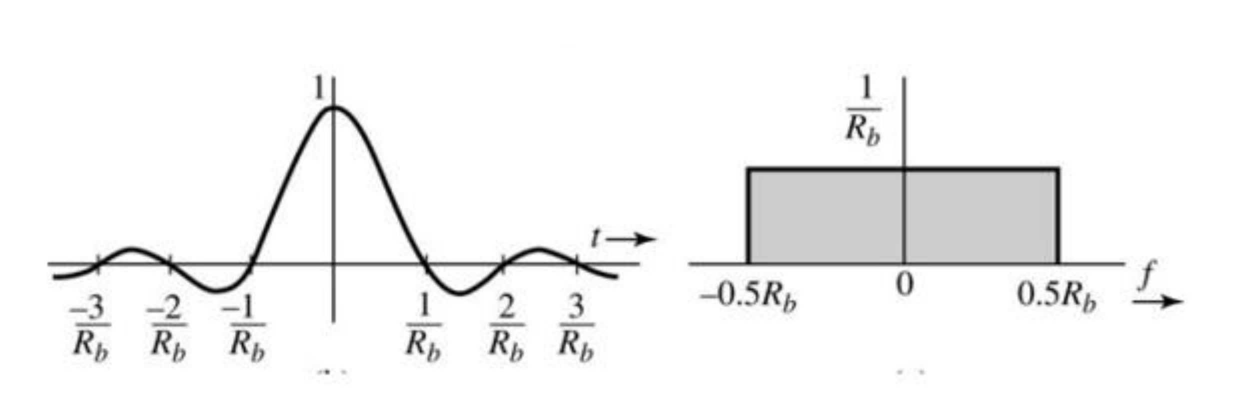
LECTURE 7: MATCHED FILTERING



ECEN 310

Communication Engineering

VICTORIA UNIVERSITY OF WELLINGTON TE HERENGA WAKA

Digital Baseband Tansmission

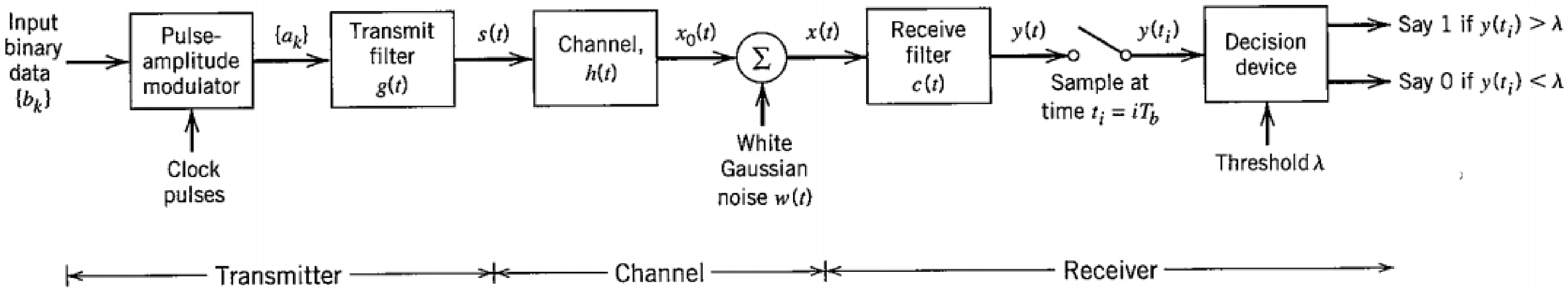
Read Proakis and Salehi

Overview

- Previously we talked about the design of the *overall* pulse $p(t)$ (ie that just prior to sampling - combination of transmit filter, channel, and receive filter responses)
- We established the zero ISI criterion for $p(t)$ - Nyquist Zero ISI Criterion
- We will now look at the effects of noise.
- We will consider in detail the design of the receive filter, specifically with the aim of minimizing the probability of error.
- This will lead to the second fundamental result in communications theory - the *matched filter receiver*

