Gruppen Di-T14 / Mi-T25

Tutorübung zu
Grundlagen: Rechnernetze und Verteilte Systeme
(SS 16)

Michael Schwarz

Institut für Informatik
Technische Universität München

26.04 / 27.04.2016

Abtastung, Rekonstruktion und Quantisierung

Natürlich vorkommende Signale sind zeitkontinuierlich und wertkontinuierlich, d. h. sie nehmen zu unendlich vielen Zeitpunkten beliebige reelle Werte an.

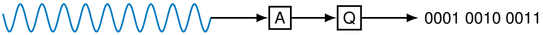
Problem für Computer:

- Endlicher Speicher
- Endliche Rechengenauigkeit

Lösung: Diskretisierung von Signalen im

- Zeitbereich (Abtastung) und
- Wertbereich (Quantisierung).

Ein zeit- und wertdiskretes Signal ist digital und wird in Wörtern fester Länge gespeichert.



Vergleiche: Nutzung von Fix- bzw. Gleitkommazahlen anstelle von reellen Zahlen entspricht einer Rundung (Quantisierung) auf eine endliche Anzahl diskreter Stufen.

Das Signal $s(t)$ wird mittels des Einheitsimpulses (Dirac-Impulses) $\delta[t]$ in äquidistanten Abständen T_a (Abtastintervall) für $n \in \mathbb{Z}$ abgetastet:

$$\hat{s}(t) = s(t) \sum_{n=-\infty}^{\infty} \delta[t - nT_a], \text{ mit } \delta[t - nT_a] = \begin{cases} 1 & t = nT_a, \\ 0 & \text{sonst.} \end{cases}$$

Da $\hat{s}(t)$ nur zu den Zeitpunkten nT_a für ganzzahlige n von Null verschieden ist, vereinbaren wir die Schreibweise $\hat{s}[n]$ für zeitdiskrete aber wertkontinuierliche Signale.

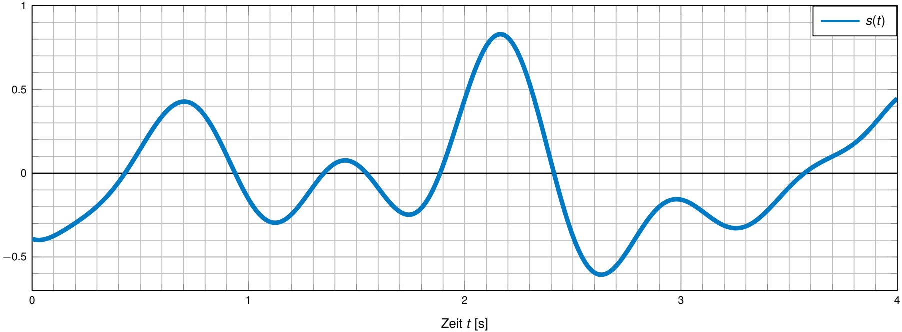


Abbildung: Zeitkontinuierliches Signal $s(t)$

Abtastung

Das Signal $s(t)$ wird mittels des Einheitsimpulses (Dirac-Impulses) $\delta[t]$ in äquidistanten Abständen T_a (Abtastintervall) für $n \in \mathbb{Z}$ abgetastet:

$$\hat{s}(t) = s(t) \sum_{n=-\infty}^{\infty} \delta[t - nT_a], \text{ mit } \delta[t - nT_a] = \begin{cases} 1 & t = nT_a, \\ 0 & \text{sonst.} \end{cases}$$

Da $\hat{s}(t)$ nur zu den Zeitpunkten nT_a für ganzzahlige n von Null verschieden ist, vereinbaren wir die Schreibweise $\hat{s}[n]$ für zeitdiskrete aber wertkontinuierliche Signale.

