Modulation & Multiplexing

A Wireless Communication System



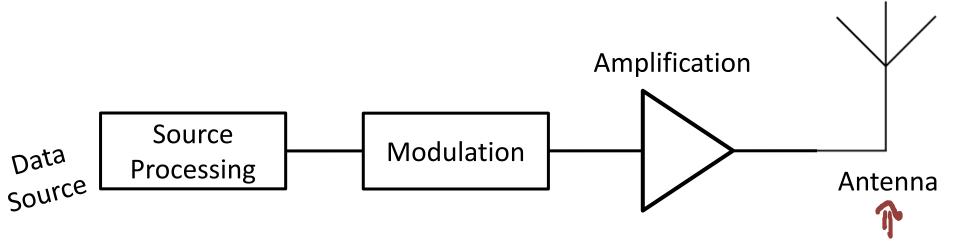
A Wireless Communication System



The role of the transmitter is to:

- convert the source data into a form suitable for transmission;
- to transmit the signal using no more power than is necessary (because high power can be dangerous, cause more interference to other signals, and consume more electrical power);
- be suitable for the channel selected and required data rate (appropriate signal frequency for range, data rate, wireless propagation characteristics, etc.).

The Transmitter Architecture



The transmitter normally has **four functions**:

- 1. Source Processing (preparation of data for reliable delivery, compression etc.);
- 2. Modulation (manipulate a signal to carry the data);
 - 3. Amplification (scale the signal up to an appropriate strength);
 - 4. Interface to the Channel (this means inserting the signal into the channel e.g., an antenna for wireless transmission).

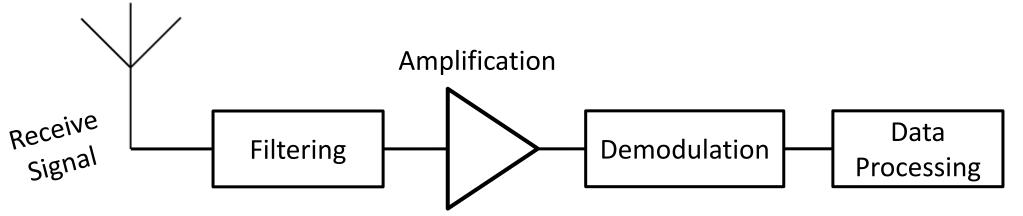
A Wireless Communication System



The role of the receiver is:

- to be as simple to implement as possible;
- to convert the received signal from the channel back into the original data (minimising the effect of any noise or corruption);
- to introduce as little distortion or noise into the received signal as possible;
- be suitable for the channel selected and required data rate (appropriate signal frequency for range, data rate, wireless propagation characteristics, etc.).

The Receiver Architecture



The receiver normally has **five functions**:

- 1. Interface to channel (for example an antenna for wireless transmission);
- 2. Filtering (this is undertaken to remove as much noise as possible that is not in the frequency band of interest to us);
- Amplification (it always helps especially important to do it as soon as possible);
- Demodulation (reverse the modulation process as efficiently as possible);
- 5. Data Processing.

Modulation

Modulation is the process of changing the parameters of the **carrier signal**, in accordance with the instantaneous values of the **modulating signal**.

Modulation is important for a number of reasons:

- To overcome equipment limitations (modulation is used to place signals in a portion of the spectrum where equipment limitations are minimal or most easily met);
- For frequency assignment (frequency bands to each radio application);
- Ease of radiation (antennas size depends on the wavelength).

We will be focusing on *Amplitude Modulation (AM)* and *Frequency Modulation (FM) communication schemes*. They have been improved on over the years, but most schemes can be traced back to these two methods.

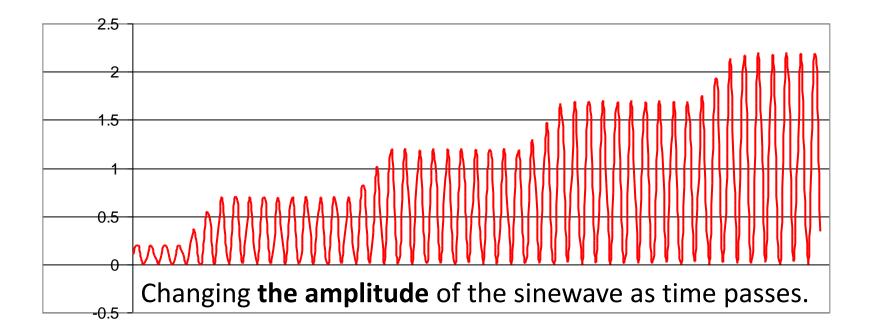
Amplitude Modulation (AM)

Modulating the Signal

A sinusoidal signal can be modulated in different ways:

$$A\sin(2\pi f t + \phi)$$

We could change the amplitude, increasing and decreasing it so as to make it represent some data. This is called **AMPLITUDE MODULATION**.



In amplitude modulation we modify the amplitude of the scaling factor according to the input signal, almost always in proportion to the input signal.

If S(t) is our information signal, then we say

$$y(t) = \left[1 + k_a s(t)\right] \left[A \sin(2\pi f_c t + \phi)\right]$$

wave and we normally assume A = 1.

This is just a sine

where we say

 $k_{\rm a}$ is our amplitude sensitivity (an amplification factor),

 $f_{\rm c}$ is our carrier frequency.

Remember: the carrier wave is nothing more than a simple up-down sinusoidal signal.



Conditions for this to work:

- The carrier frequency must be much higher than the highest frequency component of your signal.
 (highest frequency -> how fast it can potentially change)
- The amplitude sensitivity (k_a) should be chosen such that the total amplitude scaling factor is always **POSITIVE**, at all times, whatever the input signal may be. $\left[1+k_a s(t)\right] \geq 0 \quad \forall t$
- The ratio of the peak variation of our modulating signal to our carrier amplitude is called the **PERCENTAGE MODULATION**:

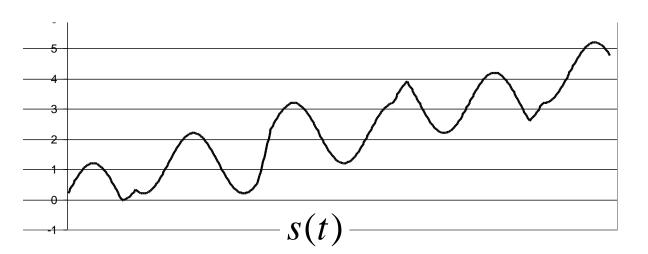
$$\frac{\left|\max\left(k_a s(t)\right)\right|}{1} \qquad \forall s(t)$$

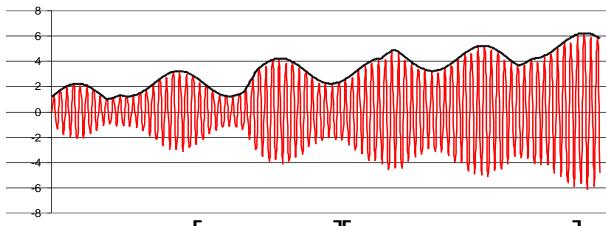
MODULATION & MULTIPLEXING

When done correctly, there are a lot of carrier periods compared to the changes in the signal s(t).

