

Puls-Code Modulation (PCM)

Zur Erfassung und rechnerbasierten Verarbeitung physikalischer Messgrößen werden spezielle Sensoren eingesetzt, mit denen der zeitliche Verlauf der jeweiligen Messgröße in ein analoges elektrisches Signal gewandelt wird. Beispielsweise wird bei Sprache, die als Schallwelle übertragen wird, ein Mikrofon als Sensor eingesetzt, mit dem die Veränderung des Schalldrucks in ein elektrisches Signal gewandelt wird. Der elektrische Signalverlauf stellt ein zeit- und wertekontinuierliches Signal dar. Um das Signal mit einem Digitalrechner verarbeiten zu können, muss das zeit- und wertekontinuierliche Signal in ein zeit- und wertediskretes Signal umgesetzt werden. Dazu setzt man einen Analog-/Digital-Umsetzer (ADU) ein. In den folgenden Abschnitten werden die zur Umsetzung notwendigen Schritte der Abtastung und Quantisierung vorgestellt. Die Vorgehensweise zur Wandlung eines analogen Signals in eine Folge quantisierter Abtastwerte bezeichnet man als Puls-Code-Modulation (PCM).

Abtastung

Zur Erfassung eines analogen, elektrischen Signalverlaufs wird das Signal zu äquidistanten Zeitpunkten $n \cdot T$ abgetastet, wie es Bild 1 beispielhaft zur Aufnahme eines Audiosignals mit einem Mikrofon als Sensor veranschaulicht.

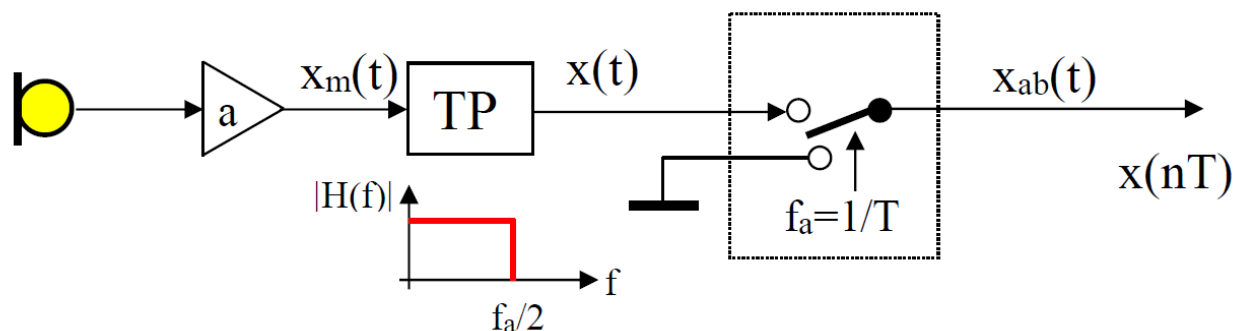


Bild 1: Abtastung des tiefpassgefilterten Mikrofonsignals

Den Kehrwert der Zeit T bezeichnet man als die Abtastfrequenz $f_a = 1/T$, die die Anzahl der Abtastwerte je Sekunde festlegt. Die konkrete Wahl des Werts der Abtastfrequenz wird durch das sogenannte Abtasttheorem, das auch als 1. Nyquistkriterium bezeichnet wird, bestimmt. Das Abtasttheorem besagt, dass ein Signal mit einer Frequenz abgetastet werden muss, die größer als das Doppelte der höchsten in dem Signal enthaltenen Frequenz ist. $f_a > 2 \cdot f_{\max}$

Da man die in einem Signal auftretende maximale Frequenz häufig nicht kennt, erfolgt vor der Abtastung eine Filterung des analogen Signals mit einem Tiefpass. Dieser verhindert weitgehend das Auftreten von Frequenzanteilen oberhalb von $f_a/2$, um das zur Festlegung von f_{\max} umgestellte Abtasttheorem $f_{\max} < f_a/2$ zu erfüllen. Da ein realer analoger Tiefpass nicht die ideale rechteckförmige Charakteristik besitzt, wie sie in Bild 1 veranschaulicht wird, wird die Grenzfrequenz f_g des Tiefpasses in der Regel zu $f_g < f_a/2$ gewählt. Damit wird insbesondere der Tatsache Rechnung getragen, dass die Flanke im Bereich der Grenzfrequenz des Tiefpasses nur eine endliche Steilheit besitzt.

Beispielsweise besitzt ein Sprachsignal Frequenzanteile im Bereich bis etwa 7 kHz. Dies bedingt eine Abtastung mit einer Frequenz, die größer als 14 kHz sein muss. Bei der Entwicklung des analogen Telefons hat man allerdings festgestellt, dass auch die Beschränkung auf den Bereich von 300 Hz bis 3,4 kHz noch ein gut verständliches Sprachsignal liefert. Daher verwendet man zur digitalen Erfassung der Sprache im Bereich der Telephonie in der Regel eine Abtastfrequenz von **8 kHz**, so dass man auch bei Einsatz eines nicht idealen Tiefpasses noch Frequenzanteile bis 3,4 kHz erfassen kann. Zur Realisierung einer qualitativ besseren Übertragung mit höherer Sprachqualität verwendet man eine Abtastfrequenz von **16 kHz**. Damit lassen sich Frequenzanteile bis etwa 7 kHz auch mit einer nicht idealen Filtercharakteristik erfassen.

