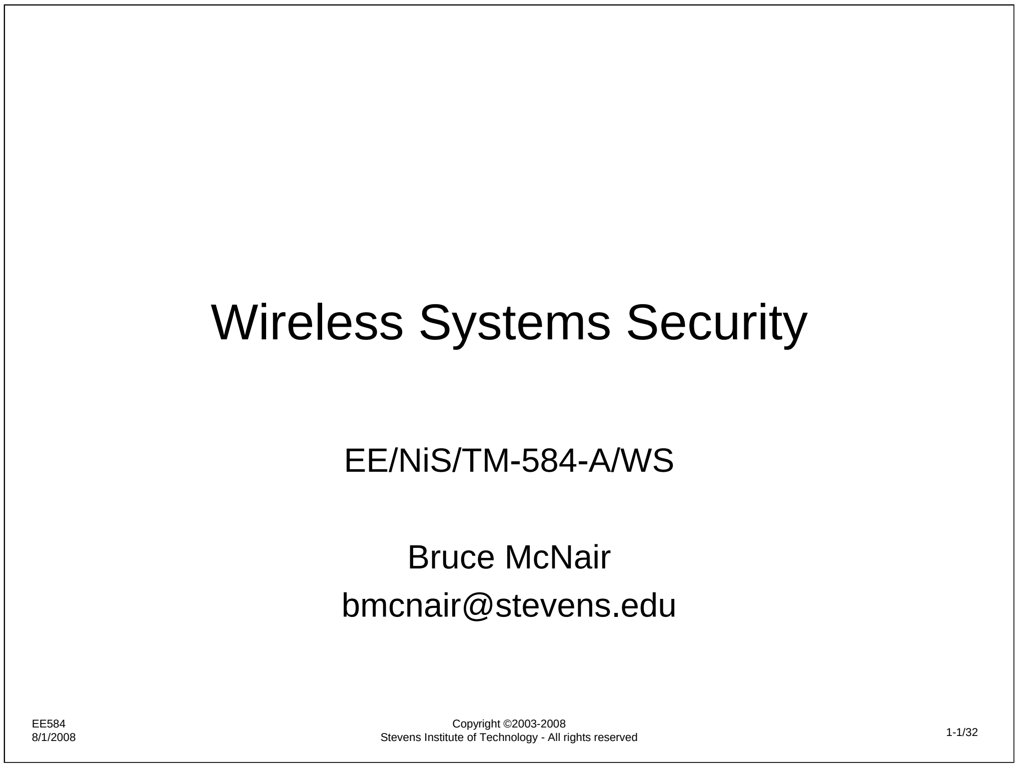


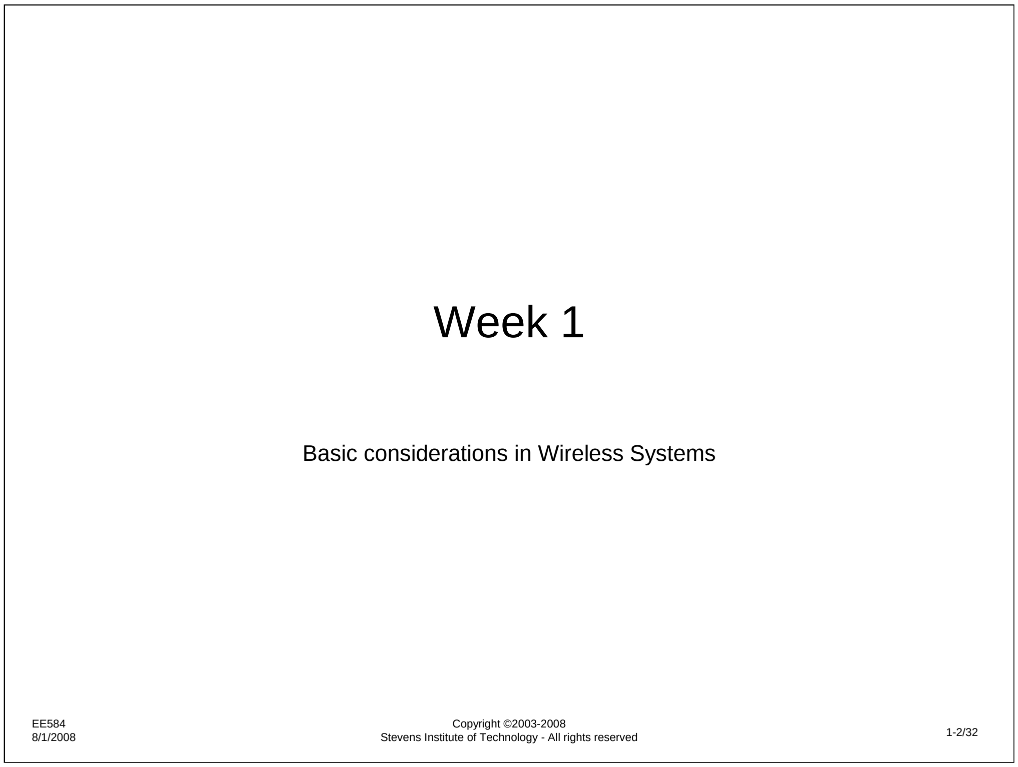
Wireless Systems Security

EE/NiS/TM-584-A/WS

Bruce McNair
bmcnair@stevens.edu

EE584
8/1/2008

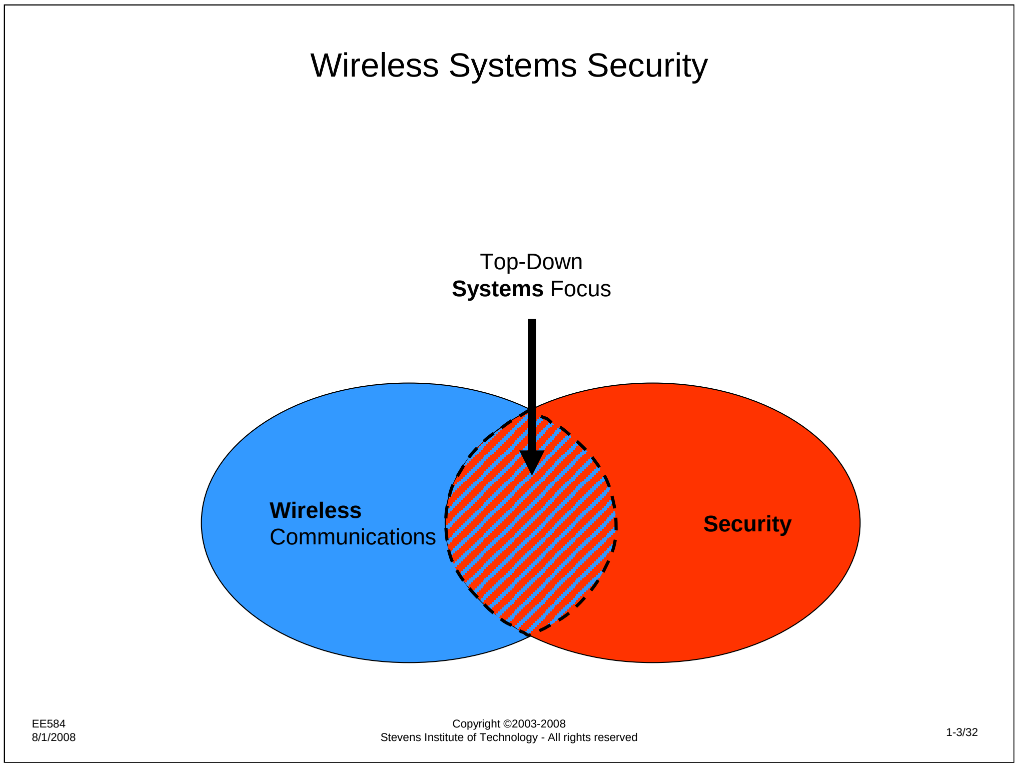
1-1/32



Week 1

Basic considerations in Wireless Systems

This week’s material will examine wireless systems; We will start with a discussion of the history of wireless and look at some of the considerations that are particularly relevant if we are concerned with security in these systems.



Wireless Systems Security

Top-Down
Systems Focus

Wireless
Communications

Security

EE584
8/1/2008

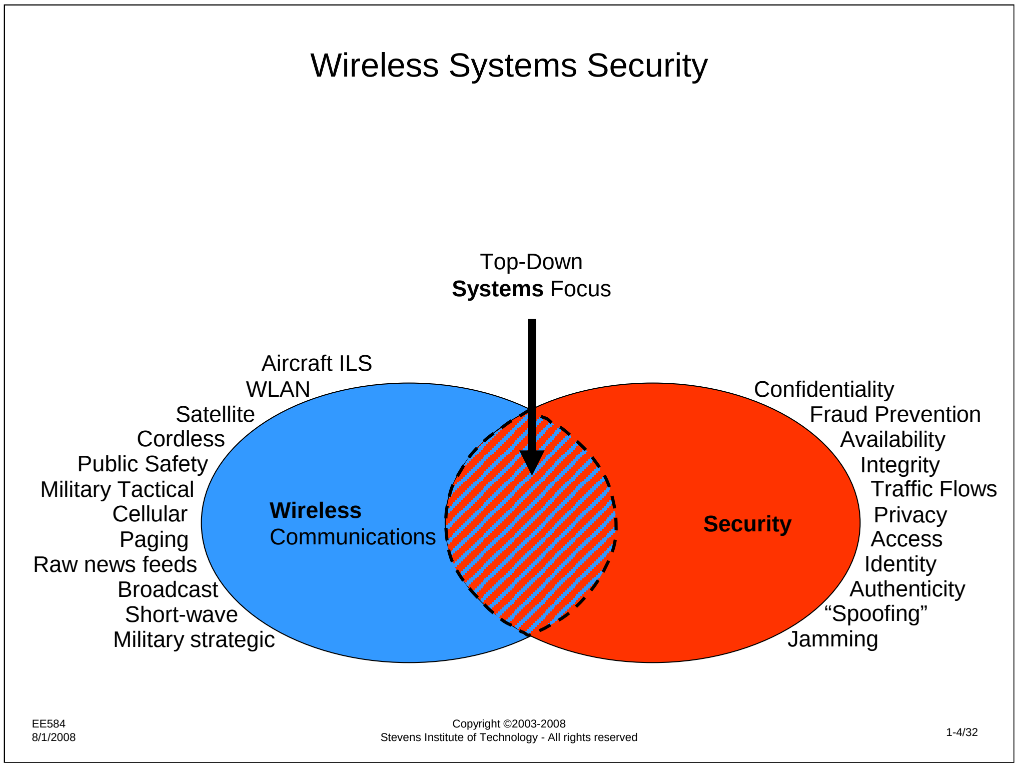
Copyright ©2003-2008
Stevens Institute of Technology - All rights reserved

1-3/32

To start, let's consider what the overall viewpoint of this course will be. We will obviously be looking at wireless systems, which is currently a popular subject for study. Obviously, we will also be looking at security, which has also gotten a lot of attention in recent years.

What is particularly interesting is the overlap of the two subject areas. In general, people examining security do not particularly look at how the unique aspects of wireless must be considered in examining security. Likewise, when studying wireless systems, the security aspects of the system design are not generally considered until issues develop.

For this course, I want to look carefully at the intersection of these two areas, taking a top-down integrated system focus, rather than getting lost in the details.



Wireless Systems Security

Top-Down
Systems Focus

Aircraft ILS
WLAN
Satellite
Cordless
Public Safety
Military Tactical
Cellular
Paging
Raw news feeds
Broadcast
Short-wave
Military strategic

Wireless
Communications

Security

Confidentiality
Fraud Prevention
Availability
Integrity
Traffic Flows
Privacy
Access
Identity
Authenticity
“Spoofing”
Jamming

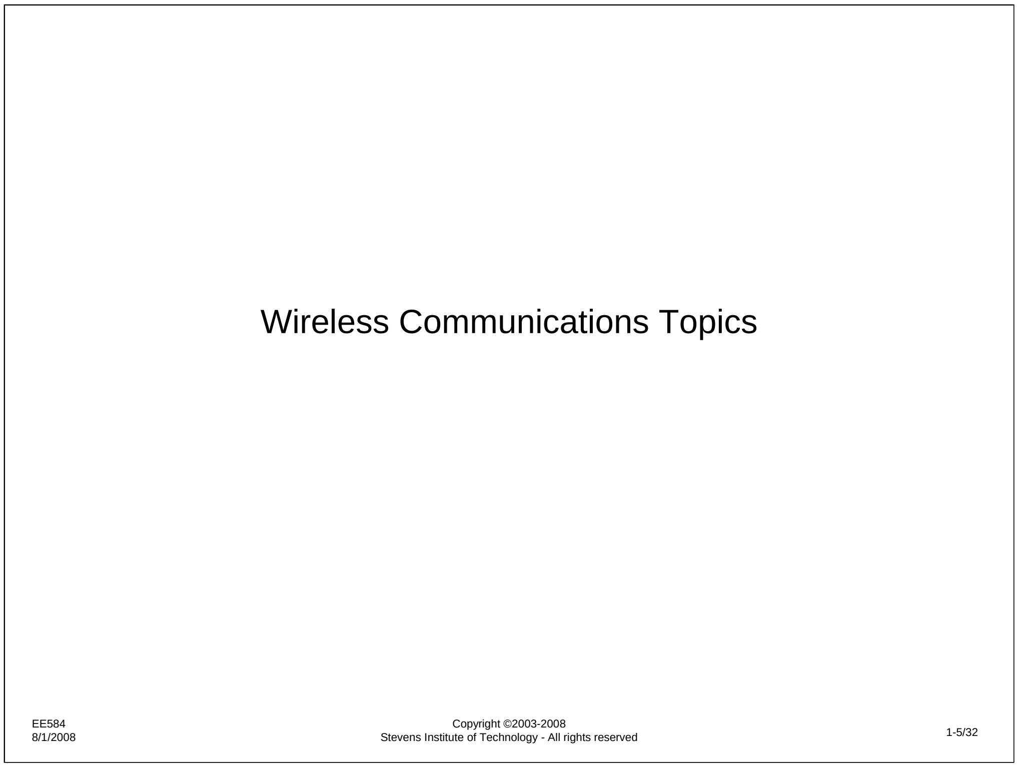
EE584
8/1/2008

1-4/32

To start, we need to consider exactly what we mean by “wireless communications” and what we mean by “security.” Listed above are a few of the many aspects of these subject areas.

