

<u>Allgemeines</u>

zurück zum Inhalt

Unter Modulation versteht man das Transferieren eines Informationen tragenden Signals in einen anderen (höheren) Frequenzbereich [3].

Es gibt verschiedene Gründe, eine solche Frequenzverschiebung vorzunehmen (nach [1]):

- Anpassung des zu übertragenden Signals (welches die niederfrequente (NF) Information enthält) an die Übertragungsstrecke (Freiraum, Kabel) durch Benutzung verschiedener Trägerfrequenzen (KW, MW, UKW)
- Mehrfachausnutzung von Übertragungsstrecken
- Erhöhung der Störsicherheit

Verwendet man ein sinusförmiges Trägersignal, so kann man dessen Amplitude, Phase oder Frequenz gezielt beeinflussen (modulieren). Dementsprechend gibt es für Sinusträger *Amplitudenmodulation* (AM), *Phasenmodulation* (PM) und *Frequenzmodulation* (FM). AM wird im MW-, KW- und LW-Bereich verwendet, Phasen- und Frequenzmodulation z. B. im UKW-Rundfunk.

Die AM hat allerdings z. B. gegenüber der FM den Nachteil, dass die Amplitude eines Signals bei der Übertragung leichter gestört werden kann als dessen Frequenz.

Bei der AM wird mit der im zeitlichen Verlauf eines niederfrequenten (NF) Signals (z. B. Sprachsignal eines Mikrofons) steckenden Information die Amplitude eines hochfrequenten Signals (Trägersignals) beeinflusst. Dieses hochfrequente Signal wird übertragen, und nach dem Empfang wird das NF-Signal - also die Information, um die es eigentlich geht - durch ein geeignetes Demodulationsverfahren aus dem modulierten Signal extrahiert.

Wir betrachten zunächst den einfachsten Fall, nämlich einen hochfrequenten Sinusträger und ein ebenso sinusförmiges modulierendes NF-Signal. An dieser Stelle soll betont werden, dass das Verhältnis von Trägerfrequenz zu maximaler Signalfrequenz in der Realität größer ist als in den vorgestellten Beispielen.

Signale und Spektren

zurück zum Inhalt

Modulierendes NF-Signal:

 $x_{NF}(t, \omega_{NF}) := cos(\omega_{NF} \cdot t)$

 $x_T(t, \omega_T) := \cos(\omega_T \cdot t)$

Trägersignal:

Das NF-Signal soll nun die Amplitude des Trägers modulieren; daher definieren wir als amplitudenmoduliertes Signal

 $x_{AM}(t, \omega_T, \omega_{NF}, M) := (1 + M \cdot x_{NF}(t, \omega_{NF})) \cdot x_T(t, \omega_T)$

Dabei ist M der sogenannte Modulationsgrad.

Die drei Signale werden im folgenden dargestellt; der Modulationsgrad kann über einen Schieberegler zwischen 0 und 1 eingestellt werden):



Betrachten wir beide Signale im Frequenzbereich, so stellen wir fest:

Während das NF-Signal nur die (Kreis-)Frequenz ω_T enthält (bei komplexer Betrachtung $\pm \omega_T$), gibt es im modulierten Signal drei Frequenzen, nämlich außer der Trägerfrequenz ω_T noch die beiden Seitenbandfrequenzen $\omega_T - \omega_{NF}$ und $\omega_T + \omega_{NF}$, wie eine einfache trigonometrische Umformung zeigt:

$$\begin{aligned} x_{AM}(t,\omega_T,\omega_{NF},M) &= \left(1 + M \cdot x_{NF}(t,\omega_{NF})\right) \cdot x_T(t,\omega_T) = \left(1 + M \cdot \cos(\omega_{NF} \cdot t)\right) \cdot \cos(\omega_T \cdot t) \\ x_{AM}(t,\omega_T,\omega_{NF},M) &= \cos(\omega_T \cdot t) + \frac{M}{2} \cdot \left[\cos\left[\left(\omega_T - \omega_{NF}\right) \cdot t\right] + \cos\left[\left(\omega_T + \omega_{NF}\right) \cdot t\right]\right] \end{aligned}$$

Diese Summe ist natürlich i. a. keine (endliche) Fourierreihe. Das modulierte Trägersignal ist ja i. a. nicht periodisch!