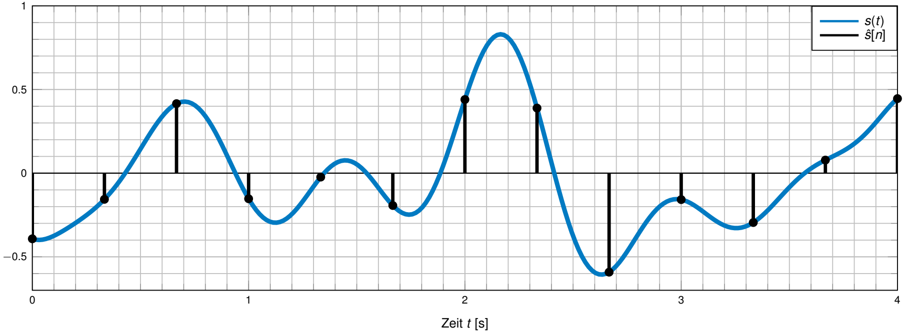


Abbildung: Zeitkontinuierliches Signal $s(t)$ und Abtastwerte $\hat{s}[n]$

Mittels der Abtastwerte $\hat{s}[n]$ ist es möglich, das ursprüngliche Signal $s(t)$ zu rekonstruieren:

$$s(t) \approx \sum_{n=-\infty}^{\infty} \hat{s}[n] \cdot \operatorname{sinc}\left(\frac{t - nT_a}{T_a}\right).$$

- Abtastwerte sind Stützstellen und
- dienen als Gewichte für eine passende Ansatzfunktion (trigonometrische Interpolation, vgl. Polynominterpolation $\rightarrow$ Numerisches Programmieren).

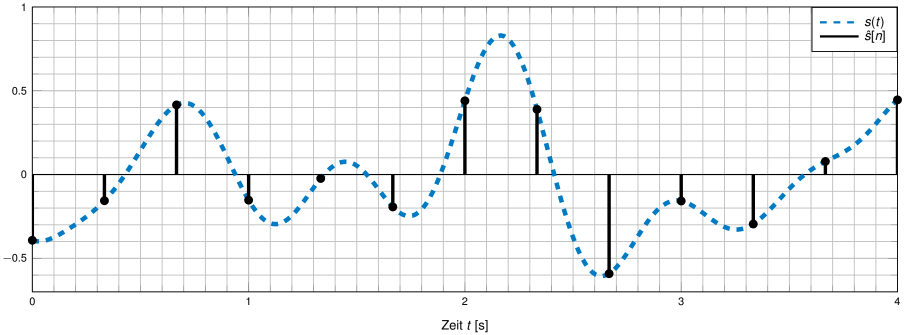


Abbildung: Abtastwerte $\hat{s}[n]$

Rekonstruktion

Mittels der Abtastwerte $\hat{s}[n]$ ist es möglich, das ursprüngliche Signal $s(t)$ zu rekonstruieren:

$$s(t) \approx \sum_{n=-\infty}^{\infty} \hat{s}[n] \cdot \operatorname{sinc}\left(\frac{t - nT_a}{T_a}\right).$$

- Abtastwerte sind Stützstellen und
- dienen als Gewichte für eine passende Ansatzfunktion (trigonometrische Interpolation, vgl. Polynominterpolation → Numerisches Programmieren).

Abbildung: Jeder Abtastwert $\hat{s}[n]$ dient als Gewicht für die Ansatzfunktion zur Rekonstruktion.

Mittels der Abtastwerte $\hat{s}[n]$ ist es möglich, das ursprüngliche Signal $s(t)$ zu rekonstruieren:

$$s(t) \approx \sum_{n=-\infty}^{\infty} \hat{s}[n] \cdot \operatorname{sinc}\left(\frac{t - nT_a}{T_a}\right).$$

- Abtastwerte sind Stützstellen und
- dienen als Gewichte für eine passende Ansatzfunktion (trigonometrische Interpolation, vgl. Polynominterpolation $\rightarrow$ Numerisches Programmieren).

