

ITS 323 – SIGNALS, BANDWIDTH AND DATA RATES

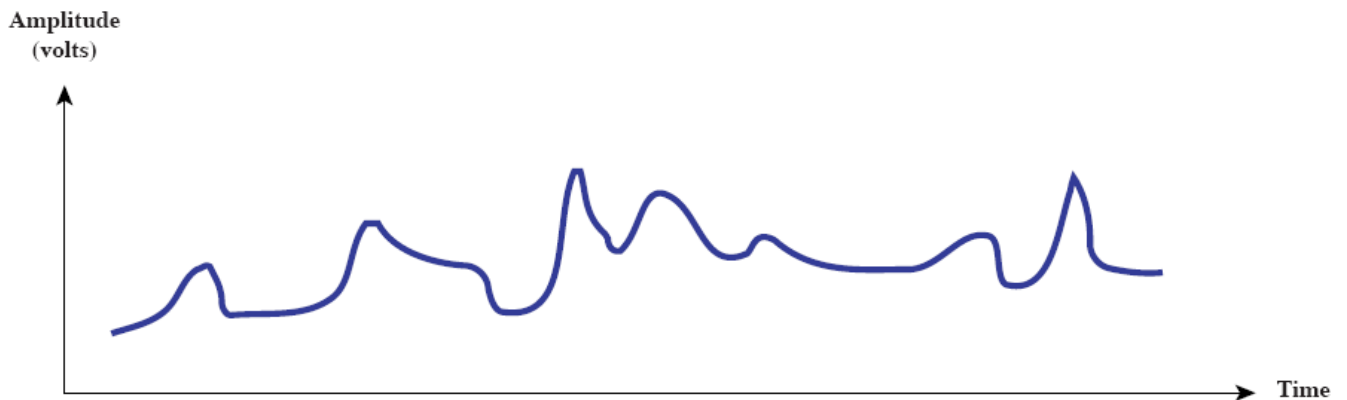
This is a summary of the main points from the Data Transmission lecture (in particular, slides 6 to 16). It does not cover anything new from the lectures, but instead tries to provide a written description for the students that did not understand the lecture.

1 Signals

Data is communicated between devices via an electromagnetic signal.

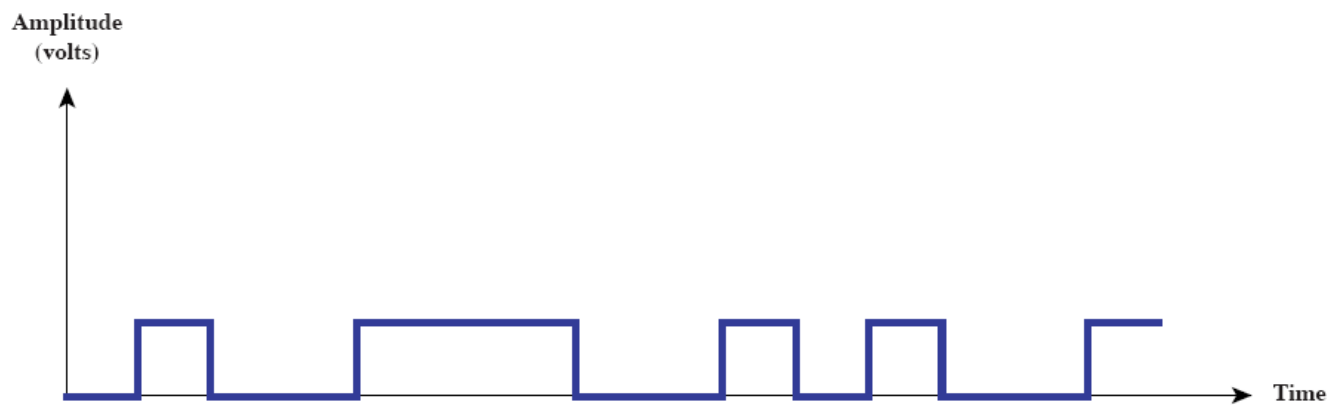
In the time domain, we can view the signal as a measure of signal energy (e.g. voltage) versus time. The relationship between energy and time is a function $s(t)$.

The signal can be considered *analog* or *digital*. The figures below illustrate examples of each.



(a) Analog

Figure 1



(b) Digital

Figure 2

2 Viewing Signals as Sine Waves

Any communication signal (analog or digital) can be represented as a set of analog sine waves. Viewing signals as sine waves makes it easier for the design, analysis and implementation of systems that need to transmit and receive signals.

