Disclaimer

This summary is part of the lecture “ETH Communication Systems” (227-0121-00) by Prof. Dr. Armin Wittneben (FS19). It is based on the lecture.

Please report errors to huettern@student.ethz.ch such that others can benefit as well.

The upstream repository can be found at https://github.com/noah95/formulasheets

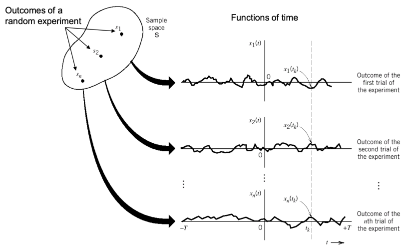
Contents

ETH Communication Systems 2019

Noah Huetter

February 5, 2020

1 Random Processes



A random process $X(t)$:

- is a sample space composed of (real valued) time functions: $\{x_1(t), x_2(t), \ldots, x_n(t)\}$
- observed at a fixed t_k is a random variable $X(t_k) = \{x_1(t_k), x_2(t_k), \ldots, x_n(t_k)\}$
- The time function $x_s(t)$ is a **realization** (sample function)
- $x_s(t_k)$ observed at t_k is a real number
- A stochastic process consists of infinitely many random variables, one for each t_k, with the CDF $F_{\{X(t_k)\}}(x) = P(X(t_k) \leq x)$

1.1 Stationary processes

A process is **Strict Sense Stationary (SSS)** if:

- $X(t)$ and $X(t+\tau)$ have same statistics $\forall \tau$
- The joint distribution function of a set of r.v. observed at times $t_1, \ldots, t_n$ is invariant to a time-shift.

$$\forall n, \tau, t_1, \ldots, t_n : \\ F_{\{X(t_1+\tau), X(t_2+\tau), \ldots, X(t_n+\tau)\}}(x_1, x_2, \ldots, x_n) = \\ F_{\{X(t_1), X(t_2), \ldots, X(t_n)\}}(x_1, x_2, \ldots, x_n)$$

Properties:

$$\forall t_k : \mu_X(t_k) = \mu_X \\ \forall t_1, t_2 : R_X(t_1, t_2) = R_X(t_2 - t_1) = R_X(\tau) \\ C_X(t_1, t_2) = \mathrm{E}\left[(X(t_1) - \mu_X)(X(t_2) - \mu_X)\right] \\ = R_X(t_2 - t_1) - \mu_X^2$$

A process is **Wide Sense Stationary (WSS)** if a r.p. has a *constant* mean and the autocorrelation depends only on the *time difference.*

$$\forall t : \mu_X(t) = \mu_X \\ \forall t_1, t_2 : R_X(t_1, t_2) = R_X(t_2 - t_1) = R_X(\tau)$$

Strict sense stationary $\Longrightarrow$ wide sense stationary.

1.2 Mean and correlation

Defined as expectation of r.v. $X(t_k)$ by observing process at time t_k.

$$\mu_X(t_k) = \mathrm{E}\left[X(t_k)\right] = \int_{-\infty}^{\infty} x f_{\{X(t_k)\}}(x) \, \mathrm{d}x$$

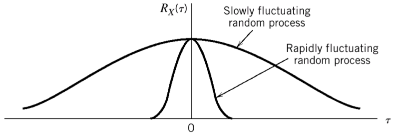
Autocorrelation function R_X and autovariance function C_X of a random process:

$$R_X(t_1, t_2) = \mathrm{E}\left[X(t_1)X(t_2)\right] \triangleq \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X_1, X_2}(x_1, x_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \\ R_{XY}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y f_{X,Y}(x, y) dx dy \\ C_X(t_1, t_2) = R_X(t_1, t_2) - m_X^2 \\ = R_X(t_2 - t_1) - m_X^2 (\text{for WSS})$$

- The mean and autocorrelation function determine the autocovariance function
- The mean and autocorrelation function only describe the first two moments of the process

Properties of the autocorrelation function:

$$\mathrm{E}\left[X^2(t)\right] = R_X(0) \qquad R_X(\tau) = R_X(-\tau) \\ |R_X(\tau)| \leq R_X(0)$$



The Cross-correlation function $R_{XY}(t, u)$ of two random processes:

$$R_{XY}(t, u) = \mathrm{E}\left[X(t)Y(u)\right] = \\ \int_{-\infty}^{\infty} x y \cdot f_{X,Y}(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

- Stationariy means $R_{XY}(t, u) = R_{XY}(\tau)$ for $\tau = t - u$
- Not generally an even function of t
- Not necessarily a maximum at $\tau = 0$
- Symmetry: $R_{XY}(\tau) = R_{YX}(-\tau)$

1.3 Ergodicity

Definition: A random process is *ergodic* in the mean if

- Time average approaches ensemble averages for increasing T
- The variance of the time average approaches zero for incr. T

$$\lim_{T \to \infty} \mu_X(T) = \mu_X \qquad \lim_{T \to \infty} \mathrm{Var}\left[\mu_X(T)\right] = 0$$

Or in other words: The same behavior averaged over time as averaged over the space of all the system's states.

1.4 Filtered processes

Stationary random process $X(t)$ is input to a linear timeinvariant (LTI) filter with impulse response $h(t)$.

$$Y(t) = \int_{-\infty}^{\infty} h(\tau_1) X(t - \tau_1) \, \mathrm{d}\tau_1 \\ S_Y(f) = |H(f)|^2 S_X(f)$$

Find mean and autocorrelation of $Y(t)$:

$$\mu_X = \mathrm{E}\left[X(t)\right] \quad R_X(\tau) = \mathrm{E}\left[X(t)X(t-\tau)\right]$$

$$\mu_Y = \mathrm{E}\left[Y(t)\right] = \mathrm{E}\left[\int_{-\infty}^{\infty} h(\tau_1) X(t - \tau_1) \, \mathrm{d}\tau_1\right]$$

Can interchange expectation and integration if stable $\int_{-\infty}^{\infty} |h(t)| \, \mathrm{d}t < \infty$ and finite mean $\mu_X < \infty$

$$\mu_Y = \int_{-\infty}^{\infty} h(\tau_1) \mathrm{E}\left[X(t - \tau_1)\right] \mathrm{d}\tau_1 = \mu_X \int_{-\infty}^{\infty} h(\tau_1) \, \mathrm{d}\tau_1$$

Autocorrelation:

$$R_Y(t, u) = \mathrm{E}\left[Y(t)Y(u)\right] = \\ \mathrm{E}\left[\int_{-\infty}^{\infty} h(\tau_1) X(t - \tau_1) \, \mathrm{d}\tau_1 \int_{-\infty}^{\infty} h(\tau_2) X(u - \tau_2) \, \mathrm{d}\tau_2\right]$$

Additional condition for interchange is finite mean-square value: $R_X(0) = \mathrm{E}\left[X^2(t)\right] < \infty$

$$R_Y(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\tau_1) h(\tau_2) R_X(\tau - \tau_1 + \tau_2) \, \mathrm{d}\tau_1 \, \mathrm{d}\tau_2$$

(WS) stationary input process $X(t)$ to a stable LTI filter $\Longrightarrow$ (WS) stationary output process $Y(t)$.

1.5 Power spectral density

$$S_X(f) = \mathscr{F}[R_X(\tau)](f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f \tau} \, \mathrm{d}\tau$$

- $S_X(0) = \int_{-\infty}^{\infty} R_X(\tau) d\tau$
- $E[X^2(t)] = R_X(0) = \int_{-\infty}^{\infty} S_X(f) df$
- $S_X(f) \geq 0 \; \forall f$
- $S_X(f) = S_X(-f) \; \forall f$, iff $X(t) \in \mathbb{R}$

1.6 Gaussian process

Consider the r.v. $Y = \int_0^T g(t) X(t) \, \mathrm{d}t$ where $g(t)$ is in an arbitraty function. If Y is gaussian distributed, then the process $X(t)$ is a *Gaussian process*

- A filtered Gaussian process remains a Gaussian process
- If $X(t)$ is a GP, the arbitrary set of r.v. $\vec{X} = [X(t_1), \ldots, X(t_n)]^T$ is jointly gaussian distributed for any n
- The joint cdf is of these r.v. is completely determined by the **means** $\mu_X(t_i) = \mathrm{E}\left[X(t_i)\right]$ and **covariances** $C_X(t_k, t_i) = \mathrm{E}\left[(X(t_k) - \mu_X(t_k))(X(t_i) - \mu_X(t_i))\right]$

Multivariative Gauss distribution:

$$f(x) = \frac{\exp\left(-\frac{1}{2}(\vec{x} - \vec{m}_x)^T \underline{\Sigma}^{-1} (\vec{x} - \vec{m}_x)\right)}{(2\pi)^{\frac{n}{2}} \det\left(\underline{\Sigma}\right)^{\frac{1}{2}}}$$

$$\underline{\Sigma} := \begin{bmatrix} \mathrm{Cov}\left(X_1, X_1\right) & \ldots & \mathrm{Cov}\left(X_1, X_n\right) \\ \vdots & \vdots & \vdots \\ \mathrm{Cov}\left(X_n, X_1\right) & \ldots & \mathrm{Cov}\left(X_n, X_n\right) \end{bmatrix}$$

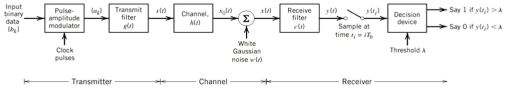
1.7 Noise

White noise is defined by its autocorrelation.

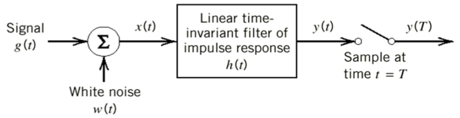
$$\boxed{R_W(\tau) = \frac{N_0}{2}\delta(t) \qquad S_W(f) = \frac{N_0}{2}}$$

2 Baseband Pulse Transmission

Digital Baseband Pulse Transmission System: Based on the sample $y(t_i)$ the receiver generates an estimate $\hat{a}_i$ of the amplitude a_i of the transmitted pulse $g(t-iT_b)$.



2.1 Matched Filter



$$y(t) = g_0(t)+n(t) = h(t)*g(t)+h(t)*w(t)$$

Maximize *pulse signal-to-noise ratio* η at sampling time $t = T$:

$$\eta = \frac{|g_0(T)|^2}{\mathrm{E}\left[n^2(t)\right]} = \frac{\left|\int\limits_{-\infty}^{\infty} H(f)G(f)e^{j2\pi fT}\,\mathrm{d}f\right|^2}{\frac{N_0}{2}\int\limits_{-\infty}^{\infty}|H(f)|^2\,\mathrm{d}f}$$

Using Schwarz's inequality:

$$\boxed{\eta \leq \frac{2}{N_0}\int\limits_{-\infty}^{\infty}|G(f)|^2\,\mathrm{d}f}$$

The quality sign (optimum) holds if $a(x) \propto b^*(x)$, i.e.

$$\boxed{H_{\text{opt}}(f) = kG^*(f)e^{-j2\pi fT} \Rightarrow h_{\text{opt}}(t) = kg(T-t)}$$

The impulse response of the optimum filter, except for the scaling factor k, is a time-reversed and delayed version of the input signal $g(t)$.

The pulse SNR of a machted filter depends only on the ratio of the signal energy E to the PSD of the white noise at the input filter.

$$\eta_{\max} = \frac{2}{N_0}\int\limits_{-\infty}^{\infty}|G(f)|^2\,\mathrm{d}f = \frac{2E}{N_0}$$

$$E = \int\limits_{-\infty}^{\infty}|g(t)|^2\,\mathrm{d}t = \int\limits_{-\infty}^{\infty}|G(f)|^2\,\mathrm{d}f$$

2.2 Error Rate

Discussed for a binary bipolar non-return-to-zero (NRZ) signal with amplitude A, bit duration T_b.

$$x(t) = \begin{cases} +A+w(t), & \text{Symbol 1 transmitted} \\ -A+w(t), & \text{Symbol 0 transmitted} \end{cases}$$

$$p_{10} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{A+\lambda}{\sqrt{N_0/T_b}}\right) = \mathrm{Q}\left(\sqrt{2}\frac{A+\lambda}{\sqrt{N_0/T_b}}\right)$$
$$= \mathbf{P}\left(y > \lambda \mid \text{symbol 0 was sent}\right)$$

The avg. prob. of symbol error P_e:

$$\boxed{P_e = \frac{p_0}{2}\,\mathrm{erfc}\left(\frac{A+\lambda}{\sqrt{N_0/T_b}}\right)+\frac{p_1}{2}\,\mathrm{erfc}\left(\frac{A-\lambda}{\sqrt{N_0/T_b}}\right)}$$

The error function:

$$\mathbf{P}\left(n > a\right) \equiv \mathrm{Q}\left(\frac{a}{\sigma_n}\right) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{1}{\sqrt{2}}\frac{a}{\sigma_n}\right)$$

Optimum decision threshold λ that maximizes P_e:

$$\boxed{\lambda_{\text{opt}} = \frac{N_0}{4AT_b}\log\left(\frac{p_0}{p_1}\right)}$$

2.3 Intersymbol Interference

Arises when the channel is *dispersive*, the magn. freq. resp. is not constant over the range of interest.

$$s(t) = \sum_k a_k\cdot g(t-kT_b)$$

$$y(t) = \mu\sum_k a_k\cdot p(t-kT_b)+n(t)$$

$$t(t_i) = \underbrace{\mu a_i}_{i\text{-th bit}} + \underbrace{\sum_{\substack{k=-\infty \\ k\neq i}}^{\infty} a_k p(i-k)T_b}_{\text{ISI}} + n(t_i)$$

2.4 Nyquist's Criterion

In order to avoid ISI, we require $p(mT_b) = 0$ for $m \neq 0$ and obtain

$$\sum_{m=-\infty}^{\infty} p(mT_b)\delta(t-mT_b) = \delta(t) \circ\!\!-\!\!\bullet P_\delta(f) = 1$$

And the nyquist criterion ($R_b = 1/T_b$ symbol rate):

$$\boxed{\sum_{n=-\infty}^{\infty} P(f-nR_b) = T_b}$$

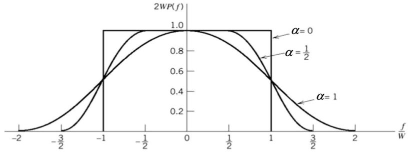
In words: The pulse function P in freq. domain copied with spacing R_b must be constant.

