Fund. of Digital Communications
Chapter 5: Demodulation, Detection

Klaus Witrisal
witrisal@tugraz.at

Signal Processing and Speech Communication Laboratory
www.spsc.tugraz.at
Graz University of Technology
January 23, 2014

Outline

■ 5-1 Correlation-Type Demodulator and Matched Filter
■ 5-2 Optimal Detection
■ 5-3 Error Probability
■ 5-4 Equalization (and Channel Distortions)

Reference: (Figures taken from these books.)
Correlation-type demodulator

- **Channel:** Additive white Gaussian noise (AWGN) is added
  \[ r(t) = s_m(t) + n(t) \]

- Transmitted signal \( \{s_m(t)\}, m = 1, 2, ..., M \) is represented by \( N \) basis functions \( \{\psi_k(t)\}, k = 1, 2, ..., N \)

- Received signal \( r(t) \) is projected onto these basis functions \( \{\psi_k(t)\} \)
  \[
  \int_0^T r(t) \psi_k(t) dt = \int_0^T [s_m(t) + n(t)] \psi_k(t) dt
  \]
  \[ r_k = s_{mk} + n_k, \quad k = 1, 2, ..., N \]

- Output vector in signal space: \( r = s_m + n \)

---

**Correlation demod. (cont’d)**

![Diagram of correlation-type demodulator](image)
Correlation demod. (cont’d)

- **Received signal**

\[
    r(t) = \sum_{k=1}^{N} s_{mk} \psi_k(t) + \sum_{k=1}^{N} n_k \psi_k(t) + n'(t)
\]

\[
    = \sum_{k=1}^{N} r_k \psi_k(t) + n'(t)
\]

- **Correlator outputs** \( r = [r_1, r_2, ... r_N]^T \) are **sufficient statistik** for the decision
  
  ♦ i.e.: there is **no additional info** in \( n'(t) \)
  
  ♦ \( n'(t) \) is part of \( n(t) \) that is **not representable** by \( \{ \psi_k(t) \} \)

- **Interpretation of** \( r \): **noise cloud** in signal space

[Proakis 2002]
Correlation demod. (cont’d)

- Characterization of the noise component
  - Mean values
    \[ E\{n_k\} = \ldots = \int_0^T E\{n(t)\} \psi_k(t) \, dt = 0 \]
  - Correlations (= covariances, since means = 0)
    \[ E\{n_k n_m\} = \ldots = \frac{N_0}{2} \delta[m - k] \]
  - i.e.: \( N \) noise components \( \{n_k\} \) are zero-mean, uncorrelated, Gaussian random variables
  - their variances \( \sigma_n^2 = N_0/2 \) (because \( ||\psi_k(t)||^2 = 1 \))

- Statistics of the correlator outputs \( \{r_k\} \)
  - Mean values
    \[ E\{r_k\} = E\{s_{mk} + n_k\} = s_{mk} \]
  - Variances (identical)
    \[ \sigma_r^2 = \sigma_n^2 = N_0/2 \]
  - conditional PDF; likelihood function
    \[ f(r_k|s_{mk}) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r_k-s_{mk})^2}{N_0}}, \quad k = 1, 2, \ldots, N \]
    \[ f(r|m) = \prod_{k=1}^N f(r_k|s_{mk}) \quad m = 1, 2, \ldots, M \]
Matched Filter Demodulator

- Uses $N$ Filters (instead of the correlators) with impulse responses

$$h_k(t) = \psi_k(T - t), \quad 0 \leq t \leq T$$

- Filter outputs

$$y_k(t) = \int_0^t r(\lambda) h_k(t - \lambda) d\lambda = ...$$

- Sampling at $t = T$

$$y_k(T) = \int_0^T r(\lambda) \psi_k(\lambda) d\lambda = r_k, \quad k = 1, 2, ..., N$$

identical to correlator outputs!

