AQA Physics A-level

Topic 13: Electronics

Key Points

Transistors

A transistor is a commonly used form of **electronic switch** that consists of three connections:

1. Gate
2. Drain
3. Source

Charge carriers enter at the source and leave through the drain, however only if a sufficient voltage is applied to the gate. **Conventional current** flow is in the opposite direction and so will show current entering at the drain and leaving from the source.

You should be aware that transistors make use of **extrinsic semiconductors**, which are a type of semiconductor that have had impurities added to them, to alter the material's conductivity in a certain way. Transistors use **n-type semiconductors**, which have had extra conduction electrons added to increase their **conductivity**. The 'n' represents negative, since the majority of the charge carriers are negatively charged electrons.

MOSFET Characteristics

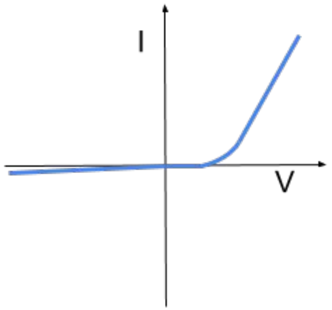
When the voltage at the gate (V_{GS}) is zero, the drain current (I_{DSS}) will take a specific value. This value will alter with changes in **temperature** and means that the **threshold voltage** (V_{th}) will fall as the temperature increases. This gives rise to MOSFET **characteristic** diagrams, which you should be able to interpret.

MOSFET transistors can be used to switch on **high current devices** such as solenoids, using a low current circuit such as a **low power logic system**. When using it for this purpose however, **protective** precautions should be made:

- If the device that is being switched on by the MOSFET is **inductive** (such as a solenoid), when the MOSFET is switched off, a consequence of **Lenz's law** is that the coil will produce a large voltage to try and maintain the **field** - this large spike in voltage could **damage** the component and so a **diode** is added in parallel to prevent this
- MOSFETS have a **high input resistance** which means that **electrostatic charge** may build up at the gate if it is unconnected, causing the switch to be inadvertently switched on - to prevent this, a **resistor** should be added between the gate and 0V (note that this resistor value should be **very high** to prevent it drawing current from the input signal)

The Zener Diode

A **p-n junction diode** is the type of diode studied in the electricity section of the A-Level course. The I-V characteristic of such a diode is shown to the right, however you should also be aware that at **very large reverse voltages**, the diode will breakdown, and current will be able to flow in the reverse direction. Usually, this reverse current flow is undesirable and so the **breakdown voltage** (V_{BR}) is designed to be as high as possible. For Zener diodes however, this is not the case.



A zener diode is a diode that has been designed to have a much smaller and specific breakdown voltage. You should be aware of its properties:

- In the **forward direction**, a zener diode has a similar I-V characteristic to a standard diode
- The **breakdown voltage** is **independent** of the reverse current (I_Z) beyond around **5mA**
- Zener diodes can be used to **compare** an input voltage with a fixed voltage, for example in an oil temperature sensor in a car - a warning light is switched on if the temperature exceeds a safe level
- They can also be used provide a **fixed voltage source**

The Photodiode

A **photodiode** is a type of **p-n junction diode** for which the current flow through it increases with an increase in light intensity. This occurs because:

- **Photons** of light are absorbed
- The photons transfer **energy** to the charge carriers in the diode
- The **charge carriers** are released, increasing the current flow

This process happens in all p-n junction diodes, and so with the exception of a photodiode for which this effect is desired, all other types must be sealed in **light-tight casings**.

Photodiodes can operate in one of **two** different **modes**:

1. Photoconductive (Reverse-current biased)
2. Photovoltaic (Forward-current biased)

Photovoltaic photodiodes are used for applications such as **photocells** whereas photoconductive photodiodes are preferred for uses in **electronics**. This is because they have a much **quicker response time** and the relationship between light intensity and current is very **linear**.

The Hall Effect Sensor

A **hall effect sensor** is a **magnetic field sensor** that converts a magnetic field into a voltage output. The basic function of it is as follows:

- The sensor, consisting of a **p-type semiconductor** being supplied with a **constant current**, is placed in an **external magnetic field**
- This field causes the charge carriers passing through the semiconductor to **deflect**
- This deflection creates a potential difference, known as the **Hall voltage**

The Hall voltage is usually **very small** even for sizeable fields, and so the sensor will often be connected to an **amplifier**. The voltage is **linearly** related to the magnetic field strength, meaning a Hall Effect sensor can be used to measure the strength of fields.

Another **use** of Hall effect sensors is to turn on an **output** when the **field strength exceeds a certain value**. For example, a sensor could be connected to a **rotating engine part**, and the output pulses from the sensor can be used to detect the speed of the part's rotation.

They are **preferred** over traditional switches since they don't involve **moving parts** making them more **reliable**, **quicker**, don't suffer from **contact bounce**, and are unaffected by **environmental** conditions. They are also used in applications such as **fuel level sensing** where traditional switches could be dangerous.

Analogue Signals

An **analogue signal** is one that can **any value** - there are an **infinite number** of different levels that an analogue signal can take. Examples of analogue signals include the output of a **light sensor** or **microphone**, both of which can take any number of different frequency inputs.

You should be aware of the different **features** of an analogue output:

- The main **parameter** of an analogue signal is **size**
- The **peak voltage** (V_{pk}) is the greatest magnitude of the positive half of the signal
- The **peak-peak voltage** ($V_{pk\text{-}pk}$) is the difference between the greatest signal in the positive and in the negative half
- The **root mean square** (r.m.s) value is another measure of the signal's amplitude

Note that the formula for calculating the rms voltage of a signal given in the oscilloscope topic of the A-Level course, only applies for **sinusoidal wave forms** and so often can't be used for many analogue signals. In these circumstances, the V_{pk} should be given as a measure of the signal's amplitude.

Analogue Sensors

You should be aware of several different **examples** of analogue sensors and what they measure:

- Thermistor - temperature
- LDR - light intensity
- Pressure sensor - pressure
- Strain gauge - mechanical deflection
- Moisture sensor - moisture
- pH probe - pH level
- Lambda sensor - oxygen levels
- Hall effect device - magnetic field strength
- Microphone - sound

It is important to note that in some of these examples, the device itself doesn’t produce a **voltage** output. Thermistors for example, experience a change in **resistance** which is later converted into an associated voltage with **additional circuitry**.

