Lecture 15: Line Codes and Pulse Shaping

John M Pauly

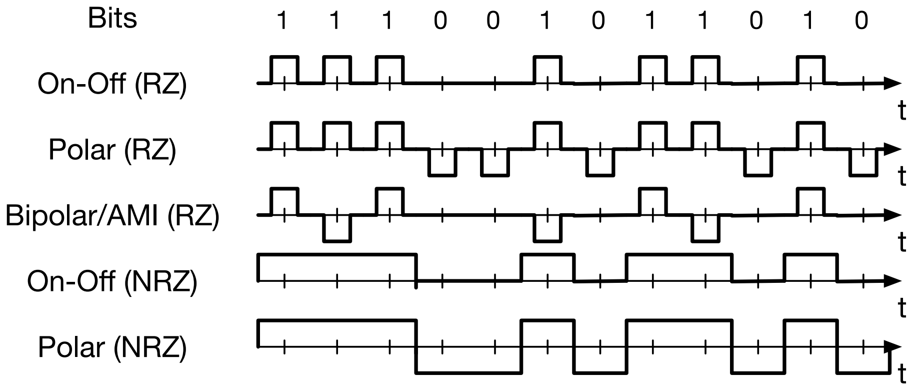
November 9, 2021

Digital Communications: Line Codes and Pulse Shaping

- Review
  - Line codes
  - Pulse width and polarity
  - Power spectral density
- Intersymbol interference (ISI)
- Pulse shaping to reduce ISI
- Embracing ISI

Based on lecture notes from John Gill

Line Code Examples, from last time ...



RZ = Return to Zero NRZ = Non-Return to Zero

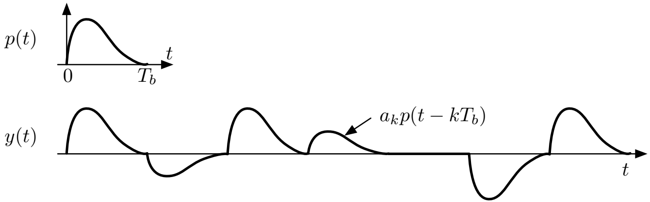
Line Code Examples, from last time ...

Features we would like

- Minimum bandwidth (NRZ)
- Easy clock recovery (RZ)
- Frequently, no DC value (bipolar pulses)

PSD of Line Codes

- The PSD of a line code depends on the shapes of the pulses that correspond to digital values. Assume the pulses $p(t)$ are amplitude modulated (PAM),



- The transmitted signal is the sum of weighted, shifted pulses.

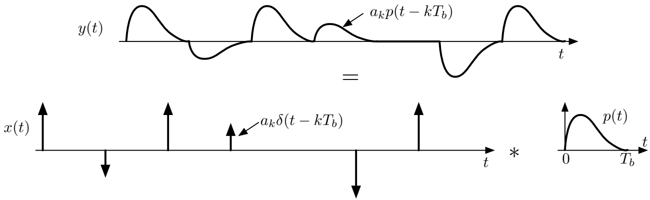
$$y(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT_b)$$

where T_b is spacing between pulses.

- Pulse may be wider than T_b, which leads to inter-symbol interference (ISI). We will look at this case shortly.

PSD of Line Codes (cont.)

- PSD depends on pulse shape, rate, and digital values $\{a_k\}$.
- We can simplify analysis by representing $y(t)$ as impulse train convolved with $p(t)$



- Then $Y(f) = P(f)X(f)$, and the PSD of $y(t)$ is

$$S_y(f) = |P(f)|^2 S_x(f)$$

- $P(f)$ depends only on the pulse, independent of digital values or rate.
- $S_x(f)$ increases linearly with rate $1/T_b$ and depends on distribution of values of $\{a_k\}$. E.g., $a_k = 1$ for all k has narrower PSD.

