

# Waveform Coding-Techniken

## Inhalt

[Einführung](#)

[Voraussetzungen](#)

[Anforderungen](#)

[Verwendete Komponenten](#)

[Konventionen](#)

[Pulsmodulation](#)

[Filterung](#)

[Probenahme](#)

[Digitalisierung von Sprachanwendungen](#)

[Quantifizierung und Codierung](#)

[Konkurrenzfähig](#)

[A-law- und u-law-Companding](#)

[Differenzial Puls Code Modulation](#)

[Adaptives DPCM](#)

[Spezifische 32 KB/s-Schritte](#)

[Zugehörige Informationen](#)

## Einführung

Obwohl Menschen gut für die analoge Kommunikation ausgestattet sind, ist die analoge Übertragung nicht besonders effizient. Wenn analoge Signale aufgrund eines Übertragungsverlusts schwach werden, ist es schwierig, die komplexe analoge Struktur von der Struktur des zufälligen Übertragungsrauschens zu trennen. Wenn Sie analoge Signale verstärken, verstärken sie auch das Rauschen, und schließlich werden analoge Verbindungen zu laut. Digitale Signale mit nur "Ein-Bit-" und "Zero-Bit"-Zuständen lassen sich leichter von Geräusch trennen. Sie können ohne Korruption verstärkt werden. Digitale Kodierung ist bei Fernverbindungen eher gegen Rauschschäden geschützt. Auch die Kommunikationssysteme der Welt haben sich in ein digitales Übertragungsformat konvertiert, das als Pulscode-Modulation (PCM) bezeichnet wird. PCM ist ein Kodierungstyp, der als "Wellenform"-Kodierung bezeichnet wird, da er eine codierte Form der ursprünglichen Sprachwellenform erstellt. Dieses Dokument beschreibt auf hoher Ebene den Umwandlungsprozess analoger Sprachsignale in digitale Signale.

## Voraussetzungen

### Anforderungen

Für dieses Dokument bestehen keine speziellen Anforderungen.

### Verwendete Komponenten

Dieses Dokument ist nicht auf bestimmte Software- und Hardwareversionen beschränkt.

## Konventionen

Weitere Informationen zu Dokumentkonventionen finden Sie in den [Cisco Technical Tips Conventions](#).

## Pulscodemodulation

PCM ist eine Wellenformcodierungsmethode, die in der ITU-T G.711-Spezifikation definiert ist.

### Filterung

Der erste Schritt zur Umwandlung des Signals von analog in digital besteht darin, die Komponente mit höherer Frequenz zu filtern. Dies erleichtert die Downstream-Konvertierung dieses Signals. Der größte Teil der Energie der gesprochenen Sprache liegt zwischen 200 oder 300 Hertz und etwa 2700 oder 2800 Hertz. Etwa 3.000 Hertz Bandbreite für Standardsprache und Standardsprachkommunikation ist vorhanden. Daher benötigen sie keine genauen Filter (sehr teuer). Eine Bandbreite von 4000 Hertz wird aus Sicht der Geräte erzeugt. Dieser Bandbegrenzer-Filter verhindert Aliasing (Antialiasing). Dies geschieht, wenn das analoge Eingangssignal nicht abgetastet wird, definiert durch das Nyquist-Kriterium als  $F_s < 2(BW)$ . Die Abtastfrequenz ist kleiner als die höchste Frequenz des analogen Eingangssignals. Dadurch entsteht eine Überlappung zwischen dem Frequenzspektrum der Proben und dem analogen Eingangssignal. Der Tiefpass-Ausgangfilter, der zur Rekonstruktion des ursprünglichen Eingangssignals verwendet wird, ist nicht intelligent genug, um diese Überlappung zu erkennen. Daher wird ein neues Signal erzeugt, das nicht von der Quelle stammt. Diese Erstellung eines falschen Signals bei der Abfrage wird Aliasing genannt.

### Probenahme

Der zweite Schritt zur Umwandlung eines analogen Sprachsignals in ein digitales Sprachsignal besteht darin, das gefilterte Eingangssignal mit konstanter Abtastfrequenz zu testen. Dies wird durch die so genannte Pulsamplitudenmodulation (PAM) erreicht. In diesem Schritt wird das ursprüngliche analoge Signal verwendet, um die Amplitude eines Impulses mit konstanter Amplitude und Frequenz zu modulieren. (Siehe Abbildung 2.)