When modulated, the signal is contained in the way the amplitude changes, and we call this the "envelope" of the curve (symmetric up/down).

This form of amplitude modulation is called conventional AM, or just AM.





$$y(t) = [1 + k_a s(t)] [\sin(2\pi f_c t + \phi)]$$
 $k_a = 1, A = 1, \phi = 0$

Imagine that we have a signal s(t) which is a single sinusoid of frequency f_s . (a combination of sinusoids can be treated the same)

Assuming A=1, $k_a=1$, $\phi=0$, we get

$$y(t) = [1 + \sin(2\pi f_S t)] [\sin(2\pi f_C t)]$$

which when multiplied through becomes

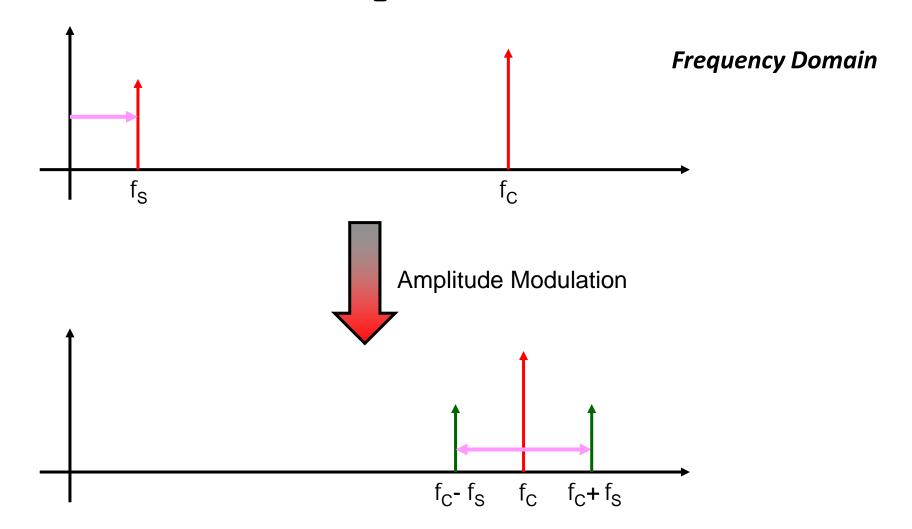
$$y(t) = \sin(2\pi f_C t) + \left[\sin(2\pi f_C t)\sin(2\pi f_S t)\right]$$

$$= \sin(2\pi f_C t) + \frac{1}{2}\left[\cos(2\pi (f_C - f_S)t) - \cos(2\pi (f_C + f_S)t)\right]$$

which has frequency components at:

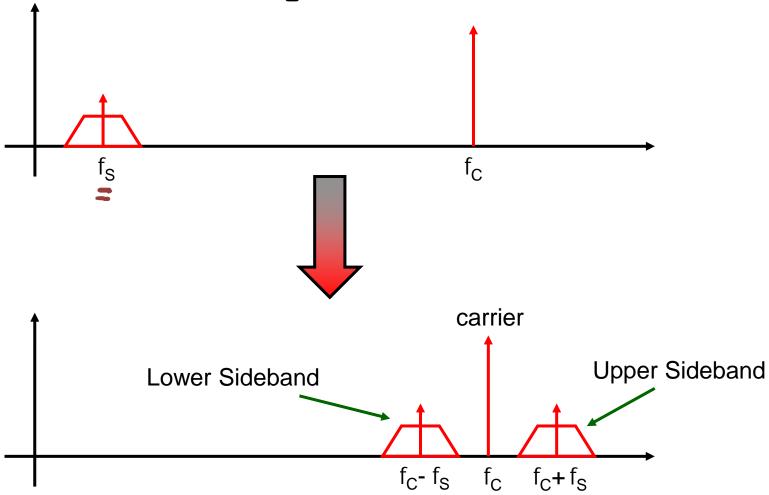
$$f_{C}-f_{S}$$
 $f_{C}+f_{S}$

AM - Bandwidth Requirement



To faithfully transmit a signal f_S , the transmitted signal requires a **bandwidth of 2** f_S .

AM - Bandwidth Requirement



Conventional AM is also known as Double-Sideband AM (DSB-AM).

Transmitted Power

The transmitted signal is (assuming normal assumptions)

$$y(t) = [1 + k_a \sin(2\pi f_S t)] [A \sin(2\pi f_C t)]$$

The transmitted signal can be broken into its frequency components, and the power calculated (you can try integrating if you wish ©)

$$y(t) = A\sin(2\pi f_C t) + \frac{1}{2}Ak_a \left[\cos(2\pi (f_C - f_S)t) - \cos(2\pi (f_C + f_S)t)\right]$$

$$P(Y) = \frac{1}{2}A^{2} + \frac{1}{2}(\frac{1}{2}Ak_{a})^{2}[1+1]$$

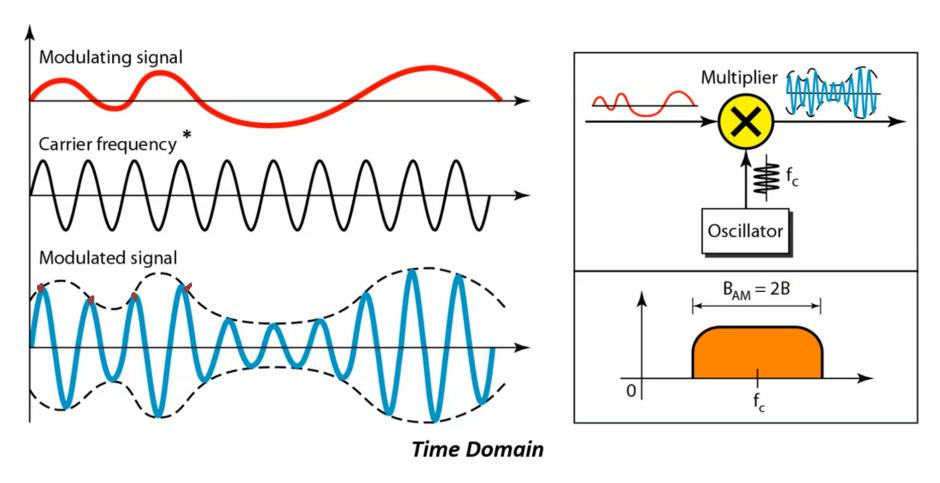
$$= \frac{1}{2}A^{2} + (\frac{1}{2}Ak_{a})^{2}$$

$$P(Y) = P(f_{C}) + P(\text{signal})$$

Note: If A=1, $k_a=1$, then only 1/3 of the power is used for the signal.

The transmitted signal contains the power of the signal AND the power of the carrier. But the carrier does not contain any information, so it is just a burden.

