Bandwidth-Efficient Correlative Trellis-Coded Modulation Schemes

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH

for the degree of
Doctor of Technical Sciences

presented by

STEFAN RAMSEIER
dipl. El. Ing. ETH
born April 26, 1961
citizen of Luzern LU

accepted on the recommendation of

Prof. Dr. F. Eggimann, referee
Prof. Dr. S.G. Wilson, co-referee

Zürich, 1991
Meinen Eltern

und

Claudia und Esther

gewidmet
Acknowledgements

I am deeply grateful to Dr. Walter Braun and Dr. Dacfey Dzung at ABB/Ascom and Professor Fritz Eggimann at ETH for their continuous interest, their numerous suggestions and their guidance throughout the whole project.

I also want to express my thanks to Professor Stephen G. Wilson of the University of Virginia for his interest in this work and for numerous suggestions improving this report.

A large note of appreciation goes to all members of the ABB/Ascom Communications Group for their helpful collaboration and for many fruitful discussions.
Abstract

A novel digital modulation scheme is proposed that allows to increase the data rate compared to uncoded modulation when the out-of-band power is fixed. If the data rate remains unchanged, this corresponds to a bandwidth reduction that is achieved by shaping the signal spectrum by the introduction of correlation in the transmitter. This correlation has to be resolved in the receiver by a maximum likelihood sequence estimator. Since the sequence of transmitted symbols can be represented by a trellis, this modulation scheme is called correlative trellis coded modulation (C-TCM). The transmitter is first optimized with respect to bandwidth-efficiency, i.e. it is designed to minimize the resulting out-of-band power. Where possible, a secondary optimization with respect to power-efficiency is performed, such that the signal-to-noise ratio required at the receiver to guarantee a certain error rate is minimum.

The optimization shows that for an encoder with memory one the well-known modulation schemes duobinary and dicode are optimum. The mappers that lead to a minimum average or peak power are presented.

The method derived for the optimization of the transmitter is applicable for any number of delay cells. In this report we present results for encoders with up to five delay cells, where new modulation schemes with time-limited and Nyquist pulse shaping filters are found. It is shown that the optimum transmitters can also be represented as partial response systems.

Considering the power-efficiency of the C-TCM scheme, it is shown that the remarkable increase of data rate of up to 300% compared to uncoded quaternary modulation is realized at the cost of about 1.5 dB additional signal power per delay cell. Comparing C-TCM with uncoded M-ary quadrature amplitude modulation, however, C-TCM allows to increase the data rate by up to 0.5 bit/s/Hz for a given performance or to reduce the signal power by about 1 dB for a given data rate.
Kurzfassung

Ein neuartiges digitales Modulationsverfahren wird beschrieben, das die Erhöhung der Datenrate gegenüber uncodierter Modulation erlaubt, wenn die Ausserbandleistung festgehalten wird. Dies entspricht einer Bandbreitenreduktion bei konstanter Datenrate, die dadurch erreicht wird, indem im Sender das Signalspektrum durch Einführen von Korrelation geformt wird; diese Korrelation muss im Empfänger mit Hilfe eines Maximum Likelihood Sequenzschätzers wieder rückgängig gemacht werden. Da die Funktion des Senders mit einem Trellis beschrieben werden kann, wird dieses Verfahren Korrelative Trelliscodierte Modulation (C-TCM) genannt. Zuerst wird der Sender bezüglich Bandbreiteneffizienz optimiert, d.h. er wird so gewählt, dass die Ausserbandleistung minimal wird. Wo möglich wird eine zweite Optimierung bezüglich Leistungseffizienz durchgeführt, damit das Signal/Rauschleistungsverhältnis am Empfänger, das zum Erreichen einer bestimmten Fehlerrate notwendig ist, minimal wird.

Die Optimierung zeigt, dass für Sender mit Gedächtnis 'eins' die bekannten Modulationsarten duobinär und dicode optimal sind. Für diesen Fall werden die Mapper, die zu einer minimalen mittleren oder Spitzenleistung führen, dargestellt.

Die Methode zur Optimierung des Senders gilt für eine beliebige Anzahl von Verzögerungselementen. In diesem Bericht präsentieren wir Resultate für Sender mit zwei bis fünf Verzögerungselementen, die neue Modulationsarten mit zeitbegrenzten und Nyquist-Pulsformungsfilters darstellen. Es wird gezeigt, dass die optimalen Sender auch als Partial Response Systeme dargestellt werden können.

Beim Betrachten der Leistungseffizienz von C-TCM wird klar, dass die bemerkenswerte Erhöhung der Datenrate von bis zu 300% gegenüber uncodierter vierwertiger Modulation mit einem Verlust von etwa 1.5 dB Signalleistung pro Verzögerungselement erkauft wird. Wenn C-TCM hingegen mit uncodierter M-wertiger Quadratramplitudencode verglichen wird, zeigt sich, dass mit C-TCM die Datenrate bei gleichbleibender Signalleistung um bis zu 0.5 bit/s/Hz erhöht, oder die Sendeleistung bei gleichbleibender Datenrate um etwa 1 dB verringert werden kann.
Contents

1 Introduction 9

2 The System Model 11
   2.1 Power Spectral Density of the Output Signal of a FSM-Transmitter 13
   2.2 Spectrum of the Output Signal of a Shift Register FSM 16
   2.3 Transfer Function of the Sampled T-Time-Limited Pulse Shaping Filter 18
   2.4 Transfer Function Of The Spectral Raised Cosine Filter 21

3 Optimization Criteria 23
   3.1 Bandwidth-Efficiency Criterion 23
   3.2 Average-Power-Efficiency Criterion 24
   3.3 Peak-Power-Efficiency Criterion 26

4 Bandwidth-Efficient Mappers 27
   4.1 General Form of the Bandwidth-Efficient Mapper 27
   4.2 Bandwidth-Efficient Mapper for the Encoder with a Single Delay Cell 28
      4.2.1 Optimum Bandwidth-Efficient Mapper for \( \alpha_1 > 0 \) 30
      4.2.2 Optimum Bandwidth-Efficient Mapper for \( \alpha_1 < 0 \) 32
      4.2.3 Summary: Optimum Bandwidth-Efficient Mapper for the Single Delay-Cell Encoder 33
   4.3 Bandwidth-Efficient Mapper for the Encoder with Two Delay Cells 34
   4.4 Bandwidth-Efficient Mapper for the Encoder with Three Delay Cells 37
   4.5 Bandwidth-Efficient Mapper for the Encoder with Four Delay Cells 38
   4.6 Bandwidth-Efficient Mapper for the Encoder with Five Delay Cells 39
   4.7 Optimum Mapper for a Quaternary Input Alphabet 41

5 Representation of the Trellis Encoder as an FIR Filter 43
   5.1 Representation of the Optimum Mapper as an FIR Filter 43
   5.2 Optimization of the FIR Filter 46
      5.2.1 Filters with a Symmetrical Impulse Response and an Odd Number of Taps 49
      5.2.2 Filters with a Symmetrical Impulse Response and an Even Number of Taps 52
      5.2.3 Filters with a Skew-Symmetrical Impulse Response and an Odd Number of Taps 55
      5.2.4 Filters with a Skew-Symmetrical Impulse Response and an Even Number of Taps 57
6 Optimization of the Pulse Shaping Filter

6.1 Optimization of the Sampled Time-Limited Pulse Shaping Filter

6.1.1 Optimum Time-Limited Pulse Shaping Filter for the Uncoded Transmitter

6.1.2 Optimum Time-Limited Pulse Shaping Filter for the Single Delay-Cell Encoder

6.1.3 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Two Delay Cells

6.1.4 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Three Delay Cells

6.1.5 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Four Delay Cells

6.1.6 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Five Delay Cells

6.1.7 Summary of Optimum Encoders with a Time-Limited Pulse Shaping Filter

6.2 Optimization of the Nyquist Pulse Shaping Filter

6.2.1 Optimization of the Nyquist Pulse Shaping Filter for the Single Delay-Cell Encoder

6.2.2 Optimization of the Nyquist Pulse Shaping Filter for the Encoder with Two and More Delay Cells

6.2.3 Summary of Optimum Encoders with a Spectral Raised Cosine Pulse Shaping Filter

7 Power-Efficiency

7.1 Average-Power-Efficient Mapper for the Single Delay-Cell Encoder

7.1.1 Average-Power-Efficient Mapper for $\eta_2$

7.1.2 Average-Power-Efficient Mapper for $\eta_3$

7.1.3 General Average-Power-Efficient Mapper for the Single Delay-Cell Encoder

7.2 Peak-Power-Efficient Mapper for the Single Delay-Cell Encoder

7.3 Performance of the Optimized System in an AWGN Channel

7.4 A Nonlinear Data Precoding Scheme

8 Conclusions

A Proof Of Theorem 1

B Proof Of Theorem 2

C Euclidean Distance Of Signals Generated By A Shift-Register Encoder

D Derivation of Eq. 5.14
## Contents

### E Bandwidth / Power Tradeoff of the FIR Filter Transmitter

- E.1 FIR Filter With a Single Delay Cell, $L = 1$ ................................................. 118
- E.2 FIR filters With Two Delay Cells, $L = 2$ ..................................................... 118

### F Trellis Encoder with a Pulse Shaping Filter Bank

- F.1 Spectrum Calculation and Optimization of a Shift-Register Encoder With a Pulse Shaping Filter Bank ................................................................. 124
- F.2 Bandwidth- and Power-Efficiency of the Pulse Shaping Filter Bank Transmitter, $\gamma = 1$ ................................................................. 125
- F.3 Bandwidth- and Power-Efficiency of the Pulse Shaping Filter Bank Transmitter, $\gamma > 1$ ................................................................. 130
Chapter 1

Introduction

The use of digital communication systems for applications such as the mobile radio services has become very popular in the last few years. There are several advantages of digital compared to analog communication, such as the possibility to transmit voice and digital data over the same channel. Furthermore, digitized voice is much easier to encrypt than an analog signal, a fact that is important in private mobile radio services.

The drawback, however, is the large bandwidth occupied by a digital radio signal. In mobile radio or satellite applications, bandwidth is at a premium. Hence there is a considerable interest in modulation schemes that preserve bandwidth. In the past, several methods to reduce the bandwidth of the transmitted signal were proposed. In general, if the alphabet size $M$ is increased (as in $M$-PSK or $M$-QAM), the bandwidth-efficiency increases at the cost of a decreased error performance. If the communication system is not optimized in the signal space, but in its baseband representation, the influence of the pulse shaping filter becomes apparent. The optimization of the time-limited pulse shaping filter was first performed by Slepian and Pollak [Sle61] and Mueller [Mue73].

Another method is to introduce coding to shape the spectrum of the transmitted signal. These coded modulation schemes result in an increased receiver complexity because the receiver has to resolve the intersymbol interference introduced by the transmitter. For systems that require a constant envelope, the continuous phase modulation (CPM) proved to be very efficient, see [And86]. CPM allows the reduction of the bandwidth more or less without affecting the performance. A different approach is the trellis coded modulation (TCM), that improves the performance without changing the bandwidth, see e.g. [Ung87] and references therein. TCM can also be used if the restriction on the constant envelope is dropped. Another method is not to transmit the current symbol itself but a linear combination of this and some previous symbols. This method leads to the partial response signalling (PRS), that was introduced by Lender [Len63] and then generalized by Kretzmer [Kre66] and Kabai and Pasupathy [Kab75]. Some of these modulation schemes are also known as line codes. They are mainly designed to introduce nulls in the spectrum of the transmitted signal or to inhibit long runs of the same symbol that may cause the loss of synchronization.

Our approach is to introduce coding in the transmitter to reduce the bandwidth. The resulting modulation scheme can be represented as a TCM system, with the difference that in contrast to the 'classical' trellis codes, our transmitter is not optimized with respect to power-, but with respect to bandwidth-efficiency. Because our modulation
introduces a correlation between the input symbols, we call it \textit{correlative trellis coded modulation (C-TCM)} to distinguish it from TCM.

In Chapter 2, the system model is presented, and the power spectral densities of the proposed transmitter signals are computed. We first consider a transmitter consisting of a finite state machine followed by a pulse shaping filter. We then simplify this model by replacing the FSM by a shift-register encoder. We also show the receiver which consists of a matched filter and a maximum likelihood sequence estimator. In Chapter 3, we define the criteria that will be used to optimize our communications system. The main aim of our work is to reduce the bandwidth of the transmitted signal, where the bandwidth-efficiency criterion is used to minimize the out-of-band power. To achieve a good performance of the overall system, the power-efficiency criterion requires that the minimum Euclidean distance of two signals is maximized. The optimization with respect to bandwidth-efficiency is performed in Chapter 4, where it is shown that for a given pulse shaping filter the optimum mapper can be expressed analytically. The results of this optimization show that there is an alternative representation of the transmitter that leads to the same results. The properties of this partial response transmitter are investigated in Chapter 5. In Chapter 6, the pulse shaping filter is optimized. We consider two different filter structures, the \textit{time-limited} and the \textit{Nyquist} pulse shaping filter. The performance of the overall system in an additive white Gaussian noise channel is analyzed in Chapter 7. In the last chapter, we present a comparative overview over the state of the art of different bandwidth- and power-efficient modulation schemes, and we demonstrate the improvement of our new correlative trellis coded modulation scheme.
Chapter 2
The System Model

In this chapter we present the complex baseband model of our communications system. The transmitter consists of a finite state machine (FSM) followed by a pulse shaping filter, and the receiver includes a matched filter and a maximum likelihood sequence estimator (MLSE). The transmitter corresponds to a generalized coded quadrature amplitude modulation scheme.

The power spectral density of the transmitted signal is calculated in Sec. 2.1. Because the analytical treatment of this general FSM transmitter is too complex, we then consider a simplified transmitter, where the FSM is specified to be a shift-register encoder. This is the transmitter that will be optimized in the remaining part of our work. The power spectral density of its output signal is calculated in Sec. 2.2. In the last two sections of this chapter we compute the transfer function of the time-limited and the spectral raised cosine filter. These two filters will be used as pulse shaping filters.

We start by introducing the general time-invariant FSM-transmitter, whose schematic diagram is shown in Fig. 2.1. The transmitted complex baseband signal $s(t)$ is generated by feeding an independent and identically distributed (iid) $M$-ary data sequence at a symbol rate $1/T$ into a finite state machine (FSM). The output of the FSM, a $U$-ary complex valued Dirac pulse, is then filtered by a linear filter with transfer function $Q(f)$. This model of the transmitter is fairly versatile, as it allows the representation of most of the well known line coding schemes such as duobinary or AMI coding as well as the representation of trellis codes. If different filters $Q_i(f)$ are used for each possible output symbol of the FSM, even continuous phase modulation schemes can be modeled (see [Bec88]).

We also consider a simplified transmitter, where the general FSM is replaced by a
shift-register followed by a mapper, as shown in Fig. 2.2. Here again the input signal $x(n)$, which is assumed to be iid and $M$-ary, enters an $M$-ary shift register with $L$ stages. The mapper then produces a $U$-ary complex-valued Dirac pulse depending on the current and the $L$ previous input symbols. This pulse is filtered by the pulse shaping filter $Q(f)$, whose output is the desired baseband signal.

The two transmitters described above shape the spectrum by the introduction of correlation into the $u(n)$ sequence, which results in intersymbol interference (ISI). The receiver has to resolve this ISI, as shown in Fig. 2.3. The received signal $r(t)$, which is assumed to be corrupted by additive white Gaussian noise, is filtered by the matched filter $Q^*(f)$. Its output is sampled at the symbol rate $1/T$. The original signal is then reconstructed by a maximum likelihood sequence estimator (MLSE), whose output $\hat{x}(n-n_r)$ is the estimate of the input symbol at time $n - n_r$, where $n_r$ denotes the decoding delay. The number of states of the MLSE is $M^L$.

Most of the results considering the bandwidth-efficiency of a transmitter hold for any pulse shaping filter $Q(f)$. For the computation of the power-efficiency, however, we have to restrict ourselves to generalized Nyquist filters, i.e. filters that are ISI-free and therefore do not introduce further correlation. We shall use two classes of pulse shaping filters. The first is the $T$-time-limited filter, whose impulse response vanishes outside the interval $[0,T]$. For practical reasons, we consider only the sampled version of this filter. The second pulse shaping filter is a special type of Nyquist filter, the spectral raised cosine (SRC) filter.

We now derive the power spectral density of the output signals of the two transmitters described above.
2.1 Power Spectral Density of the Output Signal of a FSM-Transmitter

In this section we compute the power spectral density of the signal generated by the FSM transmitter of Fig. 2.1. We assume that the FSM has \( N \) distinct states \( \sigma(n) = 0 \ldots N - 1 \), a \( U \)-ary output alphabet, and it is described by the two functions

\[
\sigma(n + 1) = g(x(n), \sigma(n)) \tag{2.1}
\]

and

\[
u(n) = h(x(n), \sigma(n)), \tag{2.2}
\]

where Eq. 2.1 tells in which state the FSM is forced to move at time instant \( n + 1 \) when it was in state \( \sigma(n) \) at time \( n \) and the source symbol \( x(n) \) is fed to the encoder. Eq. 2.2 describes how the encoded symbol \( u(n) \), the output of the FSM, depends on the source symbol \( x(n) \) and on the actual state \( \sigma(n) \) of the encoder. The sequence of the encoded complex symbols \( \{u(n)\} \) is then sent into a linear modulator (i.e. a pulse shaping filter), whose output can be written in the form

\[
s(t) = \sum_{n=-\infty}^{\infty} u(n) \cdot q(t - nT), \tag{2.3}
\]

where \( q(t) \) is the impulse response of pulse shaping filter and \( T \) is the symbol duration. We now present the power spectral density of the signal \( s(t) \) defined in Eq. 2.3.

The state sequence of a finite state machine, \( \{\sigma(n)\} \), is a homogeneous Markov chain with the \( N \times N \) transition probability matrix \( P \),

\[
P = \|p_{im}\| \quad i, m \in [0..N - 1], \tag{2.4}
\]

where

\[
p_{im} = Pr[\sigma(n + 1) = m \mid \sigma(n) = i], \quad i, m \in [0..N - 1]. \tag{2.5}
\]

This equation describes the probability that the next state is \( m \) if the present state is \( i \). The matrix \( P \) provides all the information we need about the sequence of the encoder states. To compute the power spectral density of the output signal \( s(t) \), we further need to know the probabilities \( d_{kim} \) that the symbol \( u(n) = \tilde{u}_k \) is output if a transition from state \( i \) to \( m \) occurs,

\[
d_{kim} = Pr[u(n) = \tilde{u}_k \mid \sigma(n) = i, \sigma(n + 1) = m]. \tag{2.6}
\]

With this definition we allow the mapper to generate different output symbols for identical state transitions, as used by some Ungerboeck codes. If parallel state transitions are not allowed, \( d_{kim} \) is either 0 or 1.

We assume that the Markov chain has \( N \) states and that it is irreducible and aperiodic (for a generalization refer to [Bec88]). Then the stationary probability distribution of the states \( \sigma(m) \)

\[
\pi_i = \lim_{n \to \infty} P^n_{ii} \tag{2.7}
\]
exists and can be computed by solving the set of equations

$$\pi_i \geq 0$$

$$\pi = \pi P$$

$$\sum_{i=0}^{N-1} \pi_i = 1,$$

where $\pi$ is the row vector $[\pi_0 \ldots \pi_{N-1}]$. The sequence of $P^n$ converges to the matrix $P^{\infty}$

$$P^{\infty} = \lim_{n \to \infty} P^n$$

that exists under our assumptions and has all its rows equal to the probability vector $\pi$.

Let us now evaluate four quantities that play an important role in the discussion that follows. We concentrate on the results of [Bec88] and [Big86] and omit the complete derivation.

The first quantity is the stationary probability $s_{km}$ that the symbol $u_k$ is output and a transition to state $m$ occurs:

$$s_{km} = Pr[u(n) = u_k, \sigma(n + 1) = m] = \sum_{i=0}^{N-1} \pi_i P_{im} d_{kim}.$$  \hfill (2.10)

The probabilities $s_{km}$ define a $R \times N$ matrix

$$S = [s_{km}], \quad k \in [0 \ldots R - 1], \quad m \in [0 \ldots N - 1]$$  \hfill (2.11)

where $R = N \cdot M$ is the total number of state transitions of the FSM. The second quantity is the probability $e_{ki}$ that the symbol $u_k$ is output if the present state of the encoder is $i$,

$$e_{ki} = Pr[u(n) = u_k | \sigma(n) = i] = \sum_{m=0}^{N-1} p_{im} d_{kim}.$$  \hfill (2.12)

The probabilities $e_{ki}$ define a $R \times N$ matrix

$$E = [e_{ki}], \quad k \in [0 \ldots R - 1], \quad i \in [0 \ldots N - 1]$$  \hfill (2.13)

The expected value $\mu$ of the encoder output is

$$\mu = \frac{1}{M} \cdot \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \pi_i u_{ij},$$  \hfill (2.14)

where $u_{ij}$ denotes the $U$-ary output symbol if the encoder is in state $i$ and the input symbol is $j$. Next, we define the column vector $u$ whose $R$ components are the possible complex output values of the mapper:

$$u = [u_{00}, u_{01}, \ldots, u_{0(M-1)}, u_{10}, u_{11}, \ldots, u_{(N-1)(M-1)}]',$$  \hfill (2.15)

where the prime denotes transposition. Note that $U \leq R$, with equality if and only if all $u_{ij}$ are distinct. Some simple algebraic operations show that with this definition Eq. 2.14 can be written as

$$\mu = \frac{1}{M} \cdot \pi G u,$$  \hfill (2.16)
2.1. Power Spectral Density of the Output Signal of a FSM-Transmitter

where $G$ denotes the $N \times R$ matrix

$$G = \begin{bmatrix}
1 & \ldots & 1 & 0 & \ldots & 0 & \ldots & 0 \\
0 & \ldots & 0 & 1 & \ldots & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 0 & \ldots & 0 & 1 & 1
\end{bmatrix}$$ (2.17)

with $M$ ones in each row. The absolute square of $\mu$ is

$$|\mu|^2 = \frac{1}{M^2} \cdot \mu^\dagger G' \mu G \mu,$$ (2.18)

where the dagger denotes the transposed conjugate. The mean-square value $C_0$ of the encoded symbols is

$$C_0 = \frac{1}{M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \pi_i |u_{ij}|^2 = \frac{1}{M} \cdot \mu^\dagger H \mu,$$ (2.19)

where $H$ denotes the diagonal $R \times R$ matrix

$$H = \text{diag}(\pi_0, \ldots, \pi_0; \pi_1, \ldots, \pi_1; \ldots, \pi_{N-1}, \ldots, \pi_{N-1}),$$ (2.20)

and each element $\pi_i$ occurs $M$ times.

With the quantities defined above, we can write the spectrum $W(f)$ of the output signal $s(t)$ as

$$W(f) = W_c(f) + W_d(f),$$ (2.21)

where the continuous part of the spectrum $W_c(f)$ is

$$W_c(f) = \frac{|Q(f)|^2}{T} \cdot \left\{ (C_0 - |\mu|^2) + 2\Re \left[ \mu^\dagger S(I - P^\infty) \left( \sum_{\nu=0}^{\infty} \exp(-j2\pi \nu f T)(P - P^\infty)^\nu \right) \exp(-j2\pi f T)E' \mu \right] \right\},$$ (2.22)

where we defined $0^0 = I$. Note that for some cases, e.g. for uncoded modulation or for most of the Ungerboeck codes (cf. [Big86]), the above equation collapses, i.e. the expression in brackets is zero. This means that these modulation schemes do not shape the signal power spectrum.

The discrete part of the spectrum $W_d(f)$ equals

$$W_d(f) = |\mu|^2 \cdot \frac{|Q(f)|^2}{T} \sum_{n=-\infty}^{\infty} \delta(f - n/T),$$ (2.23)

where $1/T$ denotes the symbol rate and where $\delta(x)$ denotes the Dirac function. Eq. 2.21 to 2.23 are a combination of the results of [Bec88] and [Big86]. To simplify the discussion in the following sections, we try to express the continuous part of the spectrum in the form

$$W_c(f) = \frac{|Q(f)|^2}{T} \sum_{\nu=0}^{\infty} \left[ \mu^\dagger W_{\nu,\mu} \cdot \cos(2\pi \nu f T) + j\mu^\dagger V_{\nu,\mu} \cdot \sin(2\pi \nu f T) \right],$$ (2.24)
where \( j^2 = -1 \). To accomplish this, we consider the frequency dependent parts of Eqs. 2.22 and 2.24, and then derive the matrices \( W_\nu \) and \( V_\nu \) by comparing the coefficients. Setting \( \nu = 0 \) in Eq. 2.24, we see that \( y^H W_0 y \) has to be equal to \( (C_0 - |\mu|^2) \), because this is the only term in braces of Eq. 2.22 that is independent of \( f \):

\[
W_0 = \frac{1}{M} \cdot H - \frac{1}{M^2} \cdot G' \pi' \pi G,
\]

with the matrices \( G \) and \( H \) defined in Eqs. 2.17 and 2.20. To find the matrices \( W_\nu \) and \( V_\nu \), we write the frequency dependent part of the expression in braces of Eq. 2.22 as

\[
2 \text{Re} \left[ \sum_{\nu=1}^{\infty} y^H \cdot A_\nu \cdot y \cdot \exp(-j2\nu f T) \right],
\]

where

\[
A_\nu = S(I - P^{\infty})(P - P^{\infty})^{\nu-1}E'. \tag{2.27}
\]

Comparing the coefficients of Eqs. 2.24 and 2.26 and using some simple algebraic identities, we find that

\[
W_\nu = A_\nu + A'_\nu \tag{2.28}
\]

and

\[
V_\nu = A_\nu - A'_\nu. \tag{2.29}
\]

The continuous part of the spectrum can therefore be expressed by Eq. 2.24, where the coefficient \( W_0 \) is defined by Eq. 2.25, the coefficients \( W_\nu \) by Eq. 2.28 and the coefficients \( V_\nu \) by Eq. 2.29. This representation of the spectrum is useful because the DC component and the components at multiples of the symbol rate can be identified easily. Furthermore, there is a separation of the matrices determined by the finite state machine (\( W_\nu \) and \( V_\nu \)) and the mapping to the output symbol vector \( y \).

The above equations hold for all irreducible and aperiodic Markov chains.

### 2.2 Spectrum of the Output Signal of a Shift Register FSM

In this section, we compute the power spectral density of the output signal of the shift-register transmitter of Fig. 2.2. This transmitter represents a special class of Markov chains that can be modeled by a shift register that has no feedback, i.e. it has a memory of length \( L \). The corresponding FSM consists of \( L \) \( M \)-ary delay cells whose outputs are fed into the mapper. We label the \( N = M^L \) states by the integers \( \sigma(n) \in (0..N-1) \) according to

\[
\sigma(n) = \sum_{k=1}^{L} z(n-k) \cdot M^k, \tag{2.30}
\]

where \( z(k) \in (0..M-1) \) denotes the symbol that entered the encoder at time \( k \). The state transition function \( g \) defined in Eq. 2.1 now becomes

\[
g(x(n), \sigma(n)) = M\sigma(n) + z(n) \pmod{M^L}, \tag{2.31}
\]
2.2. Spectrum of the Output Signal of a Shift Register FSM

i.e. the new state is computed by multiplying the previous state by \( M \) (shift) and adding the new input symbol, modulo the number of states.

The mapper is modeled by a device whose output may be different for any possible state transition. Since the encoder has \( M^L \) states and there are \( M \) different input symbols, the mapper produces \( R = M^{L+1} \) \( U \)-ary output symbols, some of which may be identical \( (U \leq R) \). We say that the mapper outputs the symbol \( u_{ij} \) if the symbol \( j \) is fed into the encoder when it is in state \( i \). Then the output function \( h \) defined in Eq. 2.2 is

\[
h(x(n), \sigma(n)) = u_{ij} \quad \text{if} \quad x(n) = j \text{ and } \sigma(n) = i. \tag{2.32}
\]

To compute the power spectral density of the signal that is generated by the encoder and the mapper defined above, we have to express the transition probability matrix \( P \) and the output probabilities \( d_{kim} \) in terms of the functions \( g \) and \( h \). We first consider the transition probability matrix \( P \). If all input symbols are equally probable, the probability that a transition occurs from state \( \sigma(n) = i \) to \( \sigma(n + 1) = m \) equals

\[
P_{im} = \begin{cases} \frac{1}{M}, & \text{if } m - iM \pmod{M^L} \in \{0\ldots M - 1\}, \quad i, m \in \{0\ldots N - 1\} \\ 0, & \text{otherwise.} \end{cases} \tag{2.33}
\]

The transition probability matrix is then defined by

\[
P = \|P_{im}\| \quad i, m \in \{0\ldots N - 1\}. \tag{2.34}
\]

With the mapper function \( h \) we can compute the probability \( d_{kim} \) that the symbol \( u(n) = \bar{u}_k \) is output if a transition from state \( i \) to \( m \) occurs, where \( \bar{u}_k \) is an element of the vector \( u \) (see Eq. 2.15).

\[
d_{kim} = P_{r}[u(n) = \bar{u}_k | \sigma(n) = i, \sigma(n + 1) = m] = \begin{cases} 1, & \text{if } \bar{u}_k = u_{ij}, \quad k = iM + j \text{ and } m = (iM + j) \pmod{M^L} \\ 0, & \text{otherwise.} \end{cases} \tag{2.35}
\]

As can be verified easily, the transition probability matrix \( P \) is doubly stochastic, i.e. the sum of each row and each column of \( P \) equals one. In this case (see [Kar75, p. 108]) the stationary probability matrix \( P^\infty \) is

\[
P^\infty = \frac{1}{M^L} \cdot K, \tag{2.36}
\]

where \( K \) denotes the \( M^L \times M^L \) matrix of ones, and the stationary probability distribution vector \( \pi \) is

\[
\pi = \frac{1}{N} \cdot [1, \ldots, 1]. \tag{2.37}
\]

Inserting this result in Eq. 2.25 yields

\[
W_0 = \frac{1}{NM} \cdot I - \frac{1}{(NM)^2} \cdot K, \tag{2.38}
\]

where \( I \) denotes the identity matrix, and where \( K \) denotes the all-ones matrix. A sufficient but not necessary condition for the transmitted signal to be DC-free is \( \mu = \)
\( \mathbf{y}' \cdot \mathbf{1} = 0 \). In this case, the spectrum of the output signal contains no spectral lines \((W_d(f) \equiv 0)\), and \(W_0\) simplifies to

\[
W_0 = \frac{1}{NM} \cdot \mathbf{1}. \tag{2.39}
\]

To compute the power spectral density of the output signal, it is essential to determine the powers of the matrix \((\mathbf{P} - \mathbf{P}^\infty)\) (cf. Eq. 2.22). If the encoder consists of a shift register, this matrix is nilpotent, i.e. there is an integer \(\nu_0\) such that for all integers \(\nu\) greater than or equal \(\nu_0\) the matrix \((\mathbf{P} - \mathbf{P}^\infty)^\nu\) is the all-zero matrix. In this case, the infinite sum of Eq. 2.22 reduces to a sum of \(\nu_0\) terms, as stated in the following theorem.

**Theorem 1** For a finite state machine consisting of \(L\) \(M\)-ary delay cells with the transition probability matrix \(\mathbf{P}\) defined in Eqs. 2.33 and 2.34 and the stationary probability matrix defined in Eq. 2.36,

\[
(\mathbf{P} - \mathbf{P}^\infty)^\nu = 0 \quad \forall \ \nu \geq L. \tag{2.40}
\]

The proof of this theorem is given in Appendix A. Theorem 1 states that the \(L\)-stage shift register introduces a correlation of length \(L\), as expected.

