Chapter 5 Data communication systems

5.1 Systems and links

Learning objectives:

→ Describe the main features of a variety of communication links.

→ Compare different communication links in terms of frequency, bandwidth, and other properties.

→ Explain what is meant by time-division multiplexing.

Principles of communication systems

A communication system transfers information carried by signals, which may be analogue or digital. A system that transfers information without delay to the receiver as it is supplied is called a real-time communication system. Figure 1 is a block diagram showing an example of the essential parts of a real-time radio communication system.

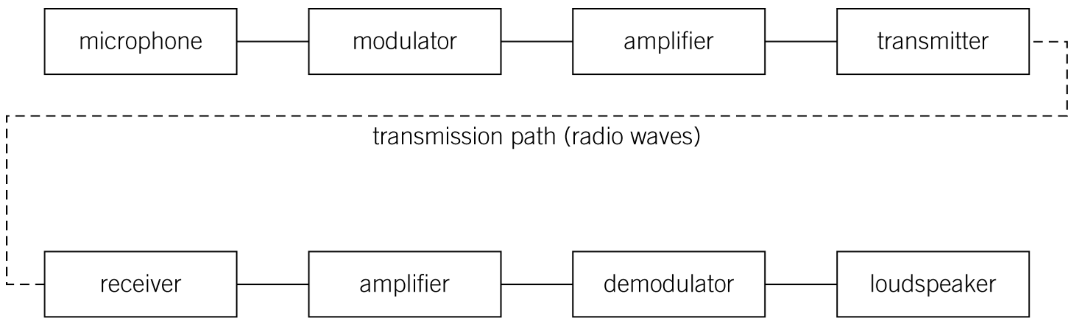

Figure 1 *A real-time communication system*

- The microphone is an **input transducer** that converts information into a signal.
- The modulator puts the signal onto a carrier wave.
- The amplifier increases the amplitude of the carrier wave and the signal.
- The transmitter sends the carrier wave and the signal along a transmission path.
- The receiver detects the signal.
- The second amplifier increases the amplitude of the carrier wave and the signal.
- The demodulator takes the signal off the carrier wave.
- The loudspeaker is an **output transducer** that converts the signal into information.

In this example, the input transducer signal is an analogue signal. It may be converted to a digital signal, in which case, pulse regenerators would need to be included in the transmission path and a digital-to-analogue converter would also need to be included before the signal is supplied to the output transducer.

Communication links and channels

Information may be transmitted along various transmission paths, including:

- twisted wire pairs and coaxial cables, as described below, carrying electrical signals
- radio and microwave links, including satellite links
- optical fibres carrying light signals.

The communication links listed above are described in detail in the following sections. The links may be compared in various ways, including:

- **bandwidth**, which is the range of frequencies that can be transmitted
- **attenuation**, which is the reduction in the signal amplitude due to physical processes such as absorption
- **noise**, which is random variations in the signal due to electrical interference
- **bit rates**, which is the number of bits per second carried by a digital signal.

Links and channels

The bandwidth of a communication link depends on the type of link. An optical fibre has a much wider bandwidth than a twisted wire pair or a coaxial cable. The bandwidth of a communication link may be wide enough to be subdivided into frequency **channels** with a different signal allocated to each channel. For digital signals, each channel can carry a number of digital signals using a process known as **time-division multiplexing**, as described at the end of this topic.

Table 1 at the end of this topic gives a summary of the comparisons.

Wire pairs

Wire pairs are used to carry low-frequency signals over short distances. Attenuation over longer distances occurs because:

- the heating effect of the current in the wires causes energy to be transferred as heat to the surroundings
- the wires act as aerials and radiate electromagnetic waves, which carry energy away from the wires. Twisting the wires together reduces this effect, but cannot eliminate it.

The wires may pick up external interference caused by electromagnetic waves from other sources and, where several twisted wire pairs are near each other, signals in one pair may be picked up by the other pairs.

Coaxial cables

A coaxial cable consists of an inner conductor in the form of a thick copper wire separated by polythene insulation from an outer copper braid conductor, as shown in Figure 2. The outer conductor is covered by a layer of plastic insulation.