Digital Signals

A **digital signal** can only take **discrete values**. It is also generally accepted that digital signals are ones that can only take **one of two values**. It is this idea that allows **binary form** to be used, which has in turn enabled the development of devices such as computers. It is useful to understand that:

- The two states of a digital signal are often just referred to as On/Off, High/Low, True/False or 1/0
- The **absolute values** of the two states rarely has importance

- A combination of 8 digital signals makes up a **byte**
- Each signal within a byte is known as a **bit**

There are two main types of circuit that make use of digital signals:

- **Combinational logic circuits** involve decisions being made based on input signals
- **Sequential logic circuits** involve counting and pulsing

Analogue-Digital Conversion

Most real-world measurements produce **analogue signals**, whereas the most useful form for **processing** data are **digital signals**. This means that an **A-D converter** (ADC) must be used. You should understand this process:

- The analogue is **sampled** and a given rate known as the **sampling rate**
- The ADC converts the signal at each point into a binary number that represents the **instantaneous amplitude** at that point
- These numbers combine to form a **digital signal output** that can be easily processed

The process of converting analogue amplitudes into binary numbers is known as **quantisation**. The sequence of numbers produced is referred to as the **digitised signal**.

The opposite of an ADC is a DAC and is when the digitised signal is converted back into an analogue, 'real-world' output.

Sampling Rate and Resolution

The **quality** of analogue to digital conversion depends on two main factors. The first factor is the **sampling rate**:

- The sampling rate is the **frequency** at which the signal is sampled
- If the sampling rate is **too low** some of the signal variation will be missed - this results in an unrepresentative digital signal
- If the sampling rate is **too high**, the electronics needed to carry out the conversion become too complex and the bandwidth will increase

The second factor that affects the quality of a digitised signal is the **resolution**:

- The resolution is the **range** of numbers that the amplitude is represented by
- It is determined by the **number of bits** in each sample
- The higher the **resolution**, the **smaller** the change in the analogue signal that will cause a noticeable change in the digitised signal

Advantages and Disadvantages

There are a number of **advantages** associated with digital sampling and digital signals:

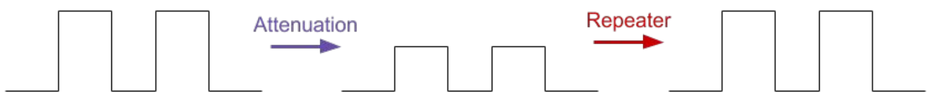
- They are **easier to process** and allow for more complex processing techniques such as the **digital filtering** of sound
- They are more **easily stored** and preserved than analogue signals
- They are less affected by **noise** since the only thing that needs to be distinguished is whether the value is a high or a low - this means that slight **fluctuations** don't ruin the signal
- They can be **encoded**
- Encoding different digital signals in different ways allow several different signals to be passed through a **single data link** without confusion

There are however associated **disadvantages**. An example is that digital signals require a much **greater bandwidth** than its equivalent analogue version. One solution to this is through compression, for example audio compression can be carried out on an MP3 file.

Digital Noise Reduction

Noise reduction is vital in order to produce a clear, easily processed digital signal. There are two main ways to reduce the effects of noise:

1. **Repeaters** can be used to amplify a weakened signal:



2. **Regenerators** can be used to regenerate a noisy signal:

The LC Resonant Filter

An **LC resonant filter** is a type of filter than be used to **process** an analogue signal. It is based on **two** components, the first being an **inductor**:

- An inductor is made up of **coils of wire**
- They **oppose current flow**, dependant on **frequency**
- A measure of an inductor's opposition to current flow at a given frequency is its inductance
- **Inductance i**s dependant on the **number of turns** of wire and the **material** it is wrapped around
- The unit of inductance is the **Henry** (H)

The second component that a LC resonant filter is based on is a **capacitor**:

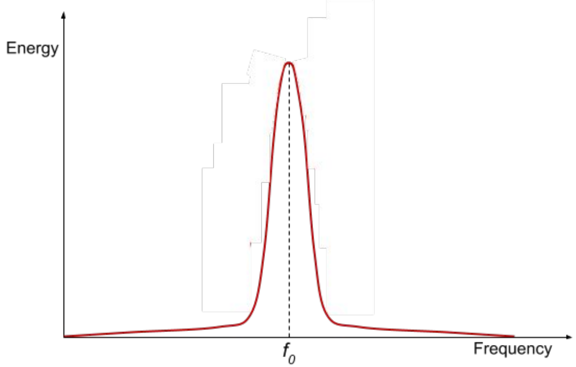
- Unlike as studied in the A Level capacitors section, the capacitors in an LC resonant filter operate under **alternating current**
- The **charge** on a capacitor can't be change **instantaneously**
- The greater the **rate** of voltage change applied across it, the greater the **initial current**

The Parallel LC Circuit

A **parallel LC circuit** behaves in different ways depending on the **frequency** applied to it:

- At very **low frequencies** XL will be small whilst Xc is **large** meaning the **net reactance** will be **low**
- At very **high frequencies** XL will be large whilst Xc is small, once again giving a **low net reactance**
- At the **resonant frequency**, the energy will continually transfer between the capacitor and inductor and there will appear to be **infinite** reactance

The consequence of this is that a **resonant filter circuit** can be created. These circuits can select a **narrow range of frequencies** and is the type used by **radios** when tuning into a specific radio broadcast.



Resonance

The resonant frequency of an LC circuit occurs when the **reactance** of the capacitor and inductor are **equal**. This means that an equation for the **resonant frequency** (f_0) can be calculated:

$$2\pi fL = \frac{1}{2\pi fC} \qquad f = \frac{1}{2\pi\sqrt{(LC)}}$$

You should also be able to compare the resonance of an LC circuit with the resonance of a **simple harmonic system** such as a mass on a spring.

$$T = 2\pi\sqrt{\frac{m}{k}} \qquad f = \frac{1}{2\pi\sqrt{(m/k)}} \qquad \textit{And so...} \qquad LC = m/k$$

Increasing mass has the **same effect** as increasing inductance (the resonant frequency is lowered) and so they can be said to be analogous. This means that it can also be said that capacitance is **analogous** to 1/k.

The Ideal Operational Amplifier

An op-amp **amplifies** the voltage from **two inputs** to produce a **single output** equal to the sum of the two magnitudes. You should understand the following:

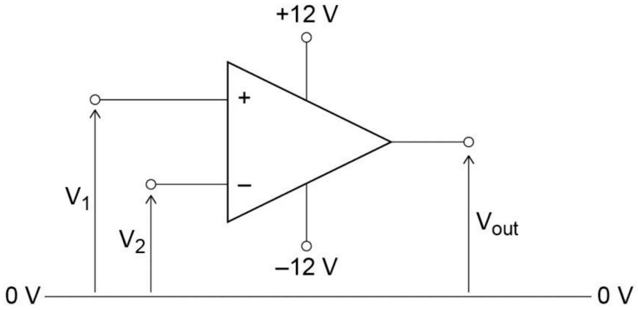
- The **output voltage** is represented by $\mathbf{V_O}$
- $\mathbf{A_{OL}}$ represents the **open-loop gain**
- An **open-loop circuit** is one where no signals from the output is fed back into the inputs
- Other types of circuits may be **closed-loop**, where **feedback** is fed back to the inputs
- An **output voltage** greater than the **supply voltage** can never be generated due to **saturation**

