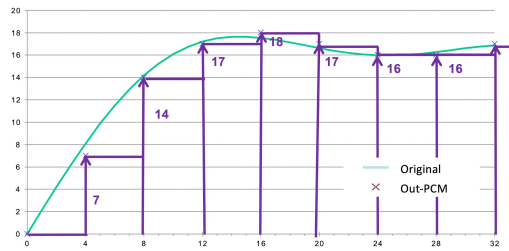


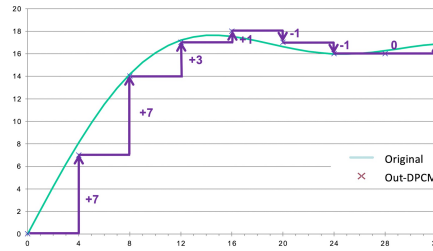
Audiocodierung

PCM (linear quantisiert)



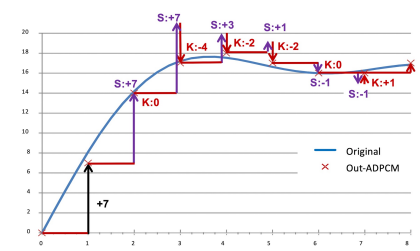
Es wird immer der exakte absolute Wert codiert.
Für jeden Abtastwert muss der gesamte Wertebereich zur Verfügung stehen.

Differential-PCM (DPCM)



Es wird lediglich die Differenz codiert. Bei kontinuierlichen Signalen ist der Wertebereich pro Abtastwert viel geringer.

Adaptive Differential-PCM (ADPCM)



Es wird die Abweichung der Differenz codiert. Dies erlaubt ein noch geringeren Wertebereich und kleinere Bitbreite für Signale mit grossen Abschnitten und stetiger Steigung.

Tonerzeugung eines reinen Sinustones

Die einzelnen Samples S_i für eine gewünschte Frequenz f kann in Abhängigkeit der Abtastrate R , und dem Skalierungsfaktor K berechnet werden:

$$S_i = K \cdot \sin\left(\frac{i \cdot 2 \pi \cdot f}{R}\right), \quad K = 2^{\text{Bps}-1} - 1 \quad (\text{Bps} = \text{Bits pro Sekunde}), \quad \text{z.B. } K = 2^{15} - 1 \quad (\text{bei } 16 \text{ Bit pro Sample}), \quad f = 1 \text{ kHz}, \quad R = 32 \text{ kHz}$$

Schalldruckpegel (Sound Pressure Level, SPL)

Logarithmische Grösse in Dezibel (dB) zur Beschreibung der Stärke eines Schallereignisses.

Schallpegel : $L_p = 20 \text{ dB} \cdot \log_{10}\left(\frac{p}{p_0}\right)$, p = Effektiver Schalldruck [Pa], p_0 = Bezugsschalldruck (Hörschwelle $p_0 = 0.00002 \text{ Pa}$)

Verdoppelung des Schalldrucks: $L_p = 20 \text{ dB} \cdot \log_{10}(2) = 6.02 \text{ dB}$ (+ 6 dB)

Menschliche Hörschwelle : Schalldruckpegel, bei dem das menschliche Gehör Töne und Geräusche gerade noch wahrnimmt.

Zeitliche Maskierung : Leise Töne vor, während und nach einem Ereignis sind nicht hörbar.

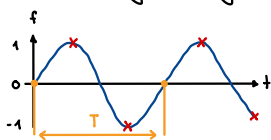
Spektrale Maskierung : Ein lauter Ton maskiert andere Töne mit leicht unterschiedlicher Frequenz.

Sub-Band Coding : Frequenz-Spektrum wird in Sub-Bänder unterteilt. Es werden nur Bits zum Quantisieren wie nötig eingesetzt. Dies verbessert die Kompression, das Quantisierungsrauschen wird allerdings erhöht. Ziel ist es, das Quantisierungsrauschen gerade unter die Maskierschwelle zu bringen.

Pulse Code Modulation (PCM)

Filterung : Hohe und tiefe Frequenzen werden entfernt. Signal wird auf hörbaren Frequenzbereich begrenzt.

Abtastung : Abtastung des Signals mit dem Abtasttheorem von Shannon : $F_{\text{abtast}} [\text{Hz}] > 2 \cdot f_{\text{max}} [\text{Hz}]$



$$T = \text{Periodendauer [s]} = \frac{1}{f}$$

$$f = \text{Frequenz [Hz]}$$

x Innerhalb einer Periode muss das Signal mind. 2-Mal abgetastet werden

$$\text{Abtastfrequenz / Samplingrate [Hz]} = \text{Samples (Anzahl Stützstellen) pro Sekunde}$$

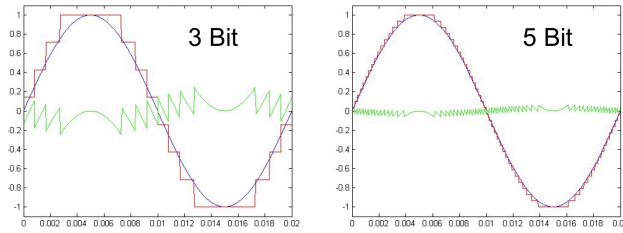
↳ Gibt an, wie oft in einer Sekunde der Audiopegel erfasst und gespeichert wird.

z.B. Bei 44.1 kHz werden 44'100 Werte für eine Sekunde Audio gespeichert.

$$\text{Bitrate / Datenrate [Bit/s]} = \text{Anzahl Kanäle} \cdot \text{Samplerate [Hz]} \cdot \text{Auflösung [Bit]}$$

↳ Steht für die Bandbreite der Audiodatei, also welche Datenmenge in einer Sekunde verarbeitet wird.