Power Spectral Density of Line Codes (review)

- In general, the PSD of a line code is

$$S_y(f) = |P(f)|^2 \, S_x(f)$$

  where

$$S_x(f) = \mathcal{F}\{R_x(t)\} = \frac{1}{T_b} \sum_{n=-\infty}^{\infty} R_n e^{-jn2\pi f T_b}$$

  so

$$S_y(f) = |P(f)|^2 \left(\frac{1}{T_b} \sum_{n=-\infty}^{\infty} R_n e^{-jn2\pi f T_b} \right)$$

  In many cases, only $R_0 \neq 0$, or just a few terms are not equal to zero.

- We would like to limit the bandwidth of the transmitted signal
  - Modify $P(f)$
  - Modify $x(t)$

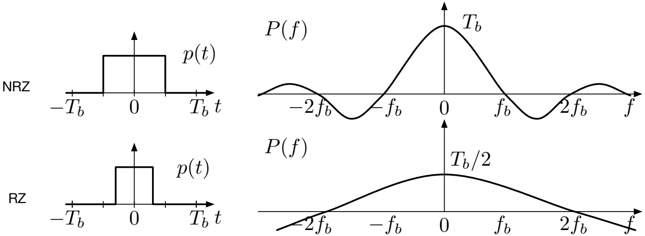
$P(f)$ for RZ and NRZ Pulses

- NRZ (100% pulse)

$$\begin{aligned} p(t) &= \Pi(t/T_b) \\ P(f) &= T_b \operatorname{sinc}(\pi T_b f) \\ |P(f)|^2 &= T_b^2 \operatorname{sinc}^2(\pi T_b f) \end{aligned}$$

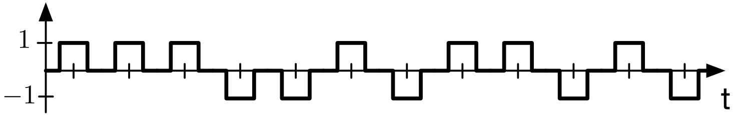
- RZ half-width:

$$\begin{aligned} p(t) &= \Pi(t/(T_b/2)) \\ P(f) &= \tfrac{1}{2} T_b \operatorname{sinc}(\tfrac{1}{2}\pi T_b f) \\ |P(f)|^2 &= \tfrac{1}{4} T_b^2 \operatorname{sinc}^2(\tfrac{1}{2}\pi T_b f) \end{aligned}$$



Modifying $x(t)$

Polar signaling. Transmit 1 as $+p(t)$, and 0 as $-p(t)$,



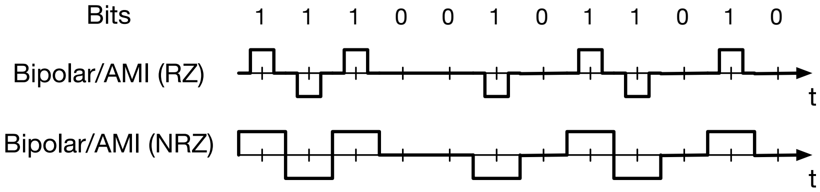
Only $R_0 \neq 0$, and

$$S_y(f) = \frac{|P(f)|^2}{T_b} R_0 = \frac{|P(f)|^2}{T_b}$$

The PSD of polar signaling depends only on the spectrum of $p(t)$.

Modifying $x(t)$, continued

Bipolar signaling. Transmit 1's as alternating $\pm p(t)$, and 0's as 0.



Bipolar signaling for full-width (NRZ) pulses.

$$S_y(f) = \frac{|P(f)|^2}{T_b} \sin^2(\pi T_b f) = T_b \operatorname{sinc}^2(\pi T_b f) \sin^2(\pi T_b f)$$

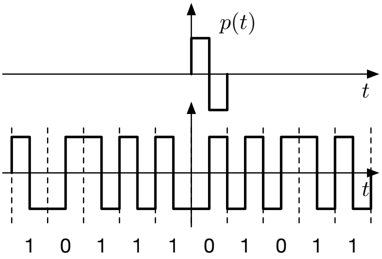
For half-width (RZ) pulses:

$$S_y(f) = \frac{T_b}{4} \operatorname{sinc}^2\left(\frac{1}{2}\pi f T_b\right) \sin^2(\pi T_b f)$$

The PSD depends on $P(f)$, but decays more rapidly due to the $\sin^2(\cdot)$ terms.