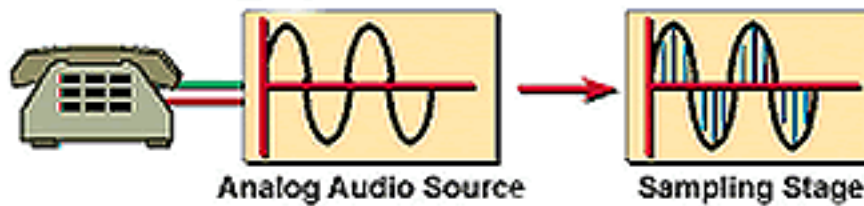
Der Pulszug bewegt sich mit konstanter Frequenz, die so genannte Abtastfrequenz. Das analoge Sprachsignal kann mit einer Million Mal pro Sekunde oder zwei- bis dreimal pro Sekunde abgetastet werden. Wie wird die Häufigkeit der Probenahme bestimmt? Ein Wissenschaftler namens Harry Nyquist entdeckte, dass das ursprüngliche analoge Signal rekonstruiert werden kann, wenn genügend Proben genommen werden. Er stellte fest, dass bei einer Abtastrate, die mindestens doppelt so hoch ist wie die höchste Frequenz des ursprünglichen analogen Sprachsignals, dieses Signal durch einen Low-Pass-Filter am Ziel rekonstruiert werden kann. Das Nyquist-Kriterium ist wie folgt angegeben:

$$F_s > 2(BW)$$

$F_s$  = Sampling frequency

$BW$  = Bandwidth of original analog voice signal

Abbildung 1: Analoges Sampling



## Digitalisierung von Sprachanwendungen

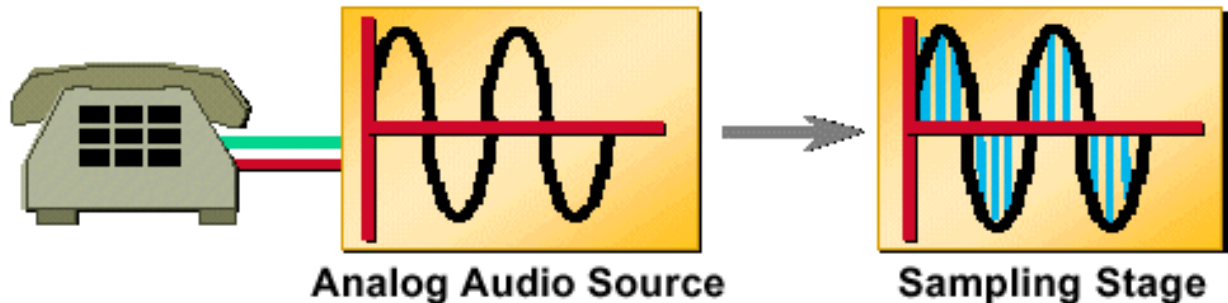
Nach dem Filtern und Abtasten (mithilfe von PAM) eines analogen Eingangs-Sprachsignals, werden diese Samples als Vorbereitung auf die Übertragung über ein Telefonnetz digitalisiert. Die Digitalisierung analoger Sprachsignale wird als PCM bezeichnet. Der einzige Unterschied zwischen PAM und PCM besteht darin, dass PCM den Prozess einen Schritt weiter fortsetzt. PCM decodiert jedes analoge Beispiel mithilfe von binären Codewörtern. PCM verfügt über einen Analog-Digital-Konverter auf der Quellseite und einen Digital-to-Analog-Konverter auf der Zielseite. PCM verwendet eine Methode, die als Quantisierung bezeichnet wird, um diese Proben zu codieren.