Ideal nyquist channel: The simplest function $P(f)$ that satisfies this is the rectangular function (ideal LPF) with $W = R_b/2$, R_b the nyquist rate.

$$P(f) = \frac{1}{2W}\,\mathrm{rect}\left(\frac{f}{2W}\right) = \begin{cases} \frac{1}{2W}, & -W \leq f \leq W \\ 0, & |f| > W \end{cases}$$

$$p(t) = \mathrm{sinc}(2Wt) \quad W = \frac{1}{2T_b} \quad E_b = \frac{A^2}{R_b}$$

Raised Cosine Spectrum: consists of flat portion and sinusodial rolloff.



$$P(f) = \begin{cases} \frac{1}{2W} & 0 \leq |f| \leq f_1 \\ \frac{1}{4W}\left(1-\sin\left[\frac{\pi(|f|-W)}{2W-2f_1}\right]\right) & |f| \in [f_1, 2W-f_1] \\ 0 & |f| > 2W-f_1 \end{cases}$$

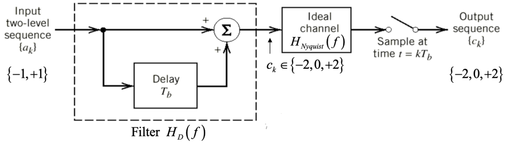
$$p(t) = \mathrm{sinc}(2Wt)\left(\frac{\cos(2\pi\alpha Wt)}{1-16\alpha^2W^2t^2}\right)$$

$$\alpha = 1-\frac{f_1}{W} \in [0,1] \quad \text{Rolloff factor}$$

Bandwidth is larger: $B_T = 2W-f_1 = W(1+\alpha)$.

2.5 Correlative-Level Coding

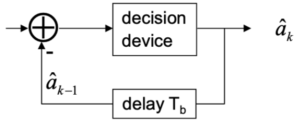
Use basepulses which introduce controlled ISI. Same BW but higher P_e



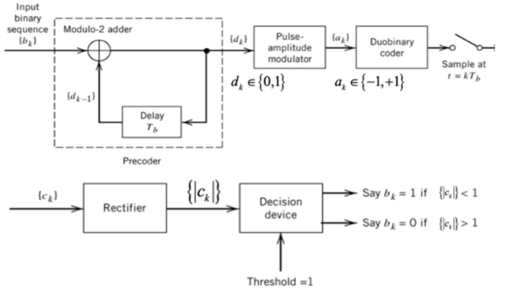
$$H_I(f) = \begin{cases} 2T_b\cos(\pi fT_b)e^{-i\pi fT_b}, & |f| < 1/2T_b \\ 0, \text{else} \end{cases}$$

$$h_I(t) = \frac{T_b^2\sin(\pi t/T_n)}{\pi t(T_b-t)}$$

Decoding



Precoding The decision feedback receiver is prone to propagating error. Using modulo-2 precoding, this can be omitted.



$$\boxed{\begin{gathered} d_k = b_k \oplus d_{k-1} \Rightarrow b_k = d_k \oplus d_{k-1} \\ c_k = \begin{cases} 0, & b_k = 1 \\ \pm 2, & b_k = 0 \end{cases} \end{gathered}}$$

2.6 Baseband M-ary PAM Transmission

In a M-ary PAM system: M possible amplitude levels. One symbol encodes $\log_2 M$ bits. Thus the signal rate T is related to the bit duration T_b of a binary PAM as:

$$T = T_b\log_2 M$$

- For same avg. P_e, an M-ary PAM requires more Tx power
- If $M \gg 2$ the Tx energy per bit must be increased by $M^2/(3\log_2 M)$ for same P_e

3 Signal Space Analysis

Continuous AWGN (Additive white gaussian noise) channel.

- All symbols m_i from source are eually likely $p_i = p(m_i) = \frac{1}{M}$
- Transmitter codes each m_i into a signal $s_i(t) \in \{s_k(t) | 1 \leq k \leq M\}$

- Channel adds AWGN $x(t) = s_i(t) + w(t)$ for $0 \leq t \leq T$
- The optimal receiver minimizes the avg. pob. of symbol error P_e

$$P_e = \sum_{i=1}^{M} p_i \mathbf{P}\left(\hat{m} \neq m_i \mid m_i\right)$$

3.1 Geometric Signal Representation

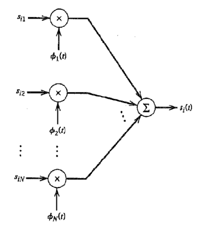
Let $\{\phi_i(t)\}_{i=1\dots N}$ be a set of othonormal basis fuctions of the signal set $\{s_i(t)\}_{i=1\dots M}$. All signals can be expressed as a finite sum. The coeff. s_{ij} are given by the projection onto $\{\phi_i(t)\}_{i=1\dots N}$.

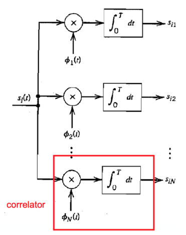
The orthonormal functions deine a N-dimensional Euclidean space - the signal space.

$$\int\limits_0^T \phi_i(t)\phi_j(t)\,\mathrm{d}t = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & - \neq j \end{cases}$$

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\phi_j(t) \quad s_{ij} = \int_0^T s_i(t)\phi_j(t)\,\mathrm{d}t$$

$$0 \leq t \leq T, \quad i = 1 \dots M, \quad j = 1 \dots N$$





$$\langle s_i(t), s_k(t)\rangle = \int_0^T s_i(t)s_k(d)\,\mathrm{d}t = \mathbf{s}_i^\top \cdot \mathbf{s}_k$$

$$\|\mathbf{s}_i\|^2 = \langle s_i(t), s_i(t)\rangle = \int_0^T s_i(t)^2$$

$$\|\mathbf{s}_i - \mathbf{s}_k\|^2 = \sum_{j=1}^{N}(s_{ij} - s_{kj})^2 = \int_0^T (s_i(t) - s_k(t))^2\,\mathrm{d}t$$

$$\cos\theta_{jk} = \frac{\mathbf{s}_i^\top \cdot \mathbf{s}_k}{\|\mathbf{s}_i\| \cdot \|\mathbf{s}_k\|} \quad E_i = \sum_{j=1}^{N} s_{ij}^2 = \|\mathbf{s}_i\|^2$$

Gram-Schmidt orthogonalization procedure: Start with a complete system $s_1(t), \dots, s_M(t)$ that generates the signal space. At each step generate a new basis function ϕ_i. The basis has only $N \leq M$ functions.

1. Build basis function ϕ_1 from s_1

$$\phi_1(t) = \frac{s_1(t)}{\sqrt{\int_0^T s_1^2(t)\,\mathrm{d}t}}$$

2. Search for a basis function from $s_2(t)$

$$s_{21} = \langle s_2(t), \phi_1(t)\rangle = \int_0^T s_2(t)\phi_1(t)\,\mathrm{d}t$$
$$g_2(t) = s_2(t) - s_{21}\phi_1(t)$$

If $g_2 = 0$, s_2 is lin. dep. on ϕ_1 and does not lead to a new basis function. Otherwise:

$$\phi_2(t) = \frac{g_2(t)}{\sqrt{\int_0^T g_2^2(t)\,\mathrm{d}t}}$$

3. Search for a basis function from $s_3(t)$

$$s_{31} = \langle s_3(t), \phi_1(t)\rangle \quad s_{32} = \langle s_3(t), \phi_2(t)\rangle$$
$$g_3(t) = s_3(t) - s_{31}\phi_1(t) - s_{32}\phi_2(t)$$

If $g_3 = 0$, s_3 is lin. dep. on ϕ_1 and ϕ_2 and does not lead to a new basis function. Otherwise:

$$\phi_3(t) = \frac{g_3(t)}{\sqrt{\int_0^T g_3^2(t)\,\mathrm{d}t}}$$

4. Search for a basis function from $s_M(t)$. Project s_M on the already determined basis functions, decompose S_M into its projection and a difference term g_M. If $g_M \neq 0$:

$$\phi_N(t) = \frac{g_M(t)}{\sqrt{\int_0^T g_M^2(t)\,\mathrm{d}t}}$$

3.2 Discrete System Model

The signal vector $\mathbf{s}$, noise vector $\mathbf{w}$ and the received signal $\mathbf{x}$.

$$\mathbf{s}_i = \begin{bmatrix} s_{i1} & \dots & s_{iN}\end{bmatrix}^\top \quad \mathbf{w} = \begin{bmatrix} w_1 & \dots & w_N\end{bmatrix}^\top$$
$$\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_N\end{bmatrix}^\top = \mathbf{s}_i + \mathbf{w}$$
$$\mathrm{E}\left[w_j\right] = 0 \quad \mathrm{E}\left[w_j \cdot w_k\right] = \delta_{jk} \quad \mathrm{Var}\left(w_j\right) = \frac{N_0}{2}$$

Theorem of Irrelevance For signal detection with AWGN, only the projection of the noise onto the basis functions of the signal set $\{s_i(t)\}_{i=1}^{M}$ affect the sufficient statistics of the detection problem. The remainder of the noise is irrelevant.

$$\mu_{X_j} = \mathrm{E}\left[X_j\right] = \mathrm{E}\left[s_{ij} + W_j\right] = s_{ij} + \mathrm{E}\left[W_j\right] = s_{ij}$$
$$\sigma_{X_j}^2 = \mathrm{Var}\left(X_j\right) = \mathrm{E}\left[(X_j - s_{ij})^2\right] = \mathrm{E}\left[W_j^2\right] = \frac{N_0}{2}$$
$$W_j = \int_0^T W(t)\phi_j(t)\,\mathrm{d}t$$

The elements X_j and X_k of the received signal vector have the covariance

$$\mathrm{Cov}\left(x_j, x_k\right) = \mathrm{E}\left[(x_j - \mu_{x_j})(x_k - \mu_{x_k})\right] = 0, \quad j \neq k$$

Thus the x_j are mutually uncorrelated. $\implies$ statistical independence.

Likelihood Function As the x_j are statistically indep. the conditional PDF of $\mathbf{x}$ given $\mathbf{s}$ (i.e. symbol m_i sent using signal s_i) follows:

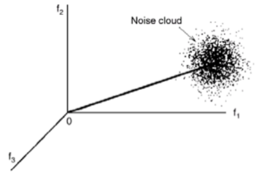
$$L(\mathbf{s}_i) := f_x(\mathbf{x}|\mathbf{s}_i) = f_W(\mathbf{w} = \mathbf{x} - \mathbf{s}_i) =$$
$$= \frac{1}{(\pi N_0)^{N/2}} \exp\left[-\frac{1}{N_0}\sum_{j=1}^{N}(x_j - s_{ij})^2\right]$$
$$l(\mathbf{s}_i) = \log L(\mathbf{s}_i) = -\frac{1}{N_0}\sum_{j=1}^{N}(x_j - s_{ij})^2 + c$$

$$c = -\frac{N}{2}\log(\pi N_0) \quad i \in \{1, \dots, M\}$$

L likelihood function, l log-likelihood function can be used because the pdf is always nonnegative and monot. incr. The constant c is indep. of hyp. $\mathbf{s}_i$ and can be discarded for the decision.

3.3 Detection and Decoding

Detection problem: Given the observation $\mathbf{x}$, determine an estimate $\hat{m}$ of the transmitted symbol m_i, s.t. the probability of error is minimized.



$$P_e(m_i|\mathbf{x}) = \mathbf{P}\left(m_i \text{not sent} \mid \mathbf{x}\right) = 1 - \mathbf{P}\left(m_i \text{sent} \mid \mathbf{x}\right)$$

The MAP (Maximum-A-Posteriori) decision rule is optimum in the minimum prob. of error sense. Set $\hat{m} = m_i$ if:

$$\mathbf{P}\left(m_i \text{sent} \mid \mathbf{x}\right) \geq \mathbf{P}\left(m_k \text{sent} \mid \mathbf{x}\right) \quad \forall k \neq i$$

Rephrased using Baye's rule, set $\hat{m} = m_i$ if (p_k a priori prob. of transmitting m_k, $f_x(\mathbf{x}|m_k)$ cond. pdf of $\mathbf{x}$ given m_k):

$$\hat{m} = \arg\max_{m_k} \frac{p_k \cdot f_x(\mathbf{x}|m_k)}{f_x(\mathbf{x})} \quad \forall k \neq i$$

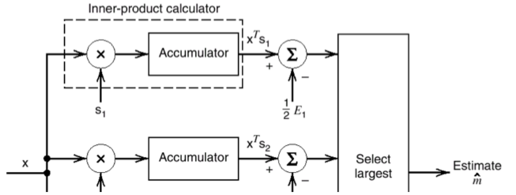
We can drop $f_x(\mathbf{x})$ as it is indep. of the symbol decision. For equiprobable source symbols, we obtain the ML decision rule: Set $\hat{m} = m_i$ if $l(m_k)$ max. for $k = i$.

Simplfied ML Rule: $\mathbf{x}$ lies in region Z_i if

$$\sum_{j=1}^{N} x_j s_{kj} - \frac{1}{2}E_k$$

is maximum for $k = i$.