System model - review

Binary PAM

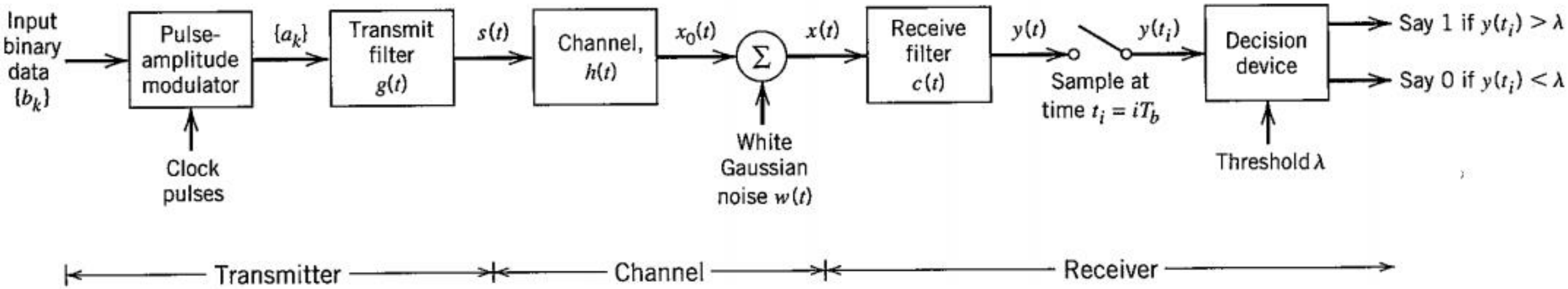


- Consider a generic *binary* PAM, where the amplitude of the pulses is varied in *discrete* manner according to the digital input data $\{b_k\}$
- The incoming binary sequence $\{b_k\}$ consists of 1's and 0's arriving every T_b seconds; these are mapped onto a symbols $\{a_k\}$; let us assume the mapping of

$$a_k = \begin{cases} +1 & \text{if } b_k = 1 \\ -1 & \text{if } b_k = 0 \end{cases}$$

- These symbols 'amplitude modulate' a *transmit filter* $g(t) \rightleftharpoons G(f)$
- Note that these two steps are analogous to line encoding (which assumed rectangular pulses), but with a pulse shape $g(t)$

Binary PAM



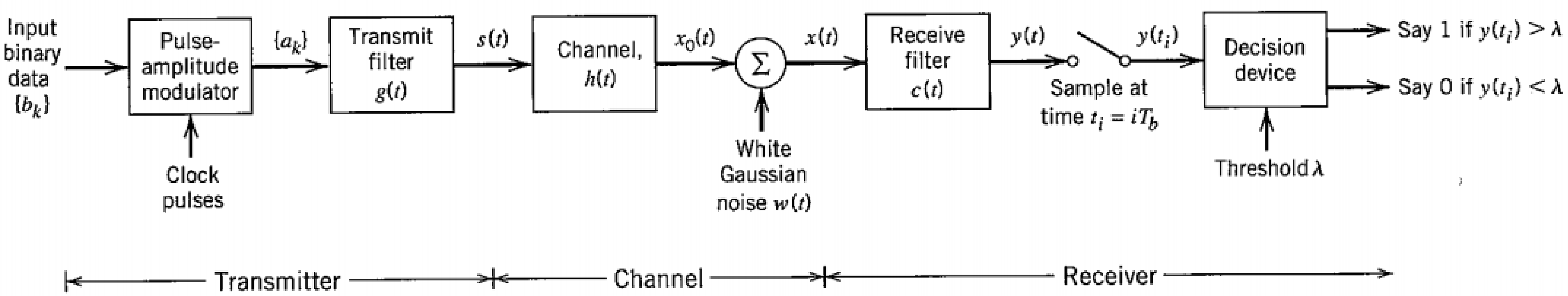
- The output of the transmit filter is a sequence of pulses