Ideal op-amps make the following **assumptions**:

- A_{OL} is **infinitely large**
- The input draws **no current** from the source of the signals
- The input has an **infinite input resistance**
- Equally it is assumed that the output current doesn't affect the output signal and so the **output resistance is zero**
- It has an **infinite bandwidth**

The Comparator

An op-amp can be used as **a voltage comparator** when placed in an **open-loop** circuit:

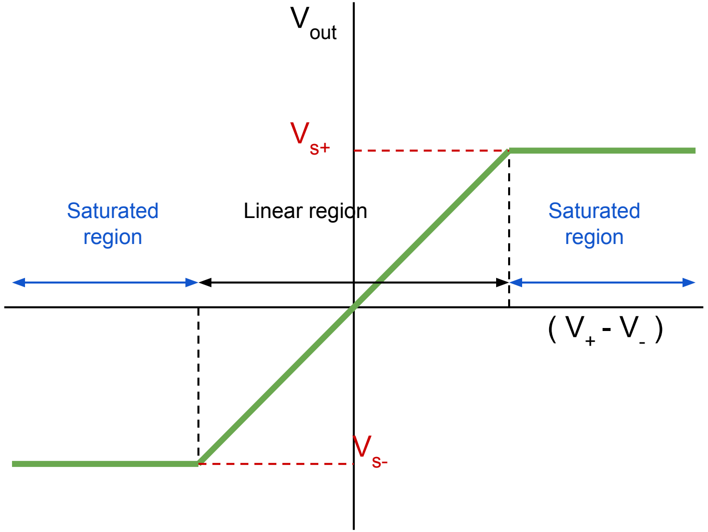


Taken from the AQA Electronics Teaching Guide

- Assuming an ideal op-amp, the theoretical output would be **infinite** (A_{OL} is theoretically infinite), however in practice, it will instead reach the supply voltage
  - Since $V_{out}= A_{OL}(V1-V2)$, if V1>V2 the V_{out} will equal +12V
    - If V2>V1, the V_{out} will equal -12V
- This means that the **magnitude** of the input values is insignificant, but that the op amp acts as a comparator to produce one of two **outputs**

Saturation

An **operational amplifier** can experience a property referred to as saturation. This occurs when the **input** applied to the op-amp is such that it attempts to produce an **output** that is greater than its **supply voltage**.



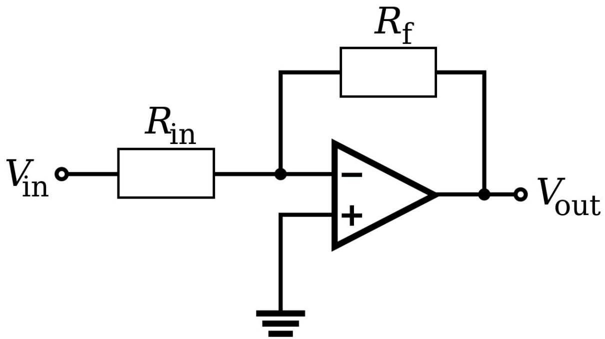
The variation of an op-amp's **output voltage**, against the **change** in its **two input voltages**, is shown to the left. From this you can observe the following:

- Regardless of any increase in the input voltage magnitude, the output voltage will **not exceed** the source voltage
- This applies to both **positive** and **negative** voltage outputs
- The gradient of the linear region is A_{OL}, which can be derived from:

$$V_{out} = A_{OL}(V1 - V2)$$

Inverting Amplifier Configuration

In an **inverting amplifier configuration**, the output voltage from the operational amplifier is fed back into the **inverting input** of the op-amp, forming a **closed-loop** circuit with **negative feedback**. Note that the non-inverting input (+) is connected to the ground and so is at 0V, and that it is assumed to be an ideal op-amp and so the open-loop gain (Aol) is infinite. This leads to the following derivation of the inverting input value:



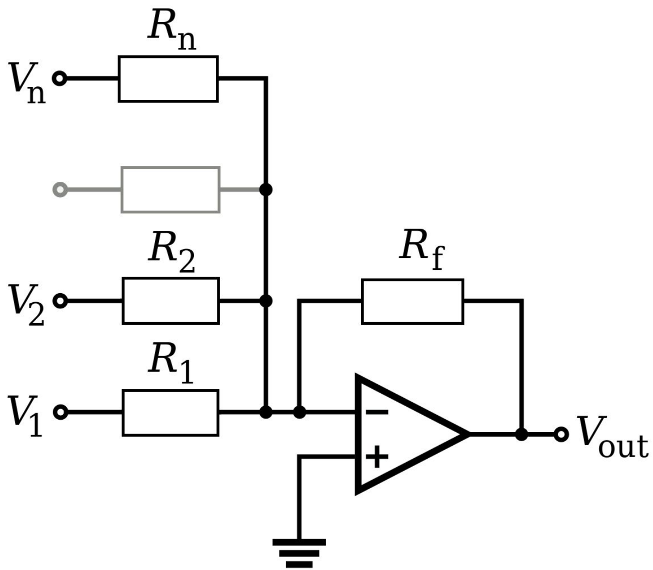
$$V_{out} = A_{ol}\,(V_+ - V_-) \qquad V_{out} = \infty\,(V_+ - V_-) \qquad V_- = \frac{-V_{out}}{\infty} \qquad \therefore\ V_- = 0V$$

Since V_- is effectively at 0V, it is referred to as a **virtual earth**. This idea is then used to derive the **transformer function** for the circuit:

$$-\frac{R_f}{R_{in}} = \frac{V_{out}}{V_{in}}$$

The Summing Amplifier

The inverting amplifier configuration can be extended to allow for **several signals** to be inputted into an amplifier and be amplified **independently**.



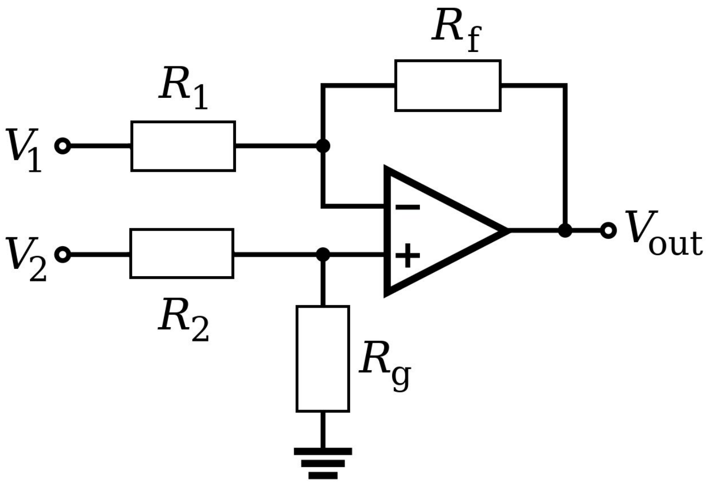
- **Virtual earth** allows the signals to be inputted without **interference**
- The **total output** from this configuration will equal the **sum** of all the individual **input signals**
- Used for purposes such as **audio mixers**

The **transfer function** for this configuration is:

$$V_{out} = -R_f \left(\frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3} \ldots \right)$$

The Difference Amplifier

Similar to the summing amplifier configuration, the difference amplifier handles multiple inputs.