Modifying $p(t)$

Split phase, or Manchester encoding,

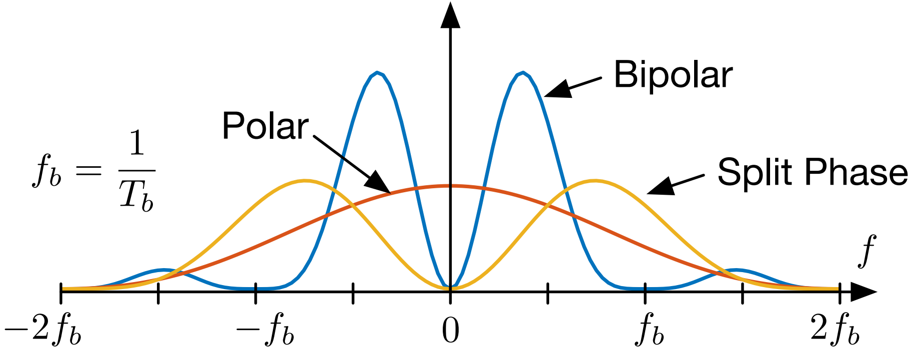


The PSD is

$$
\begin{aligned}
S_y(f) &= |P(f)|^2 \left(\frac{1}{T_b} \sum_{n=-\infty}^{\infty} R_n e^{-jn2\pi f T_b} \right) \\
&= |P(f)|^2 \frac{1}{T_b}
\end{aligned}
$$

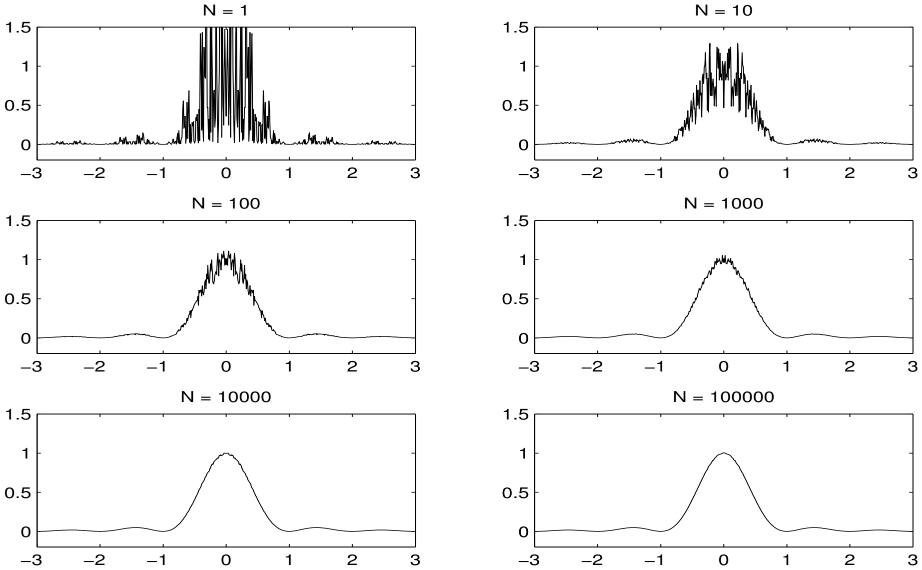
Comparison of PSD's

RZ Polar, NRZ Bipolar, and Split Phase



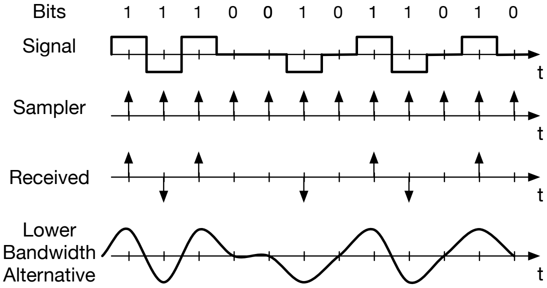
NRZ Bipolar would look like Split Phase.