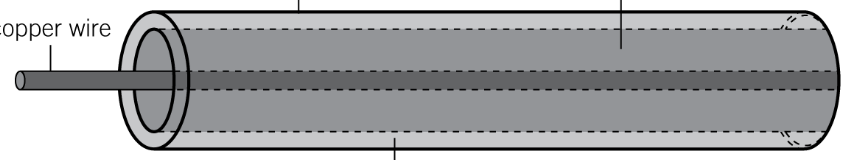

Figure 2 *A coaxial cable*

Attenuation of signals is much less along a coaxial cable than along a twisted wire pair. A typical coaxial cable has a bandwidth of about 60 MHz. Therefore, unlike a twisted wire pair, a coaxial cable of the same length is able to carry high-frequency signals up to 60 MHz without significant attenuation. For this reason, a coaxial cable can carry much more information than a wire pair can. For example:

- a coaxial cable is capable of carrying 100 million bits per second over a distance of 1 km
- a twisted wire pair is capable of transferring no more than about 10 million bits per second over a distance of 100 m.

Radio waves

Radio waves and microwaves are used for radio and TV broadcasting, satellite communications, and two-way communications including mobile phone communications. Radio waves are electromagnetic waves, so they travel at the speed of light. Electromagnetic waves radiate from a transmitter aerial when an alternating voltage of frequency greater than about 30 kHz is applied to the aerial. The frequency range covered by radio waves and microwaves is divided into wavebands for different users, as you can see in Table 1 at the end of this topic.

Transmitter aerials used for local, national, and international broadcasting stations need to be designed to transmit radio waves in all directions. The further a receiver aerial is from the transmitter aerial, the weaker is the signal it detects. This is because the waves spread out as they travel away from the transmitter, and as they travel through the atmosphere, they lose some energy because of atmospheric absorption.

Radio waves are affected by the Earth's atmosphere and surface according to frequency.

Surface waves

Low-frequency (below 3 MHz) or long-wavelength radio waves stay near the surface. They diffract and spread out from the transmitter and travel further over the sea than the land. Their range over land is about 1000 km. They are used for long-distance broadcasts and maritime communications. **Ground waves** are surface waves close to the Earth.

Sky waves

Medium and high-frequency waves reflect from the ionosphere in the upper atmosphere back to the ground, where they reflect back towards the ionosphere and undergo further reflections. These are called **sky waves**. The ionosphere is a belt of the upper atmosphere where gas atoms are ionised by solar particles to create ions and free electrons. The electron density of the ionosphere increases with height, causing the radio waves to change speed and refract. This refraction effect causes radio waves at frequencies below 30 MHz to be reflected back to the ground. Above this frequency, radio waves pass through the ionosphere. Therefore, medium and high-frequency radio signals can travel around the Earth as a result of being repeatedly reflected by the ionosphere and the ground surface.

Space waves

Radio waves of frequencies greater than 30 MHz (VHF, UHF, and microwaves) are used for line-of-sight communications and satellite links. VHF and UHF radio waves have a lower range than radio waves of lower frequencies that are transmitted at the same power, so VHF and UHF radio waves are used for local transmissions.

Microwave links

Microwaves are electromagnetic waves in the frequency range from 3 GHz to 30 GHz. They have a bandwidth of the order of GHz, so they can carry more information than radio waves can. They are used for satellite communications and for point-to-point communications between line-of-sight transmitters and receivers. For long-distance ground communications, many repeater stations are needed with line-of-sight adjacent transmitter/receiver towers.

In air, microwaves have wavelengths of about 10 cm or less. The transmitter aerial, or element, is fixed at the focal point of a parabolic reflecting dish. As a result, the waves from the transmitter element are reflected by the dish to form a parallel beam. When the beam is directed at a receiver consisting of a similar reflector with a receiving aerial at its focus, they are focused by the receiver reflector onto the receiver aerial. In this way, signals can be carried by the microwave beam along the line of sight between the transmitter and the receiver.