Below are two sine waves: the first is $s_1(t) = \sin(2\pi ft)$ and the second is $s_2(t) = (1/3)\sin(2\pi(3f)t)$.

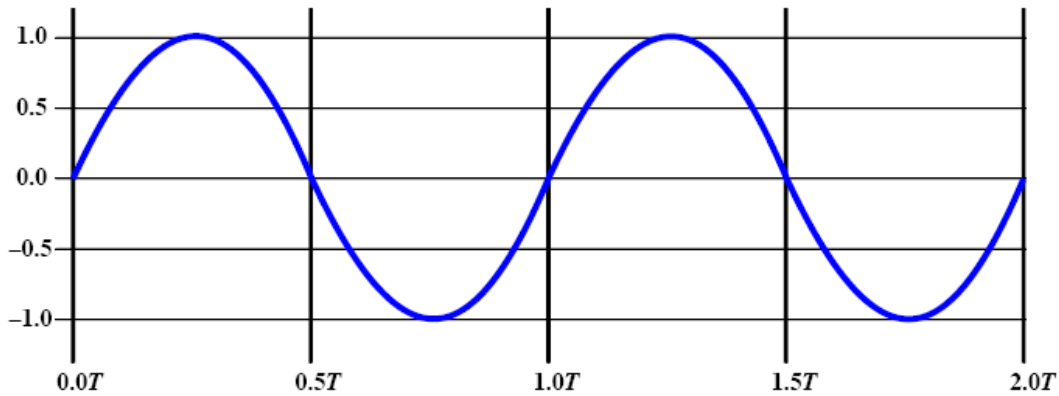


Figure 3

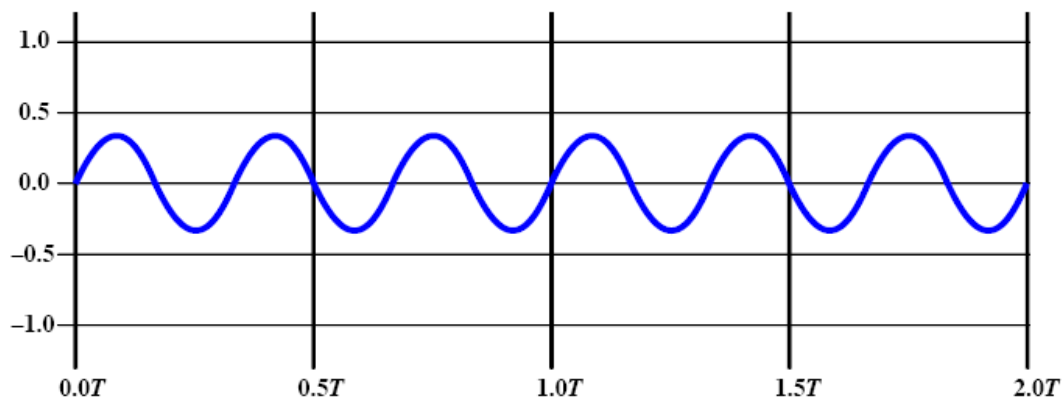


Figure 4

If we add the two sine waves together (and multiply by a factor of $4/\pi$ – this is just to bring the amplitude close to 1.0) we get the figure below:

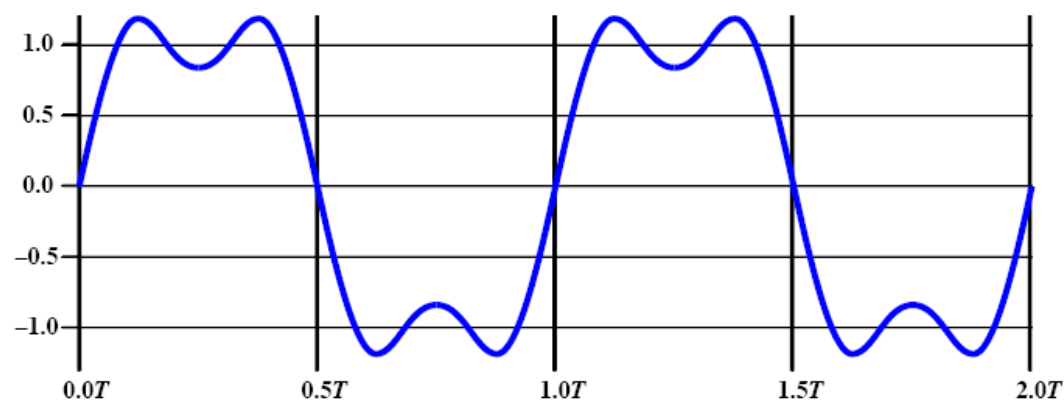


Figure 5

The resulting signal is:

$$s(t) = \left(\frac{4}{\pi}\right)(s_1(t) + s_2(t))$$

$$= \left(\frac{4}{\pi}\right)\left(\sin(2\pi ft) + \left(\frac{1}{3}\right)\sin(2\pi 3ft)\right)$$