PSD of Polar Signaling (Matlab Experiment)



Pulse Shaping

- So far, we've assumed all of the pulses are square pulses, either 50% or 100% of the pulse spacing.
- We can use other pulses, all we care is that the waveform have the right value at the samples!



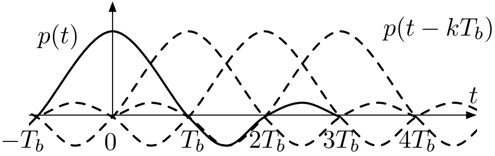
- How do we design the waveform so that it passes through the samples, and has minimum bandwidth?
- This is a lot like the sampling reconstruction problem, except we only care about the values at the sample points.

Reducing ISI: Pulse Shaping

- A time-limited pulse cannot be bandlimited
- Linear channel distortion results in spread out, overlapping pulses
- Nyquist introduced three criteria for dealing with ISI.

  The first criterion was that each pulse is zero at the sampling time of other pulses.

$$p(t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm kT_b \,,\; k = \pm 1, \pm 2, \ldots \end{cases}$$



Harry Nyquist, "Certain topics in telegraph transmission theory", Trans. AIEE, Apr. 1928

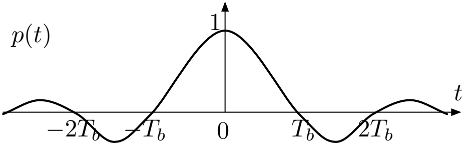
Pulse Shaping: sinc Pulse

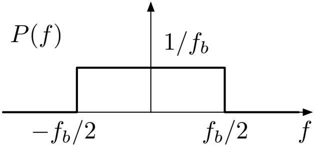
- Let $f_b = 1/T_b$. The sinc pulse $\operatorname{sinc}(\pi f_b t)$ satisfies Nyquist's first crierion for zero ISI:

$$\operatorname{sinc}(\pi f_b t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm k T_b\,,\ k = \pm 1, \pm 2, \ldots \end{cases}$$

- This pulse is bandlimited. Its Fourier transform is

$$P(f) = \frac{1}{f_b} \Pi\Big(\frac{f}{f_b}\Big)$$



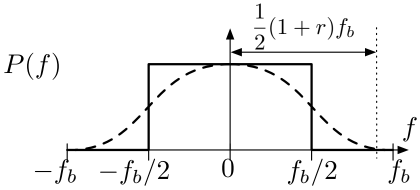


- Unfortunately, this pulse has infinite width and decays slowly.

Nyquist Pulse

Nyquist increased the width of the spectrum in order to make the pulse fall off more rapidly.

The Nyquist pulse has spectrum width $\frac{1}{2}(1+r)f_b$, where $0 < r < 1$.



We want to choose $P(f)$, and hence $p(t)$, so that

- $p(t) = 0$ for $t = nT_b$, so there is no intersymbol interference.
- $p(t) = 1$ for $t = 0$

So that it basically functions like a sinc. Can we do this?

If we sample the pulse $p(t)$ at rate $f_b = 1/T_b$, then

$$\overline{p}(t) = p(t)\,\mathrm{III}_{T_b}(t) = p(t)\delta(t) = \delta(t)\,.$$

since $p(t)$ is zero at multiples of T_b, except at zero where it is 1.
The Fourier transform of the sampled signal is

$$\overline{P}(f) = \sum_{k=-\infty}^{\infty} P(f - kf_b) = 1$$

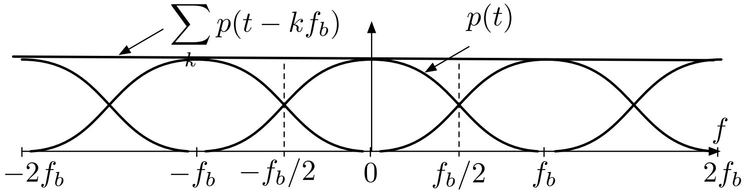
The first term is the Fourier transform of the sampled signal $p(t)\,\mathrm{III}_{T_b}(t)$, which equals the Fourier transform of $\delta(t)$, which is 1.