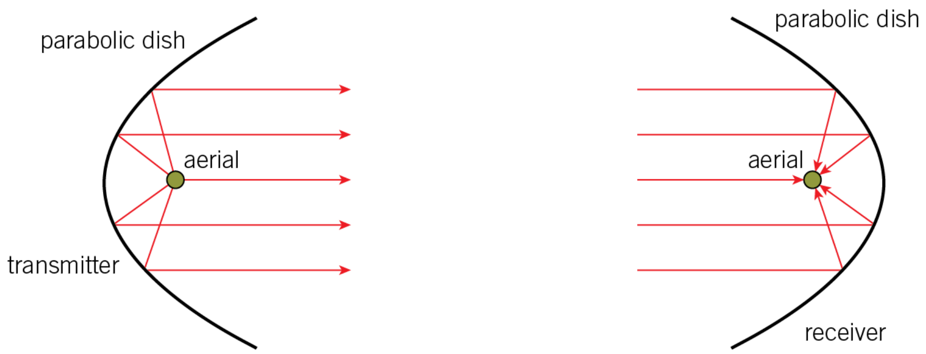

Figure 3 *A microwave link*

The parabolic shape of the reflectors ensures that the beam:

- is almost parallel so that as much power as possible is in the beam at the receiver
- is focused onto the reflector aerial so that as much power as possible is received from the beam.

Diffraction

The transmitter aerial is not a point object, so the waves diverge slightly when they reflect from the transmitter dish. Also, diffraction occurs when the waves are reflected by each dish. To minimise diffraction, the diameter of the transmitter dish must be much larger than the wavelength of the waves.

QUESTION: Describe how the signal-to-noise ratio is affected if diffraction is significant.

Compared with radio waves, microwave beams:

- spread out much less than longer-wavelength electromagnetic waves, so they weaken less
- have a much greater bandwidth, so they can carry many more channels
- pass straight through the atmosphere, unlike radio waves at frequencies below 30 MHz, so they can be used for satellite links.

Satellite communication signals such as phone signals are carried by microwave beams between geostationary satellites and ground-based receivers. The carrier frequency of the link between the ground and the satellite (i.e., the **up-link**) is different to that of the link between the satellite and the ground (i.e., the **down-link**). This is so that the transmitted signal does not swamp the reception of the signal received at the satellite (i.e., de-sense the receiver).

Long-distance communications using satellite links are more reliable than using radio waves because long-wavelength radio waves are poorly received in hilly areas, and medium- and short-wavelength radio waves rely on reflection by the ionosphere, which varies in reflectivity during the day. Satellite TV broadcasting is from geostationary satellites, which each transmit a microwave beam to dish receivers on a wide area of the Earth's surface.

Satellite uses

Satellites are placed in orbits that depend on the purpose of the satellite.

Geostationary satellites orbit the Earth directly above the equator with a period of 24 hours in the same direction of rotation as the Earth. Such a satellite remains directly over the same point on the equator. The orbital radius for such an orbit is 42 000 km.

Geostationary satellites are used for international phone links and satellite TV broadcasts. A single geostationary satellite can be used for TV broadcasting or for continuous communication between any two points on the Earth in the line of sight of the satellite. Several satellites sending signals between each other and the ground are needed for global communications. However, because the transit time of phone signals that travel long distances between geostationary satellites is unacceptably long, such signals are carried on the surface by optical fibre links as well or via a larger number of satellites in a lower orbit.

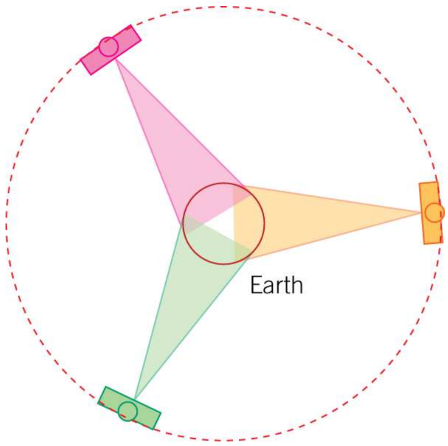

Figure 4 *Geostationary satellites*

Polar satellites pass directly over both of the Earth's poles with a time period of about 90 minutes. For such satellites:

- their orbits must be much lower than geostationary orbits
- successive passes across the equator in the same direction are further west each time. For a 90-minute orbit, each of these passes is 22.5° further west (= 360° × 1.5 hours / 24 hours).

Polar satellites are used for many purposes, including surveillance, monitoring, and weather forecasting. Because a polar satellite is in a much lower orbit than a geostationary orbit, its signals cover a much smaller area of the surface. Continuous communication outside the area covered by a single satellite is possible if there are several equally spaced satellites linked by microwave beams along the orbit.