The continuous part of the spectrum of a signal generated by a shift register with \(L\) stages and a mapper that may output different signals for each possible state transition can now be written as

\[
W_c(f) = \mathbf{Q}(f) \cdot \sum_{\nu=0}^{L} \left[ \mathbf{u}^\dagger \mathbf{W}_\nu \mathbf{u} \cdot \cos(2\pi \nu f T) + j \mathbf{u}^\dagger \mathbf{V}_\nu \mathbf{u} \cdot \sin(2\pi \nu f T) \right]. \tag{2.41}
\]

Comparing this result with Eq. 2.24, we see that the shift-register encoder introduces a memory of length \(L\), which is intuitively obvious. The matrices \(\mathbf{W}_\nu\) and \(\mathbf{V}_\nu\) are obtained by first computing the matrices \(\mathbf{S}\) (Eq. 2.10), \(\mathbf{E}\) (Eq. 2.12) and \(\mathbf{A}_\nu\) (Eq. 2.27). Equations 2.28, 2.29 and 2.38 or 2.39 then lead to the desired result.

The above results hold for any pulse shaping filter \(\mathbf{Q}(f)\). In the next two sections we shall introduce two classes of pulse shaping filters. The first is the time-limited filter, whose impulse response vanishes outside the interval \([0, T]\). The second pulse shaping filter is a special type of Nyquist filter, the spectral raised cosine (SRC) filter. We shall compute the transfer functions of these two filters.

### 2.3 Transfer Function of the Sampled \(T\)-Time-Limited Pulse Shaping Filter

In this section we compute the transfer function of the sampled, \(T\)-time-limited pulse shaping filter, where we follow the method introduced by Mueller [Mue73]. For practical reasons, we consider the *sampled* version of the filter which can easily be implemented
2.3. Transfer Function Of the Sampled Time-Limited Pulse Shaping Filter

\[ Q(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} q(t) \exp(-j2\pi fT)dt = \frac{1}{\sqrt{2\pi}} \sum_{i=-\kappa}^{\kappa-1} q_i \exp\left(-j\frac{\pi fiT}{\kappa}\right) = q' \cdot p, \]  
\[ q' = [q_{\kappa}, q_{\kappa-1}, \ldots, q_1] \]  
\[ p' = \frac{1}{\sqrt{2\pi}} \left[ \exp(j\pi fT), \exp\left(-j\frac{\pi fT^{-\kappa + 1}}{\kappa}\right), \ldots, \exp\left(-j\frac{\pi fT^{\kappa - 1}}{\kappa}\right) \right]. \]

As seen from Eq. 2.41, the frequency dependent part of the spectrum of the output signal can be expressed in terms of \(|Q(f)|^2 \cos(2\pi \nu fT)|^2 \sin(2\pi \nu fT)|^2\). For the computation of the out-of-band power that will be accomplished in the next chapters, it is convenient to introduce the functions \(\alpha_{\nu}\) and \(\beta_{\nu}\), which are defined as follows:

\[ \alpha_{\nu} = \int_{-B}^{B} |Q(f)|^2 \cos 2\pi \nu fT df \]  
\[ \beta_{\nu} = \int_{-\infty}^{\infty} |Q(f)|^2 \cos 2\pi \nu fT df. \]
These quantities are defined for any pulse shaping filter $Q(f)$ and will be referenced in this generality all over this report. The quantities corresponding to $|Q(f)|^2 \sin(2\pi \nu f T)$ are not considered here because they are equal to zero for symmetric functions $|Q(f)|^2$.

The next step is to compute $\alpha_\nu$ and $\beta_\nu$ for the special class of sampled $T$-time-limited pulse shaping filters. In this case, we can express these functions in terms of the sampled pulse shape $q$:

$$\alpha_\nu = q' \cdot \left( \int_{-B}^{B} P_p f \cos 2\pi \nu f T \, df \right) \cdot q = q' R_\nu^{(B)} q, \quad (2.48)$$

where $R_\nu^{(B)}$ is a $2\kappa \times 2\kappa$ real, symmetric matrix

$$R_\nu^{(B)} = ||r_{\nu i k}^{(B)}|| \quad (2.49)$$

with the components

$$r_{\nu i k}^{(B)} = \frac{1}{2\pi} \int_{-B}^{B} \exp \left( -\frac{j \pi f T (i - k)}{\kappa} \right) \cos(2\pi \nu f T) df, \quad -\kappa \leq i, k \leq \kappa - 1. \quad (2.50)$$

The above integral is independent of the pulse shape $q$ and can be computed analytically. For $\nu = 0$ we get

$$r_{0 i k}^{(B)} = \frac{B}{2\pi} \cdot \text{sinc} \left( \pi B T \frac{(i - k)}{\kappa} \right), \quad (2.51)$$

where $\text{sinc}(x) = \sin(x)/x$. For $\nu > 0$ we get

$$r_{\nu i k}^{(B)} = \frac{B}{4\pi} \left[ \text{sinc} \left( 2\pi \nu B T \left( 1 + \frac{(i - k)}{2\kappa \nu} \right) \right) + \text{sinc} \left( 2\pi \nu B T \left( 1 - \frac{(i - k)}{2\kappa \nu} \right) \right) \right], \quad \nu > 0. \quad (2.52)$$

Because of sampling, the spectrum of the generated signal is periodic with period $2\kappa / T$. To compute the total energy of the signal, we consider only the spectrum in the frequency interval $[-\kappa / T, \kappa / T]$. The definition of Eq. 2.47 has therefore to be changed to

$$\beta_\nu = \int_{-\kappa / T}^{\kappa / T} |Q(f)|^2 \cos 2\pi \nu f T df = q' R_{\nu}^{(\kappa / T)} q. \quad (2.53)$$

Combining Eqs. 2.51, 2.52 and 2.53 yields

$$R_{0}^{(\kappa / T)} = \frac{\kappa}{2\pi T} \cdot I \quad (2.54)$$

and

$$R_{\nu}^{(\kappa / T)} = 0, \quad \nu > 0. \quad (2.55)$$

We shall use Eqs. 2.48 and 2.53 in the subsequent chapters to determine the out-of-band power of the transmitted signal.
2.4 Transfer Function Of The Spectral Raised Cosine Filter

In this section we present the transfer function of the spectral raised cosine (SRC) filter, and we compute the quantities $\alpha_\nu$ and $\beta_\nu$ defined in Eqs. 2.46 and 2.47. The SRC filter is a Nyquist filter and its transfer function is band-limited, in contrast to the time-limited filter introduced in the last section.

The absolute square of the transfer function of the SRC filter is given by

$$|Q(f)|^2 = \begin{cases} 1 & |f| \leq \frac{1-\gamma}{2T} \\ \cos^2 \left[ \frac{\pi}{4\gamma} (2fT + \gamma - 1) \right] & \frac{1-\gamma}{2T} < |f| < \frac{1+\gamma}{2T} \\ 0 & \frac{1+\gamma}{2T} \leq |f| \end{cases}$$

(2.56)

where $\gamma \in [0 \ldots 1]$ is the rolloff factor.

We now compute the quantities $\alpha_\nu$ and $\beta_\nu$ defined in Sec. 2.3 for the SRC filter. Due to the structure of the transfer function of this filter, the integral of Eqs. 2.46 and 2.47 can be computed analytically:

$$\alpha_0 = \begin{cases} 2B & B \leq \frac{1-\gamma}{2T} \\ B + \frac{1-\gamma}{2T} + \frac{\gamma}{\pi T} \cos \left[ \frac{\pi}{2\gamma} (2BT - 1) \right] & \frac{1-\gamma}{2T} < B < \frac{1+\gamma}{2T} \\ 1/T & B \geq \frac{1+\gamma}{2T} \end{cases}$$

(2.57)

$$\alpha_\nu = \begin{cases} \frac{\sin 2\pi \nu BT}{\pi \nu T} B \leq \frac{1-\gamma}{2T} \\ \frac{\gamma}{2\pi T (1 - 2\nu \gamma)} \cdot \cos \left[ \frac{\pi}{2\gamma} (2BT (2\nu \gamma - 1) + 1) \right] + \frac{\gamma}{2\pi T (1 + 2\nu \gamma)} \cdot \cos \left[ \frac{\pi}{2\gamma} (2BT (2\nu \gamma - 1) - 1) \right] & \frac{1-\gamma}{2T} < B < \frac{1+\gamma}{2T} \\ + \frac{\sin 2\pi \nu BT}{2\pi \nu T} + \frac{(-1)^\nu \sin \pi \nu \gamma}{2\pi \nu T (4\nu^2 \gamma^2 - 1)} & B \geq \frac{1+\gamma}{2T} \end{cases}$$

(2.58)

and

$$\beta_\nu = \begin{cases} 1 & \nu = 0 \\ 0 & \text{otherwise.} \end{cases}$$

(2.59)

We shall use these quantities in the computation of the out-of-band power in the following chapters.
2. The System Model
Chapter 3
Optimization Criteria

In this chapter we introduce the criteria that will be used to optimize the transmitter shown in Fig. 2.2. We are interested in a transmitter that minimizes the out-of-band power for a given data rate. We shall therefore optimize the encoder/mapper with respect to bandwidth-efficiency; the particular bandwidth-efficiency criterion, the relative out-of-band power, will be defined in Sec. 3.1. As an indicator of the performance of our transmitter we shall consider the minimum Euclidean distance of the output signal. This secondary optimization criterion will be introduced in Sec. 3.2 (average-power-efficiency criterion) and in Sec. 3.3 (peak-power-efficiency criterion).

3.1 Bandwidth-Efficiency Criterion

The power spectral density $W(f)$ of a signal $s(t)$ is a function of the correlation of the symbols $u(n)$, the pulse shaping filter, and the data rate. To maximize the data rate at a given bandwidth, or to minimize the out-of-band power for a given data rate, we are interested in minimizing the fraction of the power that lies outside that bandwidth, or equally, to maximize the fractional inband power:

**Bandwidth-Efficiency Criterion:** The modulation scheme defined in Fig. 2.2 is bandwidth-efficient if the inband power is maximized:

$$\eta_{\text{max}} = \max_{u, Q(t)} \frac{\int_{-B}^{B} W(f) df}{\int_{-\infty}^{\infty} W(f) df},$$

where $B$ denotes the desired bandwidth. The optimum encoder/mapper that minimizes the out-of-band power therefore maximizes the objective function $\eta_{\text{max}}$. The graphical representation of the optimization criterion is shown in Fig. 3.1, where $\eta$ is the ratio of the shaded area to the total area between the curve and the frequency axis.

Assuming that the expected value of the encoder output is zero, i.e. $W_d(f) \equiv 0$, the bandwidth-efficiency criterion for the shift register encoder can be written as

$$\eta_{\text{max}} = \max_{u, Q(t)} \frac{\sum_{\nu=0}^{L} \left[ W_\nu \int_{-B}^{B} \frac{|Q(f)|^2}{T} \cdot \cos 2\pi \nu f T df \right]}{\sum_{\nu=0}^{L} \left[ W_\nu \int_{-\infty}^{\infty} \frac{|Q(f)|^2}{T} \cdot \cos 2\pi \nu f T df \right]},$$

where we used Eq. 2.41, and where we assumed that the function $|Q(f)|^2$ is even, which is always true for real pulses $q(t)$. Using the definitions of $\alpha_\nu$ and $\beta_\nu$ (see Eqs. 2.46
3. Optimization Criteria

and 2.47) we finally get

\[ \eta_{\text{max}} = \max_{\mu, Q(f)} \frac{\mu^t \left( \sum_{\nu=0}^{L} \alpha_{\nu} W_{\nu} \right) \mu}{\mu^t \left( \sum_{\nu=0}^{L} \beta_{\nu} W_{\nu} \right) \mu}. \]  

(3.3)

3.2 Average-Power-Efficiency Criterion

In an additive white Gaussian noise channel, for high SNR the error probability at the receiver is asymptotically a function of the minimum path distance. We say that a modulation scheme is average-power-efficient, if the power of the signal that assures a certain error probability is small. We therefore compute the energy of a coded signal and its minimum path distance.

In our analysis we restrict ourselves to pulse shaping filters whose impulse response is orthogonal to its copy that is shifted by an integer multiple of the symbol duration \( T \), i.e.

\[ \int_{-\infty}^{\infty} q(t) \cdot q(t - nT) dt = 0 \quad \forall n \neq 0. \]  

(3.4)

Eq. 3.4 assures that the pulse shaping filter does not introduce intersymbol interference. It holds for all generalized Nyquist pulses, which is stated in the following theorem:

---

**Theorem 2** A real pulse \( q(t) \) satisfies

\[ \int_{-\infty}^{\infty} q(t) \cdot q(t - nT) dt = 0 \quad \forall n \neq 0 \]  

(3.5)

if and only if its power spectral density satisfies

\[ |Q(f)|^2 = LP(f) + F(f), \]  

(3.6)
where $LP(f)$ is the ideal low-pass filter

$$LP(f) = \begin{cases} 
1 & \text{if } |f| < 1/2T \\
0 & \text{otherwise}
\end{cases}$$  \hfill (3.7)

and where $F(f)$ is a function that satisfies

$$\sum_{j=0}^{\infty} F(f - \frac{j}{T} + \frac{1}{2T}) + \sum_{j=0}^{\infty} F(-f + \frac{j}{T} + \frac{1}{2T}) = 0,$$  \hfill (3.8)

i.e. $\sum_{j=0}^{\infty} F(f - \frac{j}{T})$ is odd with respect to $1/2T$.

---

The proof of this theorem is given in Appendix B.

There are two classes of filters that satisfy Theorem 2, the $T$-time-limited and the Nyquist filters. The impulse response of the $T$-time-limited filter is zero outside the interval $[0... T]$; the sampled version of this filter was introduced in Sec. 2.3. A particular Nyquist filter that is often used in practice, the spectral raised cosine filter, was defined in Sec. 2.4. With the above definition, we can now formulate the

**Average-Power-Efficiency Criterion:** A modulation scheme defined in Fig. 2.2 is average-power-efficient if the ratio of minimum path distance to bit energy is maximized:

$$\max \frac{d^2_{\text{min}}}{2E_b}.$$  \hfill (3.9)

The average bit energy $E_b$ is defined by

$$E_b = \frac{|u|^2}{ML+1 \cdot \log_2(M)} \cdot \int_{-\infty}^{\infty} \hat{q}^2(t)dt.$$  \hfill (3.10)

We assumed that the $R$ output symbols $u_{ij}$ are equiprobable, which is true for equiprobable input symbols because of the regular structure of the encoder. Note that this does not imply that all elements of the $U$-ary output alphabet are equiprobable, because the $R$ output symbols need not be distinct! We have further assumed that the function $|Q(f)|^2$ is even, which is always true for real pulses $q(t)$.

The minimum Euclidean distance of the shift-register encoder/mapper is defined by

$$d_{\text{min}}^2 = \min_{s^{(0)}(n) \neq s^{(1)}(n)} \int_{-\infty}^{\infty} |s^{(0)}(t) - s^{(1)}(t)|^2 dt,$$  \hfill (3.11)

where $s^{(0)}(t)$ and $s^{(1)}(t)$ are the output signals of the pulse shaping filter for the two input sequences $x^{(0)}(n)$ and $x^{(1)}(n)$, see Fig. 2.2. As shown in Appendix C, for generalized Nyquist pulses (which include the $T$-time-limited pulses) Eq. 3.11 is equivalent to

$$d_{\text{min}}^2 = \int_{-\infty}^{\infty} q^2(t)dt \cdot \min_{(u^{(0)}(n)) \neq (u^{(1)}(n))} \sum_n |u^{(0)}(n) - u^{(1)}(n)|^2,$$  \hfill (3.12)

where $u(n)$ denotes the output symbol of the mapper at time $n$, see Fig. 2.2.
3.3 Peak-Power-Efficiency Criterion

Another limitation that has to be considered in practical systems is the peak power. The power amplifier of the transmitter has to be linear due to the non-constant envelope of the signal generated by our modulation scheme. The peak power of the transmitter is the maximum power it can generate in its linear operation range. In order to avoid distortion of the transmitted signal due to saturation of the power amplifier, we are interested in signals that do not exceed a certain peak power, which is defined as

\[
P_{\text{mas}} = \max_{t=mT} \left| \sum_{n=-\infty}^{\infty} u(n) \cdot q(t-nT) \right|^2
\]

With this equation, we can formulate the

**Peak-Power-Efficiency Criterion:** A modulation scheme defined in Section 2.1 is peak-power-optimal if the ratio of the minimum path distance to the peak power is maximized:

\[
\max \frac{d_{\text{min}}^2}{P_{\text{mas}}}
\]

The minimum path distance is defined in Eq. 3.12.
Chapter 4

Bandwidth-Efficient Mappers

In this chapter we shall apply the bandwidth-efficiency criterion defined in Sec. 3.1 to the transmitter defined in Fig. 2.2 to find the mapper that minimizes the out-of-band power for a given pulse shaping filter. We first determine the general form of the optimum mapper and then we present results for the single delay-cell encoder \((L=1)\) and for encoders with up to five delay cells \((L = 2 \ldots 5)\) separately.

4.1 General Form of the Bandwidth-Efficient Mapper

In this section we compute the optimum mapper \(u\) in terms of the input alphabet size \(M\), the number of delay cells of the encoder \(L\), the optimization bandwidth \(B\) and the pulse shaping filter \(Q(f)\). Since we want to find the optimum mapper for a given pulse shaping filter, we have to rewrite the bandwidth-efficiency criterion defined in Eq. 3.3 by dropping the maximization over \(Q(f)\):

\[
\eta_{\text{max}} = \max_u \frac{u^\dagger \left( \sum_{\nu=0}^L \alpha_\nu W_\nu \right) u}{u^\dagger \left( \sum_{\nu=0}^L \beta_\nu W_\nu \right) u}.
\] (4.1)

For the time-limited and the band-limited pulse shaping filters, \(\beta_\nu = 0\) for \(\nu > 0\) (see Eqs. 2.55 and 2.59). Furthermore, if the expected value of the encoder output is zero, i.e. if the components of the mapper \(u\) add up to zero, the matrix \(W_0\) is proportional to the identity matrix (see Eq. 2.39). If we normalize the pulse shaping filter gain such that \(\beta_0 = 1\), the bandwidth-efficiency criterion of Eq. 4.1 simplifies to

\[
\eta_{\text{max}} = NM \cdot \max_u \frac{u^\dagger \left( \sum_{\nu=0}^L \alpha_\nu W_\nu \right) u}{u^\dagger u}.
\] (4.2)

But the right hand part of Eq. 4.2 is a Rayleigh quotient, which is defined as a ratio of quadratics, see [Str88, p. 348]. Rayleigh's principle ([Str88, p. 349]) states that the Rayleigh's quotient is maximized by the eigenvector corresponding to the maximum eigenvalue of the matrix \(\sum \alpha_\nu W_\nu\). We have therefore proved the following
Theorem 3 The mapper \( u \) that minimizes the inband-power of a communications system as shown in Fig. 2.2 for a given pulse shaping filter \( Q(f) \) is the eigenvector of the matrix \( \sum_{\nu=0}^{L} a_\nu W_\nu \) corresponding to its maximum eigenvalue \( \lambda_{\max} \), where only eigenvectors whose components add up to zero are considered, and where the pulse shaping filter is normalized such that \( \int_{-\infty}^{\infty} |Q(f)|^2 df = 1 \). The inband power is then given by

\[
\eta_{\max} = MN \cdot \lambda_{\max}.
\] (4.3)

Note that the size of the relevant matrix is \( M^{L+1} \times M^{L+1} \).

4.2 Bandwidth-Efficient Mapper for the Encoder with a Single Delay Cell

In this section we want to find the bandwidth-efficient mapper of the single delay-cell encoder, as shown in Fig. 4.1. The trellis for this encoder for a binary input alphabet \( (M = 2) \) is shown in Fig. 4.2. There are two states and two branches leaving each state and therefore four possible output symbols.

![Figure 4.1: Schematic diagram of the single delay-cell encoder/modulator](image)

To find the optimum mapper for a general value of \( M \), we start with rewriting Theorem 3 for \( L = 1 \): We have to find the maximum eigenvalue of the matrix

\[
\alpha_0 W_0 + \alpha_1 W_1 = \frac{\alpha_0}{MN} \cdot I + \alpha_1 W_1,
\] (4.4)

where we used Eq. 2.39 to get the second identity. We can express the eigenvalues \( \lambda'_i \) of the above matrix in terms of the eigenvalues \( \lambda_i \) of the matrix \( W_1 \):

\[
\lambda'_i = \alpha_1 \cdot \lambda_i + \frac{\alpha_0}{MN}.
\] (4.5)

The maximum inband power of the single delay-cell encoder is then given by

\[
\eta = \alpha_0 + MN \alpha_1 \lambda_{\max},
\] (4.6)
where $\lambda_{\text{max}}$ is the maximum eigenvalue of the matrix $W_1$. Note that the optimum mapper, i.e. the eigenvector $u$, is independent of the pulse shaping filter $Q(f)$, because it depends only on $W_\nu$ and not on $\alpha_\nu$. This means that the mapper and the pulse shape can be optimized separately, and that the following results hold for all possible pulse shaping filters that are normalized such that $\beta_0 = 1$.

The eigenvalues of $W_1$ were computed for an input alphabet size $M = 2, 4, 8, 16$ and 32. For the two smallest alphabet sizes, analytical results could be found, otherwise numerical methods were used. The computation shows that there are $MN - 2M + 2$ eigenvalues $0$, $M - 1$ eigenvalues $-1/MN$ and $M - 1$ eigenvalues $+1/MN$. The corresponding out-of-band powers and power spectral densities (psd) are

\begin{align*}
\text{Case 1:} \\
\lambda_1 &= 0 \\
\eta_1 &= \alpha_0 \\
W_1(f) &= \frac{|Q(f)|^2}{T} \tag{4.7}
\end{align*}

\begin{align*}
\text{Case 2:} \\
\lambda_2 &= \frac{1}{MN} \\
\eta_2 &= \alpha_0 + \alpha_1 \\
W_2(f) &= \frac{|Q(f)|^2}{T} (1 + \cos 2\pi f T) = \frac{2|Q(f)|^2}{T} \cos^2 \pi f T \tag{4.8}
\end{align*}

\begin{align*}
\text{Case 3:} \\
\lambda_3 &= -\frac{1}{MN}
\end{align*}
\[ \eta_3 = \alpha_0 - \alpha_1 \quad (4.9) \]

\[ W_3(f) = \frac{|Q(f)|^2}{T} \left( 1 - \cos 2\pi f T \right) = \frac{2|Q(f)|^2}{T} \sin^2 \pi f T. \]

Which of these three mappers is optimum depends on the sign of \( \alpha_1 \), i.e. on the pulse shaping filter \( Q(f) \) and on the optimization bandwidth \( B \). For \( \alpha_1 > 0 \), we have \( \eta_2 > \eta_3 \) and vice versa. Eq. 4.8 defines the psd of the duobinary modulation scheme, which was introduced by Lender in 1963 ([Len63]). Its system polynomial equals \( 1 + D \) (see [Kab75]), which means that the signal is constructed by adding the current ('1') and the previous ('D') input symbol. This modulation scheme is optimum if \( \alpha_1 > 0 \). Eq. 4.9 defines the psd of the dicode modulation scheme, see [Mor87]. Such a signal is constructed by subtracting the previous from the current input symbol, i.e. its system polynomial equals \( 1 - D \). This modulation scheme is optimum if \( \alpha_1 < 0 \). For \( \alpha_1 = 0 \), we get the same power spectral density as for an uncoded transmission. Since we are interested in the reduction of the out-of-band power, we shall discard this case in the sequel.

The duobinary and the dicode modulation schemes, which were mainly introduced to create a spectral null at \( f = 0 \) or \( f = 1/T \), are known to have good spectral properties. We have shown, however, that they are optimum with respect to our criterion of bandwidth-efficiency.

If we compare the inband powers \( \eta_1 \ldots \eta_3 \), we see that the maximum inband power of the single delay-cell encoder can be written as

\[ \eta_{\text{max}} = \alpha_0 + |\alpha_1| \quad (4.10) \]

and the optimum power spectral density is

\[ W(f) = \frac{|Q(f)|^2}{T} (1 + \text{sign}(\alpha_1) \cdot \cos(2\pi f T)). \quad (4.11) \]

In the next two subsections, we shall compute the eigenvectors \( u \) of the matrix \( W_1 \) corresponding to the eigenvalues \( \lambda_2 \) and \( \lambda_3 \). This will lead to the optimum mapper for the single delay-cell encoder.

### 4.2.1 Optimum Bandwidth-Efficient Mapper for \( \alpha_1 > 0 \)

In this subsection we shall compute the eigenvectors of \( W_1 \) corresponding to the eigenvalue \( \lambda_2 = +1/MN \), which defines an optimum bandwidth-efficient mapper if the pulse shaping filter \( Q(f) \) and the desired bandwidth \( B \) are chosen such that \( \alpha_1 > 0 \), (see Eq. 2.46). The eigenvectors of \( W_1 \) corresponding to the eigenvalue \( +1/MN \) were computed analytically for an input alphabet size \( M = 2 \) and \( M = 4 \) and numerically for \( M = 8 \), 16 and 32 using a general purpose routine from a software library. (Example: For \( M=2 \) the eigenvector is \([1,-1,0,0] \). This means \{cf Eq. 2.15\} that if the current and the previous input symbol are equal, either a +1 or a −1 is output, depending on the value of the input symbol. Otherwise a zero is output.) The computation shows that there
are $M - 1$ different eigenvectors, whose components all add up to zero. The components of the eigenvectors furthermore show an interesting property:

If we write the mapper $u$ in the form

$$u = [u_{00}, u_{01}, \ldots, u_{0(M-1)}, u_{10}, u_{11}, \ldots, u_{(M-1)(M-1)}]'$$

(4.12)

the components of all the eigenvectors of the matrix $W_1$ corresponding to the eigenvalue '+1/MN' can be written as

$$u_{ij} = u_{oi} + u_{oj} - u_{00} \quad \forall i, j \in [0..M - 1].$$

(4.13)

We can now define $M$ complex variables

$$Y_i = u_{oi} - u_{00}/2 \quad i \in [0..M - 1]$$

(4.14)

such that

$$u_{ij} = Y_i + Y_j.$$  

(4.15)

Since there are $M$ variables $Y_i$, but only $M - 1$ eigenvectors, we have to introduce the additional constraint

$$\sum_{i=0}^{M-1} Y_i = 0,$$

(4.16)

which assures that all the eigenvectors are DC-free, i.e. that $\sum_i \sum_j u_{ij} = 0$. With the remaining $M - 1$ complex variables $Y_i$, we still have $2(M-1)$ degrees of freedom to select a bandwidth-efficient mapper. This means that we can choose any $M$ complex numbers $Y_i$ that satisfy Eq. 4.16 to find a mapper that leads to a maximum inband power if $\alpha_1 > 0$. We will eventually choose the $Y_i$ to optimize the power-efficiency. Taking a closer look at Eq. 4.15, we see that all these mappers generate the output symbol as a complex addition of the current ($Y_j$) and the previous ($Y_i$) mapped symbol. The original encoder/modulator of Fig. 4.1 can therefore be reduced to the encoder/modulator of Fig. 4.3. The difference between these two representations lies in the complexity of the mappers: The output alphabet of the mapper of Fig. 4.1 is $M^2$-ary, whereas the output alphabet of the mapper of Fig. 4.3 is only $M$-ary.

Figure 4.3: Optimum single delay-cell encoder/modulator for $\eta_2$

To conclude, if $\alpha_1 > 0$, the bandwidth-efficient modulation scheme of Fig. 4.1 reduces to the one of Fig. 4.3, where the mapper generates an output symbol $u'(n) = Y_i$ if the input symbol is $x(n) = i, \quad i = 0 \ldots M - 1$, with the restriction that $\sum_{i=0}^{M-1} Y_i = 0$. 
4.2.2 Optimum Bandwidth-Efficient Mapper for $\alpha_1 < 0$

In this subsection we shall repeat the procedure of the previous subsection to compute the eigenvectors of $W_1$ corresponding to the eigenvalue $\lambda_3 = -1/MN$. These eigenvectors define an optimum mapper if the pulse shaping filter $Q(f)$ and the desired bandwidth $B$ are chosen such that $\alpha_1 < 0$.

As before, the eigenvectors of $W_1$ corresponding to the eigenvalue $-1/MN$ were computed analytically or numerically for an alphabet size $M = 2, 4, 8, 16$ and $32$. There are again $M - 1$ eigenvectors, whose components all add up to zero. They also show the interesting property that their components can be written as

$$u_i = -u_0i + u_0j + u_{00} \quad \forall \ i, j \in [0..M - 1].$$

We can now define $M$ complex variables

$$Z_i = u_{0i}$$

such that

$$u_{ij} = -Z_i + Z_j + Z_0.$$  \hspace{1cm} (4.19)

Since there are only $M - 1$ eigenvectors but $M$ variables $Z_i$, we have to make sure that all eigenvectors are DC-free. This requirement is satisfied if

$$Z_0 = 0.$$  \hspace{1cm} (4.20)

![Figure 4.4: Optimum single delay-cell encoder/modulator for $\eta_3$](image)

Figure 4.4: Optimum single delay-cell encoder/modulator for $\eta_3$

Note that this mapper is bandwidth-efficient independent of the choice of $Z_i$. We can therefore use the remaining $2(M - 1)$ degrees of freedom to perform a secondary optimization with respect to power-efficiency, as will be shown in Chapter 7.

Eq. 4.19 states that this mapper subtracts the previous symbol ($Z_i$) from the current symbol ($Z_j$). Hence the original encoder/modulator of Fig. 4.1 can be transformed to the encoder/modulator of Fig. 4.4. As before, the output alphabet size of the original mapper can be reduced from $M^2$ to $M$.

To summarize, if $\alpha_1 < 0$, the bandwidth-efficient modulation scheme of Fig. 4.1 reduces to the one of Fig. 4.4, where the mapper generates an output symbol $u'(n) = Z_i$ if the input symbol is $x(n) = i, \ i = 0 \ldots M - 1$, with the restriction that $Z_0 = 0$. 

4.2.3 Summary: Optimum Bandwidth-Efficient Mapper for the Single Delay-Cell Encoder

In the last two subsections we showed that by inspection of the eigenvectors of the matrix $W_1$ the complexity of the mapper of Fig. 4.1 can be dramatically reduced. Comparing the resulting encoder/modulators for $\alpha_1 > 0$ (Fig. 4.3) and $\alpha_1 < 0$ (Fig. 4.4), it can be seen that the only difference is whether the previous symbol is added or subtracted. The two mappers can therefore be combined, as shown in Fig. 4.5.

![Figure 4.5: Optimum single delay-cell encoder/modulator](image)

Depending on the sign of $\alpha_1$, the previous symbol is added to or subtracted from the current symbol. The delay cell and the adder of Fig. 4.5 represent an FIR-filter of length one with frequency response

$$1 + \text{sign}(\alpha_1) \cdot \exp(-j2\pi fT). \quad (4.21)$$

It is therefore possible to represent the encoder/modulator of Fig. 4.5 as a mapper followed by a single filter that combines the filter of Eq. 4.21 with the pulse shaping filter $Q(f)$. The result of this operation is shown in Fig. 4.6.

![Figure 4.6: Filter representation of the optimum linear quadrature-amplitude (QAM) modulator with memory one](image)

It should be emphasized that the mappers of Figs. 4.5 and 4.6 depend on the sign of $\alpha_1$. If $\alpha_1 > 0$, the mapper generates an output symbol $u'(n) = Y_i$ if the input symbol is $x(n) = i$, $i = 0 \ldots M - 1$, with the restriction that $\sum_{i=1}^{M} Y_i = 0$. If $\alpha_1 < 0$ the mapper generates an output symbol $u'(n) = Z_i$ if the input symbol is $x(n) = i$, $i = 0 \ldots M - 1$, with the restriction that $Z_0 = 0$.

It is again noticed that all these mappers lead to a bandwidth-efficient transmitter. The remaining $2(M - 1)$ degrees of freedom will be used in in Chapter 7 for a secondary optimization of the power efficiency.
4.3 Bandwidth-Efficient Mapper for the Encoder with Two Delay Cells

In this section we deal with a communication system with an $M$-ary input alphabet, an encoder consisting of two delay cells, a mapper, and a pulse shaping filter, see Fig. 4.7. The trellis of this encoder is shown in Fig. 4.8 for a binary input alphabet, $M=2$.

