EE521 Analog and Digital Communications

February 15, 2006

Instructor: James K Beard, PhD
Office: Ft. Washington 115
Email:. jkbeard@temple.edu, jkbeard@comcast.net
Office Hours: Wednesdays 5:00 PM to 6:00PM
Location: Ft. Washington 107
Time: Wednesdays 6:00 PM – 8:30 PM
Web Page: http://temple.jkbeard.com

Texts:

- Bernard Sklar, Digital Communications, Second Edition, Prentice Hall P T R, 2001 (2004 printing), ISBN 0-13-084788-7
- Digital Communication Systems Using SystemVue, by Dr. Silage, ISBN 1-58-450850-7

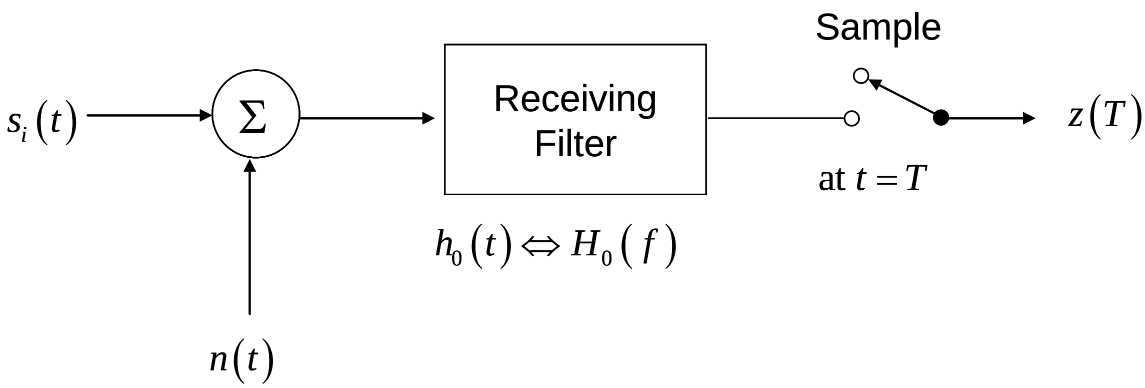
Today's Topics

- Matched Filters
- Homework
- Examples
- From Sklar Chapter 4, Bandpass Modulation and Demodulation
    - o Why Modulate
    - o Digital Bandpass Modulation Techniques
    - o Detection of Signals in Gaussian Noise
    - o Coherent Detection
    - o Non-coherent Detection
    - o Complex Envelope
    - o Error Performance for Binary Systems
- Assignment
- SystemView assignment

Matched Filters

Matched filters are a topic in both Chapter 3 (section 3.2.2, pp 122- 125) and Chapter 4 (section 4.4.2, pp. 184-188). Chapter 3 treats the analog matched filter and Chapter 4 treats the digital matched filter. We started the analog matched filter last time.

The Analog Matched Filter



The sampler output consists of signal component a_i and a noise component n_0 that has variance σ_0^2. The ratio of signal power to noise power at time T is

$$\left(\frac{S}{N}\right)_T = SNR(T) = \frac{a_i^2}{\sigma_0^2}$$

The objective is to find the filter transfer function $H_0(f)$ that maximizes the SNR.

We begin by finding the signal and noise powers. The signal power is given in terms of the inverse Fourier transform of the output as

$$a_i(t) = \int_{-\infty}^{\infty} H(f) \cdot S_i(f) \cdot \exp(+j \cdot 2\pi \cdot \cdot t) \cdot df$$

The noise output power is given in terms of an input noise power spectral density, two-sided, of $N_0/2$ as

$$\sigma_0^2 = \frac{N_0}{2} \cdot \int_{-\infty}^{\infty} |H(f)|^2 \cdot df$$

At a given time T, the SNR is given by

$$\left(\frac{S}{N}\right)_T = \frac{\left| \int_{-\infty}^{\infty} H(f) \cdot S_i(f) \cdot \exp(j \cdot 2\pi \cdot f \cdot T) \cdot df \right|^2}{\frac{N_2}{2} \cdot \int_{-\infty}^{\infty} |H(f)|^2 \cdot df}$$