1!) ist das modulierte Signal periodisch mit der Periodendauer 1. Wir können daher eine Fourierreihenentwicklung vornehmen. Dazu berechnen wir die entsprechenden komplexen Fourierkoeffizienten für $n \ge 0$ (jene für n < 0 ergeben sich daraus durch komplexe Konjugation):

$$N_{max} := 110 \qquad n := 0 .. N_{max} \qquad c_{NFI_n} := \int_0^1 x_{NF_Band}(t) \cdot e^{-j \cdot n \cdot 2 \cdot \pi \cdot t} dt \qquad c_{NF2} := \overline{c_{NF1}}$$
$$c_{AMI_n} := \int_0^1 x_{AM}(t, \omega_T, x_{NF_Band}, M) \cdot e^{-j \cdot n \cdot 2 \cdot \pi \cdot t} dt \qquad c_{AM2} := \overline{c_{AM1}}$$

In der grafischen Darstellung lassen wir die beiden Koeffizienten, die zu $-f_T$ und f_T gehören (f_T muss also ganzzahlig sein!) wegen ihrer besonderen Größe weg:



Man sieht deutlich die Aussage von oben etwas allgemeiner bestätigt: durch die Modulation werden die Frequenzanteile des NF-Signals um $\#_T$ verschoben.

Modulationsvorgang

zurück zum Inhalt

Wie geht die Amplitudenmodulation nun konkret vor sich? Prinzipiell kann dazu jede nichtlineare Kennlinie verwendet werden. Wir demonstrieren das Verfahren am Beispiel einer Diodenkennlinie (in Wirklichkeit liegt natürlich ein Vorwiderstand in Serie!).

Zunächst wird die (durchaus realistische) Diodenkennlinie definiert und grafisch dargestellt:

$$I(U) := 10^{-11} \cdot \left(e^{\frac{U}{.025}} - 1 \right)$$

U := -0.1, -0.1 + 0.001..0.7



An der Diode liegen in Serie eine Gleichspannung U_0 und die beiden Wechselspannungen:

 $u(t) = U_0 + u_s \cdot \cos(\omega_T \cdot t) + u_s \cdot \cos(\omega_{NF} \cdot t)$

Durch die Krümmung der Kennline wird der gleichmäßige Verlauf der Spannung (blau) verzerrt; die Amplitude des Stromes (rot) verändert sich. Das folgende Video zeigt diesen Sachverhalt:



Für den Fall einer quadratischen Kennlinie $I(u, c_1, c_2) = c_1 \cdot U + c_2 \cdot U^2$ erhält man mit $u_{NF}(t) = u_{NF} \cdot cos(\omega_{NF} \cdot t)$ $u_T(t) = u_T \cdot cos(\omega_T \cdot t)$ $u(t) = U_0 + u_T(t) + u_{NF}(t) = U_0 + u_T \cdot cos(\omega_T \cdot t) + u_{NF} \cdot cos(\omega_{NF} \cdot t)$

folgenden Stromverlauf:

$$i(t) = c_1 \cdot u(t) + c_2 \cdot u(t)^2$$

$$i(t) = c_1 \cdot \left(U_0 + u_T \cdot \cos(\omega_T \cdot t) + u_{NF} \cdot \cos(\omega_{NF} \cdot t)\right) + c_2 \cdot \left(U_0 + u_T \cdot \cos(\omega_T \cdot t) + u_{NF} \cdot \cos(\omega_{NF} \cdot t)\right)^2$$