Die Abtastwerte des zeitdiskreten Signals $x(n \cdot T)$ mit $n = -\infty, \dots, -2, -1, 0, 1, 2, \dots, \infty$ lassen sich aus einer mathematischen Beschreibung des analogen Signals $x(t)$ durch eine Substitution von t durch $n \cdot T$ berechnen. Damit ergibt sich beispielsweise für ein Cosinussignal die nachstehende mathematische Darstellung des zeitdiskreten Signals.

$$x(t) = \cos(2 \cdot \pi \cdot f \cdot t) \quad t \rightarrow n \cdot T \quad x(n \cdot T) = \cos(2 \cdot \pi \cdot f \cdot n \cdot T) = \cos\left(2 \cdot \pi \cdot n \cdot \frac{f}{f_a}\right)$$

Es ergibt sich eine Folge von Abtastwerten, die das Cosinussignal repräsentiert. Die zuvor angegebene mathematische Beschreibung veranschaulicht die Abhängigkeit von dem Verhältnis der Frequenz des Cosinus zur Abtastfrequenz. Damit erhält man beispielsweise bei gleichzeitiger Verdopplung der Frequenz f und der Abtastfrequenz die gleiche Folge von Abtastwerten. Ohne Kenntnis der Abtastfrequenz kann man einer Folge von Abtastwerten keinen absoluten Frequenzwert zuordnen. Dies macht deutlich, dass die Verarbeitung eines Signals in einem

Digitalrechner immer relativ zur Abtastfrequenz erfolgt. Man beschreibt das zeitdiskrete Signal in der Regel auch nur in Abhängigkeit des ganzzahligen Abtastindex n als $x(n)$. Dabei definiert n , welcher Wert aus der Folge von Abtastwerten bearbeitet wird.

Um die Eigenschaften des abgetasteten Signals im Frequenzbereich zu bestimmen, beschreibt man die Abtastung zu äquidistanten Zeitpunkten mathematisch als die Multiplikation des Signals mit einer Folge von Dirac Impulsen, wie es Bild 2 veranschaulicht.

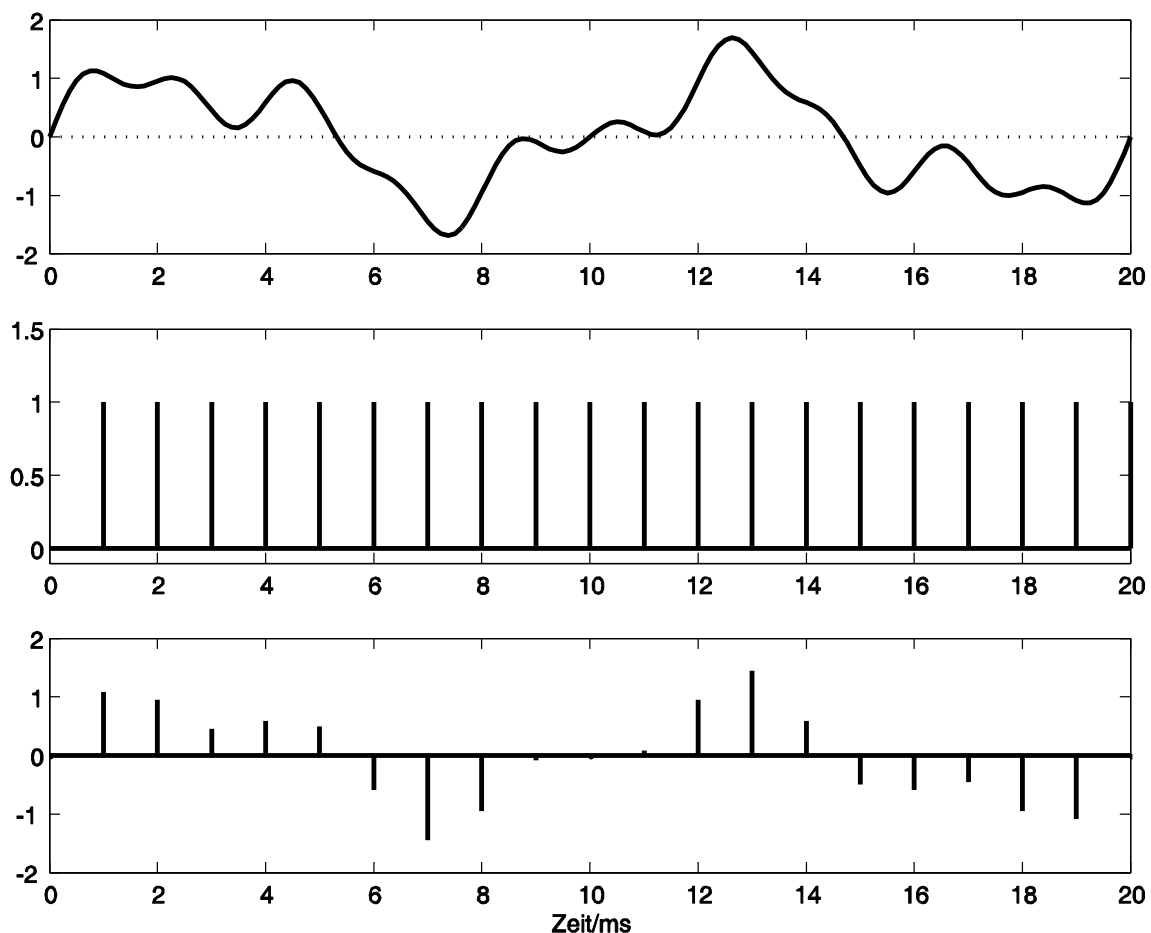


Bild 2: Analoges Signal (oben), Folge von Dirac-Impulsen (Mitte), PAM Signal (unten)

Das Signal, das aus der Folge von Dirac Impulsen besteht, nimmt zu den Abtastzeitpunkten den Wert 1 an und ist ansonsten Null. Das aus der Multiplikation resultierende Signal $x_{ab}(t)$ nimmt zu den Abtastzeitpunkten die Werte $x(n \cdot T)$ an und ist ansonsten Null. Man bezeichnet diese Vorgehensweise auch als *Pulsamplitudenmodulation* (PAM).

