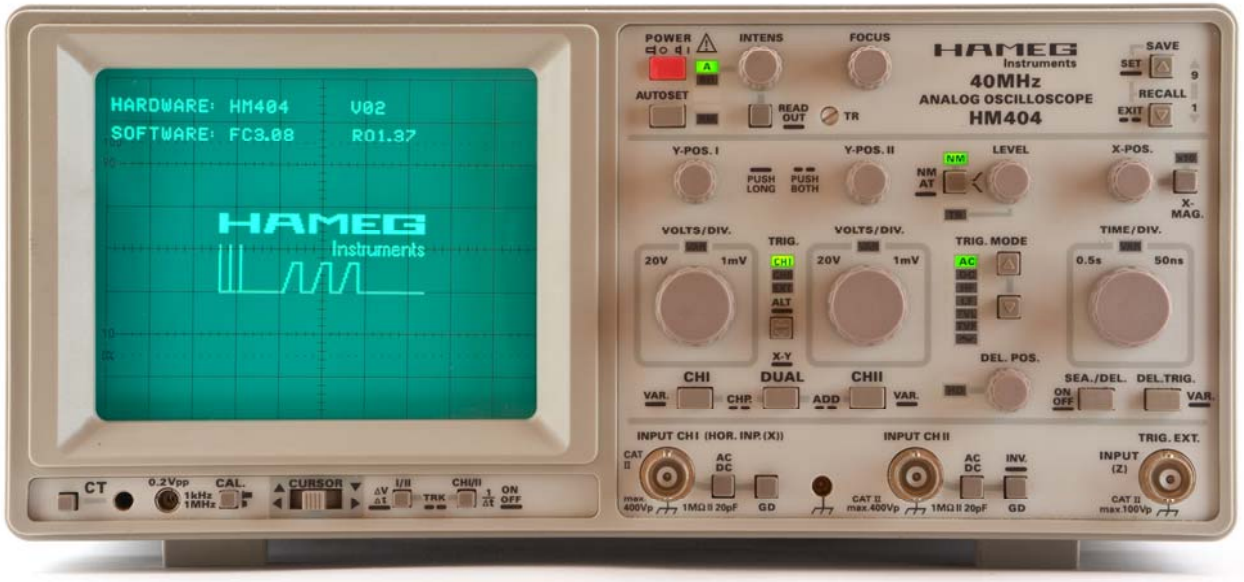
Module 4 B2
Electronic Fundamentals



AERO-Bildung

Aviation Training Center

1 SEMI-CONDUCTORS

1.1.1 Introduction to Semi-conductors

A semi-conductor is a class of crystalline solids. Its energy state is indeterminate between that of an insulator and a conductor.
A pure semi-conductor is known as an 'Intrinsic semi-conductor'. For example, 99.99% pure silicon is known as "High purity" silicon.
An intrinsic semi-conductor is a poor conductor (good insulator) and is very dependant on temperature. The resistance of a semi-conductor decreases with an increase in temperature. This is known as a "negative temperature co-efficient".

Examples of intrinsic semi-conductors are Silicon, Germanium and Gallium Arsenide.

The main semi-conductor in commercial use is Silicon. Silicon is a base material in sand; hence it is in abundant supply.

Electrons and Current Flow

An electron is a physical particle with a negative charge.

When a battery is placed across a conductor the electrons will be attracted to the positive terminal of that battery. Figure 1 illustrates the actual current flow when a power supply is connected across a conducting circuit.

Electron current flow is 'Negative-to Positive'.

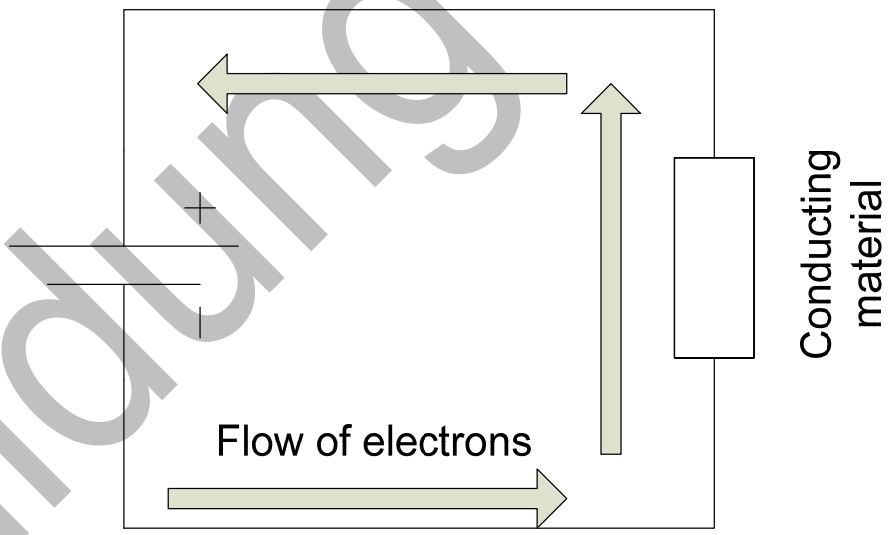


Figure 1 Actual Current Flow in an Electrical Circuit

A hole is not a physical particle. It is simply the space (missing bond) remaining after an electron vacates the outer ring of an atom during current flow.

However, a hole can be treated as a physical particle because it has a positive charge and moves in exactly the opposite direction to the flow of electrons.

When a battery is placed across a conductor the holes will be attracted to the negative terminal of that battery. This is known as conventional current flow. Figure 2 illustrates the conventional current flow when a power supply is connected across a conducting circuit.

Conventional current flow is 'Positive to Negative'

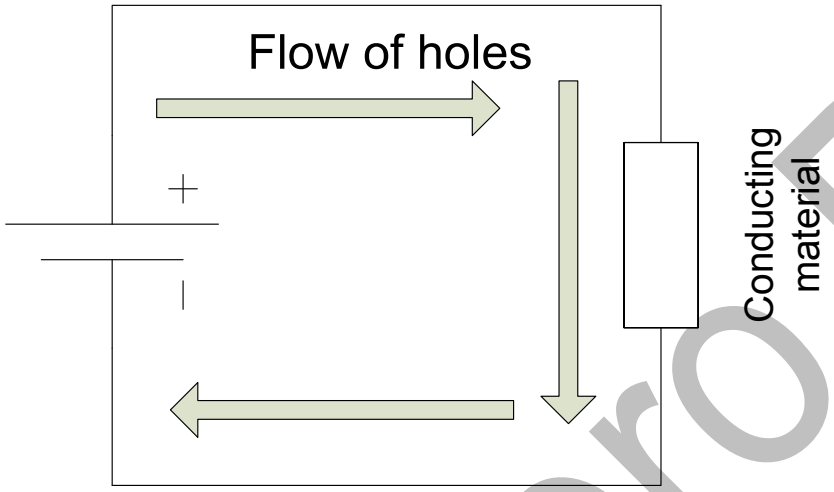


Figure 2 Conventional Current Flow in an Electrical Circuit

Atomic Structure

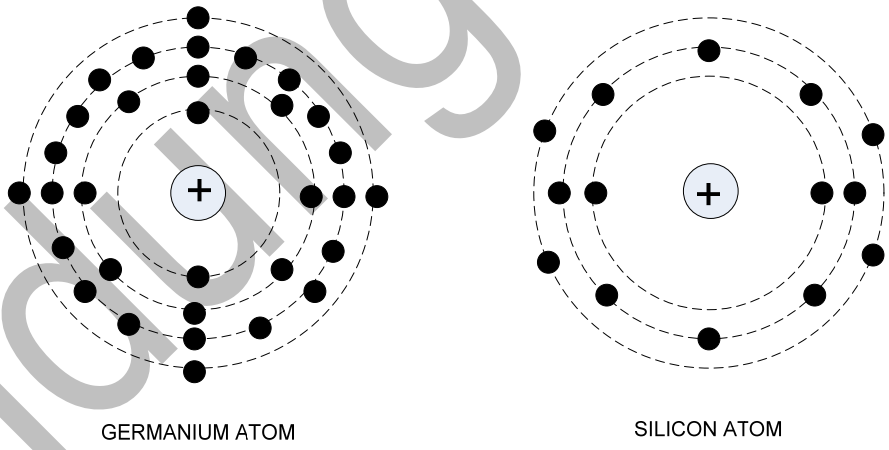


Figure 3 Atomic Structure of Germanium and Silicon

The atomic structures of the two most commonly used semi-conductors are illustrated in figure 3. It should be noted that orbiting electrons are not arranged in an orderly manner. The above is a 2-dimensional representation of particles moving on a 3-dimensional plane. The electrons do not rotate in perfect circles or on a flat plane.

The above atoms are known as "Group 4" materials. This is because they have four electrons on their outer shell. These are known as "Valence Electrons".

Valence Electrons

These electrons are farthest from the nucleus of the atom. They are less tightly bound by the attractive force of the nucleus. These outermost electrons have higher energy than those within the atom; this makes it easier for them to move between atoms.

The movement of valence electrons between atoms is the main factor in electrical conduction.

The movement of valence electrons between atoms is shown in figure 4. It should be noted that as an electron is attracted away from the atom, it leaves a hole. That hole subsequently attracts another electron, leaving another hole. For illustrative purposes, the diagram shows the valence electrons moving left to right. It can be seen that holes move from right to left (as indicated by the arrows).

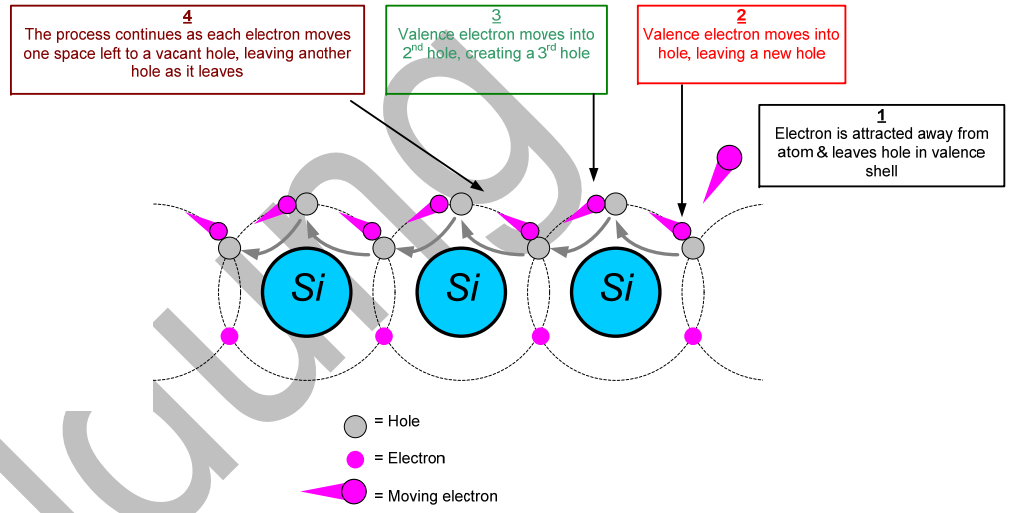


Figure 4 Movement of Valence Electrons between Atoms

Covalent Bonds

The valence electrons of a semi-conductor atom can link together with the valence electrons of other adjacent atoms. These links are known as “Covalent bonds”.

As a result of covalent bonding, each atom has a half share in eight valence electrons. This linking arrangement of multiple atoms forms a stable crystal lattice.

The two-dimensional arrangements of covalent bonds in a silicon crystal lattice are shown in figure 5.

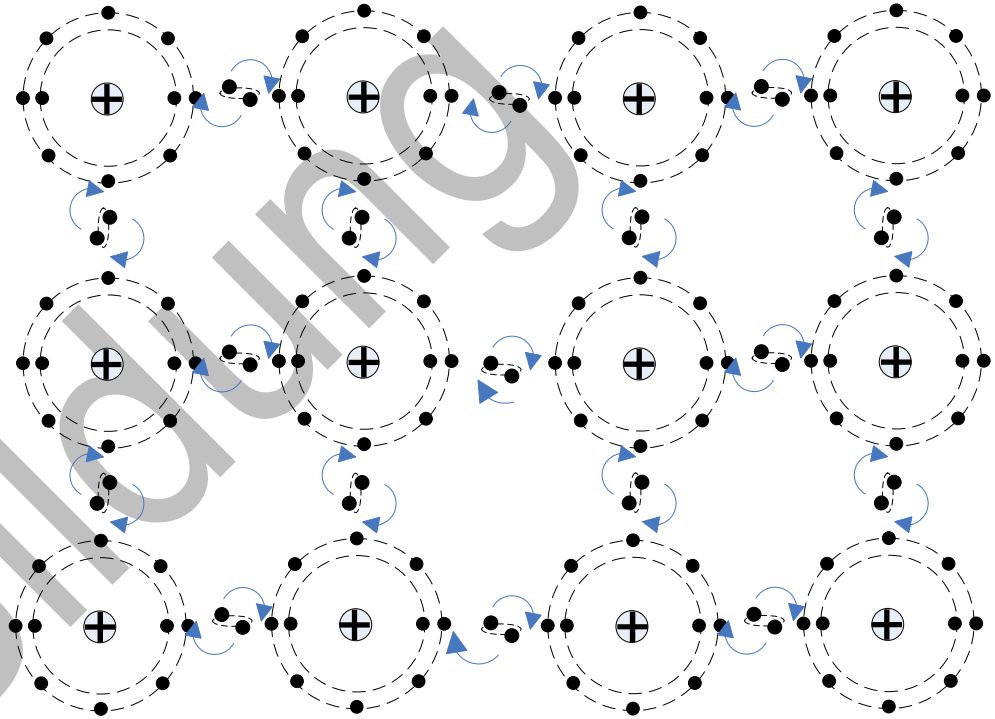


Figure 5 Co-valent Bonding of a Silicon Crystal Lattice

1.1.2 Semi-conductor Doping

Semi-conductor doping is a chemical treatment performed on an intrinsic semi-conductor to either increase or decrease conductivity.

Doping improves the ability to transmit and control electrical current through the semi-conductor.

Processes also exist to further purify an intrinsic semi-conductor prior to doping. This is often necessary in order to effectively control/measure the doping process.

Doped semi-conductors are often referred to as “Extrinsic Semi-conductors”.

Doping involves adding a small volume of impurity atoms (usually only about $1*10^6$) to the semi-conductor material.

As noted earlier, the intrinsic semi-conductor has 4 valence electrons. The impurity element is selected to have atoms with an extra electron or a missing electron (in relation to the semi-conductor atom).

The addition of an impurity atom is added with the intention of altering the number of electrons in a semi-conductor.

There are two types of extrinsic semi-conductor:

- N-Type Semi-conductor.
- P-Type Semi-conductor.

The N-Type Semi-conductor

Group 5 materials are used to provide an extra electron after doping with the intrinsic semi-conductor. These elements have 5 valence electrons. This means that one electron is free after the covalent bonding occurs. This electron is referred to as a “Donor Electron”.

These donated electrons have the effect of lowering the semi-conductor’s resistance, therefore making it more conductive. These electrons are un-bonded and free to wander the crystal lattice. The donor electrons will migrate through the atomic structure and can therefore act as current carriers when a low voltage is applied. Figure 6 shows the atomic arrangement after doping with the group 5 material.

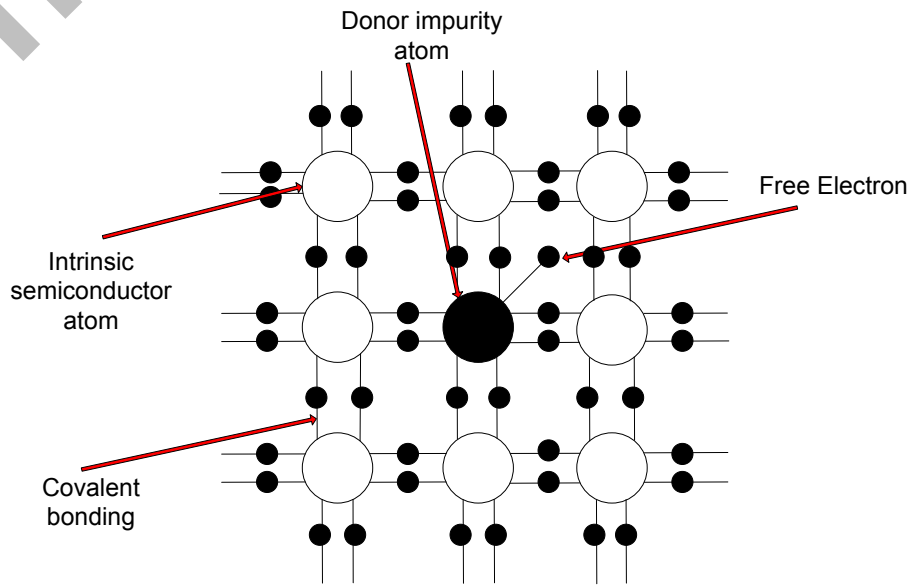


Figure 6 Atomic Arrangement after Doping with Group 5 Material

The number of free electrons can be strictly controlled by the doping process. Figure 7 shows the current flow in an N-type semi-conductor

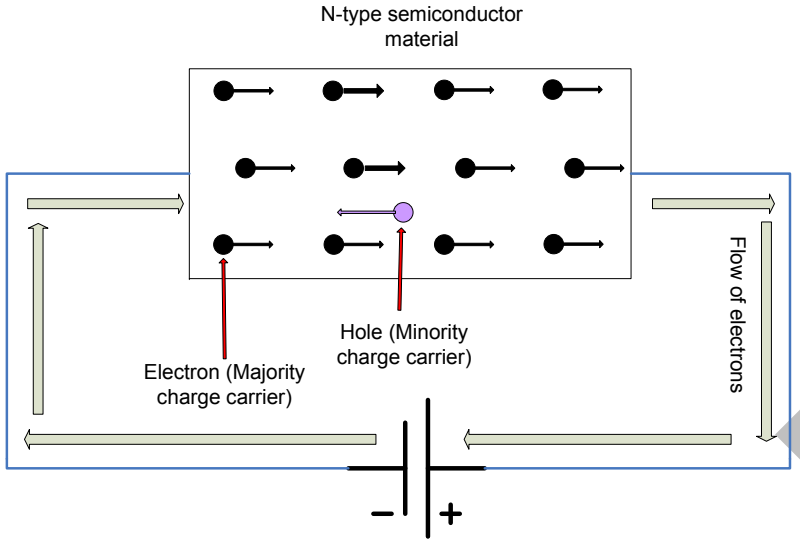


Figure 7Current Flow in an N-Type Semi-conductor

The electron is the majority carrier in an N-type semi-conductor. The term 'N-Type' is used because the electron is negatively charged.
Due to intrinsic leakage conduction, the hole (positively charged) acts as a minority carrier in the opposite direction. It should be noted that this current is negligible in comparison to the main current flow.

Phosphorus and Arsenic are good examples of group 5 materials used in this doping process.

The P-Type Semi-conductor

Group 3 materials are used to provide an extra hole after doping with the intrinsic semi-conductor. These elements have 3 valence electrons. This creates an extra hole for a valence electron already within the intrinsic material. This hole is commonly referred to as an "Acceptor".

When a power supply is applied, electrons will be attracted to the positively charged holes. The flow of electrons to these holes creates more holes elsewhere in the material. The result is a flow of charge from the positive terminal to the negative terminal. Figure 8 shows the atomic arrangement of the semi-conductor after it has been doped with the group 3 materials.

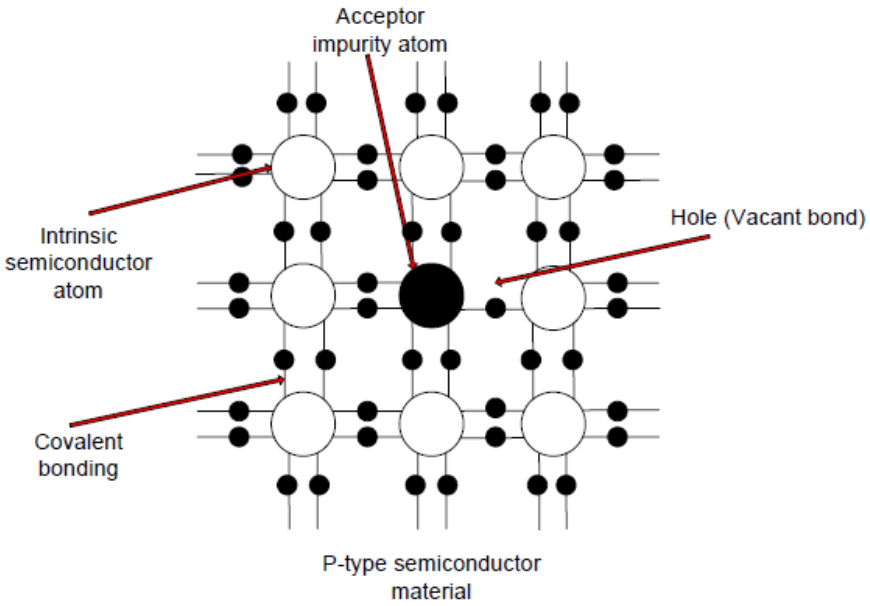


Figure 8 Atomic Arrangement after Doping with a Group 3 Material

The hole is the majority carrier in this doped semi-conductor. The term “P-type” is used because the material is positively charged.
Figure 9 shows the direction of current flow in the P-type semi-conductor.

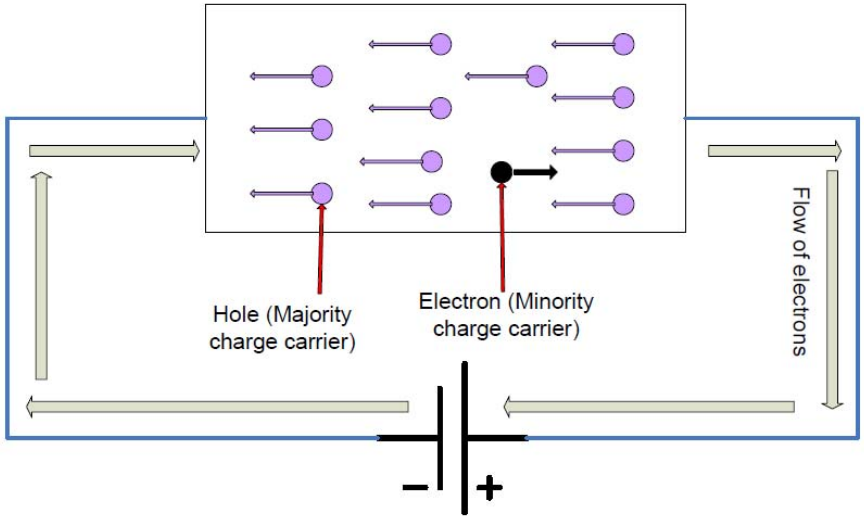


Figure 9 Current Flow in a P-Type Semi-conductor

Due to intrinsic leakage, the electron (negatively charged) acts as a minority carrier in the opposite direction.

Indium, Aluminium and Boron are good examples of group 3 materials used in this doping process.

1.1.3 Semi-conductor Materials Tutorial

All answers can be taken directly from the course text.

1. State the meaning of:
   a) Semi-conductor.
   b) An intrinsic semi-conductor.
   c) Negative temperature co-efficient.

2. Give 2 examples of an intrinsic semi-conductor.

3. State the semi-conductor that is used in most electronic production and give the main reason for its use.

4.
   a) State the properties of an electron and state the direction of actual current flow.
   b) State the properties of a “Hole” and the direction of conventional current flow
   c) Which type of current flow is most generally used when understanding and processing electrical circuits?

5. State the meaning of a group 4 material.

6. State the meaning of a Valence electron. How are they different from other electrons within the semi-conductor atom?

7. What is meant by ‘Co-valent’ bonding”?

8.
   a) What is meant by “Semi-conductor doping”?
   b) Why is an intrinsic semi-conductor purified further prior to the doping process?
   c) What does the doping process involve?
   d) What is an Extrinsic Semi-conductor*?*
   e) Describe how the impurity atoms are selected for semi-conductor doping.
   f) What are the two types of extrinsic semi-conductor?

9.
   a) Describe the properties of a semi-conductor doped with a Group 5 material.
   b) Using a diagram, illustrate the flow of current in a semi-conductor doped with a group 5 material.
   c) State the Majority and Minority carriers.
   d) Give 2 examples of group 5 materials used in this doping process.

10.
   a) Describe the properties of a semi-conductor doped with a Group 3 material
   b) Using a diagram, illustrate the flow of current in a semi-conductor doped with a Group 3 material.
   c) State the Majority and Minority carriers.
   d) Give 2 examples of Group 3 materials used in this process.

1.2 DIODES

1.2.1 The P-N Junction

A typical junction diode structure is formed using P-type and N-type materials. The P and N-type materials are in contact to form a P-N junction. Under the correct conditions, the holes and electrons are allowed to join (electron-hole combinations), allowing a current to flow across the junction.

As covered in the previous section, the P-type material consists of Holes as the Majority carrier. This means the P-type material is positively charged and will attract any free electrons.

The N-type material consists of electrons as the majority carrier. This means the N-type material is negatively charged and electrons will be attracted to any holes.

The P-N Junction without Applied Power

Under static conditions (without applied power), the P-type material has an excess of holes and a shortage of electrons (all the free electrons have filled some vacant holes, which the material still has in abundance). Inversely, the N-type material has an abundance of electrons and a shortage of holes (all the holes have been filled, leaving a significant number of free electrons). Figure 10 shows the P-N junction in the static state.

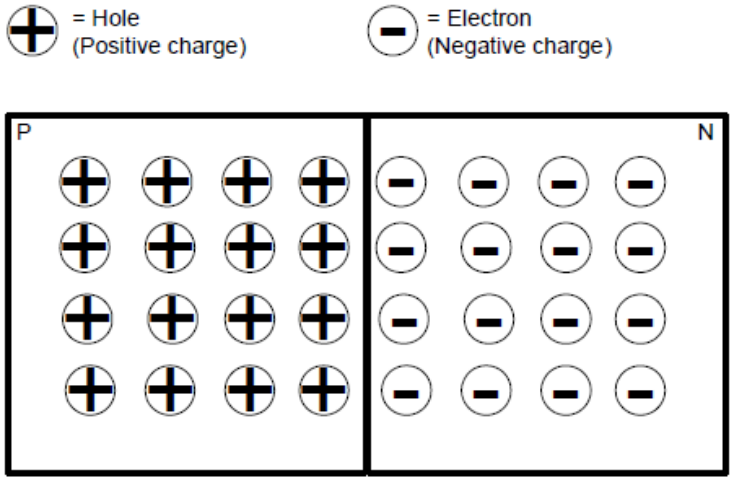


Figure 10 P-N Junction in a Static State

When the N-type material loses an electron (creates a hole), the charge becomes positive. When this occurs for many electrons, the area of the N-type material begins to reject holes (attract electrons).

When the P-type material holes become occupied with an electron, the charge becomes negative. When this occurs for many holes, the area of P-type material begins to reject electrons (attract holes).

This creates a depletion region of current carriers due to the lack of electron-hole combinations. In-turn a further lack of electron-hole combinations occurs. This arrangement of the P-N junction is shown in figure 11.

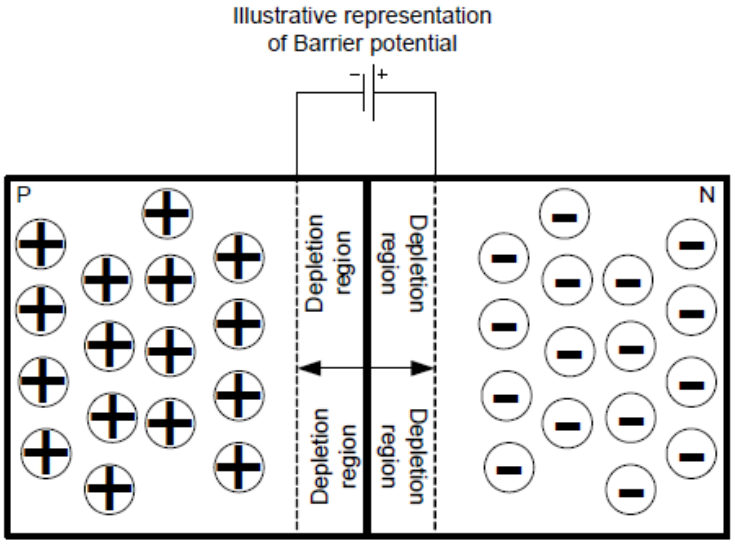


Figure 11 Depletion Region in a P-N Junction

The depletion region is also known as a barrier. Characteristically, the majority current carriers are repelled from the junction contact causing low current across the junction. The flow of low current is caused by the minority carriers in each material.

The P-N Junction with Forward Bias

An electro-magnetic force (EMF) can be applied to the junction. This generally occurs in the form of an applied voltage signal, with the positive polarity connected to the P-type material.

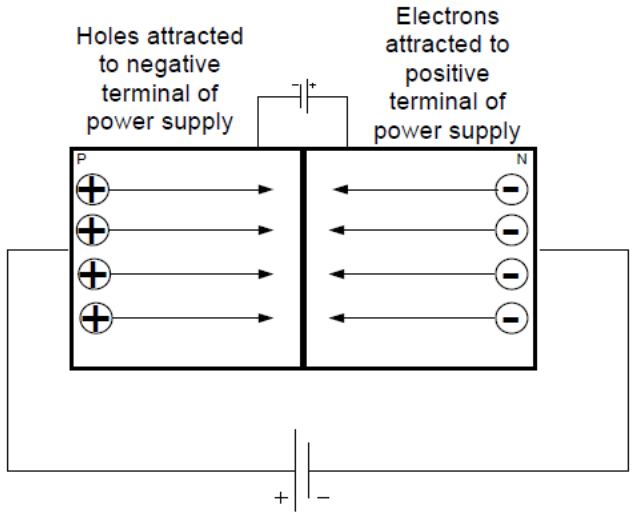


Figure 12 Effects of a Voltage applied to the P-Type Region of a P-N Junction

Figure 12 shows the effects of a voltage applied to the P-type material. If the applied EMF is great enough to overcome the barrier, electrons (from the N-type material) and holes (from the P-type material) are attracted to the barrier. These majority current carriers combine and cause a flow of (conventional) current into the N-type region. The EMF injects fresh carriers into the junction. The current is sustained by majority carriers already in the junction. This means the current flow is large.