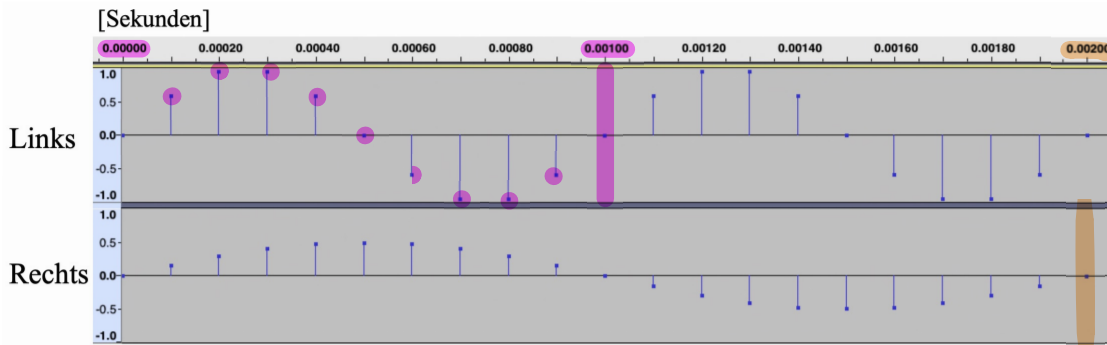
Quantisierung : Durch Abtastung ist ein Treppenförmiges Signal entstanden, welches nicht exakt dem Original entspricht. Differenz zwischen Treppenstufen und Originalsignal heisst Quantisierungsrauschen. Jede Bit-Zunahme bedeutet Verdoppelung der Anzahl Stufen und damit eine Halbierung des Quantisierungsrauschen (-6dB).



! Auflösung [Bit] gibt an, wie viel Speicher für einen Sample-Wert genutzt wird. z.B. erlauben 5 Bit (2^5) 32 Werte

Anzahl Stützstellen = Samplingrate : Frequenz

Codierung : Jedem quantisierten Messwert wird ein Wert von einer bestimmten Bitlänge zugeordnet.
 Bitstrom: Anzahl Bit pro Messwert Abtastfrequenz z.B. $44'100 \cdot 16 \text{ Bit} = 705'600 \text{ Bit/s}$



Abtastfrequenz = 10 Abtastungen (Säulen) zwischen 0.0005s und 0.001s, Wie viele Abtastungen in 1.000s?

↳ $10 \cdot 1000 = 10'000 \text{ Hz} = 10 \text{ kHz}$! $1 \text{ kHz} = 1'000 \text{ Hz}$, $1 \text{ s} = 1'000 \text{ ms}$

Frequenz oberes Signal (linker Kanal) = In 0.001s 1 Sin-Kurve, Wie viele Sin-Kurven in 1.000s?

↳ $1'000 : 1 = 1'000 \text{ Hz} = 1 \text{ kHz}$

Frequenz unteres Signal (rechter Kanal) = In 0.002s 1 Sin-Kurve, Wie viele Sin-Kurven in 1.000s?

↳ $1'000 : 2 = 500 \text{ Hz}$

Um wie viel Dezibel ist das untere Signal gedämpfter als das obere?

Oberes Signal: Frequenz = 1 kHz
 Unteres Signal: Frequenz = 500 Hz

$$\left. \begin{array}{l} \text{Oberes Signal: Frequenz} = 1 \text{ kHz} \\ \text{Unteres Signal: Frequenz} = 500 \text{ Hz} \end{array} \right\} 20 \cdot \log_{10} \left(\frac{500}{1000} \right) = -6.021 \text{ dB}, \text{ unteres Signal ist also 6 dB gedämpft.}$$

Wave File Format

Audio unkomprimiert, Enthält PCM Rohdaten, Headerinformationen vor den Daten

Grösse der Audiodatei : (Abtastfrequenz [Hz] · Dauer [s] + 1) · Auflösung [Byte] · Anzahl Kanäle + Header [Byte] = [Byte]

z.B. Header 44 Bytes, linker und rechter Kanal Auflösung je 16 Bit, Abtastfrequenz 10 kHz, Dauer 4 Sekunden und bei 4s genau das letzte Sample

↳ $(10'000 \text{ Hz} \cdot 4 \text{ s} + 1) \cdot 2 \text{ Byte} \cdot 2 + 44 \text{ Byte} = 160'048 \text{ Byte}$

16 Bit : 8 = 2 Byte