Correlation receiver



3.4 Probability of Error

$\mathbf{P}\left(A_{ik}\right) = P_2(\mathbf{s}_i, \mathbf{s}_k)$ is the pairwise error prob. that observation $\mathbf{x}$ is closer to $\mathbf{s}_k$ than to $\mathbf{s}_i$:

$$P_2(\mathbf{s}_i, \mathbf{s}_k) = \mathbf{P}\left(\|\mathbf{x} - \mathbf{s}_k\|^2 < \|\mathbf{x} - \mathbf{s}_i\|^2\right)$$

With the euclidean distance $d_{14} := \|\mathbf{s}_1 - \mathbf{s}_4\|$:

$$P_2(\mathbf{s}_1, \mathbf{s}_4) = \mathbf{P}\left(z < -\frac{1}{2}d_{14}\right) = Q\left(\frac{d_{14}}{\sqrt{2N_0}}\right)$$
$$= \frac{1}{2}\,\mathrm{erfc}\left(\frac{d_{14}}{2\sqrt{N_0}}\right)$$

The pairwise probability of error only depends on the Euclidean distance and is e.g. invariant to rotation and translation of the signal constellation

From the union bound we have

$$P_e(m_i) \leq \sum_{\substack{k=1\\k\neq i}}^{M} P_2(\mathbf{s}_i, \mathbf{s}_k)$$

P_e is the error prob. averaged over all symbols. An upper bound follows as

$$P_e = \sum_{i=1}^{M} p_i P_e(m_i) \leq \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M} p_i \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right)$$

4 Passband Data Transmission

In bandpass data transmission, information modulates a carrier and occupies a restricted bandwidth in frequency. The carrier can be modulated by changing:

- Amplitude (ASK)
- Phase (PSK)
- Frequency (FSK)

Coherent modulation is when the receiver's local oscillator is phase-synchronous to the transmitter's local oscillator.

$M = s^n$ levels for signalling information (M-ary xSK). Using M levels, symbol duration $T = nT_b$ is changed while keeping the same datarate. Bandwidth shrinks accordingly by $1/nT_b$.

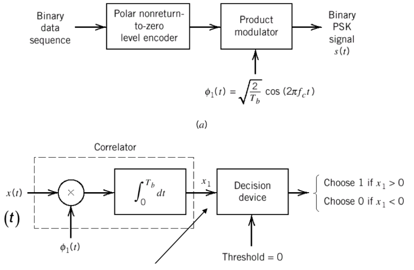
Figures of merit: Symbol error probability at given SNR, power spectral density, bandwidth efficiency $\rho = R_b/B$ [bit/s/Hz].

4.1 PSK: Coherent Phase Shift Keying

BPSK: Binary PSK

$$\phi_1(t) = \sqrt{\frac{2}{T_b}}\cos(2\pi f_c t)$$

$$s_1(t) = \sqrt{E_b}\phi_1(t) \quad s_2(t) = -\sqrt{E_b}\phi_1(t)$$



$$p_{10} = p_{01} = P_e = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)$$

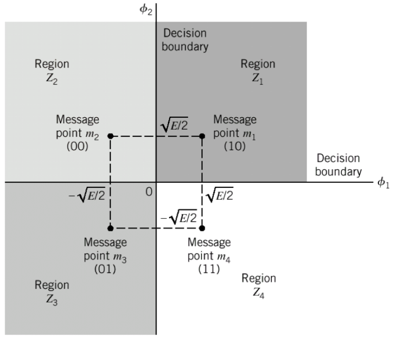
$$S(f-f_c) \approx 2E_b\operatorname{sinc}^2(T_b f)$$

QPSK: Quadriphase SK, use more than just two phase levels.

$$\phi_1(t) = \sqrt{\frac{2}{T_b}}\cos(2\pi f_c t) \quad \phi_2(t) = \sqrt{\frac{2}{T_b}}\sin(2\pi f_c t)$$

$$\mathbf{s}_{10} = \begin{bmatrix}+c\\+c\end{bmatrix} \mathbf{s}_{00} = \begin{bmatrix}-c\\+c\end{bmatrix} \mathbf{s}_{01} = \begin{bmatrix}-c\\-c\end{bmatrix} \mathbf{s}_{11} = \begin{bmatrix}+c\\-c\end{bmatrix}$$

$$c = \sqrt{E/2}$$



Every QPSK symbol carries 2 bits, hence the symbol energy is twice the energy per information bit: $E = 2E_b$. A QPSK system achieves same BER (P_e) as a BPSK at same E_b/N_0 but at *twice the bit rate*.

$$P_e = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)$$

$$S_B(f) = 4E_b\operatorname{sinc}^2(2T_b f)$$

4.2 QAM: Hybrid Amplitude/Phase Modulation

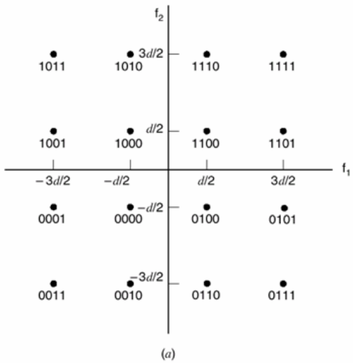
QAM: M-ary quadrature amplitude modulation, change phase and amplitude.

$d_{\min}$ is the distance between adjacent messages in the signal space.

$$\phi_1(t) = \sqrt{\frac{2}{T_b}}\cos(2\pi f_c t) \quad \phi_2(t) = \sqrt{\frac{2}{T_b}}\sin(2\pi f_c t)$$

$$\mathbf{s}_i = \frac{d_{\min}}{2}\begin{bmatrix}a_i\\b_i\end{bmatrix} a_i, b_i \text{odd integers}, i = 1, \ldots, M$$

Mapping an even number f bits per symbol (e.g. 4bits → 16 symbolds), results in a quadratic $L \times L$ square constellation with $L = \sqrt{M}$. Gray coding is often used for mapipng the bits to the QAM symbols.



(a)

$$P_c = (1-P_e')^2 \rightarrow P_e = 1-P_c = 1-(1-P_e')^2 \approx 2P_e'$$

$$P_e' = \left(1-\frac{1}{\sqrt{M}}\right)\operatorname{erfc}\left(\sqrt{\frac{d_{\min}^2}{4N_0}}\right)$$

$$E_{\mathrm{av}} = \frac{(M-1)d_{\min}^2}{6}$$

$$P_e \approx 2P_e' = 2\left(1-\frac{1}{\sqrt{M}}\right)\operatorname{erfc}\left(\sqrt{\frac{3E_{\mathrm{av}}}{2(M-1)N_0}}\right)$$

E_{av} average symbol energy.

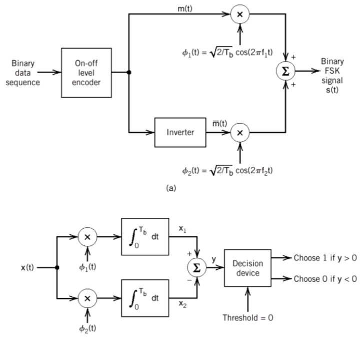
4.3 FSK: Coherent Frequency-Shift Keying

$$\phi_i(t) = \begin{cases}\sqrt{\frac{2}{T_b}}\cos(2\pi f_i t), & 0 \leq t \leq T_b\\ 0, & \text{else.}\end{cases}$$

$$f_i = \frac{n_c+i}{T_b} \quad i = 1, 2, \quad n_c \in \mathbb{N}$$

$$\mathbf{s}_1 = \sqrt{E_b}\begin{bmatrix}1\\0\end{bmatrix} \quad \mathbf{s}_2 = \sqrt{E_b}\begin{bmatrix}0\\1\end{bmatrix}$$

f_i chosen by rule to avoid phase discontinuities. The two frequencies f_1 and f_2 are $1/T_b$ Hz appart. The ϕ_i are orthogonal for $f_i = (n_c+i)/T_b$.



Distance between message points in signal space is $1/\sqrt{2}$ smaller compared to binary PSK.

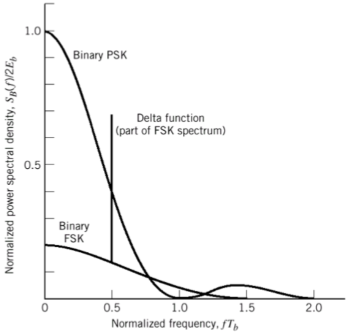
$$d_{\min} = \sqrt{2E_b}$$

$$P_e = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{(d_{\min}/2)^2}{N_0}}\right) = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right)$$

$$S_B(f) = \frac{E_b}{2T_b}\left[\delta\left(f-\frac{1}{2T_b}\right)+\delta\left(f+\frac{1}{2T_b}\right)\right]+\ldots$$

$$\frac{8E_b\cos^2(\pi T_b f)}{\pi^2(4T_b^2 f^2-1)^2}$$

PSD contains two delta pulses and decays much faster than BPSK due to continuous phase operation.



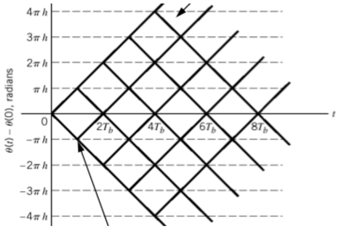
4.4 CPFSK Continuous Phase FSK

$$s(t) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t+\theta(t))$$

$$\theta(t) = \theta(0)\pm\frac{\pi h}{T_b}t, \quad 0 \leq t \leq T_b$$

$$h = T_b(f_1-f_2), \quad f_c = \frac{1}{2}(f_1+f_2)$$

h modulation index.



4.5 MSK Minimum Shift Keying

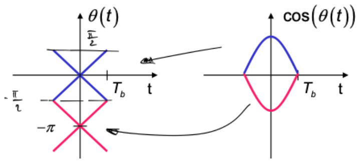
For integer valued h, the accumulated phase at end of symbol is independent of the previous and current symbol. $\implies$ no phase memory, each symbol can be decoded independently.

$$f_1 - f_2 = 0.5/T_b$$

The minimum difference, for which $s_1(t), s_2(t)$ orthogonal.

$\theta(0) = 0$	$\theta(T_b) = \pi/2$	symbol 1 transmitted
$\theta(0) = \pi$	$\theta(T_b) = \pi/2$	symbol 0 transmitted
$\theta(0) = -\pi$	$\theta(T_b) = -\pi/2$	symbol 1 transmitted
$\theta(0) = 0$	$\theta(T_b) = -\pi/2$	symbol 0 transmitted

Estimation of $\theta(0)$: Expanding $s(t)$ into two terms we get:



We can estimte $\theta(0)$ by observing

$$\sqrt{\frac{2E_b}{T_b}} \cos(\theta(t)) \cos(2\pi f_c t)$$

MSK Signal-Space representation

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos\left(\frac{\pi}{2T_b} t\right) \cos(2\pi f_c t), \quad -T_b \leq t \leq T_b$$

$$\phi_2(t) = \sqrt{\frac{2}{T_b}} \sin\left(\frac{\pi}{2T_b} t\right) \sin(2\pi f_c t), \quad 0 \leq t \leq 2T_b$$

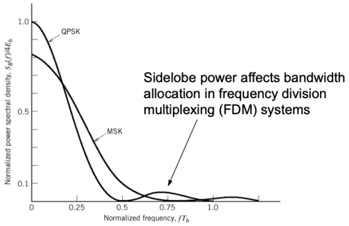
A coherent receiver has to integrate over two bit periods:

$$x_1 = \int_{-T_b}^{T_b} x(t)\phi_1(t)\,\mathrm{d}t = \sqrt{E_b}\cos(\theta(0)) + w_1$$

$$x_2 = \int_{0}^{2T_b} x(t)\phi_2(t)\,\mathrm{d}t = -\sqrt{E_b}\sin(\theta(T_b)) + w_2$$

Bit error rate The four points in the signal-space diagram correspond to two symbol, hence the BER (P_e) is the same with QPSK.

$$P_e = \frac{1}{2}\operatorname{erfc}\left(\frac{d_{\min}/2}{\sqrt{N_0}}\right) = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)$$

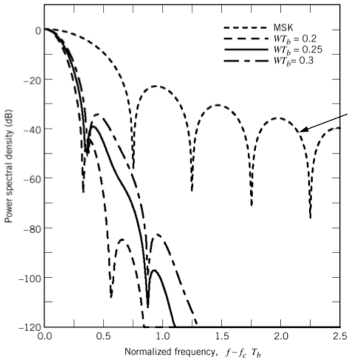


Sidelobe power affects bandwidth allocation in frequency division multiplexing (FDM) systems

4.6 GMSK

To make sidelobes of MSK smaller, filter the NRZ signal with pulse shaping function.

$$\boxed{H(f) = \exp\left[-\frac{\log 2}{2}\left(\frac{f}{W}\right)^2\right]}$$

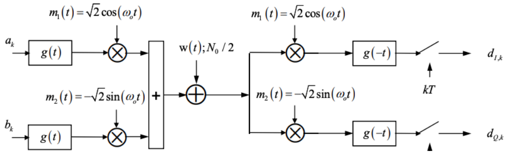


The parameter α depends on the time bandwidth product WT_b. The quantity $10\log(\alpha/2)$ expresses the degradation in dB of GMSK compared to MSK. MSK: $WT_b = \infty, \alpha = 2$

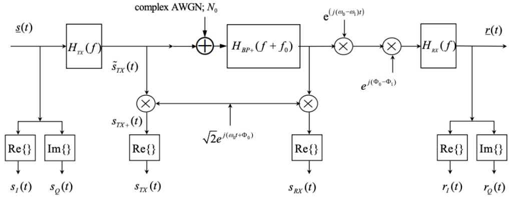
$$P_e = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{\alpha E_b}{2N_0}}\right)$$

4.7 Equivalent baseband representation

QAM: Two branches: inphase (I) and quadrature (Q)



By using complex valued signals, the transmission system can be written as an LTI system.