Die Folge von Dirac Impulsen lässt sich formal beschreiben als

$$\sum_{n=-\infty}^{\infty} \delta(t - n \cdot T)$$

Unterwirft man diese Impulsfolge einer Fourier Transformation, so ergibt sich im Spektralbereich ebenfalls eine Folge von Dirac Impulsen:

$$\sum_{n=-\infty}^{\infty} \delta(t - n \cdot T) \quad \circ \text{---} \bullet \quad \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta(f - n/T) = \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta(f - n \cdot f_a)$$

Die Dirac Impulse treten bei Vielfachen der Abtastfrequenz f_a auf.

Aus der Multiplikation des Signals mit einer Impulsfolge im Zeitbereich wird eine Faltung des Spektrums mit der entsprechenden Impulsfolge im Frequenzbereich.

$$\begin{array}{c}
 x_{ab}(t) = x(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - n \cdot T) \\
 \downarrow \\
 X_{ab}(f) = X(f) * \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta(f - n \cdot f_a)
 \end{array}$$

Dies wird in Bild 3 in einer zweiseitigen spektralen Darstellung einschließlich negativer Frequenzen veranschaulicht, in der das Betragsspektrum eines gemäß dem Abtasttheorem tiefpassgefilterten analogen Signals, die Folge von Dirac Impulsen im Frequenzbereich sowie das Betragsspektrum des abgetasteten Signals dargestellt sind.

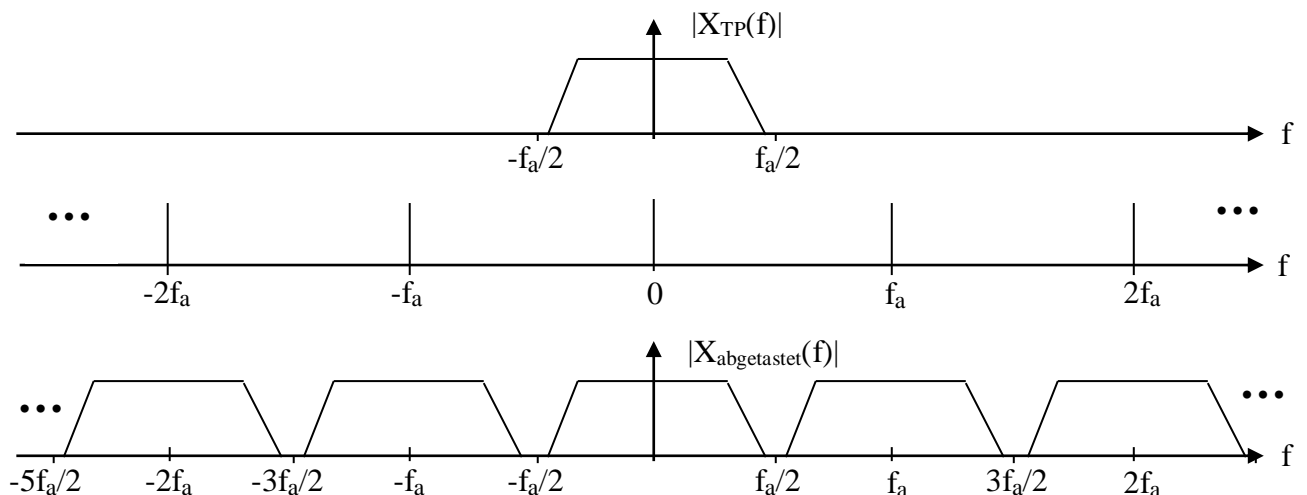


Bild 3: Wiederholtes Auftreten des TP-Spektrums nach einer Faltung des Spektrums mit einer Folge von Dirac-Impulsen

Die Abtastung im Zeitbereich führt zu einer periodischen Wiederholung des tiefpassgefilterten Spektrums bei Vielfachen der Abtastfrequenz. Das Spektrum $X_{ab}(f)$ des PAM Signals $x_{ab}(t)$ ist

unendlich ausgedehnt. Zur Rekonstruktion des analogen TP Signals aus dem PAM Signal muss das PAM Signal mit einem Tiefpass mit einer Grenzfrequenz $f_g \sim f_a/2$ gefiltert werden. Die komplette Verarbeitungskette zur Gewinnung und Übertragung eines pulsamplitudenmodulierten Signals sowie einer Rekonstruktion des analogen Signals aus dem PAM Signal ist in Bild 4 dargestellt.