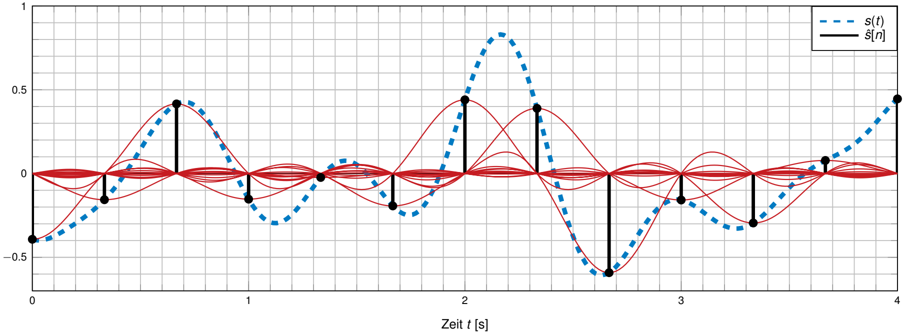


Abbildung: Jeder Abtastwert $\hat{s}[n]$ dient als Gewicht für die Ansatzfunktion zur Rekonstruktion.

Mittels der Abtastwerte $\hat{s}[n]$ ist es möglich, das ursprüngliche Signal $s(t)$ zu rekonstruieren:

$$s(t) \approx \sum_{n=-\infty}^{\infty} \hat{s}[n] \cdot \operatorname{sinc}\left(\frac{t - nT_a}{T_a}\right).$$

- Abtastwerte sind Stützstellen und
- dienen als Gewichte für eine passende Ansatzfunktion (trigonometrische Interpolation, vgl. Polynominterpolation → Numerisches Programmieren).

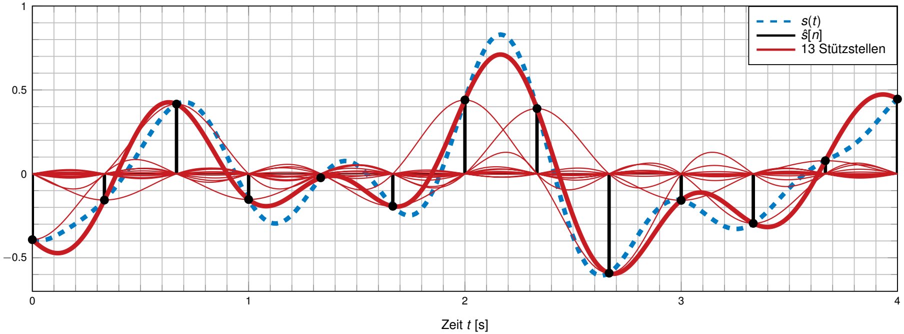


Abbildung: Die Summe der gewichteten Ansatzfunktionen nähert sich dem ursprünglichen Signal in Abhängigkeit der Anzahl der Summenglieder.

Rekonstruktion

Mittels der Abtastwerte $\hat{s}[n]$ ist es möglich, das ursprüngliche Signal $s(t)$ zu rekonstruieren:

$$s(t) \approx \sum_{n=-\infty}^{\infty} \hat{s}[n] \cdot \operatorname{sinc}\left(\frac{t - nT_a}{T_a}\right).$$

- Abtastwerte sind Stützstellen und
- dienen als Gewichte für eine passende Ansatzfunktion (trigonometrische Interpolation, vgl. Polynominterpolation → Numerisches Programmieren).

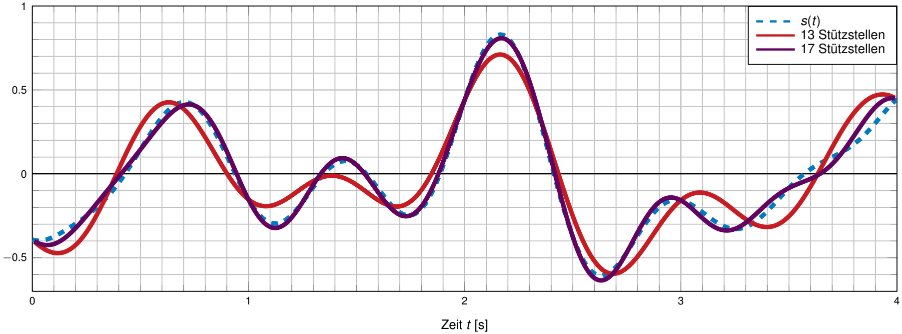


Abbildung: Die Summe der gewichteten Ansatzfunktionen nähert sich dem ursprünglichen Signal in Abhängigkeit der Anzahl der Summenglieder.

