

Überblick zu Kapitel 3 des Buches „Digitalsignalübertragung“

Im Mittelpunkt von Kapitel 3 stehen die **Impulsinterferenzen**, die beispielsweise durch Verzerrungen des Übertragungskanals entstehen oder mit einer von der Nyquistbedingung abweichenden Realisierung des Empfangsfilters zusammenhängen. In den Kapiteln 3.5 bis 3.8 werden einige **Entzerrungsverfahren** beschrieben, mit denen die Systemdegradation durch Impulsinterferenzen abgemildert werden kann.

Im Einzelnen werden behandelt:

- die Ursachen und Auswirkungen von *Impulsinterferenzen*,
- das *Augendiagramm* als geeignetes Hilfsmittel zur Beschreibung von Impulsinterferenzen,
- die *Fehlerwahrscheinlichkeitsberechnung* unter Berücksichtigung von Kanalverzerrungen,
- der Einfluss von Impulsinterferenzen bei *mehrstufiger und/oder codierter Übertragung*,
- der *optimale Nyquistentzerrer* als Beispiel für lineare Kanalverzerrung,
- die *Entscheidungsrückkopplung* (DFE) – eine effektive nichtlineare Detektorrealisierung,
- der *Korrelationsempfänger* als Beispiel für Maximum-Likelihood- bzw. MAP-Detektoren,
- der *Viterbi-Empfänger*, ein aufwandsreduzierter MAP-Entscheidungsalgorithmus.

Die Beschreibung erfolgt durchgehend im Basisband. Die Ergebnisse lassen sich jedoch problemlos auch auf die im Kapitel 1.5 behandelten Trägerfrequenzsysteme anwenden.

Geeignete Literatur: [BBC87] – [Hän97] – [Han04] – [Hay88] – [Joh92] – [Kam04] – [Lin05] – [Lük04] – [Pro01] – [PS04] – [Söd00] – [Söd01] – [ST85] – [TS87]

Die grundlegende Theorie wird auf 67 Bildschirmseiten dargelegt. Außerdem beinhaltet dieses Kapitel 13 Aufgaben und acht Zusatzaufgaben mit insgesamt 103 Teilaufgaben sowie 85 Grafiken. Einige der fünf **Lernvideos** (LV) und zehn **Interaktionsmodule** (IM) sind nachfolgend aufgeführt:

- **Anmerkungen zur Übertragungsfunktion** (LV, Kapitel 3.1, Dauer 9:08)
- **Lineare und nichtlineare Verzerrungen** (LV, Kapitel 3.1, 3-tlg: Dauer 16:30)
- **Dämpfung von Kupferkabeln** (IM, Kapitel 3.1)
- **Zeitverhalten von Kupferkabeln** (IM, Kapitel 3.1)
- **Komplementäre Gaußsche Fehlerfunktion** (IM, Kapitel 3.2)
- **Augendiagramm und Augenöffnung** (IM, Kapitel 3.2 und 3.4)
- **Lineare Nyquistentzerrung** (IM, Kapitel 3.5)
- **Entscheidungsrückkopplung – DFE** (IM, Kapitel 3.6)
- **Viterbi-Empfänger für einen Vorläufer** (IM zu Kapitel 3.8)

Weitere Informationen zum Thema sowie Aufgaben, Simulationen und Programmierübungen finden Sie im Versuch „3: Impulsinterferenzen und Entzerrung“ des Software-Praktikums *Simulation digitaler Übertragungssysteme* [Söd01] von Prof. Günter Söder. Dieses Praktikum basiert auf vier der 24 DOS-Programme des Lehrsoftwarepakets *LNTsim*.

Lehrsoftwarepaket LNTsim (Zip-Version, mehr als 50 MB)

Versuchsanleitung „Impulsinterferenzen und Entzerrung“ (PDF-Version, ca. 3.7 MB)

Definition des Begriffs „Impulsinterferenz“ (1)

Für die beiden ersten Kapitel dieses Buches wurde vorausgesetzt, dass der Detektionsgrundimpuls $g_d(t)$

- entweder auf den Zeitbereich $|t| \leq T$ begrenzt ist, oder
- äquidistante Nulldurchgänge im Symbolabstand T aufweist.

Bezeichnen wir die Abtastwerte von $g_d(t)$ bei Vielfachen der Symboldauer T (Abstand der Impulse) als die **Detektionsgrundimpulswerte**, so wurde in den Kapiteln 1 und 2 stillschweigend vorausgesetzt:

$$g_\nu = g_d(\nu T) = \begin{cases} g_0 & \text{für } \nu = 0, \\ 0 & \text{für } \nu \neq 0. \end{cases}$$

Als Konsequenz dieser Annahme hat sich daraus ergeben, dass der Nutzanteil (Index „S“)

$$d_S(t) = \sum_{\nu=-\infty}^{+\infty} a_\nu \cdot g_d(t - \nu \cdot T) \quad \text{mit } a_\nu \in \{-1, +1\}$$

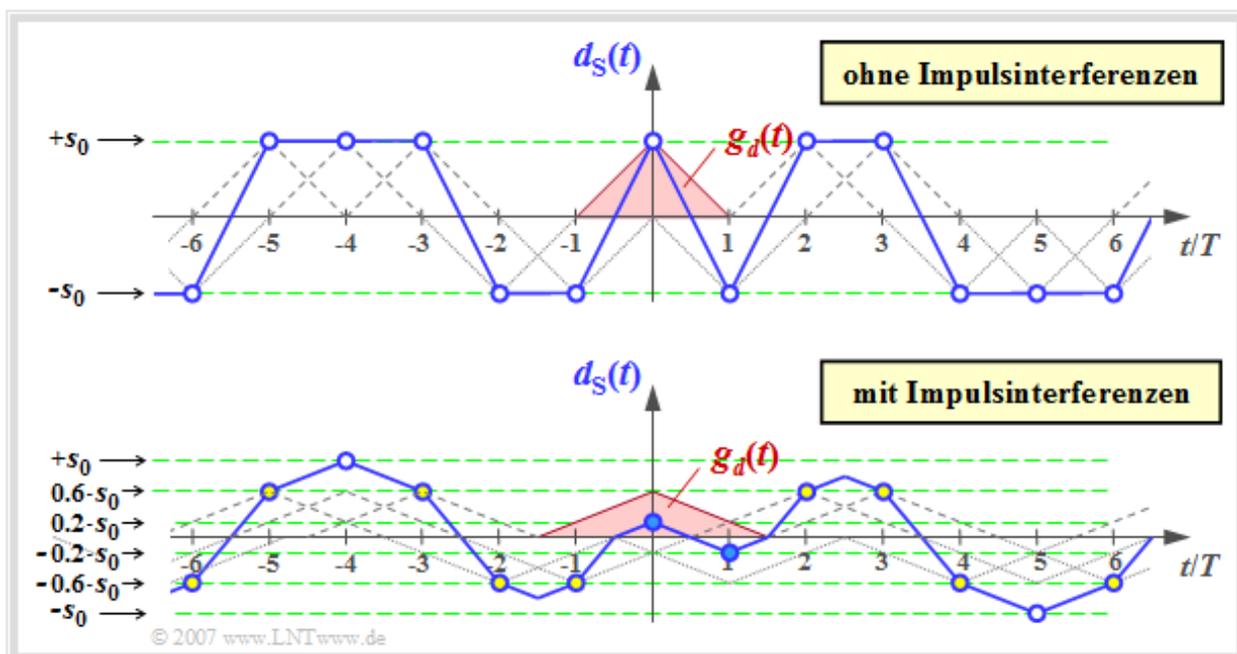
des Detektionssignals zu den Zeitpunkten νT nur zwei verschiedene Werte annehmen kann, nämlich $\pm g_0$.

Die obere der beiden Grafiken zeigt $d_S(t)$ für diesen impulsinterferenzfreien Fall mit $g_0 = s_0$ und $g_{\nu \neq 0} = 0$.

Darunter gezeichnet ist der Signalverlauf für die Detektionsgrundimpulswerte

$$g_0 = 0.6 \cdot s_0, \quad g_{-1} = g_1 = 0.2 \cdot s_0, \quad g_\nu = 0 \quad \text{für } |\nu| \geq 2,$$

die Impulsinterferenzen hervorrufen.



In beiden Bildern ist der (jeweils dreieckförmige) Detektionsgrundimpuls $g_d(t)$ rot eingezeichnet. Die Detektionszeitpunkte νT sind jeweils durch blaue Kreise markiert. Die Bildbeschreibung wird auf der nächsten Seite fortgesetzt.

Definition des Begriffs „Impulsinterferenz“ (2)

Man erkennt aus dem unteren Signalverlauf auf der **vorherigen Seite**:

- Der Detektionsgrundimpuls $g_d(t)$ ist nun im Bereich $|t| \leq 1.5T$ von Null verschieden und erfüllt somit nicht mehr die Nyquist-Bedingung der Impulsinterferenzfreiheit.
- Dies hat zur Folge, dass zu den (mit Kreisen markierten) Detektionszeitpunkten nicht nur zwei Werte ($\pm s_0$) möglich sind wie im oberen Bild. Vielmehr gilt für die Detektionsnutzabtastwerte:

$$d_S(\nu \cdot T) \in \{\pm s_0, \pm 0.6s_0, \pm 0.2s_0\}.$$

- Die Abtastwerte, die aufgrund ungünstiger Nachbarimpulse nahe an der Schwelle liegen, werden durch das AWGN-Rauschen (mit Rauscheffektivwert σ_d) häufiger verfälscht als die weiter außen liegenden Abtastwerte.
- Beispielhaft werden mit $\sigma_d/s_0 = 0.2$ die blau ausgefüllten Punkte nahe der Schwelle mit großer Wahrscheinlichkeit $p_S = Q(1) \approx 16\%$ verfälscht und die äußeren Punkte (mit weißem Kern) nur mit $p_S = Q(5) \approx 3 \cdot 10^{-7}$. Die Fehlerwahrscheinlichkeit der gelb gefüllten Punkte (alle im Abstand $0.6 \cdot s_0$ von der Null-Linie) liegt dazwischen:

$$p_S = Q(3) \approx 0.13\%.$$

Bisher wurden die Auswirkungen von Impulsinterferenzen möglichst anschaulich dargelegt. Es fehlt noch eine exakte Begriffsbestimmung.

Definition: Unter **Impulsinterferenz** (englisch: *Intersymbol Interference*, ISI) versteht man die Beeinträchtigung einer Symbolentscheidung aufgrund einer Impulsverbreiterung (Zeitdispersion) und damit verbunden eine Abhängigkeit der Fehlerwahrscheinlichkeit von den Nachbarsymbolen.

In anderen Worten:

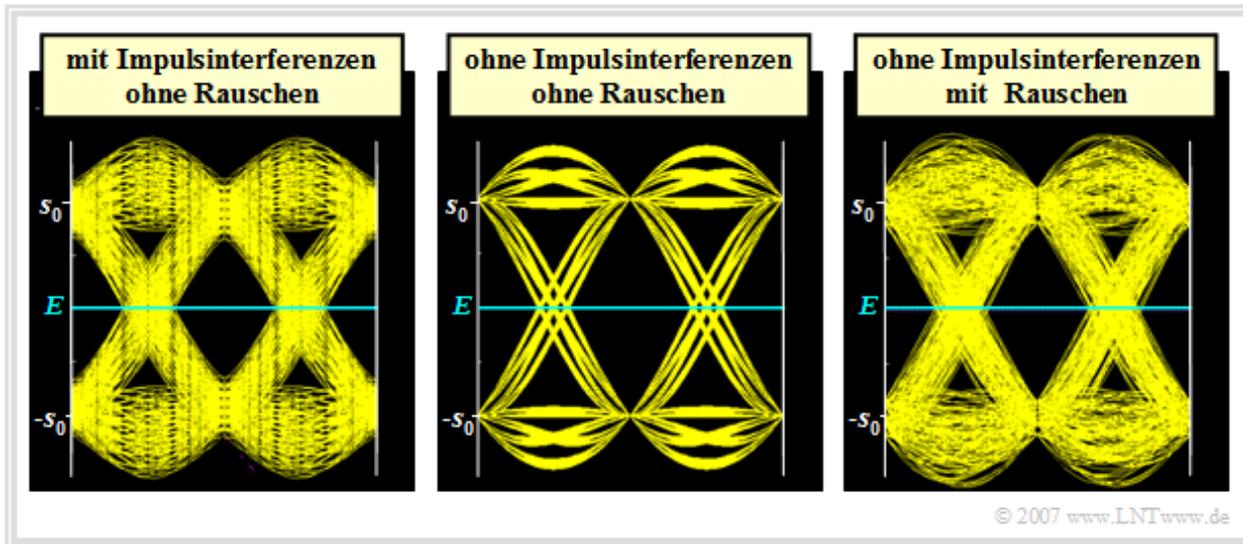
- Durch abfallende Flanken vorangegangener Impulse („Nachläufer“) und ansteigende Flanken folgender Impulse („Vorläufer“) wird der momentan anliegende Detektionsabtastwert verändert.
- Dadurch kann die Wahrscheinlichkeit einer Fehlentscheidung für das aktuelle Symbol vergrößert oder verkleinert werden, je nachdem, ob der Abstand zur Schwelle kleiner oder größer wird.
- Im statistischen Mittel – also bei Betrachtung einer (unendlich) langen Symbolfolge – führt dies stets zu einer (beträchtlichen) Erhöhung der (mittleren) Symbolfehlerwahrscheinlichkeit p_S .

Mögliche Ursachen für Impulsinterferenzen

Die nachfolgende Grafik zeigt das **Augendiagramm** für ein

- impulsinterferenzbehaftetes System ohne Rauschen (links),
- ein impulsinterferenzfreies System ohne Rauschen (Mitte),
- das gleiche impulsinterferenzfreie System mit Rauschen (rechts).

Auf die Definition, Bedeutung und Berechnung des Augendiagramms wird im **Kapitel 3.2** noch ausführlich eingegangen. Die Bilder wurden mit dem Programm „bas“ erzeugt. Hinweis zum Download dieses Programms aus *LNTsim* finden Sie am Beginn dieses Kapitels unter **Kapitelüberblick**.

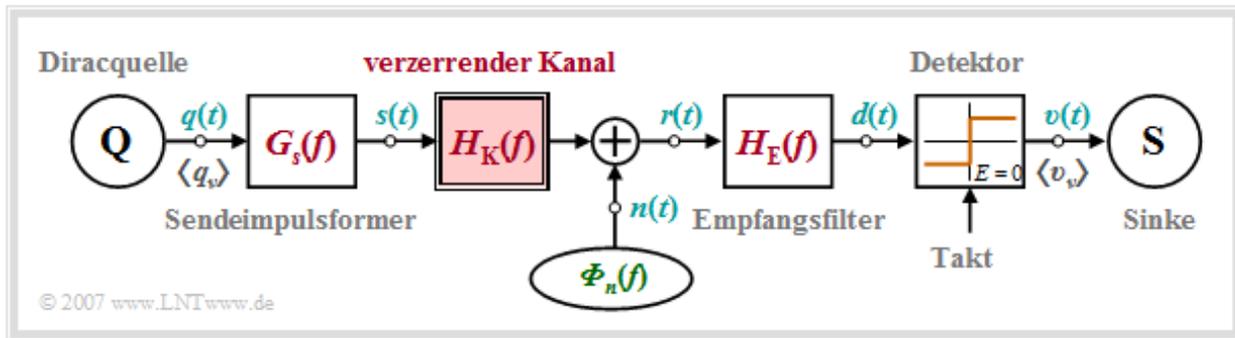


Diese Bilder können wie folgt interpretiert werden:

- Das mittlere Diagramm stammt von einem Nyquistsystem mit Cosinus-Rolloff-Charakteristik (Rolloff-Faktor $r = 0.5$). Es treten somit keine Impulsinterferenzen auf.
- Auch das rechte Augendiagramm stammt von einem impulsinterferenzfreien System (genauer gesagt: vom gleichen System wie die mittlere Grafik), obwohl hier $d(vT) = \pm s_0$ nicht zutrifft. Die Abweichungen von den Sollwerten $\pm s_0$ sind hier auf das AWGN-Rauschen zurückzuführen.
- Aus diesem letzten Punkt folgt die wichtige Erkenntnis: Die Frage, ob ein impulsinterferenzfreies oder ein impulsinterferenzbehaftetes System vorliegt, kann nur anhand des Detektionssignals (bzw. des Augendiagramms) ohne Rauschen entschieden werden.
- Das linke Diagramm weist auf Impulsinterferenzen hin, da hier kein Rauschen berücksichtigt ist. Ein Grund für diese Impulsinterferenzen könnte sein, dass der Gesamtfrequenzgang von Sender und Empfänger das **erste Nyquistkriterium** aufgrund von Toleranzen nicht exakt erfüllt.
- Impulsinterferenzen entstehen aber auch bei einem Kanal mit frequenzabhängigem Frequenzgang $H_K(f)$, wenn es dem Empfänger nicht gelingt, die Dämpfungs- und Phasenverzerrungen des Kanals vollständig (hundertprozentig) zu kompensieren.
- Letztendlich kommt es auch beim mittleren System zu Impulsinterferenzen, wenn nicht exakt in Augenmitte entschieden wird, sondern zu einem Detektionszeitpunkt $T_D \neq 0$. In diesem Fall müssen dann die Detektionsgrundimpulsweite zu $g_v = g_d(T_D + v \cdot T)$ definiert werden.

Einige Anmerkungen zum Kanalfrequenzgang

Für die weiteren Abschnitte von Kapitel 3 wird meist von folgendem Blockschaltbild ausgegangen.



Der wesentliche Unterschied gegenüber dem **Blockschaltbild zu Kapitel 1** ist der Kanalfrequenzgang $H_K(f)$, der in den Kapiteln 1 und 2 stets zu $H_K(f) = 1$ und damit als ideal angenommen wurde.

Im Folgenden gelte für den **Frequenzgang** und die **Impulsantwort** des Kanals:

$$H_K(f) = \exp \left[-a_{*(Np)} \cdot \sqrt{\frac{f}{R_B/2}} \right] \cdot \exp \left[-j \cdot a_{*(Np)} \cdot \sqrt{\frac{f}{R_B/2}} \right],$$

$$h_K(t) = \frac{a_{*(Np)}}{\sqrt{2\pi^2 \cdot R_B \cdot t^3}} \cdot \exp \left[-\frac{a_{*(Np)}^2}{2\pi \cdot R_B \cdot t} \right].$$

Hierbei gibt $a_{*(Np)}$ die **charakteristische Kabeldämpfung** bei der halben Bitrate in Neper (Np) an:

$$a_{*(Np)} = a_K(f = R_B/2) = 0.1151 \cdot a_{*(dB)}.$$

Der entsprechende dB-Wert ist um den Faktor $1/0.1151 = 8.686$ größer. Bei realisierten Systemen liegt die charakteristische Kabeldämpfung $a_{*(dB)}$ im Bereich zwischen 40 dB und 100 dB. Auf den Zusatz „(Np)“ bzw. „(dB)“ wird im Folgenden meist verzichtet.

Im **Kapitel 4** des Buches „Lineare zeitinvariante Systeme“ wird gezeigt, dass diese Gleichungen die Verhältnisse bei leitungsgebundener Übertragung über Koaxialkabel mit guter Näherung wiedergeben. Bei einer Zweidrahtleitung ist die Abweichung zwischen dieser sehr einfachen, analytisch handhabbaren Formel und den tatsächlichen Gegebenheiten etwas größer.

Eine kurze Zusammenfassung der Herleitungen in Kapitel 4 des Buches „LZI-Systeme“ folgt auf den beiden nächsten Seiten, wobei wir uns zur Vereinfachung auf ein redundanzfreies Binärsystem festlegen. Somit ist die Bitrate R_B gleich dem Kehrwert der Symboldauer T .

Frequenzgang eines Koaxialkabels (1)

Ein **Koaxialkabel** mit dem Kerndurchmesser 2.6 mm, dem Außendurchmesser 9.5 mm und der Länge l hat den folgenden Frequenzgang:

$$H_K(f) = \exp [a_K(f) + j \cdot b_K(f)] = e^{-\alpha_0 \cdot l} \cdot e^{-(\alpha_1 + j \cdot \beta_1) \cdot f \cdot l} \cdot e^{-(\alpha_2 + j \cdot \beta_2) \cdot \sqrt{f} \cdot l},$$

wobei bei diesen Abmessungen – man spricht vom Normalkoaxialkabel – folgende Parameter gelten:

$$\alpha_0 = 0.00162 \frac{\text{Np}}{\text{km}}, \quad \alpha_1 = 0.000435 \frac{\text{Np}}{\text{km} \cdot \text{MHz}}, \quad \alpha_2 = 0.2722 \frac{\text{Np}}{\text{km} \cdot \sqrt{\text{MHz}}},$$

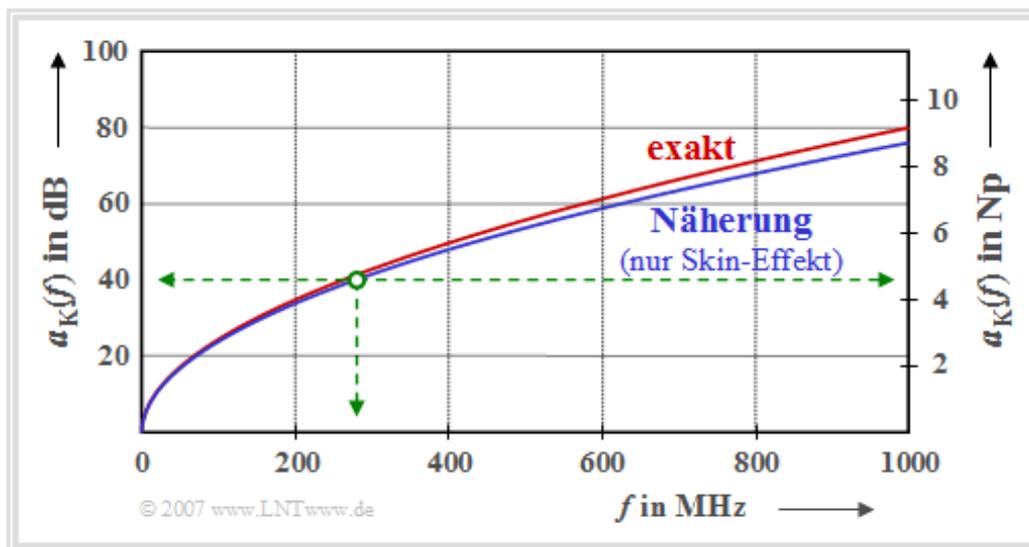
$$\beta_1 = 21.78 \frac{\text{rad}}{\text{km} \cdot \text{MHz}}, \quad \beta_2 = 0.2722 \frac{\text{rad}}{\text{km} \cdot \sqrt{\text{MHz}}}.$$

In obiger Gleichung sind die Dämpfungsparameter in „Np“ einzusetzen, die Phasenparameter in „rad“.

Die Grafik zeigt den Dämpfungsverlauf

$$a_K(f) = \alpha_0 \cdot l + \alpha_1 \cdot f \cdot l + \alpha_2 \cdot \sqrt{f} \cdot l \approx \alpha_2 \cdot \sqrt{f} \cdot l$$

für ein Koaxialkabel von einem Kilometer Länge für Frequenzen bis 1000 MHz. Die Achse ist sowohl in dB (links) als auch in Np (rechts) beschriftet. Ein Neper entspricht etwa 8.7 dB.



Die ausführliche Bildbeschreibung folgt auf der nächsten Seite.

Bereits an dieser Stelle soll darauf hingewiesen werden, dass die obige Näherung

$$a_K(f) \approx \alpha_2 \cdot \sqrt{f} \cdot l$$

nur für Koaxialkabel zulässig ist. Bei diesen können die Koeffizienten $\alpha_0 \approx 0$ und $\alpha_1 \approx 0$ vernachlässigt werden. Für eine symmetrische Zweidrahtleitung sind diese Koeffizienten sehr viel größer und die obige Näherung ist unzulässig.

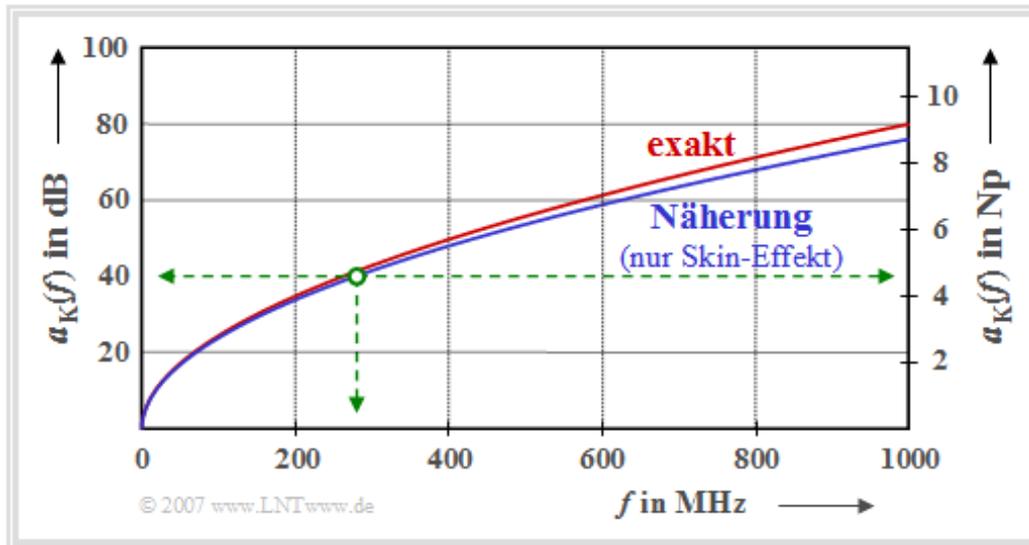
Genauere Informationen finden Sie im **Kapitel 4.3** des Buches „Lineare zeitinvariante Systeme“.

Frequenzgang eines Koaxialkabels (2)

Betrachten wir nochmals die Grafik von eben. Diese zeigt den Dämpfungsverlauf

$$a_K(f) = \alpha_0 \cdot l + \alpha_1 \cdot f \cdot l + \alpha_2 \cdot \sqrt{f} \cdot l \approx \alpha_2 \cdot \sqrt{f} \cdot l$$

für ein Koaxialkabel von einem Kilometer Länge für Frequenzen bis 1000 MHz. Die Achse ist sowohl in dB (links) als auch in Np (rechts) beschriftet. Ein Neper (Np) entspricht etwa 8.7 dB.



Man erkennt aus diesem Diagramm und den obigen Zahlenwerten:

- Der von den Ohmschen Verlusten herrührende erste Term ($\alpha_0 \cdot l$) ist vernachlässigbar. Zudem bewirkt er nur eine frequenzunabhängige Dämpfung und keine Signalverzerrung.
- Der auf die Querverluste zurückzuführende zweite Term ist proportional zur Frequenz und macht sich erst bei sehr hohen Frequenzen bemerkbar; er wird im Folgenden vernachlässigt.
- Die frequenzproportionale Phase $\beta_1 \cdot f \cdot l$ hat nur eine Signalverzögerung um die Laufzeit $\beta_1 \cdot l/2\pi$ zur Folge, jedoch keine Verzerrung. Auch diese Laufzeit wird im Folgenden außer Acht gelassen.
- Mit diesen Vereinfachungen wird somit der Frequenzgang allein durch den **Skinneffekt** bestimmt. Da die Zahlenwerte für α_2 (in Np) und β_2 (in rad) übereinstimmen, gilt somit auch:

$$H_K(f) = e^{-(\alpha_2 + j \cdot \beta_2) \cdot \sqrt{f} \cdot l} \cdot e^{-\alpha_2 \cdot l \cdot \sqrt{j \cdot 2f}}$$

- Häufig wird in der Literatur – und auch in diesem Tutorial – das Dämpfungsmaß bei der halben Bitrate benutzt, das wir als charakteristische Kabeldämpfung (in Neper) bezeichnen:

$$a_* = a_K(f = R_B/2) = a_K(f = \frac{1}{2 \cdot T}) \approx \frac{\alpha_2 \cdot l}{\sqrt{2 \cdot T}}$$

- Bei einem Binärsystem mit der halben Bitrate $R_B/2 = 280$ Mbit/s und $l = 1$ km ergibt sich a_* zu etwa 4.55 Np bzw. 40 dB (grün eingezeichnete Markierungen). Beträgt aber die halbe Bitrate nur 70 Mbit/s, so charakterisiert $a_* = 40$ dB ein Übertragungssystem mit der Kabellänge $l = 2$ km.

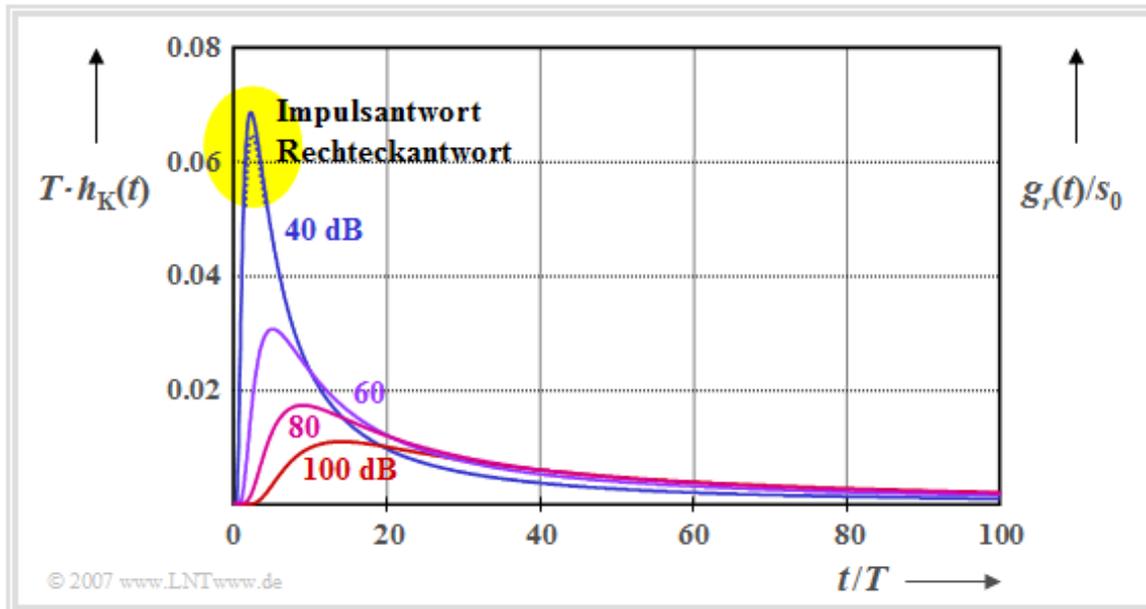
Wir verweisen an dieser Stelle auf das Interaktionsmodul **Dämpfung von Kupferkabeln**.

Impulsantwort eines Koaxialkabels

Betrachten wir nun die **Impulsantwort**, die bei einem Binärsystem ($R_B = 1/T$) wie folgt lautet:

$$h_K(t) = \frac{a_{*(Np)}/T}{\sqrt{2\pi^2 \cdot (t/T)^3}} \cdot \exp \left[-\frac{a_{*(Np)}^2}{2\pi \cdot t/T} \right].$$

Dieser Zeitverlauf ist nachfolgend für $a_* = 40$ dB, 60 dB, 80 dB und 100 dB dargestellt. Beachten Sie wieder die Umrechnung: 1 Np = 8.686 dB.



Man erkennt aus dieser Zeitbereichsdarstellung

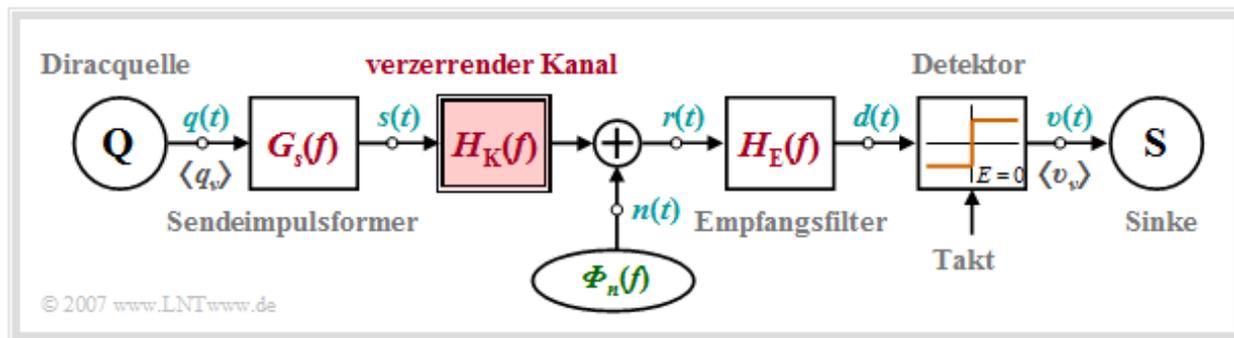
- Bereits mit der relativ kleinen charakteristischen Kabeldämpfung $a_* = 40$ dB erstreckt sich die Impulsantwort über mehr als 100 Symbolauern.
- Je größer a_* gewählt wird, desto breiter und niedriger wird die Impulsantwort. Das Integral über $h_K(t)$ von 0 bis Unendlich ist für alle Kurven gleich, da stets $H_K(f=0) = 1$ gilt.
- Der **Empfangsgrundimpuls** $g_r(t) = g_s(t) * h_K(t)$ ist nahezu formgleich mit $h_K(t)$. Die rechte Ordinatenachse zeigt $g_r(t)/s_0$, wenn $g_s(t)$ ein NRZ-Rechteckimpuls mit Höhe s_0 und Dauer T ist.
- Für $a_* \geq 60$ dB sind $h_K(t)$ und $g_r(t)$ bei geeigneter Normierung innerhalb der Zeichengenauigkeit nicht zu unterscheiden. Für $a_* = 40$ dB erkennt man eine kleine Differenz an der Spitze (gelbe Hinterlegung); $g_r(t)/s_0$ ist hier minimal kleiner als $T \cdot h_K(t)$.
- Mit der charakteristischen Dämpfung 40 dB beträgt die Impulsamplitude am Kabelende weniger als 7% der Eingangsamplitude. Bei 60 dB bzw. 100 dB sinkt dieser Wert auf 3% bzw. 1%.

In der Aufgabe A3.1 wird die hier betrachtete Impulsantwort noch eingehend analysiert. Weiterhin verweisen wir auf das folgende Interaktionsmodul:

Zeitverhalten von Kupferkabeln

Blockschaltbild und Voraussetzungen für Kapitel 3

Betrachten wir nochmals das Blockschaltbild eines Übertragungssystems mit stark verzerrendem Kanal.