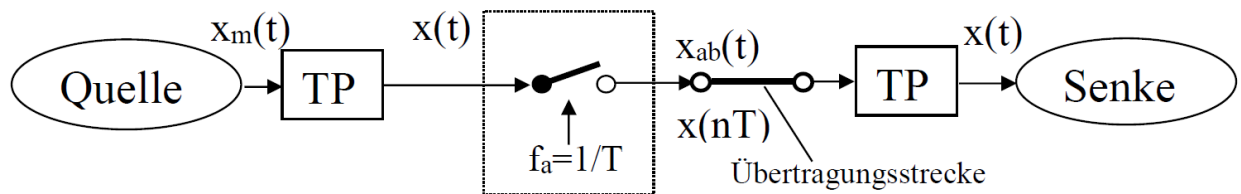


Bild 4: PAM Signalgenerierung und Rekonstruktion des TP gefilterten analogen Signals

Die Kenntnis von dem wiederholten Auftreten des Spektrums eines abgetasteten Signals kann auch herangezogen werden, um die bei einer Verletzung des Abtasttheorems auftretenden Effekte darzustellen. Es wird der Fall betrachtet, dass vor der Abtastung eines Signals keine entsprechende TP Filterung erfolgt, so dass in dem abzutastenden Signal Frequenzanteile oberhalb der halben Abtastfrequenz enthalten sind. Beispielhaft ist dazu in Bild 5 das Spektrum eines Cosinussignals, das eine Frequenz von 5 kHz besitzt, im oberen Bild dargestellt.

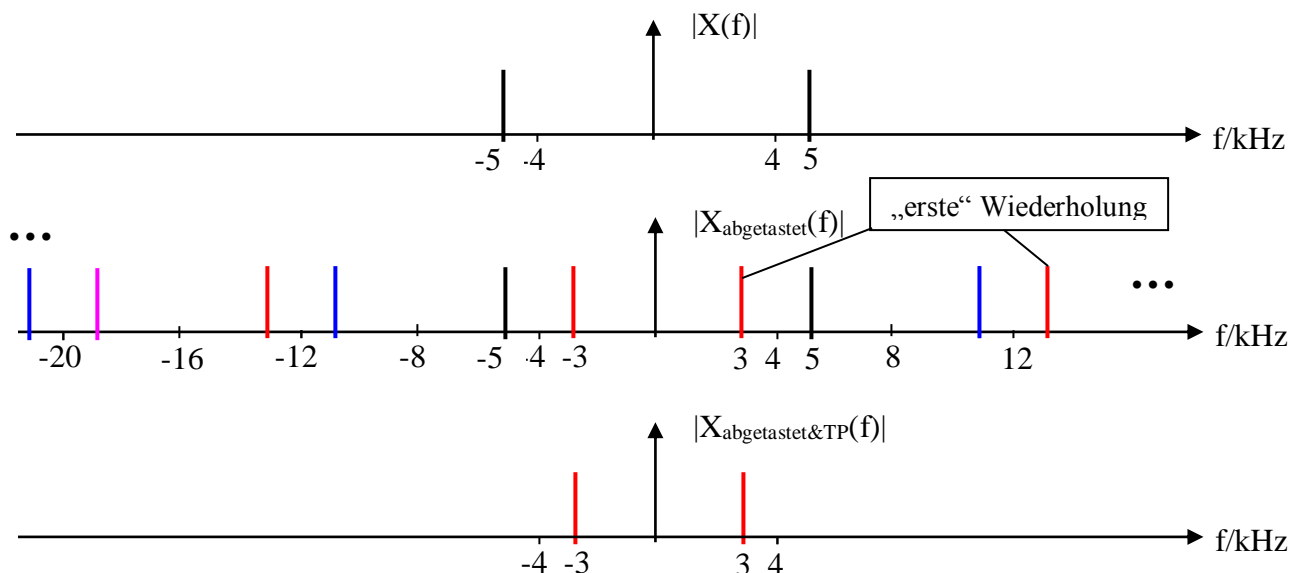


Bild 5: Spektrum eines 5 kHz Signals (oben), Spektrum des unterabgetasteten Signals (Mitte), Spektrum des TP gefilterten Signals (unten)

Das Spektrum besteht aus zwei Dirac Impulsen bei den Frequenzen -5 und +5 kHz. Nach einer Abtastung des Cosinussignals mit einer Frequenz von 8 kHz, treten weitere Impulse bei $8-5 = 3$ kHz und bei $8+5 = 13$ kHz als „erste“ Wiederholung des Spektrums auf. Als „zweite“ Wiederholung treten weitere Impulse bei $16-5 = 11$ kHz und bei $16+5 = 21$ kHz auf. Entsprechend fortgesetzt treten weitere Impulse für die weiteren Vielfachen der Abtastfrequenz sowie im negativen Frequenzbereich auf. Wird das aus der Abtastung resultierende PAM Signal mit einem korrekt gewählten Tiefpass mit $f_g \sim \frac{f_a}{2} = 4 \text{ kHz}$ gefiltert, so erhält man ein Cosinussignal, das eine Frequenz von 3 kHz besitzt. Den beobachteten Effekt kann man verallgemeinernd so beschreiben, dass Frequenzanteile, die im abzutastenden Signal oberhalb von $f_a/2$ bei $f_a/2 + \Delta f$ vorhanden sind, nach der Abtastung und Filterung bei $f_a/2 - \Delta f$ auftreten. Man spricht dabei auch von einer „Rückfaltung“ der oberhalb von $f_a/2$ liegenden Frequenzanteile. Im Allgemeinen kommt es zu einer Überlagerung der eigentlichen Frequenzanteile im Frequenzbereich unterhalb von $f_a/2$ mit den „rückgefalteten“ Komponenten. Man nennt diesen Effekt der Überlagerung von Spektralanteilen „Aliasing“. Das vor der Abtastung eingesetzte TP Filter wird daher häufig auch als *Antialiasingfilter* bezeichnet.