$$s(t) = \sum_{k=-\infty}^{\infty} a_k g(t - kT_b)$$

- The signal $s(t)$ is then transmitted over a *linear* communication channel, described by $h(t) \rightleftharpoons H(f)$
- Ignoring noise (for now), the channel output is the convolution

$$x(t) = s(t) * h(t)$$

Binary PAM

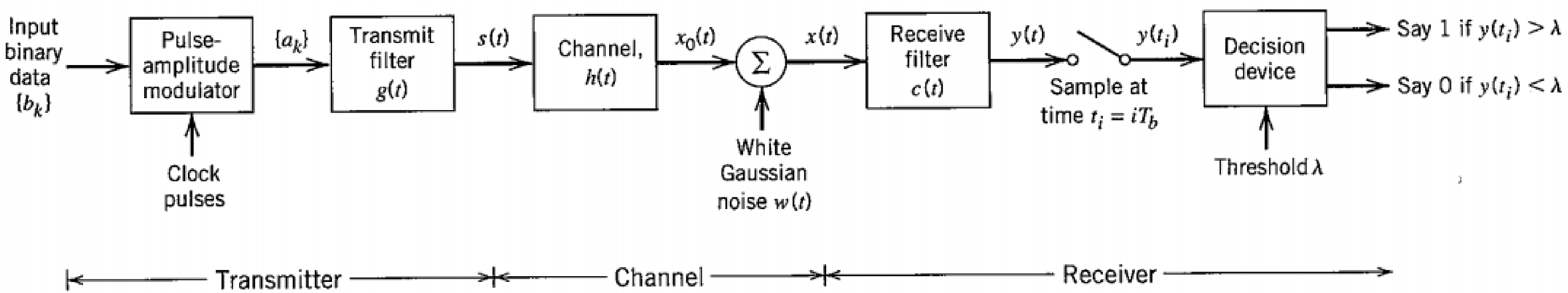


- At the receiver, channel output $x(t)$ is passed through a *receive filter*, given by $c(t) \rightleftharpoons C(f)$, resulting in

$$y(t) = x(t) * c(t)$$

- The filter output $y(t)$ is sampled synchronously - the timing instants being determined by a *timing synchronisation unit*
- The samples $y(t_k)$ are interpreted by means of *decision device* - it determines the best estimate $\hat{a}_k$ of the data symbol a_k, which is them mapped back to a bit value (1 or 0).
- Typically, this is done by comparing the sample value to a threshold: if the threshold is exceeded, the symbol is assumed to be a 1, otherwise it is interpreted as a 0

Binary PAM



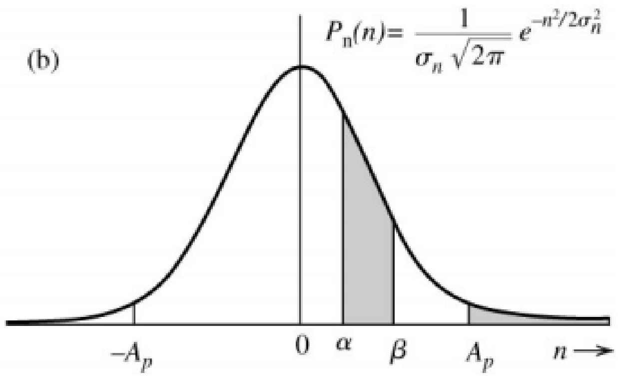
- NOTE 1: The effects of the channel are typically compensated for at the receiver by an *equaliser* (a transversal filter, just prior to the Decision Device), such that the equalised channel response is $\delta(t)$ (no distortion);
- We will discuss equalisation later on in the course, but for now we will assume that the channel has been equalised;
- NOTE 2: In his book, Lathi does not consider receive filtering in Chapter 7; Instead the discussion is simply in terms of a an overall pulse shape, which he denotes $p(t)$.
- Assuming an equalised channel, the overall pulse in the above system description is $p(t) = g(t) * c(t)$

AWGN

Detection Errors

- Typically, the noise in the system is assumed to have a normal (Gaussian) distribution

$$P_n(n) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-n^2/2\sigma_n^2}$$



- The assumption of Gaussian noise comes from the *central limit theorem* - that is, for a large number of sources of noise with different distributions, the sum is a normal distribution.

