (en me kan herinneren) voor de modulatie van CW was een trafo via dewelke het signaal naar de antenne werd gevoerd. Om deze trafo zat een extra wikkeling waar een gelijkstroom doorheen gestuurd kon worden. Door een flinke gelijkstroom door deze wikkeling te sturen raakt de kern van een trafo verzadigd. Een verzadigde kern zorgt ervoor dat er amper meer HF doorgegeven kan worden. Op deze manier kreeg de antenne dus geen vermogen meer toegevoerd. Ik meen dat dit principe toegepast werd in de tijden van vonkzenders. Het was in die tijd niet handig om het enorme vermogen dat geleverd werd door de generator zo plots in en uit te schakelen. Dat zou een zware belasting zijn voor het mechaniek. Een zender was in die tijd immers meer mechanica dan elektronica. Vaak bestond het uit een zware benzinemotor van al gauw enkele tientallen PK's (dat was in 1915

toch al een flink apparaat), met daaraan vast een generator die een spanning van enkele kilovolts gaf. Die hoge spanning werd vaak naar een (roterende) vonkbrug gevoerd. Het signaal dat eruit kwam werd dan via een LC-kring, resonierend op de grondfrequentie of een harmonische van de vonkbrugfrequentie afgestemd. Het signaal over die LC-kring werd daarna dan op de antenne gezet.



Figuur 8: roterende vonkbrug, anno 1915

3.4 demodulatie

Een standaard 'huis tuin en keukenontvanger' is niet goed in staat om CW weer te geven. Dit komt omdat deze ontvangers alleen reageren op de amplitude van een draaggolf⁵. Je hoort uit de luidspreker een tik en een beetje ruis wanneer er signaal is, en niets wanneer er geen signaal is. Maar niet de duidelijke piepjes.

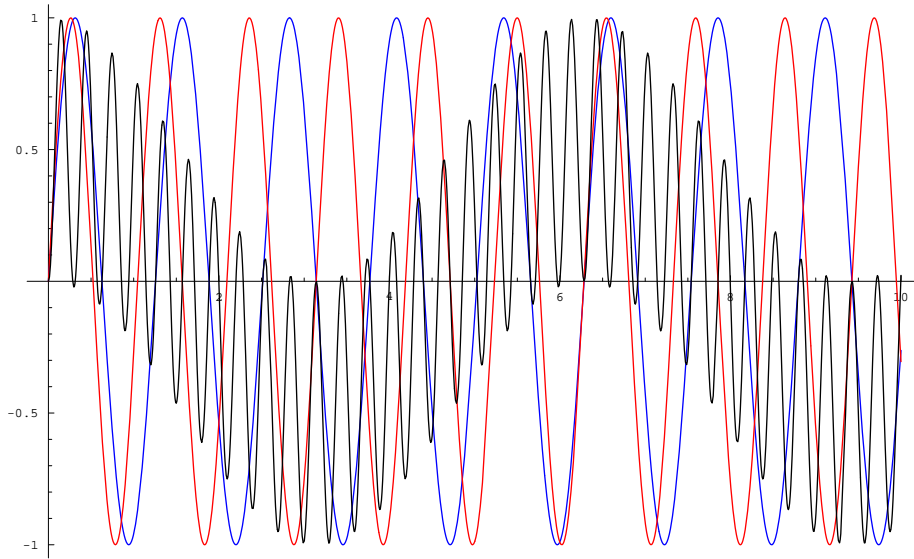
Om dat toch te krijgen moet het CW signaal (frequentie f_z gemixt worden met een signaal van een iets hogere of lagere frequentie, f_m . De verschilfrequentie hoor je dan uit de luidspreker. Een ideale mixer is een ding dat de spanning van twee signalen met elkaar vermenigvuldigt.

$$\sin f_z t \sin f_m t = \frac{1}{2} \cos(f_z - f_m)t - \frac{1}{2} \cos(f_z + f_m)t \quad (8)$$

Stel dat f_m precies $1kHz$ verschilt met f_z dan horen we netjes een $1kHz$ pieptoontje uit de luidspreker. De andere term in de vergelijking levert een onhoorbaar hoge toon op, waar we verder geen last mee hebben. Op deze manier is het dus mogelijk om een losse draaggolf hoorbaar te maken. Eigenlijk is dit bijna de enige methode die gebruikt wordt.

Nou klinkt het alsof het dus vrij moeilijk is om een CW-signaal te demoduleren. In de praktijk volstaat het al om een gewone AM-ontvanger te nemen, en hier een extra oscillator met een frequentie in de buurt van de middenfrequentie van de ontvanger aan toe te voegen. (meestal ongeveer $455kHz$). Dit signaal moet dan 'ergens' in de middenfrequent-trappen toegevoegd worden. Het mixen van het signaal hoef je niet apart te doen. De diode-detector die in de meeste

⁵zie hoofdstuk AM-ontvangst



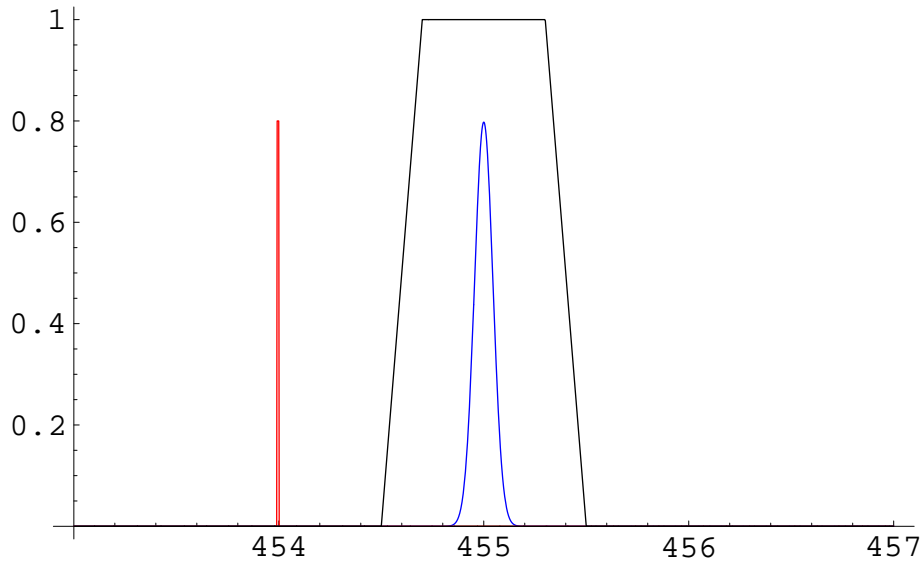
Figuur 9: Rood is f_m , blauw is f_z en zwart is het product van beiden, waarin duidelijk het audiosignaal met lagere frequentie zichtbaar is. (voor de duidelijkheid, de frequenties zijn weer VEEL lager dan realistisch gekozen)

AM-ontvangers aanwezig is zal deze 2 signalen wel mixen. Een diode-detector heeft namelijk de eigenschap ALLE signalen die hij binnenkrijgt met elkaar te mixen. Let op, er is een groot verschil tussen het toevoegen van 2 signalen aan elkaar, en het mixen. De ene komt overeen met het optellen van de 2 spanning (misschien met verzwakking, maar dat maakt niet uit). Mixen komt overeen met het vermenigvuldigen van 2 spanningen.

De simpelste methode is nog wel een mini-zendertje te maken rond de zendfrequentie van de CW-zender. Dit signaal zal ook gewoon in de ontvanger terechtkomen, met hetzelfde effect als hierboven. Er kleven echter te veel nadelen aan om dit serieus te nemen.

In een serieuze ontvanger wordt echter een extra smal bandfilter gebruikt, met een bandbreedte van 1kHz of smaller. Je stemt de ontvanger dan af zodat de zender precies binnen het filter valt, maar de rest erbuiten. (alles wat je extra ontvangt is ruis en ander gepiep, wat niet de bedoeling is). Door nu het oscillatorsignaal pas na het filter in te voegen (wat het meest gebruikelijk is), kunnen toch nog pieptonen van een normale frequentie bereikt worden. Meestal wordt er bij een dergelijk ontvanger wel een iets geavanceerder mixer gebruikt dan de simpele diodedetector. Ook al zou dit in principe al wel genoeg zijn, de diodedetector zal ook alle signalen die door het filter heen komen onderling mixen, wat tot enige problemen kan leiden. Een dubbelgebalanceerde mixer is wat dat betreft de beste oplossing.

In de beginjaren werden veelal mechanische mixers gebruikt. In dit geval



Figuur 10: Zwart is het bandfilter, op de gebruikelijke frequentie van 455kHz , en met een breedte van 600Hz . Het te ontvangen signaal is blauw. (dit heeft een kleine bandbreedte, zoals afgebeeld), en het oscillatorsignaal is rood. (een ideale oscillator heeft in principe een bandbreedte 0, in de praktijk is het gewoon heel erg smal). De verschilfrequentie is hier dus zo'n 1kHz . (het bandfilter laat hier 700Hz tot 1.3kHz toe, de rest wordt weggefilterd. Dus ook als er nog andere zenders in de buurt zitten, er zullen alleen frequenties van 700Hz tot 1.3kHz uit de luidspreker komen!)

bestond de mixer uit een snel ronddraaiende schijf met contacten erop aangebracht. Die schijf was zodanig, dat er om en om twee draden recht werden doorverbonden, of juist gekruist werden doorverbonden. Aan de ene kant werd het (gefilterde) antenne-sigitaal gezet, aan de andere kant een koptelefoon. De mechanische mixers werden echter al snel na de invoering van de elektronenbuis afgeschreven.

Dit proces komt wiskundig gezien overeen met het vermenigvuldigen met een blokgolf met amplitude 1. En omdat er in zo'n blokgolf ook een sinus op de grondfrequentie aanwezig is, werkt dit dus als een mixer. Omdat er in een blokgolf ook sinussen op 3 maal de grondgolf aanwezig zijn, is deze detector ook gevoelig voor signalen op een 3 maal zo hoge frequentie. (in de praktijk geen enkel probleem)

4 Amplitudemodulatie

4.1 algemene informatie

Er is lang gezocht naar een onderdeel dat in staat was om een elektrisch signaal te versterken. Pas rond 1920 kwamen de eerste elektronenbuizen in de handel. Er is daarvoor al wel gezocht naar een dergelijk onderdeel. Er werd bijvoorbeeld geëxperimenteerd met grote bakken koolstofkorrels met een luidsprekertje op de

bodem. Dat werkte volgens hetzelfde principe als de koolstofmicrofoon in een telefoontoestel. Maar de elektronenbuis was voor lange tijd eigenlijk het enige onderdeel dat ook echt goed en betrouwbaar werkte.

Met de introductie van de elektronenbuis werd het mogelijk om een echte oscillator te bouwen. Daarvoor gebruikte men niet al te efficiënte vonkzenders. Met een elektronenbuis is het mogelijk om een echte oscillator te maken. Het grote voordeel hiervan is dat er een veel beter signaal uit komt, en dat het ook op veel hogere frequenties kan werken dan een mechanische zender. Door deze hogere frequenties waren ook grotere afstanden mogelijk. Al duurde het wel een tijdje voordat men dit doorkreeg.

Met introductie van de elektronenbuis en de diode werd het ook mogelijk om amplitudemodulatie toe te passen. Voor het eerst werd er muziek en spraak uitgezonden in de aether. Dit was min of meer het startschot van een enorme bedrijvigheid op radiogebied. Overal gingen zendamateurs aan de gang, en het aantal omroepen en zenders steeg explosief. Was in 1920 een radio nog een bijzonder technisch ding, alleen weggelegd voor grote bedrijven, telecommatenschappen en een paar welgestelde radio-amateurs. In 1935 was de radio al heel normaal. Mede dankzij projecten als de volksempfanger in Duitsland was het een normaal gebruiksvoorwerp geworden.

Tegenwoordig is voor omroepzenders de AM-band natuurlijk niet meer zoveel gebruikt. Maar zeker op de kortegolf valt er nog een hoop AM te luisteren. En de kortegolfband is toch de band bij uitstek voor amateurs. Verder zijn modulatiemethodes als dubbel- en enkelzijband sterk gerelateerd aan amplitudemodulatie. Het is dus nog zeker belangrijk.

4.2 wiskundige representatie

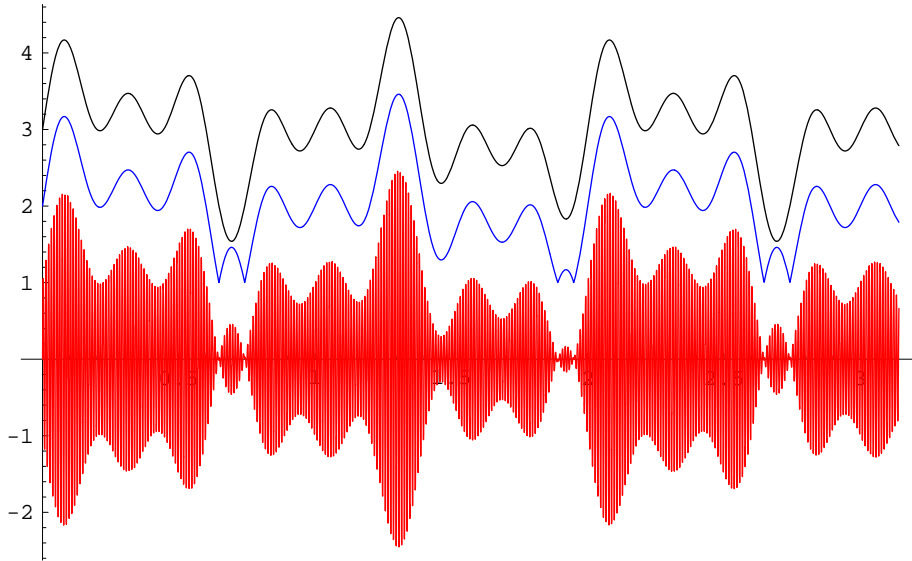
Een amplitude gemoduleerd signaal is draaggolf $x(t)$, waarvan de amplitude in grootte gevarieerd wordt met het audiosignaal $a(t)$. De frequentie is f_z . Dat levert dus het volgende signaal op:

$$x(t) = (1 + a(t)) \sin 2\pi f_z t \quad (9)$$

De term $1+$ zorgt ervoor dat er altijd een draaggolf is. Is deze term achterwege dan spreken we van dubbelzijbandmodulatie, maar daarover meer in het hoofdstuk enkelzijbandmodulatie. Omdat met een simpele diodedetector alleen de amplitude, en niet de fase van de draaggolf gemeten wordt, moet de die DC spanning erbij. Ik heb hier dus 1 genomen, maar dat is nogal afhankelijk van hoe groot het audiosignaal is. $1 + a(t)$ mag in dit geval niet kleiner dan nul worden, anders loopt het fout. Mocht het wel gebeuren, dan spreekt men van overmodulatie, en dit geeft veel vervorming.

Het hangt overigens sterk van het type modulator af wat er gebeurt bij overmodulatie. In de praktijk zal op het moment van overmodulatie meestal $x(t) = 0$ worden.

Het kan altijd interessant zijn om te zien wat het frequentiespectrum van een AM-gemoduleerde zender is. We kunnen hier in principe direct een fourierreeks van maken, maar er is een handig trucje om het direct in te zien. Stel het audiosignaal $a(t)$ bestaat slechts uit n frequentie: $a(t) = \sin 2\pi f_a t$. Ik reken het uit voor maar n frequentie, maar van daaruit valt het makkelijk te veralgemenen.