Quantisierung und Codierung

In einem Digitalrechner werden Werte als Dualzahlen mit einer bestimmten Anzahl von Bits dargestellt. Dazu werden in einem zweiten Schritt die Amplituden der Abtastwerte quantisiert. Der Wertebereich, in dem die Amplitudenwerte auftreten, wird in eine festgelegte Anzahl von $2^{N_{\text{bit}}}$ gleich breiten Intervallen unterteilt. Alle Amplitudenwerte in einem Intervall werden als eine Dualzahl mit N_{bit} Bits codiert.

Die Vorgehensweise zur Quantisierung und Codierung ist exemplarisch in Bild 6 für eine lineare Unterteilung des zu quantisierenden Amplitudenbereichs in 8 Intervalle dargestellt. Auf der x-Achse werden die Amplituden der Abtastwerte aufgetragen. Der zu quantisierende Amplitudenbereich geht von $-A_{\text{max}}$ bis $+A_{\text{max}}$. Ein Quantisierungsintervall besitzt dann die Breite

$$\Delta x = \frac{A_{\text{max}} - (-A_{\text{max}})}{2^{N_{\text{bit}}}} = \frac{2 \cdot A_{\text{max}}}{2^3} = \frac{A_{\text{max}}}{4}$$

Die treppenförmige Abbildungskennlinie, die man als Quantisierungskennlinie bezeichnet, ergibt sich aus einer Abbildung der Werte in einem Quantisierungsintervall auf den Wert in der Mitte des Intervalls. Alle Werte in einem Intervall werden in diesem Fall als Dualzahl mit 3 Bits codiert.

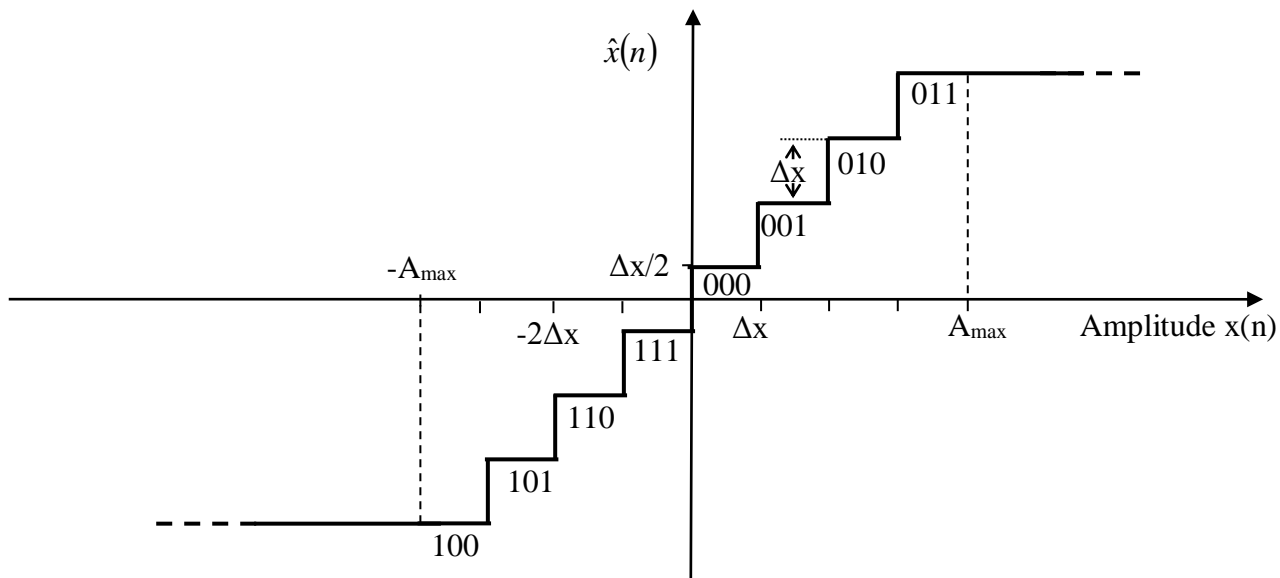


Bild 6: Quantisierungs- und Abbildungskennlinie für eine PCM bei einer Wortlänge von 3 Bit

Die Quantisierungskennlinie lässt sich mathematisch beschreiben durch

$$\hat{x}(n) = \text{sign}(x(n)) \cdot \left[\text{int} \left(\frac{|x(n)|}{\Delta x} \right) + \frac{1}{2} \right] \cdot \Delta x,$$

wobei $\text{sign}(\arg)$ das Vorzeichen von \arg repräsentiert und

$\text{int}(\arg)$ die Bestimmung der nächstkleineren ganzen Zahl von \arg definiert.

Die Abfolge von Abtastung, Quantisierung und Codierung bezeichnet man als *Pulsmodulation* (PCM). Das Ausgangssignal der PCM ist zeit- und wertediskret. Die praktische Realisierung der PCM erfolgt in einem Analog-Digital Umsetzer (ADU).