Rekonstruktion

Wann ist eine verlustfreie Rekonstruktion möglich?

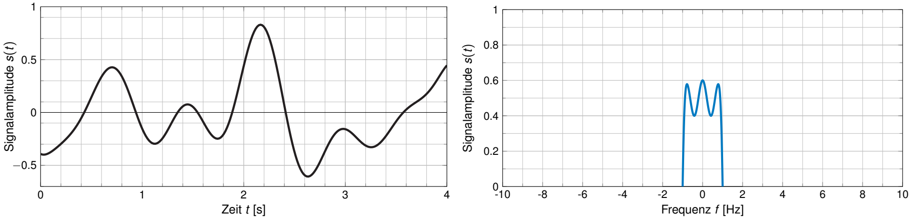
- Die Multiplikation im Zeitbereich entspricht einer Faltung im Frequenzbereich:

$$s(t) \cdot \delta[t - nT] \quad \circ\!\!-\!\!\bullet \quad \frac{1}{T} S(f) * \delta[f - n/T].$$

- Diese Faltung mit Einheitsimpulsen entspricht einer Verschiebung von $S(f)$ entlang der Abszisse.

Folglich entspricht die Abtastung des Signals $s(t)$ in Abständen T_a der periodischen Wiederholung seines Spektrums $S(f)$ in Abständen von $f_a = {}^1/_{T_a}$.

Beispiel: Abtastung eines auf B bandbegrenzten Signals $s(t)$ mit der Abtastfrequenz $f_a = 3B$:



(a) Ursprüngliches Signal $s(t)$

(b) Zugehöriges Spektrum $S(f)$ (schematisch)

Rekonstruktion

Wann ist eine verlustfreie Rekonstruktion möglich?

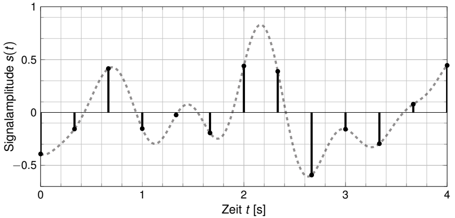
- Die Multiplikation im Zeitbereich entspricht einer Faltung im Frequenzbereich:

$$s(t) \cdot \delta[t - nT] \quad \circ\!\!-\!\!\bullet \quad \frac{1}{T} S(f) * \delta[f - n/T].$$

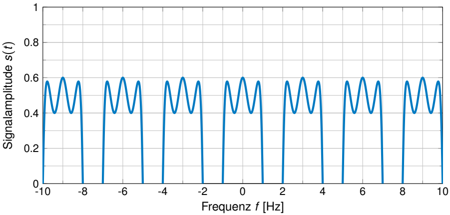
- Diese Faltung mit Einheitsimpulsen entspricht einer Verschiebung von $S(f)$ entlang der Abszisse.

Folglich entspricht die Abtastung des Signals $s(t)$ in Abständen T_a der periodischen Wiederholung seines Spektrums $S(f)$ in Abständen von $f_a = {}^1/_{T_a}$.

Beispiel: Abtastung eines auf B bandbegrenzten Signals $s(t)$ mit der Abtastfrequenz $f_a = 3B$:



(a) Abgetastetes Signal $\hat{s}[n]$



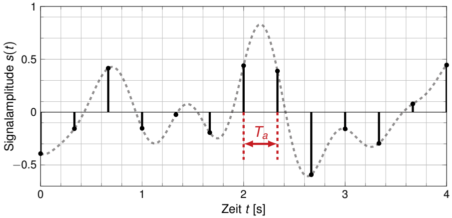
(b) Zugehöriges Spektrum $\hat{S}(f)$ (schematisch)

Abtasttheorem von Shannon und Nyquist

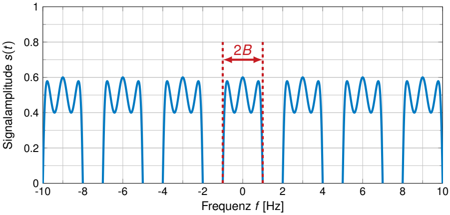
Ein auf $|f| \leq B$ bandbegrenztes Signal $s(t)$ ist vollständig durch äquidistante Abtastwerte $\hat{s}[n]$ beschrieben, sofern diese nicht weiter als $T_a \leq {}^1/_{2B}$ auseinander liegen. Die Abtastfrequenz, welche eine vollständige Signalrekonstruktion erlaubt, ist folglich durch

$$f_a > 2B$$

nach unten beschränkt.