Figuur 11: Zwart is $a(t)$, het audiosignaal. Blauw is de amplitude van $x(t)$. Rood is het licht overgemoduleerde AM-signaal $x(t)$. Let op, op punten waar het audiosignaal te sterk negatief is, daar klapt blauw om, met sterke vervorming tot gevolg

meniseren.

$$x(t) = (1 + \sin 2\pi f_a t) \sin 2\pi f_z t = \sin 2\pi f_z t + \frac{1}{2} \cos 2\pi(f_z - f_a)t - \frac{1}{2} \cos 2\pi(f_z + f_a)t$$

En dan hetzelfde trucje met een cosinus: $a(t) = \cos 2\pi f_a t$

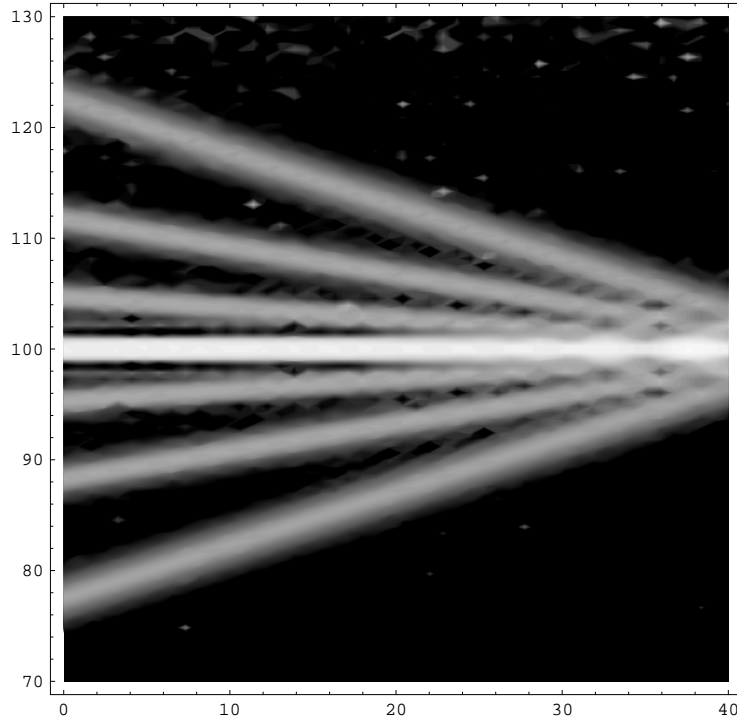
$$x(t) = (1 + \cos 2\pi f_a t) \sin 2\pi f_z t = \sin 2\pi f_z t + \frac{1}{2} \sin 2\pi(f_z - f_a)t + \frac{1}{2} \sin 2\pi(f_z + f_a)t$$

Je ziet dus dat een enkele toon wordt omgezet naar 2 signalen die symmetrisch rond de zendfrequentie f_z liggen op $f_z - f_a$ en $f_z + f_a$. Zouden we aan het audiosignaal meerdere sinussen of cosinussen toevoegen, dan zouden ze allemaal netjes op dezelfde manier terugkomen symmetrisch rond f_z . En dat is erg nuttig. Als we de fourierreeks, of fouriergetransformeerde van het audiosignaal weten, dan kunnen we die daarna gewoon symmetrisch rond f_z plakken. Neem nu voor $a(t)$ een algemeen audiosignaal, met $\hat{a}_{sin}(s)$ en $\hat{a}_{cos}(s)$ zijn fouriergetransformeerde. En gebruik makend van 3 krijgen we dan de volgende formule:

$$\begin{aligned} x(t) &= \sin 2\pi f_z t (1 + a(t)) = \sin 2\pi f_z t \left(1 + \int_{s=0}^{s=\infty} \hat{a}_{sin}(s) \sin ts + \hat{a}_{cos}(s) \cos ts \, ds \right) \\ &= \sin 2\pi f_z t + \int_{s=0}^{s=\infty} \hat{a}_{sin}(s) (\cos(f_z - s)t - \cos(f_z + s)t) + \hat{a}_{cos}(s) (\sin(f_z - s)t + \sin(f_z + s)t) \, ds \end{aligned}$$

Een feest van een formule, en ook niet zo heel erg praktisch. Het laat in ieder geval wel goed zien hoe het frequentiespectrum van het HF-signaal $x(t)$ eruit ziet. Zo is het bijvoorbeeld duidelijk dat de HF bandbreedte 2x zo breed is als de audiobandbreedte. Is namelijk de audiobandbreedte bijvoorbeeld 4.5kHz (zoals bij alle omroepers), dan wil dat zeggen dat $\hat{a}_{sin}(s)$ en $\hat{a}_{cos}(s)$ nul zijn voor $s > 4.5kHz$. Dus voor $s > 4.5kHz$, dan is het stuk binnen de integraal 0, en dus

ook de 4 termen die het HF-spectrum maken. En dus zal er ook alleen maar een HF-signaal zijn dat loopt van $f_z + s_{max} = f_z + 4.5kHz$ en $f_z - s_{max} = f_z - 4.5kHz$ (want $\hat{a}_{sin}(s)$ en $\hat{a}_{cos}(s)$ zijn natuurlijk tijdonafhankelijk). Dat levert dus een bandbreedte op van $9kHz$. 2 maal de bandbreedte van het audio signaal.



Figuur 12: Het spectrum van de toon $a(t)$ uit het vorige plaatje, waarvan de frequentie steeds kleiner wordt. Links is de frequentie nog vrij hoog, en dus is het spectrum (dat uit 3 bij elkaar gevoegde tonen bestaat) erg breed. Rechts is de frequentie van de toon ongeveer naar nul gedaald, en het spectrum is daar zeer klein.

Het laatste belangrijke punt aan de AM-modulatie is de zogenaamde modulatie diepte. Dit is toch te belangrijk om niet genoemd te hebben. De modulatie diepte m , meestal gemeten in procenten, geeft aan hoe sterk een signaal gemoduleerd is. Als A_{max} de maximum bereikte amplitude bij een uitzending is, en A_{min} het minimum, dan geldt:

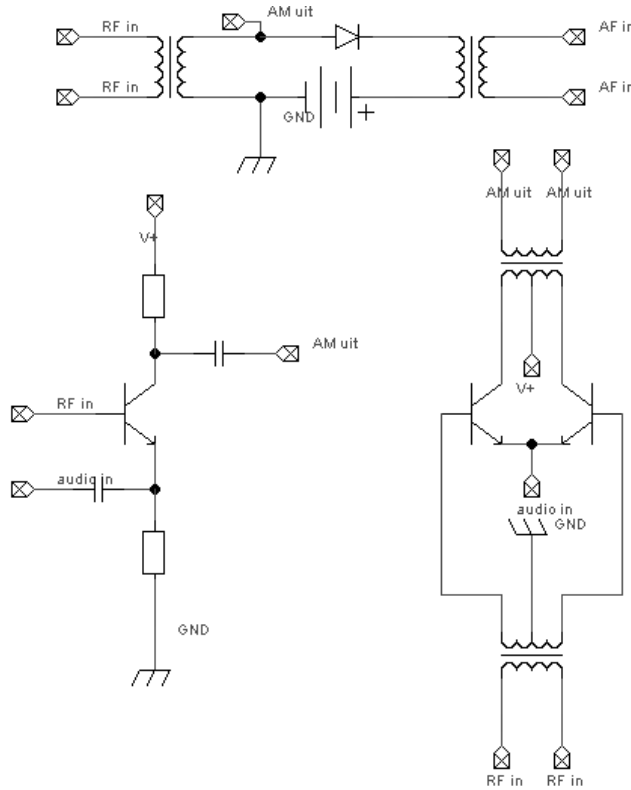
$$m = \frac{A_{max} - A_{min}}{A_{max} + A_{min}} \quad (10)$$

4.3 modulatie

Ook voor de modulatie van AM zijn er vele methoden mogelijk, van zeer simpel tot vrij ingewikkeld. De simpelste goed werkende methode is om een sinusspanning te maken op de frequentie waarop men wil zenden. Men kan de grootte van deze spanning dan door middel van een diode afkappen tot de amplitude die gewenst is. Denk eraan dat de simpelste methode, met een diode, een sterke

belasting aan de voorgaande trap teweeg brengt. Deze moet dus wel een voldoende hoge uitgangsweerstand hebben. (ook weet niet te hoog, want dan gaan de interne capaciteiten van de diode teveel een rol spelen)

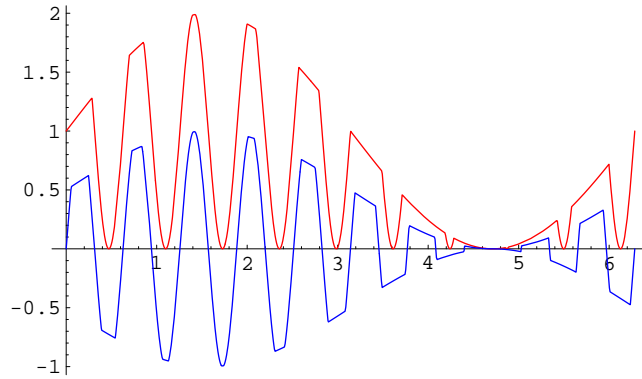
Min of meer hetzelfde trucje kan men ook doen met gebruik van transistoren. Dan wordt de basis-emitter diode gebruikt om hetzelfde effect te verkrijgen. Eventueel kan men twee transistoren of diodes gebalanceerd inzetten voor dezelfde taak. Door het gebalanceerd te doen valt het ook goed te gebruiken voor hogere vermogens.



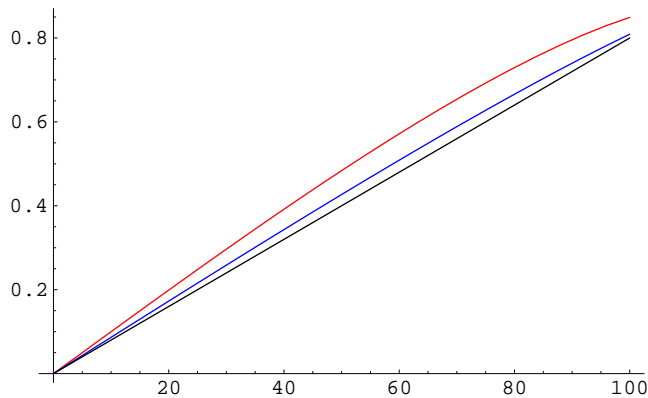
Figuur 13: Principeschema's van enkele simpele modulators. Boven: de simpelste diode mixer, de batterij is voor het instellen van de rust-amplitude. Links: een ongebalanceerde modulator met transistor. Rechts: een gebalanceerde modulator met transistoren.

Men dient er wel rekening mee te houden dat de ongebalanceerde methodes een zeer sterke vervorming van de sinusspanning geven. De gebalanceerde methode geeft minder vervorming, maar nog steeds moet men er rekening mee houden dat er geen zuivere sinus meer uit komt. Daarom zal men extra aandacht moeten schenken aan het filteren van ongewenste frequentiecomponenten op de uitgang van de modulator. (overigens moet bij iedere zender het uitgaande signaal gefilterd worden, dus zo heel erg is het niet). Het is echter niet alleen het HF-signaal dat vervormt. Ook het audiosignaal vervormt. (weer vooral bij de ongebalanceerde methode). Ongefilterd lijkt het audio signaal er netjes in

te zitten, als het signaal gefilterd wordt, dan blijkt er toch vervorming te zijn opgetreden. De componenten die het signaal zo perfect de goede vorm laten volgen blijken dus deels op een andere frequentie te zitten! Dit komt omdat een sterk afgeknotte sinus in verhouding meer amplitude op de grondfrequentie bezit dan een amper afgeknotte sinus. Het wordt duidelijk in de volgende twee plaatjes:



Figuur 14: Hier is een signaal van een HF signaal dat 100% gemoduleerd werd met een audiosignaal van $\frac{1}{10}$ van de HF frequentie. In het rood het signaal gegeven door een ongebalanceerde diodemixer, en in het blauw het signaal van een gebalanceerde diodemixer.



Figuur 15: In dit plaatje is op de X-as het aantal procent modulatie neergezet, en op de Y-as de hoeveelheid signaal dat na filtering aanwezig is. Rood is weer de ongebalanceerde mixer, en blauw de gebalanceerde.

Zelfs de gebalanceerde mixer heeft bij volledige modulatie nog wel enkele procenten vervorming. Met deze methodes valt dus geen echt perfect signaal te bereiken (al zal het voor de meesten ruimschoots voldoen). Het bereiken van een perfect signaal is erg lastig, zo niet onmogelijk. Men kan hiervoor

proberen om, net zoals bij audio, de amplitude van het signaal op de antenne te meten, en dit terugkoppelen naar de modulator. Net als een opamp doet met spanning. Eenvoudig is dit echter niet. Men moet erg rekening houden met de tegenkoppelfactor. Anders gaat het op andere frequenties genereren, met meestal een rookpluimpje tot gevolg, en anders wel een telecomautoriteit op de stoep. Ik beschouw het echter te ingewikkeld om hier veel aandacht aan te besteden.

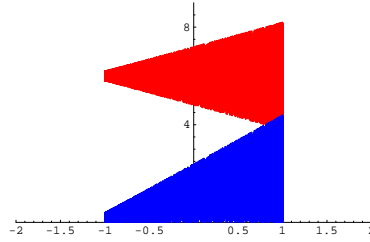
De dubbelgebalanceerde modulatoren zullen in 90% van de gevallen ook best voldoen. Het signaal van de dubbelgebalanceerde mixer valt ook nog sterk te verbeteren door de RF-trafo te laten resoneren op de ontvangsfrequentie. Dan moet je wel goed alle impedanties op elkaar afstemmen, omdat er anders erg veel signaal verloren gaat. Deze methode is wel wat moeilijker om toe te passen als je pas na de HF-eindtrap mixt, vanwege het slechtere rendement. Maar als je het mixt voor de eindtrap, dan heb je het probleem van de vervorming van HF-endiversterverker. Gelukkig zijn er hier en daar genoeg mooie schema's te vinden.

Een ander belangrijk punt is het meten. Een AM zender valt het best te meten met een scoop die een ADD en een XY-modus bevat. Het is ook mogelijk met een HF-probe in gewone modus, maar ik gebruik zelf liever de eerste methode, die ik dan ook zal behandelen. Je hebt een frequentiegenerator nodig en een scoop met een bandbreedte van minimaal de zendfrequentie.

De opstelling is de volgende:

- Zet de scoop op een sinus van bijvoorbeeld 300Hz . Te hoog en te laag kan fasedraaiingen geven die roet in het eten gooien.
- Sluit de frequentiegenerator aan op de zender, en op de Y-ingang van de oscilloscoop (die in XY-modus staat)
- Sluit de de zender aan op een dummyload, en meet met de X-ingang van de oscilloscoop het signaal over de dummyload.
- Zet de oscilloscoop in XY-modus. Als het goed is is er nu een meer of minder vervormde driehoek zichtbaar.
- Nu zet men de scoop op functie ADD, zodat signaal X en signaal Y opgeteld worden. Je kunt de signalen zodanig optellen dat de onderste of bovenste zijde van de driehoek precies horizontaal valt. Stel het zodanig in dat de driehoek overal zo goed mogelijk horizontaal ligt.
- lees af hoeveel V afwijking er is van de horizontale lijn (V_{dist}), en hoe groot het totale signaal is zonder modulatie V_{tot} . De vervorming is nu: $\frac{V_{dist}}{V_{tot}}$.

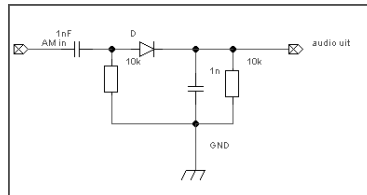
Als je op deze manier de vervorming meet dan is ook gelijk goed zichtbaar bij welke signaalgroottes de meeste vervorming optreedt, zodat je ook veel gerichter maatregelen kan nemen. In principe valt dit ook te doen met HF-probes, maar dan kan je niet meer tegelijkertijd kijken of het signaal bijvoorbeeld overal een goede sinusvorm heeft.



Figuur 16: Dit is het beeld dat je op de oscilloscoop zou moeten zien. Rood is zonder ADD, en blauw is met ADD. Dit is bij een modulatie van zo'n 85%

4.4 demodulatie

Eigenlijk is de demodulatie van een AM signaal erg simpel. Je kan met een diode simpelweg de amplitude van het signaal aftasten, en dat is genoeg. Er is op de

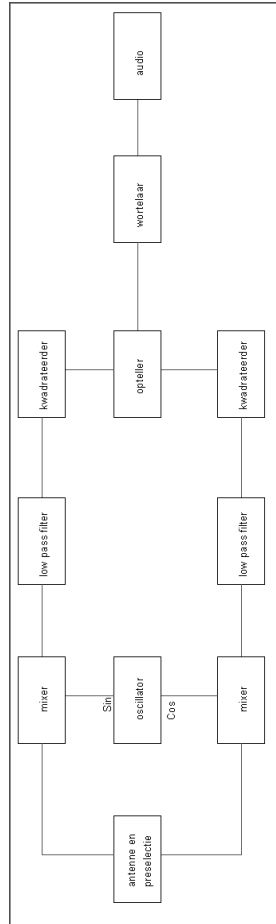


Figuur 17: simpele AM demodulator

uitgang nog een klein beetje HF-signaal aanwezig, (ondanks het RC-filter). Dat moet in de audiotrappen nog verder worden onderdrukt, omdat de eindtrap van de audioversterker hier niet altijd even goed op zal reageren. Maar omdat de frequentie van het HF signaal zo veel hoger is dan de frequentie van het audiosignaal is dit totaal geen probleem. Er zijn vele hiervan afgeleide schema's. In veel simpele eenkringsontvangers wordt in plaats van een diode een transistor gebruikt. De basis-emitter diode zorgt dan weer voor de niet-lineariteit die het signaal detecteert. Ook kan een dergelijke type detector met 2 diodes uitgevoerd worden, om de dubbele audiosterkte te verkrijgen. Allemaal kleine variaties op dezelfde methode. In principe is ieder onderdeel dat niet-lineair ⁶ is in staat tot het detecteren van een AM signaal.

Tegenwoordig is men in staat om AM ook op een andere manier te demoduleren, met een zogenaamde quadratuur-demodulator. Omdat tegenwoordig het liefst alles in een chip wordt gestopt, probeert men ook een 1-chips AM-ontvanger te maken. Dat betekent bijvoorbeeld dat er geen LC-kringen gebruikt kunnen worden om het antennesignaal te filteren. Aan de andere kant is het weer wel mogelijk om allerlei ingewikkelde berekeningen op het signaal toe te passen, die analoog vrij lastig zijn, en vervorming zouden geven. Ik zal het hier verder behandelen als een analoge schakeling, maar in principe kan de oscillator /footnoteDoor gebruik te maken van DDS, Digital Direct Synthesis. Dit is een methode waarbij een redelijk kloppende sinus en of cosinusspanning

⁶Alle onderdelen die geen zuivere weerstand capaciteit of inductie hebben, zoals allerhande halfgeleiders en buizen.