Additive White Gaussian Noise

- The primary spectral characteristic of thermal noise is that its power spectral density $S(f)$ is *constant* for all frequencies of interest in a communication system

$$S_n(f) = \frac{N_0}{2}$$

  - the term *white* is used to indicate presence in all frequencies, as in white light containing equal amounts of frequencies in entire visible spectrum
  - the constant N_0 has units of watts/Hz
  - the factor of 2 is included to indicate that $S(f)$ is a *two-sided* power spectral density

- The autocorrelation function is the inverse Fourier Transform of $S_n(f)$, and thus

$$R_n(\tau) = \frac{N_0}{2}\delta(\tau)$$

AWGN

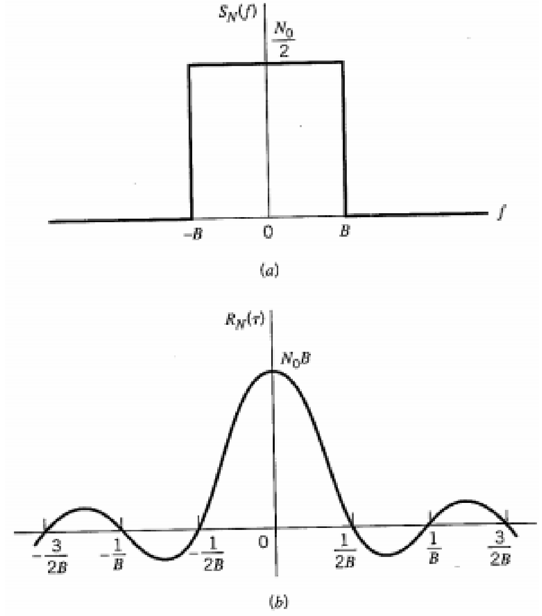
- Note that a Gaussian process $X(t)$ applied to a stable linear filter $H(f)$ results in a random process $Y(t)$ which is also Gaussian
- The spectral power density of $Y(t)$ is

$$S_Y(f) = |H(f)|^2 S_X(f)$$

Filtered AWGN

- Consider white Gaussian noise $n(t)$ with power spectral density $N_0/2$ passing through an ideal low-pass filter of bandwidth B and unity passband amplitude.

$$S_N(f) = \begin{cases} \frac{N_0}{2} & |f| < B \\ 0 & |f| > B \end{cases}$$



the autocorrelation function of $n(t)$ is

$$R_N(f) = \int_{-B}^{B} \frac{N_0}{2} e^{j2\pi f\tau} df = N_0 B \operatorname{sinc}(2B\tau)$$

- note that the autocorrelation function has zero crossings at $\tau = \pm k/2B$
- thus if $n(t)$ is sampled at $2B$ samples/sec the samples are uncorrelated, and being Gaussian, also statistically independent

SNR and Eb/No

- The basic figure of merit in analogue communications is *average signal power to average noise power ratio*: S/N (SNR)
- In digital communications, we typically use a scaled version of SNR, namely E_b/N_0
  - E_b is the energy per bit $= ST_b$ where T_b is the bit (pulse) duration
  - N_0 is the noise power divided by the bandwidth W : $N_0 = N/W$
- Denoting the bit *rate* by $R = 1/T_b$, we have