One point to consider is that if the frequencies of the individual sine waves are integer multiples of each other, then the frequency of the resultant signal $s(t)$ is the lowest frequency (called the *fundamental frequency*). In our example, the frequencies of the sine waves are f and $3f$. Therefore the frequency of the resultant signal, that is the fundamental frequency of $s(t)$ is f . (Remember, f is just a variable. If $f = 2$ Hz, then the frequency of $s(t)$ is 2Hz, the frequency of $s_1(t)$ is 2Hz, and the frequency of $s_2(t)$ is 6Hz.)

3 But what is a sine wave?

Since we know that any communication signal (even a digital signal) can be viewed as a set of sine waves, it is important to remember the general structure of a sine wave. It can be written as:

$$s(t) = A \sin(2\pi ft + \phi)$$

where:

- A , amplitude: the maximum strength of the signal [volts]
- f , frequency: rate at which the signal repeats [Hertz]
 - T , period: duration of a single cycle ($1/f$) [seconds]
 - λ , wavelength: v/f , where v is normally the speed of light.
- t , time: the current point in time [seconds] (remember $s(t)$ is a function of the signal strength versus time)
- ϕ , phase: a measure of the shift of the signal from some origin [radians]
 - [You do not need to know this, but someone may be interested: Why is phase measured as an angle? Remember from basic physics that we can relate frequency (in Hertz) to an angular frequency known as ω (measured in radians): $\omega = 2\pi f$. So in fact the basic form of a sine wave is represented in radians, e.g. $\sin(\omega t + \phi)$. It is just easier for us to talk about frequency in Hertz and phase in radians]

The figures below illustrate the impact of varying some of these parameters.