Figuur 18: prinsieschema quadratuurdemodulator

gemaakt kan worden door middel van chip die op een clockfrequentie werkt die slechts een paar maal hoger is dan de uitgangsfrequentie gedigitaliseerd worden, en ook alles dat na de mixers komt. Het is ondoenlijk om het antenne signaal digitaal te samplen, en daar dan nog eens allerlei berekening op los te laten. Voor een frequentie van $10MHz$ zou je een A-D convertor met een samplingsrate van $20MHz$ en een resolutie van minimaal $16bits$ nodig hebben. Dat is tot nu toe nog steeds vele malen duurder dan dat analoog te doen. Na het mixen is meestal een bandbreedte van $20kHz$ genoeg voor AM, en $300kHz$ genoeg voor FM. Beiden zijn goed haalbaar. Er zijn dan ook sinds kort complete 1-chipsontvangers op de markt die volgens dit principe werken.

Ik zal nu stuk voor stuk alle stappen in het prinsieschema afgaan.

Allereerst komen we de antenne en preselectiekring tegen. In principe zou er voor deze ontvanger geen preselectie nodig zijn. Maar de praktijk leert toch dat het ZEER nuttig is om dat wel te doen. Er komt anders ontzettend veel signaal op de mixers, wat voor veel kruismodulatie en andere problemen zorgt.

Daarna krijgen we een oscillator, en niet zomaar eentje, want hij moet een sinusspanning en een cosinusspanning leveren, liefst ook bij gelijke amplitude. Dit is analoog behoorlijk lastig. Maar het is wel goed mogelijk om een 4 maal hogere frequentie op te wekken, en dit dan in een paar digitale 2-delers omhoog delen naar de goede frequentie, met faseverschuiving. Let wel op dat er dan ook harmonischen van de grondfrequentie sterk aanwezig zijn in het signaal waarmee je mixt. De ontvang is dan dus ook gevoelig op 3, 5, 7, enz. maal de ontvangstfrequentie. (door toepassen van preselectie is dat geen groot probleem). Ook kan men met RC-filters de fase draaien, maar dat levert alleen geen gelijke amplitude op voor de sinus en cosinusspanning. Dat valt op te lossen door er een superheterodyne van te maken. Dan kan je mixen op de (vrij hoge) middenfrequentie (zie hoofdstuk superheterodyne)

Daarna wordt het antenne signaal in de twee mixers gemixt. Het is belangrijk dat dit met een gebalanceerde of dubbelgebalanceerde mixer gebeurt, omdat een simpele ongebalanceerde mixer alle signalen zal detecteren. Daardoor komen alle sterke zenders die binnen de preselectie vallen altijd als audio door. Niet heel sterk, wel goed hoorbaar. Ze worden dan gemixt met de draaggolven die door de zenders worden uitgezonden, in plaats van met het oscillatorsignaal. In (9) hebben we al gezien hoe het audiosignaal eruit ziet. (dit is de handigste representatie, de andere representatie, met fouriergetransformeerde moet ook werken, maar levert een wel erg ingewikkelde berekening op)

De mixer geeft dus de volgende signalen, als f_z de frequentie van de zender is, f_m de frequentie van de oscillator, en $S + Sa(t)$ de modulatie van de zender, waarbij S een factor voor de sterkte van de zender is.

$$\begin{aligned} I(t) &= (S + Sa(t)) \sin 2\pi f_z t \sin 2\pi f_m t \\ Q(t) &= (S + Sa(t)) \sin 2\pi f_z t \cos 2\pi f_m t \end{aligned}$$

$$\begin{aligned} I(t) &= \frac{1}{2}(S + Sa(t))(\cos 2\pi(f_z - f_m)t - \cos 2\pi(f_z + f_m)t) \\ Q(t) &= \frac{1}{2}(S + Sa(t))(\sin 2\pi(f_z - f_m)t + \sin 2\pi(f_z + f_m)t) \end{aligned}$$

Met het low-pass filter zal ik een beetje wiskundig sjoemelen om de berekening simpel te houden. In principe bestaat $a(t)$ ook uit vele frequentiecomponenten, die allemaal weer gemixt zijn met de cosinussen en sinussen die erachter staan. Ik ga er nu echter eventjes vanuit dat als de frequentie $(f_z - f_m)$ door het low-pass filter komt, dat dan het complete audiosignaal er ook doorheen komt. (in feite is de aanname dat het audiosignaal bandbreedte 0 heeft).

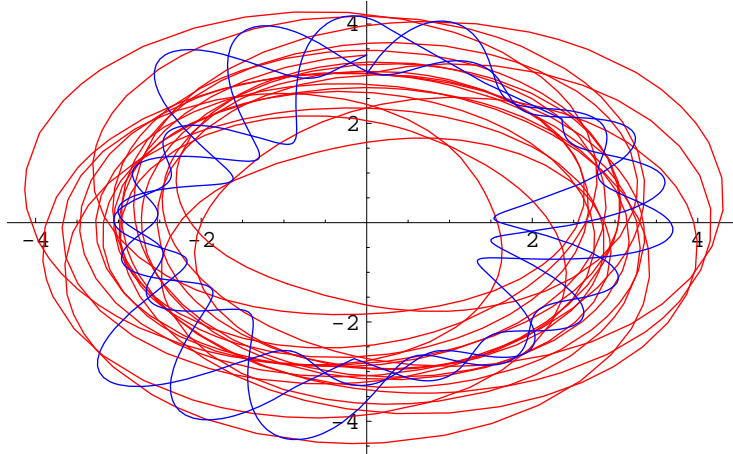
Het filter doet het volgende, het laat alleen frequentiecomponenten door met frequentie $f < f_f$ door. Overigens: f_f de afsnijfrequentie van het filter (in de praktijk bijvoorbeeld $6kHz$). Het is zo al te zien dat de enige component in de I en Q signalen die kans maakt om door te komen het signaal is met frequentie $f_z - f_m$. Dus alleen zenders waarvan de zendfrequentie f_z minder dan $6kHz$ verschilt met de mengefrequentie komen erdoor, de rest wordt allemaal geblokkeerd. We hebben hier dus een filter gemaakt dat ZONDER resonantiekringen toch bepaalde zenders eruit kan filteren. (digitaal is deze methode van filtering ook zeer eenvoudig).

Na filtering zien het I en Q signaal er dus als volgt uit:

$$I(t) = \frac{1}{2}(S + Sa(t)) \cos 2\pi(f_z - f_m)t$$

$$Q(t) = \frac{1}{2}(S + Sa(t)) \sin 2\pi(f_z - f_m)t$$

Dit lijkt heel erg op de parametervergelijking van een cirkel met straal $S + Sa(t)$,



Figuur 19: Het plaatje wat men op de oscilloscoop zou zien wanneer men het IQ-signaal in XY-modus bekijkt. De modulatie is zoals in het eerste plaatje (maar nu zonder overmodulatie). Rood is bij een hoge $f_z - f_m$, en Blauw bij een vrij lage $f_z - f_m$ (de oscillator staat dan bijna op de frequentie van de zender)

als je $I(t)$ op de X-as zet, en $Q(t)$ op de Y-as, dan zie je op de oscilloscoop in XY-modus inderdaad een cirkel ontstaan waarvan de straal met het audiosignaal mee varieert. (in de praktijk wordt het geen mooie cirkel, omdat $a(t)$ al veranderd is voordat er een complete cirkel is getekend). Dit is het belangrijkste testsignaal dat je hebt om zo'n demodulator te testen.

Dan nu nog even over de fout die gemaakt wordt door het niet meerekenen van het frequentiespectrum van $a(t)$. Het signaal van een AM-zender bestaat uit twee symmetrische zijbanden rond de zendfrequentie f_z . Door het omlaag mixen zitten die zijbanden gegroepeerd rond $f_z - f_m$. Stel dat de complete zijband een breedte heeft van $4.5kHz$, $f_z - f_m = 3kHz$ en $f_f = 6kHz$. Dan is er dus een zijband die zich van 3 tot $7.5kHz$ uitstrekt, en eentje die zich van 3 tot $-1.5 = 1.5kHz$. Nou komt dus niet de hele hoge zijband erdoor, die wordt na $3kHz$ audiofrequentie afgekapt. Dus tot $3kHz$ gaat alles nog perfect zoals het bedoeld was. De hogere frequentiecomponenten komen nog maar half door, en kunnen voor vervorming zorgen, maar dat is niet zo erg.

Omdat we als luisteraar genteresseerd zijn in $a(t)$, hebben we er veel aan als we $S + Sa(t)$ hebben. Dit laatste signaal is namelijk het audiosignaal, versterkt of verzwakt met factor S, en met constante S erbij opgeteld. De gelijkspanning S kan later ook direct gebruikt worden om de signaalsterkte te meten.

Eerst nog wel even de moeilijkheid hoe we $S + Sa(t)$ uit het signaal halen. Zoals opgemerkt stellen de I en Q signalen een cirkel voor, met als straal $S +$

$Sa(t)$. We berekenen dus de cirkel van de straal, en de formule daarvoor is via pythagoras:

$$\begin{aligned} R(t) &= \sqrt{I(t)^2 + Q(t)^2} = \\ &= \sqrt{(S + Sa(t))^2 \cos^2 2\pi(f_z - f_m)t + (S + Sa(t))^2 \sin^2 2\pi(f_z - f_m)t} = \\ &= (S + Sa(t))\sqrt{\cos^2 2\pi(f_z - f_m)t + \sin^2 2\pi(f_z - f_m)t} = (S + Sa(t)) \end{aligned}$$

Het is nu dus wel duidelijk waar alle stappen na de low-pass filters voor dienen. Het kwadrateren kan met dubbelgebalanceerde mixers geschieden. Het optellen kan natuurlijk gewoon met twee weerstanden, en het wortelen kan ook met een gebalanceerde mixer en een comperator.

Merk overigens op dat mixen met een signaal met dezelfde frequentie als de zender ($f_z - f_m = 0$), en in fase alleen een I-signaal zou geven, en dat het I-signaal dan al direct $S + Sa(t)$ zal bedragen. (het hele Q-signaal mag dus vergeten worden):

$$I(t) = \frac{1}{2}(S + Sa(t)) \cos 2\pi(f_z - f_m)t = \frac{1}{2}S + \frac{1}{2}Sa(t)$$

Dit is het principe achter de synchrodyne. De synchrodyne is een ontvanger die meestal bestaat uit een oscillator die ook meteen mixt. Doordat de oscillator ook tegelijk mixt, staat er ook een antenne-signaal op de oscillator. Daardoor heeft de oscillator de neiging om meegetrokken te worden naar de draaggolf van van de zender, en in fase daarmee te oscilleren. Dat levert ook een mooi AM signaal op. Maar het stemt lastig af, want het meetrekken van de oscillator gebeurt maar in een erg klein gebied. Stem je ernaast af, dan hoor je een hard gepiep. Meestal wordt deze methode van demodulatie bij eenkringers en andere (zeer) simpele ontvangers gebruikt.

5 Enkelzijband

5.1 algemene informatie

Enkelzijbandmodulatie is in de jaren 30 ontwikkeld om zoveel mogelijk telefoonsignalen door n enkel draadje te sturen. Het voordeel aan enkelzijband is namelijk dat dit maar de helft van de bandbreedte in beslag neemt van gewone AM modulatie. Het is later ook bij amateurs meer in zwang gekomen door het hogere rendement, er is namelijk 6 keer meer nuttig vermogen dan bij AM modulatie. Daar staan dan echter wel weer grote nadelen tegenover, en dan vooral de moeilijkheid om het te detecteren. Voor spraak gaat het allemaal wel. Maar muziek luisteren via enkelzijband is eigenlijk geen doen. Daarom is het nooit door de omroepen gebruikt. Alhoewel... Er zijn plannen geweest om de twee verschillende zijbanden van een AM signaal beiden apart te moduleren met 2 verschillende signalen om stereo op de middengolf te verkrijgen. Dit is echter nooit wat geworden in Europa, al had het makkelijk gekund.

5.2 wiskundige interpretatie

Er bestaan twee verschillende soorten EZB modulatie, USB en LSB, upper resp. lower side band. Het verschil zit m erin dat bij USB de hoge zijband van een

AM signaal wordt, en bij LSB alleen de lage. Je zou kunnen zeggen dat bij USB modulatie bij ieder frequentiecomponent de zendfrequentie f_z wordt opgeteld, en bij LSB wordt van f_z de betreffende frequentie afgetrokken. Stel weer dat we een de fouriergetransformeerde van een signaal $a(t)$ hebben: $\hat{a}_{sin}(s)$ en $\hat{a}_{cos}(s)$. Dan ziet het USB signaal er zo uit:

$$x(t) = \int_{s=0}^{s=\infty} \hat{a}_{sin}(f_z + s) \sin ts + \hat{a}_{cos}(f_z + s) \cos ts \quad (11)$$

en voor het LSB signaal het volgende:

$$x(t) = \int_{s=0}^{s=\infty} \hat{a}_{sin}(f_z - s) \sin ts + \hat{a}_{cos}(f_z - s) \cos ts$$

Stel dat we bijvoorbeeld slechts een sinustoon van $200Hz$ USB-uitzenden, $a(t) = \sin 400\pi t$, dan krijgen we uit de zender dus een signaal van $f_z + 200Hz$. En dat is dan ook al het signaal dat wordt uitgezonden. Iedere toon krijg de zendfrequentie erbij opgeteld. In de praktijk blijkt alleen bij losse sinustonen dit een signaal op te leveren waar je iets aan kan meten. Alle ingewikkelder signalen zien er chaotisch uit, en lijken op het eerste gezicht niet op het signaal wat je erin stopt.

Het grootste probleem aan deze modulatiemethode is dat de ontvangstzijde moet zien te achterhalen wat de zendfrequentie f_z is. In het geval er een losse sinustoon wordt uitgezonden is het onmogelijk dit te achterhalen, omdat alles wat wordt uitgezonden een sinus met frequentie die f_z hoger is. Voor spraak is het op het gehoor redelijk te achterhalen. stemt men te hoog af, dan hoort men de stem te laag klinken, en andersom, daarna gokt men de zendfrequentie dus op zo'n manier dat het verstaanbaar klinkt. Maar echt helemaal goed zit je niet snel, daar heb je een referentiefrequentie of iets dergelijks voor nodig, dat je niet hebt. Alleen voor TV uitzendingen wordt wel een dergelijk systeem gebruikt. Daar wordt wel een draaggolf gebruikt, en een klein beetje lage zijband, en een complete hoge zijband.

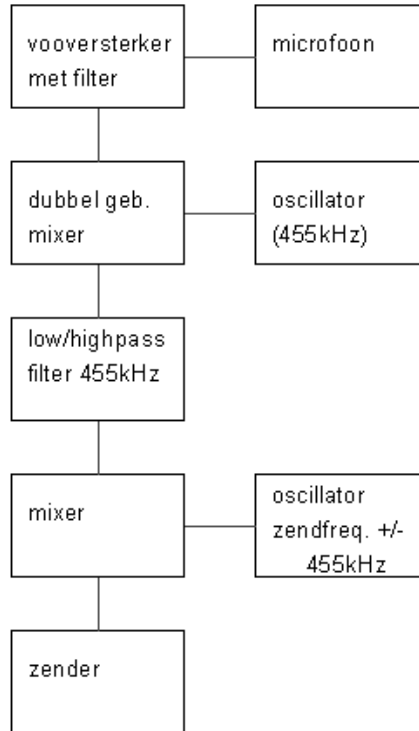
5.3 modulatiemethoden

Voor enkelzijband bestaan er 3 modulatiemethoden: de filtermethode, de fasemethode en de weavermethode. Ik zal ze alledrie behandelen. Het handigste is om te beginnen met de filtermethode, omdat dit de eenvoudigste is, en het meeste inzicht geeft in het enkelzijbandsignaal. De fase- en filtermethode maken ingenieus gebruik van een beetje goniometrie, en zijn dus wat ingewikkelder. En dus wel zo interessant.

Allereerst de filtermethode. Bij deze methode wordt het signaal in een dubbelgebalanceerde mixer naar een vaste tussenfrequentie f_m gemixt. Bijvoorbeeld $455kHz$. Ik ga er even van uit dat het audiosignaal bestaat uit een toon met frequentie f_a ($a(t) = \sin 2\pi f_a t$). Dan staat op de uitgang van de dubbelgebalanceerde mixer:

$$x(t) = a(t) \sin 2\pi f_z t = \frac{1}{2}(\cos 2\pi(f_z - f_a)t - \cos 2\pi(f_z + f_a)t)$$

Het is dus duidelijk, we hebben twee zijbanden gecentreerd rond f_z . Om de ene zijband te selecteren hoeven we alleen de andere weg te filteren. Er bestaan



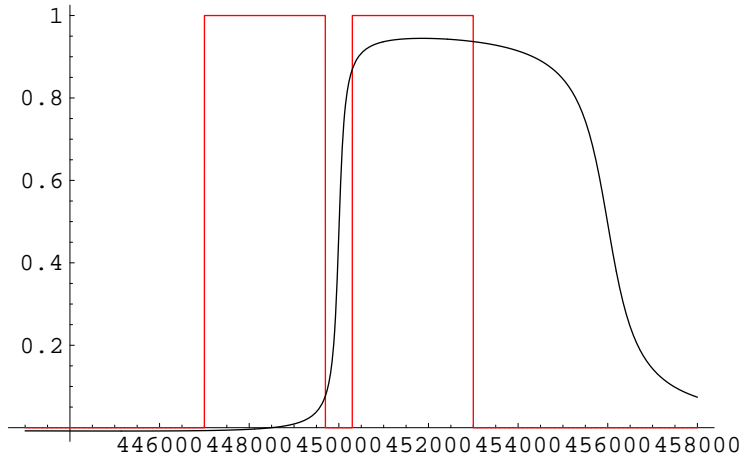
Figuur 20: Blokschema van de filtermethode voor SSB

hiervoor mechanische filters die zeer stijle flanken hebben.(d.w.z. filters waar buiten hun doorlaat-gebied de amplitude heel erg snel afneemt voor frequenties die er steeds verder buiten liggen). We hebben dan een filter nodig om de hoge zijband te selecteren, en een ander filter om de lage zijband te selecteren. de ene laat boven de $455kHz$, en de andere juist eronder.