Betrachten wir den letzten Term etwas genauer

$$(U_0 + u_T \cdot \cos(\omega_T \cdot t) + u_{NF} \cdot \cos(\omega_{NF} \cdot t))^2 = U_0^2 + u_T^2 \cdot \cos(\omega_T \cdot t)^2 + u_{NF}^2 \cdot \cos(\omega_{NF} \cdot t)^2 \dots + 2 \cdot U_0 \cdot u_T \cdot \cos(\omega_T \cdot t) \dots + 2 \cdot U_0 \cdot u_{NF} \cdot \cos(\omega_{NF} \cdot t) \dots + 2 \cdot u_T \cdot u_{NF} \cdot \cos(\omega_T \cdot t) \cdot \cos(\omega_{NF} \cdot t)$$

so können wir daraus unter Anwendung der passenden Summensätze folgendes Teilergebnis ableiten:

$$(U_0 + u_T \cdot \cos(\omega_T \cdot t) + u_{NF} \cdot \cos(\omega_{NF} \cdot t))^2 = U_0^2 + \frac{u_{NF}^2}{2} + \frac{u_T^2}{2} \dots$$

$$+ 2 \cdot U_0 \cdot u_{NF} \cdot \cos(\omega_{NF} \cdot t) \dots$$

$$+ \frac{u_{NF}^2}{2} \cdot \cos(2 \cdot \omega_{NF} \cdot t) \dots$$

$$+ 2 \cdot U_0 \cdot u_T \cdot \cos(\omega_T \cdot t) \dots$$

$$+ \frac{u_T^2}{2} \cdot \cos(2 \cdot \omega_T \cdot t) \dots$$

$$+ u_T \cdot u_{NF} \cdot \cos[(\omega_T - \omega_{NF}) \cdot t] \dots$$

$$+ u_T \cdot u_{NF} \cdot \cos[(\omega_T + \omega_{NF}) \cdot t] \dots$$

Für den Stromverlauf ergibt sich damit insgesamt

$$\begin{split} i(t) &= c_1 \cdot U_0 + c_2 \cdot \left(U_0^2 + \frac{u_{NF}^2}{2} + \frac{u_T^2}{2} \right) \dots \\ &+ \left(c_1 + 2 \cdot c_2 \cdot U_0 \right) \cdot u_{NF} \cdot \cos(\omega_{NF} \cdot t) \dots \\ &+ c_2 \cdot \frac{u_{NF}^2}{2} \cdot \cos(2 \cdot \omega_{NF} \cdot t) \dots \\ &+ \left(c_1 + 2 \cdot c_2 \cdot U_0 \right) \cdot u_T \cdot \cos(\omega_T \cdot t) \dots \\ &+ c_2 \cdot \frac{u_T^2}{2} \cdot \cos(2 \cdot \omega_T \cdot t) \dots \\ &+ c_2 \cdot u_T \cdot u_{NF} \cdot \cos\left[\left(\omega_T - \omega_{NF} \right) \cdot t \right] \dots \\ &+ c_2 \cdot u_T \cdot u_{NF} \cdot \cos\left[\left(\omega_T + \omega_{NF} \right) \cdot t \right] \dots \end{split}$$

In dieser Darstellung sind die einzelnen Frequenzanteile sofort zu erkennen: i(t) enthält

- einen Gleichanteil
- die NF-Frequenz
- die doppelte NF-Frequenz
- die doppelte Trägerfrequenz
- die Trägerfrequenz
- und die beiden Seitenbänder bei $\omega_T \omega_{NF}$ und $\omega_T + \omega_{NF}$

wobei die drei für uns wesentlichen Terme am Schluss stehen. Da die Abstände zu den unerwünschten Frequenzen groß sind, lassen sich diese leicht wegfiltern.

Nebenbemerkung: Jede nichtlineare Kennlinie lässt sich natürlich lokal (d. h. bei kleinen Signalamplituden) durch eine quadratische Kennlinie näherungsweise ersetzen (z. B. Taylorentwicklung bis zur Ordnung 2).

Ein Zahlenbeispiel soll das abgeleitete Ergebnis verdeutlichen:

Trägerfrequenz und Modulationsfrequenz:

 $f_T = 50 \qquad \qquad f_{NF} = 4$

Definition der Zeitsignale:

 $U_0 := 3$ $u_T := 2$ $u_{NF} := 1$

 $u_T(t) := u_T \cdot cos(\omega_T \cdot t) \qquad u_{NF}(t) := u_{NF} \cdot cos(\omega_{NF} \cdot t) \qquad u(t) := U_0 + u_{NF}(t) + u_T(t)$ Definition der Kennlinie: $c_I := 1 \qquad c_2 := 0.5 \qquad i(t) := c_I \cdot u(t) + c_2 \cdot u(t)^2$