Aufgrund des in der Grafik rot hervorgehobenen Kanalfrequenzgangs $H_K(f)$ ergeben sich auch für die anderen Systemkomponenten gewisse Einschränkungen:

- Der Empfangsgrundimpuls $g_r(t) = g_s(t) * h_K(t)$ erstreckt sich über Hunderte von Bits (siehe letzte Seite). Deshalb kann das Empfangsfilter $H_E(f)$ nicht als **Matched-Filter** angesetzt werden, da so die Dauer des Detektionsgrundimpulses $g_d(t)$ gegenüber $g_r(t)$ nochmals etwa verdoppelt würde.
- Vielmehr muss $H_E(f)$ die enormen Dämpfungsverzerrungen (α_2 -Term) und Phasenverzerrungen (β_2 -Term) des koaxialen Kanals $H_K(f)$ kompensieren, wenn wie in den Kapiteln 3.1 bis 3.5 von einem einfachen Schwellenwertentscheider ausgegangen wird.
- Diese lineare Form der Signalverzerrung kann durch aufwändigere Entscheiderstrategien – wie in den Kapiteln 3.6 bis 3.8 (Entscheidungsrückkopplung, Korrelationsempfänger, Viterbi-Empfänger) beschrieben – unterstützt werden. Bei leitungsgebundener Übertragung kann aber aufgrund der sehr starken Verzerrungen auf eine lineare Signalverzerrung nicht vollständig verzichtet werden.
- Das Rauschen $n(t)$ wird weiterhin als additiv, weiß und gaußverteilt (AWGN) angesetzt, was bei einem Koaxialkabel gerechtfertigt ist. Bei einer Zweidrahtleitung ist das Nebensprechen von benachbarten Kupferadern die dominante Störung, wie im Kapitel **ISDN (Integrated Services Digital Network)** im Buch „Beispiele von Nachrichtensystemen“ ausführlich dargelegt wird.

Zunächst betrachten wir in den Kapiteln 3.2 und 3.3 die binäre bipolare redundanzfreie Übertragung, wobei die Bitrate $R_B = 1/T$ beträgt. Dabei wird stets vorausgesetzt:

- Der Sendegrundimpuls $g_s(t)$ ist NRZ-rechteckförmig mit Amplitude s_0 und Dauer T . Somit ist das Sendesignal $s(t)$ zu allen Zeiten gleich $\pm s_0$. Die Spektralfunktion lautet: $G_s(f) = s_0 \cdot T \cdot \text{si}(\pi f T)$.
- Eine Aufteilung entsprechend der Wurzel-Wurzel-Charakteristik macht bei leitungsgebundener Übertragung keinen Sinn. Es würden bereits beim Sender zu starke Impulsinterferenzen auftreten.

Der Einfluss von Impulsinterferenzen bei mehrstufiger und/oder codierter Digitalsignalübertragung wird im Kapitel 3.4 behandelt.

Gaußförmiges Empfangsfilter (1)

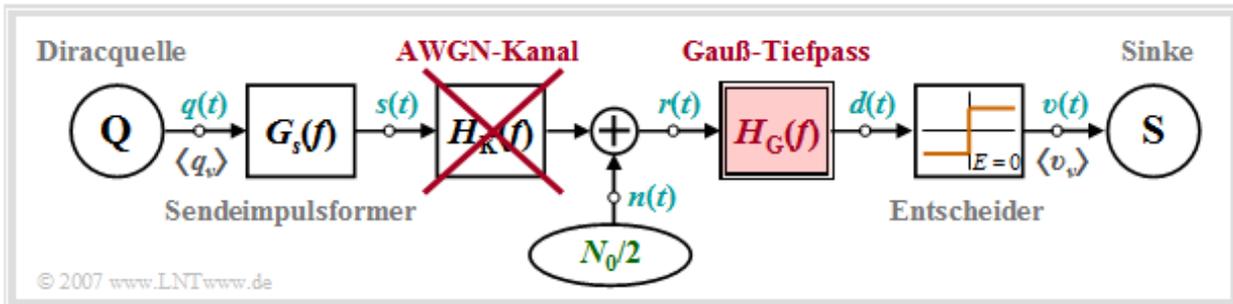
Zur quantitativen Berücksichtigung der **Impulsinterferenzen** wird folgende Konfiguration angenommen:

- Rechteckförmiger NRZ-Sendegrundimpuls $g_s(t)$ mit der Höhe s_0 und der Dauer T ,
- Gaußförmiges Empfangsfilter mit der Grenzfrequenz f_G :

$$H_E(f) = H_G(f) = \exp \left[-\frac{\pi \cdot f^2}{(2f_G)^2} \right] \quad \bullet \text{---} \circ \quad h_E(t) = h_G(t) = \exp \left[-\pi \cdot (2f_G t)^2 \right].$$

- AWGN-Kanal, das heißt, es gilt $H_K(f) = 1$ und $\Phi_n(f) = N_0/2$.

Für das gesamte Kapitel 3.2 wird somit von nachfolgendem **Blockschaltbild** ausgegangen.



Aufgrund der getroffenen Voraussetzungen gilt für den **Detektionsgrundimpuls**:

$$g_d(t) = g_s(t) \star h_G(t) = 2f_G \cdot s_0 \cdot \int_{t-T/2}^{t+T/2} \exp \left[-\pi \cdot (2 \cdot f_G \cdot \tau)^2 \right] d\tau.$$

Die Integration führt zu folgendem Ergebnis:

$$g_d(t) = s_0 \cdot \left[Q \left(2 \cdot \sqrt{2\pi} \cdot f_G \cdot (t - T/2) \right) - Q \left(2 \cdot \sqrt{2\pi} \cdot f_G \cdot (t + T/2) \right) \right],$$

$$g_d(t) = s_0/2 \cdot \left[\operatorname{erfc} \left(2 \cdot \sqrt{\pi} \cdot f_G \cdot (t - T/2) \right) - \operatorname{erfc} \left(2 \cdot \sqrt{\pi} \cdot f_G \cdot (t + T/2) \right) \right].$$

Hierbei sind zwei Varianten der komplementären Gaußschen Fehlerfunktion verwendet, nämlich

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-u^2/2} du, \quad \operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-u^2} du.$$

Das Modul **Komplementäre Gaußsche Fehlerfunktion** liefert die Zahlenwerte von $Q(x)$ und $\operatorname{erfc}(x)$.

Die **Rauschleistung** am Ausgang des gaußförmigen Empfangsfilters $H_G(f)$ ist gleich

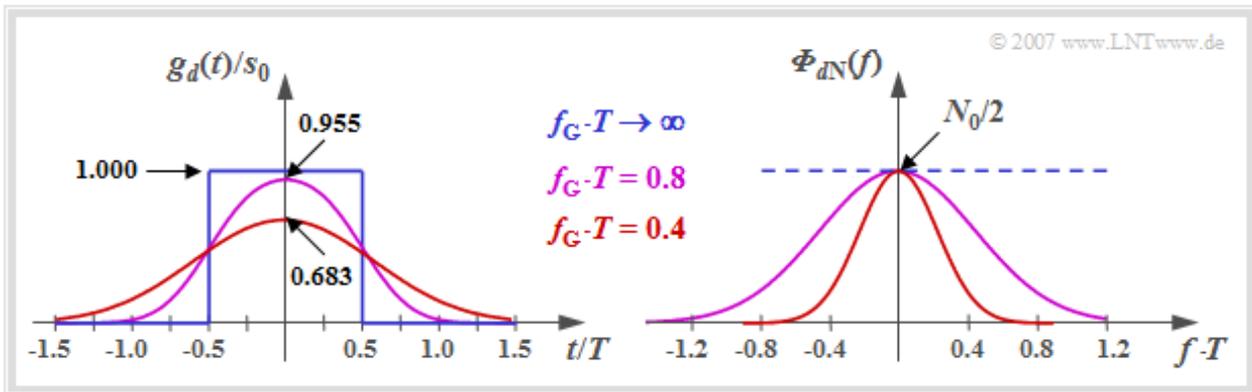
$$\sigma_d^2 = \frac{N_0}{2} \cdot \int_{-\infty}^{+\infty} |H_G(f)|^2 df = \frac{N_0 \cdot f_G}{\sqrt{2}}.$$

Aus diesen beiden Gleichungen erkennt man bereits:

- Je kleiner die Grenzfrequenz f_G des Gauß-Tiefpasses ist, desto kleiner ist der Rauscheffektivwert σ_d und umso besser ist demzufolge das Rauschverhalten.
- Eine kleine Grenzfrequenz führt aber zu einer starken Abweichung des Detektionsgrundimpulses $g_d(t)$ von der Rechteckform und damit zu nicht vernachlässigbaren Impulsinterferenzen.

Gaußförmiges Empfangsfilter (2)

Die nachfolgend linke Grafik zeigt den Detektionsgrundimpuls $g_d(t)$ am Ausgang eines Gaußtieffpasses mit der Grenzfrequenz f_G , wenn am Eingang ein NRZ-Rechteckimpuls (blauer Kurvenverlauf) anliegt.



Man erkennt aus dieser Darstellung:

- Der Gaußtieffpass $H_G(f)$ bewirkt, dass der Detektionsimpuls $g_d(t)$ gegenüber dem Sendesignal $g_s(t)$ verkleinert und verbreitert wird. Man spricht von *Zeitdispersion*.
- Diese Impulsverformung ist umso stärker, je kleiner die Grenzfrequenz f_G ist. Beispielsweise wird mit $f_G \cdot T = 0.4$ (rote Kurve) das Impulsmaximum bereits auf etwa 68% herabgesetzt.
- Im Grenzfall $f_G \cdot T \rightarrow \infty$ hat der Gaußtieffpass keine Wirkung $\Rightarrow g_d(t) = g_s(t)$. Allerdings ist in diesem Fall keinerlei Rauschbegrenzung wirksam, wie aus dem rechten Bild hervorgeht.

Die **Grafik auf der folgenden Seite** zeigt das Detektionssignal $d(t)$ nach dem Gaußtieffpass (vor dem Entscheider) für zwei verschiedene Grenzfrequenzen:

- Der obere Signalverlauf gilt für die (normierte) Grenzfrequenz $f_E \cdot T = 0.8$.
- Für den unteren Signalverlauf ist die Grenzfrequenz nur halb so groß: $f_E \cdot T = 0.4$.

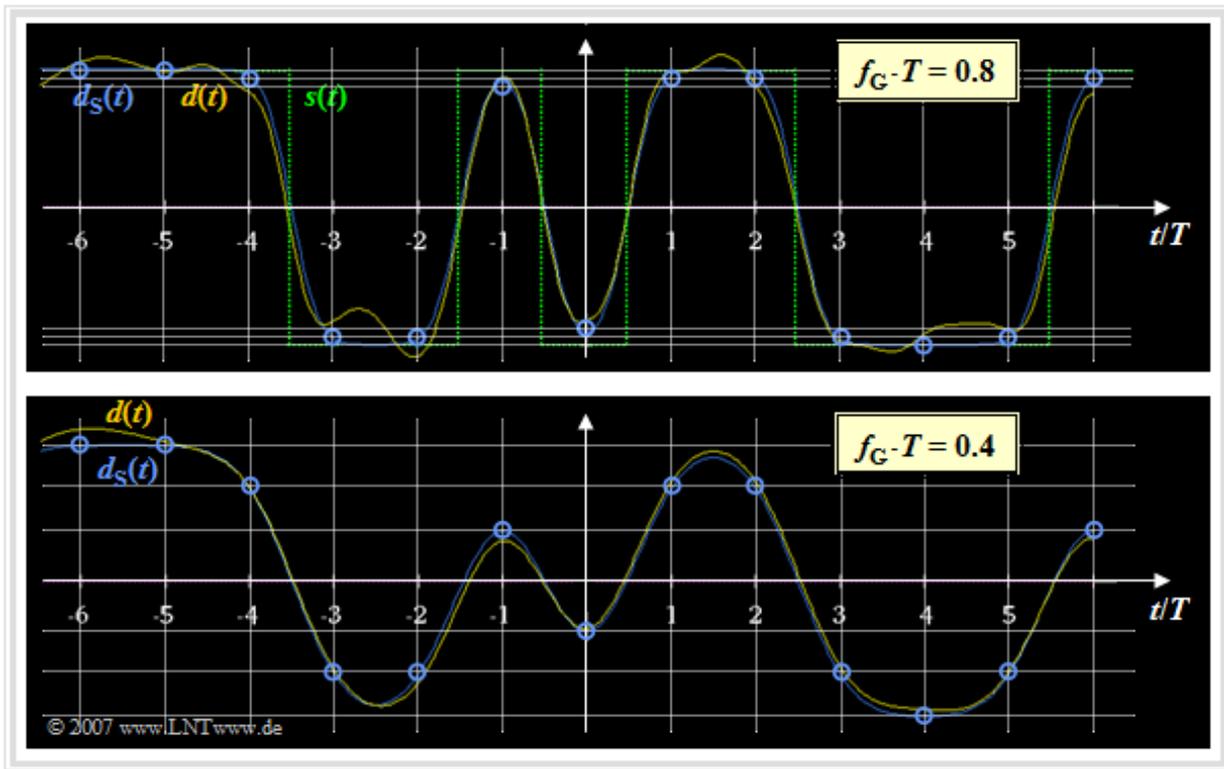
Dargestellt sind in beiden Diagrammen gleichermaßen:

- der Anteil $d_S(t)$ ohne Berücksichtigung des Rauschens (blau),
- das gesamte Detektionssignal $d(t)$ inklusive der Rauschkomponente (gelb),
- das Sendesignal $s(t)$ als Referenzsignal (grün gepunktet).

Leider sind die verschiedenen Signalverläufe dieses Bildschirmabzugs sehr schwer zu erkennen, besonders in der PDF-Version. Die Bildbeschreibung folgt auf der nächsten Seite.

Gaußförmiges Empfangsfilter (3)

Die Grafik zeigt das Detektionssignal $d(t)$ nach dem Gaußtiefpass (also vor dem Entscheider) für zwei verschiedene (normierte) Grenzfrequenzen, nämlich $f_G \cdot T = 0.8$ und $f_G \cdot T = 0.4$.



Durch einen Vergleich dieser Bilder lassen sich folgende Aussagen verifizieren:

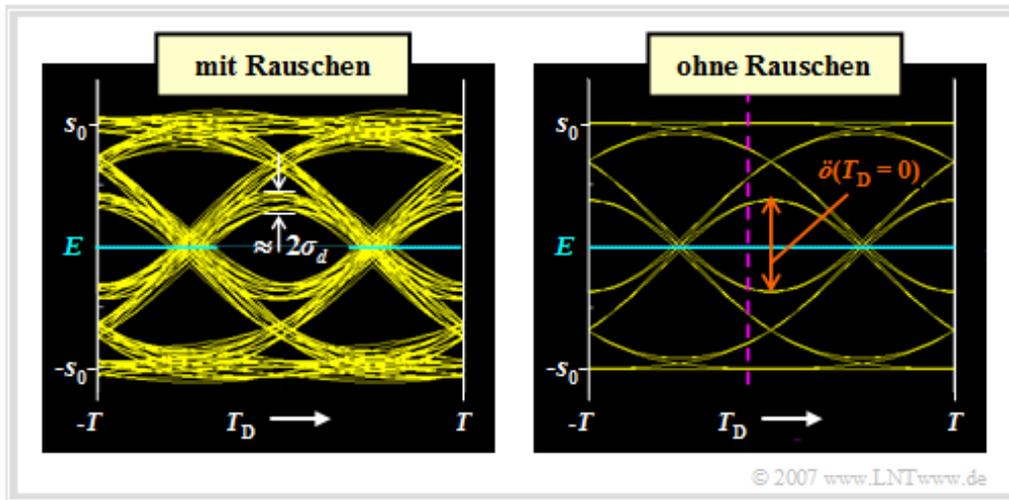
- Mit der Grenzfrequenz $f_G \cdot T = 0.8$ (obere Grafik) ergeben sich zu den Detektionszeitpunkten (bei Vielfachen von T) nur geringfügige Impulsinterferenzen. Durch den Gaußtiefpass werden hier in erster Linie die Ecken des Sendesignals $s(t)$ abgerundet.
- Dagegen sind im unteren Bild ($f_G \cdot T = 0.4$) die Auswirkungen der Impulsinterferenzen deutlich zu erkennen. Zu den Detektionszeitpunkten νT kann das blau dargestellte Detektionsnutzsignal $d_S(t)$ sechs verschiedene Werte annehmen (eingezeichnete Rasterlinien).
- Der Rauschanteil $d_N(t)$ – erkennbar als Differenz zwischen der gelben und der blauen Kurve – ist mit $f_G \cdot T = 0.8$ im statistischen Mittel größer als mit $f_G \cdot T = 0.4$.
- Dieses Ergebnis kann mit der rechten Grafik auf der **letzten Seite** erklärt werden, die das Leistungsdichtespektrum der Rauschkomponente $d_N(t)$ zeigt:

$$\Phi_{dN}(f) = N_0/2 \cdot |H_G(f)|^2 = N_0/2 \cdot \exp \left[-\frac{2\pi f^2}{(2f_G)^2} \right].$$

Das Integral über $\Phi_{dN}(f)$ – also die Rauschleistung σ_d^2 – ist für $f_G \cdot T = 0.8$ (violette Kurve) doppelt so groß als mit der kleineren Grenzfrequenz $f_G \cdot T = 0.4$ (rote Kurve).

Definition und Aussagen des Augendiagramms

Der oben dargelegte Sachverhalt lässt sich auch am Augendiagramm erklären. Wir gehen von einem redundanzfreien binären bipolaren NRZ-Rechtecksignal $s(t)$ und dem Gaußtiefpass mit $f_G \cdot T = 0.4$ aus. Dargestellt sind die Augendiagramme nach dem Gaußtiefpass, links mit Berücksichtigung des Rauschens \Rightarrow Signal $d(t)$ und rechts ohne Berücksichtigung des Rauschens \Rightarrow Signal $d_S(t)$.



Definition: Unter dem **Augendiagramm** (im Englischen: *Eye Pattern*) versteht man die Summe aller übereinander gezeichneten Ausschnitte des Detektionssignals, deren Dauer ein ganzzahliges Vielfaches der Symboldauer T ist.

Dieses Diagramm hat eine gewisse Ähnlichkeit mit einem Auge, was zu seiner Namensgebung geführt hat. Diese Darstellung erlaubt wichtige Aussagen über die Qualität eines digitalen Übertragungssystems:

- Nur das Augendiagramm des Signals $d(t)$ kann messtechnisch auf einem Oszilloskop dargestellt werden, das mit dem Taktsignal getriggert wird. Aus diesem Augendiagramm (linke Grafik) kann beispielsweise der Rauscheffektivwert σ_d abgelesen – besser gesagt: abgeschätzt – werden.
- Das Augendiagramm ohne Rauschen (rechte Grafik) bezieht sich auf das Detektionsnutzsignal $d_S(t)$ und kann nur mittels einer Rechnersimulation ermittelt werden. Für ein realisiertes System ist dieses Augendiagramm nicht darstellbar, da der Rauschanteil $d_N(t)$ nicht eliminiert werden kann.
- Bei beiden Diagrammen wurden jeweils 2048 Augenlinien gezeichnet. In der rechten Grafik sind jedoch nur $2^5 = 32$ Augenlinien unterscheidbar, da der vorliegende Detektionsgrundimpuls $g_d(t)$ auf den Zeitbereich $|t| \leq 2T$ beschränkt ist (siehe **frühere Grafik** mit $f_G \cdot T = 0.4$, rote Kurve).
- Die inneren Augenlinien bestimmen die **vertikale Augenöffnung** $\ddot{o}(T_D)$. Je kleiner diese ist, desto größer ist der Einfluss von Impulsinterferenzen. Bei einem (impulsinterferenzfreien) Nyquistsystem ist die vertikale Augenöffnung maximal. Normiert auf die Sendeamplitude gilt hier $\ddot{o}(T_D)/s_0 = 2$.
- Bei symmetrischem Grundimpuls ist der Detektionszeitpunkt $T_D = 0$ optimal. Mit einem anderen Wert (z.B. $T_D = -T/10$) wäre $\ddot{o}(T_D)/s_0$ etwas kleiner und damit die Fehlerwahrscheinlichkeit deutlich größer. Dieser Fall ist in der Grafik durch die violett-gestrichelte Vertikale angedeutet.

Mittlere Fehlerwahrscheinlichkeit (1)

Wir gehen wie bei den bisherigen Grafiken im Kapitel 3.2 von folgenden Voraussetzungen aus:

- NRZ-Rechtecke mit Amplitude s_0 , AWGN-Rauschen mit N_0 , wobei

$$10 \cdot \lg \frac{s_0^2 \cdot T}{N_0} \approx 13 \text{ dB} \Rightarrow \frac{N_0}{s_0^2 \cdot T} = 0.05.$$

- Gaußförmiges Empfangsfilter mit Grenzfrequenz $f_G \cdot T = 0.4$:

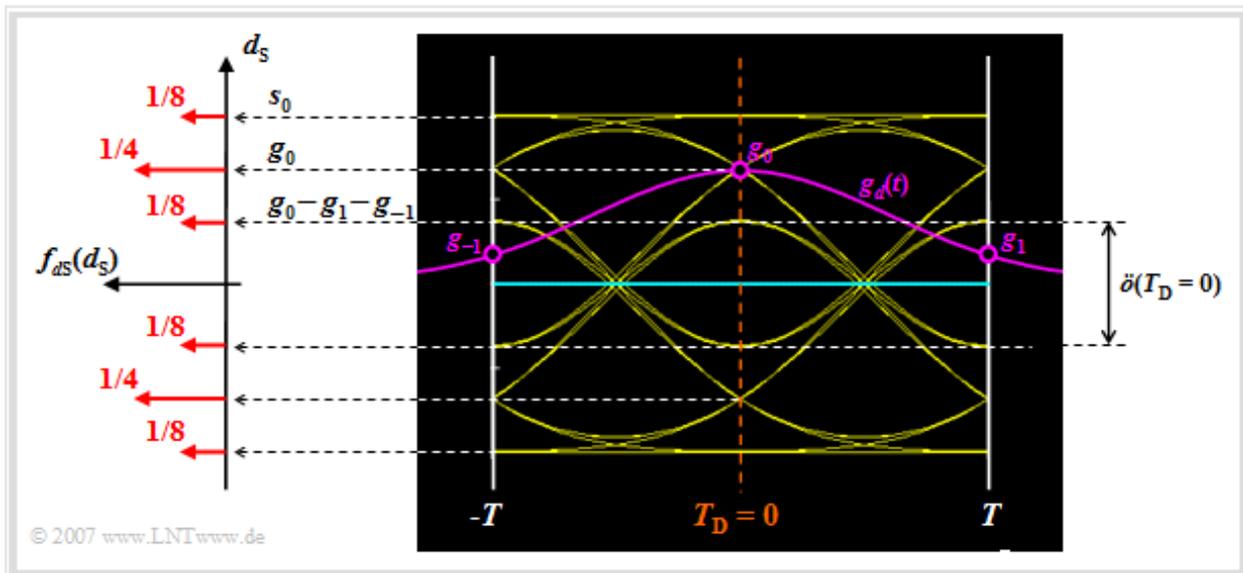
$$\sigma_d^2 = \frac{(N_0/T) \cdot (f_G \cdot T)}{\sqrt{2}} = \frac{0.05 \cdot s_0^2 \cdot 0.4}{\sqrt{2}} \Rightarrow \sigma_d = \sqrt{0.0141} \cdot s_0 \approx 0.119 \cdot s_0.$$

- Für die Detektionsgrundimpulswerte gilt:

$$g_0 = g_d(t = 0) \approx 0.68 \cdot s_0,$$

$$g_1 = g_d(t = T) \approx 0.16 \cdot s_0, \quad g_{-1} = g_d(t = -T) \approx 0.16 \cdot s_0.$$

Alle anderen Grundimpulswerte können vernachlässigt werden.

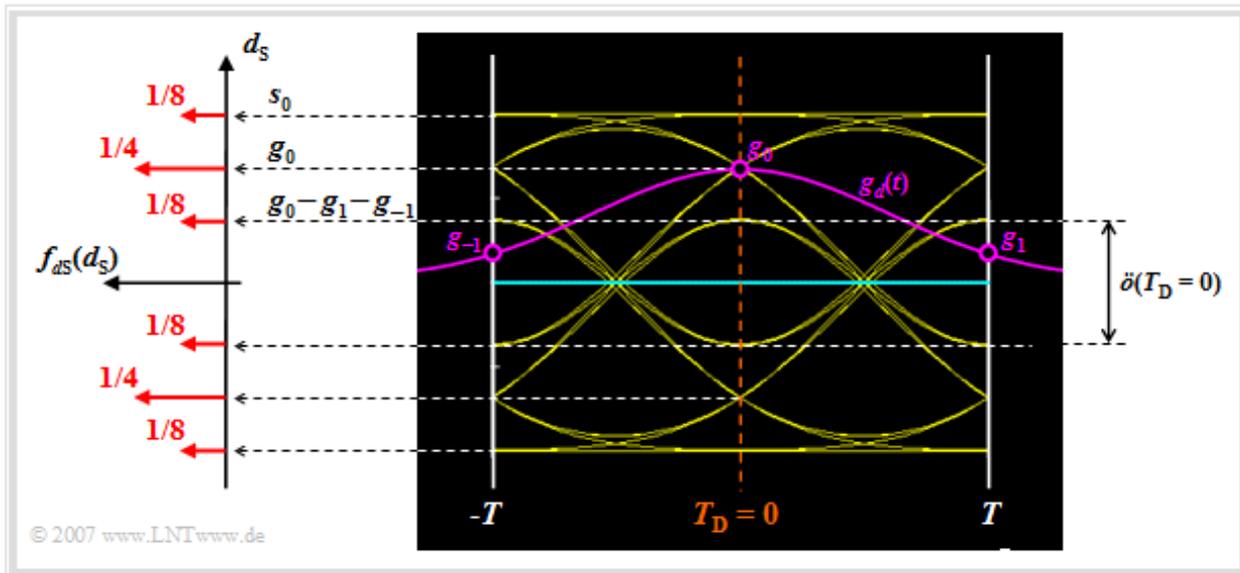


Analysieren wir nun die möglichen Werte für das Detektionsnutzsignal zu den Detektionszeitpunkten:

- Von den insgesamt 32 Augenlinien schneiden vier die Ordinate $t = 0$ bei $g_0 + 2 \cdot g_1 = s_0$. Diese Linien gehören zu den Amplitudenkoeffizienten „... , +1, +1, +1, ...“.
- Die vier Augenlinien, die jeweils die Amplitudenkoeffizienten „... , -1, +1, -1, ...“ repräsentieren, ergeben den Detektionsnutzabtwert $d_S(T_D = 0) = g_0 - 2 \cdot g_1 = 0.36 \cdot s_0$.
- Dagegen tritt der Nutzabtwert $d_S(T_D = 0) = g_0 = 0.68 \cdot s_0$ doppelt so häufig auf. Dieser geht auf die Amplitudenkoeffizienten „... , +1, +1, -1, ...“ oder „... , -1, +1, +1, ...“ zurück.
- Für die 16 Augenlinien, welche die Ordinate $t = 0$ unterhalb der Entscheidungsschwelle $E = 0$ schneiden, ergeben sich genau spiegelbildliche Verhältnisse.

Die Konsequenzen dieser Analyse werden auf der nächsten Seite beschrieben.

Mittlere Fehlerwahrscheinlichkeit (2)



Die möglichen Werte $d_S(T_D)$ und deren **Auftrittswahrscheinlichkeiten** findet man in obiger Grafik in der Wahrscheinlichkeitsdichtefunktion (WDF) der Detektionsnutzabtastwerte wieder:

$$f_{d_S}(d_S) = 1/8 \cdot \delta(d_S - s_0) + 1/4 \cdot \delta(d_S - 0.68 \cdot s_0) + 1/8 \cdot \delta(d_S - 0.36 \cdot s_0) + 1/8 \cdot \delta(d_S + s_0) + 1/4 \cdot \delta(d_S + 0.68 \cdot s_0) + 1/8 \cdot \delta(d_S + 0.36 \cdot s_0).$$

Damit kann die (mittlere) **Symbolfehlerwahrscheinlichkeit** des impulsinterferenzbehafteten Systems angegeben werden. Unter Ausnutzung der Symmetrie erhält man mit $\sigma_d/s_0 = 0.119$:

$$\begin{aligned} p_S &= 1/4 \cdot Q\left(\frac{s_0}{\sigma_d}\right) + 1/2 \cdot Q\left(\frac{0.68 \cdot s_0}{\sigma_d}\right) + 1/4 \cdot Q\left(\frac{0.36 \cdot s_0}{\sigma_d}\right) \\ &\approx 1/4 \cdot Q(8.40) + 1/2 \cdot Q(5.71) + 1/4 \cdot Q(3.02) \approx \\ &\approx 1/4 \cdot 2.20 \cdot 10^{-17} + 1/2 \cdot 1.65 \cdot 10^{-9} + 1/4 \cdot 1.26 \cdot 10^{-3} \approx 3.14 \cdot 10^{-4}. \end{aligned}$$

Anhand dieses Zahlenbeispiels erkennt man, dass

- bei Vorhandensein von Impulsinterferenzen die (mittlere) Fehlerwahrscheinlichkeit p_S wesentlich durch die inneren Augenlinien bestimmt wird,
- der Rechenaufwand zur Bestimmung der Fehlerwahrscheinlichkeit p_S sehr groß ist, insbesondere dann, wenn die Impulsinterferenzen von sehr vielen Grundimpulswerten g_v herrühren.

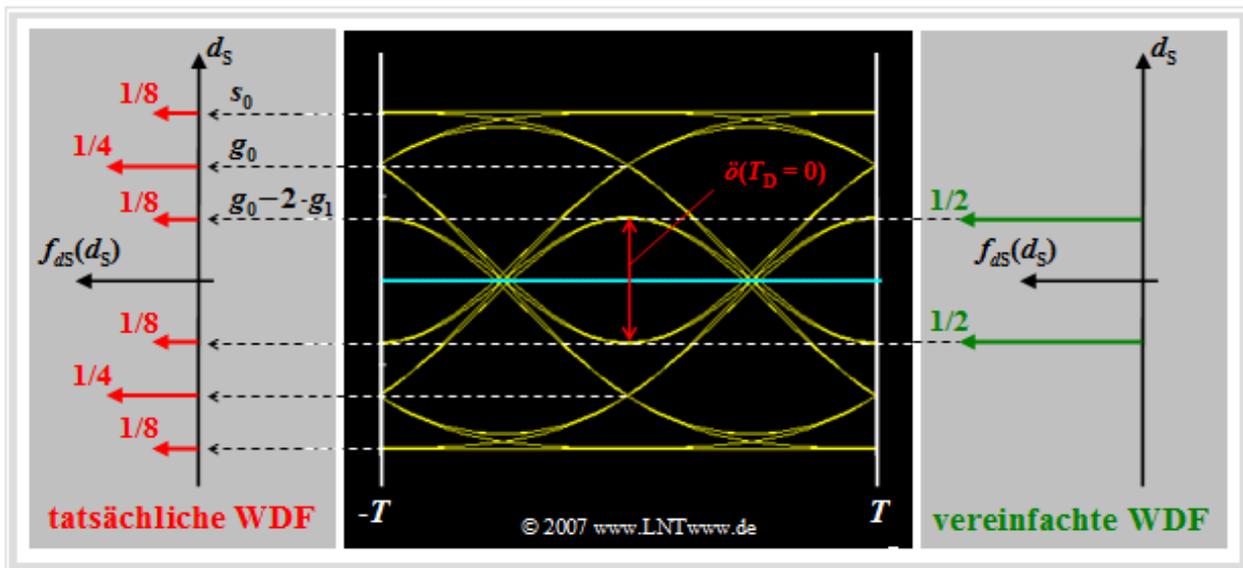
Beispiel: Gilt wie hier $g_2 = g_{-2} = \dots = 0$, so muss zur Berechnung von p_S nur über drei Terme (falls $g_{-1} = g_1$) bzw. über vier Terme (andernfalls) gemittelt werden, wenn die Symmetrie bezüglich der Entscheidungsschwelle $E = 0$ berücksichtigt wird. Sind dagegen die Grundimpulswerte g_{-5}, \dots, g_5 von Null verschieden und $E \neq 0$, so ist eine Mittelung über bis zu $2^{11} = 2048$ Augenlinien erforderlich.

Ungünstigste Fehlerwahrscheinlichkeit (1)

Als eine sehr einfache Näherung für die tatsächliche Fehlerwahrscheinlichkeit p_S verwendet man häufig die **ungünstigste Fehlerwahrscheinlichkeit** (englisch: *Worst-Case Error Probability*)

$$p_U = Q \left(\frac{\ddot{o}(T_D)/2}{\sigma_d} \right),$$

für deren Berechnung stets von den ungünstigsten Symbolfolgen ausgegangen wird. Das bedeutet, dass hier die tatsächliche WDF der Nutzabtastwerte (in der Grafik links eingezeichnet) durch eine vereinfachte WDF mit nur den beiden inneren Diracfunktionen (in der Grafik rechts dargestellt) ersetzt wird.