Matched Filter Demodulator (cont’d)
Matched Filter Demodulator (cont’d)

- Property of the matched filter: **maximization of the output SNR** at sampling time \( t = T \)
- Derivation: **arbitrary** impulse response \( h(t), 0 \leq t \leq T \)
  \[
y(t) = \ldots
y(T) = \ldots
\]
- Signal-to-noise ratio (SNR)
  \[
  \left( \frac{S}{N} \right)_0 = \frac{y_s^2(T)}{\mathbb{E}\{y_n^2(T)\}}
\]
- Maximum SNR is found with Cauchy-Schwarz inequ.
  \[
  h(t) = C_s(T - t) \quad \text{und} \quad \left( \frac{S}{N} \right)_0 = \frac{2E_s}{N_0}
\]

Derivation of the Matched Filter Demodulator (cont’d)

- SNR of the sampled filter output
  \[
  \left( \frac{S}{N} \right)_0 = \frac{y_s^2(T)}{\mathbb{E}\{y_n^2(T)\}} = \ldots = \frac{\left[ \int_0^T h(\lambda)s(T - \lambda) \, d\lambda \right]^2}{\frac{N_0}{2} \int_0^T h^2(T - t) \, dt}
\]
- denominator is energy of the filter IR \( \rightarrow \) hold constant
- maximization of the numerator using Cauchy-Schwarz:
  \[
  \left[ \int_{-\infty}^{\infty} g_1(t)g_2(t) \, dt \right]^2 \leq \int_{-\infty}^{\infty} g_1^2(t) \, dt \int_{-\infty}^{\infty} g_2^2(t) \, dt
  \]
  **Equality** (i.e. maximum) holds for \( g_1(t) = Cg_2(t) \)

This yields \( h(t) = C_s(T - t) \) to obtain max. SNR
5-2 Optimum Detection
(= Entscheidung)

- Signal model: received signal vector with i.i.d. noise
  \[ r = s_m + n \]
  
  - \( s_m \) ... points in \( N \)-dimensional signal space
  - \( n \) ... random vector with i.i.d. components
    \( \sim \mathcal{N}(0, N_0/2) \)

- Find: decision minimizing the error probability, based on the observed vector \( r \)

- Solution: decision rule based on a posteriori probabilities:
  \[ P(\text{signal } s_m \text{ was transmitted } | r) \quad m = 1, 2, ..., M \]
  
  - choose \( s_m \) that maximizes \( P(s_m|r) \) for \( m = 1, 2, ..., M \)

Optimum Detection (cont’d)

Maximizing the posterior probability:

- Maximum a posteriori probability (MAP) criterion
- finding the posterior probability using Bayes’ rule:
  \[ P(s_m|r) = \frac{f(r|s_m)P(s_m)}{f(r)} \]

  - \( f(r|s_m) \) ... conditional PDF of \( r \) given \( s_m \) was transmitted (likelihood function)
  - \( P(s_m) \) ... a priori probability
  - PDF \( f(r) = \sum_{m=1}^{M} f(r|s_m)P(s_m) \) is independent of \( s_m \), hence irrelevant

- Metrik for MAP criterion: \( PM(r, s_m) = f(r|s_m)P(s_m) \)
Optimum Detection (cont’d)

Maximum Likelihood (ML) Criterion

- equivalent to MAP; for equal a priori probabilities
  \[ P(s_m) = \frac{1}{M}, \quad \forall m = 1, 2, \ldots, M \]

- decision means finding the largest (the maximum) likelihood function \( f(r|s_m) \) (at the observed point \( r \) by choosing \( s_m \))

ML criterion for AWGN channel (logarithm of \( f(r|s_m) \)):

\[
\ln[f(r|s_m)] = -\frac{N}{2} \ln(\pi N_0) - \frac{1}{N_0} \sum_{k=1}^{N} (r_k - s_{mk})^2
\]

note: first term is irrelevant; \( 1/N_0 \) too:

- ML criterion means minimizing the distance!

\[
D(r, s_m) = \sum_{k=1}^{N} (r_k - s_{mk})^2 = \|r - s_m\|^2
\]

- \( D(r, s_m) \) ... distance metric

- Equivalently: maximizing the correlation metric:
  \[
  C(r, s_m) = 2r^T s_m - \|s_m\|^2
  \]
Optimum Detection – Example