AM Process



^{*}Remember: the carrier wave is nothing more than a simple sinusoidal signal

- A carrier signal is modulated only in amplitude (signal strength).
- The modulating signal is the envelope of the carrier.
- The required bandwidth is 2B, where B is the bandwidth of the modulating signal.
- Since on both sides of the carrier freq. $f_{\rm c}$, the spectrum is identical, we can discard one half, thus requiring a smaller bandwidth for transmission.

Advantages of AM

Simplicity:

- Easy to modulate (control the amplifier at the output).
- Easy to demodulate in time domain using envelope detector.

Well understood:

- The oldest scheme and most intuitive.
- The waveform is also visually intuitive.

Well defined spectrum:

the bandwidth required is twice the maximum frequency component in the signal and is located above/below the carrier.

Disadvantages of AM

- Requires a comparatively large bandwidth (twice the bandwidth of the original signal).
- Power hungry: two sidebands and the carrier to transmit, the carrier power does not convey any information.
- Any noise or interference that adds to the amplitude of the signal will directly affect the information signal.

Conclusions

- Amplitude Modulation is a very useful communication scheme.
- More complex versions exist that provide superior performance but require more complicated receivers and are a little more difficult to understand.
- AM schemes are still in use today, and are unlikely to disappear.

FM & PM Modulation

FM & PM Modulation

<u>**Definition**</u>: Frequency Modulation is where the *frequency of the carrier* signal is modified according to our input data.

<u>**Definition**</u>: Phase Modulation is where the *phase component of the carrier* signal is modified according to our input data.

Phase modulation (PM) and frequency modulation (FM) are similar processes where we are actually modifying the angle of the sinusoid. Both PM and FM are forms of angle modulation.

23

Modulating the Signal

A sinusoidal signal can be modulated in different ways:

$$A\sin(2\pi f t + \phi)$$

I could change the frequency, increasing it and decreasing it so as to make it represent some data. This is called **FREQUENCY MODULATION**.



Changing the frequency of the sinewave as time passes.

FM Modulation

In frequency modulation we change the frequency. The modulating signal is a frequency term and we have a scaling factor k_f which we call the frequency sensitivity:

$$f(t) = f_C + k_f s(t)$$
signal

Constant phase-offsets carry no information, so when we're discussing FM signals, we generally ignore the offset component for purposes of clarity.

So an example of the instantaneous frequency under FM would be

$$f(t) = f_C + 500\cos(200\pi t)$$
signal

 $k_{\rm f}$ = frequency sensitivity

Mondulation Index, β

In general, the maximum frequency deviation from the centre frequency is given by $\Delta f_{\rm max} = k_f \left| s(t) \right|_{\rm MAX}$

And for a sinusoid of amplitude A, this is reduced to

$$\Delta f_{\text{max}} = k_f A$$

Similar to AM, we have a term that we call the modulation index (β), and we define it as the ratio of Δf to the modulating signal frequency (for a single sinusoid):

$$\beta_{\rm FM} = \frac{\Delta f_{\rm max}}{f_S} = \frac{k_f A}{f_S}$$

We have an input of signal f_S , how much variation do we add to the carrier f_C ? A value of 1 means that a 100 Hz signal will shift the carrier by 100 Hz.

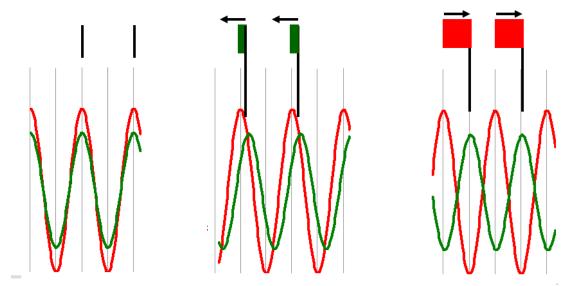
A typical value for β is in the order of 5.

Modulating the Signal

A sinusoidal signal can be modulated in different ways:

$$A\sin(2\pi f t + \phi)$$

I could change the phase of the carrier, increasing it and decreasing it so as to make it represent some data. This is harder to imagine, but you could compare it to the original signal and notice the difference in the position of the peaks. This is called **PHASE MODULATION**.

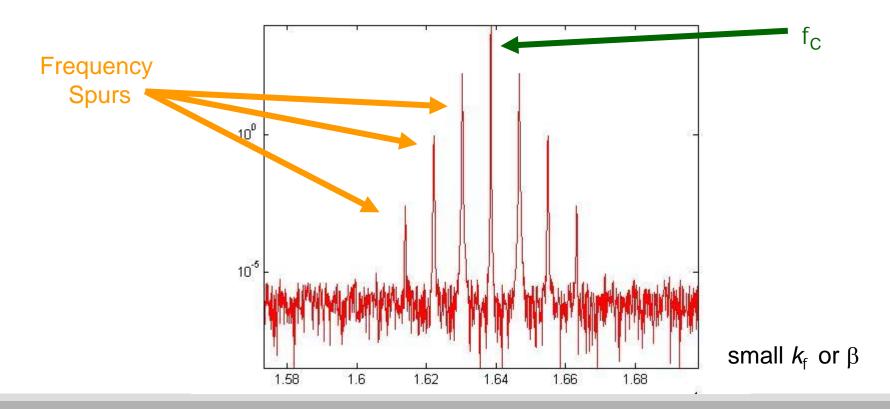


Changing the phase of the sinewave as time passes.

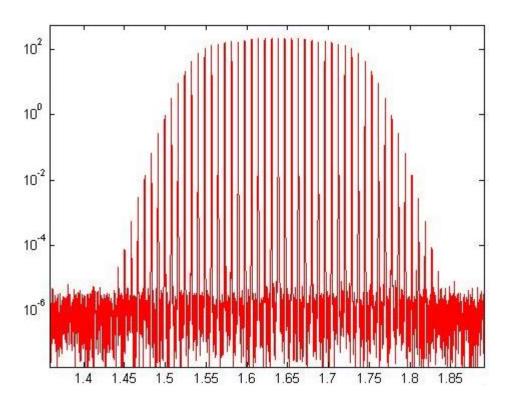
FM/PM Spectral Response

$$c_{\text{FM}}(t) = A\cos(2\pi f_C t + \beta\sin(2\pi f_S t))$$

An FM/PM modulated signal (wideband or narrowband) has a complex spectral response. There is no one single frequency component.



FM/PM Spectral Response

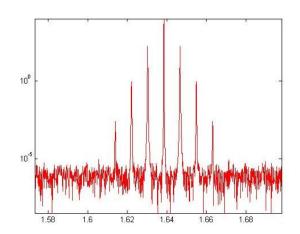


larger k_f or β

As the modulation index increases, we get a wider bandwidth and we get a lot more frequency spikes within that bandwidth.

The relationship between the number, position, and power of these individual frequency components is complex and non-linear. But there are some short-cuts to help us.

