1 Modulation

Last lecture we talked about how a digital transmitter sends one of $M$ symbols, that is, linear combinations, of a small number $N$ of orthogonal waveforms. The TX’s choice from the $M$ possible linear combinations lets it send one of $\log_2 M$ bits per symbol. This section talks specifically about the choices for the orthogonal waveforms and the symbols.

1.1 PAM

Three “types” of modulations only use one (orthogonal) waveform, $\phi_0(t) = \cos(2\pi f_c t) p(t)$, where $p(t)$ is the pulse shape. (It has nothing to be orthogonal to, except for waveforms sent during other symbol periods.) A linear combination, then, is just $a_{i,0}p(t)$, for some constant $a_{i,0}$. A few different modulations are so named by their choice of $M$ different choices of $a_{i,0}$, for $i = 0, \ldots, M - 1$.

- On-off-keying (OOK): $M = 2$, and we choose $a_{0,0} = 0$, and $a_{1,0} = 1$. When sending a “0” bit, the transmitter doesn’t send anything. It only actually transmits energy when sending a “1” bit.

- Binary PAM, a.k.a. binary phase-shift-keying (BPSK): $M = 2$, and we choose $a_{0,0} = -1$, and $a_{1,0} = 1$. Now, the amplitude of the sinusoid being transmitted is switched from a +1 to a −1, or vice versa, when the bit switches. This is also a phase shift of 180°, which is why it can be called “phase shift” keying.
• **M-ary PAM**: Binary PAM is extended in M-ary PAM to include \( M \) (for \( M \) equal to some power of 2) equally spaced amplitudes, centered on zero. So, for example, \( M = 4 \) PAM would have \( a_{0,0} = -3 \), \( a_{1,0} = -1 \), \( a_{2,0} = +1 \), \( a_{3,0} = +3 \). In general, for \( M \)-ary PAM, \( a_{i,0} = 2^i - M + 1 \).

Note that **differential phase shift keying** (DPSK) is an implementation variant of BPSK, which we will discuss next lecture.

### 1.2 M-ary QAM and PSK

Many types of modulations use the following two orthogonal waveforms:

\[
\phi_0(t) = \cos(2\pi f_c t)p(t) \\
\phi_1(t) = \sin(2\pi f_c t)p(t)
\]

where \( p(t) \), again, is the pulse shape. These symbols \( a_i = [a_{i,0}, a_{i,1}]^T \) can then be plotted on a 2-D graph, typically with the \( \phi_0(t) \) amplitude plotted on the horizontal axis, and the \( \phi_1(t) \) amplitude plotted on the vertical axis.

There are two main “types” of modulations which use these two orthogonal waveforms:

• **Phase-shift keying (PSK)**: PSK places all \( a_i \) uniformly on the unit circle. That is, \( \|a_i\| = 1 \) for all \( i \). For \( M = 2 \), this is the same as BPSK. For \( M = 4 \), the symbols are spaced every 90 degrees and thus the constellation looks like the four corners of a square, and is also called quadrature phase shift keying (QPSK). For \( M = 8 \), symbols are every 45 degrees apart.

• **QAM**: Generally, any modulation that uses \( \phi_0(t) \) and \( \phi_1(t) \) above can be considered as QAM. So PSK is a sometimes a subtype of QAM. But QAM is not limited to \( \|a_i\| = 1 \). For example, square QAM has \( M = 2^{2k} \) for some integer \( k \geq 4 \), where symbols are arranged in a square grid, centered at zero. Note \( M = 4 \) square QAM is the same as QPSK. But \( M = 16 \) and \( M = 64 \) are also common values of \( M \) for square QAM.

We drew five different constellation diagrams of PSK and QAM modulations in the Lecture 11 notes, on page 4.

Note that for QPSK or square \( M \)-QAM, you send \( \log_2 M \) bits total per symbol. We can also look at the modulation as sending two independent signals, one on the in-phase (horizontal axis) and one on the quadrature (vertical axis), each sending \( \frac{1}{2}\log_2 M \) bits per symbol.

Note that OQPSK and \( \pi/4 \) QPSK are variations on QPSK, that have equivalent fidelity performance and identical bandwidth.
as QPSK, but they are “constant envelope” modulations, which we will discuss next lecture.

**Bandwidth:** Note that for PAM, PSK, and QAM, assuming SRRC pulse shaping with rolloff factor $\alpha$, the null-to-null bandwidth of the signal is $B = (1 + \alpha)/T_s$, where $1/T_s$ is the symbol rate.

### 1.3 FSK

Binary FSK modulations use the following orthogonal waveforms:

$$
\begin{align*}
\phi_0(t) &= \cos (2\pi [f_c - n\Delta f] t) p(t) \\
\phi_1(t) &= \cos (2\pi [f_c + n\Delta f] t) p(t)
\end{align*}
$$

for some positive integer $n$, where $\Delta f = \frac{1}{4T_s}$. We showed that for $n = 2$, that is, when the frequency difference (“offset”) is $1/T_s$, that the two are orthogonal. Standard binary FSK uses $n = 2$.