This state of current flow (applied to the P-type region) is known to be forward bias because the conventional current flow is positive (P-type, majority carriers are holes) to negative (N-type, majority carriers are electrons).

The barrier potential for germanium is 0.2V - 0.3V. The barrier potential for silicon is 0.6V - 0.7V.

The P-N Junction with Reverse Bias

If the external EMF is reversed in polarity (the positive terminal connected to the N-type region, the negative terminal connected to the P-type region) there will be no current flow (except that due to intrinsic conduction of minority carriers).

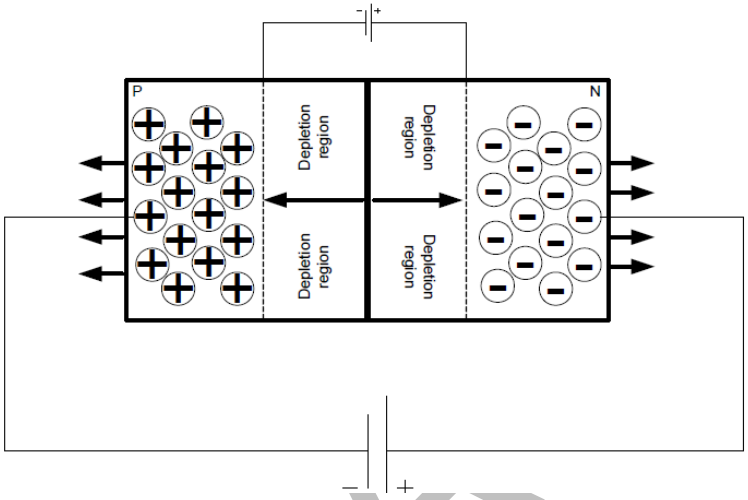


Figure 13 Effects of a Voltage applied to the N-Type Region of a P-N Junction

Figure 13 shows the effect of reverse bias on a P-N diode. The holes in the P-type region will be attracted to the electrons in the negative terminal of the power supply. The electrons in the N-type region will be attracted to the positive terminal of the power supply. There can be no electron-hole combinations at the junction causing the depletion region to increase (widen).

The application of an excessive reverse bias potential to a P-N junction will cause catastrophic destruction to the diode.

1.2.2 Rectivier/ Normal Diode

The Forward Bias Diode

The circuit diagram of a standard forward bias diode is shown in figure 14.
It is important to note that the input to the P-type region on a diode is called the “Anode”. The output of the diode from the N-type region is known as the “Cathode”. The physical layout of a diode is also shown on figure 14.

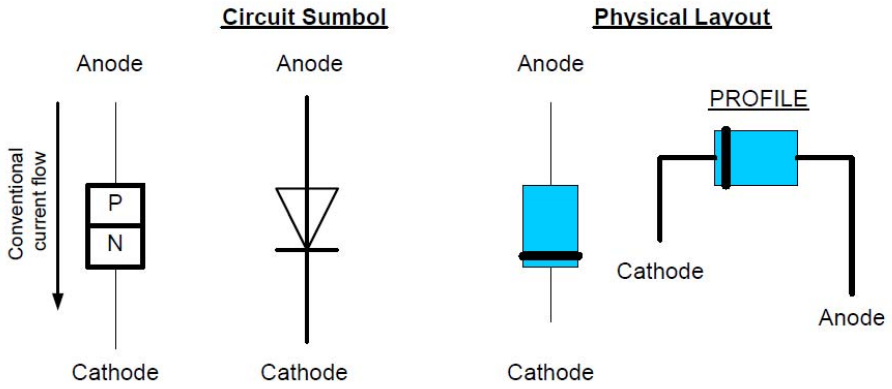


Figure 14 The Standard P-N Diode

The Cathode is identified by a black band on the main body of the diode.

Diode Uses

The most important point to note about a diode is that under the correct conditions it allows current to flow in one direction only (Anode to cathode or P to N).

The main use of the standard P-N diode is to convert alternating current (AC) to direct current (DC). This action is known as rectification and is discussed later.

Another use of the standard P-N diode is to block current flow in the reverse direction (Cathode-Anode). This is mainly used on digital systems to prevent current flow into an integrated circuit.

Diode Characteristics

The diode characteristic curve for silicon is shown in figure 15.

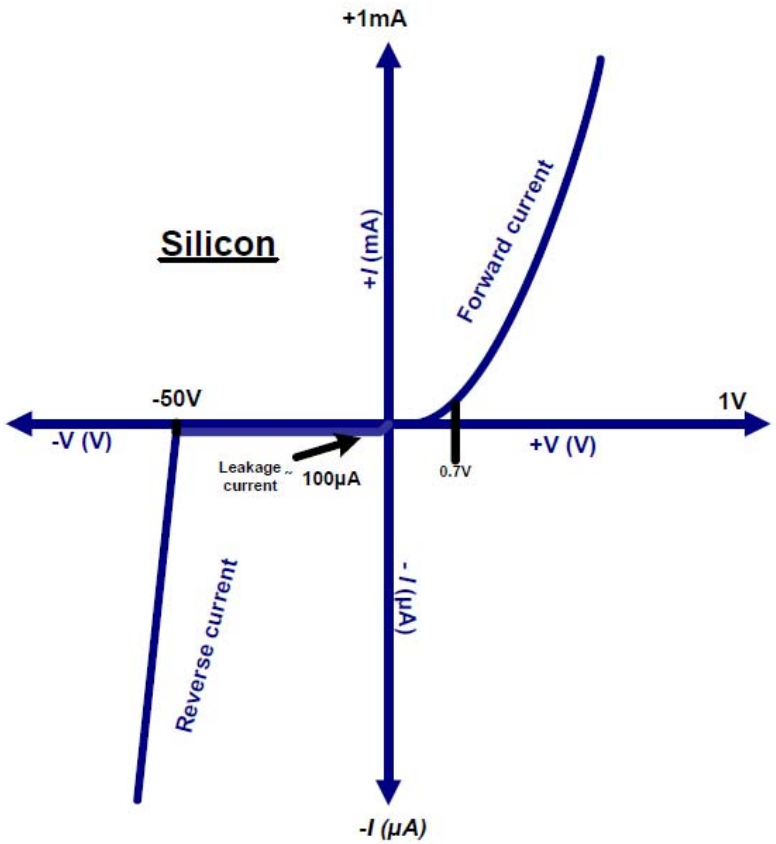


Figure 15 Diode Characteristic Curve of a Silicon Diode

The diode forward bias switch-on voltages are stated below:

- Silicon = 0.7V
- Germanium = 0.3V

When operating in forward bias, the voltage must overcome the barrier before the current can increase. Once the barrier is overcome the current flow increases exponentially with an increase in voltage. From the graphs, it can be noted that the measurements are in milliamps.

During reverse bias operation, the electrons and holes move away from the junction as the depletion region increases. A negligible current is permitted to flow due to leakage from minority carriers (usually no more than 100µA).

Avalanche Breakdown

As the reverse voltage slowly increases, the negative current flow remains relatively stable. Once the reverse voltage limit is exceeded, the leakage current rises to a magnitude that can be limited only by resistance.

The maximum inverse voltage (given on the specifications of practical diodes) dictates the maximum reverse voltage that can be applied to the diode before avalanche breakdown will occur.

Avalanche breakdown occurs when an excessive external reverse bias voltage is applied to the diode.

Assume that a single conduction electron gains sufficient energy from the applied power source to accelerate it towards the positively charged end of the P-N junction. As it is propelled across the junction it collides with an atom. This releases enough energy to dislodge another valence electron for conduction.
There are now two free high-energy electrons. Assume each of these electrons collides with another atom. This has now freed another two valence electrons from their bonds. This continues with each collision, effectively doubling the number of electrons each time.
This results in a rapid increase of reverse current as the number of free electrons continues to multiply. This is known as the **avalanche effect**.

Other Characteristics of a Standard Diode

There are two other main factors that affect the operating conditions of a diode:

- Maximum forward current.
- Maximum operating temperature.

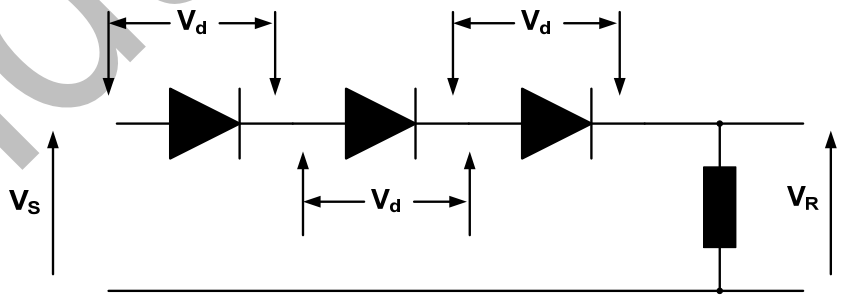
The maximum forward current should be limited (using an appropriate resistor) to prevent the junction from heating up during normal operation. When a voltage is applied to the diode and a current is allowed to flow, power is dissipated across the circuit. This power dissipation will cause the diode to heat up and produce an increase in leakage current as the junction atoms collide.

The maximum operating temperature must be observed for the same reason.

Diodes in Series

When diodes are connected in series the current is the same through each diode. It should be noted that there will be a drop across each diode equal to the bias voltage (0.7V Silicon). The maximum forward current of the diode should not be exceeded.

An example of a set of diodes in series is shown in figure 16.



$$V_R = V_S - 3V_d$$

Figure 16 Diodes in Series

Kirchoff's voltage law applies where the total drop across the diodes subtracts from the supply to leave the remaining voltage to power the load.

1.2.3 Half Wave Rectification

A Half-wave rectifier is a circuit that converts an alternating sine wave into a pulsating DC signal that consists of one-half of a sine wave for each input cycle.

The circuit in figure 17 shows the circuit of a half-wave rectifier.

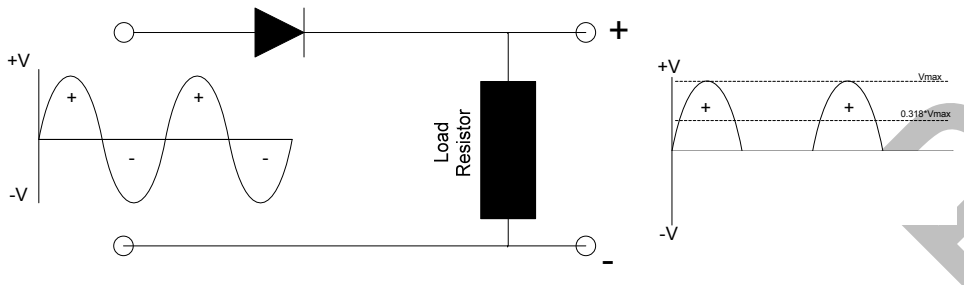


Figure 17 Half Wave Rectification

An AC supply can be connected to the diode. When the supply is positive (with respect to the “P” region), the circuit is forward biased. This is referred to as the positive half-cycle. When the supply goes negative for the remainder of the cycle (the negative half-cycle), the current is blocked due the reverse bias on the junction. Therefore, the junction will only pass current on the positive half of the input cycle.

The formula to calculate the average Voltage of a half-wave rectified load is shown in equation 1a.

$$V_{AV} = \frac{V_{\max}}{\pi}$$

Equation 1a: Average Voltage of a Half-Wave Rectified Load

The improved (and quicker) method of obtaining the average voltage is shown in equation 1b.

$$V_{AV} = 0.318 \times V_{\max}$$

Equation 1b: Quick Calculation

The output from the half-wave rectifier will produce a half-sine for every cycle at the input. The output is always positive (though still variable) and is now DC.

The frequency is still equal to that of the supply frequency because the time period of the half-cycles remains the same as the time period of the input sinusoidal signal.

The diode barrier potential also has an effect on the output of the half-wave rectifier. The description above demonstrates an ideal diode; therefore it did not discuss the effects of the barrier voltage.

When the diode is operating under forward bias, the input signal must overcome the diode barrier voltage (0.7V for Silicon) before it will conduct. Therefore the output voltage will always be 0.7V less than the input voltage. This is expressed in equation 1c.

$$V_{OUT} = V_{IN} - 0.7$$

Equation 1c: Effect of Diode Barrier Potential on the Half-Wave Rectifier

The Peak-inverse voltage (PIV) also has an effect on the operation of the half-wave rectifier. The negative half cycle of the input will apply a reverse bias to the diode. The highest value of reverse bias will be at the peak of each negative alternation. The diode must be capable of withstanding this repetitive value of reverse bias.

1.2.4 Full Wave Rectification

A Full-wave rectifier is a circuit that converts an alternating sine wave into a pulsating DC signal that consists of both halves of a sine wave for each input cycle.

The positive half-cycle of the sine wave is applied at point A (figure 18). The current is forced to the load resistor through the only forward bias diode available at its position. The arrows show the path of the current on the positive half cycle.

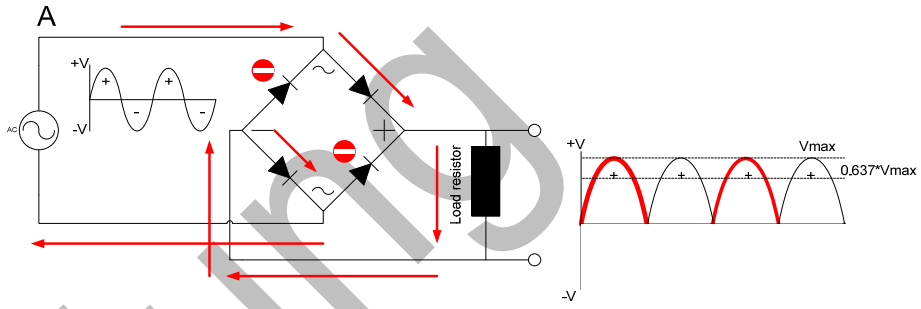


Figure 18 The Full Wave Bridge Rectifier Positive Path

The negative half cycle of the sinusoid is applied at point B (figure 19). The arrows show the path of the current on the negative half cycle.

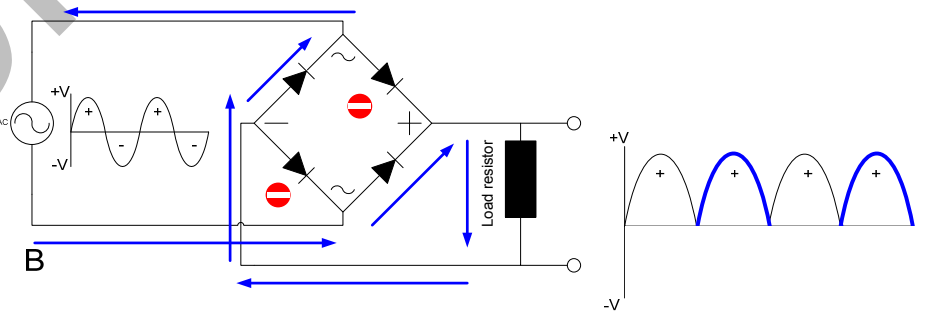


Figure 19 The Fullwave Rectifier Negative Path

The direction of the current through the load is the same for both cycles. Therefore the load signal is DC.

The full period of the applied sinusoid signal is produced at the load resistor. Therefore, the average voltage at the load is twice the value of the half-wave rectified circuit. Equation 2 shows the formulae to calculate the average voltage of a full-wave rectified signal.

$$V_{AV} = 0.637 \times V_{\max}$$

OR

$$V_{AV} = \frac{2V_{\max}}{\pi}$$

Equation 2: Average Voltage of a Full-Wave Rectified Load

Full-wave rectification is more efficient than half-wave rectification because the full cycle of the sine wave is converted to DC. It should be noted, that the DC is still variable. This is undesirable to apply in circuits that require a stable DC supply.

A capacitor can be added across the load resistor to further stabilise the DC output. The addition to the circuit is shown in figure 20.

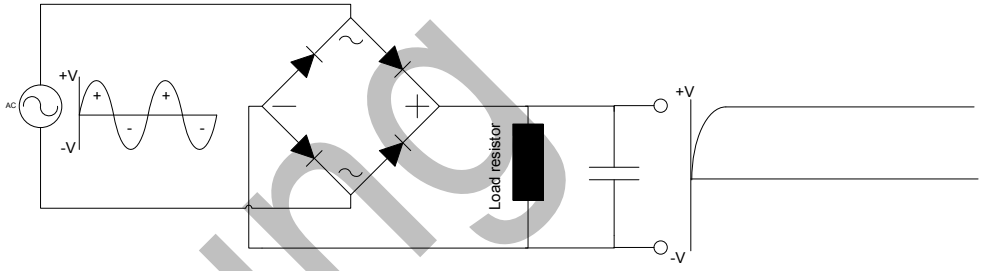


Figure 20 Capacitor Addition to Bridge Rectifier

It should be noted that the DC output will not be “perfect” (as illustrated). An element of ripple will be evident at the output.

The value of the capacitor will determine the time taken to charge and discharge. The charge/discharge process is the cause of the ripple.

Figure 21 shows the conversion from a full wave rectified DC signal to stable DC. The DC ripple can be minimised by the designer’s careful consideration of the capacitor value.

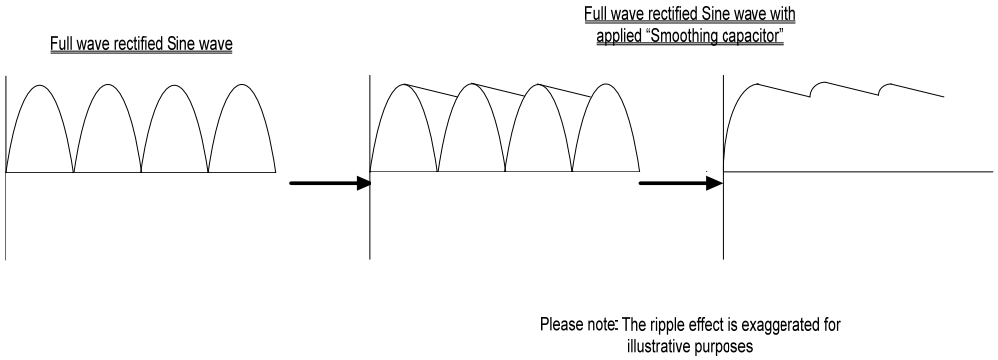


Figure 21 Rectified DC to Stabilised DC

The positive cycle will charge the capacitor until it reaches its maximum value. The voltage of the positive cycle will then decrease. During this time the capacitor will begin to discharge.

Before the capacitor can discharge completely, the negative cycle will re-charge the capacitor to its maximum value. The process will then begin again.

Other methods of full-wave rectification exist for single-phase signals. Methods also exist to rectify three phase signals.

1.2.5 Diode Testing

In most cases, a digital multi-meter (DMM) is the fastest way to check for a defective diode.

A functional diode will show extremely high resistance under reverse bias (or open circuit). It will show very low resistance under forward bias.

Faulty (Open)

A faulty open diode will read extremely high resistance in both forward and reverse bias. It can be caused by a defective junction on the diode material or an improper connection when mounted to a circuit board. This is the most common type of fault on a diode.

Faulty (Short)

A short circuit diode will show as zero resistance in both forward and reverse bias.

DMM Testing

Many digital multi-meters have a continuity function position designed to test the functionality of a diode. It is usually marked by a small diode symbol on the main dial. When this function is in use, the DMM will apply a bias voltage to the diode (Usually 2-3V).

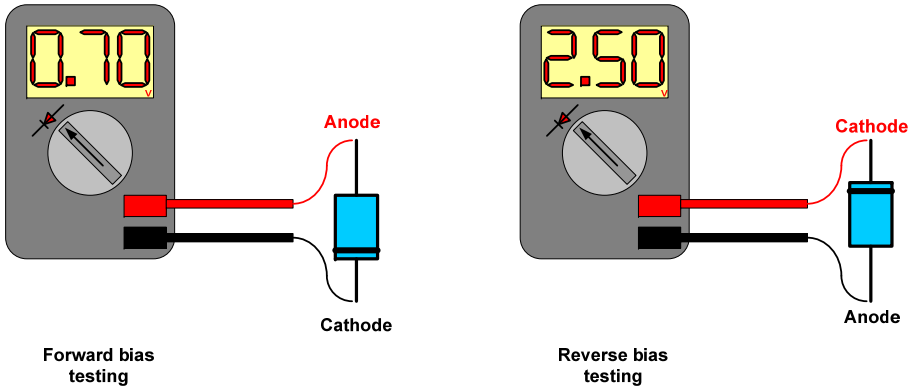


Figure 22 Functional Diode DMM Testing

Functioning Diode – DMM Test

The Cathode is marked by a black band on a diode. The red cable (used as positive) should be connected to the Anode to forward bias the diode. The black cable (Ground) should be connected to the Cathode. The user should read between 0.6V -0 .9V (the expected result is 0.7V for silicon). The diode may also emit a loud beep from the speaker when forward biased.

To reverse bias the diode, the user should connect the red wire to the Cathode. The black wire is connected to the Anode. The user should read a voltage based on the DMM supply. For illustrative purposes, 2.5V will be used (shown in figure 22). The DMM may also display 'OL'. These readings indicate a very high path of resistance (very low current = negligible voltage drop).

Defective Diode DMM Test

If a diode has an open circuit fault, the meter will display 2.5V or 'OL' when the diode is both forward and reverse biased. This is shown in figure 23.

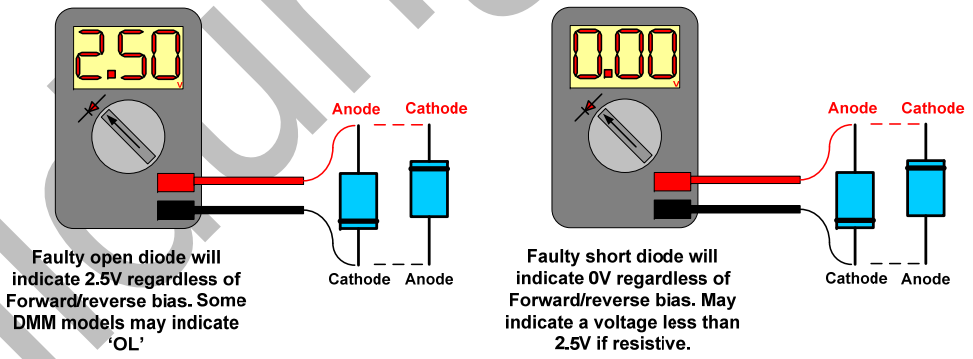


Figure 23 Faulty Diode DMM Test

If a diode has a short circuit fault, the DMM will read 0V when forward/reverse biased. The diode may not be a complete short and may exhibit a small internal resistance. This means that it will read a voltage larger than the maximum 0.9V in forward bias, but lower than 2.5V desired in reverse bias (typical values at around 1.1V-1.5V in both directions).

Static Resistance Measurements

An easy method of testing a diode is to use an ohmmeter. The forward 24 and reverse resistances 25 can be checked at a voltage determined by the battery potential of the ohmmeter and the resistance range at which the meter is set.

a) Forward resistance check

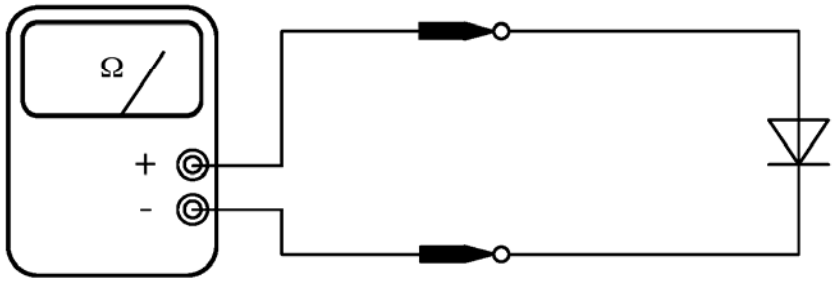


Figure 24 Testing a Diode (forward check)

b) Reverse resistance check

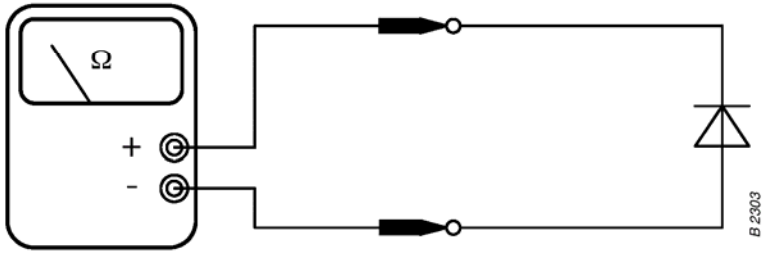


Figure 25 Testing a Diode (reverse check)

When the test leads of the ohmmeter are connected to the diode, a resistance will be measured that is different from the resistance indicated if the leads are reversed.

The smaller value is called the forward resistance, and the larger value is called the reverse resistance. If the ratio of reverse-to-forward resistance is approximately greater than 10 to 1, the diode should be capable of functioning as a rectifier.

This is a very limited test, which does not take into account the action of the diode at voltages of different magnitudes and frequencies.

Tutorial Sheet 2.1: P-N Junction Diodes

1. Describe the behaviour of the P-N junction under the following conditions:
   - Initial contact (no applied power).
   - Forward bias.
   - Reverse bias.

2a. Demonstrate (with illustrations) the construction of a standard P-N diode in terms of:
   - Conventional Current flow.
   - P-N junction layout.
   - Anode.
   - Cathode.
   - Equivalent circuit diagram.

2b. Describe the main uses of a P-N junction diode?

3a. Three Silicon diodes are connected in series with a resistive load (R_1) of 10kΩ. The supply voltage (V_S) is 10V. Calculate:
   - The Voltage drop across a single diode.
   - The Voltage dropped across all three diodes.
   - The remaining Voltage across R_1 (V_{R1}).
   - The total current draw.

3b. Three Silicon diodes are connected in series with a resistive load (R_2) of 100Ω. The circuit is drawing 29mA of current.

Using ohm's law and known power formulae, calculate:
   - The Voltage across the resistor R_2 (V_{R2}).
   - The Supply voltage (V_S).
   - The Power dissipated in the circuit.

4. Three Silicon diodes are connected in parallel. The Supply current is 6A. What is the current through each diode?

5a. Draw the circuit diagram of a Half-wave rectifier.

5b. Show the output (with respect to the 10V AC input) of a Half-wave rectifying circuit (the barrier potential of the P-N junction must be considered).

5c. State the average output Voltage of a 240V Half-wave rectified signal.

6a. Draw the circuit diagram of a Full-Wave bridge rectifier.

6b. Show the conventional current path of the positive half-cycle of the AC input signal.

6c. Show the conventional current path of the negative half-cycle of the AC input signal.

6d. Show the output DC signal (with respect to the 10V AC input signal) of a Full-wave rectifier (the barrier potential of a P-N junction must be considered).

6e. State the average output Voltage of a 240V Full-wave rectified signal.

6f. Explain the effect of adding a Capacitor to the Full-wave rectifier circuit.

6g. Research and illustrate another type of Full-Wave rectifier circuit using P-N diodes. Explain its operation.

.

1.2.6 Light Emitting Diodes (LEDs)

The circuit diagram and physical construction of an LED is shown in figure 26.

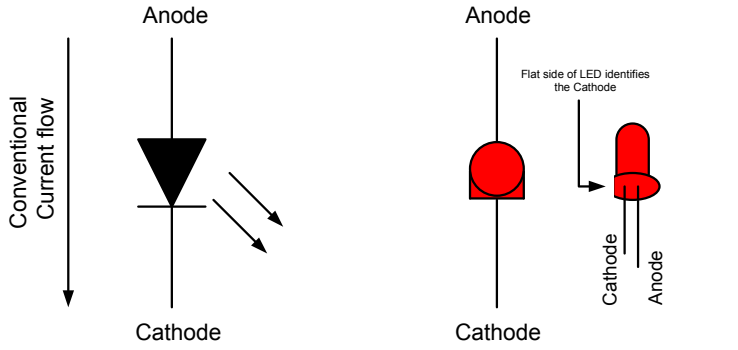


Figure 26 Light Emitting Diode (LED)

The forward and reverse conditions of the LED are identical to that of the standard P-N diode.

The exception can be found under the forward bias operation of the LED.

As previously shown, under forward bias conditions, electron-hole combinations occur as electrons cross the P-N junction. When the holes and electrons combine, energy is released in the form of heat and light.

A large area on a layer of the semi-conductor is left exposed to permit photons to escape as visible light. This process is known as electro-luminescence. This is illustrated in figure 27.

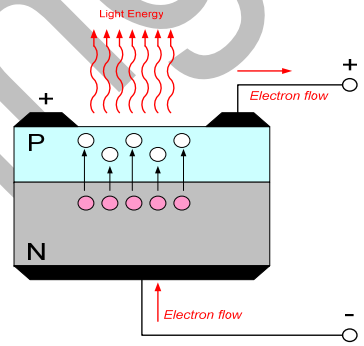


Figure 27 Electro-Luminescence in a Led

Silicon and Germanium are not used because they primarily produce energy in the form of heat. Gallium Arsenide (GaAs), Gallium Arsenide Phosphide (GaAsP) and Gallium Phosphide (GaP) are used as primary LED semi-conductor materials.

Gallium Arsenide (GaAs) emits infra-red (IR) radiation. Gallium Arsenide Phosphide (GaAsP) emits red or yellow visible light. Gallium Phosphide (Gap) emits red or green light.

The amount of visible light generated is directly proportional to the forward current. The colour of light generated by each material may also be controlled by the doping procedures. A guide to a good operating current to produce sufficient visible light is around 30mA.

A common use for LED's is in seven segment displays and readout displays.

1.2.7 Zener Diode

The Zener effect describes the process where a diode's applied voltage reaches a level large enough to allow electrons to break their covalent bonds.

In a standard P-N diode, the leakage current remains constant up to a certain applied reverse bias. When that limit is exceeded the leakage rises to a level that is limited only by circuit resistance/internal impedance.