Power of an FM Signal



$$c_{\text{FM}}(t) = \underbrace{A\cos(2\pi f_C t + \beta \sin(2\pi f_S t))}$$

Since the value of the amplitude of the sine wave in an FM/PM modulated signal does not change, the transmitted power is a constant irrespective of the modulation index.

To calculate the power of an FM/PM transmitted signal, consider the case where there is no modulation and take that as your power:

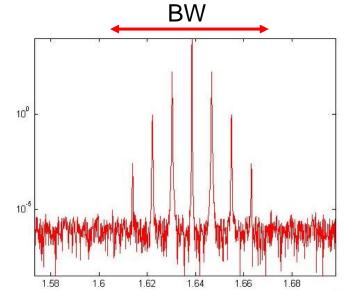
$$P_{\text{PM}} = P_{\text{FM}} = P_{\text{CARRIER}} = \frac{1}{2} A_{\text{C}}^2$$
 watts

Carson's Rule

bandwidth
$$\approx 2(\Delta f_{\text{max}} + f_S) \approx 2(\beta + 1)f_S$$

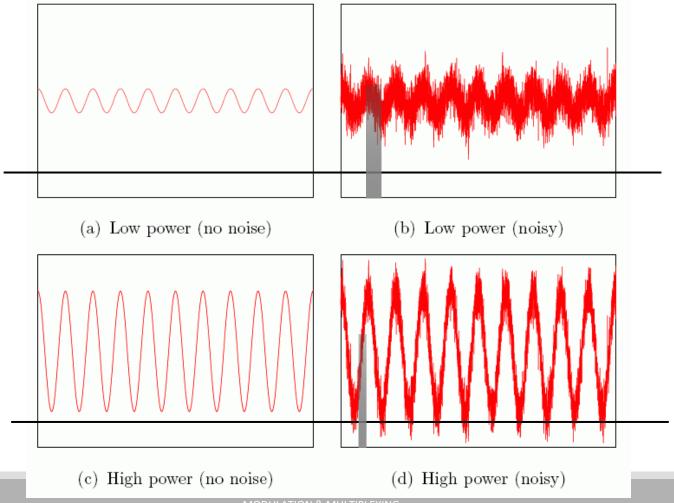
Carson's rule states that almost all the power (~98%) in an FM modulated signal lies within a bandwidth, given above, centered on the carrier frequency (where f_s is the maximum frequency in the modulating signal).

This is just a rule of thumb. It was invented around 1915-1920. It is an approximation that was very useful when you couldn't easily calculate Bessel functions, either because computers didn't exist, or you needed an approximate answer quickly.



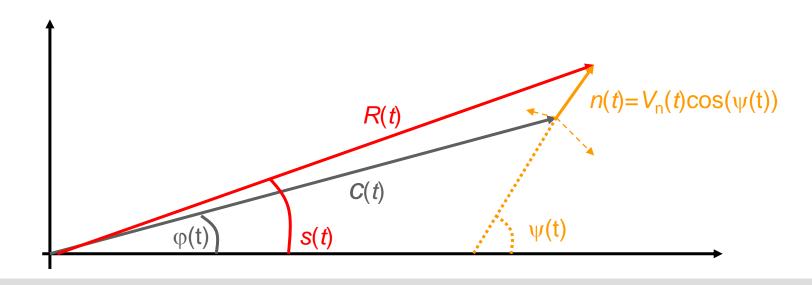
Noise in FM/PM

Noise superimposes itself on top of the waveform, changing its amplitude. Of course though, this also looks a little bit like it's changing the frequency, too. We call this jitter. But the more power we provide, the less effect it has.

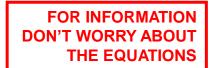


Geometrically, it's simple trigonometry.

If we say we have a carrier signal, and we superimpose a random signal on top of it, say angle ψ , magnitude V_n , then the resulting signal (R) has a slightly bigger vector length and a slightly different angle. However, the bigger the original carrier vector, the smaller the effect noise has.



FM: Signal to Noise Ratio



The SNR can be calculated as

$$SNR = 10 \log_{10} \left[\frac{2(A_C k_f)^2 P_S}{P_N} \right] dB$$

Increasing the modulation sensitivity k_f (or β) increases the SNR by the square of the improvement (so a very powerful reason to use a big modulation index).

A big modulation index means that we have a very wide bandwidth.

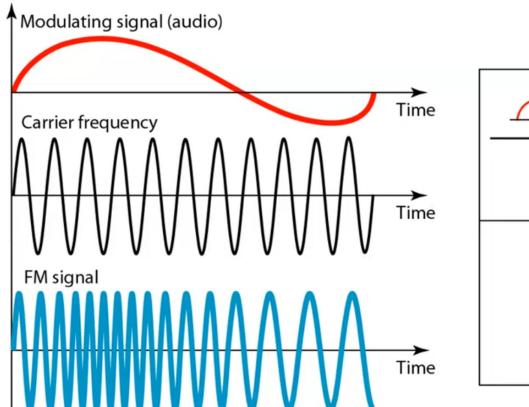
bandwidth
$$\approx 2(\Delta f_{\text{max}} + f_S) \approx 2(\beta + 1)f_S$$

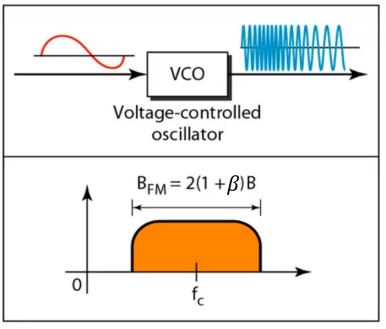
So doubling k_f (or β), improves SNR by a factor of 4, and increases the bandwidth required by a factor of 2.

Using more bandwidth improves quality, even if transmit power stays the same.

FM Process







- The modulating signal changes the frequency (f_c) of the carrier signal
- The total bandwidth **required** for FM can be determined from the bandwidth (B) of the audio signal: $B_{FM} = 2(1 + \beta)B$, where β is usually ~5.

FM/PM Review

Advantages

- Robust FM/PM displays better resistance to noise and interference than AM.
- Flexible Allows for trade-off of performance against bandwidth.
- Easy power control.

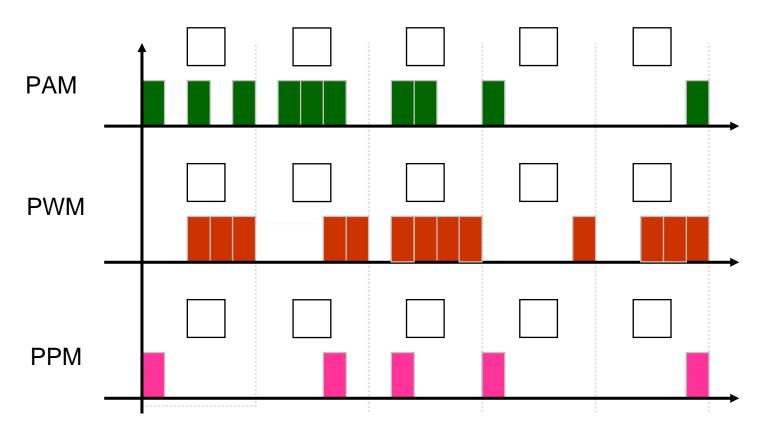
Disadvantages

- Bandwidth FM/PM requires much more bandwidth than AM.
- Complexity FM/PM has a higher complexity of implementation than AM.