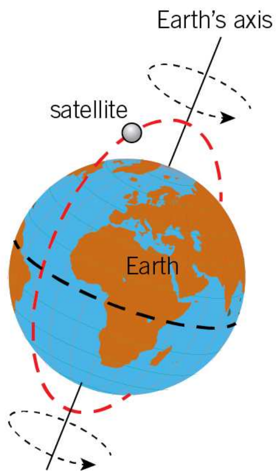

Figure 5 *A polar satellite*

Optical fibres

Optical fibres are used in communications to transmit infrared or visible light pulses coded as digital signals. Infrared light of wavelength about 1200 nm (i.e., frequency of the order of 250 THz) is usually used to carry optical signals in optical fibres.

Noise can be removed from such signals using regenerator circuits at intervals. Look back at Topic 2.1 if you need to. The fibres used are **step-index fibres**, which means that each fibre consists of a core of transparent material surrounded by a cladding of transparent material of lower refractive index. Light rays travelling along a fibre that reach the core–cladding boundary are totally internally reflected at the boundary, as long as the fibre does not bend too much. Figure 6 shows how a light ray travels along a step-index fibre as a result of successive reflections at the core–cladding boundary.

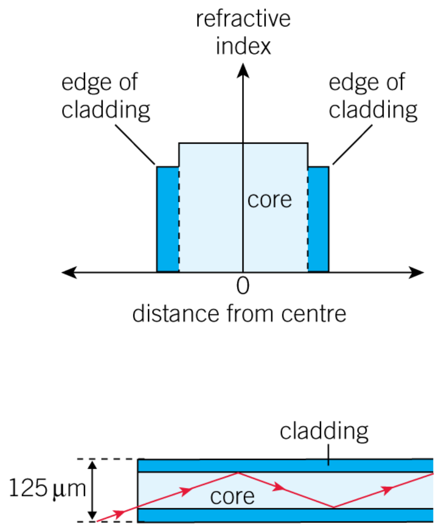

Figure 6 *A step-index fibre*

Link

Optical fibres were looked at in Topic 5.3, Total internal reflection, in Year 1 of the *AQA Physics* student book.

Light pulses from a laser source can be sent along an optical fibre at a rate of 10^8 pulses per second. The speed of light in a fibre depends on the light frequency. Light from an LED has a range of frequencies, so a pulse of light from an LED becomes longer as it travels along a fibre because its different frequencies travel at different speeds. This does not happen with a laser light, because it has a single frequency.

The duration of a single pulse, or bit, can in theory be as short as 10^{-14} s, corresponding to the sampling frequency needed for light at frequencies of about 10^{14} Hz. This would allow bits from hundreds of thousands of different signals to be transmitted between successive bits from any one signal. However, at present, laser pulses of duration less than about 10^{-10} s cannot be generated.

Notes

- A light pulse also spreads out because of **modal dispersion**, also called multi-path dispersion. This happens because a light ray that is repeatedly reflected at the core–cladding boundary travels further than a light ray that travels straight along the fibre. The leading edge of the pulse is defined by the light rays that travel along the shortest path. The trailing edge is defined by the light rays that travel along the longest path. The longer the length of the fibre, the greater the difference between the shortest and longest path, and therefore the longer the pulse. The effect is reduced considerably by using **mono-mode fibre**. Such a fibre has a very thin core of about 5 μm in diameter. Because the core is so thin, modal dispersion is effectively eliminated because light travelling along the core effectively travels along the shortest path.
- A laser pulse sent along a step-index fibre can become distorted due to absorption and scattering by the fibre material as well as loss through the core–cladding boundary if the fibre is bent too much. All these effects reduce the amplitude of the pulse.

Advantages of using optical fibres for transmission compared with other forms of transmission include:

- a greater transmission capacity in terms of bits per second
- a wider bandwidth (i.e., range of frequencies that can be transmitted)
- less signal power loss
- cheaper, lighter, and smaller than other cables
- high security because cross-talk is negligible and they do not radiate electromagnetic waves
- no interference from external electromagnetic waves.

Time-division multiplexing

More than one signal can be transmitted along a single channel by **multiplexing** the signals. This can be achieved for digital signals by reducing the duration of each bit and dividing the time interval between one bit of a signal and the next into a number of time slots of equal duration. Each slot is then used to carry a bit from a different signal. This technique is known as **time-division multiplexing**.