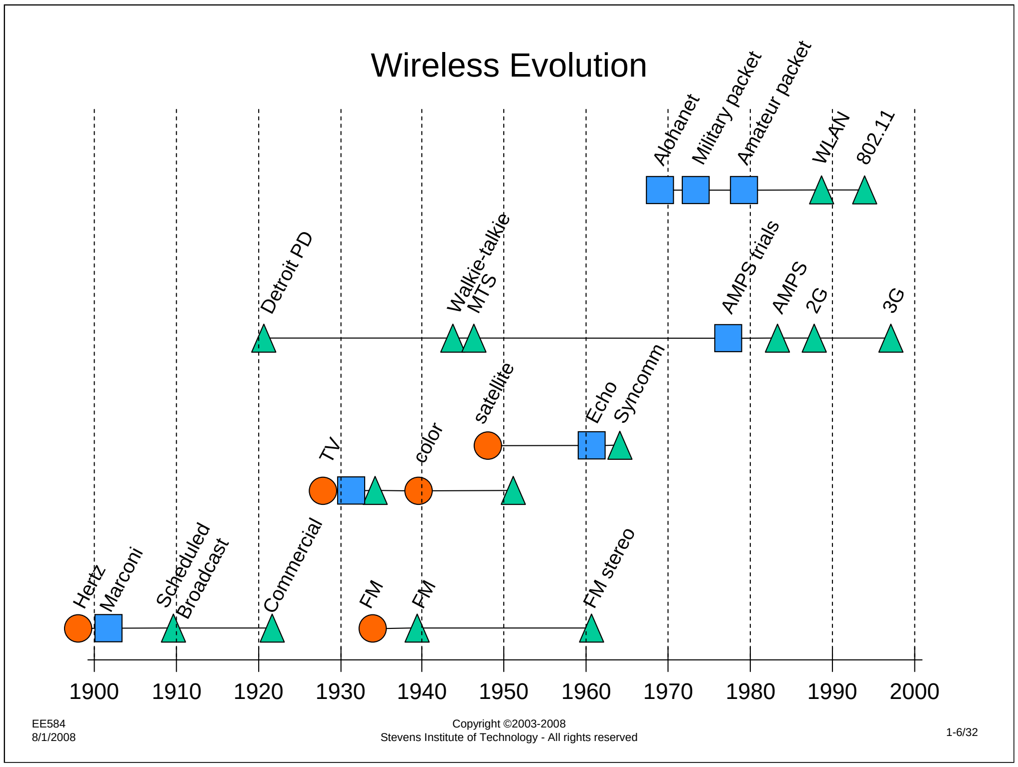
Some of the wireless systems are commonplace. We all use or regularly encounter cordless phones, cellular phones, and wireless LANs. What we might not consider, but certainly are wireless systems, are things like the Aircraft Instrument Landing Systems that enable aircraft to land at night or in inclement weather. Broadcast TV and satellite communications are clearly wireless systems, but we might not consider that there are security considerations with these systems. We will look at these and several other types of wireless systems in this course.

When we think of security, things tend to get less well defined. We often think of keeping information confidential, that is, restricted from being read by people who are not authorized, but there are many aspects of security beyond this one. Making sure information is available when it is needed is a security consideration, as is the need to keep information correct, assuring its integrity. Some users may be concerned about intruders knowing the volume of traffic they are sending at any given time or to whom they are communicating. These are examples of traffic flow security. We will examine these and several other aspects of security throughout this course.



Wireless Communications Topics

To begin, let’s discuss the wireless aspects of system design.



Wireless Evolution

EE584
8/1/2008

Copyright ©2003-2008
Stevens Institute of Technology - All rights reserved

1-6/32

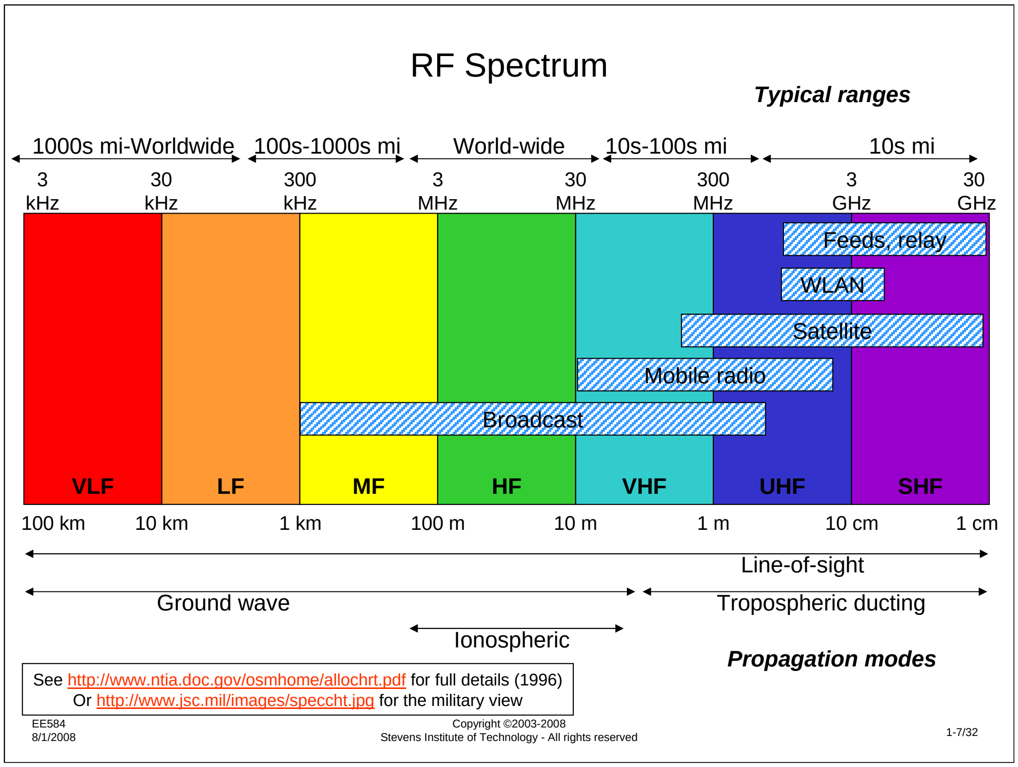
There has been a dramatic evolution in wireless communications since it was invented a little over 100 years ago. In the early days, we saw slow deployment of the technology, from the discovery of electromatic propagation by Hertz in the late 1800, the initial trans-Atlantic experiments by Marconi, and deployment of commercial radio stations.

TV and FM broadcast evolved a bit more rapidly, probably due to the existence of radio broadcast as a well understood medium.

In part because of the new, personalized technology, but also because of the size and cost, mobile wireless communications at first evolved slowly, but accelerated rapidly after the initial deployment of AMPS. Another part of this accelerated deployment was the revolution in component density that advances in electronics were making possible.

Probably the most dramatic evolution has been the deployment of wireless LANs. Here, because of the level of integration and the market size, military users, business users, and home users have access to similar capabilities.

From a security perspective, the rapid evolution of wireless systems suggests that systems might get deployed before designers have an opportunity to fully consider the security issues. This will be part of this course’s focus.



RF Spectrum

Typical ranges

1000s mi-Worldwide | 100s-1000s mi | World-wide | 10s-100s mi | 10s mi

3 kHz | 30 kHz | 300 kHz | 3 MHz | 30 MHz | 300 MHz | 3 GHz | 30 GHz

Feeds, relay

WLAN

Satellite

Mobile radio

Broadcast

VLF | LF | MF | HF | VHF | UHF | SHF

100 km | 10 km | 1 km | 100 m | 10 m | 1 m | 10 cm | 1 cm

Line-of-sight

Ground wave

Tropospheric ducting

Ionospheric

Propagation modes

See http://www.ntia.doc.gov/osmhome/allochrt.pdf for full details (1996)
Or http://www.jsc.mil/images/speccht.jpg for the military view

EE584
8/1/2008

Copyright ©2003-2008
Stevens Institute of Technology - All rights reserved

1-7/32

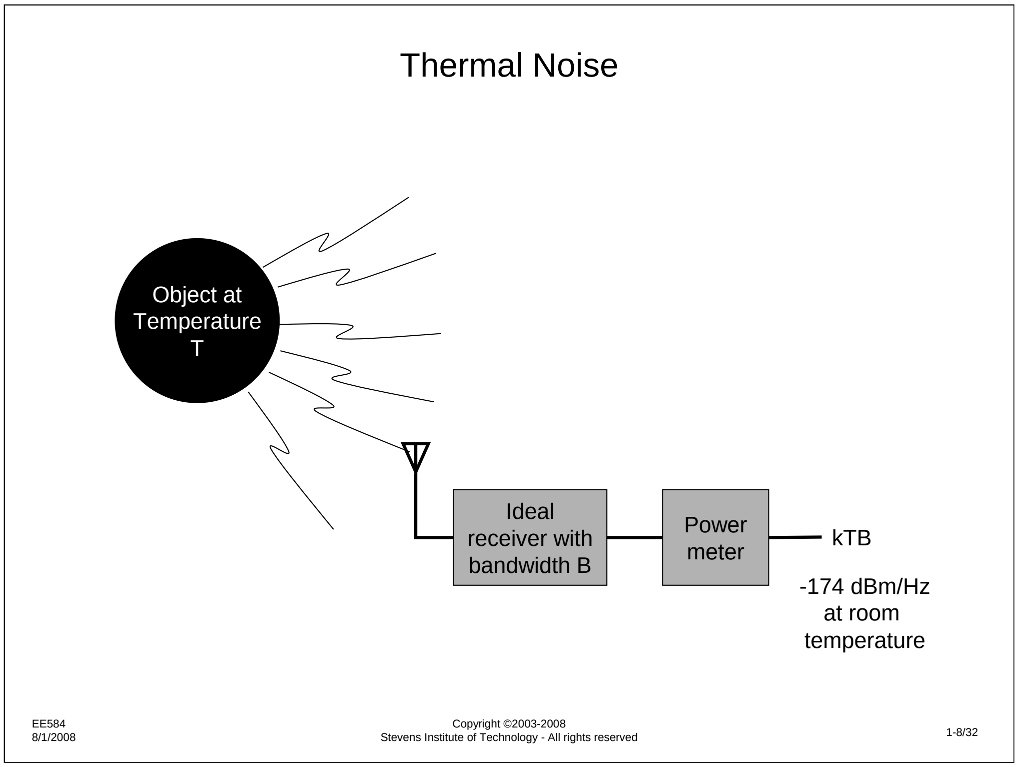
This chart is intended to grossly summarize information that could fill 10 textbooks so that we can talk about general wireless terminology.

Spectrum allocations to a service, an application or an individual are stated in terms of the operating frequency. I have shown the 7 most popular “decades” of frequency. Originally, spectrum was allocated based on wavelength and several services still refer to a band of operation in terms of its wavelength. Beyond this purpose, however, it is useful to keep wavelength in mind, since it directly relates to antenna sizes, transmission line considerations, and the physical geometry of devices that we might design.

I have listed the common propagation modes and typical transmission range for the various parts of the spectrum. These are not hard and fast rules, but merely guidelines about the range and propagation mode one might expect at different operating frequencies. For instance, in the HF region, line-of-sight and ground wave propagation are present, but the strongest signal will often be one that has been reflected from the ionosphere from thousands of miles away. On the other hand, at 150 MHz, it is very unlikely that the ionosphere will be capable of reflecting a signal. On the other hand, weather conditions may create tropospheric conditions that favor “ducting” of the signal.

Finally, I have illustrated the typical ranges of operating frequencies for various services or applications. Particular systems will be allocated a small band of frequencies within the range indicated.

Discussion topic: What are the security implications of where a system is operating in the RF spectrum?



Thermal Noise

Object at Temperature T

Ideal receiver with bandwidth B

Power meter

kTB

-174 dBm/Hz at room temperature

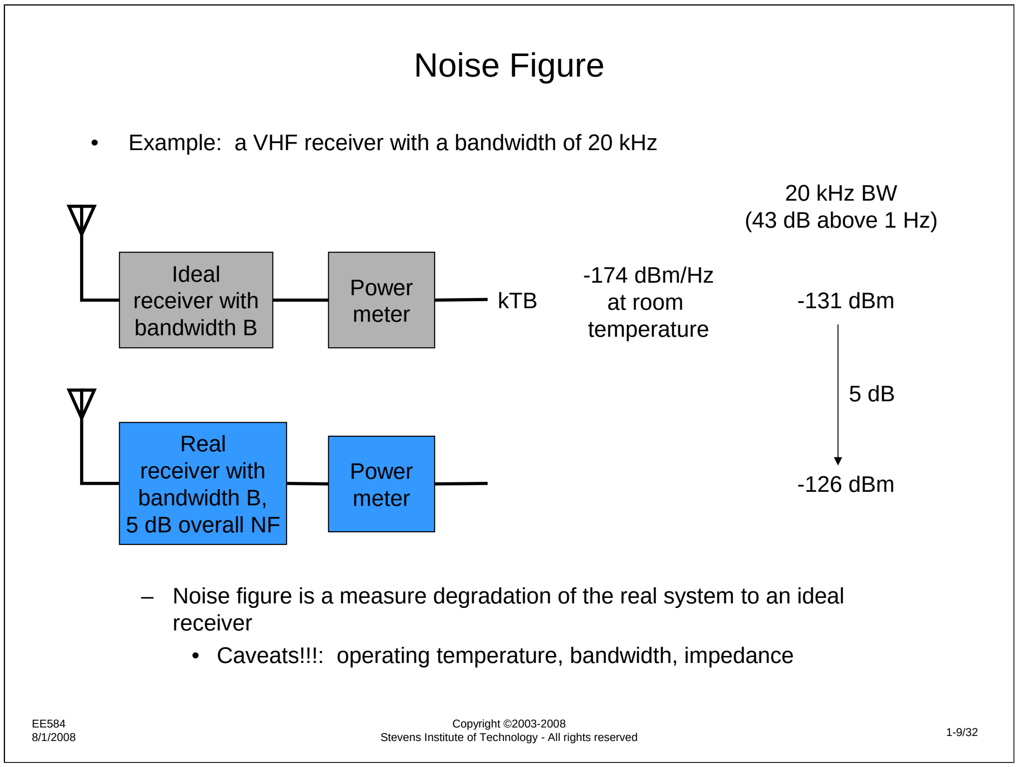
EE584
8/1/2008

1-8/32

As we start to discuss wireless communications system, the first design consideration we will examine is the noise performance of the system. I start here for two reasons: first, noise is the fundamental limitation of many wireless communications systems, determining other factors such as bandwidth and power; second, as we will see later, noise can be used by the system designer to their advantage in addressing some security considerations.