![Figure 4.7: Schematic diagram of the two delay-cells encoder/modulator](image)

Our goal is to find a mapper that combines the present and the last two symbols to generate one of $R = M^3$ possible output signals $u_{ij}$ such that the inband power of the signal $s(t)$ is maximized. We derive the criterion for the output symbols $u_{ij}$ that assures an optimal spectrum. We show that in contrast to the single delay-cell encoder the optimum mapper of the encoder with two delay-cells depends on the pulse shaping filter.

![Figure 4.8: Trellis of the binary-input two delay-cells encoder](image)

To find the optimum mapper for the encoder with two delay cells, we rewrite Theo-
rem 3 for $L = 2$: We have to compute the maximum eigenvalue of the matrix

$$W_2a_2 + W_1a_1 + W_0a_0 = W_2a_2 + W_1a_1 + \frac{a_0}{MN} \cdot I. \quad (4.22)$$

The computation shows that there are six distinct eigenvalues that correspond to the following values for the inband power:

$$\begin{align*}
\eta_1 &= \alpha_0 \\
\eta_2 &= \alpha_0 + \alpha_1 \\
\eta_3 &= \alpha_0 - \alpha_1 \\
\eta_4 &= \alpha_0 - \alpha_2 \\
\eta_5 &= \alpha_0 + \frac{\alpha_2}{2} - \frac{1}{2}\sqrt{\alpha_2^2 + 8\alpha_1^2} \\
\eta_6 &= \alpha_0 + \frac{\alpha_2}{2} + \frac{1}{2}\sqrt{\alpha_2^2 + 8\alpha_1^2} 
\end{align*} \quad (4.23)$$

where we normalized the pulse shaping filter such that $\beta_0 = 1$. The eigenvectors corresponding to the six eigenvalues are

$$\begin{align*}
\mathbf{u}_1 &= [-r_1 - r_2, r_1 + r_2, 3r_1, -3r_1, r_1 + r_2, -r_1 - r_2, 3r_1, 3r_1]^T \\
\mathbf{u}_2 &= [1, 0, -1, 0, 0, -1, 0, 1]^T \\
\mathbf{u}_3 &= [0, 1, 0, -1, -1, 0, 1, 0]^T \\
\mathbf{u}_4 &= [0, 1, 0, 1, -1, 0, -1, 0]^T \\
\mathbf{u}_5 &= [2b, b + c, -2c, -b - c, b + c, 2c, -b - c, -2b]^T \\
\mathbf{u}_6 &= [2b, b - a, 2a, a - b, b - a, -2a, a - b, -2b]^T 
\end{align*} \quad (4.24)$$

where $r_1$ and $r_2$ are any two real numbers, and where the real numbers $a$, $b$ and $c$ depend on the pulse shaping filter and on the optimization bandwidth:

$$\begin{align*}
a &= \sqrt{\alpha_2^2 + 8\alpha_1^2} - 3\alpha_1 \\
b &= \alpha_2 - \alpha_1 \\
c &= \sqrt{\alpha_2^2 + 8\alpha_1^2} + 3\alpha_1 
\end{align*} \quad (4.25)$$

The corresponding power spectral densities are

$$\begin{align*}
W_1(f) &= \frac{|Q(f)|^2}{T} \\
W_2(f) &= \frac{|Q(f)|^2}{T}(1 + \cos 2\pi f T) = 2\frac{|Q(f)|^2}{T} \cos^2 \pi f T \\
W_3(f) &= \frac{|Q(f)|^2}{T}(1 - \cos 2\pi f T) = 2\frac{|Q(f)|^2}{T} \sin^2 \pi f T \\
W_4(f) &= \frac{|Q(f)|^2}{T}(1 - \cos 4\pi f T) \\
W_5(f) &= \frac{|Q(f)|^2}{T} \left(3\left(b^2 + c^2\right) + bc + 2(b^2 - c^2) \cos 2\pi f T\right) \quad (4.26)
\end{align*}$$
Comparing the power spectral densities (psd) of the encoders with a single and with two delay cells (Eqs. 4.7, 4.8, 4.9 and 4.26), it can be seen that the psd $W_1 \ldots W_3$ are identical. Among the other three power spectral densities, $W_4$ belongs to another well known modulation scheme. It is called modified duobinary and was first introduced by Lender [Len68] in 1968. Its system polynomial equals $1 - D^2$, i.e. we subtract the second to last from the current symbol. $W_5(f)$ and $W_6(f)$ correspond to new modulation schemes that have not been found in the literature.

To find the optimum mapper, we have to find out which of the six objective functions of Eq. 4.23 is maximum for a given pulse shaping filter and a given optimization bandwidth. Since we assume that an additional delay cell decreases the out-of-band power, the optimum modulation scheme of interest corresponds to $\eta_4 \ldots \eta_6$. Because it is obvious that $\eta_5 \leq \eta_6$, the maximum inband power of the encoder with two delay cells is therefore given by

$$\eta_{\text{max}}(B) = \max_{i,Q(f)} \eta_i(B, Q(f)), \quad i \in [4, 6].$$

This maximization will be carried out in Chapter 6, where the pulse shaping filter will be optimized. For illustration purposes, Fig. 4.9 shows a possible realization of the transmitter corresponding to $\eta_6$ for a binary input alphabet. We see that the components of the vector $u(n)$ appear as output symbols of the mapper.

Figure 4.9: Transmitter with Two Delay Cells Corresponding to $\eta_6$, Binary Input

The results presented in this section were computed for a binary input alphabet, $M = 2$. The symbolic programming packages used to find these analytical results failed
at the attempt to compute the eigenvectors of the matrix \( \alpha_2 W_2 + \alpha_1 W_1 \) for \( M = 4 \) due to the size of this matrix, which is \( 64 \times 64 \). However, it was possible to find the structure of this mapper intuitively, and the result could be verified analytically. This mapper will be presented in Sec. 4.7.

### 4.4 Bandwidth-Efficient Mapper for the Encoder with Three Delay Cells

According to Theorem 3, the optimum encoder with three delay-cells is proportional to the eigenvector corresponding to the largest eigenvalue of the matrix \( \sum_{i=0}^{3} \alpha_i W_i \). The attempt to compute the eigenvalues and eigenvectors of this matrix with the symbolic programming packages MATHEMATICA and MACSYMA failed even for the simplest case with a binary input alphabet. Because of the size of the matrix \((16 \times 16)\), these programs could not find the eigenvalues. Hence we optimized the mapper numerically for a given pulse shape and a fixed optimization bandwidth. Inspecting the numerical result, a special structure of the eigenvector could be observed. Using this result, the size of the original problem could be reduced dramatically, such that finally the analytical form of all eigenvectors could be found.

As stated in Sec. 4.3, the eigenvalues of the encoder with a single delay-cell are also eigenvalues of the encoder with two delay cells. The same is true for three delay cells, i.e. the eigenvalues found for \( L = 2 \) are also solutions for \( L = 3 \). The uncoded case \((\eta_1, \text{see Eq. 4.23})\) corresponds to a four-fold eigenvalue, whereas the eigenvalues of the duobinary \((\eta_2)\) and of the AMI-coded mapper \((\eta_6)\) each have multiplicity two. The other mappers \((\eta_4, \eta_5)\) all correspond to single eigenvalues. In addition, four new eigenvalues are found:

\[
\begin{align*}
\lambda_7 &= \frac{y+1}{2R} - \frac{y^2 - 2y + 4x^2 - 8x + 5}{2R} \\
\lambda_8 &= \frac{y+1}{2R} + \frac{y^2 - 2y + 4x^2 - 8x + 5}{2R} \\
\lambda_9 &= \frac{y+1}{2R} - \frac{y^2 - 2y + 4x^2 - 8x + 5}{2R} \\
\lambda_{10} &= \frac{y+1}{2R} + \frac{y^2 - 2y + 4x^2 - 8x + 5}{2R} \quad \text{(4.28)}
\end{align*}
\]

where \( R = M^{L+1} \), \( x = \frac{\alpha_3}{\alpha_1} \) and \( y = \frac{\alpha_3}{\alpha_1} \). The resulting inband powers \( \eta_7 \ldots \eta_{10} \) are:

\[
\begin{align*}
\eta_7 &= \alpha_0 - \frac{\alpha_3}{2} - \frac{\alpha_3}{2} - \frac{1}{2} \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2} \\
\eta_8 &= \alpha_0 - \frac{\alpha_3}{2} + \frac{\alpha_3}{2} - \frac{1}{2} \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2} \\
\eta_9 &= \alpha_0 + \frac{\alpha_3}{2} + \frac{\alpha_3}{2} - \frac{1}{2} \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2} \\
\eta_{10} &= \alpha_0 + \frac{\alpha_3}{2} - \frac{\alpha_3}{2} + \frac{1}{2} \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2}. \quad \text{(4.29)}
\end{align*}
\]

The mappers corresponding to these eigenvalues are defined as follows:

\[
\begin{align*}
\mathbf{u}_7 &= [0, c, r_1, r_1 + c, -r_1, c - r_1, 0, c, -c, 0, r_1 - c, r_1, -c - r_1, -r_1, -c, 0]' \quad \text{(4.30)}
\end{align*}
\]
\[ Y_8 = [0, d, r_1, r_1 + d, -r_1, d - r_1, 0, d, -d, 0, r_1 - d, r_1, -d - r_1, -r_1, -d, 0]' \]
\[ Y_9 = [e, r_1, e - r_1, 0, e - r_1, 0, e - 2r_1, -r_1, r_1, 2r_1 - e, 0, r_1 - e, 0, r_1 - e, -r_1, -e]' \]
\[ Y_{10} = [f, r_1, f - r_1, 0, f - 2r_1, -r_1, f, 0, r_1 - f, 0, r_1 - f, -r_1, -f]' \]

where \( r_1 \) is any real number, and where \( c, d, e \) and \( f \) are defined by

\[
c = \frac{r_1}{2(\alpha_2 - \alpha_1)} \left(-\alpha_1 + \alpha_3 + \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2}\right)
\]
\[
d = \frac{r_1}{2(\alpha_2 - \alpha_1)} \left(-\alpha_1 + \alpha_3 - \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2}\right)
\]
\[
e = \frac{r_1}{2(\alpha_2 + \alpha_1)} \left(\alpha_1 + 2\alpha_2 + \alpha_3 - \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2}\right)
\]
\[
f = \frac{r_1}{2(\alpha_2 + \alpha_1)} \left(\alpha_1 + 2\alpha_2 + \alpha_3 + \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2}\right).
\]

To determine the optimum mapper, we have to compute \( \eta_1..\eta_{10} \) for a given pulse shape and a given optimization bandwidth. Then the mapper corresponding to the largest \( \eta \) is selected. Because we expect that the best mapper is one of \( u_7..u_{10} \) and because for any pulse shape \( \eta_6 \geq \eta_7 \) and \( \eta_{10} \geq \eta_9 \), the optimum mapper is either \( u_9 \) or \( u_{10} \), depending on the pulse shaping filter and the optimization bandwidth. The optimization of the pulse shaping filter will be performed in Chapter 6, and the structure of the mapper with a quaternary input alphabet \((M = 4)\) will be presented in Sec. 4.7.

### 4.5 Bandwidth-Efficient Mapper for the Encoder with Four Delay Cells

Theorem 3 states that the optimum mapper of the four-delay-cell encoder is proportional to the eigenvector of the largest eigenvalue of the matrix \( \sum_{\nu=0}^{\nu} \alpha_{\nu} W_{\nu} \). The eigenvalues were computed with the same combined numerical/analytical method as described in Section 4.4. As expected, the ten eigenvalues of the solution for \( L = 3 \) are also eigenvalues of the matrix for \( L = 4 \). Furthermore, five new eigenvalues are found. The corresponding out-of-band powers are:

\[
\eta_{11} = \alpha_0 - \frac{\alpha_2 + \alpha_4}{2} - \frac{1}{2} \sqrt{\alpha_4^2 - 2\alpha_2\alpha_4 + 4\alpha_3^2 - 8\alpha_1\alpha_3 + 4\alpha_2^2 + 4\alpha_1^2}
\]
\[
\eta_{12} = \alpha_0 - \frac{\alpha_2 + \alpha_4}{2} + \frac{1}{2} \sqrt{\alpha_4^2 - 2\alpha_2\alpha_4 + 4\alpha_3^2 - 8\alpha_1\alpha_3 + 4\alpha_2^2 + 4\alpha_1^2}
\]
\[
\eta_{13} = \alpha_0 + \frac{\alpha_2 + \alpha_4}{3} + 2\sqrt{\frac{p}{3}} \cdot \cos(\varphi/3)
\]
\[
\eta_{14} = \alpha_0 + \frac{\alpha_2 + \alpha_4}{3} + 2\sqrt{\frac{p}{3}} \cdot \cos(\varphi/3 + \frac{2\pi}{3})
\]
\[
\eta_{15} = \alpha_0 + \frac{\alpha_2 + \alpha_4}{3} + 2\sqrt{\frac{p}{3}} \cdot \cos(\varphi/3 + \frac{4\pi}{3}),
\]

where

\[
p = \frac{1}{3} \left(\alpha_4^2 - \alpha_2\alpha_4 + 3\alpha_3^2 + 6\alpha_1\alpha_3 + 7\alpha_2^2 + 9\alpha_1^2\right)
\]
4.6. Bandwidth-Efficient Mapper for the Encoder with Five Delay Cells

\[ q = \frac{1}{27} \left( 2\alpha_4^3 - 3\alpha_2\alpha_4^2 + \alpha_4 \left( 9\alpha_2^3 + 18\alpha_1\alpha_3 + 15\alpha_2^2 - 27\alpha_1^2 \right) + 9\alpha_2\alpha_3^2 + 126\alpha_1\alpha_2\alpha_3 - 34\alpha_2^3 + 135\alpha_2^2\alpha_2 \right) \] (4.34)

and

\[ \cos \varphi = \frac{q}{2\sqrt{27}}. \] (4.35)

The eigenvalues corresponding to \( \eta_{13} \ldots \eta_{15} \) are real if

\[ D = \left( \frac{q}{2} \right)^2 - \left( \frac{p}{3} \right)^3 < 0. \] (4.36)

As will be shown in Chapter 6, for all pulse shapes that were considered, \( \eta_{13} \) proved to be maximum. The optimum mapper has the form

\[ W_{13} = \begin{bmatrix} a, b, c, -a + b + c, 3a - 2b - 2c, 2a - b - c, 2b - c, a - b - c, c, \\ -a + b + c, 2c - a, -2a + 2b + 2c, 2a - 2b - c, a - b - c, a - 2b, -b, b, 2b - a, \\ b + c - a, c + 2b - 2a, 2a - b - 2c, a - 2c, a - b - c, -c, -a + b + c, \\ -2a + 2b + c, -a + b + 2c, -3a + 2b + 2c, a - b - c, -c, -b, -a \end{bmatrix}, \] (4.37)

where \( a, b, c \) are the components of the eigenvector \( x = [a, b, c]' \) of the matrix \( W_{13} \)

\[ W_{13} = \begin{bmatrix} \alpha_4 + 2\alpha_3 + \alpha_1 & -\alpha_3 - \alpha_2 + \alpha_1 & -\alpha_3 + \alpha_2 \\ \alpha_3 + 2\alpha_2 & -\alpha_3 - \alpha_2 + \alpha_1 & -\alpha_2 + \alpha_1 \\ \alpha_4 + \alpha_3 - \alpha_2 + 2\alpha_1 & -\alpha_4 + \alpha_2 & -\alpha_3 + 2\alpha_2 - 2\alpha_1 \end{bmatrix} \] (4.38)

corresponding to the largest eigenvalue of \( W_{13} \).

The optimum mapper for a quaternary input alphabet (\( M = 4 \)) will be given in Sec. 4.7.

4.6 Bandwidth-Efficient Mapper for the Encoder with Five Delay Cells

As stated in Theorem 3, the optimum mapper of the encoder with five delay-cells is proportional to the eigenvector of the maximum eigenvalue of the matrix \( \sum_{\nu=0}^{\nu=5} \alpha_v W_v \).

As before, we used a combined numerical/analytical approach to reduce the size of the original problem and finally obtained two 3 \( \times \) 3 matrices that could be solved analytically.

In addition to the 15 eigenvalues of the encoder with four delay-cells, six new eigenvalues were found. They correspond to the inband powers \( \eta_{16} \ldots \eta_{21} \), that are defined by

\[ \eta_{16} = \alpha_0 + \frac{\alpha_1 + \alpha_3 + \alpha_5}{3} + 2\sqrt{\frac{p_{16}}{3}} \cdot \cos(\varphi_{16}/3) \]
\[ \eta_{17} = \alpha_0 + \frac{\alpha_1 + \alpha_3 + \alpha_5}{3} + 2\sqrt{\frac{p_{17}}{3}} \cdot \cos(\varphi/3 + 2\pi/3) \]
4. Bandwidth-Efficient Mappers

\[ \eta_{18} = \alpha_0 + \frac{\alpha_1 + \alpha_3 + \alpha_5}{3} + 2\sqrt{\frac{p_{16}}{3}} \cdot \cos\left(\varphi/3 + \frac{4\pi}{3}\right) \]

\[ \eta_{19} = \alpha_0 - \frac{\alpha_1 + \alpha_3 + \alpha_5}{3} + 2\sqrt{\frac{p_{19}}{3}} \cdot \cos\left(\varphi_{19}/3\right) \]  

\[ \eta_{20} = \alpha_0 - \frac{\alpha_1 + \alpha_3 + \alpha_5}{3} + 2\sqrt{\frac{p_{19}}{3}} \cdot \cos\left(\varphi_{19}/3 + \frac{2\pi}{3}\right) \]

\[ \eta_{21} = \alpha_0 - \frac{\alpha_1 + \alpha_3 + \alpha_5}{3} + 2\sqrt{\frac{p_{19}}{3}} \cdot \cos\left(\varphi_{19}/3 + \frac{4\pi}{3}\right), \]

where

\[ p_{16} = \frac{1}{3} \left( \alpha_5^2 - (\alpha_3 + \alpha_1)\alpha_5 + 3\alpha_4^2 + 6\alpha_1\alpha_4 + 4\alpha_3^2 + (6\alpha_2 - \alpha_1)\alpha_3 + 6\alpha_2^2 + 6\alpha_1\alpha_2 + 7\alpha_1^2 \right) \]

\[ q_{16} = \frac{1}{27} \left( 2\alpha_5^3 - (3\alpha_3 + 3\alpha_1)\alpha_5^2 + (9\alpha_4^2 + 18\alpha_1\alpha_4 + 6\alpha_3^2 + (18\alpha_2 + 12\alpha_1)\alpha_3 - 9\alpha_2^2 - 36\alpha_1\alpha_2 - 12\alpha_2^2)\alpha_5 + (9\alpha_3 - 18\alpha_1)\alpha_4^2 + ((54\alpha_2 + 72\alpha_1)\alpha_3 + 54\alpha_2^2 + 54\alpha_1\alpha_2 - 36\alpha_2^2)\alpha_4 - 16\alpha_3^2 + (6\alpha_1 - 36\alpha_2)\alpha_5^2 + (9\alpha_2^2 + 90\alpha_1\alpha_2 + 69\alpha_1^2)\alpha_5 + 72\alpha_1\alpha_2^2 + 72\alpha_1^2\alpha_2 - 7\alpha_1^3 \right) \]

\[ \cos(\varphi_{16}) = \frac{q_{16}}{2\sqrt{p_{16}^4}} \]

and

\[ p_{19} = \frac{1}{3} \left( \alpha_5^2 - (\alpha_3 + \alpha_1)\alpha_5 + 3\alpha_4^2 - 6\alpha_1\alpha_4 + 4\alpha_3^2 - (6\alpha_2 + \alpha_1)\alpha_3 + 6\alpha_2^2 - 6\alpha_1\alpha_2 + 7\alpha_1^2 \right) \]

\[ q_{19} = -\frac{1}{27} \left( 2\alpha_5^3 + (3\alpha_3 + 3\alpha_1)\alpha_5^2 + (9\alpha_4^2 - 18\alpha_1\alpha_4 + 6\alpha_3^2 + (18\alpha_2 + 12\alpha_1)\alpha_3 - 9\alpha_2^2 + 36\alpha_1\alpha_2 - 12\alpha_2^2)\alpha_5 + (9\alpha_3 - 18\alpha_1)\alpha_4^2 + ((54\alpha_2 - 72\alpha_1)\alpha_3 + 54\alpha_2^2 + 54\alpha_1\alpha_2 + 36\alpha_2^2)\alpha_4 - 16\alpha_3^2 + (6\alpha_1 + 36\alpha_2)\alpha_5^2 + (9\alpha_2^2 - 90\alpha_1\alpha_2 + 69\alpha_1^2)\alpha_5 + 72\alpha_1\alpha_2^2 - 72\alpha_1^2\alpha_2 - 7\alpha_1^3 \right) \]

\[ \cos(\varphi_{19}) = \frac{q_{19}}{2\sqrt{p_{19}^4}} \]

For all pulse shapes that were considered, either \( \eta_{16} \) or \( \eta_{19} \) is maximum. If \( \eta_{16} > \eta_{19} \), the optimum mapper is

\[ u_{16} = [a, b, c, -a + b + c, 2a - b - c, a - c, -a + b, 0, 2a - b - c, a - c, a - b, 0, 3a - 2b - 2c, 2a - b - 2c, 2a - b - c, a - b - c, c, -a + b + c, -a + 2c, -2a + b + 2c, a - b - c, c, a - b, 0, c - b, c - a, a - b, 0, c - b, c - a, a - b, 0, a - b - c, a - b, a - b, a - b, a - b, a - b, 0, -a + b, a - c, b - c, a - b, a - c, -a + b, a - c, b - c, 0, -a + b, 2a - b - 2c, a - b - c, -a + b, -a + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -2a + b + c, -a + b, -a + c, -2a + b + c, a - b - c, -c, -b, -a]' \]
4.7. Optimum Mapper for a Quaternary Input Alphabet

where \( a, b, c \) are the components of the eigenvector \( \mathbf{x} = [a, b, c]' \) of the matrix \( \mathbf{W}_{16} \)

\[
\mathbf{W}_{16} = \begin{bmatrix}
\alpha_5 + 2\alpha_4 + \alpha_3 + \alpha_1 & -\alpha_5 - \alpha_4 + \alpha_2 + \alpha_1 & -\alpha_4 + \alpha_2 \\
\alpha_4 + 2\alpha_3 + \alpha_2 & -\alpha_4 - \alpha_3 + \alpha_1 & -\alpha_3 + \alpha_1 \\
\alpha_5 + \alpha_4 + \alpha_2 + \alpha_1 & -\alpha_5 + \alpha_1 & -\alpha_4 + \alpha_3 - \alpha_1 
\end{bmatrix}
\]

(4.47)
corresponding to the largest eigenvalue of \( \mathbf{W}_{16} \). If \( \eta_{16} \) \( < \eta_{19} \), the optimum mapper is

\[
\mathbf{H}_{19} = [0, a, b, a + b, c, a + c, b + c, a + b + c, -c, a - c, b - c, a + b - c, 0, a, b, a + b, -b, a - b, 0, a, -a + b, a - b + c, c, a + c, -b - c, a - b - c, -c, a - c, -b, a - b, b - c, -a, 0, -a + b, b - a - b, -b, -a, 0, -a - b + c, -b + c, -a + c, -c, -a - b, -b, -a, 0]',
\]

(4.48)
where \( a, b, c \) are the components of the eigenvector \( \mathbf{x} = [a, b, c]' \) of the matrix \( \mathbf{W}_{19} \)

\[
\mathbf{W}_{19} = \begin{bmatrix}
-\alpha_5 & -\alpha_4 + \alpha_1 & -\alpha_3 + \alpha_2 \\
-\alpha_4 + \alpha_1 & -\alpha_3 & -\alpha_2 + \alpha_1 \\
-\alpha_3 + \alpha_2 & -\alpha_2 + \alpha_1 & -\alpha_1 
\end{bmatrix}
\]

(4.49)
corresponding to the largest eigenvalue of \( \mathbf{W}_{19} \).

Numerical results of the mapper with a quaternary input alphabet \( (M = 4) \) will be given in Sec. 4.7.

4.7 Optimum Mapper for a Quaternary Input Alphabet

The results of the previous sections are valid for encoders with a binary input alphabet. We now show how these results can be used to compute the optimum mapper with a quaternary input alphabet.

The optimum mapper for any input alphabet size \( M \) is proportional to the eigenvector of the largest eigenvalue of the matrix \( \sum_{\nu=0}^{L} \alpha_{\nu} \mathbf{W}_{\nu} \) (see Theorem 3). Because of the size of this matrix, which is \( M^{L+1} \times M^{L+1} \), no analytical results could be found for \( M > 2 \) and \( L \geq 2 \). We therefore try to express the optimum mapper for a quaternary input alphabet in terms of the binary mapper.

The optimum mapper for the binary input alphabet \( (M = 2) \) is always real, because the matrix \( \sum_{\nu=0}^{L} \alpha_{\nu} \mathbf{W}_{\nu} \) is symmetric. An intuitive procedure to construct the quaternary mapper is to extend the binary real mapper to the complex plane. We do this by splitting up the quaternary input data stream into two binary data streams by feeding the even input bits into the first and the odd input bits into the second binary encoder (see Fig. 4.10). Each of these two binary encoders/mappers has the structure that was computed for \( M = 2 \). The output of the first encoder then represents the real part and
the output of the second encoder the imaginary part of the baseband signal. In formal notation, we can express the quaternary mapper \( u' \) in terms of the binary mapper \( u \) as follows:

\[
u'(\sigma) = u(\sigma_r) + j \cdot u(\sigma_i), \tag{4.50}\]

where \( j^2 = -1 \) and where \( \sigma_r \) and \( \sigma_i \) are the states of binary shift registers. They are defined as

\[
\sigma_r = \sum_{k=0}^{L} b_{2k} 2^k \quad \sigma_i = \sum_{k=0}^{L} b_{2k+1} 2^k \quad \sigma_r, \sigma_i \in [0, 2^{L+1} - 1], \tag{4.51}\]

where the coefficients \( b_k \) are the binary representation of the state \( \sigma \):

\[
\sigma = \sum_{k=0}^{2L+1} b_k 2^k \quad \sigma \in [0, 4^{L+1} - 1]. \tag{4.52}\]

For \( L \) up to three, the optimality of the mapper \( u' \) was verified analytically, i.e. it was shown that \( u' \) is indeed an eigenvector of the matrix \( \sum_{\nu=0}^{L} \alpha_{\nu} W_{\nu} \). A numerical optimization procedure showed that it is also the eigenvector corresponding to the largest eigenvalue of this matrix. This is the justification of the intuitive idea of finding the optimum mapper for a quaternary input alphabet.
Chapter 5

Representation of the Trellis Encoder as an FIR Filter

In this chapter we show that the optimum mappers of the encoders for one to five delay cells can be represented as a finite impulse response (FIR) filter. With this approach, the structure of the transmitter reduces to a cascade of two filters. We further present a method that can be used to optimize the filter coefficients directly, such that the complexity of the original problem can be drastically reduced.

5.1 Representation of the Optimum Mapper as an FIR Filter

As we have already seen in Sec. 4.2, the optimum single delay-cell encoder can be represented as an FIR filter with two taps. This is the motivation of the attempt to express the optimum mappers for two and more delay cells in terms of FIR filter coefficients. It is not obvious that this problem can be accomplished, because the mapper coefficients depend on the pulse shaping filter and on the optimization bandwidth; even so, the encoder/mapper may be nonlinear.

We consider an FIR filter with \( L \) delay cells and \( L + 1 \) real coefficients \( g_0 \ldots g_L \), as shown in Fig. 5.1. We then derive the FIR-filter that is equivalent to the encoder/mapper by comparing the filter and the mapper coefficients. Starting with a binary input alphabet and two delay cells, we generalize this result for encoders with up to five delay cells.

Fig. 4.9 shows the mapper \( u_6 \) of the encoder with two delay cells. As will be shown in Chapter 6, this mapper is optimum if a time-limited or a Nyquist pulse shaping filter \( Q(f) \) is chosen. To find the filter coefficients \( g_0, g_1 \) and \( g_2 \) corresponding to this mapper, we compute all 8 possible output filter signals \( u_f(n) \) as function of the input signals \( x(n), x(n-1) \) and \( x(n-2) \) and compare them with the output signals \( u_m(n) \) of the mapper (see Table 5.1).

The output of the mapper and the output of the FIR filter have to be equal, i.e. \( u_f(n) = u_m(n) \). For \( L = 2 \) we can write this equation in matrix notation

\[
 X_2 \cdot g = u_6, \quad (5.1)
\]
where $g$ is the vector of the FIR filter coefficients

$$g = [g_0 \ldots g_L]'$$

(5.2)

and where $X_2$ is the matrix constructed from the three left-most columns of Table 5.1. In general, $X_L$ is the $2^{L+1} \times (L + 1)$ matrix with the elements

$$x_{L_{ij}} = -1\left[\frac{i-1}{2^L}\right] \quad i = 1 \ldots 2^{L+1}, \quad j = 1 \ldots L + 1,$$

(5.3)

where $[x]$ denotes the smallest integer that is greater than or equal to $x$. For $L = 2$ and the mapper $\mu_6$ we get a system with eight equations and three unknowns. The solutions

<table>
<thead>
<tr>
<th>$x(n)$</th>
<th>$x(n-1)$</th>
<th>$x(n-2)$</th>
<th>$u_f(n)$</th>
<th>$u_m(n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-g_0 - g_1 - g_2$</td>
<td>$2b$</td>
</tr>
<tr>
<td>$+1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$+g_0 - g_1 - g_2$</td>
<td>$b-a$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$+1$</td>
<td>$-1$</td>
<td>$-g_0 + g_1 - g_2$</td>
<td>$2a$</td>
</tr>
<tr>
<td>$+1$</td>
<td>$+1$</td>
<td>$-1$</td>
<td>$+g_0 + g_1 - g_2$</td>
<td>$a-b$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$-1$</td>
<td>$+1$</td>
<td>$-g_0 - g_1 + g_2$</td>
<td>$b-a$</td>
</tr>
<tr>
<td>$+1$</td>
<td>$-1$</td>
<td>$+1$</td>
<td>$+g_0 - g_1 + g_2$</td>
<td>$-2a$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$+1$</td>
<td>$+1$</td>
<td>$-g_0 + g_1 + g_2$</td>
<td>$a-b$</td>
</tr>
<tr>
<td>$+1$</td>
<td>$+1$</td>
<td>$+1$</td>
<td>$+g_0 + g_1 + g_2$</td>
<td>$-2b$</td>
</tr>
</tbody>
</table>

Table 5.1: Mapper and FIR-filter coefficients for $\mu_6$
are

\begin{align}
  g_0 &= -\frac{a+b}{2} \\
  g_1 &= a - b \\
  g_2 &= \frac{a+b}{2}.
\end{align}

The resulting FIR-filter with 3 taps is shown in Fig. 5.2.

\begin{figure}[h]
\centering
\includegraphics[width=0.7\textwidth]{fig5_2}
\caption{FIR-filter corresponding to \( \eta_0 \)}
\end{figure}

The same procedure was repeated for several mappers with up to five delay cells. The resulting filter coefficients for the mappers \( \eta_2, \eta_3, \eta_6, \eta_8, \eta_{10}, \eta_{13}, \eta_{16} \) and \( \eta_{19} \) are shown in Table 5.2, where \( r_1 \) is an arbitrary real number, and where \( a \) and \( b \) are defined in Eq. 4.25, \( d \) and \( f \) are defined in Eq. 4.31, \( [a_{13}, b_{13}, c_{13}]^T \) is the eigenvector of the matrix \( W_{13} \) defined in Eq. 4.38, \( [a_{16}, b_{16}, c_{16}]^T \) is the eigenvector of the matrix \( W_{16} \) defined in Eq. 4.47 and \( [a_{19}, b_{19}, c_{19}]^T \) is the eigenvector of the matrix \( W_{19} \) defined in Eq. 4.49. In Table 5.2 we present only mappers that lead to minimum out-of-band power for a time-limited or Nyquist pulse shaping filter \( Q(f) \), as will be shown in Chapter 6.