## Quantifizierung und Codierung

Abbildung 2: Pulscode-Modulation - Nyquist Theorem

# Pulse Code Modulation—Nyquist Theorem

Voice Bandwidth =  
200 Hz to 3400 Hz



## Codec Technique

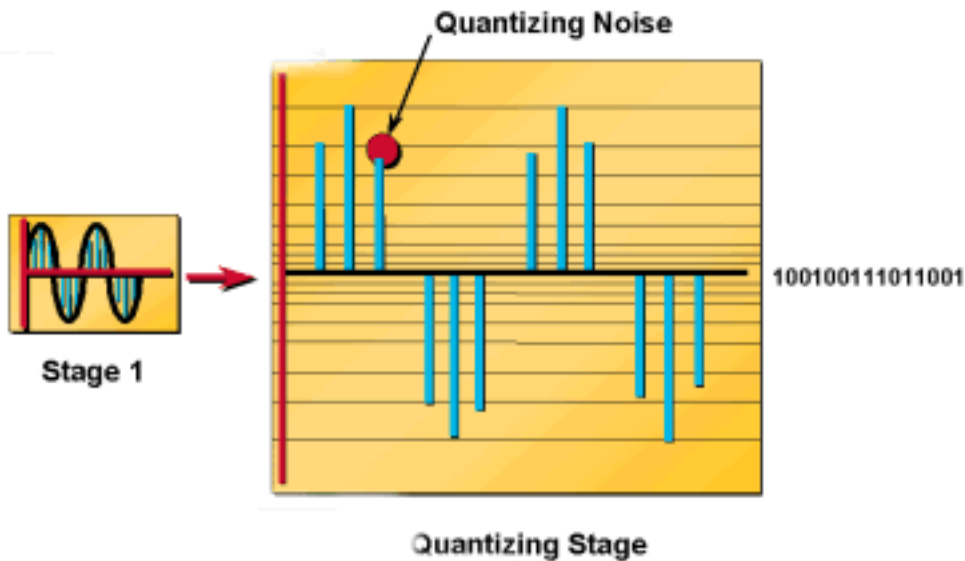
Quantisierung ist der Prozess, bei dem jeder analoge Beispielwert in einen eigenen Wert konvertiert wird, dem ein eindeutiges digitales Codewort zugewiesen werden kann.

Wenn die Eingangssignalproben in die Quantisierungsphase eintreten, werden sie einem Quantisierungsintervall zugewiesen. Alle Quantisierungsintervalle sind in gleichem Abstand (einheitliche Quantisierung) im dynamischen Bereich des Eingangs-Analogsignals. Jedem Quantisierungsintervall wird ein separater Wert in Form eines binären Codeworts zugewiesen. Die standardmäßige Wortgröße beträgt acht Bit. Wenn ein Eingangsanalogsignal 8000 Mal pro Sekunde abgetastet wird und jeder Probe ein Codewort mit einer Länge von 8 Bits zugewiesen wird, beträgt die maximale Übertragungsbitrate für Telefonesysteme mit PCM 64.000 Bit pro Sekunde. Abbildung 2 zeigt, wie die Bitrate für ein PCM-System abgeleitet wird.

Jeder Eingangsprobe wird ein Quantisierungsintervall zugewiesen, das seiner Amplitudenhöhe am nächsten kommt. Wird einer Eingabeprobe kein Quantisierungsintervall zugewiesen, das ihrer tatsächlichen Höhe entspricht, wird ein Fehler in den PCM-Prozess eingefügt. Dieser Fehler wird als Quantisierungsgeräusch bezeichnet. Die Quantifizierung des Rauschens entspricht dem zufälligen Rauschen, das das Signal-Rausch-Verhältnis (SNR) eines Sprachsignals beeinflusst. SNR ist ein Maß für die Signalstärke im Verhältnis zum Hintergrundgeräusch. Das Verhältnis wird in der Regel in Dezibel (dB) gemessen. Wenn die Stärke des eingehenden Signals in Mikrovolt  $V_s$  und der Rauschpegel auch in Mikrovolt  $V_n$  ist, wird das Signal-Rausch-Verhältnis  $S/N$  in Dezibel durch die Formel  $S/N = 20 \log_{10}(V_s/V_n)$  angegeben. SNR wird in Dezibel (dB) gemessen. Je höher die SNR, desto besser die Sprachqualität. Das Mengengeräuschpegel reduziert die SNR eines Signals. Daher beeinträchtigt eine erhöhte Quantisierung des Rauschens die Qualität eines Sprachsignals. Abbildung 3 zeigt, wie ein Quantisierungsgeräusch erzeugt wird. Zu Codierungszwecken gibt ein N-Bit-Wort  $2^N$ -Quantisierungsetiketten aus.