Thermal noise is the most common noise we must deal with in a wireless communications system. Everything is subject to thermal noise, due to the random motion of electrons in objects at temperatures above 0 degrees Kelvin. The noise power attributed to thermal noise is “kTB” – Boltzmann’s constant times the Kelvin temperature times the measurement bandwidth. Since thermal noise is truly random, its autocorrelation is zero for any nonzero time offset, making its spectrum uniform. A useful representation of kTB is –174 dBm/Hz – that is 174 dB below 1 milliwatt when measured in a 1 Hz bandwidth. From this value, we can easily scale to whatever bandwidth we need for a communications system.

Two items to note: this value for kTB noise is for room temperature (20 degrees Celsius) and assumes an ideal receiver.



Noise Figure

- Example: a VHF receiver with a bandwidth of 20 kHz

20 kHz BW
(43 dB above 1 Hz)
Ideal receiver with bandwidth B
Power meter
kTB
-174 dBm/Hz at room temperature
-131 dBm
5 dB
Real receiver with bandwidth B, 5 dB overall NF
Power meter
-126 dBm

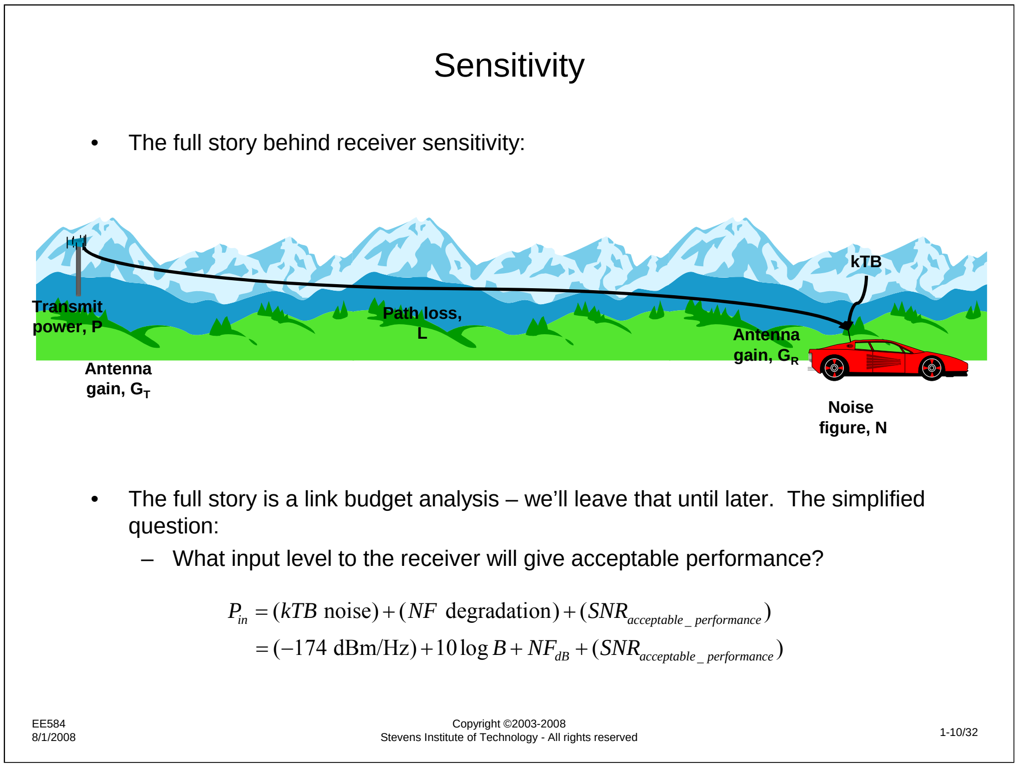
– Noise figure is a measure degradation of the real system to an ideal receiver
    - Caveats!!!: operating temperature, bandwidth, impedance

EE584
8/1/2008

1-9/32

So-called kTB noise is the ideal noise level in a system. In fact, a real system is never ideal, but encounters a higher level of noise which we measure with the system's "noise figure."

The concept of a system's noise figure is quite simple: it is the amount of excess noise generated above what an ideal system would create. In the example above, we compare an ideal receiver with a 20 kHz bandwidth at room temperature (with an input impedance that is the same as the real receiver). An ideal analysis might tell us that the perfect receiver could produce a particular bit error rate with a particular transmit power and path loss. We can compare the real receiver and, assuming we have calibrated for any other imperfections (actual receiver bandwidth, performance of demodulator, etc.) we can determine at what transmit power level and path loss we get the expected performance. The difference in signal level tells us that the real receiver is not performing as well as the ideal receiver. The difference in dB is the system noise figure.



Sensitivity

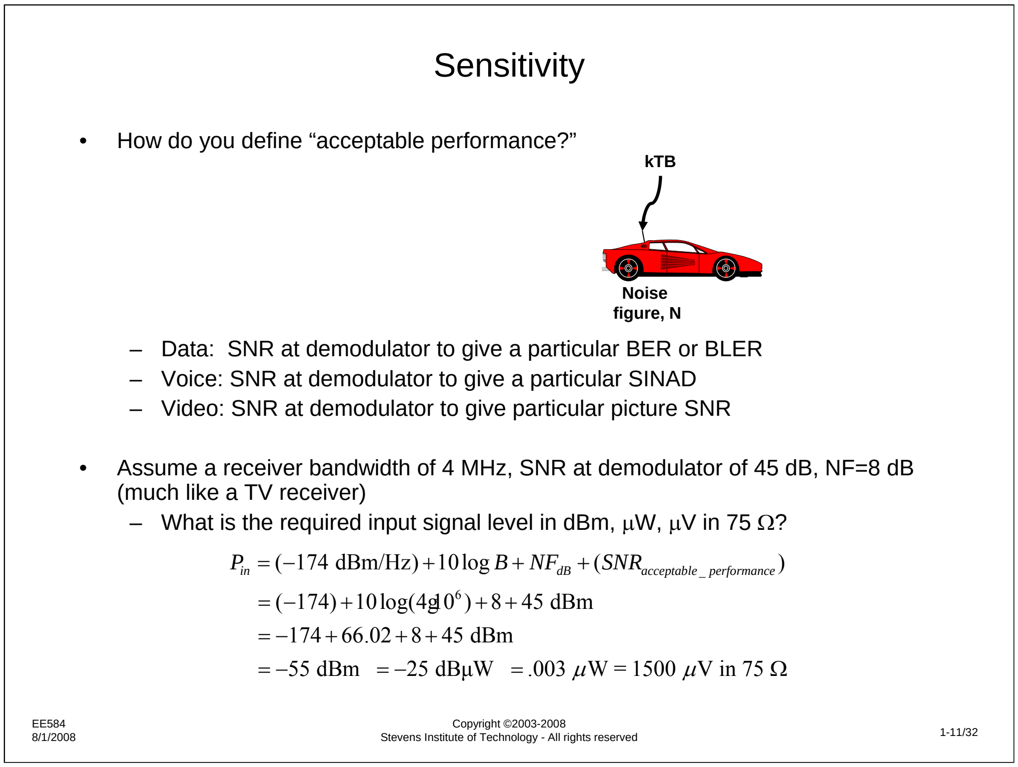
- The full story behind receiver sensitivity:

- The full story is a link budget analysis – we'll leave that until later. The simplified question:
  - What input level to the receiver will give acceptable performance?

$$P_{in} = (kTB \text{ noise}) + (NF \text{ degradation}) + (SNR_{acceptable\_performance})$$
$$= (-174 \text{ dBm/Hz}) + 10\log B + NF_{dB} + (SNR_{acceptable\_performance})$$

We use the ideal noise level in combination with the system noise figure to address the "sensitivity" of a receiver to a signal.

To define receiver sensitivity, we take the receiver out of the environment and describe a lab-based experiment: considering the noise figure, receiver bandwidth, and SNR needed, what is the minimum input signal level that would suffice? While this is a straightforward definition, there are two pitfalls – specifying how one arrives at the SNR for an acceptable performance level and determining the receiver bandwidth. Generally, the first would be specified by a standard or customer requirements. The second can be complicated to determine for some systems.



Sensitivity

- How do you define “acceptable performance?”

  – Data: SNR at demodulator to give a particular BER or BLER
  – Voice: SNR at demodulator to give a particular SINAD
  – Video: SNR at demodulator to give particular picture SNR

- Assume a receiver bandwidth of 4 MHz, SNR at demodulator of 45 dB, NF=8 dB (much like a TV receiver)
  – What is the required input signal level in dBm, μW, μV in 75 Ω?

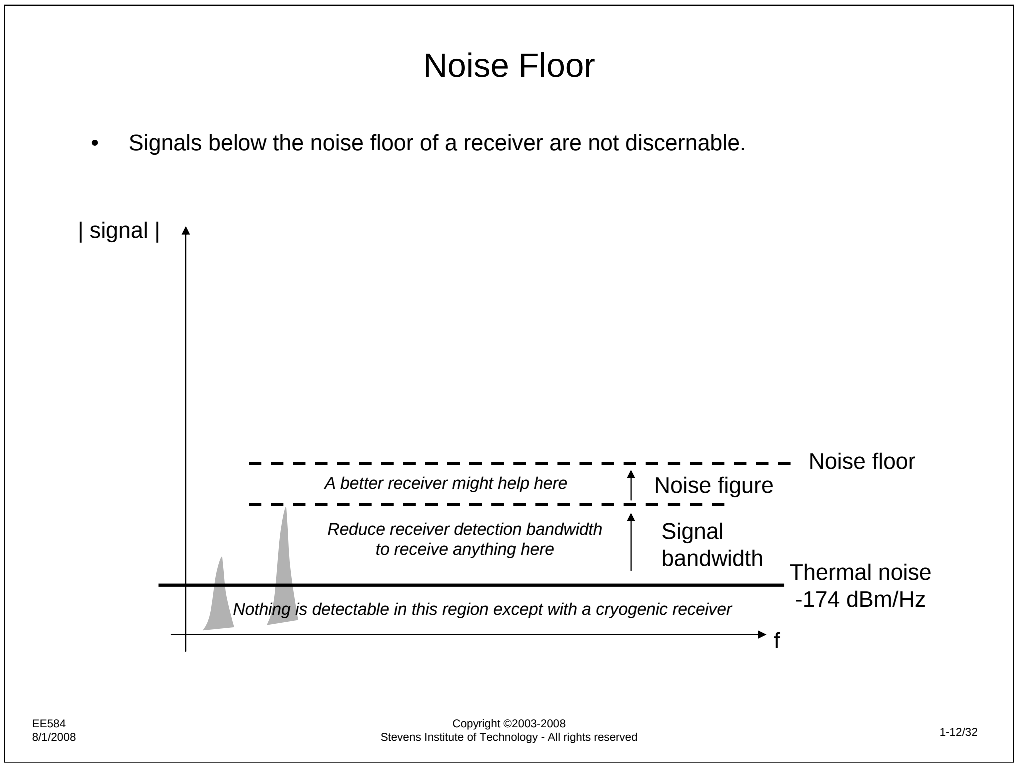
$$
\begin{aligned}
P_{in} &= (-174\ \text{dBm/Hz}) + 10\log B + NF_{dB} + (SNR_{acceptable\_performance}) \\
&= (-174) + 10\log(4g10^{6}) + 8 + 45\ \text{dBm} \\
&= -174 + 66.02 + 8 + 45\ \text{dBm} \\
&= -55\ \text{dBm} \quad = -25\ \text{dB}\mu\text{W} \quad = .003\ \mu\text{W} = 1500\ \mu\text{V in } 75\ \Omega
\end{aligned}
$$

EE584
8/1/2008

1-11/32

Typically, there will be standard measures for “acceptable performance.” For data, Bit Error Rate or Block Error Rate are typical metrics. A system might specify that the BLER is to be less than 1 in 100 or BER must be better than 10^{-6}. For analog speech communications systems, output SINAD (Signal to Interference, Noise and Distortion) is a typical measure. A system specification may state that the output SINAD must be better than 10 dB when the input signal is at the specified level of receiver sensitivity.

Taking the parameters for a typical TV broadcast system, we can calculate the receiver sensitivity, normally specified in dBm or microvolts. As shown, a signal level of –55 dBm would be needed to produce the desired video SNR. By comparison, an AMPS receiver, operating in a 30 kHz bandwidth, the same noise figure, and a 10 dB carrier to noise ratio at the input to the demodulator would have a sensitivity of –111 dBm. The reason for the 60 dB difference in signal level is twofold – first, because the required SNR for voice is lower than that for video, we get 35 dB improvement. Second, the video signal is 4 MHz wide, while the speech signal is 30 kHz wide. This is more than a 21 dB difference.