Note that the filter coefficients corresponding to \( \eta_2, \eta_6, \eta_{10}, \eta_{13} \) and \( \eta_{16} \) are real and symmetric, whereas the coefficients corresponding to \( \eta_3, \eta_8 \) and \( \eta_{19} \) are real and skew-symmetric. With this FIR filter representation, the optimum encoder can be easily implemented.

So far, we have shown that the mappers for a binary input alphabet and up to five delay cells can be represented as an FIR filter. In Sec. 4.7 it was verified that for a quaternary input alphabet, the optimum encoder with up to five delay cells can be represented with two binary encoders, one for the inphase and the other for the
quadrature channel. With the same reasoning, it can be shown that the optimum FIR filter for a quaternary input alphabet has the same coefficients as shown in Table 5.2 when the four possible input symbols are 1, \(j\), \(-1\) and \(-j\), where \(j = \sqrt{-1}\). The multipliers of the filter shown in Fig. 5.1 then have to perform the multiplication of the complex symbol \(x(n - \nu T)\) with the real coefficient \(g_\nu\).

### 5.2 Optimization of the FIR Filter

In the last section, we showed that the optimum mapper can be represented as an FIR filter. Here, we optimize the coefficients of the FIR filter directly. This method will lead to the same results as described above, but it will be less complex: The size of the matrix involved in the optimization of the mapper was \(M^{L+1} \times M^{L+1}\), whereas the size of the FIR-matrix is about \((L + 1)/2 \times (L + 1)/2\).

In the past, several methods were published to reduce the bandwidth of a signal by filtering it with an FIR filter, see eg. [Mue73], [Che82],[Sun83]. These papers treat filters with several taps per symbol, and the length of the impulse response is one or more symbols. Here we consider a transmitter that consists of the cascade of a \(T\)-spaced FIR filter with \(L + 1\) taps (where \(T\) denotes the symbol duration) and a pulse shaping filter \(Q(f)\), as shown in Fig. 5.3.

We assume that the input symbols \(x(n)\) are independent and identically distributed with values \(-1\) and \(+1\) (binary input alphabet) or \(-1\), \(-j\), 1 and \(j\) (quaternary input alphabet) where \(j = \sqrt{-1}\). These symbols enter the transmitter at a rate \(1/T\) and are filtered by an FIR-filter of length \(L\). This filter has \(L + 1\) complex coefficients \(g_0 \ldots g_L\). The output of the FIR filter, a \(T\)-spaced series of complex Dirac pulses enters the pulse shaping filter \(Q(f)\), whose output signal \(s(t)\) is our desired baseband signal.

This system model is similar to the partial response signal (PRS) model described
5.2. Optimization of the FIR Filter

Figure 5.3: Schematic diagram of the partial response transmitter

in [Kab75]. We assume that the pulse shaping filter $Q(f)$ is a Nyquist filter, such that the intersymbol interference is introduced by the FIR filter only. In contrast to most of the popular PRS schemes, we do not restrict ourselves to integer FIR filter coefficients.

Because the signal $s(t)$ is generated by the cascade of an FIR and a pulse shaping filter, its power spectral density can be written as

$$W(f) = \frac{|Q(f)|^2}{T} \cdot |G(f)|^2,$$

where $G(f)$ denotes the transfer function of the FIR filter. In general, $G(f)$ equals

$$G(f) = \sum_{\nu=0}^{L} g_\nu \exp(-j2\pi\nu fT) = \underline{p}(f) \cdot \underline{g},$$

where the dagger denotes the conjugate transpose, and where $\underline{g}$ is the vector with the filter coefficients

$$\underline{g} = [g_0, \ldots, g_L]^t,$$

and $\underline{p}(f)$ is the vector

$$\underline{p}(f) = [1, \exp(j2\pi fT), \ldots, \exp(j2\pi LfT)]^t.$$

The absolute square of the transfer function then equals

$$|G(f)|^2 = \underline{p}(f)\underline{p}(f)^t \underline{g},$$
5. Representation of the Trellis Encoder as an FIR Filter

Note that the transfer function is periodic with period $1/T$. The inband power $P(B)$, i.e. the power that lies within the bandwidth $B$ can be expressed using Eqs. 5.5 and 5.9 as

$$P(B) = \int_{-B}^{B} W(f)df = g^\dagger \left( \int_{-B}^{B} \frac{|Q(f)|^2}{T} P(f) P^\dagger(f) df \right) g = \frac{g^\dagger H(B) g}{T}, \quad (5.10)$$

where we define the $(L+1) \times (L+1)$ matrix $H(B) = ||h_{\nu\mu}(B)||$ with the components

$$h_{\nu\mu}(B) = \int_{-B}^{B} |Q(f)|^2 \exp(-j2\pi(\nu - \mu) fT) df, \quad \nu, \mu = 0 \ldots L. \quad (5.11)$$

The total power of the transmitted signal $P(\infty)$ is computed by letting $B \to \infty$:

$$P(\infty) = \int_{-\infty}^{\infty} W(f)df = \frac{g^\dagger H(\infty) g}{T}, \quad (5.12)$$

where $H(\infty)$ is the matrix $||h_{\nu\mu}(\infty)||$ with components

$$h_{\nu\mu}(\infty) = \int_{-\infty}^{\infty} |Q(f)|^2 \exp(-j2\pi(\nu - \mu) fT) df, \quad \nu, \mu = 0 \ldots L. \quad (5.13)$$

Note that for real pulses, $h_{\nu\mu}(B) = \alpha_{\nu-\mu}$, and $h_{\nu\mu}(\infty) = \beta_{\nu-\mu}$, where $\alpha_\nu$ and $\beta_\nu$ are defined in Eqs. 2.46 and 2.47, respectively. In Appendix D it is shown that for Nyquist pulses

$$h_{\nu\mu}(\infty) = \int_{-\infty}^{\infty} |Q(f)|^2 \exp(-j2\pi(\nu - \mu) fT) df = \delta_{\nu\mu} \cdot \int_{-\infty}^{\infty} |Q(f)|^2 df. \quad (5.14)$$

Normalizing the pulse shaping filter such that $\int_{-\infty}^{\infty} |Q(f)|^2 df = 1$, the elements of the matrix $H(\infty)$ can be written as $h_{\nu\mu}(\infty) = \delta_{\nu\mu}$, such that $H(\infty)$ equals the identity matrix. Hence we have shown that for a normalized Nyquist pulse shaping filter, the total power $P(\infty)$ reduces to

$$P(\infty) = \frac{g^\dagger g}{T}. \quad (5.15)$$

Using the bandwidth-efficiency criterion of Eq. 3.1 yields

$$\eta = \frac{\int_{-B}^{B} W(f)df}{\int_{-\infty}^{\infty} W(f)df} = \frac{P(B)}{P(\infty)} = \frac{g^\dagger H(B) g}{g^\dagger g}. \quad (5.16)$$

The last expression of Eq. 5.16 has the form of a Rayleigh coefficient, i.e. it is maximized if $\mathbf{g}$ is the eigenvector of $H(B)$ corresponding to the maximum eigenvalue, see [Str88].

Let us now have a closer look at the structure of the matrix $H(B)$, cf. Eq. 5.11. Because the elements $h_{\nu\mu}(B)$ depend only on the difference $\nu - \mu$, $H(B)$ is a Toeplitz matrix. Eq. 5.11 shows that $H$ is symmetric. Because $Q(f)$ is even for all the filters we consider, $H$ is real. Hence $H(B)$ is a real symmetric Toeplitz matrix. This fact can be used to compute the eigenvectors of $H(B)$. The eigenvectors of this matrix are real, and Makhoul [Mak81] showed that they are either symmetric or skew-symmetric if the eigenvalues are distinct. Combining all these results, we have shown that the inband power of the transmitter of Fig. 5.3 is maximized by an FIR filter with real symmetric
or skew-symmetric coefficients. This fact allows us to reformulate the problem by considering only this class of filters, such that the complexity of the original problem can be drastically reduced.

We now have to distinguish between filters with an even and an odd number of symmetric or skew-symmetric taps, such that we finally consider four different filter structures. Note that all these filters have a constant group delay (see [Rab75, pp. 81-83]). For all these filters, we can express the inband power as

\[ \eta = \frac{a^{(n)} H^{(n)}(B) a^{(n)}}{a^{(n)} H^{(n)}(\infty) a^{(n)}} \]  

(5.17)

where the vectors \( a^{(n)} \) represent the modified FIR filter coefficients, and where the matrices \( H^{(n)}(B) = ||h^{(n)}(B)|| \) and \( H^{(n)}(\infty) = ||h^{(n)}(\infty)|| \) show the influence of the spectral shaping. The superscript \( (n) \) is used to distinguish the four different filter structures. The modified FIR filter coefficients \( a^{(n)} \) were obtained from the original coefficients \( a \) by using the symmetry or the skew-symmetry property of the different filters. This transformation will be explained in detail in the next four subsections where the coefficients of the four filter structures will be optimized.

Note that it is not obvious whether the symmetric or the skew-symmetric filter leads to a smaller out-of-band power. The decision which of the two filter structures is optimum can only be made for a given pulse shaping filter \( Q(f) \) and a normalized optimization bandwidth \( BT \), as will be shown in Chapter 6.

5.2.1 Filters with a Symmetrical Impulse Response and an Odd Number of Taps

In this subsection, we compute the optimum FIR filter with an odd number \((L + 1)\) of symmetric taps. This filter has an even number of delay cells \( L \), and the condition for the symmetric taps can be written as \( g_\nu = g_{L+1-\nu}, \quad \nu = 0 \ldots L + 1 \). According to [Rab75, Eq. 3.25], the transfer function of this filter can be written as

\[ G(e^{j\omega}) = e^{-j\omega L/2} \sum_{\nu=0}^{L/2} a_\nu^{(1)} \cos(\nu \omega), \quad \omega = 2\pi f T, \]  

(5.18)

where we have \( L/2 + 1 \) coefficients \( a_\nu^{(1)} \) that are expressed in terms of the \( L + 1 \) filter taps \( g_\nu \) as follows:

\[ a_0^{(1)} = g_{L/2} \]  

(5.19)

\[ a_\nu^{(1)} = 2g_{L/2-\nu}, \quad \nu = 1 \ldots L/2. \]

The term \( e^{-j\omega L/2} \) in Eq. 5.18 shows that the delay of this filter is \( L/2 \) samples, which is an integer number. We write \( G(f) \) as

\[ G(f) = e^{-j\pi L f T} \sum_{\nu=0}^{L/2} a_\nu^{(1)} \cos(2\pi \nu f T). \]  

(5.20)
Computing the absolute square of the transfer function of Eq. 5.20, yields

\[
|G(f)|^2 = \left( \sum_{\nu=0}^{L/2} a^{(1)}_{\nu} \cos(2\pi\nu f T) \right)^2 = \sum_{i=0}^{L/2} \sum_{j=0}^{L/2} a^{(1)}_{i} a^{(1)}_{j} \cos(2\pi i f T) \cos(2\pi j f T)
\]

\[
= \frac{1}{2} \sum_{\nu=0}^{L/2} \sum_{j=0}^{L/2} a^{(1)}_{i} a^{(1)}_{j} \left( \cos 2\pi (i + j) f T + \cos 2\pi (i - j) f T \right)
\]

\[
= a^{(1)} G^{(1)}(f) a^{(1)}^T,
\]

where the vector \( a^{(1)} \) is defined as

\[
a^{(1)} = [a^{(1)}_0, a^{(1)}_1, \ldots, a^{(1)}_{L/2}]^T,
\]

and where the matrix \( G^{(1)}(f) \) is defined as

\[
G^{(1)}(f) = ||a^{(1)}||
\]

with the coefficients

\[
g^{(1)}_{ij} = \frac{1}{2} \left[ \cos 2\pi (i + j) f T + \cos 2\pi (i - j) f T \right], \quad i, j = 0 \ldots L/2.
\]

Since the transmitter shown in Fig. 5.3 consists of a cascade of the FIR filter \( G(f) \) and the pulse shaping filter \( Q(f) \), the power spectral density of the output signal equals the product of the squared transfer function of two filters. The inband power \( P(B) \), i.e. the signal power that lies within the bandwidth \( B \), can be computed using Eqs. 5.21 and 5.5 as

\[
P(B) = \int_{-B}^{B} W(f) df = \frac{a^{(1)} H^{(1)}(B) a^{(1)}}{T},
\]

where the matrix \( H^{(1)}(B) \) equals \( ||h^{(1)}|| \) with the components

\[
h^{(1)}_{ij} = \frac{1}{2} (\alpha_{i+j} + \alpha_{|i-j|}), \quad i, j = 0 \ldots L/2,
\]

and where we defined \( \alpha_{\nu} \) as in Eq. 2.46:

\[
\alpha_{\nu} = \int_{-B}^{B} |Q(f)|^2 \cos 2\pi \nu f T df.
\]

Letting \( B \to \infty \), the total power \( P(\infty) \) of the transmitted signal can be computed as

\[
P(\infty) = \int_{-\infty}^{\infty} W(f) df = \frac{a^{(1)} H^{(1)}(\infty) a^{(1)}}{T},
\]

where the matrix \( H^{(1)}(\infty) \) is defined as

\[
H^{(1)}(\infty) = \text{diag}[1, \frac{1}{2}, \frac{1}{2}, \ldots]
\]
and where we assumed that the normalized pulse shaping filter is either Nyquist or
time-limited (see Eq. 5.14), so that
\[
\int_{-\infty}^{\infty} |Q(f)|^2 \cos 2\pi \nu f T = \delta_{\nu_0}.
\] (5.30)

Our goal is to find the coefficients of the FIR filter that maximize the ratio of the inband
power to the total power. Combining Eqs. 5.25 and 5.28 with the bandwidth-efficiency
criterion of Eq. 3.1 yields
\[
\eta = \frac{P(B)}{P(\infty)} = \frac{\varrho(\nu)^{\nu_0}}{\varrho(\nu)^{\nu_1}}.
\] (5.31)

The maximization of \( \eta \) leads to a generalized eigenvalue problem. Since \( H^{(1)}(\infty) \) is a
diagonal matrix, we can represent \( \eta \) as the solution of an ordinary eigenvalue problem,
see [Str88, p. 344]. We take the "square root" of the matrix \( H^{(1)}(\infty) \):
\[
H^{(1)}(\infty) = D' D
\] (5.32)
with \( D \) and \( D^{-1} \) equal to
\[
D = \text{diag}[1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \ldots] \quad D^{-1} = \text{diag}[1, \sqrt{2}, \sqrt{2}, \ldots].
\] (5.33)

Making the substitutions
\[
\tilde{a}^{(1)} = D a^{(1)} \quad \text{and} \quad a^{(1)} = D^{-1} \tilde{a}^{(1)},
\] (5.34)
we can write Eq. 5.31 as
\[
\eta = \frac{\varrho^{(1)^T} H^{(1)}(B) \varrho^{(1)}}{\varrho^{(1)^T} D' D \varrho^{(1)}} = \frac{\tilde{a}^{(1)^T} (D^{-1})' H^{(1)}(B) D^{-1} \tilde{a}^{(1)}}{\tilde{a}^{(1)^T} (D^{-1})' D' \tilde{a}^{(1)}} = \frac{\tilde{a}^{(1)^T} (D^{-1})' H^{(1)}(B) D^{-1} \tilde{a}^{(1)}}{\tilde{a}^{(1)^T} \tilde{a}^{(1)}}
\] (5.35)
which has the form of a Rayleigh coefficient. It is therefore maximized by the eigenvector
of the matrix \( (D^{-1})' H^{(1)}(B) D^{-1} \) corresponding to the maximum eigenvalue. The filter
coefficients \( g_v \) are then found with the help of Eqs. 5.34 and 5.19.

To illustrate this method, we compute the optimum coefficients for the filter with
three taps (\( L = 2 \)). In this case, the matrices are
\[
H^{(1)}(B) = \begin{bmatrix}
\alpha_0 & \alpha_1 \\
\alpha_1 & (\alpha_2 + \alpha_0)/2
\end{bmatrix} \quad H^{(1)}(\infty) = \begin{bmatrix}
1 & 0 \\
0 & 1/2
\end{bmatrix},
\] (5.36)
\[
D = \begin{bmatrix}
1 & 0 \\
0 & 1/\sqrt{2}
\end{bmatrix} \quad D^{-1} = \begin{bmatrix}
1 & 0 \\
0 & \sqrt{2}
\end{bmatrix}
\] (5.37)
and
\[
(D^{-1})' H^{(1)}(B) D^{-1} = \begin{bmatrix}
\alpha_0 & \sqrt{2}\alpha_1 \\
\sqrt{2}\alpha_1 & \alpha_2 + \alpha_0
\end{bmatrix}.
\] (5.38)
The two eigenvalues were found to be
\[
\lambda = \alpha_0 + \frac{\alpha_2}{2} \pm \frac{1}{2} \sqrt{\alpha_2^2 + 8\alpha_0^2}.
\] (5.39)
5. Representation of the Trellis Encoder as an FIR Filter

These equal \( \eta_5 \) and \( \eta_6 \) found in Sec. 4.3. Since the square root in Eq. 5.39 is not negative, the eigenvalue with the plus sign is the larger one. The corresponding eigenvector \( \tilde{a}^{(1)} \) equals

\[
\tilde{a}^{(1)} = \left[ 4a_1, \sqrt{2} \left( \alpha_2 + \sqrt{\alpha_2^2 + 8\alpha_1^2} \right) \right]',
\]

and using Eq. 5.34 we can substitute \( \tilde{a}^{(1)} \) by \( a^{(1)} \):

\[
a^{(1)} = \left[ 4a_1, 2 \left( \alpha_2 + \sqrt{\alpha_2^2 + 8\alpha_1^2} \right) \right]'.
\]

Using Eq. 5.19 we find the optimum filter coefficients

\[
g_0 = g_2 = \alpha_2 + \sqrt{\alpha_2^2 + 8\alpha_1^2},
g_1 = 4\alpha_1.
\]

These coefficients are a scalar multiple of those shown in Table 5.2 for \( \eta_6 \) and therefore result in the same out-of-band power.

To quantify the improvement of the direct computation of the filter coefficients, we have to compare this method with the method used to compute the coefficients of Table 5.2. The optimization of the mapper shown in Sec. 4.3 involved the computation of the eigenvectors of an 8 x 8 matrix, and the coefficients of the FIR filter were then obtained by solving a linear system of three equations. The direct computation of the filter coefficients introduced in this section was much less complex; it could be performed by computing the eigenvectors of a 2 x 2 matrix.

With the same method we also computed the symmetrical coefficients of the filter with five taps \( (L = 4) \). The resulting eigenvalues are equal to \( \eta_13 \ldots \eta_15 \) of Eq. 4.32. Computing the eigenvector corresponding to the maximum eigenvalue \( (\eta_13) \), we find the same coefficients as shown in Table 5.2 for \( \eta_13 \).

5.2.2 Filters with a Symmetrical Impulse Response and an Even Number of Taps

In this subsection, we apply the same procedure as before to compute the optimum symmetrical filter with an even number of taps, i.e. with an odd number of delay cells \( L \). The symmetry condition for the filter coefficients is again given by \( g_\nu = g_{L+1-\nu}, \) \( \nu = 0 \ldots L + 1 \). Since we have an even number of taps, the center of symmetry of the impulse response lies midway between two samples. The transfer function of this filter equals (see [Rab75, Eq. 3.28])

\[
G(e^{j\omega}) = e^{-j\omega L/2} \left( \sum_{\nu=0}^{(L-1)/2} a^{(2)}_\nu \cos \left[ \omega \left( \nu + \frac{1}{2} \right) \right] \right),
\]

where the coefficients \( a^{(2)}_\nu \) are functions of the filter coefficients \( g_\nu \),

\[
a^{(2)}_\nu = 2g_{(L-1)/2-\nu}, \quad \nu = 0 \ldots (L - 1)/2.
\]
In this case, the delay of the filter is \( L/2 \) samples, as indicated by the term \( e^{-j\pi L/2} \) in the transfer function. This delay is not an integer number of samples, because \( L \) is odd. We write \( G(f) \) as

\[
G(f) = e^{-j\pi L f T} \sum_{\nu=0}^{(L-1)/2} a_{\nu}^{(2)} \cos(2\pi(\nu + 1/2)f T).
\] (5.45)

The absolute square of this function equals

\[
|G(f)|^2 = \sum_{i=0}^{(L-1)/2} \sum_{j=0}^{(L-1)/2} a_i^{(2)} a_j^{(2)} \cos 2\pi(i + 1/2)f T \cos 2\pi(j + 1/2)f T
\]

\[
= \frac{1}{2} \sum_{i=0}^{(L-1)/2} \sum_{j=0}^{(L-1)/2} a_i^{(2)} a_j^{(2)} (\cos 2\pi(i + j + 1)f T + \cos 2\pi(i - j)f T)
\]

\[
= g^{(2)'} G^{(2)}(f) a^{(2)},
\] (5.46)

where

\[
a^{(2)} = [a_0^{(2)}, a_1^{(2)}, \ldots, a_{(L-1)/2}^{(2)}]^t
\] (5.47)

and the matrix \( G^{(2)}(f) \) equals \( \|g^{(2)}_{ij}\| \) with the components

\[
g^{(2)}_{ij} = \frac{1}{2}(\cos 2\pi(i + j + 1)f T + \cos 2\pi(i - j)f T), \quad i, j = 0 \ldots (L - 1)/2.
\] (5.48)

In analogy to the previous subsection, the inband power \( P(B) \) can be written as

\[
\int_{-B}^{B} W(f) df = \frac{1}{2T} a^{(2)'} H^{(2)}(B) a^{(2)},
\] (5.49)

where the components of the matrix \( H^{(2)}(B) = \|h^{(2)}_{ij}\| \) equal

\[
h^{(2)}_{ij} = \alpha_{i+j+1} + \alpha_{i-j}, \quad i, j = 0 \ldots (L - 1)/2
\] (5.50)

and where the coefficients \( \alpha_{\nu} \) are defined in Eq. 5.27. The total power \( P(\infty) \) of the signal equals

\[
P(\infty) = \int_{-\infty}^{\infty} W(f) df = \frac{a^{(2)'} H^{(2)}(\infty) a^{(2)}}{T},
\] (5.51)

where

\[
H^{(2)}(\infty) = \frac{1}{2} \cdot I
\] (5.52)

and where \( I \) denotes the identity matrix. We again assume that the pulse shaping filter \( Q(f) \) is chosen such that Eq. 5.30 holds. The bandwidth-efficiency criterion of Eq. 3.1 can be written together with Eqs. 5.49, 5.51 and 5.52 as

\[
\eta = \frac{a^{(2)'} H^{(2)}(B) a^{(2)}}{a^{(2)'} a^{(2)}}.
\] (5.53)

This equation has the form of a Rayleigh coefficient, i.e. it is minimized by the eigenvector of the matrix \( H^{(2)}(B) \) corresponding to the maximum eigenvalue. The filter coefficients can then be obtained by using Eq. 5.44.
This procedure shall be explained by the example of the filter with a single delay cell, \( L = 1 \). In this case, the matrix \( H(2)(B) \) reduces to a scalar

\[
H(2)(B) = \alpha_0 + \alpha_1
\]

(5.54)

whose 'eigenvalue' equals \( \alpha_0 + \alpha_1 \), and whose 'eigenvector' equals any real number. Using Eq. 5.44 and normalizing the filter coefficients, we find that

\[
g_0 = g_1 = \frac{1}{\sqrt{2}}.
\]

(5.55)

This result is the same as shown in Table 5.2 for \( \mu_2 \), but it was obtained much more easily than with the optimization of the mapper. As another example, we compute the optimum symmetric coefficients for the FIR filter with 4 taps, \( L = 3 \). According to Eq. 5.50, the matrix \( H(2)(B) \) equals

\[
H(2)(B) = \begin{bmatrix}
\alpha_1 + \alpha_0 & \alpha_2 + \alpha_1 \\
\alpha_2 + \alpha_1 & \alpha_3 + \alpha_0
\end{bmatrix}.
\]

(5.56)

The two eigenvalues are

\[
\lambda = \alpha_0 + \frac{\alpha_1}{2} + \frac{\alpha_3}{2} \pm \frac{1}{2} \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2}.
\]

(5.57)

They are equal to \( \eta_9 \) and \( \eta_{10} \) found in Sec. 4.4, where the optimum encoder with three delay cells was computed. The eigenvector corresponding to the eigenvalue with the plus sign equals

\[
g^{(2)} = \left[ 4(\alpha_1 + \alpha_2), 2 \left( -\alpha_1 + \alpha_3 + \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2} \right) \right].
\]

(5.58)

Using Eq. 5.44, we find the filter coefficients

\[
g_0 = g_3 = -\alpha_1 + \alpha_3 + \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 + 8\alpha_1\alpha_2 + 5\alpha_1^2}
g_1 = g_2 = 2(\alpha_1 + \alpha_2)
\]

(5.59)

which is the same result as shown in Table 5.2 for \( \mu_{10} \), which can be verified using Eq. 4.31. Here we could again show that the direct optimization of the filter coefficients and the computation of the coefficients after the optimization of the mapper lead to the same results. The direct optimization of the FIR filter coefficients, however, was accomplished by computing the eigenvector of a \( 2 \times 2 \) matrix, whereas the optimization of the mapper was much more complicated.

For a filter with 6 taps \( (L = 5) \), we find the same eigenvalues \( \eta_{16} \ldots \eta_{18} \) that were computed for the optimum encoder with five delay cells, see Eq. 4.39. The resulting filter coefficients corresponding to the maximum eigenvalue are those shown in Table 5.2 for \( \mu_{18} \).
5.2. Optimization of the FIR Filter

5.2.3 Filters with a Skew-Symmetrical Impulse Response and an Odd Number of Taps

In this subsection we optimize the FIR filter with an odd number of taps, i.e. with an even number of delay cells \( L \). We further assume that the filter coefficients are skew-symmetric, i.e. \( g_{\nu} = -g_{L+1-\nu} \). Since there is an odd number of taps, this condition implies that \( g_{L/2} = 0 \). The transfer function of such a filter with real coefficients can be written as (see [Rab75, Eq. 3.31])

\[
G(e^{j\omega}) = e^{-j\omega L/2} e^{j\pi/2} \left( \sum_{\nu=0}^{L/2-1} a_{\nu}^{(3)} \sin(\omega(\nu + 1)) \right),
\]

where the relation between \( a_{\nu}^{(3)} \) and the filter coefficients \( g_{\nu} \) is given by

\[
a_{\nu}^{(3)} = 2g_{L/2-1-\nu} \quad \nu = 0 \ldots L/2 - 1.
\]

This filter again has a delay of \( L/2 \) samples, i.e. it has a constant group delay, but it introduces an additional phase shift of \( \pi/2 \), as indicated by the term \( e^{-j\omega L/2} e^{j\pi/2} \) of the transfer function. We write \( G(f) \) as

\[
G(f) = e^{-j\pi f T} e^{-j\pi/2} \sum_{\nu=0}^{L/2-1} a_{\nu}^{(3)} \sin(2\pi(\nu + 1)f T).
\]

The power spectral density of the output signal of the FIR filter then equals

\[
|G(f)|^2 = \sum_{i=0}^{L/2-1} \sum_{j=0}^{L/2-1} a_{i}^{(3)} a_{j}^{(3)} \sin 2\pi (i + 1)f T \sin 2\pi (j + 1)f T
\]

\[
= \frac{1}{2} \sum_{i=0}^{L/2-1} \sum_{j=0}^{L/2-1} a_{i}^{(3)} a_{j}^{(3)} \left( \cos(2\pi(i - j)f T) - \cos 2\pi(i + j + 2)f T \right) \quad (5.63)
\]

\[
= a^{(3)'G^{(3)}(f)a^{(3)}},
\]

where the vector \( a^{(3)} \) is defined as

\[
a^{(3)} = [a_{0}^{(3)}, a_{1}^{(3)}, \ldots, a_{L/2-1}^{(3)}]',
\]

and where the matrix \( G^{(3)}(f) = \|g_{ij}^{(3)}\| \) equals

\[
g_{ij}^{(3)} = \frac{1}{2} \left( \cos 2\pi(i - j)f T - \cos 2\pi(i + j + 2)f T \right), \quad i, j = 0 \ldots L/2 - 1. \quad (5.66)
\]

The inband power \( P(B) \) is now given by

\[
\int_{-B}^{B} W(f) df = \frac{1}{2T} a^{(3)'H^{(3)}(B)a^{(3)}}, \quad (5.67)
\]

where the matrix \( H^{(3)}(B) = \|h_{ij}^{(3)}\| \) is defined as

\[
h_{ij}^{(3)} = a_{i-j} - a_{i+j+2}, \quad i, j = 0 \ldots L/2 - 1. \quad (5.68)
\]
Since we only consider pulse shaping filters that satisfy Eq. 5.30, the total power \( P(\infty) \) of the signal can be written as

\[
P(\infty) = \int_{-\infty}^{\infty} W(f) df = \frac{\mathbf{a}^{(3)'H^{(3)}(\infty)a^{(3)}}}{T},
\]

(5.69)

where the matrix \( H^{(3)}(\infty) \) equals

\[
H^{(3)}(\infty) = \frac{1}{2} \mathbf{I},
\]

(5.70)

and where \( \mathbf{I} \) denotes the identity matrix. The ratio of the inband power to the total power can be computed combining Eqs. 5.67, 5.69 and 5.70:

\[
\eta = \frac{\mathbf{a}^{(3)'H^{(3)}(B)a^{(3)}}}{\mathbf{a}^{(3)'a^{(3)}}}.
\]

(5.71)

Here we again find that this equation has the form of a Rayleigh coefficient, i.e. it is maximized by the eigenvector of the matrix \( H^{(3)}(B) \) corresponding to the maximum eigenvalue. The filter coefficients are those obtained by Eq. 5.61, remembering that \( g_{L/2} = 0 \).

With this method we compute the optimum FIR filter with skew-symmetric coefficients and three taps \( (L = 2) \). According to Eq. 5.68, the matrix \( H^{(3)}(B) \) equals

\[
H^{(3)}(B) = a_0 - a_2,
\]

(5.72)

i.e. the matrix reduces to a scalar. The 'eigenvalue' equals this scalar, and all real numbers are 'eigenvectors'. Using Eq. 5.61 and normalizing the filter coefficients yields

\[
\begin{align*}
g_0 &= \frac{1}{\sqrt{2}} \\
g_1 &= 0 \\
g_2 &= -\frac{1}{\sqrt{2}}
\end{align*}
\]

(5.73)

This solution is known as modified duobinary modulation and it was found in Sec. 4.3 as the mapper \( \Psi_4 \).