Zum Vergleich wird einerseits die zugehörige Fourierreihe (ohne Gleichanteil) berechnet

$$i_{Fourier}(t) := 2 \cdot Re\left(\sum_{n=1}^{N_{max}} c_{iI_n} \cdot e^{j \cdot 2 \cdot \pi \cdot n \cdot t}\right)$$

und andererseits eine "tiefpassgefilterte" Fourierdarstellung, welche dieselben Fourierkoeffizienten verwendet, jedoch nur die Frequenzen f_T , $f_T - f_{NF}$, $f_T + f_{NF}$ enthält und dementsprechend den erwarteten symmetrischen Verlauf besitzt (siehe unten!):



Zum Schluss soll uns noch die wichtige Frage beschäftigen, wie man aus dem empfangenen modulierten Signal das ursprüngliche NF-Signal zurückgewinnen kann.

Demodulation

zurück zum Inhalt

Dafür gibt es im wesentlichen zwei Möglichkeiten:

- die Synchrondemodulation
- die Hüllkurvendemodulation

Die *Synchrondemodulation* ist besonders mathematisch interessant. Dabei wird das modulierte Signal mit dem Trägersignal (synchron) multipliziert, was (ähnlich wie bei der Multiplikation des NF-Signals mit dem Träger) eine neuerliche Verschiebung der beiden Seitenbänder um die Trägerfrequenz nach "links" und "rechts" bewirkt. Dadurch entsteht (neben noch höherfrequenten Spektralteilen) das ursprüngliche Spektrum an der ursprünglichen Position. Filtert man die falschen Spektralanteile weg, bleibt das Originalspektrum (allerdings verkleinert) übrig, was natürlich im Zeitbereich dem Originalsignal entspricht. Zur Synchrondemodulation muss vor der Demodulation der Träger aus dem Signal rückgewonnen oder am Empfangsort exakt noch einmal (eben synchron) erzeugt werden.

Das Prinzip wird zuerst wieder am einfachsten Beispiel gezeigt:

Modulierendes NF-Signal:
$$x_{NF}(t, \omega_{NF}) = cos(\omega_{NF} \cdot t)$$
Trägersignal: $x_T(t, \omega_T) = cos(\omega_T \cdot t)$ Moduliertes Signal: $x_{AM}(t, \omega_T, \omega_{NF}, M) = (1 + M \cdot x_{NF}(t, \omega_{NF})) \cdot x_T(t, \omega_T)$

Wie bereits früher gezeigt kann man dafür schreiben

oder

$$x_{AM}(t,\omega_T,\omega_{NF},M) = \cos(\omega_T \cdot t) + \frac{M}{2} \cdot \left[\cos\left[(\omega_T - \omega_{NF}) \cdot t\right] + \cos\left[(\omega_T + \omega_{NF}) \cdot t\right]\right]$$

Synchronmultipliziertes Signal:

$$x_{Sync_mult}(t, \omega_T, \omega_{NF}, M) = x_{AM}(t, \omega_T, \omega_{NF}, M) \cdot x_T(t, \omega_T)$$

$$x_{Sync_mult}(t, \omega_T, \omega_{NF}, M) = (1 + M \cdot x_{NF}(t, \omega_{NF})) \cdot x_T(t, \omega_T)^2$$

Unter Verwendung des Ergebnisses

$$x_{AM}(t,\omega_T,\omega_{NF},M) = \cos(\omega_T \cdot t) + \frac{M}{2} \cdot \left[\cos\left[\left(\omega_T - \omega_{NF}\right) \cdot t\right] + \cos\left[\left(\omega_T + \omega_{NF}\right) \cdot t\right]\right]$$

formen wir den Term weiter um:

$$x_{Sync_mult}(t, \omega_T, \omega_{NF}, M) = cos(\omega_T \cdot t)^2 + \frac{M}{2} \cdot \left[cos\left[(\omega_T - \omega_{NF}) \cdot t\right] + cos\left[(\omega_T + \omega_{NF}) \cdot t\right]\right] \cdot cos(\omega_T \cdot t)$$