Important names and notation:

$\tilde{s}_{\mathrm{TX}}(t)$	complex envelope of s_{TX}
$s_{\mathrm{TX}+}(t)$	analytic signal (pre-envelope of s_{TX})
$s_{\mathrm{TX}}(t)$	physical passband signal

$$S_{TX+}(f) = \begin{cases} 2S_{\mathrm{TX}}(f) & f > 0 \\ 0 & \text{else} \end{cases}$$

$$H_{BP+}(f) = \begin{cases} H_{\mathrm{BP}}(f) & f > 0 \\ 0 & \text{else} \end{cases}$$

$$H_{\mathrm{TX}}(f) = 0 \forall |f| \geq f_0 \quad H_{\mathrm{RX}}(f) = 0 \forall |f| \geq B$$

$$B = \max(0, f_0 - |f_0 - f_1|)$$

Complex Envelope Inphase and quadrature components.

$$\tilde{s}_{\mathrm{TX}}(t) = \tilde{s}_{\mathrm{TX,I}}(t) + j\tilde{s}_{\mathrm{TX,Q}}(t) \quad \tilde{S}_{\mathrm{TX}}(f) = 0 \forall |f| > f_0$$

Analytic signal Pre-envelope. Has one-sided spectrum, scaling factor to preserve power values in passband and equivalent baseband.

$$s_{\mathrm{TX}+}(t) = \tilde{s}_{\mathrm{TX}}(t)\sqrt{2}\exp(j\omega_0 t)\exp(j\phi_0)$$

$$S_{\mathrm{TX}+}(f) = \sqrt{2}\exp(j\phi_0)\tilde{S}_{\mathrm{TX}}(f - f_0)$$

$$S_{\mathrm{TX}+}(f) = 0 \forall f < 0$$

$$\tilde{S}_{\mathrm{TX}}(f) = \frac{1}{\sqrt{2}}\exp(-j\phi_0) S_{\mathrm{TX}+}(f + f_0)$$

Physical passband signal

$$s_{\mathrm{TX}}(t) = \mathrm{Re}\{s_{\mathrm{TX}+}(t)\}$$

$$S_{\mathrm{TX}}(f) = \frac{1}{2}\left(S_{\mathrm{TX}+}(f) + S^*_{\mathrm{TX}+}(-f)\right)$$

$$s_{\mathrm{TX}}(t) = \sqrt{2}\tilde{s}_{\mathrm{TX,I}}(t)\cos(\omega_0 t + \phi_0) - \sqrt{2}\tilde{s}_{\mathrm{TX,Q}}(t)\sin(\omega_0 t + \phi_0)$$

$$s_{\mathrm{TX}}(t) = \left\{\sqrt{2}\sqrt{\tilde{s}^2_{\mathrm{TX,I}}(t) + \tilde{s}^2_{\mathrm{TX,Q}}(t)}\right\}\cos(\omega_0 t + \phi_0 + \phi(t))$$

$$\phi(t) = \mathrm{atan2}(\tilde{s}_{\mathrm{TX,Q}}(t), \tilde{s}_{\mathrm{TX,I}}(t))$$

Summary

$$\boxed{x(t) = \mathrm{Re}\{x_+(t)\} \quad x_+(t) = \tilde{x}(t)\sqrt{2}\mathrm{e}^{j2\pi f t}}$$

$x(t)$	physical passband signal
$x_+(t)$	analytic signal (pre-envelope of $x(t)$)
$\tilde{x}(t)$	compex envelope of $x(t)$

4.8 Noncoherent Detection

Carrier phase θ at the receiver becomes a random variable.

4.9 ML detection with unknown phase shift

$$L(\mathbf{s}_i) \triangleq f_X(\mathbf{x}|\mathbf{s}_i) = \frac{1}{(\pi N_0)^{N/2}}\exp\left[-\frac{1}{N_0}\sum_{j=0}^{N}(x_j - s_{ij})^2\right]$$

The ML receiver selects the hypothesis, which maximizes the likelihood function

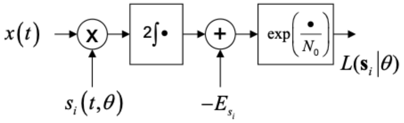
$$\hat{i} = \arg\max_i (L(\mathbf{s}_i))$$

Expanding the sum in the exponent, the likelihood function can be calculated from the output of a correlator bank.

$$\boxed{L(\mathbf{s}_i) = c\exp\left[\frac{2}{N_0}\int x(t)s_i(t)\,\mathrm{d}t - \frac{1}{N_0}E_{s_i}\right]}$$

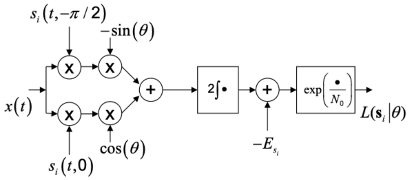
With $E_{s_i} = \sum_j s_{ij}^2 = \int s_i^2\,\mathrm{d}t$

For a known phase offset, the modified receiver correlates with a rotate version of each hypothesis.



Two-branch correlator

$$s_i(t,\theta) = s_i(t,\theta = 0)\cos\theta - s_i(t,\theta = -\pi/2)\sin\theta$$



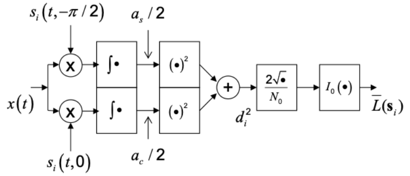
Equi-Energy Signals with unknown phase offset Shifting the integrator to each branch and obtain equi-energy signals with known phase offset:

$$L(\mathbf{s}_i|\theta) = \exp\left(\frac{1}{N_0}\left(a_c\cos\theta - a_s\sin\theta\right)\right)$$
$$= \exp\left(\frac{1}{N_0}\sqrt{a_c^2+a_s^2}\cos(\theta+\phi)\right)$$
$$\phi = \angle(a_c + ja_s)$$

With unknown phase offset, we have to average the likelihood function across all phase offsets θ.

$$\overline{L(\mathbf{s}_i)} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\left(\frac{1}{N_0}\sqrt{a_c^2+a_s^2}\cos(\theta+\phi)\right)\,\mathrm{d}\theta$$
$$= I_0\left(\frac{1}{N_0}\sqrt{a_c^2+a_s^2}\right)$$

I_0 is the modified Bessel function of order zero.



As I_0 is monotonously increasing, a simplified decision rule follows as

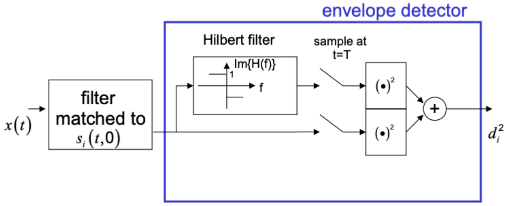
$$\hat{i} = \arg\max_i d_i^2$$

Note that we need a two-branch correlator for each hypothesis $s_i(t)$.

Instead of the two-branch correlator we can use two matched filter - sampler pairs to calculate the decision variable.

We can determine the decision variable with one matched filter and a Hilbert tranformer. The matched filter - envelope detector pair is called a noncoherent matched filter.

$$s_i(t,\theta=-\pi/2) \circ\!\!-\!\!\bullet S_i(f,\theta=-\pi/2)$$
$$= -j\,\mathrm{sgn}(f)S_i(f,\theta=0)$$

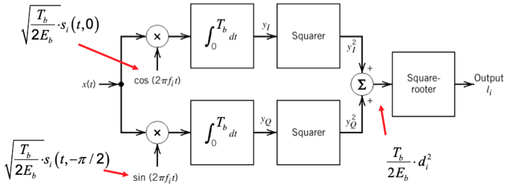


4.10 Noncoherent FSK

Signal $x(t)$ at the receiver with unknown carrier phase offset θ:

$$x(t) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_i t + \theta) + w(t), \quad i = 1,2, 0 \le t \le T_b$$

The signals s_1 and s_2 each require such a branch. A comparator subsequently compares the two outputs l_i to devide between the hypothesis s_1 and s_2.

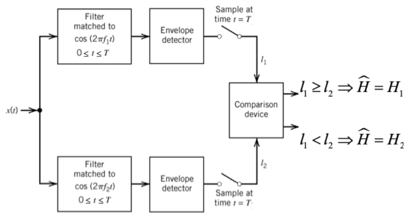


We have

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$$

This corresponds to a degradation of at least 3dB compared to coherent MSK. Less degradation compared to BFSK.

Another implementation is with matched bandpass filters to f_1 and f_2 followed by envelope detectors, samplers and a comparison device.



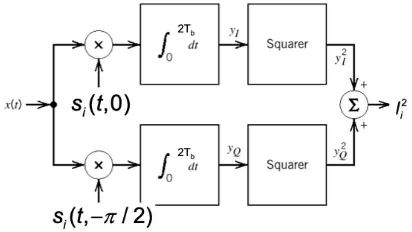
4.11 DPSK Differential PSK

Differential precoding at transmitter: Symbol 0 $\Longrightarrow$ π phasejump, Symbol 1 $\Longrightarrow$ no phase-jump. Assumption: θ does not change significantly between two adjecent sampling instances.

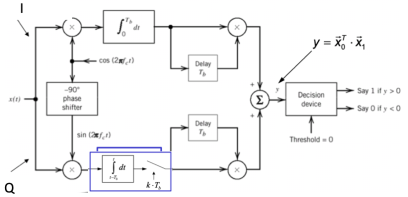
$$s_1(t,\theta) = \begin{cases}\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta), & 0 \le t \le T_b\\ \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta), & T_b \le t \le 2T_b\end{cases}$$

$$s_2(t,\theta) = \begin{cases}\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \theta), & 0 \le t \le T_b\\ \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t + \pi + \theta), & T_b \le t \le 2T_b\end{cases}$$

Noncoherent detector for DPSK:



Quadrature implementation of simplified detector:

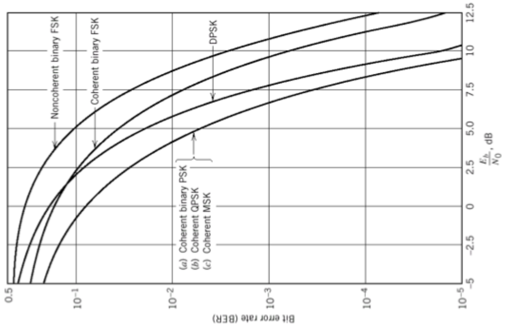


DPSK is a special case of noncoherent, orthogonal modulation with $T = 2T_b$ and $E = 2E_b$. The bit error rate is given by:

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$$

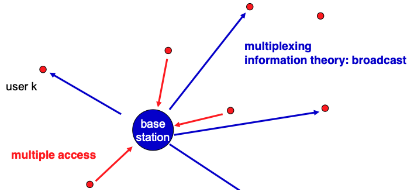
4.12 Performance comparison

Modulation	P_e
Coherent BPSK Coherent QPSK Coherent MSK	$\frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)$
Coherent binary FSK	$\frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right)$
DPSK	$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$
Noncoherent binary FSK	$\frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$



5 Multi User Radio Communications

5.1 Multiple Access techniques



Accomodation of several users in the same wireless environment.

- FDMA Frequency domain multiple access
- TDMA Time domain multiple access
- CDMA Code division multiple access
- SDMA Spatial division multiple access

5.2 Radio Communication over line-of-sight (LOS)

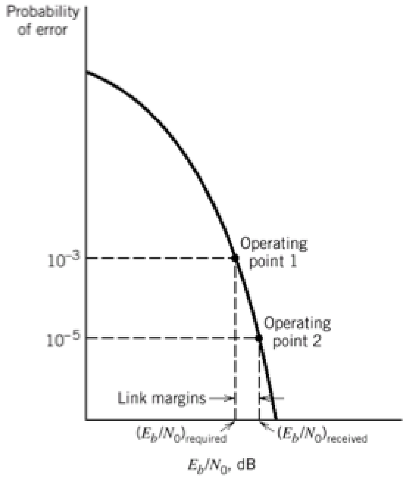
Free-space (line of sight) communication on up- and down-link. AWGN model appropriate. Used MA techniques:

- **FDMA** Non linearity of transponder causes interference between frequency subband. Transponder is operated in lin. regime below max output power. Reduced power efficiency.
- **TDMA** Can operate at close to full power efficiency. Commonly used.
- **SDMA** Multiple antennas allow beam forming to different lcoations.

Link budget Link (power) budget: Budgeting of all gains and losses. Accounting of resources available to transmitter and receiver, sources of loss of power, sources of noise. Allows performance estimation of LOS links.

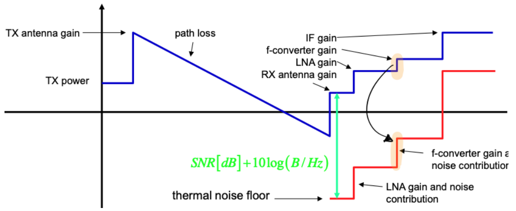
Link Margin Curve relates P_e to E_b/N_0. Max accepted P_e leads to Op1 and minimal required

$(E_b/N_0)_{\text{req}}$. Actually received $(E_b/N_0)_{\text{rec}}$ define Op2.



Link margin: Difference. Provides Protection agains unexpected changes. Large margin: high reliability but low efficiency.

$$M = 10\log\left(\frac{E_b}{N_0}\right)_{\text{rec}} - 10\log\left(\frac{E_b}{N_0}\right)_{\text{req}} \qquad [M] = \text{dB}$$



5.3 Antenna characterization

Receiver is located in the farfield of the transmitter. D largest dimenstion of antenna.

$$d_f \gg \frac{2D^2}{\lambda}$$

Idealized reference antenna radiates uniformly in all directions. Power density as a function of distance scales according to free-space propagation. Φ Radiation intensity in watts per unit solid angle.

$$\rho(d) = \frac{P_t}{4\pi d^2} \qquad [\rho] = \frac{\text{W}}{\text{m}^2} \qquad \Phi = d^2\rho(d)$$

Total radiated power P and average power per unit solid angle P_{av}:

$$P = \int \Phi(\theta,\phi)\,\mathrm{d}\Omega \quad P_{av} = \frac{1}{4\pi}\int \Phi(\theta,\phi)\,\mathrm{d}\Omega = \frac{P}{4\pi}$$

$$[P] = \text{W} \quad [P_{av}] = \frac{\text{W}}{\text{steradian}}$$

Directivity gain $g(\theta,\phi)$ Ratio of radiation intensity in a specific direction to the avg. radiated power.