We also computed the optimum FIR filter coefficients for the skew-symmetrical filter with 5 taps, \( L = 4 \). We found the same eigenvalues \( \eta_{11} \) and \( \eta_{12} \) that were computed in Sec. 4.5. The filter coefficients are

\[
\begin{align*}
g_0 &= -g_4 = -\alpha_2 + \alpha_4 - \sqrt{\alpha_4^2 - 2\alpha_2 \alpha_4 + 4\alpha_3^2 - 8\alpha_1 \alpha_3 + \alpha_2^2 + 4\alpha_3^2} \\
g_1 &= -g_3 = 2(\alpha_3 - \alpha_1) \\
g_2 &= 0.
\end{align*}
\]

(5.74)

As will be shown in Chapter 6, this filter is less bandwidth-efficient than the filter with five taps and symmetrical coefficients \( (\Psi_{13} \) in Table 5.2) if the pulse shaping filter \( Q(f) \) is a time-limited or a Nyquist filter.
5.2.4 Filters with a Skew-Symmetrical Impulse Response and an Even Number of Taps

Here we derive the optimum skew-symmetric coefficients of the FIR filter with an even number \((L+1)\) of taps. This filter has an odd number of delay cells, \(L\). The condition for the coefficients to be skew-symmetric is \(g_{\nu} = -g_{L+1-\nu}\). According to [Rab75, Eq 3.34], the transfer function of this filter equals

\[
G(e^{j\omega}) = e^{-j\omega L/2} e^{j\pi/2} \left( \sum_{\nu=0}^{(L-1)/2} a_{\nu}^{(4)} \sin(\omega(\nu + \frac{1}{2})) \right),
\]

where the relation between \(a_{\nu}^{(4)}\) and the filter coefficients \(g_{\nu}\) is

\[
a_{\nu}^{(4)} = 2g_{(L-1)/2-\nu}, \quad \nu = 0 \ldots (L-1)/2.
\]

This filter has a delay of \(L/2\) samples, which is not an integer number, because \(L\) is odd. We write \(G(f)\) as

\[
G(f) = e^{-j\pi f T} e^{j\pi/2} \left( \sum_{\nu=0}^{(L-1)/2} a_{\nu}^{(4)} \sin 2\pi(\nu + \frac{1}{2})f T \right).
\]

The power spectral density of the output signal of this filter for a white input signal is given by the absolute square of the transfer function

\[
|G(f)|^2 = \sum_{i=0}^{(L-1)/2} \sum_{j=0}^{(L-1)/2} a_{i}^{(4)} a_{j}^{(4)} \sin 2\pi(i + \frac{1}{2})f T \sin 2\pi(j + \frac{1}{2})f T
\]

\[
= \frac{1}{2} \sum_{i=0}^{(L-1)/2} \sum_{j=0}^{(L-1)/2} a_{i}^{(4)} a_{j}^{(4)} (\cos 2\pi(i - j)f T - \cos 2\pi(i + j + 1)f T)
\]

\[
= \mathbf{a}^{(4')} \mathbf{G}^{(4)}(f) \mathbf{a}^{(4)},
\]

where we define the vector \(\mathbf{a}^{(4)}\) as

\[
\mathbf{a}^{(4)} = [a_{0}^{(4)}, a_{1}^{(4)}, \ldots, a_{(L-1)/2}^{(4)}]
\]

and the matrix \(\mathbf{G}^{(4)}(f) = ||g_{ij}^{(4)}||\) as

\[
g_{ij}^{(4)} = \frac{1}{2} [\cos 2\pi(i - j)f T - \cos 2\pi(i + j + 1)f T], \quad \nu = 0 \ldots (L-1)/2.
\]

The inband power \(P(B)\) of the signal equals

\[
P(B) = \int_{-B}^{B} W(f)df = \frac{1}{2T} \mathbf{a}^{(4')} \mathbf{H}^{(4)}(B) \mathbf{a}^{(4)},
\]

where the matrix \(\mathbf{H}^{(4)}(B) = ||h_{ij}^{(4)}||\) is defined as

\[
h_{ij}^{(4)} = \alpha_{i-j} - \alpha_{i+j+1}.
\]
The total power $P(\infty)$ of the signal can be expressed as

$$P(\infty) = \int_{-\infty}^{\infty} W(f) df = \frac{\mathbf{a}^{(4)} \mathbf{H}^{(4)}(\omega) \mathbf{a}^{(4)}}{T}$$  \hspace{1cm} (5.83)$$

with

$$\mathbf{H}^{(4)}(\omega) = \frac{1}{2} \cdot I,$$  \hspace{1cm} (5.84)$$

where $I$ denotes the identity matrix, and where we again considered only time-limited and Nyquist pulse shaping filters, such that Eq. 5.30 holds. The bandwidth-efficiency criterion of Eq. 3.1 can be written together with Eqs. 5.81 and 5.83 as

$$\eta = \frac{\int_{-B}^{B} W(f) df}{\int_{-\infty}^{\infty} W(f) df} = \frac{\mathbf{a}^{(4)} \mathbf{H}^{(4)}(B) \mathbf{a}^{(4)}}{\mathbf{a}^{(4)} \mathbf{a}^{(4)}}.$$  \hspace{1cm} (5.85)$$

As expected, it has the form of a Rayleigh coefficient, which is maximized by the eigenvector of the matrix $\mathbf{H}^{(4)}(B)$ corresponding to the maximum eigenvector.

Computing the optimum skew-symmetrical filter with two taps results in the well-known dicode modulation scheme; the corresponding filter coefficients are shown in Table 5.2 for $u_3$.

For the FIR filter with 4 skew-symmetrical coefficients, we find the matrix $\mathbf{H}^{(4)}(B)$

$$\mathbf{H}^{(4)}(B) = \begin{bmatrix} \alpha_0 - \alpha_1 & \alpha_1 - \alpha_2 \\ \alpha_1 - \alpha_2 & \alpha_0 - \alpha_3 \end{bmatrix},$$  \hspace{1cm} (5.86)$$

whose eigenvalues are

$$\lambda = \alpha_0 - \frac{\alpha_1}{2} - \frac{\alpha_3}{2} \pm \frac{1}{2} \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2}. $$  \hspace{1cm} (5.87)$$

They are equal to $\eta_7$ and $\eta_8$ found in Sec. 4.4 when the mapper with three delay cells was optimized. The eigenvector corresponding to the larger eigenvalue equals

$$\mathbf{a}^{(4)} = \left[ 4 (\alpha_2 - \alpha_1), 2 \left( -\alpha_1 + \alpha_3 - \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2} \right) \right]' $$  \hspace{1cm} (5.88)$$

and we find the filter coefficients using Eq. 5.76

$$g_0 = -g_3 = -\alpha_1 + \alpha_3 - \sqrt{\alpha_3^2 - 2\alpha_1\alpha_3 + 4\alpha_2^2 - 8\alpha_1\alpha_2 + 5\alpha_1^2} $$  \hspace{1cm} (5.89)$$

$$g_1 = -g_2 = 2 (\alpha_2 - \alpha_1). $$

These coefficients are the same as those shown in Table 5.2 for $u_3$, as can be verified using Eq. 4.30. Here again the complexity of the computation of the filter coefficients was much smaller than the complexity of the optimization of the mapper.

For the FIR filter with 6 taps, we find the eigenvalues that are equal to $\eta_{19} \ldots \eta_{21}$. The filter coefficients were computed the usual way, and for the maximum eigenvalue we got the same result as shown in Table 5.2 for $u_{19}$. 
5.2.5 Optimization of the Linear Phase FIR Filters

In the previous four subsections, we have shown how to compute the coefficients of an FIR filter such that the inband power of the signal generated by the transmitter of Fig. 5.3 is maximized. For a given pulse shaping filter \( Q(f) \) and the desired optimization bandwidth \( B \), the filter coefficients for both the symmetric and the skew-symmetric filter have to be computed and the filter with the smaller resulting out-of-band power has to be selected. Note that all these filters produce a delay of \( L/2 \) symbols, where \( L + 1 \) is the number of filter taps. The skew-symmetric filters furthermore produce a phase shift of \( \pi/2 \).

The optimum filter coefficients were obtained by computing the eigenvectors of a matrix. The size of these eigenvectors is shown in Table 5.3 for symmetric and skew-symmetric filters with an even and an odd number of taps:

<table>
<thead>
<tr>
<th>Number of Taps</th>
<th>even</th>
<th>odd</th>
</tr>
</thead>
<tbody>
<tr>
<td>symmetric</td>
<td>((L + 1)/2)</td>
<td>(L/2 + 1)</td>
</tr>
<tr>
<td>skew-symmetric</td>
<td>((L + 1)/2)</td>
<td>(L/2)</td>
</tr>
</tbody>
</table>

Table 5.3: Size of the eigenvectors of the optimum FIR filter

For some of the optimum filters, the coefficients were independent of the pulse shaping filter; they correspond to the well-known modulation schemes duobinary, dicode and modified duobinary. The coefficients of the remaining filters depend on both the pulse shaping filter \( Q(f) \) and on the optimization bandwidth \( B \). Note that all 21 eigenvalues found for the optimization of the mapper in Chapter 4 correspond to a filter with constant group delay, because the filter coefficients are either symmetric or skew-symmetric. For filters with even more taps, numerical results can be obtained by using the powerful software routines for computing eigenvectors. In this case, the quantities \( \alpha \) defined in Eq. 5.30 have to be computed for a given pulse shaping filter and a desired optimization bandwidth. Then, the eigenvectors of the matrices \((D^{-1})'H^{(1)}(B)D^{-1}\) (Eqs. 5.26 and 5.33) and \(H^{(3)}(B)\) (Eq. 5.68) or \(H^{(2)}(B)\) (Eq. 5.50) and \(H^{(4)}(B)\) (Eq. 5.82) are computed, depending on the number of taps. These eigenvectors are related to the filter coefficients as defined by Eqs. 5.34 and 5.19, 5.44, 5.61 or 5.76.

Note that whether the symmetric or the skew-symmetric filter leads to a smaller out-of-band power depends on the pulse shaping filter \( Q(f) \) and on the optimization bandwidth \( B \). The results for \( T\)-time-limited and SRC pulse shaping filters will be presented in Chapter 6.

Another interesting aspect of the comparison of the encoder/mapper and the FIR-filter appears by taking a closer look at Eqs. 2.41, 5.5 and 5.6. The first equation shows the spectrum of the shift-register encoder, whereas the other two equations define the spectrum of the FIR-filter. These two spectra are essentially the same, which means that both transmitters may generate signals with identical power spectral densities. The mapper of the shift-register encoder, however, might be chosen such that the resulting system is non-linear, whereas the FIR filter is always linear. This means that the higher
order statistics of these two transmitters might differ, although their spectra (second-order statistics) are identical. The shift-register encoder which is optimum in the sense of bandwidth-efficiency, however, is linear, because it can be represented as a FIR-filter. This statement was proved for a binary and a quaternary input alphabet and up to five delay cells, but we suspect, that it also holds for encoders with even more delay cells.
Chapter 6

Optimization of the Pulse Shaping Filter

In Chapter 4 we derived the optimum mapper for the communication system of Fig. 2.2 for up to five delay cells \((L = 1 \ldots 5)\), and in Chapter 5 we presented a method that allows the numerical optimization of encoders with even more delay cells. It could be shown that the mapper of the single delay-cell encoder is independent of the pulse shaping filter, whereas the mapper for two and more delay cells depends on both the optimization bandwidth \(B\) and the pulse shaping filter \(Q(f)\). The goal of this chapter is to compute the pulse shaping filter that maximizes the inband power of the entire transmitter for a given optimization bandwidth. We shall first derive the optimum time-limited pulse shaping filter and then compare these results with a Nyquist pulse shaping filter.

6.1 Optimization of the Sampled Time-Limited Pulse Shaping Filter

In this section we derive the optimum sampled time-limited pulse shaping filter for encoders with up to five delay cells. For the single delay-cell encoder, the optimum pulse shape can be represented by an eigenvector, whereas for encoders with two and more delay-cells only purely numerical results could be found. We start with the derivation of the optimum time-limited filter for an uncoded transmitter.

6.1.1 Optimum Time-Limited Pulse Shaping Filter for the Uncoded Transmitter

In this subsection the optimum time-limited pulse shaping filter for the uncoded transmitter will be presented. This result will be compared with the resulting out-of-band power of the encoder/mapper/modulator with up to five delay cells, and it will help us to quantify the gain that can be achieved by the introduction of coding.

To find the filter that maximizes the inband power of an uncoded transmitter, we
rewrite the bandwidth-efficiency criterion of Eq. 3.3 for \( L = 0 \): We have to maximize

\[
\eta_{\text{max}} = \max_{\mathbf{u}} \frac{\mathbf{u}^\dagger \alpha_0 \mathbf{W}_0 \mathbf{u}}{\mathbf{u}^\dagger Q(t) \mathbf{u} \beta_0 \mathbf{W}_0 \mathbf{u}}.
\]  

(6.1)

Inserting the definition of \( \mathbf{W}_0 \) (Eq. 2.39) in the above equation yields

\[
\eta_{\text{max}} = \max_{\mathbf{u}} \frac{\alpha_0}{Q(t) \beta_0},
\]

(6.2)

i.e. the mapper \( \mathbf{u} \) no longer appears in the maximization. This fact is not surprising since the mapper of an uncoded transmitter only causes a one-to-one mapping in the complex plane and therefore does not affect the spectrum of the output signal. In the time-continuous case, the impulse response of the optimum pulse shaping filter is a \textit{prolate spheroidal wave function} (PSWF), as was already shown by Slepian and Pollak [Sle61]. We shall use this terminology in the sequel for all optimized time-limited pulses, although we restrict ourselves to \textit{sampled} pulses for practical reasons.

Using the definitions of \( \alpha_0 \) (Eq. 2.48) and \( \beta_0 \) (Eq. 2.53), we can write Eq. 6.2 for a sampled \( T \)-time-limited pulse shaping filter as

\[
\eta_{\text{max}} = \max_{\mathbf{q}} \frac{\mathbf{q}^\dagger \mathbf{R}_0^{(B)} \mathbf{q}}{\mathbf{q}^\dagger \mathbf{R}_0^{(\kappa/T)} \mathbf{q}} = \max_{\mathbf{q}} \frac{\mathbf{q}^\dagger \mathbf{R}_0^{(B)} \mathbf{q}}{\mathbf{q}^\dagger \frac{\kappa}{2\pi T} \cdot \mathbf{I} \cdot \mathbf{q}} = \frac{2\pi T}{\kappa} \max_{\mathbf{q}} \frac{\mathbf{q}^\dagger \mathbf{R}_0^{(B)} \mathbf{q}}{\mathbf{q}^\dagger \mathbf{q}},
\]

(6.3)

where the quantities appearing in this equation were defined in Sec. 2.3.

Figure 6.1: Optimized pulse shape for an uncoded transmitter, \( BT=0.9 \)

The right-hand part of Eq. 6.3 is a Rayleigh quotient, and it is therefore maximized by the eigenvector of the matrix \( \mathbf{R}_0^{(B)} \) corresponding to the maximum eigenvalue. The inband power is then proportional to this eigenvalue.
6.1. Optimization of the Sampled Time-Limited Pulse Shaping Filter

We computed the eigenvectors of the matrix $R_0^{(B)}$ numerically for 16 samples per symbol ($\kappa = 8$) for normalized optimization bandwidths $BT = 0.9$ and 1.5. The resulting pulse shapes are shown in Figs. 6.1 and 6.2. It can be seen that the optimum pulse is symmetric and is nonzero at the end of the interval. This does not contradict the well known fact that signal steps in the time domain widen the spectrum of that signal, because that statement describes an asymptotic behaviour of the spectrum, whereas we want to maximize the power within a finite normalized bandwidth $BT$. Comparing the pulse shapes for $BT = 0.9$ and 1.5 shows indeed that the steps at the beginning and at the end of the pulse get smaller for an increasing optimization bandwidth.

The out-of-band power that can be achieved with an uncoded transmitter is shown in Fig. 6.3. The two solid lines represent the simulated out-of-band power of the two transmitters optimized for $BT = 0.9$ and 1.5. The dashed line in Fig. 6.3 shows the envelope of all possible uncoded transmitters. It was drawn by computing the out-of-band power of a transmitter optimized for normalized bandwidths $BT$ in the range of 0 to 2.

6.1.2 Optimum Time-Limited Pulse Shaping Filter for the Single Delay-Cell Encoder

To compute the optimum sampled time-limited pulse shaping filter for the encoders described in Chapter 4, we express the functions $\alpha_\nu$ and $\beta_\nu$ in terms of the matrices $R_\nu^{(B)}$ and $R_\nu^{(s/T)}$ defined in Sec. 2.3:

$$\alpha_\nu = q'R_\nu^{(B)}q, \quad \beta_\nu = q'R_\nu^{(s/T)}q.$$  \tag{6.4}
Figure 6.3: Out-of-band power of an optimized, uncoded transmitter
The optimum pulse shaping filter \( q \) is then obtained by combining Eq. 6.4 with the out-of-band power \( \eta_i \) defined in Chapter 4 and by maximizing over \( q \):

\[
\eta_{i\text{max}} = \max_q \eta_i. \tag{6.5}
\]

For the encoder with a single delay cell, the maximum inband power is either \( \eta_2 = \alpha_0 + \alpha_1 \) (Eq. 4.8) or \( \eta_2 = \alpha_0 - \alpha_1 \) (Eq. 4.9), depending on the sign of \( \alpha_1 \). This statement holds for all possible pulse shaping filters that are normalized such that \( \beta_0 = 1 \). According to Eqs. 6.4 and 6.5, the optimum pulse shaping filter for the duobinary modulation (\( \eta_2 \)) is defined by

\[
\frac{\alpha_0 + \alpha_1}{\beta_0} = \max_q \left[ \frac{q' \left[ R_0^{(B)} + R_1^{(B)} \right]}{q' \left[ R_0^{(\kappa/T)} + R_1^{(\kappa/T)} \right]} \right] \tag{6.6}
\]

i.e. the optimum pulse shape is given by the eigenvector of the matrix \( R_0^{(B)} - R_1^{(B)} \) corresponding to its maximum eigenvalue. The optimum pulse shapes of these two modulation schemes look similar to those in Figs. 6.1 and 6.2. The achievable out-of-band power is shown in Fig. 6.4. The solid line shows the out-of-band power of an uncoded transmitter, the dotted line corresponds to the duobinary and the dashed line to the dicode modulation scheme. We see that depending on the normalized optimization bandwidth \( BT \) either duobinary or dicode leads to a smaller out-of-band power. The gain that can be achieved by the introduction of coding reaches from a few % up to 40 % and largely depends on the optimization bandwidth. The bandwidth-efficiency in \( \text{bit/sec/Hz} \) for a given out-of-band power and a quaternary input alphabet is computed by dividing the bit rate by the bandwidth, i.e. \( 1/T/B = 1/ BT \). If an uncoded QPSK system is optimized for an out-of-band power of -20dB, the bandwidth-efficiency is 0.89 \( \text{bit/sec/Hz} \). If a single delay-cell encoder is optimized for the same parameters, the bandwidth-efficiency increases by 25% to 1.11 \( \text{bit/sec/Hz} \).

### 6.1.3 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Two Delay Cells

In this subsection we compute the optimum time-limited pulse shaping filter for the encoder with two delay cells. According to Eq. 4.27, the maximum inband power is either \( \eta_4 \) or \( \eta_6 \). For \( \eta_4 \), the modified duobinary modulation scheme, we can apply the same procedure as in the previous subsection:

\[
\eta_{4\text{max}} = \max_q \left[ \frac{\alpha_0 - \alpha_2}{\beta_0} \right] = \max_q \left[ \frac{q' \left[ R_0^{(B)} - R_2^{(B)} \right]}{q' \left[ R_0^{(\kappa/T)} \right]} \right] = \frac{2\pi T}{\kappa} \max_q \left[ \frac{q' \left[ R_0^{(B)} - R_2^{(B)} \right]}{q' q} \right]. \tag{6.8}
\]
Figure 6.4: Out-of-band power of uncoded, duobinary, and dicode transmitters
The optimum pulse shaping filter $q$ can therefore be computed as the eigenvector of the matrix $R_0^{(B)} - R_2^{(B)}$ corresponding to the maximum eigenvector. For $\eta_6$, this method fails because $\eta_6$ is a nonlinear function of the $\alpha_i$, see Eq. 4.23. All we can do is to express $\eta_6$ as functions of $q$ and $BT$ and to find the maximum using a numerical optimization routine on a computer. The correctness of this method was verified by computing the previous four objective functions the same way, where the results of the two different methods agree. The objective function $\eta_6$ equals

$$
\eta_6 = \frac{q' \left[ R_0^{(B)} + \frac{1}{2} R_2^{(B)} \right] q + \frac{1}{2} \sqrt{\left( q' R_3^{(B)} q \right)^2 + 8 \left( q' R_1^{(B)} q \right)^2}}{q' R_0^{(\mu/T)} q}.
$$

(6.9)

Despite the different structure of Eq. 6.9, the optimum pulse shapes look similar to those of Figs. 6.1 and 6.2. Fig. 6.5 shows the achievable out-of-band power corresponding to $\eta_4$ (solid line) and $\eta_6$ (dotted line). It can be seen, that for normalized optimization bandwidths $BT \in [0 \ldots 2]$, $\eta_6$ is always optimum.

The optimum bandwidth-efficient modulation scheme with two delay cells is drawn in Fig. 4.9, where $a$ and $b$ are defined in Eq. 4.25 and are computed for the optimum $q$ using Eq. 6.4.

If we optimize a system with two delay-cells and a single time-limited pulse shaping filter for an out-of-band power of -20dB and quaternary input symbols, we can transmit 1.28 bit/sec/Hz. The increase of data rate equals 44% with respect to an uncoded and still 15% with respect to an optimum single delay-cell encoder.

### 6.1.4 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Three Delay Cells

To find the optimum time-limited pulse shaping filter for the encoder with three delay cells, we proceed as described in Sec. 6.1.2. The optimum sampled filter is given by the vector $q$ that maximizes

$$
\eta_8 = \frac{q' \left[ R_0^{(B)} - \frac{1}{2} R_1^{(B)} - \frac{1}{2} R_3^{(B)} \right] q}{q' R_0^{(\mu/T)} q} + \frac{\sqrt{\left( q' R_3^{(B)} q \right)^2 - 2 \left( q' R_1^{(B)} q \right) \left( q' R_3^{(B)} q \right) + 4 \left( q' R_2^{(B)} q \right)^2 - 8 \left( q' R_1^{(B)} q \right) \left( q' R_2^{(B)} q \right) + 5 \left( q' R_1^{(B)} q \right)^2}}{2q' R_0^{(\mu/T)} q}
$$

(6.10)

or

$$
\eta_{10} = \frac{q' \left[ R_0^{(B)} + \frac{1}{2} R_1^{(B)} + \frac{1}{2} R_3^{(B)} \right] q}{q' R_0^{(\mu/T)} q} + \frac{\sqrt{\left( q' R_3^{(B)} q \right)^2 - 2 \left( q' R_1^{(B)} q \right) \left( q' R_3^{(B)} q \right) + 4 \left( q' R_2^{(B)} q \right)^2 + 8 \left( q' R_1^{(B)} q \right) \left( q' R_2^{(B)} q \right) + 5 \left( q' R_1^{(B)} q \right)^2}}{2q' R_0^{(\mu/T)} q}
$$

(6.11)
Figure 6.5: Out-of-band power of an encoder with two delay-cells, time-limited pulses
The result of this numerical maximization is shown in Fig. 6.6, where the envelope of the out-of-band power of $\eta_8$ (solid line) and of $\eta_{10}$ (dotted line) are shown. It can be seen that, depending on the normalized optimization bandwidth $BT$, either $\eta_8$ or $\eta_{10}$ is optimum. The staircase-like structure of the two curves stems from the spectral nulls that are introduced by the coding.

![Achievable out-of-band power with 3 delay cells, time-limited pulses](image)

Figure 6.6: Achievable out-of-band power with 3 delay cells, time-limited pulses

### 6.1.5 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Four Delay Cells

The optimum time-limited pulse shaping filters for $\eta_{11}$ to $\eta_{15}$ given the optimization bandwidth $B$ are computed using Eqs. 6.4 and 6.5. The results of this numerical opti-
mization are shown in Fig. 6.7, where the *envelope* of the achievable out-of-band power is shown for $\eta_{_{11}}$ to $\eta_{_{18}}$.

It can be seen that for normalized optimization bandwidths $BT \in [0...2]$ the mapper $\mu_{_{13}}$ is always optimum.

### 6.1.6 Optimum Time-Limited Pulse Shaping Filter for the Encoder with Five Delay Cells

The optimum time-limited pulse shaping filter for the encoder with five delay cells ($\eta_{_{16}}$ to $\eta_{_{21}}$) is also obtained using Eqs. 6.4 and 6.5. The results of this numerical optimization are shown in Fig. 6.8, where the *envelope* of the achievable out-of-band power is shown for $\eta_{_{16}}$ to $\eta_{_{21}}$.

We can see that either $\eta_{_{18}}$ or $\eta_{_{19}}$ are optimum, depending on the bandwidth.

### 6.1.7 Summary of Optimum Encoders with a Time-Limited Pulse Shaping Filter

Here we summarize the results of the previous 6 subsections. Fig. 6.9 shows the achievable out-of-band powers for encoders with memory $L = 0...5$. The staircase-like structure of the curves for $L = 1...5$ stems from the spectral nulls that are introduced by the coding.

To quantify the gain of data rate that can be achieved with these new modulation schemes, we compared the data rates of systems with a quaternary input alphabet and an out-of-band power of -20 dB. The results of this comparison are shown in Table 6.1. It is remarkable that the data rate can be increased by as much as 70% when an encoder with five delay cells is used.

<table>
<thead>
<tr>
<th>Delay Cells</th>
<th>Bandwidth-Efficiency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bit/sec/Hz</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.89</td>
<td>100</td>
</tr>
<tr>
<td>1</td>
<td>1.11</td>
<td>125</td>
</tr>
<tr>
<td>2</td>
<td>1.28</td>
<td>144</td>
</tr>
<tr>
<td>3</td>
<td>1.39</td>
<td>156</td>
</tr>
<tr>
<td>4</td>
<td>1.47</td>
<td>165</td>
</tr>
<tr>
<td>5</td>
<td>1.51</td>
<td>170</td>
</tr>
</tbody>
</table>

Table 6.1: Bandwidth-Efficiencies at -20dB out-of-band power, PSWF, $M=4$
Figure 6.7: Achievable out-of-band power with 4 delay cells, time-limited pulses
Figure 6.8: Optimized transmitters with 5 delay-cells, PSWF
6.1. Optimization of the Sampled Time-Limited Pulse Shaping Filter

Figure 6.9: Achievable out-of-band power with a time-limited pulse shaping filter
6.2 Optimization of the Nyquist Pulse Shaping Filter

In this section we derive the Nyquist pulse shaping filter that minimizes the out-of-band power of the communication system of Fig. 2.2. For the single delay-cell encoder, it is shown that the ideal low-pass filter is optimum. Because this filter is not realizable, we consider the spectral raised cosine filter, for which good approximations exist. The same filter will be used to show the performance of the encoders with two to five delay cells.

We start by writing the power spectral density of the output signal of a Nyquist filter as

\[ |Q(f)|^2 = \begin{cases} 1 + F(f) & |f| < \frac{1}{2T} \\ F(f) & \frac{1}{2T} < |f| < \frac{1}{T} \\ 0 & |f| > \frac{1}{T} \end{cases} \tag{6.12} \]

where the function \( F(f) \) is even with respect to \( f = 0 \) and odd with respect to \( 1/2T \). Furthermore, \( F(f) \) has to be positive and less than or equal to one in the interval \([1/2T, 1/T]\), such that the power spectral density \( |Q(f)|^2 \) is positive for all frequencies \( f \):

\[
F(-f) = F(f) \\
F(-f + \frac{1}{2T}) = -F(f + \frac{1}{2T}) \tag{6.13}
\]

\[ 0 \leq F(f) \leq 1 \quad \frac{1}{2T} < |f| < \frac{1}{T}. \]

According to Eq. 2.41, the power spectral density of the output signal of the transmitter of Fig. 2.2 equals

\[ W(f) = \frac{|Q(f)|^2}{T} \cdot \sum_{\nu=0}^{L} a_\nu \cos(2\pi \nu f T) + b_\nu \sin(2\pi \nu f T), \tag{6.14} \]

where the coefficients

\[ a_\nu = u^\dagger \mathbf{W}_\nu u \tag{6.15} \]

and

\[ b_\nu = u^\dagger \mathbf{V}_\nu u \tag{6.16} \]

may be functions of the pulse shaping filter \( Q(f) \) and the optimization bandwidth \( B \), but independent of \( f \). Combining the bandwidth-efficiency criterion (Eq. 3.3) and Eq. 6.14 yields the inband power:

\[
\eta = \frac{\int_{-B}^{B} W(f) df}{\int_{-\infty}^{\infty} W(f) df} = \frac{\int_{-B}^{B} |Q(f)|^2 + \sum_{\nu=0}^{L} a_\nu \cos(2\pi \nu f T) df}{\int_{-\infty}^{\infty} |Q(f)|^2 df}, \tag{6.17} \]

where we used the fact that for Nyquist pulses (see Appendix D)

\[ \int_{-\infty}^{\infty} |Q(f)|^2 \cos(2\pi \nu f T) df = \delta_{\nu0} \int_{-\infty}^{\infty} |Q(f)|^2 df \tag{6.18} \]
and that for real pulses $|Q(f)|^2$ is symmetric, i.e.

$$\int_{-B}^{B} |Q(f)|^2 \sin(2\pi \nu f T) df = 0.$$  \hspace{1cm} (6.19)

For bandwidths $B$ less than $1/2T$, Eq. 6.17 can be written as

$$\eta = \frac{\int_{-B}^{B} (1 + F(f)) \sum_{\nu=0}^{L} a_{\nu} \cos(2\pi \nu f T) df}{2a_0} \quad B < \frac{1}{2T}$$

(6.20)

because $F(f)$ is even with respect to $f = 0$ and odd with respect to $f = 1/2T$. Dropping the factors that are independent of $F(f)$, we have to find the function $F(f)$ that maximizes

$$\max_{F(f)} \int_{0}^{B} F(f) \sum_{\nu=0}^{L} a_{\nu} \cos(2\pi \nu f T) df \quad B < \frac{1}{2T}$$

(6.21)

for all $F(f)$ that satisfy Eq. 6.13. For bandwidths $B$ greater than $1/2T$, Eq. 6.17 can be written as

$$\eta = \frac{2T}{a_0} \sum_{\nu=0}^{L} a_{\nu} \left[ \int_{0}^{\frac{B}{2T}} (1 + F(f)) \cos(2\pi \nu f T) df + \int_{\frac{B}{2T}}^{B} F(f) \cos(2\pi \nu f T) df \right] \quad B > \frac{1}{2T}.$$  

(6.22)

Dropping the constant factors yields Eq. 6.21, without restriction on $B$. Hence for a given mapper, i.e. for given coefficients $a_{\nu}$, the Nyquist pulse shaping filter that maximizes the inband power is given by Eq. 6.12, where $F(f)$ is chosen such that

$$\max_{F(f)} \sum_{\nu=0}^{L} a_{\nu} \int_{0}^{B} F(f) \cos(2\pi \nu f T) df$$

(6.23)

where $F(f)$ satisfies Eq. 6.13.

### 6.2.1 Optimization of the Nyquist Pulse Shaping Filter for the Single Delay-Cell Encoder

In this subsection we shall derive the optimum Nyquist pulse shaping filter for the single delay-cell encoder. As already stated in Sec. 6.1.2, the maximum inband power of this encoder is either $a_0 + a_1$ (Eq. 6.8) or $a_0 - a_1$ (Eq. 6.9), depending on the sign of $a_1$.

If $a_1 > 0$, the coefficients $a_{\nu}$ defined in Eq. 6.15 equal

$$a_0 = 1, \quad a_1 = 1.$$  \hspace{1cm} (6.24)

Combining this equation with Eq. 6.23, we have to find the function $Q(f)$ that satisfies Eq. 6.13 such that

$$\max_{F(f)} \int_{0}^{B} F(f)(1 + \cos(2\pi f T)) df.$$  \hspace{1cm} (6.25)

The factor $(1 + \cos(2\pi f T)) = 2 \cos^2(\pi f T)$ is not negative, but according to Eq.6.13, $F(f)$ is less than or equal to zero for $f < 1/2T$. If we also consider the second equation of Eq. 6.13, we see that the function $F(f) = 0$ maximizes Eq. 6.25. This means that the
ideal low-pass filter is the optimum pulse shaping filter for the single delay-cell encoder if $\alpha_1 > 0$. If $\alpha_1 < 0$, $a_0 = 1$ and $a_1 = -1$, such that $F(f)$ has to be chosen such that

$$
\max_{F(f)} \int_0^B F(f)(1 - \cos(2\pi f T))df.
$$

(6.26)

Because the factor $1 - \cos(2\pi f T) = 2 \sin^2(\pi f T)$ is also non-negative, the ideal low-pass filter also maximizes the inband power for $\alpha_1 < 0$. Since for an ideal low-pass filter, $\alpha_1 > 0$, we have shown that the out-of-band power of a single delay-cell encoder of Fig. 2.2 is maximized by a duobinary encoder followed by an ideal low-pass filter. The achievable out-of-band power of this system is shown in Fig. 6.11 (solid line).