Dan hebben we dus een mooi SSB signaal op $455kHz$, maar daar heb je nog weinig aan, want je wil toch graag zelf je zendfrequentie nog wel een beetje in kunnen stellen. Om dat te doen moet dit signaal nog een keer gemixt worden, en nu met een frequentie $455kHz$ hoger of lager dan de uiteindelijke zendfrequentie. Doordat je de tweede keer mixt ontstaan er wel weer twee mengproducten, maar als je bijvoorbeeld uit wil zenden op $6MHz$, dan zitten het ene mengproduct dus op $6MHz$, en het andere op $6,910MHz$. Dat verschil is zo groot dat het probleemloos weg te filteren valt.

De fasemethode is wat ingewikkelder. Allereerst wordt het signaal weer gefilterd zodat het weer netjes van $300Hz$ tot $3kHz$ loopt. Maar daarna komt meteen het moeilijkste stuk. Het signaal moet namelijk 90 graden gedraaid worden. Een schema wat dat doet heet een hilberttransformator. (de hilberttransformatie is namelijk een methode om alle cosinussen om te zetten in sinussen en andersom). Ik zal hier de meest gebruikte methode beschrijven: het polyfase netwerk.

We zullen weer uitgaan van een audio-signaal dat uit slechts n frequentie



Figuur 21: Het rode signaal is het signaal afkomstig uit de dubbelgebalanceerde mixer, 2 zijbanden, de hoge van $455.3kHz$ tot $458kHz$, en de lage van $449.7kHz$ tot $447kHz$. De zwarte lijn is het filter, en dat selecteert hier dus alleen de hoogste zijband. De lage zijband is door het filter bijna compleet onderdrukt.

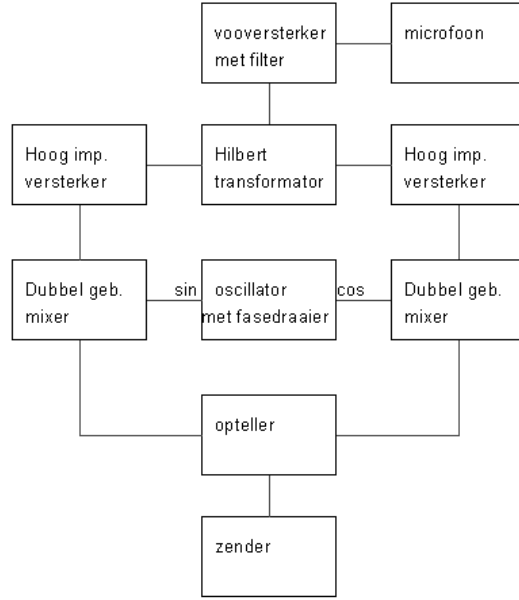
bestaat. $a(t) = \sin 2\pi f_a t$. Door deze toon 4 maal 90 graden in fase te draaien kan ik 4 verschillende signalen bekomen:

$$\begin{aligned} a_1(t) &= \sin 2\pi f_a t \\ a_2(t) &= \cos 2\pi f_a t \\ a_3(t) &= -\sin 2\pi f_a t \\ a_4(t) &= -\cos 2\pi f_a t \end{aligned}$$

Nou maak ik gebruik van de schakeling in de figuur om de fases te draaien. Stel ik sluit hier de signalen a_1 tot a_4 op aan. Wat eruit komt, a'_1 tot a'_4 moet zeker weten dezelfde faseverhoudingen hebben (90 graden ertussen). Want iedere keer wordt het signaal op twee naast elkaar liggende pinnen (dus 90 graden verschillend) op dezelfde manier via condensator en weerstand bij elkaar gevoegd. Wat niet bekend is is de faseverhouding van a'_1 t.o.v. a_1 (blijkt onbelangrijk), en de sterkte van a'_1 t.o.v. a_1 , is zeer belangrijk. Dit zal ik straks uitrekenen. Nou definieer ik:

$$\begin{aligned} b_1(t) &= \sin 2\pi f_a t \\ b_2(t) &= -\cos 2\pi f_a t \\ b_3(t) &= -\sin 2\pi f_a t \\ b_4(t) &= \cos 2\pi f_a t \end{aligned}$$

Deze schelen weer onderling allemaal precies 90 graden. En ze voldoen ook weer allemaal aan dezelfde eigenschappen die ik hierboven ook voor a_1 tot a_4 aannam. Er is echter een groot verschil in hoe ze door het filter heen komen. Daarom reken ik eerst uit wat het filter doet met de signalen. Daarvoor gebruik ik dan wel de complexe notatie van een sinusgolf en impedanties om het wat sneller uit te rekenen. Omdat de verschillende condensatoren en weerstanden elkaar niet beïnvloeden kijk ik alleen naar een setje van 1 condensator en 1 weerstand waar 2 golven op staan die onderling 90 graden verschoven zijn. $a_1(t) = e^{i2\pi t}$



Figuur 22: Blokschema van de fasemethode voor SSB

en $a_2(t) = ie^{2\pi t}$. (met $i = \sqrt{-1}$). Ik trek eerst overal a_2 vanaf, zodat ik kan rekenen alsof a_2 aan de aarde ligt, hetgeen makkelijker is.

$$\begin{aligned}
 a_2'(t) - ie^{i2\pi f_a t} &= \frac{R}{R - \frac{i}{2\pi f_a C}} (1 - i) e^{i2\pi f_a t} \\
 a_2'(t) &= \left(\frac{2\pi f_a RC - i2\pi f_a RC}{2\pi f_a RC - i} + \frac{i2\pi f_a RC + 1}{2\pi f_a RC - i} \right) e^{2\pi f_a t} e^{i2\pi f_a t} \\
 a_2'(t) &= \frac{2\pi f_a RC + 1}{2\pi f_a RC - i} e^{i2\pi f_a t}
 \end{aligned}$$

De amplitude van a_2' is dus:

$$A = \frac{|2\pi f_a RC + 1|}{\sqrt{(2\pi f_a RC)^2 + 1}}$$

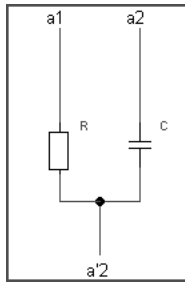
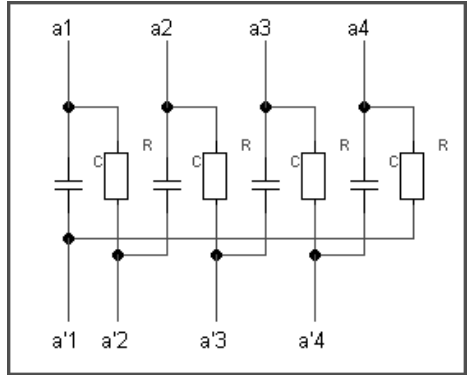
Op eenzelfde manier valt ook af te leiden dat als de fase van de a_2 juist 90 graden terug is t.o.v. a_1 , zoals bij b_1 tot b_4 , de amplitude gegeven is door:

$$B = \frac{|2\pi f_a RC - 1|}{\sqrt{(2\pi f_a RC)^2 + 1}}$$

Het signaal $a_1'(t)$ wordt nu dus bijvoorbeeld:

$$a_1'(t) = a_1(t) A e^{i\theta}$$

met θ de faseverschuiving t.o.v. van $a_1(t)$. Maar deze factor is voor a_1 tot a_4 gelijk. Ik hoef niet te weten hoeveel die faseverschuiving is, zolang de onderlinge verhouding voor a_1' tot a_4' nog steeds 90 graden is. Het enige wat belangrijk is is dat de grootte van alle 4 die signalen verandert na het filter. Dit laatste is een zeer belangrijk gegeven. Want als $2\pi f_a RC = 1$, dan vallen de b -signalen dus helemaal weg voor die bepaalde frequentie. Voor andere frequenties bestaat



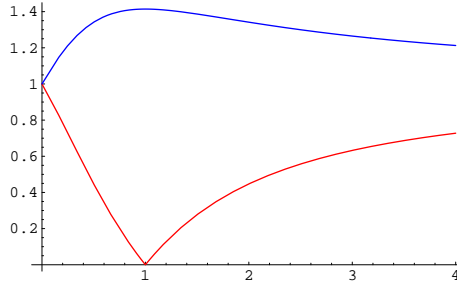
ze nog wel. Terwijl A juist extra groot is voor deze frequenties. We zijn dus in staat om a_1 tot a_4 door te laten, en b_1 tot b_4 juist tegen te houden. Dit kunnen we effectief gebruiken door op de ingang een combinatie van deze beide signalen te zetten. Noem nu c_1 tot c_4 de signalen die we daadwerkelijk toevoegen de schakeling. Dan kies ik ze zo:

$$\begin{aligned}
 c_1(t) &= a_1(t) + b_2(t) = \sin 2\pi f_a t - \cos 2\pi f_a t \\
 c_2(t) &= a_2(t) + b_3(t) = \cos 2\pi f_a t - \sin 2\pi f_a t = -c_1(t) \\
 c_3(t) &= a_3(t) + b_4(t) = -\sin 2\pi f_a t + \cos 2\pi f_a t = -c_1(t) \\
 c_4(t) &= a_4(t) + b_1(t) = -\cos 2\pi f_a t + \sin 2\pi f_a t = c_1(t)
 \end{aligned}$$

We blijken dus genoeg te hebben aan twee 180 graden verschoven signalen ($c_1(t)$ en $-c_1(t)$). Maar daar valt vrij simpel aan te komen. Een inverterende versterker is genoeg. De signalen b_1 tot b_4 vallen allemaal weg aan de uitgang voor de goede frequentie. En voor die frequentie houden we dus alleen a_1 tot a_4 over.

Voor de frequentie waarvoor geldt: $2\pi f_a RC = 1$ houden we nu 4 netjes 90 graden verschoven signalen over. voor andere frequenties wordt B al snel weer groter. Maar daar hebben we geen last meer van als er bijvoorbeeld 6 van deze filters na elkaar worden geplaatst, met ieder net een iets andere frequentie. In dat geval is het mogelijk om bijna alle verkeerde signalen tegen te houden in het gebied van $300Hz$ tot $3kHz$. Maar dan moet het audiosignaal wat erin gaat ook al wel goed gefilterd zijn, want voor frequenties die erbuiten vallen neemt de amplitude snel toe.

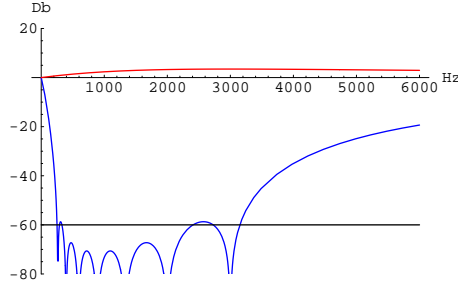
Bij meerdere van deze schakelingen achter elkaar mag je simpelweg de amplitudes met elkaar vermenigvuldigen. Stel dat A_1 tot A_6 en B_1 tot B_6 de verschillende amplitudes uit een 6-voudig filter zijn. Dan krijg ik dus voor $A_{tot} = A_1 A_2 A_4 A_4 A_5 A_6$, en idem voor B_{tot} . Let op, A_1 tot B_6 zijn eigenlijk



Figuur 23: Het blauwe signaal stelt A voor, en het rode signaal stelt B voor. Dit alles voor $2\pi RC = 1$. Duidelijk is te zien hoe B verdwijnt voor $f_a = 1$ ($2\pi f_a RC = 1$)

allemaal functies de frequentie f_a , en ze hebben als parameter $2\pi RC$. Om een goede modulator te maken is het nodig om over het volledige bereik van $300Hz$ tot $3kHz$ de B_{tot} zo klein mogelijk te houden. Er valt exact uit te rekenen hoe groot de weerstanden en condensatoren dan precies gekozen moeten worden. Maar in de praktijk is het het handigst om het gewoon bij standaardwaarden voor de condensatoren en weerstanden te houden. Dat scheelt misschien $5dB$ of een 10% in bandbreedte, echt significant is het niet. Voor een filter dat bestaat uit 6 elementen kan men mooi weerstanden met een vaste waarde van $10k\Omega$ nemen, en condensatoren met een waarde beginnend bij $2\pi \cdot 3000 \cdot 10000C = 1$, dus $C_1 = 4.7nF$. En als men daarna voor iedere volgende stap de capaciteit met ongeveer 1.5 vermenigvuldigd, dan krijgt men: $C_2 = 6.8nF, C_3 = 10nF, C_4 = 15nF, C_5 = 22nF, C_6 = 33nF$. Dan zien we dat C_6 iets te laag uitkomt, want die zouden we eigenlijk rond de $3kHz$ willen hebben. Om dat te compenseren zou voor de weerstanden een waarde van $12k\Omega$ gekozen kunnen worden. (Men zou natuurlijk ook kunnen kiezen voor allemaal condensatoren van dezelfde waarde en weerstanden van verschillende waarden) De demping die men verkrijgt bij deze waarden krijg heb ik uitgezet in een grafiek. Het is mogelijk om het nog ietsje beter te doen door condensator of weerstandswaarden ietsje slimmer te kiezen, zodat alle toppen even hoog uitkomen. Het verschil is toch marginaal. Waarschijnlijk zelfs niet eens merkbaar door de toleranties van de componenten. Let goed op dat men hier dus wel condensatoren en weerstanden met een extra lage tolerantie moet kiezen. Kleine veranderingen zorgen meteen voor grote fouten. Dus het liefst ook onderdelen met zeer kleine toleranties, en alle weerstanden en condensatoren matchen voor zover je daartoe in staat bent. De uitgangsweerstand van deze schakeling is ook behoorlijk hoog, dus het is ook noodzakelijk om dit met iets hoogohms uit te koppelen.

Dit was het lastigste deel van de fasedraaiende SSB-modulator. Voor iedere spectrale component van het audiosignaal met frequentie f_a die eruit ziet als $a(t) = \sin 2\pi f_a t$ hebben we nu dus $a_s(t) = \sin 2\pi f_a t$ en $a_c(t) = \cos 2\pi f_a t$. We mixen dit met een sinus en een cosinus tot de eindfrequentie, of weer tot een tussenfrequentie. Dat kan allebei, en eigenlijk maakt het niet zo knap veel uit. Als we eerst naar een tussenfrequentie van bijvoorbeeld weer $455kHz$ gaan, dan valt van daaruit hetzelfde trucje toe te passen als ook is gebeurd bij de



Figuur 24: Het blauwe signaal stelt weer het A_{tot} -signaal voor, het signaal dat erdoor moet komen. Let ook op, dit wordt zelfs lichtelijk versterkt door dit passieve filter. Rood stelt het B_{tot} -signaal voor. zwart is een referentielijn op $-60dB$. De onderdrukking zit dus rond de -60 dB.

filtermethode. Ik zal er daarom maar vanuit gaan dat we direct naar de zendfrequentie gaan. Al blijkt het in de praktijk vaak handiger om eerst naar een tussenfrequentie te gaan. In dat geval moet je het signaal nog eens mixen.

We mixen dus het ene signaal met een cosinus en het andere signaal met een sinus.

$$\begin{aligned}
 x_1(t) &= a_s(t) \sin 2\pi f_z t = \sin 2\pi f_a t + \theta \sin 2\pi f_z t \\
 x_2(t) &= a_c(t) \cos 2\pi f_z t = \cos 2\pi f_a t + \theta \cos 2\pi f_z t \\
 x_1(t) &= \frac{1}{2} \cos 2\pi(f_z - f_a)t - \frac{1}{2} \cos 2\pi(f_z + f_a)t \\
 x_2(t) &= \frac{1}{2} \cos 2\pi(f_z - f_a)t + \frac{1}{2} \cos 2\pi(f_z + f_a)t \\
 x_1(t) + x_2(t) &= x_{LSB}(t) = \cos 2\pi(f_z - f_a)t \\
 x_1(t) - x_2(t) &= x_{USB}(t) = \cos 2\pi(f_z + f_a)t
 \end{aligned}$$

Het valt eenvoudig te controleren dat dit ook klopt voor $a(t) = \cos 2\pi f_a t$ en dus: $a_s(t) = \cos 2\pi f_a t$ en $a_c(t) = -\sin 2\pi f_a t$. Omdat dit ook klopt gaat het goed voor iedere losse frequentie component. En omdat alle bewerkingen weer lineair waren naar $a(t)$ klopt het dus ook voor een compleet audiosignaal.