$$\frac{E_b}{N_0} = \frac{ST_b}{N/W} = \frac{S}{N}\left(\frac{W}{R}\right)$$

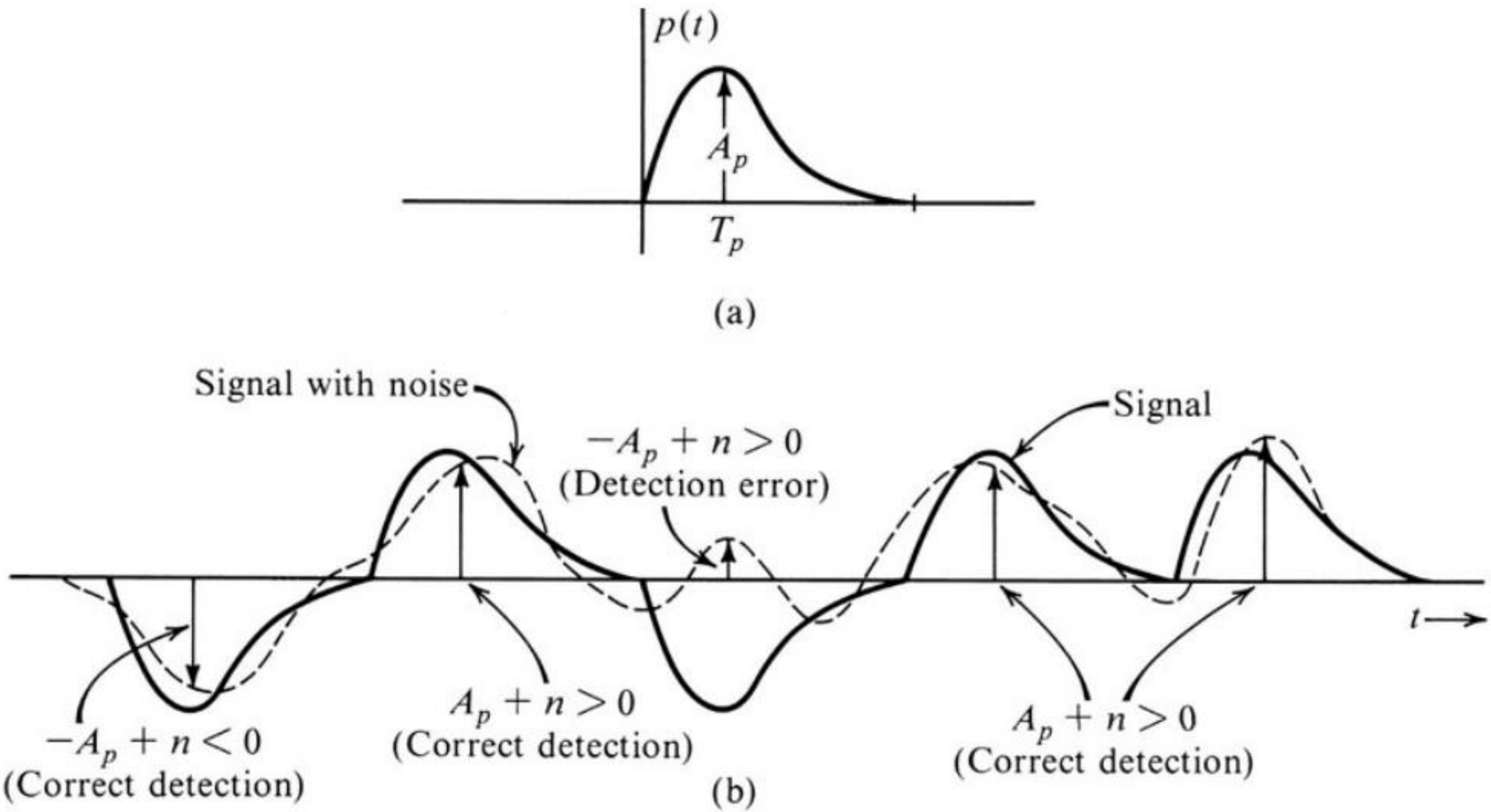
- Thus E_b/N_0 is SNR normalized by the bandwidth and bit rate
- One of the key performance metrics in digital communications is a plot of bit-error rate P_e versus E_b/N_0
- It is typically plotted as $\log P_e$ vs $10\log_{10}(E_b/N_0)$, and due to its shape is referred to as *waterfall* curve

