

Zur Trägerableitung bei höherstufigen QAM-Signalen

On Carrier Regeneration for Higher Level QAM-Signals

Von Michael H. W. Hoffmann*

Übersicht:

Die Trägerrückgewinnung eines QAM-Signals erfolgt üblicherweise mit Hilfe eines Regelkreises. Wesentlicher Bestandteil dieses Kreises ist ein Phasendetektor. Ausgehend von den Anforderungen an die Eigenschaften eines guten Phasendetektors werden zunächst bekannte Verfahren zur Trägerregeneration untersucht. Als Alternative wird dann ein neues Verfahren vorgestellt, das den Anforderungen insbesondere bei Anwendung auf höherstufige QAM-Signale besser genügt. Das Verfahren stützt sich auf bedingte Entscheidungen bezüglich Betrag und Phase der komplexen Hüllkurve des Eingangssignals. Es wird eine Realisierung vorgeschlagen, die sich auch gut zur Anwendung bei hohen Datenraten eignet.

Abstract:

Carrier regeneration for QAM signals is normally performed in a control loop. An essential part of this loop is a phase detector whose output signal is a function of the difference in phase between the signal of the local oscillator and the unmodulated carrier of the received signal. Known methods for carrier regeneration are studied based on the functional characteristics required of a good phase detector. Then, as an alternative, a new method is introduced which provides a better performance particularly with respect to higher level QAM systems. The method uses conditional decisions referring to the absolute value and the phase of the complex envelope of the received signal. A hardware implementation adapted to high data transfer rates is suggested.

Für die Dokumentation:

Quadraturamplitudenmodulation / Trägerregeneration / Synchronisation

1. Einleitung

Nachrichtentechnische Verfahren zur Informationsübertragung mittels quadraturamplitudenmodulierter Signale (QAM-Signale) haben in den letzten Jahren zunehmend an Bedeutung gewonnen. Üblicherweise benötigt man bei der Demodulation von QAM-Signalen am Empfangsort den phasensynchronen, unmodulierten Träger des empfangenen Signals. Viele Demodulatoren, insbesondere solche zur Verarbeitung hoher Datenraten, benutzen zur Trägerableitung eine Phasenregelschleife. Deren Regelgröße wird in einem Phasendetektor gewonnen, der seine Information aus der komplexen Hüllkurve des modulierten Signals an den Abtastzeitpunkten gewinnt.

In herkömmlichen Detektoren wird durch unabhängige Entscheidungen über die kartesischen Komponenten

der komplexen Hüllkurve festgestellt, welcher mögliche Sendezustand (Sollzustand) der komplexen Hüllkurve zum Abtastzeitpunkt (Empfangszustand) am nächsten kommt. Daraus gewinnt man im Mittel eine Funktion der Phasenabweichung zwischen dem Träger des empfangenen Signals und dem Lokaloszillatorsignal.

Die Eigenschaften solcher Phasendetektoren sind gut bei Anwendung auf niederstufige QAM-Signale. Bei höherstufigen QAM-Signalen werden jedoch insbesondere die Akquisitionseigenschaften schlechter: Außer verlängerten Synchronisationszeiten tritt auch eine zunehmende Wahrscheinlichkeit der Synchronisation an ungeeigneten Phasenwinkeln auf (false locking).

Im folgenden wird daher nach einer allgemeinen Beschreibung solcher Schaltungen und des Standes der Technik ein neu entwickeltes Konzept zur Trägerablei-

* ANT Nachrichtentechnik GmbH, Backnang

tung vorgestellt werden, das mit bedingten Entscheidungen über die Polarkomponenten der komplexen Hüllkurve arbeitet. Es führt zu Phasendetektorcharakteristiken, die gutes Akquisitionsverhalten mit guten Eigenschaften nach erfolgter Synchronisation kombinieren.

2. Demodulation von QAM-Signalen

2.1 Signaldarstellung

Bekanntlich lassen sich quadraturamplitudenmodulierte Signale wie folgt darstellen:

$$u_{\text{QAM}}(t) = K \{a(t) \cos(\omega t + \varphi) + b(t) \sin(\omega t + \varphi)\}. \quad (1)$$

Hierbei sind $\cos(\omega t + \varphi)$ und $\sin(\omega t + \varphi)$ die beiden in Quadratur stehenden Komponenten des Signalträgers; $a(t)$ und $b(t)$ sind die zu übertragende Information. K ist ein konstanter Amplitudenfaktor.

$a(t)$ und $b(t)$ übertragen Daten eines Abtastsystems. Daher dürfen die Funktionen $a(t)$ und $b(t)$ an den äquidistanten Abtastzeitpunkten $t_k = kT$ nur Werte a_k und b_k aus einem Wertevorrat mit endlich vielen Werten annehmen [1, 2]. In einem 16QAM-System beispielsweise sind die Werte a_k und b_k Elemente der Menge $\{-3, -1, +1, +3\}$.

2.2 Quadraturempfänger

Man demoduliert QAM-Signale üblicherweise mit Hilfe eines Quadraturempfängers gemäß Bild 1 [2]. Ein spannungsgesteuerter Oszillator (VCO) hat die Aufgabe, ein zum Träger des Eingangssignals $u_{\text{QAM}}(t)$ in phasensynchroner Relation stehendes Signal zu erzeugen. Ein Phasenschieber PS bildet aus dem VCO-Signal zwei um 90° phasenverschobene Signale:

$$u_{\text{O}}(t) = \hat{u}_{\text{O}} \cos(\omega t + \varphi + \psi), \quad (2)$$

$$u_{\text{Q}}(t) = \hat{u}_{\text{O}} \sin(\omega t + \varphi + \psi). \quad (3)$$

Die Phase ψ ist von dem aus einem Regelfilter RF stammenden VCO-Steuersignal u_{R} abhängig.

Die Abwärtsmischer M1 bzw. M2 erzeugen aus dem Eingangssignal $u_{\text{QAM}}(t)$ und den Signalen $u_{\text{O}}(t)$ bzw. $u_{\text{Q}}(t)$ die Basisbandsignale

$$u_x(t) = K \{a(t) \cos \psi - b(t) \sin \psi\}. \quad (4)$$

$$u_y(t) = K \{a(t) \sin \psi + b(t) \cos \psi\}. \quad (5)$$

Die in $u_x(t)$ und $u_y(t)$ enthaltene Information über die Taktfrequenz $1/T$ und ihre Phasenlage kann in einem Taktsignal phasenrichtig abzuleiten (siehe hierzu beispielsweise [3]).