For example, if the time between successive bits of a signal is 7.8 μs, and each bit is a pulse of duration 0.2 μs, the time interval between successive bits of the same signal can be divided into 39 time slots. Therefore, 39 different signals could be carried in one frequency channel. In practice, several time slots would be used to synchronise the signals. However, the frequency channel could carry 30 different signals as long as it has a bandwidth of at least 5 MHz so that it can transmit bits at a rate of 5 MHz (= 1 bit every 0.2 μs). More signals can be carried if the pulse duration is shortened.

Table 1 *Summary of the advantages and disadvantages of each type of communication link*

Type of link	Twisted wire pair	TV coaxial cable	Radio / microwave link	Optical fibre
Typical attenuation	$\approx 5 \times$ weaker at 10 MHz per 100 m	$\approx 10 \times$ weaker at 100 MHz per 100 m	$\approx 10^8 \times$ weaker at 1 GHz per km	$\approx 0.95 \times$ weaker for infrared per km

Type of link	Twisted wire pair	TV coaxial cable	Radio / microwave link	Optical fibre
Bandwidth	500 kHz	50 MHz	FM: 200 kHz microwaves: 1 GHz	~ GHz for 1 km (bandwidth decreases with increasing distance)
Noise	electrical interference from electric motors, etc., cross-over from other wires	electrical interference from electric motors etc.	electrical storms, cosmic sources	no electrical interference, noise is removed by regenerators
Bit rate / 10^6 bits per second	10	100	digital TV: 20 microwave beam: 70	> 10^6
Repeater / regenerator spacing	short distances	up to 2–3 km	line of sight for microwaves	up to 100 km or more
Security	low, unless coded digital signals used			high
Cross-over	causes noise	insignificant	if atmospheric conditions extend local signals	insignificant
Cost / convenience	low / high	medium / medium	high / low	low / high

Summary questions

1 State one advantage and one disadvantage of an optical fibre link compared with a microwave link between a transmitter and a receiver 30 km away.

2 Explain what is meant by a geostationary satellite, and describe how such satellites can be used for global communications.

3 **a** Explain why a coaxial cable can transmit a digital signal of bandwidth 10 MHz, whereas a twisted wire pair cannot.

b Explain why a signal transmitted along a coaxial cable cannot be detected if the cable is too long.

4 **a** Explain what is meant by time-division multiplexing.

b Describe how a mobile phone frequency channel can carry up to eight different calls.

c A digital signal consists of 16-bit words transmitted every 0.10 ms. Each bit is a pulse lasting 0.125 μs. Calculate how many of these signals could be transmitted using time-division multiplexing in a single frequency channel, and calculate the bit rate in the channel.

5.2 Amplitude modulation and frequency modulation

Learning objectives:

→ Describe and compare amplitude and frequency modulation.
→ State and compare the bandwidths of amplitude- and frequency-modulated signals.
→ Explain what is meant by pulse code modulation.

Modulation principles

When you are listening directly to someone, their vocal cords are vibrating and producing sound waves that are modulated (i.e., varied by the vocal cords) to carry information. The sound waves travel through the air and are detected by your ears.

When you listen to a radio programme, your radio is detecting radio waves of a particular frequency transmitted by a radio station. In the studio at the radio station, sound waves are converted to electrical waves, called **audio waves**, using a microphone. These waves are then used to modulate the radio waves sent out by the radio station. The radio waves from the transmitter are called **carrier waves**. They have much higher frequencies than the upper frequency limit of audio waves, which is about 15 kHz. Radio stations send out modulated carrier waves rather than audio-frequency radio waves because such audio-frequency radio waves would:

- need very long aerials
- have a very short range compared with carrier waves of the same intensity
- be impossible to select from the waves sent out by other radio stations.

Audio waves are an example of an **information signal**. Other examples of information signals are TV signals, mobile phone signals, and data transfer to and from computers. Information signals can be carried by electromagnetic waves, or by electrical waves in the form of alternating voltages. The signal, which could be either analogue or digital is used as described below to modulate the carrier wave.

An **analogue signal** can have any value from zero to an upper limit. It can be carried by a wave by using the signal in one of two ways: **amplitude modulation (AM)** or **frequency modulation (FM)**.

Amplitude modulation

The amplitude of a carrier wave of constant frequency is modulated by the amplitude of the signal to be carried. Figure 1 outlines how an analogue signal is produced using AM.