Für die halbe vertikale Augenöffnung gilt mit den Grundimpulsweiten $g_\nu = g_d(T_D + \nu \cdot T)$ allgemein:

$$\ddot{o}(T_D)/2 = g_0 - \sum_{\nu=1}^n |g_\nu| - \sum_{\nu=1}^v |g_{-\nu}|.$$

Diese Gleichung kann wie folgt interpretiert werden:

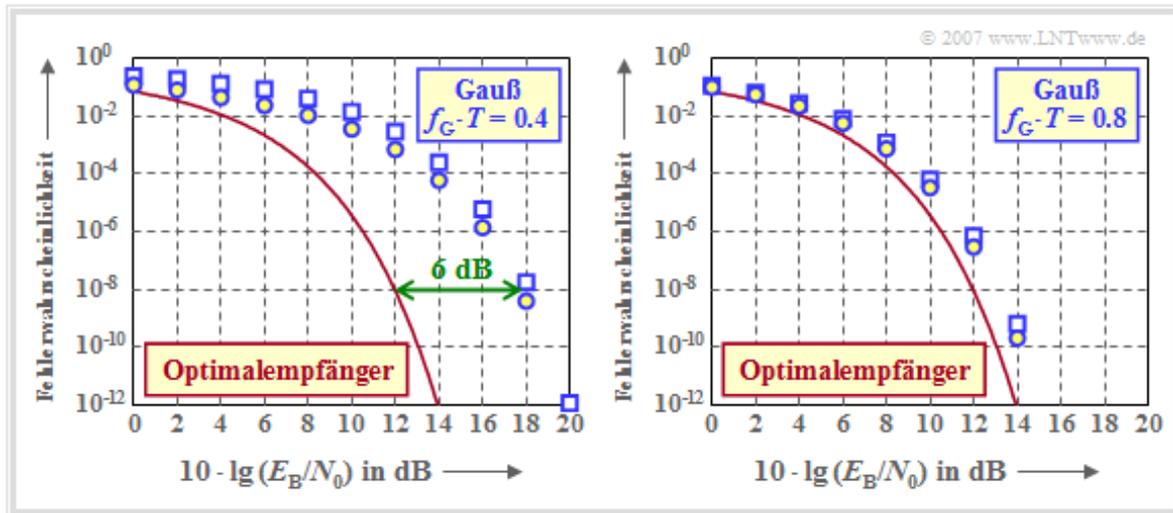
- $g_0 = g_d(T_D)$ ist der so genannte **Hauptwert** des Grundimpulses. Bei Nyquistsystemen gilt stets $\ddot{o}(T_D)/2 = g_0$. Mit Ausnahme von **Kapitel 3.6** wird im Folgenden stets $T_D = 0$ gesetzt.
- Die erste Summe beschreibt die Impulsinterferenzen der n **Nachläufer** vorangegangener Impulse. Stillschweigend vorausgesetzt wird $g_\nu = 0$ für $\nu > n$.
- Die zweite Summe berücksichtigt den Einfluss der v **Vorläufer** nachfolgender Impulse unter der Voraussetzung $g_{-\nu} = 0$ für $\nu > v$.
- Sind alle Impulsvor- und -nachläufer positiv, so lauten die beiden ungünstigsten Symbolfolgen „... , -1, -1, +1, -1, -1, ...” und „... , +1, +1, -1, +1, +1, ...”. Dies trifft zum Beispiel für das hier betrachtete gaußförmige Empfangsfilter zu.
- Sind einige Grundimpulsweiten negativ, so wird dies in obiger Gleichung durch die Betragsbildung berücksichtigt. Es ergeben sich dann andere „Worst-case“-Folgen als gerade genannt.

Ungünstigste Fehlerwahrscheinlichkeit (2)

Die Grafik zeigt die Fehlerwahrscheinlichkeiten in Abhängigkeit des Quotienten E_B/N_0 , nämlich

- die mittlere Fehlerwahrscheinlichkeit p_S bei gaußförmigem Empfangsfilter (blaue Kreise),
- die ungünstigste Fehlerwahrscheinlichkeit p_U bei gaußförmigem Empfangsfilter (blaue Rechtecke),
- die Fehlerwahrscheinlichkeit ($p_S = p_U$) des Optimalempfängers gemäß **Kapitel 1.2** (rote Kurve).

Die Energie pro Bit ist dabei gleich $E_B = s_0^2 \cdot T$ (NRZ-Rechteck-Sendeimpulse).



Die linke Grafik gilt für die (normierte) Grenzfrequenz $f_G \cdot T = 0.4$, die rechte für ein breitbandigeres Empfangsfilter mit $f_G \cdot T = 0.8$. Diese Ergebnisse können wie folgt interpretiert werden:

- Die ungünstigste Fehlerwahrscheinlichkeit p_U ist stets eine obere Schranke für die tatsächliche Symbolfehlerwahrscheinlichkeit p_S . Je kleiner der Einfluss der Impulsinterferenzen ist (große Grenzfrequenz), um so näher liegen p_S und p_U zusammen. Beim Optimalempfänger gilt $p_S = p_U$.
- Bei gaußförmigem Empfangsfilter mit $f_G \cdot T \geq 0.3$ werden die Impulsinterferenzen allein durch die Nachbarimpulse hervorgerufen ($g_2 = g_3 = \dots \approx 0$), so dass für p_S auch eine untere Schranke angegeben werden kann:

$$p_U/4 \leq p_S \leq p_U.$$

- Die starken Impulsinterferenzen eines gaußförmigen Empfangsfilters mit $f_G \cdot T = 0.4$ führen dazu, dass gegenüber dem Optimalempfänger ein um 6 dB größeres E_B/N_0 aufgewendet werden muss (vierfache Leistung), damit die Fehlerwahrscheinlichkeit den Wert 10^{-8} nicht überschreitet.
- Der horizontale Abstand zwischen der blauen p_S -Kurve (markiert durch Kreise) und der roten Vergleichskurve ist aber nicht konstant. Bei $p_S = 10^{-2}$ beträgt der Abstand nur 4 dB.
- Die rechte Grafik zeigt, dass mit $f_G \cdot T = 0.8$ der Abstand zum Vergleichssystem weniger als 1 dB beträgt. Auf der nächsten Seite wird gezeigt, dass bei einem gaußförmigen Empfangsfilter die (normierte) Grenzfrequenz $f_G \cdot T = 0.8$ näherungsweise das Optimum darstellt.

Optimierung der Grenzfrequenz (1)

Für die Systemoptimierung und den Systemvergleich erweist es sich als zweckmäßig, anstelle von ρ_U das **ungünstigste Signal-zu-Rausch-Leistungsverhältnis** (S/N-Verhältnis) zu verwenden:

$$\rho_U = \frac{[\ddot{o}(T_D)]^2}{\sigma_d^2} \Rightarrow p_U = Q(\sqrt{\rho_U}) .$$

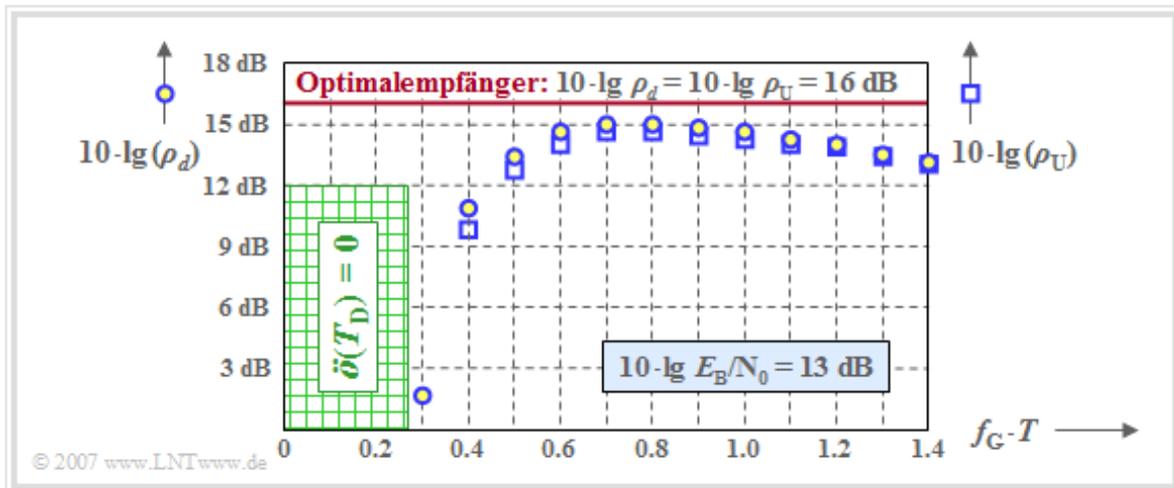
Die mittlere Symbolfehlerwahrscheinlichkeit p_S kann formal über die Q-Funktion ebenfalls durch ein S/N-Verhältnis ausgedrückt werden:

$$\rho_d = [Q^{-1}(p_S)]^2 .$$

Die Grafik zeigt die beiden Größen in logarithmischer Form ($10 \cdot \lg \rho_d$ bzw. $10 \cdot \lg \rho_U$) abhängig von der normierten Grenzfrequenz $f_G \cdot T$ eines gaußförmigen Empfangsfilters, wobei $10 \cdot \lg E_B/N_0 = 13$ dB zugrunde liegt. Zum Vergleich ist als rote horizontale Linie auch das Ergebnis für den Optimalempfänger eingezeichnet. Hier gilt gemäß **Kapitel 1.2**:

$$\rho_d = \rho_U = \frac{2 \cdot E_B}{N_0} \Rightarrow 10 \cdot \lg \rho_d = 10 \cdot \lg \rho_U \approx 16 \text{ dB} .$$

Die gelb gefüllten Kreise kennzeichnen $10 \cdot \lg \rho_d$ und die Quadrate $10 \cdot \lg \rho_U$.



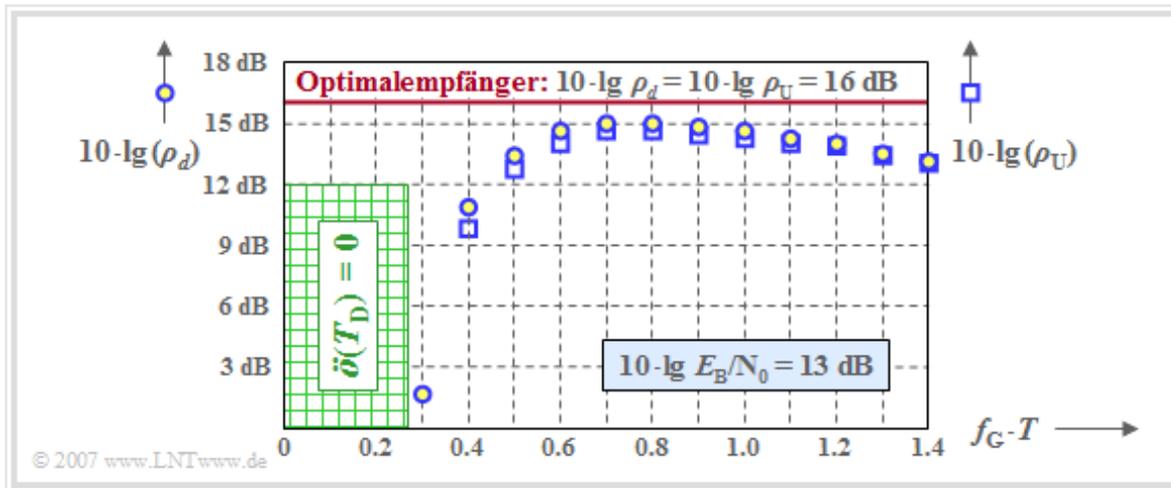
Die Bildbeschreibung folgt auf der nächsten Seite.

Optimierung der Grenzfrequenz (2)

Die Grafik zeigt die beiden soeben definierten Größen in logarithmischer Darstellung,

- $10 \cdot \lg \rho_d \Rightarrow$ gelb gefüllte Kreise \Rightarrow „mittleres“ Detektions-SNR,
- $10 \cdot \lg \rho_U \Rightarrow$ blau umrandete Quadrate \Rightarrow ungünstigstes SNR,

abhängig von der Grenzfrequenz $f_G \cdot T$ eines gaußförmigen Empfangsfilters für $10 \cdot \lg E_B/N_0 = 13$ dB. Die rote horizontale Linie zeigt das Ergebnis für den Optimalempfänger.



Man erkennt aus der Grafik:

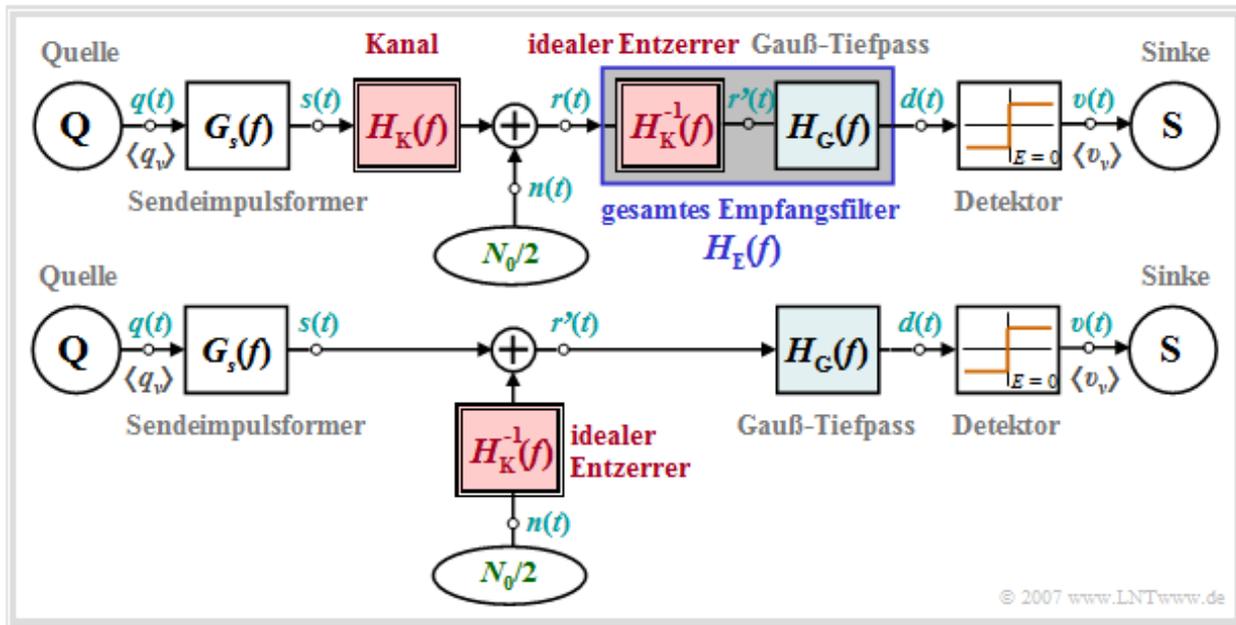
- Das Optimierungskriterium ρ_d führt näherungsweise zur optimalen Grenzfrequenz $f_G \cdot T = 0.8$. Eine kleinere Grenzfrequenz hat stärkere Impulsinterferenzen zur Folge (kleinere Augenöffnung), eine größere Grenzfrequenz bewirkt einen größeren Rauscheffektivwert σ_d .
- Ein gaußförmiges Empfangsfilter mit $f_G \cdot T \approx 0.8$ führt zum Störabstand $10 \cdot \lg \rho_d \approx 15$ dB und damit zur Fehlerwahrscheinlichkeit $p_S \approx 10^{-8}$. Zum Vergleich: Für den optimalen Empfänger (an den Sender angepasste Impulsantwort) ergeben sich $10 \cdot \lg \rho_d \approx 16$ dB und $p_S \approx 10^{-10}$.
- Die Grafik zeigt aber auch, dass das sehr viel einfachere Optimierungskriterien ρ_U (bzw. p_U) näherungsweise zur gleichen optimalen Grenzfrequenz $f_G \cdot T = 0.8$ führt. Für diese Grenzfrequenz erhält man $10 \cdot \lg \rho_U \approx 14.7$ dB sowie die ungünstigste Fehlerwahrscheinlichkeit $p_U \approx 3 \cdot 10^{-8}$.
- Ist die Grenzfrequenz $f_G \cdot T < 0.27$, so ergibt sich für die vertikale Augenöffnung $\delta(T_D) = 0$. Man spricht von einem *geschlossenen Auge*. Dies hat zur Folge, dass einige ungünstige Impulsfolgen auch dann falsch entschieden würden, wenn überhaupt kein Rauschen vorhanden wäre.
- Weitere Untersuchungen haben gezeigt, dass das Optimierungskriterium ρ_U auch bei kleinerem E_B/N_0 ausreichend ist. Bei einem verzerrungsfreien Kanal, d.h. $H_K(f) = 1$, ergibt sich somit die optimale Grenzfrequenz stets zu $f_G \cdot T \approx 0.8$, zumindest bei realitätsnaher Betrachtungsweise.

Alle Aussagen von Kapitel 3.2 können mit folgendem Interaktionsmodul nachvollzogen werden:

Augendiagramm und Augenöffnung

Idealer Kanalentzerrer (1)

Bei einem Übertragungssystem, dessen Kanalfrequenzgang $H_K(f)$ starke Verzerrungen hervorruft, gehen wir von folgendem Blockschaltbild (obere Grafik) und äquivalentem Ersatzschaltbild (untere Grafik) aus.



Zu diesen Darstellungen ist Folgendes anzumerken:

- Das Empfangsfilter $H_E(f)$ wird – zumindest gedanklich – aus einem **idealen Kanalentzerrer** $1/H_K(f)$ und einem Tiefpass $H_G(f)$ zusammengesetzt. Hierfür verwenden wir in diesem Kapitel beispielhaft einen **Gaußtiefpass** mit der Grenzfrequenz f_G .
- Verschiebt man nun den idealen Entzerrer – wiederum rein gedanklich – auf die linke Seite der Rauschadditionsstelle, so ändert sich bezüglich dem S/N-Verhältnis an der Senke und bezüglich der Fehlerwahrscheinlichkeit nichts gegenüber dem oben gezeichneten Blockschaltbild.
- Aus dem unteren Ersatzschaltbild erkennt man, dass sich durch den Kanalfrequenzgang $H_K(f)$ bezüglich des Detektionsnutzsignals $d_S(t)$ nichts ändert, wenn man diesen mit $1/H_K(f)$ vollständig kompensiert. Das Nutzsignal hat somit die genau gleiche Form wie im Kapitel 3.2 berechnet.
- Die Degradation durch den Kanalfrequenzgang $H_K(f)$ zeigt sich vielmehr durch eine signifikante Erhöhung der **Detektionsstörleistung**, also der Varianz des Signals $d_N(t)$:

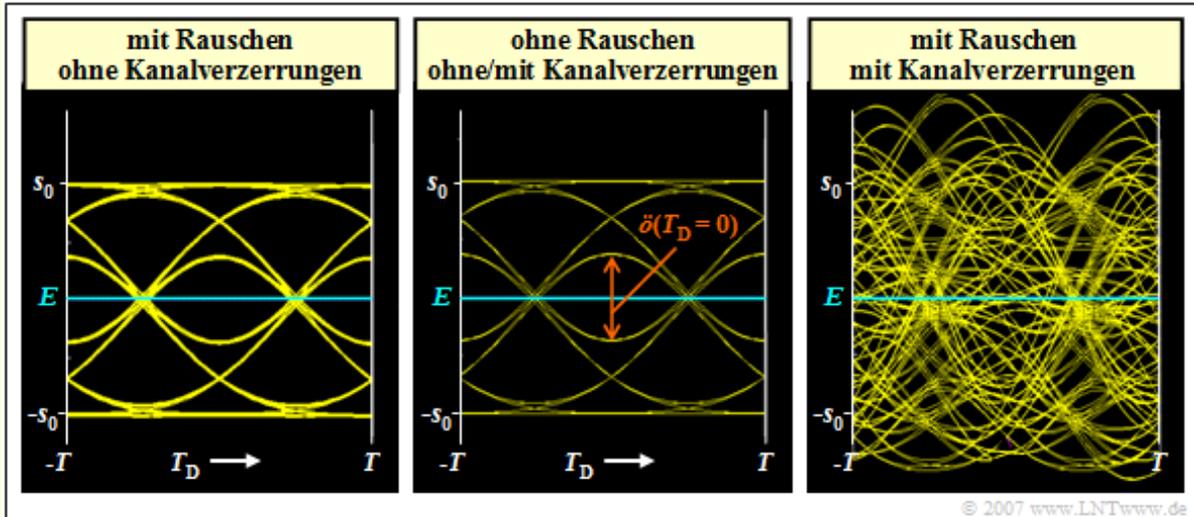
$$\sigma_d^2 = \frac{N_0}{2} \cdot \int_{-\infty}^{+\infty} |H_E(f)|^2 df = \frac{N_0}{2} \cdot \int_{-\infty}^{+\infty} \frac{1}{|H_K(f)|^2} \cdot |H_G(f)|^2 df.$$

- Voraussetzung für endliches σ_d^2 ist, dass der Tiefpass $H_G(f)$ das Rauschen $n(t)$ bei (sehr) hohen Frequenzen stärker abschwächt, als es vom idealen Entzerrer $1/H_K(f)$ angehoben wird.

Anmerkung: Der Kanalfrequenzgang $H_K(f)$ muss nach Betrag und Phase entzerrt werden, allerdings nur in einem von $H_G(f)$ vorgegebenen eingeschränkten Frequenzbereich. Eine vollständige Phasentzerrung ist aber nur auf Kosten einer (frequenzunabhängigen) Laufzeit möglich, die im Folgenden nicht weiter berücksichtigt wird.

Idealer Kanalentzerrer (2)

Beispiel: Wir betrachten wieder ein Binärsystem mit NRZ-Rechteckimpulsen und gaußförmigem Empfangsfilter mit der (normierten) Grenzfrequenz $f_G \cdot T = 0.4$. Die mittlere Grafik zeigt für diesen Fall das Augendiagramm des Detektionsnutzsignals $d_S(t)$ – also ohne Berücksichtigung des Rauschens. Dieses ist identisch mit dem in **Kapitel 3.2** mehrfach dargestellten Augendiagramm.



Das linke Augendiagramm ergibt sich bei idealem Kanal, also für $H_K(f) = 1$. Es berücksichtigt das AWGN-Rauschen, das aber hier mit $10 \cdot \lg E_B/N_0 = 30$ dB deutlich geringer angenommen wurde als im Kapitel 3.2. Für diese Konfiguration wurde per Simulation ermittelt:

$$10 \cdot \lg \rho_U \approx 26.8 \text{ dB} \Rightarrow p_U < 10^{-40}.$$

Dagegen gilt das rechte Diagramm für ein **Koaxialkabel**, wobei die charakteristische Kabeldämpfung $a_* = 40$ dB beträgt. Hierfür ergeben sich deutlich ungünstigere Systemgrößen:

$$10 \cdot \lg \rho_U \approx -4.6 \text{ dB} \Rightarrow p_U \approx 0.28.$$

Dieses Ergebnis kann wie folgt interpretiert werden:

- Unter der Voraussetzung eines idealen Kanalentzerrers ergibt sich das gleiche „Augendiagramm ohne Rauschen“ wie beim idealen Kanal $H_K(f) = 1$ (siehe mittlere Grafik).
- Durch die Kanalverzerrung $1/H_K(f)$ wird der Rauschanteil extrem verstärkt. Im rechten Beispiel ist wegen der starken Verzerrung eine weitgehende Entzerrung erforderlich. Die Rauschleistung σ_d^2 ist um den Faktor 1300 größer als links (keine Verzerrung \Rightarrow keine Entzerrung).
- Eine akzeptable Fehlerwahrscheinlichkeit ergibt sich nur bei kleinerer Rauschleistungsdichte N_0 . Beispielsweise erhält man mit $10 \cdot \lg E_B/N_0 = 50$ dB (statt 30 dB) folgendes Ergebnis:

$$10 \cdot \lg \rho_U \approx 15.4 \text{ dB} \Rightarrow p_U \approx 2 \cdot 10^{-9}.$$

Erhöhung der Rauschleistung durch lineare Entzerrung (1)

Die Augendiagramme auf der letzten Seite dokumentieren eindrucksvoll die Erhöhung der Rauschleistung σ_d^2 bei unveränderter vertikaler Augenöffnung, wenn man den Kanalfrequenzgang $H_K(f)$ empfangsseitig durch dessen Inverse kompensiert. Dieses Ergebnis soll nun anhand der Rauschleistungsdichte $\Phi_{dN}(f)$ nach dem Empfangsfilter (vor dem Entscheider) interpretiert werden, wobei folgende Einstellungen gelten:

- Der Kanal sei ein **Koaxialkabel** mit dem Betragsfrequenzgang

$$|H_K(f)| = \exp \left[-a_* \cdot \sqrt{2fT} \right] \quad \text{mit } a_* = 1.7 \text{ Np (entsprechend 15 dB).}$$

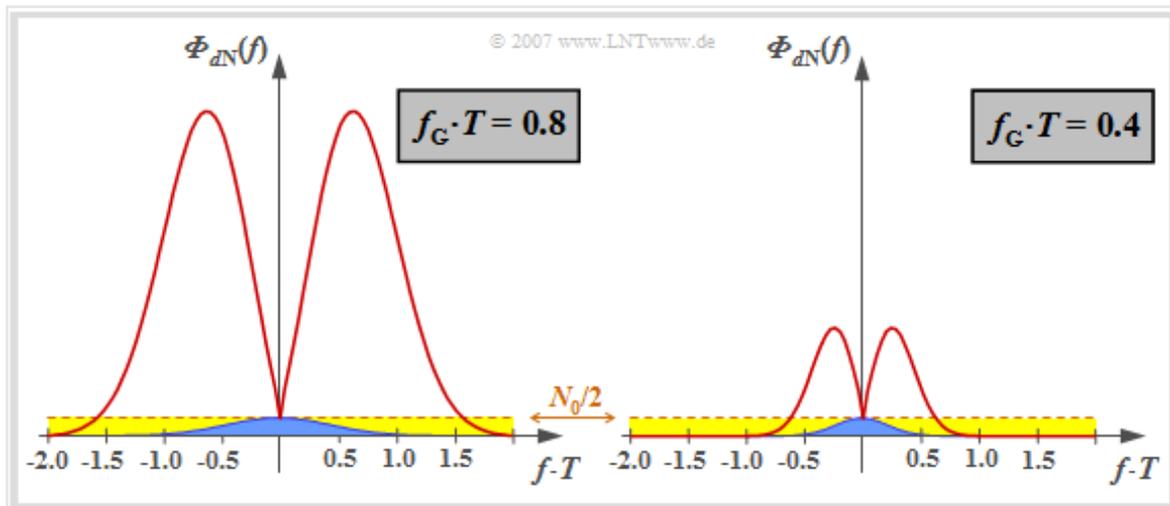
- Der **ideale Kanalentzerrer** $H_K^{-1}(f)$ kompensiert den Kanalfrequenzgang vollständig. Über die Realisierung der Dämpfungs- und Phasenentzerrung wird hier keine Aussage getroffen.
- Zur Rauschleistungsbegrenzung wird ein **Gaußtieffpass** eingesetzt:

$$|H_G(f)| = \exp \left[-\pi \cdot \left(\frac{f}{2f_G} \right)^2 \right] \quad \text{mit } f_G = 0.8/T \text{ bzw. } f_G = 0.4/T.$$

Damit gilt für die Rauschleistungsdichte vor dem Entscheider:

$$\Phi_{dN}(f) = \frac{N_0}{2} \cdot \frac{|H_G(f)|^2}{|H_K(f)|^2} = \frac{N_0}{2} \cdot \exp \left[2 \cdot a_* \cdot \sqrt{2fT} - \pi/2 \cdot (f/f_G)^2 \right].$$

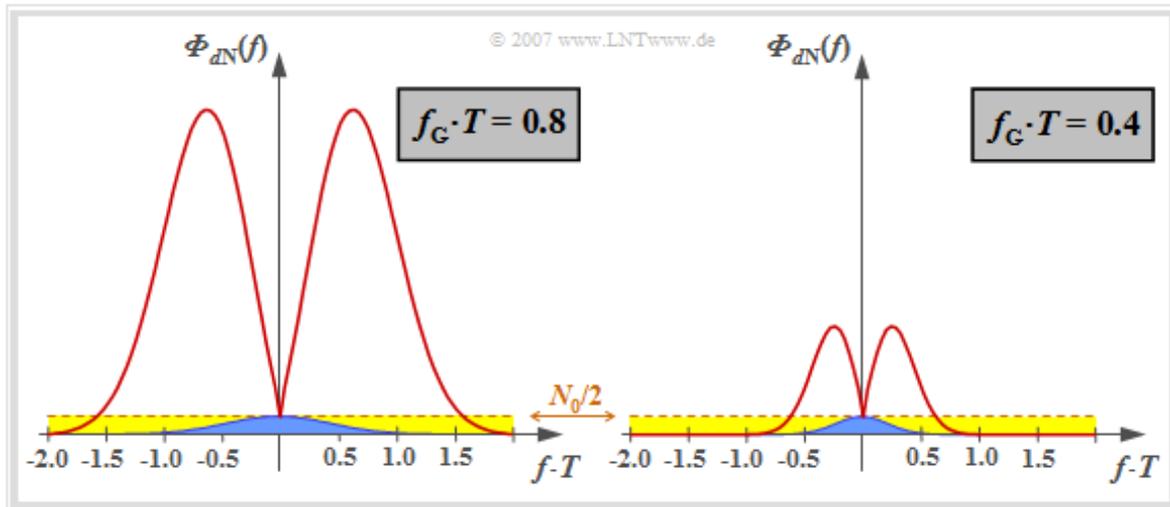
Dieser Verlauf ist nachfolgend für die beiden (normierten) Grenzfrequenzen $f_G \cdot T = 0.8$ (links) bzw. $f_G \cdot T = 0.4$ (rechts) dargestellt. Die Interpretation erfolgt auf der nächsten Seite.



Beachten Sie, dass hier aus Darstellungsgründen die charakteristische Kabeldämpfung mit $a_* = 15$ dB (entsprechend 1.7 Np) deutlich kleiner gewählt ist als beim **rechten Augendiagramm** auf der letzten Seite (gültig für $a_* = 40$ dB).

Erhöhung der Rauschleistung durch lineare Entzerrung (2)

Die Grafik zeigt nochmals die Rauschleistungsdichte $\Phi_{dN}(f)$ für zwei verschiedene Grenzfrequenzen.



Betrachten wir zunächst die linke Grafik für die (normierte) Grenzfrequenz $f_G \cdot T = 0.8$, die nach den Berechnungen im **Kapitel 3.2** für $H_K(f) = 1$ das Optimum darstellt.

- Gelb hinterlegt ist die konstante Rauschleistungsdichte $N_0/2$ am Empfängereingang. Bei idealem Kanal wird diese durch den Gaußtieffpass begrenzt und ergibt die Detektionsrauschleistung σ_d^2 (in der Grafik durch die blaue Fläche gekennzeichnet).
- Werden – wie bei leitungsgebundener Übertragung üblich – höhere Frequenzen stark gedämpft, so steigt $|H_E(f)|^2$ aufgrund des idealen Kanalentzerrers sehr stark an, bevor für $f \cdot T \geq 0.6$ (nur gültig für $a_* = 15$ dB und $f_G \cdot T = 0.8$) der dämpfende Einfluss des Gaußfilters wirksam wird.
- Die Rauschleistung σ_d^2 ist nun gleich der Fläche unter der roten Kurve, die etwa um den Faktor 28 größer ist als die blaue Fläche. Die Auswirkungen dieser unterschiedlichen Rauschleistungen erkennt man auch in den Augendiagrammen auf der letzten Seite, allerdings für $a_* = 40$ dB.

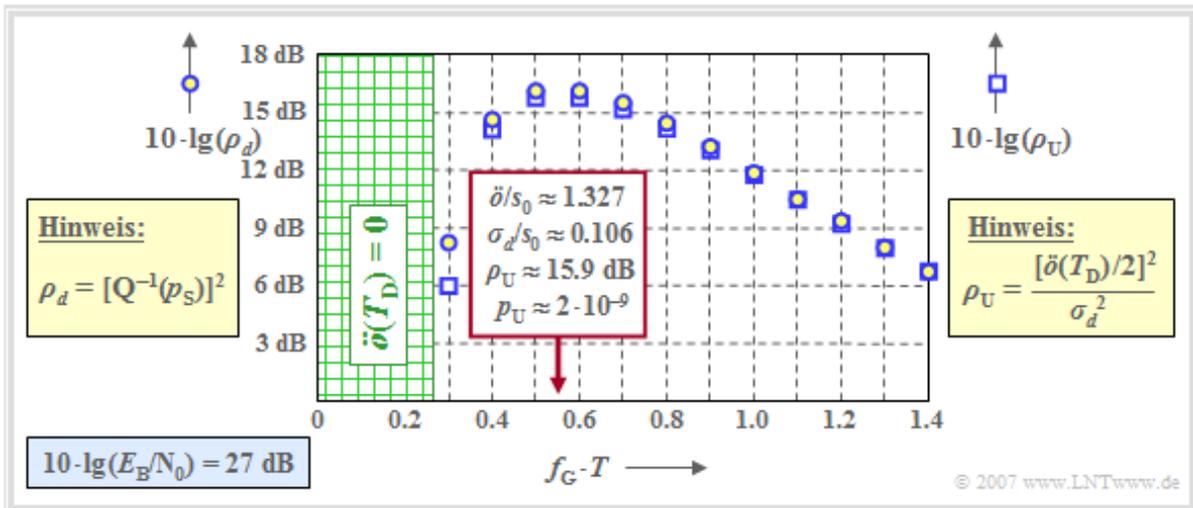
Die rechte Grafik zeigt die Rauschleistungsdichte $\Phi_{dN}(f)$ für die normierte Grenzfrequenz $f_G \cdot T = 0.4$. Hier wird die Rauschleistung durch den idealen Kanalentzerrer nur noch um den Faktor 9 vergrößert (Verhältnis zwischen der Fläche unter der roten Kurve und der blauen Fläche).

Aus obiger Grafik und den bisherigen Erläuterungen geht bereits hervor, dass bei verzerrendem Kanal $\Rightarrow H_K(f) \neq 1$ die Grenzfrequenz $f_G \cdot T = 0.8$ nicht mehr optimal sein wird.

Optimierung der Grenzfrequenz (1)

Die Grafik zeigt die Störabstände $10 \cdot \lg \rho_d$ (als Maß für die mittlere Fehlerwahrscheinlichkeit p_S) sowie $10 \cdot \lg \rho_U$ (als Maß für die ungünstigste Fehlerwahrscheinlichkeit p_U) in Abhängigkeit der Grenzfrequenz f_G des gaußförmigen Gesamtfrequenzgangs $H_G(f) = H_K(f) \cdot H_E(f)$. Dieses Bild gilt für

- einen koaxialen Übertragungskanal mit der charakteristischen Kabeldämpfung $a_* = 15$ dB,
- AWGN-Rauschen mit $10 \cdot \lg E_B/N_0 = 27$ dB, wobei $E_B = s_0^2 \cdot T$ zu setzen ist.



Man erkennt aus dieser Darstellung und durch Vergleich mit der entsprechenden **Grafik** in Kapitel 3.2, die für $H_K(f) = 1$ und $10 \cdot \lg E_B/N_0 = 13$ dB gegolten hat:

- Auch bei stark verzerrendem Kanal ist ρ_U eine geeignete untere Schranke für ρ_d (das heißt, es ist stets $\rho_d \geq \rho_U$) und dementsprechend p_U eine sinnvolle obere Schranke für $p_S \Rightarrow p_U \geq p_S$.
- Bei der hier betrachteten Kabeldämpfung $a_* = 15$ dB ist die Grenzfrequenz $f_G \cdot T = 0.55$ in etwa optimal und es gilt $\ddot{o}(T_D)/s_0 \approx 1.327$ sowie $\sigma_d/s_0 \approx 0.106$. Daraus ergeben sich der (ungünstigste) Störabstand $10 \cdot \lg \rho_U \approx 15.9$ dB und die worst-case-Fehlerwahrscheinlichkeit $p_U \approx 2 \cdot 10^{-9}$.