Damit können $u_x(t)$ und $u_y(t)$ zu den Zeitpunkten $t_k = kT$ abgetastet werden, und man erhält

$$x = u_x(kT) = K \{a_k \cos \psi - b_k \sin \psi\}, \quad (6)$$

$$y = u_y(kT) = K \{a_k \sin \psi + b_k \cos \psi\}. \quad (7)$$

Diese Gleichungen machen augenfällig, daß bei $\psi = 0$ oder $\psi = n2\pi$ (n ganzzahlig), also bei phasenrichtiger Anwendung des lokalen Trägersignals, in den Größen x und y die demodulierte Information zur Verfügung steht. Die somit wesentliche Aufgabe der Nachregelung der Phase ψ auf 0° oder dazu äquivalente Phasen wird in einem Phasenregelkreis wahrgenommen. Dieser besteht aus einem Phasendetektor PD, dem Regelfilter RF, den Mischern M1, M2, dem Oszillator VCO und dem Phasenschieber PS. Im Phasendetektor wird die Abweichung zwischen Soll- und Istphase festgestellt. Das Regelfilter dient zur Kontrolle der dynamischen Eigenschaften des Regelkreises.

In differenzcodierten Übertragungssystemen genügt es, die Phase ψ auf Vielfache von $\pi/2$ nachzuregeln (siehe (6) und (7)).

2.3 Zustandsdiagramme

Die komplexe Hüllkurve [4] von $u_{\text{QAM}}(t)$ ist

$$v_{\text{QAM}}(t) = K \{a(t) + j b(t)\} \exp(j(\omega t + \varphi)). \quad (8)$$

Normiert man K auf 1, und stellt man die komplexe Hüllkurve in einem mit dem Winkel $\omega t + \varphi$ rotierenden Koordinatensystem dar, dann ist $a(t) + j b(t)$ ein Signalzeiger, der $v_{\text{QAM}}(t)$ repräsentiert. Bild 2 zeigt die möglichen Positionen der Spitzen eines 16QAM-Signalzeigers zum Abtastpunkt (Zustandsdiagramm).

Aus (6) und (7) ergibt sich, daß die abgetasteten Mischsignale x und y des Quadraturempfängers gemäß Bild 1 für $\psi = 0$ die Komponenten der komplexen Hüllkurve des Empfangssignals zum Abtastzeitpunkt in einem Koordinatensystem liefern, welches phasensynchron mit der Trägerphase mitrotiert (Bezugskoordinatensystem, Sendereferenz). Die Punkte (x, y) fallen im Bezugssystem mit den gesendeten Datenpaaren (a_k, b_k) (Sollzustände) zusammen.

Durch Umformung von (8) weist man nach, daß in einem mit dem Winkel $\omega t + \varphi - \psi(t)$ rotierenden Koordinatensystem die komplexe Hüllkurve durch den Signalzeiger

$$[a(t) \cos \psi(t) - b(t) \sin \psi(t)] + j[a(t) \sin \psi(t) + b(t) \cos \psi(t)]$$

repräsentiert wird. Aus (4) und (5) folgt daher, daß die Mischsignale u_x und u_y die Komponenten der komplexen Hüllkurve des Empfangssystems in einem Koordinaten-

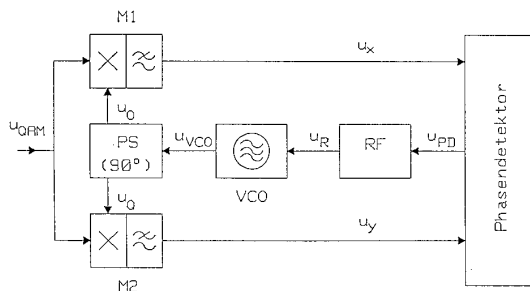


Bild 1: Quadraturempfänger

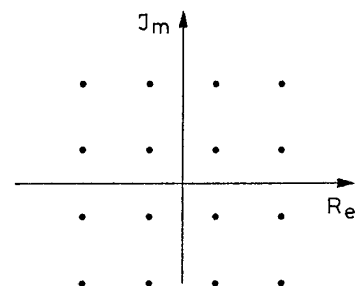


Bild 2: Komplexe Zustandsebene, 16QAM

system ergeben, das um ψ gegenüber dem Bezugssystem phasenverschoben ist (lokales Referenzsystem).

Die Darstellung des durch (x, y) gegebenen Signalzeigers im Bezugssystem (Zustandsdiagramm) liefert somit eine Aussage über die mögliche momentane Phasenverschiebung zwischen dem Lokaloszillatorsignal und dem unmodulierten Empfangssignalträger.

2.4 Phasendetektoren

Zur Gewinnung eines Regelsignals, das die phasenrichtige Synchronisation des Lokaloszillators auf den Empfangsträger ermöglicht, wird ein Phasendetektor benötigt. Die üblichen Phasendetektoren für QAM-Signale bauen auf Entwicklungen für QPSK-Demodulatoren auf. Im folgenden wird ein typischer QPSK-Detektor kurz umrissen. Dazu betrachte man das Zustandsdiagramm nach Bild 3.

Die Sollzustände sind als dicke Punkte eingetragen. Sie liegen an den Koordinaten (a_k, b_k) mit $a_k, b_k \in \{-1, +1\}$. Zusätzlich sind sogenannte Entscheidungsschwellen eingetragen. Das sind die strichlierten Geraden, die vertikal oder horizontal verlaufen und durch die Gleichungen $y = \hat{y}$ und $x = \hat{x}$ festgelegt sind, mit $\hat{x}, \hat{y} \in \{-1, 0, +1\}$.

Die Entscheidungsschwellen unterteilen die Zustands-ebene in quadratische Gebiete, die „Entscheidungsgebiete“. Jedes dieser Quadrate kann eindeutig durch ein Koordinatenpaar, $(\hat{x}, \hat{y})$ charakterisiert werden. In Bild 3 sind die einzelnen Gebiete durch kleine Kreissymbole markiert, deren Mittelpunkte mit den Koordinaten $(\hat{x}, \hat{y})$ zusammenfallen.