Directivity D, maximal directivity gain over all directions

Power gain G with $\eta_{\text{radiation}} \in [0,1]$ the radiation efficiency factor.

EIRP Effective isotropically radiated power referenced to an isotropic source.

Beamwidth Angle between the two directions in which the radiation intensity is one-half the maximum. Higher power gain leads to narrower beamwidth.

Effective Apperture A Ratio of power available at the antenna terminals to the power per unit area of the appropriately polarized incident electromagnetic wave.

Apperture efficiency η_{ap} whith A_{ph} physical area.

$$g(\theta,\phi) = \frac{\Phi(\theta,\phi)}{P_{av}} = \frac{\Phi(\theta,\phi)}{P/(4\pi)} \qquad D = \max_{\theta,\phi} g(\theta,\phi)$$

$$G = \eta_{\text{radiation}} D \qquad \text{EIRP} = P_t G_t$$

$$A = (\lambda^2/4\pi)G \qquad \eta_{ap} = A/A_{ph}$$

Frii's Free-Space Equation Power captured by receiver at distance d in LoS:

$$P_r = \left(\frac{\text{EIRP}}{4\pi d^2}\right) A_r = \frac{P_t G_t A_r}{4\pi d^2} = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2$$

Path loss PL is the difference between transmit signal power and receive signal power.

$$\text{PL} = 10\log\frac{P_t}{P_r} = -10\log(G_t G_r) + 10\log\left(\frac{4\pi d}{\lambda}\right)^2$$

5.4 Noise figure

Spot noise figure $F(f)$ ratio of total available output noise power per unit bandwidth to portion thereof due to source alone. G power gain. **Lower is better**

$$F(f) = \frac{S_{NO}(f)}{G(f)S_{NS}(f)} = \frac{P_S S_{NO}}{P_O S_{NS}} = \frac{\text{SNR}_{\text{Source}}(f)}{\text{SNR}_{\text{Output}}(f)}$$

Signal to noise ratio after receiver amplifier is

$$\text{SNR}_{\text{in}} - F$$

If two-port is noise free:

$$S_{NO}(f) = G(f)S_{NS}(f)$$

For physical systems

$$S_{NO}(f) > G(f)S_{NS}(f)$$

Equivalent Noise Temperature T_e. For low noise devices T_e is a better measure bcs F is close to unity. Two-port device matched to source impedance is considered. N_1 available input noise power:

$$N_1 = \left(\frac{\sqrt{4kTR_sB}}{2R_s}\right)^2 R_s = kTB.$$

Total output noise power N_2 and noise figure:

$$N_2 = GN_1 + N_d = Gk(T+T_e)B$$

$$F = \frac{N_2}{N_2 - N_d} = \frac{T+T_e}{T}$$

Equivalent noise temperature T_e

$$T_e = T(F-1)$$

Cascade of Two-Port Networks Use factor not dB! Best if Lowest F first in chain.

$$F = F_1 + \frac{F_2-1}{G_1} + \frac{F_3-1}{G_1G_2} + \frac{F_4-1}{G_1G_2G_3} + \cdots$$

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1G_2} + \frac{T_{e4}}{G_1G_2G_3} + \cdots$$

5.5 Radio Communication over multipath channnels

In mobile radio systems the transmitters and receivers are mobile. This leads to a stochastic channel. Multipath phenomenenon leads to **Fading**.

Narrowband fading Complex envelopes ot $\tilde{s}$ Tx, $\tilde{s_o}$ Rx signal and $\tilde{h}$ timevarying impulse response of channel. The Rayleigh fading model for NLOS conditions is modeled as zero-mean complex Gaussian random process. Characterized by autocorrelation function $R_{\tilde{h}}$ and Doppler spectrum $S_{\tilde{H}}$.

$$\tilde{s}_o(t) = \int_{-\infty}^{\infty} \tilde{s}(t-\tau)\tilde{h}(\tau;t)\,\mathrm{d}\tau$$

$$\tilde{h}(\tau;t) = \tilde{h}(t)\delta(\tau) \implies \tilde{s}_o(t) = \tilde{h}(t)\tilde{s}(t)$$

$$R_{\tilde{h}}(\Delta t) = \mathrm{E}\left[\tilde{h}^*(t)\tilde{h}(t+\Delta t)\right] \circ\!\!-\!\!\bullet S_{\tilde{H}}(\nu)$$

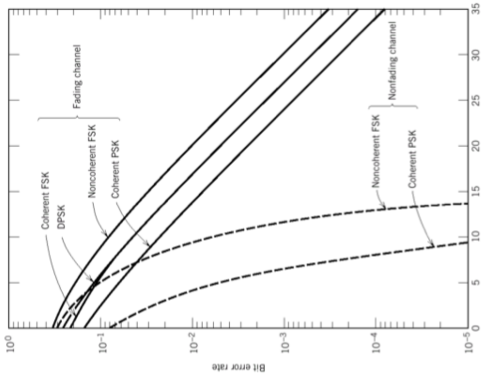
Coherent BPSK over slow rayleigh fading channel Input-output relation is $\tilde{s}_o(t) = \alpha\exp(-j\phi)\tilde{s}(t) + \tilde{w}(t)$ with α and ϕ Rayleigh and uniformly distributed r.v.

$$P_{e|\tilde{h}}(\gamma) = \frac{1}{2}\operatorname{erfc}\sqrt{\gamma} \qquad \gamma = \frac{\alpha^2 E_b}{N_0}$$

Averaging over all channel realizations:

$$P_e(\gamma_0) = \int_0^\infty P_{e|\tilde{h}}(\gamma)f(\gamma)\,\mathrm{d}\gamma = \frac{1}{2}\left(1 - \sqrt{\frac{\gamma_0}{1-\gamma_0}}\right)$$

$$\gamma_0 = \mathrm{E}[\gamma] = \frac{E_b}{N_0}\mathrm{E}\left[\alpha^2\right]$$



At high SNR deep fades dominate performance $P_e \propto 1/SNR$.

Diversity Availability of independently faded copies of the transmit signal at the receiver. Diversity order L number of available indep. faded versions of the same signal. At high SNR, L-th order diversity allows $P_e \propto 1/SNR^L$

5.6 Summary

$$N_0 = kTF \quad P_n = N_0 B \quad \text{SNR}_{\text{dB}} = P_{r,\text{dB}} - P_{n,\text{dB}}$$

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2$$

$$P_r = RE_{b,min} M \quad A = G_r \frac{\lambda^2}{4\pi} = \eta_{ap} A_{ph}$$

N_0 Noise power spectral density [W/Hz]
F Noise Figure
B Bandwidth [Hz]
P_n Noise power [W]
P_r/P_t Receiver/Transmitter power [W]
G_r/G_t Receiver/Transmitter antenna gain
d Distance between receiver/transmitter
R Datarate
$E_{b,min}$ Minimum energy per bit
M Link margin
A Effective apperture
η_{ap} Apperture efficiency
A_{ph} Physical area of antenna

6 Information Theory

Recap: p Bit error prob. for BNRZ channel and amplitude A.

$$p = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{A^2}{N_0}}\right)$$

6.1 Uncertainty, Information and Entropy

A source emits a message S. S is a r.v. taking values in a finite alphabet $S = \{s_0, \dots, s_{K-1}\}$.

$$\mathbf{P}(S = s_k) = p_k \qquad \sum_{k=0}^{K-1} p_k = 1$$

The definition of information I is:

$$I(s_k) \triangleq -\log p_k$$

$$\begin{array}{lll}
I(s_k) = 0 & \text{for} & p_k = 1 \\
I(s_k) \geq 0 & \text{for} & 0 \leq p_k \leq 1 \\
I(s_k) > I(s_i) & \text{for} & p_k < p_i \\
I(s_k s_i) = I(s_k) + I(s_i) & \text{if stat. indep.} & p_{ki} = p_k p_i
\end{array}$$

When I is specified in bits, logarithms are to base 2.

The definition of entropy H is:

$$H(\mathcal{S}) \triangleq E[I(\mathcal{S})] = \sum_{k=0}^{K-1} p_k I(s_k) = -\sum_{k=0}^{K-1} p_k \log p_k$$

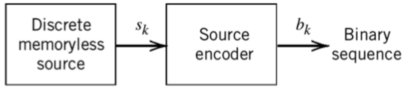
$$\begin{array}{lll}
0 \leq H(\mathcal{S}) \leq \log_2 K & & \\
H(\mathcal{S}) = 0 & \text{iff} & p_k = 1 \quad \text{for one } k \\
H(\mathcal{S}) = \log_2 K & \text{iff} & p_k = \frac{1}{K} \forall k
\end{array}$$

Extended source Divide a seq. of $n \cdot M$ successive source symbols into M blocks. Consider each block of n symbols as a single "super symbol" taking on values in $\mathcal{S}^n$. The entropy of the extended source is:

$$H(\mathcal{S}^n) = n \cdot H(\mathcal{S})$$

6.2 Source Coding Theorem

How is the information of a source efficiently represented? Requirements: Code words must be binary, unique decodability.



The average code word length of a code whose kth code-word is of length L_k:

$$\bar{L} = \sum_{k=0}^{K-1} p_k L_k$$

The average code word length is lower bounded by the entropy of the source.

$$\bar{L} \geq H(\mathcal{S}) \triangleq L_{min}$$

Coding efficiency is a meausre of code quality:

$$\eta = \frac{L_{min}}{\bar{L}} \leq 1$$

6.3 Data Compation

Goal: Elmininate redundancy. Seek for codes that approach Shannon's lower bound on the avg code-word length.

Prefix Codes: No code-word is a prefix of another code-word. Leads to implicit recognition of end of code word. For each discrete, memoryless source there exists a prefix code s.t.:

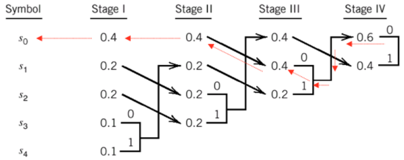
$$H(\mathcal{S}) \leq \bar{L} < H(\mathcal{S}) + 1$$
$$H(\mathcal{S}^n) \leq \bar{L}_n < H(\mathcal{S}^n) + 1$$
$$\Leftrightarrow nH(\mathcal{S}) \leq \bar{L}_n < nH(\mathcal{S}) + 1$$
$$\Leftrightarrow H(\mathcal{S}) \leq \frac{\bar{L}_n}{n} < H(\mathcal{S}) + \frac{1}{n}$$

$\frac{\bar{L}_n}{n}$ effective nmber of bit per source symbol.

Huffman Coding yields a prefix code that minimizes the avg. code-word length when the source is memoryless.

1. Assign a 0 and 1 to the symbols of lowest probability
2. Replace the two symbols by a new pseudo-symbol whose prob. is the sum of the two probs.
3. Repeat 1. and 2. until only a single pseudo-symbol left

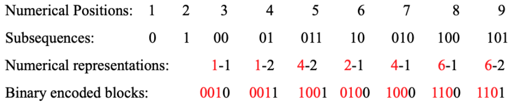
The code sequence for each symbol is found by backtracking from last symbol and tracing the 0s and 1s.



Drawbacks:

- Need to know probabilities a priori
- Redundancy due to memory in source can only be removed by using large extension codes, increasing complexity

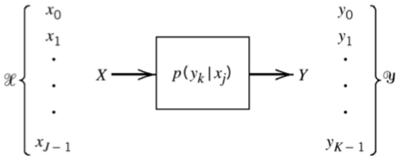
Lempel-Ziv Coding Adaptive algorithm of low complexity that captures the source statistic and memory in the source intrinsically.