- Eine kleinere Grenzfrequenz würde zu einer deutlich kleineren Augenöffnung führen, ohne dass dadurch auch σ_d gleichermaßen verkleinert wird. Beispielsweise gilt mit $f_G \cdot T = 0.4$:

$$\ddot{o}/s_0 \approx 0.735, \quad \sigma_d/s_0 \approx 0.072 \quad \Rightarrow \quad 10 \cdot \lg \rho_U \approx 14.1 \text{ dB} \quad \Rightarrow \quad p_U \approx 1.8 \cdot 10^{-7}.$$

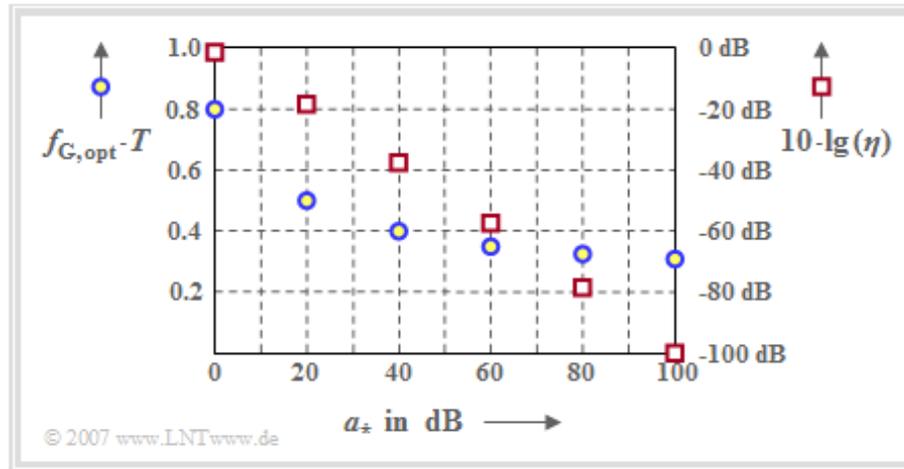
- Ist die Grenzfrequenz f_G zu groß, so wird das Rauschen weniger effektiv begrenzt. Beispielsweise lauten die Werte für die Grenzfrequenz $f_G \cdot T = 0.8$:

$$\ddot{o}/s_0 \approx 1.819, \quad \sigma_d/s_0 \approx 0.178 \quad \Rightarrow \quad 10 \cdot \lg \rho_U \approx 14.2 \text{ dB} \quad \Rightarrow \quad p_U \approx 1.7 \cdot 10^{-7}.$$

- Das Optimum mit $10 \cdot \lg \rho_d \approx 16.2$ dB und $10 \cdot \lg \rho_U \approx 15.9$ dB ist deutlich ausgeprägter als bei idealem Kanal. Bei einem Vergleich der Störabstände ist allerdings zu berücksichtigen, dass hier $10 \cdot \lg E_B/N_0 = 27$ dB zugrunde liegt; im Kapitel 3.2 wurde stets von 13 dB ausgegangen.

Optimierung der Grenzfrequenz (2)

Die optimale Grenzfrequenz $f_{G, \text{opt}}$ hängt signifikant von der Stärke der Verzerrungen des Koaxialkabels ab, wie aus der folgenden Grafik hervorgeht (blaue Kreise mit gelber Füllung, linke Achsenbeschriftung). Genauer gesagt: ausschließlich von der charakteristischen Kabeldämpfung a_* bei der halben Bitrate.



Man erkennt aus dieser Darstellung:

- Je größer die charakteristische Kabeldämpfung a_* ist und damit auch der Einfluss des Rauschens, um so niedriger ist die optimale Grenzfrequenz $f_{G, \text{opt}}$.
- Allerdings ist $f_{G, \text{opt}}$ stets größer als $0.27/T$. Andernfalls würde sich ein geschlossenes Auge ergeben, gleichbedeutend mit der „Worst–case“-Fehlerwahrscheinlichkeit $p_U = 0.5$.

Mit roten Rechtecken ist in der Grafik der **Systemwirkungsgrad**

$$\eta = \frac{\rho_d}{\rho_{d, \text{max} | A}} = \frac{\rho_d}{2 \cdot s_0^2 \cdot T / N_0}$$

als Funktion der charakteristischen Kabeldämpfung a_* dargestellt. η gibt das Verhältnis des mit der betrachteten Konfiguration erreichbaren SNR ρ_d zum maximal möglichen S/N-Verhältnis an, wobei als Nebenbedingung der Optimierung von **Spitzenwertbegrenzung** ausgegangen wird (*Anmerkung:* In Kapitel 1.4 wurde dieser Systemwirkungsgrad mit η_A bezeichnet).

Wegen der NRZ–Sendeimpulse gilt $E_B = s_0^2 \cdot T$. Ersetzt man p_S durch p_U und damit ρ_d durch ρ_U , so lautet obige Gleichung:

$$\eta \approx \frac{\rho_U}{2 \cdot E_B / N_0}$$

Die Diskussion des Kurvenverlaufs $\eta = \eta(a_*)$ folgt auf der nächsten Seite.

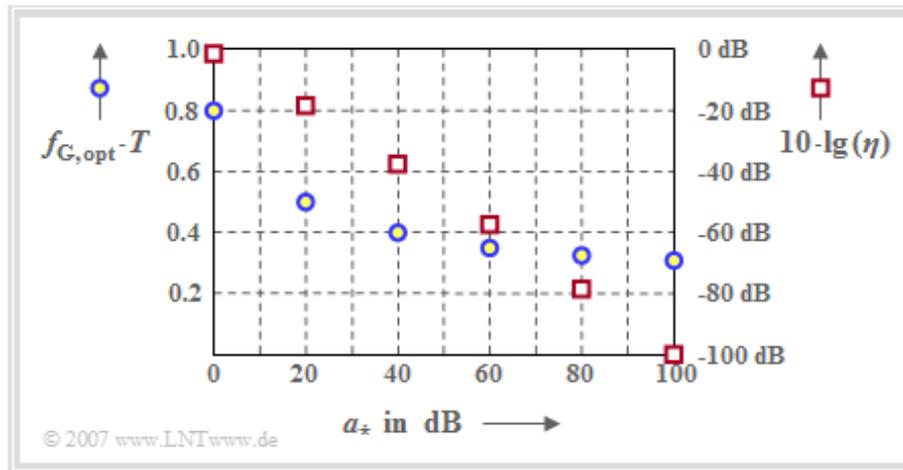
Systemvergleich mittels Systemwirkungsgrad

Die roten Rechtecke geben den **Systemwirkungsgrad**

$$\eta = \frac{\rho_d}{2 \cdot E_B/N_0} \approx \frac{\rho_U}{2 \cdot E_B/N_0}$$

in Abhängigkeit von der charakteristischen Kabeldämpfung a_* an. Diese Darstellung gilt für

- ein Binärsystem mit NRZ-Sendeimpulsen,
- einen koaxialen Übertragungskanal mit der charakteristischen Dämpfung a_* ,
- einen gaußförmigen Gesamtfrequenzgang mit jeweils optimierter Grenzfrequenz (blaue Kreise).



Wie nun an einigen Zahlenbeispielen verdeutlicht werden soll, vermeidet die Darstellung $\eta = \eta(a_*)$ einige Probleme, die sich aus dem großen Wertebereich von S/N-Verhältnissen ergeben:

- $10 \cdot \lg \eta(a_* = 0 \text{ dB}) = -1.4 \text{ dB}$ sagt aus, dass der bestmögliche Gaußtieffpass ($f_G \cdot T = 0.8$) bei $H_K(f) = 1$ um 1.4 dB schlechter ist als der optimale (Matched-Filter-) Empfänger.
- Gehen wir von $H_K(f) = 1$ und $10 \cdot \lg(E_B/N_0) = 10 \text{ dB}$ aus, so besagt die obige Gleichung auch, dass diese Gaußkonfiguration zu folgender (worst-case) Fehlerwahrscheinlichkeit führen wird:

$$\begin{aligned} 10 \cdot \lg \rho_U &= 10 \cdot \lg E_B/N_0 + 10 \cdot \lg(2) + 10 \cdot \lg(\eta) \approx \\ &\approx 10 \text{ dB} + 3 \text{ dB} - 1.4 \text{ dB} = 11.6 \text{ dB} \Rightarrow p_U \approx 7 \cdot 10^{-5} . \end{aligned}$$

- Soll diese (ungünstigste) Fehlerwahrscheinlichkeit $p_U = 7 \cdot 10^{-5} \Rightarrow 10 \cdot \lg \rho_U = 11.6 \text{ dB}$ bei einem Kanal mit der charakteristischen Kabeldämpfung $a_* = 80 \text{ dB}$ nicht überschritten werden, so muss demnach für das Verhältnis E_B/N_0 gelten:

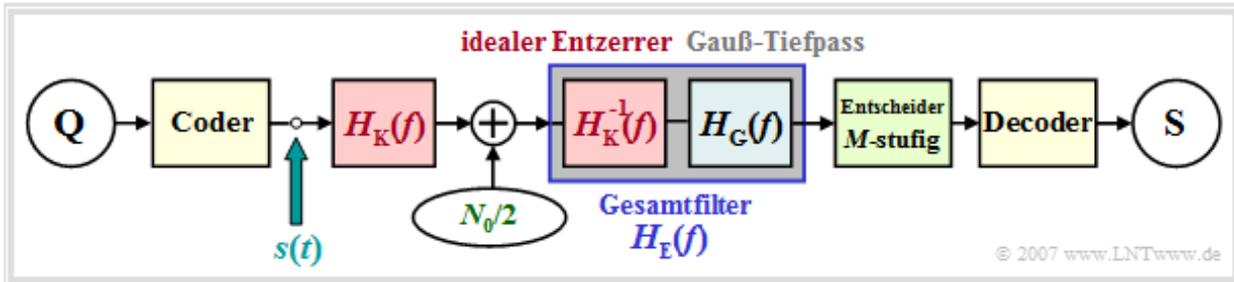
$$10 \cdot \lg E_B/N_0 \geq 11.6 \text{ dB} - 3 \text{ dB} - (-78.2) \text{ dB} = 86.8 \text{ dB} \Rightarrow E_B/N_0 \approx 5 \cdot 10^8 .$$

- Um dies zu erreichen, muss allerdings die Grenzfrequenz des Gaußtieffpasses entsprechend den blauen Kreisen in obiger Gleichung auf $f_G = 0.33/T$ herabgesetzt werden.

Augenöffnung bei redundanzfreien Mehrstufigensystemen (1)

Wir gehen weiterhin von folgenden Voraussetzungen aus:

- NRZ-Rechteck-Sendeimpulse,
- Koaxialkabel und AWGN-Rauschen,
- ideale Kanalverzerrung, sowie
- ein Gaußtieffpass zur Rauschleistungsbegrenzung.



Im Unterschied zu **Kapitel 3.3** ist das weiterhin redundanzfreie Sendesignal $s(t)$ nun nicht mehr binär, sondern M -stufig, was sich nur im Wertevorrat der **Amplitudenkoeffizienten** auswirkt:

$$s(t) = \sum_{\nu=-\infty}^{+\infty} a_{\nu} \cdot g_s(t - \nu \cdot T) \quad \text{mit } a_{\nu} \in \{a_1, \dots, a_{\mu}, \dots, a_M\}.$$

Dementsprechend besitzt der Entscheider nun nicht mehr nur eine, sondern $M - 1$ Entscheidungsschwellen und im Augendiagramm sind bei geöffnetem Auge $M - 1$ Augenöffnungen erkennbar.

Vergleicht man die Augendiagramme (ohne Rauschen)

- eines binären ($M = 2$),
- eines ternären ($M = 3$), und
- eines quaternären ($M = 4$)

Übertragungssysteme bei gleichem vorgegebenen Detektionsgrundimpuls $g_d(t)$ und gleicher Symboldauer T , so erhält man für die halbe vertikale Augenöffnung allgemein:

$$\ddot{o}(T_D)/2 = \frac{g_0}{M - 1} - \sum_{\nu=1}^{\infty} |g_{-\nu}| - \sum_{\nu=1}^{\infty} |g_{\nu}|.$$

Hierbei bezeichnet $g_0 = g_d(t = 0)$ wie im Kapitel 3.3 den Hauptwert, während die beiden Summen in obiger Gleichung

- die Vorläufer g_1, g_2, \dots (zweiter Term), und
- die Nachläufer g_{-1}, g_{-2}, \dots (dritter Term)

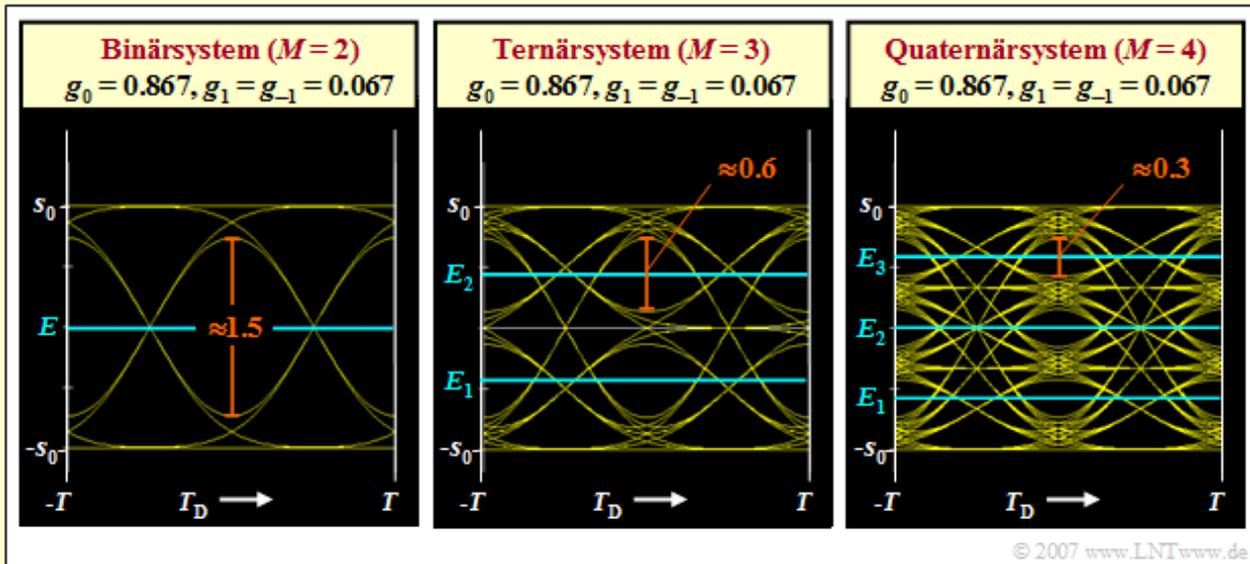
berücksichtigen. Dabei gilt stets $g_{\nu} = g_d(t = \nu \cdot T)$.

Auf der nächsten Seite wird diese Gleichung an einem Beispiel verdeutlicht.

Augenöffnung bei redundanzfreien Mehrstuflensystemen (2)

Beispiel: Nachfolgend sehen Sie drei Augendiagramme für die Stufenzahlen $M = 2$, $M = 3$ und $M = 4$. Das binäre Augendiagramm gilt für einen Gaußtieffpass mit der Grenzfrequenz $f_G \cdot T \approx 0.6$. Mit dem Hauptwert $g_0 = 0.867 \cdot s_0$ und den beiden Ausläufern $g_{-1} = g_1 = 0.067 \cdot s_0$ ergibt sich in diesem Fall für die vertikale Augenöffnung (Rundung auf eine Nachkommastelle):

$$\ddot{o}(T_D) = 2 \cdot (g_0 - 2 \cdot g_1) \approx 1.5 \cdot s_0.$$



Bei den Mehrstuflensystemen ist die Augenöffnung per se um den Faktor $1/(M - 1)$ kleiner. Dadurch wird hier die Augenöffnung durch die (gleich großen) Vor- und Nachläufer (relativ gesehen) stärker verringert als beim Binärsystem. Man erhält bei gleichen Grundimpulswerten für

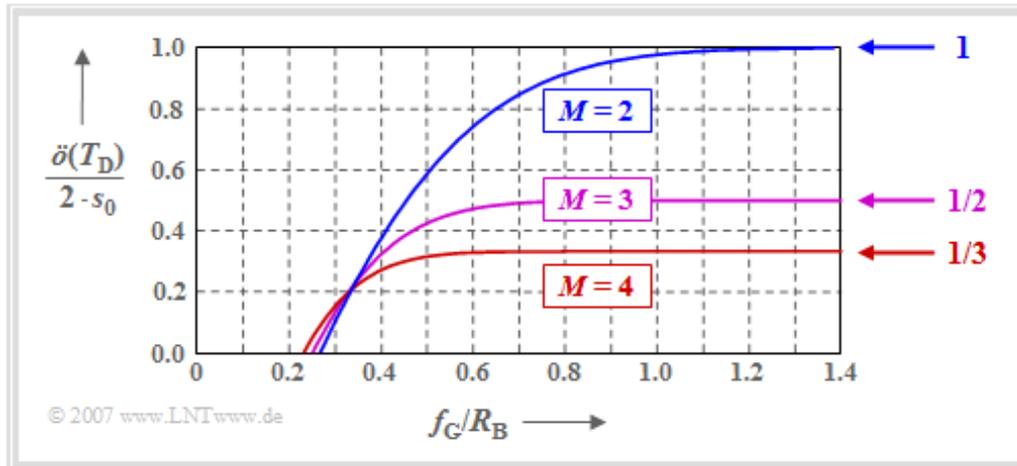
$$M = 3 : \ddot{o}(T_D) = 2 \cdot (g_0/2 - 2 \cdot g_1) \approx 0.6 \cdot s_0,$$

$$M = 4 : \ddot{o}(T_D) = 2 \cdot (g_0/3 - 2 \cdot g_1) \approx 0.3 \cdot s_0.$$

Anzumerken ist, dass auch für diese beiden Mehrstuflensysteme jeweils die normierte Grenzfrequenz $f_G \cdot T = 0.6$ zugrundeliegt. Bei einem Systemvergleich ist allerdings zu beachten, dass sich durch die größere Stufenzahl auch der Informationsfluss erhöht. Das heißt, dass die Mehrstuflensysteme besser sind, als es diese Grafiken aussagen. Mehr darüber auf der nächsten Seite.

Vergleich zwischen Binär- und Quaternärsystem (1)

Der auf der letzten Seite angestellte Vergleich ist nicht fair, da nicht von gleichem Informationsfluss ausgegangen wurde. Ein Systemvergleich bei **konstanter äquivalenter Bitrate** R_B muss vielmehr auch berücksichtigen, dass bei den (redundanzfreien) Mehrstufigensystemen die Symboldauer T um den Faktor $\text{ld}(M)$ größer ist als beim Binärsystem, was sich günstig auf die Impulsinterferenzen auswirkt.



Die Grafik zeigt die (auf s_0 normierte) halbe Augenöffnung in Abhängigkeit des Quotienten f_G/R_B des gaußförmigen Empfangsfilters. In der **Aufgabe Z3.4** wird diese in analytischer Form wie folgt berechnet:

$$\begin{aligned} \ddot{o}_{\text{norm}} = \frac{\ddot{o}(T_D)}{2 \cdot s_0} &= \frac{M}{M-1} \cdot \frac{g_0}{s_0} - 1 = \\ &= \frac{1}{M-1} \cdot \left[1 - 2 \cdot M \cdot Q \left(\sqrt{2\pi} \cdot \text{ld}(M) \cdot f_G/R_B \right) \right]. \end{aligned}$$

Man erkennt aus obiger Grafik:

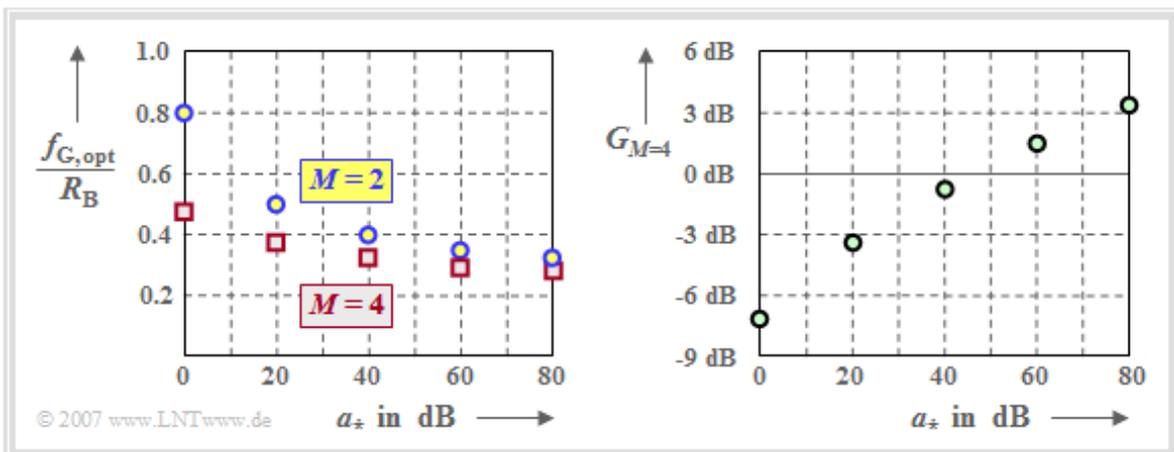
- Bei breitbandigem Filter (das heißt: für großes f_G) ist das Binärsystem den Mehrstufigensystemen deutlich überlegen. Die normierte halbe Augenöffnung beträgt im Grenzfall $\ddot{o}_{\text{norm}} = 1$ (für $M = 2$), $\ddot{o}_{\text{norm}} = 1/2$ (für $M = 3$) bzw. $\ddot{o}_{\text{norm}} = 1/3$ (für $M = 4$).
- Wie aus obiger Grafik hervorgeht, führt für Grenzfrequenzen $f_G/R_B < 0.35$ die Stufenzahl $M = 4$ (rote Kurve) zu einer größeren Augenöffnung als $M = 2$ (blaue Kurve). Das Ternärsystem ($M = 3$, violette Kurve) liegt fast im gesamten Bereich zwischen dem Binär- und dem Quaternärsystem.
- Besonders erwähnenswert ist, dass sich beim Quaternärsystem erst mit einer Grenzfrequenz $f_G/R_B < 0.23$ ein geschlossenes Auge ergibt (was zu sehr großen Fehlerwahrscheinlichkeiten führt), während die Binärübertragung bereits für $f_G/R_B < 0.27$ nicht mehr möglich ist.

Vergleich zwischen Binär- und Quaternärsystem (2)

Vergleichen wir nun die **optimalen Grenzfrequenzen** des Gaußfilters, die sich für $M = 2$ bzw. $M = 4$ ergeben. Dem Vergleich liegt ein koaxialer Übertragungskanal mit der charakteristischen Kabeldämpfung a_* zugrunde. Je größer dieser Kanalparameter ist (das heißt auch: wie größer die Kabellänge ist), desto stärker wird das Rauschen durch die erforderliche Entzerrung beim Empfänger verstärkt.

Interpretieren wir zunächst die linke Grafik:

- Bei verzerrungsfreiem Kanal ($a_* = 0$ dB) ergeben sich die optimalen Grenzfrequenzen zu 0.8 (für $M = 2$) bzw. 0.48 (für $M = 4$) – jeweils normiert auf die äquivalente Bitrate. Entsprechend dem Kurvenverlauf $\ddot{o}(T_D)/(2s_0)$ auf der letzten Seite ist hier das Binärsystem dem Quaternärsystem deutlich überlegen.
- Mit der charakteristischen Kabeldämpfung $a_* = 80$ dB erhält man für das Binärsystem ($M = 2$) die optimale Grenzfrequenz $f_{G,opt} = 0.33/T$. Für das Quaternärsystem (für $M = 4$) ergibt sich ein kleinerer Wert: $f_{G,opt} = 0.28/T$.



Das optimierte Binärsystem ist aber trotz größerer Augenöffnung nicht immer besser als das optimierte Quaternärsystem, da auch die Rauschleistung zu berücksichtigen ist. Diese wird mit kleiner werdenden Grenzfrequenz ebenfalls kleiner.

Die rechte Grafik zeigt den **Störabstandsgewinn** des Quaternärsystems gegenüber dem Binärsystem,

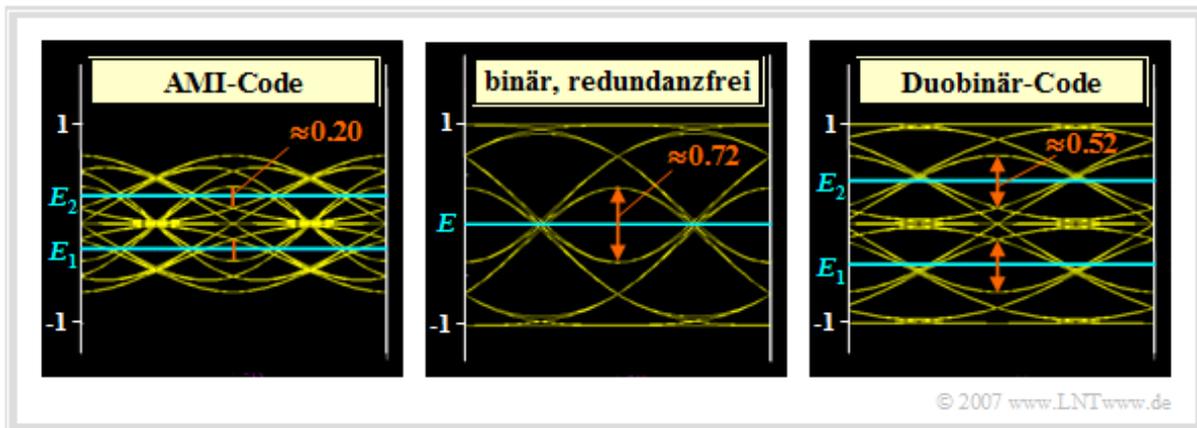
$$G_{M=4} = 10 \cdot \lg \rho_{U, M=4} - 10 \cdot \lg \rho_{U, M=2},$$

wenn die Grenzfrequenzen gemäß der linken Grafik jeweils optimal gewählt werden. Demnach gilt:

- Für $a_* < 50$ dB ist das Binärsystem optimal. Beim verzerrungsfreien Kanal ($a_* = 0$ dB) ergibt sich ein um ca. 7 dB größeres SNR als mit $M = 4$.
- Dagegen ergeben sich für $a_* \geq 50$ dB günstigere Verhältnisse für die Quaternärübertragung. Bei 80 dB Kabeldämpfung ist der Störabstandsgewinn gegenüber $M = 2$ größer als 3 dB.

Augenöffnung bei den Pseudoternär codes (1)

In Kapitel 2.4 wurden die **Pseudoternär codes** allgemein beschrieben und es wurden für diese die Augendiagramme bei Nyquistimpulsformung angegeben. In der Grafik auf dieser Seite sehen Sie die Augendiagramme – jeweils ohne Rauschen – für den AMI-Code (links) und den Duobinär code (rechts) im Vergleich zum redundanzfreien Binär code (Mitte). Die Amplitude ist jeweils zu $s_0 = 1$ normiert.



Alle Augendiagramme gelten für ein gaußförmiges Empfangsfilter mit der Grenzfrequenz $f_G \cdot T = 0.4$, woraus sich folgende (normierte) Grundimpuls werte ergeben:

$$g_0 \approx 0.68, \quad g_1 = g_{-1} \approx 0.16, \quad g_2 = g_{-2} = \dots \approx 0.$$

Beim redundanzfreien Binär system (mittlere Grafik) erhält man somit für die Augenöffnung

$$\ddot{o}(T_D) = 2 \cdot (g_0 - 2 \cdot g_1) = 0.72 \quad \Rightarrow \quad \ddot{o}_{\text{norm}} = \frac{\ddot{o}(T_D)}{2 \cdot s_0} = 0.36$$

im Vergleich zu $\ddot{o}(T_D) = 2$ bzw. $\ddot{o}_{\text{norm}} = 1$ beim binären Nyquist system.

Da die Pseudoternär codes mit der gleichen Symbolrate arbeiten wie das redundanzfreie Binär system,

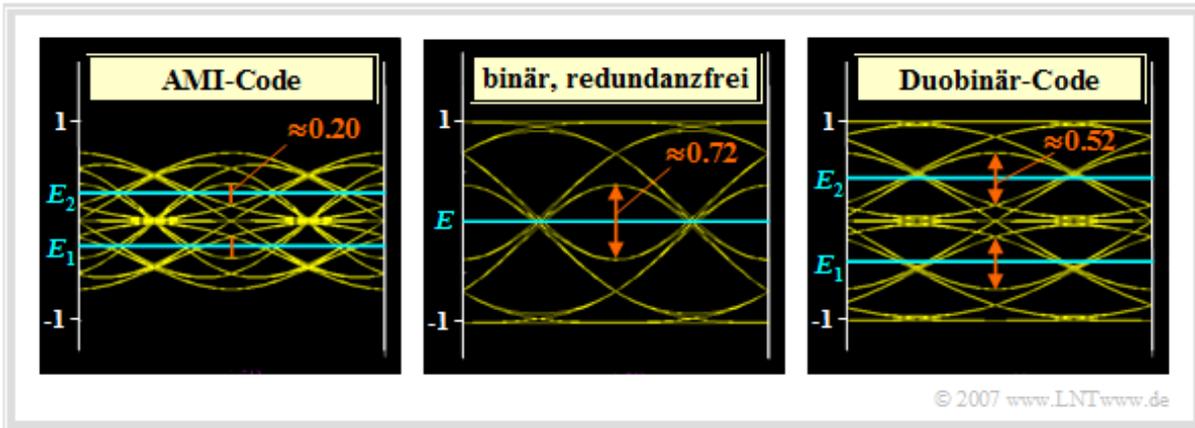
- sind die Detektionsgrundimpuls werte g_v und auch der Rauscheffektivwert σ_d in allen Fällen gleich,
- ist die (halbe) Augenöffnung für die Systemoptimierung ebenso geeignet wie das S/N-Verhältnis

$$\rho_U = [\ddot{o}(T_D)/2]^2 / \sigma_d^2 \text{ und die daraus resultierende (ungünstigste) Fehlerwahrscheinlichkeit } p_U.$$

Die Beschreibung der beiden äußeren Augendiagramme folgt auf der nächsten Seite.

Augenöffnung bei den Pseudoternär codes (2)

Bei den Pseudoternär codes sind jeweils zwei Augenöffnungen zu erkennen und man benötigt für die ternäre Entscheidung zwei Schwellenwerte E_1 und E_2 .



Interpretieren wir nun das Augendiagramm bei **AMI-Codierung**:

- Die obere Begrenzung des oberen Auges gehört zur Symbolfolge „... , -1, +1, -1, ... ” und liegt demzufolge bei $d_{\text{oben}} = g_0 - 2 \cdot g_1$.
- Die untere Begrenzungslinie $d_{\text{unten}} = g_1$ geht auf die Symbolfolge „... , 0, 0, +1, ... ” bzw. auf die Folge „... , +1, 0, 0, ... ” zurück. Hierbei ist berücksichtigt, dass die Folge „... , +1, 0, +1, ... ” durch die AMI-Codierregel ausgeschlossen wird.
- Damit gilt für die Augenöffnung des AMI-Codes:

$$\ddot{o}(T_D) = d_{\text{oben}} - d_{\text{unten}} = g_0 - 3 \cdot g_1 = 0.20 \Rightarrow \ddot{o}_{\text{norm}} = \frac{\ddot{o}(T_D)}{2 \cdot s_0} = 10 \%$$

- Die obere Entscheiderschwelle E_2 sowie die untere Entscheiderschwelle E_1 liegen bei

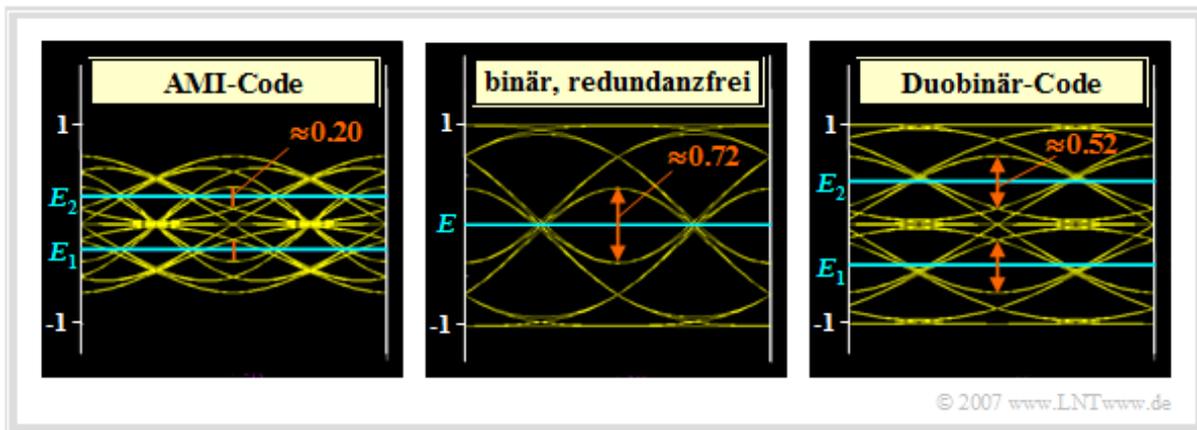
$$E_2 = 1/2 \cdot (d_{\text{oben}} + d_{\text{unten}}) = 1/2 \cdot (g_0 - g_1) = 0.27, \quad E_1 = -0.27.$$

Beim **Duobinär code** (rechte Grafik) tritt die besonders ungünstige alternierende Symbolfolge nicht auf und man erhält für die Augenöffnung sowie die obere Entscheiderschwelle:

$$d_{\text{oben}} = g_0, \quad d_{\text{unten}} = g_1 \Rightarrow \ddot{o}(T_D) = g_0 - g_1 = 0.52 \Rightarrow \ddot{o}_{\text{norm}} = 26 \%, \\ E_2 = 1/2 \cdot (g_0 + g_1) = 0.42, \quad E_1 = -0.42.$$

Hinweis: Augendiagramm und Augenöffnung bei AMI- und Duobinär codierung können ebenfalls mit dem Interaktionsmodul **Augendiagramm und Augenöffnung** angezeigt werden. Die angegebenen Fehlerwahrscheinlichkeiten gelten allerdings nur für den verzerrungsfreien Kanal ($a_* = 0$ dB).