A theorem in advanced calculus is *Schwarz's inequality*, which we give in general terms here as

$$\left| \int_a^b f_1(x) \cdot f_2(x) \cdot dx \right|^2 \le \left(\int_a^b |f_1(x)|^2 \cdot dx \right) \cdot \left(\int_a^b |f_2(x)|^2 \cdot dx \right)$$

Equality can be seen to hold when one of the functions is proportional to the complex conjugate of the other. *Cauchy's inequality* is the equivalent inequality for sums.

From Schwartz's inequality we have

$$\left|\int_{-\infty}^{\infty} H(f)\cdot S_i(f)\cdot \exp(j\cdot 2\pi\cdot f\cdot T)\cdot df\right|^2 \le \left(\int_{-\infty}^{\infty}\left|H(f)\right|^2\cdot df\right)\cdot\left(\int_{-\infty}^{\infty}\left|S_i(f)\right|^2\cdot df\right)$$

with equality only when

$$H(f) = C\cdot S_i^*(f)\cdot \exp(-j\cdot 2\pi\cdot f\cdot T)$$

where C is an arbitrary real or complex constant. The inverse Fourier transform gives us

$$h(t) = C\cdot s_i(T-t)$$

This is called a *matched filter* because the transfer function is matched to the signal.

Another point that we need to understand is taken from the maximum SNR, which we get by substituting the base inequality into the equation for SNR. We get

$$\left(\frac{S}{N}\right)_T \le \frac{\int_{-\infty}^{\infty}\left|S(f)\right|^2\cdot df}{N_0/2} = \frac{E_S}{N_0/2}$$

where E_S is the total signal energy. The point to be taken is that the maximum SNR is only a function of the total signal energy, and is independent of how this energy is distributed in the spectrum.

The Sampled Matched Filter

For sampled signals, the development is similar except for the use of sums in place of integrals and the use of Cauchy's inequality in place of Schwartz' inequality. Without loss of generality, we take the filter impulse response as a convolution.

The sampled matched filter is derived as follows. Consider the output of a digital filter

$$a_i(j) = \sum_{k=0}^{\infty} h(k)\cdot s_i(j-k)$$

The filter can be a convolution (finite impulse response, or FIR) or a recursive (infinite impulse response, or IIR) digital filter.

We begin by looking at the output of the digital filter for noise, independent from sample to sample, that has a variance of σ_0^2. The noise output is

$$hn(j) = \sum_{k=0}^{\infty} h(k)\cdot n(j-k)$$

and the variance of this noise output is

$$\sigma_H^2 = E\left(hn(j)^2\right) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} h(p)\cdot h(q)\cdot E\left(n(j-p)\cdot n(j-q)\right)$$

$$= \sigma_0^2\cdot\sum_{k=0}^{\infty}\left(h(k)\right)^2$$

We write the SNR for sample m as

$$\left(\frac{S}{N}\right)_m = \frac{\left|\sum_{k=0}^{\infty} h(k)\cdot s_i(m-k)\right|^2}{\sigma_0^2 \cdot \sum_{k=0}^{\infty}\left(h(k)\right)^2}$$

We will use Cauchy's inequality. In general terms, Cauchy's inequality is

$$\left[\sum_{k=1}^{n} a(k)\cdot b(k)\right]^2 \le \left(\sum_{k=1}^{n}\left(a(k)\right)^2\right)\cdot\left(\sum_{k=1}^{n}\left(b(k)\right)^2\right)$$

with equality only when the *b(k)* are proportional to the *a(k)*. Taking *a(k)* as *h(k)* and *b(k)* as $s_i(m-k)$ and noting that the inequality holds as we let *n* increase without bound, we have