Noise Floor

- Signals below the noise floor of a receiver are not discernable.

| signal |

Noise floor

A better receiver might help here

Noise figure

Reduce receiver detection bandwidth to receive anything here

Signal bandwidth

Thermal noise -174 dBm/Hz

Nothing is detectable in this region except with a cryogenic receiver

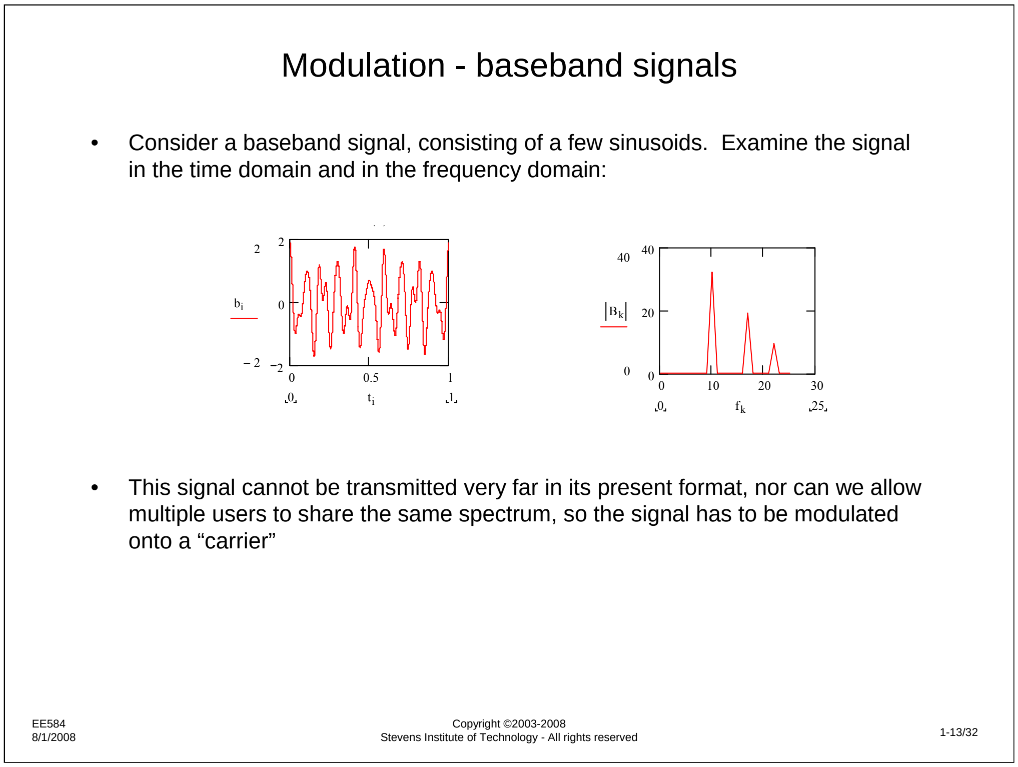
f

EE584
8/1/2008

1-12/32

The noise floor of a receiver is the level of noise generated by the receiver, kTB plus the noise figure. For a given receiver, no signal below the noise floor can be detected. If the noise figure were improved and/or the receiver bandwidth were reduced, lower level signals would be detectable, but then the receiver noise floor would be reduced, since these are the parameters that determine where the noise floor will be.

We will use noise floor later as we examine how different modulation schemes interact with the noise in the presence of enemies who may be trying to intercept a signal or interfere with it.



Modulation - baseband signals

- Consider a baseband signal, consisting of a few sinusoids. Examine the signal in the time domain and in the frequency domain:

2
2
b_i
0
− 2
−2
0
0.5
1
0
t_i
1
40
40
$|B_k|$
20
0
0
0
10
20
30
0
f_k
25

- This signal cannot be transmitted very far in its present format, nor can we allow multiple users to share the same spectrum, so the signal has to be modulated onto a “carrier”

EE584
8/1/2008

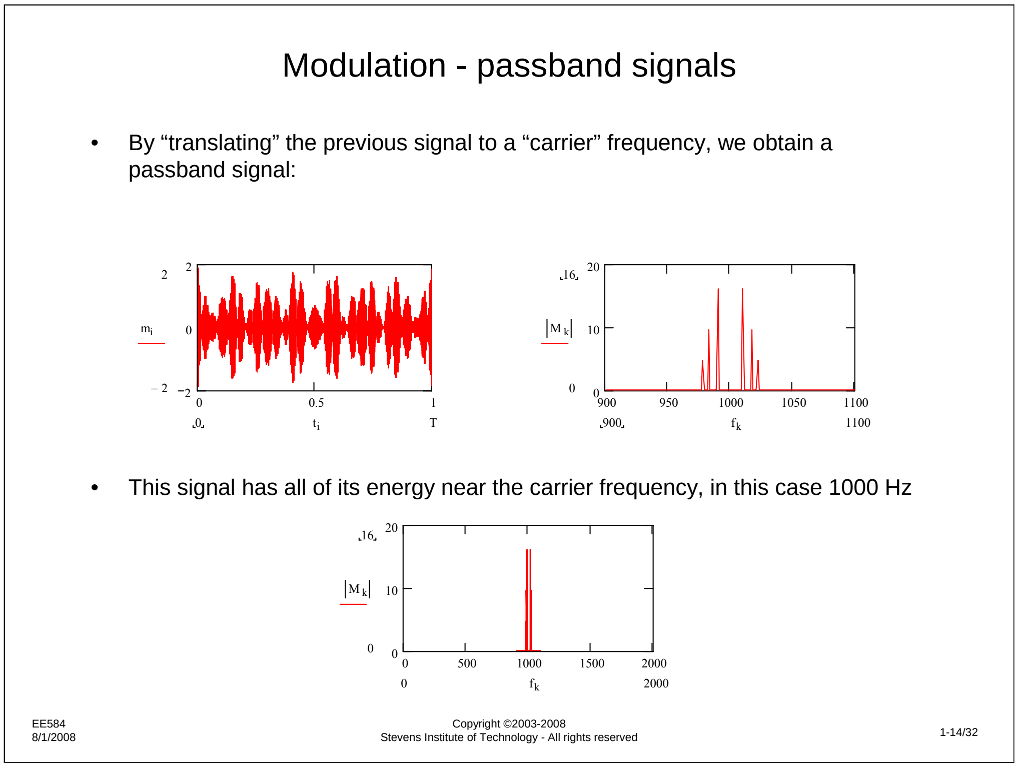
1-13/32

To be able to transmit a signal over a wireless media, we must change it into a format that is suitable for (potentially long distance) transmission over the media. This is the function of modulation.

A baseband signal is a signal that has its energy concentrated near zero frequency. The signal might be narrowband, e.g., a 3 kHz voice signal, or it could be wideband (or broadband), e.g., a 5 MHz video signal, but the important aspect of a baseband signal is that it has negligible energy above some upper limit.

Note: for the next several topics, I will be showing the positive frequency components of the signal. In fact, since the signals we deal with will typically be real-valued signals, their spectrum must be symmetrical around zero, so there are corresponding negative frequency components. Until we discuss systems that are sensitive to the phase of the signal, we can neglect the negative frequencies.

I have chosen a set of baseband frequencies and amplitudes to show the overall shape of the signal clearly in the frequency domain. You will see this shape on the next slide, when we look at passband signals.



Modulation - passband signals

- By “translating” the previous signal to a “carrier” frequency, we obtain a passband signal:

- This signal has all of its energy near the carrier frequency, in this case 1000 Hz

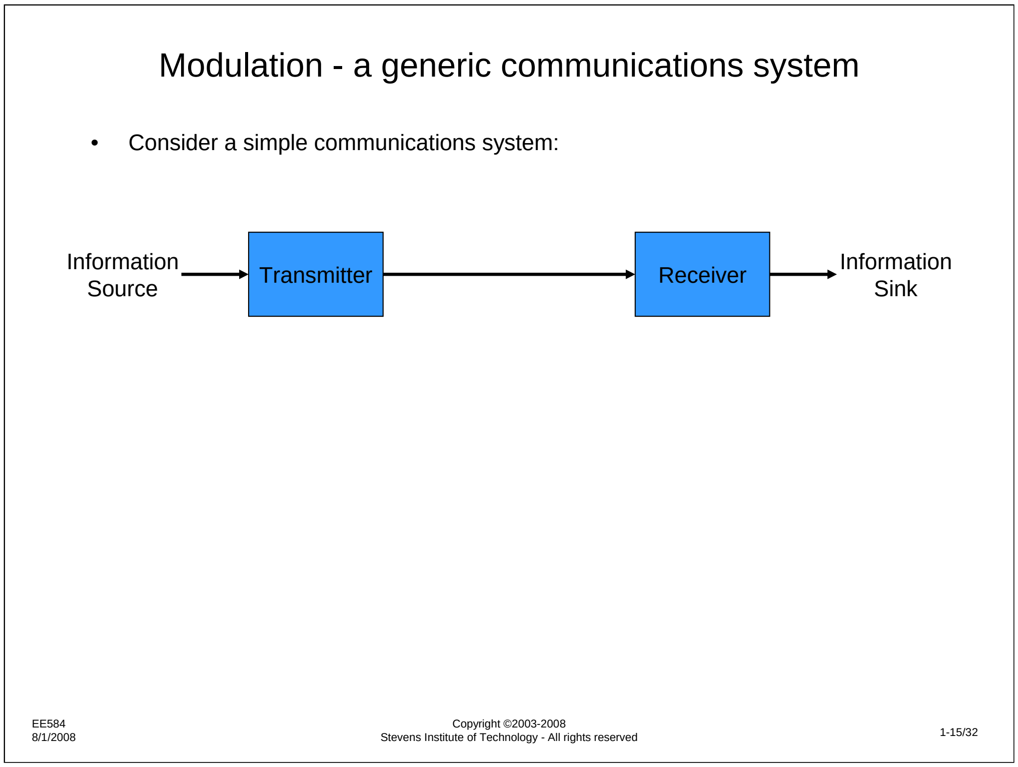
EE584
8/1/2008

1-14/32

In contrast, a passband signal has energy concentrated around some non-zero frequency. A passband signal has zero energy below some lower frequency limit and above some upper frequency limit. Generally, the extent of the lower to the upper frequency limit of the signal is referred to as its “bandwidth.” The center frequency of the passband signal is generally referred to as the carrier frequency. For the most part, the carrier frequency of RF systems will be much greater than the bandwidth, but in later sessions, we will see some instances where we may be dealing with signals that have center frequencies that are close to the signal bandwidth.

Here you will notice that the shape of the previous baseband signal spectrum is preserved through the translation I have performed on the signal frequency. If you look carefully, you will see that the passband signal I have generated is symmetrical around the 1000 Hz carrier. In fact, this is because the baseband signal was symmetrical around zero frequency. As you will see later, the method used to translate the signal to the carrier frequency preserves the spectrum.

However, while this process of translating the baseband signal to a passband signal happens to preserve the spectrum, this is not a requirement when we create a passband signal for transmission.



Modulation - a generic communications system

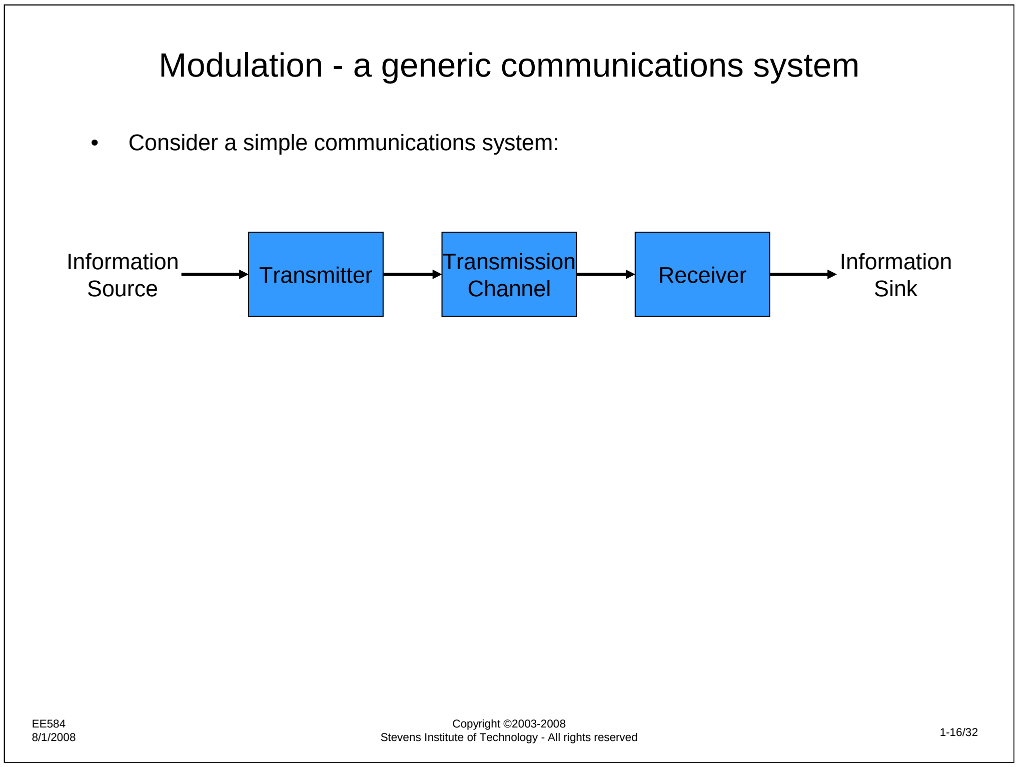
- Consider a simple communications system:

Information Source → Transmitter → Receiver → Information Sink

EE584
8/1/2008

1-15/32

Here we have almost the simplest type of communications system imaginable - the information source generates information that we want to deliver to the information sink. The source might be anything from human speech to computer data to a video signal. To make the example relevant to this course, we have to design some type of transmitter and receiver.