Given: binary, antipodal PAM

- \( s_1(t) = -s_2(t) = g_T(t) \)
- \( s_1 = -s_2 = \sqrt{E_b} \) (\( E_b \) ... energy per bit)
- \( P(s_1) = p, P(s_2) = 1 - p \)

Find: metrics for MAP detection in AWGN

- Result: threshold detection can be performed!
  \[
  r \geq \frac{s_1}{s_2} \gamma
  \]

Likelihood functions for binary antipodal PAM:

[Proakis 2002]
5-3 Error Probability

- Assumption: binary antipodal signals with 
  \( P(s_1) = P(s_2) = 0.5 \) and \( s_1 = -s_2 = \sqrt{E_b} \)

- therefore we can use a threshold detector with threshold \( \gamma = 0 \)

- error probability for \( s_1(t) \)
  \[
  P(e|s_1) = \int_{-\infty}^{0} f(r|s_1) \, dr = ... = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)
  \]

  \( Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-x^2/2} \, dx \) ... integral over Gaussian function

Error Probability (cont’d)

- average error probability per bit:
  \[
  P_b = P(s_1)P(e|s_1) + P(s_2)P(e|s_2)
  = \frac{1}{2} P(e|s_1) + \frac{1}{2} P(e|s_2)
  = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)
  \]

- using the distance between signal constellations:
  \[
  P_b = Q\left(\sqrt{\frac{d_{12}^2}{2N_0}}\right)
  \]

  holds for any binary signals with \( P(s_1) = P(s_2) = 0.5 \)

  e.g. bin. antipodal: \( d_{12} = 2\sqrt{E_b} \); orthogonal: \( d_{12} = \sqrt{2E_b} \)
Error Probability (cont’d)

- $P_b$ for antipodal and orthogonal binary signals

- Approximate symbol error rate (SER) for higher-order modulation schemes

$$P_M \approx N_e Q \left( \sqrt{\frac{d_{\text{min}}^2}{2N_0}} \right)$$

where

- $N_e$ ... average number of nearest neighbors (in signal space)
- $d_{\text{min}}$ ... (min.) distance to those neighboring signal vectors
Error Probability (cont’d)

- for \(M\)-ary PAM

\[
N_e = 2 \frac{M - 1}{M}
\]

\[
d_{\min} = 2 \sqrt{\frac{3 \log_2 M}{M^2 - 1}} \sqrt{E_b}
\]

\[
P_M \approx 2 \frac{M - 1}{M} \times Q\left(\sqrt{\frac{6 \log_2 M}{M^2 - 1}} \sqrt{\frac{E_b}{N_0}}\right)
\]

- BER for Gray coding

\[
P_b \approx P_M / k
\]

Error Probability (cont’d)

- for \(M\)-ary PSK

\[
N_e = 2
\]

\[
d_{\min} = 2 \sqrt{E_s} \sin \frac{\pi}{M}
\]

\[
= 2 \sqrt{kE_b} \sin \frac{\pi}{M}
\]

\[
P_M \approx 2Q\left(\sqrt{\frac{2kE_b}{N_0}} \sin \frac{\pi}{M}\right)
\]

- BER for Gray coding

\[
P_b \approx P_M / k
\]
Error Probability (cont’d)

- for $M$-ary orthogonal sign.
  
  $N_e = M - 1$

  $d_{\text{min}} = \sqrt{2E_s} = \sqrt{2kE_b}$

  $P_M \leq (M - 1)Q\left(\sqrt{\frac{kE_b}{N_0}}\right)$

- Upper bound:
  
  $P_M < e^{-k(\frac{E_b}{N_0} - 2 \ln 2)/2}$

  i.e. $P_M \rightarrow 0$ for $k \rightarrow \infty$ if

  $\frac{E_b}{N_0} > 2 \ln 2 = 1.4 \text{ dB}$

- in fact: $\frac{E_b}{N_0} > -1.6 \text{ dB}$ sufficient: Shannon limit

Comparison of modulation schemes:

- shows $R/W$: transmission rate normalized with bandwidth

- as a function of SNR per bit: $\frac{E_b}{N_0}$

- to reach SER of $10^{-5}$

- Shannon limit is approached for orthogonal signals if $k \rightarrow \infty$

Figure 7.66 Comparison of several modulation methods at $10^{-5}$ symbol error probability.