$$\left[\sum_{k=0}^{\infty} h(k)\cdot s_i(m-k)\right]^2 \le \left(\sum_{k=1}^{\infty}\left(h(k)\right)^2\right)\cdot\left(\sum_{k=1}^{\infty}\left(s_i(m-k)\right)^2\right)$$

with equality only when

$$h(k) = C\cdot s_i(m-k)$$

where *C* is an arbitrary constant, which can be set for normalization or other convenience in the software. Note that, as with the analog matched filter, the matched filter is a replica of the signal, reversed in time. Clearly, for a realizable filter (i.e., $h(k)=0$ for $k<0$) the value of *m* must be selected as the length of a finite signal or more. For signals that have exponential tails, the value of *m* is a parameter in determining approximation error in the matched filter.

The maximum SNR achieved by a digital matched filter is

$$\left(\frac{S}{N}\right)_m \le \frac{\sum_{k=0}^{\infty}\left(s_i(m-k)\right)^2}{\sigma_0^2} = \frac{E_s}{\sigma_0^2}$$

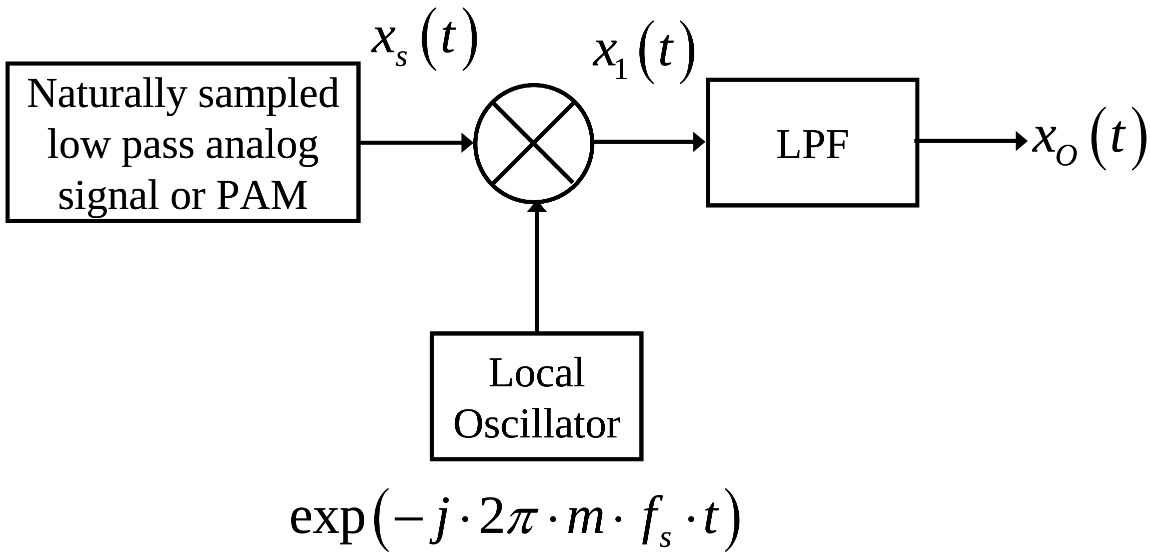
As with the analog matched filter, the maximum SNR is only a function of the total signal energy.