Modulation - a generic communications system

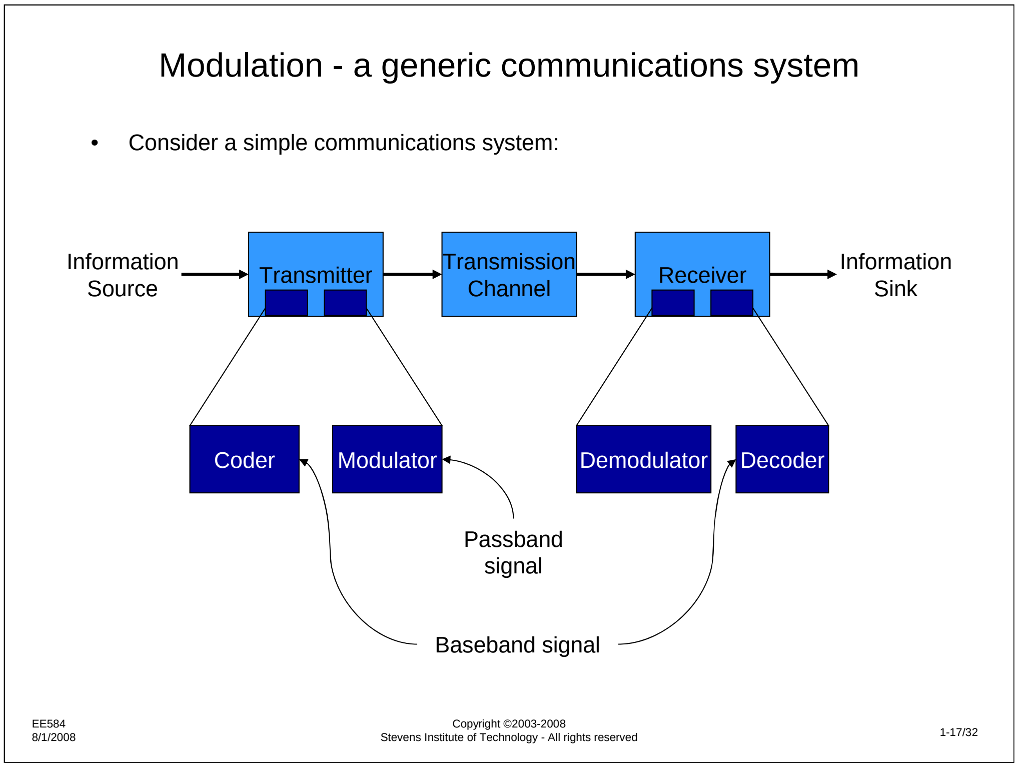
- Consider a simple communications system:

Information Source → Transmitter → Transmission Channel → Receiver → Information Sink

EE584
8/1/2008

1-16/32

Practically, the signal from the transmitter will need to be transmitted through a channel. Since we are designing RF communications systems, we will assume that the channel has a variety of impairments - noise, interference, distortion, fading, etc. Later on, we will get into more detail about what the impairments might be and how they influence our design of the transmitter and receiver.



Modulation - a generic communications system

- Consider a simple communications system:

Information Source

Transmitter

Transmission Channel

Receiver

Information Sink

Coder

Modulator

Demodulator

Decoder

Passband signal

Baseband signal

EE584
8/1/2008

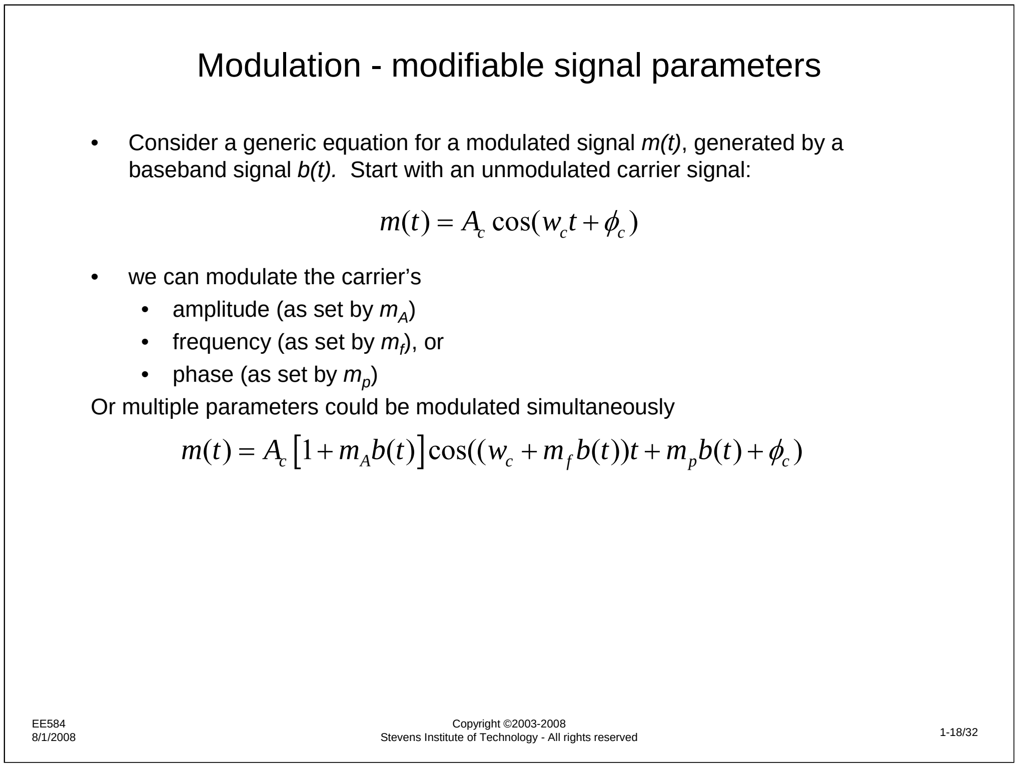
1-17/32

Now, we come back to the previous discussion of baseband and passband signals…

In general, the input source information must be conditioned in some way to deal with the impairments the signal will face in the transmission channel. We might convert an analog source (e.g., human speech) to a digital representation so that transmission errors can be controlled more effectively. Or, we might have a digital source (e.g., computer data) that needs to have redundancy added to protect it from the noise and interference that exists on the channel.

For the most part, we can consider this pre-processing of the input information to be a coding operation that occurs in the transmitter's coder block. At the receiver, there is a corresponding process, the decoder, which converts the coded information back to the original format for delivery to the information sink. The output of the coder and the input to the decoder will generally be baseband signals.

These baseband signals are transformed into passband signals through the processing performed by the modulator and, conversely, the passband signal at the receiver is converted to baseband through the processing of the demodulator. In subsequent slides, we will see how the various parameters of the passband signal can be modified to convey the baseband information.



Modulation - modifiable signal parameters

- Consider a generic equation for a modulated signal *m(t)*, generated by a baseband signal *b(t).* Start with an unmodulated carrier signal:

$$m(t) = A_c \cos(w_c t + \phi_c)$$

- we can modulate the carrier's
  - amplitude (as set by m_A)
  - frequency (as set by m_f), or
  - phase (as set by m_p)

Or multiple parameters could be modulated simultaneously

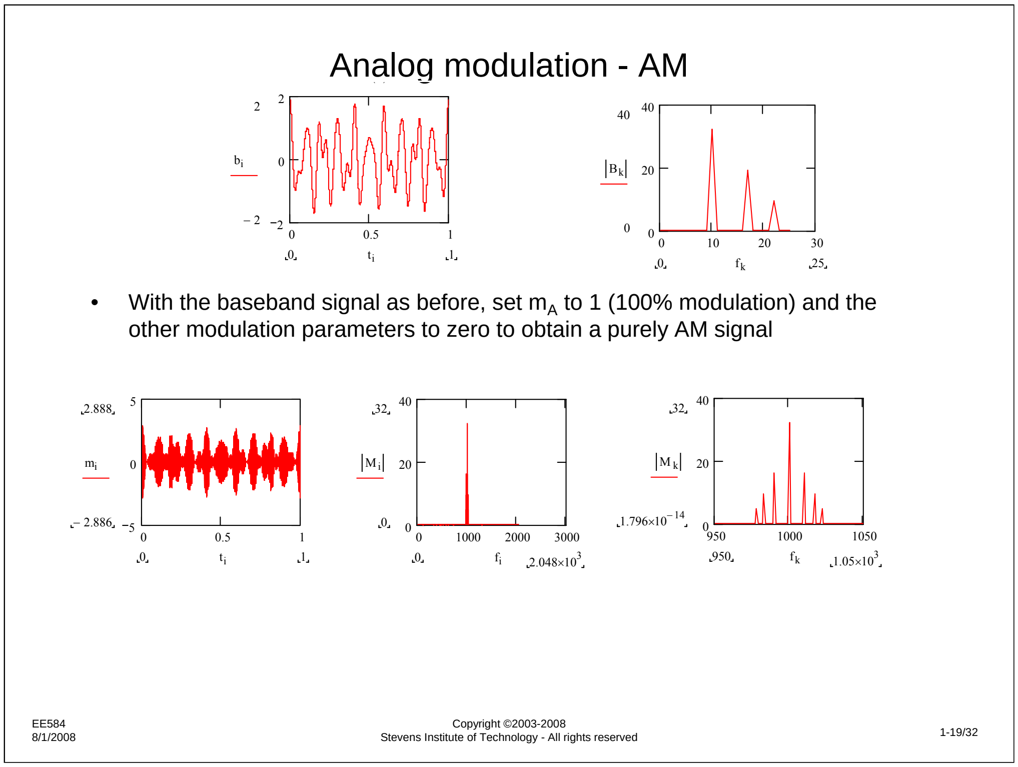
$$m(t) = A_c \left[1 + m_A b(t)\right] \cos((w_c + m_f b(t))t + m_p b(t) + \phi_c)$$

Here, I have written the equation for an unmodulated sinusoidal carrier - there are three parameters for the signal - its amplitude, its frequency and its phase.

Each of these parameters could be modified or modulated signal by a modulating signal b(t). We will define three constants to define the level of modulation for each parameter.

While this equation describes the theoretical modulation of the carrier, you will see that there are different ways of physically realizing the modulation process. Likewise, there will be different ways to demodulate (or detect) the modulated signal. Most important, each modulation technique creates different considerations in the design of the transmitter and receiver, which we will discuss in depth.

Lastly, in the current discussion, I have not put any constraints on b(t) - is it an analog signal or a digital signal? We will treat the two separately, but you will see that they are closely related to each other.



Analog modulation - AM

- With the baseband signal as before, set m_A to 1 (100% modulation) and the other modulation parameters to zero to obtain a purely AM signal

EE584
8/1/2008

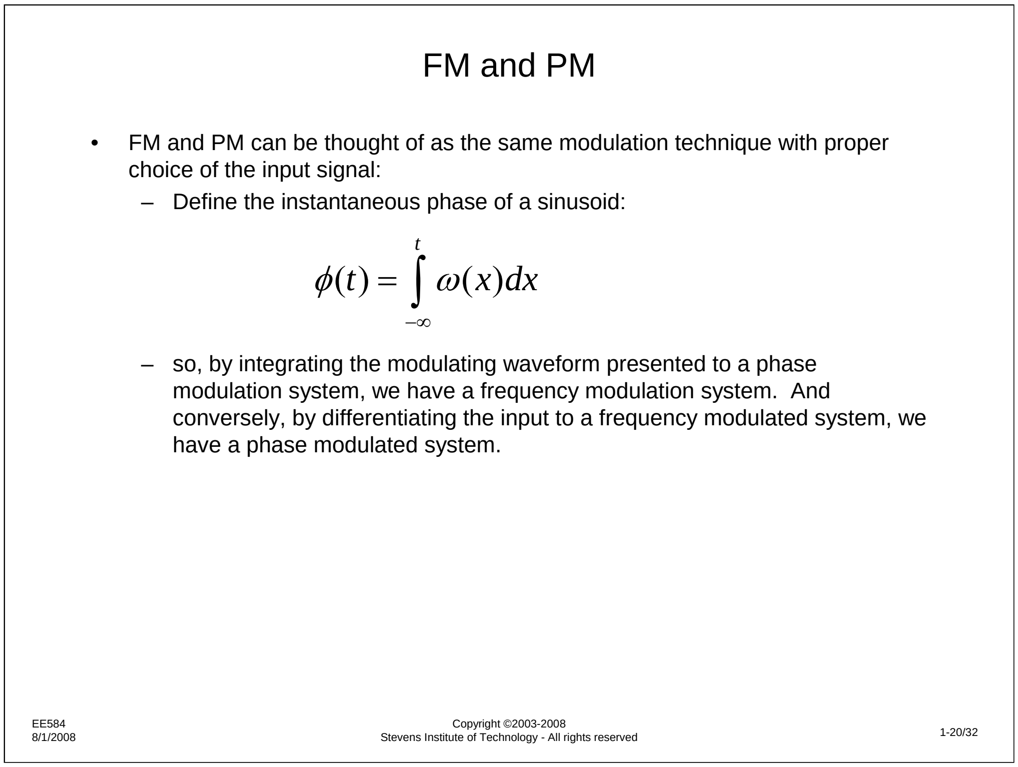
Copyright ©2003-2008
Stevens Institute of Technology - All rights reserved

1-19/32

Let’s assume for now that the input to the modulator is an analog signal, perhaps speech, or perhaps the color and luminance signals from a video camera. Later we will deal with signals that are either inherently digital or have been converted into digital signals.

Let’s again consider the waveforms we used to discuss baseband signals.

Modulating the carrier’s amplitude, we see a replication of the baseband signal in the amplitude of the carrier. In the frequency domain, we see that there is energy at, and perhaps near the carrier frequency. Expanding this, near the carrier frequency, we see the baseband signal spectrum is generated above and below the carrier frequency, but there is a strong component of the carrier present. In fact 50% of the total transmit power is present in the carrier. While this carrier being present permits a simpler receiver design, it wastes half the transmit power, since the carrier actually conveys no information. Also, it turns out that the two copies of the baseband signal, one above and one below the carrier, are identical, so they both convey the same information. While it is possible to transmit this signal more efficiently, we won’t get into this right away. This mode of AM transmission is known as Single Sideband with Suppressed Carrier (SSB-SC) or, more typically, Single Sideband (SSB).