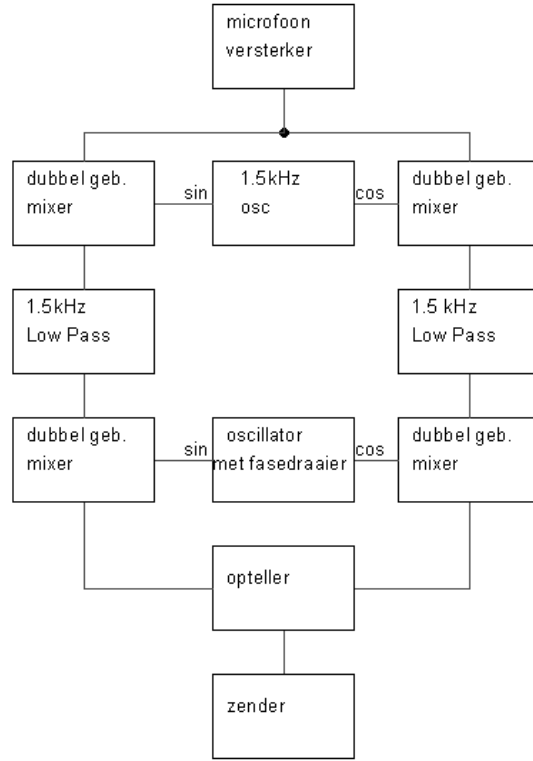
Het kiezen tussen LSB of USB, (lower of upper sideband) kan dus door $x_2(t)$ erbij op te tellen of af te trekken. Nou is aftrekken nog wel een enigszins vervelende operatie. Gelukkig komt uit het polyfasenetwerk ook het signaal $-a_c(t)$. Door hiermee te mixen in plaats van met $a_c(t)$ krijg je $-x_2(t)$. En dus komt het kiezen tussen LSB en USB neer op het kiezen tussen de ene fase, en de tegenover liggende fase uit het polyfasenetwerk. Weer erg handig allemaal.

De laatste SSB modulator die ik zal behandelen is de weaver modulator. Deze maakt weer gebruik van een hoop goniometrie. Allereerst hoeft het audiosignaal bij deze modulator niet van te voren al gefilterd te worden. Dat gebeurt in de modulator zelf met behulp van 2 filters met $1.5kHz$ breedte.

We gaan weer van start met $a(t) = \sin f_a t$. Als eerste wordt dit weer gemixt met een signaal van $1.5kHz = f_m$. (deze frequentie mag ook hoger zijn. Het dubbele van f_m is de totale bandbreedte. Noem $I(t)$ en $Q(t)$ de signalen die ontstaan door met een sinus van f_m te mixen resp. een cosinus.

$$\begin{aligned}
 I(t) &= \sin 2\pi f_a t \sin 2\pi f_m t = \frac{1}{2} \cos 2\pi(f_a - f_m)t - \frac{1}{2} \cos 2\pi(f_a + f_m)t \\
 Q(t) &= \sin 2\pi f_a t \cos 2\pi f_m t = \frac{1}{2} \sin 2\pi(f_a - f_m)t + \frac{1}{2} \sin 2\pi(f_a + f_m)t
 \end{aligned}$$

Dit wordt door een low pass filter gehaald met afsnijfrequentie f_m ($1.5kHz$).



Figuur 25: Blokschema van de weavermethode voor SSB

Hierdoor vallen alle termen met $(f_a + f_m)$ weg. (f_a is natuurlijk altijd positief). Alleen de termen met $(f_a - f_m)$ blijven over, alleen wanneer $f_a < 2f_m$. (omdat frequenties in principe altijd positief zijn selecteert het filter alles waarvan de frequentie $-f_m < f_{in} < f_m$). We houden over: $I'(t)$ en $Q'(t)$.

$$\begin{aligned} I'(t) &= \frac{1}{2} \cos 2\pi(f_a - f_m)t \\ Q'(t) &= \frac{1}{2} \sin 2\pi(f_a - f_m)t \end{aligned}$$

Dit wordt dan weer met de zendfrequentie, of met een tussenfrequentie gemixt. Maar in ieder geval een frequentie waar weer een sinus en een cosinus van voorhanden zijn. De frequentie hiervan is f_z , en we krijgen nu $x_I(t)$ en $x_Q(t)$:

$$\begin{aligned} x_I(t) &= I'(t) \sin 2\pi f_z t = \frac{1}{2} \cos 2\pi(f_a - f_m)t \sin 2\pi f_z t \\ x_Q(t) &= Q'(t) \cos 2\pi f_z t = \frac{1}{2} \sin 2\pi(f_a - f_m)t \cos 2\pi f_z t \end{aligned}$$

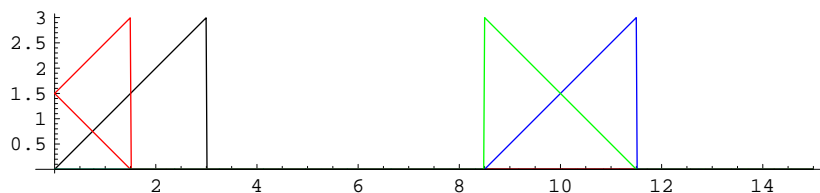
$$\begin{aligned} x_I(t) &= \frac{1}{4} \sin 2\pi(f_z - f_a + f_m)t + \frac{1}{4} \sin 2\pi(f_z + f_a - f_m)t \\ x_Q(t) &= \frac{1}{4} \sin 2\pi(f_a - f_m - f_z)t + \frac{1}{4} \sin 2\pi(f_a + f_m + f_z)t \end{aligned}$$

$$\begin{aligned} x_{USB}(t) &= x_I(t) + x_Q(t) = \frac{1}{2} \sin 2\pi(f_z + f_a - f_m)t \\ x_{LSB}(t) &= x_I(t) - x_Q(t) = \frac{1}{2} \sin 2\pi(f_z - f_a + f_m)t \end{aligned}$$

Dit geeft dus weer ongeveer hetzelfde resultaat als eerst. Maar er is wel een klein

verschil, dat is de frequentie f_m die erbij wordt geteld of eraf wordt getrokken. De zendfrequentie wijkt nu dus altijd precies f_m af van de frequentie f_z . Maar dat zijn maar details. Om nu van USB naar LSB om te schakelen is het niet mogelijk om een andere uitgang van het polyfasenetwerk te kiezen. Het is wel mogelijk om de sinus en de cosinus van n van beide oscillatoren om te draaien. Op die manier bereik je hetzelfde effect.

Wil men deze methode toepassen dan moet er vooral veel aandacht geschonken worden aan de twee filters. Ze moeten zo identiek mogelijk zijn, en zeer sterk naar beneden lopen. Het valt ook wel aan te raden om de audiofrequenties onder de $300Hz$ toch maar af te kappen, en de kantelfrequenties van de filters net iets lager te nemen dan f_m . Dit alles om niet in de problemen te komen met het overlappen van de gemixte signalen.



Figuur 26: Stel dat het zwarte driehoekje het spectrum is van het audiosignaal. Dan wordt dat door de eerste mixer met $1.5kHz$, en door het filter 'rond de $0kHz$ gevouwen'. (dankzij de fase valt nog wel te onderscheiden wat van een hoge frequentie komt en wat niet). Daarna wordt het gemixt met f_z van $10kHz$. Dan is blauw het USB signaal en groen het LSB signaal.

5.4 demodulatie

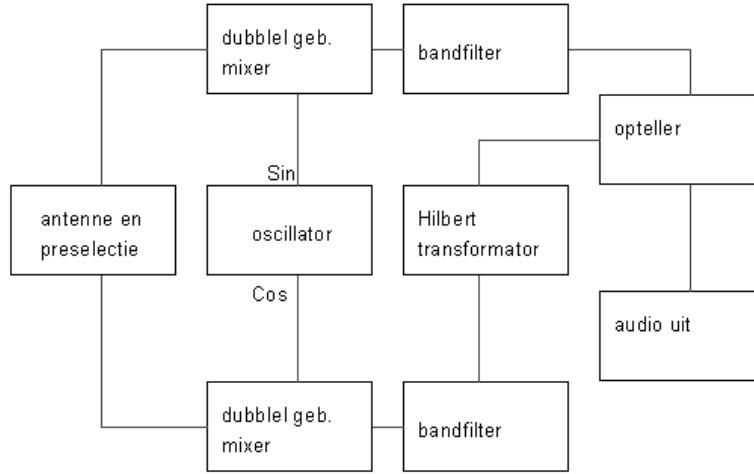
Voor het demoduleren van een SSB signaal worden over het algemeen precies dezelfde methoden gebruikt als voor het moduleren van een SSB signaal. Kwam het moduleren overeen met het optellen van de zendfrequentie bij iedere spectrale component van het audiosignaal, het demoduleren komt overeen met het aftrekken van de zendfrequentie van het SSB signaal. Omdat dat zo mooi analoog aan elkaar is, kunnen in principe ook dezelfde methodes toegepast worden. Het meest algemeen is wat ik eerst de filtermethode heb genoemd. In een goede ontvanger is een bandfilter aanwezig met een breedte van ongeveer $3kHz$. Dit filtert precies het SSB signaal eruit. Dit valt te demoduleren door simpelweg te mixen met de zendfrequentie (of vaak de middenfrequentie). Stel dat we een USB gemoduleerd enkelzijbandsignaal hebben van een losse toon met frequentie f_a . Dan is zendersignaal: $x(t) = -\cos 2\pi f_z + f_a t$. Dit gemixt met de zendfrequentie geeft:

$$x(t) \sin 2\pi f_z t = -\cos 2\pi f_z + f_a t \sin 2\pi f_z t = \frac{1}{2} \sin 2\pi f_a t + \frac{1}{2} \sin 2\pi(2f_z + f_a)t$$

En het audiosignaal komt er dus weer mooi uitrollen, samen met een signaal van hoge frequentie dat zonder probleem eruit gefilterd kan worden. Voor dit doel volstaat een simpele diodemixer.

Let erop dat de mengoscillator precies op f_z moet worden afgestemd. Gebeurt dit niet, dan worden alle frequenties in het audio-frequenties een stukje opgeschoven,

met enorme vervorming tot gevolg. Het is dan ook noodzaak om de mengfrequentie zeer nauwkeurig te regelbaar te maken, zodat je zeer dicht bij f_z in de buurt kan komen.



Figuur 27: SSB-quadratuurdemodulator met hilberttransformator

En dan de implementatie voor quadratuurontvangers. Een quadratuurontvanger is in principe vlak boven en onder de mengfrequentie f_m gevoelig. Het is echter mogelijk om te onderscheiden welk signaal er van een frequentie boven de mengfrequentie afkomstig is, en welk van een signaal onder de mengfrequentie.

Stel nu dat we twee SSB zenders hebben, de een zend een toon uit met frequentie f_a en is USB gemoduleerd, de ander zend een toon uit met frequentie f_b en is LSB gemoduleerd. Beide zenders hebben als zendfrequentie f_z . De oscillator van mijn quadratuurontvanger stel ik ook af op $f_z = f_m$. Het antennesignaal is nu dus: $x(t) = \cos 2\pi(f_z - f_b)t - \cos 2\pi(f_z + f_a)t$ Voor $I(t)$ en $Q(t)$ vind ik dan:

$$\begin{aligned} I(t) &= x(t) \sin 2\pi f_z t = (\cos 2\pi(f_z - f_b)t - \cos 2\pi(f_z + f_a)t) \sin 2\pi f_z t \\ Q(t) &= x(t) \cos 2\pi f_z t = (\cos 2\pi(f_z - f_b)t - \cos 2\pi(f_z + f_a)t) \cos 2\pi f_z t \end{aligned}$$

Na filtering van de hoge componenten (en $2x$ versterking) vind ik dan:

$$\begin{aligned} I(t) &= \sin 2\pi f_b t + \sin 2\pi f_a t \\ Q(t) &= \cos 2\pi f_b t - \cos 2\pi f_a t \end{aligned}$$

En dan gaat nu het $Q(t)$ signaal door de hilberttransformator (wordt $Q_H(t)$), dewelke het signaal 90 graden vooruit schuift.

$$Q_H(t) = -\sin 2\pi f_b t + \sin 2\pi f_a t$$

Het optellen van beide Q_H en I geeft duidelijk het LSB signaal. Zouden we de 90 graden terug schuivende uitgang van de hilberttransformator gebruiken, dan krijgen we $-Q_H$, en dit opgeteld bij I geeft mooi het USB signaal. Met

deze methode is het dus mogelijk om te selecteren welke zijband je wilt hebben. Een belangrijk punt is wel dat er weer een scherp $300Hz$ tot $3kHz$ bandfilter nodig is. (dit in tegenstelling tot de quadratuurdemodulator voor AM, waar een simpel lows-pass filter voldoende was)

Dit demoduleren kan overigens ook op min of meer dezelfde manier als een weavermodulator gedaan worden. In principe wordt dan exact dezelfde schakeling gebruikt, maar dan met de $1.5kHz$ en de f_z oscillatoren omgewisseld. (in eerste instantie wordt er dan ook een quadratuursignaal gemaakt, dus in principe is dit ook een methode voor quadratuurdemodulatie, die veel eenvoudiger is).

6 FM

6.1 algemene informatie

De laatste veelgebruikte modulatiemethode is frequentiemodulatie, kortweg FM. Dit is tegenwoordig veruit de populairste vorm van omroepradio omdat het superieur klinkt aan AM modulatie. Maar dat ligt niet aan de modulatie, maar aan de grote bandbreedte van een FM signaal. De bandbreedte van een FM omroepsignaal ligt in de grootte orde van $150kHz$. Veel meer dan de $40kHz$ die voor HIFI-stereo AM modulatie nodig zou zijn. (AM modulatie met beide zijbanden verschillend, om stereo te verkrijgen) Frequentiestabiliteit was echter in de begintijd van de FM ontvangst nogal een probleem, daarom, en om ruis tegen te gaan, werd voor een grote bandbreedte gekozen.

Zoals de naam al zegt wordt bij FM de frequentie gemoduleerd. Dit in tegenstelling tot AM, (en SSB wat ervan afgeleid is). Dit maakt zenders veel eenvoudiger, omdat de amplitude van het signaal niet nauwkeurig geregeld hoeft te worden. Het signaal is altijd maximaal, en dat maakt het mogelijk om de zender in klasse C in te stellen. Met het grote voordeel van een hoog rendement en relatief goedkope en simpele onderdelen. Daar tegenover staat dan weer dat het een stuk lastiger is om te demoduleren.

6.2 wiskundige interpretatie

Het FM signaal bestaat dus uit een signaal waarvan de frequentie lichtjes schommelt rond de zendfrequentie. Het lukt om te zeggen dat het signaal er zo uit ziet (bij modulatie met toon met frequentie f_a):

$$x(t) \neq \sin 2\pi(f_z + \alpha \sin 2\pi f_a t)t$$

Dit is echter helemaal fout. Aangezien bij toenemende tijd het effect van het modulerende signaal hier ook toeneemt. Ik zal allereerst de functie van een fasegemoduleerd signaal geven. (PM-modulatie). Deze modulatie lijkt zeer sterk op frequentiemodulatie, zoals we later zullen zien. Laat $a(t)$ het audiosignaal zijn, en α een constante die aangeeft hoe sterk gemoduleerd wordt.

$$x_{PM}(t) = \sin 2\pi(f_z t + \alpha a(t)) \quad (12)$$

Voor frequentiemodulatie is het ietsje lastiger. De fase van het signaal neemt steeds toe in de tijd, met snelheid f_z zonder modulatie. Bij lineair toenemende fase krijg je een standaard sinusfunctie, met frequentie f_z . Maar met modulatie,

dan neemt de fase ietsje sneller of slomer toe, en wel met $f_z + \alpha a(t)$. In een infinitesimaal klein tijdseenheidje dt zal dus de fase zijn toegenomen met $f_z dt + \alpha a(t) dt$. En dus kan men voor de fase $\Theta(t)$ schrijven:

$$\Theta_{FM}(t) = \int_{x=0}^{x=t} f_z + \alpha a(x) dx = f_z t + \alpha \int_{x=0}^{x=t} a(x) dx$$

Met x een dummy-tijdsvariabele. Dit levert dan de uiteindelijke formule voor een FM signaal:

$$x(t) = \sin 2\pi\Theta(t) = \sin 2\pi(f_z t + \alpha \int_{x=0}^{x=t} a(x) dx) \quad (13)$$

Nu bestaat er een relatie tussen het fasegemoduleerde signaal, en het frequentiegemoduleerde signaal. Voor het fasegemoduleerde signaal geldt duidelijk:

$$\Theta_{PM}(t) = f_z t + \alpha a(t) \quad (14)$$

Het enige verschil zit in de integraal. Maar ook hier bieden fouriergetransformeerden weer een de uitkomst om het resultaat voor fasemodulatie $a(T)$ om te praten naar $\int_{t=0}^{t=T} a(t) dt$. (let op de naamsverandering x naar t , waar t loopt van 0 tot T).

$$\begin{aligned} \frac{\Theta_{FM}(T) - f_z T}{\alpha} &= \int_{t=0}^{t=T} a(t) dt \\ \int_{t=0}^{t=T} a(t) dt &= \int_{t=0}^{t=T} \int_{s=0}^{s=\infty} \hat{a}_{sin}(s) \sin ts + \hat{a}_{cos}(s) \cos ts \, ds dt = \\ &= \int_{s=0}^{s=\infty} -\hat{a}_{sin}(s) \frac{1}{s} \cos Ts + \hat{a}_{sin}(s) \frac{1}{s} + \hat{a}_{cos}(s) \frac{1}{s} \sin Ts \, ds \end{aligned}$$

Vergelijken met het resultaat voor fasemodulatie, met een signaal $b(t)$:

$$\begin{aligned} \frac{\Theta_{PM}(T) - f_z T}{\alpha} &= b(T) \\ b(T) &= \int_{s=0}^{s=\infty} \hat{b}_{sin}(s) \sin Ts + \hat{b}_{cos}(s) \cos Ts \, ds \end{aligned}$$

levert dat fasemodulatie hetzelfde signaal geeft als men een signaal

$$\begin{aligned} \hat{b}_{sin}(s) &= \frac{\hat{a}_{cos}(s)}{s} \\ \hat{b}_{cos}(s) &= -\frac{\hat{a}_{sin}(s)}{s} \end{aligned}$$

Voor $s \neq 0$ en voor $s = 0$:

$$\hat{b}_{cos}(0) = \int_{s=0}^{s=\infty} \frac{\hat{a}_{sin}(s)}{s}$$

Het komt dus overeen met het 90 graden terugdraaien in fase, en het dempen met een factor $1/s$ (met s de frequentie). Het FM signaal lijkt dus sterk op een vervormd PM signaal. We zouden echter ook van te voren deze bewerking precies omgekeerd gedaan kunnen hebben. In dat geval zouden we een vervormd

PM signaal hebben, maar een PM signaal dat als een normaal FM signaal klinkt. Kortom, eigenlijk zijn beide types modulatie compleet identiek, op een standaard high-pass/low-pass filter na. (90 graden fasedraaiing, en verzwakking met $1/s$ komt precies overeen met de stroom door een spoel)

Alles wat er dus over een FM signaal wordt gezegd geldt dus ook voor een PM signaal, zij het ietsje aangepast. Daarom zal ik het verder hebben over PM modulatie, in de wetenschap dat dit toch bijna identiek is aan FM.