Grenzfrequenzoptimierung bei Pseudoternärcodierung

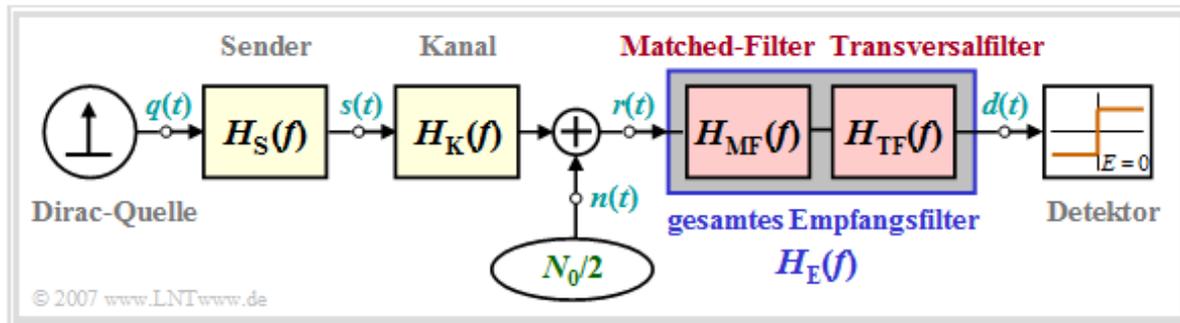


Unter Berücksichtigung eines koaxialen Übertragungskanal und der damit notwendigen Kanalverzerrung sind folgende Aussagen möglich:

- Der AMI-Code führt stets zu einem schlechteren Störabstand als der redundanzfreie Binärcode, wenn der Gesamtfrequenzgang gaußförmig verläuft. Mit der charakteristischen Kabeldämpfung $a_* = 80$ dB beträgt der Störabstandsverlust ca. 11 dB.
- Dieser Verlust ist darauf zurückzuführen, dass trotz ternärer Codierung die Symbolrate gegenüber dem binären Vergleichssystem nicht vermindert wird. Dies hat zur Folge, dass beim AMI-Code bereits eine Grenzfrequenz $f_G \cdot T < 0.36$ zu einem geschlossenen Auge führt.
- Dagegen ergibt sich beim Duobinärcode ein geschlossenes Auge erst ab $f_G \cdot T < 0.22$. Dadurch ist auch die optimale Grenzfrequenz kleiner als beim redundanzfreien Binärsystem. Bei 80 dB Kabeldämpfung ist der Duobinärcode in Kombination mit $f_G \cdot T = 0.28$ um 3.3 dB besser.
- Allerdings ist zu berücksichtigen: Der AMI-Code ist gleichsignalfrei und kann damit auch über einen Telefonkanal mit $H_K(f = 0) = 0$ übertragen werden. Dies ist der entscheidende Grund, dass der AMI-Code zum Beispiel bei **ISDN** (*Integrated Services Digital Network*) eingesetzt wird.
- Alle Ergebnisse in diesem Kapitel gelten jedoch unter der Bedingung $H_K(f = 0) = 1$. Soll ein redundanzfreies Signal oder das duobinär-codierte Signal über einen gleichsignalundurchlässigen Kanal übertragen werden, so ist eine aufwändige Gleichsignalwiedergewinnung erforderlich, die stets ebenfalls mit einer Degradation des S/N-Verhältnisses verbunden ist [ST85].

Struktur des optimalen Nyquistentzerrers

In diesem Abschnitt gehen wir von folgendem Blockschaltbild eines Binärsystems aus.



Hierzu ist anzumerken:

- Die **Diracquelle** liefert die zu übertragende Nachricht (Amplitudenkoeffizienten a_v) in binärer bipolarer Form. Sie wird als redundanzfrei vorausgesetzt.
- Die Sendeimpulsform $g_s(t)$ wird durch den **Senderfrequenzgang** $H_S(f)$ berücksichtigt. Bei allen Beispielen ist $H_S(f) = \text{si}(\pi f T)$ zugrunde gelegt.
- Bei manchen Herleitungen werden Sender und Kanal – hierfür wird meist ein **Koaxialkabel** angenommen – durch den gemeinsamen Frequenzgang $H_{SK}(f) = H_S(f) \cdot H_K(f)$ zusammengefasst.
- Das Empfangsfilter $H_E(f)$ setzt sich multiplikativ aus dem **Matched-Filter** $H_{MF}(f) = H_{SK}^*(f)$ und dem **Transversalfilter** $H_{TF}(f)$ zusammen, zumindest kann es gedanklich so aufgespalten werden.
- Der Gesamtfrequenzgang zwischen der Diracquelle und dem Schwellenwertentscheider soll die **erste Nyquistbedingung** erfüllen. Es muss also gelten:

$$H_S(f) \cdot H_K(f) \cdot H_{MF}(f) \cdot H_{TF}(f) = H_{Nyq}(f).$$

- Mit dieser Bedingung ergibt sich die maximale Augenöffnung (keine Impulsinterferenzen). Deshalb gelten für das **Detektions-SNR** und den **Systemwirkungsgrad** bei binärer Signalisierung:

$$\rho_d = \frac{2 \cdot s_0^2 \cdot T}{\sigma_d^2} = \frac{2 \cdot s_0^2 \cdot T}{N_0} \cdot \frac{1}{\sigma_{d, \text{norm}}^2} \Rightarrow \eta = \frac{\rho_d}{\rho_{d, \text{max}}} = \frac{\rho_d}{2 \cdot s_0^2 \cdot T / N_0} = \frac{1}{\sigma_{d, \text{norm}}^2}.$$

- Die Optimierungsaufgabe beschränkt sich also darauf, das Empfangsfilter $H_E(f)$ so zu bestimmen, dass die **normierte Rauschleistung** vor dem Entscheider den kleinstmöglichen Wert annimmt:

$$\sigma_{d, \text{norm}}^2 = \frac{\sigma_d^2}{N_0/T} = T \cdot \int_{-\infty}^{+\infty} |H_E(f)|^2 df \stackrel{!}{=} \text{Minimum}.$$

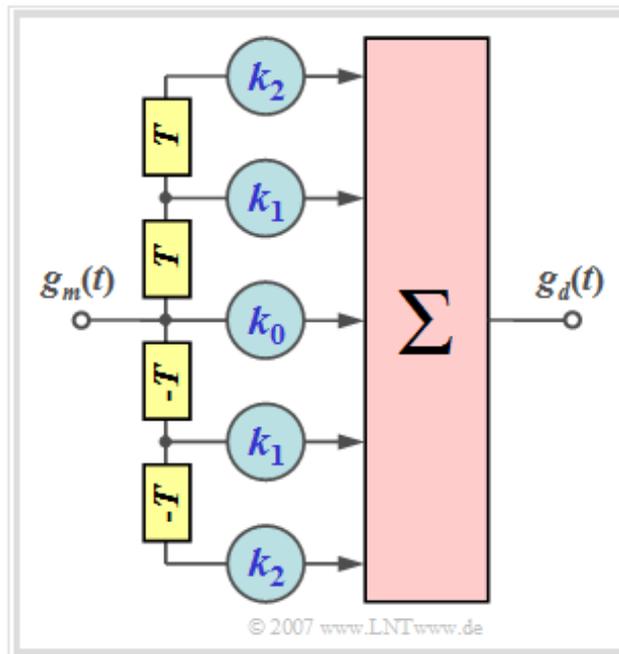
- Wir bezeichnen die Konfiguration als **Optimale Nyquistentzerrung** (ONE). Obwohl diese auch – und besonders effektiv – bei Mehrstufensystemen anwendbar ist, setzen wir zunächst $M = 2$.

Wirkungsweise des Transversalfilters (1)

Verdeutlichen wir uns zunächst die Aufgabe des symmetrischen Transversalfilters

$$H_{TF}(f) \quad \bullet \text{---} \circ \quad h_{TF}(t) = \sum_{\lambda=-N}^{+N} k_{\lambda} \cdot \delta(t - \lambda \cdot T).$$

N gibt die Ordnung des Filters an. Für die Filterkoeffizienten gilt $k_{-\lambda} = k_{\lambda}$. Dieses Filter ist somit durch die Koeffizienten k_0, \dots, k_N vollständig bestimmt. Die Grafik zeigt ein Filter zweiter Ordnung ($N = 2$).



Für den Eingangsimpuls $g_m(t)$ setzen wir ohne Einschränkung der Allgemeingültigkeit voraus, dass dieser

- symmetrisch um $t = 0$ ist (Ausgang des Matched-Filters),
- zu den Zeiten vT und $-vT$ den Wert $g_m(v)$ besitzt.

Damit sind die Eingangsimpulswerte:

$$\dots, g_m(3), g_m(2), g_m(1), g_m(0), g_m(1), g_m(2), g_m(3), \dots$$

Für den Detektionsgrundimpuls $g_d(t)$ am Filterausgang ergeben sich demzufolge zu den Zeitpunkten vT mit den Abkürzungen $g_0 = g_d(t = 0)$, $g_1 = g_d(t = \pm T)$, $g_2 = g_d(t = \pm 2T)$ folgende Werte:

$$\begin{aligned} t = 0: \quad g_0 &= k_0 \cdot g_m(0) + k_1 \cdot 2 \cdot g_m(1) + k_2 \cdot 2 \cdot g_m(2), \\ t = \pm T: \quad g_1 &= k_0 \cdot g_m(1) + k_1 \cdot [g_m(0) + g_m(2)] + k_2 \cdot [g_m(1) + g_m(3)], \\ t = \pm 2T: \quad g_2 &= k_0 \cdot g_m(2) + k_1 \cdot [g_m(1) + g_m(3)] + k_2 \cdot [g_m(2) + g_m(4)]. \end{aligned}$$

Aus diesem System mit drei linear unabhängigen Gleichungen kann man nun die Filterkoeffizienten k_0 , k_1 und k_2 so bestimmen, dass der Detektionsgrundimpuls $g_d(t)$ durch die normierten Stützstellen

$$\dots, g_3, g_2 = 0, g_1 = 0, g_0 = 1, g_1 = 0, g_2 = 0, g_3, \dots$$

vollständig gegeben ist. Auf der nächsten Seite wird die Optimierung der Filterkoeffizienten an einem einfachen Beispiel verdeutlicht.

Wirkungsweise des Transversalfilters (2)

Beispiel: Wir gehen von dem symmetrischen Eingangssignal entsprechend dem oberen Diagramm aus. Mit der Abkürzung $g_m(v) = g_m(\pm v \cdot T)$ gibt es folgende Abtastwerte im Abstand der Symboldauer T :

$$g_m(t) = \exp\left(-\sqrt{2} \cdot |t|/T\right)$$

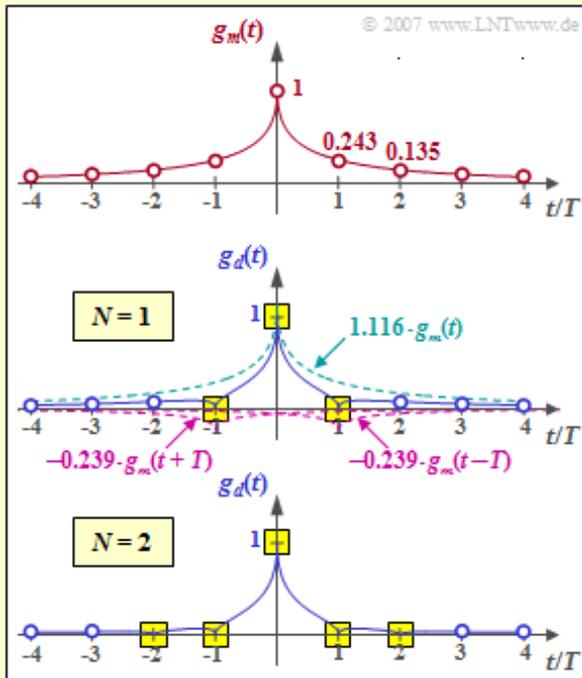
$$\Rightarrow g_m(0) = 1, \quad g_m(1) = 0.243, \quad g_m(2) = 0.135, \quad g_m(3) = 0.086, \quad g_m(4) = 0.059.$$

Für den Ausgangsimpuls soll $g_d(0) = 1$ und $g_d(\pm T) = 0$ gelten. Hierzu eignet sich ein **Laufzeitfilter erster Ordnung** mit den Koeffizienten k_0 und k_1 , die folgende Bedingungen erfüllen müssen:

$$t = \pm T: \quad g_1 = k_0 \cdot 0.243 + k_1 \cdot [1.000 + 0.135] = 0 \quad \Rightarrow \quad k_1 = -0.214 \cdot k_0,$$

$$t = 0: \quad g_0 = k_0 \cdot 1.000 + k_1 \cdot 2 \cdot 0.243 = 1 \quad \Rightarrow \quad 0.896 \cdot k_0 = 1.$$

Daraus erhält man die optimalen Filterkoeffizienten $k_0 = 1.116$ und $k_1 = 0.239$. Das mittlere Diagramm zeigt, dass damit der erste Vorläufer und der erste Nachläufer kompensiert werden können und zugleich $g_d(0) = 1$ gilt (gelbe Hinterlegung). Die weiteren Detektionsgrundimpulswerte (blaue Kreise) sind aber von 0 verschieden und bewirken Impulsinterferenzen.



Das untere Diagramm zeigt, dass mit einem Filter zweiter Ordnung ($N = 2$) Nulldurchgänge bei $\pm T$ und bei $\pm 2T$ erzwungen werden, wenn die Koeffizienten $k_0 = 1.127$, $k_1 = 0.219$ und $k_2 = 0.075$ geeignet gewählt sind. Das Gleichungssystem zur Bestimmung der optimalen Koeffizienten lautet dabei:

$$t = 0: \quad g_0 = k_0 \cdot 1.000 + k_1 \cdot 2 \cdot 0.243 + k_2 \cdot 2 \cdot 0.135 = 1,$$

$$t = \pm T: \quad g_1 = k_0 \cdot 0.243 + k_1 \cdot [1.000 + 0.135] + k_2 \cdot [0.243 + 0.086] = 0,$$

$$t = \pm 2T: \quad g_2 = k_0 \cdot 0.135 + k_1 \cdot [0.243 + 0.086] + k_2 \cdot [1.000 + 0.059] = 0.$$

Die Ergebnisse können wie folgt verallgemeinert werden:

- Mit einem Laufzeitfilter N -ter Ordnung können der Hauptwert $g_d(0)$ zu 1 (normiert) sowie die ersten N Nachläufer und die ersten N Vorläufer zu Null gemacht werden.
- Weitere Vor- und Nachläufer ($|v| > N$) lassen sich so nicht kompensieren. Es ist auch möglich, dass diese außerhalb des Kompensationsbereichs vergrößert werden oder sogar neu entstehen.
- Im Grenzübergang $N \rightarrow \infty$ (in der Praxis heißt das: ein Filter mit sehr vielen Koeffizienten) ist eine vollständige Nyquistentzerrung und damit eine impulsinterferenzfreie Übertragung möglich.

Beschreibung im Frequenzbereich (1)

Die Tatsache, dass sich der **optimale Nyquistentzerrer** multiplikativ aus

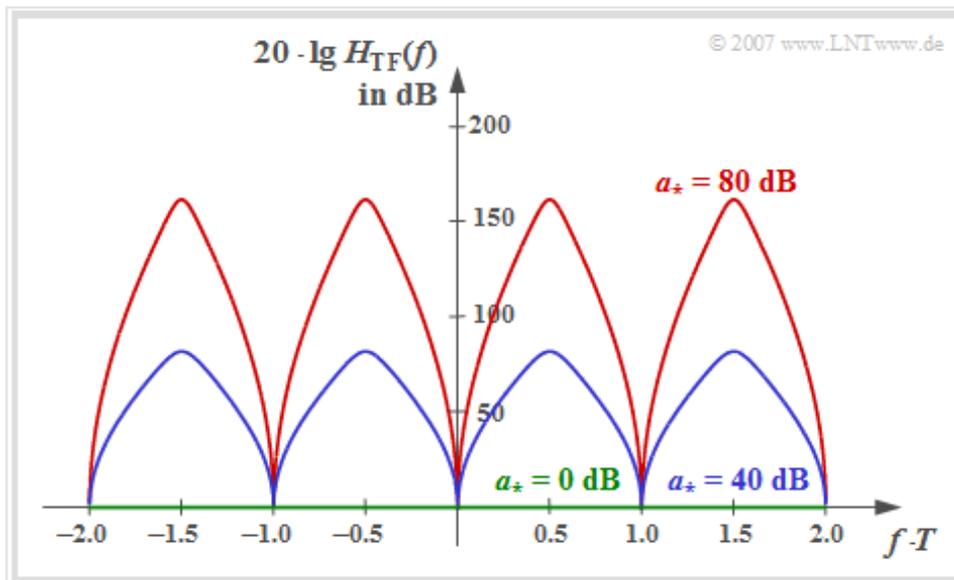
- dem Matched-Filter $H_{MF}(f) = H_S^*(f) \cdot H_K^*(f)$ – also angepasst an den Empfangsgrundimpuls –
- und einem Transversalfilter $H_{TF}(f)$ mit unendlich vielen Filterkoeffizienten

zusammensetzt, folgt aus dem ersten Nyquistkriterium. Durch Anwendung der *Variationsrechnung* erhält man den Frequenzgang des Transversalfilters (siehe [ST85]):

$$H_{TF}(f) = \frac{1}{\sum_{\kappa=-\infty}^{+\infty} |H_{SK}(f - \frac{\kappa}{T})|^2} \quad \text{mit} \quad H_{SK}(f) = H_S(f) \cdot H_K(f).$$

Die Grafik zeigt diesen Verlauf in logarithmierter Form für rechteckförmige NRZ-Sendeimpulse und ein Koaxialkabel mit der charakteristischen Kabeldämpfung

- $a_* = 0 \text{ dB} \Rightarrow$ grüne Null-Linie,
- $a_* = 40 \text{ dB} \Rightarrow$ blauer Funktionsverlauf,
- $a_* = 80 \text{ dB} \Rightarrow$ roter Funktionsverlauf.



Man erkennt aus obiger Gleichung und dieser Skizze:

- $H_{TF}(f)$ ist **reell**, woraus sich die symmetrische Struktur des Transversalfilters ergibt: $k_{-\lambda} = k_{\lambda}$.
- $H_{TF}(f)$ ist eine mit der Frequenz $1/T$ **periodische** Funktion.
- Die Koeffizienten ergeben sich somit aus der **Fourierreihe** (angewandt auf die Spektralfunktion):

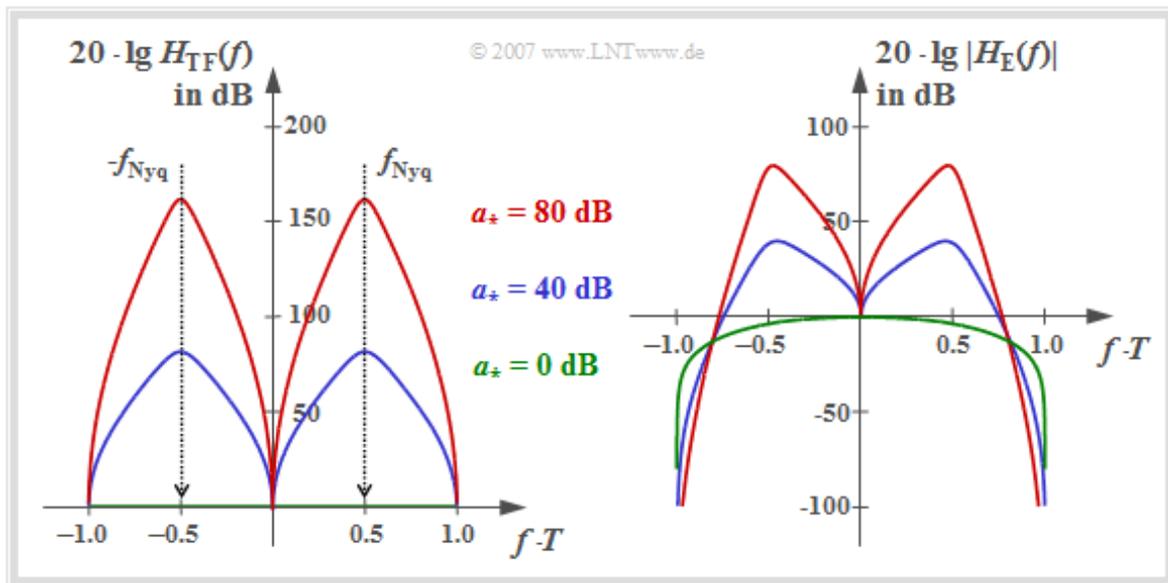
$$k_{\lambda} = T \cdot \int_{-1/(2T)}^{+1/(2T)} \frac{\cos(2\pi f \lambda T)}{\sum_{\kappa=-\infty}^{+\infty} |H_{SK}(f - \kappa/T)|^2} df \Rightarrow H_{TF}(f) = \sum_{\lambda=-\infty}^{+\infty} k_{\lambda} \cdot e^{-j2\pi f \lambda T}.$$

Die Bildbeschreibung wird auf der nächsten Seite fortgesetzt.

Beschreibung im Frequenzbereich (2)

Die linke Grafik zeigt den Verlauf $20 \cdot \lg H_{TF}(f)$ im Bereich $|f| \leq 1/T$. Rechts ist der Frequenzgang $20 \cdot \lg |H_E(f)|$ des gesamten Empfangsfilters einschließlich Matched-Filter dargestellt. Es gilt:

$$H_E(f) = H_{MF}(f) \cdot H_{TF}(f) = \frac{H_{SK}^*(f)}{\sum_{\kappa=-\infty}^{+\infty} |H_{SK}(f - \kappa/T)|^2}$$



Zu diesen Darstellungen ist anzumerken:

- Der Transversalfilter-Frequenzgang $H_{TF}(f)$ ist symmetrisch zur Nyquistfrequenz $f_{Nyq} = 1/(2T)$. Diese Symmetrie ist beim Empfangsfilter-Gesamtfrequenzgang $H_E(f)$ nicht mehr gegeben.
- Die Maxima der Frequenzgänge $H_{TF}(f)$ und $|H_E(f)|$ hängen signifikant von der charakteristischen Kabeldämpfung ab. Es gilt:

$$a_* = 40 \text{ dB: } \text{Max}[H_{TF}(f)] \approx 80 \text{ dB, } \text{Max}[|H_E(f)|] \approx 40 \text{ dB,}$$

$$a_* = 80 \text{ dB: } \text{Max}[H_{TF}(f)] \approx 160 \text{ dB, } \text{Max}[|H_E(f)|] \approx 80 \text{ dB.}$$

Für $a_* = 0 \text{ dB}$ (idealer Kanal) kann auf das Transversalfilter verzichtet werden und es gilt, wie bereits im **Kapitel 1.2** hergeleitet:

$$H_E(f) = H_S(f) = \text{si}(\pi fT).$$

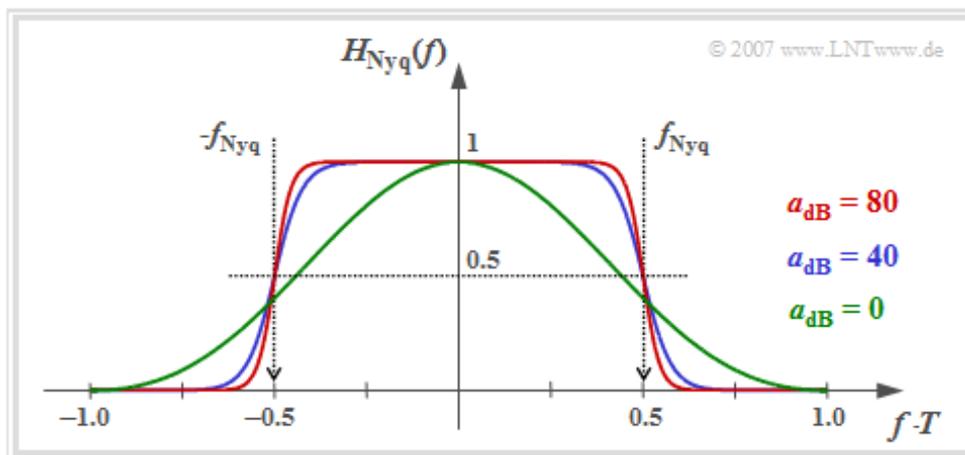
Approximation des optimalen Nyquistentzerrers

Betrachten wir nun den Gesamtfrequenzgang zwischen der Diracquelle und dem Entscheider. Dieser setzt sich multiplikativ aus den Frequenzgängen von Sender, Kanal und Empfänger zusammen. Entsprechend der Herleitung muss der Gesamtfrequenzgang die Nyquistbedingung erfüllen:

$$H_{\text{Nyq}}(f) = H_S(f) \cdot H_K(f) \cdot H_E(f) = \frac{|H_{\text{SK}}(f)|^2}{\sum_{\kappa=-\infty}^{+\infty} |H_{\text{SK}}(f - \kappa/T)|^2}$$

Die Grafik zeigt folgende **Eigenschaften des optimalen Nyquistfilters**:

- Ist die Kabeldämpfung hinreichend groß ($a_* > 10$ dB), so kann der Gesamtfrequenzgang mit sehr guter Näherung durch einen **Cosinus-Rolloff-Tiefpass** beschrieben werden.
- Je größer a_* ist, desto kleiner ist der Rolloff-Faktor und um so steiler verläuft der Flankenabfall. Für die charakteristische Kabeldämpfung $a_* = 40$ dB ergibt sich $r \approx 0.4$, für 80 dB ist $r \approx 0.18$.
- Oberhalb der Frequenz $f_{\text{Nyq}} \cdot (1 + r)$ besitzt $H_{\text{Nyq}}(f)$ keine Anteile. Bei idealem Kanal – also für $a_* = 0$ dB – reicht $H_{\text{Nyq}}(f) = \text{sinc}^2(\pi f T)$ allerdings theoretisch bis ins Unendliche (grüne Kurve).



Mit dem folgenden Interaktionsmodul können Sie sich den Cosinus-Rolloff-Tiefpass im Frequenz- und Zeitbereich verdeutlichen:

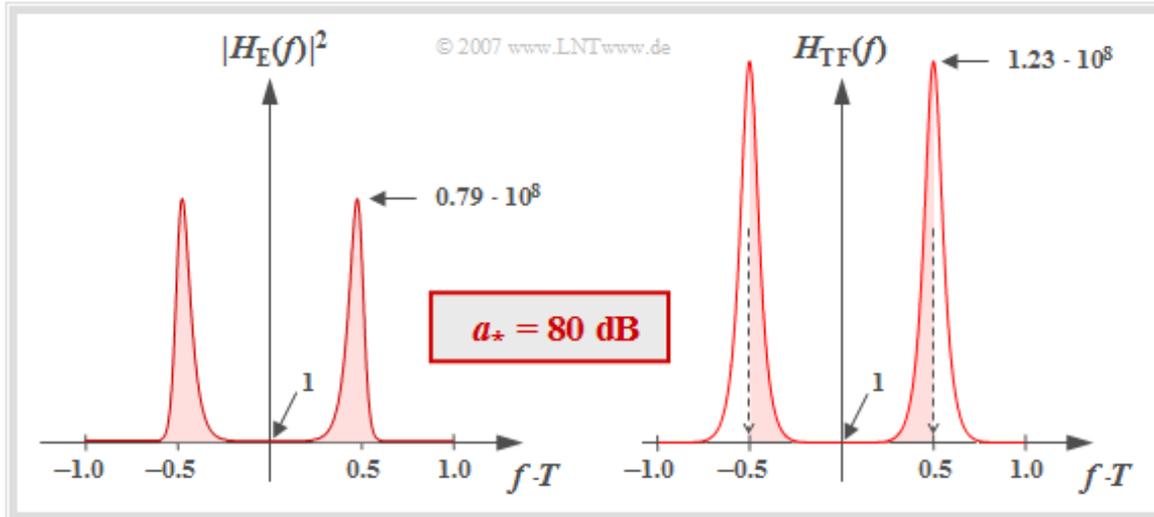
Tiefpässe im Frequenz- und Zeitbereich

Berechnung der normierten Störleistung

Betrachten wir nun noch die (normierte) Störleistung am Entscheider. Für diese gilt:

$$\sigma_{d, \text{norm}}^2 = \frac{\sigma_d^2}{N_0/(2T)} = T \cdot \int_{-1/(2T)}^{+1/(2T)} |H_E(f)|^2 df.$$

Das linke Bild zeigt $|H_E(f)|^2$ im linearen Maßstab für die charakteristische Kabeldämpfung $a_* = 80$ dB.



Beachten Sie, dass $|H_E(f=0)| = 1$ ist. Da die Frequenz auf $1/T$ normiert wurde, entspricht die normierte Störleistung genau der (rot hinterlegten) Fläche unter dieser Kurve. Die numerische Auswertung ergibt:

$$\sigma_{d, \text{norm}}^2 = 1.68 \cdot 10^7 \Rightarrow 10 \cdot \lg \sigma_{d, \text{norm}}^2 \approx 72.25 \text{ dB}.$$

Es kann gezeigt werden, dass die normierte Störleistung auch mit dem Transversalfilter–Frequenzgang $H_{TF}(f)$ berechnet werden kann, wie in der rechten Grafik dargestellt:

$$\sigma_{d, \text{norm}}^2 = T \cdot \int_{-1/(2T)}^{+1/(2T)} H_{TF}(f) df \quad (= k_0).$$

Die roten Flächen sind in beiden Bildern exakt gleich. Man erkennt auch, dass der mittlere Koeffizient k_0 gleich der normierten Störleistung ist. In der zweiten Spalte der nachfolgenden Tabelle ist $10 \cdot \lg(k_0)$ in Abhängigkeit der charakteristischen Kabeldämpfung angegeben. Aufgrund der gewählten Normierung gilt diese Tabelle auch für redundanzfreie Mehrstufensysteme; M bezeichnet hierbei die Stufenzahl.

$a_*/\log_2(M)$	$10 \cdot \lg(k_0)$	k_1/k_0	k_2/k_0	k_3/k_0	k_4/k_0	k_5/k_0	k_6/k_0	k_7/k_0
0 dB	0 dB	0	0	0	0	0	0	0
20 dB	16.6 dB	-0.590	0.168	-0.060	0.013	-0.007	0.001	-0.002
40 dB	34.6 dB	-0.793	0.438	-0.212	0.100	-0.047	0.022	-0.010
60 dB	53.3 dB	-0.878	0.617	-0.381	0.224	-0.129	0.074	-0.042
80 dB	72.3 dB	-0.919	0.727	-0.518	0.348	-0.228	0.148	-0.095

Die Koeffizienten k_1, k_2, k_3, \dots des Transversalfilters weisen für $a_* \neq 0$ alternierende Vorzeichen auf. Für $a_* = 40$ dB sind vier Koeffizienten betragsmäßig größer als $k_0/10$, für $a_* = 80$ dB sogar sieben.

Vergleich anhand des Systemwirkungsgrades

Für einen Systemvergleich eignet sich der **Systemwirkungsgrad**, der das erreichbare Detektions-SNR ρ_d in Bezug zum maximalen SNR $\rho_{d, \max}$ setzt, das allerdings nur bei idealem Kanal $H_K(f) = 1$ erreichbar ist. Für den Systemwirkungsgrad gilt bei M -stufiger Übertragung und optimaler Nyquistentzerrung:

$$\eta = \frac{\rho_d}{s_0^2 \cdot T / N_0} = \frac{\log_2 M}{(M - 1)^2 \cdot k_0}$$

Die (normierte) Störleistung k_0 kann aus der **Tabelle** auf der letzten Seite abgelesen werden. Beachten Sie die Normierung der charakteristischen Kabeldämpfung a_* in der ersten Spalte.

Die folgende Tabelle aus [ST85] ermöglicht einen Systemvergleich für $a_* = 80$ dB. Verglichen werden

- der **gaußförmige Gesamtfrequenzgang** (GTP) entsprechend Kapitel 3.4,
- der **optimale Nyquistentzerrer** (ONE) entsprechend Kapitel 3.5.

© 2007 www.LNTwww.de	Systemwirkungsgrad η				
	$M=2$	$M=3$	$M=4$	$M=8$	$M=16$
Gaußtiefpass (GTP) mit optimierter Grenzfrequenz	-78.5 dB	-75.8 dB	-74.8 dB	-74.2 dB	-75.1 dB
Optimaler Nyquistentzerrer (ONE)	-72.3 dB	-60.7 dB	-56.6 dB	-52.5 dB	-52.1 dB

Das Ergebnis dieses Vergleichs kann wie folgt zusammengefasst werden:

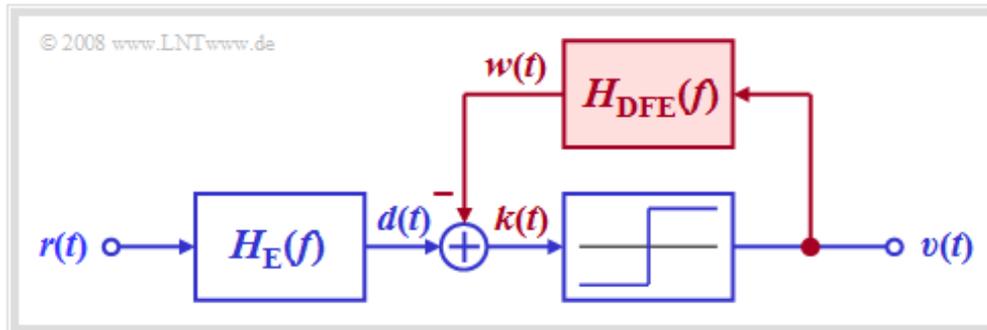
- Im binären Fall ($M = 2$) ist das impulsinterferenzfreie System (ONE) um etwa 6 dB besser als das impulsinterferenzbehaftete System (GTP).
- Wendet man die optimale Nyquistentzerrung bei Mehrstufensystemen an, so ist gegenüber „GTP“ ein weiterer, deutlicher Störabstandsgewinn möglich. Für $M = 4$ ist dieser Gewinn etwa 18.2 dB.
- Das schmalbandige GTP-System kann allerdings deutlich verbessert werden, wenn man einen Empfänger mit Entscheidungsrückkopplung verwendet. Dieser wird im Kapitel 3.6 behandelt.

Hinweis: Alle Ergebnisse von Kapitel 3.5 lassen sich mit folgendem Interaktionsmodul nachvollziehen:

Lineare Nyquistentzerrung

Prinzip und Blockschaltbild

Eine Möglichkeit zur Verminderung von Impulsinterferenzen bietet die **Entscheidungsrückkopplung** (engl.: *Decision Feedback Equalization* – abgekürzt DFE). In der deutschsprachigen Literatur wird diese manchmal auch als *Quantisierte Rückkopplung* (QR) bezeichnet.