- Produces the **difference** between two signals
- The gain for the two inputs is usually **equal** so R1 and R2, and Rf and Rg, are usually the same

The **uses** of the difference amplifier configuration include:

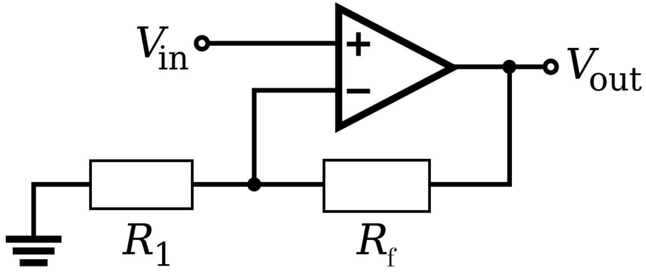
- **Microphones:** 50Hz interference picked up through microphone leads cannot be filtered out without losing part of the original signal, so a difference amplifier is used
- **ECG machines:** Interference can be eliminated by connecting an electrode to a position where only the interference will be picked up, and then using a difference amplifier to remove it from the actual signal

The **transfer function** for this configuration is:

$$V_{out} = (V_+ - V_-)\frac{R_f}{R_1}$$

The Non-Inverting Amplifier

For some purposes, it is important that the signal being amplified retains its **polarity**. For these purposes, the **non-inverting amplifier** is used:



- Unlike with the inverting amplifier configuration, the non-inverting amplifier **doesn't have** a **virtual earth**
- The gain of the amplifier can **never** be lower than **1**
- The circuit doesn't draw any **current** since the input signal is connected directly to the op-amp (which is modelled to have **infinite input resistance**)

The **transfer function** for this configuration is:

$$V_{out} = V_{in}\left(1 + \frac{R_f}{R_1}\right)$$

An important observation, is that if the Rf becomes 0 and/or R1 becomes infinite, a **unity gain buffer circuit** is produced:

- The configuration will have a **gain of 1**
- **No amplification** will occur
- Used to **interface** between a signal source and device with low input resistance without disturbing the signal

Real Operational Amplifiers

Throughout the prior analysis of the different **amplifier configurations**, we have made the assumption that the operation amplifiers behave as if they were **ideal**. In reality, their **characteristics** are slightly different.

	Ideal	Real
Open-loop gain (A_{OL})	Infinite	10^6
Input resistance (R_{in})	Infinite - draws no current	10^{12}
Output resistance (R_{out})	Zero	100 Ω
Output voltage (V_{out})	-VSVout +VS	-VS<Vout< +VS,Vout 0
Bandwidth	Infinite - can operate at any input frequency	Open-loop - Around 15 Hz Closed-loop - Around 2.5 MHz

- Op-Amp characteristics are affected by their **temperature** and **environment**
- The **real values** are significantly large enough/small enough that it is still suitable to assume that they are infinite/zero
- The **output resistance** of a real op-amp means that there is a **limit** on the current that it can deliver

Another key difference to be aware of is that when the inverting and non-inverting terminals are the same (or both grounded), an ideal op-amp would output a value of 0 V, however a real op-amp will output a small voltage called the **offset voltage**. Often, this value must be **cancelled out** for the circuit to function as required.

Bandwidth

The **bandwidth** of an op-amp is the **range of frequencies** at which it operates. In general, op-amps have a **very small** bandwidth due to the **large gains** they generate - changing the output from one voltage extreme to the other cannot be done quickly. There are **two types** of bandwidth you should be familiar with:

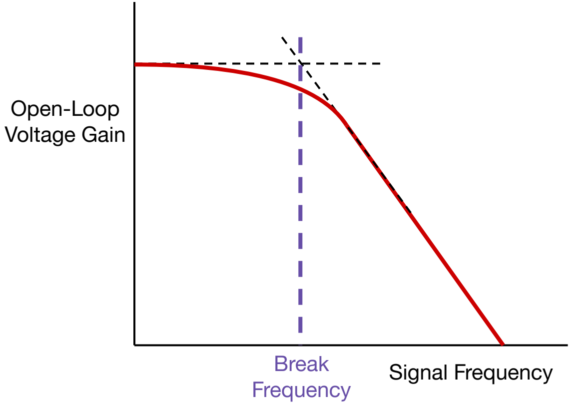
1. **Open-Loop Bandwidth**: Op-Amps generally have a value between 10Hz and 30Hz
2. **Closed-Loop Bandwidth**: Op-Amps generally have a value between 2MHz and 3MHz

The closed-loop bandwidth is achieved when the op-amp is in a circuit involving a **unity gain amplifier** configuration. The set-up will result in a **lower gain** but a **higher bandwidth frequency**, as can be observed in the above figures.

The closed-loop bandwidth may also be referred to as the **Gain-Bandwidth Product** (GBP.)

Frequency Response Curves

A **frequency response curve** can be constructed for an op-amp to show how the **voltage gain** varies with the input signal's **frequency**. Due to the magnitude of the values involved, the scale used when plotting these diagrams is usually **logarithmic**.



- The **break frequency** is the value at which the voltage gain begins to change
- To locate the break frequency from a diagram you need to use **straight line approximations**, as shown to the left
- The point where the two construction lines **meet** is the break frequency
- If the signal frequency starts at 0, the break frequency also gives the value of the op-amp's **bandwidth**

For a **closed loop circuit**, the break frequency will be significantly higher than for an open-loop circuit, since the gain will be less. It is important to note that the relationship between voltage gain and bandwidth is **linear**, and so:

$$\textit{Gain} \times \textit{Bandwidth} = \textit{Constant}$$

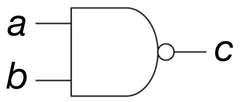
Logic Circuits

Logic circuits are circuits that are designed to process **inputs** and generate the required **outputs** through the use of **logic gates**. Logic gates produce **digital outputs** from **digital inputs**. This means that the inputs and outputs are either 'high' or 'low'.

A 'high' input is produced by a voltage, whereas a 'low' input is usually 0V. 'High' inputs can be represented with a '1' whereas 'low' inputs are represented with a '0'.

The truth tables showing the combinations of inputs (a and b) and their corresponding outputs (c) for three of the main types of logic gates are found below.