Im ungestörten Fall liegen die Komponenten x und y des empfangenen und in Quadratur abwärts gemischten Signals auf einem Kreis mit Mittelpunkt im Ursprung der Zustandsebene und mit Radius $1/\sqrt{2}$. Ein Vergleich der Signalkomponenten x bzw. y mit den mittleren Entscheidungsschwellen $\hat{x} = 0$ bzw. $\hat{y} = 0$ erlaubt eine Schätzung, welchen erlaubten Zustandswerten a_k bzw. b_k die empfangenen Signalkomponenten x und y am nächsten kommen („hard decision“ in kartesischen Koordinaten). Diese geschätzten Zustandswerte werden im folgenden mit $\hat{x}$ und $\hat{y}$ bezeichnet. Sie bestimmen den erlaubten Zustand $(\hat{x}, \hat{y})$ mit der geringsten euklidischen Distanz zu (x, y) .

Mit Hilfe weiterer Komparatoren wird festgestellt, in welchem Entscheidungsgebiet mit Charakterisierung $(\hat{x}, \hat{y})$ sich (x, y) befindet: (x, y) wird zu $(\hat{x}, \hat{y})$ „entschieden“ („soft decision“ in kartesischen Koordinaten).

Bestimmt nun beispielsweise $(\hat{x}, \hat{y})$ überwiegend Entscheidungsgebiete, die in Bild 3 durch $\oplus$ gekennzeichnet sind, dann ist die lokale Referenzebene um einen Drehwinkel $-45^\circ \leq \psi < 0^\circ$ gegenüber der Sendereferenz gedreht. Der Drehwinkel muß folglich im mathematisch positiven Sinn korrigiert werden. Wird hingegen überwiegend auf durch $\ominus$ markierte Gebiete entschieden, dann muß die lokale Referenz im mathematisch negativen Sinn gedreht werden.

Bildet man also ein Signal, das immer den Wert -1 hat, wenn (x, y) in einem durch $\ominus$ gekennzeichneten Gebiet liegt, und das immer den Wert $+1$ hat, wenn (x, y) in einem durch $\oplus$ gekennzeichneten Gebiet liegt, dann entspricht das Vorzeichen dieses Signals im Mittel der Korrekturrichtung für den Drehwinkel, um den die

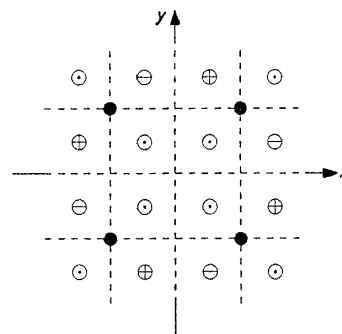


Bild 3: Entscheidungsgebiete für QPSK (4QAM)

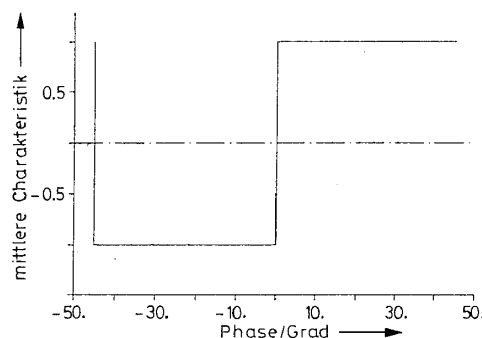


Bild 4: Mittlere Phasendetektorkennlinie v

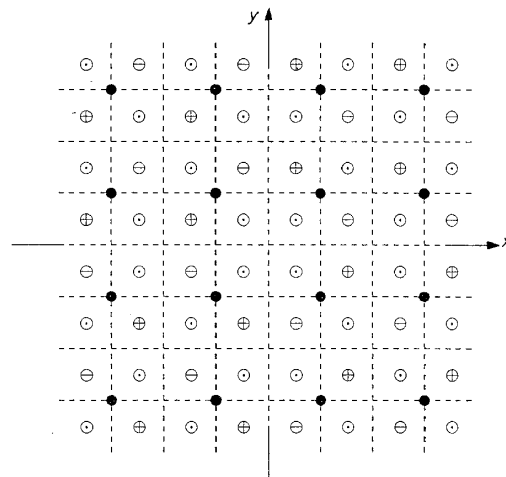


Bild 5: Entscheidungsgebiete für 16QAM

lokale Referenzebene gedreht werden muß, um mit der Sendereferenzebene zusammenzufallen. Ein Beispiel für ein solches Signal ist [2]:

$$v = \frac{1}{2} \cdot [\operatorname{sgn}(y) \operatorname{sgn}(\hat{x} - \hat{x}) - \operatorname{sgn}(x) \operatorname{sgn}(\hat{y} - \hat{y})]. \quad (9)$$

Bild 4 zeigt den Graphen von $\langle v \rangle$, das ist v gemittelt über alle (gleichverteilten) erlaubten Sendezustände, aufgetragen über ψ . ψ ist der Phasenwinkel, um den die Sendereferenz gegenüber der lokalen Referenz gedreht ist.

In Bild 4 ist eine volle Periode der Kennlinie von -45° bis $+45^\circ$ aufgetragen. Die Phasensynchronisation

der lokalen Referenzebene auf die Sendereferenz ist also nur bis auf Vielfache von 90° möglich. Dies hat seine Ursache in der 90° -Symmetrie der erlaubten Sendezustände.

Eine Erweiterung dieses Konzeptes auf höherstufige QAM-Verfahren wird in [2] beschrieben. Sie wird am Beispiel eines 16QAM-Systems erläutert. In Bild 5 ist wieder das Zustandsdiagramm aufgetragen. Die Sollzustände sind als dicke Punkte gekennzeichnet. Ihre Koordinaten liegen bei (a_k, b_k) mit $a_k, b_k \in \{-3, -1, +1, +3\}$. Außerdem sind horizontale und vertikale Entscheidungsschwellen strichliert eingetragen. Sie werden durch $x = \tilde{x}$ und durch $y = \tilde{y}$ mit $\tilde{x}, \tilde{y} \in \{-3, -2, -1, 0, +1, +2, +3\}$ bestimmt. Wie in Bild 3 wird in den Entscheidungsgebieten durch die Markierungen $\oplus$, $\ominus$ bzw. $\odot$ eine positive, negative oder unbestimmte Winkelkorrekturrichtung zugeordnet.