**Bandwidth:** For Binary FSK, when a SRRC pulse is used in $p(t)$, the transmitted signal bandwidth is given by

$$B = 2n\Delta f + (1 + \alpha)R_s$$

where $R_s = R_b = 1/T_b$. So, for Binary FSK,

$$B = 4\frac{1}{4T_s} + (1 + \alpha)R_s = (2 + \alpha)R_s$$

### 1.4 MSK

You can also show that orthogonality is achieved for any integer $n$ in (1). Since putting the two sinusoids closer together reduces the bandwidth of the signal, this is a good idea for spectral efficiency. When $n = 1$ binary FSK is called *minimum shift keying (MSK)*.

But MSK is actually the same as OQPSK with a half-cosine pulse shape. This is not something we discuss until next lecture.

### 1.5 Receiver Complexity Options

One issue in receivers is the costs of coherent vs. non-coherent reception.

**Def’n:** *Coherent Reception*

A receiver is coherent if it is phase-synchronous, that is, it estimates and then uses the phase of the incoming signal.

Coherent reception requires a phase-locked loop (PLL) for each different carrier frequency in the signal. This can be a problem for FSK receivers that have high $M$. Also, accurate phase-locked loops can be difficult to achieve for mobile radios for cases with high Doppler.
Differential reception is another technique to aid receivers which operate without accurate phase. When Doppler is present, and the coherent receiver ends up with a consistent phase error of 180 degrees, a standard receiver will flip every subsequent bit estimate. A differential receiver encodes only the change in phase. In the same case, when the phase is flipped by 180 degrees, one bit error will be made because of the flip, but the subsequent bits will be all correct.

2 Fidelity

Now, let’s compare the fidelity (probability of bit error and probability of symbol error) across different modulations and receiver complexity options. For the same energy per bit, using a different modulation would result in a different fidelity.

Recall $N = FkT_0B$. We also define $N_0 = FkT_0$, so that $N = N_0B$. The units of $N_0$ are Watts per Hertz, or equivalently, Joules.

The signal power at the receiver can also be written in terms of the energy per symbol or energy per bit. Since energy is power times time, the energy used to send a symbol duration $T_s$ is $E_s = ST_s$, where $S$ is the received power $P_r$, and $E_s$ has units Joules per symbol. To calculate the energy needed to receive one bit, we calculate $E_b = ST_b/\log_2 M$ Joules per bit. To shorten this, we define $T_b = T_s/\log_2 M$ and then $E_b = ST_b$ or $E_b = S/R$, where $R = 1/T_b$ is the bit rate.

The probability of bit error is a function of the ratio $E_b/N_0$. For example, for BPSK,

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

where $Q(z)$ is the tail probability of a Gaussian r.v., given by

$$Q(z) = \frac{1}{2}\text{erfc}\left(\frac{z}{\sqrt{2}}\right)$$

Put this into Matlab or your calculator, it is a useful function. You might also use the table or graph on pages 645-646 of the Rappaport book.

Other symbol error rates and bit error rates for a variety of modulations is shown in Table 2.

What are the most important? In my opinion: BPSK, DPSK, QPSK, 16-QAM, 64-QAM, 2-non-co-FSK, 2-co-FSK. For the homework, you will make a table, for each modulation, listing $P_{\text{bit error}}$ formula, and bandwidth. You will also plot $P_{\text{bit error}}$ as a function of $E_b/N_0$. Both will be very useful to have in your portfolio so that you have them available on the exam 2. An example is shown in Figure 1.
Table 1: Summary of probability of bit and symbol error formulas for several modulations.