(a) Abgetastetes Signal $\hat{s}[n]$



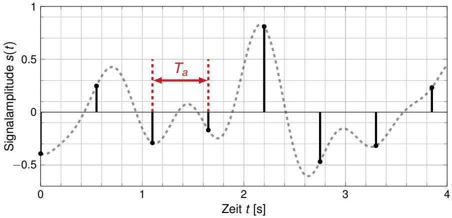
(b) Zugehöriges Spektrum $\hat{S}(f)$ (schematisch)

Abtasttheorem von Shannon und Nyquist

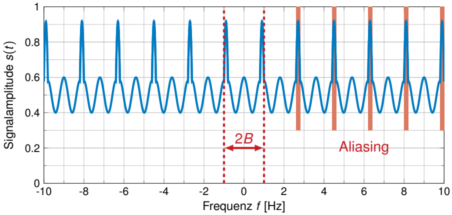
Ein auf $|f| \leq B$ bandbegrenztes Signal $s(t)$ ist vollständig durch äquidistante Abtastwerte $\hat{s}[n]$ beschrieben, sofern diese nicht weiter als $T_a \leq {}^{1}\!/_{2B}$ auseinander liegen. Die Abtastfrequenz, welche eine vollständige Signalrekonstruktion erlaubt, ist folglich durch

$$f_a > 2B$$

nach unten beschränkt.



(a) Abgetastetes Signal $\hat{s}[n]$



(b) Zugehöriges Spektrum $\hat{S}(f)$ (schematisch)

- Wählt man $f_a < 2B$, so überlappen sich die periodischen Wiederholungen des Spektrums
- Diesen Effekt bezeichnet man als Aliasing
- Eine verlustfreie Rekonstruktion ist in diesem Fall nicht möglich

Die Abtastwerte $\hat{s}[n] \in \mathbb{R}$ sind noch kontinuierlich im Wertebereich und können i. A. nicht exakt gespeichert werden.

Lösung: Quantisierung

- Die Unterscheidung von $M = 2^N$ Signalstufen erfordert Codewörter von N bit
- Jeder Signalstufe wird dabei ein bestimmtes Codewort zugeordnet
- Die Signalstufen werden im Quantisierungsintervall $I_Q = [a, b]$ „sinnvoll“ verteilt
- Was ist „sinnvoll“?

Quantisierung

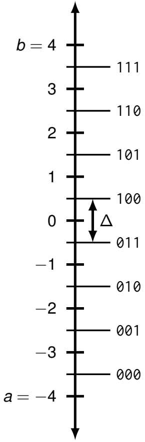
Die Abtastwerte $\hat{s}[n] \in \mathbb{R}$ sind noch kontinuierlich im Wertebereich und können i. A. nicht exakt gespeichert werden.

Lösung: Quantisierung

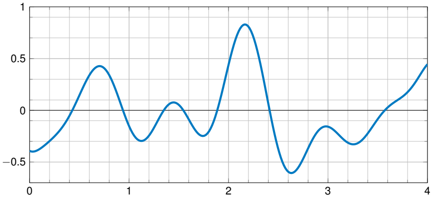
- Die Unterscheidung von $M = 2^N$ Signalstufen erfordert Codewörter von N bit
- Jeder Signalstufe wird dabei ein bestimmtes Codewort zugeordnet
- Die Signalstufen werden im Quantisierungsintervall $I_Q = [a, b]$ „sinnvoll" verteilt
- Was ist „sinnvoll"?