When operating in forward bias, the Zener diode behaves in an identical fashion to a standard P-N diode. The rush of current can begin when the barrier potential is exceeded.

When under reverse bias operating conditions (detailed earlier), a standard P-N diode would be destroyed if the reverse limit is exceeded. The Zener diode can utilise the Avalanche effect. The Zener diode is shown in figure 28.

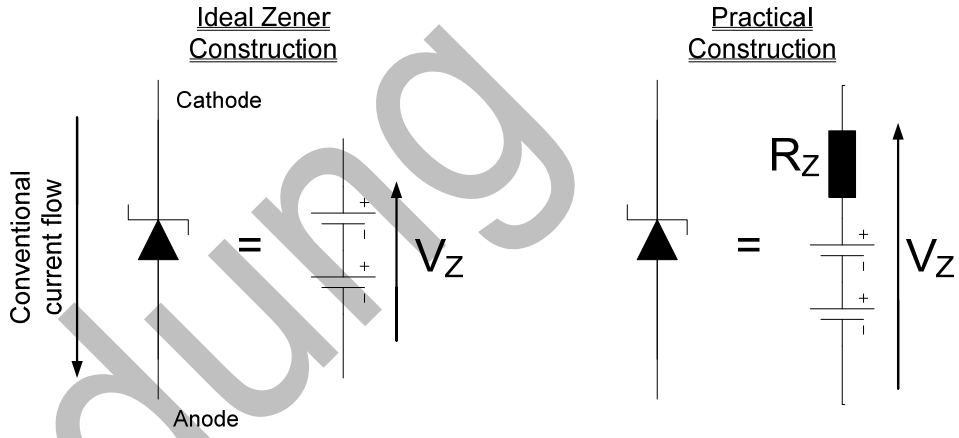


Figure 28 The Zener Diode

The ideal equivalent circuit shows that if the Zener diode is connected to a varying (unstable) input voltage, it will act as a stable (fixed) signal if the output is taken across it.

The practical circuit shows that the Zener diode has "Internal impedance" due to the manufacture and materials used in the diode construction. This will limit the current through the Zener diode.

The reverse breakdown voltage limit is designed to be lower for a Zener diode (3→100 Volts) than for a standard P-N diode (Silicon ≈ 1000V).

The operating conditions of a Zener diode are shown in figure 29.

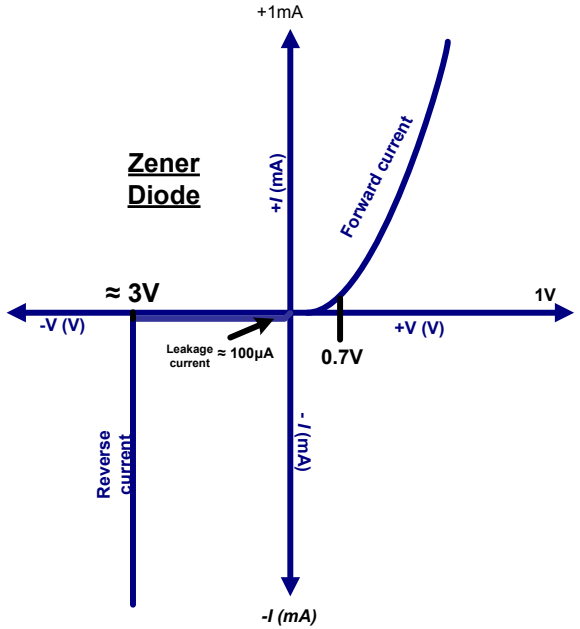


Figure 29 Operating Conditions of a Zener Diode

It can be seen from the operating conditions that the "rush of current" behaves the same for the Zener diode when in reverse bias, as it does when in forward bias.

The difference is that the "Rush of current" occurs at a higher value when in reverse bias, than in forward bias. Eg:

- Forward bias potential 0.7V.
- Reverse bias potential 3V→100V.

The Zener Diode as a Voltage Stabiliser

The Zener diode can be used as a voltage stabiliser. A large reverse current increase can be produced from a small reverse voltage increase.

It is this characteristic that allows a small voltage to be fixed from a larger supply. The design of a Zener diode stabiliser circuit is shown in figure 30.

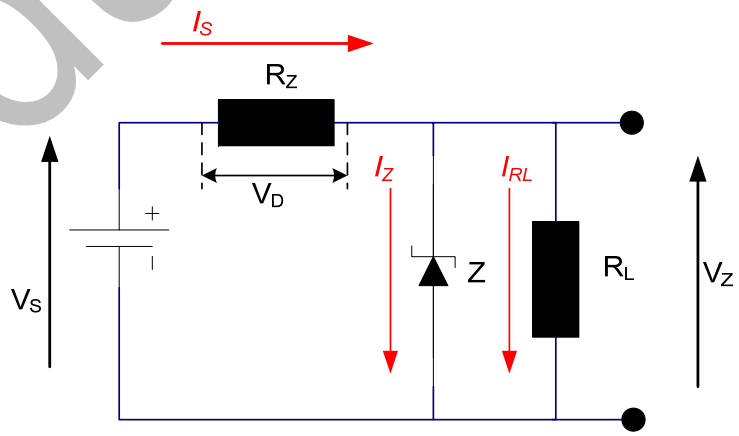


Figure 30 Voltage Stabiliser Circuit

Example:

A 20V stabilised signal is needed from a 50V supply dissipating 4W of power.

This means that the circuit must drop 30V across the series resistor (R_S). To calculate the value of the series resistor we must first use the power formula;

$$I_Z = \frac{P_T}{V_Z}$$

Where;

I_Z = Current through the Zener diode
P_T= Power dissipated in the circuit (Watts)
V_Z = Desired Voltage across the Zener diode

So:

$$I_Z = \frac{4W}{20V} = 0.2A = 200mA$$

The current through the Zener diode (therefore the series resistor current) is 200mA.

The series resistor value can now be calculated using ohm's law.

$$R_S = \frac{V_D}{I_Z}$$

Where;

V_D = Voltage dropped across Series resistor.
R_S = Series resistance.

So:

$$R_S = \frac{30V}{0.2A} = 150\Omega$$

A 150Ω resistor is required to stabilise 20V across a Zener diode from a 50V supply.

Provided the breakdown current is maintained, the Zener can act as a reservoir of current. If the supply voltage increases, the current through the series resistor increases. This maintains the Voltage across the Zener diode.
If the supply falls, the current decreases (with the drop across the resistor), again maintaining the voltage across the diode.

A minimum value of reverse current must be maintained to keep the diode in regulation. This value is known as I_{ZK}.

The maximum value of reverse current, above which the diode may be damaged, is known as I_{ZM}.

The data sheet supplied with the diode will detail the Zener test voltage, V_{ZT}, at a value of reverse current called I_{ZT}.

The load across a Zener diode should also be observed. It should be noted that current will follow the path of least resistance.

The Zener stabiliser operates more efficiently when its parallel load resistance is a high value.

If the load resistance is allowed to fall, its current will increase (at the expense of the Zener diode). This will mean that the avalanche effect operating the Zener will cease if its voltage falls below the breakdown limit. Equation 5 shows the method of calculating the load current using ohm's law. Equation 6 shows the method of calculating the Zener current using the obtained load current

$$I_{RL} = \frac{V_Z}{R_L}$$

Where;

I_{RL} = Load current.
V_Z = Zener voltage (also equal to load voltage).
R_L = Load resistance.

Equation 5: OhHMs Law used to calcualte Load Current

$$I_Z = I_S - I_{RL}$$

Where;

I_Z = Zener current.
I_S = Total supplied current.

Equation 6: Zener Current (Calculated From Load Current)

Experimentation on the Zener diode will show that as the Load resistor current increases, the Zener current decreases by the same amount. (The voltage across the Load and Zener will remain fixed if it is above the breakdown limit).

If the load resistor current decreases, the current through the Zener diode will increase by the same amount. Figure 31 shows the effect of lowering the load resistance (increasing load current at the expense of the Zener diode).

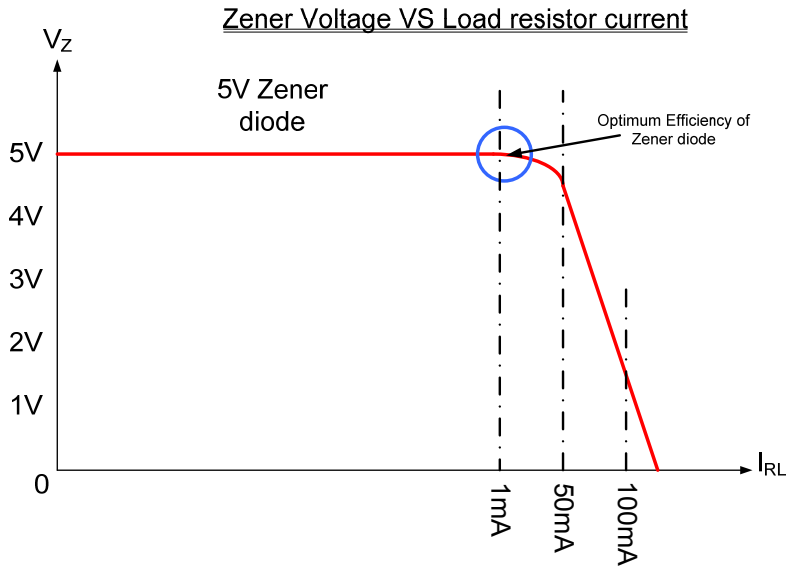


Figure 31 Effect of Lowering Load Resistance

A method to restore the tolerance of the Zener breakdown voltage is to choose a higher series resistor when low-resistance loads are applied across the Zener diode.

A good example application of the Zener diode as a voltage stabiliser is in a circuit designed to protect against Electro-static discharge.

1.2.8 The Shottky Diode

The circuit diagram for a Shottky diode is shown in figure 32.

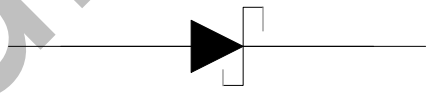


Figure 32 Shottky Diode

The Shottky diode fuses metal to a semiconductor to improve switching speeds for application in logic circuits. The mobility of electrons is greater than holes, so the N-type doped semiconductor is used.

Aluminium has three valance electrons, making it an acceptor atom. When Aluminium is fused to an N-type semiconductor, some of the atoms migrate into the silicon making a very thin P-region. Almost all of the holes combine with electrons. This means that nearly 100% of the current carriers are electrons as they are abundant in the N-type material.

As a consequence, minority carriers (holes in this case) do not contribute to any current flow process. The diode has a smaller depletion region, thus lower capacitance and a smaller forward conducting barrier voltage (0.2V-0.4V).

1.2.9 Varactor Diode

The circuit diagram for a Varactor diode is shown in figure 33.

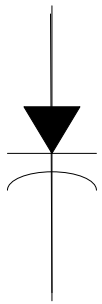


Figure 33 Varactor Diode

Varactor diodes are known also as variable capacitors. This is because P-N junction capacitance varies with the amount of reverse bias voltage applied to it. This means that the capacitance can be changed simply by changing the reverse bias.

The Varactor diode is mainly used in electronic tuning circuits, such as TV's and radios. It is essentially a reverse bias P-N junction that uses the capacitance of the induced the depletion region. It repels the majority carriers; therefore it will not conduct. This characteristic causes it to behave like the insulating dielectric material of a capacitor (The material between the plates).

The P and N regions are conductive, so they behave like the plates on a capacitor. From module 3 (Capacitors), it can be seen that capacitance is affected by the area of the plates (A), the distance between them (thickness of the dielectric, D) and the dielectric constants (ε_o and ε_R). This is shown in Equation 7.

$$C = \frac{\varepsilon_O \varepsilon_R A}{D}$$

Equation 7: Capacitance Formula

As the reverse bias voltage increases, the depletion region at the P-N junction widens. This has the effect of increasing the dielectric thickness which would decrease the value of the capacitor.

As the reverse bias decreases, the depletion region narrows, therefore decreasing the depletion region and increasing the capacitance. This is shown in figure 34.

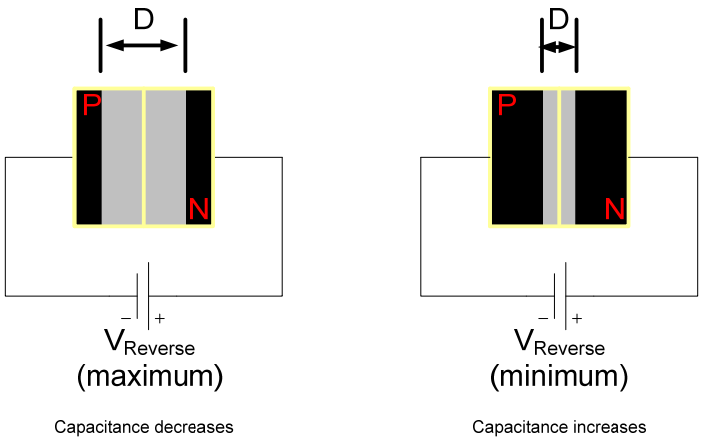


Figure 34 Depletion Region on a Varactor Diode

The capacitance parameters are generally controlled by the doping methods of the depletion region and the construction of the diode. The Varactor diodes have a typical capacitance range of a few pico-farads to a few hundred pico-farads.

1.2.10 Photo-Diodes

The circuit symbol for a Photo-diode is shown in figure 35.

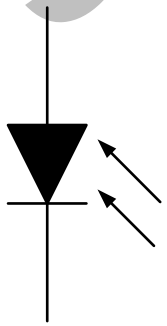


Figure 35 Photo-Diode

The photo-diode is a P-N junction device that operates in reverse bias. The diode has a transparent area that allows light to strike the junction.

When operating under reverse bias, a standard rectifier diode has a small leakage current caused by minority carriers.

Reverse bias current is generated thermally by electron-hole combinations in the depletion region.
The electron-hole pairs are carried across the junction by the applied reverse bias voltage.

In a rectifier diode, increased electron-hole combinations causes a rise in temperature. This rise in temperature causes a rise in the reverse current.
In a Photo-diode, the reverse current increases with the intensity of light at the exposed P-N junction.

When there is no supplied light, the reverse current (I_λ) is negligible. This is known as dark current. An increase in light intensity produces an increase in reverse current. Light intensity is expressed in Irradiance (mW/cm^2). This is shown in figure 40.

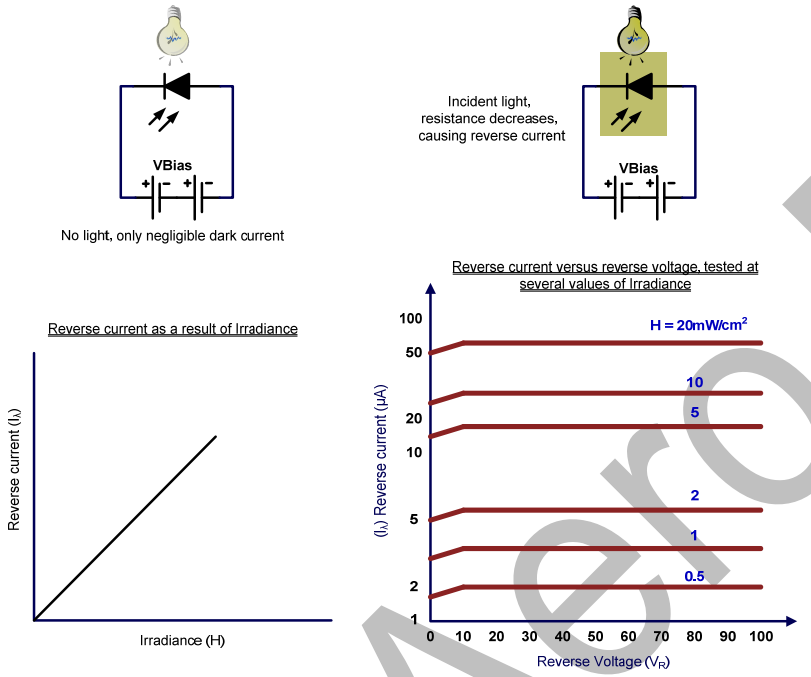


Figure 36 Supplied Light on a Photo-Diode

1.2.11 Voltage-Sensitive Components

Voltage-sensitive components are referred to as 'varistors'. The name varistor stands for 'variable resistor' but because of its resistance/ voltage characteristic it is also called 'voltage-dependent resistor' (VDR). Detail a) shows the schematic symbol of a VDR.

As shown in the resistance/voltage characteristic diagram (detail b)), the resistance of a VDR decreases when the applied voltage increases, and the resistance increases when the voltage decreases. This characteristic makes the VDR an appropriate component to absorb transient voltages occurring in electronic circuits.

Varistors are solid-state components made from semiconductor materials, such as silicon carbide (SiC), selenium (Se) and zinc oxide (ZnO) as main ingredients Figure 38, detail a)).

The semiconductor material greatly used determines the voltage-dependent characteristic and the ability of the VDR to absorb transient voltages (detail b)).

Generally, the insulation between SiC particles (in case of SiC varistors), and those of ZnO grains (in case of ZnO varistors) is maintained while the voltage applied to varistors is low. But, when a voltage beyond a specified voltage value (the varistor voltage) has been applied, it destroys the insulation and causes a large current flow, that means the resistance becomes low. The insulation is restored when current flow sharply decreases and the applied voltage falls below the varistor voltage value.

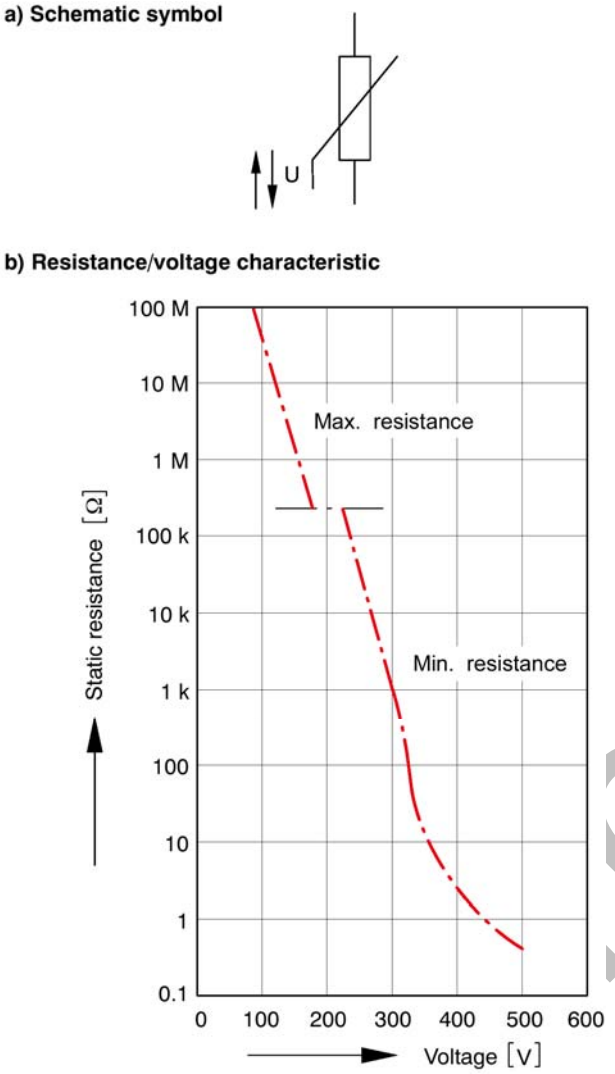


Figure 37 Varistor (ZnO)

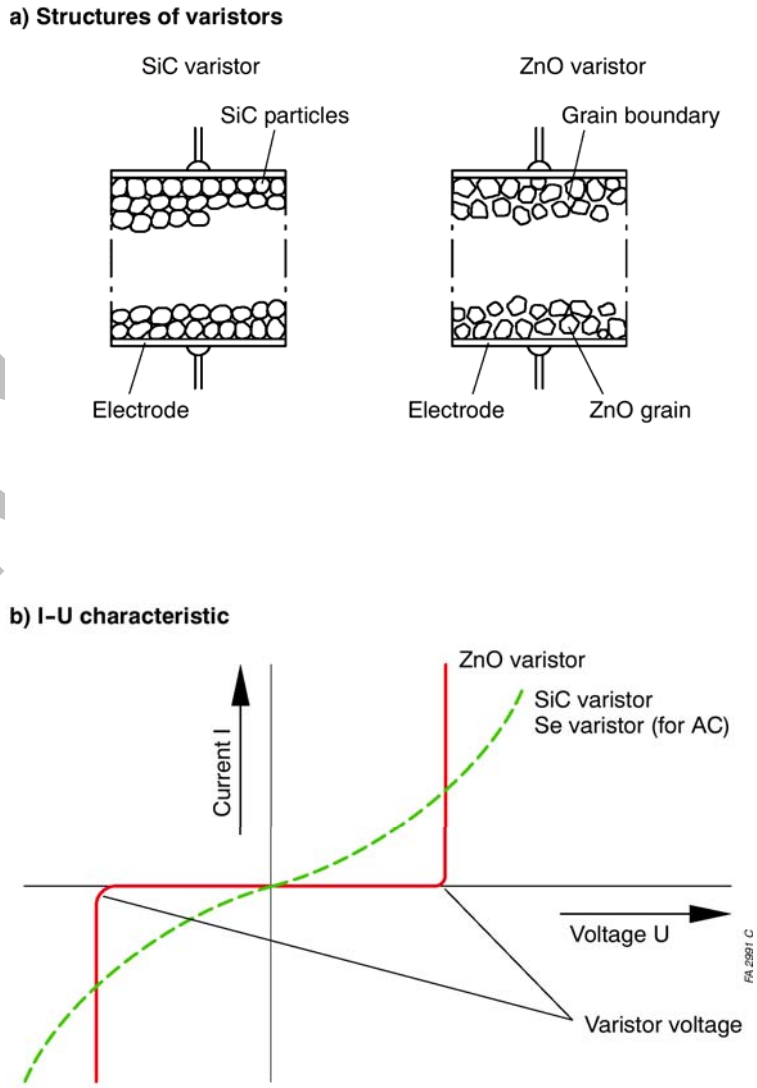


Figure 38 Current/Voltage Characteristics

1.2.12 Photo -Conductive Cells

Photo conductive cells, commonly known as light-dependent-resistors (LDR), are passive electronic components used for the detection of light (optoelectronic radiation). They consist of semiconductor material, such as

- ☐ cadmium sulphide (CdS)
- ☐ cadmium selenide (CdSe)
- ☐ lead sulphide (PbS).

In Figure 39, detail a) a photo-conductive-cell is represented. Detail b) shows the schematic symbol.

Photo-conductive-cells can be regarded as resistors which change their resistance according to the intensity of light illumination (detail c). The photo-conductive resistance decreases when the light intensity increases, and vice versa, when the light intensity decreases the resistance increases. Since photo-conductive cells have no PN junctions, they can be used for DC and AC circuits. They provide high photo-sensitivity to light, employing wavelengths from ultraviolet to infrared.

These types of optoelectronic detectors are

- ☐ highly sensitive to light radiation
- ☐ robust
- ☐ cheap, compared to other optoelectronic detectors.

Application of Photo -Conductive Cells

Generally, photo conductive cells are employed in electronic circuits as

- ☐ dimming switches
- ☐ automatic flame -guards
- ☐ lighting control
- ☐ photometers.

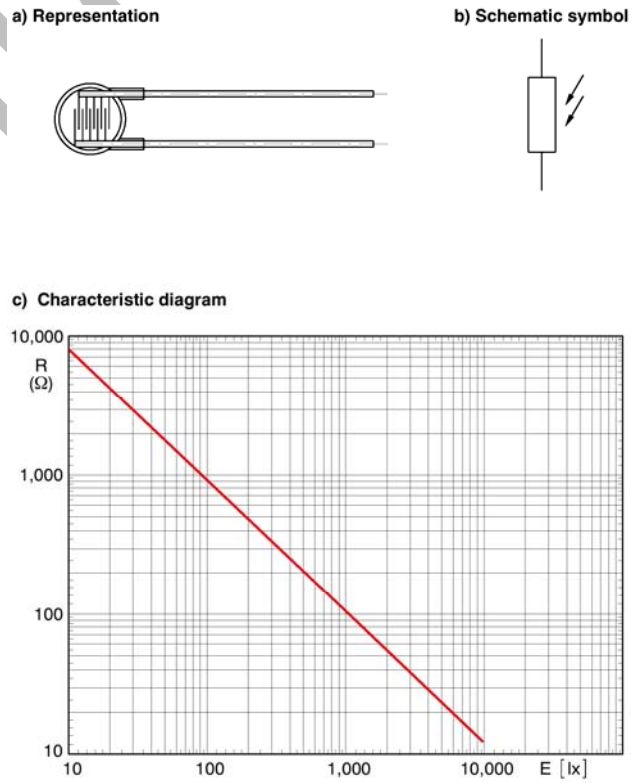


Figure 39 Photo-Conductive Cell

1.2.13 Thyristor

Thyristors are constructed using four separate layers of semiconductor material. They are called the Silicon Controlled Rectifier (SCR).

The construction, equivalent circuit and circuit symbol are shown in figure 40.

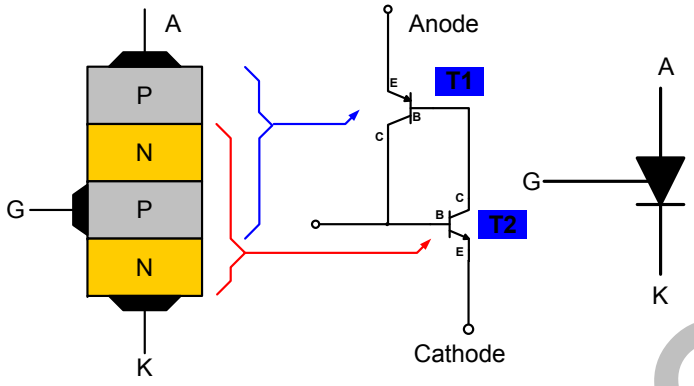


Figure 40 Silicon Controlled Rectifier

The physical construction of the SCR is thought of as an NPN transistor interconnected with a PNP transistor. It has two P regions and two N regions. The two middle layers of the transistor are shared.
The SCR is a Uni-directional device (like a rectifier diode) with three terminals. The current conduction is controlled by the gate (G). Normal forward bias operation applies when a positive voltage is applied to the gate.

Normal transistor operation is covered later in these notes. When a burst of current is applied to the gate, it becomes the initial base of the transistor junction **T2**. This action switches **T2** on; pulling a larger current from the collector to the emitter (this is the normal behaviour of an NPN transistor).

This also has an effect on the PNP transistor, illustrated as **T1**. The base of **T1** is connected to the collector of **T2**. The large current flow into the collector of **T2** forms the conducting path for the base current of **T1**. This action switches **T1** on, allowing current to flow from the emitter into the base. In turn, **T2** pulls a larger current from the collector of **T1** into its base.

The entire SCR diode current sustains itself as the collector current of **T1** is large enough to support the base current of **T2**, even when the gate voltage is removed. This base current is the minimum required value to keep **T2** switched on. This becomes the minimum holding current.

The SCR will remain switched on until the electron flow from the Cathode to the Anode (Conventional current flow of Anode-Cathode) is greater than or equal to a specified holding current. Once the active current drops below the holding value, the SCR will turn off. It can be switched on again with another voltage applied to the gate.

The SCR will also switch off if a reverse bias is applied to the diode or the supply is removed completely.

The diagram in figure 41 shows the SCR used to rectify and control the average power delivered to a load. It will half-wave rectify a sinusoidal load like a normal diode. A trigger pulse is applied to the SCR gate during the positive half-cycle of the input signal. If the SCR is triggered earlier in the half-cycle, the load receives more average power. If the SCR is triggered later in the half-cycle, less average power is delivered to the load.

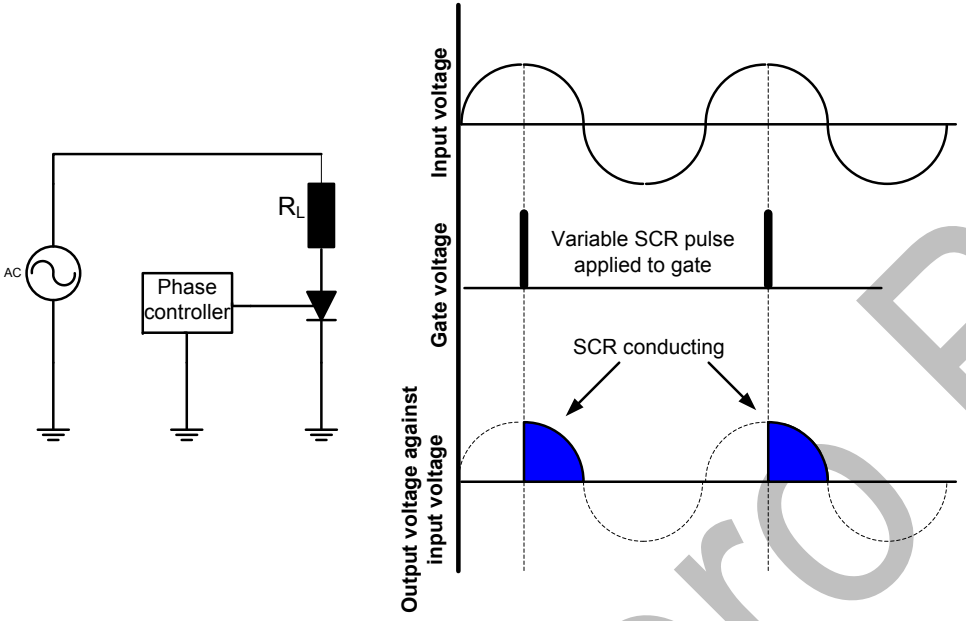


Figure 41 Thyristor Average Power Controller

1.2.14 The Triac

The Triac consists of two Tyristors connected to allow bi-directional current flow. Both thyristors share a common gate. The circuit symbol for a Triac is shown in figure 42.

The simplest method of describing the Triac shows one thyristor is forward bias during forward current (the other is reverse bias and not operating), the other thyristor is forward bias during reverse current. This is shown in figure 42.