Die Grafik zeigt den entsprechenden Empfänger. Man erkennt anhand des Blockschaltbildes:

- Ohne die rot eingezeichnete Signalarückführung ergäbe sich ein herkömmlicher Digitalempfänger mit Schwellenwertentscheidung entsprechend **Kapitel 3.3**. Für die nachfolgende Beschreibung wird wieder angenommen, dass sich das gesamte Empfangsfilter $H_E(f)$ aus dem Kanalentzerrer $1/H_K(f)$ und einem Gaußtieffpass $H_G(f)$ zur Rauschleistungsbegrenzung zusammensetzt.
- Beim Empfänger mit Entscheidungsrückkopplung wird vom rechteckförmigen Ausgangssignal $v(t)$ über ein lineares Netzwerk mit dem Frequenzgang $H_{DFE}(f)$ ein **Kompensationssignal** $w(t)$ gewonnen und an den Eingang des Schwellenwertentscheiders zurückgeführt.
- Dieses Signal $w(t)$ wird vom vorentzerrten Signal $d(t)$ subtrahiert. Bei geeigneter Dimensionierung des Rückkopplungsnetzwerkes weist somit das **korrigierte Signal** $k(t) = d(t) - w(t)$ keine (oder zumindest deutlich geringere) Impulsnachläufer auf als das Signal $d(t)$. Die Impulsvorläufer können dagegen aus Kausalitätsgründen nicht beeinflusst werden.
- Da bei diesem Empfänger mit Entscheidungsrückkopplung das Kompensationssignal $w(t)$ vom rauschfreien Sinkensignal $v(t)$ abgeleitet wird, ist die Signalentzerrung nicht mit einer Erhöhung der Rauschleistung verbunden wie bei linearer Entzerrung. Vielmehr besitzt das korrigierte Signal $k(t)$ den gleichen Rauscheffektivwert σ_d wie das Signal $d(t)$.

Hinweis: Die Signalverläufe dieses nichtlinearen Entzerrungsverfahrens „DFE“ sowie die zugehörigen Fehlerwahrscheinlichkeiten – gültig für einen verzerrungsfreien Kanal – können mit dem folgenden Interaktionsmodul angezeigt werden:

Entscheidungsrückkopplung

Ideale Entscheidungsrückkopplung

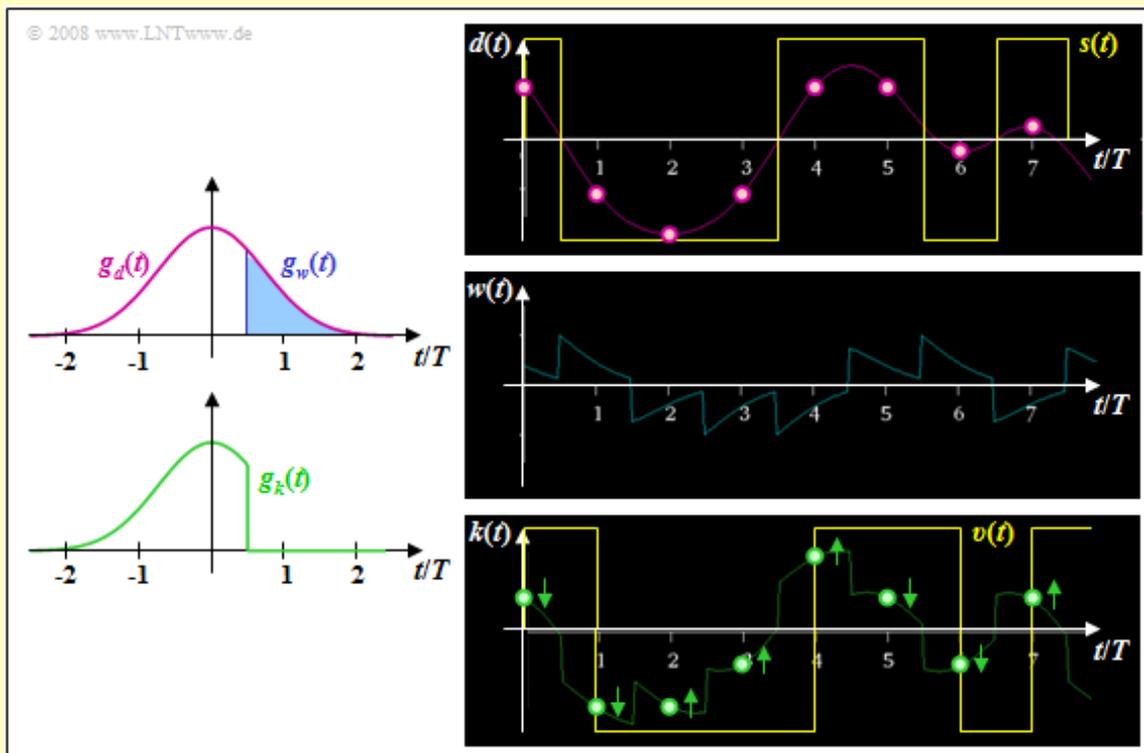
Wir behandeln zunächst die ideale DFE–Realisierung anhand der Grundimpulse.

Definition: Eine **ideale Entscheidungsrückkopplung** liegt vor, wenn am Entscheider der folgende Grundimpuls anliegt:

$$g_k(t) = \begin{cases} g_d(t) & \text{für } t < T_D + T_V, \\ 0 & \text{für } t \geq T_D + T_V. \end{cases}$$

Das bedeutet, dass im Idealfall der Kompensationsgrundimpuls $g_w(t)$ den linear vorentzerrten Impuls $g_d(t)$ für alle Zeiten $t > T_D + T_V$ exakt nachbilden muss. Die aus Realisierungsgründen erforderliche Verzögerungszeit T_V muss stets kleiner als die Symboldauer T sein; im Folgenden gelte stets $T_V = T/2$.

Beispiel: Der Gesamtfrequenzgang $H_K(f) \cdot H_E(f) = H_G(f)$ sei gaußförmig mit der Grenzfrequenz $f_G = 0.3/T$. Bei NRZ–Rechteckimpulsen ergibt sich dann der skizzierte Detektionsgrundimpuls $g_d(t)$.



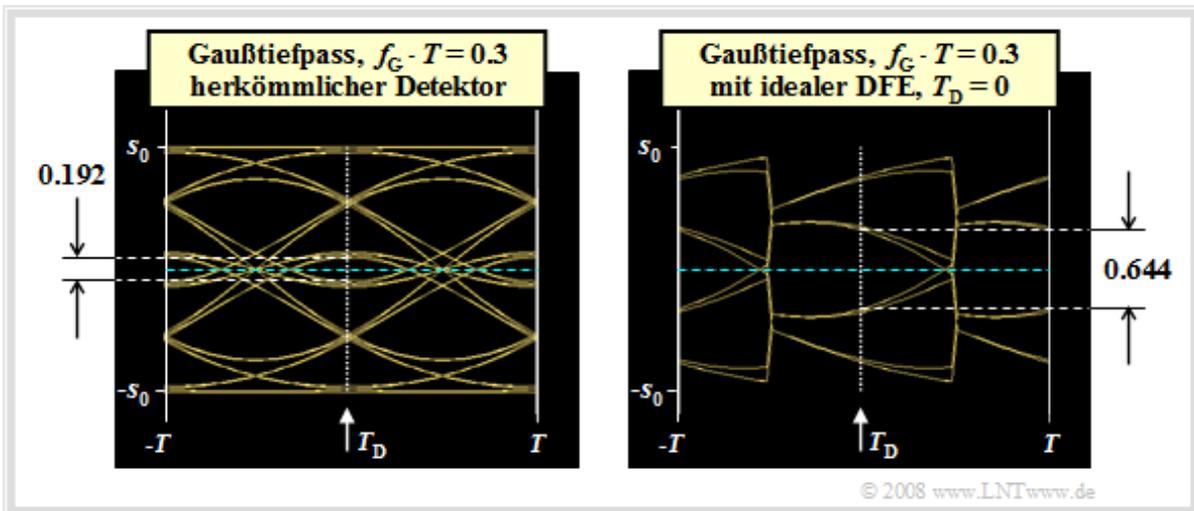
Links dargestellt sind auch die Grundimpulse $g_w(t)$ und $g_k(t)$ bei idealer Entscheidungsrückkopplung, wobei der Detektionszeitpunkt $T_D = 0$ und die Verzögerungszeit $T_V = T/2$ zugrunde liegen.

Die rechten Bilder aus [Söd01] – alle ohne Berücksichtigung des Rauschens – machen deutlich, dass durch die Kompensation aller Impulsnachläufer mittels des Korrektursignals $w(t)$ die Abstände der Nutzabstandswerte $d_S(vT)$ von der Entscheiderschwelle $E = 0$ verändert werden. Besonders geringe Abstände wie beispielsweise zu den Zeitpunkten $t = 6T$ und $t = 7T$ werden deutlich vergrößert und damit deren Fehlerwahrscheinlichkeiten stark verringert (Pfeile weggehend von der Schwelle).

Dagegen werden die im Signal $d(t)$ weit vom Schwellenwert $E = 0$ entfernten Detektionsabstastwerte zur Schwelle hin verschoben und deren Verfälschungswahrscheinlichkeit leicht erhöht. Dies erkennt man zum Beispiel für den Zeitpunkt $t = 5T$.

Augenöffnung und Fehlerwahrscheinlichkeit bei DFE

Betrachten wir nun die Augendiagramme ohne DFE (linke Grafik) und mit idealer DFE (rechte Grafik).



Dabei wird von den gleichen Voraussetzungen wie auf der letzten Seite ausgegangen, so dass folgende Grundimpulswerte vorliegen:

$$g_0 = g_d(t = 0) = 0.548 \cdot s_0, \quad g_1 = g_d(t = T) = 0.214 \cdot s_0 = g_{-1},$$

$$g_2 = g_d(t = 2T) = 0.012 \cdot s_0 = g_{-2}, \quad g_3 = g_{-3} = \dots \approx 0.$$

Diese beiden Augendiagramme können wie folgt interpretiert werden:

- Beim herkömmlichen Empfänger (ohne DFE) gilt bei binärer bipolarer redundanzfreier Codierung unter Berücksichtigung der Symmetrie:

$$\begin{aligned} \ddot{o}(T_D = 0) &= 2 \cdot [g_0 - |g_{-1}| - |g_{-2}| - |g_1| - |g_2|] = \\ &= 2 \cdot [g_0 - 2 \cdot g_1 - 2 \cdot g_2] = 0.192 \cdot s_0. \end{aligned}$$

- Dagegen werden bei idealer DFE die beiden Nachläufer g_1 und g_2 vollständig kompensiert und man erhält für die vertikale Augenöffnung:

$$\ddot{o}(T_D = 0) = 2 \cdot [g_0 - g_{-1} - g_{-2}] = 0.644 \cdot s_0.$$

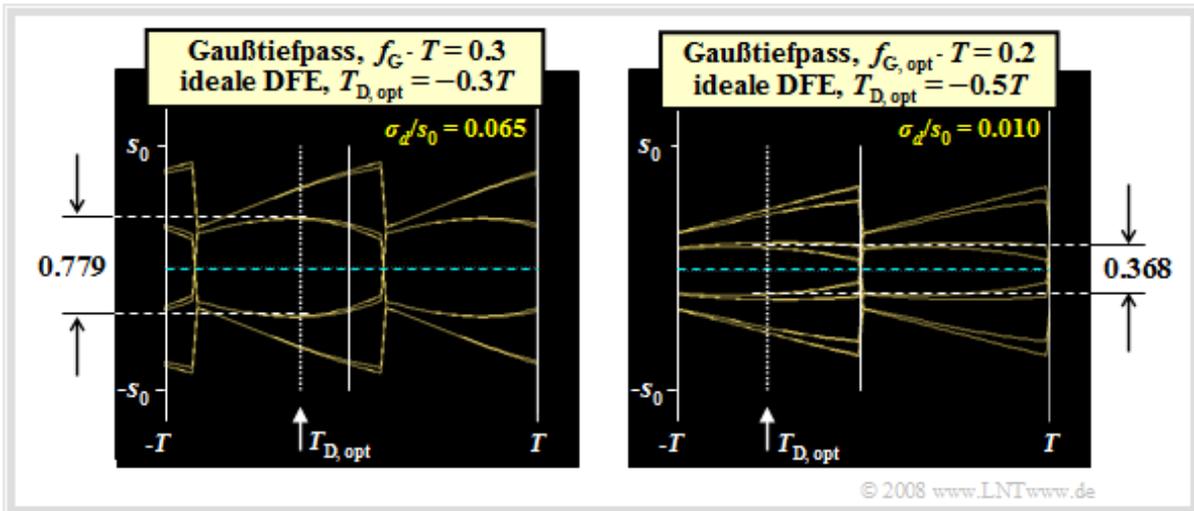
- Da das Korrektursignal $w(t)$ aus dem entschiedenem und damit rauschfreien Signal $v(t)$ abgeleitet wird, wird der Rauscheffektivwert durch die Entscheidungsrückkopplung nicht verändert. Der Störabstandsgewinn durch die DFE ist somit im betrachteten Beispiel gleich

$$G_{DFE} = 20 \cdot \lg \frac{0.644}{0.192} \approx 10.5 \text{ dB}.$$

Bei einem Koaxialkabel mit charakteristischer Kabeldämpfung $a_* = 80 \text{ dB}$ und $10 \cdot \lg E_B/N_0 = 80 \text{ dB}$ bedeutet dieser Störabstandsgewinn beispielsweise, dass die ungünstigste Fehlerwahrscheinlichkeit p_U durch die DFE von 7% auf ca. $4 \cdot 10^{-7}$ verkleinert wird – eine durchaus beachtenswerte Verbesserung.

Optimierung eines Übertragungssystems mit DFE

Die letzte Seite hat deutlich gemacht, dass die Entscheidungsrückkopplung bereits dann einen enormen Störabstandsgewinn bewirkt, wenn von einer festen Grenzfrequenz f_G und dem Detektionszeitpunkt $T_D = 0$ ausgegangen wird. Das System lässt sich aber weiter verbessern, wenn die beiden Parameter f_G und T_D gemeinsam optimiert werden.



Betrachten wir die Augendiagramme ohne Rauschen für $f_G \cdot T = 0.3$ (links) und $f_G \cdot T = 0.2$ (rechts). Für die nachfolgenden Berechnungen werden weiterhin die charakteristische Kabeldämpfung $a_* = 80$ dB sowie der AWGN-Parameter $10 \cdot \lg E_B/N_0 = 80$ dB (mit $E_B = s_0^2 \cdot T$) vorausgesetzt, so dass sich der normierte Rauscheffektivwert zu $\sigma_d/s_0 = 0.065$ (für $f_G \cdot T = 0.3$) bzw. $\sigma_d/s_0 = 0.010$ (für $f_G \cdot T = 0.2$) ergibt. Die Optimierungsergebnisse lassen sich wie folgt zusammenfassen:

- Mit $f_G \cdot T = 0.3$ kann durch Verschiebung des Detektionszeitpunktes auf $T_{D,opt} = -0.3T$ die Augenöffnung auf $\delta(T_D) = 0.779 \cdot s_0$ vergrößert werden. Daraus resultiert gegenüber $T_D = 0$ (vergleiche letzte Seite) ein weiterer Störabstandsgewinn von $20 \cdot \lg(0.779/0.644) \approx 1.65$ dB und die Fehlerwahrscheinlichkeit ergibt sich nun zu $p_U \approx 1.3 \cdot 10^{-9}$ (gegenüber $4 \cdot 10^{-7}$).
- Bei einem DFE-Empfänger kann man zusätzlich die Grenzfrequenz weiter herabsetzen. So ergibt sich mit $f_G \cdot T = 0.2$ und $T_D = 0$ die zwar kleine, aber immerhin von 0 verschiedene Augenöffnung $\delta(T_D) = 0.152 \cdot s_0$, die zusammen mit dem sehr günstigen Rauscheffektivwert $\sigma_d/s_0 = 0.010$ zum (ungünstigsten) Störabstand 17.6 dB und zur Fehlerwahrscheinlichkeit $p_U = 1.6 \cdot 10^{-14}$ führt.
- Durch Kombination der Grenzfrequenz $f_G \cdot T = 0.2$ mit dem Detektionszeitpunkt $T_D = -0.5T$ erhält man schließlich die bei den getroffenen Voraussetzungen optimale Systemkonfiguration mit der Augenöffnung $\delta(T_D) = 0.368 \cdot s_0$ und dem (ungünstigsten) Störabstand $10 \cdot \lg \rho_U = 25.3$ dB. Die Fehlerwahrscheinlichkeit ist damit (praktisch) gleich 0.

Realisierungsaspekte der Entscheidungsrückkopplung (1)

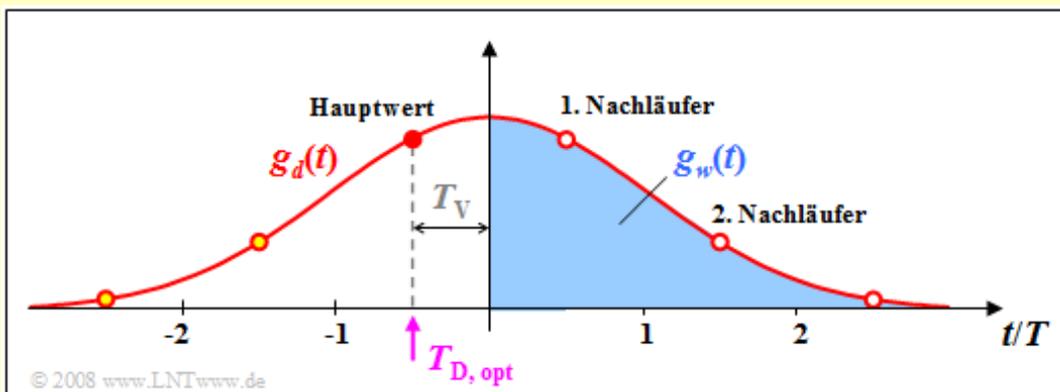
Als ein wesentliches Ergebnis von Kapitel 3.5 und Kapitel 3.6 empfiehlt sich folgende Vorgehensweise: Für ein Übertragungssystem über Kupferleitungen (Koaxialkabel, Zweidrahtleitung) sind aufgrund des erreichbaren Signal–zu–Rauschabstandes am Entscheider folgende Systemvarianten besonders geeignet:

- ein **Mehrstuflensystem** (z.B. $M = 4$) und die optimale **Nyquistentzerrung** zur Kompensation der starken Impulsinterferenzen, hervorgerufen durch die linearen Kanalverzerrungen.
- ein **Binärsystem** mit relativ kleiner Bandbreite des Gesamtfrequenzganges $H_G(f) = H_K(f) \cdot H_E(f)$ und ein nichtlinearer Detektor mit **Entscheidungsrückkopplung**.

Beide Systemvarianten liefern bei idealisierten Bedingungen vergleichbar gute Resultate. Zu beachten ist allerdings, dass es bei beiden Systemen durch Realisierungsungenauigkeiten zu großen Degradationen kommen kann, die hier am Beispiel des DFE–Systems genannt werden:

- Da über das Fernsprechnetz kein Gleichsignal übertragen werden kann, für unsere Berechnungen aber $H_K(f = 0) = 1$ angenommen wird, ist am Empfänger eine *Gleichsignalwiedergewinnung* erforderlich. Diese Aussage trifft in gleicher Weise für das quaternäre Nyquistsystem zu.
- Beim DFE–System muss der Kompensationsimpuls den vorentzerrten Grundimpuls $g_d(t)$ exakt nachbilden. Dies ist insbesondere dann schwierig, wenn $g_d(t)$ sehr breit ist (kleine Grenzfrequenz, zum Beispiel $f_G \cdot T = 0.2$) und die Optimierung den Detektionszeitpunkt $T_{D, \text{opt}} = -T/2$ liefert.
- Kommt es aufgrund eines sehr großen Rauschwertes zu einer Fehlentscheidung, so werden auch die nachfolgenden Symbole mit großer Wahrscheinlichkeit verfälscht. Allerdings gibt es immer wieder Symbolfolgen, die diese Fehlerfortpflanzung unterbrechen.

Beispiel: Die Grafik zeigt den Grundimpuls $g_d(t)$ für die Grenzfrequenz $f_G \cdot T = 0.2$ (rote Kurve) und den Kompensationsimpuls $g_w(t)$ für $T_D = -T/2$ (blau gefüllt). Hierbei ist wieder eine Verzögerungszeit $T_V = T/2$ zwischen Entscheidung und Beginn der Signalkorrektur berücksichtigt.



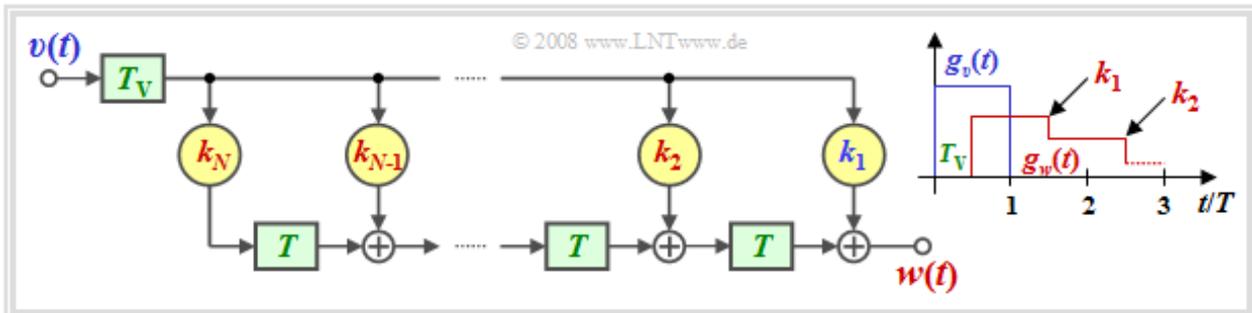
Man erkennt, dass bereits für $T_D = -T/2$ der erste Nachläufer $g_d(T_D + T) = g_d(T/2)$ genau so groß ist wie der Hauptwert $g_d(T_D) = g_d(-T/2)$. Gelingt es nicht, tatsächlich alle Nachläufer zu kompensieren, so ergibt sich schnell ein geschlossenes Auge und damit die Fehlerwahrscheinlichkeit $p_U = 50\%$.

Realisierungsaspekte der Entscheidungsrückkopplung (2)

Für eine schaltungstechnische Realisierung genügt es, wenn der korrigierte Grundimpuls $g_k(t)$ lediglich zu den äquidistanten Detektionszeitpunkten $T_D + \nu \cdot T$ zu Null wird. Eine Realisierungsmöglichkeit stellt somit ein Laufzeitfilter gemäß der nachfolgenden Grafik dar,

- dessen Ordnung N (Anzahl der Filterkoeffizienten), und
- dessen Filterkoeffizienten k_ν (mit $\nu = 1, \dots, N$)

durch den Grundimpuls $g_d(t)$ sowie den Detektionszeitpunkt T_D festgelegt sind.



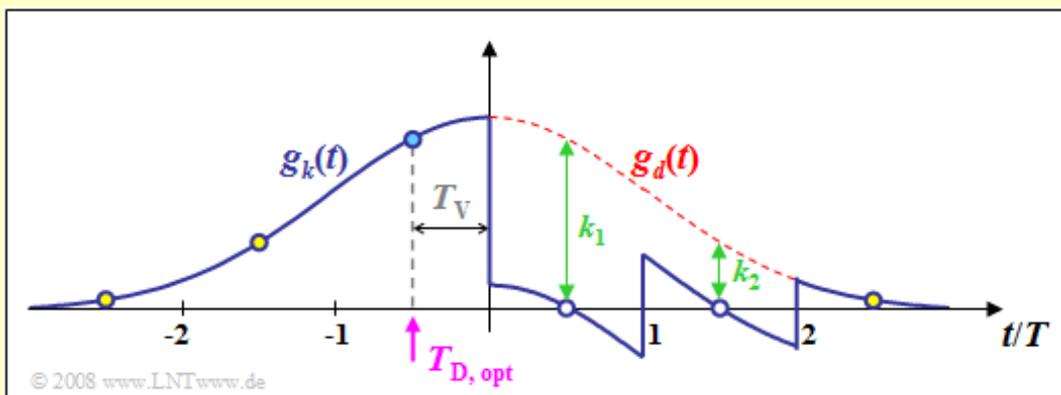
Diese DFE-Realisierung weist folgende Eigenschaften auf

- Da das Ausgangssignal $v(t)$ rechteckförmig ist, ist der Kompensationsimpuls $g_w(t)$ treppenförmig.
- Bei richtiger Dimensionierung der Filterkoeffizienten k_ν gilt für $\nu = 1, \dots, N$:

$$g_w(T_D + \nu \cdot T) = g_d(T_D + \nu \cdot T) \Rightarrow g_k(T_D + \nu \cdot T) = 0.$$

- Zum Detektionszeitpunkt T_D ergibt sich die genau gleiche vertikale Augenöffnung wie bei idealer DFE. Nachteilig ist eine kleinere horizontale Augenöffnung.

Beispiel: Die Grafik zeigt die Grundimpulse $g_d(t)$ und $g_k(t)$ bei der Entscheidungsrückkopplung mit einem Laufzeitfilter zweiter Ordnung. Es gelten die gleichen Voraussetzungen wie im Beispiel auf der letzten Seite: $f_G \cdot T = 0.2$, $T_D = -T/2$, $T_V = T/2$.

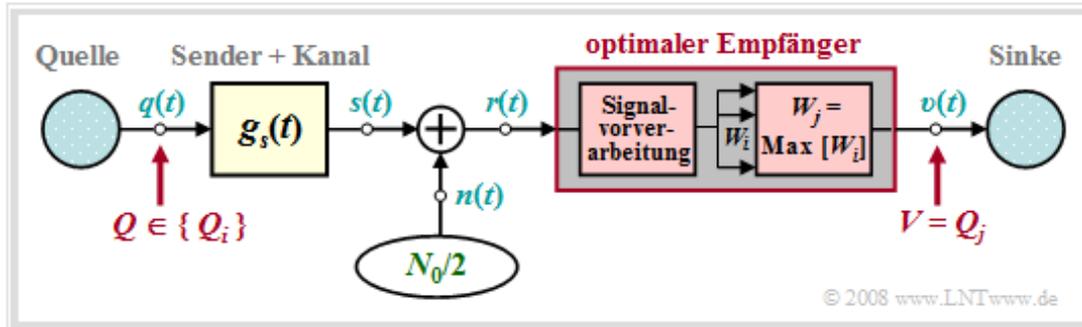


Wegen der Ordnung $N = 2$ werden hier allerdings nur die beiden ersten Nachläufer $g_d(0.5T)$ und $g_d(1.5T)$ kompensiert. Der dritte Nachläufer $g_d(2.5T)$ könnte durch einen weiteren Filterkoeffizienten k_3 zu 0 gemacht werden. Dagegen können die Impulsvorläufer $g_d(-1.5T)$ und $g_d(-2.5T)$ prinzipiell nicht kompensiert werden.

Betrachtetes Szenario im Kapitel 3.7

Alle bisher beschriebenen Digitalempfänger treffen stets symbolweise Entscheidungen. Werden dagegen mehrere Symbole gleichzeitig entschieden, so können bei der Detektion statistische Bindungen zwischen den Empfangssignalabtastrwerten berücksichtigt werden, was eine geringere Fehlerwahrscheinlichkeit zur Folge hat – allerdings auf Kosten einer zusätzlichen Laufzeit.

In diesem – teilweise auch im nächsten Kapitel – wird von folgendem Übertragungsmodell ausgegangen:



Gegenüber den letzten Kapiteln 3.5 und 3.6 ergeben sich folgende Unterschiede:

- $Q \in \{Q_i\}$ mit $i = 0, \dots, M-1$ bezeichnet eine zeitlich begrenzte Quellensymbolfolge $\langle q_v \rangle$, deren Symbole vom optimalen Empfänger gemeinsam entschieden werden sollen.
- Beschreibt Q eine Folge von N redundanzfreien Binärsymbolen, so ist $M = 2^N$ zu setzen. Dagegen gibt M bei symbolweiser Entscheidung die Stufenzahl der digitalen Quelle an.
- Im obigen Modell werden eventuelle Kanalverzerrungen dem Sender hinzugefügt und sind somit bereits im Grundimpuls $g_s(t)$ und im Signal $s(t)$ enthalten. Diese Maßnahme dient lediglich einer einfacheren Darstellung und stellt keine Einschränkung dar.
- Der optimale Empfänger sucht unter Kenntnis des aktuell anliegenden Empfangssignals $r(t)$ aus der Menge $\{Q_0, \dots, Q_{M-1}\}$ der möglichen Quellensymbolfolgen die am wahrscheinlichsten gesendete Folge $\{Q_j\}$ und gibt diese als Sinkensymbolfolge V aus.
- Vor dem eigentlichen Entscheidungsalgorithmus muss durch eine geeignete Signalvorverarbeitung aus dem Empfangssignal $r(t)$ für jede mögliche Folge Q_i ein Zahlenwert W_i abgeleitet werden. Je größer W_i ist, desto größer ist die Rückschlusswahrscheinlichkeit, dass Q_i gesendet wurde.
- Die Signalvorverarbeitung muss für die erforderliche Rauschleistungsbegrenzung und – bei starken Kanalverzerrungen – für eine ausreichende Vorentzerrung der entstandenen Impulsinterferenzen sorgen. Außerdem beinhaltet die Vorverarbeitung auch die Abtastung zur Zeitdiskretisierung.

MAP– und Maximum–Likelihood–Entscheidungsregel (1)

Man bezeichnet den (uneingeschränkt) optimalen Empfänger als MAP–Empfänger, wobei „MAP“ für „Maximum–a–posteriori“ steht.

Definition: Der **MAP–Empfänger** ermittelt die M Rückschlusswahrscheinlichkeiten $\Pr(Q_i|r(t))$ und setzt seine Ausgangsfolge V gemäß der Entscheidungsregel ($i = 0, \dots, M - 1, i \neq j$):

$$\Pr(Q_j|r(t)) > \Pr(Q_i|r(t)).$$

Die **Rückschlusswahrscheinlichkeit** $\Pr(Q_i|r(t))$ gibt an, mit welcher Wahrscheinlichkeit die Folge Q_i gesendet wurde, wenn das Empfangssignal $r(t)$ am Entscheider anliegt. Mit dem Satz von **Bayes** kann diese Wahrscheinlichkeit wie folgt berechnet werden:

$$\Pr(Q_i|r(t)) = \frac{\Pr(r(t)|Q_i) \cdot \Pr(Q_i)}{\Pr(r(t))}.$$

Die MAP–Entscheidungsregel lässt sich somit wie folgt umformulieren bzw. vereinfachen. Man setze die Sinkensymbolfolge $V = Q_j$, falls für alle $i \neq j$ gilt:

$$\begin{aligned} \frac{\Pr(r(t)|Q_j) \cdot \Pr(Q_j)}{\Pr(r(t))} &> \frac{\Pr(r(t)|Q_i) \cdot \Pr(Q_i)}{\Pr(r(t))} \\ \Rightarrow \Pr(r(t)|Q_j) \cdot \Pr(Q_j) &> \Pr(r(t)|Q_i) \cdot \Pr(Q_i). \end{aligned}$$

Eine weitere Vereinfachung dieser MAP–Entscheidungsregel führt zum ML–Empfänger, wobei „ML“ für „Maximum–Likelihood“ steht.

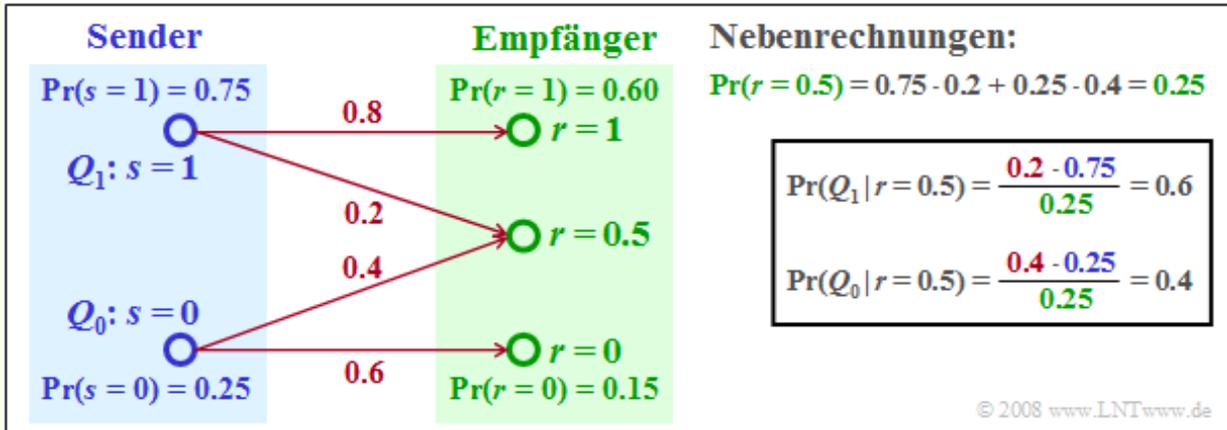
Definition: Der **Maximum–Likelihood–Empfänger** – abgekürzt ML – entscheidet nach den bedingten Vorwärtswahrscheinlichkeiten $\Pr(r(t)|Q_i)$ und setzt die Folge $V = Q_j$, falls für alle $i \neq j$ gilt:

$$\Pr(r(t)|Q_j) > \Pr(r(t)|Q_i).$$

Ein Vergleich dieser beiden Definitionen zeigt, dass bei gleichwahrscheinlichen Quellensymbolen der ML– und der MAP–Empfänger gleiche Entscheidungsregeln befolgen und somit vollkommen äquivalent sind. Bei nicht gleichwahrscheinlichen Symbolen ist der ML– dem MAP–Empfänger unterlegen, da er für die Detektion nicht alle zur Verfügung stehenden Informationen nutzt.

MAP- und Maximum-Likelihood-Entscheidungsregel (2)

Beispiel: Zur Verdeutlichung von ML- und MAP-Entscheidungsregel konstruieren wir nun ein sehr einfaches Beispiel mit nur zwei Quellensymbolen ($M = 2$). Diese beiden möglichen Symbole Q_0 und Q_1 werden durch die Sendesignale $s = 0$ bzw. $s = 1$ dargestellt. Dagegen kann das Empfangssignal – warum auch immer – drei verschiedene Werte annehmen, nämlich $r = 0$, $r = 1$ und $r = 0.5$.



Die Empfangswerte $r = 0$ und $r = 1$ werden sowohl vom ML- als auch vom MAP-Entscheider den Senderwerten $s = 0$ (Q_0) bzw. $s = 1$ (Q_1) zugeordnet. Dagegen werden die beiden Entscheider bezüglich des Empfangswertes $r = 0.5$ zu einem anderen Ergebnis kommen:

- Die Maximum-Likelihood-Entscheidungsregel führt zum Quellensymbol Q_0 , wegen

$$\Pr(r = 0.5 | Q_0) = 0.4 > \Pr(r = 0.5 | Q_1) = 0.2.$$
- Die MAP-Entscheidung führt dagegen zum Quellensymbol Q_1 , da entsprechend der Grafik gilt:

$$\Pr(Q_1 | r = 0.5) = 0.6 > \Pr(Q_0 | r = 0.5) = 0.4.$$

ML-Entscheidung bei Gaußscher Störung

Wir setzen nun voraus, dass sich das Empfangssignal $r(t)$ additiv aus einem Nutzsignal $s(t)$ und einem Störanteil $n(t)$ zusammensetzt, wobei die Störung als gaußverteilt und weiß angenommen wird (Beispiel: AWGN-Rauschen):

$$r(t) = s(t) + n(t).$$

Eventuelle Kanalverzerrungen werden zur Vereinfachung bereits dem Signal $s(t)$ beaufschlagt.