Die Quantisierungskennlinie beinhaltet auch die Vorgehensweise bei der Rekonstruktion eines analogen Signals aus der Folge von binären Codewörtern. Ein Codewort, das aus N_{bit} Bits besteht, wird wieder auf den Amplitudenwert in der Mitte des zugehörigen Intervalls abgebildet. Dabei tritt ein Quantisierungsfehler $e(n)$ auf, der sich als Differenz des quantisierten Amplitudenwerts und des ursprünglichen Werts ergibt. $e(n) = \hat{x}(n) - x(n)$

Die Quantisierungsfehler liegen dabei in dem Intervall $-\frac{\Delta x}{2} \leq e(n) \leq \frac{\Delta x}{2}$. Sie überlagern sich dem ursprünglichen Signal als sogenanntes Quantisierungsrauschen. Bei akustischen Signalen wird das Rauschen bei einer zu geringen Bitanzahl hörbar. Die Quantisierungsfehler, die bei einer

Quantisierung des bereits in Bild 6 verwendeten Signalabschnitts auftreten, werden in Bild 7 veranschaulicht. Dabei wird eine Quantisierung des Amplitudenbereichs von -2 bis +2 mit $N_{\text{bit}} = 3$

Bit vorgenommen, so dass sich die Breite eines Quantisierungsintervalls zu $\Delta x = \frac{2 \cdot 2}{2^3} = \frac{1}{2}$ ergibt

und die Quantisierungsfehler im Intervall $-\frac{1}{4} \leq e(n) \leq \frac{1}{4}$ auftreten.

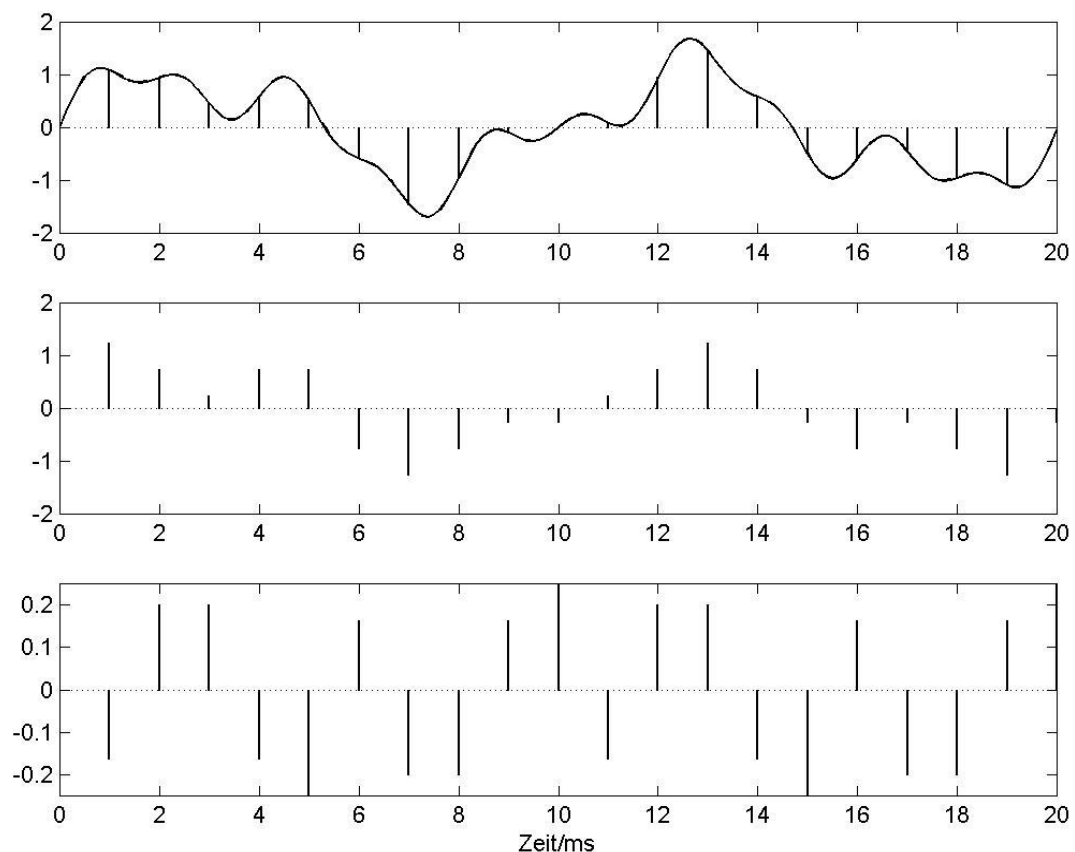


Bild 7: Analoges und PAM Signal (oben), PCM Signal (Mitte), Quantisierungsfehler (unten)

Um den Einfluss des Quantisierungsrauschens quantitativ zu beschreiben, betrachtet man das Verhältnis der Leistungen des Signals und des Rauschens. Man bezeichnet das Verhältnis daher auch als Signal/Rauschleistungsverhältnis (SNR = signal-to-noise ratio). Die Leistung eines Signals lässt sich bei Kenntnis der Auftrittswahrscheinlichkeiten aller Amplitudenwerte als Erwartungswert

der quadrierten Amplitude berechnen:
$$S = E\{x^2\} = \int_{x=-\infty}^{\infty} p(x) \cdot x^2 dx$$

Man bezeichnet die Funktion $p(x)$, die die Wahrscheinlichkeit des Auftretens der Amplitude x beschreibt, auch als *Verteilungsdichtefunktion*.