Het zou interessant zijn om te weten wat het spectrum van een PM zender is. Daaruit valt dan ook het spectrum van een FM-zender af te leiden, door gebruik te maken van het omzetten zoals hierboven beschreven. We gaan op dezelfde wijze te werk als bij de AM modulatie. Omdat PM lastiger is dan AM zal ik het alleen uitrekenen voor een simpele sinustoon. Voor een algemeen audiosignaal is het veel te complex en ook niet echt nuttig om het echt uit te rekenen. Dat zal ik dan ook niet doen, een simpel piepje is al ruimschoots lastig genoeg.

Om het overzichtelijk te houden zal ik de factoren 2π achterwege laten.

$$\begin{aligned}x(t) &= \sin(f_z t + \alpha \sin f_a t) \\x(t) &= \frac{1}{2} \cos(f_z t) \sin(\alpha \sin f_a t) + \frac{1}{2} \sin(f_z t) \cos(\alpha \sin f_a t)\end{aligned}$$

Hier komt echter het probleem. De factoren $\sin(\alpha \sin f_a t)$ en $\cos(\alpha \sin f_a t)$ vallen niet met goniometrie weg te werken. De enige manier om dit te doen is door middel van een reeksontwikkeling⁷. Voor kleine modulatiediepten ($\alpha \ll 1$) is een goede benadering om te stellen dat $\sin(\alpha \sin f_a t) = \alpha \sin f_a t$ en $\cos(\alpha \sin f_a t) = 1$. Dit is nuttig voor smalbandFM, vanwege de kleine bandbreedte (kleine frequentiezwaai). Het spectrum hiervan is wel exact te berekenen. Men krijgt dan:

$$\begin{aligned}x(t) &= \frac{\alpha}{2} \cos(f_z t) \sin f_a t + \frac{1}{2} \sin(f_z t) \\x(t) &= \frac{\alpha}{4} \sin(f_z t + f_a t) - \frac{\alpha}{4} \sin(f_z t - f_a t) + \frac{1}{2} \sin f_z t\end{aligned}$$

Dit lijkt wel heel erg op het AM signaal. Er is echter een groot verschil, en dat is dat is de tegenfase van de twee zijbanden, dit in tegenstelling tot het AM signaal, waar beiden positief zijn. Door dit min-teken zijn ze niet meer te detecteren door middel van een gewone diodedetector. De beide mengproducten heffen elkaar dan precies op, terwijl ze elkaar normaal juist versterken. In principe zou het met een goede EZB-detector wel gewoon mogelijk moeten zijn, maar er zijn voor FM-demodulatie handiger methoden.

De bandbreedte van een smalbandFM zender is dus ongeveer 2 maal de bandbreedte van het audiosignaal. We zijn hier echter begonnen met een PM-signaal signaal en n of andere mysterieuze factor α . Dit moet nog omgezet worden naar iets dat bruikbaar is voor het berekenen van de frequentiezwaai van een FM-zender. Ik zal er nu verder weer vanuit gaan dat we praten voor een audiospectrum van $300Hz$ tot $3kHz$. Ten eerste moet de factor $\alpha \leq 0.15$ zijn om de benadering te laten kloppen. ($\sin x \approx x$ voor $-0.15 < x < 0.15$ op ongeveer 1%) Hier valt met enige moeite de frequentieverandering uit te

⁷Bij deze methode (om precies te zijn de Taylorreeks), wordt een functie benaderd door het maken van een polynoom met (meestal) oneindig veel termen

berekenen. Het signaal zag er zo uit:

$$x(t) = \sin(f_z t + \alpha \sin f_a t)$$

De fase is dus:

$$\begin{aligned}\Theta(t) &= f_z t + \alpha \sin f_a t \\ f(t) &= \frac{d\Theta}{dt} = f_z + \alpha f_a \cos f_a t\end{aligned}$$

De frequentie zwaait dus $\alpha f_a \leq 0.15 f_a$ rond de zendfrequentie. De frequentiezwaai van een PM gemoduleerde zender bedraagt dus MAXIMAAL ongeveer $0.15 f_a$. Wordt het groter, dan komen er nog meer zijbanden, wat ik straks verder uit zal leggen. Wordt er echter minder gemoduleerd, α kleiner, dan zullen de zijbanden erg zwak zijn. Dus dit is wel een redelijk streefgetal. Dit was echter nog allemaal uitgerekend voor PM-modulatie. Wat betekent dit dan voor FM-modulatie?

Bij FM-modulatie is de frequentiezwaai meestal constant. Je hebt over het algemeen een oscillator waarvan de frequentie afhankelijk is van de spanning, en je zet er gewoon een audiosignaal op. De frequentiezwaai van de zender is dus voor iedere frequentie gelijk aan $\Delta f = 2\alpha f_a$. (f_a is dus de frequentie van het audiosignaal, en Δf hoever de zendfrequentie heen en weer gaat) De laagste frequentie geeft de grootste frequentiezwaai, we moeten de frequentiezwaai dus aanpassen aan de laagste frequentie om $\alpha < 0.15$ te houden. Dat levert bij een audio-bereik van $300 Hz$ op: $\Delta f = 90 Hz$. Berekent men nu α voor een signaal van $3 kHz$, dan komt men erachter dat $\alpha = 0.015$. Even terug naar:

$$x(t) = \frac{\alpha}{4} \sin(f_z t + f_a t) - \frac{\alpha}{4} \sin(f_z t - f_a t) + \frac{1}{2} \sin f_z t$$

Het blijkt dat de zijbanden wel heel erg zwak worden voor de hogere tonen. Daarom worden de hoge tonen meestal wat verder versterkt dan de lage, zodat ook voor de hoge tonen $\alpha = 0.15$. Dit voorkomt veel ruis. De ruis zou anders in de hoge tonen een onevenredig groot zijn. Ook kan vaak α nog wel een klein beetje groter worden om ruis tegen te gaan.

De exacte uitkomst voor een PM signaal is een stuk ingewikkelder, en maakt gebruik van de Taylorreeks de sinus en cosinus:

$$\begin{aligned}\sin x &= \sum_{n=0}^{n=\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\ \cos x &= \sum_{n=0}^{n=\infty} \frac{(-1)^n}{(2n)!} x^{2n}\end{aligned}$$

Waarbij x vervangen wordt door $\alpha \sin f_a t$. Hierdoor wordt de formule het volgende:

$$x(t) = \frac{1}{2} \cos(f_z t) \sum_{n=0}^{n=\infty} \frac{(-1)^n}{(2n+1)!} (\alpha \sin f_a t)^{2n+1} + \frac{1}{2} \sin(f_z t) \sum_{n=0}^{n=\infty} \frac{(-1)^n}{(2n)!} (\alpha \sin f_a t)^{2n}$$

(door in de sommatie niet verder te gaan dan $n = 0$ zie je dat je het smalband-FM signaal krijgt)

Allereerst is het nuttig om de sommaties wat verder uit te werken. Dit is namelijk nog heel er onoverzichtelijk. Men kan de machten van sinus en cosinus uitschrijven zodat men inziet welke frequentiecomponenten er zichtbaar zijn.

$$(\sin x)^{2n+1} = \sum_{i=0}^{i=n} \frac{1}{4^n} \left(\binom{2n}{i} + \binom{2n}{i-1} \right) \sin((2(n-i)+1)x)$$

$$(\sin x)^{2n} = -\binom{2n-1}{n} + \sum_{i=0}^{i=n} \frac{-1^{n-i}}{2 \cdot 4^n} \left(\binom{2n-1}{i} + \binom{2n-1}{i-1} \right) \cos(2(n-i)x)$$

In iedere macht zitten dus verschillende frequentiecomponenten. Het is nu de truc om zo te herschikken in de taylorreeksen dat in de taylorreeks de $\sin x$ zo ver mogelijk naar buiten komt. In dat geval houdt je namelijk voor de taylorreeksen het volgende over:

$$\sin(\alpha \sin f_a t) = \sum_{n=0}^{n=\infty} a_n \sin(2n+1)f_a t$$

$$\cos(\alpha \sin f_a t) = \sum_{n=0}^{n=\infty} b_n \cos(2n)f_a t$$

Met a_n en b_n op hun beurt weer reeksen, die van α afhangen:

$$a_n = \sum_{i=n}^{i=\infty} \frac{(-1)^i}{(2i+1)!} \frac{\alpha^{2i+1}}{4^i} \left(\binom{2i}{i-n} + \binom{2i}{i-n-1} \right)$$

$$b_n = \sum_{i=n}^{i=\infty} \frac{(-1)^i}{(2i)!} \frac{\alpha^{2i}}{2 \cdot 4^i} \left(\binom{2i-1}{i-n} + \binom{2i-1}{i-n-1} \right)$$

$$b_0 = \sum_{i=n}^{i=\infty} \frac{(-1)^i}{(2i)!} \frac{\alpha^{2i}}{2 \cdot 4^i} \binom{2i-1}{i}$$

Zodat $x(t)$ valt om te schrijven naar:

$$x(t) = \frac{1}{2} \cos(f_z t) \sum_{n=0}^{n=\infty} a_n \sin(2n+1)f_a t + \frac{1}{2} \sin(f_z t) \sum_{n=0}^{n=\infty} b_n \cos(2n)f_a t$$

Uiteindelijk valt dit dan natuurlijk om te schrijven naar:

$$x(t) = \frac{1}{4} \sum_{n=0}^{n=\infty} a_n \left[-\sin[(f_z - (2n+1)f_a)t] + \sin[(f_z + (2n+1)f_a)t] \right] + b_n \left[\sin[(f_z - 2nf_a)t] + \sin[(f_z + 2nf_a)t] \right] \quad (15)$$

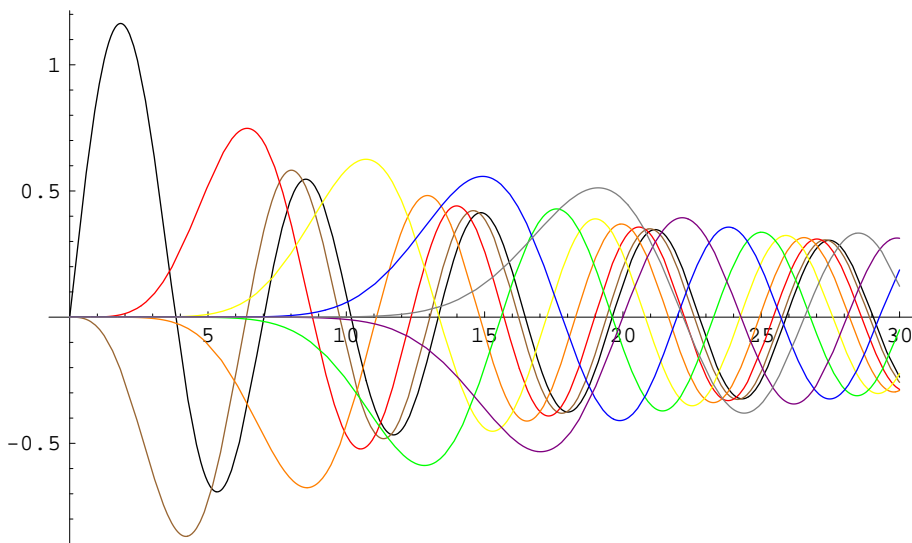
Je ziet hierin duidelijk dat er in principe voor iedere gehele n op $f_z + n$ en $f_z - n$ een signaal uitgezonden wordt. Gelukkig worden de reeksen a_n en b_n vrij snel klein (afhankelijk van α), zodat het spectrum toch nog vrij smal is. Bij de benadering voor kleine α die we eerst maakten zijn alleen $a_1 \neq 0$ en $b_0 \neq 0$, en alle anderen wel gelijk aan 0. Vul je dat in (15), dan zie je dat dit klopt. In de praktijk is het een flink werk om de verschillende a_n en b_n uit te rekenen voor de

verschillende α 's die kunnen optreden in een realistisch audiosignaal. Dit geeft in ieder geval een goed idee van wat er gebeurt.

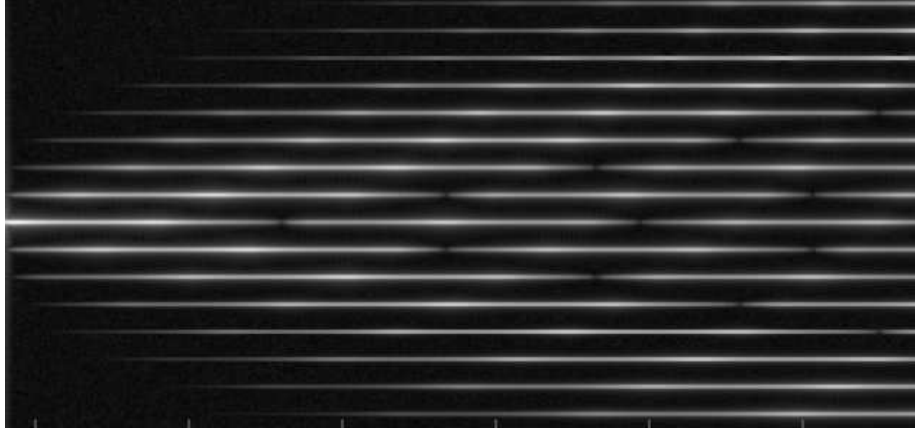
Het zou fijn zijn om voor FM ook weer een regel af te leiden voor de bandbreedte B van het signaal. Dit blijkt echter niet eenvoudig, omdat de $a_n(\alpha)$ en $b_n(\alpha)$ coëfficiënten niet zomaar bepaald kunnen worden uit α . Er is gelukkig nog wel de regel van Carson. Δf is de frequentiezwaai, $f_{a,max}$ de hoogste audiofrequentie:

$$B \cong 2(\Delta f + f_{a,max}) \quad (16)$$

Overigens is er bij het omzetten naar een FM signaal een NIET-lineaire operatie uitgevoerd. Dat wil zeggen dat je dus NIET het HF-signaal $x(t)$ voor alle verschillende audiofrequenties mag uitrekenen om het daarna simpelweg op te tellen. Als $a(t)$ het audiosignaal is, dan zou je na een tijdje $\sin \alpha a(t)$ en $\cos \alpha a(t)$ uit moeten rekenen, het geen duidelijk niet lineair is. Dit in tegenstelling tot alle andere methoden van modulatie die tot nu toe behandeld zijn. (de benadering voor kleine α was trouwens WEL lineair, omdat je daar $\sin \alpha a(t) \cong \alpha a(t)$ en $\cos \alpha a(t) \cong 1$ stelt, en dit laatste is duidelijk wel lineair.



Figuur 28: a_0 tot en met a_8 (met als kleur voor n de betreffende weerstandskleur) geplot voor oplopende α (x-as). Duidelijk is te zien hoe bij oplopende modulatieindex de termen voor steeds hogere n mee gaan tellen (en daarmee dus ook de bandbreedte). Om het vermogen te behouden is bij die grotere bandbreedte de amplitude van iedere individuele frequentiecomponent kleiner.



Figuur 29: Het frequentiespectrum van een PM-gemoduleerde toon, waarbij α steeds verder toeneemt. (de frequentie op de Y-as, en α op de X-as) Duidelijk is te zien dat er steeds meer frequentiecomponenten een rol spelen bij hogere frequenties, en ook is het uitdoven van bepaalde frequenties bij bepaalde α goed te zien.