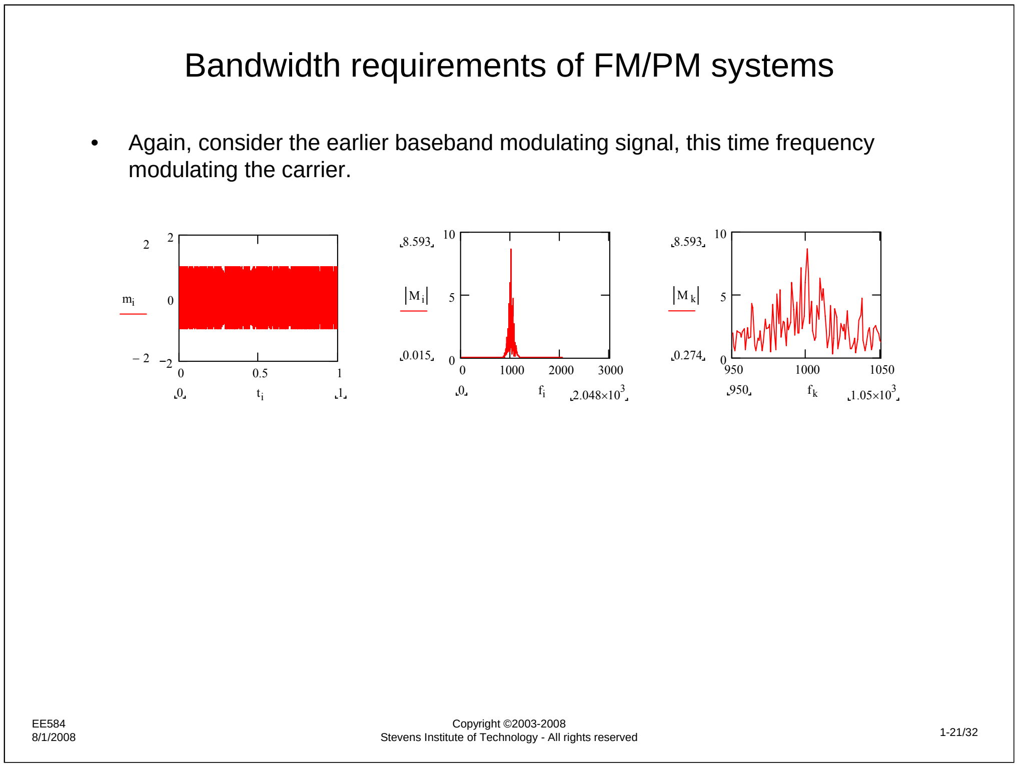
FM and PM

- FM and PM can be thought of as the same modulation technique with proper choice of the input signal:
  - Define the instantaneous phase of a sinusoid:

$$\phi(t) = \int_{-\infty}^{t} \omega(x)dx$$

  - so, by integrating the modulating waveform presented to a phase modulation system, we have a frequency modulation system. And conversely, by differentiating the input to a frequency modulated system, we have a phase modulated system.

Since the frequency and phase of a sinusoid are so closely related to each other (the instantaneous frequency is the derivative of the phase and the phase is the integral of the instantaneous frequency), there is generally an equivalent waveform that could be input to a PM system to generate FM and vice versa. For instance, to generate a sine modulated FM signal, we could phase modulate with a cosine. A more interesting example is an FM system modulated with a square wave, versus a PM system modulated with a triangular wave. In either case, the resulting signals are the same.



Bandwidth requirements of FM/PM systems

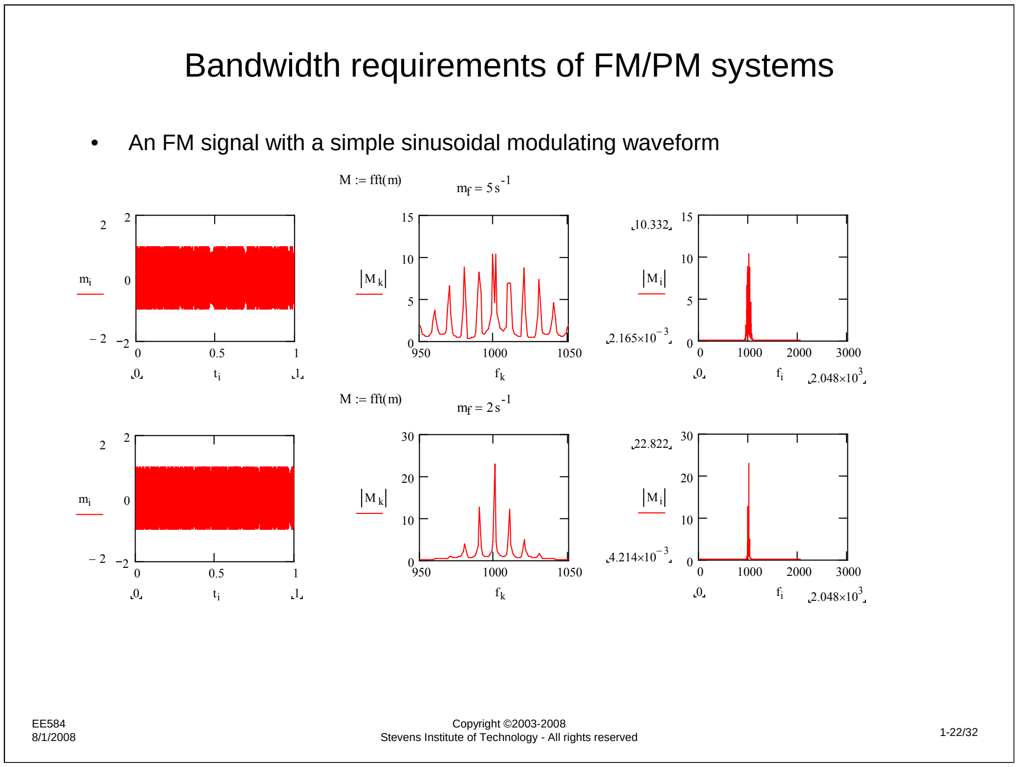
- Again, consider the earlier baseband modulating signal, this time frequency modulating the carrier.

EE584
8/1/2008

1-21/32

The first thing to note here is the constant amplitude of the modulated signal. This makes sense, since we are not modulating the amplitude of the carrier. If we looked closely at the zero crossings of the FM carrier, we would see that they become slightly closer and slightly farther apart, depending on the amplitude of the modulating signal.

What is of greater interest is the spectrum of the FM signal. It is obviously much broader than it was for the AM signal. In fact, if you look at the energy near the carrier, it almost appears random. This is partly because of the modulating waveform I chose previously.



Bandwidth requirements of FM/PM systems

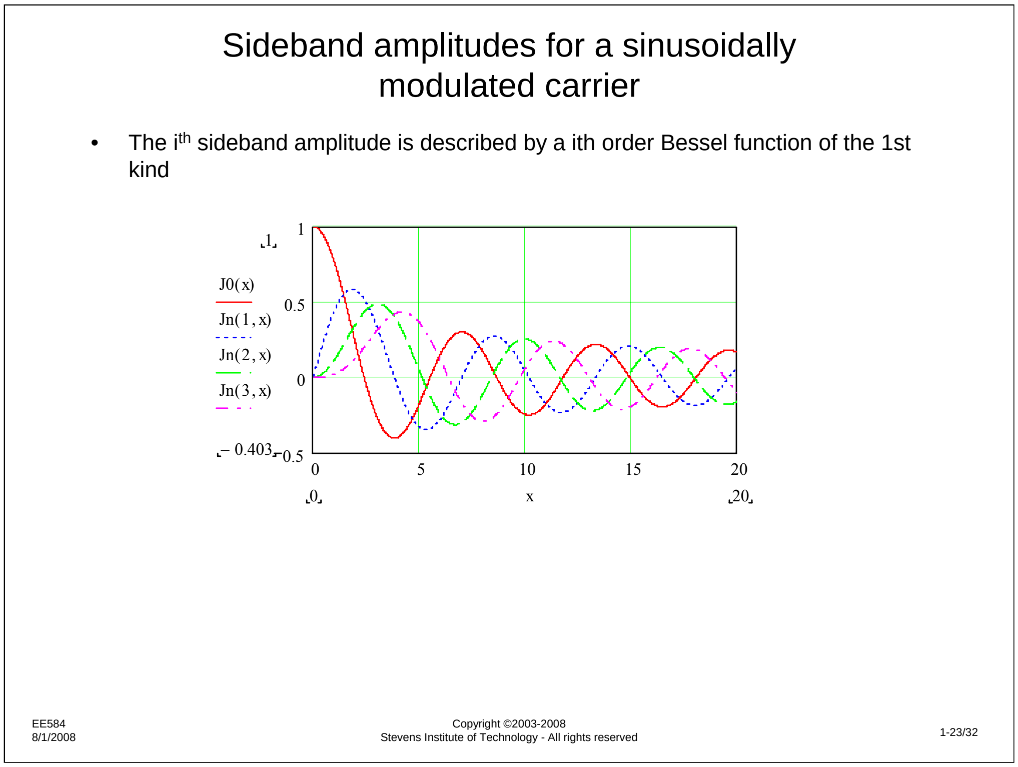
- An FM signal with a simple sinusoidal modulating waveform

EE584
8/1/2008

1-22/32

With a simple sinusoid modulating the carrier, again, the amplitude remains constant. However, the spectrum becomes much simpler. Here, the modulating sinusoid is at 10 Hz. We see that there are sidebands around the carrier spaced every 10 Hz. However, as we adjust the modulating factor (listed above the center plot) we see that the sidebands are higher amplitude for a higher modulating factor.

In fact, it can be shown that the level of the sidebands are related in a complex fashion to the modulation. For the case of simple sinusoidal modulation, the amplitude of the carrier and each sideband is defined by a series of Bessel functions, related to the modulating frequency and the modulating factor.



Sideband amplitudes for a sinusoidally modulated carrier

- The i^{th} sideband amplitude is described by a ith order Bessel function of the 1st kind

The x axis represents the “level” of modulation of the carrier - for FM, the modulation index is defined to be the frequency deviation that occurs for the peak sine wave value divided by the modulating frequency of the sinusoid. The y axis represents the amplitude of each sideband, including the 0^{th} sideband, the carrier. So, J0(x), the 0th order Bessel function of the first kind is the carrier amplitude. J1(x), the first order Bessel function of the first kind defines the amplitude of the first sideband - in the case of a sinusoidal modulating signal, this sideband is separated from the carrier by the modulating frequency. Likewise, J2(x) (or Jn(2,x) in the Mathcad notation), represents the second sideband.

A few things worth noting:

(1) The sidebands are symmetrical around the carrier, so the first lower sideband is the same amplitude as the first upper sideband

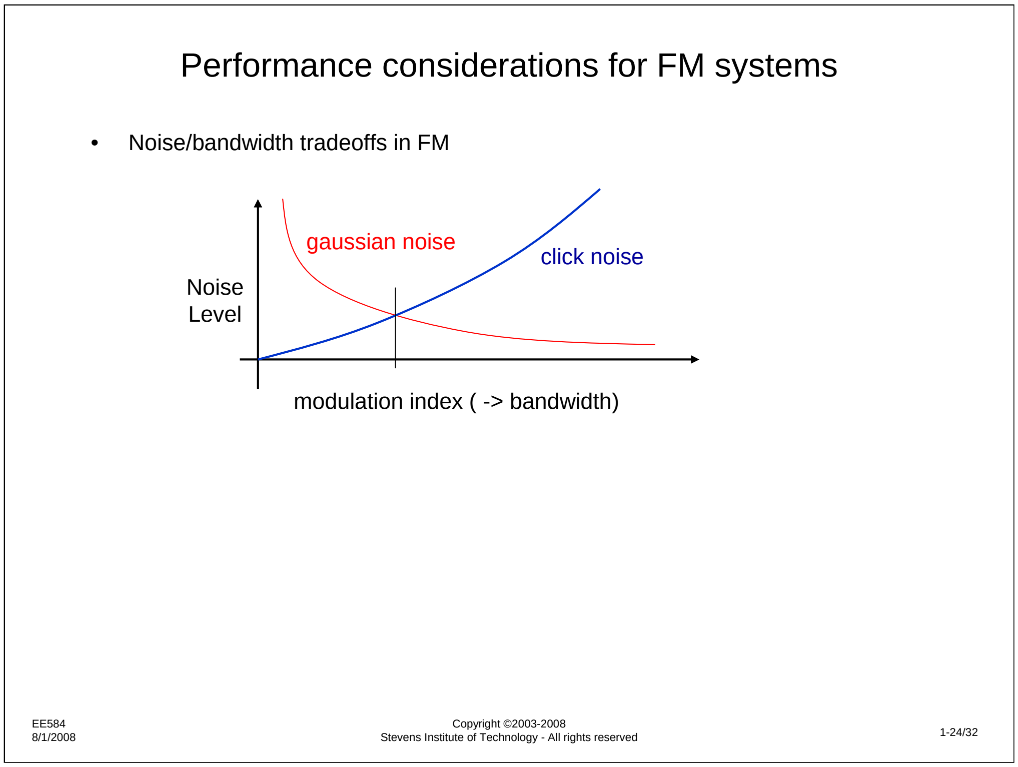
(2) J0(x) goes through zero for some values of x. This means that there are some levels of modulation where the carrier disappears.

(3) For even moderate levels of modulation, there are a LARGE number of non-zero sidebands

(4) As the level of modulation increases, the number of significant sidebands increases.

(5) For very low levels of modulation, only the first few sidebands are significant, which is referred to as narrowband FM (NBFM)

(6) cellular AMPS, VHF military tactical analog systems, VHF/UHF police/fire/emergency services, and most FM amateur radio systems use NBFM



Performance considerations for FM systems

- Noise/bandwidth tradeoffs in FM

EE584
8/1/2008

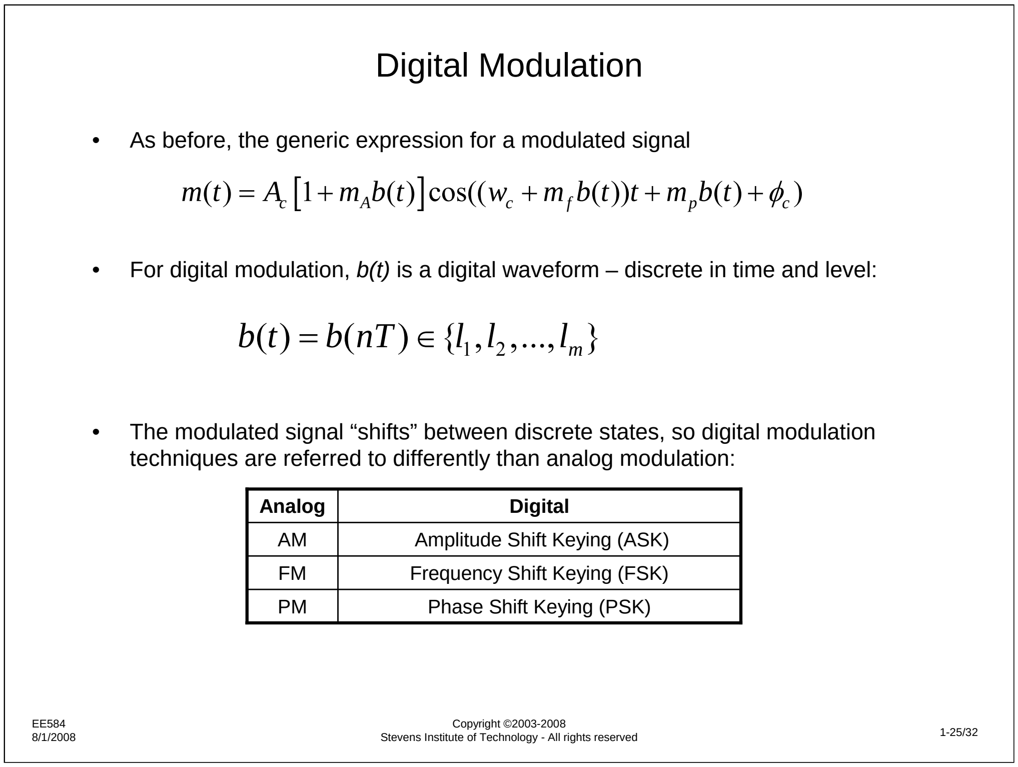
1-24/32

In FM systems, the “intensity” of modulation corresponds to the amount the carrier frequency is shifted for the same input signal level.