$$\begin{aligned} x_{Sync_mult}(t, \omega_T, \omega_{NF}, M) &= \frac{1}{2} \cdot \left(1 + \cos(2 \cdot \omega_T \cdot t)\right) \dots \\ &+ \frac{M}{4} \cdot \left[\cos(\omega_{NF} \cdot t) + \cos\left[\left(2 \cdot \omega_T - \omega_{NF}\right) \cdot t\right]\right] \dots \\ &+ \frac{M}{4} \cdot \left[\cos(\omega_{NF} \cdot t) + \cos\left[\left(2 \cdot \omega_T + \omega_{NF}\right) \cdot t\right]\right] \end{aligned}$$

Durch die synchrone Multiplikation mit dem Träger entsteht also das ursprüngliche Spektrum (multipliziert mit dem Faktor $\frac{M}{2}$). Außerdem werden Spektralanteile bei den weit davon entfernten Frequenzen $2 \cdot f_T$, $2 \cdot f_T - f_{NF}$ und $2 \cdot f_T + f_{NF}$ erzeugt, die sich aber leicht herausfiltern lassen.

Jetzt verwenden wir als NF-Signal wieder die Sinusüberlagerung von früher; wir definieren das synchron-multiplizierte Signal

$$x_{Sync}(t, \omega_T, x_{NF_Band}, M) \coloneqq x_{AM}(t, \omega_T, x_{NF_Band}, M) \cdot x_T(t, \omega_T)$$

und berechnen seine komplexen Fourierkoeffizienten:

$$c_{Sync}{}_{n} := \int_{0}^{1} x_{Sync}(t, \omega_{T}, x_{NF}Band, M) \cdot e^{-j \cdot n \cdot 2 \cdot \pi \cdot t} dt \qquad c_{Sync}{}_{n} := \overline{c_{Sync}}{}_{n}$$

Robert Salvador

 $c_{Sync2_n} := c_{Sync_n}$





Mit "Synchronmultipliziertes Signal vor Tiefpassfilterung" ist gemeint, dass alle Spektralanteile enthalten sind (die Summe reicht von 0 bis N_{max} ; bei "nach Tiefpassfilterung" verzichten wir auf die Spektralbereiche um $\pm 2 \cdot f_T$. Die Summenbildung endet hier bei f_{max} . Diese Tiefpassfilterung kommt einer Mittelwertbildung gleich, wie aus dem Bild oben deutlich hervorgeht. Damit ist das ursprüngliche NF-Signal bis auf einen Gleichanteil und eine Zahlenfaktor exakt wiedergewonnen. Dies zeigt auch das nächste Diagramm, in dem die Differenz zwischen der ursprünglichen Einhüllenden und dem verstärkten demodulierten Signal dargestellt ist. Die Unterschiede bewegen sich im Bereich der Rechenungenauigkeit.



Bei der zweiten Methode, der *Hüllkurvendemodulation* geht man einen ähnlichen Weg. Statt einer Multiplikation mit dem Träger wird hier jedoch der Betrag (Vollweg- oder auch Einweggleichrichtung) des übertragenen Signals gebildet; danach werden die hochfrequenten Anteile herausgefiltert, was wiederum einer Mittelwertbildung gleichkommt.

Eine prinzipielle Schaltung, die das leistet, wird weiter unten vorgestellt.

Zunächst zur Demonstration, bei der praktisch derselbe Weg eingeschlagen wird wie bei der Synchrondemodulation:

Betragsbildung (Einweggleichrichtung):
$$x_{H\ddot{u}ll}(t, \omega_T, x_{NF}Band, M) := temp \leftarrow x_{AM}(t, \omega_T, x_{NF}Band, M)$$

 $temp \leftarrow temp \cdot \frac{sign(temp) + 1}{2}$
 $temp$





$$i_D(t) = i_R(t) + i_C(t) = \frac{u_a(t)}{R} + \frac{d}{dt}q_C(t) = \frac{u_a(t)}{R} + C \cdot \frac{d}{dt}u_C(t) = \frac{u_a(t)}{R} + C \cdot \frac{d}{dt}u_a(t)$$