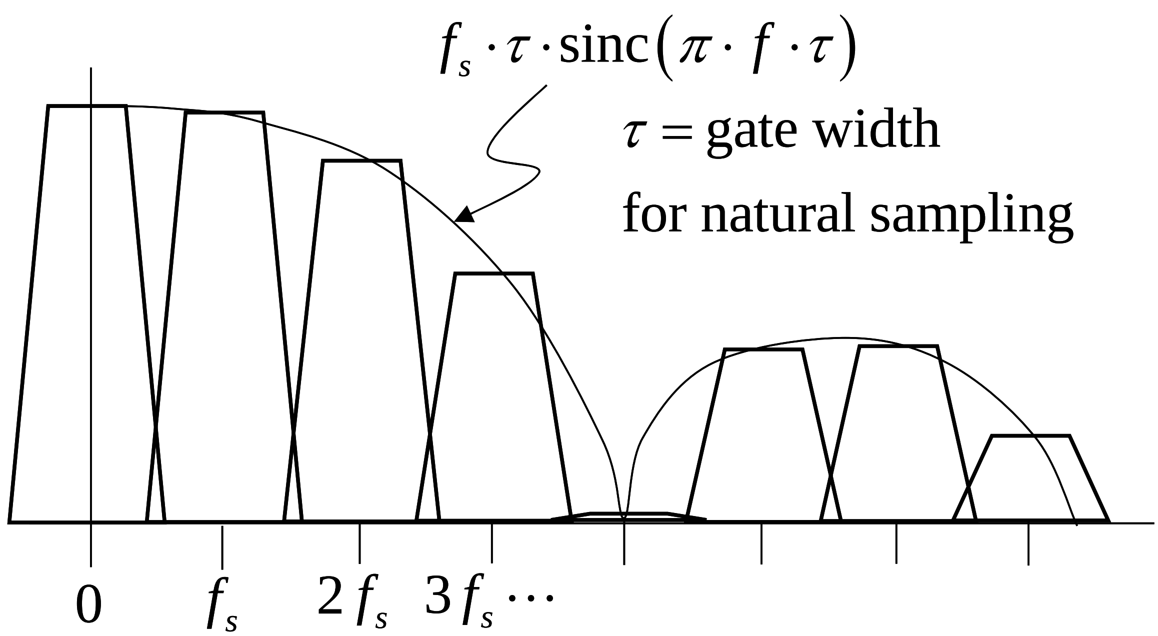
Homework

Sklar Problem 2.4 page 101

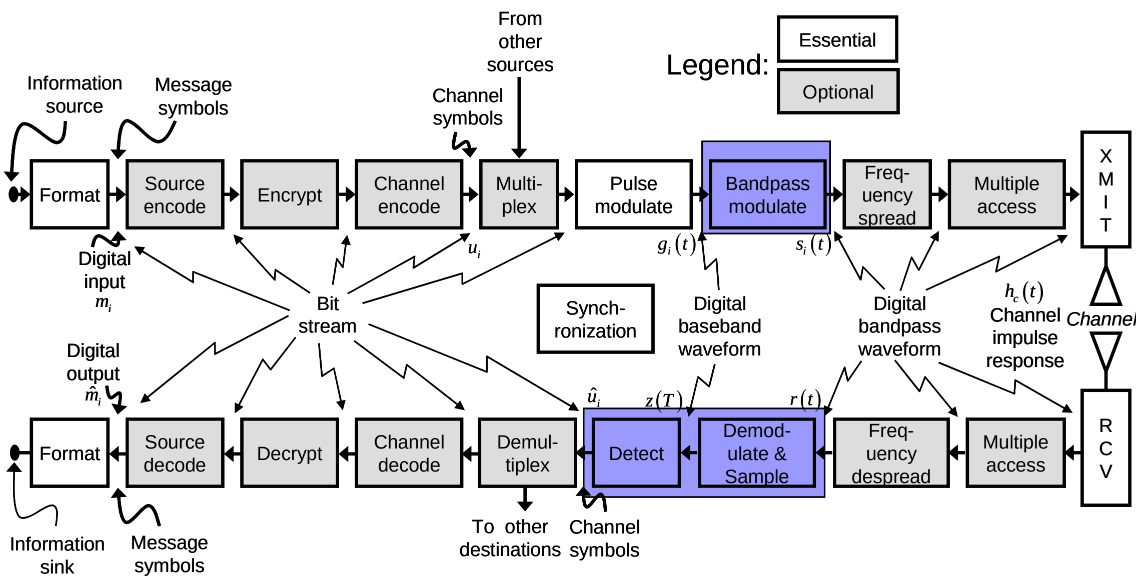
The basic principle is the frequency aliasing inherent in the sampling operation. Natural sampling (2.4.1.2 pp. 66-68) is used, so that the sin(x)/x frequency response of the natural sampling aperture will be applied to the signal spectrum. Contrast with impulse sampling, 2.1.1.1, which has a sampling aperture width of effectively zero and has no rolloff.