Likewise, the signal level output from the demodulator is greater as the modulation level (frequency deviation) is increased.

I’ll leave the details for a communications theory course, but it can be shown that the ultimate signal quality (output SNR) for an FM system increases with a greater frequency deviation for the same input signal frequency. Stated in other words, as the modulation index of FM (the carrier frequency deviation divided by the modulating frequency) increases, the ultimate output SNR for a given carrier to noise ratio (CNR) increases. This is known as the FM improvement effect. But, as we saw earlier, increasing the modulation level (modulation index) increases the bandwidth of the FM signal, so there is a direct tradeoff between signal quality and bandwidth. We can look at this FM improvement effect as if the noise level is fixed (actually based on the bandwidth of the receiver - more on this later) and the increasing output signal level for higher levels of modulation continues to improve the output SNR.

This is not the entire story, however. While the FM improvement is beneficial at high CNR, there is a price to pay at low CNR. Although the high modulation index decreases the effect of the background gaussian noise, we have to contend with something called FM “Click Noise”. Again, we’ll leave the details to a course in communications theory, but the FM demodulator trades the hissing gaussian noise for impulse noise clicks. The rate at which clicks occurs depends on the CNR, but it also depends on the peak frequency deviation - clicks are more likely to occur as the received signal’s instantaneous frequency differs from the center frequency of the receiver more.



Digital Modulation

- As before, the generic expression for a modulated signal

$$m(t) = A_c\left[1 + m_A b(t)\right]\cos((w_c + m_f b(t))t + m_p b(t) + \phi_c)$$

- For digital modulation, *b(t)* is a digital waveform – discrete in time and level:

$$b(t) = b(nT) \in \{l_1, l_2, ..., l_m\}$$

- The modulated signal “shifts” between discrete states, so digital modulation techniques are referred to differently than analog modulation:

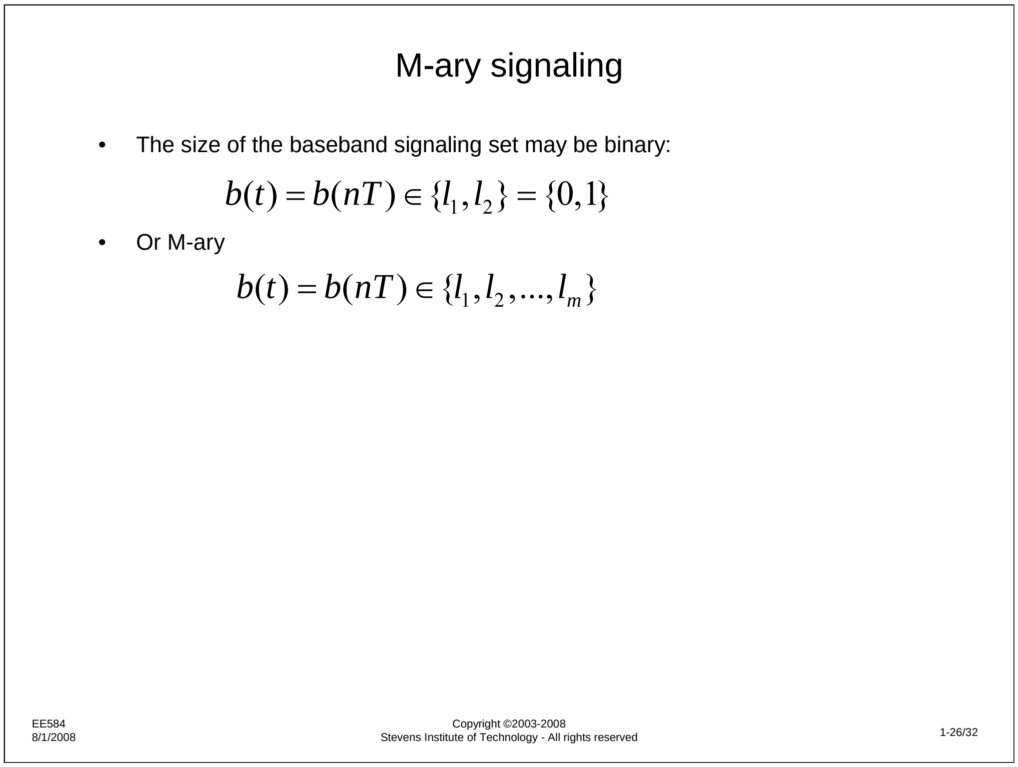
Analog	Digital
AM	Amplitude Shift Keying (ASK)
FM	Frequency Shift Keying (FSK)
PM	Phase Shift Keying (PSK)

Digital modulation relies on the same modifications of the carrier as analog modulation – the distinction is that the baseband modulating signal is digital

The digital baseband waveform changes at discrete points in time (e.g., it is sampled) and is only allowed to assume discrete values (e.g., it is quantized).

Quantization of the baseband signal allows the receiver to make specific decisions about which signal was sent, even in the presence of noise.

What is the advantage of having the modulated signal switch levels at discrete points of time?



M-ary signaling

- The size of the baseband signaling set may be binary:

$$b(t) = b(nT) \in \{l_1, l_2\} = \{0,1\}$$

- Or M-ary

$$b(t) = b(nT) \in \{l_1, l_2, ..., l_m\}$$

EE584
8/1/2008

Copyright ©2003-2008
Stevens Institute of Technology - All rights reserved

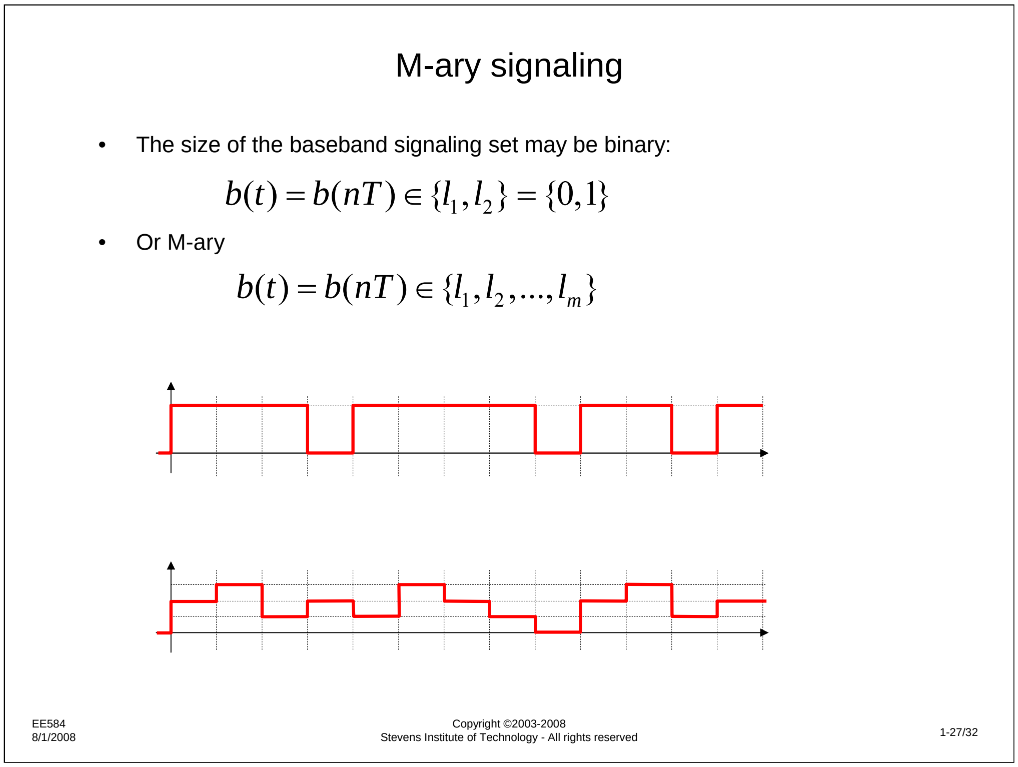
1-26/32

From the previous expression for the baseband waveform, b(t), the signal is allowed to take on a discrete set of levels: l(sub)1,…,l(sub)n.

Frequently, this set is constrained to two levels. Without loss of generality, these levels can be labeled “0” and “1”.

Most frequently, when an m-ary signal is sent with more than two levels, the number of levels, m, is a power of two. This allows the signal to be represented as a binary signal elsewhere in the communications system without complicated translation.

A note on signal representation – when describing a binary signal, it is customary to define the signal in terms of the logic levels – 0 and 1, off and on, FALSE and TRUE. When we consider modulating the signal, it is usually easiest to think of the signal as having a value of –1 or +1. Likewise, if a 4-ary signal has logic levels of {0,1,2,3}, it will be easier to deal with a signal that is symmetrical around zero volts. We might also represent the signal with a maximum range of –1 to +1, so the signal levels become {-1, -1/3, 1/3, +1}. Or, to keep the representation of the signals as simple as possible, {-¾ , -¼ ¼, ¾}. All of these are logically equivalent.



M-ary signaling

- The size of the baseband signaling set may be binary:

$$b(t) = b(nT) \in \{l_1, l_2\} = \{0,1\}$$

- Or M-ary

$$b(t) = b(nT) \in \{l_1, l_2, \ldots, l_m\}$$

EE584
8/1/2008

From the previous expression for the baseband waveform, b(t), the signal is allowed to take on a discrete set of levels: l(sub)1,…,l(sub)n.

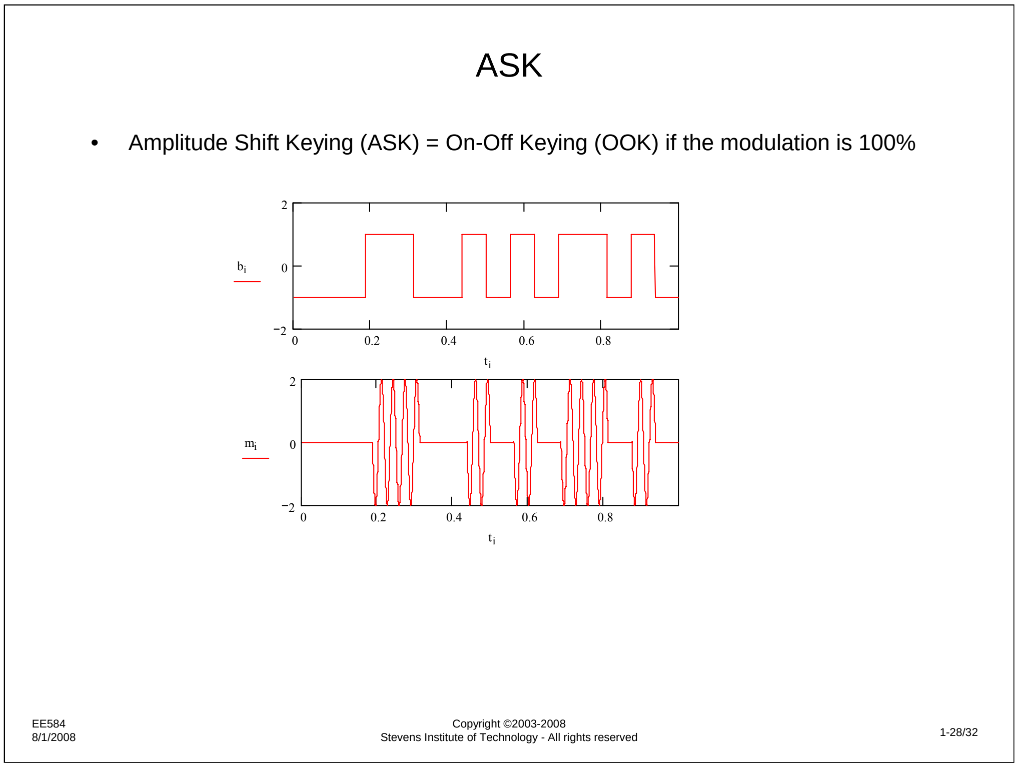
Frequently, this set is constrained to two levels. Without loss of generality, these levels can be labeled “0” and “1”.

Most frequently, when an m-ary signal is sent with more than two levels, the number of levels, m, is a power of two. This allows the signal to be represented as a binary signal elsewhere in the communications system without complicated translation.

The graphs illustrate the baseband modulating waveform for binary (upper) and m-ary (for m=4, sometimes called quatenary).

For the same signaling rate (baud), the 4-level system transmits twice as much information as the binary system, but there is a cost associated with this increased information rate.

The 4-level has levels that are ¼ the spacing of the binary system so, for a given noise level and average signal power level, there is a greater susceptibility to noise.