Dies ergibt, oben eingesetzt:

$$u_e(t) = \left(\frac{u_a(t)}{R} + C \cdot \frac{d}{dt}u_a(t)\right) \cdot R_D + u_a(t)$$

Dabei ist natürlich zu beachten, dass der Diodenwiderstand R_D nicht konstant ist sondern von der Diodenspannung bzw. indirekt von Ein- und Ausgangsspannung abhängt:

$$R_D = R_D(u_D) = \frac{u_D(t)}{i_D(t)} = \frac{u_e(t) - u_a(t)}{i_D(t)}$$

Nun wird die Zeit diskretisiert und der Differenzialquotient durch einen Differenzenquotienten ersetzt, sodass eine Differenzengleichung entsteht:

$$u_{e_n} = \frac{R_D}{R} \cdot u_{a_n} + R_D \cdot C \cdot \frac{u_{a_n} - u_{a_{n-1}}}{\Delta t} + u_{a_n}$$
$$u_{a_n} = \frac{u_{e_n} \cdot R \cdot \Delta t + R \cdot C \cdot R_D \cdot u_{a_{n-1}}}{R \cdot \Delta t + R_D \cdot (\Delta t + R \cdot C)}$$

Bei bekannter Diodenkennlinie kann man den Diodenwiderstand als Funktion der Diodenspannung definieren und in die Gleichung einsetzen:

$$u_{a_n} = \frac{u_{e_n} \cdot R \cdot \Delta t + \tau \cdot R_D \left(u_{e_{n-1}} - u_{a_{n-1}}\right) \cdot u_{a_{n-1}}}{R \cdot \Delta t + R_D \left(u_{e_{n-1}} - u_{a_{n-1}}\right) \cdot \left(\Delta t + \tau\right)}$$

Als Diodenkennlinie wird jene von oben verwendet. Damit kann der Diodenwiderstand definiert werden:

$$R_D(U) := wenn\left(I(U) > 0, \frac{U}{I(U)}, 2.5 \cdot 10^9\right)$$

U := -0.1, -0.1 + 0.001..0.8



Zeit := 1

Wahl der Schaltungsparameter:

$$R := 10000 \qquad \qquad C := 2.2 \cdot 10^{-6} \qquad \qquad \tau := R \cdot C \qquad \qquad \tau = 0.022$$

Berechnung des Zeit-Diskretisierungsintervalls und der Punktezahl:

$$T_T := \frac{1}{f_T} \qquad \Delta t := \min\left(\frac{\tau}{100}, \frac{T_T}{100}\right) \qquad \qquad N := \frac{Zeit}{\Delta t} \qquad \qquad \Delta t = 2 \times 10^{-4} \qquad \qquad N = 5 \times 10^3$$

Definition der Eingangsspannung

$$n := 0..N \qquad u_{e_n} := x_{AM} (n \cdot \Delta t, \omega_T, x_{NF}_{Band}, M)$$

Rekursive Berechnung des Ausgangssignals:

$$u_{a_0} := u_{e_0} \qquad u_{a_{n+1}} := \frac{u_{e_n} \cdot R \cdot \Delta t + \tau \cdot R_D (u_{e_n} - u_{a_n}) \cdot u_{a_n}}{R \cdot \Delta t + R_D (u_{e_n} - u_{a_n}) \cdot (\Delta t + \tau)} \qquad m := 0 \dots N \qquad t := 0,0.001 \dots Zeit$$



Die hochfrequenten Signalanteile werden (bei richtiger Dimensionierung von *R* und *C*) stark verkleinert, das Originalsignal daher einigermaßen erkennbar rekonstruiert. Das Ergebnis ist aber dennoch unbefriedigend. Gerade hier ist aber zu beachten, dass in der Realität, wie schon erwähnt, die Trägerfrequenz gegenüber der höchsten Modulationsfrequenz wesentlich größer ist als bei unseren Demonstrationen. Dadurch werden die zeitlichen "Lücken" zwischen den einzelnen Maxima viel kürzer und lassen sich durch Tiefpassfilterung viel besser "auffüllen". Die Präzision bei der Demodulation steigt damit beträchtlich.