For actual systems, the ideal low-pass filter is impractical for several reasons, such as its sensitivity to timing jitter. We therefore show the performance of a spectral raised cosine (SRC) filter with rolloff parameter $\gamma$, which is defined in Eq. 2.56. This filter is not realizable either, but good approximations exist. For the SRC filter, the out-of-band power of the resulting optimum transmitter can be computed analytically using the closed form expressions for $a_{\nu}$ defined in Eqs. 2.57 and 2.58.

To quantify the gain that can be achieved with the single delay-cell encoder, Fig. 6.10 shows the out-of-band power of an uncoded transmitter with a SRC pulse shaping filter with rolloff factors 0 (ideal low-pass filter), 0.3, 0.7 and 1.0.

The out-of-band power of the single delay-cell encoder is shown in Fig. 6.11 for the same rolloff factors. The sign of $\alpha_1$ had to be computed to decide whether the duobinary or the dicode mapper has to be chosen. If we want to optimize the system for the -20dB out-of-band power, for the first two cases (rolloff = 0.3 and 0.7), $\alpha_1 > 0$, so that the mapper of Fig. 4.3 (duobinary) is chosen, whereas for a rolloff factor of 1.0, $\alpha_1 < 0$, so that we use the encoder of Fig. 4.4 (dicode). It is remarkable that the SRC filter with rolloff = 0.3 has the same -20dB bandwidth as the ideal low-pass filter.

The bandwidth-efficiency of an uncoded QPSK transmitter with a spectral raised cosine filter and a rolloff parameter of 0.3 is 1.76 bit/sec/Hz if an out-of-band power of -20dB has to be achieved. With the introduction of an optimum single delay-cell encoder, the bandwidth-efficiency is raised by 36% to 2.43 bit/sec/Hz.

### 6.2.2 Optimization of the Nyquist Pulse Shaping Filter for the Encoder With Two and More Delay Cells

For the encoder with two delay cells, the maximum inband power is either $\eta_4$ or $\eta_6$ (cf. Eq. 4.27). For the duobinary modulation ($\eta_4$), the same procedure as in the last subsection also leads to the ideal low-pass filter as the optimum Nyquist pulse shaping filter. For $\eta_6$, the mapper of the optimum encoder depends on the pulse shaping filter $Q(f)$ and on the normalized optimization bandwidth $BT$. Since the coefficients $a_{\nu}$ defined in Eq. 6.15 represent the mapper, they also depend on $Q(f)$ and $BT$. Combining the mapper coefficients with Eq. 6.23 yields a function of $F(f)$ that is too complicated to be maximized analytically. We suspect, however, that the ideal low-pass filter is again the optimum filter, since we tested various filters and no counter-example has been found so far. The same holds for encoders with more than two delay cells. For practical reasons we shall show the out-of-band power of several transmitters with a SRC pulse shaping filter and a rolloff factor of 0.3 in the remaining part of this section.
Figure 6.10: Out-of-band power of an uncoded transmitter with a spectral raised cosine (SRC) pulse shaping filter
Figure 6.11: Out-of-band power of an optimized transmitter with one delay-cell
6.2. Optimization of the Nyquist Pulse Shaping Filter

We start with Fig. 6.12, where the out-of-band power for \( \eta_4 \) and \( \eta_6 \) is shown. It can be seen that for the chosen rolloff factor 0.3, \( \eta_6 \) is always better. The out-of-band power of the mappers \( \eta_6 \) and \( \eta_{10} \) of the encoder with three delay cells are depicted in Fig. 6.13. In this case, \( \eta_{10} \) proves to be optimum. The six mappers \( \eta_{11} \ldots \eta_{15} \) of the encoder with four delay cells are shown in Fig. 6.14. Here, the optimum mapper is \( \eta_{13} \). In Fig. 6.15, we present the mappers of the encoder with five delay-cells, where the optimum mapper is \( \eta_{16} \).

6.2.3 Summary of Optimum Encoders with a Spectral Raised Cosine Pulse Shaping Filter

In this section we summarize the results of the previous subsections. Fig. 6.16 shows the achievable out-of-band powers for the mappers \( \eta_1, \eta_2, \eta_6, \eta_{10}, \eta_{13} \) and \( \eta_{16} \), i.e. the best mappers for the encoders with memory \( L = 0 \ldots 5 \), for a SRC pulse shaping filter with a rolloff factor of 0.3.

To show the increase of the data rate that can be achieved with these optimum modulation schemes, Table 6.2 shows the bandwidth-efficiency of systems with a quaternary input alphabet and an out-of-band power of -20 dB. It is remarkable that the data can be transmitted three times as fast when an encoder with five delay cells is used.

<table>
<thead>
<tr>
<th>Delay Cells</th>
<th>Bandwidth-Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bit/sec/Hz</td>
</tr>
<tr>
<td>0</td>
<td>1.76</td>
</tr>
<tr>
<td>1</td>
<td>2.43</td>
</tr>
<tr>
<td>2</td>
<td>3.14</td>
</tr>
<tr>
<td>3</td>
<td>3.85</td>
</tr>
<tr>
<td>4</td>
<td>4.65</td>
</tr>
<tr>
<td>5</td>
<td>5.40</td>
</tr>
</tbody>
</table>

Table 6.2: Bandwidth-Efficiencies at -20dB out-of-band power, SRC, \( M=4 \)
Figure 6.12: Optimized transmitters with 2 delay-cells, SRC, rolloff = 0.3
Figure 6.13: Optimized transmitters with 3 delay-cells, SRC, rolloff = 0.3
Figure 6.14: Optimized transmitters with 4 delay-cells, SRC, rolloff = 0.3
6.2. Optimization of the Nyquist Pulse Shaping Filter

Figure 6.15: Optimized transmitters with 5 delay-cells, SRC, rolloff = 0.3
6. Optimization of the Pulse Shaping Filter

Figure 6.16: Out-of-band power with a SRC pulse shaping filter, rolloff = 0.3
Chapter 7

Power-Efficiency

In the previous chapters, the communication system of Fig. 2.2 was optimized with respect to bandwidth-efficiency. Here we investigate the power-efficiency, i.e. we want to determine the performance of the system in an additive white Gaussian noise (AWGN) channel. This will be accomplished by computing the minimum Euclidean distance of the generated signals and verifying the results by computer simulation.

As we have seen in Sec. 4.2, the bandwidth-optimization of the single delay-cell encoder still left some degrees of freedom for the mapper that can be used to perform a secondary optimization with respect to power-efficiency. For the encoders with two and more cells, the bandwidth-efficient mappers are uniquely determined, such that no further optimization can be performed. All we can do is to compute the power-efficiency of these mappers, and to compare them with an uncoded transmitter.

We first consider the single delay-cell encoder and optimize the mapper with respect to average and peak-power-efficiency.

7.1 Average-Power-Efficient Mapper for the Single Delay-Cell Encoder

The results of Sec. 4.2 showed that the optimum single delay-cell encoder consists of a mapper, an FIR filter with two taps and a pulse shaping filter, see Fig. 4.5. For the mapper, there are still $2(M-1)$ degrees of freedom left, where $M$ is the input alphabet size. Depending on the sign of $\alpha_1$, there are two different mappers that have to be considered. In this section, we first rewrite the average power-efficiency criterion defined in Sec. 3.2 and apply it then to the two different mappers.

The average bit energy of the signal transmitted by the single delay-cell encoder is (see Eq. 3.10)

$$E_b = \frac{\sum_{i,j} |h_{ij}|^2}{M^{L+1} \cdot \log_2(M)} \cdot \int_{-\infty}^{\infty} q^2(t)dt.$$  \hspace{1cm} (7.1)

According to Eq. 3.12, the Euclidean distance of two output signals is proportional to the square of the difference of the two corresponding mapper output sequences. Two paths with minimum mutual distance stemming from the same node of the trellis can be generated by feeding the single delay-cell encoder with two input sequences that differ only in one symbol. In this case, the two corresponding mapper output sequences
differ in two symbols, because the encoder has a memory of one symbol. For a single delay-cell encoder, Eq. 3.12 can therefore be written as

$$d_{\text{min}}^2 = \min_{i,j,k,l, j \neq k} \left( |u_{ij} - u_{ik}|^2 + |u_{jl} - u_{kl}|^2 \right) \cdot \int_{-\infty}^{\infty} q^2(t) \, dt,$$

(7.2)

where \( u_{ij} \) is an element of the optimum mapper \( u \)

$$u = [u_{00}, u_{01}, \ldots, u_{0(M-1)}, u_{10}, u_{11}, \ldots, u_{(N-1)(M-1)}].$$

(7.3)

We assumed that the two input sequences were \( \{\ldots, i, j, \ldots\} \) and \( \{\ldots, i, k, \ldots\} \), and that the pulse \( q(t) \) satisfies the generalized Nyquist criterion. The average-power-efficiency criterion for single delay-cell encoders can now be obtained by combining Eqs. 3.9, 7.1 and 7.2:

$$\max \frac{d_{\text{min}}^2}{2E_b} = \max \frac{M^L + 1 \cdot \log_2(M) \cdot \left( |u_{ij} - u_{ik}|^2 + |u_{jl} - u_{kl}|^2 \right)}{2 \sum_{i,j} |u_{ij}|^2}. \quad (7.4)$$

In the next two subsections this criterion will be applied to the two different mappers corresponding to \( \eta_2 \) and \( \eta_3 \), which are optimum for pulse shapes \( \alpha_1 > 0 \) and \( \alpha_1 < 0 \), respectively.

### 7.1.1 Average-Power-Efficient Mapper for \( \eta_2 \)

Here we compute the average-power-efficient mapper corresponding to \( \eta_2 \). In this case our mapper is defined by the \( M \) complex variables \( Y_i \) introduced in Sec. 4.2.1. It can be seen from Eq. 4.15 that \( u_{ij} = u_{ij} \), such that the minimum distance simplifies to

$$d_{\text{min}}^2 = \min_{i,j,k,l, j \neq k} \left( |u_{ij} - u_{ik}|^2 + |u_{jl} - u_{kl}|^2 \right) \cdot \int_{-\infty}^{\infty} q^2(t) \, dt,$$

(7.5)

where the two terms on the right hand side define the symbol distances between the symbols \( u_{ij} \) and \( u_{ik} \) or \( u_{ij} \) and \( u_{ik} \). Because the above equation must hold for all states \( i, j, k \) and \( l \), we can write

$$d_{\text{min}}^2 = \min_{i,j,k,l, j \neq k} 2 \cdot |u_{ij} - u_{ik}|^2 \cdot \int_{-\infty}^{\infty} q^2(t) \, dt.$$

(7.6)

This means that to find the minimum path distance, it is sufficient to assure a minimum symbol distance for the symbols of all branches that stem from the same node. Applying Eq. 4.15, it can be shown that this minimum distance reduces to

$$d_{\text{min}}^2 = \min_{i \neq j} 2 \cdot |Y_i - Y_j|^2 \cdot \int_{-\infty}^{\infty} q^2(t) \, dt.$$

(7.7)

With the same reasoning, the average bit energy reduces to

$$E_b = 2 \frac{\sum_{i=0}^{M-1} |Y_i|^2}{M^L \cdot \log_2(M)} \cdot \int_{-\infty}^{\infty} q^2(t) \, dt$$

(7.8)

such that the average-power-efficiency criterion can be written as

$$\max \frac{d_{\text{min}}^2}{2E_b} = \max \frac{M^L \cdot \log_2(M) \cdot \min_{i \neq j} |Y_i - Y_j|^2}{2 \sum_{i=0}^{M-1} |Y_i|^2},$$

(7.9)
where the maximization goes over the $M$ complex numbers $Y_i$ such that $\sum_{i=0}^{M-1} Y_i = 0$.

The problem defined by Eq. 7.9 is well known: Find $M$ points in the complex plane whose mutual distance exceeds a certain threshold and whose energy (defined as the sum of the squared distances from the center of gravity) is minimum. In other words, find the densest packing of $M$ circles on a two-dimensional plane. This problem was investigated some years ago to find optimum signal sets for $M$-ary QAM (see [For84]), i.e. we can make use of the known power-optimum QAM signal sets. Fig. 7.1 shows the best known constellations for $M = 2 \ldots 32$, where the cross marks the center of gravity. These constellations are all hexagonal, but for $M = 4$ there are also square and rhombus constellations with the same energy. For all these constellations, $\sum_{i=0}^{M-1} Y_i = 0$ automatically holds.

![Figure 7.1: Optimum signal sets for uncoded $M$-ary modulation](image)

The bandwidth- and average-power-efficient single delay-cell encoder for $\alpha_1 > 0$ therefore consists of the encoder shown in Fig. 4.5 with $\alpha_1 > 0$ and with the $M$-ary mapper according to Fig. 7.1.

### 7.1.2 Average-Power-Efficient Mapper For $\eta_3$

Here we compute the average-power-efficient mapper corresponding to $\eta_3$. In this case our mapper is defined by the $M$ complex variables $Z_i$ introduced in Sec. 4.2.2. It can be seen from Eq. 4.19 that because the vector $u$ is DC-free, $u_{ij} = -u_{ji}$, such that the minimum distance simplifies to

$$d_{\text{min}}^2 = \min_{i \neq j} 2 \cdot |Z_i - Z_j|^2 \cdot \int_{-\infty}^{\infty} q^2(t) \, dt.$$

(7.10)
With the same reasoning, the average bit energy reduces to

\[ E_b = 2 \left( \frac{M \sum_{i=0}^{M-1} |Z_i|^2 - |\sum_{i=0}^{M-1} Z_i|^2}{M^{L+1} \log_2(M)} \right) \cdot \int_{-\infty}^{\infty} q^2(t) dt. \]  

(7.11)

Defining the transformation

\[ Y_i = Z_i - \frac{1}{M} \sum_{j=0}^{M-1} Z_j \quad \leftrightarrow \quad Z_i = Y_i - Y_0 \]  

(7.12)

and the constraints

\[ Z_0 = 0 \quad \leftrightarrow \quad \sum_{i=0}^{M-1} Y_i = 0, \]  

(7.13)

the average-power-efficiency criterion can be written as

\[ \text{max} \frac{d_{\text{min}}^2}{2E_b} = \text{max} \frac{M^L \log_2(M) \cdot \min_{i \neq j} |Y_i - Y_j|^2}{2 \sum_{i=0}^{M-1} |Y_i|^2}, \]  

(7.14)

where the maximization goes over the \( M \) complex numbers \( Y_i \) such that \( \sum_{i=0}^{M-1} Y_i = 0 \). This equation is the same as Eq. 7.9, i.e. in the 'Y-space', the optimum mapper for \( \alpha_1 < 0 \) is also defined by Fig. 7.1. The inverse transformation to the 'Z-space' is performed by setting

\[ Z_i = Y_i - Y_0. \]  

(7.15)

This means that we have to subtract \( Y_0 \) from the output of the mapper of Fig. 7.1. Since the encoder anyway subtracts the previous from the current output symbol of the mapper, this correction is superfluous. The bandwidth- and average-power-efficient single delay-cell encoder for \( \alpha_1 < 0 \) therefore consists of the encoder shown in Fig. 4.5 with \( \alpha_1 < 0 \) and with the \( M \)-ary mapper according to Fig. 7.1.

### 7.1.3 General Average-Power-Efficient Mapper for the Single Delay-Cell Encoder

We have shown that the mapper of an average-power-efficient communication system is defined by Fig. 7.1, independent of the sign of \( \alpha_1 \). This means that the general bandwidth- and average-power-efficient single delay-cell encoder consists of the encoder shown in Fig. 4.5, where the mapper is chosen according to Fig. 7.1. The performance of the overall system will be investigated in Sec. 7.3.

### 7.2 Peak-Power-Efficient Mapper for the Single Delay-Cell Encoder

In this section we want to minimize the peak power of the optimum single delay-cell encoder. As in the last section, we have to distinguish between the two different mappers for \( \alpha_1 > 0 \) and \( \alpha_1 < 0 \). With the same transformations as in Eq. 7.12, we can show that
the optimum mapper does not depend on the sign of $\alpha_1$, such that we only compute the optimum mapper for $\alpha_1 > 0$.

According to Eq. 4.14 we can choose any $M$ complex numbers $Y_i$ to form the optimum mapper $u$, with the only restriction that $\sum_{i=1}^{M} Y_i = 0$. We now try to express the peak power of a single-delay-cell transmitter in terms of the complex numbers $Y_i$ and of the pulse shape $q(t)$. We use the fact that there is an input sequence $\{i, i, i, \ldots\}$ that produces the output sequence $\{u_n, u_{n+1}, u_{n+2}, \ldots\}$. If the input symbol $i$ is chosen such that

$$\max_i Y_i,$$  \hspace{1cm} (7.16)

the mapper produces a sequence of output symbols each having maximum amplitude. The peak power defined in Eq. 3.13 can now be written as

$$P_{\text{max}} = 2 \max_i |Y_i|^2 \cdot \max_{t=nT} \left( \sum_{n=-\infty}^{\infty} |q(t - nT)| \right)^2,$$  \hspace{1cm} (7.17)

where one term depends on the mapper and the other term depends on the pulse shaping filter. Optimizing the mapper, we consider the second term to be constant. Combining Eqs. 3.14, 7.7 and 7.17, and dropping the constant factors yields the peak-power-efficiency criterion:

$$\max \frac{d_{\text{min}}^2}{P_{\text{max}}} = \max_{i,j} \frac{|Y_i - Y_j|^2}{|Y_i|^2}, \quad i, j \in 1..M, \ i \neq j.$$  \hspace{1cm} (7.18)
Equation 7.18 defines an interesting geometrical problem that appeared in mathematical magazines during the late sixties and early seventies:

Determine the minimum diameter D of a circle which can contain M equal non-overlapping circles of diameter d.

This problem is important in packaging and in the design of ropes and conductor cables. In our terminology, the diameter d of the small circle is proportional to the minimum path distance and the term \((D - d)/2\), the greatest distance between the origin and the center of the outermost circle, is proportional to the maximum bit energy. Fig. 7.2 shows the best known configurations for \(M = 2, 4, 8\) and 16, where the desired signal points are the unmarked centers of the small circles. The first three configurations were found by Kravitz [Kra67], and the configuration for \(M = 16\) was found by Goldberg [Gol71]. The optimality of the first three configurations was proved by Pirl [Pir69], but no proof for \(M = 16\) has been found so far.

Table 7.1 gives the optimum values for \(Y_i\) for \(M = 2, 4, 8\) and 16 in polar coordinates, i.e. the coordinates \(\rho \exp(j\theta)\) of the centers of the circles where the diameter of the small circles \(d = 1\).

We finally get the following result: The general bandwidth- and peak-power-efficient single delay-cell encoder consists of the encoder shown in Fig. 4.5, where the mapper is chosen according to Fig. 7.2.

7.3 Performance of the Optimized System in an AWGN Channel

In this section we analyze the performance of the optimum system in an additive white Gaussian noise (AWGN) channel. We start with an uncoded system; these results will be used as a reference. We then consider a method to compute the performance analytically and finally show the results of a computer simulation.

It is well known (see [Pro83, p. 146]) that in additive white Gaussian noise, the bit error rate of a BPSK system is

\[
P_b = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_b}{N_0}} \right),
\]

(7.19)

where \(E_b\) denotes the bit energy, and where \(N_0\) is the one-sided spectral noise density. For a coded communication system, the performance is dominated by the minimum Euclidean distance \(d_{min}^2\). In this case, the bit error probability is approximated by

\[
P_b \approx \frac{1}{2} N(d_{min}^2) \text{erfc} \left( \sqrt{\frac{d_{min}^2}{N_0}} \right),
\]

(7.20)

where \(N(d_{min}^2)\) denotes the number of erroneous bits of all error paths with distance \(d_{min}^2\) to the correct path.

The minimum distance of a given transmitter was determined by an exhaustive search over all pairs of paths with a length of up to 8 symbols. The curve of the minimum distance contains 'U-shaped' pieces and depends on the optimization bandwidth, as is
7.3. Performance of the Optimized System in an AWGN Channel

<table>
<thead>
<tr>
<th>Input Alphabet Size $M$</th>
<th>$\rho$</th>
<th>$\vartheta$</th>
<th>$d_{\text{min}}^2/2E_{\text{b, max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.5</td>
<td>0°</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>180°</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.7071</td>
<td>0°</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.7071</td>
<td>90°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.7071</td>
<td>180°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.7071</td>
<td>270°</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0°</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>0°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>51.429°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>102.857°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>154.286°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>205.714°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>257.143°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1524</td>
<td>308.571°</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1.8077</td>
<td>289.431°</td>
<td>0.612</td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>321.544°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>353.685°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>25.772°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>57.886°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>90.000°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>122.114°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>154.228°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>186.342°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>218.456°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8077</td>
<td>250.569°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.9058</td>
<td>270.000°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8151</td>
<td>52.162°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8602</td>
<td>338.992°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8151</td>
<td>127.838°</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8602</td>
<td>211.078°</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1: Peak-Power-Efficient Signal Constellations
Figure 7.3: Minimum distance of the transmitter with $L = 2$, $T$-time-limited pulse shaping filter
shown in Fig. 7.3 for $M = 2$ and $L = 2$. The 'U-shaped' pieces belong to different paths in the code trellis that are responsible for the minimum distance. Comparing Figs. 6.5 and 7.3, it can be seen that for optimization bandwidths where the out-of-band power decreases only slowly, the minimum distance increases and vice versa. With this comparison, the bandwidth-power tradeoff becomes apparent.

<table>
<thead>
<tr>
<th># delay cells</th>
<th>BT</th>
<th>$d^2_{\text{min}}/2E_b$</th>
<th>$E_b/N_0$ comp.</th>
<th>$E_b/N_0$ sim.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.89</td>
<td>2.00</td>
<td>8.8 dB</td>
<td>9.1 dB</td>
</tr>
<tr>
<td>2</td>
<td>0.78</td>
<td>1.17</td>
<td>11.1 dB</td>
<td>10.8 dB</td>
</tr>
<tr>
<td>3</td>
<td>0.71</td>
<td>0.77</td>
<td>12.9 dB</td>
<td>12.6 dB</td>
</tr>
<tr>
<td>4</td>
<td>0.68</td>
<td>0.54</td>
<td>14.5 dB</td>
<td>14.3 dB</td>
</tr>
<tr>
<td>5</td>
<td>0.66</td>
<td>0.40</td>
<td>15.8 dB</td>
<td>15.5 dB</td>
</tr>
</tbody>
</table>

Table 7.2: Optimization bandwidths for a -20dB out-of-band power, PSWF pulses

In the sequel, we restrict ourselves to those transmitters that lead to an out-of-band power of -20 dB, which corresponds to the 99% bandwidth. The corresponding optimization bandwidths were obtained from Fig. 6.9 (time-limited pulses) and from Fig. 6.16 (spectral raised cosine pulses) and are listed in Tables 7.2 and 7.3, respectively. These tables also show the minimum distance of the corresponding system. They furthermore list the signal-to-noise ratio needed to achieve a bit error probability of $10^{-4}$, where we show both the results that were computed with Eq. 7.20 and the result of a computer simulation.

<table>
<thead>
<tr>
<th># delay cells</th>
<th>BT</th>
<th>$d^2_{\text{min}}/2E_b$</th>
<th>$E_b/N_0$ comp.</th>
<th>$E_b/N_0$ sim.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.42</td>
<td>2.00</td>
<td>8.8 dB</td>
<td>9.1 dB</td>
</tr>
<tr>
<td>2</td>
<td>0.32</td>
<td>1.21</td>
<td>11.0 dB</td>
<td>10.9 dB</td>
</tr>
<tr>
<td>3</td>
<td>0.25</td>
<td>0.79</td>
<td>12.8 dB</td>
<td>12.7 dB</td>
</tr>
<tr>
<td>4</td>
<td>0.21</td>
<td>0.56</td>
<td>14.3 dB</td>
<td>14.2 dB</td>
</tr>
<tr>
<td>5</td>
<td>0.18</td>
<td>0.41</td>
<td>15.7 dB</td>
<td>15.8 dB</td>
</tr>
</tbody>
</table>

Table 7.3: Optimization bandwidths for a -20dB out-of-band power, SRC pulses

Note that there is only a minor difference between the transmitter with the time-limited and the one with the SRC pulse shaping filter, and that the simulation results agree pretty well with the results that were computed with the approximation of Eq. 7.20. The simulation results are also shown in Fig. 7.4 for the transmitter with the time-limited and in Fig. 7.5 for the transmitter with the SRC pulse shaping filter. Note that the bandwidth reduction is realized at the cost of about 1.5 dB additional signal power per delay cell.

The single delay-cell encoder should be considered in some more detail. It is well known, that the duobinary and the dicode modulation schemes are catastrophic in the sense that there exist pairs of paths of infinite length with a finite Euclidean distance. In practice, such modulation schemes are therefore not used. In our simulation, we increased the survivor memory of the maximum likelihood sequence estimator to combat this problem. A more elegant way is to use a precoder, which eliminates this problem, see [For72]. The disadvantage of this method is the fact that each error event leads to two bit errors.
Figure 7.4: Bit error rate in AWGN with a $T$-time-limited pulse shaping filter
Figure 7.5: Bit error rate in AWGN with a spectral raised cosine pulse shaping filter
In the next section we shall investigate how a generalized precoder can be used for all C-TCM schemes.

7.4 A Nonlinear Data Precoding Scheme

In this section we present a nonlinear precoding scheme that eliminates the intersymbol interference caused by our modulation scheme, without affecting the spectral properties of the transmitted signal. The advantage of such a system is the possibility to decode the received data on a symbol-by-symbol basis, i.e. the receiver structure becomes very simple as there is no need for a maximum likelihood sequence estimator. The drawback of this method is a performance loss of several dB, as will be shown in the sequel.

In the early seventies, Harashima and Miyakawa [Har72] and Tomlinson [Tom71] independently developed a method for precoding data that are transmitted over a channel that introduces intersymbol interference. This precoding scheme was designed to eliminate the influence of the channel such that at the receiver the data are free of intersymbol interference and can be decoded by a symbol detector.

Our idea is to use this precoder together with our C-TCM transmitter. Looking at the FIR realization of this transmitter (see Fig. 5.3), it can be seen that the FIR-filter part is identical to the ISI-channel considered by Tomlinson. Hence we can combine the precoder and the FIR filter to generate our baseband signal. The schematic diagram of the modified transmitter is shown in Fig. 7.6. The M-ary input symbol \(x(n)\) is first processed by a nonlinear recursive filter. The nonlinearity of this filter consists of the \(\text{mod } N\) device defined such that it subtracts or adds \(N\) until its output signal is in the interval \((-N/2, N/2]\). The real number \(N\) has to be chosen in a suitable manner. The output signal of the nonlinear device is then filtered by the FIR filter whose coefficients are chosen as described in Chapter 5 and normalized such that \(g_0 = 1\). In Fig. 7.6, the precoder and the FIR filter are combined to get the minimum number of multipliers and adders. The resulting output signal \(u_p(t)\) is proportional to the input signal \(x(n) \pmod N\). Its power spectral density is hard to compute; for large \(N\) it is about identical to the spectrum of the FIR-filter, see [Cal89]. The transmitted baseband signal \(s_p(t)\) is then generated by filtering the signal \(u_p(n)\) with the pulse shaping filter \(Q(f)\).
7.4. A Nonlinear Data Precoding Scheme

The receiver corresponding to this transmitter is shown in Fig. 7.7. The input signal \( r(t) \), which is assumed to be the transmitted signal \( s_p(t) \) corrupted by additive white Gaussian noise, is first filtered by the matched filter \( Q^*(f) \), and then sampled with period \( T \). These samples are then reduced (mod \( N \)) (as described above), and the detector generates the estimate \( \hat{x}(n) \). Comparing this receiver with the MLSE-receiver of Fig. 2.3, it can be seen that its complexity is drastically reduced, and that it causes no decoding delay.

Note that for all selections of the filter coefficients \( g_v \) that are not integers, the output of the nonlinear device of the transmitter is approximately uniform in the interval \((-N/2, N/2] \). Because this signal is the input to the FIR filter, it is no longer possible to use a MLSE decoder, because it would have an infinite number of states. For the simple optimum modulation schemes duobinary, dicode and modified duobinary, however, it is possible to choose \( N = 2 \) such that the combined precoder / FIR filter can be represented with a trellis with a finite number of states. If we choose \( N = 2 \) for the duobinary modulation scheme \((g_0 = g_1 = 1)\), the Tomlinson filter reduces to the well-known standard precoder for the duobinary channel, see [For72].

We now compare three different realizations of the single delay-cell encoder with a time-limited pulse shaping filter and a normalized optimization bandwidth \( BT = 0.89 \). First we consider the original transmitter with a maximum likelihood sequence estimator whose survivor memory has the length of 100 symbols (compared to about 5 symbols of a traditional decoder). With this decoder, error events with an Euclidean distance \( d^2 \leq d_{\text{min}}^2/2 \) of up to length 100 can be decoded correctly. The second realization consists of a precoder and the original transmitter, i.e. it corresponds to the transmitter of Fig. 7.6 with \( g_0 = g_1 = 1 \) and \( N = 2 \). This transmitter can be represented with a trellis with two states. The receiver is an ordinary MLSE receiver adapted to this trellis. The third transmitter is the same as the second one, but the receiver is the simple symbol detector of Fig. 7.7. The out-of-band power of the different transmitters is identical, i.e. the precoder does not shape the spectrum. The out-of-band power is shown in Fig. 6.9 for \( L = 1 \). The power-efficiency of the three systems, however, looks quite different. Fig. 7.8 shows the performance of BPSK (solid line), and of the three realizations described above (dotted, dashed and dash-dotted in this order).

Note that both systems with a MLSE receiver show about the same performance. The transmitter without precoder is slightly better, because the precoder causes two bit errors per error event. The symbol detector loses about 2.5 dB with respect to the other systems and about another dB with respect to uncoded BPSK.

To quantify the performance of the symbol detector of the transmitter with two and more delay cells, we performed computer simulations for the systems with precoder, where the optimization bandwidths are those defined in Tables 7.2 and 7.3. The results
Figure 7.8: Bit error rate in AWGN for different realizations of the single delay-cell system
are shown in Figs. 7.9 and 7.10 for the time-limited and for the SRC pulse shaping filter, respectively. The signal-to-noise ratios required for a bit error probability of $10^{-4}$ are listed in Table 7.4. Comparing Tables 7.2, 7.3 and 7.4, it can be seen that the symbol detector loses between 3 and 10 dB with respect to the MLSE decoder.

<table>
<thead>
<tr>
<th># delay cells</th>
<th>PSWF BT</th>
<th>$E_b/N_0$</th>
<th>SRC BT</th>
<th>$E_b/N_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.89</td>
<td>11.8 dB</td>
<td>0.42</td>
<td>11.8 dB</td>
</tr>
<tr>
<td>2</td>
<td>0.78</td>
<td>16.1 dB</td>
<td>0.32</td>
<td>16.9 dB</td>
</tr>
<tr>
<td>3</td>
<td>0.71</td>
<td>18.2 dB</td>
<td>0.25</td>
<td>19.9 dB</td>
</tr>
<tr>
<td>4</td>
<td>0.68</td>
<td>20.2 dB</td>
<td>0.21</td>
<td>22.4 dB</td>
</tr>
<tr>
<td>5</td>
<td>0.66</td>
<td>22.0 dB</td>
<td>0.18</td>
<td>25.8 dB</td>
</tr>
</tbody>
</table>

Table 7.4: Performance of the precoder / symbol detector, $P_b = 10^{-4}$
Figure 7.9: Bit error rate in AWGN for the symbol detector, $T$-time-limited pulse shaping filter
7.4. A Nonlinear Data Precoding Scheme

Figure 7.10: Bit error rate in AWGN for the symbol detector, spectral raised cosine (SRC) pulse shaping filter
Chapter 8

Conclusions

A novel correlative trellis coded modulation (C-TCM) scheme is introduced and optimized with respect to bandwidth- and power-efficiency. The transmitter consists of an $M$-ary shift register with $L$ stages, a mapper and a pulse shaping filter. The mapper and the pulse shaping filter are optimized such that the out-of-band power of the transmitted signal is minimum. Two different classes of pulse shaping filters are considered: time-limited and band-limited (Nyquist) filters.