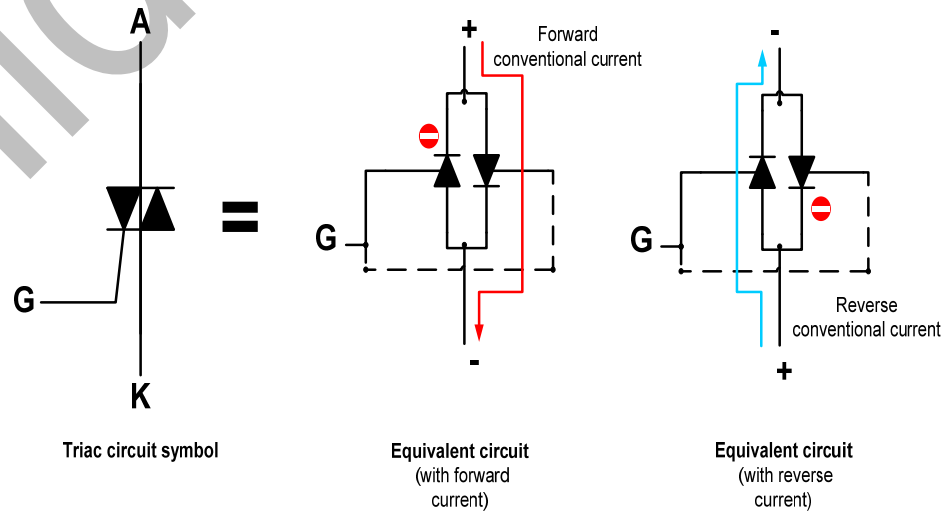


Figure 42 The Triac

Triacs can also be used to control the amount of active power delivered to a load using the same method as the thyristor. The gate trigger can be applied to deliver a controlled portion of each half-cycle (both positive and negative) to the load. The diagram for the full wave power controller is shown in figure 43.

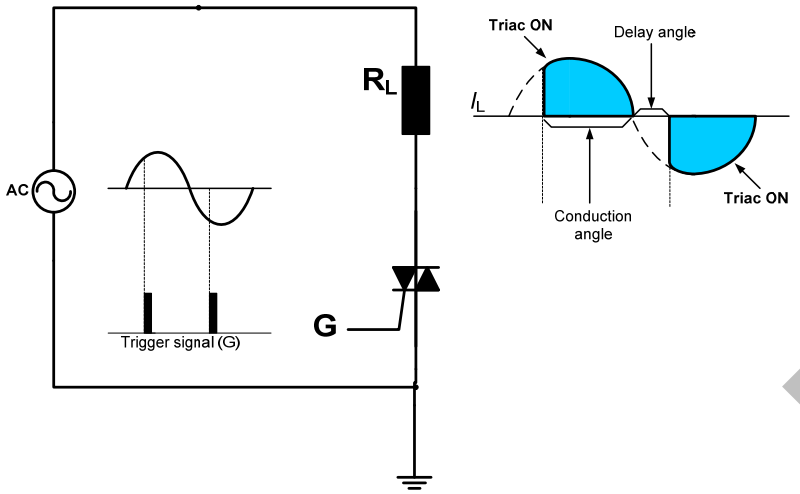


Figure 43 Full Wave Phase Controller

Throughout the positive half-cycle of the input, the diode is off until the gate is triggered. This is known as the delay angle. When this occurs the diode is allowed to conduct for the remaining portion of the half-cycle. This is known as the conduction angle.

The Triac is bi-directional, so a reverse current is allowed through the diode. This means that the triggered phase angle control can also be applied to the negative half-cycle of the input.

1.2.15 The Diac

The Diac is a bi-directional device that operates without a gate. It can operate if a sufficient voltage, known as the break-over potential, is applied across it. The circuit symbol for the Diac is shown in figure 44.

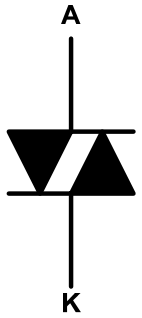


Figure 44 The Diac

1.2.16 Tutorial Sheet 2.2: Other Diodes

1a. Show the circuit diagram of an LED, include the conventional current flow.

1b. What is the cause of emitted light in an LED?

1c. What Semi-conductors are used in the construction of an LED?

2a. A Zener diode exhibits a 50mV change in V_Z for a 2mA change in I_Z. What is the Zener impedance?

2b. A certain Zener diode has an impedance of 5Ω. The data sheet supplied with the diode gives V_{ZT} = 6.8V at I_{ZT} = 20mA.

What is the voltage across the Zener diode, V_Z, when the current is 30mA?

What is the V_Z when I_Z = 10mA?

2c. Use figure 34 as a circuit design, the load resistor (R_L) can be ignored. Determine the minimum and maximum voltages that can be regulated by the Zener diode if the minimum operating current, I_{ZK} = 1mA, V_{ZT} = 5.1V at I_{ZT} = 7mA and the maximum operating current, I_{ZM} = 15mA.

The Zener impedance is 10Ω. The series resistor, R_Z is 680Ω.

2d. Use figure 34 as a circuit design. Determine the minimum and maximum load currents for which the Zener diode will maintain regulation.

The supply voltage is 24V. The series resistor R_Z is 470Ω. V_Z = 12V, I_{ZK} = 3mA, I_{ZM} = 90mA.

Assume that Z_Z = 0Ω and V_Z remains a constant 12V over a range of current values for simplicity. What is the minimum R_L that can be used?

3a. Name the three types of Thyristor.

3b. Describe the operation of one of the three Thyristors.

4. Describe the operation of the Schottky diode.

5. Describe the operation of the Varacter diode.

6. Describe the operation of the Photo-diode.

Tutorial Sheet 2.1 Solutions

1. Directly referenced from the text.

2a. Directly referenced from the text.

2b. Directly referenced from the text.

3a. The voltage drop across a single diode is 0.7V. The voltage drop across all 3 diodes is 2.1V (0.7V*3). The remaining voltage across R_1 is;

$$V_{R1} = V_S - (3 \times V_D)$$

so;

$$V_{R1} = 10V - 2.1V = 7.9V$$

The current draw from the supply is;

$$I_S = \frac{VR_1}{R_1} = \frac{7.9V}{10k\Omega} = 790\mu A$$

3b. The voltage across the resistor, R_2, is developed using the supply current;

$$V_{R2} = 29mA \times 100\Omega = 2.9V$$

The supply voltage is determined using the calculated drop across R_2 and the drop across the 3 diodes.

$$V_S = V_{R2} + (3 \times V_D)$$
$$V_S = 2.9V + (3 \times 0.7V)$$
$$V_S = 2.9V + 2.1V = 5V$$

The total power dissipated in the circuit is;

$$P_T = V_S \times I_S = 5V \times 29mA = 145mW$$

4. The three diodes are in parallel so the current will divide equally between them;

$$I_D = \frac{I_T}{3} = \frac{6A}{3} = 2A$$

5a. Directly referenced from the text.

5b. Directly referenced from the text.

5c. The average value of a 240V Half-wave rectified AC signal is;

$$V_{AV} = 0.318 \times V_{\max}$$
$$V_{AV} = 0.318 \times 240V = 76.32V$$

6a. Directly referenced from the text.

6b. Directly referenced from the text.

6c. Directly referenced from the text.

6d. Directly referenced from the text.

6e. The average voltage of a 240V full-wave rectified signal is;

$$V_{AV} = 0.637 \times V_{\max}$$
$$V_{AV} = 0.637 \times 240v = 152.88V$$

6f. Directly referenced from the text.

6g. Personal research.

The desired clamping frequency is;

$$F_{CLAMP} = 10F_{RC}$$
$$F_{CLAMP} = 10 \times 10Hz = 100Hz$$

Tutorial Sheet 2.2: Solutions

1a Directly referenced from the text.

1b. Directly referenced from the text.

1c. Directly referenced from the text.

2a. $Z_Z = \frac{\Delta V_Z}{\Delta I_Z} = \frac{50mV}{2mA} = 25\Omega$, the term 'Δ' denotes the change in signal values.

2b. Figure 11a can be used from the text. The 30mA current is a 10mA increase (ΔI_Z) above I_{ZT} = 20mA.

So;

$$\Delta I_Z = +10mA$$

Meaning;

$$\Delta V_Z = (\Delta I_Z \times Z_Z) = (10mA) \times (5\Omega) = +50mV$$

The change in voltage due to the increase in current above the I_{ZT} value causes the Zener terminal voltage to increase. The Zener voltage for I_Z = 30mA is;

$$V_Z = 6.8V + \Delta V_Z = 6.8V + 50mV = 6.85V$$

The Zener voltage for I_Z = 10mA can also be given. The 10mA current is a 10mA decrease below I_{ZT} = 20mA.

So;

$$\Delta I_Z = -10mA$$

Meaning

$$\Delta V_Z = (\Delta I_Z \times Z_Z) = (-10mA) \times (5\Omega) = -50mV$$

The change in Zener voltage due to the decrease in current below I_{ZT} causes the Zener terminal voltage to decrease. The Zener voltage for I_Z = 10mA is;

$$V_Z = 6.8V - \Delta V_Z = 6.8V - 50mV = 6.75V$$

2c. Using figure 11c as a design, with the Zener impedance at 10Ω and a 680Ω series resistor;

At I_{ZK}, the output voltage is;

$$V_{OUT} = V_{ZT} - \Delta V_Z$$

so;

$$V_{OUT} = 5.1V - (\Delta I_Z \times Z_Z)$$

The value of V_{ZT} at I_{ZT} has been given so it can be assumed that these will be the normal operating values. This means that the value of I_{ZT} must change to I_{ZK}. This means that the Zener current will change from 7mA (I_{ZT}) to 1mA (I_{ZK}). This can be considered as shown;

$$V_{OUT} = 5.1V - ((I_{ZT} - I_{ZK}) \times Z_Z)$$

so;

$$V_{OUT} = 5.1V - ((6mA) \times 10\Omega)$$

finally

$$V_{OUT} = 5.1V - 0.06V = 5.04V$$

The minimum input voltage can be determined;

$$V_{IN(\min)} = (I_{ZK} \times R_Z) + V_{OUT}$$

so;

$$V_{IN(\min)} = ((1mA)(680\Omega)) + 5.04V = 5.72V$$

At I_{ZK}, the change in Zener current is

15mA – 7mA = 8mA, so the maximum output voltage is;

$$V_{OUT} = 5.1V + \Delta V_Z$$

so;

$$V_{OUT} = 5.1V + (\Delta_{IZ} \times Z_Z)$$

so;

$$V_{OUT} = 5.1V + ((8mA) \times (10\Omega)$$

finally

$$V_{OUT} = 5.1V + 0.08V = 5.18V$$

The maximum input voltage can now be calculated;

$$V_{IN(\max)} = (I_{ZM} \times R_Z) + V_{OUT}$$

so;

$$V_{IN(\max)} = ((15mA)(680\Omega)) + 5.18V = 15.38V$$

2d. When I_L = 0A, I_Z is at its maximum. I_Z is also the total circuit current (I_T) in this condition.

$$I_Z = \frac{V_{IN} - V_Z}{R_Z} = \frac{24V - 12V}{470\Omega} = 25.5mA$$

This is significantly less than the maximum operating current, I_{ZM}, at 90mA, so it is acceptable to use. The minimum operating load current through R_L is 0A ($I_{L(min)}$ = 0A).

The maximum value of I_L will occur through R_L when I_Z is at its minimum (equal to I_{ZK}). This means that $I_{L(max)}$ is determined as follows;

$$I_{L(\max)} = I_T - I_{Z(\min)} = 25.5mA - 3mA = 22.5mA$$

The minimum value of R_L must be;

$$R_{L(\min)} = \frac{V_Z}{I_{L(\max)}} = \frac{12V}{22.5mA} = 533\Omega$$

3a. Directly referenced from the text.

3b. Directly referenced from the text.

4. Directly referenced from the text.

5. Directly referenced from the text.

6. Directly referenced from the text.

1.3 TRANSISTORS

1.3.1 The NPN Transistor

A transistor is a semi-conductive device used for amplification and automatic/digital switching.

The Bi-polar transistor crystal is constructed using two P-N junctions.

The construction of an NPN transistor is shown in figure 45.

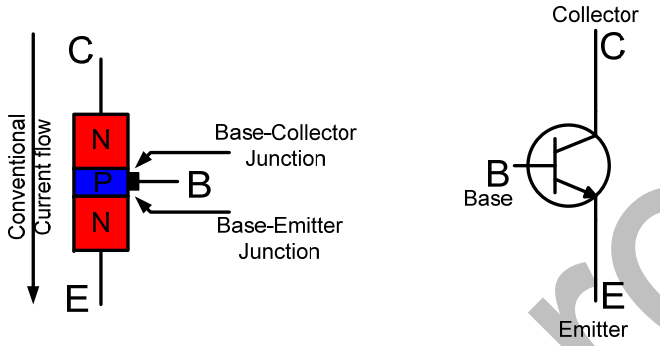


Figure 45 NPN Transistor Construction

Electrons are the majority carrier in the NPN transistor. The inner "P" region (known as the base) is thinner and very lightly doped in comparison to the "N" regions. The arrow from the device indicates the direction of conventional current flow.

NPN Characteristics

For the transistor to be on, the Base-Emitter junction must be forward biased. Much like the diode, the B-E junction must have a forward potential of at least 0.7V if it is a silicon device.

The base-emitter junction allows electron-hole combinations when operating in forward bias. This current is small (negligible) due to doping levels in the base and emitter regions. The thin base (P) region has a limited number of available holes when compared to the significant number of free electrons that have been applied to the Emitter. The remaining free electrons are attracted to the Collector supply voltage so they diffuse over the B-C junction.

Figure 46 shows the electron flow in the NPN transistor for illustrative purposes. It should be noted that conventional current flow is in the exact opposite direction to electron flow.

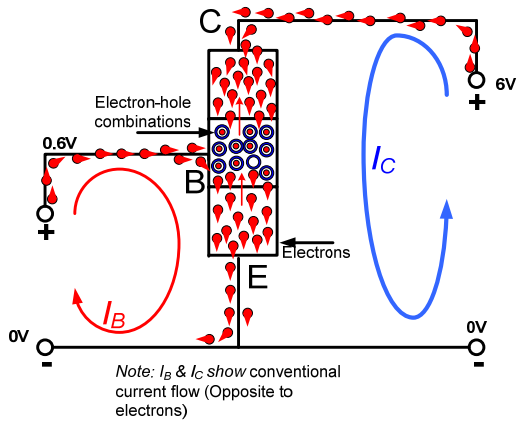


Figure 46 Electron Flow in an NPN Junction

The Base-emitter junction is forward bias in relation to the Base. The Base-Collector junction is reverse bias in relation to the Base. This is shown in figure 47.

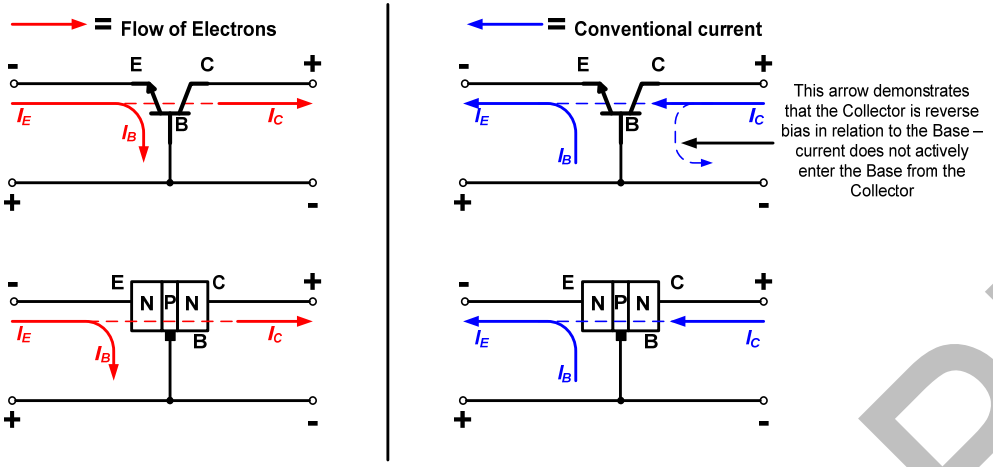


Figure 47 NPN Bias – In Relation to the Base

The small conventional current flow in the B-E junction "pulls" a much larger current from the Collector. This is because the Base-Emitter combinations cross into the more heavily doped (N-type) Collector.

Consequently, a small change in Base current (I_B) causes a larger change in Collector current (I_C). The ratio of Collector current to Base current is known as 'B'. Equation 1 shows the method of calculating the value of B.

$$B = \frac{I_C}{I_B}$$

and

$$I_C = B \times I_B$$

Equation 1: Method of Calculating B and I_C

The ratio may also be referred to as H_{FE} in some Electronic texts. The value of B is also sensitive to temperature. If the temperature is constant, the value of B is directly related to I_C. If the temperature is increased, the value of B will also increase. If the temperature is decreased, the value of B will also decrease.

It can also be seen that the conventional current output of the transistor is the Emitter (I_E). The Emitter current is the sum of the Base and Collector currents. Equation 2 shows the method of calculating the **actual** emitter current.

$$I_E = I_C + I_B$$

Equation 2: Method of Calculating Emitter Current (I_E)

Figure 48 shows the conventional current flow of an NPN transistor.

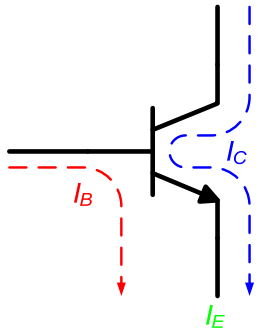


Figure 48 Conventional Current Flow of an NPN Transistor

1.3.2 The PNP Transistor

The construction of a PNP transistor is shown in figure 49.

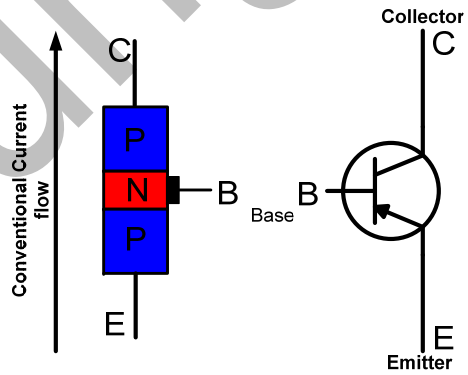


Figure 49 PNP Transistor Construction

As with the NPN transistor, the Base-Emitter junction must operate under forward bias to be active. To operate under this condition with Silicon, the bias must exceed 0.7V. Holes are the majority carriers in the PNP transistor.

The Emitter is still forward biased with respect to the base. However, in the case of the PNP transistor, the Emitter must be at a higher voltage than the Base (NPN transistor has the Base at a higher voltage than the Emitter).

The Collector is still reverse biased with respect to the Base. The Collector must be at a lower voltage than the Base. This is shown in figure 50.

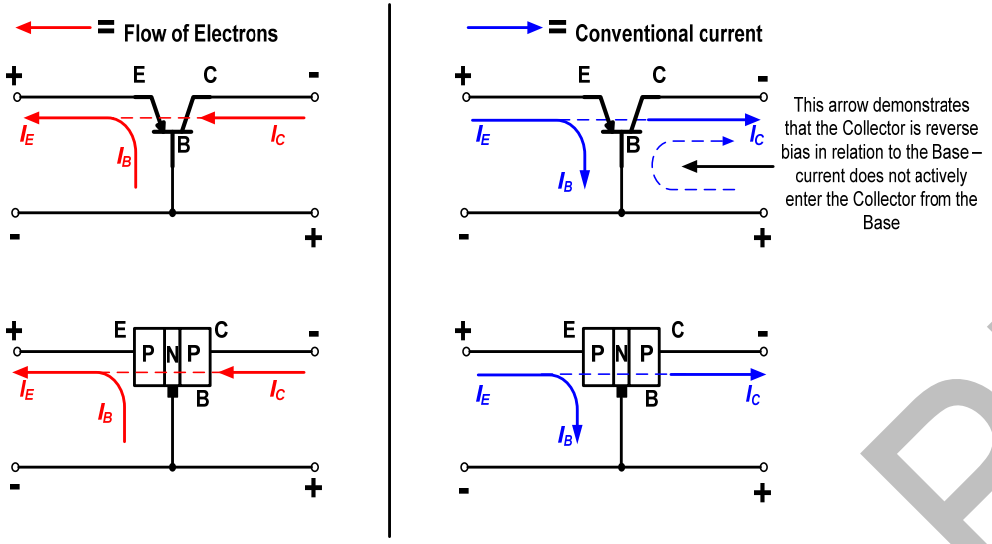


Figure 50 PNP Bias in Relation to the Base

Figure 51 shows the Conventional current flow in a PNP transistor.

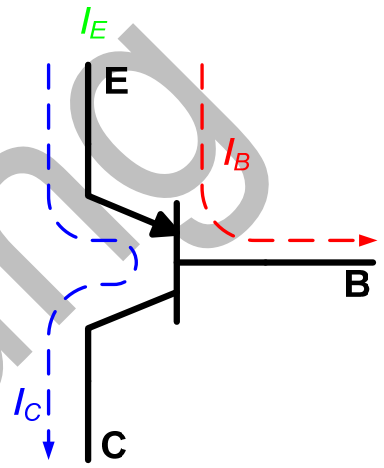


Figure 51 Conventional Current Flow in a PNP Transistor

It can be seen from the conventional current input to the PNP transistor that the Base and Collector currents are equal to the Emitter current when added together. The method of calculating the Emitter current is as shown in Equation 2.

$$I_E = I_C + I_B$$

Equation 2: Method of Calculating Emitter Current (I_E)

1.3.3 Gain

Gain is the measured amount that an electrical signal is increased or decreased. It is used to measure the level of amplification. This gain value is generated from the ratio of the output to the input.

Attenuation is used to describe loss from an electrical output signal (in relation to the input). In these circumstances the ratio of output to input is less than '1'.

Gain and Attenuation is a useful method of calculating the effectiveness of an amplifier. The primary objective of amplification is to produce a larger voltage/current/power signal from a smaller signal. The circuit symbol for an amplifier is shown in figure 3a.

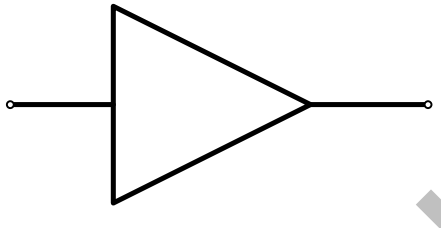


Figure 52 Circuit Symbol of a Standard Amplifier

The ratio of a single amplifier is simply the output voltage (AC) divided by the input voltage (AC). Due to the fact that this is a ratio, there are no units. This is shown in equation 3a.

$$Gain = \frac{V_{OUT}}{V_{IN}}$$

Equation 3a: Gain Ratio

This equation is used to generate the commercial unit of gain. This is measured in the Decibel (dB) unit. The equation to convert the gain ratio in Decibels is shown in equation 3b.

$$A_v(dB) = 20\log_{10}(G)$$

Equation 3b: Gain in dB

Example:

A Voltage amplifier has an AC input of 0.2V. It has an output of 5V. To calculate the gain ratio.

$$Gain = \frac{V_{OUT}}{V_{IN}} = \frac{5V}{0.2v} = 25$$

This amplifier has a gain ratio of 25. This means that the output is 25 times larger than the input. This can now be converted to dB.

$$A_v(dB) = 20\log_{10}(25) = 27.95dB$$

This is useful when many amplifiers are cascaded in series to create a multi-stage amplifier. A multi-stage amplifier is shown in figure 3b. The character 'G_n' represents the number of amplifiers.

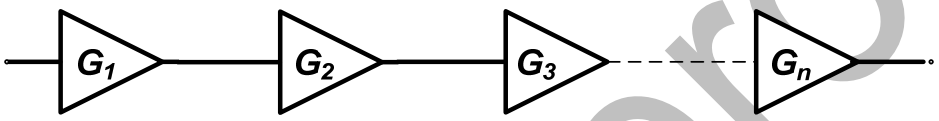


Figure 3b: Multi-Stage Amplifier

The total gain ratio is the product of all the amplifiers. This is shown in equation 3c. This means that the gain in dB is the sum of all the amplifiers dB values. This is shown in equation 3d.

$$G_{TOTAL} = G_1 \times G_2 \times G_3 \times .....G_n$$

Equation 3c: Multi-Stage Amp Total Gain

$$A_{v(TOTAL)}(dB) = A_v1(dB) + A_v2(dB) + A_v3(dB).....A_vn(dB)$$

Equation 3d: Multi-Stage Amp Total in dB

Example:

A cascaded amplifier has the following gains.

G_V1 = 10, G_V2 = 15 and G_V3 = 20. The overall gain ratio can be calculated as shown.

$$G_{TOTAL} = 10 \times 15 \times 20 = 3000$$

The gain in dB is calculated as shown for each stage and in total.

$$A_v1 = 20\log_{10}(10) = 20dB$$

$$A_v2 = 20\log_{10}(15) = 23.5dB$$

$$A_v3 = 20\log_{10}(20) = 26dB$$

So;

$$A_{v(Total)} = 20dB + 23.5dB + 26dB = 69.5dB$$

The gain behaviour of an amplifier is very sensitive to frequency, so the gain (dB) is used when calculating the best frequency to operate an amplifier.

The frequency boundaries for the best gain are measured at 3 decibels down from the desired operating gain. This is used when calculating the system bandwidth. The 3dB 'down-point' is the point at which the system consumes half the power that it would if it was running at the desired gain value. This is illustrated in figure 53.

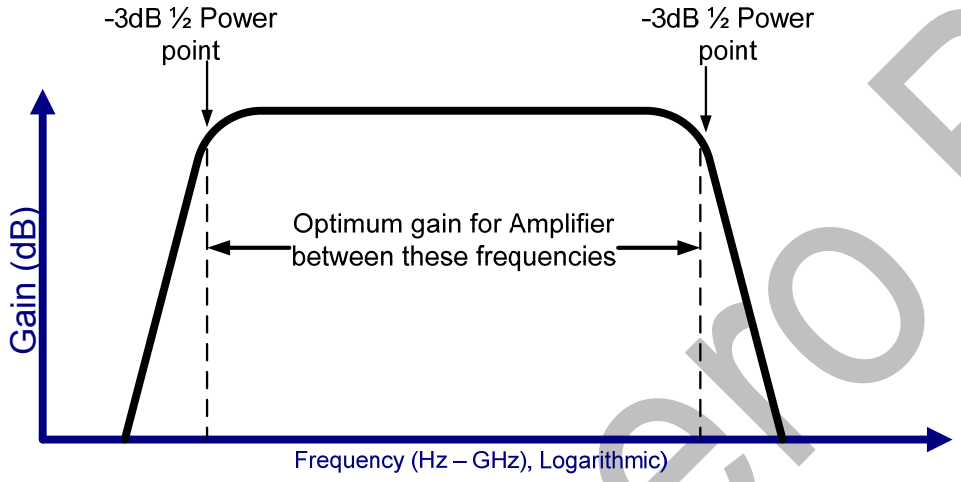
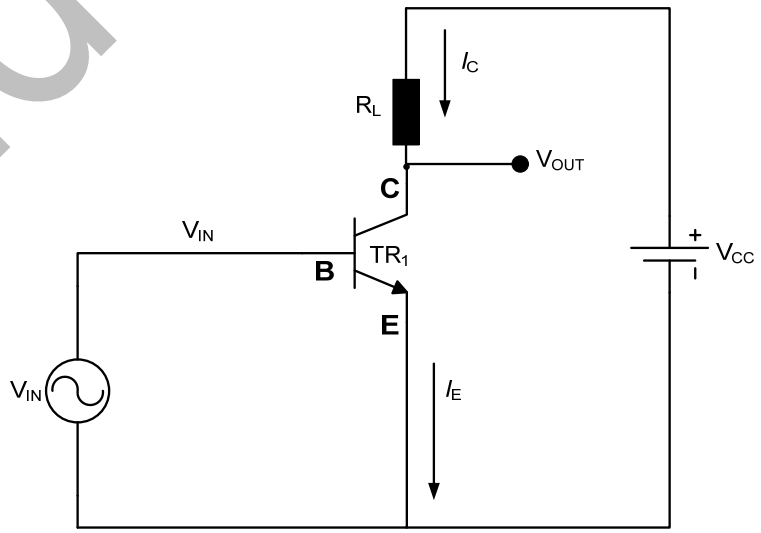


Figure 53 3dB Half-Power Point

1.3.4 THE TRANSISTOR AS A VOLTAGE AMPLIFIER

DC bias allows a transistor to operate as an amplifier. The circuit describes a current amplifier. This can be converted to a Voltage amplifier with a simple adjustment.

To obtain an output voltage, a load resistor must be added in series with the collector of the transistor. Figure 54 shows the transistor when converted to a simple voltage amplifier.



V_{IN} = Input Voltage (AC)
V_{OUT}= Output signal (AC)
V_{CC}= Collector Supply (DC)

Figure 54 Simple Voltage Amplifier

The base-collector junction is already very high resistance (Kilo-Ohms) because it operates in reverse bias (with respect to the base, as described earlier). That means that the addition of a few hundred Ohms' resistance will have little overall effect on the Collector current.

If an AC signal voltage is applied, it will cause an AC signal current at the Collector of the transistor. The addition of a series resistor will convert that Collector current to a Collector voltage.

The application of a load will cause the output voltage (V_{OUT}) to appear in opposite polarity to the input voltage (V_{IN}). The output signal is shown in comparison to the input signal is shown in figure 55.

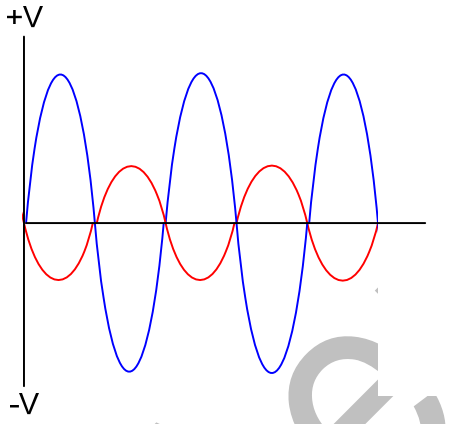


Figure 55 V_{OUT} Compared to V_{IN}

It should be noted that figure 55 is for illustrative purposes. The signal 'V_{OUT}' will be much larger in comparison than the signal 'V_{IN}'. This is due to the gain of the transistor. There will also be a DC offset to V_{OUT} determined by the DC voltage drop across R_L. It should be noted that the voltage drop across R_L is calculated using ohm's law ($V_{RL} = I_C \times R_L$).