This is an interesting result!

Nyquist Pulse (cont.)

Since we are sampling below the Nyquist rate $(1+r)f_b$, the shifted transforms overlap.

Nyquist's criterion require the spectrum overlaps add to 1 for all f.



The result is the same as if we had used a sinc for $p(t)$, and summed them up. From the sampling perspective these do the same thing.

For parameter r with $0 < r < 1$, the resulting pulse has bandwidth

$$B_r = \frac{1}{2}(f_b + rf_b)$$

The parameter r is called *roll-off factor*

Nyquist Pulse (cont.)

There are many pulse spectra satisfying this condition. e.g., trapezoid:

$$P(f) = \begin{cases} 1 & |f| < \frac{1}{2}(1-r)f_b \\ 1 - \frac{|f|-(1-r)f_b}{2f_b} & \frac{1}{2}(1-r)f_b < |f| < \frac{1}{2}(1+r)f_b \\ 0 & |f| > \frac{1}{2}(1-r)f_b \end{cases}$$

A trapezoid is the difference of two triangles. Thus the pulse with trapezoidal Fourier transform is the difference of two sinc^2 pulses.

Example: for $r = \frac{1}{2}$,

$$P(f) = \frac{3}{2}\Lambda\left(\frac{f}{\frac{3}{2}f_b}\right) - \frac{1}{2}\Lambda\left(\frac{f}{\frac{1}{2}f_b}\right)$$

so the pulse is

$$p(t) = \frac{9}{4}\,\mathrm{sinc}^2(\tfrac{3}{2}f_b t) - \frac{1}{4}\,\mathrm{sinc}^2(\tfrac{1}{2}f_b t)$$

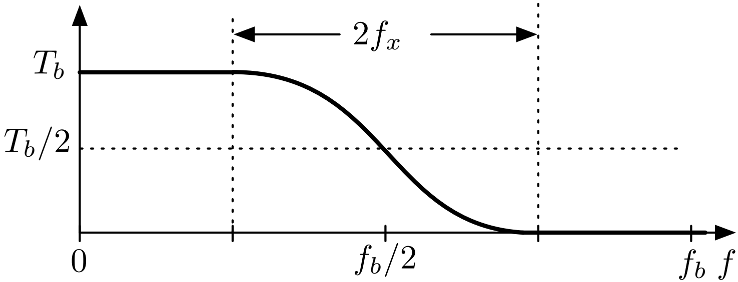
This pulse falls off as $1/t^2$.

Nyquist Pulse (cont.)

Nyquist chose a pulse with a “vestigial” raised cosine transform.

This transform is smoother than a trapezoid, so the pulse decays more rapidly.

The Nyquist pulse is parametrized by r. Let $f_x = rf_b/2$.

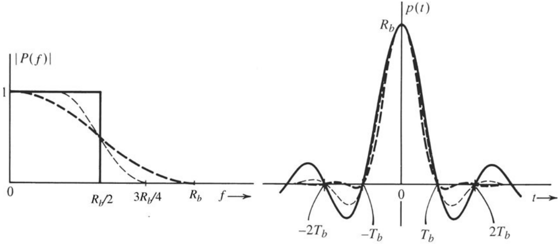


Nyquist Pulse (cont.)

Nyquist pulse spectrum is raised cosine pulse with flat porch.

$$P(f) = \begin{cases} 1 & |f| < \frac{1}{2}f_b - f_x \\ \frac{1}{2}\left(1 - \sin\pi\left(\frac{f - \frac{1}{2}f_b}{2f_x}\right)\right) & |f| - \frac{1}{2}f_b| < f_x \\ 0 & |f| > \frac{1}{2}f_b + f_x \end{cases}$$

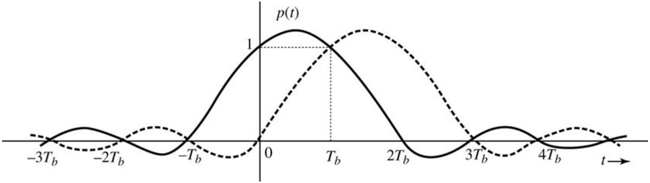
The transform $P(f)$ is differentiable, so the pulse decays as $1/t^2$.