Numerical Positions:	1	2	3	4	5	6	7	8	9
Subsequences:	0	1	00	01	011	10	010	100	101
Numerical representations:			1-1	1-2	4-2	2-1	4-1	6-1	6-2
Binary encoded blocks:			0010	0011	1001	0100	1000	1100	1101

- Constructed by parsing the source data stream into segments other than 0 and 1 that are shortest subsequences not encountered previously
- Segment 0 and 1 are asigned indices 1 and 2
- N stored subsequences are indexed from 3 to $N+2$
- A new sequence can always be composed from an old seubsequence (root subsequence) and a 0 and a 1 (innovation symbol)

6.4 Discrete Memoryless Channel



X, Y are statistically dependant r.v. Discrete if the input and output alphabet ($\mathcal{X}$, $\mathcal{Y}$) are of finite size Memoryless if the current output depends on the current input only. It is fully described by

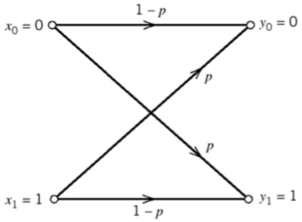
- input alphabet $\mathcal{X} = \{x_0, \dots, x_{J-1}\}$
- output alphabet $\mathcal{Y} = \{y_0, \dots, y_{K-1}\}$
- transition probabilities $p(y_k|x_j) = \mathbf{P}(Y = y_k \mid X = x_j)$

Transition Matrix

$$\mathbf{P} = \begin{bmatrix} p(y_0|x_0) & \cdots & p(y_{K-1}|x_0) \\ \vdots & \ddots & \vdots \\ p(y_0|x_{J-1}) & \cdots & p(y_{K-1}|x_{J-1}) \end{bmatrix}$$

$$p(y_k) = \mathbf{P}(Y = y_k) = \sum_{j=0}^{J-1} p(y_k|x_j) p(x_j)$$

Binary, Symmetric Channel



$J = K = 2$, transition probability p. Error probability is p.

$$C_{\text{BSC}} = 1 + p \log_2 p + (1-p) \log_2 (1-p)$$

6.5 Mutual Information

Conditional entropy of X given Y is a measure for the uncertainty about X if Y is known:

$$H(X|Y) = -\mathrm{E}_{X,Y}[\log_2 p(X|Y)] =$$
$$-\sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2 p(x_j|y_k)$$

Remember: $p(x_j, y_k) = p(x_j|y_k)\cdot p(y_k)$

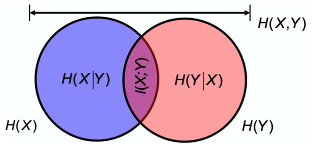
Mutual information is the reduction of the uncertainty about X achieved by observing Y.

$$\boxed{I(X;Y) \triangleq H(X) - H(X|Y)}$$

$$I(X;Y) = I(Y;X)$$
$$H(X) - H(X|Y) = H(Y) - H(Y|X)$$
$$I(X;Y) \geq 0$$
$$H(X) \geq H(X|Y)$$
$$I(X;Y) = 0 \Leftrightarrow X, Y \text{independent}$$

Joint entropy of X and Y is defined as:

$$H(X,Y) = -\mathrm{E}_{X,Y}[\log_2 p(X,Y)] =$$
$$-\sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2 p(x_j, y_k)$$
$$I(X;Y) + H(X,Y) = H(X) + H(Y)$$



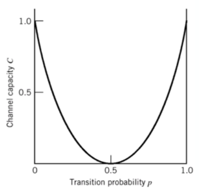
6.6 Channel Capacity

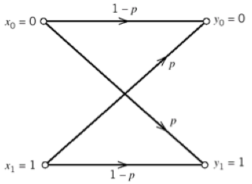
A channel is a statistical model with input X and output Y. Mutual information depends also on $p(x)$.

$$I(X;Y) = H(Y) - H(Y|X) =$$
$$\sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log\left(\frac{p(y_k|x_j)}{p(y_k)}\right) =$$
$$\sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(y_k|x_j)p(x_j)\log\left(\frac{p(y_k|x_j)}{\sum_{j=0}^{J-1} p(y_k|x_j)p(x_j)}\right)$$

Channel capacity is the maximum mutual information over all possible input distributions:

$$\boxed{C \triangleq \max_{\{p(x_j)\}} I(X;Y) \quad C = B\log_2(1+\mathrm{SNR})}$$





6.7 Channel Coding Theorem

Shannon's Channel Coding Theorem

- Consider discrete, memoryless source emitting values in $\mathcal{S}$
- One symbol emitted every T_s seconds - information rate $H(\mathcal{S})/T_s$
- One coded symbol transmitter every T_c seconds

> Theorem If $H(\mathcal{S})/T_s < C/T_c$ there exists a channel code yielding an arbitrarily small block (message) error probability as the channel code-word length goes to infinity. For $H(\mathcal{S})/T_s \geq C/T_c$ such a code does not exist.

6.8 Differential Entropy

Idea: Source and channels with continuous alphabets.

Differential entropy:

$$\boxed{h(X) = -\int_{-\infty}^{\infty} f_X(x)\log f_X(x)\,\mathrm{d}x}$$

The entropy $H(X)$ goes to ∞ in the limit of $\delta x \to 0$. But the mutual information is well defined:

$$\boxed{I(X;Y) = h(X) - h(X|Y)}$$

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ be a Gaussian random variable with variance σ^2. Then the diff. entropy is uniquely determined by its variance.

$$h(X) = \frac{1}{2}\log(2\pi e\sigma^2)$$

6.9 Information Capacity Theorem

Information capacity is the maximum of the mutual information between input X_i and output Y_i overa ll distributions of X_i fulfilling the contraint:

$$\boxed{C_D = \max_{\{f_X(x)\}} \{I(X_i;Y_i) : \mathrm{E}[X_i^2] = E_s\}}$$

Where E_s is interpreted as the average energy per symbol.

$$I(X_i;Y_i) = h(Yi) - h(Y_i|X_i) = h(X_i+N_i) - h(X_i+N_i|X$$
$$h(X_i+N_i) = \frac{1}{2}\log(2\pi e(E_s+\sigma^2))$$
$$h(N_i) = \frac{1}{2}\log(2\pi e\sigma^2)$$
$$C_D = \frac{1}{2}\log_2\left(1+\frac{E_s}{\sigma^2}\right)$$

6.10 Implications of the Inf. Capacity Thm.

ToDo: Not yet covered in lecture

6.11 Colored Noise Channel

ToDo: Not yet covered in lecture

7 Data Link Layer

7.1 Channel Coding

The channel encoder takes the information bit sequence m_i as input and outputs coded bits c_j. Received bits are denoted by r_j. We assume a discrete memoryless channel with noise. Purpose of encoding: Change BER from problematic to acceptable for a fixed E_b/N_0 or reduce required E_b/N_0 for a fixed BER.

FEC (Forward error correction) Decoder exploits redundancy to correct errors and decide on the message bits

Error detection Decoder exploits redundancy to detect errors, doesn't correct them

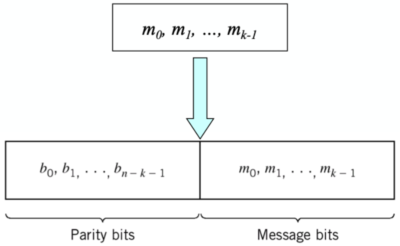
Block codes have no memory in the encoder and **Concolutional codes** have memory in the encoder.

7.2 Linear Block Codes

(n, k) linear block code takes k information bits and produces n coded bits.

- Any to code words added (mod 2) produce a third code word in the code
- The all-zero code word is part of the code

Systematic Linear Block Code Unaltered Message bits are extended with $n-k$ parity bits that are linear sums of the message bits.



$$b_i = p_{0,i}m_0 + p_{1,i}m_1 + \cdots + p_{k-1,i}m_{k-1}$$
$$p_{j,i} \in \{0, 1\}$$

In matrix notation:

$$\boxed{\begin{gathered}\mathbf{b} = [b_0, b_1, \ldots, b_{n-k-1}] \\ \mathbf{m} = [m_0, m_1, \ldots, m_{k-1}] \\ \mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0,n-k-1} \\ p_{10} & p_{11} & \cdots & p_{1,n-k-1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} \end{bmatrix} \\ \mathbf{b} = \mathbf{mP}\end{gathered}}$$

With the generator matrix $\mathbf{G}$:

$$\boxed{\begin{gathered}\mathbf{c} = [c_0, c_1, \ldots, c_{n-1}] = [\mathbf{b}|\mathbf{m}] = \mathbf{m}[\mathbf{P}|\mathbf{I}_k] \\ \mathbf{G} = [\mathbf{P}|\mathbf{I}_k] \quad \mathbf{c} = \mathbf{mG} \\ \mathbf{c}_i + \mathbf{c}_j = \mathbf{m}_i\mathbf{G} + \mathbf{m}_j\mathbf{G} = (\mathbf{m}_i + \mathbf{m}_j)\mathbf{G}\end{gathered}}$$

The parity check matrix $\mathbf{H}$:

$$\boxed{\begin{gathered}\mathbf{H} = [\mathbf{I}_{n-k}|\mathbf{P}^\top] \\ \mathbf{HG}^\top = 0 \quad \mathbf{cH}^\top = \mathbf{mGH}^\top = 0\end{gathered}}$$

Multiplying a code word with the parity check matric results in the zero vector.

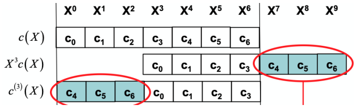
7.3 Cyclic Codes

Encoding and syndrome calculation with low complexity shift-registers. Practical decoding methods due to algebraic structure. **Cyclic property**: Any cyclic shift of a code word is also a code word. Description of a code via code polynomial:

$$\mathbf{c} = [c_0, c_1, \dots, c_{n-1}]$$

$$c(X) = c_0 + c_1X + c_2X^2 + \dots + c_{n-1}X^{n-1}$$

A cyclic shift is done by multiplication with X^j and modulo X^n+1. Or take the remainder of $c(X)X^j : (X^n+1)$ with $\mod n$ arithmetic. (E.g. with $n = 7$, $X^3/X^7 = X^{-4} = X^3 \mod 7$).



Polynomials and their order:

$m(X)$	Message polynomial	$\le k-1$
$g(X)$	Generator polynomial	$\le n-k$
$c(X) = m \cdot g$	Code polynomial	$\le n-1$
$s(X) = r \mod g$	Syndrome	

If $g(X)$ is a factor of X^n+1 then the code is cyclic:

$$g(X)h(X) = X^n+1$$

$h(X)$ is the parity check polynomial, as for all $c(X)$:

$$c(X)h(x) \mod (X^n+1)$$
$$= m(X)g(X)h(X) \mod (X^n+1) = 0$$

Generator Polynomial of degree $n-k$ that is a factor of X^n+1. Further is expanded as

$$g(X) = 1 + \sum_{i=1}^{n-k-1} g_iX^i + X^{n-k} \qquad g_i \in \{0,1\}$$

Parity Check Polynomial $h(X)$

$$g(X)h(X) = X^n+1$$

Systematic Cyclic Code Idea: Complement the shifted message polynomial such, that the resulting code word is a mutliple of $g(X)$. 1. Mult message by X^{n-k} 2. divide by $g(X)$ and get remainder $b(X)/g(X)$ 3. get systematic code $\tilde{c}(X)$.

$$m(X) = m_0 + m_1X + \dots + m_{k-1}C^{k-1}$$
$$\frac{X^{n-k}m(X)}{g(X)} = \tilde{x}(X) + \frac{b(X)}{g(X)}$$
$$\tilde{c}(X) = \tilde{m}(X)g(X) = b(X) + X^{n-k}m(X)$$

7.4 Minimum Distance Considerations

Hamming distance $d(c_1, c_2)$ number of locations in which two code words differ.

Hamming weight $w(c)$ number of non zero elements in code vector. The following hold only if linear block code:

$$d(c_1, c_2) = d(c_1 + c_2, 0) = w(c_1 + c_2)$$

Minimum distance $d_{\min}$ the smallest Hamming distance between any pair of code vectors in the code. If lin. code $d_{\min} = w_{\min}$

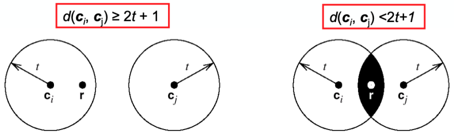
Given an $(n, k, d_{\min})$ block code and t the number of locations where a bit toggled after transmission. Error **detection**: All error patterns with

$$t \le d_{\min} - 1$$

can be detected. Error **correction** capability: Error patterns with weights

$$t \le \left\lfloor \frac{1}{2}(d_{\min} - 1) \right\rfloor$$

can be corrected surely.



Hamming bound: For given $d_{\min}$ and code word length n, good codes have large num. of possible code words 2^k, i.e. large code rate $r = k/n$. The number of code words for a binary code must satisfy:

$$2^k \left(1 + \binom{n}{1} + \dots + \binom{n}{t_0}\right) \le 2^n \qquad t_0 = \left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor$$

Perfect Code is if a binary code $(n, k, d_{\min})$ satisfies the hamming bound with equality.

7.5 Example: Hamming code

Hamming code are (n, k) lin. block codes with $m \ge 2$ and:

$$n = 2^m - 1 \quad k = 2^m - m - 1 = n - m \quad m = n - k$$

With $m = 3$ the $(7, 4)$ linear block code is:

$$\mathbf{G} = \left[\begin{array}{ccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \quad \underbrace{}_{\mathbf{P}} \; \underbrace{}_{\mathbf{I}_k}$$

Properties of Hamming codes: $d_{\min} = 3$ and $t_0 = 1$. They satisfy the Hamming bound with equality. They are Hamming codes are single-error corercting binary perfect codes.

7.6 Decoding Principles

Received bit vector $\mathbf{r} = \mathbf{c} + \mathbf{e}$ where $\mathbf{e}$ is the error vector/pattern.

The **syndrome** is defined as the projection of $\mathbf{r}$ onto $\mathbf{H}$:

$$\mathbf{s} = \mathbf{r}\mathbf{H}^\top$$

- The syndrom of $\mathbf{r}$ depends only on the error pattern $\mathbf{e}$
- All error patterns $\mathbf{e}_j^i$ that differ by a code word have the same syndrome $\mathbf{s}^i$. Thus 2^k distinct error patterns lead to the same syndrome $\mathbf{s}^i$.

Syndrom of cyclic codes

$$s(X) = r(X) \mod g(X)$$

Standard array and coset leader Construct a table with $N = n - k - 1$ and e_i the most probable $\mathbf{e}$ vectors (i.e. these with the least weight):

$c_1 = 0$	c_2	$\cdots$	c_{n-1}
e_1	$c_2 + e_1$	$\cdots$	$c_{n-1} + e_1$
e_2	$c_2 + e_2$	$\cdots$	$c_{n-1} + e_2$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
e_N	$c_2 + e_N$	$\cdots$	$c_{n-1} + e_N$

e_i are called the coset leaders. The syndrom vector points to a table entry. To obtain $\mathbf{c}$, XOR $\mathbf{r}$ with e_i.

7.7 Maximum Likelihood Decoding

ML decoding for discrete memoryless channel is minimum Hamming distance decoding:

$$\hat{\mathbf{c}} = \arg\max_{\mathbf{c} \in C} \mathbf{P}(\mathbf{r} \mid \mathbf{c}) \to \hat{\mathbf{c}} = \arg\min_{\mathbf{c} \in C} wt(\mathbf{e} = \mathbf{r} + \mathbf{c})$$

Syndrome decoding is equal to ML decoding, if each coset leader has the largest probability of occurrence among all error patterns in a coset.