Sampling replicates the system spectra. The distance between replicated spectra is the sample rate. The block diagram of Figure P2.1 mixes one of these replicated spectra to baseband. The low pass filter attenuates all but the baseband signal out of the multiplier to recover the signal.





Topics from Sklar Chapter 4, Bandpass Modulation and Demodulation



Why Modulate?

When our communication channel is a cable, we can transmit baseband signals. When we have an RF spectrum, the signals in the channel are modulated RF signals. Consideration in use of open-air RF channels include

- SNR, determined by
  - Transmit power.
  - Antenna gains.
  - Distance between transmitter and receiver.
  - Obstructions and other effects in the path between the transmitter and receiver.
  - Sources of noise other than receiver noise such as electrical machinery and other uses of the spectrum.
  - RF losses in the transmitter and receiver..
  - Other effects.
- Sharing the spectrum with other uses.
  - Multiple users in our communications design.
  - Other users not related to our use of the spectrum.
- Available bandwidth – a very significant restriction involving international agreements and allocations.
- The symbol rate we use to achieve our communications goals.

Since the bit error as a function of SNR is a function of the modulation technique, we devote a lot attention to this area. Functions of modulation include

- Preparation for wireless transmission
  - Formulation of the digital bandpass waveform.
  - Modulation on a carrier at RF in the allocated frequency band.
- Optional additional steps include:

- Multiple access, obtained by multiplexing in time or frequency, or through interleaving symbols in blocks of data (Chapter 11, next semester).
- Spread spectrum, obtained by additional modulation (Chapter 12, next semester).

Digital Bandpass Modulation Techniques

Baseband and bandpass techniques are laid out in Sklar's Figure 4.1 page 170. Please refer to that figure.

When we talked about sampling techniques last time, we added quadrature demodulation and digital bandpass sampling because these techniques are commonly used today.

Concept: Signal as a phasor

If you perform a coherent complex demodulation of a signal, the real and imaginary samples can be taken as the (x,y) components in a plot. The vector from the origin to this point is called a phasor. The length of the vector is the amplitude of the signal and its phase relative to the vector from the origin o (1,0) is the phase of the signal relative to that of the L.O. signal.

Phase Shift Keying

The signal is of the form

$$s_i(t) = \sqrt{\frac{2E}{T}} \cdot \cos\left(\omega_0 \cdot t + \phi_i(t)\right)$$

where

$$\begin{aligned} E &= \text{ energy in a pulse} \\ T &= \text{ duration time of each pulse} \\ \omega_0 &= \text{ center frequency, radians per second} \\ \varphi_i(t) &= \text{ phase in radians for pulse } i \end{aligned}$$

The data is encoded on the phase, which has two or more values. The phase is of the form

$$\phi_i(t) = \frac{2\pi \cdot i}{M}, \; i = 1, \ldots, M$$

For BPSK, M=2; for QPSK, M=4. For larger M we call this M-ary PSK. This type of modulation is easy to draw as a phasor. What is the matched filter for this signal?

Frequency Shift Keying

Frequency shift keying (FSK) signals are of the form

$$s_i(t) = \sqrt{\frac{2E}{T}} \cdot \cos\left(\omega_i \cdot t + \phi_i\right)$$

The data is encoded in the frequencies ω_i and the phases ϕ_i are either constant or arranged for phase continuity between pulses.

There are a number of different types of FSK. When the phases are continuous between pulses, we have continuous-phase FSK (CPFSP). When the frequencies and pulse width are properly chosen, a filter matched to one pulse will have a zero response to pulses at other frequencies, and we have orthogonal FSK. This criteria can be seen to be

$$\omega_i = \omega_0 + \frac{2\pi}{T} \cdot \left(i - \frac{M}{2} \right)$$

This means that, in a matched filter, the "wrong" signal will be averaged over an integral multiple of complete cycles, and will average to zero. If, in addition, if the center frequency ω_0 is set so that the condition is met that all of the frequencies are multiples of $2\pi \cdot i/T$ then ϕ can be a constant.