Beispiel: Lineare Quantisierung mit mathematischem Runden

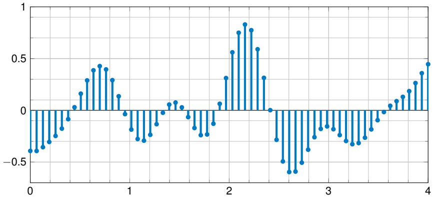
- Optimal, wenn alle Werte innerhalb I_Q mit gleicher Wahrscheinlichkeit auftreten
- Stufenbreite $\Delta = \dfrac{b-a}{M}$
- Innerhalb I_Q beträgt der maximale Quantisierungsfehler $q_{max} = \Delta/2$
- Signalwerte außerhalb I_Q werden auf die größte bzw. kleinste Signalstufe abgebildet $\Rightarrow$ Außerhalb I_Q ist der Quantisierungsfehler unbeschränkt



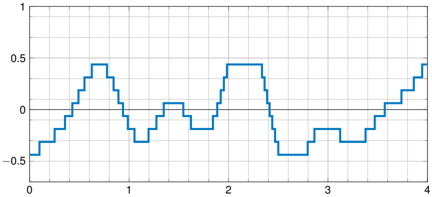
Signaltypen (Übersicht)



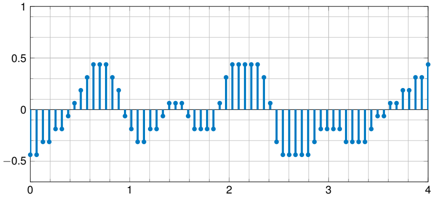
(a) Analog $s(t)$



(b) Zeitdiskret und wertkontinuierlich $\hat{s}[n]$



(c) Zeitkontinuierlich und wertdiskret $\bar{s}(t)$



(d) Digital $s[n]$

Beispiel:

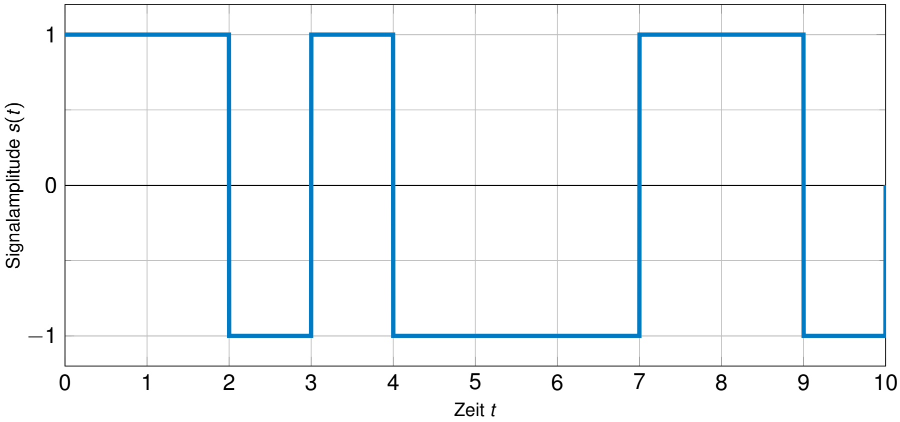


Abbildung: Idealisiertes Sendesignal

Beispiel:

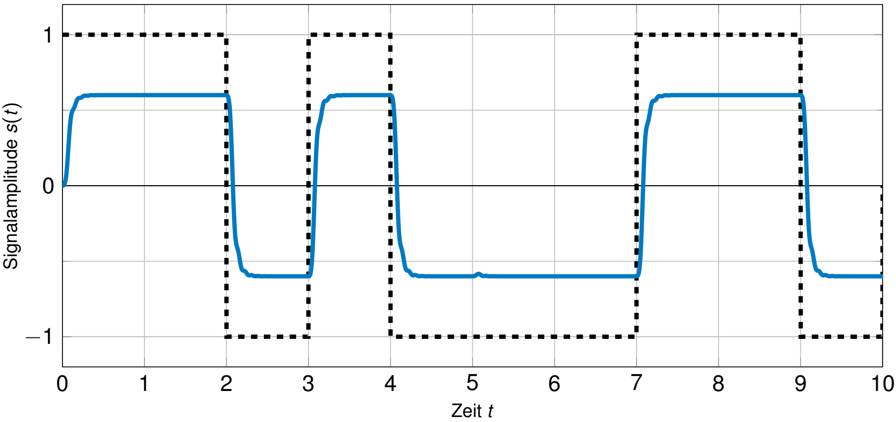


Abbildung: Sendesignal nach Dämpfung und Tiefpasseinflüssen durch den Kanal

Beispiel:

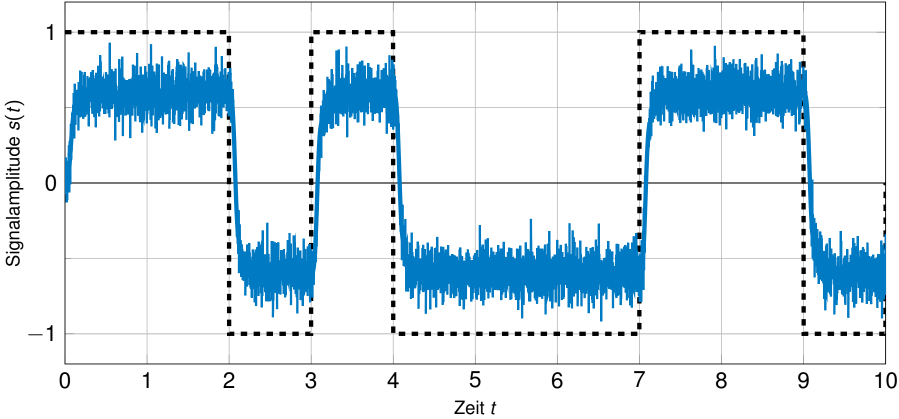


Abbildung: Sendesignal nach Dämpfung und Tiefpasseinflüssen durch den Kanal sowie mit AWGN

Rauschfreier, *M*-ärer Kanal

Angenommen es können nicht nur zwei sondern $M = 2^N$ unterscheidbare Symbole übertragen werden. Wie ändert sich die erzielbare Datenrate?

Wir erinnern uns an Quantisierung und Entropie:

- Mit einer Wortbreite von N bit lassen sich $M = 2^N$ diskrete Signalstufen darstellen.
- Emittiert eine Quelle alle Zeichen (Signalstufen) mit der gleichen Wahrscheinlichkeit, so ist die Entropie (und damit die mittlere Information) der Quelle maximal.

Folglich erhalten wir für die Übertragungsrate

- über einen Kanal der Bandbreite B
- die maximale Rate $R_{\text{max}} = 2B \log_2(M)$ bit.

Hartleys Gesetz

Auf einem Kanal der Bandbreite B mit M unterscheidbaren Signalstufen ist die Kanalkapazität durch

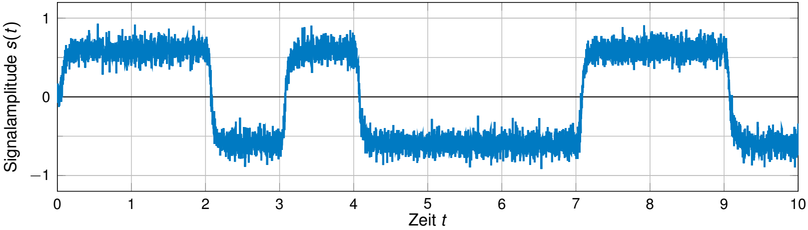
$$C_H = 2B \log_2(M) \text{ bit}$$

nach oben begrenzt.

Interessant: Wenn wir beliebig viele Signalstufen voneinander unterscheiden könnten, wäre die erzielbare Datenrate unbegrenzt! Wo ist das Problem?