7.8 Error Probabilities

Given a BSC with transistion prob. p, an (n, k, d) linear block code, α_j the number of coset leaders with weight j, the error probability is given by:

$$P_e = 1 - \underbrace{\sum_{j=0}^{n} \alpha_j p^j (1-p)^{n-j}}_{\mathbf{P}(\text{error pat=coset leader})} \qquad \alpha_0 = 1$$

Error Detection: $2^n - 2^k$ error patterns are detectable. $2^k - 1$ undetectable error patterns (zero syndrome). Probability P_u of an undetectable error in BSC with p, (n, k, d) lin. block code, w_j number of code words with weight j, P_u is the prob. that an error pattern itself is a code word:

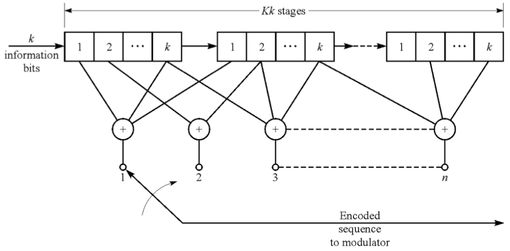
$$P_u = \sum_{j=1}^{n} w_j p^j (1-p)^{n-j}$$

8 Convolutional Codes

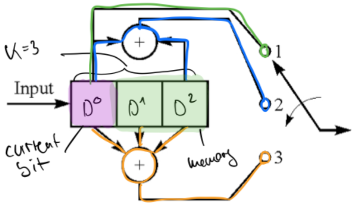
8.1 Convolutional Encoder

Consists of a shift register that shifts k-bit words at a time. Is a linear code. n Generators used to generate encoded sequence. For Nk input bits with known header and trailer ensures defined SR content and defines a $(n(N+K-1), Nk)$ lin. block code.

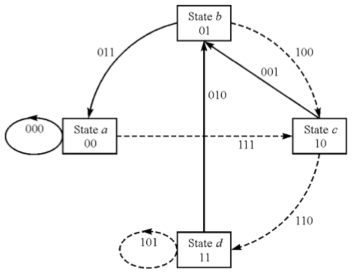
k stage length, mostly $k = 1$
K Contraint length: Number of k-bit stages
k/n Approximate code length
n Number of generators



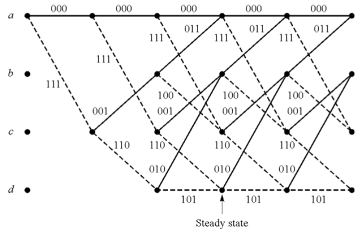
Example $(n = 3, k = 1, K = 3)$ $(3, 1, 3)$ conv. encoder. Rate $R_c = k/n = 1/3$. Systematic: Coded bit 1 is information bit. $g2 = [101] = 1 + 0 \cdot D + 1 \cdot D^2$



Corresponding **state diagram**. Solid arrows represent new message bit 0, dashed 1. Code word is denoted next to edges.



The **trellis diagram** denotes the states vertically and the progression of input message bits horizontally. Again solid (dashed) edges = message 0 (1) and code word aside edge.



Hamming weight: Weight of concatenated codewords of path through trellis
Free Hamming distance: Minimum weight of codewords.

Transfer function can be derived from state diagram. Transitions labeled with weight D^m where m is the Hamming weight of the associated coded bits. For each state there is an equilibrium condition:

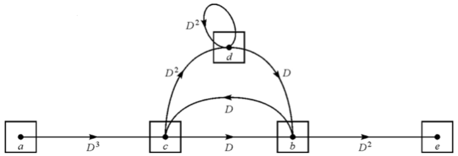
$$\begin{aligned} b: &\quad DX_c + DX_d - X_b = 0 \\ c: &\quad X_a D^3 + X_b D - X_c = 0 \\ d: &\quad D^2 X_d + D^2 X_c - X_d = 0 \\ e: &\quad D^2 X_b = X_e \end{aligned}$$

In Matrix form:

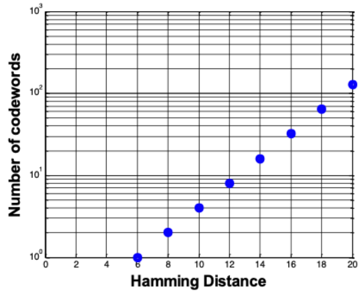
$$\begin{bmatrix} D & D & -1 & 0 \\ D^2-1 & D^2 & 0 & 0 \\ 0 & -1 & D & D^3 \\ 0 & 0 & D^2 & 0 \end{bmatrix} \begin{bmatrix} X_d \\ X_c \\ X_b \\ X_a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ X_e \end{bmatrix}$$

After gaussian elimination yields the transfer function

$$\begin{aligned} T(D) &= \frac{X_e}{X_a} = \frac{D^6}{1-2D^2} \\ &= D^6 + 2D^8 + 4D^{10} + 8D^{12} + L \end{aligned}$$



Polynomial division results in **distance distribution**. $8D^{12}$: 8 codewords with distance 12.



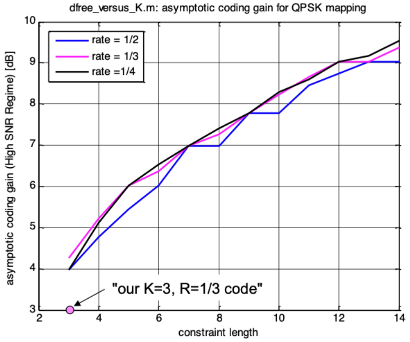
Error Probability: Pairwise prob. of error between two codewords with Hamming distance d_H:

$$P_2(d_H) = Q\left(\sqrt{\frac{2E_b}{N_0} R_c d_H}\right) \qquad E_c = E_b R_c$$

E_c Energy per coded bit

Coding Gain is asymptotic and determines performance in the high SNR regime.

$$G_a = 10\log_{10}(R_c d_{H,\text{free}})$$



8.2 Viterbi Decoder

Conventions
$m = K - 1$ memory depth
(n,k,m) codes comprise 2^{mk} states
u, v input/coded bits
$\mathbf{v^{(i)}}$ Encoded sequence
$\mathbf{r}$ Received sequence
$p(\mathbf{r}|\mathbf{v}^{(i)})$ ML metric of trial path

$$\begin{aligned} \mathbf{v^{(i)}} &= \left[\mathbf{v}_0^{(i)} \quad \dots \quad \mathbf{v}_{h+m-1}^{(i)}\right] \\ \mathbf{r} &= \left[\mathbf{r}_0 \quad \dots \quad \mathbf{r}_{h+m-1}\right] \\ p(\mathbf{r}|\mathbf{v}^{(i)}) &= \prod_k p(\mathbf{r}_k|\mathbf{v}_k^{(i)}) \end{aligned}$$

$M(\mathbf{r}|\mathbf{v}^{(i)})$ branch metric

$$\begin{aligned} M(\mathbf{r}|\mathbf{v}^{(i)}) \triangleq \log p(\mathbf{r}|\mathbf{v}^{(i)}) &= \sum_k \log p(\mathbf{r}_k|\mathbf{v}_k^{(i)}) \\ &= \sum_k M(\mathbf{r}_k|\mathbf{v}_k^{(i)}) \end{aligned}$$

BSC: Special case for binary sym. channel

$$\begin{aligned} \mathbf{P}\left(r_{l,j}=1 \mid v_{l,j}=1\right) &= \mathbf{P}\left(r_{l,j}=0 \mid v_{l,j}=0\right) = 1-p \\ \mathbf{P}\left(r_{l,j}=0 \mid v_{l,j}=1\right) &= \mathbf{P}\left(r_{l,j}=1 \mid v_{l,j}=0\right) = p \end{aligned}$$

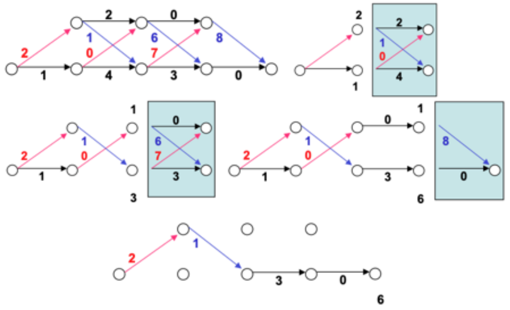
AWGN: Special case for AWGN channel

$$\log \mathbf{P}\left(r_{l,j}=a \mid v_{l,j}=b\right) = -|a-b|^2$$

Algorithm

1. Begin at time $t = m$, compute partial metric for single path entering each state. Store path (survivor) and its metric for each state.
2. Increase t by 1
   Branch metric: Compute branch metric for all 2^k paths entering a state
   Add: Compute the partial path metrics by adding the branch metric entering that state to the metric of the connecting survivor at the previous time
   Compare the partial metrics for each state for all 2^k paths entering that state
   Select the path with best metric (survivor), store it with its metric and eliminate all other paths
3. Iterate until $t = h+m+1$

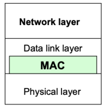
Example



9 Multiple Access Protocols

9.1 Basics of Channel Access

In broadcast networks, one channel is shared by multiple users, therefore coordination is required. The medium access control (MAC) subplayer controls access to the shared channel.



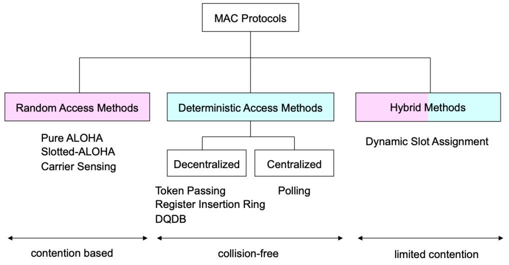
Static channel allocation: Frequency- (FDM), Time- (TDM), Code- (CDM), Space-Division (SDM) multiplexing.

Dynamic channel allocation: Time variable traffic from different sources. Collided frames will be retransmitted and can be avoided using suitable coordination among stations.

Assumptions:

- Station Model: N independent stations generating traffic.
- A single channel is available for all transmissions
- Collision: If two frames overlap in time the resulting signal is distorted.
- Cont. Time Alloc: Tx can be performed in any time instant
- Slotted Time: Time is divided into discrete slots. Tx begins at start of slot.
- Carrier Sense: Stations listen to channel before Tx

9.2 MAC Protocol Classification



9.3 ALOHA Family Protocols

Data can be sent at any time. If collision, a random waiting time is passed and the data is retransmitted.

D transfer time / frame length $[s]$
g offered load $[frames/s]$
G offered load
S throughput $[frames]$ per frame duration
P_0 prob. of successful transmission

$$G = gD \qquad S = GP_0$$

Slotted ALOHA

$$P_0 = \mathrm{e}^{-G} \qquad S = G\mathrm{e}^{-G} \leq \frac{1}{\mathrm{e}} \quad \text{eqty with} G = 1$$

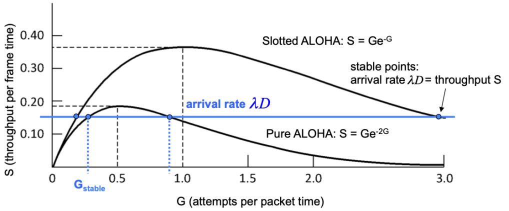
Unslotted (Pure) ALOHA

$$P_0 = \mathrm{e}^{-2G} \qquad S = G\mathrm{e}^{-2G} \leq \frac{1}{2\mathrm{e}} \quad \text{eqty with} G = 0.5$$

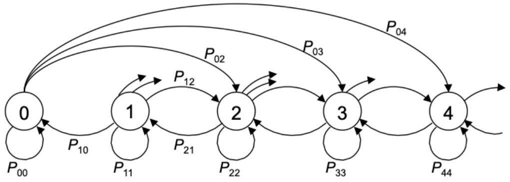
Prob. that k frames are transmittied during time T with large number of stations is poisson with $\lambda = gT, m_k = \sigma_k^2 = gT$:

$$P_0(k|T) = \frac{(gT)^k \mathrm{e}^{-gT}}{k!} \qquad T = \begin{cases} D, \text{slotted}, \\ 2D, \text{unslotted}. \end{cases}$$

$$P_0(k=0|T) = \mathrm{e}^{-gT} \qquad P_0(k=1|T) = gT\mathrm{e}^{-gT}$$



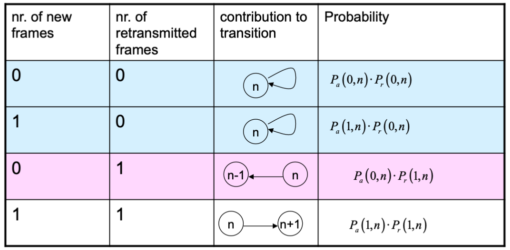
Markov Chain Model For low number of stations $N < 10$ a markov chain can be used for calculating probabilities. A station is backlogged if it encountered a collision during Tx and has to retransmit. The MC state represents the number of backlogged stations.