This shows $r = 0$, a sinc, $r = 1/2$, and $r = 1$, a raised cosine, which is a well known window function. The sidelobes of $p(t)$ decrease rapidly with r.

Controlled ISI (Partial Response Signaling)

The second Nyquist approach embraced ISI. The interference is known if we know the adjacent bits! This allows us to achieve a lower bandwidth by using an even wider pulse. The ISI can be canceled using knowledge of the pulse shape.



In this case $p(t)$ is equal to 1 for $t = 0$ and $t = T_b$.

The signal at time T_b is then depends on both $x(0)$ and $x(T_b)$.

- If both are 1, then the output $y(t) = 2$
- if one is 1, and the other is -1, the output is 0
- if both are -1, the output is -2.

Given a starting value for a_k, where $x(t) = \sum_k a_k \delta(t - kT_b)$ we can subtract off the ISI term by term, and recover the bit sequence $\{a_k\}$

Partial Response Signaling (cont.)

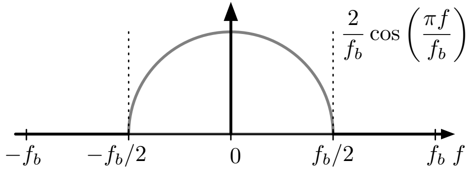
The ideal duobinary pulse is the sum two shifted sinc's

$$p(t) = \frac{\sin(\pi f_b t)}{\pi f_b t} + \frac{\sin(\pi f_b(t - T_b))}{\pi f_b(t - T_b)t} = \frac{\sin \pi f_b t}{\pi f_b t(1 - f_b t)}$$

This makes sense, it will be 1 at 0 and T_b and zero elsewhere.
The Fourier transform of $p(t)$ is

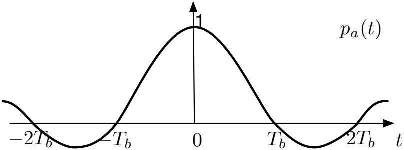
$$P(f) = \frac{2}{f_b} \cos\left(\frac{\pi f}{f_b}\right) \Pi\left(\frac{f}{f_b}\right) e^{-j\pi f / f_b}$$

The spectrum is confined to the theoretical minimum of $f_b/2$.



Zero-ISI, Duobinary, Modified Duobinary Pulses

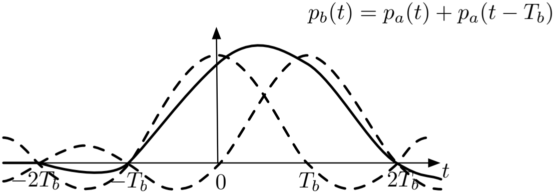
Suppose $p_a(t)$ satisfies Nyquist's first criterion (zero ISI).



Then

$$p_b(t) = p_a(t) + p_a(t - T_b)$$

is a duobinary pulse with controlled ISI.



By shift theorem,

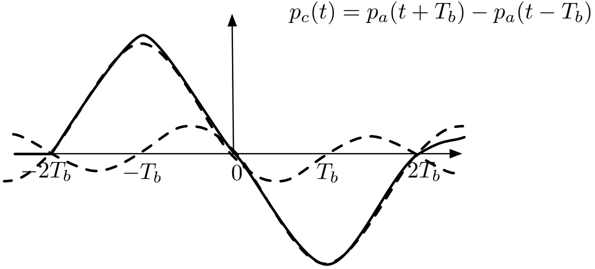
$$P_b(f) = P_a(f)(1 + e^{-j2\pi T_b f})$$

Since $P_b(f_b/2) = 0$, most (or all) of the pulse energy is below $f_b/2$.