Vergleicht man Bild 5 mit Bild 3, so stellt man fest, daß die Bilder im Bereich $\{(x, y) | -2 < x < 2, -2 < y < 2\}$ übereinstimmen. Den Rest von Bild 5 kann man dadurch konstruieren, daß man innerhalb eines Quadranten die Umgebung eines weiter vom Ursprung entfernten Sollzustandes auf den innersten Sollzustand abbildet.

Ist (x, y) das empfangene und abwärtsgemischte Signal, dann wird durch Vergleich mit den Schwellen jedem empfangenen Zustand ein Entscheidungsgebiet $(\hat{x}, \hat{y})$ zugeordnet. Außerdem wird durch Vergleich mit den Schwellen abgeschätzt, welcher erlaubte Zustand (a_k, b_k) dem Paar (x, y) am nächsten kommt. Diese Schätzung wird wieder $(\hat{x}, \hat{y})$ genannt. Ein Beispiel für eine Funktion, welche die in Bild 5 bestimmten Eigenschaften hat, ist wieder durch eine Funktion v gemäß (9) gegeben. Bei Anwendung auf höher- als 4stufige QAM-Signale haben die Operatoren $\hat{\cdot}$ und $\tilde{\cdot}$ in dieser Gleichung allerdings insofern andere Bedeutung, als es unterschiedlich viele Entscheidungsschwellen und Sendezustände gibt.

3. Gütekriterien für Phasendetektoren

Gängige Schaltungen zur Trägerrückgewinnung unterscheiden sich im wesentlichen durch die Art des Phasendetektors. Will man also Vor- und Nachteile der einzelnen Schaltungskonzeptionen gegeneinander abwägen, dann sind Gütekriterien zur Beurteilung der Phasendetektoren erforderlich. Man gewinnt diese durch eine mathematische Beschreibung des durch Bild 1 gegebenen Phasenregelkreises.

Die Signale u_{QAM} und u_{VCO} werden im Phasendetektor PD miteinander verglichen. Sein Ausgangssignal u_{PD} ist

$$u_{PD} = A_g(\psi) + u_{PD0}. \quad (10)$$

Hierbei ist g eine periodische Funktion, deren Maximal- und Minimalwerte $+1$ und -1 sind. u_{PD0} kann also im Extremfall im $\pm A$ um den Mittelwert $u_{PD0} = 0$ variieren. Im folgenden wird angenommen, daß $u_{PD0} = 0$ gilt. Außerdem wird vorausgesetzt, daß $g(\psi)$ eine in ψ ungerade Funktion ist. $g(\psi)$ ist im Normalfall noch von weiteren, in $u_{QAM}(t)$ enthaltenen Parametern abhängig, die hier aus Gründen der Übersichtlichkeit nicht explizit aufgeführt werden.

Der (stationäre) Funktionszusammenhang $g(\psi)$ wird üblicherweise als Charakteristik des Phasendetektors, ihr Graph als Phasendetektorkennlinie bezeichnet.

Zur Verbesserung des dynamischen Regelverhaltens wird u_{PD} in einem Regelfilter RF in seinem Frequenzgang zu u_R verändert. In den meisten Fällen wird ein Filter mit der Übertragungsfunktion

$$H(s) := \frac{U_R(s)}{U_{PD}(s)} = \frac{L\{u_R(t)\}}{L\{u_{PD}(t)\}} \quad (11)$$

$$= \frac{1 + s\tau_2}{b + s\tau_1} \quad (12)$$

eingesetzt, wobei b , τ_1 und τ_2 Filterkonstanten sind.

Das Ausgangssignal $u_R(t)$ des Filters wird benutzt, um die Momentanfrequenz des Oszillators VCO zu regeln:

$$\omega_{VCO}(t) = \omega_0 + k_0 u_R(t). \quad (13)$$

$\omega_0/2\pi$ ist die Oszillatorruhefrequenz und k_0 ist die Modulationsempfindlichkeit des Oszillators. Für das Ausgangssignal des Oszillators ergibt sich somit

$$u_{VCO} = \cos \Phi_0(t) \quad (14)$$

mit

$$\Phi_0(t) = \int_0^t \omega_{VCO}(\tau) d\tau + \phi_0 = \omega_0 t + k_0 \int_0^t u_R(\tau) d\tau + \phi_0. \quad (15)$$

ϕ_0 ist die Phase des Oszillatorsignals zum Zeitpunkt $t = 0$. Damit kann die in (10) benötigte Winkeldifferenz ψ als

$$\psi(t) = \omega t + \varphi - \Phi_0(t) \quad (16)$$

bestimmt werden, oder mit $\omega - \omega_0 = \Delta\omega$

$$\psi(t) = \Delta\omega t + \varphi - \phi_0 - k_0 \int_0^t u_R(\tau) d\tau. \quad (17)$$

Aus (10) bis (17) erhält man

$$\tau_1 \dot{\psi} + [b + \tau_2 k_0 A g'(\psi)] \psi + k_0 A g(\psi) = b \Delta\omega. \quad (18)$$

Gleichung (18) ist eine nichtlineare Differentialgleichung zweiter Ordnung, die sich nicht allgemein lösen läßt. Es lassen sich aber einige allgemeine Aussagen über stabile Lösungen gewinnen. Setzt man voraus, daß ψ_∞ eine mögliche Lösung ist, die sich nach hinreichend langer Zeit nicht mehr verändert, dann erhält man durch Einsetzen in (18) als notwendige Bedingung für ψ_∞

$$g(\psi_\infty) = \frac{b \Delta\omega}{k_0 A}. \quad (19)$$

Da g beschränkt ist, folgt

$$|\Delta\omega| < \left| \frac{k_0 A}{b} \right| =: \Delta\omega_H. \quad (20)$$

$\Delta\omega_H$, der sogenannte Haltebereich, ist die maximale Ablage der Kreisfrequenz des lokalen Oszillatorsignals von der Trägerkreisfrequenz des QAM-Signals, die von dem Regelkreis noch ausgeglegt werden kann.