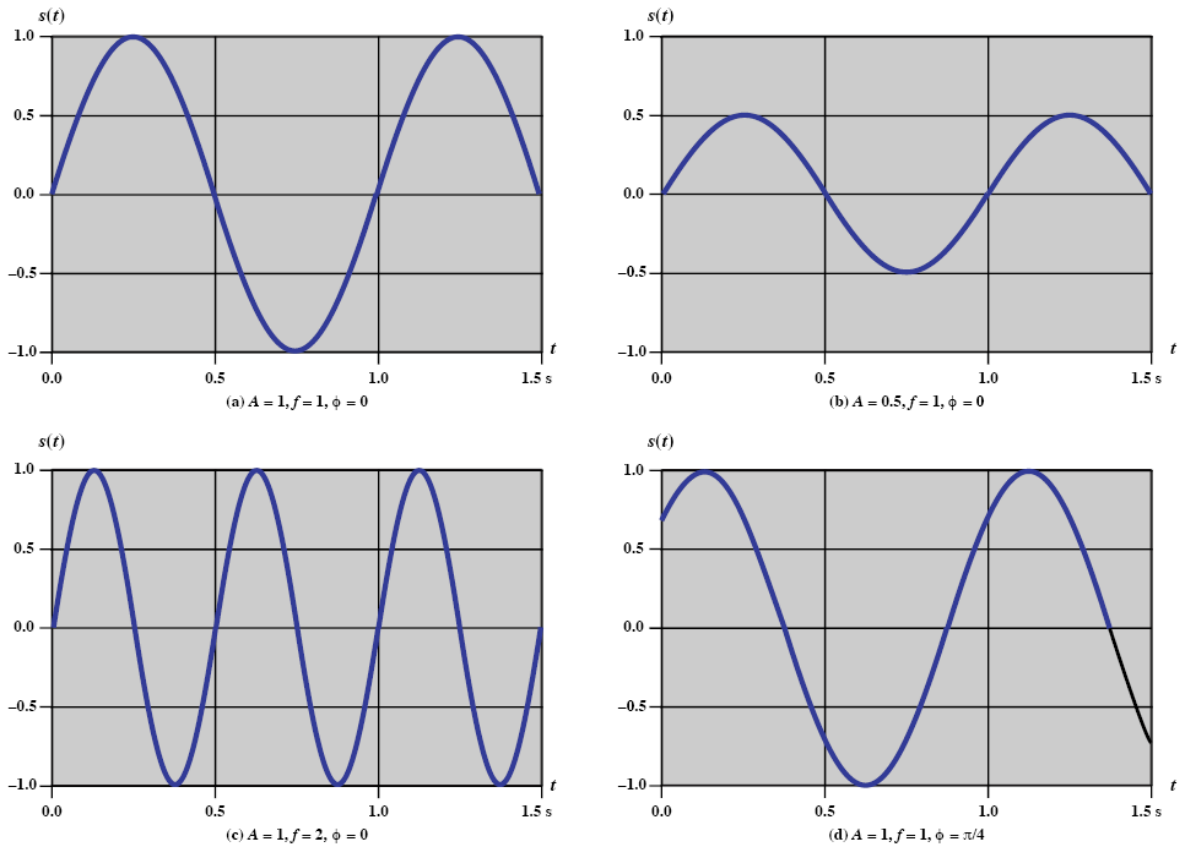


Figure 6

4 The Frequency Domain

As a way to look at signals in a more convenient form (as well as to allow mathematical analysis of signals easier, and hence make the design of systems easier), we can represent a signal in the time domain (e.g. the figures shown previously) or the frequency domain.

The frequency domain function of a signal ($S(f)$) shows the peak amplitude (A) of the sine waves the make up the signal at each frequency.

For example in Figure 5 the signal is made up of two sine waves. The first has a frequency component of $1f$. It has an amplitude of $4/\pi$. The second sine wave has a frequency of $3f$. It has an amplitude of $4/\pi \times 1/3$. Therefore, the frequency domain signal can be plotted as below.

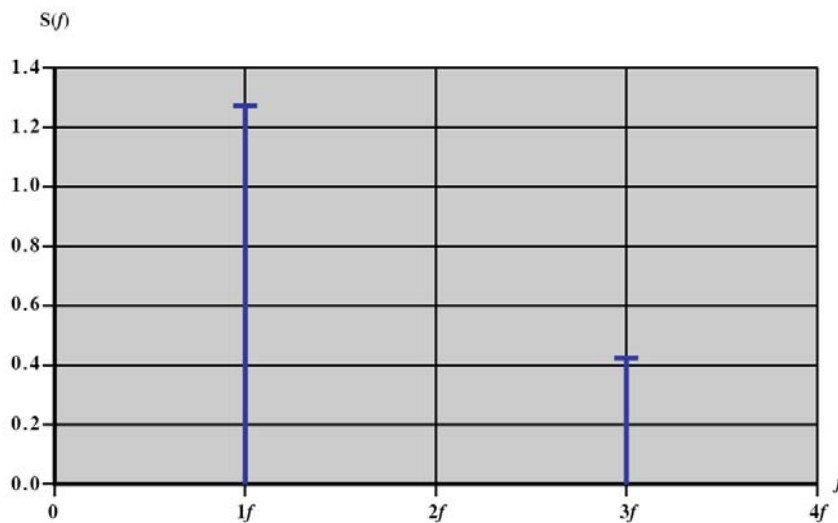


Figure 7

In other words, Figure 5 and Figure 7 convey the same information: they both indicate the amplitude of the sine waves and the frequency of the sine waves. If we combined Figure 3 and Figure 4 into one figure, then that would also show this same information.

5 Spectrum, Bandwidth and Data Rates

Spectrum is the range of frequencies of a signal. In our example above, the range of frequencies from $1f$ to $3f$. (Remember, f is just a variable. If $f = 2$ Hz, then the spectrum of the signal is from 2Hz to 6Hz.)

Absolute bandwidth is the width of the spectrum (e.g. the largest frequency component minus the smallest frequency component). In our example the absolute bandwidth is $2f$ (or if $f = 2$ Hz, then the absolute bandwidth is 4Hz).

The problem with absolute bandwidth is that for some signals (especially digital signals), in theory there are an infinite number of frequency components, that is, the absolute bandwidth is infinite.

So in fact we often deal with a 'real' bandwidth which is the set of frequencies transmitted within some fixed range. In practice, the medium (e.g. cable) and transmission device has limitations as to the frequency components it can carry. The bandwidth of the medium therefore limits the 'real' bandwidth of our signal.

We will see some examples shortly...

6 Returning to Component Sine Waves

Lets look in more detail at our example signal made up of sine waves. Below are four signals, in each one we have more component sine waves.