Note that the diagram for FSK, Figure 4.5 (b) on page 174, is not a phasor plot, but a three-dimensional plot. This figure diagrams the responses of three matched filters for an orthogonal FSK signal with *M*=3.

Amplitude Shift Keying

Amplitude shift keying (ASK) has the general form

$$s_i(t) = \sqrt{\frac{2E_i}{T}} \cdot \cos(\omega_0 \cdot t + \phi)$$

and the signal is coded on the amplitudes represented by the pulse energies E_i. The amplitudes most often used are zero and full power for binary amplitude shift keying. This simple form is called on-off keying and you recognize it as the simplest form of modulation, used in early Morse code and primitive radar. It's not used much anymore because it's not efficient with modern transmitters and it doesn't properly exploit bandwidth efficiencies available with coherent decoding. This is revealed when you look at the matched filter for the zero amplitude pulse.

Amplitude Phase Keying

Modulation of both the amplitude and phase gives us a signal with the general form

$$s_i(t) = \sqrt{\frac{2E_i(t)}{T}} \cdot \cos(\omega_0 \cdot t + \phi_i(t))$$

This waveform is most easily appreciated when it is viewed as a phasor. The values of amplitude and phase can be a set of *M* points spaced efficiently inside a circle representing maximum transmitter power, and allow a larger value for *M* than might otherwise be obtained. The effective pulse energy is determined by the minimum distance between the points, not the transmitter power, so we have a tradeoff between available SNR and *M*.

Waveform Amplitude Coefficient

The waveform amplitude coefficient is usually written

$$A = \sqrt{\frac{2E}{T}}$$

This equation relates the sinusoidal peak amplitude to the pulse duration *T* and the pulse energy *E*. The physical units of *A* are the square root of watts, or volts per root ohm, or amperes times root ohms, where a root ohm is the square root of ohms. It takes on physical meaning when related to the impedance of the transmission line or load.

Detection of Signals in Gaussian Noise

Decision Regions

Gaussian noise is two-dimensional on a phasor plot. Use of the phasor plot is useful in diagramming coherent detection techniques as determination of regions in the plot in which a given bit or symbol is defined as the result of the detection operation.

Correlation Receiver

For the phase and phase-amplitude modulations, a phasor plot is a simple way to diagram the decision regions. For other signals, we should use a more abstract space in which each dimension represents the output of a filter matched to the signal for one of the *M* pulses. This “filter space” is valid for all pulses and, in fact, is what is used in the receiver for the decision process. A receiver that uses multiple parallel matched filters for decoding M-ary coded pulses is a correlation receiver.

Gaussian noise propagates to the output of a correlation receiver, forming an M-dimensional Gaussian distribution about the point in M-space representing the output for a particular signal without noise.

Likelihood Ratio Decision Threshold

The probability density function of the output of the *M* channels of a correlation filter for a given signal is as follows. We designate the output of the correlation filter for signal s_i by the M-vector $\underline{a}_1$ and the covariance of the noise outputs as the M by M covariance matrix *R*. When the matched filters for the M channels are for orthogonal signals, the noises out of the M channels are uncorrelated and *R* is diagonal (why?). The probability density function for the outputs of the M channels, given that the signal is s_i, is

$$p\left(\underline{z}\middle|s_i\right)=\frac{1}{\left(2\pi\right)^{M/2}\cdot\left|R\right|^{1/2}}\cdot\exp\left(-\frac{1}{2}\cdot\left(\underline{z}-\underline{a}_i\right)^T\cdot R^{-1}\cdot\left(\underline{z}-\underline{a}_i\right)\right)$$

An alternate term for the probability density function is the *likelihood function*. When we have a signal s_j in noise, where *j* is not necessarily equal to *i*, and we substitute the actual values of the *M* receiver outputs for $\underline{z}$, we call this the likelihood function of signal *j* relative to signal *i*.