Man kann durch eine einfache Rechnung zeigen, daß nicht jede Lösung von (19) auch eine asymptotisch stabile Lösung von (18) ist. Dazu wird angenommen, daß $\psi(t) = \psi_\infty + \varepsilon(t)$ sei. ε sei zum Zeitpunkt $t=0$ betragsmäßig sehr klein. Durch Einsetzen in (18) und Entwickeln nach ε ergibt sich dann

$$\tau_1 \dot{\varepsilon} + [b + \tau_2 k_0 A g'(\psi_\infty)] \varepsilon + k_0 A g'(\psi_\infty) \varepsilon = 0. \quad (21)$$

Diese Gleichung kann allgemein gelöst werden. Es stellt sich dabei heraus, daß ε nur dann in einer Um-

gebung von ψ_∞ asymptotisch gegen 0 geht, daß also nur dann eine stabile Lösung vorliegt, wenn

$$k_0 A g'(\psi_\infty) > 0 \quad (22)$$

gilt.

Bei positivem k_0 und A und bei $\Delta\omega = 0$ sind also die stabilen Synchronisationspunkte die Nullstellen der Phasendetektorcharakteristik mit positiver Steigung!

Die Existenz einer stabilen Lösung besagt noch nicht, daß eine Frequenzablage $|\Delta\omega| < \Delta\omega_H$ auch tatsächlich zu 0 ausgeregelt wird. Durch Verallgemeinerung eines von A. J. Viterbi [5] auf Sinus-Phasendetektoren angewendeten Verfahrens kann gezeigt werden, daß für ausregelbare Frequenzablagen $\Delta\omega$ folgende Bedingung erfüllt sein muß:

$$|\Delta\omega| < \Delta\omega_P, \quad (23)$$

$$\Delta\omega_P := \frac{\pi b}{\tau_1} + 2 k_0 A \sqrt{\frac{\tau_2}{\tau_1} \left\{ \frac{g^2}{b} + \frac{\tilde{g}}{\tau_2 k_0 A} \right\}}, \quad (24)$$

$$\overline{g^2} := \frac{1}{2\pi P} \int_{-\pi P}^{\pi P} g^2(\psi) d\psi, \quad (25)$$

$$\tilde{g} := \frac{1}{2\pi P^2} \int_{-\pi P}^{\pi P} \psi g(\psi) d\psi, \quad (26)$$

$$2\pi P := \text{Periode der Kennlinie } g(\psi). \quad (27)$$

Diese Bedingung gilt streng nur, wenn der Mittelwert von g über eine Periode verschwindet. Für den allgemeineren Fall läßt sich eine ähnliche Bedingung ableiten. $\Delta\omega_P$ heißt Akquisitions- oder Ziehbereich.

Die Bedingung nach den Gln. (23) bis (27) zeigt, daß bei fester Parametereinstellung der Ziehbereich um so größer wird, je größer die Fläche zwischen dem Graphen von $g(\psi)$ und der ψ -Achse ist und je größer der gewogene Mittelwert von $g(\psi)$ zwischen 0 und πP ist. Dabei ist zu beachten, daß $|g(\psi)|_{\max} = 1$ ist.

Zwar kann durch Wahl von $|b| \ll 1$ der Ziehbereich sehr groß werden, es wird dann aber auch unter Umständen die Zeit bis zum Erreichen der Synchronisation sehr groß. Eine Vergrößerung des Ziehbereichs durch Kennlinienauswahl gemäß (23) bis (27) ist daher zu bevorzugen.

Um eine möglichst kurze Synchronisationszeit zu erreichen, werden Regelkreise meist so dimensioniert, daß zu erwartende Frequenzablagen $\Delta\omega$ innerhalb des sogenannten Fangbereiches liegen:

$$|\Delta\omega| < \Delta\omega_L. \quad (28)$$

Dieser ist so definiert, daß die Winkeldifferenz ψ innerhalb einer Periode auf den asymptotischen Wert ψ_∞ nachgeregelt wird.

Für Synchronisationskreise mit Filtern entsprechend (12) kann der Fangbereich in Verallgemeinerung eines in [6] angegebenen Verfahrens als

$$\Delta\omega_L := k_0 A \frac{\tau_2}{\tau_1} \quad (29)$$

bestimmt werden.

Durch Rechnersimulationen kann gezeigt werden, daß bei ansonsten gleichem Verhalten in unmittelbarer Umgebung des Synchronisationspunktes die Synchronisations-

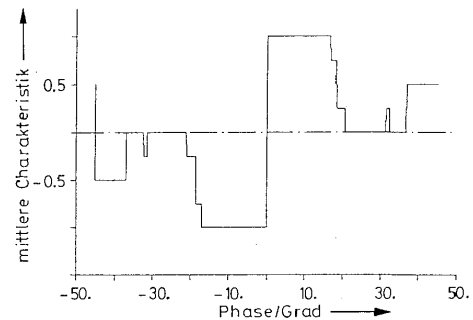


Bild 6: Mittlere Phasendetektorkennlinie v bei 16QAM

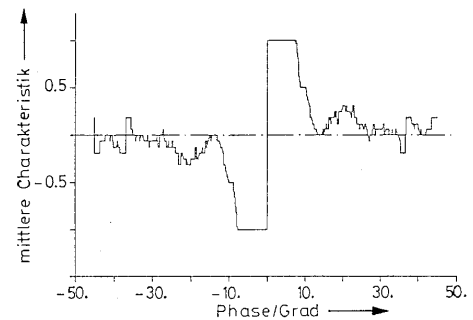


Bild 7: Mittlere Phasendetektorkennlinie v bei 64QAM

zeit innerhalb des Fangbereiches um so kürzer ist, je größer die Fläche zwischen $g(\psi)$ und der ψ -Achse ist.

Zur Erzielung kurzer Synchronisationszeiten, großer Akquisitionsbereiche und zur Vermeidung von Falschsynchronisationspunkten sind daher Phasendetektorkennlinien zu suchen, die

1. innerhalb eines Periodizitätsintervalls genau eine Nullstelle mit positiver Steigung aufweisen,
2. eine möglichst große Fläche zwischen der Phasendetektorkennlinie und der ψ -Achse haben.