Error Performance

Detection Errors

- The received pulse, assuming binary transmission, is $\pm p(t) + n(t)$
- At each sampling instant, sample value is $y_i = \pm A_p + n$ where we let $A_p = p(nT)$ is the pulse amplitude at the time of sampling
- The decision device must estimate the transmitted symbol $\hat{a}_k$ based on the sample value
    - Decide $b_k = 1$, (ie $a_k = 1$) if $y_i > 0$
    - Decide $b_k = 0$, (ie $a_k = -1$) if $y_i < 0$

Detection Errors



Detection Errors

- The decision error will occur if either
  - $b_k = 1$ was sent but $y_i = A_p + n < 0$
  - $b_k = 0$ was sent but $y_i = -A_p + n > 0$
- In the second case, the probability of error is

$$P(\epsilon | b_k = 0) = P(n > A_p) = \frac{1}{\sqrt{2\pi}\sigma} \int_{A_p}^{\infty} e^{-y^2/2\sigma^2} \, dy = Q\left(\frac{A_p}{\sigma}\right)$$

- By symmetry, the error given 1 is

$$P(\epsilon | b_k = 1) = P(n < -A_p) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{-A_p} e^{-y^2/2\sigma^2} \, dy = Q\left(\frac{A_p}{\sigma}\right)$$

- In the above we performed change of variable $\tilde{y} = y/\sigma$ to obtain the standard form for the Q function

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-y^2/2} dy$$

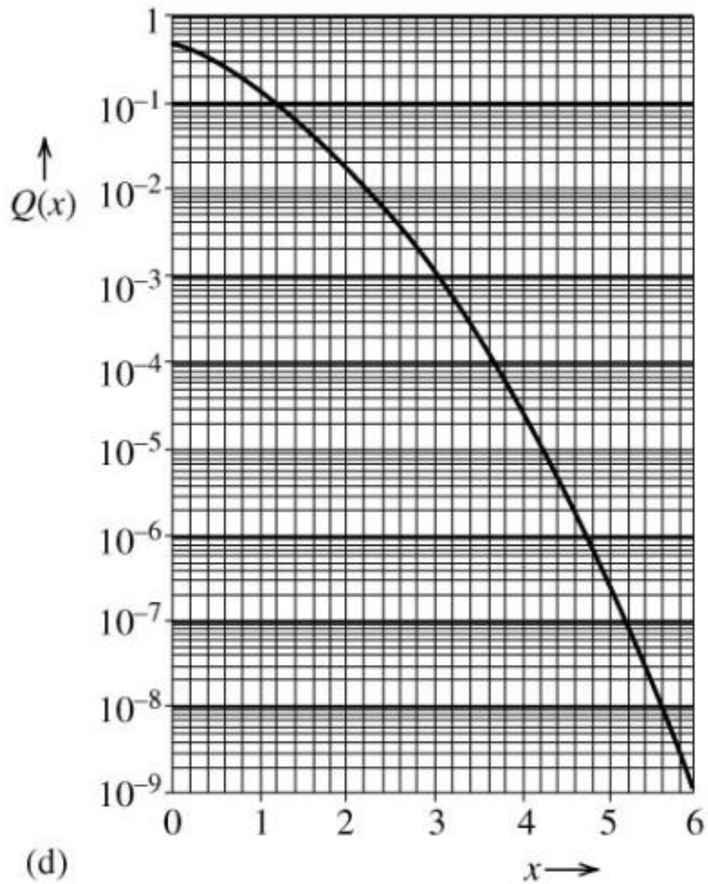
Detection Errors

- Thus, the total probability error is the average of the two

$$P(\epsilon) = P(\epsilon|b_k = 1)P(b_k = 1) + P(\epsilon|b_k = 0)P(b_k = 0)$$
$$= 2\frac{1}{2}Q\left(\frac{A_p}{\sigma}\right) = Q\left(\frac{A_p}{\sigma}\right)$$

where we assumed equal probability of the input bits $P(b_k = 0) = P(b_k = 1) = \frac{1}{2}$