[Proakis 2002]
Channel Capacity

[Proakis02, Sec. 9.3]

- **Channel capacity** for AWGN channel (Shannon)

\[
C = W \log_2 \left( 1 + \frac{P}{N_0 W} \right)
\]

- $C$ ... channel capacity [bit/s]
- $W$ ... bandwidth [Hz]
- $P$ ... signal power [W]
- $N_0$ ... noise PSD [W/Hz]

- **increase power** $P \rightarrow$ (slow) increase of capacity
- **no limit but lots of power required**

- **increase bandwidth** $W$
- at fixed power $P$

\[
C = W \log_2 \left( 1 + \frac{P}{N_0 W} \right)
\]

- **consider limit** $W \to \infty$

\[
\lim_{W \to \infty} C = \frac{P}{N_0} \log_2 e = 1.44 \frac{P}{N_0}
\]

- **there is a hard limit**
- **need to increase $P$ too**
Channel Capacity (cont’d)

- In practice we must have \( R < C \)
  \[
  R < W \log_2 \left( 1 + \frac{P}{N_0 W} \right)
  \]
  define
- \( r = \frac{R}{W} \) “spectral rate”
- \( E_b = \frac{P}{R} \) energy per bit
  we obtain
  \[
  r < \log_2 \left( 1 + r \frac{E_b}{N_0} \right)
  \]

5-4 Equalization

[Proakis 2002, Section 8.6]
- Nyquist criterion must be fulfilled for ISI-free transmission
- Cascade of linear filters
  \[
  G_T(f) C(f) G_R(f) = X_{rc}(f)
  \]
- \( X_{rc}(f) \) ... Fourier transform of raised-cosine pulse (ISI free!)
Channel Distortion (cont’d)

Design considerations; assume channel $C(f)$ known

- Cascade $G_T(f)C(f)G_R(f) = X_{rc}(f)$ fulfills Nyquist
- Noise $n(t)$ is AWGN (flat spectrum)

→ Receiver filter $G_R(f)$ should be matched to cascade $G_T(f)C(f)$

$$G_T(f) = \sqrt{\frac{X_{rc}(f)}{C(f)}} e^{-j2\pi f t_0} \quad G_R(f) = \sqrt{X_{rc}(f)} e^{-j2\pi f t_r}$$

Equalization

- Channel is unknown or changes with time
  - e.g. telephone channel – routing changes
  - e.g. mobile radio – multipath propagation

→ design $G_T(f)$ and $G_R(f)$ as for AWGN channel:
  - root raised cosine frequency response

$$G_T(f) = \sqrt{X_{rc}(f)} e^{-j2\pi f t_0} \quad |G_R(f)| \cdot |G_T(f)| = X_{rc}(f)$$

- Output of receiver filter

$$r(t) = y(t) + n(t) = \sum_{k=-\infty}^{\infty} a[k]x(t - kT) + n(t)$$

$a[k] \in \{a_m\}_{m=1}^{M}$ ... PAM (or QAM) symbols

$x(t) = g_T(t) * c(t) * g_R(t)$ ... cascaded system response
Equalization (cont’d)

Equivalent output with **periodic sampling**:

\[
y[k] = y(kT) = \sum_{l=-L}^{L} x_l a[k - l] \quad \text{(noise-free)}
\]

\[
x_k = x(kT) \quad \text{... equiv. FIR discrete-time channel filter}
\]

![Diagram showing the equalization process](Diagram.png)