Bild 4 zeigt die Phasendetektorcharakteristik, die bei Anwendung von v gemäß (9) auf ein QPSK-System und Mittelung über alle Sendezustände entsteht. Bild 6 zeigt die entsprechende Phasendetektorkennlinie bei Anwendung von v auf ein 16QAM-System, und Bild 7 zeigt die Kennlinie bei Anwendung auf 64QAM. Vergleich mit den oben gestellten Anforderungen an eine gute Phasendetektorcharakteristik zeigt, daß die durch (9) entstehende Charakteristik gute Resultate bei QPSK, brauchbare Resultate bei 16QAM und weniger befriedigende Resultate bei 64QAM zeigt.

4. Phasendetektoren mit bedingter Entscheidungsstrategie

Den Grund für das wenig befriedigende Verhalten des oben beschriebenen Phasendetektors macht man sich leicht klar, wenn man den Fall betrachtet, daß ein Zeichen mit maximaler Amplitude gesendet wird, beispielsweise das Zeichen (3, 3) im Falle eines 16QAM-Systems. Angenommen, die lokale Referenzebene sei um $\psi \approx 45^\circ$ gegenüber der Sendereferenz verschoben. Wenn

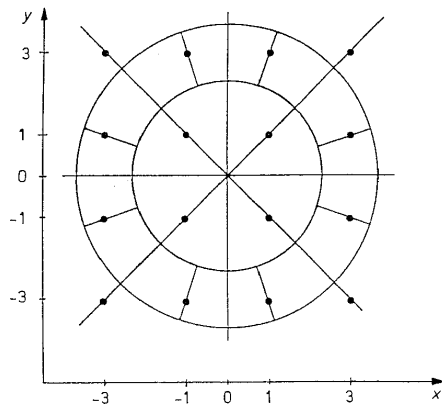


Bild 8: Polare Entscheidungsgebiete für 16QAM

nun ψ nach und nach auf 0° verkleinert wird, dann erzeugt die in (9) gegebene Gleichung alternierend positive und negative Werte, obwohl doch zu einem empfangenen Wert dieser Amplitude der nächstgelegene Sollwert eindeutig in Richtung kleinerer Winkelwerte liegt: Die Information über die Amplitude der Empfangswerte wird offenbar nicht ausgenutzt.

Der Phasendetektor hat die Aufgabe ein Signal zur Winkelkorrektur zu erzeugen. Es liegt daher nahe, diese Aufgabe in einem angepaßten, nämlich einem polaren Koordinatensystem zu lösen und dabei die Information über die Signalamplitude mitauszunutzen. Detektoren, die nach diesem Prinzip arbeiten, wurden erstmalig in [7] vorgeschlagen.

Das Verfahren wird im folgenden allgemein abgehandelt. Zur Veranschaulichung dient Bild 8, welches die Ebene der empfangenen Zustände für ein 16QAM-System zeigt.

Zunächst werden die kartesischen Komponenten (x, y) des empfangenen und in Quadratur abwärts gemischten Signals in Polarkoordinaten (R, α) umgewandelt. Die erste Komponente des Polarkoordinatenpaares wird im folgenden als „Amplitude“, die zweite als „Winkel“ oder „Phase“ bezeichnet. Weiter werden zu jedem Sollzustand mit den kartesischen Koordinaten (a_k, b_k) die äquivalenten Polarkoordinaten $(R(k), \alpha(k))$ bestimmt. Man ordnet die $R(k)$ der Größe nach und erhält so einen Wertevorrat

$$\tilde{\alpha}_{i,0} \leq \alpha_{i,1} < \tilde{\alpha}_{i,1} < \alpha_{i,2} < \tilde{\alpha}_{i,2} < \dots < \tilde{\alpha}_{i,q_i-1} < \alpha_{i,q_i} < \tilde{\alpha}_{i,0} + 2\pi = \tilde{\alpha}_{i,q_i}. \quad (36)$$

$\mathfrak{R}$ mit r verschiedenen Werten R_i , den „Nominalamplituden“, mit

$$0 \leq R_1 < \dots < R_i < \dots < R_r. \quad (30)$$

Im Beispiel gemäß Bild 8 ist $R_1 = \sqrt{2}$, $R_2 = \sqrt{10}$ und $R_3 = \sqrt{18}$.

Um entscheiden zu können, welcher Nominalamplitude ein empfangener Amplitudenwert R am nächsten kommt, bestimmt man „Amplitudenschwellen“ $\tilde{R}_i$ mit

$$0 = \tilde{R}_0 \leq R_1 < \tilde{R}_1 < R_2 < \tilde{R}_2 < \dots < \tilde{R}_{r-1} < R_r < \tilde{R}_r = \infty. \quad (31)$$

Eine mögliche Wahl für diese Schwellen ist

$$\tilde{R}_i = (R_i + R_{i+1})/2 \quad \text{für } i = 1, \dots, r-1. \quad (32)$$

Die Amplitudenschwellen sind in Bild 8 als Kreise eingetragen. Sie unterteilen die Ebene in Kreis- und Kreisringflächen, welche die Nominalamplituden R_i repräsentieren, und die im folgenden Amplitudengebiete genannt werden.

Durch Vergleich des Amplitudenwertes R des empfangenen Signals mit den Schwellen $\tilde{R}_i$ wird R zu $\tilde{R}$ unterschieden: $\tilde{R}$ ist derjenige Amplitudenwert R_k für den

$$\tilde{R}_{k-1} \leq R < \tilde{R}_k \quad (33)$$

gilt, wobei der aktuelle Wert von k durch

$$\tilde{R}_{k-1} \leq R < \tilde{R}_k \quad (34)$$

gegeben ist.

Die (erste) Entscheidung der Amplitudenkoordinate auf $\tilde{R}$ erfolgt ohne Vorbedingung.

Als nächster Schritt ist zu bestimmen, wie die relative Phasenlage des empfangenen Signals zum „nächstgelegenen“ Sollzustand ist, der die Amplitude $\tilde{R}$ hat. Diese zweite Entscheidung ist also abhängig von der ersten Entscheidung.