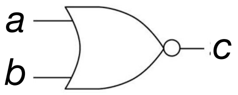
NAND Gate



If both inputs are on, the output is off. Otherwise, the output is on.

a	b	c
0	0	1
0	1	1
1	0	1
1	1	0

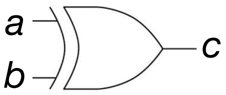
NOR Gate



If both inputs are off, the output is on. Otherwise, the output is off.

a	b	c
0	0	1
0	1	0
1	0	0
1	1	0

EOR Gate



The output is on if one of the inputs is on. If both inputs are on or off, the output is off.

a	b	c
0	0	0
0	1	1
1	0	1
1	1	0

Truth Table Analysis

There are **two** main methods of analysing logic circuits that you should be able to carry out. The first of these is known as **truth table analysis**, which is carried out as follows:

1. Assign a different letter to the output of every logic gate in the circuit
2. Construct a truth table consisting of the the **whole system inputs** on the far left, the **internal outputs** in the middle and the **whole system output** on the far right
3. Fill in all possible combinations for the whole system inputs in the far left columns
4. Fill in the corresponding internal outputs for each combination of inputs
5. Fill in the corresponding whole system outputs

This approach requires a careful step-by-step process. To help speed up the process, you may find it beneficial to write out a statement in words for each logic gate to help you work out the correct output for any given combination.

'If both A and B are on, C is off, otherwise C is on.'

Boolean Algebra

A second method of **logic circuit analysis**, is through the use of boolean algebra. This involves reducing the logic gates into algebraic relationships. You need to know the following notation:

A AND **B** *is represented by* $\mathbf{A \cdot B}$

A OR **B** *is represented by* $\mathbf{A + B}$

NOT **A** *is represented by* $\overline{\mathbf{A}}$

You also need to be aware of **three key rules** that can be applied to the above combinations:

1. Commutative

$A \cdot B = B \cdot A$

$A + B = B + A$

2. Associative

$A \cdot (B \cdot C) = (A \cdot B) \cdot C$

$A + (B + C) = (A + B) + C$

3. Distributive

$A \cdot (B + C) = (A \cdot B) + (A \cdot C)$

$A + (B \cdot C) = (A + B) \cdot (A + C)$

This allows you to form the following simplifications:

$A \cdot A = A$

$A \cdot \overline{A} = 0$

$A + \overline{A} = 1$

$A \cdot 0 = 0$

$A + 0 = A$

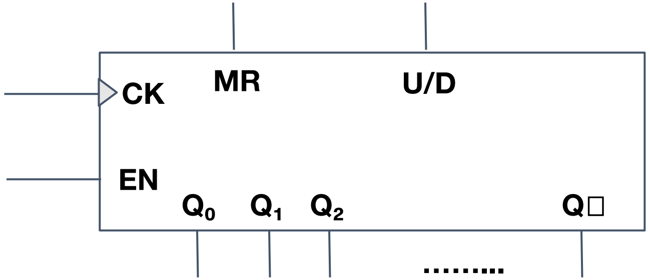
$A + 1 = 1$

$A + A = A$

$A \cdot 1 = A$

Counting Circuits

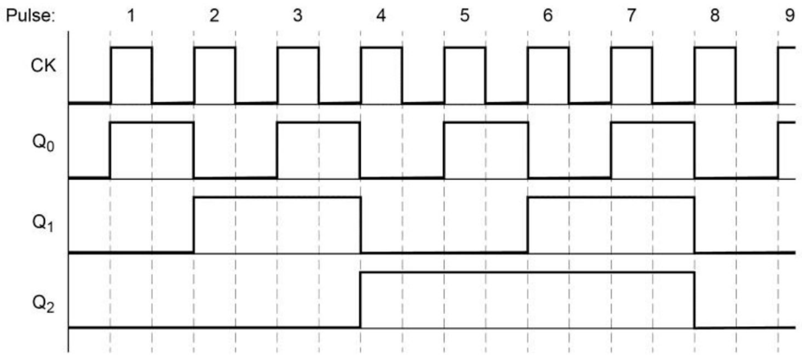
A **static logic circuit** takes inputs and generates a fixed output based on the inputs. A **sequential logic circuit** however can process **time-dependant signals** and follow pre-defined **sequences**. This opens up the possible uses of logic circuits to things such as **counting**, **timing** and **sequencing**. You should understand the different inputs and outputs of a standard counting chip:



- **Q_0 - Q□: Outputs**
- **CK**: **Clock input** - the pulses that are being counted enter through this input
- **EN**: **Enable input**: counter only functions when this input is high
- **MR**: **Master reset** - if a high input enters here, the counter will reset to zero
- **U/D**: **Up/Down input** - the counter will counter upwards when logic 1 and downwards when logic 0

Timing Diagrams

Unlike for static logic circuits, counters operate with time-changing signals and so it is not useful to represent them with logic tables. Instead, timing diagrams are used.



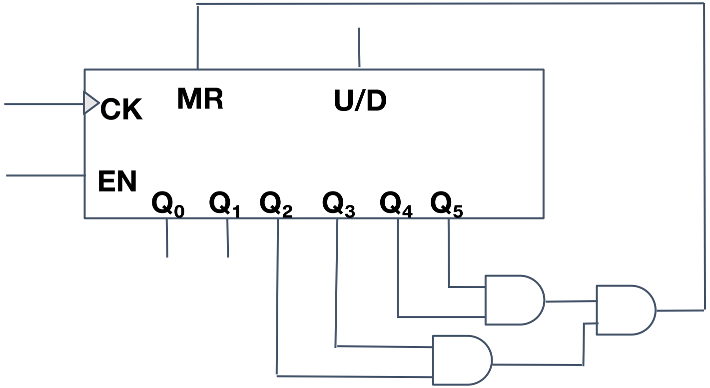
Taken from the AQA A Level Physics, Electronics Teaching Guide

1. **Frequency** of output pulses is always **less** than that of the **clock pulse**
2. Changes to the outputs can only occur when the clock pulse changes from **0 to 1**
3. The counter will restart on the 8th clock pulse

(**Rising-edge triggered** 0 to 1, falling-edge triggered 1 to 0)

The Modulo-n Counter

The **terminal count** of a counter, is the number after which the counter reset. This is usually achieved by connecting the **final output** to the **master reset** port. However, in many systems, the terminal count will be a number that is not a power of 2, and so **logic gates** can be combined to produce a terminal count of any number within the range of the counter. An example is a **minute counter**, where the counter will count up to 59 and then reset on the 60th pulse:

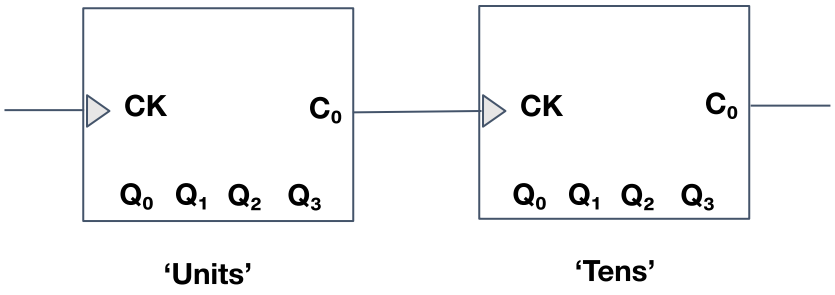


- The EN and U/D are held high
- The CK input receives pulses at a frequency of 1 second
- The counter's 60th pulse results in Q_2, Q_3, Q_4 and Q_5 all being high since the sum of their numerical equivalents is:
  $4 + 8 + 16 + 32 = 60$
- The logic gates shown are AND gates, so only when all 4 outputs are high will the input to MR be high

The BCD Counter

The **Binary-Coded-Decimal** counter is a type of **Modulo-n Counter** that drives **multi-digit decimal displays**. Each digit is a number between 0 and 9 and so each digit must be driven by a **4 bit** binary counter chip. To allow multi-digit displays, BCD counter chips have a C_0 output:

- The C_0 output remains **high** throughout counts 1-8
- On the 9th (and final) count, C_0 will drop to **low**
- When the counter resets (on the 10th pulse), C_0 will return to being high
- This triggers a pulse into the **CK** of the next digit counter



The Johnson Counter

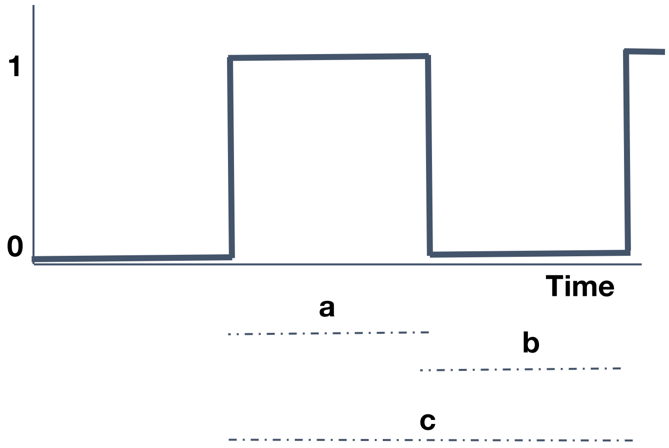
The **Johnson Counter**, also known as a **decade counter**, is a counter that doesn't produce binary outputs. It consists of **10 outputs**, which **sequentially** turn high:

- Only **one** output is high at any given time
- The **Q_0** output represents zero and so when the counter **resets**, it is high
- Each pulse that enters the **CK input** triggers the next output to turn high (and the previous output to return to low)

Johnson counters are useful for devices that perform actions in a **set order**, with each action lasting a set length of time.

Pulse Definitions

When studying pulses, the key information is related to the **length of time** and **frequency** of each pulse. You need to know the following basic and derived definitions of a pulse:



- **a:** On Time or Mark (t_1 ; M)
- **b:** Off Time or Space (t_0 ; S)
- **c:** Period (t_p)

Using these quantities, you can then **calculate** a further set of useful parameters:

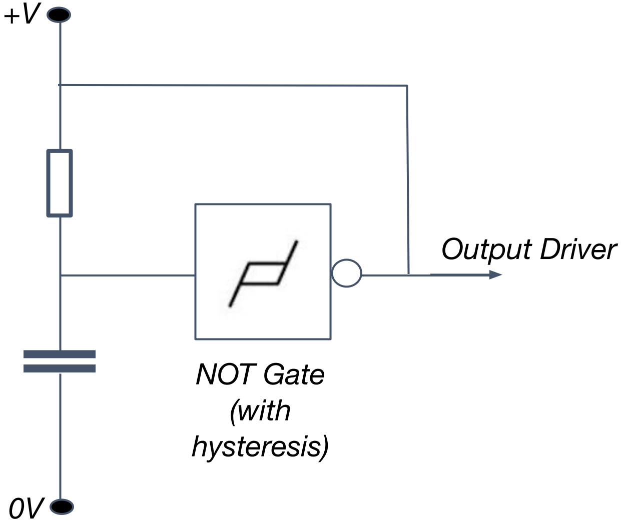
$$\frac{1}{t_p} = \text{pulse repetition frequency (PRF) or pulse rate}$$

$$\frac{M}{S} = \text{mark-to-space ratio}$$

$$\frac{t_1}{t_1 + t_0} \times 100 = \text{duty cycle (\%)}$$

Astables

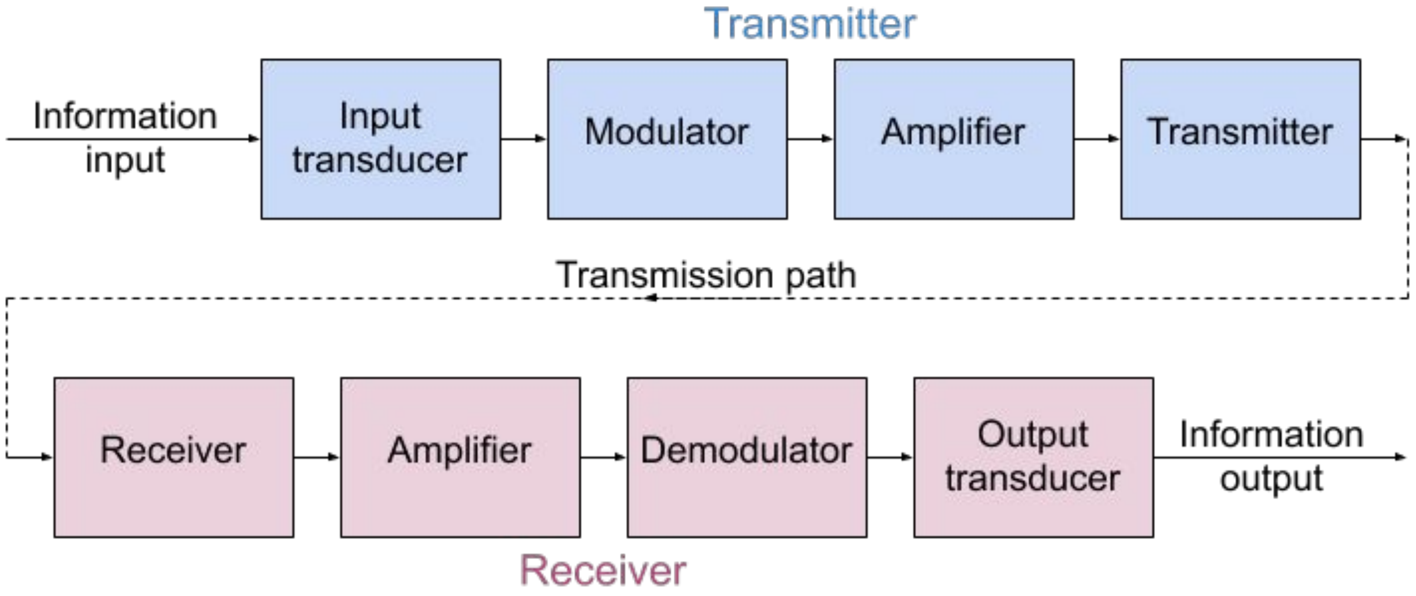
To **generate pulses** used for purposes such as counter chips, a special type of circuit known as an **astable circuit** is used. An astable circuit is one that **doesn't** have a stable state - this means it doesn't rest in an on or off state. These circuits are generally packed into a chip, such as a **555 astable chip**. These contain an **RC circuit**:



1. The **capacitor** is not charged and so the **NOT gate** input is 0, and the output is 1
2. This means the capacitor will begin to **charge** up through the **resistor**
3. When the charge reaches a certain level, the NOT gate input will change to 1, resulting in the output becoming 0
4. This means the capacitor will **discharge** back through the **same resistor**
5. When the charge falls below the threshold value, the process will **repeat**, resulting in a sequence of on and off pulses

Since the charging and discharging occurs through the same resistor, the on and off times are **identical**.

Data Communication System



Data Communication System

You should understand the processes that occur in a data communication system:

1. The **input transducer** converts the information/data into an electrical signal (an example of this would be a microphone)
2. The **modulator** processes the signal so that it is in a suitable form for the transmission path that it is being sent down
3. The **transmitter amplifier** amplifies the signal to a suitable level
4. The **transmitting device** sends the signal into the intended transmission path (an example of this would be an aerial)
5. The **transmission path** is the medium that the signal travels through
6. The **receiving device** receives the signal and converts it back into an electrical form
7. The **receiver amplifier** amplifies the signal to a suitable level for processing
8. The **demodulator** obtains the original signal in a suitable form
9. The **output amplifier** boosts the signal to a level that can drive the desired output transducer
10. The **output transducer** outputs the information in a suitable format

Copper Cables

Copper cables are one of the most common forms of **transmission media**. They are mainly used over **short distances** and for **low data rates**. There are **three types** you need to be aware of, the first being a **coaxial cable**:

- Used for radio, television and other **high frequency signals** over a short distance
- Composed of an inner **copper core** that is insulated with a **dielectric** and then wrapped in a **braided copper screen** and an **outer insulating jacket**
- The purpose of the copper screen is to provide **electrostatic shielding**, reducing the likelihood of **external interference**
- Has an upper frequency limit of around 1 GHz
- At its upper frequency it has a high level of signal attenuation
- Coaxial cables aren’t a very secure transmission medium

Copper Cables

The second type of commonly used cable is a **twisted-pair cable**:

- Used for **computer networking**
- Consists of **several pairs** of conductors that have been twisted together and then surrounded by an outer insulating jackets
- Each pair has **good noise immunity**
- Has an upper frequency of around **100-250 MHz**
- At its upper frequency, the signal strength halves roughly every **12.5 metres**
- **Cheap**, making it good for commercial use
- **More secure** than a coaxial cable, but still not a massively secure transmission medium

The final type of copper cable you should know about is a **plain copper wire**:

- Consists of single strands of copper wire
- Very **low noise immunity**
- **Radiates energy** and so can’t be used at high frequencies
- Very **insecure**
- Very **low cost**

Optical Fibres

Optical fibres are a transmission medium used for non free-space communication links due to their large bandwidth. The basic structure of an optical fibre is:

- A **central glass fibre** core through which the signal is passed
- **Cladding** to allow TIR to occur and provide tensile strength
- **Protective sheath** to protect the fibre from scratches and other damage

Advantages of using optical fibres as a transmission medium include:

- Each fibre is **very small** meaning **hundreds** of **independent** fibres can be bundled in the same cable
- Can handle **very high signal frequencies**
- Data rates of up to **270 Gbps** can be achieved by a single fibre
- They are **very secure**

Optical Fibre Modes

There are two main types of optical fibre. The first is a **multi-mode** optical fibre:

- Uses **total internal reflection** to keep the signal within the core and to transmit it from one end to the other
- The core has a **high refractive index** whereas the cladding has a **low refractive index**
- **Attenuation** occurs over long distances, but to a much lesser extent than copper cables
- **Pulse broadening** can occur due to different rays entering at slightly different angles and so taking slightly different length paths through the core

The second type of optical fibre is a **mono-mode fibre**:

- The core diameter is made as **narrow** as possible so that the number of different paths that can be taken is significantly minimised
- The diameter is of the order of **8μm**, which is comparable to the **wavelength** of light
- The light will **propagate** the signal in ways other than through TIR
- Mono-mode fibres experience much **less attenuation** than multi-mode fibres

EM Waves in Free Space

Electromagnetic waves can be used to transfer a signal between a transmitter and receiver when there is **no connection** between them. You should be familiar with the **advantages and disadvantages** of this method:

- **Costs** associated with wireless transmission can **vary** massively, with some applications being fairly inexpensive and others, such as those requiring **satellites**, being very costly
- Very **easily intercepted** so not a **secure** method of transmission
- Data transmission must be **greater** than around **60 kHz** since for an aerial to be effective its magnitude must be around half the wavelength - at small frequencies, the wavelength is too large to be practical
- Transmission frequencies can't usually exceed **100 GHz** since receivers for very high frequencies are difficult and costly to produce

Longwave Signals

The first type of electromagnetic signals you should be familiar with are **long wave signals:**

- Include wavelengths of around **1 km - 2 km**
- Frequencies are of the order of **150-300 kHz**
- BBC **radio broadcast** used long wave signals for a long period of time and radio 4 still make use of this type of signal
- The National Physical Laboratory 60 kHz MSF signal is in the longwave range - it is this signal that provides the time for many **radio controlled clocks** and watches
- They can travel **very long distances** since they are more **easily diffracted** around objects
- The r**efractive indices** of the atmosphere and Earth's surface mean they can propagate as **ground waves**
- Means that transmitter and receiver don't need to be in **line of sight** of each other

Shortwave Signals

The second type of electromagnetic signals you should be familiar with are **short wave signals:**

- Include wavelengths of around **10 m - 100 m**
- Frequencies are of the order of **3-30 MHz**
- Used to broadcast **consumer programmes**
- Used for long-distance communications to **ships** and **aircraft**
- Like long wave signals they can travel **very long distances**, but for different reasons
- Shortwave signals form **sky waves**
- The **ionosphere** reflects waves with frequencies above around 500 kHz, meaning the short waves are **reflected** back towards the Earth's surface, allowing them to travel vast distances
- The reflection also means that once again the transmitter and receiver don't need to be in direct **line of sight** of each other