**Maximum likelihood sequence detection (MLSD)**

- **optimum** detector (for AWGN: \( r[k] = y[k] + n[k] \)):
  - choose information sequence \( a := \{a[k]\}_{k=1}^{K} \)
  - based on the received sequence \( r := \{r[k]\}_{k=1}^{K+L-1} \)
  - each \( a[k] \in \{a_m\}_{m=1}^{M} \)
  - hence there are \( M^K \) possible sequences \( \{a^{(m)}\}_{m=1}^{M^L} \)
  - joint likelihood function: (w/ \( y^{(m)}[k] = a^{(m)}[k] \ast x_k \))

\[
f(r|a^{(m)}) = \prod_{k=1}^{K+L-1} f(r[k]|a^{(m)}[k])
\]

\[
= \prod_{k=1}^{K+L-1} \frac{1}{\sqrt{\pi N_0}} e^{(r[k]-y^{(m)}[k])^2/N_0}
\]
Maximum likelihood sequence detection (MLSD) (cont’d)

joint likelihood function (cont’d):

\[ f(r|a^{(m)}) = \frac{1}{(\pi N_0)^{K+L/2}} \exp \left[ -\frac{1}{N_0} \sum_{k=1}^{K+L-1} (r[k] - y^{(m)}[k])^2 \right] \]

equivalent distance metric (squared distance)

\[ D(r, a^{(m)}) = \sum_{k=1}^{K+L-1} (r[k] - y^{(m)}[k])^2; \quad y^{(m)}[k] = \sum_{l=-L_1}^{L_2} x_l a^{(m)}[k - l] \]

MLSD: channel \( \{ x_k \} \) is known; find for \( m = 1, 2, \ldots, M^K \)

\[ \hat{m} = \arg \max_{m} f(r|a^{(m)}) = \arg \min_{m} D(r, a^{(m)}) \]

Viterbi (MLSD)

can be done sequentially using the Viterbi algorithm

- distances are computed recursively:

\[ D(r_{(1:k)}, a^{(m)}_{(1:k)}) = D(r_{(1:k-1)}, a^{(m)}_{(1:k-1)}) + (r[k] - \sum_{l=-L_1}^{L_2} x_l a^{(m)}[k - l])^2 \]

- Complexity:
  - symbols \( a[k] \) are \( M \)-ary PAM symbols
  - FIR channel has memory \( L - 1 = L_1 + L_2 \)
    → compute \( M^L \) branch metrics at each stage \( k \), yielding \( M^{L-1} \) surviving sequences (trellis diagram)
    → complexity: \( (K + L - 1)M^L \) (scales linearly with \( K \))
  - while \( M^L \ll M^K \), still \( M \) and \( L \) must be small!
Viterbi example (MLSD)

- Channel model: $L = 2$ (1 memory); $M = 2$ (binary)
  \[ r[k] = x_0 a[k] + x_1 a[k-1] + n[k]; \quad a[k] \in \{-1, +1\} \]

- state variable (memory) $a[k-1] \in \{-1, +1\}$

- information sequence: \{\(a[k]\}\}_{k=1}^{K} = a$; each $a[k] \in \{\pm 1\}$; i.e. $2^K$ possible sequences denoted $a^{(m)}$

- observation sequence: \{\(r[k]\}\}_{k=1}^{K+1} = r$

- likelihood functions (for $m = 1, ..., M^K$)
  \[
  f(r|a^{(m)}) \propto \exp\left[ -\frac{1}{N_0} \sum_{k=1}^{K+1} (r[k] - x_0 a^{(m)}[k] - x_1 a^{(m)}[k-1])^2 \right]
  \]

Viterbi example (MLSD) (cont’d)

- MLSD maximizes likelihood fct. by testing any $a^{(m)}$, which is equivalent to minimizing:
  \[
  D(r, a^{(m)}) = \sum_{k=1}^{K+L-1} (r[k] - x_0 a^{(m)}[k] - x_1 a^{(m)}[k-1])^2
  \]

- at stage $k$
  \[
  D(r_{1:k}, a^{(m)}_{1:k}) = \sum_{l=1}^{k-1} (r[l] - x_0 a^{(m)}[l] - x_1 a^{(m)}[l-1])^2 \\
  + (r_k - x_0 a^{(m)}[k] - x_1 a^{(m)}[k-1])^2 \\
  = D(r_{1:k-1}, a^{(m)}_{1:k-1}) + \text{branch metric}
  \]

- Viterbi requires:
  - $M^L = 2^2 = 4$ branch metrics at each state
  - $M^{L-1} = 2^1 = 2$ surviving paths at each state
  \[
  \rightarrow (K + 1)4 \ll M^K = 2^K \text{ metric computations}
  \]
Linear Equalizers

- Use linear filters to **compensate for channel distortion**
- filter has adjustable parameters
  - **preset equalizers**: measure channel; set parameters
  - **adaptive equalizers**: update parameters during data transmission

![Block diagram of a system with an equalizer.](image)

Zero-Forcing Linear Equalizer

How to select equalizer $G_E(f)$?