Hierin ist ein wesentlicher Unterschied zu den Verfahren mit nicht-bedingter Entscheidungsstrategie zu sehen!

Durch diese bedingte Vorgehensweise wird der nächstgelegene Sollzustand *nicht* als der Zustand bestimmt, dessen *euklidische Distanz* zu dem Empfangszustand am geringsten ist. Vielmehr wird als Distanz zwischen empfangenem Signal (R, α) und einem beliebigen Sollzustand mit $R_k = \tilde{R}$ eine Metrikfunktion des Betrags der Winkeldifferenz modulo 2π definiert! Diese Distanz wird im folgenden Winkeldistanz ϑ genannt. Im einfachsten Fall wird der Betrag selbst als Winkeldistanz definiert.

Zur praktischen Bestimmung der Winkeldistanz geht man wie folgt vor: Für alle Sollzustände mit Amplitudenwert R_i werden die Winkelkomponenten $\alpha_{i,j}$ bestimmt und der Größe nach geordnet. So erhält man zu jedem Amplitudenwert R_i im allgemeinen q_i verschiedene Winkelwerte mit

$$-\pi/2 \leq \alpha_{i,1} < \dots < \alpha_{i,j} < \dots < \alpha_{i,q_i}. \quad (35)$$

Im Beispiel gemäß Bild 8 gibt es zu R_1 $q_1 = 4$, zu R_2 $q_2 = 8$ und zu R_3 $q_3 = 4$ verschiedene Winkelwerte.

Es werden nun „Winkelschwellen“ $\alpha_{i,j}$ eingeführt:

Die Winkelschwellen werden so bestimmt, daß ihre Winkeldistanz zum nächst größeren und zum nächst kleineren Nominalwinkel gleich groß ist. Im einfachsten Fall ist

$$\tilde{\alpha}_{i,j} = \begin{cases} (\alpha_{i,j} + \alpha_{i,j+1})/2 & \text{für } j = 1, \dots, q_i-1, \\ (\alpha_{i,j} + \alpha_{i,1} + 2\pi)/2 & \text{für } j = q_i. \end{cases} \quad (37)$$

In Bild 8 sind innerhalb der durch die Nominalamplituden R_i bestimmten Amplitudengebiete die Winkel $\alpha_{i,j}$ und die Winkelschwellen $\tilde{\alpha}_{i,j}$ als Geradenstücke eingetragen, deren gedachte Verlängerung durch den Ursprung geht. Diese Geradenstücke erzeugen zusammen mit den Amplitudenschwellen kreissektor- und kreisringsektorförmige Entscheidungsgebiete.

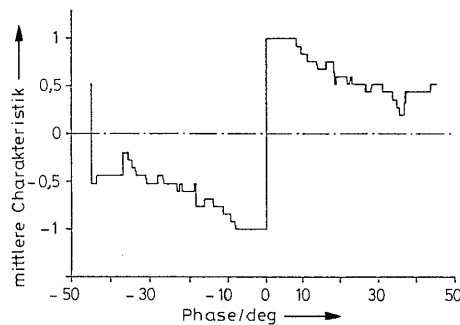
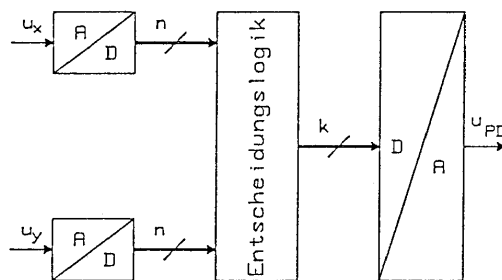
Bild 9: $g(\psi)$ bei 64QAM

Bild 10: Mögliche Hardwarerealisierung des neuen Verfahrens

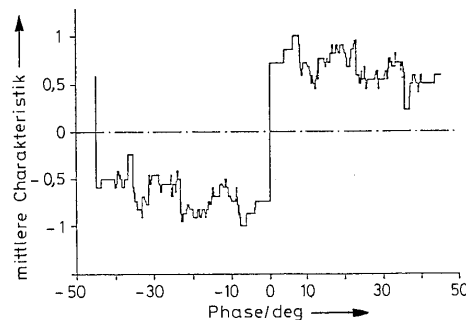


Bild 11: Phasendetektorkennlinie mit digitalisiertem Verfahren bei 64QAM

Nach Entscheidung der empfangenen Signalkomponente R auf $\hat{R} = R_k$ kann durch Vergleich der empfangenen Komponente α mit den Winkeln $\alpha_{k,j}$ bzw. den Winkelschwellen $\hat{\alpha}_{k,j}$ entschieden werden, welcher Sollwinkel am wahrscheinlichsten gesendet wurde und in welchem Entscheidungsgebiet sich der empfangene Zustand befindet.

Man gelangt nun in einfacher Weise zu einer Phasendetektorkennlinie, wenn man

$$v = \text{sgn}(\vartheta) \quad (38)$$

und

$$g(\psi) = \langle v \rangle \quad (39)$$

bildet. ϑ ist dabei die oben definierte Winkeldistanz. Das Ergebnis ist jedoch bei höherstufigen QAM-Verfahren noch nicht befriedigend. Dies kann am Beispiel des in

Bild 8 gezeigten Entscheidungsdiagramms erklärt werden. In dem Kreisring, in dem die Sollzustände mit $R_i = 1/\sqrt{10}$ liegen, gibt es innerhalb eines Quadranten, also innerhalb eines Periodizitätsbereiches, mehr als einen erlaubten Zustand. Der Winkeldistanzbereich, für den eine korrekte Entscheidung getroffen werden kann, ist somit kleiner als eine Periodizitätslänge. Dagegen ist in den beiden anderen Amplitudenbereichen (wenigstens im ungestörten Fall) immer eine korrekte Entscheidung möglich.