6.3 modulatie

Het moduleren van een FM of PM signaal is eigenlijk vrij simpel. Alles wat gedaan moet worden is het veranderen van de frequentie van een oscillator. Het handigste is dan om een VCO⁸ te nemen. De meest gebruikelijke methode om een VCO te maken is het gebruik van een varicap-diode. Een varicap-diode is een soort diode die in sperrichting gebruikt wordt. De (normaal parasitaire) capaciteit tussen de twee aansluitingen wordt dan gebruikt in een LC-kring. Hoe hoger de sperspanning over de diode, hoe kleiner de capaciteit. De resonantiefrequentie van de LC-kring is $1/\sqrt{LC}$. En de capaciteit van een varicap hangt vaak lineair⁹ af van de spanning. Dat zou impliceren dat de frequentie niet lineair afhankelijk is van de spanning, wat toch wel de bedoeling is. Gelukkig is de frequentiezwaai over het algemeen echter zeer klein in vergelijking met de zendfrequentie. (op $100MHz$ wordt meestal uitgezonden met een frequentiezwaai van $180kHz$). Door de kleine frequentiezwaai kunnen we de oscillatorfrequentie best als min of meer lineair afhankelijk van de spanning veronderstellen.

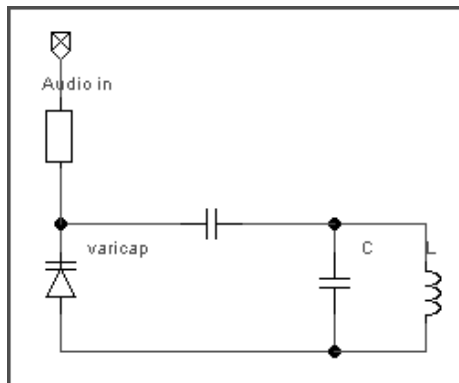
In de meeste zenders is er ook een PLL-schakeling toegepast. Dit zorgt ervoor dat er eigenlijk geen sprake is van echte FM-modulatie, maar van PM-modulatie. Maar zoals gezien maakt dit (met de goede voorzorgen) niet uit voor het audiosignaal.

6.4 demodulatie

Bij bijna iedere vorm van FM-demodulatie wordt eerst een discriminator toegepast. De ideale discriminator is een zeer snelle comparator die precies omschakelt wan-

⁸Voltage Controlled Oscillator, oscillator waarvan de frequentie afhankelijk is van een voltage dat je ergens op de schakeling zet

⁹heel erg lineair is het vaak niet, maar het is in ieder geval niet met een $C = 1/U^2$ verband. (alleen dit zou een echt lineaire frequentie-afhankelijkheid maken)



Figuur 30: Een LC kring waarop een varicap-diode is aangesloten. De resonantiefrequentie hiervan is dus afhankelijk van het audiosignaal.

neer het FM-sigitaal boven of onder de 0V zit. Het uitgangssigitaal is dan een blokgolf, die 1 is wanneer het ingangssigitaal in de positieve halve periode zit, en 0 in de negatieve halve periode. De fase-informatie blijft dus behouden, zoals de bedoeling is. Het uitgangssigitaal lijkt sterk op het ingangssigitaal, het is echter een blokgolf met constante amplitude. En dankzij de constante amplitude die die blokgolf heeft, komt er dus geen AM-modulatie door. Die AM-modulatie is ook niet de bedoeling om door te komen. Als de amplitude verandert, dan is dat het gevolg van ruis, of van vervorming in de ontvanger. Het is dus ook alleen maar goed dat die gefilterd wordt.

Na de discriminator komt de eigenlijke demodulator van het FM-sigitaal. Ik zal 3 methoden behandelen om een FM-sigitaal te demoduleren. Allereerst flankdetectie, dan de standaard FM-detector, en daarna nog "pulsdetectie".

Flankdetectie maakt gebruik van een resonantiekkring. De doorlaatkromme van een resonantiekkring wordt gegeven door:

$$A_{uit} = \frac{A_{in}}{\alpha - (f - f_r)^2}$$

met A_{uit} de amplitude van het sigitaal over de resonantiekkring, A_{in} het ingangssigitaal, α een constante die samenhangt met de Q-factor van de resonantiekkring, en f_r de resonantiefrequentie van de kring en f de frequentie van het sigitaal dat erin gaat.

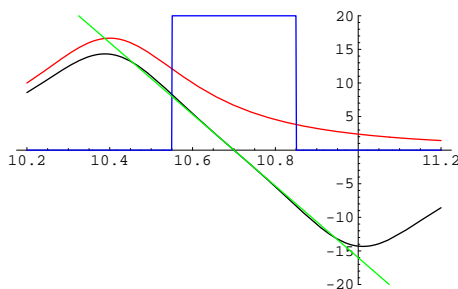
Zoals je ziet is de amplitude van het sigitaal over de resonantiekkring afhankelijk van de frequentie. Dankzij de discriminator is A_{in} van constante grootte. En dus hangt A_{uit} alleen af van de van de frequentie van het ingangssigitaal. Door nu een FM-sigitaal aan ingang te hangen krijgen we dus een amplitude die varieert afhankelijk van de frequentie van het FM-sigitaal. De amplitude A_{uit} van het sigitaal op de LC-kring kan men eenvoudig detecteren met een diode.

Nou zouden we het liefst willen dat de A_{uit} recht evenredig is met de frequentie f . Zoals je ziet is dit niet echt het geval. Tenzij je een sigitaal neemt dat redelijk ver van f_r vandaan zit, en dat niet al te sterk FM-gemoduleerd is. Maar echt mooi lineair zal het nooit worden met deze schakeling.

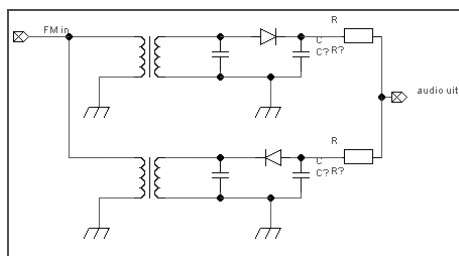
De meest gebruikelijke methode is het om twee flankdetectoren op verschillende frequenties te gebruiken, en de signalen hiervan van elkaar af te trekken. Stel dat we twee flankdetectoren hebben, op $f_{r,1}$ en $f_{r,2}$. We kiezen die frequenties zo dat de signalen die we willen detecteren precies tussen $f_{r,1}$ en $f_{r,2}$ in liggen. Door beide signalen te detecteren en van elkaar af te trekken krijgen we wel een redelijk rechte kromme. Dit is technisch niet zo heel erg lastig, maar geeft toch een goed signaal.

$$A_{uit} = \frac{A_{in}}{\alpha - (f - f_{r,1})^2} + \frac{A_{in}}{\alpha - (f - f_{r,2})^2}$$

Hoewel je het niet meteen aan de formule ziet, zal dit voor een goed gekozen $f_{r,1}$, $f_{r,2}$, en α behoorlijk mooi lineair zijn tussen $f_{r,1}$ en $f_{r,2}$.



Figuur 31: De lineariteit van de dubbele flankdetector (zwart) is duidelijk stukken beter dan die van de enkele (rood). De rest van de FM-ontvanger zorgt ervoor dat al het signaal dat op de ingang komt het blauwe signaal is. Op het gebied waar signaal is, is de detector vrijwel lineair (groen).

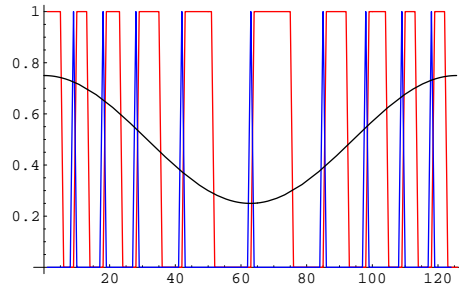


Figuur 32: De demodulator met dubbele flankdetectie.

De laatste methode die ik zal bespreken is een methode voor volledig lineaire detectie. Deze methode wordt amper toegepast (ik stuitte erop in een aflevering van RadioBulletin uit eind jaren 50). In de meeste ontvanger wordt voor FM een middenfrequentie van $10.7MHz$ gebruikt. Dat zou voor pulsdetectie in principe ook wel kunnen, maar het is handiger om op een frequentie van rond de $400kHz$ te werken. (met een frequentiezwaai van $250kHz$ tot $550kHz$ levert dit veel meer signaal op). Dus na de MF trappen op $10.7MHz$ volgt er nog een mixer met een frequentie van $10.3MHz$, zodat het signaal zich in het gewenste frequentiegebied bevindt. Er volgt een low-pass filter, en een discriminator, zodat

we nu een mooi blokvormig signaal in het genoemde frequentiegebied hebben. De truc is nu om op de stijgende periode van de blokgolf een kort pulsje te genereren, met vaste lengte T . De duur van T moet (veel) korter zijn dan periodelijd van de hoogste frequenties die kunnen optreden. In dit voorbeeld zou voor T een waarde van $0.5\mu s$ of minder dus goed zijn.

Het aantal pulsjes per seconde is nu gelijk aan de frequentie van het signaal. De gemiddelde spanning van de pulsjes is (omdat ze allemaal dezelfde lengte hebben) dan ook recht evenredig met de frequentie, en dus ook met de modulatie. Het is dus een vervormingsloze FM-detector.



Figuur 33: Rood is het FM signaal dat uit de discriminator komt. Blauw zijn de pulsen, zwart is de modulatie.

7 Quadratuur Amplitude Modulatie

7.1 algemene informatie

Quadratuur Amplitude Modulatie (QAM) is eigenlijk geen echte modulatiesoort. Het is een verzamelnaam voor signalen opgewekt met een QAM modulator. Het wordt tegenwoordig voor zeer verschillende soorten digitale communicatie gebruikt, van internet tot digitenne.

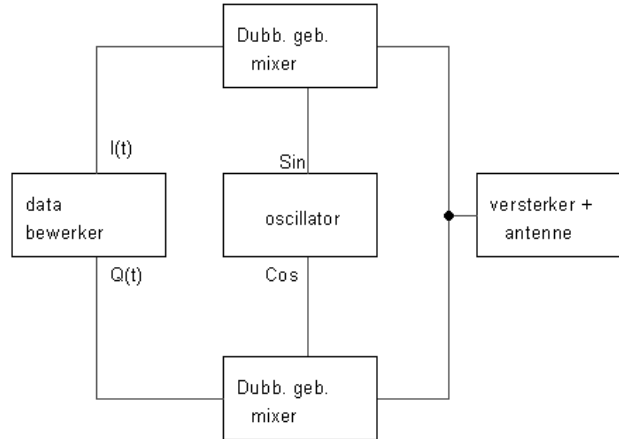
7.2 wiskundige beschrijving

Het principe van een QAM modulator is het volgende: Je neemt twee DSB-zenders op 1 frequentie, waarbij de ene zender 90 graden uit fase loopt met de ander. De twee signalen $I(t)$ en $Q(t)$ kan je zelf kiezen al naar gelang het soort modulatie je wil hebben. In principe komt dit overeen met een 2-dimensionale ruimte¹⁰, de I-Q ruimte, die alle informatie bevat die er uitgezonden kan worden met een zender. Het is dus in principe ook mogelijk om alle vormen van modulatie uit dit ding te halen. In veel gevallen zijn er echter simpeler methoden om dit te doen, en wordt er dus een ander soort modulator toegepast.

Het uitgangssignaal ziet er dus als volgt uit:

$$x(t) = I(t) \sin f_z t + Q(t) \cos f_z t$$

¹⁰Het is niet zomaar 2-dimensionaal omdat ik het leuk vind, maar het is dmv de Fouriertransformatie te bewijzen dat je aan de I- en Q- dimensie genoeg hebt om ieder mogelijk signaal te beschrijven, en dat het niet mogelijk is om bijvoorbeeld een derde dimensie te bedenken. Ieder signaal in die 3e dimensie zou te schrijven zijn als de som van een signaal in de I- en een in de Q- ruimte.



Figuur 34: principe QAMmodulator

Voor een AM-signaal, geldt bijvoorbeeld:

$$\begin{aligned} I_{AM}(t) &= 1 + a(t) \\ Q_{AM}(t) &= 0 \end{aligned}$$

Voor SSB, met $a_{\perp}(t)$ het 90 graden gedraaide audiosignaal $a(t)$:

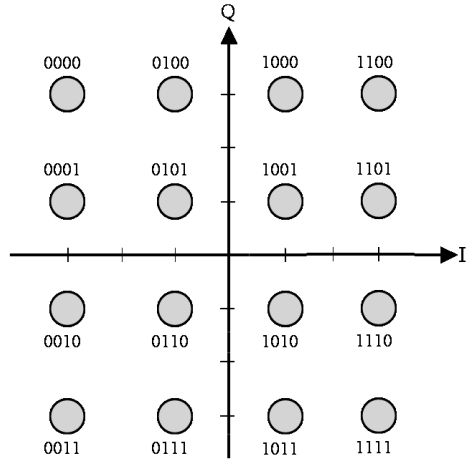
$$\begin{aligned} I_{SSB}(t) &= a(t) \\ Q_{SSB}(t) &= a_{\perp}(t) \end{aligned}$$

En voor FM kan men schrijven:

$$\begin{aligned} I_{FM}(T) &= \sin \int_{t=0}^{t=T} a(t) \\ Q_{FM}(T) &= \cos \int_{t=0}^{t=T} a(t) \end{aligned}$$

Dat is natuurlijk allemaal wel leuk en aardig, maar voor digitale modulatie wordt het wel echt interessant om het zo voor te stellen. (en meestal praat men over digitale QAM modulatie, als men het over QAM heeft.) Stel bijvoorbeeld dat we 4 bits tegelijk willen uitzenden. Dan zenden we bijvoorbeeld de eerste 2 bits uit in de I-ruimte, en de tweede 2 in de Q-ruimte. De eerste 2 bits kunnen de waardes 1,2,3 en 4 aannemen. Daarmee laat ik en amplitude voor I overeenkomen van respectievelijk $-2, -1, 1$ en 2 . Idem voor de tweede 2 bits, die ik in het Q-signaal stop. Zend ik dus bijvoorbeeld 1101 uit, dan krijg ik: $I = 2$ en $Q = -1$. Dat komt dus overeen met $x(t) = 2 \sin f_z t - \cos f_z t$.

De keuze om 4 bits tegelijk uit te zenden was volkomen arbitrair, het kunnen er net zo goed 8 zijn. Ook de plaats op de I en op de Q-as kunnen heel anders gekozen worden. (zodat de efficiency wat hoger wordt). De plaats in de IQ ruimte, waar de verschillende bits liggen wordt wel de 'constellation' genoemd. In het geval van 4 bits, zitten er 16 punten in de constellation, en men spreekt dan van QAM16. Wordt er 1 byte tegelijk verzonden, dan spreekt men van QAM256. Vaak wordt er ook gekozen voor een constellation waar de bits in het rond zitten. Hierdoor is het maximum vermogen wat gelijkjer versdeeld over de



Figuur 35: Constellation voor QAM16

verschillende mogelijke bits. De amplitude van het signaal is namelijk:

$$A(t) = \sqrt{I(t)^2 + Q(t)^2}$$

Een ander punt waar een constellation aan moet voldoen is dat de verschillende punten zo ver mogelijk uit elkaar liggen in de vermogensruimte. Een vermogen wordt gegeven door $P = U^2/R$, en dus $P \propto U^2$, want R is natuurlijk constant. Het vermogensverschil ΔP tussen twee punten $x_1 = (I_1, Q_1)$ en $x_2 = (I_2, Q_2)$ in de IQ-ruimte kan gedefinieerd worden als:

$$\Delta P(x_1, x_2) = (I_1 - I_2)^2 + (Q_1 - Q_2)^2$$

Het gemiddelde vermogen van (witte) ruis is een constante. Wordt er een bepaald signaal uitgezonden, dan zal de ontvanger dat signaal ontvangen, plus een gegeven hoeveelheid ruis. Die ruis is o.a. afhankelijk van de antenne en ontvanger. Wordt er een bepaald vast punt in de IQ-ruimte uitgezonden (x_z), dan zal op de ontvanger een punt in de IQ-ruimte ontvangen (x_o) dat daar lichtelijk van afwijkt dankzij de ruis. Gemiddeld zal het vermogensverschil tussen het te ontvangen punt en het uitgezonden punt precies het ruisvermogen bedragen. Stellen we dus dat $\Delta P = \langle P_{ruis} \rangle$, de gemiddelde ruis en $d(x_o, x_z)$ de afstand in de IQ-ruimte tussen het ontvangstsignaal, en het zendersignaal. Dan vinden we:

$$\langle P_{ruis} \rangle = \langle (I_o - I_z)^2 + (Q_o - Q_z)^2 \rangle = \langle d(x_o, x_z)^2 \rangle$$

Gemiddeld is dus het kwadraat van de afstand van het ontvangen signaal tot het uitgezonden P_{ruis} . Men kan dan zeggen dat de gemiddelde fout-afstand $d(x_o, x_z) \approx \sqrt{P_{ruis}}$. (dit klopt niet precies maar dat maakt niet uit). Om de fout dus zo goed mogelijk te elimineren is het nodig om de verschillende punten zo ver mogelijk uit elkaar te kiezen, want als punten dicht bij elkaar liggen, dan zal de kans op fouten door ruis het grootst zijn. Het is dus belangrijk om alle punten zo ver mogelijk uit elkaar te kiezen. Daarom worden ook nogal eens

andere constellations gebruikt, waar de afstand tussen de punten meer gelijk is verdeeld.