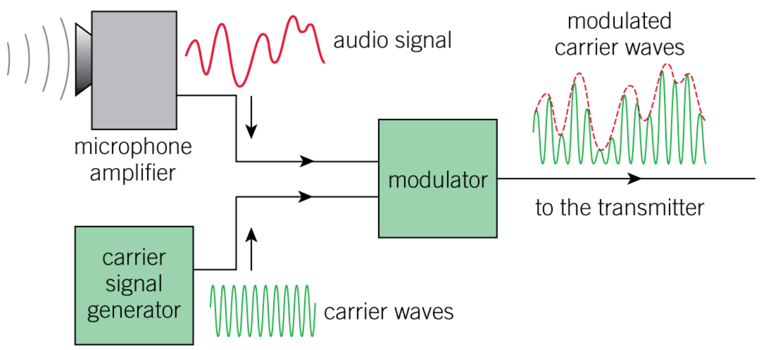

Figure 1 *Amplitude modulation*

Frequency modulation

The frequency of a carrier wave is modulated by the amplitude of the signal to be carried. Figure 2 outlines how an analogue signal is produced using FM.

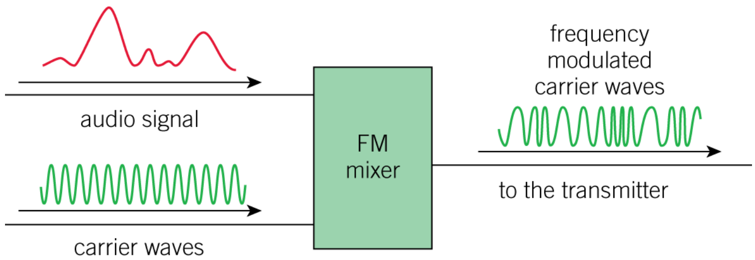

Figure 2 *Frequency modulation*

A **digital signal** is a signal that has only two values. It can be carried by a wave by using the signal to switch the amplitude or the frequency between two different levels.

More about pulse code modulation

AM and FM are terms usually applied only to analogue signals and not to digital signals. An analogue signal may be transmitted by converting it to a digital signal, which is then transmitted using pulse code modulation (PCM), as explained in Topic 2.1.

Advantages of using PCM to transmit a signal digitally compared with FM or AM transmission are:

- The reconstructed signal has no noise and is of a higher quality than an FM or AM signal.
- More than one signal can be transmitted using time-division multiplexing on a single frequency channel.
- A PCM signal can be coded to prevent it being read by unauthorised users.

Disadvantages of using PCM to transmit a signal digitally compared with FM and AM transmission are:

- PCM signals need extra processing circuits for encoding and decoding.
- The higher the sampling frequency or the longer the word length, the greater the bandwidth of a PCM signal.

Bandwidth

Consider a carrier wave of frequency f_c modulated by a signal of constant frequency f_m. The carrier wave is a wave of constant frequency. Before it is used to carry a signal, its frequency spectrum is a single line at frequency f_c. To modulate it, the signal to be carried is added to the carrier wave.

For AM, the frequency spectrum of the modulated carrier wave consists of the same line as before and two lines at either side at frequencies $f_1 = f_c - f_m$ and $f_2 = f_c + f_m$, as shown in Figure 3.