DC Bias Applied to a Voltage Amplifier

A potential divider network is used to control the bias to the Base of the transistor. This bias allows current to flow into the Base and generate current carriers to the Emitter. Figure 56 shows the divider-controlled voltage amplifier.

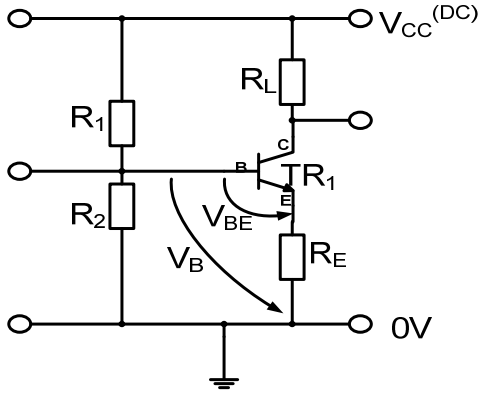


Figure 56 Potential Divider Network

The use of the PD network ensures that two supplies are not required to create the Collector supply and the Base bias. (The Base-Collector region requires a supply. The Base-Emitter also requires a supply). This would not be practical for circuits with multiple transistors.

R_1 and R_2

These resistors are used to form a potential divider network from the main supply rail. They create the bias voltage to the Base of the transistor.

The Base voltage (V_B) is the same as the voltage across V_{R2}. The values of resistors R_1 and R_2 determine V_{R2} and V_B. It should be noted that Base voltage (V_B) is not always equal to V_{BE}. V_{BE} is the Base-Emitter voltage. V_{BE} is fixed at 0.7V.

The Base voltage (V_B and hence V_{R2}) can be calculated using equation 4a. Figure 57 proves that the ratio of resistor values controls the Base voltage.

$$V_B = \frac{R_2}{R_1 + R_2} \times V_{CC}$$

Equation 4a: Calculating V_B and V_{R2}

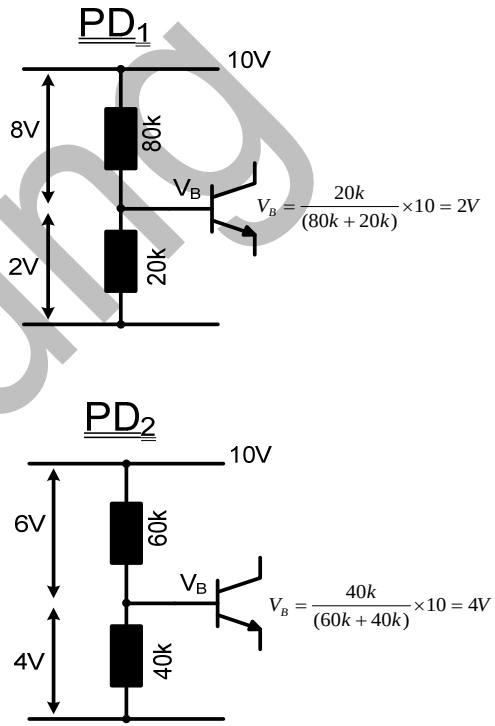


Figure 57 PD Ratio Effect on V_B

The diagram PD_1 shows the Base voltage (V_B) with resistor values of 80kΩ (R_1) and 20kΩ (R_2). The DC supply (V_{CC}) is 10V. The voltage across R_1 is shown to be 8V. The voltage across R_2 (therefore V_B) is 2V.

The diagram PD_2 shows the Base voltage (V_B) when the resistor values of the potential divider have been changed to 60kΩ (R_1) and 40kΩ (R_2). The Base voltage is now shown to be 4V.

When the exact voltage drop across R_1 is known, the Base current can be calculated using ohm's law. Equation 4b shows how to calculate the current into the Base (I_B).

$$I_B = I_{R1} - I_{R2}$$

Equation 4b: Calculating Base Current (I_B)

Emitter Resistor (R_E) (Temperature stabilisation)

The resistor R_E is used to prevent the Emitter from over-heating. This prevents thermal runaway. Figure 58 shows the Emitter resistor.

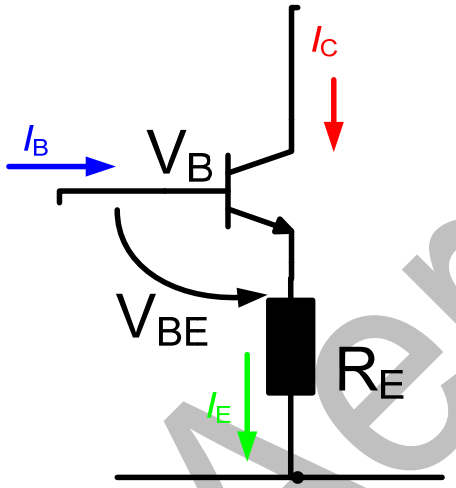


Figure 58 Emitter Resistor (R_E)

The voltage across R_E can be calculated using V_B and V_{BE} (fixed at 0.7V). Equation 4c shows the method of calculating the Emitter voltage (V_{RE}).

$$V_{RE} = V_B - V_{BE}$$

Equation 4c: Emitter Voltage (V_{RE})

The Emitter current (I_E) can be calculated (from ohm's law) if the Base and Collector currents (I_B and I_C) are not known. This method is shown in equation 4d.

$$I_E = \frac{V_{RE}}{R_E}$$

Equation 4d: Emitter Current (I_E)

It can also be seen that Base current (µA) is negligible when compared to Collector current (mA). For analysis purposes, it can be assumed that Emitter current (I_E) is approximately equal to Collector current (I_C). Equation 4e shows the method to determine the approximate collector current.

$$I_E \approx I_C$$

Equation 4e: Approximate Collector Current

Collector Characteristics

The simple amplifier circuit can be used to test for Collector characteristics. The objective is to show how the Collector current (I_C) varies with changes to the Collector-Emitter voltage (V_{CE}).

These changes in V_{CE} can be conducted for fixed levels of Base current, I_B. This is shown in figure 59.

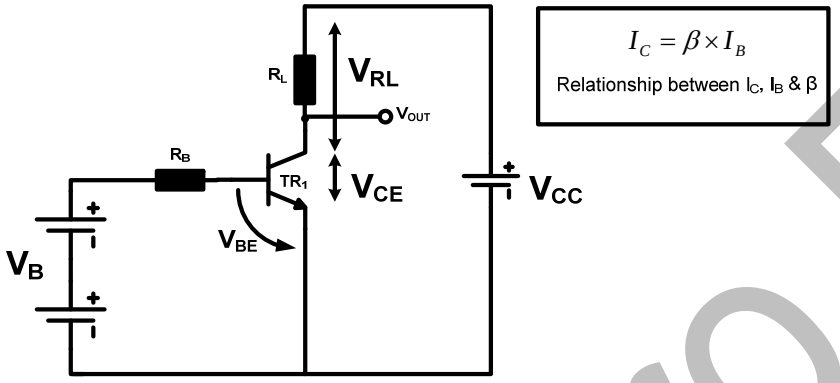


Figure 59 Simple Amplifier with Fixed Base Current

If the Base voltage (V_B) is fixed to generate a specific level of Base current and the Collector supply (V_{CC}) is fixed to zero, the Base-Emitter junction is forward biased (0.7V). The Emitter and Collector are both 0V, meaning the Collector is now forward biased with respect to the Base (this is because the Base is at a higher potential than the Collector).

The Base current conducts through the Base-Emitter junction (B-E) due to the low resistance path to ground. The Collector is 0V, causing I_C to be zero.

When both junctions are forward biased, the transistor is in the 'Saturation' region. This is shown on point A→B on figure 60.

If V_{CC} increases, the Collector current increases. As a consequence, V_{CE} also gradually increases. The Base-Collector junction is still forward biased because V_{CE} is less than the Base-Emitter voltage (V_{BE} = 0.7V).

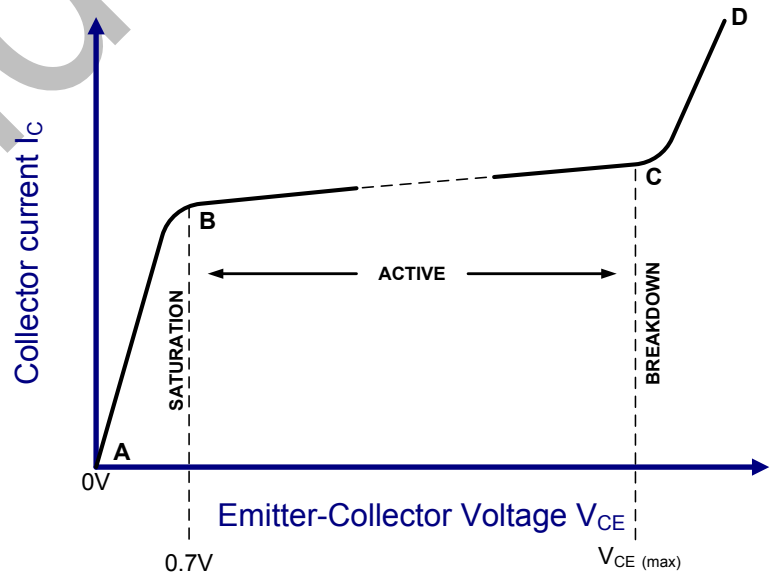


Figure 60 I_C Vs. V_{CE} for a Single Test Value of I_B

Under ideal conditions, the Transistor will enter the 'Active' operation (also known as linear) when V_{CE} exceeds 0.7V. This is because the B-C junction has become reverse biased (The Base is now at a lower potential than the Collector).

Once the B-C junction becomes reverse biased, the Collector current can remain relatively constant, even as V_{CE} continues to increase. I_C will remain constant provided I_B and the load resistor R_L do not change.

The Collector current will not be exactly constant. It will actually increase slightly with an increase in V_{CE}.

This occurs because the Base-Collector depletion region widens (due to normal reverse bias operation). There are fewer holes in the Base for electron-hole pairs. This causes an increase in the Collector/Base gain (B). Equation 1 (shown earlier) shows the relationship between Collector and Base currents. The 'Active' region is shown between points B→C on figure 60. This region is used to carry out transistor amplification.

Once the V_{CE} voltage reaches a sufficiently high value, the reverse biased B-C junction begins to breakdown. The Collector current rapidly increases. This is shown between points C→D on figure 60. Operation in the 'Breakdown' region would cause catastrophic destruction to the transistor.

The collector current characteristic curves can be plotted for many different values of I_B. This is illustrated in figure 61.

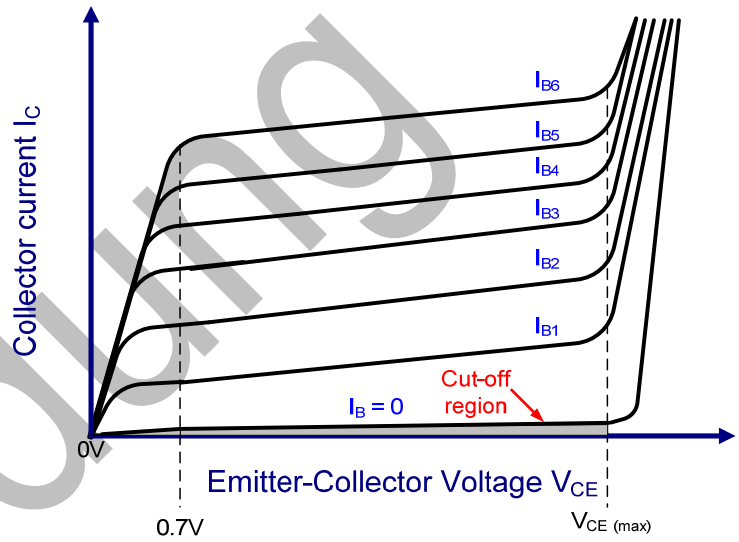


Figure 61 I_C Vs. V_{CE} for Multiple Values of I_B

Cut-off Considerations

The transistor is operating in the Cut-off region when $I_B = 0$. When this occurs no real amplification can take place. It should be noted that a negligible leakage current is generated from the Collector. This is mainly caused by thermally produced carriers. This value can be known as I_{CEO}. This has been exaggerated in figure 61 for illustrative purposes. Under Cut-off conditions, both the B-E junction and the B-C junction are reverse biased.

The Cut-off value is defined as the 'non-conducting' state of a transistor.

Saturation Considerations

Saturation is the state of a Bi-polar junction transistor in which the Collector current has reached a maximum value and is independent of a further increase in Base current.

Under Saturation conditions, I_C increases as I_B increases. The voltage drop across R_L increases, causing V_{CE} to decrease.

V_{CE} eventually drops to the saturation value, known as $V_{CE(SAT)}$. At this point, I_C cannot increase further, even with an increase in I_B. The relationship between I_C, I_B and B (equation 1) becomes invalid. $V_{CE(SAT)}$ usually occurs below the 'knee' of the Collector curve. It is a negligible value (assumed to be 0V).

Quiescence (Q-Point)

The Q-point is used to determine the I_C and V_{CE} values for any given Base current.

The Base current is established using the Base voltage as a bias. A 'DC load-line' can be drawn onto the transistor characteristic curve. It is described as the point where the Base current curve intersects the DC load line. The co-ordinates of the Q-point are the values of I_C and V_{CE}. This is shown in figure 62.

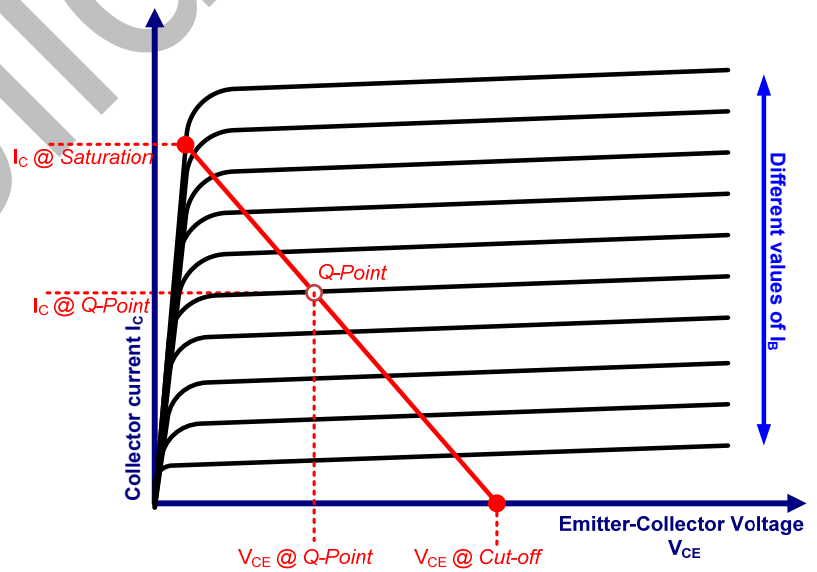


Figure 62 Quiescent Point Using a Load-Line

The load-line is determined by R_L and V_{CC}, not the transistor in use. A straight line is drawn between the Saturation and Cut-off regions of the Collector curve. Any value of I_C and its V_{CE} value will operate on this line.

Example:

When the transistor operates in the Cut-off region, there is negligible Collector current (Assumed to be zero). This means that V_{CE} becomes the same value as the supply voltage (V_{CC}). The term V_{CC} actually stands for Collector @ Cut-off (**Voltage Collector** at **Cut-off)**. For this example, V_{CE} becomes 30V.

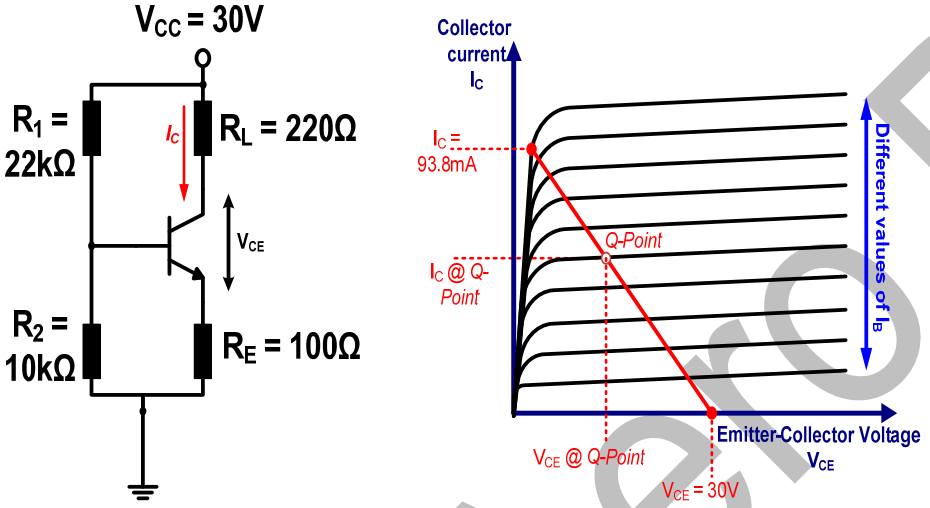


Figure 63 Example Q-Point

The Saturation point can also be calculated. It should be noted that V_{CE} is assumed to be 0V. This means that the full V_{CC} drop is across R_L + R_E.

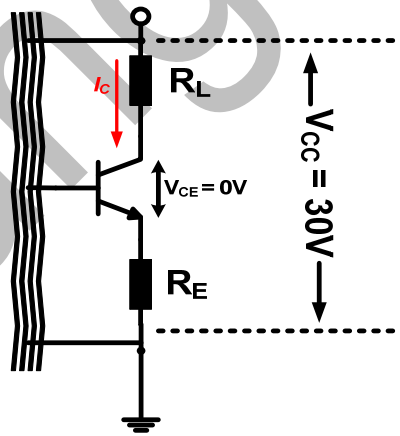


Figure 64 Example Transistor Output Voltages

The Saturation Collector current ($I_{C(SAT)}$) can be calculated using Ohm's law.

$$I_{C(SAT)} = \frac{V_{CC}}{(R_L + R_E)} = \frac{30V}{(220\Omega + 100\Omega)} = 93.8mA$$

This is the maximum possible value of Collector current. It cannot be increased further without changing the supply (V_{CC}), R_L or R_E. A straight line can now be drawn between V_{CE} =30V and I_C = 93.8mA. The Q-point will occur wherever the load-line intersects the curve for a selected value of Base current.

AC Signal Operation

The DC bias conditions (described above) must be set-up to allow an AC signal to be applied to the Transistor. The AC signals are identified using lower-case letters, eg. I_c to describe AC Collector current.

The AC signal amplifier is shown in figure 65. There are a few adjustments made to allow the action of AC signal amplification.

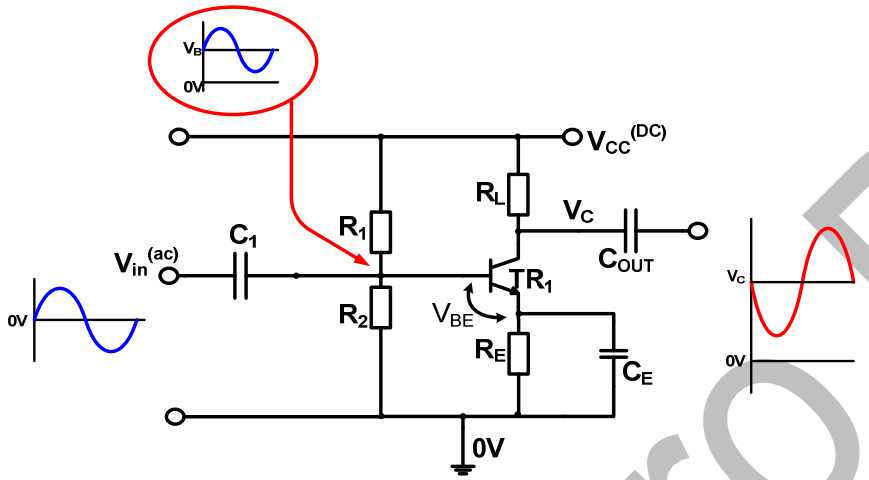


Figure 65 AC Voltage Amplifier

Capacitors C_1 and C_{OUT} have been added to the input of the Base and Collector output (V_C) respectively. These have been added to block any external DC bias from interfering with the circuit. This is more important when amplifiers are cascaded together. These are known as 'Coupling capacitors'.

The capacitor, C_E, has been added to provide an alternative path for the Emitter AC signal. The current through R_E must remain constant to maintain the DC bias ($I_E \approx I_C$). If an AC signal is allowed through R_E, the current will change, therefore changing the B-E voltage. V_{BE} must be at least 0.7V to keep the transistor switched 'ON'.

The applied AC signal causes the Base current to vary at the same frequency above and below its DC bias value. The variation of Base current causes a variation in Collector current.

The gain of the transistor causes the variation in Collector current to be larger than the variation in Base current. The ratio of AC Collector current, I_c, to the AC Base current, I_b, is shown as $ß_{ac}$ or hfe. This is shown in equation 4f.

$$\beta_{ac} = \frac{I_c}{I_b}$$

Equation 4f: AC Ratio of Collector Current to Base Current

The relationship is the same in AC conditions as shown previously in DC conditions (shown in equation 1). However, the value of β_{ac} can differ slightly from B (DC conditions).

The variation in Collector current causes the voltage drop across R_L to vary also. This will affect the Collector voltage, V_C (used as the output, shown in figure 65).

As I_C increases, the voltage drop across R_L increases also. This causes the output at V_C to decrease. This occurs because

$$V_C = V_{CC} - R_L .$$

As the Collector current decreases, the voltage drop across R_L decreases. This causes an increase in V_C at the output.

This means that there is a 180^O phase difference between collector current and collector voltage. Therefore the Base voltage, V_b, and Collector voltage V_c are opposite in polarity. This is known as 'Anti-phase' or 'Inversion'. This has been illustrated in figure 55 (earlier). Figure 66 shows the output of the AC signal amplifier when compared to the input (if viewed on an oscilloscope).

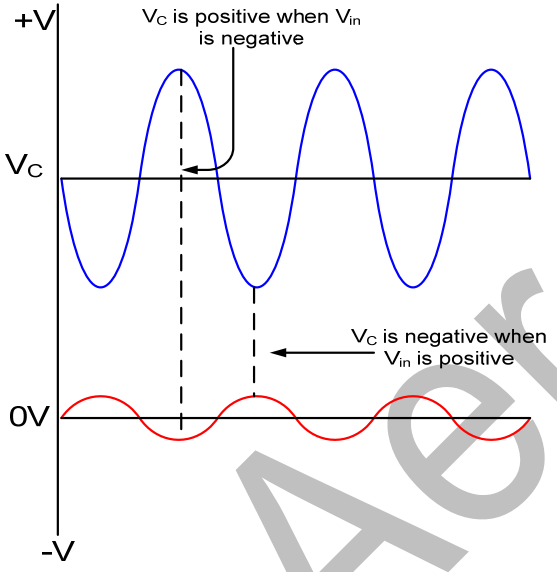


Figure 66 V_{OUT} Compared to V_{IN}

Gain Ratio on an AC Amplifier

As shown earlier in equation 3a, the gain ratio, G, is $\frac{V_{out}}{V_{in}}$. In this case, V_{in} is the AC voltage at the input to C_1 and the Base. The AC Collector voltage, V_c and C_{OUT} are defined as the output, V_{out}.

The B-E junction is forward biased, therefore the signal operation at the Emitter is approximately equal to the signal operation at the Base.

Since $V_b \approx V_e$, the gain is approximately $\frac{V_c}{V_e} = \frac{(I_c \times R_L)}{(I_e \times R_E)}$. As with the DC properties, I_e is approximately equal to I_c, so they cancel out to leave the method of calculating the gain. This is shown in equation 4g.

$$G \approx \frac{R_L}{R_E}$$

Equation 4g: Gain in an AC-Amplifier

Example

A 50mV rms signal voltage is applied to the Base.

(a) Determine the AC output signal voltage from the amplifier.

(b) Determine the DC Collector offset that the output AC signal is varying.

(c) Show the output waveform.

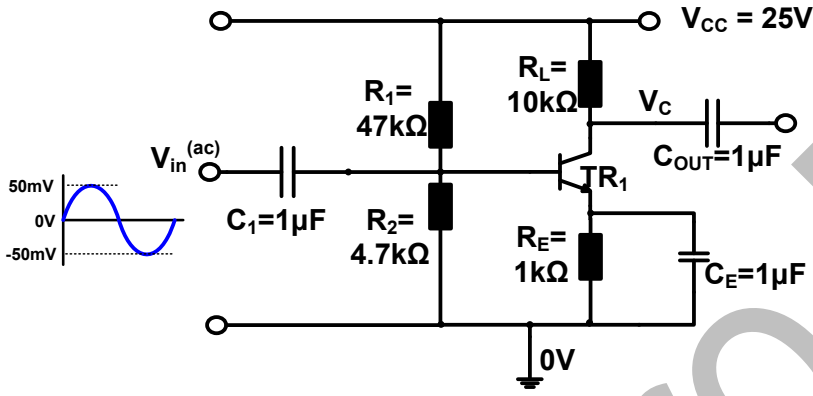


Figure 67 Emitter Amplification Circuit

(a) The signal gain ratio can used to determine the output signal voltage.

$$Gain = \frac{R_L}{R_E} = \frac{10k\Omega}{1k\Omega} = 10$$

From equation 3a, $Gain = \frac{V_{OUT}}{V_{IN}}$

This can be transposed mathematically so that;

$$V_{OUT} = GAIN \times V_{IN} = (10) \times (50mV) = 500mV(rms)$$

The AC output signal voltage is 500mV (rms)

(b) The Collector DC offset can now be calculated using the potential divider network equation (shown in equation 4a).

$$V_B = \frac{R_2}{R_1 + R_2} \times V_{CC} = \frac{4.7k\Omega}{51.7k\Omega} \times 25V = 2.27V$$

This voltage can now be used to calculate the Emitter voltage;

$$V_E = V_B - V_{BE} = 2.27V - 0.7V = 1.57V$$

This voltage can now be used to calculate the Emitter current (which is approximately equal to Collector current, shown in Equation 4e);

$$I_C \approx I_E = \frac{V_E}{R_E} = \frac{1.57V}{1k\Omega} = 1.57mA$$

This can now be used to calculate the voltage drop across R_L.

$$V_{RL} = I_C \times R_L = (1.57mA) \times (10k\Omega) = 15.7V$$

The voltage drop across R_L can be used to calculate the DC offset at the Collector, V_C.

$$V_C = V_{CC} - V_{RL} = 25V - 15.7V = 9.3V$$

The DC offset at the Collector is 9.3V.

(c) To show the $500mV_{(rms)}$ output waveform riding the 9.3V DC Collector voltage, the peak voltage of the output must be calculated.

The method to calculate the RMS voltage from an ac signal is shown below;

$$V_{rms} = V_{Peak} \times 0.7071$$

Using mathematical transposition, the peak signal can be determined;

$$V_{Peak} = \frac{V_{rms}}{0.7071}$$

The peak ac value can now be calculated;

$$VPeak = \frac{500mVrms}{0.7071} = 707mV$$

The output signal will alternate above and below 9.3V by 0.707V.

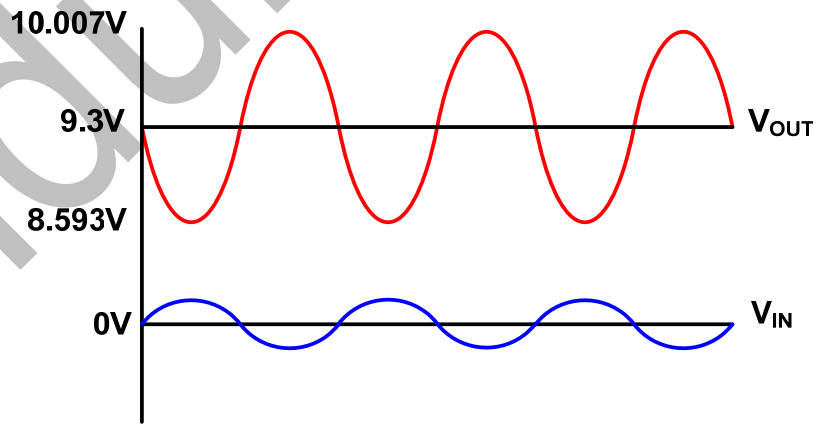


Figure 68 Signal Amplification

AC Signal Operation Applied to the Load Line

The load-line can also be used to assess an AC signal. As before, the Q-point can be found from the DC Collector characteristic curves. Once the Q-point is found, the variations of Base current can be plotted.

Example:

The Q-Point DC values are set to I_B = 40µA, I_C = 4mA and V_{CE} = 8V. The AC signal varies I_B to 50µA maximum and 30µA minimum. The corresponding values for I_C and V_{CE} can be seen on the Collector curve.

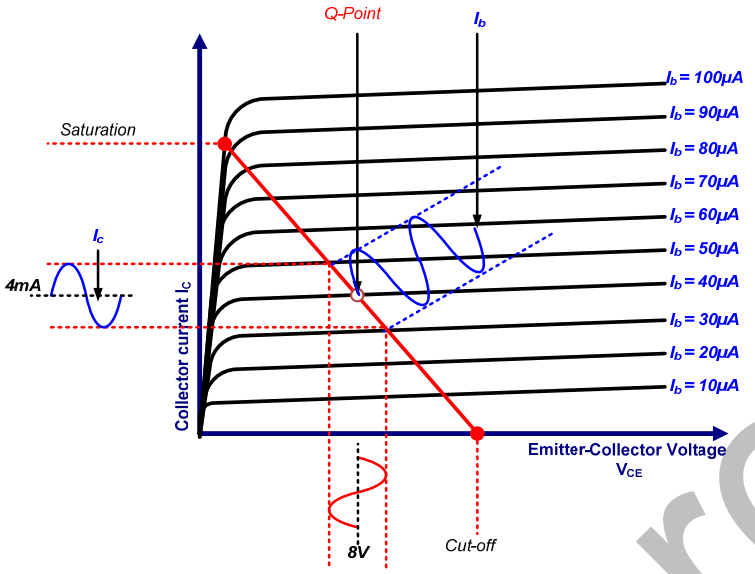


Figure 69 Transistor Clipping

The amplifier will operate as described above provided the variations do not reach the Cut-off or Saturation regions. This will occur if the Base current becomes too large. The Collector current (and output voltage) will show 'clipping' as the amplifier is above its maximum operating limits. This is shown in figure 70.