Rauschen

- Rauschen macht es schwer, Signalstufen auseinanderzuhalten
- Je feiner die Signalstufen gewählt werden, desto schwieriger wird dies



Maß für die Stärke des Rauschens:

$$\text{SNR} = \frac{\text{Signalleistung}}{\text{Rauschleistung}} = \frac{P_S}{P_N}$$

Das Signal to Noise Ratio (SNR) wird in der Einheit dB angegeben:

$$\text{SNR}\,\text{dB} = 10 \cdot \log_{10}(\text{SNR})$$

Beispiel: $P_S = 1\,\text{mW}$, $P_N = 0.5\,\text{mW}$

$$\text{SNR} = 10 \cdot \log_{10}\left(\frac{1}{0.5}\right)\,\text{dB} \approx 3.0\,\text{dB}$$

Shannon-Hartley-Theorem

Auf einem Kanal der Bandbreite B mit additiven weißen Rauschen mit Rauschleistung P_N und Signalleistung P_S beträgt die obere Schranke für die erreichbare Datenrate

$$C_S = B \log_2\left(1 + \frac{P_S}{P_N}\right) \text{ bit.}$$

Herleitung des Theorems: siehe Shannons Veröffentlichung Communication in the Presence of Noise von 1949 [1].

Vergleich mit Hartleys Gesetz (aufschlussreich!):

$$C_H = 2B \log_2(M) = 2B \log_2\left(\frac{b-a}{\Delta}\right).$$

- Die Intervallgrenzen a, b beziehen sich hier auf das unquantisierte Signal
- Mit $\alpha = a + \Delta/2$ und $\beta = b - \Delta/2$ als minimale bzw. maximale quantisierte Signalamplitude erhalten wir

$$C_H = 2B \log_2\left(\frac{\beta - \alpha + \Delta}{\Delta}\right) = B \log_2\left(\left(1 + \frac{\beta - \alpha}{\Delta}\right)^2\right). \tag{3}$$

Wird (3) ausmultipliziert, kommt aber etwas anderes raus!

- C_S **berücksichtigt nur additives Rauschen des Kanals, aber keine Quantisierungsfehler.**
- C_H **berücksichtigt nur die Signalstufen und damit die Quantisierung („Quantisierungsrauschen"), aber keine Kanaleinflüsse.**
- Der fehlende gemischte Term, wenn man (3) ausmultipliziert und mit C_S vergleicht, liegt in der Unabhängigkeitsannahme des Nutzsignals und des Rauschens begründet ($\mathrm{E}[x\eta] = \mathrm{E}[x]\mathrm{E}[\eta]$).
  Der Quantisierungsfehler ist natürlich nicht unabhängig vom Eingangssignal – aus diesem Grund lässt sich (3) nicht ohne Näherung in die selbe Form wie C_S bringen.

Zusammenfassung

Die Kanalkapazität C ist durch zwei Faktoren beschränkt:

- **Die Anzahl M der unterscheidbaren Symbole**
  Selbst ein rauschfreier Kanal hilft nichts, wenn wir nur zwei Symbole nutzen (können).
- **Signal-to-Noise Ratio (SNR)**
  Ist das Signal-zu-Rausch-Verhältnis SNR zu gering, muss ggf. der Abstand Δ zwischen den Signalstufen erhöht und damit die Anzahl unterscheidbarer Symbole verringert werden, um eine zuverlässige Unterscheidung gewährleisten zu können.

Für die tatsächliche Kanalkapazität C gilt also folgende obere Schranke:

$$C < \min\{C_H, C_S\} = \min\left\{2B\log_2(M), B\log_2\left(1 + \mathrm{SNR}\right)\right\} \text{ bit.}$$

Anmerkungen:

- Das ist nur ein Modell – mit stark vereinfachenden Annahmen.
- Wie man einen Kanalcode mit genau der richtigen Menge Redundanz konstruieren kann, so dass C maximiert wird, ist ein offenes Problem der Informationstheorie. (← Challenge!)
- Wir sprechen hier von Datenraten im informationstheoretischen Sinn, d. h. die zu übertragenden Daten liegen redundanzfrei vor. Dies ist in praktischen Systemen nie gewährleistet:
  - Nutzdaten werden vor dem Senden nicht zwangsläufig (und niemals optimal) komprimiert
  - Zusätzlich zu den Nutzdaten werden Kontrollinformationen (Header) benötigt (→ später)

⇒ Die tatsächlich erzielbare Netto-Datenrate liegt unterhalb der informationstheoretischen Schranke.

Übertragungsrate

$$r = \frac{N}{T}$$

Ein Zitat

“Never underestimate the bandwidth of a station wagon full of tapes hurtling down the highway.”

– Andrew S. Tanenbaum