Abbildung 3: Analog zur digitalen Konvertierung

# Pulse Code Modulation— Analog to Digital Conversion



Eine Möglichkeit, das Quantisierungsgeräusch zu reduzieren, besteht darin, die Menge der Quantisierungsintervalle zu erhöhen. Die Differenz zwischen der Amplitudenhöhe des Eingangssignals und dem Quantisierungsintervall nimmt ab, wenn die Quantisierungsintervalle erhöht werden (Erhöhungen der Intervalle verringern das Quantisierungsgeräusch). Die Anzahl der Codewörter muss jedoch auch proportional zur Erhöhung der Quantisierungsintervalle erhöht werden. Dieser Prozess führt zu zusätzlichen Problemen, die sich mit der Kapazität eines PCM-Systems zur Behandlung von mehr Codewörtern befassen.

SNR (einschließlich Quantisierungsgeräusch) ist der wichtigste Einzelfaktor, der die Sprachqualität bei gleichmäßiger Quantisierung beeinflusst. Bei der Uniform-Quantisierung werden im gesamten dynamischen Bereich eines analogen Eingangssignals gleichmäßige Quantisierungsstufen verwendet. Aus diesem Grund verfügen niedrige Signale über eine kleine SNR-Funktion (geringe Sprachqualität) und hohe Signale über eine große SNR-Funktion (hohe Sprachqualität). Da die meisten erzeugten Sprachsignale von geringer Qualität sind, stellt eine bessere Sprachqualität bei höherem Signalpegel eine sehr ineffiziente Methode zur Digitalisierung von Sprachsignalen dar. Zur Verbesserung der Sprachqualität bei niedrigeren Signalwerten wird die gleichmäßige Quantisierung (einheitliches PCM) durch einen nicht einheitlichen Quantisierungsprozess namens "Companding" ersetzt.

## Konkurrenzfähig

Das "Companding" bezieht sich auf den Prozess, zunächst ein analoges Signal an der Quelle zu komprimieren und dieses Signal dann wieder auf seine ursprüngliche Größe zu erweitern, wenn es sein Ziel erreicht. Der Begriff Companding wird durch die Kombination der beiden Begriffe, Komprimierung und Erweiterung, in einem Wort erstellt. Zum Zeitpunkt des Companding werden analoge Eingangssignalproben in logarithmische Segmente komprimiert. Jedes Segment wird dann mit einheitlicher Quantisierung quantifiziert und codiert. Der Komprimierungsprozess ist logarithmisch. Die Komprimierung erhöht sich, wenn die Stichprobensignale zunehmen. Mit anderen Worten, die größeren Probensignale werden mehr komprimiert als die kleineren Probensignale. Dies führt dazu, dass das Quantisierungsgeräusch mit zunehmendem Probensignal zunimmt. Eine logarithmische Steigerung des Quantisierungsrauschens im

gesamten dynamischen Bereich eines Eingangsstichprobensignals hält die SNR-Konstante in diesem dynamischen Bereich aufrecht. Die ITU-T-Standards für das Companding werden als A-law und u-law bezeichnet.

## A-law- und u-law-Companding

A-law und u-law sind vom Beratungsausschuss für internationale Telefonie und Telegrafie (CCITT) G.711 definierte Audiokomprimierungsschemata (Codecs), die 16-Bit-lineare PCM-Daten auf acht Bit logarithmischer Daten komprimieren.