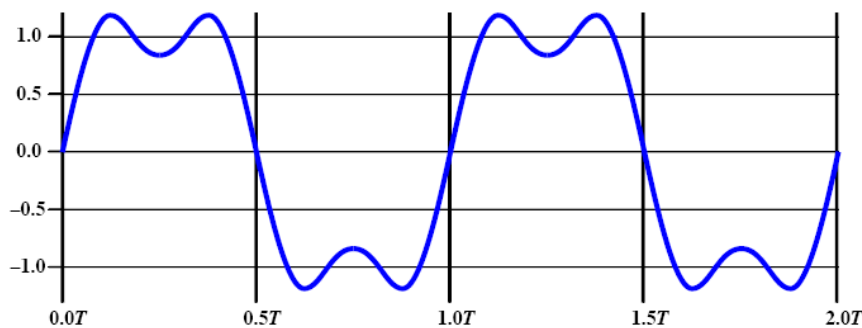


Figure 8

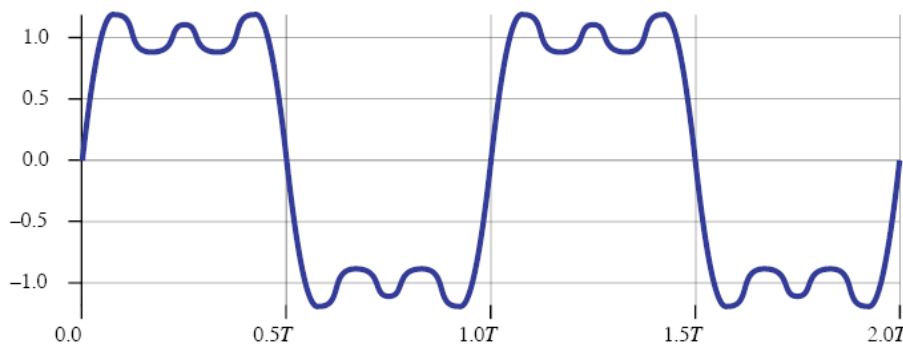


Figure 9

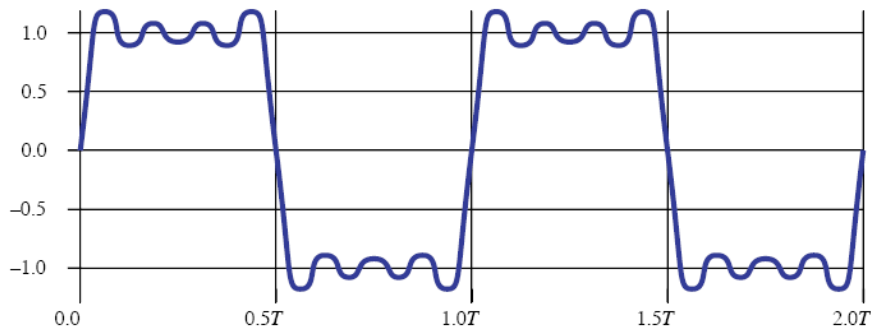


Figure 10

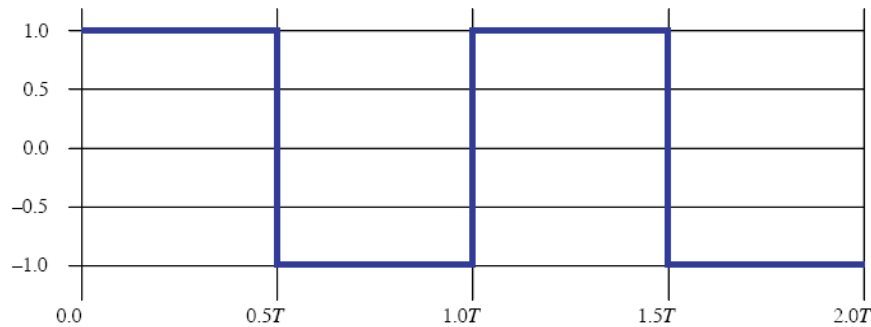


Figure 11

The above four figures represent the following four signals:

$$s(t) = \frac{4}{\pi} [\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t)]$$

$$s(t) = \frac{4}{\pi} [\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t) + \frac{1}{5} \sin(2\pi(5f)t)]$$

$$s(t) = \frac{4}{\pi} [\sin(2\pi f t) + \frac{1}{3} \sin(2\pi(3f)t) + \frac{1}{5} \sin(2\pi(5f)t) + \frac{1}{7} \sin(2\pi(7f)t)]$$

$$s(t) = \frac{4}{\pi} \sum_{\text{odd } k} (\frac{1}{k}) \sin(2\pi(kf)t)$$

One thing we notice: the more sine wave components we add, the closer this signal looks to a square wave. In fact, if we add an infinite number of sine waves, we will get a square wave (that is, a digital signal). Of course if there are an infinite number of frequency components, then the absolute bandwidth will be infinite.