Detection Errors



(d)

Matched Filter

- Let $\rho = \frac{A_p}{\sigma}$, and so $P(\epsilon) = Q\left(\frac{A_p}{\sigma}\right) = Q(\rho)$
- Because $Q(\rho)$ decreases monotonically with ρ, we can minimize the probability of error P_e by maximizing ρ
- In doing so, we will look at the optimum receive filter $c(t) \leftrightharpoons C(f)$
- We will see that the ρ is maximized if $c(t)$ is *matched* to the *incoming* pulse shape $p'(t) = g(t) * h(t)$ (that is the combination of transmit filter and the channel)

$$c(t) = k'p'(t_m - t)$$

where t_m is the sampling instant.

Matched Filter

- Recalling that $p(t)$ is the pulse shape prior to sampling, we will seek $C(f)$ which maximizes
$$\rho^2 = \frac{p^2(t_m)}{\sigma^2}$$
which *represents the SNR at the sampling time* t_m.
- The pulse shape $p(t)$ is the inverse FT
$$p(t) = \int_{-\infty}^{\infty} P'(f)C(f)e^{j2\pi ft}df$$
- The instantaneous power of the filter output as
$$|p(t)|^2 = \left| \int_{-\infty}^{\infty} P'(f)C(f)e^{j2\pi ft_m}df \right|^2$$

Matched Filter

- Now lets consider the average noise power after filtering
- If the power spectral density of $n(t)$ is $S_N(f)$ ($= \frac{N_0}{2}$ if white), the then filtered noise has a power spectral density

$$S_{N'}(t) = S_N(f)|C(f)|^2$$

- The average power of the filtered noise is thus

$$\sigma^2 = \int_{-\infty}^{\infty} S_{N'}(f)\ df = \int_{-\infty}^{\infty} S_n(f)|C(f)|^2\ df$$

- putting it all together, we have the signal to noise ratio

$$\rho^2 = \frac{\left|\int_{-\infty}^{\infty} P'(f)C(f)e^{j2\pi ft}df\right|^2}{\int_{-\infty}^{\infty} S_n(f)|C(f)|^2 df} \tag{1}$$

Proof: Cauchy-Schwartz's Inequality

Theorem (Cauchy-Schwartz's Inequality)

For two complex functions $\phi_1(x)$ and $\phi_2(x)$ in the real variable x, satisfying the condition

$$\int_{-\infty}^{\infty} |\phi_1(x)|^2 dx < \infty \text{ and } \int_{-\infty}^{\infty} |\phi_2(x)|^2 dx < \infty$$

we have that

$$\left| \int_{-\infty}^{\infty} \phi_1(x)\phi_2(x) dx \right|^2 \leq \int_{-\infty}^{\infty} |\phi_1(x)|^2 dx \int_{-\infty}^{\infty} |\phi_2(x)|^2 dx$$

with equality holding if, and only if,

$$\phi_1(x) = k\phi_2^*(x)$$

Proof: Matched Filter

- Invoking Cauchy-Schwartz's Inequality on (1), with

$$\phi_1(x) = X(f) = C(f)\sqrt{S_n(f)} \qquad \phi_2(x) = Y(f) = \frac{P'(f)e^{j2\pi ft}}{\sqrt{S_n(f)}}$$

we have

$$\begin{aligned}
\rho^2 &= \frac{\int_{-\infty}^{\infty} |X(f)Y(f)\, df|^2}{\int_{-\infty}^{\infty} |X(f)|^2\, df} \\
&\le \frac{\int_{-\infty}^{\infty} |X(f)|^2\, df \int_{-\infty}^{\infty} |Y(f)|^2\, df}{\int_{-\infty}^{\infty} |X(f)|^2\, df} \\
&= \int_{-\infty}^{\infty} |Y(f)|^2\, df = \frac{\int_{-\infty}^{\infty} |P'(f)|^2\, df}{\int_{-\infty}^{\infty} S_n(f)\, df} && (2) \\
& && (3)
\end{aligned}$$

with equality iff $X(f) = k\,[Y(f)]^*$ or

$$C(f)\sqrt{S_n(f)} = k\left[\frac{kP'(f)e^{j2\pi fT}}{\sqrt{S_n(f)}}\right]^* = \frac{kP(-f)e^{-j2\pi fT}}{\sqrt{S_n(f)}}$$