The logarithm of the likelihood function, with terms not a function of *i* or *j* dropped, is proportional to

$$J=\left(\underline{z}-\underline{a}_i\right)^T\cdot R^{-1}\cdot\left(\underline{z}-\underline{a}_i\right)$$

This quantity, or sometimes its square root, is called the *Bhatacharya distance* between the points in M-space $\underline{z}$ and $\underline{a}_i$. Comparing the Bhatacharya distance between the point *z* and points for two different values of *i* is equivalent to comparing the ratio of the likelihood functions to one, and is sometimes called a *likelihood ratio test*.

Detections are defined by partitioning the M-space into regions where the Bhatacharya distance is smallest for each value of i, and values of $\underline{z}$ in each region are declared detections of the respective signal s_i.

Coherent Detection

Coherent detection is possible when the timing and phase of the received signal is known. This requires a synchronization function in the receiver that determines the transmitter timing and carrier phase, and tracks it during the period that communications signals are received. The synchronization function usually involves special signals transmitted specifically for that purpose at the beginning of communications and, sometimes, periodically during communications. A simple example is the stop bits in RS-232 communications. This function is the subject of Sklar's Chapter 10 and will be a topic next semester.

Coherent Detection of PSK Waveforms

MPSK waveforms must be detected coherently, except differential PSK. This means that a matched filter for the specific phase of the waveform, synchronized in time, is used to provide the signal provided for the detection decision. This has two major advantages over non-coherent detection processing:

- Since the phase is known, only one channel, not complex demodulation, is needed to provide the signal for detection; addition of a second channel would double the noise power, so coherent detection has an inherent 3 dB SNR advantage.
- Since the timing is known, only the signal at the time of arrival is processed. This removes uncertainty about the time of detection. Since matched filter outputs at offset times are not processed, and only the signal at peak SNR is processed, losses due to use of off-peak signals are avoided.

Coherent Detection of FSK

Detection of FSK signals is done by simultaneous channels using filters matched to the M frequencies used in the FSK modulation. Coherent detection has the same two major advantages.

Non-coherent Detection

Detection of Differential PSK

Differential PSK waveforms are processed without phase information by comparing the phase of successive pulses. The information is coded in the phase change, not in the absolute phase. As a result, two noisy signals are subtracted, so that an inherent 3 dB SNR disadvantage over coherent detection is incurred.

Noncoherent Detection of FSK

FSK signals can be processed for detection without phase information by using quadrature demodulation and using the signal energy in each channel for the detection process. Since quadrature detection has two channels and a single matched filter for coherent detection produces a real signal, we see a 3 dB SNR disadvantage.

Complex Modulation from Baseband

We showed that a complex signal could be found from a bandpass signal by quadrature demodulation. Quadrature demodulation consists of two operations:

- Multiplication of the bandpass signal by a complex exponential that translates the center frequency to zero frequency
- A two-channel (i.e. complex) low-pass filter that removes unwanted components of the result, such as bandpass signals centered at negative twice the original center frequency, and spurious signals generated by imperfections in the process.

If the signal is sampled at I.F., any anti-aliasing filtration must be done at I.F., and the requirements on sampling aperture and sample time jitter are more stringent because of the more carrier at I.F. varies more quickly than a baseband signal of the same amplitude.

If sampling at I.F. is at a higher rate than required to accommodate the signal because of practical considerations in the anti-aliasing filter, a digital complex low-pass filter can be used at baseband to reduce the bandwidth for processing to just that required for the signal. Since this bandwidth is reduced over what is required for sampling, this digital output can be decimated to produce samples for processing at the Nyquist rate instead of the higher rate required to accommodate the I.F. signal.