Die notwendige Rauschleistungsbegrenzung wird durch einen Integrator realisiert; dies entspricht einer Mittelung der Rauschwerte im Zeitbereich. Begrenzt man das Integrationsintervall auf den Bereich t_1 bis t_2 , so kann man für jede Quellensymbolfolge Q_i eine Größe W_i ableiten, die ein Maß für die bedingte Wahrscheinlichkeit $\Pr(r(t)|Q_i)$ darstellt:

$$W_i = \int_{t_1}^{t_2} r(t) \cdot s_i(t) dt - \frac{1}{2} \cdot \int_{t_1}^{t_2} s_i^2(t) dt = I_i - \frac{E_i}{2}.$$

Diese Entscheidungsgröße W_i kann über die **k -dimensionale Verbundwahrscheinlichkeitsdichte** der Störungen (mit $k \rightarrow \infty$) und einigen Grenzübergängen hergeleitet werden. Die Gleichung lässt sich wie folgt interpretieren:

- Die Integration dient der Rauschleistungsbegrenzung. Werden vom ML-Detektor N Binärsymbole gleichzeitig entschieden, so ist bei verzerrungsfreiem Kanal $t_1 = 0$ und $t_2 = NT$ zu setzen.
- Der erste Term der obigen Entscheidungsgröße W_i ist gleich der über das endliche Zeitintervall NT gebildeten **Energie-Kreuzkorrelationsfunktion** zwischen $r(t)$ und $s_i(t)$ an der Stelle $\tau = 0$:

$$I_i = \varphi_{r, s_i}(\tau = 0) = \int_0^{NT} r(t) \cdot s_i(t) dt.$$

- Der zweite Term gibt die halbe Energie des betrachteten Nutzsignals $s_i(t)$ an, die zu subtrahieren ist. Die Energie ist gleich der AKF des Nutzsignals an der Stelle $\tau = 0$:

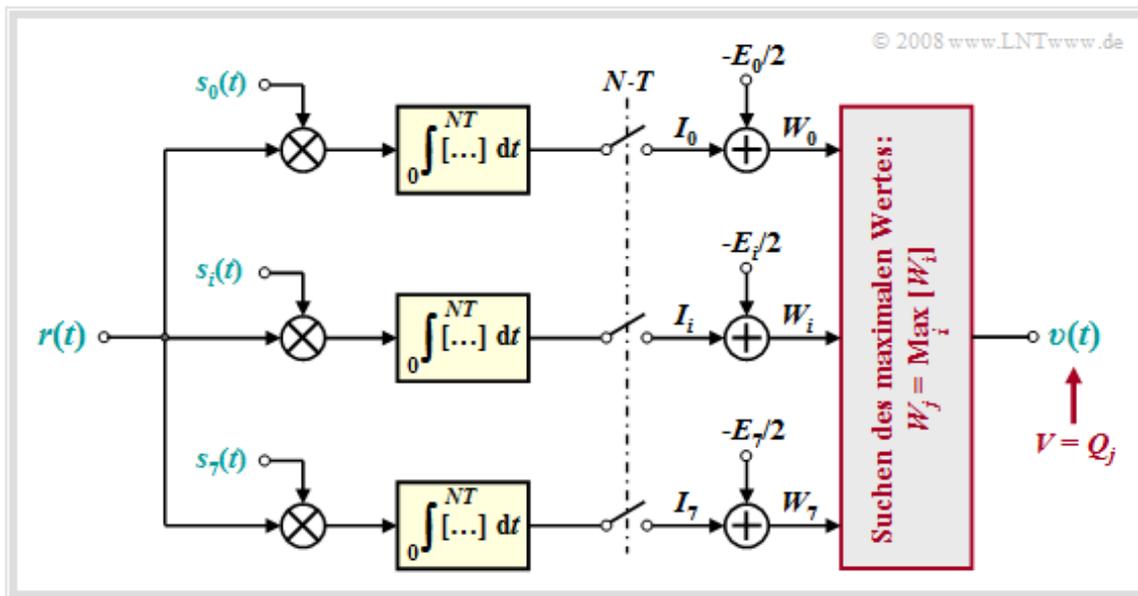
$$E_i = \varphi_{s_i}(\tau = 0) = \int_0^{NT} s_i^2(t) dt.$$

- Bei verzerrendem Kanal ist die Impulsantwort $h_K(t)$ nicht diracförmig, sondern beispielsweise auf den Bereich $-T_K \leq t \leq +T_K$ ausgedehnt. In diesem Fall muss für die beiden Integrationsgrenzen $t_1 = -T_K$ und $t_2 = NT + T_K$ eingesetzt werden.

Korrelationsempfänger

Es gibt verschiedene schaltungstechnische Implementierungen des Maximum-Likelihood-Empfängers. Beispielsweise können die erforderlichen Integrale durch lineare Filterung und anschließender Abtastung gewonnen werden. Man bezeichnet diese Realisierungsform als **Matched-Filter-Empfänger**, da hier die Impulsantworten der M parallelen Filter formgleich mit den Nutzsignalen $s_0(t), \dots, s_{M-1}(t)$ sind.

Die M Entscheidungsgrößen I_i sind dann gleich den Faltungsprodukten $r(t) * s_i(t)$ zum Zeitpunkt $t = 0$. Beispielsweise erlaubt der im **Kapitel 1.4** ausführlich beschriebene „optimale Binärempfänger“ eine Maximum-Likelihood-Entscheidung mit den ML-Parametern $M = 2$ und $N = 1$.



Wir beschränken uns hier auf den sog. **Korrelationsempfänger** entsprechend obigem Blockschaltbild. Zur Vereinfachung werden $N = 3$, $t_1 = 0$, $t_2 = 3T$ sowie $M = 2^3 = 8$ vorausgesetzt. Man erkennt:

- Der Korrelationsempfänger bildet insgesamt $M = 8$ Kreuzkorrelationsfunktionen zwischen dem Empfangssignal $r(t) = s_k(t) + n(t)$ und den möglichen Sendesignalen $s_i(t)$, $i = 0, \dots, M-1$. Vorausgesetzt ist für diese Beschreibung, dass das Nutzsignal $s_k(t)$ gesendet wurde.
- Der Korrelationsempfänger sucht nun den maximalen Wert W_j aller Korrelationswerte und gibt die dazugehörige Folge Q_j als Sinkensymbolfolge V aus. Formal lässt sich die ML-Entscheidungsregel wie folgt ausdrücken:

$$V = Q_j, \text{ falls } W_j > W_i \text{ für alle } i \neq j.$$

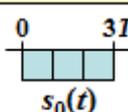
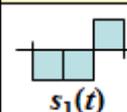
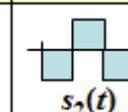
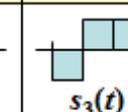
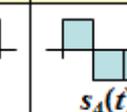
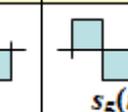
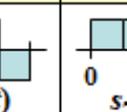
- Setzt man weiter voraus, dass alle Sendesignale $s_i(t)$ die genau gleiche Energie besitzen, so kann man auf die Subtraktion von $E_i/2$ in allen Zweigen verzichten. In diesem Fall werden folgende Korrelationswerte miteinander verglichen ($i = 0, \dots, M-1$):

$$I_i = \int_0^{NT} s_j(t) \cdot s_i(t) dt + \int_0^{NT} n(t) \cdot s_i(t) dt.$$

- Mit großer Wahrscheinlichkeit ist $I_j = I_k$ größer als alle anderen Vergleichswerte $I_i \neq k$. Ist das Rauschen $n(t)$ allerdings zu groß, so kann auch der ML-Empfänger eine Fehlentscheidung treffen.

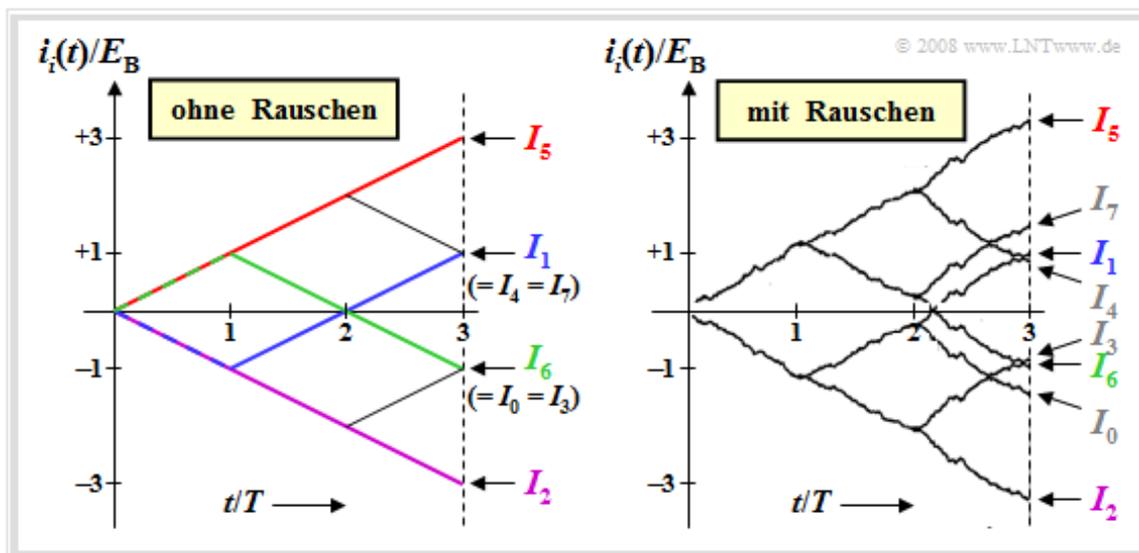
Darstellung des Korrelationsempfängers im Baumdiagramm (1)

Verdeutlichen wir uns die Funktionsweise des Korrelationsempfängers im **Baumdiagramm**, wobei die $2^3 = 8$ möglichen Quellensymbolfolgen Q_i der Länge $N = 3$ durch bipolare rechteckförmige Sendesignale $s_i(t)$ repräsentiert werden. Die möglichen Quellensymbolfolgen $Q_0 = „LLL“$, ... , $Q_7 = „HHH“$ und die dazugehörigen Sendesignale $s_0(t)$, ... , $s_7(t)$ sind nachfolgend aufgeführt.

$Q_0: LLL$	$Q_1: LLH$	$Q_2: LHL$	$Q_3: LHH$	$Q_4: HLL$	$Q_5: HLH$	$Q_6: HHL$	$Q_7: HHH$
							
$s_0(t)$	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$	$s_5(t)$	$s_6(t)$	$s_7(t)$

© 2008 www.LNTwww.de

Wegen den bipolaren Amplitudenkoeffizienten und der Rechteckform sind alle Signalenergien E_0, \dots, E_7 gleich $N \cdot E_B$, wobei E_B die Energie eines Einzelimpulses der Dauer T angibt. Deshalb kann auf die Subtraktion des Terms $E_i/2$ in allen Zweigen verzichtet werden. Eine auf den Korrelationswerten I_i basierende Entscheidung liefert ebenso zuverlässige Ergebnisse wie die Maximierung der korrigierten Werte W_i .



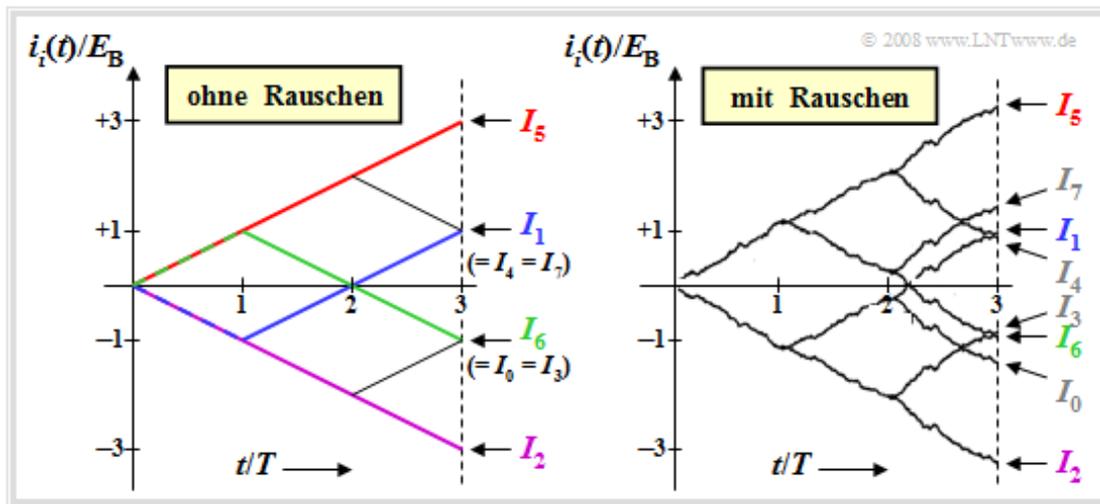
In der linken Grafik sind die fortlaufenden Integralwerte dargestellt, wobei vom tatsächlich gesendeten Signal $s_5(t)$ und dem rauschfreien Fall ausgegangen wird:

$$i_i(t) = \int_0^t r(\tau) \cdot s_i(\tau) d\tau = \int_0^t s_5(\tau) \cdot s_i(\tau) d\tau \Rightarrow I_i = i_i(3T).$$

Das rechte Baumdiagramm berücksichtigt AWGN-Rauschen $n(t)$ mit der Varianz $\sigma_n^2 = 4 \cdot E_B/T$.

Die Interpretation dieser Grafik folgt auf der nächsten Seite.

Darstellung des Korrelationsempfängers im Baumdiagramm (2)



Es folgt zunächst die Beschreibung des rauschfreien Baumdiagramms (linke Grafik):

- Wegen der Rechteckform der Signale $s_i(t)$ sind alle Funktionsverläufe $i_i(t)$ geradlinig. Die auf $3T$ normierten Endwerte sind $+3$, $+1$, -1 und -3 .
- Der maximale Endwert ist $I_5 = 3E_B$ (roter Kurvenverlauf), da tatsächlich das Signal $s_5(t)$ gesendet wurde. Ohne Rauschen trifft der Korrelationsempfänger somit natürlich die richtige Entscheidung.
- Der blaue Kurvenzug $i_1(t)$ führt zum Endwert $I_1 = -E_B + E_B + E_B = E_B$, da sich $s_1(t)$ von $s_5(t)$ nur im ersten Bit unterscheidet. Die Vergleichswerte I_4 und I_7 sind ebenfalls gleich E_B .
- Da sich $s_0(t)$, $s_3(t)$, und $s_6(t)$ von $s_5(t)$ in zwei Bit unterscheiden, gilt $I_0 = I_3 = I_6 = -E_B$. Die grüne Kurve zeigt $i_6(t)$, das zunächst ansteigt (erstes Bit stimmt überein) und dann über zwei Bit abfällt.
- Die violette Kurve führt zum Endwert $I_2 = -3E_B$. Das zugehörige Signal $s_2(t)$ unterscheidet sich von $s_5(t)$ in allen drei Symbolen und es gilt $s_2(t) = -s_5(t)$.

Im rechten Baumdiagramm sind die Funktionsverläufe aufgrund des Rauschanteils $n(t)$ nicht mehr gerade und es ergeben sich auch etwas andere Endwerte als ohne Rauschen. Im betrachteten Beispiel entscheidet der Korrelationsempfänger mit großer Wahrscheinlichkeit richtig, da die Differenz zwischen I_5 und dem nächstgrößeren Wert I_7 mit $1.65E_B$ verhältnismäßig groß ist.

Die Fehlerwahrscheinlichkeit ist in dem hier betrachteten Beispiel allerdings nicht besser als die des Matched-Filter-Empfängers mit symbolweiser Entscheidung. Entsprechend **Kapitel 1.4** gilt auch hier:

$$p_B = Q\left(\sqrt{2 \cdot E_B/N_0}\right) = 1/2 \cdot \operatorname{erfc}\left(\sqrt{E_B/N_0}\right).$$

Weist das Eingangssignal aber statistische Bindungen auf, so führt die gemeinsame Entscheidung von N Bit zu einer spürbaren Verbesserung, da der ML-Empfänger dies berücksichtigt. Solche Bindungen können entweder durch sendeseitige Codierung bewusst erzeugt werden (siehe Buch „Kanalcodierung“) oder durch (lineare) Kanalverzerrungen ungewollt entstehen.

Bei Vorhandensein solcher Impulsinterferenzen ist die Berechnung der Fehlerwahrscheinlichkeit deutlich schwieriger. Es können jedoch vergleichbare Näherungen wie beim Viterbi-Empfänger angegeben werden, die am Ende von **Kapitel 3.8** zu finden sind.

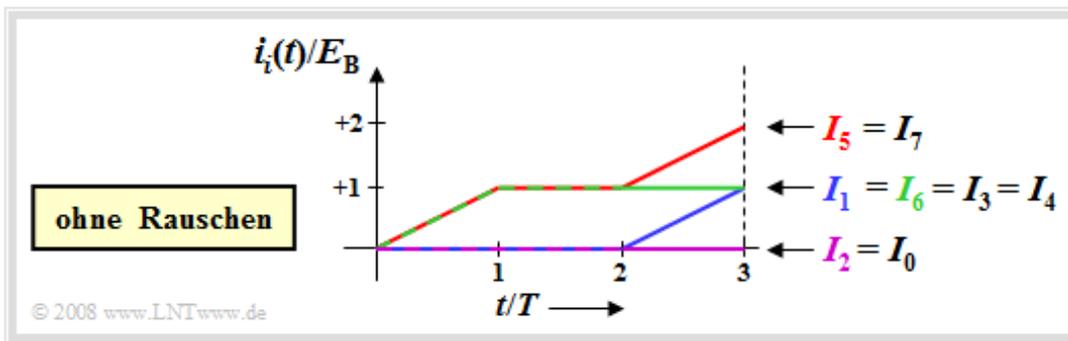
Darstellung des Korrelationsempfängers im Baumdiagramm (3)

Bei unipolarer Signalisierung führt eine auf I_i basierende Entscheidung nicht zum richtigen Ergebnis. Hier muss vielmehr auf die korrigierten Vergleichswerte $W_i = I_i - E_i/2$ zurückgegriffen werden.

Q_0 : LLL	Q_1 : LLH	Q_2 : LHL	Q_3 : LHH	Q_4 : HLL	Q_5 : HLH	Q_6 : HHL	Q_7 : HHH
$s_0(t)$	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$	$s_5(t)$	$s_6(t)$	$s_7(t)$
$E_0/E_B = 0$	$E_1/E_B = 1$	$E_2/E_B = 1$	$E_3/E_B = 2$	$E_4/E_B = 1$	$E_5/E_B = 2$	$E_6/E_B = 2$	$E_7/E_B = 3$

© 2008 www.LNTwww.de

Nachfolgend sehen Sie das Baumdiagramm für den rauschfreien Fall, wobei wie auf der letzten Seite vorausgesetzt ist, dass $s_5(t)$ gesendet wurde. Auch die Farben korrespondieren mit der Grafik auf der letzten Seite.



Für dieses Beispiel ergeben sich folgende Vergleichswerte, jeweils normiert auf E_B :

$$I_5 = I_7 = 2, \quad I_1 = I_3 = I_4 = I_6 = 1, \quad I_0 = I_2 = 0,$$

$$W_5 = 1, \quad W_1 = W_4 = W_7 = 0.5, \quad W_0 = W_3 = W_6 = 0, \quad W_2 = -0.5.$$

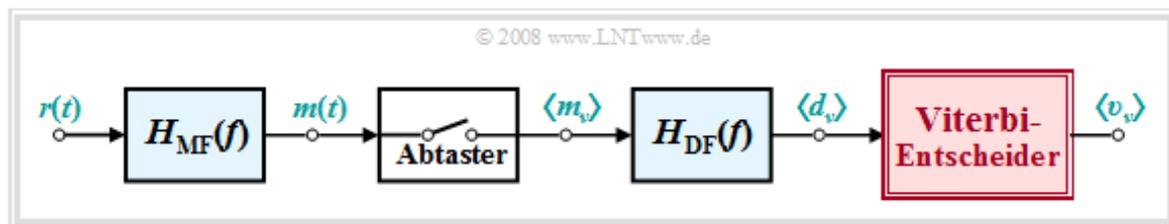
Bei einem Vergleich hinsichtlich der maximalen I_i -Werte wären die Quellensymbolfolgen Q_5 und Q_7 gleichwertig. Bei Berücksichtigung der unterschiedlichen Energien ($E_5 = 2, E_7 = 3$) ist dagegen $W_5 > W_7$.

Blockschaltbild und Voraussetzungen für Kapitel 3.8 (1)

Der **Korrelationsempfänger** ist im Sinne der **Maximum-Likelihood-Entscheidungsregel** optimal, das heißt, er führt bei gleichwahrscheinlichen Quellensymbolen zur minimalen Fehlerwahrscheinlichkeit. Nachteilig ist:

- Der Realisierungsaufwand steigt exponentiell mit der Länge N der zu detektierenden Symbolfolge.
- Da die Folge gemeinsam entschieden wird, kommt es bei großem N zu langen Verzögerungen.

In den 1970er Jahren hat **Andrew J. Viterbi** einen ML-Empfänger vorgeschlagen, der die Detektion von Teilen der empfangenen Nachricht erlaubt und bei dem sich der Realisierungsaufwand auch bei unendlich langen Folgen in Grenzen hält.



Zu den einzelnen Komponenten des Blockschaltbildes ist anzumerken:

- Das an den Empfangsgrundimpuls und die Störung angepasste Matched-Filter $H_{MF}(f)$ dient der **Störleistungsbegrenzung**. Das MF-Ausgangssignal $m(t)$ bzw. die Folge $\langle m_v \rangle$ der äquidistanten Signalwerte nach der Abtastung besitzt das bestmögliche Signal-zu-Stör-Leistungsverhältnis.
- Aufgabe des Dekorrelationsfilters $H_{DF}(f)$ ist es, aus der Folge $\langle m_v \rangle$ die Detektionsabtwerte $d_v = d_{Sv} + d_{Nv}$ zu gewinnen, deren Störanteile d_{Nv} unkorreliert sind. Dieses Filter wird deshalb auch **Whitening-Filter** genannt.
- Der **Viterbi-Entscheider**, der im Mittelpunkt der folgenden Betrachtungen steht, gewinnt aus der Folge $\langle d_v \rangle$ seiner wertkontinuierlichen Eingangswerte die binäre Ausgangsfolge $\langle v_v \rangle$ entsprechend der Maximum-Likelihood-Regel mit der kleinstmöglichen Fehlerwahrscheinlichkeit $\Pr(v_v \neq q_v)$.

Die Beschreibung wird auf der nächsten Seite fortgesetzt.

Blockschaltbild und Voraussetzungen für Kapitel 3.8 (2)

Um den **Viterbi-Algorithmus** möglichst einfach beschreiben und veranschaulichen zu können, werden hier einige vereinfachende Voraussetzungen getroffen:

- Die Amplitudenkoeffizienten seien **unipolar** $\Rightarrow a_v \in \{0, 1\}$. Anzumerken ist, dass es bei der Verwendung bipolarer Koeffizienten $a_v \in \{-1, +1\}$ nur weniger Modifikationen bedarf.
- Der Grundimpuls $g_d(t)$ besteht nur aus dem **Hauptwert** $g_0 = g_d(t = T_D)$ und **einem Vorläufer** $g_{-1} = g_d(t = T_D - T)$.

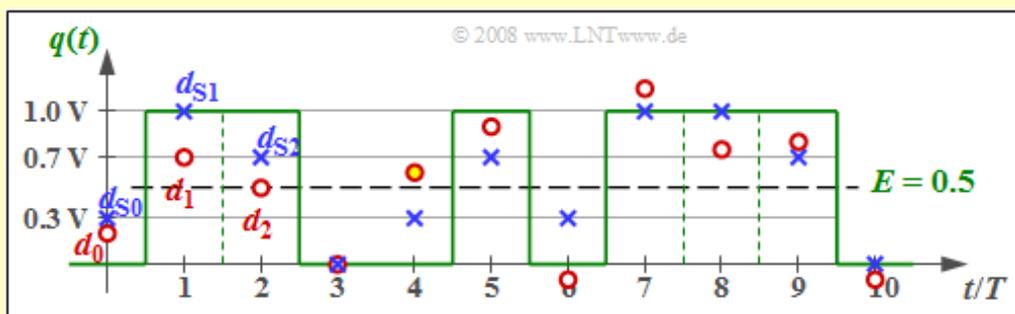
- Damit ergeben sich für die wertkontinuierlichen **Detektionsabtastrwerte**

$$d_v = a_v \cdot g_0 + a_{v+1} \cdot g_{-1} + d_{Nv},$$

wobei die Rauschkomponente d_{Nv} als gaußverteilt angenommen wird (Streuung σ_d).

Bei bipolarer Signalisierung ist der Algorithmus nicht aufwändiger. Dagegen steigt der Rechenaufwand, wenn der Detektionsgrundimpuls breiter wird und mehr als nur einen Vorläufer g_{-1} aufweist. Die Vernachlässigung von Nachläufern stellt keine grundlegende Einschränkung dar, weil jeder Impuls $g_d(t)$ diese Bedingung durch geeignete Wahl des Detektionszeitpunktes T_D erfüllen kann. Anzumerken ist weiter, dass im Folgenden alle Signalwerte auf 1 normiert werden.

Beispiel: In der Grafik sind die Detektionsnutzabtastrwerte d_{sv} als (blaue) Kreuze eingetragen, wobei die zugehörigen Amplitudenkoeffizienten $a_1 = 1, a_2 = 1, a_3 = 0, \dots$ aus dem grün eingezeichneten Quellensignal $q(t)$ abgelesen werden können. Die Grundimpulswerte sind in diesem Beispiel zu $g_0 = 0.7$ und $g_{-1} = 0.3$ angenommen. Aus der Grafik ist weiter zu erkennen, dass d_{sv} nur vier verschiedene Werte, nämlich $0, g_0, g_{-1}$ und $g_0 + g_{-1}$, annehmen kann.



Die am Viterbi-Entscheider anstehenden Abtastrwerte (rote Punkte) sind $d_0 = 0.2, d_1 = 0.7, d_2 = 0.5, d_3 = 0, \dots$, wobei die Differenzen $d_{Nv} = d_v - d_{sv}$ von einer AWGN-Rauschquelle herrühren.

Ein Schwellenwertentscheider (mit der Schwelle bei $E = 0.5$) würde bei diesen dargestellten zehn Bit mindestens eine Fehlentscheidung treffen (bei $v = 4$), und eventuell eine weitere bei $v = 2$, falls d_2 geringfügig kleiner ist als der Schwellenwert $E = 0.5$. Dagegen wird der Viterbi-Empfänger diese Folge der Länge 10 richtig entscheiden, wie auf den nächsten Seiten gezeigt werden wird.

Fehlergrößen und Gesamtfehlergrößen (1)

Wie in Kapitel 3.7 gibt $Q \in \{Q_i\}$ die begrenzte, aus N Binärsymbolen bestehende Quellensymbolfolge an. Die Anzahl der möglichen Symbolfolgen Q_i ist somit 2^N . V bezeichnet wieder die Sinkensymbolfolge der Länge N , die vom Viterbi-Entscheider gleich der wahrscheinlichsten Folge Q_j gesetzt wird.

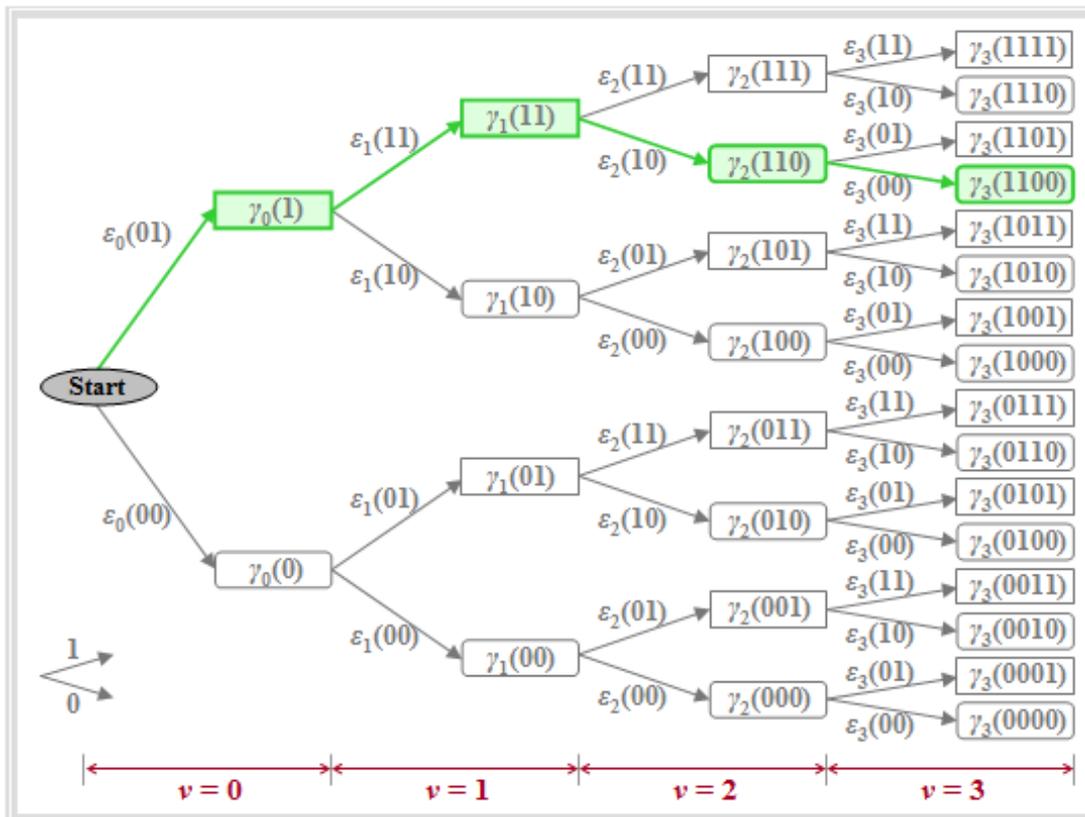
Definition: Die **Fehlergröße** bezeichnet die quadratische Abweichung zwischen dem tatsächlichen, verrauschten Abtastwert d_v und dem zur Folge Q_i gehörenden Nutzabtastwert $d_{Sv}(i)$:

$$\varepsilon_v(i) = |d_v - d_{Sv}(i)|^2.$$

Die **Gesamtfehlergröße** $\gamma_v(i)$ kennzeichnet die Summe aller Fehlergrößen bis zum Zeitpunkt v :

$$\gamma_v(i) = \sum_{k=0}^v \varepsilon_k(i).$$

Man bezeichnet $\gamma_v(i)$ auch als *Metrik*. Anzumerken ist, dass diese Definitionen an einem Grundimpuls mit Hauptwert g_0 und nur einem Vorläufer g_{-1} angepasst sind. Bei v Vorläufern müsste die obige Summe bei $k = 1 - v$ beginnen. Der Parameter $i \in \{0, \dots, 2^{N+1}-1\}$ wird meist binär dargestellt, so dass i gleichzeitig die Amplitudenkoeffizienten a_1, \dots, a_{v+1} (jeweils entweder 0 oder 1) angibt.



Die Grafik verdeutlicht die oben definierten Größen in einer Baumstruktur, woraus zu erkennen ist, dass die Gesamtfehlergrößen iterativ berechnet werden können:

$$\gamma_v(i) = \gamma_{v-1}(i') + \varepsilon_v(i'').$$

i , i' und i'' sind unterschiedliche Laufvariable. Die Bildbeschreibung folgt auf der nächsten Seite.