For the first signal, the absolute bandwidth is $3f - 1f = 2f$

For the second signal, the absolute bandwidth is $5f - 1f = 4f$

For the third signal, the absolute bandwidth is $7f - 1f = 6f$

For the fourth signal (square wave), the absolute bandwidth is $\infty f - 1f = \infty$ (infinity)

Of course you can do the calculations of bandwidth if we set the variable $f = 2\text{Hz}$!

7 Data Rates and Bandwidth

Now let's consider sending data (bits, 0 and 1) with signals. Looking at Figure 8, we see that there is a "high" signal (when the signal energy, or voltage, is close to +1.0), and there is a "low" signal (when close to -1.0). We want to send 0's and 1's, so let's say that a high signal corresponds to a 1 and a low signal corresponds to a 0. When a receiver receives the signal it will measure the received voltage: if it is close to +1.0, then it will interpret it as a bit 1; if it is close to -1.0, then it will interpret it as a bit 0. This is a very simple way to send bits.

So now lets consider the data rate: how many bits can we send per second?

We have said that the real bandwidth of a signal is usually limited by the medium. Lets assume our transmission medium allows us to send up to 4MHz. That is, our bandwidth is 4MHz. Remember, the bandwidth of Figure 8 is $3f - 1f = 2f = 4\text{MHz}$. Therefore, $f=2\text{MHz}$. The frequency of our signal is 2MHz.

If the frequency f is 2MHz, then the period T is $0.5\mu\text{sec}$ (since $T = 1/f$). Looking at Figure 8 again, we see that 4 bits can be transmitted in the entire picture: a 1 (high), a 0 (low), 1, 0. And the total time it takes to transmit these four bits (1010) is $2.0T$, or $1\mu\text{sec}$. So if we can send 4 bits in $1\mu\text{sec}$, then it means we can send 4,000,000 bits in 1 second.

Our data rate that can be achieved with this signal is 4Mb/s (when the bandwidth is 4MHz with Figure 8).

Now lets consider another case: the bandwidth is 4MHz but instead we use Figure 9. With this signal the absolute bandwidth is $5f - 1f = 4f = 4\text{MHz}$, so $f = 1\text{MHz}$. So we have to transmit this signal at $f = 1\text{MHz}$ to use a bandwidth of 4MHz. So lets do our data rate calculation again:

$$T = 1/f = 1\mu\text{sec}$$

$$4 \text{ bits sent within } 2.0T = 2\mu\text{sec}$$

Data rate is 2Mb/s

So lets compare: using the same bandwidth (4MHz), with the signal of Figure 8 we can send at 4Mb/s, but with the signal of Figure 9 we can only send at 2Mb/s. So the first signal gives a faster data rate and so is better (or is it? We will describe some problems with it shortly).

Now another example: allow the bandwidth to be 8MHz, and use the signal of Figure 9. $f = 2\text{MHz}$ and you should find the data rate is now 4Mb/s.

So lets summarise so far:

- Increasing the bandwidth, increases the data rate
 - With Figure 9, going from 4MHz to 8MHz, increases data rate from 2Mb/s to 4Mb/s
- Increasing the frequency, increases the data rate
 - With BW = 4MHz, going from 1MHz (Figure 9) to 2MHz (Figure 8), increases the data rate from 2Mb/s to 4Mb/s

So to achieve a high data rate, we just transmit with high bandwidth and at a high frequency ...

But there are limitations. In practice, the greater the bandwidth/frequency, the greater the cost. Why? Because equipment is more complex at higher frequencies, and also bandwidth available to systems is limited by both the capabilities of transmitters/receivers and by government regulations (wireless LAN transmission is only allowed at certain frequencies within a small bandwidth – this is so the transmission does not interfere with other uses such as cellular phones, amateur radio, TV, and so on).

And finally, using a smaller bandwidth increases the chances of errors at the receiver. Why? Compare the Figures 8 to 11 again. Try to work out why a receiver may experience more errors with the lower bandwidth signals, than with the higher bandwidth signals...