For the single delay-cell encoder, the well-known modulation schemes duobinary and dicode prove to be optimum. For these modulation schemes, a secondary optimization of the mapper is performed with respect to power-efficiency. The signal set of the average-power-efficient mapper is shown to be hexagonal, whereas the signal set of the peak-power-efficient mapper is 'circular'. For encoders with two to five delay cells, new modulation schemes are found. They allow a remarkable increase of the data rate by up to 300% compared to uncoded binary or quaternary modulation. This gain is realized at the cost of about 1.5 dB additional signal power per delay cell required to achieve the same performance in an AWGN channel.

To show the improvement of C-TCM over traditional modulation schemes, we compare BPSK, QPSK, CPM, M-ary QAM, TCM, convolutionally encoded QAM and C-TCM on the basis of an out-of-band power of -20dB (which corresponds to the 99% bandwidth) and a bit error rate of $10^{-4}$ in an AWGN channel. Except for CPM, we use a spectral raised cosine pulse shaping filter with a rolloff factor of 0.3. For CPM, a binary raised cosine pulse and a modulation index $h = 1/2$ are used. For each of these modulation schemes, we compute the bandwidth-efficiency $\eta$ in bits/s/Hz, and we determine the signal-to-noise ratio required for the desired bit error rate by computer simulation. As an absolute reference, we choose the Shannon bound [Sha48], which can be written as

$$\frac{E_b}{N_0} \geq \frac{2^n - 1}{\eta}. \tag{8.1}$$

The result of this comparison is shown in the bandwidth-efficiency versus signal-to-noise ratio plot of Fig 8.1. The small numbers beside some points on this plot denote the number of states of the trellis decoder, which is a measure for the receiver complexity. Points without numbers can be decoded with a symbol detector.

For the parameters mentioned above, uncoded BPSK leads to a bandwidth-efficiency of 0.88 bits/s/Hz at an $E_b/N_0$ of 8.4 dB. The bandwidth-efficiency of QPSK is twice the one of BPSK without change of the $E_b/N_0$, because QPSK uses one complex dimension
Figure 8.1: Bandwidth- and power-efficiency of several modulation schemes, bit error rate $P_b = 10^{-4}$, 99%-bandwidth
in the signal space. If a convolutional code is used with a QPSK scheme, the bandwidth-

efficiency is reduced by the rate of the code, whereas the required $E_b/N_0$ is decreased by
5 to 7 dB, as is shown for the (2,1,7) and the (4,1,14) code. In this case, the performance is increased at the cost of a decreased bandwidth-efficiency. In Fig. 8.1, this corresponds to a shift towards the lower left corner. A different result is shown for trellis codes. Here, the bandwidth-efficiency remains unchanged relative to uncoded modulation, but the required $E_b/N_0$ is decreased by up to 4 dB, which is manifested in a shift to the left in our plot. The gain of the TCM scheme is realized with an increased complexity of the receiver. Another power/bandwidth tradeoff becomes apparent, if the alphabet size of uncoded modulation schemes is increased. In this case, the bandwidth-efficiency increases, but the performance decreases. This is shown in Fig. 8.1 as a shift towards the upper right corner. The uncoded modulation schemes with a non-constant envelope, such as 8AMPM, 16QAM and 64 QAM roughly lie on a line parallel to the Shannon bound, whereas 8PSK loses about 2 dB because all signal points lie on a circle and the constellation is therefore less compact than that of QAM. CPM also shows a loss with respect to QAM, which confirms the fact that the restriction of the constant envelope leads to a decreased bandwidth/power-efficiency.

The performance of C-TCM is also shown in Fig. 8.1 for encoders with one to five delay cells and a quaternary input alphabet, corresponding to four to 64 states of the trellis. As with M-ary QAM, the increased bandwidth-efficiency is achieved at the cost of an increased $E_b/N_0$. The points roughly lie on a line parallel to the Shannon bound. Compared to M-ary QAM, C-TCM allows to increase the bandwidth-efficiency by up to 0.5 bit/s/Hz or to decrease the signal power by about 1 dB, where the increased receiver complexity is the price that has to be paid. Fig. 8.1 shows that the main application of C-TCM is in systems where a high bandwidth-efficiency of several bits/s/Hz is required.

At the beginning of our work we expected that C-TCM would lead to a bigger increase of data rate compared to uncoded modulation than it actually does. These results are somewhat disappointing, and one might guess that a more general transmitter model should lead to modulation schemes that lie closer to the Shannon bound. To understand the mechanism of the bandwidth/power tradeoff, the FIR filter transmitter is investigated in Appendix E with respect to this aspect. It is shown that bandwidth can indeed be traded for performance, but the resulting systems also lie at intermediate points on the dashed C-TCM line of Fig. 8.1. Hence, it seems that the structure of this FIR filter transmitter limits the gain compared to uncoded modulation.

To overcome this problem, we therefore try to modify the C-TCM transmitter in the sense that the mapper and the pulse shaping filter are combined to a pulse shaping filter bank, as shown in Appendix F. In this case, the output of the mapper is a pulse shape instead of a complex Dirac pulse. Hence, different pulse shapes can be generated for different contents of the shift register. However, optimization of the pulse shaping filter bank with respect to our bandwidth-efficiency criterion shows that these additional degrees of freedom do not result in a modulation scheme that lies closer to the Shannon bound, which is quite surprising. Hence the original C-TCM transmitter with a single SRC pulse shaping filter still leads to the best results.

It seems that we have reached a limit which is due to the optimization criterion, which tries primarily to minimize the out-of-band power. In this work we have shown how much bandwidth or performance can be gained with a fairly general coded transmitter
that is optimized with respect to bandwidth-efficiency. To find even better modulation schemes, we would rather have to use an optimization criterion which takes power-efficiency explicitly into account, similar to the work of Ungerboeck, or by using an intelligent combination of both bandwidth- and power-efficiency.
Appendix A

Proof Of Theorem 1

We prove that \((P - P^n)^n = 0\), \(\forall n \geq L\), where
\[
P = \|p_{im}\|, \quad i, m \in [0..M^L - 1]
\] (A.1)

and
\[
p_{im} = \begin{cases} \frac{1}{M}, & \text{if } m = iM \pmod{M^L} + n, \quad n \in [0, M - 1] \\ 0, & \text{otherwise} \end{cases}
\] (A.2)

and
\[
P^\infty = \frac{1}{M^L} \cdot K,
\] (A.3)

where \(K\) denotes the \(M^L \times M^L\) all-ones matrix. By definition, \(P^k \cdot P^\infty = P^\infty\), and \((P^\infty)^k = P^\infty\). Using these identities, the binomial series yields
\[
(P - P^n)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} P^k (P^\infty)^{n-k}
\]
\[
= \sum_{k=0}^{n-1} \binom{n}{k} (-1)^{n-k} P^k (P^\infty)^{n-k} + P^n
\]
\[
= P^\infty \cdot \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} 1^k - P^\infty + P^n
\]
\[
= P^\infty (1 - 1)^n - P^\infty + P^n
\]
\[
= P^n - P^\infty.
\] (A.4)

These algebraic manipulations show that it is equivalent to prove that \(P^L = P^\infty\). We first prove by induction, that
\[
p_{im}^{(k)} = \begin{cases} M^{-k}, & \text{if } m = iM^k \pmod{M^L} + p, \quad p \in [0, M^k - 1] \\ 0, & \text{otherwise} \end{cases}
\] (A.5)

where \(p_{im}^{(k)}\) is an element of the matrix \(P^k\). For \(k=1\), Eq. A.5 is equivalent to Eq. A.2. We now compute \(p_{im}^{(k+1)}\) by multiplying the matrices \(P^k\) and \(P\):
\[
p_{im}^{(k+1)} = \sum_{l=0}^{M^L-1} p_{il} \cdot p_{lm}^{(k)}
\] (A.6)

107
and therefore

\[ p_{im}^{(k+1)} = \sum_l^* M^{-(k+1)}, \quad (A.7) \]

where \( \sum_l^* \) denotes the sum over all \( l \in [0, M^L - 1] \) restricted by the two conditions \( l = iM \pmod{M^L} + n \) and \( m = iM^k \pmod{M^L} + p \), where \( n \in [0, M - 1] \) and \( p \in [0, M^k - 1] \). Combining these two restrictions, we can find a condition on \( m \):

\[
m = iM^{k+1} + nM^k + p \pmod{M^L} = iM^{k+1} + q \pmod{M^L}, \quad q \in [0, M^{k+1} - 1], \quad (A.8)
\]

where \( q = nM^k + p \). Because \( n \) and \( p \) define a non-overlapping, contiguous interval for \( q \), the sum of Eq. A.7 contains only one element. Eq. A.7 can therefore be rewritten:

\[
p_{im}^{(k+1)} = \begin{cases} M^{-(k+1)}, & \text{if } iM^{(k+1)} \pmod{M^L} + q, \quad q \in [0, M^{k+1} - 1] \\ 0, & \text{otherwise,} \end{cases} \quad (A.9)
\]

which proves that Eq. A.5 is correct. It now remains to compute \( p_{im}^L \):

\[
p_{im}^L = \begin{cases} M^{-L}, & \text{if } iM^L \pmod{M^L} \leq m \leq (i+1)M^L - 1 \pmod{M^L} \\ 0, & \text{otherwise,} \end{cases} \quad (A.10)
\]

which is equivalent to

\[
p_{im}^L = \begin{cases} M^{-L}, & \text{if } 0 \leq m \leq M^L - 1 \\ 0, & \text{otherwise.} \end{cases} \quad (A.11)
\]

But Eq. A.11 states that \( p_{im}^L = M^{-L} \) \( \forall i, m \in [0, M^L - 1] \) and therefore \( P^L = M^{-L} \cdot K \). This is equivalent to \( P^L = P^\infty \), which, combined with equation Eq. A.4, completes the proof.
Appendix B

Proof Of Theorem 2

In this appendix, we show that a pulse \( q(t) \) satisfies

\[
\int_{-\infty}^{\infty} q(t) \cdot q(t - nT) dt = 0 \quad \forall n \neq 0 \tag{B.1}
\]

if and only if its power spectral density \( |Q(f)|^2 \) equals

\[
|Q(f)|^2 = LP(f) + F(f), \tag{B.2}
\]

where \( LP(f) \) is the ideal low-pass filter

\[
LP(f) = \begin{cases} 
1 & \text{if } |f| < 1/2T \\
0 & \text{otherwise}
\end{cases} \tag{B.3}
\]

with the impulse response

\[
lp(t) = \text{sinc}(t) \tag{B.4}
\]

and where \( F(f) \) is a function that satisfies

\[
\sum_{j=0}^{\infty} F(f - \frac{j}{T} + \frac{1}{2T}) + \sum_{j=0}^{\infty} F(-f - \frac{j}{T} + \frac{1}{2T}) = 0, \tag{B.5}
\]

i.e. \( \sum_{j=0}^{\infty} F(f - \frac{j}{T}) \) is odd with respect to \( 1/2T \).

We start by defining the function

\[
p(t) = \int_{-\infty}^{\infty} q(s)q(s - t) ds \tag{B.6}
\]

and its Fourier transform

\[
P(f) = |Q(f)|^2. \tag{B.7}
\]

The function \( p(t) \) is real and symmetric because \( q(t) \) is real. Eq. B.1 can now be written as

\[
p(nT) = \int_{-\infty}^{\infty} q(t)q(t - nT) dt = \delta_{n0}, \tag{B.8}
\]

where \( \delta_{n0} \) is the Kronecker symbol which is always zero, except for \( n = 0 \) where it is 1. Because \( p(t) \) is real and symmetric, \( p(nT) \) can be written without loss of generality as

\[
p(nT) = \text{sinc}(nT) + f(nT), \quad f(nT) = 0 \quad \forall n, \tag{B.9}
\]

where we introduced the new function \( f(t) \) which is real and symmetric. We now show that
Lemma 1  For all real, symmetric functions \( f(t) \)
\[
f(nT) = 0 \quad \forall n \iff \sum_{n} b_n f(nT) = 0 \quad \forall b. \tag{B.10}
\]

Proof: Sufficiency:
\[
f(nT) = 0 \quad \forall n \implies \sum_{n} b_n f(nT) = 0 \quad \text{for any } b. \tag{B.11}
\]

Necessity:
Assume that there is an \( n' \) such that \( f(n'T) \neq 0 \) and that there are no restrictions on \( f(nT), n \neq n' \). This implies that \( \sum b_n f(nT) \neq 0 \) if \( b(n' \neq n) \neq 0 \) and \( b(n' = n) = 0 \), which proves the necessity.

By taking the Fourier transform of \( f(t) \), we now write the condition \( \sum_{n} b_n f(nT) = 0 \) \( \forall b \) as
\[
\sum_{n} b_n f(nT) = \frac{1}{2\pi} \sum_{n} b_n \int_{-\infty}^{\infty} F(f) \exp(j2\pi n fT) df = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(f) \sum_{n} b_n \exp(j2\pi n fT) df. \tag{B.12}
\]

\( F(f) \) is real and even because \( f(t) \) is real and even. Writing \( b = b_r + jb_i \), it can be shown that
\[
\sum_{n} b_n f(nT) = 0 \quad \forall b \iff \int_{-\infty}^{\infty} F(f) \sum_{n} b_n \cos(2\pi n fT) df = 0 \quad \forall \text{real } b. \tag{B.13}
\]

The Fourier transform of \( b \),
\[
B(f) = \sum_{n} b_n \cos(2\pi n fT) \tag{B.14}
\]
is real, symmetric and periodic. Then Eq. B.13 yields
\[
f(nT) = 0 \quad \forall n \iff \int_{-\infty}^{\infty} F(f) B(f) df = 0 \tag{B.15}
\]

for all real, symmetric and periodic functions \( B(f) \). Using the fact that \( B(f) \) is periodic with period \( 1/T \), Eq. B.15 can be written as
\[
f(nT) = 0 \quad \forall n \iff \int_{0}^{\frac{1}{T}} B(f) \sum_{j=0}^{\infty} F(f - j/T) df = 0 \tag{B.16}
\]

for all real, symmetric and periodic functions \( B(f) \). A sufficient criterion for Eq. B.16 is the condition
\[
\sum_{j=0}^{\infty} F(f - j/T) = 0, \tag{B.17}
\]

which is definition of the generalized Nyquist pulse \( F(f) \), see [Gib65]. Taking a closer look at Eq. B.16, it can be seen that there is a less stringent condition on \( F(f) \) which is both necessary and sufficient. We use the fact that \( B(f) \) is even w.r. to \( f = 0 \).
and periodic with period $1/T$. This implies that $B(f)$ is also even w.r. to $f = 1/2T$. Eq. B.16 is therefore satisfied if and only if $\sum_{j=0}^{\infty} F(f - j/T)$ is odd w.r. to $f = 1/2T$:

$$f(nT) = 0 \quad \forall n \iff \sum_{j=0}^{\infty} F(f - j/T + 1/2T) + \sum_{j=0}^{\infty} F(-f + j/T + 1/2T) = 0. \quad (B.18)$$

Combining Eq. B.18 with Eqs. B.7 and B.9 completes the proof.

We consider two special classes of functions $Q(f)$ that satisfy Eqs. B.1 and B.2: the time-limited and the Nyquist pulses. For pulses that are limited to the time interval $[0,T]$, Eq. B.1 trivially holds. We now show that of all pulses whose power spectral densities are limited to the interval $[-1/T, 1/T]$, only the Nyquist pulses satisfy Eq. B.2.

For pulses which are defined by

$$F(f) = 0, \quad |f| > 1/T, \quad (B.19)$$

the right-hand side of Eq. B.16 can be rewritten by dropping all terms for $j \neq 0$:

$$F(f + \frac{1}{2T}) + F(-f + \frac{1}{2T}) = 0. \quad (B.20)$$

But Eq. B.20 is the definition of a function that is odd with respect to the point $1/2T$. Combining Eqs. B.20 and B.2 shows that the Fourier transform of all pulses that satisfy Eqs. B.1 and B.19 are constructed by combining an ideal low-pass filter with any filter whose Fourier transform is odd with respect to $1/2T$. But this is exactly the definition of a Nyquist pulse [Nyl28].
B. Proof Of Theorem 2
Appendix C

Euclidean Distance Of Signals Generated By A Shift-Register Encoder

In this appendix, we express the Euclidean distance of two output signals of a shift-register encoder/mapper as a function of the difference of the output symbols of the mapper. We start by the definition of the Euclidean distance of two signals:

\[ d^2 = \int |s^{(0)}(t) - s^{(1)}(t)|^2 dt. \]  

(C.1)

The signals \( s^{(0)}(t) \) and \( s^{(1)}(t) \) that are generated by a shift-register encoder/mapper can be written as

\[ s^{(0)}(t) = \sum_{n=-\infty}^{\infty} u^{(0)}(n) \cdot q(t - nT), \]  

(C.2)

and

\[ s^{(1)}(t) = \sum_{n=-\infty}^{\infty} u^{(1)}(n) \cdot q(t - nT), \]  

(C.3)

where \( \{u^{(0)}(n)\} \) and \( \{u^{(1)}(n)\} \) are the output sequences of the mapper. The difference of the signals \( s^{(0)}(t) \) and \( s^{(1)}(t) \) can now be expressed as

\[ s^{(0)}(t) - s^{(1)}(t) = \sum_n \Delta u_n \cdot q(t - nT), \]  

(C.4)

where \( \Delta u_n \) is defined as the difference between \( u^{(0)}(n) \) and \( u^{(1)}(n) \):

\[ \Delta u_n = u^{(0)}(n) - u^{(1)}(n). \]  

(C.5)

Combining Eqs. C.1 and C.4 yields

\[ d^2 = \sum_i \int_{-\infty}^{\infty} |\Delta u_i \cdot q(t - iT)|^2 dt + \sum_{i \neq j} \Delta u_i \Delta u_j \cdot \int_{-\infty}^{\infty} q(t - iT)q(t - jT)dt, \]  

(C.6)

where we used the fact that \( q(t) \) is real. Applying Theorem 2, the second term of Eq. C.6 vanishes for all generalized Nyquist pulses. The Euclidean distance of two output signals can therefore be expressed in terms of the difference of the output sequences of the mapper:

\[ d^2 = \sum_i |\Delta u_i|^2 \cdot \int_{-\infty}^{\infty} q^2(t)dt. \]  

(C.7)
C. Euclidean Distance Of Signals Generated By A Shift-Register Encoder
Appendix D

Derivation of Eq. 5.14

In this section, we show that for Nyquist pulses,

\[
\int_{-\infty}^{\infty} |Q(f)|^2 \exp(-j2\pi\nu fT) df = \delta_{\nu 0} \cdot \int_{-\infty}^{\infty} |Q(f)|^2 df. \tag{D.1}
\]

We split the integrand \(|Q(f)|^2 \exp(-j2\pi\nu fT)\) into two parts,

\[
|Q(f)|^2 \exp(-j2\pi\nu fT) = Q(f) \cdot Q^*(f) \exp(-j2\pi\nu fT). \tag{D.2}
\]

The inverse Fourier transform of the first term is

\[
F_1(f) = Q(f) \rightarrow q(t) = f_1(t) \tag{D.3}
\]

and the inverse Fourier transform of \(\exp(+j2\pi\nu fT)\) is

\[
F_2(f) = \exp(+j2\pi\nu fT) \rightarrow \delta(t + \nu T) = f_2(t). \tag{D.4}
\]

Using the convolution theorem of the Fourier transform

\[
F_1(f)F_2(f) \rightarrow \int_{-\infty}^{\infty} f_1(\tau)f_2(t - \tau)d\tau \tag{D.5}
\]

we find that

\[
F_3(f) = Q(f)\exp(+j2\pi\nu fT) \rightarrow \int_{-\infty}^{\infty} q(\tau)\delta(t + \nu T - \tau)d\tau = q(t + \nu T) = f_3(t). \tag{D.6}
\]

Applying Parseval’s theorem

\[
\int_{-\infty}^{\infty} F_1(f)F_3^*(f) df = \int_{-\infty}^{\infty} f_1(t)f_3(t) dt \tag{D.7}
\]

yields

\[
\int_{-\infty}^{\infty} |Q(f)|^2 \exp(-j2\pi\nu fT) df = \int_{-\infty}^{\infty} q(t) \cdot q(t + \nu T)dt. \tag{D.8}
\]

According to Theorem 2, the right hand side of the above equation is zero for \(\nu \neq 0\), such that applying Parseval’s theorem again we have finally shown that Eq. D.1 holds.
D. Derivation of Eq. 5.14
Appendix E

Bandwidth / Power Tradeoff of the
FIR Filter Transmitter

In Chapter 7 we showed that the reduction of the out-of-band power of the C-TCM transmitter results in a decreased performance due to the reduced minimum Euclidean distance. Here, we investigate the same FIR filter transmitter, and we show that bandwidth-efficiency can be traded for power-efficiency, but that we cannot get closer to the Shannon bound than the optimum bandwidth-efficient transmitter.

In contrast to Chapter 5, we do not optimize the filter coefficients here, but we try to express the out-of-band power in terms of a distance measure, and observe the bandwidth behaviour as the distance measure is varied. This allows us to see how bandwidth is traded for performance. We assume that the input alphabet is binary with elements +1 and −1.

The bit error rate of a coded system in an AWGN channel can be approximated as a function of the minimum Euclidean distance \(d_{\text{min}}^2\). This distance, however, is a complicated function of the filter coefficients, and no closed-form expression could be found in the literature. To find the minimum distance we have to compare all pairs of paths stemming from the same node of the trellis that merge after a certain number of transitions. The pair of paths with minimum mutual distance is responsible for the asymptotical error performance of the system. Because there is no closed-form expression for the minimum distance, we define another distance measure that will prove to be both useful and simple.

As already mentioned, the correct path and an error path stem from the same node. This means that the Euclidean distance of the path segment where the paths diverge \(d_{\text{div}}^2\) is a lower bound on the total distance. Because this reasoning also holds for the distance \(d_{\text{merge}}^2\) of the merging path segments, a lower bound for \(d_{\text{min}}^2\) is given by

\[
d_{\text{min}}^2 \geq \min (d_{\text{div}}^2 + d_{\text{merge}}^2).
\]  

(E.1)

A similar distance measure was used by Ungerboeck [Ung87] who designed trellis codes that maximize both \(d_{\text{div}}^2\) and \(d_{\text{merge}}^2\). We now express \(d_{\text{div}}^2\) and \(d_{\text{merge}}^2\) in terms of the filter coefficients \(g_{\nu}, \nu = 0 \ldots L\). We assume that two paths diverge at a certain node. This means that the two paths are identical up to this node, such that the contents of the corresponding filter delay cells are also identical. The distance \(d_{\text{div}}^2\) therefore only
depends on the first coefficient $g_0$, such that for a binary input alphabet

$$d_{\text{div}}^2 = (2g_0)^2.$$  \hspace{1cm} (E.2)

Now let us consider the merging distance. After the two paths have merged, they are identical, as well as the contents of the delay cells. Hence, the contents of the delay cells before the merger differ only in the last cell with the coefficient $g_L$. The merging distance is therefore equal to

$$d_{\text{merge}}^2 = (2g_L)^2,$$  \hspace{1cm} (E.3)

where we made the same assumptions for the input symbols as above. Combining these results yields a lower bound on the minimum distance

$$d_{\text{min}}^2 \geq 4 \left( g_0^2 + g_L^2 \right).$$  \hspace{1cm} (E.4)

We shall use this distance measure in the following sections to show the dependence of the out-of-band power on the Euclidean distance. We start with the filter with two coefficients ($L = 1$) and present results for $L = 2$ and $L = 3$.

### E.1 FIR Filter With a Single Delay Cell, L = 1

The FIR filter with a single delay cell introduces a memory of length one, such that the length of error paths is greater than or equal to two. The shortest paths therefore consist exactly of a diverging and a merging pair of path segments, such that Eq. E.4 holds with equality. Normalizing the distance by dividing it by twice the bit energy yields

$$\frac{d_{\text{min}}^2}{2E_b} = 4 \left( \frac{1}{2} \cdot \left( \frac{g_0^2 + g_1^2}{E_b} \right) \right) = 2,$$  \hspace{1cm} (E.5)

where $E_b = \sum_{\nu=0}^{L} g_{\nu}^2$. This normalized distance is the same as for an uncoded BPSK system. The performance of this filter, however, is worse than that of BPSK, because there also exist longer paths with the same distance. At a bit error rate of $10^{-4}$, the loss in SNR is about 1 dB.

Because the normalized minimum distance of this transmitter is independent of the filter coefficients, we cannot trade bandwidth for performance.

### E.2 FIR filters With Two Delay Cells, L = 2

The length of the shortest error path (not necessarily the one with minimum distance) of the encoder with two delay cells is three, except for the degenerate case where $g_0 = 0$ or $g_2 = 0$. Hence, in addition to the diverging and the merging path segments there is at least another pair of path segments that contributes to the minimum distance. If this distance is lower-bounded by Eq. E.4, the normalized distance yields

$$\frac{d_{\text{min}}^2}{2E_b} \geq 2 \cdot \frac{g_0^2 + g_L^2}{g_0^2 + g_1^2 + g_2^2},$$  \hspace{1cm} (E.6)

with equality if and only if the minimum distance of the path segments between the diverging and the merging paths is zero. We now attempt to show the relation between
the out-of-band power and the minimum distance of the transmitted signal. As the optimization of the filter coefficients showed, the filter coefficients are either symmetric or skew-symmetric. Hence, for a transmitter with two delay cells \( g_0^2 = g_2^2 \), such that \( d_{\text{div}}^2 = d_{\text{merge}}^2 = d^2/2 \). This implies that the absolute value of the filter coefficients \( |g_0| = |g_2| \). Normalizing the coefficients to unit energy yields \( |g_0| = |g_2| = d/\sqrt{2} \) and \( |g_1| = \sqrt{1 - d^2} \). The influence of the signs of the coefficients will be discussed later. Hence, we have found a method to define the FIR filter with 2 delay cells with a single parameter \( d \), where the only assumption is that the coefficients be symmetric or skew symmetric. The inband power of this transmitter equals

\[
\eta = \frac{g^\dagger H(B)g}{g^\dagger g}, \tag{E.7}
\]

where

\[
H(B) = \begin{bmatrix}
\alpha_0 & \alpha_1 & \alpha_2 \\
\alpha_1 & \alpha_0 & \alpha_1 \\
\alpha_2 & \alpha_1 & \alpha_0
\end{bmatrix}, \tag{E.8}
\]

where the coefficients \( \alpha_\nu \) are defined in Eq. 2.46. Writing Eq. E.7 with the components \( g_0 \ldots g_2 \) yields

\[
\eta = \alpha_0 + 2\alpha_1(g_0g_1 + g_1g_2) + 2\alpha_2g_0g_2, \tag{E.9}
\]

where we used the fact that \( g_0^2 + g_1^2 + g_2^2 = 1 \). Fig. E.1 shows the inband power of a transmitter with a SRC pulse shaping filter with a rolloff factor of 0.3 as a function of the bandwidth \( B \), where four different values for the parameter \( d \) are chosen.

![Inband Power Graph](image)

Figure E.1: Inband power for different parameters \( d \)

To eliminate the dependence on \( B \), we choose \( B \) such that an out-of-band power of -20 dB results, i.e. \( B \) is the 99%-bandwidth. Now we can express the bandwidth-efficiency \( \eta \) in bits/sec/Hz in terms of the parameter \( d \) by solving

\[
\frac{g^\dagger H(B)g}{g^\dagger g} = 0.99. \tag{E.10}
\]
The result of this operation is shown Fig. E.2, where we assume that all filter coefficients are positive. The shape of the curve of this figure is somewhat surprising, because we would expect that the bandwidth-efficiency decreases with an increasing minimum distance. The problems of this representation is that the parameter $d$ is not the minimum distance itself, but only an approximation. As will be shown in the sequel this approximation is not very accurate.

![Figure E.2: Bandwidth-efficiency of the FIR encoder with two taps](image)

The decreasing part of the curve of Fig. E.2 is intuitively clear. The increasing part can be explained by the fact that for $d = 0$ we get $g_0 = g_2 = 0, g_1 = 1$ which results in BPSK. Since we know that the bandwidth-efficiency can be increased by the introduction of correlation, it is obvious that $\eta$ has to be increasing for $d$ being somewhere in the interval $[0, 1]$.

Now the question about the accuracy of the approximation of the minimum distance (Eq. E.4) arises. To answer this question, we compute the minimum distance for a given $d$ for all pairs of paths up to a certain depth. This depth is chosen such that the distance of the smallest unmerged path is bigger than the minimum distance. With this approach, the minimum distance can be computed exactly. The result is shown in Table E.1 for $d \in [0.5, 0.7]$

It can be seen that in contrast to our assumption the minimum distance decreases for an increasing parameter $d$ in the interval that is considered. (This interval is chosen such that a spectral efficiency results that is significantly better than that of uncoded BPSK.) This is explained by the fact that the mutual distance of the intermediate path segments decreases faster than the distance of the diverging and the merging path segments increases.

For $d = 0.645$ we get the optimum bandwidth-efficient transmitter defined in Table 7.3 for $L = 2$. The bandwidth- and power-efficiencies of this transmitter are also shown in Fig. 8.1 in the C-TCM curve with the square denoted with the figure '8'. For a decreasing value of $d$ the resulting transmitter moves in Fig. 8.1 on the dashed line
Hence with such a FIR filter transmitter, we can indeed trade bandwidth- for power-efficiency, but the resulting systems are at best on a line parallel to the Shannon bound.

For completeness, we also considered configurations for which not all coefficients of the FIR filter are positive. The bandwidth-efficiency of these systems compared to the optimum system, however, is reduced by about a factor of 2, such that these modulation schemes are not of practical interest.

To conclude, for a FIR filter with three symmetric coefficients and an out-of-band power of -20 dB we can trade bandwidth- for power-efficiency, but we cannot get closer to the Shannon bound than the optimum system that was found in Sec. 6.2.2. It can be shown that the same conclusion holds for a FIR transmitter with 3 delay cells, and we do not expect that this changes for filters with more than 3 delay cells.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$\eta$ (bit/s/Hz)</th>
<th>$d_{\text{min}}^2 / 2E_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>2.54</td>
<td>1.55</td>
</tr>
<tr>
<td>0.52</td>
<td>2.66</td>
<td>1.49</td>
</tr>
<tr>
<td>0.54</td>
<td>2.77</td>
<td>1.43</td>
</tr>
<tr>
<td>0.56</td>
<td>2.87</td>
<td>1.38</td>
</tr>
<tr>
<td>0.58</td>
<td>2.96</td>
<td>1.33</td>
</tr>
<tr>
<td>0.60</td>
<td>3.04</td>
<td>1.28</td>
</tr>
<tr>
<td>0.62</td>
<td>3.10</td>
<td>1.25</td>
</tr>
<tr>
<td>0.64</td>
<td>3.14</td>
<td>1.22</td>
</tr>
<tr>
<td>0.66</td>
<td>3.15</td>
<td>1.20</td>
</tr>
<tr>
<td>0.68</td>
<td>3.10</td>
<td>1.18</td>
</tr>
<tr>
<td>0.70</td>
<td>2.84</td>
<td>1.17</td>
</tr>
</tbody>
</table>

Table E.1: Spectral efficiency and minimum distance as function of $d$
E. Bandwidth / Power Tradeoff
Appendix F

Trellis Encoder with a Pulse Shaping Filter Bank

In this appendix we investigate a trellis encoder with a pulse shaping filter bank. In contrast to our original transmitter (see Fig. 2.2), we combine the mapper and the pulse shaping filter to a filter bank, see Fig. F.1. We assume that iid M-ary input symbols $x(n)$ enter the shift register at time intervals $T$. At every such interval, the filter bank outputs a pulse shape $q_{ij}(t)$ depending on the current and on the $L$ previous input symbols, where $i$ denotes the state of the shift-register, and where $j$ is the current input symbol, as defined in Sec. 2.1. I.e. the filter bank consists of $R = M^{L+1}$ pulse shapes. The original transmitter is a special case of Fig. F.1, where the pulse shapes are chosen such that they are a complex multiple of a base pulse $q(t)$,

$$q_{ij}(t) = u_{ij} \cdot q(t), \quad (F.1)$$

where $u_{ij}$ is the complex output of the mapper of Fig. 2.2.