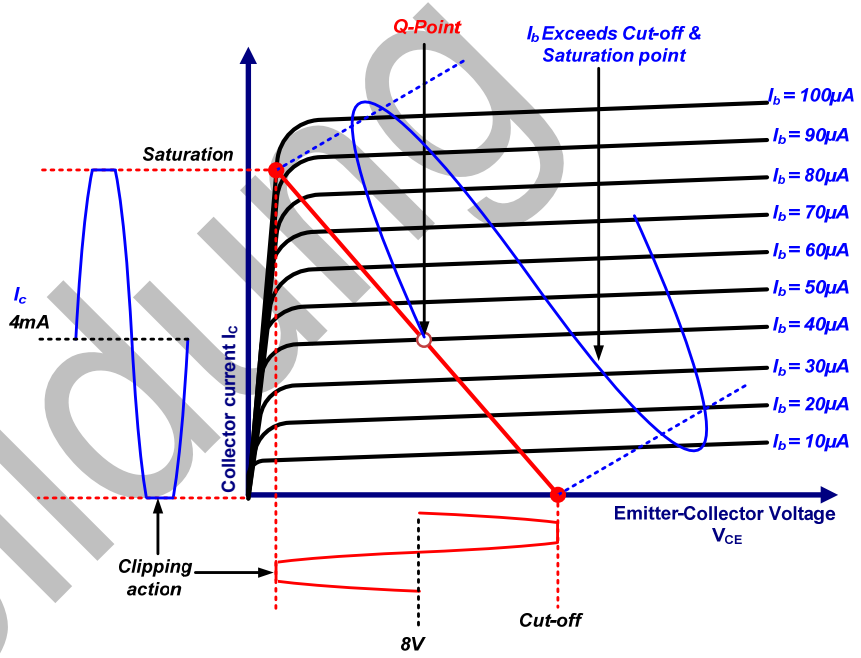


Figure 70 Clipping at Saturation and Cut-off

1.3.5 The Transistor as a Switch

The Saturation and Cut-off regions are useful when the Transistor is operated as an electronic switch. When operating in the Cut-off region, the B-E junction is not forward-biased (The V_{BE} is below 0.7V). There is effectively an open circuit across the Collector and Emitter. This is shown in figure 71.

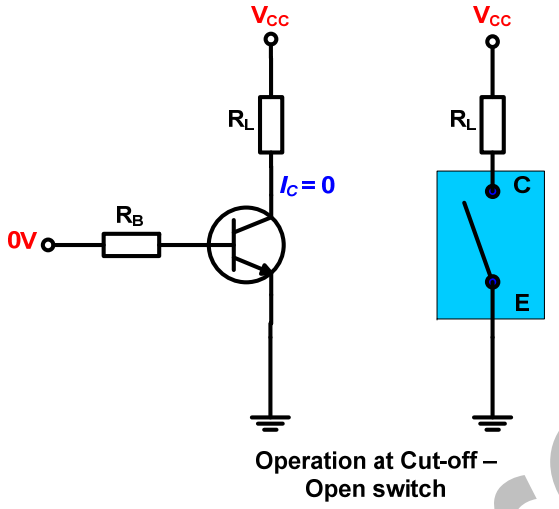


Figure 71 Transistor in Cut-off

In this condition, all currents are approximately zero (with the exception of carrier leakage). This means that $V_{CE} \approx V_{CC}$. The transistor is effectively switched 'OFF'.

When operating in the Saturation region, the B-E junction and the B-C junction are forward biased. The Base current is large enough to allow the Collector current to reach its saturated value. This is shown in figure 72.

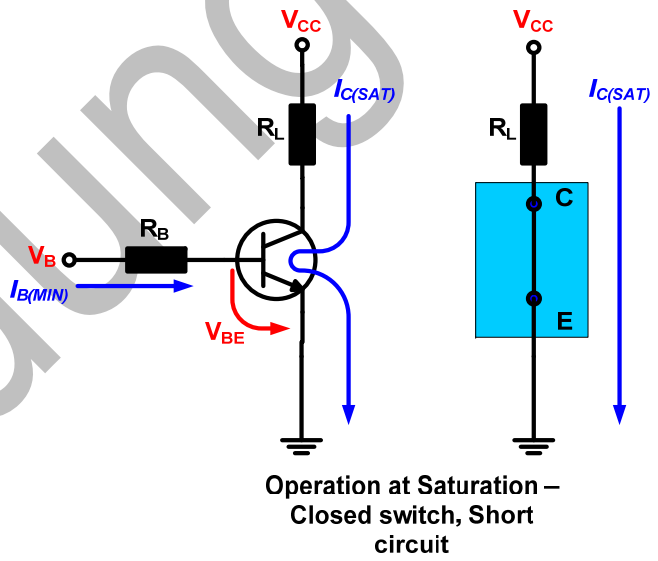


Figure 72 Transistor in Saturation

In this condition there is effectively a short circuit between the Collector and the Emitter. There is a small voltage drop caused by $V_{CE(SAT)}$. This is assumed to be zero for analytical purposes.

As shown earlier, the supply, V_{CC}, is connected to ground due to the short-circuit status across the transistor. All currents are present and the transistor is 'ON'.

The Collector current at saturation can be calculated as the V_{CC} drop is across R_L. This is shown in equation 5a.

$$I_{C(SAT)} = \frac{V_{CC}}{R_L}$$

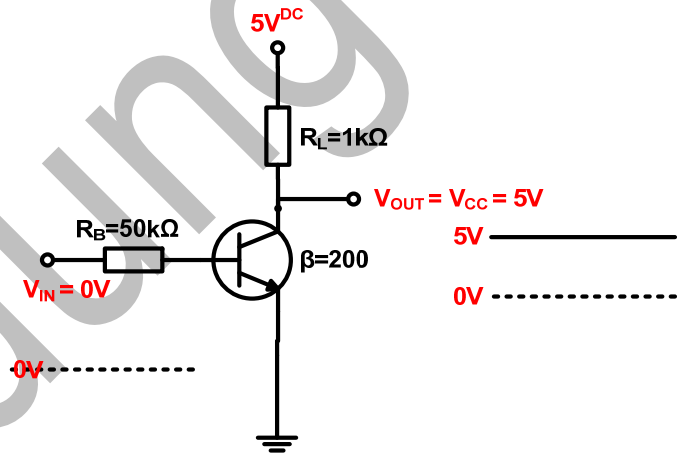
Equation 5a: Saturation Current

The relationship between I_C, I_B and B can also be transposed to calculate the minimum required Base current to operate the transistor in Saturation. This is shown in equation 5b.

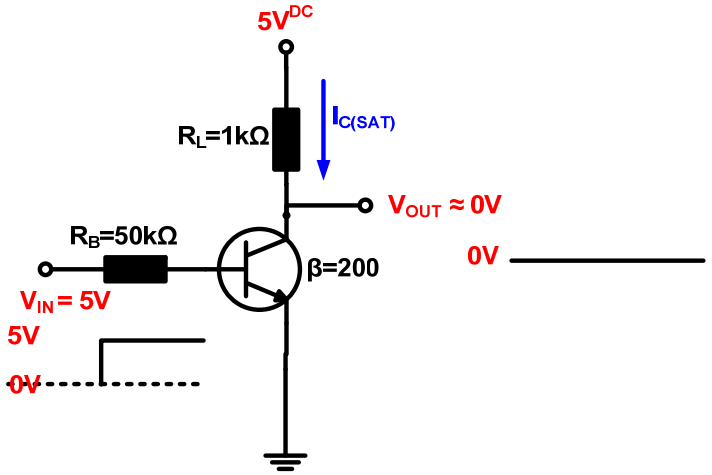
$$I_{B(MIN)} = \frac{I_C}{B}$$

Equation 5b: Minimum Required Base Current to Produce Saturation

Example:



When V_{IN} = 0V, the Base current, I_B =0A. The Collector current, I_C = 0A, so there is no voltage drop across R_L. This means that V_{CC} can found at V_{OUT}.



When V_{IN} rises to 5V, the Base current can be calculated as:

$$I_B = \frac{5V}{50k\Omega} = 100\mu A$$

If the relationship between I_C, I_B and B were valid, the Collector current would be expected to be:

$$I_C = B \times I_B = 200 \times 100\mu A = 20mA$$

If ohm's law were to be used to calculate the output voltage, it would be found that:

$$V_{OUT} = I_C \times R_L = 20mA \times 1k\Omega = 20V$$

This is not possible because V_{CC} is only 5V. The transistor is trying to amplify based on the gain of the transistor, β and the Base current.

The transistor is acting in Saturation. This means that the relationship between I_C, I_B and B is no longer valid.

As the gain relationship is not valid, the saturated Collector current is calculated using R_L;

$$I_{C(SAT)} = \frac{V_{CC}}{R_L} = \frac{5V}{1k\Omega} = 5mA$$

The supply is dropped almost completely across R_L; this means that the output voltage is approximately 0V. It will actually be about 100mV due to internal transistor impedance. This operation can be used in digital logic circuits as an inverter.

1.3.6 Coupled Amplifiers

There are three main types of coupled amplifiers using the Bi-polar junction transistor;

- The Common-Emitter amplifier.
- The Common-Collector amplifier.
- The Common-Base amplifier.

1.3.7 Common-Emitter Amplifiers

The Common-Emitter (CE) configuration is a Bi-polar junction transistor (BJT) arrangement where the Emitter is the common terminal.

Common-Emitter amplifiers can be compared to other amplifiers using the following criteria:

- Voltage gain.
- Current gain.
- Power gain.
- Output phase (with respect to input phase).
- Input resistance.

The input and output both have an Emitter connection. This is why the amplifier is known as 'Common Emitter'.

The input is connected from the Base to the Emitter. The output is connected across the Collector and Emitter.

Voltage Gain

The Common-Emitter amplifier has already been covered to illustrate basic transistor operation. More detail is required to describe the operation of the Emitter Capacitor, C_E.

The addition of the capacitor, C_E, does not affect the DC Emitter voltage. This is because the DC signal at the Emitter will view C_E as an open circuit.

As stated earlier, the capacitor is added to provide a path for an AC signal. This protects the DC bias across the B-E junction (resistor R_E) and ensures the transistor remains 'ON'. This is shown in figure 73.

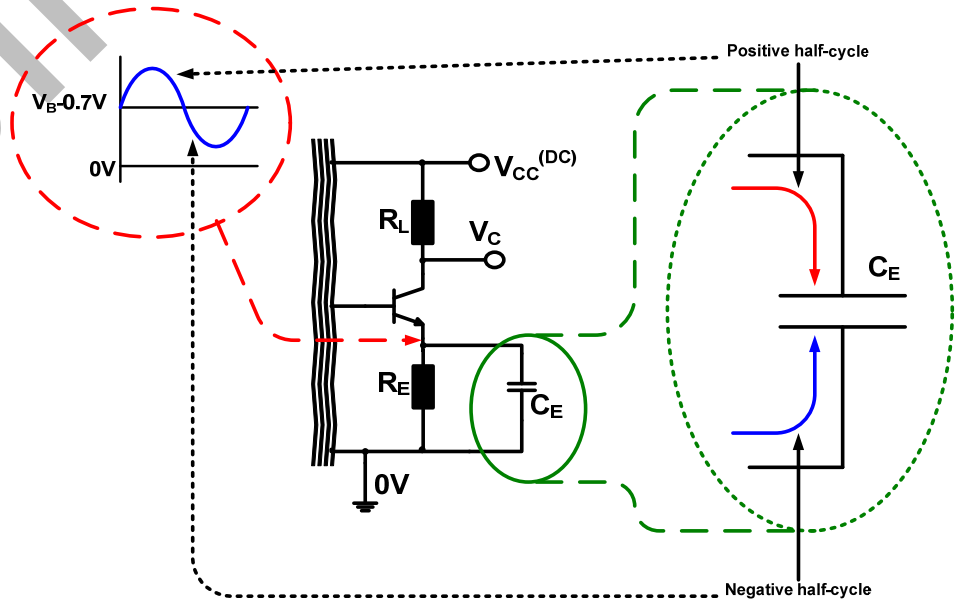


Figure 73 Emitter Bypass Capacitor

The capacitor construction is known as a 'Bypass Capacitor'. The Bypass capacitor effectively shorts the AC signal to ground. The Emitter is operating at AC ground level (in relation to the AC signal) but not DC ground (It should be noted that the AC input signal varies around V_B, as shown earlier).

Equation 6a shows the more detailed method of calculating gain based on resistance (Compared to equation 4g). Without the bypass capacitor, the internal Emitter resistance, r_e is significantly less than R_E, so it could be ignored (as shown in equation 4g).

$$Gain = \frac{R_L}{r_e + R_E}$$

Equation 6a: Gain without Bypass Capacitor

The Bypass capacitor increases voltage gain. Once the Bypass capacitor is connected across R_E, it provides an alternative route for the AC signal. This action means that R_E can now be ignored and r_e must be considered because it is the only factor remaining. Equation 6b shows the same formula without R_E.

$$Gain = \frac{R_L}{r_e}$$

Equation 6b: Gain with Bypass Capacitor Connected

The value of r_e is dependant on the Emitter current and the fixed thermal voltage produced by normal transistor operation. This value is fixed at 25mV. Equation 6c shows the method of calculating r_e.

$$r_e = \frac{25mV}{I_E}$$

Equation 6c: Method of Calculating R_e

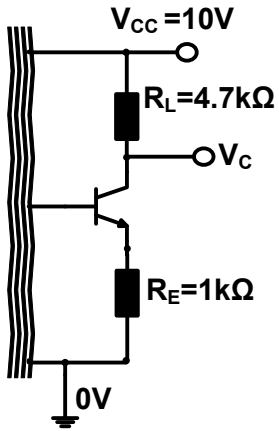
Example:

A voltage amplifier is operating without a bypass capacitor. Express the gain in dB (include the value of r_e for illustrative purposes, even though it is small when compared to R_E).

- V_B = 2V
- V_{BE} = 0.7V

If a bypass capacitor is connected across R_E, express the new dB gain.

What is the difference in gain with a bypass capacitor connected?



The value of V_E must be found;

$$V_E = V_B - V_{BE} = 2V - 0.7V = 1.3V$$

Now the value of I_E can be calculated;

$$I_E = \frac{V_E}{R_E} = \frac{1.3V}{1k\Omega} = 1.3mA$$

The value of r_e can now be determined using the transistor thermal operating Voltage.

$$r_e = \frac{25mV}{I_E} = \frac{25mV}{1.3mA} = 19.23\Omega$$

The voltage gain can now be calculated using:

- R_E = 1kΩ
- R_e = 19.23Ω

It should be noted that R_E is significantly greater than r_e. This would mean that r_e could be ignored during the calculation. For the purpose of illustration it will be included.

$$Gain = \frac{R_L}{(r_e + R_E)} = \frac{4.7k\Omega}{(19.23\Omega + 1000\Omega)} = 4.611$$

This value is to be expressed in dB;

$$A_v(dB) = 20\log_{10}(4.611) = 13.28dB$$

The Gain without a Bypass capacitor is 13.28dB.

If a Bypass capacitor is connected across R_E, the new gain can be calculated. The value of R_E can be ignored.

$$Gain = \frac{R_L}{r_e} = \frac{4.7k\Omega}{19.23\Omega} = 244$$

This value is to be expressed in dB;

$$A_v(dB) = 20\log_{10}(244) = 47.8dB$$

The gain with a bypass capacitor attached is 47.8dB.

- Gain with bypass capacitor = 47.8dB.
- Gain without bypass capacitor = 13.28dB.

$$A_v = A_{v(with)} - A_{v(without)} = 47.8dB - 13.28dB = 34.5dB$$

The addition of a bypass capacitor improves the voltage gain by 34.5dB.

Current Gain

If the bias resistors R_1 & R_2 are large enough, the current through them is negligible. This means that the input supply current should be approximately equal to the Base current. The Current gain is simply $\frac{I_c}{I_b}$ as stated earlier.

Power Gain

The power gain of a Common-Emitter amplifier is the product of the Voltage gain and the current gain. This is shown in equation 6d.

$$A_p = A_v \times A_i$$

Equation 6d: Power Gain in a Common-Emitter Amplifier

This means that the Common-Emitter amplifier is a Voltage amplifier and a Current amplifier.

Phase Inversion

The Common-Emitter amplifier is also characterised by the phase inversion at the output when compared to the input. This is known as 180^{O} phase difference.

Input Resistance

The DC input resistance is the Potential Divider network (shown earlier). The input resistance (viewed from the perspective of the ac input at the Base) is shown in figure 74.

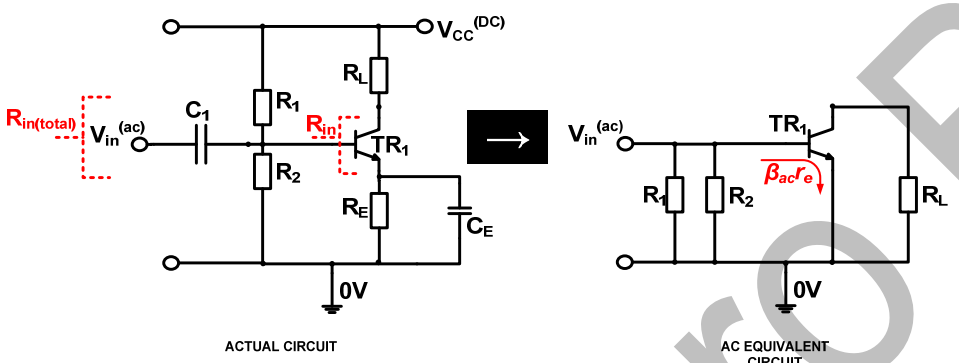


Figure 74 Input Resistance and AC Equivalent Circuit

The diagram shown in figure 74 is true for a Voltage Amplifier with a Bypass capacitor connected. For this reason, R_E is not included in the AC equivalent circuit.

It should be noted that from the perspective of the AC signal applied the Base, R_1 & R_2 are taking separate paths. Therefore the AC signal sees them in parallel. The same is also true for the Emitter output and R_L.

The input resistance can be derived from Ohm's law and is shown in equation 6e.

$$R_{in} = \frac{V_b}{I_b}$$

Note that: $V_b = I_e \times r_e$

And: $I_e \cong \beta_{ac} \times I_b$

The new values of I_e and V_b can be substituted so that:

$$R_{in} = \frac{\beta_{ac} \times I_b \times r_e}{I_b}$$

The values of I_b cancel out to leave:

$$R_{in} = \beta_{ac} \times r_e$$

Equation 6e: Input Resistance as seen from the AC Base Input Signal

From the perspective of the transistor Base, R_{in} is the input resistance. The actual total resistance, $R_{in(total)}$, includes the bias resistors (R_1 & R_2). This is shown in figure 74.

The AC signal will also see the DC supply, V_{CC}, as a path to ground. This occurs because there is no AC signal at the V_{CC} supply terminal. Therefore V_{CC} will act as AC ground.

This means that R_1 and R_2 appear to be in parallel. R_2 is connected to actual ground and R_1 is connected to AC ground (V_{CC} terminal). R_{in} will also be seen in parallel with $R_1||R_2$. The expression for the total resistance viewed from the perspective of the AC input signal is shown in equation 6f. The Collector resistance, R_L, does not have an effect on the input resistance if the circuit.

$$R_{in(total)} = R_1 \| R_2 \| R_{in}$$

Equation 6f: Total Input Resistance of the AC Equivalent Circuit

1.3.8 The Common-Collector Amplfier

The Common-Collector (CC) arrangement is a BJT configuration where the Collector is the common terminal.

The Common-Collector Amplifier is also known as an 'Emitter-Follower'. There is no load resistor at the Collector of the transistor.

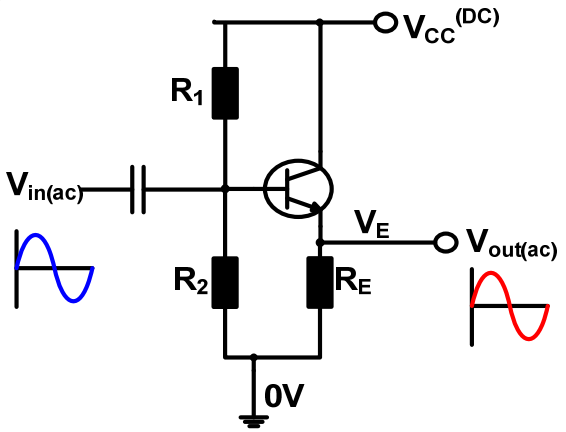


Figure 75 Common Collector Amplifier

The input is connected between the Base and Collector. The output is connected between the Collector and Emitter. A Common–Collector circuit is shown in figure 75.

Voltage Gain

As previously discussed, the Voltage gain is $A_v = \frac{V_{out}}{V_{in}}$. The voltage gain is derived in equation 6g.

Note that: $V_{out} = I_e \times R_E$

And: $V_{in} = I_e \times (r_e + R_E)$

Therefore the gain is:

$$A_v = \frac{V_{out}}{V_{in}} = \frac{I_e \times R_E}{I_e \times (r_e + R_E)}$$

The Emitter current, I_e, will cancel out to leave:

$$A_v = \frac{V_{out}}{V_{in}} = \frac{R_E}{r_e + R_E}$$

Equation 6g: Voltage Gain of an Emitter Follower

The inclusion of r_e into the calculation means that the gain will always be less than 1. However, because r_e is always significantly less than R_E, the gain is considered to be approximately 1.

Current Gain

The current gain of the Emitter-follower is identical to the Common-Emitter amplifier, assuming the bias resistors are large enough. The Current Gain is considered to be $A_i \approx \frac{I_c}{I_b}$.

Power Gain

The power gain is considered to be the product of the voltage gain and the current gain. In the case of the Emitter follower, the voltage gain is approximately 1. Therefore the power gain is approximately equal to the current gain. The Power gain is considered to be $A_p \approx A_i$.

Output phase

The output is in phase with the input. This is because it is taken from the Emitter (which is in-phase with the Base and the input).

Since the voltage gain is approximately 1 and the output is in phase with the input, the AC voltage signal at the output appears to be almost identical to the input. For this reason the Common-Collector amplifier is considered as an 'Emitter-Follower'.

Input resistance

The Emitter follower is seen to have a high-input resistance. This is a useful feature that allows it to be applied as a buffer to minimise loading when one circuit is driving another.

In the case of the Common-Collector amplifier, the Emitter resistor has not been connected to a Bypass capacitor. This means that R_E must be considered to determine the input resistance, R_{in}. This is shown in equation 6h.

$$R_{in} = \beta_{ac} \times (r_e + R_E)$$

Note that r_e is significantly less than R_E

$$r_e <<< R_E$$

$$R_{in} \approx \beta_{ac} \times R_E$$

Equation 6h: Common-Collector Emitter Resistance

Due to the application of the much larger R_E (compared to r_e), the Emitter follower has a much larger total input resistance. The total input resistance, $R_{in(tot)}$ is calculated using the same method as the Common-Emitter amplifier (shown in equation 6f). The Total input resistance is considered to be $R_{in(total)} = R_1 \| R_2 \| R_{in}$

1.3.9 The Darlington Pair

The Darlington pair is an Emitter-follower amplifier. As previously discussed the output of a Common-Collector amplifier appears to be approximately identical to the input. The only limiting factor of the current gain is the value of β. This is also a major factor in determining the input resistance, R_{in}. The Darlington pair is shown in figure 76.

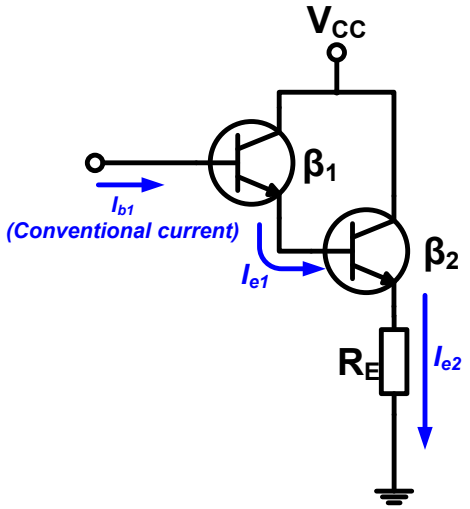


Figure 76 The Darlington Pair

The Darlington pair is used to boost the input resistance. The Collectors of both transistors are connected together. The

Emitter of the first transistor drives the Base of the second transistor.

The Emitter current of the first transistor is $I_{e1} \cong \beta_1 \times I_{b1}$. The Emitter current becomes the Base current of the second transistor, producing Emitter current from the second transistor. This is expressed as $I_{e2} \cong \beta_2 \times I_{e1} \cong \beta_2 \times \beta_1 \times I_{b1}$. The final value of the input resistance is shown in equation 6i.

$$R_{in} = \beta_1 \times \beta_2 \times R_E$$

Equation 6i: Improved Input Resistance

1.3.10 The Common-Base Amplfier

The Common-Base (CB) arrangement is a BJT configuration where the Base is the common terminal.

The Common-Base amplifier provides high voltage gain with unity current gain (current gain of 1). The C-B amplifier has low input resistance and is applied to high frequency sources with low output resistances.

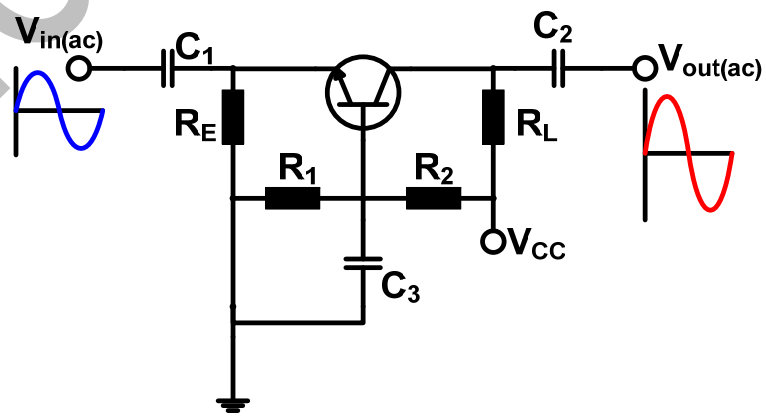


Figure 77 Common-Base Amplifier

The Common-Base amplifier is shown in figure 77. The Base is considered to be at AC signal ground. The input is between the Emitter and Base. The output is between the Base and Collector.

Voltage Gain

As stated, the input is applied to the Emitter, V_e, with the output taken from the Collector, V_c. The method used to determine the voltage gain is derived in equation 6j.

Note that $A_v = \frac{V_{out}}{V_{in}} = \frac{V_c}{V_e} = \frac{I_c \times R_L}{I_e \times r_e} \cong \frac{I_e \times R_L}{I_e \times r_e}$

The Emitter current, I_e, cancels out to leave:

$$A_v = \frac{R_L}{r_e}$$

Eq 6j: Common-Base Voltage Gain

It should be noted that this equation matches that of the Common-Emitter amplifier with a connected bypass capacitor.

Current Gain

The current gain is simply the output current, I_c, divided by the input current, I_e. The Collector current is approximately equal to the Emitter current. Therefore the current gain is always 1 (unity). The current gain is expressed as $A_i \cong 1$.

Power Gain

As stated, the current gain is ≈1 for a Common Collector amplifier. Therefore the power gain is expressed as $A_p = A_v \times 1$ $A_p \cong A_v$.

Output Phase

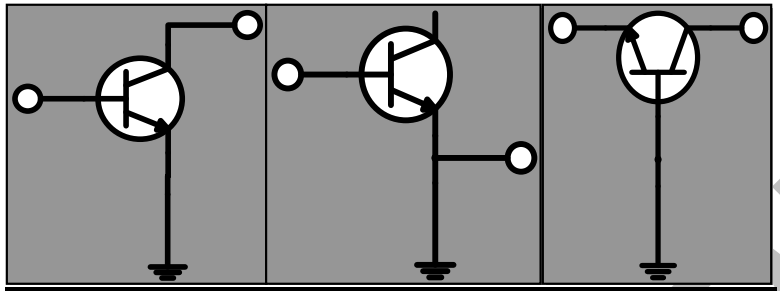
The output is in phase with the input.

Input resistance

When the input resistance, R_{in}, is taken from the perspective at the input it can be found that it is approximately r_e. This is because R_E is viewed to be in parallel with r_e. The value of R_{in} is also equal to $R_{in(total)}$. It is expressed as $R_{in(total)} = R_{in} = r_e$.

1.3.11 Coupled Amplifier Comparison

The three types of coupled amplification have been compared in figure 78.



	Common-Emitter	Common Collector	Common Base
Voltage Gain, A_v	High, R_L/r_e	Low, ≈ 1	High, R_L/r_e
Current Gain, A_i	High, β_{ac}	High, β_{ac}	Low, ≈1
Power Gain, A_p	Very High, $A_v A_i$	High, A_i	High, A_v
Input Resistance, R_{in}	Low, $\beta_{ac} r_e$	High, $\beta_{ac} R_E$	Very low, r_e
Phase of output (with respect ot input)	180^O inversion	No inversion	No inversion

Figure 78 Table of Comparisons

Multi-Stage Amplifiers

A two stage physical amplifier has been shown below in figure 79.