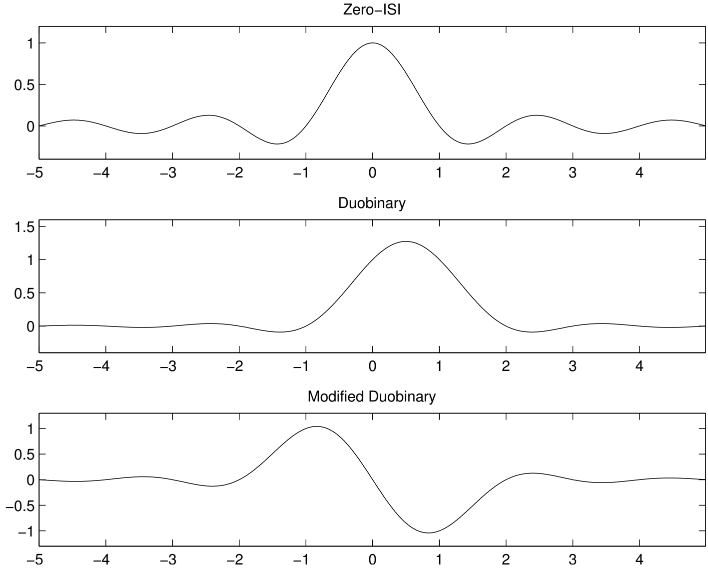
We can eliminate unwanted DC component using modified duobinary, where $p_c(-T_b) = 1$, $p_c(T_b) = -1$, and $p_c(nT_B) = 0$ for other integers n.

$$p_c(t) = p_a(t + T_b) - p_a(t - T_b) \;\Rightarrow\; P_c(f) = 2jP_a(f)\sin 2\pi T_b f$$

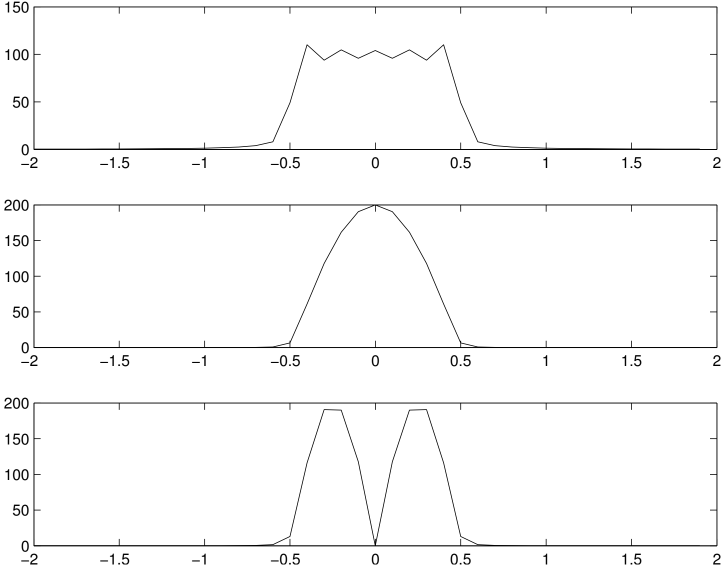
The transform of $p_c(t)$ has nulls at 0 and $\pm f_b/2$.



Zero-ISI, Duobinary, Modified Duobinary Pulses (cont.)



Zero-ISI, Duobinary, Modified Duobinary Pulses (cont.)



Partial Response Signaling Detection

- Suppose that sequence 0010110 is transmitted (first bit is startup digit).

Digit x_k	0	0	1	0	1	1	0
Bipolar amplitude	-1	-1	1	-1	1	1	-1
Combined amplitude		-2	0	0	0	2	0
Decode sequence		0	1	0	1	1	0

- Partial response signaling is susceptible to *error propagation*.
- If a nonzero value is misdetected, zeros will be misdetected until the next nonzero value.
- Note that the signal now has multiple levels (-2, 0, 2). We could also do that directly, as we will see next week.
- The need for partial response detection can be eliminated by differential encoding (from two classes ago), where the input stream is preprocessed. Then a receiver can directly read out the data bits.

Next Time

- Friday : APRS FSK Lab
- Monday : Eye Diagrams, M-ary encoding, digital carrier systems
- Wednesday : Error Correction, Parity, CRC codes
- Friday : Final project discussion