<table>
<thead>
<tr>
<th>Name</th>
<th>$P_{\text{[symbol error]}}$</th>
<th>$P_{\text{[bit error]}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$</td>
<td>same</td>
</tr>
<tr>
<td>OOK</td>
<td>$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$</td>
<td>same</td>
</tr>
<tr>
<td>DPSK</td>
<td>$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$</td>
<td>same</td>
</tr>
<tr>
<td>M-PAM</td>
<td>$\frac{2(M-1)}{M}Q\left(\sqrt{\frac{6\log_2 M E_b}{N_0}}\right)$</td>
<td>$\approx \frac{1}{\log_2 M}P_{\text{[symbol error]}}$</td>
</tr>
<tr>
<td>QPSK</td>
<td>$2Q\left(\sqrt{2\log_2 M \sin^2(\pi/M)\frac{E_b}{N_0}}\right)$</td>
<td>$\approx \frac{1}{\log_2 M}P_{\text{[symbol error]}}$</td>
</tr>
<tr>
<td>M-PSK</td>
<td>$\leq 2Q\left(\sqrt{2\log_2 M \sin^2(\pi/M)\frac{E_b}{N_0}}\right)$</td>
<td>$\approx \frac{1}{\log_2 M}P_{\text{[symbol error]}}$</td>
</tr>
<tr>
<td>Square M-QAM</td>
<td>$\approx \frac{4}{\log_2 M} \frac{M-1}{\sqrt{\log_2 M}}Q\left(\frac{3\log_2 M E_b}{N_0}\right)$</td>
<td>$\approx \frac{1}{\log_2 M}P_{\text{[symbol error]}}$</td>
</tr>
<tr>
<td>2-non-co-FSK</td>
<td>$= \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$</td>
<td>same</td>
</tr>
<tr>
<td>M-non-co-FSK</td>
<td>$= \sum_{n=1}^{M-1} \frac{(M-1)(-1)^{n+1}}{n+1} \exp\left[-\frac{n\log_2 M E_b}{n+1 N_0}\right]$</td>
<td>$= \frac{M/2}{M-1}P_{\text{[symbol error]}}$</td>
</tr>
<tr>
<td>2-co-FSK</td>
<td>$= Q\left(\sqrt{\frac{E_b}{N_0}}\right)$</td>
<td>same</td>
</tr>
<tr>
<td>M-co-FSK</td>
<td>$\leq (M-1)Q\left(\sqrt{\log_2 M \frac{E_b}{N_0}}\right)$</td>
<td>$= \frac{M/2}{M-1}P_{\text{[symbol error]}}$</td>
</tr>
</tbody>
</table>

Figure 1: Comparison of probability of bit error for BPSK and Differential BPSK.
3 Link Budgets with Digital Modulation

Link budgeting was discussed in the first part of this course assuming that the required SNR was known. But actually, we deal with a required $\frac{E_b}{N_0}$. Note

$$\frac{E_b}{N_0} = \frac{S}{N_0 R_b} = \frac{S}{N R_b}$$

Typically the ratio of bandwidth to bit rate is known. For example, what is $\frac{B}{R_b}$ for SRRC pulse shaping?

![Link Budgeting including Modulation: This relationship graph summarizes the relationships between important variables in the link budget and the choice of modulation.](image)

3.1 Shannon-Hartley Channel Capacity

We’ve talked about $S/N$, the signal power ($S$) to noise power ($N$) ratio. In the link budget section, we provided a $S/N$ required for a given modulation. Here, we talk about the fundamental limits for bandwidth efficiency for a given $S/N$.

**Def'n**: Bandwidth efficiency

The bandwidth efficiency of a digital communication system is ratio of $\eta_B = \frac{R}{B}$, where $R$ is the bits per second achieved on the link, and $B$ is the signal bandwidth occupied by the signal. Bandwidth efficiency has units of bits per second per Hertz.

The limit on bandwidth efficiency is given by Claude Shannon [1], who extended work by Ralph Hartley (a UofU alum!). The
Shannon-Hartley theorem is,

\[
\frac{R_{\text{max}}}{B} = \log_2 \left( 1 + \frac{S}{N} \right)
\]

where \( R_{\text{max}} \) is the maximum possible bit rate which can be achieved on the channel for the given signal to noise ratio.

**Example: Maximum for various \( S/N \) (dB)**

What is the maximum bandwidth efficiency of a link with \( S/N(\text{dB}) = 10, 15, 20 \)? What maximum bit rate can be achieved per 30 kHz of spectrum (one AMPS/USDC channel)? Is this enough for 3G?

**Solution:** \( S/N(\text{dB}) = 10, 15, 20 \) translates into \( S/N = 10, 31.6, 100 \). Then

\[
\begin{align*}
\frac{R_{\text{max}}}{B} &= \log_2 (11) = 3.46 \\
\frac{R_{\text{max}}}{B} &= \log_2 (32.6) = 5.03 \\
\frac{R_{\text{max}}}{B} &= \log_2 (101) = 6.66
\end{align*}
\]

With 30 kHz, multiplying, we have \( R_{\text{max}} = 104, 151, 200 \) kbps. No, this isn’t enough for 3G cellular, which says it aims to achieve up to 14 Mbps.

Shannon’s bound is great as engineers, we can come up with a quick answer for what we cannot do. But it doesn’t necessarily tell us what we can achieve. The problem is that it is very difficult to get within a couple of dB of Shannon’s bound. So, then we have to resort to the performance equations for particular modulation types.

### 3.2 Examples

**Example: 200 kHz**

What bit rate can be achieved on a 200 kHz if the \( S/N \) ratio is 20 dB?

**Example: 1.25 MHz**

What bit rate can be achieved on a 200 kHz if the \( S/N \) ratio is 12 dB?

**References**