Om de bandbreedte van een QAM signaal te berekenen gaan we weer terug naar:

$$x(t) = I(t) \sin f_z t + Q(t) \cos f_z t$$

In principe valt dit te zien als twee dubbelzijband signalen, en de bandbreedte van een DZB signaal bedraagt 2 maal de bandbreedte van het modulerende signaal. En dus is de bandbreedte B gelijk aan: $B = 2\max(B(I), B(Q))$, met $B(I)$ de bandbreedte van het I-signaal, en $B(Q)$ van het Q-signaal. Maar pas op, dit gaat niet altijd op, door $I(t)$ en $Q(t)$ slim te kiezen (zoals bij SSB) heffen twee zijbanden elkaar precies op, en dan is de bandbreedte slechts $B = \max(B(I), B(Q))$. Het meest algemeen is dus:

$$\max(B(I), B(Q)) \leq B \leq 2\max(B(I), B(Q))$$

Om de bandbreedte van een QAM-zender zoveel mogelijk te beperken moeten dus het Q en I signaal een zo klein mogelijke bandbreedte krijgen. Dat kan weer gewoon met een (al dan niet analoog) low-pass filter gedaan worden. Een nadeel van dat low-pass filter is dat $I(t)$ en $Q(t)$ steeds vrij traag veranderen. Daarom moet de ontvanger steeds precies op het goede moment kijken op welk punt in de IQ-ruimte het zendersignaal zich bevindt. Kijkt de ontvanger te vroeg, dan ziet de ontvanger het zendersignaal nog op een moment dat het niet aangekomen is op het bedoelde punt. Dat resulteert in fouten. Dit probleem treedt natuurlijk ook op wanneer de ontvanger een te kleine bandbreedte heeft. Een al te smalle bandbreedte leidt vanzelf tot fouten.

De hoeveelheid data die kan worden verzonden met een QAM signaal valt nu ook vrij makkelijk te berekenen. Wanneer een n aantal samples worden verzonden per seconde, dan is bij ideaal filteren de bandbreedte van het I en Q signaal nHz . En dus is de totale bandbreedte: $2nHz$. Stel dat er m bits worden uitgezonden per sample, dan hebben we nm bits per seconde. (die overigens niet allemaal nuttig gebruikt kunnen worden, een groot deel zal gereserveerd moeten worden voor foutcorrectie) Praktisch voorbeeld: Stel, ik luister naar een QAM-64 (6-bits) zender met een bandbreedte van 9kHz. Dan is de datasnelheid dus $\frac{9000}{2} \log_2 64 = 27000bps = 27kbps$. Dit klopt wel ongeveer met de DRM uitzendingen op midden en kortegolf.

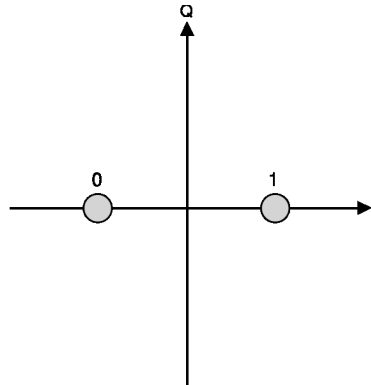
7.3 modulatie

Voor de modulatie van een algemeen QAM signaal wordt eigenlijk meestal een schema gebruikt, waarvan ik het principe hierboven al gegeven heb. Het zou in principe ook mogelijk zijn om slechts de fasehoek, en de amplitude te regelen, dit is meestal echter minder eenvoudig. Verder valt er over de modulatie vrij weinig te zeggen.

7.4 demodulatie

Bij het demoduleren van een QAM-signaal zit je met hetzelfde probleem als bij een SSB-signaal. Er wordt in een QAM signaal namelijk geen draaggolf

uitgezonden, en het is daarom lastig om de draaggolf terug te vinden. En in tegenstelling tot bij een SSB-signaal is het wel strikt noodzakelijk om de draaggolf precies terug te vinden. (SSB is ook nog wel verstaanbaar als je een beetje mis zit). Er zijn verschillende manieren om de draaggolf terug te krijgen voor alle verschillende modulatiemethoden. Over het algemeen zijn ze vrij ingewikkeld en te specifiek om hier goed te behandelen. Ik zal als voorbeeld slechts een PSK (phase shift keying) signaal nemen. Dit is de simpelste vorm van digitale QAM (QAM2). Het is zo simpel dat het eigenlijk geen echte QAM is, want er zitten maar 2 punten in de constellation, 1 op $I = 1$ en de andere op $I = -1$.



Figuur 36: PSK constellation

$$x_{PSK} = D(t) \sin 2\pi f_z t + \theta$$

Met $D(t)$ is 1 voor het zenden van een 1, en $I(t) = -1$ voor het zenden van een 0. De fasehoek θ lijkt volkomen triviaal, maar is straks toch van belang, het is de fasehoek tussen ontvanger en zender. Door de oscillator van de ontvanger een faseverschuiving te laten ondergaan veranderd θ . (dit is misschien een beetje onlogisch gedefinieerd, maar het werkt het makkelijkst)

Stel nu dat de oscillator van de demodulator is afgestemd op f_m , dan zijn de signalen die de databewerking binnekomen het volgende:

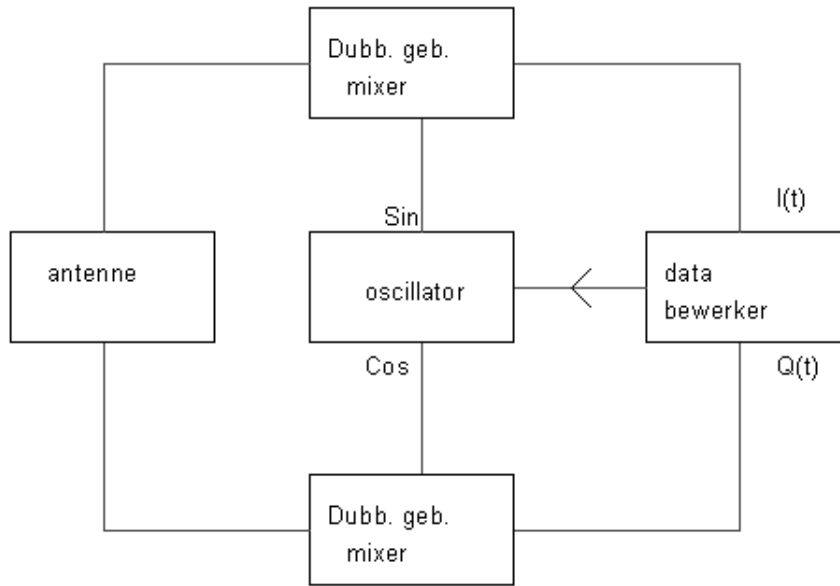
$$\begin{aligned} I(t) &= D(t) (\sin 2\pi f_z t + \theta) \sin 2\pi f_m t \\ Q(t) &= D(t) (\sin 2\pi f_z t + \theta) \cos 2\pi f_m t \end{aligned}$$

$$\begin{aligned} I(t) &= D(t) \cos 2\pi (f_z - f_m)t + \theta \\ Q(t) &= D(t) \sin 2\pi (f_z - f_m)t + \theta \end{aligned}$$

Waarbij de onderste twee signalen al meteen door een low-pass filter zijn gegaan, en verdubbeld.

Alleen wanneer f_z dus precies f_m is, en $\theta = 0$, dan verschijnt het signaal netjes op de I-ingang van de ontvanger. Zodra dat niet zo is, dan gaat het fout. Het probleem is dus duidelijk, het elimineren van $f_z - f_m$, en θ . Laat ik nu een andere hoek, ϕ , definiëren als de fouthoek, en wel als volgt:

$$\begin{aligned} \phi(t) &= 2\pi (f_z - f_m)t + \theta \\ I(t) &= D(t) \cos \phi(t) \\ Q(t) &= D(t) \sin \phi(t) \end{aligned}$$



Figuur 37: algemeen type QAM demodulator

De truc is nu dus om ϕ zo klein mogelijk te houden. (of een veelvoud van $2/\pi$, maar dat maakt verder niet uit) Er zijn 2 veelgebruikte methoden om dat te doen, de eerste orde methode en de tweede orde methode ¹¹ Bij de eerste orde methode wordt de hoek θ steeds zodanig gecorrigeerd dat $\phi = 0$. De tweede orde methode zorgt dat f_m zodanig wordt veranderd dat $\phi = 0$. (een soort PLL systeem).

De eerste orde methode wordt vooral veel gebruikt voor digitale demodulatie, omdat het bij een analoge oscillator nogal lastig is om de fase plots te veranderen. De fasefout ϕ kan vrij eenvoudig gemeten worden. Daartoe vermenigvuldigen we $I(t)$ en $Q(t)$. (hierin is A de amplitude van het ontvangen signaal, om het eenvoudig te houden heb ik die tot nu toe weggelaten)

$$I(t)Q(t) = A^2 D(t)^2 \cos \phi(t) \sin \phi(t)$$

$$I(t)Q(t) = A^2 \frac{1}{2} \sin 2\phi(t) \approx A^2 \phi(t)$$

Merk echter op dat $A^2 = I(t)^2 + Q(t)^2$

$$\phi(t) = \frac{I(t)Q(t)}{I(t)^2 + Q(t)^2}$$

Men kan dit dus direct gebruiken om de oscillator die het f_m signaal opwekt deze fasefout precies tegen te laten gaan, door de oscillator met een grootte van $\phi(t)$ terug te schuiven. Op deze manier kan men dus $D(t)$ vinden. Er zal bij een kleine mis-afstemming van f_m wel steeds een kleine fasefout zijn, maar die is

¹¹er bestaat ook nog een derde orde methode, waar de frequentieverandering $\frac{df_m}{dt}$ wordt geregeld. Dit is van belang bij (snel) bewegende ontvangers of zenders, waar het dopplereffect een rol van betekenis speelt

over het algemeen klein. Men heeft nu dus voor $I(t)$ precies het $D(t)$ gekregen. Dit was de eerste orde methode.

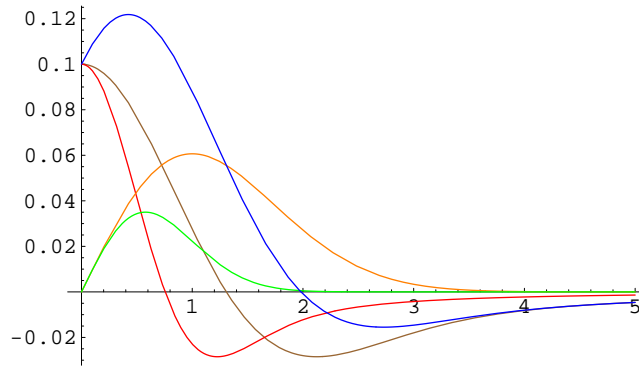
Bij de tweede orde methode wordt op dezelfde manier de fasefout $\phi(t)$ berekend. Nu wordt echter niet meer direct de oscillator van fase veranderd, maar wordt de frequentie f_m veranderd, totdat de fasefout gelijk aan nul is. Dit lijkt erg op een PLL-systeem, maar ik zal er toch nog even op in gaan, vanwege kleine verschillen. Laat $f_z - f_m(t) = \nu(t)$, de frequentiefout op tijdstip t . Bij de tweede orde methode is de frequentie-verandering afhankelijk van de fasehoek, met een bepaalde constante α , afhankelijk van hoe het gebouwd werd.

$$\begin{aligned}\phi(t) &= \nu(t)t + \theta \\ \frac{d\nu}{dt} &= \alpha\phi(t) \\ \frac{d\phi}{dt} &= \frac{d}{dt}(\nu(t)t + \theta) = t\frac{d\nu}{dt} + \nu(t) = \alpha t\phi(t) + \nu(t) \\ \frac{d^2\phi}{dt^2} &= 2\alpha\phi(t) + \alpha t\frac{d\phi}{dt}\end{aligned}$$

Deze differentiaalvergelijking valt jammer genoeg niet exact op te lossen zonder gebruik te maken van een exotische functie als de $ERF(T) = \int_0^t e^{-t^2} dt$. De oplossing is dan:

$$\phi(t) = \frac{\nu t}{\alpha} e^{-\frac{\alpha t^2}{2}} + \theta \left(1 + \sqrt{\frac{\pi\alpha}{2}} t e^{-\frac{\alpha t^2}{2}} ERF\left(\sqrt{\frac{\alpha}{2}} t\right)\right)$$

) Met θ weer de fasefout op $t = 0$, en ν de frequentiefout op $t = 0$. Deze functie is erg complex en is ook maar een benadering van wat er echt gebeurt. (o.a. omdat er een eind terug al gebruikt is dat $0.5 \sin 2\phi \approx \phi$, en omdat het een erg eenvoudig model is, waar de terugkoppeling direct en zonder vertraging werkt) Toch is het handig om te kijken wat er gebeurt, omdat dit soort problemen redelijk algemeen is, en ook vaker voorkomt in de elektronica. Je ziet duidelijk



Figuur 38: $\phi(t)$, de fasefout, bij verschillende beginvoorwaarden α , θ en ν .
 Bruin: -1;0.1;0 Rood: -3;0.1;0 Oranje: -1;0;-0.1 Groen: -1;0;-0.1 Blauw: -1;0.1;0.1

dat de fasefout steeds naar 0 toe neigt. Het zou misschien voor de hand liggen om α zo groot mogelijk te kiezen om de fasefout zo snel mogelijk weg te krijgen. Maar dat maakt het erg storingsgevoelig, en vertragingen in loop kunnen het

ook om zeep helpen. Te klein is daarentegen ook niet goed, omdat dan ϕ erg groot kan worden, en wanneer ϕ groot wordt dan klopt de uitkomst weer niet meer. Dit is dus de tweede orde methode om de $D(t)$ te verkrijgen.

We hebben nu dus $D(t)$ binnen. Maar dat is nog niet alles, om de data er ook echt uit te halen is het nodig om te weten op welke tijdstippen er bits verzonden worden, en wanneer de omschakelperiode is. Over het algemeen zal $D(t)$ sterk gefilterd zijn om bandbreedte te besparen. Het zal overigens meestal wel bekend zijn welke datarate (f_d bits per seconde) er gebruikt wordt. Er zijn vele methoden te bedenken om dit te doen. Het is bijvoorbeeld mogelijk om te kijken naar $D(t)^2$. Merk op dat er in $D(t)$ een vrij sterke component moet zitten met frequentie f_d . Na goede filtering is dit dan ook meteen de hoogste frequentiecomponent in het $D(t)$ signaal. Dit betekent dat wanneer $D(t)^2$ bekeken wordt de hoogste (en ook zeer sterke) frequentie $2f_d$ bedraagt. Door $D(t)^2$ sterk te filteren rond $2f_d$ kan dit signaal een veel grotere sterkte krijgen dan het ruisiger $D(t)$ signaal. Het is dus goed te doen om precies de pieken in het gefilterde $D(t)^2$ signaal te ontdekken. Wanneer dit signaal zijn hoogste punt bereikt, is het ook het beste moment om naar $D(t)$ te kijken, omdat dit dan als het goed is ook op zijn hoogst is.

Er zit nog wel een addertje onder het gras. Het is niet mogelijk om te bepalen of $D(t)$ nu precies een 1 of -1 uitzend. Doordat men $D(t)^2$ bekijkt wordt er helemaal niet gekeken of $D(t)$ nu -1 of 1 is. Het zal dan ook niet duidelijk zijn of er nu een 1 of een 0 wordt uitgezonden. Daarvoor moet met de data op zo'n manier aanpassen dat het voor de ontvanger zelf te bepalen valt of er een 1 of 0 wordt uitgezonden. Het veranderen van $D(t)$ zal altijd een duidelijk op te merken signaal zijn. Daarom wordt er vaak gebruik van DPSK (differential phase shift keying). Dan wordt er een 1 uitgezonden wanneer $D(t)$ omklapt, en er wordt een 0 uitgezonden, wanneer $D(t)$ gelijk blijft.

Merk op dat deze demodulatiemethode in principe ook gewoon op een AM of DSB-signaal kan worden toegepast. Dat levert precies hetzelfde op. Zolang alleen het I-signaal maar gemoduleerd wordt, en er geen Q-signaal wordt uitgezonden, dan valt op deze manier de draaggolf te vinden. Een iets aangepaste methode wordt gebruikt voor de modulatie van een TV-zender. De modulatiemethode die daar wordt toegepast is restzijbandmodulatie, VSB (Vestibul Sideband Modulatioen). In principe zend een TV-zender een DSB-signaal uit, waarvan de onderste zijband voor een groot deel is weggefilterd om bandbreedte en vermogen te sparen. De bovenste zijband bedraagt zo'n 5MHz, en de onderste ongeveer 1MHz. Men kan hieruit weer de draaggolf reconstrueren door het I en Q-signaal te filteren, zodat beiden een breedte hebben van iets minder dan die 1MHz. Voor het I en Q signaal zien nu dus de bovenste en onderste zijband er hetzelfde uit. (indien de mixer al redelijk in de buurt staat afgestemd) Door weer dezelfde methode te gebruiken als voor PSK kan men precies de draaggolf reconstrueren. Dit is zeer belangrijk, want in een analoog TV-signaal is veel belangrijke informatie opgeslagen in de laagste frequenties, en deze mogen absoluut niet verschoven worden.