Error Performance for Binary Systems

Probability of Bit Error for Coherent BPSK

The signal for BPSK is of the form

$$s_i(t) = \sqrt{\frac{2E}{T}} \cdot \cos\left(\omega_0 \cdot t + (i-1) \cdot \pi\right)$$

The SNR from a matched filter is

$$SNR_E = \frac{2E_b}{N_0}$$

For BPSK, the signals for the two phases have opposite arithmetic signs (this characteristic of a binary signal is called the antipodal property) so we are looking at the probability that the signal will change sign. Thus the probability of bit error is

$$P_{B-BPSK} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

where *Q(x)* is the complementary Gaussian probability distribution function,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_x^{\infty} \exp\left(-\frac{x^2}{2}\right) \cdot dx = \frac{1 - \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)}{2}$$

where *erf(x)* is the error function used in mathematics,

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \cdot \int_0^z \exp\left(-t^2\right) \cdot dt$$

Probability of Bit Error for Coherent Detection of Differentially Encoded PSK

Differential phase encoding has obvious advantages in robustness because slipping of phase by 180 degrees in either the medium or the synchronization will not affect the decoded signal. The probability of error is equal to the probability that one, but not both, phases are decoded incorrectly. The result is

$$P_{B-DPSK} = 2 \cdot P_{B-BPSK} \cdot \left(1 - P_{B-BPSK}\right)$$

and is given by Sklar as Equation (4.80) on page 212.

Probability of Bit Error for Coherently Detected Binary Orthogonal FSK

Decoding of binary FSK signals is done by comparing the outputs of two matched filters and taking the largest. As such, there is a 3 dB SNR penalty over BPSK, and the probability of bit error is

$$P_{B-OFSK} = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

Probability of Bit Error for Non-Coherently Detected Binary Orthogonal FSK

When non-coherent FSK detection is used, a real plus a quadrature channel is necessary because phase is unknown in non-coherent demodulation. The square of the noise signal has a Poisson (exponential) distribution. We look at the sum of the powers out of the two channels, and we are comparing two signals with a Poisson distribution. With the signal present, the exact form for the probability distribution is more complex (non-central chi-square, or Rician) but is well-approximated as an offset Poisson distribution for high SNR. The bit error is (Sklar, Equation (4.96) page 216)

$$P_{B-NCBFSK} = \frac{1}{2} \cdot \exp\left(-\frac{E_b}{2N_0}\right)$$

Probability of Bit Error for Binary DPSK

DPSK decoding can be modeled as taking a difference in two complex signals, computing the power, and thresholding the result. We have the same signals as non-coherent binary orthogonal FSK except that we have two signal amplitudes instead of one, for a 3 dB improvement in SNR. The result is

$$P_{B-BDPSK} = \frac{1}{2} \cdot \exp\left(-\frac{E_b}{N_0}\right)$$

Assignment for next time

- Load the Full Version of SystemView distributed on CD-ROM today and examine the samples and demos of the Sklar installation
- Read Sklar Chapter 4: 176-219
- Do problems 4.1, 4.2 page 101
- Take one final cut at Problem 2.4 page 101 in SystemView.

ALERT: Quiz Scheduled for February 22

Quiz will be designed for students to finish in one hour.
Rules are

- Open notes
- Open book
- No cell phones or RF-capable messaging
- No laptop computer
- Ask questions by writing them on a slip of paper and handing it to me. Answers will be on the whiteboard for all to see and use.

The quiz will begin at about 7:00 PM, and I will collect the quizzes at 8:25 PM.

ALERT: Term Project Assignment on February 22

Before the quiz on February 22, I will assign each of you an individual Term Project to be executed in SystemView. I will ask for progress reports from time to time during the semester. On April 26, the last day of class, we will schedule each of you about 15 minutes to present and demonstrate your term project. Also due at that time are

- Your slides or other presentation material.
- The SystemView file or files, and any data files used or produced by your demonstration.
- A report on your project, its design, and its execution. This short report will document the design and implementation of your term project and point out the issues that you discovered and the problems that you solved. A template will be posted on the web site for your report in the next few weeks. This template will include an outline and instructions that will provide you with information on how to prepare the report. Your presentation should follow the same basic outline and present the material in the report.