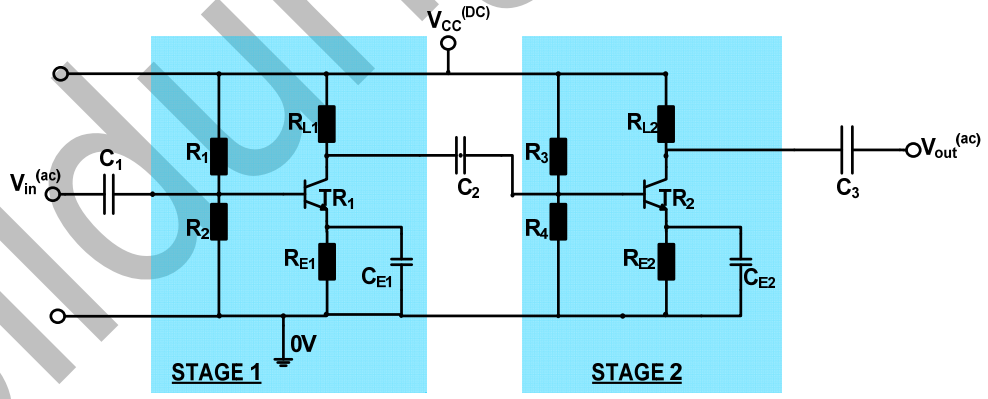


Figure 79 Two Stage C-E Amplifier

It should be noted that both stages are identical Common-emitter amplifiers. The output of the first amplifier is coupled with a capacitor to the input of the second amplifier.

Loading

When considering the gain of the first stage, the designer should be aware of the loading effect that the second stage will have on the output of the first stage.

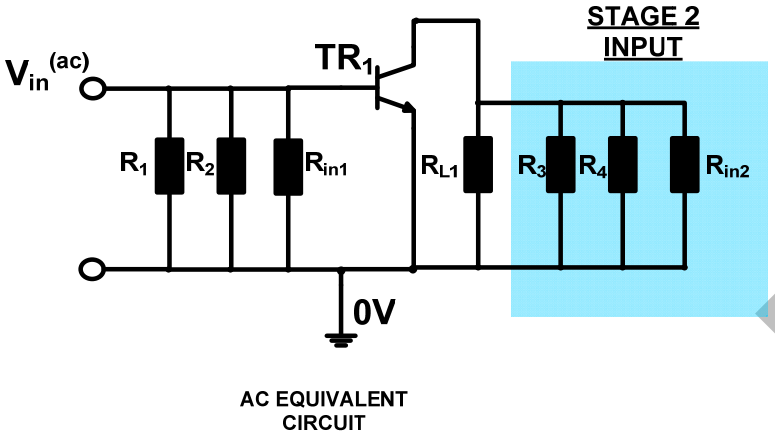


Figure 80 First Stage AC Equivalent Circuit

The coupling capacitor, C_2, behaves as a short circuit to an AC signal. This means that the total input resistance of the second stage will represent an AC load on the output of the first stage.

From the AC output perspective (of the first stage) the second stage bias resistors, R_3 & R_4, contribute to the output resistance of the first stage. The transistor Base resistance of TR_2, R_{in2}, must also be considered (as discussed in equation 79).

As shown earlier, the Collector resistor, R_L, does not have an effect on the input resistance but it does have affect the output signal. The AC equivalent circuit of the first stage has been shown in figure 80 for illustrative purposes.

1.3.12 Amplifier Classes

There are three main classes of signal amplifier.

- Class A amplifier.
- Class B amplifier.
- Class C Amplifier.

Class 'A' Operation

A 'Class A' amplifier describes the operation where the shape of an output signal is an amplified replica of its input signal.

These circuits are always biased to operate in the 'active' (linear) amplification region of the transistor characteristic curve (Illustrated in figure 70).

If an output signal does not approach the Saturation/Cut-off regions it is operating in the linear region of the transistor characteristic curve. The Saturation and Cut-off regions are the physical limits of signal amplification.

If an amplifier only uses a small percentage of these maximum limits, it is known as a 'Small-signal amplifier'.

When the output signal is larger and approaches the Saturation and Cut-off limits, it is known as a 'Large-signal amplifier'.

When the main objective is power amplification, the amplifier will be operated as a large signal amplifier. Figure 81 illustrates 'Class A' operation of an inverting amplifier.

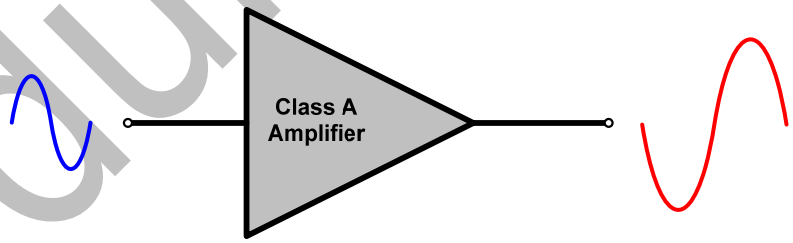


Figure 81 Class 'A' Operation

Centred Q-Point

For maximum 'Class A' operation, the Q-point (shown in figure 71) can be 'centred'. This is shown on figure 82. The Q-point can be centred using the 'DC load-line'. The Q-point can be adjusted using the Base current (as discussed earlier).

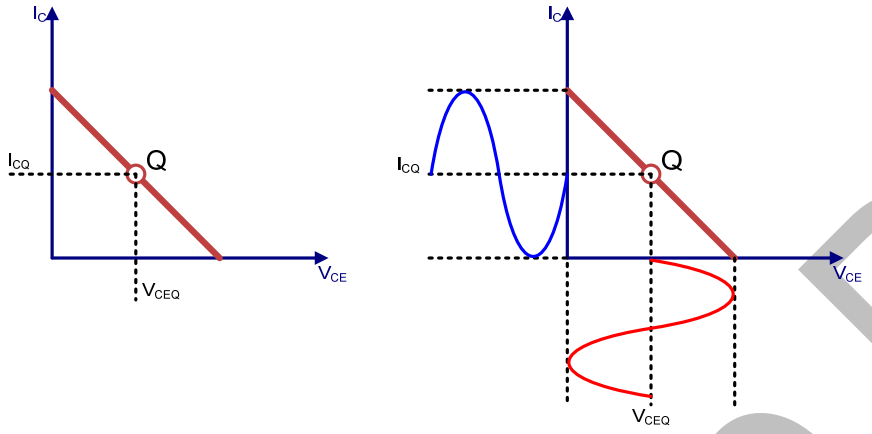


Figure 82 The 'Centred' Q-Point

The Collector current, I_C, can then alternate from its 'Q-point' value, I_{CQ}, to its maximum value (at the Saturation point, but not beyond it).

The Collector current can also alternate to its minimum value (at the Cut-off point, but not beyond it).

The Collector-Emitter voltage, V_{CE}, can swing from its 'Q-point' value, V_{CEQ}, to its maximum value at Cut-off and down to its minimum value at Saturation.

Therefore, the maximum DC values that should be selected when designing a linear amplifier should be practically less than I_{CQ} and V_{CEQ}. This ensures that they do not enter the Saturation/Cut-off regions.

Non-centred Q-point

The output value is limited if the Q-point value is 'non-centred'. Figure 83 shows a DC load-line where the Q-point is fixed towards the Cut-off region.

The output alternation of the ac signal is limited (clipped) by the Cut-off effect. This is shown in figure 83. The Collector current will alternate above I_{CQ} as normal, but clipping will occur when the signal alternates to zero. The Collector-Emitter voltage will alternate from V_{CEQ} up to its *Cut-off* value and clip the signal.

A large signal that alternates around the Q-point will experience greater clipping effects than a small signal.

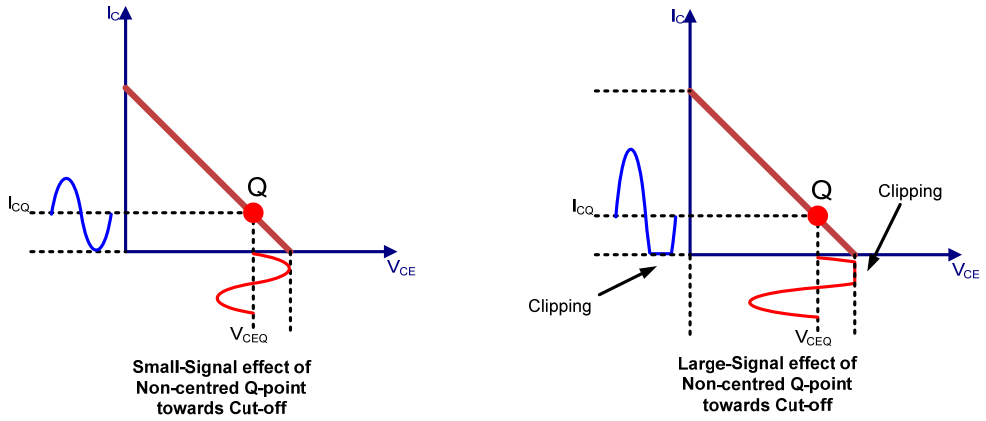


Figure 83 Q-Point Towards Cut-Off

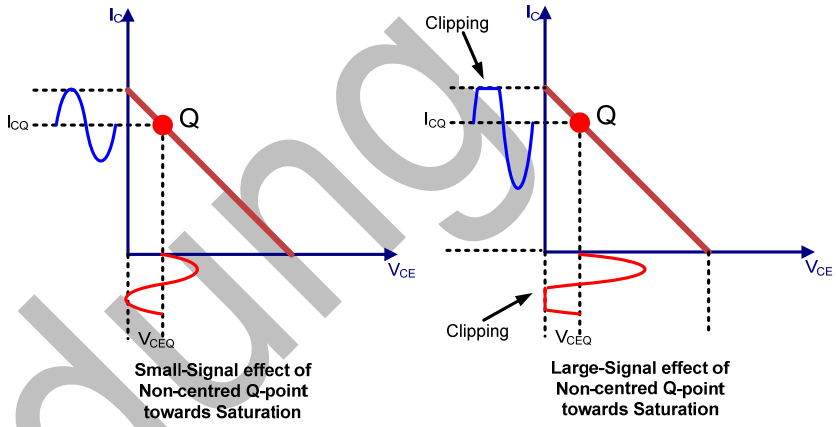


Figure 84 Q-Point Towards Saturation

The Q-point can also be fixed towards the Saturation region. This is shown in figure 84.

In this case the Collector current positive alternation is limited by the Saturation region. The Collector-Emitter voltage negative alternation will also be limited by Saturation.

Power Gain

As discussed earlier, the current gain of an amplifier depends on the β of the transistor. Therefore the power gain can be calculated as $A_p \cong \beta \times A_v$.

DC Quiescent Power

The power dissipation of a transistor with no input ac signal is simply the product of its quiescent current and quiescent voltage. The bias to the transistor is in place and it is operating at its Q-point. The DC quiescent power is shown in equation 7a.

$$P_{DQ} = I_{CQ} \times V_{CEQ}$$

Equation 7a: Dc Quiescent Power

Output Power

The output power is the product of the rms Collector voltage and the rms Collector current. This is correct for any Q-point value. This can be shown as $P_{out} = V_{ce} \times I_c$.

When the Q-point is centred, the maximum Collector current change is equal to I_{CQ}. This is because I_{CQ} is exactly half the value separating the *Saturation* and *Cut-off* regions. This is the same for V_{CEQ}. Equation 7b shows the method to calculate the maximum AC output power from a 'Class A' amplifier under signal conditions.

$$P_{out} = (0.7071 \times V_{CEQ}) \times (0.7071 \times I_{CQ})$$

OR

$$P_{out(\max)} = \frac{V_{CEQ} \times I_{CQ}}{2}$$

Equation 7b: Maximum AC Output Power of A 'Class A' Amplifier

It should be noted that this value is half the value of the DC quiescent power dissipation.

Efficiency

The efficiency of a 'Class A' amplifier is the ratio of AC output power to DC input power.

The DC input power is the product of the voltage across the supply and the current drawn from the supply ($P_{DC} = V_{CC} \times I_{CC}$).

When the Q-point is centred, the average supply current, I_{CC}, is the same as I_{CQ}.

When the transistor operates at Cut-off, the Collector-Emitter voltage is V_{CC}. Since the Q-point centre is exactly between the Cut-off and *Saturation* regions, the value of V_{CEQ} is exactly half the supply voltage, V_{CC} ($V_{CEQ} = \frac{V_{CC}}{2} or V_{CC} = 2 \times V_{CEQ}$).

Therefore the maximum efficiency will occur when the Q-point is centred. This is derived in equation 7c.

$$Eff_{(\max)} = \frac{P_{out}}{P_{DC}} = \frac{(\frac{V_{CEQ} \times I_{CQ}}{2})}{V_{CC} \times I_{CC}} = \frac{(\frac{V_{CEQ} \times I_{CQ}}{2})}{2 \times V_{CEQ} \times I_{CQ}}$$

The values of V_{CEQ} and I_{CQ} will cancel out to leave:

$$Eff_{(\max)} = \frac{P_{out}}{P_{DC}} = \frac{(\frac{1}{2})}{2} = \frac{1}{4} = 0.25 = 25\%$$

Equation 7c: Maximum Efficiency in a 'Class A' Amplifier

This means that the highest possible efficiency for the 'Class A' amplifier is 25%. This can be achieved only when the Q-point is centred on the DC load-line.

Class 'B' Operation

The 'Class B' amplifier is also known as a 'Push-Pull' amplifier. The amplifier is biased to operate in the active (linear) region of the characteristic curve for 180^{O}. It is also designed to operate in the Cut-off region for the remaining 180^{O} of the input signal.

The 'Class B' amplifier is more efficient than a 'Class A' amplifier because there is greater output power for a given value of input power. Figure 85 shows the 'Class B' amplifier.

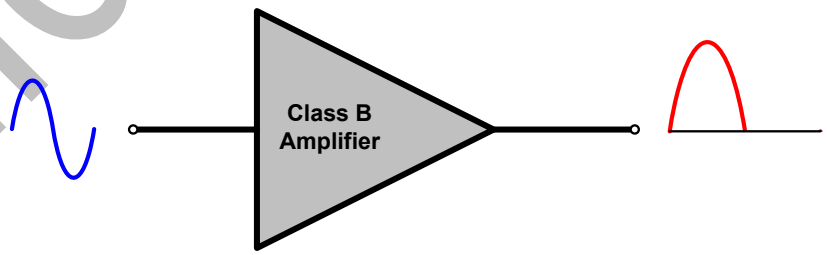


Figure 85 The 'Class B' Amplifier

Q-point

The amplifier is biased to position the Q-point at Cut-off on the load-line. At this point, the DC Collector current, $I_C \approx 0$ and $V_{CE} = V_{CE(Cut\text{-}off)}$.

The positive-half cycle of the input signal brings the amplifier out of the Cut-off region and into the active region when the transistor conducts.

The transistor is switched off when the negative half-cycle of the input drops below V_{BE} (0.7V). This is shown in figure 86 with an Emitter-follower circuit.

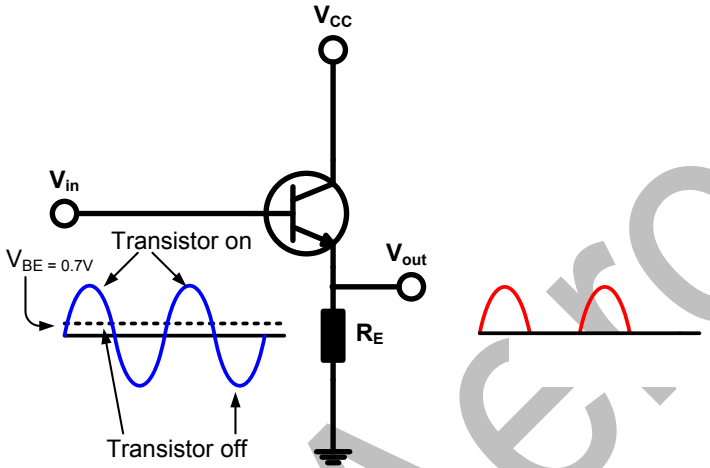


Figure 86 'Class B' Operation

Push-Pull Operation

Push-pull implementation can be implemented using two Emitter-follower circuits. This is shown in figure 87. This is known as a complimentary amplifier because one Emitter-follower uses an 'NPN' transistor. The other uses a 'PNP' device. The 'PNP' transistor will conduct on opposite alternations of the input signal relative to the 'NPN' device.

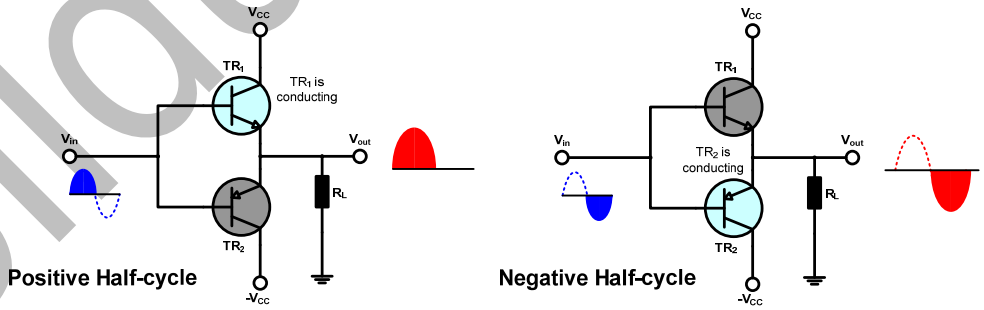


Figure 87 The Push-Pull Amplifier

There is no DC bias applied to the Base (V_B = 0V). This means that only the value of the signal voltage will allow the transistor to conduct.

Transistor TR_1 (NPN) conducts during the positive half-cycle of the input signal. Transistor TR_2 (PNP) conducts during the negative half-cycle of the input.

Crossover Distortion

The effect of the normal forward bias operation for a P-N junction causes Crossover distortion in a 'Class B' amplifier. When there is no DC bias present at the transistor Base, the input signal must become greater than V_{BE} (0.7V) before the transistor can conduct.

This causes a time-interval at the output between the positive and negative half-cycles where both transistors are off. This is shown in figure 7h.

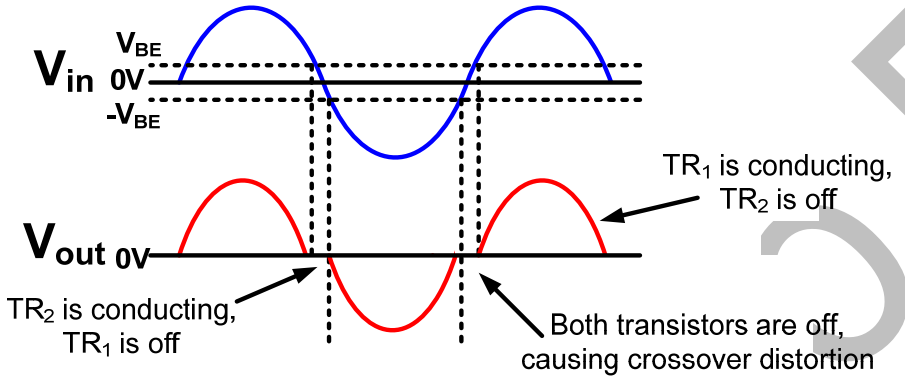


Figure 88 Crossover Distortion

The time interval is the time taken to change from 0.7V (forward bias of TR_1, V_{BE}) to -0.7V (Forward bias TR_2, $-V_{BE}$).

Bias Applied to a Push-Pull Amplifier

To eliminate the crossover distortion, both transistors can be biased slightly above Cut-off when there is no input signal. This ensures both transistors remain above the V_{BE} value.

This is achieved using a voltage divider arrangement. The necessary bias required to switch on the Base-Emitter junction of a transistor is identical to the bias required to allow current flow through a diode. This characteristic is used to create an accurate, stable bias to both transistors. This is shown in figure 89.

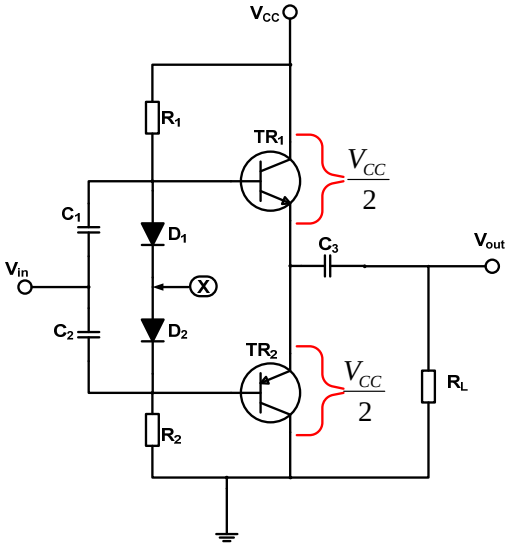


Figure 89 Crossover Elimination

The Resistors R_1 and R_2 should be equal values. This means the DC voltage at point 'X' (with respect to 0V) is $\frac{V_{CC}}{2}$.

The voltage drop across D_1 should be identical to the V_{BE} of TR_1. The voltage across D_2 should be identical to the V_{BE} of TR_2. This means that the voltage at the Emitters of both resistors is $\frac{V_{CC}}{2}$. From this it is seen that $V_{CEQ1} = V_{CEQ2} = \frac{V_{CC}}{2}$. The 'Class B' amplifier, by nature, is at *Cut-off*. This means that $I_{CQ} \cong 0$.

AC Signal Operation

Under optimum conditions, TR_1 and TR_2 are alternately driven from near Cut-off to near *Saturation*. As stated earlier, under AC conditions the supply, V_{CC}, will be viewed as a path to ground by the AC input signal.During the positive half-cycle of the input signal, the TR_1 Emitter is driven from its Q-point value $\frac{V_{CC}}{2}$ to near V_{CC}. This produces a positive peak output voltage of $V_{CEQ} = \frac{V_{CC}}{2}$.

This is shown in figure 90.

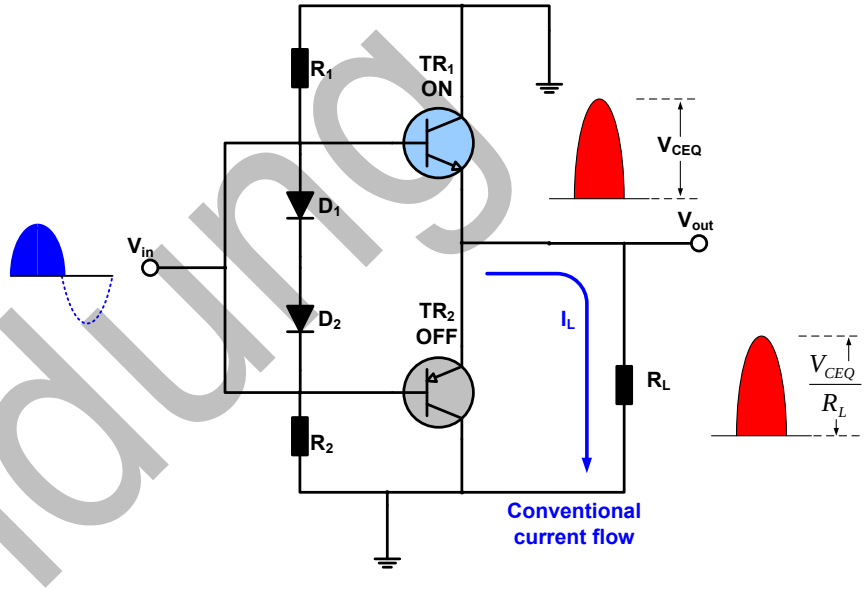


Figure 90 Tr_1 Conducting With Optimum Conditions

The output current from TR_1 will then swing from its Q-point current (Which should be approximately zero) to near is Saturation value.

During the negative half-cycle of the input signal, the Emitter of TR_2 is driven from its Q-point value of $\frac{V_{CC}}{2}$ to near zero. This causes a negative peak output voltage of $V_{CEQ} = \frac{V_{CC}}{2}$. This is shown in figure 91.

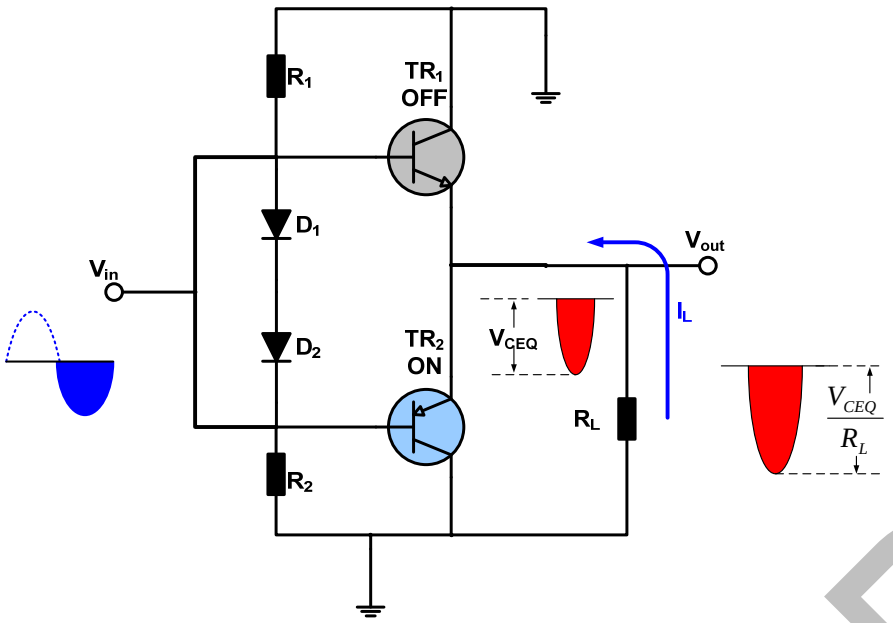


Figure 91 Tr_2 Conducting With Optimum Conditions

The current from TR_2 also swings from approximately zero to its Saturation value.

In both cases, the peak voltage across each transistor is V_{CEQ}. The AC *Saturation* current is found to be $I_{c(sat)} = \frac{V_{CEQ}}{R_L}$. The Emitter current of each transistor is $I_e \cong I_c$. The output current is the Emitter current. Therefore the peak output current is also $\frac{V_{CEQ}}{R_L}$.

Maximum Output Power

It has been shown that the maximum peak output current is $I_{c(sat)}$. The maximum peak output voltage is V_{CEQ}. The maximum average output power is derived in the equation 7d.

$$P_{out} = V_{rms(out)} \times I_{rms(out)}$$

Where:

$$V_{rms(out)} = 0.7071 \times V_{(peak)} = 0.7071 \times V_{CEQ}$$

And:

$$I_{rms(out)} = 0.7071 \times I_{(peak)} = 0.7071 \times I_{c(sat)}$$

Note that $0.7071 \times 0.7071 = 0.5$, the power formula becomes:

$$P_{out} = 0.5 \times (V_{CEQ} \times I_{c(sat)})$$

Also note that $V_{CEQ} = \frac{V_{CC}}{2}$:

$$P_{out} = \frac{0.5 \times (V_{CC} \times I_{c(sat)})}{2}$$

The power formula now becomes:

$$P_{out(\max)} = 0.25 \times (V_{CC} \times I_{c(sat)})$$

Equation 7d: Max Average Output Power

Input Power

The final value for input power is expressed in equation 7e. The input power is drawn directly from the supply and is expressed as $P_{DC} = V_{CC} \times I_{CC}$.

Each transistor only draws current for half of one frequency cycle. Therefore, the supply current is a half-wave signal and is expressed as $I_{CC} = \frac{I_{c(sat)}}{\pi}$.

$$P_{DC} = \frac{V_{CC} \times I_{c(sat)}}{\pi}$$

Equation 7e: Input Power

Efficiency

As stated previously, the 'Class B' amplifier is more efficient than the 'Class A' amplifier.

The efficiency is still determined by the maximum output power ratio to input power ($Efficiency = \frac{P_{out}}{P_{DC}}$). The efficiency is expressed in equation 7f.

$$Eff_{(\max)} = \frac{P_{out(\max)}}{P_{DC}} = \frac{0.25 \times V_{CC} \times I_{c(sat)}}{V_{CC} \times (\frac{I_{c(sat)}}{\pi})} = 0.25 \times \pi$$

$$Eff_{(\max)} = 0.25\pi = 0.785$$

Equation 7f: Efficiency of a 'Class B' Amplifier

The ideal maximum efficiency of a 'Class B' amplifier is 78.5%. This is a vast improvement on the 25% efficiency of a 'Class A' amplifier. These values are taken when the amplifier maximum is at the Saturation limits.

It should be noted that actual maximum efficiency values are slightly lower. This is because the amplifier maximum values must be kept below the Saturation limits to maintain signal integrity.

Class 'C' Operation

The 'Class C' amplifier is designed to conduct for less than 180^{O} of the input cycle. More power can be obtained from 'Class C' operation making it more efficient than either 'Class A' or 'Class B' amplifiers.

The output signal waveform is severely distorted so 'Class C' amplifiers are limited to tuned amplifiers and radio frequencies (RF). An inverting 'Class C' amplifier is shown in figure 92.

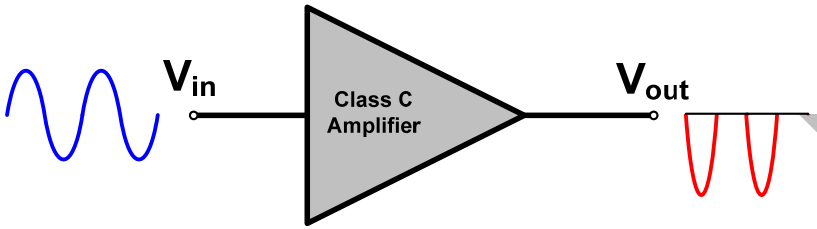


Figure 92 'Class C' Operation

A standard 'Class C' Common-Emitter amplifier with a resistive load is shown in figure 93. It is biased below the Cut-off region by the negative V_{BB} DC voltage.