Microwaves and Telemetry

The third type of electromagnetic signals you should be familiar with are **microwave signals:**

- Include wavelengths of around **3mm - 150mm**
- Frequencies are of the order of **2-100 GHz**
- Used for **3G** mobile data and **Bluetooth** communication
- Used by **satellites**
- Require **line of sight** between the transmitter and receiver
- Aerials are more **complex** due to the higher frequencies

The final form of EM signals you should be aware of, are **telemetry signals**:

- Narrow frequency range of **433-866MHz**
- Controls and monitors data capture from **remote equipment**
- Used for purposes such as remote controlled **garage doors** and **flood warning systems**

Satellite Systems

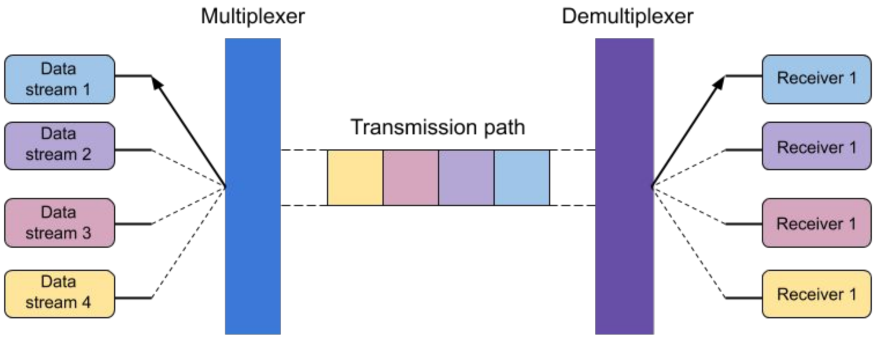
A key form of communication system is the use of **artificial satellites**. The satellite is launched into **space** and is responsible for receiving signals from Earth and then re-transmitting them. You should know about the following features:

- **Low Earth orbits** (LEOs) have an altitude of between **160km** and **1600 km** above the Earth's surface
- **Geostationary orbits** have an altitude of **35,786 km** and have a time period of 24 hours
- Non-geostationary orbits are **cheaper** due to their lower altitudes, but are only visible for part of the Earth's 24 hour period, and so have limited uses and often have to be used in pairs
- The transmission path from the Earth is known as the **up-link,** and the return path is known as the **down-link**
- Down-link signals require large amounts of **amplification** due to the long distances involved
- The up-link and down-link frequencies must be different to avoid the amplified down-link signals overwhelming the lower power up-link signals
- Satellites usually use waves in the microwave range of the **EM spectrum**

Multiplexing

The **transmission rates** of data can be improved if **multiple data streams** can use the same transmission path. This process is known as **multiplexing** and there are two types of achieving it: **time-division** multiplexing and **frequency-division** multiplexing. You only need to be familiar with **time-division multiplexing**:

- Data is split into packets, each with the identifying information, **error correction** and **synchronization** bits
- The multiplexer connects each stream in turn
- The **demultiplexer** is always in sync and connects each data stream to the correct receiver
- It is a **cheap** way of increasing data traffic rates, and doesn't involve any increase in **bandwidth**
- The synchronization data is required for the data to be properly pieced back together



Amplitude Modulation

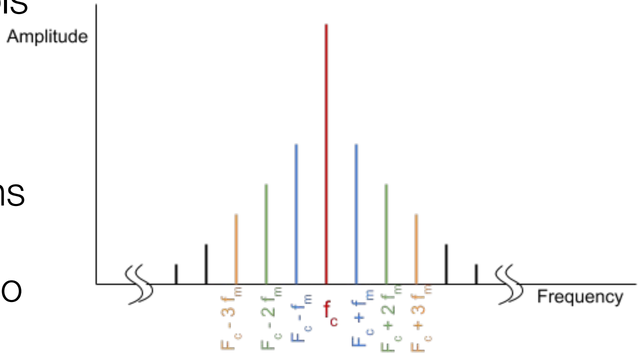
Modulation is essential when signals are transmitted as EM signals through free space. A lot of signals have a frequency that is **too low** for practical purposes, and so the process of modulation involves **translating** the information onto a wave with a much **higher frequency**.The first type of modulation you should be aware of is **amplitude modulation**:

- In amplitude modulation, the signal **amplitude** controls the amplitude of a **high frequency carrier wave**
- The modulated signal has a much higher frequency, but the amplitude at each point matches the amplitude of the original signal
- The modulated signal has a **main frequency** (fc) , alongside two additional frequencies (fU and fL)
- The additional frequencies are referred to as **side frequencies**, and the upper is calculated from the **sum** of the carrier frequency and the information signal frequency, whereas the the lower is calculated from the **difference**

Frequency Modulation

The second type of signal modulation you need to know about is **frequency modulation** (FM):

- In frequency modulation, the **amplitude** of the signal controls the **frequency** of the **carrier wave**
    - The process of FM is much more complex than for AM
        - FM often requires **more bandwidth** over AM
- The amplitude of a frequency modulated wave signal remains **constant**
- Unlike in AM, for FM the side frequencies extend outwards to **infinity** either side of the main frequency



- The side frequencies are produced by adding **integer multiples** of the signal frequency to the carrier frequency
    - The exact shape of the spectrum depends on the **frequency deviation** (Δf)
- If Δf is very small, it is **narrow band FM** and the spectrum will look much like an AM signal
    - Most FM signals are **wideband** and so their sidebands extend to +/-∞

AM and FM Comparison

You should be able to explain differences between AM and FM:

- Most types of **signal noise** affect signal amplitude rather than frequency, which means that FM systems are much less susceptible to unwanted noise
- FM receivers can remove noise through a process known as **limiting** - a limiter circuit ensures that the amplitude of an FM signal is always constant
- FM is less susceptible to **fading** which is where reflected and refracted waves arrive at the same receiver and interfere - interference affects amplitude and so is less of a problem form FM signals
- FM signals however do require **line-of-sight** between the transmitter and receiver due to their high frequencies
- AM is a much **simpler** process than FM

Pulse Code Modulation

Ordinarily when an **analogue signal** is **sampled** and converted into a digital **binary** signal, the n individual **bits** of digital data would require n separate **transmission paths** to be transmitted. For an electromagnetic signal for example, this would require n different **carriers**. To solve this problem, pulse code modulation is used:

- A **sample and hold circuit** samples an analogue signal at fixed intervals and the voltage is held until each **conversion** is completed
- The **parallel data signals** are converted by a **parallel to serial converter**, to produce a serial form of the data that is passed into the transmission path
- At the receiving end, the process is carried out again but in **reverse**, to convert the signal back into **analogue form**
- PCM is used in applications such as for **digital audio** and CDs