Bei einem natürlichen Signal kann man annehmen, dass die Quantisierungsfehler im Bereich $-\frac{\Delta x}{2} \leq e(n) \leq \frac{\Delta x}{2}$ mit gleich großer Wahrscheinlichkeit auftreten, da sich die Lage eines Amplitudenwerts in einem Quantisierungsintervall zufällig ergibt. Damit nimmt die Verteilungsdichtefunktion $p(e)$ das in Bild 8 dargestellt Aussehen an.

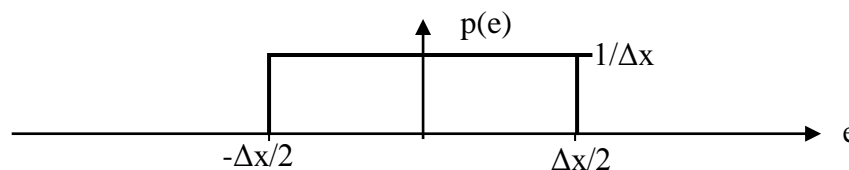


Bild 8: Verteilungsdichtefunktion des Quantisierungsfehlers

Eine Verteilungsdichtefunktion besitzt die grundlegende Eigenschaft, dass die Fläche unter der Funktion gleich Eins ist: $\int_{-\infty}^{\infty} p(x) dx = 1$. Damit ergibt sich die konstante Wahrscheinlichkeit des Auftretens eines Quantisierungsfehlers im Intervall $-\frac{\Delta x}{2} \leq e(n) \leq \frac{\Delta x}{2}$ zu $p(e) = \frac{1}{\Delta x}$. Die Leistung

N des Quantisierungsrauschens lässt sich damit berechnen zu

$$N = \int_{x=-\infty}^{\infty} e^2 \cdot p(e) de = \int_{-\Delta x/2}^{\Delta x/2} \frac{1}{\Delta x} \cdot e^2 de = \frac{1}{\Delta x} \left[\frac{e^3}{3} \right]_{-\Delta x/2}^{\Delta x/2} = \frac{1}{3 \cdot \Delta x} \cdot \left(\frac{\Delta x^3}{8} + \frac{\Delta x^3}{8} \right) = \frac{\Delta x^2}{12}$$

Nimmt man auch für das Signal ein gleichwahrscheinliches Auftreten der Amplitudenwerte im Quantisierungsbereich $-A_{\max} \leq x(n) \leq +A_{\max}$ an, so lässt sich die Leistung des Signals berechnen zu

$$S = \int_{x=-\infty}^{\infty} x^2 \cdot p(x) dx = \int_{-A_{\max}}^{A_{\max}} \frac{1}{2 \cdot A_{\max}} \cdot x^2 dx = \frac{1}{2 \cdot A_{\max}} \left[\frac{x^3}{3} \right]_{-A_{\max}}^{A_{\max}} = \frac{1}{6 \cdot A_{\max}} \cdot (2 \cdot A_{\max}^3) = \frac{A_{\max}^2}{3}$$

Damit ergibt sich das Signal/Rauschleistungsverhältnis in dB zu

$$SNR = 10 \cdot \log_{10} \left(\frac{S}{N} \right) = 10 \cdot \log_{10} \left(\frac{A_{\max}^2 \cdot 12}{3 \cdot \Delta x^2} \right)$$

$$\text{Mit } \Delta x = \frac{2 \cdot A_{\max}}{2^{N_{bit}}} \Rightarrow SNR = 10 \cdot \log_{10} \left(\frac{4 \cdot A_{\max}^2}{4 \cdot A_{\max}^2} \cdot 2^{2N_{bit}} \right) = 10 \cdot \log_{10} (2^{2N_{bit}})$$

$$SNR = 10 \cdot 2 \cdot N_{bit} \cdot \log_{10}(2) = N_{bit} \cdot 20 \cdot 0,301 = N_{bit} \cdot 6,02 \text{ dB}$$

Das Signal/Rauschleistungsverhältnis besitzt unter der Annahme des gleichwahrscheinlichen Auftretens aller Amplitudenwerte im gesamten Quantisierungsbereich eine lineare Abhängigkeit von der Bitanzahl N_{bit} . Damit ergibt sich beispielsweise die häufiger zu findende Angabe eines Signal/Rauschleistungsverhältnisses von 96 dB ($\approx 16 \cdot 6,02\text{dB}$) als Qualitätsangabe bei CDs (compact disks), bei denen ein Audiosignal in 2^{16} Intervallen, d.h. mit einer Wortlänge von 16 Bit, quantisiert wird. Zur Codierung von Sprache mit Hilfe der Puls-Code-Modulation benötigt man eine Quantisierung mit $N_{\text{bit}} = 12 \text{ Bit}$, um bei einer Rekonstruktion des Signals eine gute Sprachqualität zu gewährleisten. Damit ergeben sich Datenraten von

$$\begin{aligned} - & 8000 \frac{1}{s} \cdot 12 \text{ Bit} = 96000 \frac{\text{Bit}}{s} \quad \text{im Fall von Telefonsprache und} \\ - & 16000 \frac{1}{s} \cdot 12 \text{ Bit} = 192000 \frac{\text{Bit}}{s} \quad \text{im Fall von Breitbandsprache.} \end{aligned}$$

bei einer Verwendung der PCM zur Codierung von Sprache.