Digital Communication Schemes

Encoding Digital Signal

Let's imagine we have a binary system, for purposes of display. There are many ways to interpret a series of high voltages and low voltages:



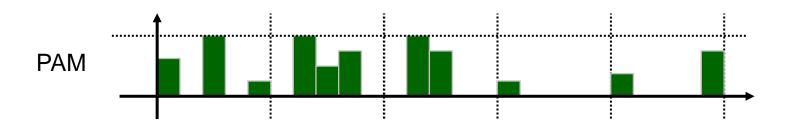
Encoding Digital Signal: PAM

PAM: Pulse Amplitude Modulation

In a normally encoded system, as we normally understand it, the data consists of a sequence of pulses of pre-agreed height and of pre-agreed location (timing) and width.

So each pulse is of fixed width and comes at an expected, agreed time but the amplitude varies. A special case is binary where the signal is either a fixed height or zero.

The amplitude of the pulse corresponds to a valid level in our agreed m-nary system. In a binary system there are only 2 levels. In a multi-level system, they can be at other discrete amplitudes.



Encoding Digital Signal: PPM

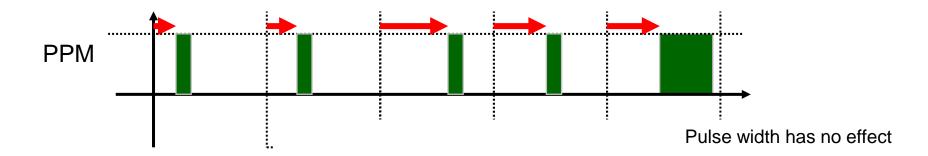
PPM: Pulse Position Modulation

The value of the signal is determined by the position of a pulse within a specific time-frame.

The pulse can be of fixed width but it is the timing offset from the start of the timing period that corresponds to the value encoded. The timing could be determined by an analog value, but it is normally controlled by a digital signal.

This system works well when you can only turn something on/off, for example a fibre-optical cable.

There is a problem in ensuring the timing of the receiver and the transmitter are the same as if we're a little skewed then we add an offset to the data.



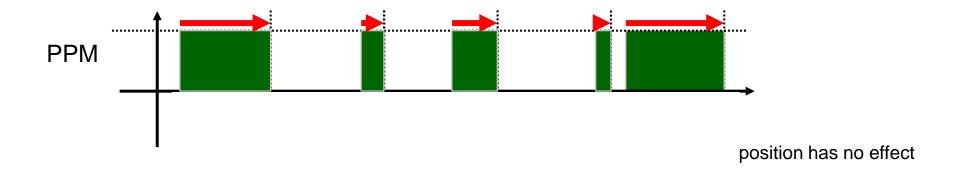
Encoding Digital Signal: PWM

PWM: Pulse Width Modulation

The value of the signal is determined by the width of the pulse within a given time frame. There will only be one pulse during that time-frame and it will either start or finish with that pulse.

The pulse is of fixed amplitude and so again suits systems with poor modulation control (on/off). The width of the pulse could be varied as an analog signal but the width variations are normally discrete values.

Very simple scheme to implement and detect. To get the value, just calculate the time up and divide by the total time. The agreement to start/finish with a pulse makes it very easy to detect the pulse.



Comparison

	PAM	PPM	PWM
Information Density	highest	lowest	lowest
Complexity of Implementation	low-med	lowest	lowest
Key Value Characteristics	amplitude	position	width
Vulnerabilities in analog channel	noise	multipath	almost immune
Importance	very	rare-ish	common

Digital Modulation: ASK

In AM modulation we have a carrier signal with a frequency f_c

$$c(t) = A\sin(2\pi f_c t + \phi)$$

If we have an information signal, we can implement AM modulation by adjusting the amplitude according to some scaled version of the information signal.

$$y_{AM}(t) = A[1 + k_a s(t)] \sin(2\pi f_c t + \phi)$$

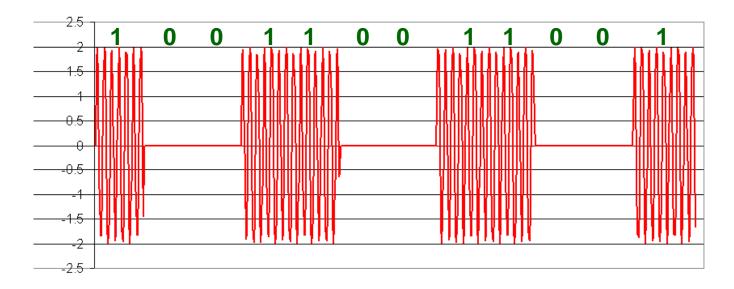
For digital modulation, the information signal is, however, a series of discrete steps, so the amplitude will change to fixed values.

There is one important factor. In modern communication schemes we normally use **DSB-AM-SC**, and sometimes **SSB**. This is for power and bandwidth efficiency. With modern technology the more complicated receivers are not a concern:

$$y_{\text{ASK}}(t) = A[s(t)] [\sin(2\pi f_c t + \phi)]$$

Digital Modulation: ASK

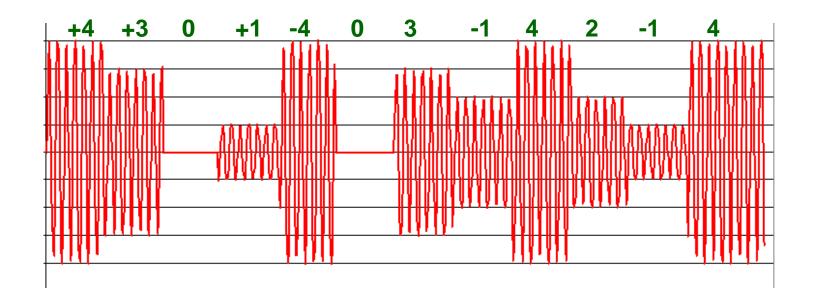
But the digital data are of a fixed number of levels, so the amplitude will vary accordingly. Imagine we have a binary data signal (PAM/PWM/PPM will behave the same here). The output would look like:



When we have only two levels (on/off) we call this technique **ON-OFF KEYING** (**OOK**). Imagine it like Morse code – the carrier is present, or it is not.