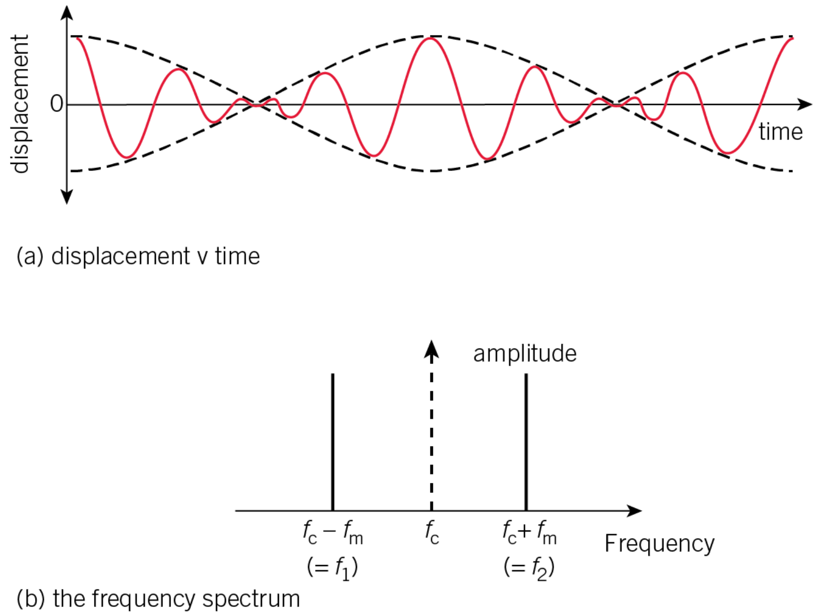

Figure 3 *A modulated carrier wave*

Two side frequencies are produced because any waveform can be resolved into sinusoidal components (or can be synthesised from sinusoidal components). The modulated carrier wave can be resolved into a sine wave of frequency $f_c - f_m$, a sine wave of frequency f_c, and a sine wave of frequency $f_c + f_m$.

In general, the information signal frequency varies between zero and an upper limit f_M, and the frequency spectrum consists of a line corresponding to the carrier frequency f_c and two sidebands of width f_M either side of the carrier frequency.

- The range of frequencies is called the **bandwidth.** This is equal to the difference between the upper and lower sideband frequencies and is therefore equal to $2f_M$.
- The relative amplitude of the sidebands and the carrier depends on the relative amplitudes of the audio and the carrier waves. No sidebands are produced if there is no audio signal.

For FM, the frequency spectrum is much wider than for AM. For a constant modulation frequency f_m, a large number of side frequencies at $\pm f_m$, $\pm 2f_m$, $\pm 3f_m$, and so on, either side of the carrier frequency f_c are produced at amplitudes that decrease the further the side frequency is from the carrier frequency. In practice:

- the information signal frequency varies between zero and an upper limit f_M, therefore producing a large number of sidebands of width $\pm f_M$, $\pm 2f_M$, and $\pm 3f_M$
- the outer sidebands are filtered out at frequencies above a **maximum deviation** Δf from the carrier frequency.

Therefore, the bandwidth of an FM signal is approximately equal to $2(\Delta f + f_M)$.

the bandwidth of an AM signal $= 2f_M$

the bandwidth of an FM signal $= 2(\Delta f + f_M)$

Notes

A radio or TV channel is a continuous band of frequencies of radio waves allocated to carry the channel signal. The width of the allocated band is the channel bandwidth. The signal bandwidth is restricted to fit the channel bandwidth. See Figure 4.

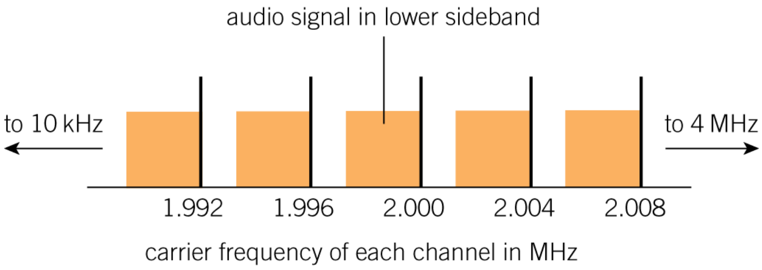

Figure 4 *Frequency channels*

The bandwidth of an FM signal is much greater than that of an AM signal.

- For AM radio signals, the channel bandwidth is restricted to 9 kHz to prevent frequency overlap from nearby stations. As a result, the upper frequency limit of the audio waves carried by AM signals is restricted before modulation to 4.5 kHz.
- As explained above, FM signals have further upper and lower sidebands at multiples of the audio frequencies (whereas AM signals only have one upper and one lower sideband). Because of these extra sidebands, FM radio bandwidths are about 200 kHz compared with 9 kHz for AM bandwidths. The much wider FM bandwidth means that audio waves of frequencies up to 15 kHz can be carried by FM signals.

The relative advantages of AM and FM transmission

Quality

The quality of a signal is affected by **noise**, which is unwanted variations in the signal amplitude that are superimposed on the transmitted signal. Noise has various causes, including the random motion of conduction electrons in resistors in electrical circuits and in radio and TV communications by atmospheric effects. A signal is difficult to detect if too much noise is present. Figure 5 shows the waveform of an analogue signal that is difficult to discern because so much noise is present.