Matched Filter

- We have shown that SNR ρ^2 is maximized iff

$$C(f) = \frac{kP'(-f)e^{-j2\pi fT}}{S_n(f)}$$

- This optimum receive filter is referred to as the *matched filter*
- It depends on
  - noise PSD $S_n(f)$
  - sampling time t_m
  - incoming pulse shape $P(f)$
- In the case of white noise $S_n(f) = \frac{N_0}{2}$, and matched filtering, from (2) we have the optimum SNR

$$\rho^2 = \frac{2}{N_0} \int_{-\infty}^{\infty} |P(f)|^2 \, df = \frac{2E_p}{N_0}$$

  where E_p is the energy of the pulse $p(t)$.
- The matched filter is then simply

$$C(f) = k' P'(-f) e^{-j2\pi f t_m}$$

Matched Filter

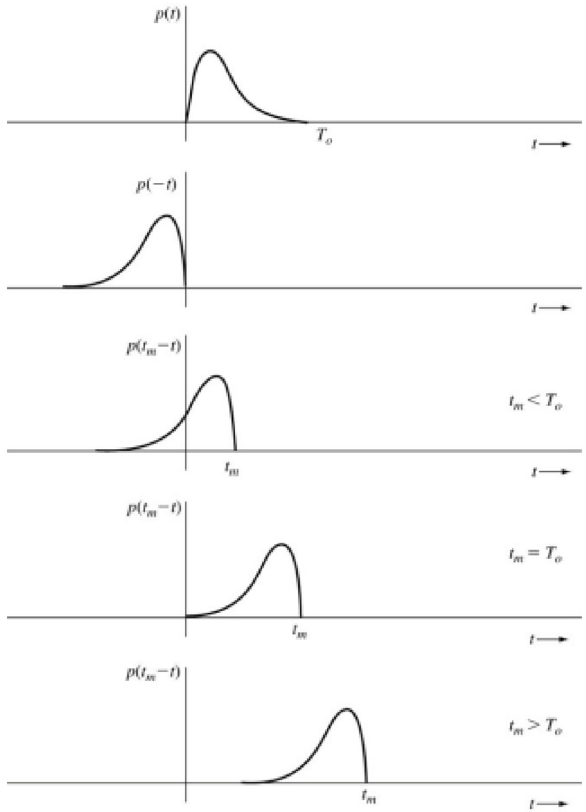
- Taking the inverse FT we obtain the impulse response

$$c(t) = \mathcal{F}^{-1}\left\{k' p(-f) e^{-j2\pi f t_m}\right\}$$
$$= k' p'(t_m - t)$$

- In order to ensure a causal filter with the minimum possible delay, we let $t_m = T_0$

$$c(t) = p'(T_0 - t)$$

$$C(f) = P'(-f) = e^{-j2\pi f T_0}$$



Note: in reality, the $C(f)$ would be matched to the transmit filter $G(f)$, and an equalizer would be used to compensate for the effects of the channel $H(f)$.

Matched Filter

- Returning to the probability of error $P_e = Q(\rho)$, we now have for the matched filter, where $\rho_{\mathsf{max}} = \sqrt{\frac{2E_p}{N_0}}$, and so

$$P_e = Q(\rho_{\mathsf{max}}) = Q\left(\sqrt{\frac{2E_p}{N_0}}\right)$$

- The system error performance is thus a function of the SNR, regardless of the pulse shape used (provided no ISI and ideal sampling instants)