### **a-law-Compander**

Die Begrenzung der linearen Stichprobenwerte auf 12 Größenbits, die A-law-Komprimierung wird durch diese Gleichung definiert, wobei A der Kompressionsparameter ( $A=87,7$  in Europa) und x die normalisierte Ganzzahl, die komprimiert werden soll.

$$F(x) = \begin{cases} \frac{A * |x|}{1 + \ln(A)} & 0 \leq |x| < \frac{1}{A} \\ \frac{\text{sgn}(x) * (1 + \ln(A|x|))}{1 + \ln(A)} & \frac{1}{A} \leq |x| \leq 1 \end{cases}$$

### **Nicht-gesetzlicher Compander**

Die Begrenzung der linearen Stichprobenwerte auf dreizehn Größenbits, wird die u-law-Komprimierung (u-law und Mu-law werden in diesem Dokument synonym verwendet) durch diese Gleichung definiert, wobei m der Kompressionsparameter ist ( $m=255$  in den USA und Japan) und x die normalisierte Ganzzahl, die komprimiert werden soll.

$$F(x) = \frac{\text{sgn}(x) * \ln(1 + \mu|x|)}{\ln(1 + \mu)} \quad 0 \leq |x| \leq 1$$

Der A-Law-Standard wird in erster Linie von Europa und der übrigen Welt verwendet. u-law wird von Nordamerika und Japan verwendet.

## Ähnlichkeiten zwischen A-law und U-law

- Beide sind lineare Näherungen der logarithmischen Ein-/Ausgangsbeziehung.
- Beide werden mithilfe von 8-Bit-Codetexten implementiert (256 Ebenen, eine für jedes Quantisierungsintervall). Acht-Bit-Codewörter ermöglichen eine Bitrate von 64 Kbit/s (Kbit/s). Dies wird berechnet, indem die Abtastrate (die doppelte Eingangsfrequenz) mit der Größe des Codewortes ( $2 \times 4 \text{ kHz} \times 8 \text{ Bit} = 64 \text{ Kbit/s}$ ) multipliziert wird.
- Beide unterteilen einen dynamischen Bereich in insgesamt 16 Segmente: Acht positive und acht negative Segmente. Jedes Segment ist doppelt so lang wie das vorherige. Innerhalb jedes Segments wird eine einheitliche Quantisierung verwendet.
- Beide verwenden einen ähnlichen Ansatz zum Kodieren des 8-Bit-Worts: First (MSB)

identifiziert Polarität. Bits 2, 3 und 4 Identifikationssegment. Die letzten vier Bits quantifizieren das Segment sind die niedrigeren Signalpegel als A-law.

### Unterschiede zwischen A-law und U-law

- Unterschiedliche lineare Annäherungen führen zu unterschiedlichen Längen und Pisten.
- Die numerische Zuweisung der Bitpositionen im 8-Bit-Codewort zu Segmenten und die Quantisierungsebenen innerhalb der Segmente sind unterschiedlich.
- A-law bietet einen größeren Dynamikbereich als U-law.
- Das u-law bietet eine bessere Signal-/Verzerrungsleistung für Signale auf niedriger Ebene als das A-law.
- Für ein einheitliches PCM-Äquivalent sind 13-Bit-Werte erforderlich. Für ein einheitliches PCM-Äquivalent sind 14 Bit für das u-law erforderlich.
- Eine internationale Verbindung muss A-law verwenden, u zu A Umwandlung ist die Verantwortung des U-law-Landes.

### Differenzial Puls Code Modulation

Zum Zeitpunkt des PCM-Prozesses sind die Unterschiede zwischen den Eingangsstichprobensignalen minimal. Differenzial PCM (DPCM) ist so konzipiert, dass dieser Unterschied berechnet und dann dieses kleine Differenzsignal anstatt des gesamten Eingangsstichprobe-Signals übertragen wird. Da der Unterschied zwischen den Eingangsproben kleiner ist als eine gesamte Eingangsprobe, wird die Anzahl der für die Übertragung erforderlichen Bits verringert. Dies ermöglicht eine Reduzierung des Durchsatzes für die Übertragung von Sprachsignalen. Mithilfe von DPCM kann die Bitrate der Sprachübertragung auf 48 Kbit/s reduziert werden.