Fehlergrößen und Gesamtfehlergrößen (2)

Zu dieser Grafik, die **hier** nochmals eingebildet werden kann, ist Folgendes anzumerken:

- Die Knoten des Baumdiagramms stehen für die *Gesamtfehlergrößen* $\gamma_v(i)$; deren Anzahl wird mit jedem Iterationsschritt verdoppelt. Zum Zeitpunkt v gibt es 2^{v+1} solcher Knoten. Beispielsweise sind für $v = 3$ genau $2^4 = 16$ Knoten zu erkennen.
- Die zu den Gesamtfehlergrößen $\gamma_v(i)$ gehörigen *Amplitudenkoeffizienten* ergeben sich, wenn man den Weg vom Anfangsknoten bis zum betrachteten Knoten verfolgt. Es wird vereinbart, dass einem nach oben gerichteten Zweig der Amplitudenkoeffizient 1 und einem nach unten gerichteten Zweig eine 0 zugeordnet wird.
- Beispielsweise kennzeichnet der grün hinterlegte Knoten $\gamma_3(1100)$ die Gesamtfehlergröße unter der hypothetischen Annahme, dass die Symbole $a_1 = 1, a_2 = 1, a_3 = 0, a_4 = 0$ gesendet wurden. Diese Zuordnung kann auch aus den Richtungen der Pfeile im Baumdiagramm abgelesen werden: zunächst zweimal nach oben, dann zweimal nach unten.
- Aufgrund des Vorläufers muss bereits zum Zeitpunkt $v = 3$ der Koeffizient a_4 mitberücksichtigt werden. Alle Knoten $\gamma_v(i)$, die unter der Voraussetzung $a_{v+1} = 1$ berechnet werden, sind im Baumdiagramm durch Rechtecke dargestellt, während die Hypothese $a_{v+1} = 0$ jeweils durch ein abgerundetes Rechteck symbolisiert ist, z. B. $\gamma_2(110)$ oder $\gamma_3(1100)$.
- Die Zweige im Baumdiagramm sind den *Fehlergrößen* $\varepsilon_v(i)$ zugeordnet. Beim vorausgesetzten Grundimpuls (nur g_0 und g_{-1} sind ungleich 0) gibt es zu jedem Zeitpunkt mit Ausnahme des Startzustandes ($v = 0$) genau vier unterschiedliche Größen:

$$\begin{aligned}\varepsilon_v(00) &= |d_v|^2, & \varepsilon_v(01) &= |d_v - g_{-1}|^2, \\ \varepsilon_v(10) &= |d_v - g_0|^2, & \varepsilon_v(11) &= |d_v - g_0 - g_{-1}|^2.\end{aligned}$$

- Die Gesamtfehlergröße γ_v ist gleich der Summe aus dem vorausgegangenen Knoten γ_{v-1} und dem dazwischenliegenden Zweig ε_v . Beispielsweise gilt für die hervorgehobenen Knoten:

$$\gamma_1(11) = \gamma_0(1) + \varepsilon_1(11), \quad \gamma_2(110) = \gamma_1(11) + \varepsilon_2(10), \quad \gamma_3(1100) = \gamma_2(110) + \varepsilon_3(00).$$

- Bei den ersten Knoten $\gamma_0(0)$ und $\gamma_0(1)$ wird berücksichtigt, dass vor der eigentlichen Übertragung (a_1, a_2, \dots) vereinbarungsgemäß stets das Symbol $a_0 = 0$ übertragen wird. Daraus folgt:

$$\gamma_0(0) = \varepsilon_0(00) = |d_0|^2, \quad \gamma_0(1) = \varepsilon_0(01) = |d_0 - g_{-1}|^2.$$

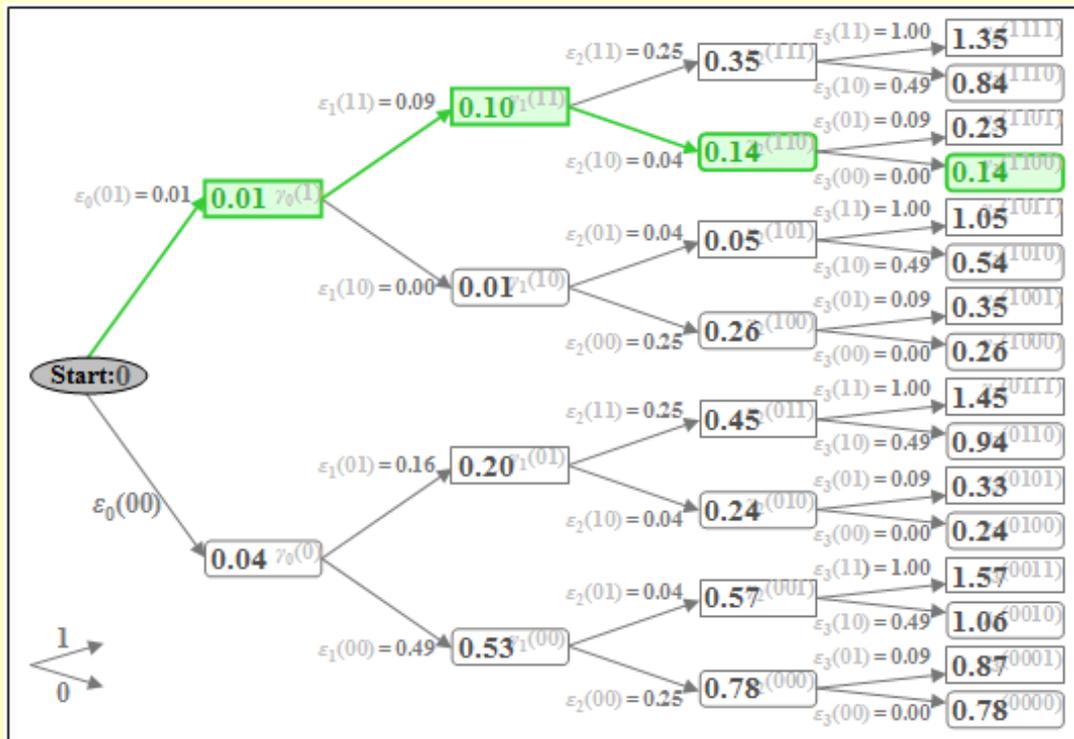
Das nachfolgende Beispiel wird hoffentlich diese etwas ermüdenden Aussagen verdeutlichen.

Fehlergrößen und Gesamtfehlergrößen (3)

Beispiel: Wir betrachten wie im **letzten Beispiel** die unipolare Quellensymbolfolge der Länge $N = 3$, mit den Amplitudenkoeffizienten $a_1 = 1, a_2 = 1, a_3 = 0$ und die weiteren Parameterwerte

$$g_0 = 0.7, \quad g_{-1} = 0.3, \quad d_0 = 0.2, \quad d_1 = 0.7, \quad d_2 = 0.5, \quad d_3 = 0.$$

Nachfolgend ist das Baumdiagramm mit den Knoten $\gamma_v(i)$ für die Zeitpunkte $v = 0$ bis $v = 3$ dargestellt. Die Berechnung der sich ergebenden Fehlergrößen $\varepsilon_v(i)$ folgt auf der nächsten Seite.



Die minimale Gesamtfehlergröße $\gamma_3(1100)$ ist gleich 0.14. Daraus ergeben sich die Koeffizienten der nach den vorliegenden Signalwerten d_0, d_1, d_2, d_3 mit größter Wahrscheinlichkeit gesendeten Folge zu $a_1 = 1, a_2 = 1$ und $a_3 = 0$ (grün eingezeichneter Pfad). Weiter ist zu sagen:

- Ist die Folgenlänge $N = 3$ (das heißt: nur drei Symbole werden durch den Viterbi-Empfänger gemeinsam entschieden), so ist auch die Entscheidung $a_4 = 0$ mit Sicherheit die richtige, da alle Koeffizienten $a_{v>3}$ als 0 vorausgesetzt wurden.
- Bei längerer Folge ($N > 3$) kann aus dem minimalen Wert $\gamma_3(1100)$ nicht unbedingt geschlossen werden, dass $a_1 = 1, a_2 = 1, a_3 = 0$ Teil der wahrscheinlichsten Folge ist. Bei Berücksichtigung weiterer Abtastwerte (d_4, d_5, \dots) könnte sich dieses vorläufige Ergebnis durchaus noch ändern.

Die Berechnung der Fehlergrößen für $v = 0$ bis $v = 3$ wird auf der nächsten Zeile gezeigt.

Fehlergrößen und Gesamtfehlergrößen (4)

Fortsetzung des Beispiels: Nachzutragen ist die Berechnung der Fehlergrößen für die Zeitpunkte $\nu = 0$ bis $\nu = 3$. Die unipolare Quellensymbolfolge hat die Länge $N = 3$. Die Amplitudenkoeffizienten seien $a_1 = 1, a_2 = 1, a_3 = 0$. Weiterhin gilt:

$$g_0 = 0.7, \quad g_{-1} = 0.3, \quad d_0 = 0.2, \quad d_1 = 0.7, \quad d_2 = 0.5, \quad d_3 = 0.$$

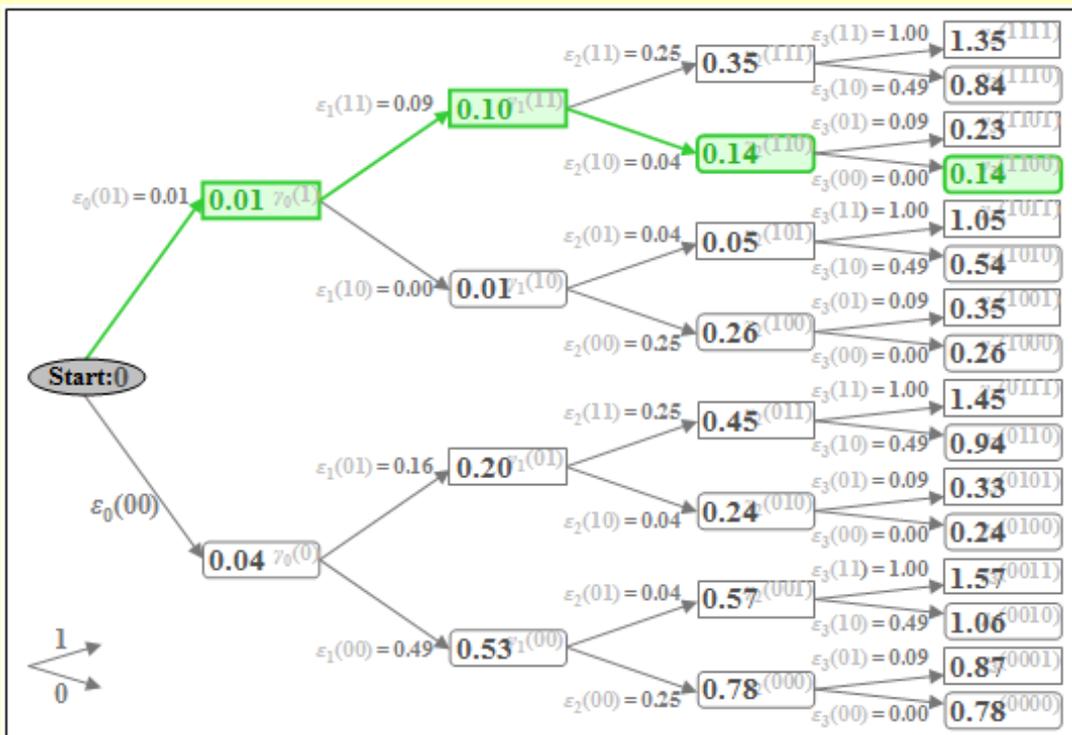
Das Baumdiagramm mit den Knoten $\gamma_\nu(i)$ ist unten dargestellt. Für die Zeitpunkte $\nu = 0$ bis $\nu = 3$ gilt:

$$\begin{aligned} \nu = 0 : \quad \varepsilon_0(00) &= [0.2 - (0 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.04, \\ \varepsilon_0(01) &= [0.2 - (0 \cdot 0.7 + 1 \cdot 0.3)]^2 = 0.01, \end{aligned}$$

$$\begin{aligned} \nu = 1 : \quad \varepsilon_1(00) &= [0.7 - (0 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.49, \\ \varepsilon_1(01) &= [0.7 - (0 \cdot 0.7 + 1 \cdot 0.3)]^2 = 0.16, \\ \varepsilon_1(10) &= [0.7 - (1 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.00, \\ \varepsilon_1(11) &= [0.7 - (1 \cdot 0.7 + 1 \cdot 0.3)]^2 = 0.09, \end{aligned}$$

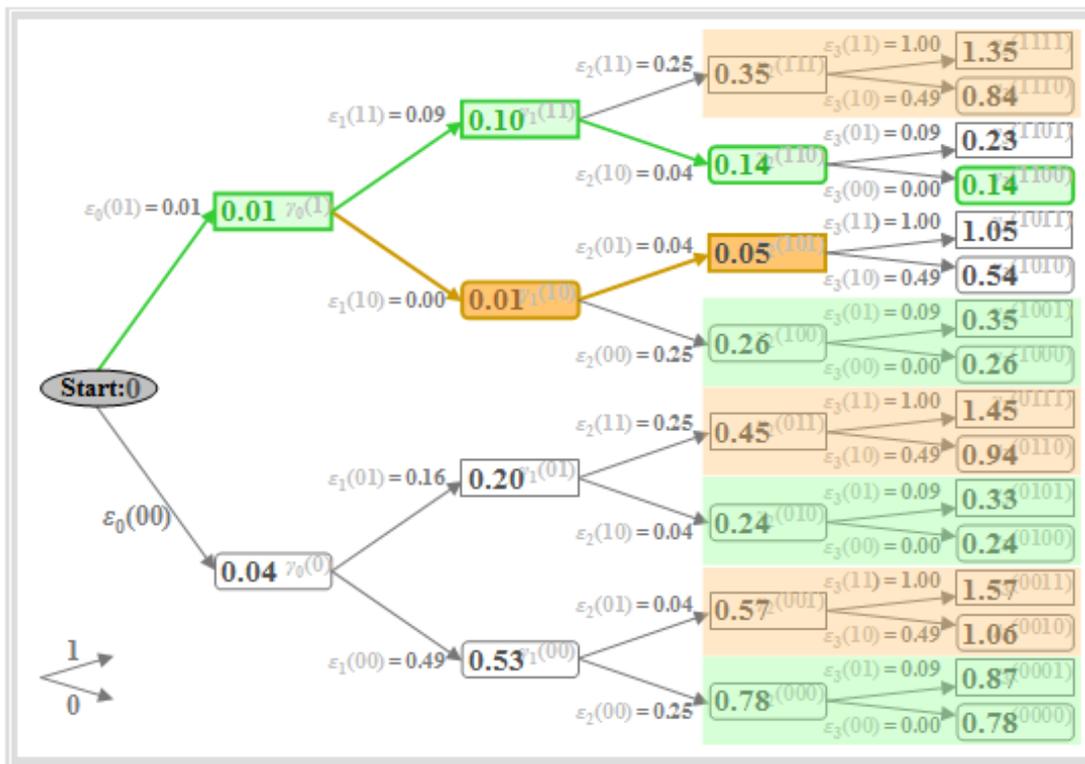
$$\begin{aligned} \nu = 2 : \quad \varepsilon_2(00) &= [0.5 - (0 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.25, \\ \varepsilon_2(01) &= [0.5 - (0 \cdot 0.7 + 1 \cdot 0.3)]^2 = 0.04, \\ \varepsilon_2(10) &= [0.5 - (1 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.04, \\ \varepsilon_2(11) &= [0.5 - (1 \cdot 0.7 + 1 \cdot 0.3)]^2 = 0.25, \end{aligned}$$

$$\begin{aligned} \nu = 3 : \quad \varepsilon_3(00) &= [0.0 - (0 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.00, \\ \varepsilon_3(01) &= [0.0 - (0 \cdot 0.7 + 1 \cdot 0.3)]^2 = 0.09, \\ \varepsilon_3(10) &= [0.0 - (1 \cdot 0.7 + 0 \cdot 0.3)]^2 = 0.49, \\ \varepsilon_3(11) &= [0.0 - (1 \cdot 0.7 + 1 \cdot 0.3)]^2 = 1.00. \end{aligned}$$



Minimale Gesamtfehlergröße und Trellisdiagramm (1)

Wir gehen weiterhin von den Zahlenwerten des letzten Beispiels aus.



Wichtige Eigenschaften des optimalen Entscheiders (Maximum-Likelihood) lassen sich aus obiger Grafik anhand des Zeitpunktes $v = 2$ erkennen:

- Die minimale Gesamtfehlergröße ist $\gamma_2(101) = 0.05$ (braun hervorgehoben). Das bedeutet: Eine Entscheidung zu diesem Zeitpunkt – basierend auf den Werten d_0, d_1 und d_2 – wäre zugunsten der Folge „101“ anstelle der gesendeten Folge „110“ ausgegangen.
- Daraus folgt: Für eine optimale Entscheidung sollte eine zu frühe endgültige Festlegung unbedingt vermieden werden. Allerdings kann man zu jedem Zeitpunkt v bereits mehrere Teilsymbolfolgen ausschließen, die zu späteren Zeitpunkten nicht mehr berücksichtigt werden müssen.
- Zu $\gamma_2(001)$, $\gamma_2(011)$ und $\gamma_2(111)$ werden jeweils genau die gleichen Fehlergrößen hinzuaddiert. Da diese drei Metriken alle größer sind als $\gamma_2(101) = 0.05$, steht bereits bei $v = 2$ fest, dass „001“, „011“ sowie „111“ nicht Bestandteil der wahrscheinlichsten Folge sein können. Diese Knoten und alle ihre Nachfahren müssen deshalb nicht weiter beachtet werden (braune Überdeckungen).
- Gleiches gilt für die abgerundeten Knoten: $\gamma_2(000)$, $\gamma_2(010)$ und $\gamma_2(100)$ sind größer als der grün markierte Knoten $\gamma_2(110) = 0.14$, so dass auch diese und ihre Nachfahren ab dem Zeitpunkt $v = 3$ nicht mehr berücksichtigt werden müssen (grüne Überdeckungen).
- Man stellt fest: Von den acht Knoten bei $v = 2$ müssen nur zwei weiterverfolgt werden, nämlich das abgerundete Rechteck $\gamma_2(110) = 0.14$ und das Rechteck $\gamma_2(101) = 0.05$. Diese beschreiben die minimalen Gesamtfehlergrößen unter der Annahme, dass $a_3 = 0$ bzw. $a_3 = 1$ sein wird.

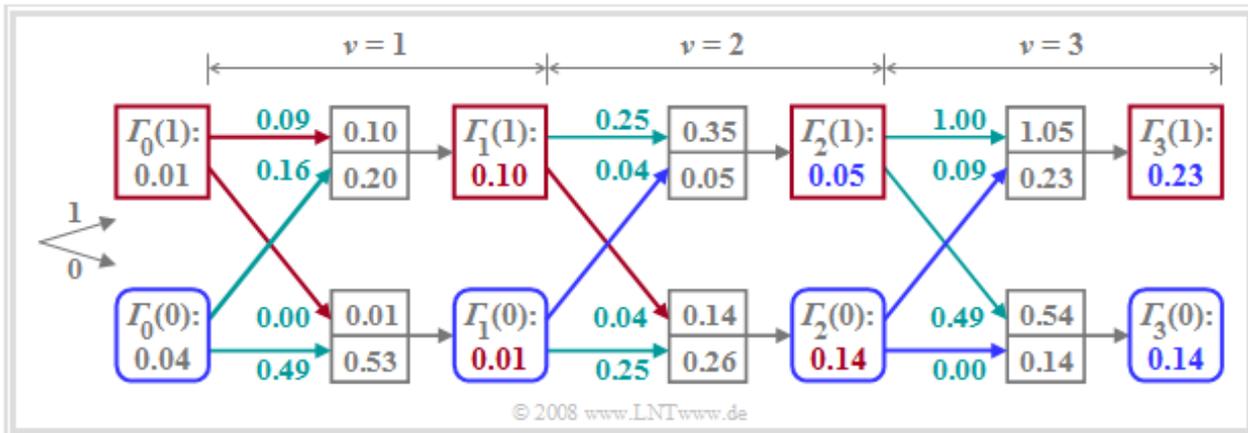
Die Beschreibung wird auf der nächsten Seite fortgesetzt.

Minimale Gesamtfehlergröße und Trellisdiagramm (2)

Verallgemeinern wir nun das Ergebnis dieses Beispiels. Unter der weiterhin gültigen Annahme, dass der Grundimpuls neben dem Hauptwert (g_0) nur einen Vorläufer (g_{-1}) aufweist, ergeben sich die beiden **minimalen Gesamtfehlergrößen** zum Zeitpunkt v formal zu

$$\begin{aligned} \Gamma_v(0) &= \text{Min} [\Gamma_{v-1}(0) + \varepsilon_v(00), \Gamma_{v-1}(1) + \varepsilon_v(10)] , \\ \Gamma_v(1) &= \text{Min} [\Gamma_{v-1}(0) + \varepsilon_v(01), \Gamma_{v-1}(1) + \varepsilon_v(11)] . \end{aligned}$$

Das Verfahren der *Gesamtfehlerminimierung* lässt sich im **Trellisdiagramm** anschaulich darstellen.



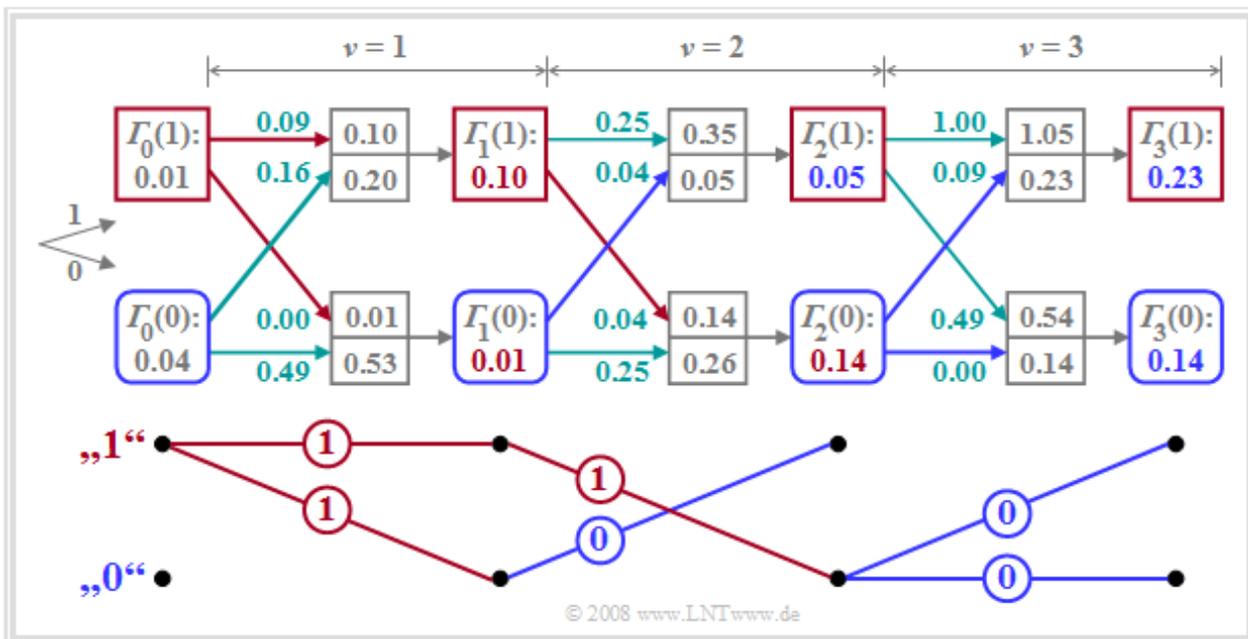
Ein Vergleich mit den Bildern auf den letzten Seiten zeigt:

- Der untere Zweig stellt die minimale Gesamtfehlergröße $\Gamma_v(0)$ dar, die zu jedem Zeitpunkt v unter der Hypothese berechnet wird, dass $a_{v+1} = 0$ gelten wird (blaue abgerundete Quadrate).
- Dagegen beschreibt der obere Zweig die minimalen Gesamtfehlergrößen $\Gamma_v(1)$ unter der Annahme $a_{v+1} = 1$ (rote Quadrate). Auch hier sind die Zahlenwerte an das bisherige Beispiel angepasst.
- Außer $\Gamma_v(0)$ und $\Gamma_v(1)$ muss der ML-Entscheider auch die dazugehörigen Symbolfolgen (Pfade) abspeichern. Diese Zweige sind in der Grafik rot bzw. blau hervorgehoben.
- Falls Γ_v aus dem Knoten $\Gamma_{v-1}(0)$ hervorgeht – also wenn der untere der beiden ankommenden Zweige hervorgehoben ist – so ist das dazugehörige Symbol „0“, andernfalls die „1“.
- Zur Zeit $v = 3$ ergeben sich z. B. sowohl $\Gamma_3(0) = 0.14$ als auch $\Gamma_3(1) = 0.23$ aus dem Vorgänger $\Gamma_2(0)$, so dass beide ausgewählte Pfade jeweils auf das Symbol „0“ verweisen (blaue Zweige).

Vereinfachtes Trellisdiagramm

Der Vorteil des Trellisdiagramms besteht darin, dass sich die Anzahl der Knoten und Zweige nicht bei jedem Iterationsschritt verdoppeln. Durch die Auswahl der minimalen Gesamtfehlergrößen werden nur noch diejenigen Symbolfolgen weiter betrachtet, die als Teil der wahrscheinlichsten Folge überhaupt noch in Frage kommen.

Das Trellisdiagramm lässt sich weiter vereinfachen, indem man nur noch die ausgewählten Zweige einzeichnet. Dies ist im unteren Teil der Grafik an unserem Zahlenbeispiel verdeutlicht. Zur Erinnerung: Die tatsächlich gesendeten Amplitudenkoeffizienten seien $a_1 = 1, a_2 = 1$ und $a_3 = 0$, und das oben gezeichnete Trellisdiagramm wurden unter der Annahme berechnet, dass aufgrund der Impulswerte $g_{-1} = 0.3$ und $g_0 = 0.7$ und des AWGN-Rauschens die Eingangswerte $d_0 = 0.2, d_1 = 0.7, d_2 = 0.5$ und $d_3 = 0$ am ML-Entscheider anliegen.



Dieses **vereinfachte Trellisdiagramm** erlaubt folgende Aussagen:

- Die Entscheidung über $a_1 = 1$ kann sofort zum Zeitpunkt $v = 1$ getroffen werden, da sich unter beiden Hypothesen – nämlich das nachfolgende Symbol ist $a_2 = 0$ bzw. das nachfolgende Symbol ist $a_2 = 1$ – das gleiche Resultat $a_1 = 1$ ergibt.
- Dagegen kann die endgültige Entscheidung über a_2 zum Zeitpunkt $v = 2$ noch nicht getroffen werden. Unter der Annahme, dass $a_3 = 0$ sein wird, müsste man sich für $a_2 = 1$ entscheiden, während die Hypothese $a_3 = 1$ zur Festlegung auf $a_2 = 0$ führen würde.
- Zur Zeit $v = 3$ ist die Entscheidung für $a_1 = 1, a_2 = 1, a_3 = 0$ endgültig, da beide durchgehenden Pfade diese (die in diesem Fall richtige) Folge suggerieren. Würde man die Entscheidung auf den Zeitpunkt $v = 4$ verschieben, so hätte man nicht mehr nutzbare Information über a_1, a_2 und a_3 .

Alle Aussagen dieses Kapitels können mit dem folgenden Interaktionsmodul überprüft werden:

Viterbi-Empfänger für einen Vorläufer

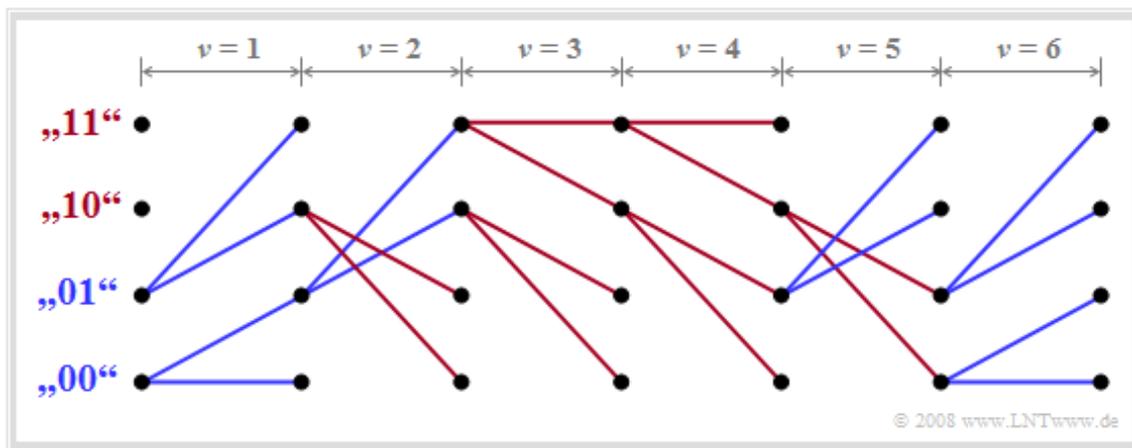
Erweiterung auf zwei Vorläufer

Wird der Grundimpuls durch die Abtastwerte g_0, g_{-1} und g_{-2} beschrieben, so gibt es im Trellisdiagramm zu jedem Zeitpunkt v genau vier Metriken $\Gamma_v(00), \dots, \Gamma_v(11)$ zu bestimmen. $\Gamma_v(01)$ beschreibt dann beispielsweise die minimale Gesamtfehlergröße für die Detektion des Symbols a_v unter der Hypothese, dass $a_{v+1} = 0$ und $a_{v+2} = 1$ sein werden, und es gilt:

$$\Gamma_v(01) = \text{Min} [\Gamma_{v-1}(00) + \varepsilon_v(001), \Gamma_{v-1}(10) + \varepsilon_v(101)] .$$

In der **Aufgabe A3.12** wird noch im Detail auf die Berechnung der Fehlergrößen und die Minimierung der Gesamtfehlergrößen eingegangen. Hier betrachten wir nur ein beispielhaftes Trellisdiagramm, das die (fehlerfreie) Detektion von folgender Symbolfolge wiedergibt:

$$a_1 = 0, a_2 = 0, a_3 = 1, a_4 = 1, a_5 = 1, a_6 = 0, \dots$$



Dieses **vereinfachte Trellisdiagramm** erlaubt folgende Aussagen:

- Alle von $\Gamma_{v-1}(00)$ bzw. $\Gamma_{v-1}(01)$ abgehende Zweige sind dem Symbol „0“ zugeordnet und blau gezeichnet. Die von den zwei oberen Zuständen abgehenden roten Zweige kennzeichnen die „1“.
- Verfolgt man die durchgehenden Pfade, so erkennt man die angegebene Folge. Da zum Zeitpunkt $v = 6$ nur blaue Zweige ankommen, liegen hier die ersten sechs Bit der Folge endgültig fest.
- Teilfolgen könnten aber auch bereits zu den Zeiten $v = 1, v = 3$ und $v = 4$ ausgegeben werden, da sich zu diesen Zeiten für alle vier Zustände die gleichen Teilpfade ergeben.
- Dagegen darf bei $v = 2$ und $v = 5$ nicht sofort entschieden werden. Beispielsweise ist zum Zeitpunkt $v = 5$ nur sicher, dass entweder $a_5 = 0, a_6 = 1$ oder $a_5 = 1, a_6 = 0$ sein werden.

Fehlerwahrscheinlichkeit bei Maximum-Likelihood-Entscheidung (1)

Die exakte Berechnung der Fehlerwahrscheinlichkeit p_B bei ML-Entscheidung (beispielsweise mit dem Korrelations- oder dem Viterbi-Empfänger) ist sehr aufwändig. Dabei müssen

- die Differenzenergien zwischen allen möglichen Symbolfolgen Q_i, Q_j ermittelt werden, wobei die Fehlerwahrscheinlichkeit im Wesentlichen durch die minimale Differenzenergie bestimmt wird;
- auch die Einflüsse von Matched-Filter $H_{MF}(f)$ und Dekorrelationsfilter $H_{DF}(f)$ berücksichtigt und der Effektivwert σ_d des Detektionsstörsignals bestimmt werden.

Eine einfache Näherung für die (mittlere) **Fehlerwahrscheinlichkeit bei ML-Entscheidung** lautet:

$$p_{ML} = Q\left(\frac{g_{\max}}{\sigma_d}\right) \quad \text{mit } g_{\max} = \text{Max } |g_\nu|.$$

Diese Näherung gilt nur bei **bipolarer Signalisierung**. Bei unipolaren Amplitudenkoeffizienten muss das Argument der Q-Funktion halbiert werden.

Für die folgende Interpretation und das anschließende Beispiel auf der nächsten Seite wird vorausgesetzt, dass $v + 1$ Grundimpulswerte (inclusive des Hauptwertes) von 0 verschieden sind. Dann gilt:

- Der Viterbi-Entscheider muss alle diese Grundimpulswerte berücksichtigen. Das bedeutet, dass ein Trellisdiagramm mit 2^v Zuständen zu bearbeiten ist.
- Voraussetzung für die Gültigkeit obiger Gleichung ist die Unkorreliertheit der Störungen am Entscheider, die durch das Dekorrelationsfilter erreicht wird.
- Für den Vergleich mit Schwellenwertentscheider (p_{SE}) bzw. Entscheidungsrückkopplung (p_{DFE}) wird der Effektivwert σ_d des Detektionsstörsignals als konstant vorausgesetzt.
- Zu berücksichtigen ist aber, dass die Optimierung des ML-Systems zu sehr schmalbandigen Filtern führt, da alle Impulsausläufer durch den ML-Algorithmus herausgerechnet werden können.
- Unter der Voraussetzung konstanter Rauschleistungsdichte N_0 (am Empfängereingang) ist deshalb der Störeffektivwert (am Entscheider) beim ML-System kleiner als bei den anderen Varianten.
- Das bedeutet: Der Störabstandsgewinn durch die Maximum-Likelihood-Entscheidung ist unter Umständen noch größer, als es das nachfolgende Beispiel (mit $\sigma_d = \text{const.}$) ausdrückt.

Fehlerwahrscheinlichkeit bei Maximum-Likelihood-Entscheidung (2)

Beispiel: Wir betrachten die Grundimpulswerte $g_{-1} = 0.2$, $g_0 = 0.6$ und $g_1 = 0.2$ und gehen vom konstanten Störeffektivwert $\sigma_d = 0.1$ aus. Zur Vereinfachung sind alle Größen normiert.

Bei einem Binärempfänger mit **einfacher Schwellenwertentscheidung** entsprechend **Kapitel 3.2** ergibt sich für die (mittlere) Fehlerwahrscheinlichkeit bei bipolarer Signalisierung:

$$\begin{aligned} p_{\text{SE}} &= \frac{1}{4} \cdot Q\left(\frac{g_0 + g_1 + g_{-1}}{\sigma_d}\right) + \frac{1}{4} \cdot Q\left(\frac{g_0 + g_1 - g_{-1}}{\sigma_d}\right) \\ &= \frac{1}{4} \cdot Q\left(\frac{g_0 - g_1 + g_{-1}}{\sigma_d}\right) + \frac{1}{4} \cdot Q\left(\frac{g_0 - g_1 - g_{-1}}{\sigma_d}\right) \approx \frac{Q(2)}{4} = 0.57\%. \end{aligned}$$

Für den Empfänger mit **idealer Entscheidungsrückkopplung** erhält man unter Berücksichtigung von $g_1 = 0$ (Kompensation des Nachläufers):

$$p_{\text{DFE}} = \frac{1}{2} \cdot Q\left(\frac{g_0 + g_{-1}}{\sigma_d}\right) + \frac{1}{2} \cdot Q\left(\frac{g_0 - g_{-1}}{\sigma_d}\right) \approx \frac{Q(4)}{4} = 0.16 \cdot 10^{-4}.$$

Dagegen führt die Anwendung der **Maximum-Likelihood-Entscheidung** (mit Korrelations- bzw. Viterbi-Empfänger) zur sehr viel kleineren Fehlerwahrscheinlichkeit

$$p_{\text{ML}} = Q\left(\frac{g_0}{\sigma_d}\right) \approx Q(6) = 10^{-9}.$$

Dies entspricht gegenüber den beiden anderen Systemen einem Störabstandsgewinn von ca. 3 dB (DFE) bzw. 7.5 dB (SE). Das Ergebnis dieser einfachen Näherung wurde durch Simulationen im Wesentlichen bestätigt.

Um den oben beschriebenen Viterbi-Algorithmus direkt anwenden zu können, müssen die (normierten) Grundimpulswerte $g_0 = 0.2$, $g_{-1} = 0.6$ und $g_{-2} = 0.2$ eingestellt werden. Eine Zeitverschiebung um Vielfache der Symboldauer T gegenüber dem aus Darstellungsgründen gewählten Koordinatensystem ändert nämlich die Leistungsmerkmale der Viterbi-Entscheidung nicht.

Die ML-Fehlerwahrscheinlichkeit nach obiger Gleichung richtet sich allein nach dem größten aller Grundimpulswerte. Dabei kann es durchaus sein, dass „Vorläufer“ größer sind als der Hauptwert.

Aus der obigen Näherung erkennt man weiter, dass eine ML-Entscheidung nur bei Vorhandensein von Impulsinterferenzen von Vorteil ist. Bei Nyquistentzerrung (das heißt: nur der Grundimpulswert g_0 ist von 0 verschieden) arbeitet auch der Maximum-Likelihood-Empfänger symbolweise und mit der gleichen Fehlerwahrscheinlichkeit $Q(g_0/\sigma_d)$ wie ein herkömmlicher Empfänger.