Die Bestimmung des Signal/Rauschleistungsverhältnisses mit $SNR = N_{\text{bit}} \cdot 6,02 \text{ dB}$ ist an zwei Bedingungen gebunden, die für Sprache in der Regel nicht erfüllt sind. Die erste Bedingung ist die volle Ausnutzung des Quantisierungsbereichs. Ein Analog-/Digitalumsetzer wird normalerweise so konfiguriert, dass es nicht zu einer Überschreitung des Quantisierungsbereichs (Übersteuerung) kommen sollte. Eine Übersteuerung kann zu sehr großen Quantisierungsfehlern führen, die sich auch akustisch störend bemerkbar machen. Daher wird man im Fall von Sprache die Quantisierung so konfigurieren, dass es auch bei einem „lauten“ Sprecher nicht zur Übersteuerung kommt. Andererseits führt dies bei einem „leisen“ Sprecher dazu, dass möglicherweise nur ein kleiner Teil des Quantisierungsbereichs genutzt wird.

Auch die zweite Bedingung, dass alle Amplitudenwerte mit gleicher Wahrscheinlichkeit auftreten, ist bei Sprache nicht erfüllt. Tatsächlich treten bei Sprachsignalen kleine Amplitudenwerte wesentlich häufiger auf als große. Eine näherungsweise Darstellung der Verteilungsdichtefunktion für Sprachsignale findet sich in Bild 9. Diese Funktion besitzt die Charakteristik einer Laplace oder Gamma Verteilung. Dabei ist die logarithmische Skalierung der Ordinate zu beachten.

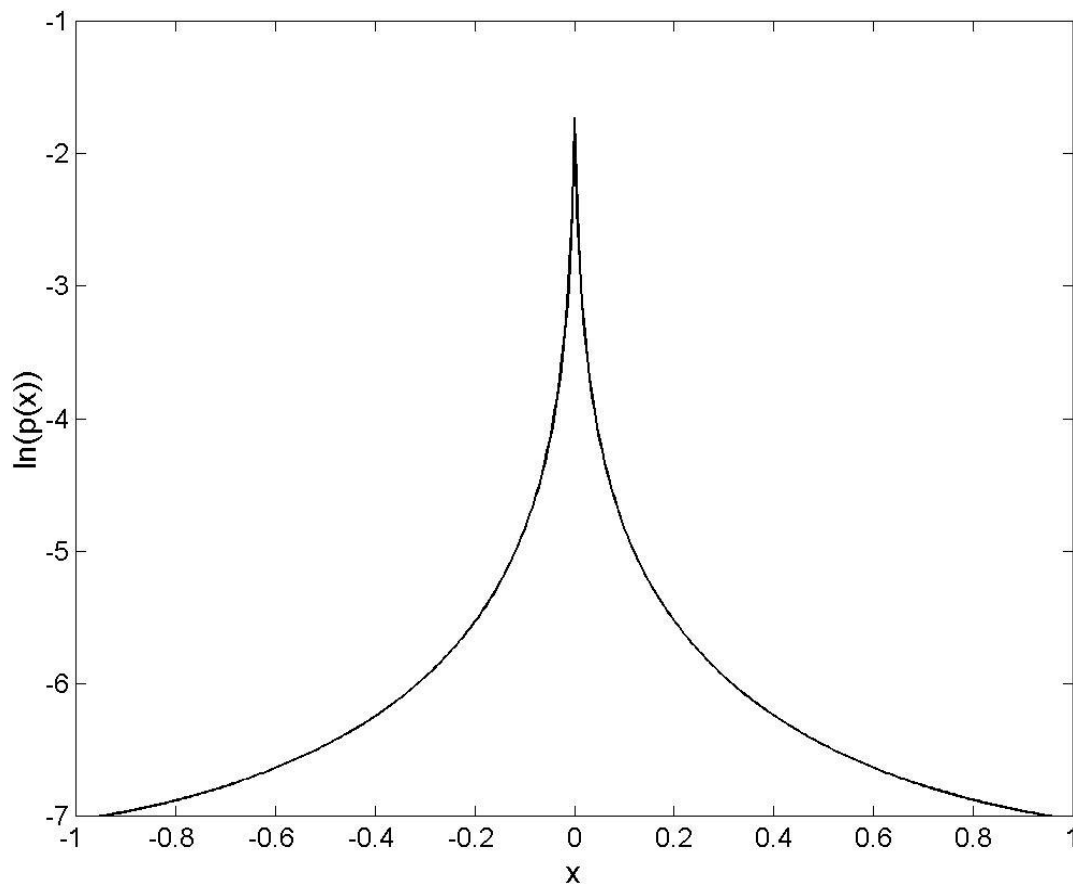


Bild 9: Gamma Funktion zur Beschreibung der Verteilungsdichtefunktion bei Sprachsignalen

Da die kleinen Amplitudenwerte wesentlich häufiger als große auftreten, resultiert daraus auch eine im Vergleich zur Gleichverteilung wesentlich geringere Signalleistung. Für Sprache erhält man daher bei linearer Quantisierung ein SNR, das um etwa 6 bis 7 dB schlechter ist als bei einem Signal, dessen Amplitudenwerte eine Gleichverteilung über den gleichen Amplitudenbereich aufweisen.

Zusammenfassend kann man festhalten, dass die häufig angegebene näherungsweise Beschreibung $SNR = N_{bit} \cdot 6,02 \text{ dB}$ als Maß für die Güte der Quantisierung die in der Praxis anzutreffenden Signal/Rauschleistungsverhältnisse bei Sprach- und Audiosignalen nicht wiedergibt. Meist stellt man Werte des SNR fest, die um etwa 6 bis 20 dB kleiner sind als die der näherungsweise Beschreibung.