![Figure F.1: Schematic diagram of the encoder with a pulse shaping filter bank](image)

We assume that the impulse response of all filters $q_{ij}(t)$ have length $\gamma T$, where $\gamma$ is a positive integer. We further assume that all pulses are orthogonal to shifted copies of
themselves and to shifted copies of all other pulses, but they need not be orthogonal to unshifted pulses, i.e.

$$\int_{-\infty}^{\infty} q_{ij}(t) q_{kl}(t + nT) = \delta_{n0} \quad \forall \ (ij), (kl)$$  \quad (F.2)

Note that for $\gamma = 1$ the above criterion trivially holds.

This restriction assures that overlapping pulses do not introduce further memory, i.e. the correlation of the output signal depends only on the structure of the shift-register. For pulses that do not satisfy Eq. F.2, the block diagram of Fig. F.1 can be expanded by adding $\gamma - 1$ delay cells and by using an equivalent $\gamma' = 1$. We then get $M^L+\gamma$ pulse shapes $q'_{kl}(t)$ that are linear combinations of the (shifted) versions of the pulse shapes $q_{ij}(t)$ of the original filter bank. Because these two representations are equivalent, we restrict ourselves to the original transmitter, i.e. to filters that satisfy Eq. F.2.

In the next sections, we first compute the power spectral density of the output signal of the transmitter of Fig. F.1, and optimize the filter bank with respect to bandwidth-efficiency. We then show the performance of the optimized systems and compare them to the encoder with a single pulse shaping filter. Because we can use the same methods as introduced earlier, we concentrate mainly on the results.

**F.1 Spectrum Calculation and Optimization of a Shift-Register Encoder With a Pulse Shaping Filter Bank**

In this section we compute the power spectral density of the output signal of a shift-register encoder with $L$ delay cells followed by a pulse shaping filter bank, as shown in Fig. F.1. We follow the results presented in [Bec88] and adapt them to our notation. The power spectral density of the output signal can then be written as

$$W(f) = \frac{1}{T} \left[ \sum_{l=0}^{L} Q_l(f) W_\nu Q(f) \cos(2\pi \nu fT) + j \cdot Q_l(f) V_\nu Q(f) \sin(2\pi \nu fT) \right], \quad (F.3)$$

where the matrices $W_\nu$ and $V_\nu$ are defined in Eqs. 2.28 and 2.29, and where $Q(f)$ is the column vector containing the Fourier transforms of all pulses $q_{ij}(t)$ in the same manner as in Sec. 2.1 the vector $u$ contained the output symbols of the mapper $u_{ij}$, see Eq. 2.15.

We now assume that the different pulse shapes consist of $\kappa$ samples per time interval $T$, i.e. there are $\gamma \kappa$ samples per pulse. We further assume that the pulses satisfy Eq. F.2. Then the relative inband power can be written as

$$\eta = \frac{\int_{-B}^{B} W(f) df}{\int_{-\infty}^{\infty} W(f) df} = \frac{q^i C q}{q^i q}, \quad (F.4)$$

where the column vector $q$ contains all samples of all pulse shaping filters

$$q = \left[ q_1(0), \ldots, q_1 \left( \frac{\gamma \kappa - 1}{\kappa} T \right), q_2(0), \ldots, q_2 \left( \frac{\gamma \kappa - 1}{\kappa} T \right), \ldots, q_R(0), \ldots, q_R \left( \frac{\gamma \kappa - 1}{\kappa} T \right) \right]'$$  \quad (F.5)
where $R = M^{L+1}$. The matrix $C$ can be shown to be

$$C = B \cdot \left[ \sum_{\nu=0}^{L} (A^{(+)}_\nu + A^{(-)}_\nu) \times W_\nu - (A^{(+)}_\nu - A^{(-)}_\nu) \times V_\nu \right],$$  \hspace{1cm} (F.6)$$

where $\times$ denotes the Kronecker matrix product $A \times B = \|a_{ij}B\|$. The matrices $A^{(+)}_\nu$ and $A^{(-)}_\nu$ are defined as

$$A^{(+)}_\nu = \|a^{(+)}_{ij}\| \quad \text{and} \quad A^{(-)}_\nu = \|a^{(-)}_{ij}\|,$$  \hspace{1cm} (F.7)$$

where

$$a^{(+)}_{ij} = \text{sinc} \left( 2\pi BT \left( \nu + \frac{(i-j)}{\kappa} \right) \right), \quad i,j = 0 \ldots \gamma \kappa - 1$$  \hspace{1cm} (F.8)$$

and

$$a^{(-)}_{ij} = \text{sinc} \left( 2\pi BT \left( \nu - \frac{(i-j)}{\kappa} \right) \right), \quad i,j = 0 \ldots \gamma \kappa - 1$$  \hspace{1cm} (F.9)$$

Eq. F.4 is a Rayleigh coefficient, i.e. the inband power of the transmitter of Fig. F.1 is maximized by the eigenvector of $C$ corresponding to the maximum eigenvalue. The optimum pulse shapes, the resulting inband power and the error performance are presented in the next section.

---

### F.2 Bandwidth- and Power-Efficiency of the Pulse Shaping Filter Bank Transmitter, $\gamma = 1$

Here we present the bandwidth-efficiency and the performance of the pulse shaping filter bank transmitter, where the length of the impulse response equals one symbol, i.e. $\gamma = 1$. For the binary encoder with a single delay cell, the filter bank consists of four different pulse shapes. They were obtained by computing the eigenvector corresponding to the maximum eigenvalue of the matrix $C$ defined in Eq. F.6 for $M = 2$ and $L = 1$. The resulting pulse shapes are real; they are shown in Fig. F.2 for a normalized optimization bandwidth $BT = 0.56$ and $\kappa = 8$ samples per symbol.

If the previous input symbol was $-1$, either the solid or the dotted pulse shape is output, depending on the current input symbol. If the previous input symbol was $+1$, either the dashed or the dashed-dotted pulse is transmitted. This means that the resulting pulse shape is smooth with only small steps, because due to the memory of the transmitter e.g. the dashed pulse will never follow the solid pulse. This leads to a decreased out-of-band power, as shown in Fig. F.3 for zero to four delay cells.

<table>
<thead>
<tr>
<th># delay cells</th>
<th>$BT$</th>
<th>bandwidth-efficiency</th>
<th>$E_b/N_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.09</td>
<td>0.92 bit/s/Hz</td>
<td>8.4 dB</td>
</tr>
<tr>
<td>1</td>
<td>0.56</td>
<td>1.79 bit/s/Hz</td>
<td>8.4 dB</td>
</tr>
<tr>
<td>2</td>
<td>0.38</td>
<td>2.63 bit/s/Hz</td>
<td>10.3 dB</td>
</tr>
<tr>
<td>3</td>
<td>0.28</td>
<td>3.57 bit/s/Hz</td>
<td>12.2 dB</td>
</tr>
<tr>
<td>4</td>
<td>0.22</td>
<td>4.55 bit/s/Hz</td>
<td>14.2 dB</td>
</tr>
</tbody>
</table>

Table F.1: Bandwidth-efficiency and performance of the filter bank encoder, $\gamma = 1$
Note that there is a remarkable reduction of the out-of-band power compared to the optimum uncoded transmitter (solid line). For four delay cells the numerical results are only reliable up to an out-of-band power of about -25dB and for five delay cells they are not reliable at all because of the size of the matrix $C$, which for $L = 5$ is $512 \times 512$. In the sequel we therefore consider only encoders with up to four delay cells.

The reduction of the out-of-band power, however, is again realized at the cost of power-efficiency. The results of a simulation on an AWGN channel are shown in Fig. F.4 for up to four delay cells, where the optimization bandwidths were again chosen such that an out-of-band power of -20dB results (this corresponds to the 99% bandwidth). The single delay-cell decoder shows about the same performance as the uncoded transmitter. Any additional delay cell, however, decreases the performance by about 2 dB. The achievable bandwidth-efficiency in $\text{bit/s/Hz}$ for the 99% bandwidth and the signal-to-noise ratio required for an error rate of $10^{-4}$ are listed in Table F.1.

As already shown in Fig. 4.10, we assumed that the input alphabet is quaternary, where the even and the odd bits are transmitted on the $I$- and on the $Q$-channel,
F.2. Bandwidth- and Power-Efficiency, $\gamma = 1$

Figure F.3: Achievable out-of-band power of the pulse shaping filter bank transmitter, $\gamma = 1$

respectively.

Fig. F.5 compares these results (solid line) with those of the encoder with a sampled time-limited single pulse shaping filter (dashed line), as listed in Tables 6.1 and 7.2. It is remarkable, how much bandwidth-efficiency can be increased with a pulse shaping filter bank with respect to a single time-limited pulse shaping filter, with roughly the same performance. It is disappointing, however, that an uncoded QAM-transmitter with a spectral raised cosine pulse shaping filter and rolloff factor 0.3 (dashed-dotted line) is only slightly worse than the filter bank transmitter, whose complexity is much higher. Furthermore, the C-TCM transmitter with the same single SRC filter (dotted line) is still better than the filter bank transmitter. This drawback can be compensated for by considering filters with longer impulse responses, as will be shown in the next section.
Figure F.4: Bit error rate in AWGN of the pulse shaping filter bank transmitter, $\gamma = 1$
Figure F.5: Bandwidth- vs. power-efficiency, filter bank transmitter
F.3 Bandwidth- and Power-Efficiency of the Pulse Shaping Filter Bank Transmitter, $\gamma > 1$

In this section we investigate a transmitter with a pulse shaping filter bank whose pulses are longer than one symbol. We assume that the pulses are orthogonal in the sense defined by Eq. F.2. As described above, otherwise the shift register can be extended by $\gamma - 1$ cells, and the length of the pulse shapes reduces to $\gamma' = 1$, such that the results of the previous section apply.

![Diagram of pulse shapes for an encoder with two delay-cells, $\gamma = 3$](image)

Figure F.6: Pulse shapes for an encoder with two delay-cells, $\gamma = 3$

The inband power of a pulse shaping filter bank transmitter is now still defined by Eq. F.4, but the pulses have to satisfy Eq. F.2. Therefore the optimum pulses cannot be found by the computation of an eigenvector, but by maximizing the expression of Eq. F.4 under the constraints of Eq. F.2. This was performed numerically using a general purpose maximization routine on a computer. As an example, we show the eight different pulse shapes of the encoder with two delay cells and $\gamma = 3$ in Fig. F.6. Note that the optimum pulse shapes are more or less a scalar multiple of a base pulse that looks like a truncated $\sin(x)/x$ pulse. Keeping this in mind, it is not surprising that with such a transmitter roughly the same bandwidth-efficiency and performance is obtained as with an encoder with a single spectral raised cosine pulse shaping filter.
To summarize, we can say that the introduction of a pulse shaping filter bank increases the complexity, but does not gain any power- or bandwidth-efficiency compared to the C-TCM transmitter with an single spectral raised cosine pulse shaping filter. This fact is surprising, as we expected that with the increased degrees of freedom introduced by the filter bank, a transmitter that operates closer to the Shannon bound should be found. As shown, this statement only holds compared to time-limited, but not to SRC pulse shaping filters. Hence, among all those different transmitters that were investigated the C-TCM transmitter with a single SRC pulse shaping filter still leads to the best results.
F. Trellis Encoder with a Pulse Shaping Filter Bank
### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g^{(1)} )</td>
<td>FIR filter vector, symmetric, odd coeff.</td>
<td>Eq. 5.22</td>
</tr>
<tr>
<td>( g^{(1)} )</td>
<td>FIR filter vector, symmetric, odd coeff.</td>
<td>Eq. 5.34</td>
</tr>
<tr>
<td>( g^{(2)} )</td>
<td>FIR filter vector, symmetric, even coeff.</td>
<td>Eq. 5.47</td>
</tr>
<tr>
<td>( g^{(3)} )</td>
<td>FIR filter vector, skew-symm., odd coeff.</td>
<td>Eq. 5.65</td>
</tr>
<tr>
<td>( g^{(4)} )</td>
<td>FIR filter vector, skew-symm., even coeff.</td>
<td>Eq. 5.79</td>
</tr>
<tr>
<td>( A_{\nu} )</td>
<td>( R \times R ) encoder matrices</td>
<td>Eq. 2.27</td>
</tr>
<tr>
<td>( A_{\nu}^{(+)} ), ( A_{\nu}^{(-)} )</td>
<td>filter bank encoder matrices</td>
<td>Eq. F.6</td>
</tr>
<tr>
<td>( B )</td>
<td>bandwidth</td>
<td>Eq. 3.1</td>
</tr>
<tr>
<td>( C )</td>
<td>filter bank encoder matrix</td>
<td>Eq. F.6</td>
</tr>
<tr>
<td>( C_0 )</td>
<td>mean square value of output symbols</td>
<td>Eq. 2.19</td>
</tr>
<tr>
<td>( d_{\text{min}} )</td>
<td>minimum Euclidean distance</td>
<td>Eq. 3.11</td>
</tr>
<tr>
<td>( d_{\text{km}} )</td>
<td>output probability</td>
<td>Eq. 2.6</td>
</tr>
<tr>
<td>( D )</td>
<td>matrix decomposition of ( H^{(1)}(\infty) )</td>
<td>Eq. 5.33</td>
</tr>
<tr>
<td>( E )</td>
<td>( R \times N ) encoder matrix</td>
<td>Eq. 2.13</td>
</tr>
<tr>
<td>( E_b )</td>
<td>average bit energy</td>
<td>Eq. 3.10</td>
</tr>
<tr>
<td>( F(f) )</td>
<td>part of pulse shaping filter</td>
<td>Eq. 3.6</td>
</tr>
<tr>
<td>( g(x(n),a(n)) )</td>
<td>state transition function of the finite state machine</td>
<td>Eq. 2.1</td>
</tr>
<tr>
<td>( g )</td>
<td>Vector of FIR filter coefficients</td>
<td>Eq. 5.2</td>
</tr>
<tr>
<td>( G )</td>
<td>( N \times R ) encoder matrix</td>
<td>Eq. 2.17</td>
</tr>
<tr>
<td>( G(f) )</td>
<td>transfer function of FIR filter</td>
<td>Eq. 5.20</td>
</tr>
<tr>
<td>( h(x(n),\sigma(n)) )</td>
<td>output function of the finite state machine</td>
<td>Eq. 2.2</td>
</tr>
<tr>
<td>( H )</td>
<td>( R \times R ) encoder matrix</td>
<td>Eq. 2.20</td>
</tr>
<tr>
<td>( H(B) )</td>
<td>FIR filter matrix</td>
<td>Eq. 5.11</td>
</tr>
<tr>
<td>( H^{(1)}(B) )</td>
<td>FIR filter matrix, symmetric, odd coeff.</td>
<td>Eq. 5.26</td>
</tr>
<tr>
<td>( H^{(2)}(B) )</td>
<td>FIR filter matrix, symmetric, even coeff.</td>
<td>Eq. 5.50</td>
</tr>
<tr>
<td>( H^{(3)}(B) )</td>
<td>FIR filter matrix, skew-symm., odd coeff.</td>
<td>Eq. 5.68</td>
</tr>
<tr>
<td>( H^{(4)}(B) )</td>
<td>FIR filter matrix, skew-symm., even coeff.</td>
<td>Eq. 5.82</td>
</tr>
<tr>
<td>( H(\infty) )</td>
<td>FIR filter matrix</td>
<td>Eq. 5.13</td>
</tr>
<tr>
<td>( H^{(1)}(\infty) )</td>
<td>FIR filter matrix, symmetric, odd coeff.</td>
<td>Eq. 5.29</td>
</tr>
<tr>
<td>( H^{(2)}(\infty) )</td>
<td>FIR filter matrix, symmetric, even coeff.</td>
<td>Eq. 5.52</td>
</tr>
<tr>
<td>( H^{(3)}(\infty) )</td>
<td>FIR filter matrix, skew-symm., odd coeff.</td>
<td>Eq. 5.70</td>
</tr>
<tr>
<td>( H^{(4)}(\infty) )</td>
<td>FIR filter matrix, skew-symm., even coeff.</td>
<td>Eq. 5.84</td>
</tr>
<tr>
<td>( I )</td>
<td>identity matrix</td>
<td>P. 17</td>
</tr>
<tr>
<td>( K )</td>
<td>matrix of ones</td>
<td>P. 17</td>
</tr>
<tr>
<td>( L )</td>
<td>number of shift register stages</td>
<td>P. 12</td>
</tr>
<tr>
<td>( LP(f) )</td>
<td>ideal low pass filter</td>
<td>Eq. 3.7</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>$M$</td>
<td>input alphabet size</td>
<td>P. 11</td>
</tr>
<tr>
<td>$n$</td>
<td>discrete time</td>
<td>P. 13</td>
</tr>
<tr>
<td>$n_r$</td>
<td>decoding delay</td>
<td>P. 12</td>
</tr>
<tr>
<td>$N$</td>
<td>number of encoder states</td>
<td>P. 13</td>
</tr>
<tr>
<td>$N'$</td>
<td>precoder modulo number</td>
<td>Fig. 7.6</td>
</tr>
<tr>
<td>$p$</td>
<td>vector of frequency integrals</td>
<td>Eq. 2.45</td>
</tr>
<tr>
<td>$P$</td>
<td>$N \times N$ state transition matrix</td>
<td>Eq. 2.4</td>
</tr>
<tr>
<td>$P^\infty$</td>
<td>$N \times N$ stationary probability matrix</td>
<td>Eq. 2.9</td>
</tr>
<tr>
<td>$P_{\text{max}}$</td>
<td>peak power of transmitted signal</td>
<td>Eq. 3.13</td>
</tr>
<tr>
<td>$q(t)$</td>
<td>impulse response of pulse shaping filter</td>
<td>Eq. 2.3</td>
</tr>
<tr>
<td>$q_i$</td>
<td>$i$th sample of pulse shape</td>
<td>Eq. 2.42</td>
</tr>
<tr>
<td>$q$</td>
<td>vectorized pulse shape</td>
<td>Eq. 2.44</td>
</tr>
<tr>
<td>$Q(f)$</td>
<td>transfer function of pulse shaping filter</td>
<td>P. 11</td>
</tr>
<tr>
<td>$Q^*(f)$</td>
<td>transfer function of matched filter</td>
<td>P. 12</td>
</tr>
<tr>
<td>$r(t)$</td>
<td>received signal</td>
<td>P. 12</td>
</tr>
<tr>
<td>$R$</td>
<td>total number of state transitions</td>
<td>P. 14</td>
</tr>
<tr>
<td>$R_{\nu}^{(B)}$</td>
<td>matrix of frequency integrals</td>
<td>Eq. 2.49</td>
</tr>
<tr>
<td>$s(t)$</td>
<td>output of pulse shaping filter at time $t$</td>
<td>Eq. 2.3</td>
</tr>
<tr>
<td>$S$</td>
<td>$R \times N$ encoder matrix</td>
<td>Eq. 2.11</td>
</tr>
<tr>
<td>$T$</td>
<td>symbol duration</td>
<td>Eq. 2.3</td>
</tr>
<tr>
<td>$u_{ij}$</td>
<td>output symbol if $\sigma(n) = \sigma_i$ and $x(n) = j$</td>
<td>P. 14</td>
</tr>
<tr>
<td>$u$</td>
<td>vector of output symbols $u_{ij}$</td>
<td>Eq. 2.15</td>
</tr>
<tr>
<td>$u(n)$</td>
<td>$R$-ary output symbol of mapper at time $n$</td>
<td>Eq. 2.2</td>
</tr>
<tr>
<td>$u'(n)$</td>
<td>output symbol of modified mapper at time $n$</td>
<td>P. 31</td>
</tr>
<tr>
<td>$U$</td>
<td>mapper output alphabet size</td>
<td>P. 11</td>
</tr>
<tr>
<td>$V_\nu$</td>
<td>$R \times R$ encoder matrices</td>
<td>Eq. 2.29</td>
</tr>
<tr>
<td>$W(f)$</td>
<td>power density spectrum of $s(t)$</td>
<td>Eq. 2.21</td>
</tr>
<tr>
<td>$W_c(f)$</td>
<td>continuous part of power density spectrum</td>
<td>Eq. 2.22</td>
</tr>
<tr>
<td>$W_d(f)$</td>
<td>discrete part of power density spectrum</td>
<td>Eq. 2.23</td>
</tr>
<tr>
<td>$W_\nu$</td>
<td>$R \times R$ encoder matrices</td>
<td>Eqs. 2.25, 2.29</td>
</tr>
<tr>
<td>$x(n)$</td>
<td>$M$-ary input symbol to encoder at time $n$</td>
<td>Eq. 2.1</td>
</tr>
<tr>
<td>$\hat{x}(n)$</td>
<td>estimated input symbol at time $n$</td>
<td>P. 12</td>
</tr>
<tr>
<td>$Y_i$</td>
<td>complex mapper output variable</td>
<td>Eq. 4.14</td>
</tr>
<tr>
<td>$Z_i$</td>
<td>complex mapper output variable</td>
<td>Eq. 4.18</td>
</tr>
<tr>
<td>$\alpha_\nu$</td>
<td>spectrum coefficients</td>
<td>Eq. 2.46</td>
</tr>
<tr>
<td>$\beta_\nu$</td>
<td>spectrum coefficients</td>
<td>Eq. 2.47</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>rolloff factor of spectral raised cosine</td>
<td>Eq. 2.56</td>
</tr>
<tr>
<td>$\delta(x)$</td>
<td>Dirac function</td>
<td>Eq. 2.23</td>
</tr>
<tr>
<td>$\eta$</td>
<td>bandwidth-efficiency</td>
<td>P. 103</td>
</tr>
<tr>
<td>$\eta_{\text{max}}$</td>
<td>bandwidth-efficiency ratio</td>
<td>Eq. 3.1</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>half number of samples per symbol</td>
<td>Eq. 2.42</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>eigenvalue</td>
<td>Th. 3</td>
</tr>
<tr>
<td>$\mu$</td>
<td>average output symbol</td>
<td>Eq. 2.14</td>
</tr>
<tr>
<td>$\pi$</td>
<td>stationary state probability distribution</td>
<td>Eq. 2.8</td>
</tr>
<tr>
<td>$\sigma(n)$</td>
<td>state of encoder at time $n$</td>
<td>Eq. 2.1</td>
</tr>
<tr>
<td>$\omega$</td>
<td>circular frequency</td>
<td>Eq. 5.18</td>
</tr>
</tbody>
</table>
Glossary

AMPM  amplitude and phase modulation
AWGN  additive white Gaussian noise
BPSK  binary phase shift keying
CPM   continuous phase modulation
C-TCM  correlative trellis coded modulation
FIR    finite impulse response
FSM    finite state machine
iid    independent and identically distributed
ISI    intersymbol interference
MLSE   maximum likelihood sequence estimator
MUX    multiplexer
psd    power spectral density
PRS    partial response signaling
PSK    phase shift keying
PSWF   prolate spheroidal wave function
QAM    quadrature amplitude modulation
QPSK   quaternary phase shift keying
TCM    trellis coded modulation
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematic diagram of the finite state machine encoder</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>Schematic diagram of the L-stage shift-register encoder/modulator</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>Schematic diagram of the receiver</td>
<td>12</td>
</tr>
<tr>
<td>2.4</td>
<td>Sampled pulse shape</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Inband power</td>
<td>24</td>
</tr>
<tr>
<td>4.1</td>
<td>Schematic diagram of the single delay-cell encoder/modulator</td>
<td>28</td>
</tr>
<tr>
<td>4.2</td>
<td>Trellis of the binary single delay-cell encoder</td>
<td>29</td>
</tr>
<tr>
<td>4.3</td>
<td>Optimum single delay-cell encoder/modulator for $\eta_2$</td>
<td>31</td>
</tr>
<tr>
<td>4.4</td>
<td>Optimum single delay-cell encoder/modulator for $\eta_3$</td>
<td>32</td>
</tr>
<tr>
<td>4.5</td>
<td>Optimum single delay-cell encoder/modulator</td>
<td>33</td>
</tr>
<tr>
<td>4.6</td>
<td>Filter representation of the optimum linear quadrature-amplitude (QAM) modulator with memory one</td>
<td>33</td>
</tr>
<tr>
<td>4.7</td>
<td>Schematic diagram of the two delay-cells encoder/modulator</td>
<td>34</td>
</tr>
<tr>
<td>4.8</td>
<td>Trellis of the binary-input two delay-cells encoder</td>
<td>34</td>
</tr>
<tr>
<td>4.9</td>
<td>Transmitter with Two Delay Cells Corresponding to $\eta_6$, Binary Input</td>
<td>36</td>
</tr>
<tr>
<td>4.10</td>
<td>Schematic diagram of the optimum encoder with a quaternary input alphabet</td>
<td>42</td>
</tr>
<tr>
<td>5.1</td>
<td>FIR-filter representation of the transmitter</td>
<td>44</td>
</tr>
<tr>
<td>5.2</td>
<td>FIR-filter corresponding to $\eta_6$</td>
<td>45</td>
</tr>
<tr>
<td>5.3</td>
<td>Schematic diagram of the partial response transmitter</td>
<td>47</td>
</tr>
<tr>
<td>6.1</td>
<td>Optimized pulse shape for an uncoded transmitter, $BT=0.9$</td>
<td>62</td>
</tr>
<tr>
<td>6.2</td>
<td>Optimized pulse shape for an uncoded transmitter, $BT=1.5$</td>
<td>63</td>
</tr>
<tr>
<td>6.3</td>
<td>Out-of-band power of an optimized, uncoded transmitter</td>
<td>64</td>
</tr>
<tr>
<td>6.4</td>
<td>Out-of-band power of uncoded, duobinary, and dicode transmitters</td>
<td>66</td>
</tr>
<tr>
<td>6.5</td>
<td>Out-of-band power of an encoder with two delay-cells, time-limited pulses</td>
<td>68</td>
</tr>
<tr>
<td>6.6</td>
<td>Achievable out-of-band power with 3 delay cells, time-limited pulses</td>
<td>69</td>
</tr>
<tr>
<td>6.7</td>
<td>Achievable out-of-band power with 4 delay cells, time-limited pulses</td>
<td>71</td>
</tr>
<tr>
<td>6.8</td>
<td>Optimized transmitters with 5 delay-cells, PSWF</td>
<td>72</td>
</tr>
<tr>
<td>6.9</td>
<td>Achievable out-of-band power with a time-limited pulse shaping filter</td>
<td>73</td>
</tr>
<tr>
<td>6.10</td>
<td>Out-of-band power of an uncoded transmitter with a spectral raised cosine (SRC) pulse shaping filter</td>
<td>77</td>
</tr>
<tr>
<td>6.11</td>
<td>Out-of-band power of an optimized transmitter with one delay-cell</td>
<td>78</td>
</tr>
<tr>
<td>6.12</td>
<td>Optimized transmitters with 2 delay-cells, SRC, rolloff = 0.3</td>
<td>80</td>
</tr>
<tr>
<td>6.13</td>
<td>Optimized transmitters with 3 delay-cells, SRC, rolloff = 0.3</td>
<td>81</td>
</tr>
<tr>
<td>6.14</td>
<td>Optimized transmitters with 4 delay-cells, SRC, rolloff = 0.3</td>
<td>82</td>
</tr>
</tbody>
</table>
6.15 Optimized transmitters with 5 delay-cells, SRC, rolloff = 0.3 83
6.16 Out-of-band power with a SRC pulse shaping filter, rolloff = 0.3 84

7.1 Optimum signal sets for uncoded M-ary modulation 87
7.2 Peak-power-efficient signal sets 89
7.3 Minimum distance of the transmitter with L = 2, T-time-limited pulse shaping filter 92
7.4 Bit error rate in AWGN with a T-time-limited pulse shaping filter 94
7.5 Bit error rate in AWGN with a spectral raised cosine pulse shaping filter 95
7.6 Transmitter with nonlinear precoder 96
7.7 Nonlinear receiver with symbol detector 97
7.8 Bit error rate in AWGN for different realizations of the single delay-cell system 98
7.9 Bit error rate in AWGN for the symbol detector, T-time-limited pulse shaping filter 100
7.10 Bit error rate in AWGN for the symbol detector, spectral raised cosine (SRC) pulse shaping filter 101

8.1 Bandwidth- and power-efficiency of several modulation schemes, bit error rate $P_b = 10^{-4}$, 99%-bandwidth 104

E.1 Inband power for different parameters d 119
E.2 Bandwidth-efficiency of the FIR encoder with two taps 120

F.1 Schematic diagram of the encoder with a pulse shaping filter bank 123
F.2 Pulse shapes for a single delay-cell encoder, $\gamma = 1$ 126
F.3 Achievable out-of-band power of the pulse shaping filter bank transmitter, $\gamma = 1$ 127
F.4 Bit error rate in AWGN of the pulse shaping filter bank transmitter, $\gamma = 1$ 128
F.5 Bandwidth- vs. power-efficiency, filter bank transmitter 129
F.6 Pulse shapes for an encoder with two delay-cells, $\gamma = 3$ 130
List of Tables

5.1 Mapper and FIR-filter coefficients for \( u_6 \) ........................................... 44
5.2 Coefficients of the optimum FIR filter ......................................................... 46
5.3 Size of the eigenvectors of the optimum FIR filter ........................................ 59

6.1 Bandwidth-Efficiencies at -20dB out-of-band power, PSWF, \( M=4 \) .... 70
6.2 Bandwidth-Efficiencies at -20dB out-of-band power, SRC, \( M=4 \) .... 79

7.1 Peak-Power-Efficient Signal Constellations .................................................. 91
7.2 Optimization bandwidths for a -20dB out-of-band power, PSWF pulses .. 93
7.3 Optimization bandwidths for a -20dB out-of-band power, SRC pulses ... 93
7.4 Performance of the precoder / symbol detector, \( P_b = 10^{-4} \) .......... 99

E.1 Spectral efficiency and minimum distance as function of \( d \) ................. 121

F.1 Bandwidth-efficiency and performance of the filter bank encoder, \( \gamma = 1 \) 125

139
Bibliography


[Bec88] H. Beckmann and J. Huber, “Mittleres Leistungsspektrum gedächtnisbehafte-


[Cal89] A.R. Calderbank and J.E. Mazo, “Baseband Line Codes Via Spectral Factor-


[Che82] P.R. Chevillat and G. Ungerboeck, “Optimum FIR Transmitter and Receiver
Filters for Data Transmission over Band-Limited Channels,” IEEE Trans. on

in the Presence of Intersymbol Interference,” IEEE Trans. on Information The-

Journal on Selected Areas in Communications, vol. SAC-2, pp. 632–647,
September 1984.

[Gib65] R.A. Gibby and J.W. Smith, “Some Extensions of Nyquist’s Telegraph Trans-
September 1965.


[Har72] H. Harashima and H. Miyakawa, “Matched-Transmission Technique for Chan-


Curriculum Vitae

I was born in Lucerne, Switzerland on 26 April 1961. I attended the primary school and the Kantonsschule in Lucerne, where I passed the final examination in the field of mathematics and sciences (Matura Typ C) in 1980. In the same year, I started my studies in Electrical Engineering at the Swiss Federal Institute of Technology (ETH), Zürich, and received the Diploma in 1986. From 1986 to 1989 I attended the postgraduate program (Nachdiplomstudium) in communications at the ETH, consisting of several courses and a thesis which was supervised by Prof. Dr. F. Eggimann. In 1986 I joined the Communications Group of the ABB Research Center; the group was transferred to Ascom Tech in 1991. At ABB/Ascom I designed, analyzed and implemented algorithms for a communications system over non-Gaussian, time-varying channels, such as the communication over power lines. In summer 1988 I started the research project "Bandwidth-Efficient Modulation Schemes" that culminated in this Ph.D. thesis.