Digital Modulation: ASK

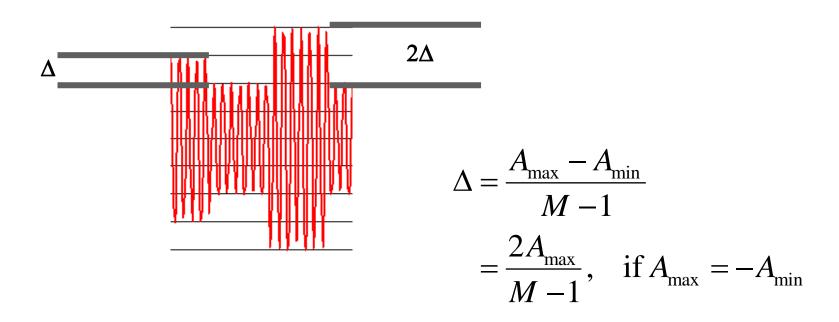
It is possible to have multiple levels. Negative value levels have the same amplitude but a 180° phase shift. With a suitable receiver, this phase shift can be detected and we can extract negative values as well as positive ones.



When we have multiple levels, we call this technique **AMPLITUDE SHIFT KEYING (ASK).** It's like ordinary AM but the information signal only has discrete levels.

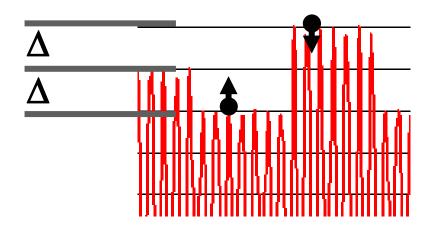
ASK: discrete levels

In our digital communication scheme we have M levels or symbols. This means that our carrier **amplitude** is split into M different possible levels.



As we are using suppressed-carrier AM, we can have negative input values, thus the **acceptable range of inputs is from -1 to +1**, scaled by the amplitude of our signal. So if there are five levels (that means 4 separating spaces), that means that the amplitude difference is the peak-peak variation divided by (5-1).

ASK: Noise



$$\Delta = \frac{2A_{\text{max}}}{M-1}$$

If a demodulated signal is distorted by noise such that it gets rounded to a different level, then we have an error.

So if we get an amplitude distortion by more than $\pm \frac{1}{2}\Delta$ then we will get an error.

So this defines our maximum allowable noise voltage, or even error voltage:

$$v_{\text{error}} = v_{\text{received}} - v_{\text{expected}} < \frac{\Delta}{2} < \frac{A_{\text{max}}}{M-1}$$

Note: in practice an error needs to last some time to cause a bit to be misread.

Why use Digital Comms Schemes?

When done correctly, digital communication schemes have the following advantages:

- Most of our data is digital, and digital communication schemes are naturally compatible with digital data.
- Digital modulation schemes are very resilient to noise and can guarantee a **noise-free communication** scheme (CD vs. tape).
- Digital communications schemes allow us to use signal processing to enhance performance, encryption to increase throughput, and error-correction to detect and fix any errors that do occur.
- While digital systems require a CPU and software algorithms, if you need high data rates, the equivalent analog system is very difficult to design. The discrete levels of the digital modulation greatly simplifies the hardware implementation, thus **making high data-rate designs easier to build**.

Multiplexing

Multiple Users

In many cases we will have multiple users who want to access the same channel, for example in a given region we may have users who want to use:

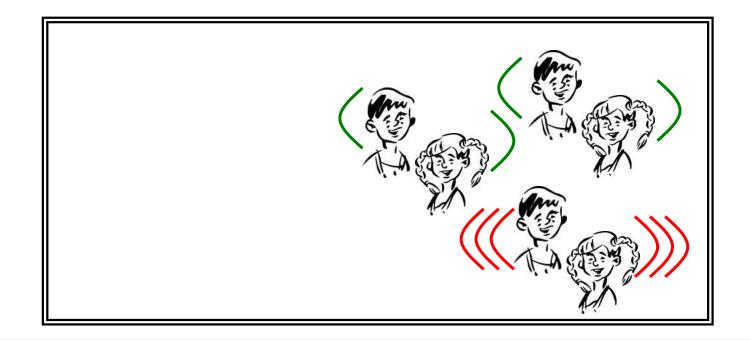
- Electromagnetic spectrum (wireless);
- Over-the-air audio;
- Light or photonic communication.

If un-coordinated, the different users will interfere with each other, making everyone's communication difficult, if not impossible.

Co-Existence: Example 1

Consider you are in a large room, 3 pairs of people talking just to each other and all grouped close together. One couple are literally shouting, making it hard to hear (that's **interference**).

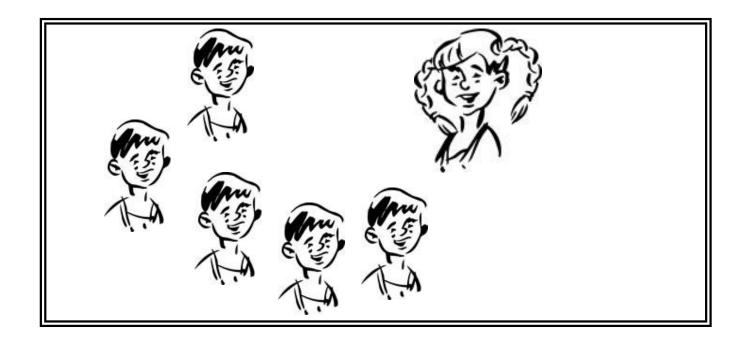
What should we do to enable reliable communication for all people???



Co-Existence: Example 2

Consider there is a class of 5 students each wanting to talk to the lecturer. If they all talk at once, there'll be confusion and **interference**.

What should we do to enable reliable communication for all people???



Possible Strategies

There are different strategies for communicating across a common channel while minimising **interference**:

- Spatial Separation: keep users apart so they don't interfere.
- **Time Separation**: agreed to speak one-at-a-time.
- Frequency Separation: transmit at different frequencies (radio, TV).

These 3 techniques define much of our communication architecture, and we'll go through a few examples now.

Spatial Co-ordination: Radio



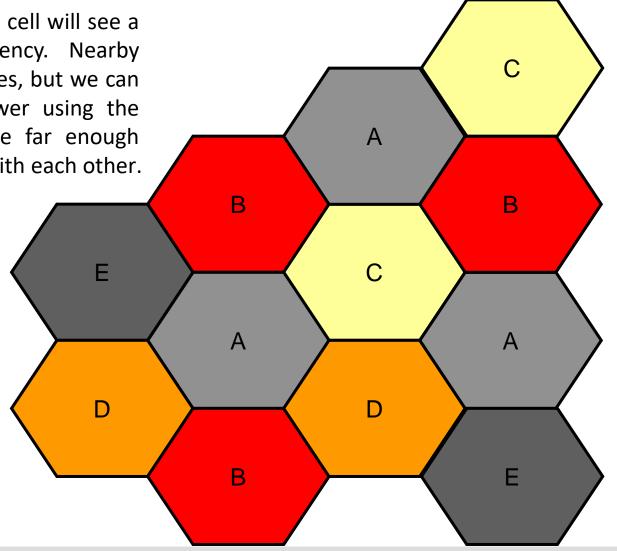
Many stations share similar frequencies, but to avoid interference, their power levels are restricted so that their range is small and they don't interfere: the more power they use, the bigger the circles.