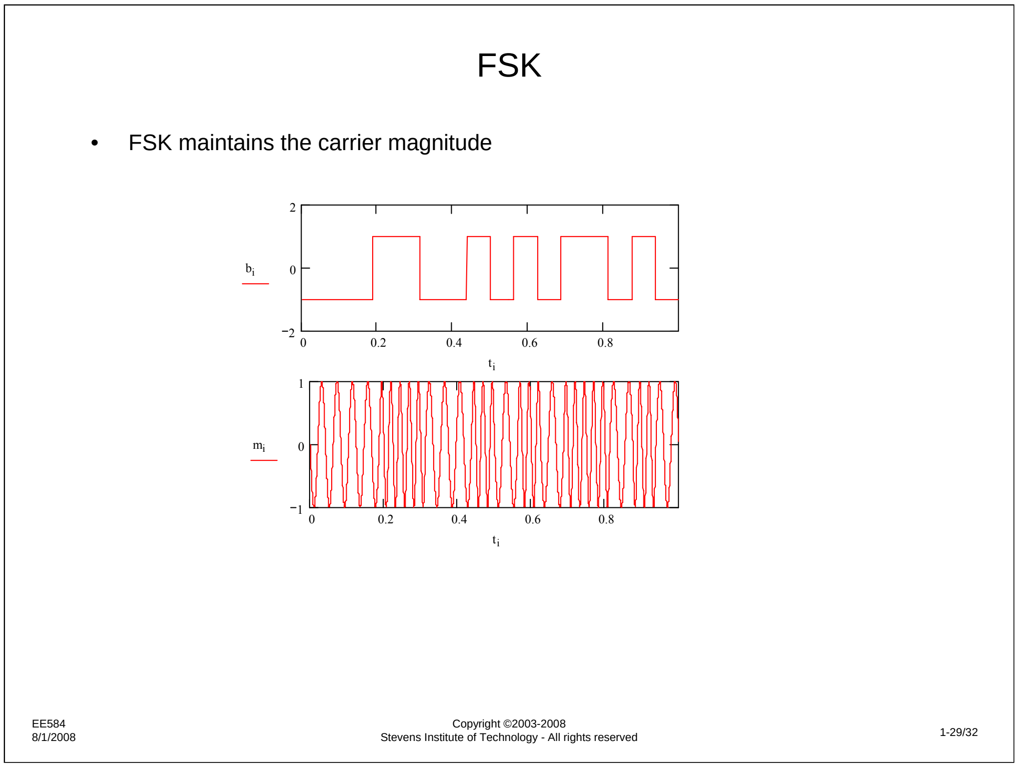


ASK

- Amplitude Shift Keying (ASK) = On-Off Keying (OOK) if the modulation is 100%

Here, I have generated a random binary sequence to serve as a modulating signal. I have also reduced the carrier frequency to a very low frequency so you can see the individual cycles of the waveform. This is going to be more important later.

By modulating at 100% (modulation index of 1), the modulated signal varies between full power and zero power. While this makes the difference between a 0 and 1 maximum, it creates a problem – how do you distinguish between a “0“ and the absence of a signal?



FSK

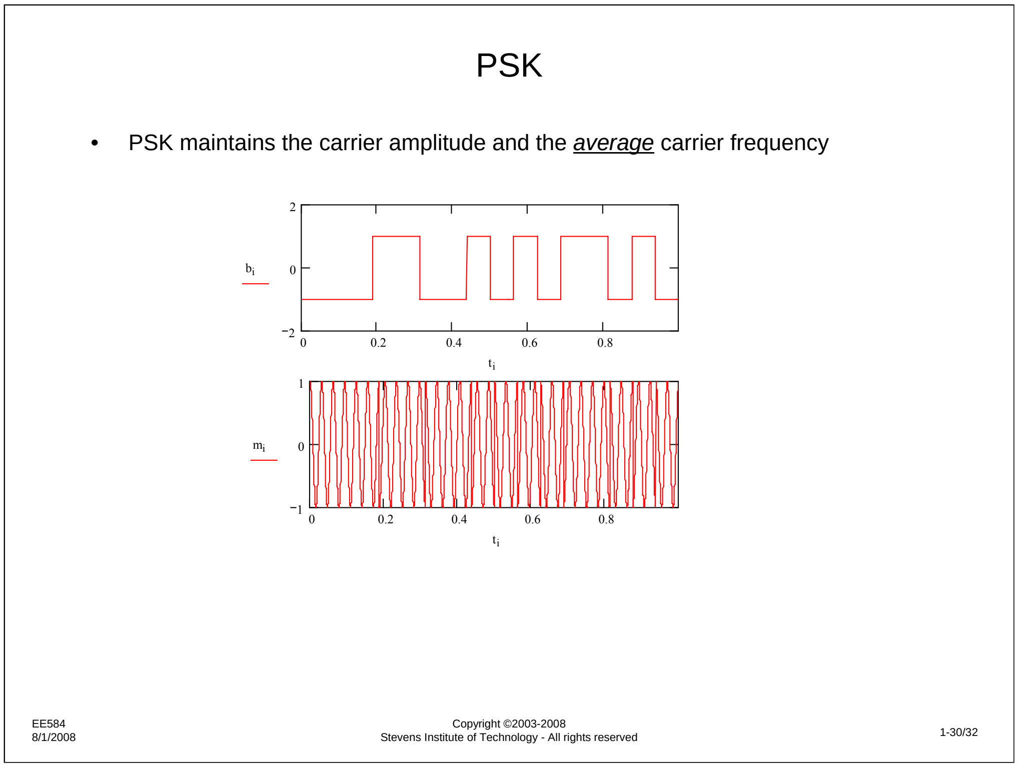
- FSK maintains the carrier magnitude

EE584
8/1/2008

1-29/32

Again, I have used the same random binary sequence, this time to frequency modulate the carrier. Since the carrier frequency is low enough to see the individual cycles and the deviation is great enough to see the difference between the mark (1) and space (0) frequency, the modulation can clearly be seen on the carrier waveform.

As it was the case for ASK, I have chosen the overall frequency deviation carefully to maintain the continuity of the waveform. After we see the PSK waveform, I will show how the carrier frequency and modulating rate are related to each other.



PSK

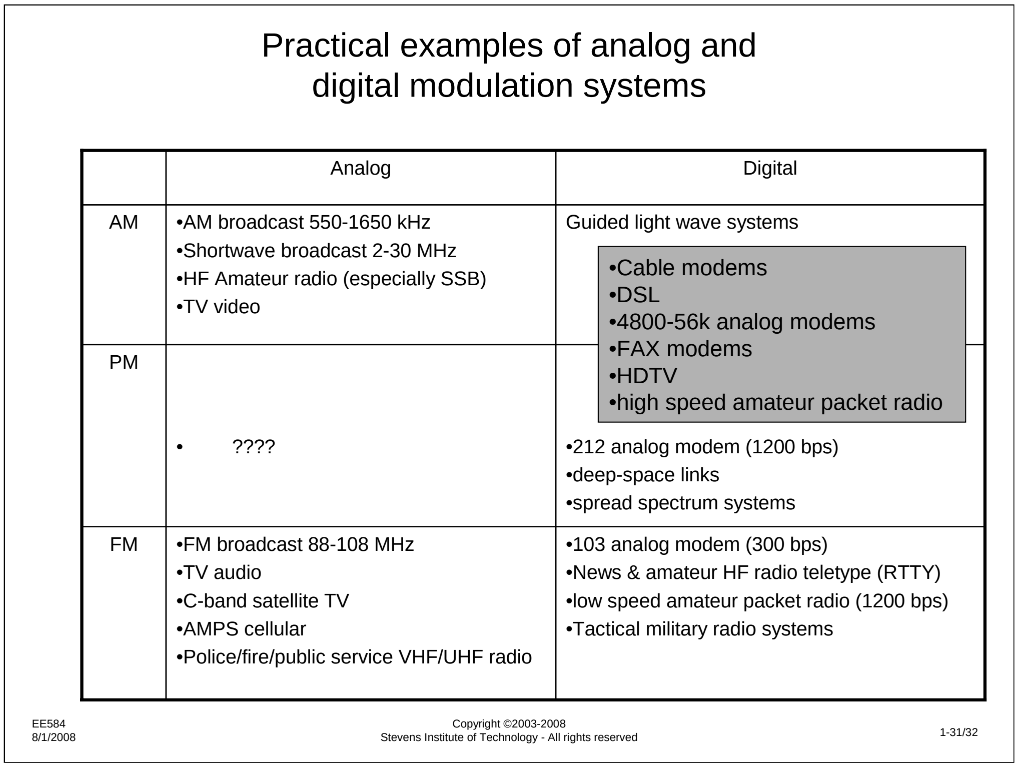
- PSK maintains the carrier amplitude and the *average* carrier frequency

EE584
8/1/2008

1-30/32

In the PSK case, it can be seen that the overall amplitude is constant and the frequency within a bit period is constant. At the bit transitions, the discontinuities in the waveform can be seen, due the instanteous phase shift in the signal.

If the phase shift, carrier frequency and bit period are chosen properly, there will be no first order discontinuity in the waveform, but there will be an instanteous change in slope.



Practical examples of analog and digital modulation systems

	Analog	Digital
AM	•AM broadcast 550-1650 kHz •Shortwave broadcast 2-30 MHz •HF Amateur radio (especially SSB) •TV video	Guided light wave systems •Cable modems •DSL •4800-56k analog modems •FAX modems •HDTV •high speed amateur packet radio
PM	• ????	•212 analog modem (1200 bps) •deep-space links •spread spectrum systems
FM	•FM broadcast 88-108 MHz •TV audio •C-band satellite TV •AMPS cellular •Police/fire/public service VHF/UHF radio	•103 analog modem (300 bps) •News & amateur HF radio teletype (RTTY) •low speed amateur packet radio (1200 bps) •Tactical military radio systems

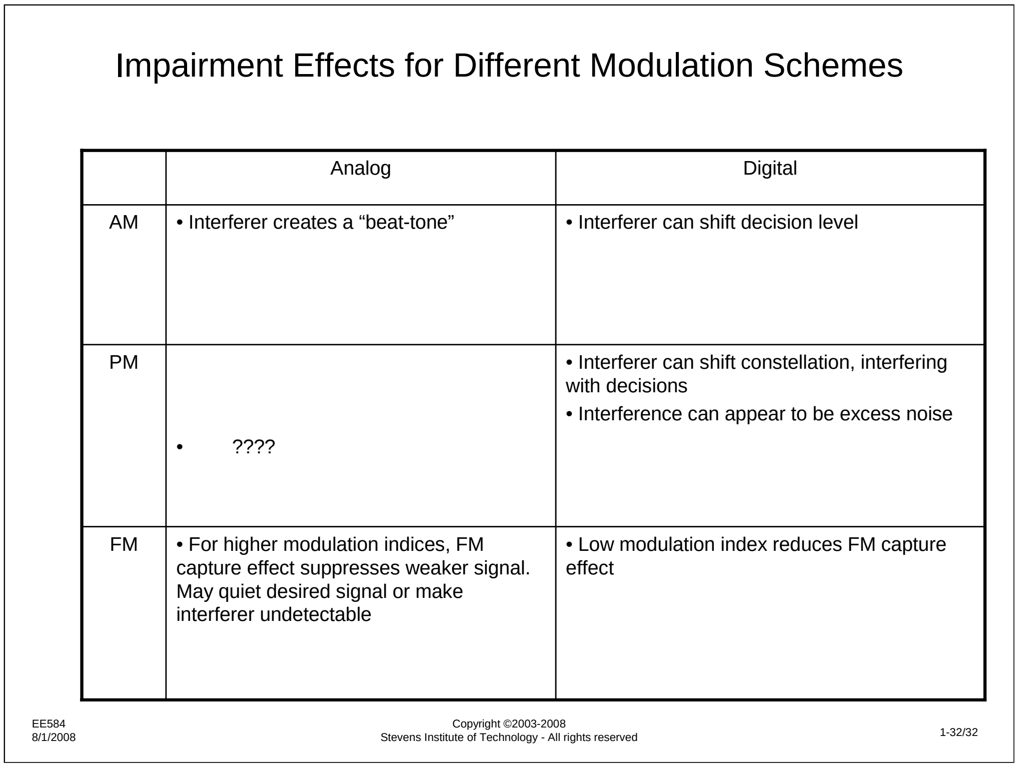
To illustrate the variety of communications systems and the modulation techniques they use, I have created this table.

A few things are worth noting:

(1) I couldn't think of or find any practical analog PM systems. The main reason is that there is no real difference between analog FM and PM. By integrating or differentiating, you could convert one to the other. It is generally easier to modulate and demodulate an FM signal, so system implementers have typically just used FM if they are building an analog angular modulation system

(2) There aren't a lot of AM digital systems - as discussed earlier, there is the problem of drift of the digital reference level as the signal amplitude changes, which is a real problem with RF systems. There is also a serious tradeoff between bandwidth efficiency, SNR and data rate. If a system designer is going to use a system that is amplitude sensitive and requires carrier recovery, there are more efficient systems to use. The one exception isn't applicable to RF - guided light wave systems.

(3) The grey box that spans digital AM and PM systems includes QAM systems, the state-of-the-art in digital communications systems. If you are going to pay for the inconvenience of needing to do carrier recovery, or if your system is sensitive to signal amplitudes, this is the most efficient way to transmit bits. Most of the latest data communications systems are using this modulation technique.



Impairment Effects for Different Modulation Schemes

	Analog	Digital
AM	• Interferer creates a "beat-tone"	• Interferer can shift decision level
PM	• ????	• Interferer can shift constellation, interfering with decisions • Interference can appear to be excess noise
FM	• For higher modulation indices, FM capture effect suppresses weaker signal. May quiet desired signal or make interferer undetectable	• Low modulation index reduces FM capture effect

As we examine the impairment effects for different modulation schemes, keep in mind how these effects might be exploited by an attacker. With this mindset, we can determine if the modulation gives the attacker any particular advantage that must be mitigated, or if there is some inherent advantage to one modulation scheme over the other that might make it more desirable for a secure communication systems.

As an example of the interaction of modulation with security considerations, note how analog FM has a "capture effect," inherent in the modulation. If there is a desired signal at an amplitude A1 and an interferer at an amplitude A2, when A1 > A2, the impact of the interferer can be greatly diminished by the capture effect. For high modulation indices, as little as 1 dB difference between the desired and interferer could result in nearly complete suppression of the interferer. On the other hand, for lower modulation indices, as is used by systems like AMPS cellular, military tactical radio, amateur radio, and public services systems, the capture effect is diminished. For optimal error rate, FSK digital systems also use a low modulation index, so they do not benefit from the capture effect, either.