Wie berechnet DPCM den Unterschied zwischen dem aktuellen Probensignal und einer vorherigen Probe? Der erste Teil von DPCM funktioniert genau wie PCM (deshalb wird er Differenzial PCM genannt). Das Eingangssignal wird mit konstanter Abtastfrequenz (doppelt so hohe Eingangsfrequenz) abgetastet. Anschließend werden diese Beispiele mithilfe des PAM-Prozesses moduliert. An diesem Punkt übernimmt der DPCM-Prozess. Das gesampelte Eingangssignal wird in einem so genannten Vorhersager gespeichert. Der Vorhersager nimmt das gespeicherte Stichprobensignal und sendet es durch ein Differenzierungsmerkmal. Das Differenzierungsmerkmal vergleicht das vorherige Probensignal mit dem aktuellen Probensignal und sendet diesen Unterschied an die Quantifizierungs- und Kodierungsphase des PCM (diese Phase kann gleichförmig quantifiziert oder mit A-law oder u-law konform sein). Nach Quantifizierung und Kodierung wird das Differenzsignal an sein endgültiges Ziel übertragen. Am Empfangs-Ende des Netzwerks wird alles umgekehrt. Zuerst wird das Differenzsignal dequantifiziert. Dann wird dieses Differenzsignal einem in einem Voraussager gespeicherten Mustersignal hinzugefügt und an einen Low-Pass-Filter gesendet, der das ursprüngliche Eingangssignal rekonstruiert.

DPCM ist eine gute Möglichkeit, die Bitrate für Sprachübertragungen zu reduzieren. Es verursacht jedoch auch andere Probleme im Zusammenhang mit der Sprachqualität. DPCM quantifiziert und codiert den Unterschied zwischen einem früheren Probeneingangssignal und einem aktuellen Probeneingangssignal. DPCM quantifiziert das Differenzsignal mittels einheitlicher Quantisierung. Eine einheitliche Quantisierung erzeugt eine SNR, die für kleine Stichprobensignale klein und für große Eingangsstichprobensignale groß ist. Daher ist die Sprachqualität bei höheren Signalen besser. Dieses Szenario ist sehr ineffizient, da die meisten von der menschlichen Stimme

erzeugten Signale klein sind. Die Sprachqualität muss sich auf kleine Signale konzentrieren. Um dieses Problem zu lösen, wird adaptives DPCM entwickelt.

## Adaptives DPCM

Adaptive DPCM (ADPCM) ist eine in der ITU-T G.726-Spezifikation definierte Wellenformcodierungsmethode.

ADPCM passt die Quantisierungsstufen des Differenzsignals an, das zum Zeitpunkt des DPCM-Prozesses generiert wurde. Wie passt ADPCM diese Quantisierungsstufen an? Ist das Differenzsignal gering, erhöht ADPCM die Größe der Quantisierungsstufen. Ist das Differenzsignal hoch, verringert ADPCM die Größe der Quantisierungsstufen. ADPCM passt also die Quantisierungsstufe an die Größe des Eingangsdifferenz-Signals an. Dadurch wird eine SNR erzeugt, die im gesamten dynamischen Bereich des Differenzsignals einheitlich ist. Bei Verwendung von ADPCM wird die Bitrate für die Sprachübertragung auf 32 Kbit/s reduziert. Dies entspricht der Hälfte der Bitrate von A-law oder einem fehlerhaften PCM. ADPCM produziert "Mautqualität"-Sprachverbindungen wie A-law oder u-law PCM. Coder muss über eine Feedback-Schleife verfügen, wobei Encoder-Ausgabebits zur Neukalibrierung des Quantizers verwendet werden.

## Spezifische 32 KB/s-Schritte

Anwendbar als ITU-Standard G.726.

- Umwandlung von A-law- oder Mu-law-PCM-Proben in eine lineare PCM-Probe.
- Berechnen Sie den prognostizierten Wert des nächsten Beispiels.
- Den Unterschied zwischen der tatsächlichen Stichprobe und dem prognostizierten Wert messen.
- Code-Unterschied: vier Bit, senden Sie diese Bits.
- vier Bit an die Prognose zurückgeben.
- Füttern Sie vier Bit zurück zum Quantizer.

## Zugehörige Informationen

- [Unterstützung von Sprachtechnologie](#)
- [Produkt-Support für Sprach- und Unified Communications](#)
- [Fehlerbehebung bei Cisco IP-Telefonie](#)
- [Technischer Support - Cisco Systems](#)