Mobile Phone Networks & Spatial Multiplexing

By using cells, a number in each cell will see a basestation with a single frequency. Nearby cells will use different frequencies, but we can arrange cells so that each tower using the same frequency (e.g., A-A-A) are far enough apart that they don't interfere with each other.

In this way we get efficient reuse of frequencies, while minimising our interference issues. This means we can get 1000's of time more data nationwide for each piece of spectrum, than normal.

Cell sizes are not all the same, and it depends on how many users you need to support in a cell.



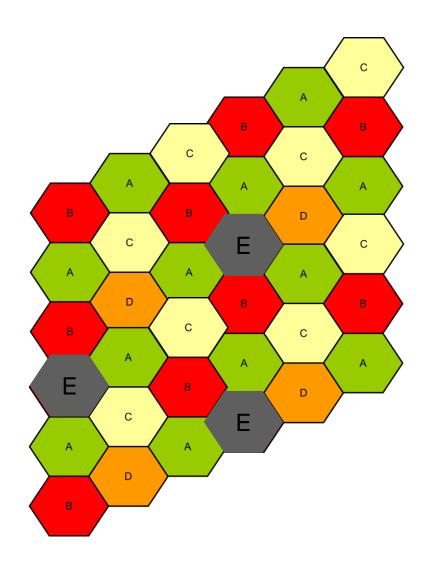
Mobile Phone Networks & Spatial Multiplexing

As we move to higher frequencies, the propagation losses increase, that means signals don't travel as far before they become too small.

This means that cells get smaller, even down to the size of a house or a street.

Small cells still have the same data capacity, so this means small dense cells allow the network to offer more data / sq.km, than before.

Unfortunately, it will cost them a lot more to build the network.



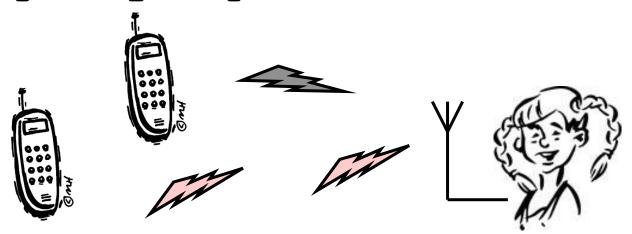
Frequency Separation

The electromagnetic spectrum exists. There is only one of it in a specific area. If more than one user transmits a wireless signal, there is the potential for both wireless signals to be received and interfere with each other.

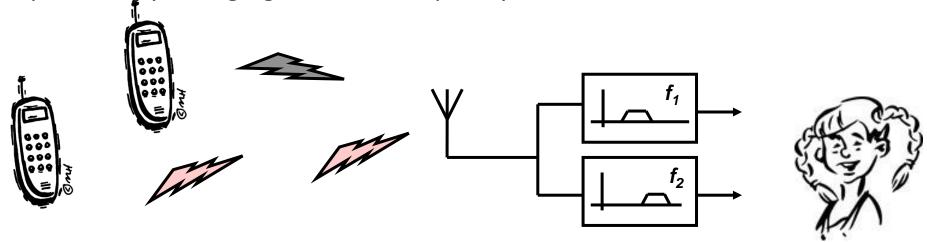
What we can do is use different frequencies and then at the receiver use a band-pass filter. The filter will only allow a certain set of frequencies to pass. The other signals are blocked.

So that means I can tune into one radio station and then another by just changing the filter frequency. All the stations are being picked up by the antenna, but if we do the filtering right, they don't interfere with each other.

Frequency Separation



So the antenna picks up a range of electromagnetic signals. We can select the frequency we want by using a **band-pass filter**. We can pick out different frequencies by changing the filter frequency.

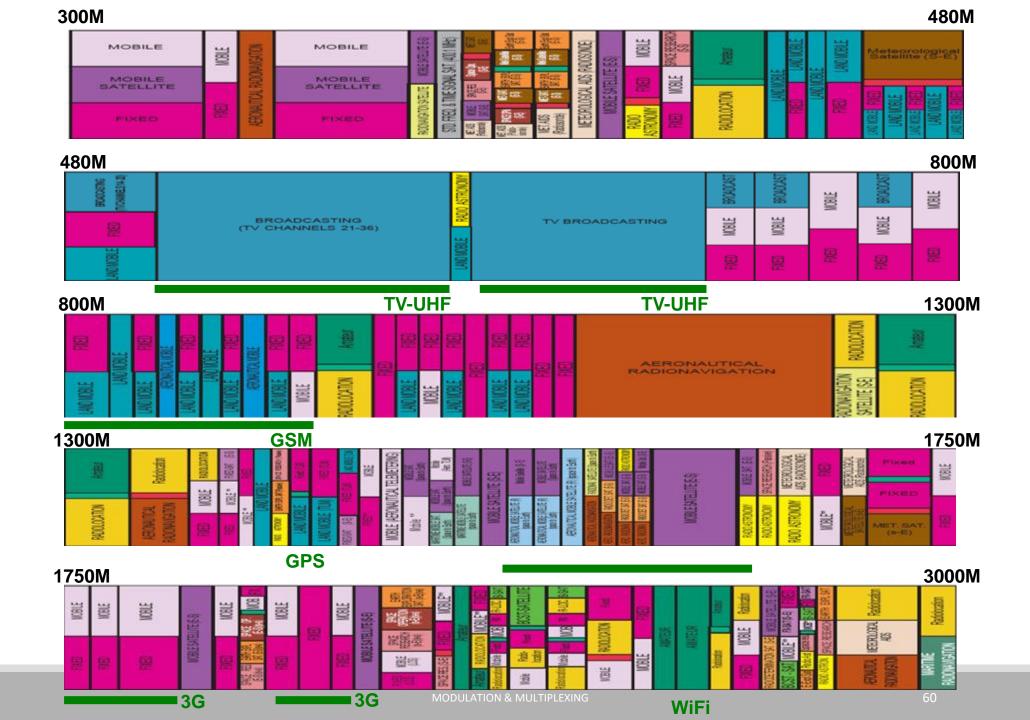


Frequency Spectrum

The most useful sections of electromagnetic (EM) spectrum is at lower frequencies (0.1 to 10 GHz) as these travel well in the atmosphere.

However, there are a lot of competing uses for the spectrum. This means that the spectrum is scattered among thousands of uses and that it is extremely difficult to get access to spectrum.