- $G_R(f)$ is matched to $G_T(f)$ to fulfill zero-ISI
- Hence $G_E(f)$ must compensate channel distortion

$$G_E(f) = \frac{1}{C(f)} = \frac{1}{|C(f)|} e^{-j\Theta_c(f)}$$

→ inverse channel filter

- **forces ISI to zero** (at sampling times)
  detector input: $r[k] = y[k] + n[k]$

- noise variance (for AWGN); **noise gain**

$$\sigma_n^2 = \int_{-\infty}^{\infty} S_n(f)|G_R(f)|^2|G_E(f)|^2 df = \frac{N_0}{2} \int_{-W}^{W} \frac{|X_{re}(f)|}{|C(f)|^2} df$$
Linear transversal filter

Practical implementation of the linear equalizer

- use FIR filter with adjustable taps \( \{c_n\} \)
- delays \( \tau \): \( \tau = T \) ... symbol spaced; aliasing, if \( 1/T < 2W \)
- \( \tau < T \) ... fractionally-spaced equalizer (often: \( \tau = T/2 \))

![Diagram of linear transversal filter]

Zero-Forcing Equalizer (cont’d)

How to determine taps \( \{c_n\} \) for the zero-forcing equalizer?

- Equalized output pulse

\[
q(t) = \sum_{n=-N}^{N} c_n x(t - n\tau)
\]

for \( x(t) = g_T(t) * c(t) * g_R(t) \); number of taps \( 2N + 1 \geq L \)

- enforce zero-ISI at sampled output:

\[
q(kT) = \sum_{n=-N}^{N} c_n x(kT - n\tau) = \begin{cases} 
1, & k = 0 \\
0, & k = \pm 1, \pm 2, \ldots, \pm N 
\end{cases}
\]

- system of \( 2N + 1 \) linear equations: \( q = Xc \)

for coefficients \( \{c_n\} \to \) approx. solution \( c = X^\dagger q \)
Minimum mean-square-error equalizer

- ZF equalizer **ignores noise!** → noise gain
- solution:
  - relax zero-ISI condition
  - consider residual ISI plus noise at equalizer output
  → use **minimum mean-square-error (MMSE)** criterion
- sampled equalizer output
  \[ z(kT) = \sum_{n=-N}^{N} c_n r(kT - n\tau) = c^T r[k] \]
- mean-square-error (MSE)
  \[ \text{MSE} = E\{(z(kT) - a[k])^2\} = E\{(c^T r[k] - a[k])^2\} \]

MMSE equalizer (cont’d)

Find the equalizer \( c \) that minimizes the MSE

\[
E\{[c^T r[k] - a[k]]^2\} = E\{[c^T r[k] - a[k]]r^T[k]c - a[k]]\}
= E\{c^T r[k]r^T[k]c\} - 2E\{c^T r[k]a[k]\} + E\{a[k]^2\}
= c^T K_r c - 2c^T k_{ar} + E\{a[k]^2\}
\]

- with \( K_r = E\{r[k]r^T[k]\} \) and \( k_{ar} = E\{r[k]a[k]\} \)
- minimization: find derivative with \( c \); set to zero
  \[
  K_r c - k_{ar} = 0 \quad \Rightarrow \quad c = K_r^{-1} k_{ar}
  \]
- correlation matrix \( K_r \) and correlation sequence \( k_{ar} \)
  **must be estimated** → training sequences
Adaptive equalizers

- perform the MMSE equalization recursively
- search along gradient for MSE (e.g. LMS algorithm)
- discussed in detail in VL: Adaptive Systems