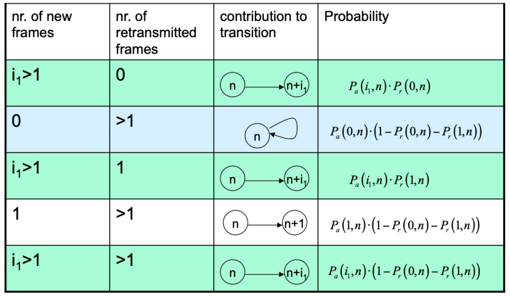
$$P = \begin{bmatrix} P_{0,0} & \dots & P_{0,m} \\ \vdots & \ddots & \vdots \\ P_{m,0} & \dots & P_{m,m} \end{bmatrix} \qquad \mathbf{p}_{j+1} = P^\top \mathbf{p}_j$$

$$j \to \infty \quad \mathbf{p}_{j+1} = \mathbf{p}_j = P^\top \mathbf{p}_j$$

Where $j \to \infty$ denotes the steady state that can be calculated using a eigenvalue problem.



nr. of new frames	nr. of retransmitted frames	contribution to transition	Probability
0	0	n ↺	$P_a(0,n) \cdot P_r(0,n)$
1	0	n ↺	$P_a(1,n) \cdot P_r(0,n)$
0	1	n-1 ← n	$P_a(0,n) \cdot P_r(1,n)$
1	1	n → n+1	$P_a(1,n) \cdot P_r(1,n)$

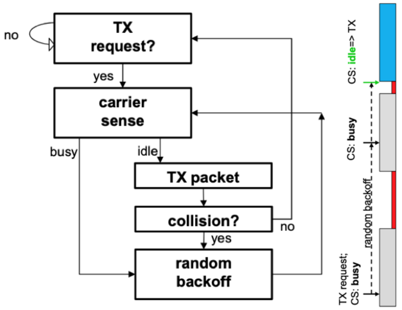


nr. of new frames	nr. of retransmitted frames	contribution to transition	Probability
i_1>1	0	n → n+i_1	$P_a(i_1,n) \cdot P_r(0,n)$
0	>1	n ↺	$P_a(0,n) \cdot (1 - P_r(0,n) - P_r(1,n))$
i_1>1	1	n → n+i_1	$P_a(i_1,n) \cdot P_r(1,n)$
1	>1	n → n+1	$P_a(1,n) \cdot (1 - P_r(0,n) - P_r(1,n))$
i_1>1	>1	n → n+i_1	$P_a(i_1,n) \cdot (1 - P_r(0,n) - P_r(1,n))$

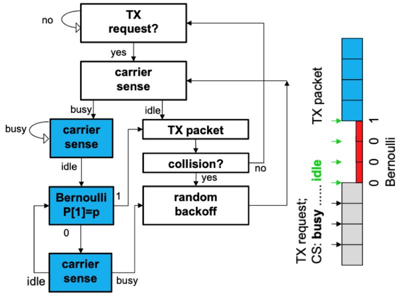
9.4 CSMA Carrier Sense Multiple Access

Listen for a ongoing transmission and act accordingly.

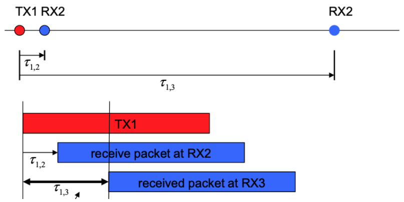
Nonpersistent Wait random time after check channel again and loop. May not need to continually sense the channel.



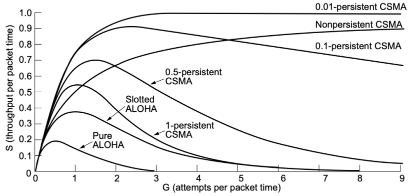
p-persistent Wait for idle, Tx with probability p, repeat until Tx or other Tx then start again. 1-persistent has $p = 1$ and sends immediately as soon as channel is free.



Vulnerable Period

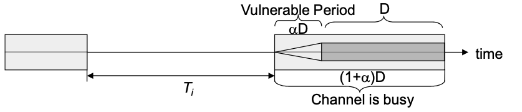


Channel Utilization



Throughput Many stations $m \to \infty$, transmission with fixed duration D, channel access Poisson distributed with rate g. Normalized offered load $G = gD$. Possible channel states: Idle, successful Tx, collision.

$\tau_{\max} = \alpha D$ Worst case vulnerable period
$\tau_{\max}$ Max. propagation delay
$P(k|T)$ Prob. of k frames in T
$D + \tau_{\max} = (1+\alpha)D$ duration of channel access without collision



$$p_T(T_i) = g\exp(-gT_i) \qquad E[T] = \int_0^\infty tp_T(t)\,\mathrm{d}t = 1/g$$

Prob. that no other packet is Tx in vulnerable period:

$$P(\text{success}) = \mathrm{e}^{-\alpha G}$$

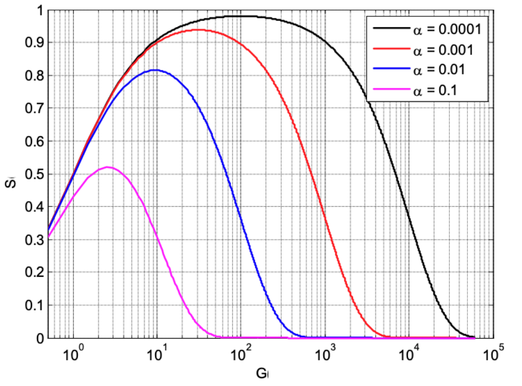
Mean duration of a Tx cycle: $T + (1+\alpha)D$

$$s \approx \frac{P(\text{success})}{T + (1+\alpha)D} = \frac{\mathrm{e}^{-\alpha G}}{T + (1+\alpha)D}$$

Normalized throughput per packet duration D

$$S = sD = \frac{\mathrm{e}^{-\alpha G}}{\frac{1}{G} + 1 + \alpha}$$

$$\lim_{\alpha \to 0} S = \frac{G}{G+1} \qquad \lim_{G \to \infty} S = 0$$



9.5 CSMA/CD collision detect

Detect a collision and stop Tx.

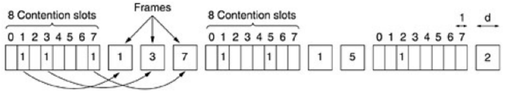
9.6 Binary exponential backoff

Based on CSMA/CD, three channel states: idle, contention, success.

- After collision, time is divided into discrete slots: length of each slot is equal to worst-case round-trip propagaion time 2τ
- After first collision, each station waits either 0 or 1 slo times before trying again
- After each further collision the backoff window is doubled (up to max 1024)
- In general after i collisions, a random number between 0 and $2^i - 1$ is chosen, and that number of slots is skipped
- after 16 collisinos, the controller reports failure to higher layer

9.7 Collision-Free Protocols

Bit-Map Protocol: Contention slots and frames. Each station wanting to send, transmits 1 during contention slots. Frames can then be sent in order of address. Addresses can be rotated to prevent starvation.



9.8 Limited-Contention Protocols

Idea: Use contention at low load (to provide low delay) but use a collision-free technique at high load (to provide good channel efficiency).

Assumptions:

- We allow k stations to contend for channel access
- each station has a probbility p of transmitting during each slot

Probability of successful transmission is $\mathbf{P}(\text{success}) = kp(1-p)^{k-1}$. Optimum value for $k = 1/p$.

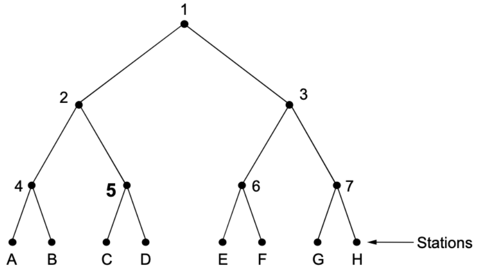
$$\mathbf{P}(\text{Success with opt. } k \mid p) = (1-p)^{1/p-1}$$

$$\mathbf{P}(\text{Success with opt. } p \mid k) = \left(\frac{k-1}{k}\right)^{k-1}$$

Adaptive Tree Walk Protocol

- First, all stations are allowed to tx
- If collision, during slot 1 only stations under node 2 may compete
- If one acquires channel, the slot following the frame is reserved for those statinos under node 3. If collision again, go down to 4

The heavier the load, the farther down the tree the seach should begin



10 Math

10.1 General

$$\cos(a)\cos(b)+\sin(a)\sin(b)=\cos(a-b)$$

$$\cos(a+b)\cos(a-b)=\frac{1}{2}\left[\cos(2a)+\cos(2b)\right]$$

$$\cos(a)\cos(b)=\frac{1}{2}(\cos(a-b)+\cos(a+b))$$

$$\sin(a)\sin(b)=\frac{1}{2}(\cos(a-b)-\cos(a+b))$$

$$\cos(a)\sin(b)=\frac{1}{2}(\sin(a+b)-\sin(a-b))$$

$$\operatorname{sinc}(x)=\frac{\sin(\pi x)}{\pi x}$$

$$\sin(x)=\frac{e^{ix}-e^{-ix}}{2i} \qquad \cos(x)=\frac{e^{ix}+e^{-ix}}{2}$$

10.2 Fourier Transform

Source: Haykin, Communication systems, 4th ed.

$\operatorname{rect}\left(\frac{t}{T}\right)$	$\circ\!\!-\!\!\bullet$	$T\operatorname{sinc}(fT)$
$\operatorname{sinc}(2Wt)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{2W}\operatorname{rect}\left(\frac{f}{2W}\right)$
$\exp(-at)u(t), a>0$	$\circ\!\!-\!\!\bullet$	$\frac{1}{a+j2\pi f}$
$\exp(-a\lvert t\rvert), a>0$	$\circ\!\!-\!\!\bullet$	$\frac{2a}{a^2+(2\pi f)^2}$
$\exp(-\pi t^2)$	$\circ\!\!-\!\!\bullet$	$\exp(-\pi f^2)$
$\begin{cases}1-\frac{\lvert t\rvert}{T}, & \lvert t\rvert<T\\ 0, & \lvert t\rvert\geq T\end{cases}$	$\circ\!\!-\!\!\bullet$	$T\operatorname{sinc}^2(fT)$
$\delta(t)$	$\circ\!\!-\!\!\bullet$	1
1	$\circ\!\!-\!\!\bullet$	$\delta(f)$
$\delta(t-t_0)$	$\circ\!\!-\!\!\bullet$	$\exp(-j2\pi ft_0)$
$\exp(j2\pi f_c t)$	$\circ\!\!-\!\!\bullet$	$\delta(f-f_c)$
$\cos(2\pi f_c t)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{2}[\delta(f-f_c)+\delta(f+f_c)]$
$\sin(2\pi f_c t)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{2i}[\delta(f-f_c)+\delta(f+f_c)]$
$\operatorname{sgn}(t)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{j\pi f}$
$\frac{1}{\pi t}$	$\circ\!\!-\!\!\bullet$	$-j\operatorname{sgn}(f)$
$u(t)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{2}\delta(f)+\frac{1}{j2\pi f}$
$\sum_{i=-\infty}^{\infty}\delta(t-iT_0)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{T_0}\sum_{n=-\infty}^{\infty}\delta\left(f-\frac{n}{T_0}\right)$

$u(t)$ unit step function
$\delta(t)$ delta function
$\operatorname{rect}(t)$ rectangular function of unit amplitude and unit duration centered on the origin
$\operatorname{sgn}(t)$ signum function
$\operatorname{sinc}(t)$ sinc function

Relations

$\alpha f(t)+\beta g(t)$	$\circ\!\!-\!\!\bullet$	$\alpha F(f)+\beta G(f)$
$f^*(t)$	$\circ\!\!-\!\!\bullet$	$F^*(-f)$
$f(at)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{\lvert a\rvert}F\left(\frac{f}{a}\right)$
$f(t-a)$	$\circ\!\!-\!\!\bullet$	$\mathrm{e}^{-j2\pi fa}F(f)$
$e^{j2\pi f_0 f}$	$\circ\!\!-\!\!\bullet$	$F(f-f_0)$
$f^{(n)}$	$\circ\!\!-\!\!\bullet$	$(j2\pi f)^n F(f)$
$t^n f(t)$	$\circ\!\!-\!\!\bullet$	$j^n F^{(n)}(f)$
$\int_{-\infty}^{t}x(\tau)\,\mathrm{d}\tau$	$\circ\!\!-\!\!\bullet$	$\frac{1}{j2\pi f}F(f)+\pi F(0)\delta(f)$
$\frac{1}{t}x(t)+\pi x(0)\delta(t)$	$\circ\!\!-\!\!\bullet$	$\int_{-\infty}^{f}X(s)\,\mathrm{d}s$
$(f*g)(t)$	$\circ\!\!-\!\!\bullet$	$F(f)\cdot G(f)$
$f(t)\cdot g(t)$	$\circ\!\!-\!\!\bullet$	$\frac{1}{2\pi}F(f)*G(f)$

$f^{(n)}$ n^{th} derivation
f^* complex conjugate

10.3 Sums

$$\sum_{k=0}^{n}q^k k = \frac{nq^{n+2}-(n+1)q^{n+1}+q}{(q-1)^2}$$

$$\sum_{k=0}^{n}k = \frac{n(n+1)}{2}$$

$$\sum_{k=0}^{n}k^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{k=0}^{n}k^3 = \frac{n^2(n+1)^2}{4}$$

$$\sum_{k=0}^{\infty}q^k = \frac{1}{1-q}$$

$$\sum_{k=1}^{\infty}q^k = \frac{q}{1-q}$$

10.4 Probability

$$\mathbf{P}\left(X>x\right)=\frac{1}{2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}\sigma_X}\right)$$

$$\operatorname{erf}(x)=\frac{1}{\sqrt{\pi}}\int_{-x}^{x}\mathrm{e}^{-t^2}\,\mathrm{d}t=\frac{2}{\sqrt{\pi}}\int_{0}^{x}\mathrm{e}^{-t^2}\,\mathrm{d}t$$

$$\operatorname{erfc}(x)=1-\operatorname{erf}x=2Q(\sqrt{2}x)=\frac{2}{\sqrt{\pi}}\int_{x}^{\infty}\mathrm{e}^{-t^2}\,\mathrm{d}t$$

$$Q(x)=\frac{1}{\sqrt{2\pi}}\int_{x}^{\infty}\exp(-y^2/2)\,\mathrm{d}y\leq\exp(-x^2/2)$$

$$Q(x)=\frac{1}{2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right)$$