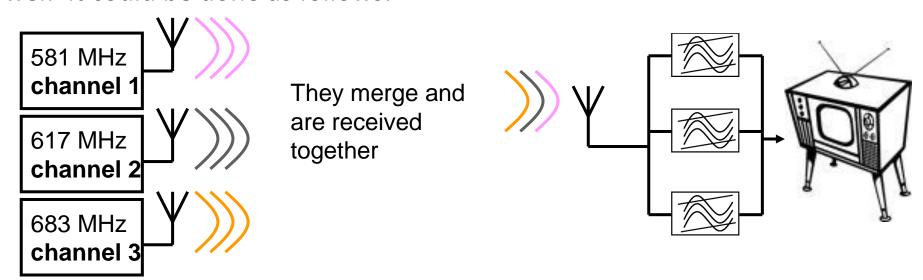
The national regulators control the spectrum. It is illegal to transmit unless you have a license or you are in an un-licensed band.



FDM Diagram

If we have multiple users or uses for our bandwidth, the process of splitting up the bandwidth to the users is called **FREQUENCY DIVISION MULTIPLEXING (FDM)**.

Imagine that we transmit multiple TV stations from the same transmitter tower. It could be done as follows:



FDM Advantages/Disadvantages

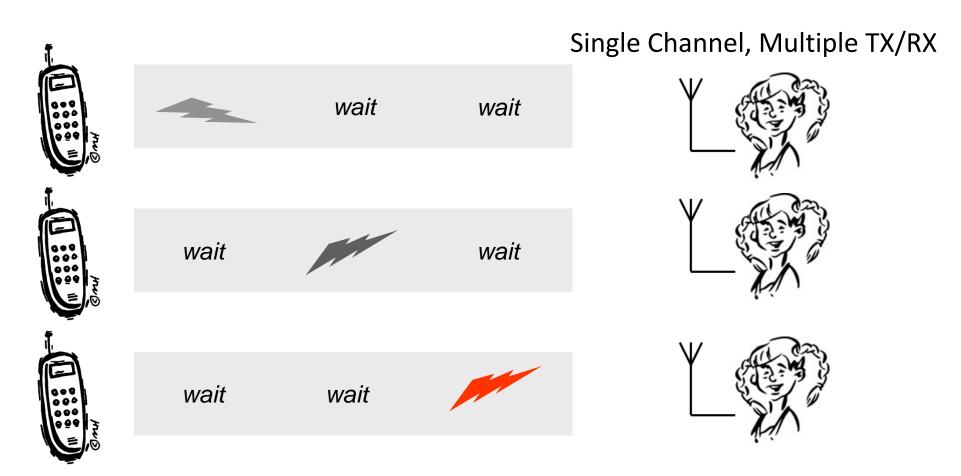
Advantages

- •The communication link is continuous. A receiver can theoretically listen/receive all transmissions at the same time. The filter just selects the transmission to interpret. Multiple filters would allow for multiple signals to be received at once.
- •You don't need to **know anything about the transmitter** implementation in order to demultiplex.
- •Implementation can be simple.

Disadvantages

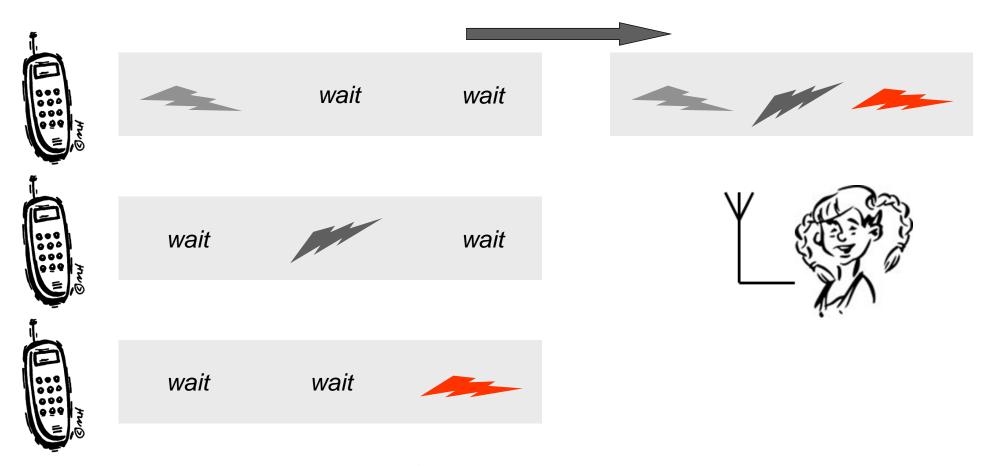
- •If the guard bands are small then the **filters become expensive** and hard-to-build.
- The guard bands are wasted spectrum that are not being used.

TIME DIVISION MULTIPLEXING (TDM)



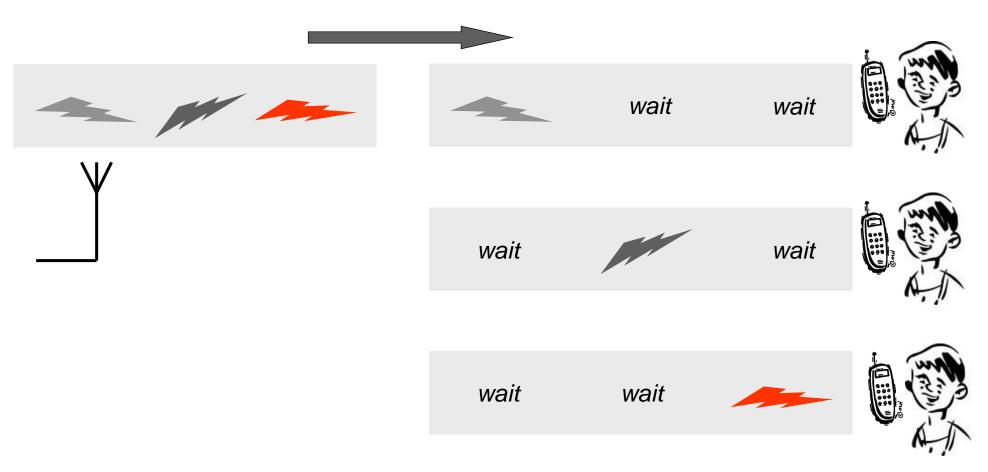
Each radio has a time for communicating. The time slot can vary from user to user. Only within that time slot is communication allowed. If there's no data to send, there is no transmission.

TDM: Multiple TX, single RX



Each transmitter has a time slot for communicating. The time slot can vary from user to user. Only within that time-slot is communication allowed. If there's no data to send, there is no transmission.

TDM: Singe TX, Multiple RX



The transmitter sends out each users signal in a sequence. The users know their position in the sequence, and just listen to the data that is meant for them.

Duplexing

Simplex communication:

- Signals are transmitted in one direction only;
- Only requires a single communication channel;
- For example: television or radio broadcasts.

Half duplex communication:

- Signals are transmitted in just one direction at a time;
- Only requires a single communication channel;
- For example: Walkie talkie.

Full duplex communication:

- Signals can be transmitted in both directions simultaneously;
- Generally requires two channels.

Simulated full duplex:

- Frequency division duplexing (FDD), similar to FDM, but give each user separate frequency subchannels for transmitting and receiving;
- Time division duplexing (TDD), similar to TDM, but give each user separate timeslot subchannels for transmit and receive.