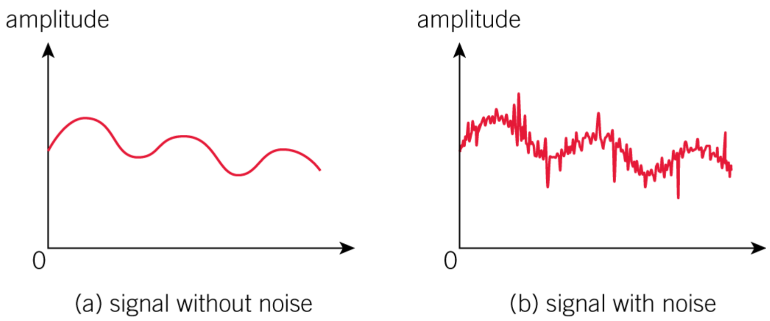

Figure 5 *A noisy signal*

FM radio transmission is less affected by noise, signal distortion, and other unwanted amplitude variations such as atmospheric effects. This is because amplitude variations do not affect the frequency modulation of the signal. As a result, received FM signals are of higher quality than received AM signals. Also, because FM signals carry a higher range of audio frequencies than AM signals, FM signals reproduce higher-quality sounds (e.g., for music programmes).

Number of channels

FM radio stations use carrier frequencies that are much higher than those used by AM stations. This is because FM bandwidths are much greater than AM bandwidths. A single FM channel with a bandwidth of 200 kHz using carrier waves of frequencies 100 kHz would use up the entire frequency spectrum between 0 and 200 kHz. By allocating FM stations to the VHF waveband, which has a frequency range from 30 to 300 MHz, FM stations can operate at different carrier frequencies (e.g., 98.2 MHz, 98.6 MHz) without their bandwidths overlapping.

Note

AM transmitters and receivers are simpler and cheaper than FM transmitters and receivers.

Range

An AM transmitter has a much greater range than an FM transmitter broadcasting at the same transmitter power. A typical FM station has a range of about 30 km by line of sight. This is because the **attenuation** (i.e., reduction of intensity) of an electromagnetic wave travelling through air increases with increase of frequency. In other words, high-frequency radio waves are attenuated more than low-frequency radio waves. So the signal from a single transmitter can reach a much larger area if it is an AM signal rather than an FM signal. To reach the same area, a large number of FM transmitters would be needed to replace one AM transmitter. For this reason, national and international radio stations use AM transmission, and local radio stations use FM transmission. Table 1 shows the wavebands that are used by each type of radio station.

Table 1

Waveband	Frequency range	Wavelength range	Uses
Low frequency (LF) / long wavelength (LW)	< 300 kHz	> 1 km	International long-wavelength AM radio
Medium frequency (MF) / medium wavelength (MW)	300 kHz to 3 MHz	100 m to 1 km	AM medium-wavelength radio
High frequency (HF)	3 to 30 MHz	10 to 100 m	AM short-wavelength radio
Very high frequency (VHF)	30 to 300 MHz	1 to 10 m	Local FM radio
Ultra high frequency (UHF)	300 to 3000 MHz	0.1 to 1 m	TV broadcasting, mobile phones, person-to-person radios
Microwave	> 3000 MHz (= 3 GHz)	0.1 mm to 0.1 m	Satellite links, satellite TV

Summary questions

1 List the following communication systems in order of increasing frequency:

FM radio International radio National radio Satellite TV

2 a With the aid of diagrams, explain the difference between an analogue signal that is amplitude modulated and an analogue signal that is frequency modulated.

b Using Figure 3, estimate the frequency of the information signal if the carrier wave has a frequency of 200 kHz.

3 A radio transmitter broadcasts a frequency-modulated signal of bandwidth 250 kHz at a carrier frequency of 101 MHz. Explain what is meant by the bandwidth of a signal, and calculate the upper and lower limits of the allocated frequency channel, assuming only the lower sideband is used.

4 a FM radio stations have a bandwidth of 250 kHz and broadcast in the VHF waveband. State one reason why FM radio is not allocated to the HF waveband or to the UHF waveband.

b An FM radio station transmits its signal in a frequency channel of bandwidth 240 kHz at a carrier frequency of 101.25 MHz. The upper limit of the modulation frequency is 15 kHz. Calculate:

i the maximum deviation of the information signal frequency from the carrier frequency

ii the carrier frequency of the next higher-frequency channel.