Daher ist es sinnvoll, zur Konstruktion einer Phasendetektorkennlinie folgende Definition zu benutzen:

$$v = w(\hat{R}) \text{sgn}(\vartheta) \quad (40)$$

$$g(\psi) = \langle v \rangle. \quad (41)$$

Dabei ist $w(\hat{R})$ eine nichtnegative "Gewichtsfunktion", die einen großen Zahlenwert hat, wenn wenige erlaubte Zustände im Amplitudengebiet $\hat{R}$ vorkommen, und umgekehrt einen kleinen Zahlenwert hat, wenn mehr erlaubte Zustände im Amplitudengebiet $\hat{R}$ existieren. Bild 9 zeigt die mittlere Kennlinie für einen Phasendetektor, der nach diesem Prinzip für ein 64QAM-System berechnet wurde. Dabei wurde

$$w(\hat{R}) = \begin{cases} 4 & \text{für } \hat{R} = \begin{cases} R_1 = \sqrt{2} \\ R_3 = \sqrt{18} \\ R_9 = \sqrt{98} \end{cases} \\ 1 & \text{sonst} \end{cases} \quad (42)$$

eingesetzt. Das Resultat erfüllt weitgehend die Forderungen an eine gute Phasendetektorcharakteristik.

Die analoge Realisierung des angegebenen Verfahrens würde den Einsatz von Schwellwertentscheidern mit nicht äquidistanten Schwellen und von Koordinatenumsetzern erforderlich machen. Günstiger ist daher eine digitalisierte Lösung, zumal ja auch die gesendete Information digitaler Natur ist. Daher wird das angegebene Verfahren wie folgt abgewandelt.

Statt die (analogen) kartesischen Signalkomponenten (x, y) in die (analoge) polare Darstellung (R, α) umzurechnen und dann gemäß obigem Verfahren weiterzuverarbeiten, wird (x, y) digitalisiert zu $(\hat{x}, \hat{y})$ mit einer Auflösung von je n bit pro Komponente. $(\hat{x}, \hat{y})$ kann als Näherung für (x, y) aufgefaßt werden. Je größer n ist, desto genauer ist die Näherung. Selbstverständlich muß n wenigstens so groß sein, daß alle verschiedenen erlaubten Sendewerte a_k und b_k dargestellt werden können.

Nun wird die (digitale) kartesische Darstellung $(\hat{x}, \hat{y})$ umgewandelt in die polare Darstellung $(\hat{R}, \hat{\alpha})$. Die weitere theoretische Vorgehensweise ist völlig analog zum oben beschriebenen Verfahren, es werden nur $\hat{R}$ und $\hat{\alpha}$ statt R und α verwandt.

Der Vorteil dieser Vorgehensweise ist nun der: Weil es nur endlich viele verschiedene Werte $\hat{R}$ und $\hat{\alpha}$ gibt können auch nur endlich viele verschiedene Winkelwerte ϑ auftreten. Folglich können sämtliche Funktionswerte der Phasendetektorkennlinie $v(\vartheta)$ vor der Anwendung des Verfahrens auf ein physikalisches Eingangssignal bestimmt und in einer Tabelle abgelegt werden. Damit wird das Verfahren einer besonders einfachen Hardwarerealisierung zugänglich.

Bild 10 zeigt eine Realisierungsmöglichkeit im Blockschaltbild. Die Signalkomponenten u_x und u_y werden

in A/D-Umsetzern in eine n -bit-Darstellung umgesetzt. / dieser n bit entsprechen dann im Fall erfolgreicher Synchronisation der demodulierten Information. In einem logischen Schaltkreis, der beispielsweise festverdrahtet oder aber als Speicher angelegt sein kann, wird die Umsetzung der Funktionstabelle in Hardware realisiert. Der Funktionswert wird in einer k -bit-Darstellung ausgegeben und in einem D/A-Umsetzer in das gewünschte Detektorausgangssignal umgewandelt.

Daß selbst durch eine relativ grobe digitale Darstellung der Signalkomponenten u_x und u_y noch gute Ergebnisse erzielt werden können, zeigt **Bild 11**. Hier wird für ein 64QAM-System mit $n=5$ eine Kennlinie erzielt, die die wesentlichen Anforderungen an gute Detektoren erfüllt.

5. Zusammenfassung

In den vorangehenden Abschnitten wurden Schaltungen zur Trägerableitung bei höherstufigen QAM-Signalen auf allgemeine Phasensynchronisationsschaltungen zurückgeführt. Durch Verallgemeinerung bekannter Methoden wurden Anforderungen an den Phasendetektor dieser Schaltung formuliert. Bekannte Schaltungen wurden, gemessen an diesen Anforderungen und bei Anwendung

auf höherstufige QAM-Signale, als weniger zufriedenstellend erkannt. Durch konsequente Anwendung einer Beschreibung in einem angepaßten Koordinatensystem und durch Einführung bedingter Entscheidungen konnten neue Phasendetektorfunktionen gefunden werden, welche sich günstig in Hardware umsetzen lassen und die auch bei hohen Datenraten gutes Akquisitionsverhalten mit guten Eigenschaften nach erfolgter Synchronisation kombinieren.

Literatur:

- [1] Foschini, G. J.; Gitlin, R. D.; Weinstein, S. B.: On the Selection of a Two-Dimensional Signal Constellation in the Presence of Phase Jitter and Gaussian Noise. *Bell Syst. Techn. J.* 52 (July/August 1973) 6, S. 927—965.
- [2] Leclert, A.; Vandamme, P.: Universal Recovery Loop for QASK and PSK Signal Sets. *IEEE Trans. COM-31* (1983) 1, S. 130—136.
- [3] Le-Ngoc, T.; Feher, K.: A Digital Approach to Symbol Timing Recovery Systems. *IEEE Trans. COM-28* (1980) 12, S. 1993—1999.
- [4] Rice, S. O.: Envelopes of Narrowband Signals. *Proc. IEEE* 70 (1982) 7, S. 692—699.
- [5] Viterbi, A. J.: *Principles of Coherent Communications*. New York: McGraw-Hill Book Company, 1966, chapter 3.
- [6] Best, R.: *Theorie und Anwendungen des Phase-locked Loops*. Aarau, Stuttgart: AT-Verlag, 1981, Kapitel 3.
- [7] Hoffmann, M. H. W.: Carrier Recovery for M-QAM-Signals. *European Conference on Radio-Relay Systems (ECRR)* (November 1986), S. 247—253.

Dr. M. H. W. Hoffmann, ANT Nachrichtentechnik GmbH, Gerberstraße 34, D-7150 Backnang

(Eingegangen am 20. 2. 1988, überarbeitete Fassung am 22. 6. 1988)