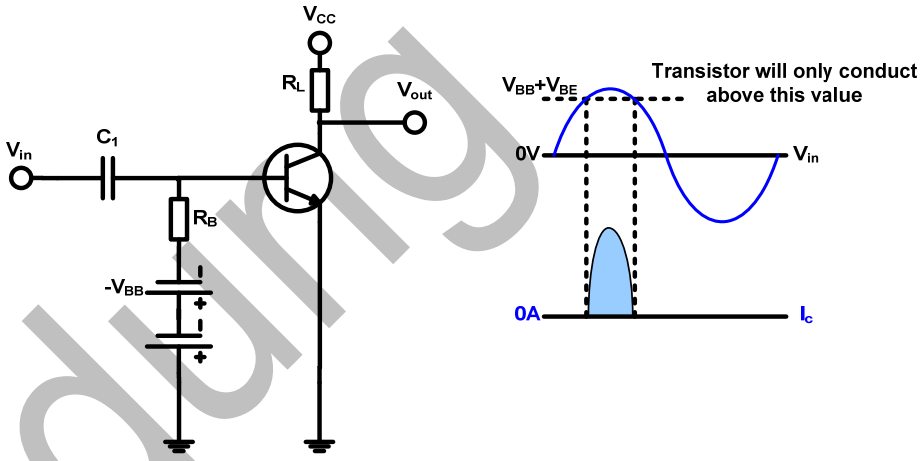


Figure 93 A 'Class C' Amplifier Circuit

The maximum value of the AC input signal is slightly greater than the V_{BB} + V_{BE}. This ensures the AC Base voltage is able to exceed the Base-Emitter junction potential and switch on the transistor. This will occur only for the time duration of the positive peak of the AC signal.

If the entire DC load-line is utilised, the maximum Collector current will be $I_{c(sat)}$. The minimum Collector voltage will be $V_{CE(sat)}$.

Power Dissipation

The power dissipation for a 'Class C' amplifier is low because the output signal is only active for a small time interval compared to the input signal. Figure 94 shows the Collector current pulses.

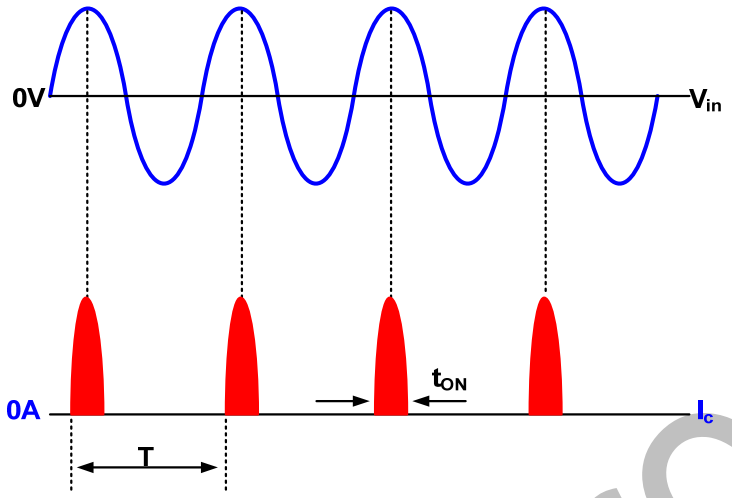


Figure 94 Collector Current Pulses in a 'Class C' Amplifer

The time interval between each pulse is 'T'. This is equal to the time period of the input signal.

The time duration of each pulse is 't_{ON}'. This is the time that the transistor is '*on*'. The transistor is '*off*' for the rest of that input cycle.

The average power dissipation depends on the 'Mark-Space' ratio (The time ratio of 't_{ON}' to 'T') and the power dissipated during 't_{ON}', which is typically very low.

Tuned Operation

The Collector voltage at the output is not a replica of the input in a 'Class C' amplifier. This means that a resistive load has no value in linear amplification.

The resistive 'Class C' amplifier must be replaced with a 'Parallel Resonant arrangement' (known as a tank). This is shown in figure 95.

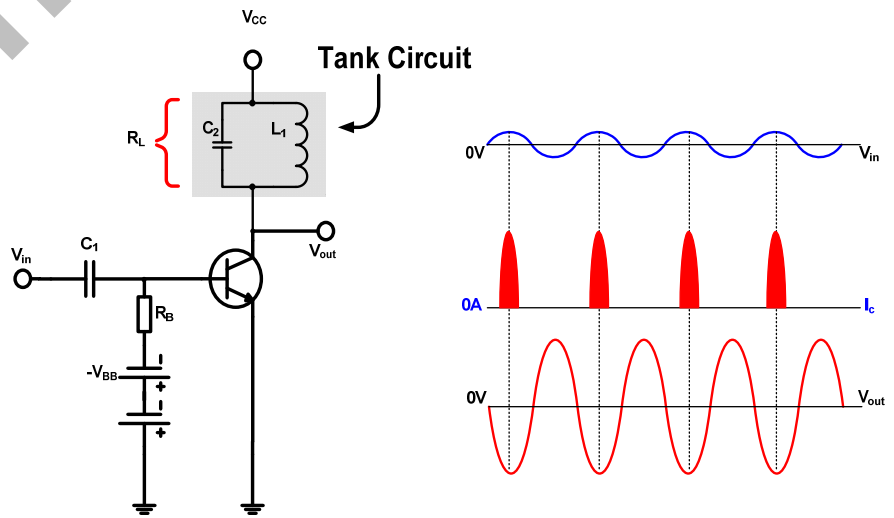


Figure 95 'Class C' Tank Circuit

The short-pulse of the Collector Current on each cycle energizes and sustains the oscillation of the tank circuit. The oscillation of the Tank circuit produces a sinusoidal output voltage. The resonant (natural) frequency of the Tank circuit is shown in equation 7g.

$$f_r = \frac{1}{2 \times \pi \times \sqrt{LC}}$$

Equation 7g: Tank Circuit Resonant Frequency

Without the regular pulses of Collector current into the tank, the output amplitude of each cycle would be less than the amplitude of the previous cycle. This is caused by the energy dissipation due to the resistance of the tank circuit. The Oscillation would eventually die out after an interval of time. This is shown in figure 96.

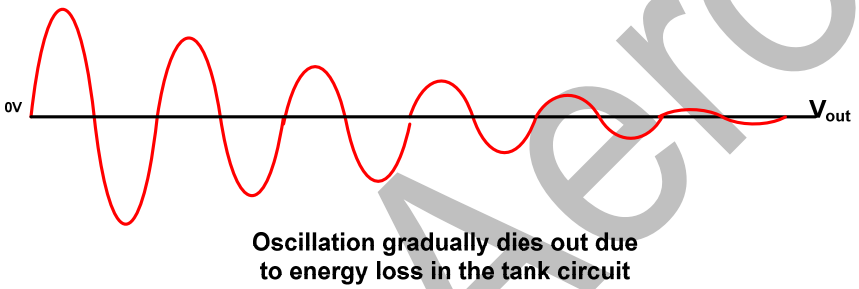


Figure 96 Transient Degradation of Output Signal

The regular pulses of Collector current re-energise the resonant tank circuit to sustain constant oscillation amplitude. The values of the capacitor and inductor coil can be chosen (using equation 7g) to match the frequency of the input signal. This means the tank voltage is tuned to the input signal and is re-energised on each cycle of the input. This is shown in figure 97.

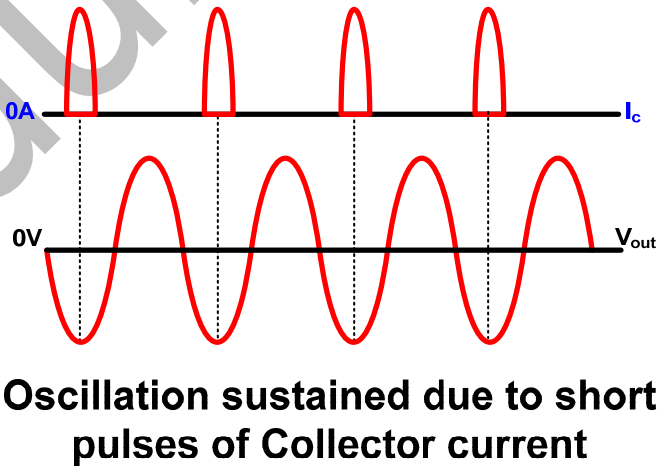


Figure 97 Tank Circuit is Tuned to Input Frequency

Maximum Output Power

The voltage across the tank circuit has a peak-peak value of $2xV_{CC}$. The maximum output power has been expressed in equation 7h.

$$P_{out} = \frac{V^2_{rms}}{R_L} = \frac{(0.7071 \times V_{CC})^2}{R_L}$$

$$P_{out} = \frac{0.5 \times V_{CC}{}^2}{R_L}$$

Equation 7h: Maximum Output Power of a 'Class C' Amplifier

The value R_L is the equivalent parallel resistance of the Collector tank circuit. It represents the parallel combination of the coil resistance and the capacitor resistance. It usually has a small value. The efficiency is expressed in equation 7i.

Total power supplied to an amplifier

$$P_T = P_{out} + P_{D(avg)}$$

Where $P_{D(avg)}$ is the average power dissipation

$$Eff = \frac{P_{out}}{P_{out} + P_{D(avg)}}$$

Equation 7i: Efficiency of a 'Class C' Amplifier

When P_{out} is significantly greater than $P_{D(avg)}$ ($P_{out} >>> P_{D(avg)}$), the 'Class C' amplifier will approach 100% efficiency.

Feedback

Feedback is the process of returning a portion of a circuit's output signal to the input. The objectives are to set-up and maintain specific operating conditions at the output.

The introduction of feedback between the output and input will cause an element of attenuation (loss) to the original output signal.

Feedback can also introduce a time-delay. This time delay is the time taken to reach the input from the output. When analysed, the time-delay will be apparent in the form of a 'Phase-shift'. This is illustrated in figure 99.

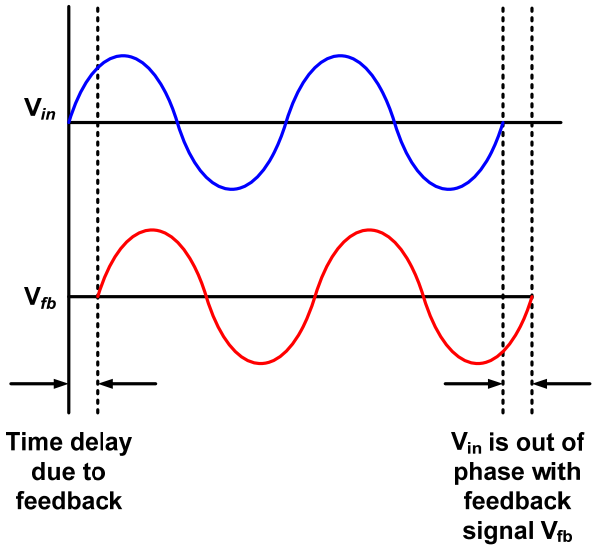


Figure 98 Phase Shift due to Time Delay

The basic required feedback construction is shown in figure 99. There are two main types of feedback:

- Positive Feedback.
- Negative Feedback.

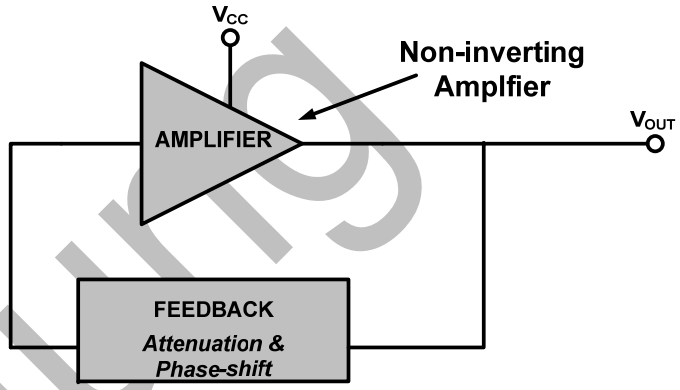


Figure 99 Feedback Construction

Positive Feedback

Positive feedback occurs when a portion of the output voltage from an amplifier is fed back to the input with no time delay (no signal phase shift). The output is then reinforced by the feedback.

The in-phase feedback voltage is amplified to produce the output voltage and the feedback voltage. This is shown in figure 100.

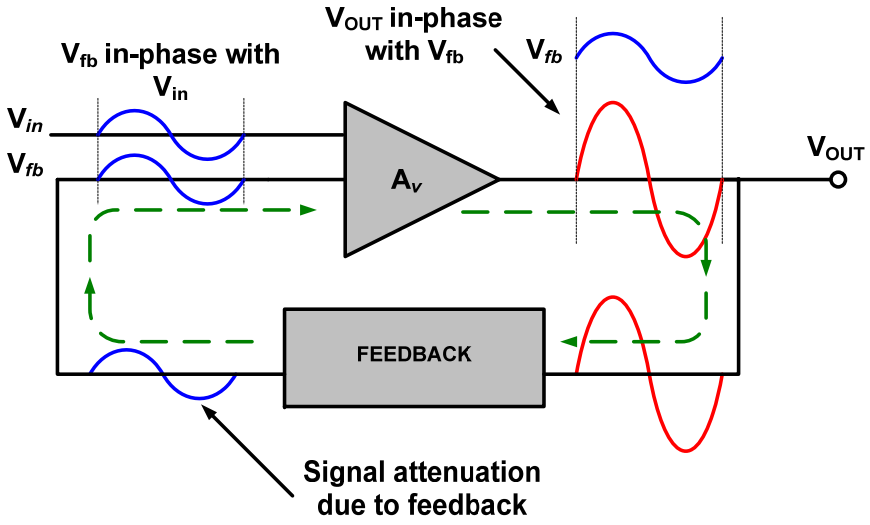


Figure 100 Positive Feedback

A loop is created in which the signal sustains itself with no input signal. A continuous sinusoid is produced. This is called Oscillation.

Negative Feedback

Negative feedback returns a portion of the output signal to the input. The returned signal has a phase-angle that opposes (subtracts from) the input signal.

Negative feedback is illustrated in figure 101. The feedback makes the output signal up to 180^{O} out of phase with the input (Input & output are anti-phase).

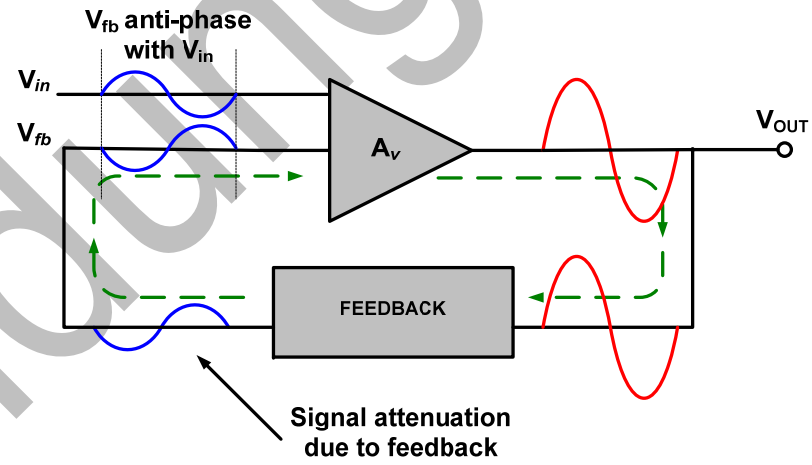


Figure 101 Negative Feedback

Negative feedback is extremely useful, especially in Op-amp applications. It allows an amplifier output to remain temporarily stable if a temporary change occurs at the input. The polarity of the output depends on the result of the subtraction between the input and feedback signals.

1.3.13 Oscillators

An Oscillator is an electronic circuit consisting of an amplifier and a positive feedback network.

An Oscillator produces a time-varying output signal without the necessity of an external input signal.

An Oscillator will produce a repetitive waveform on its output with only a stable DC supply voltage substituting an input.

The output is not necessarily sinusoidal, depending on the type of Oscillator. A basic oscillator is illustrated in figure 102.

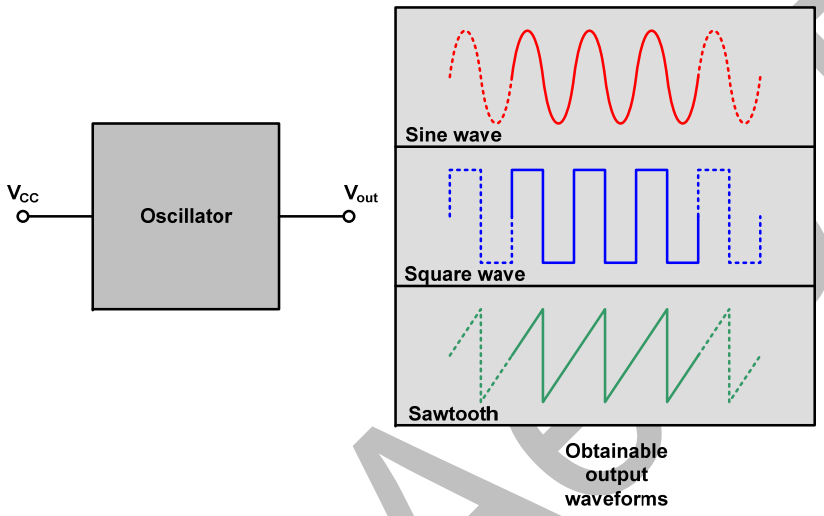


Figure 102 Basic Oscillator

Oscillation is based on the principle of positive feedback. A repetitive input signal is not required.

An Oscillator converts electrical energy in the form of DC to electrical energy in the form of AC.

A basic Oscillator consists of a transistor-based amplifier for gain and positive feedback circuit that provides attenuation and 0° phase-shift. This is illustrated in figure 103.

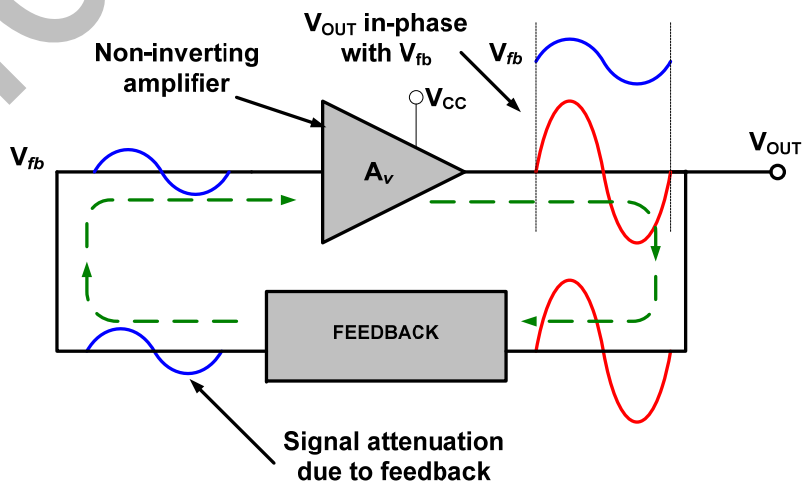


Figure 103 Oscillation using Positive Feedback

The in-phase feedback voltage is amplified to produce the output voltage and the source of the feedback voltage. The output signal is able to sustain itself without an input signal creating a continuous loop and allowing sinusoidal Oscillation to occur at the output.

Conditions for Oscillation

There are two main conditions necessary for a sustained state of Oscillation:

- Feedback loop phase-shift must be 0^{O}.
- Loop gain must be at least 1 (unity gain).

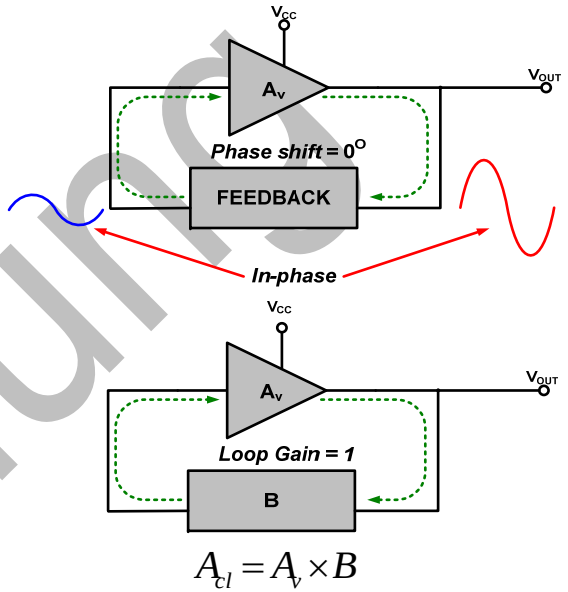


Figure 104 Conditions for Oscillation

The conditions for Oscillation are shown in figure 104. The loop gain, A_{cl}, is the complete voltage gain around the closed feedback loop. The voltage gain around the closed loop is the product of the amplifier gain, A_v, and the attenuation of the feedback circuit, B. This is shown in equation 9a.

$$A_{cl} = A_v \times B$$

Equation 9a: Closed Loop Gain

If a sinusoidal wave is desired, a loop gain greater than 1 (>1) will cause the output to rapidly saturate on both the positive and negative peaks. This will produce unacceptable distortion.

This can be avoided if the gain is controlled to maintain a loop gain of exactly 1 once oscillation has started.

Example:

If the feedback attenuation is 0.01, the amplifier gain must be exactly 100 to overcome the attenuation and avoid the signal distortion (0.01 x 100). If the amplifier gain is greater than 100, the oscillator will limit both peaks of the output waveform.

Start-Up Conditions

There are also requirements to allow the Oscillator to start-up when the DC supply is switched on. The previous conditions only describe the requirements to sustain oscillation.

For oscillation to begin, the voltage gain around the feedback loop must be greater than 1 to allow the amplitude of the output to build up to the desired value.

The gain must then decrease to exactly 1 to secure the output at the desired value and sustain oscillation. This is illustrated in figure 105.

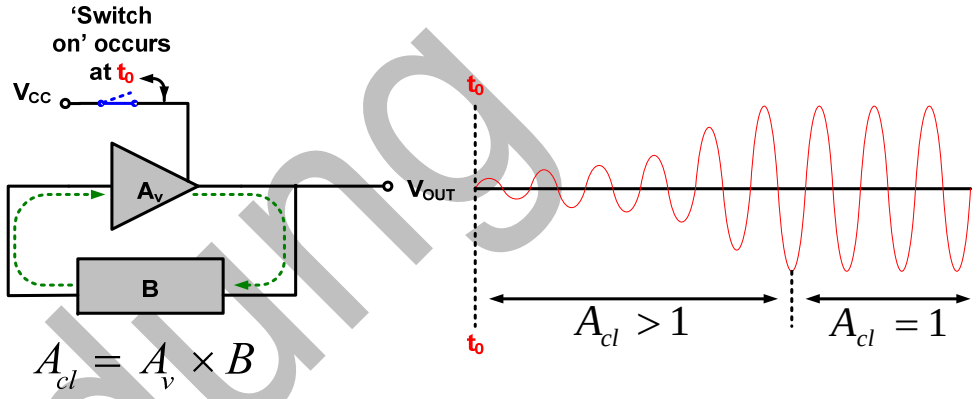


Figure 105 Start-Up Oscillation

The designer must consider the following: If the Oscillator is initially 'off' (no DC supply) and there is no output voltage, how does the feedback signal originate to begin the oscillation process?

When the power is initially switched 'on', a small positive feedback voltage develops from thermally-produced broadband noise in the resistors (or other passive components) or from 'turn-on' transient signals.

A 'turn-on' transient signal is a small signal, usually generated from the power 'spike' (sudden rush of current) when the power supply is initially activated. A transient signal will die away over a period of time.

The feedback circuit will only allow a voltage with a frequency equal to the desired oscillation frequency to appear in phase on the amplifier's input.

This initial feedback voltage is then amplified and continuously reinforced by the feedback circuit causing a build-up at the output signal. Oscillators can be used to generate reference signals in digital/analogue systems or clock signals in digital systems.

The RC Oscillator

The RC oscillator uses a network of resistors and capacitors as a feedback circuit. The basic RC Oscillator is shown in figure 106.

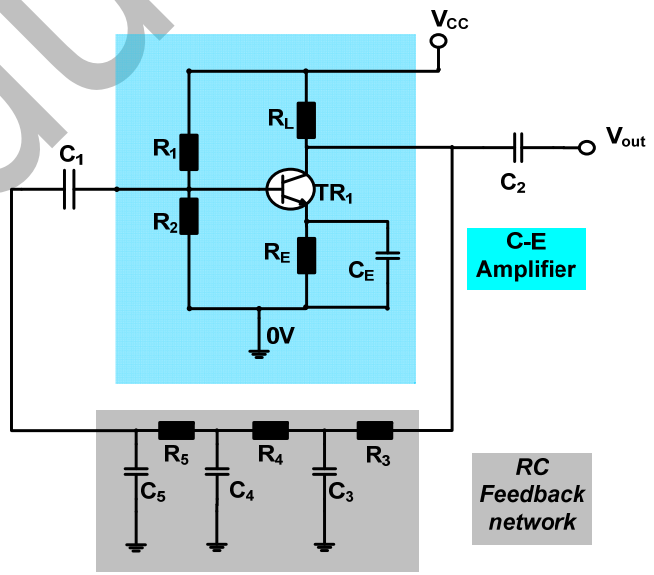


Figure 106 The RC Oscillator

Three RC 'Lag' networks can be established to have a total phase lag of 180^{O}. The Common-Emitter transistor also has a lag of 180^{O}. The feedback signal will reach the input to the amplifier with a full 360^{O} lag from the original input signal.

A 360^{O} lag is a full cycle behind the initial input signal. This is effectively 0^{O} (no practical shift in phase) when the feedback signal is re-applied to the input. This is illustrated in figure 108.

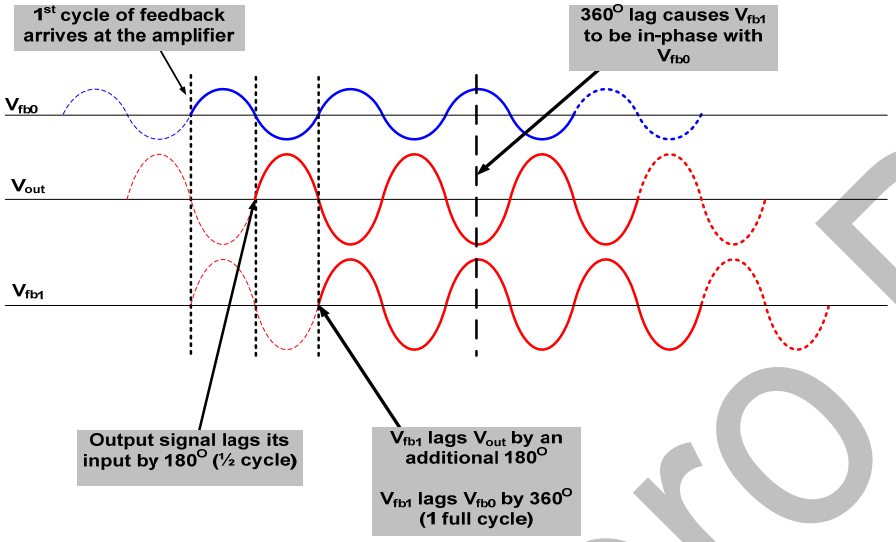


Figure 107 Full Cycle Lag

The attenuation of the RC network and the gain of the amplifier must be unity at the frequency of oscillation. This circuit will produce a sinusoidal oscillation.

The Colpitts Oscillator

Another type of tuned oscillator is the Colpitts oscillator. It is named after its inventor. This oscillator uses an LC (Inductor coil and capacitor) feedback network to achieve the necessary phase-shift. The LC network also acts as a filter that passes only the signals operating at the desired oscillation frequency. The Colpitts Oscillator is shown in figure 108.

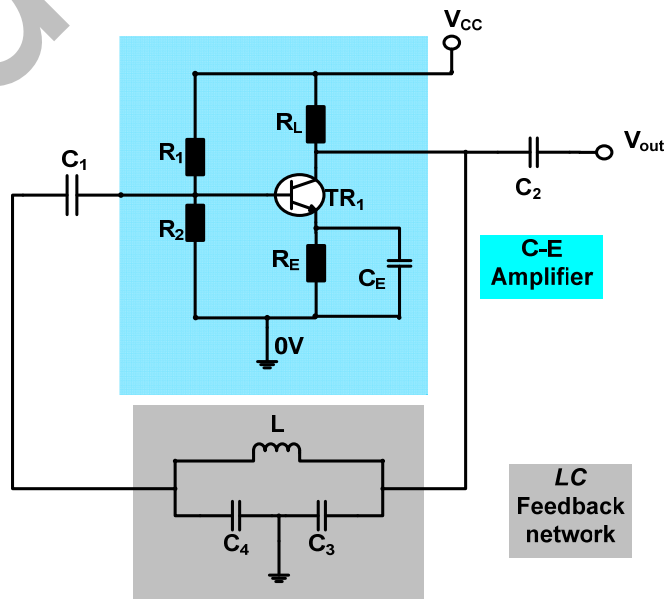


Figure 108 The Colpitts Oscillator

The frequency of oscillation can be determined using the values of the capacitors, C_3 & C_4 and the inductor coil, L.

The capacitors are connected in series. This effectively increases the distance between the plates causing the overall capacitance to decrease. The method of calculating the total capacitance is shown in equation 9b.

$$C_T = \frac{C_3 \times C_4}{C_3 + C_4}$$

Equation 9b: Total Capacitance in the Colpitts Oscillator

This value can then be used to determine the oscillation frequency. The method used to calculate the oscillation frequency is shown in equation 9c.

$$F_{OSC} \cong \frac{1}{2 \times \pi \times \sqrt{L \times C_T}}$$

Equation 9c: Oscillation Frequency

Other types of oscillators can be constructed using different combinations of coils and capacitors. These Oscillators are known as:

- The Hartley Oscillator.
- The Clapp Oscillator.

All Oscillators discussed can also be constructed using Operational amplifiers. Op-amp oscillators can also be refined and constructed in other methods and shall be discussed later.

The Crystal Oscillator

A crystal oscillator is a tuned circuit that uses a quartz crystal as the resonant tank circuit. Crystal oscillators offer greater frequency stability than other oscillators.

Quartz exhibits a property called the 'Piezo-electric effect'. When a changing mechanical stress is applied across the crystal to cause it to vibrate, a voltage is developed that matches the frequency of the mechanical vibration.

When an AC voltage is applied across the crystal, it vibrates at a frequency that matches the signal frequency. The circuit symbol for a crystal oscillator is shown in figure 109.

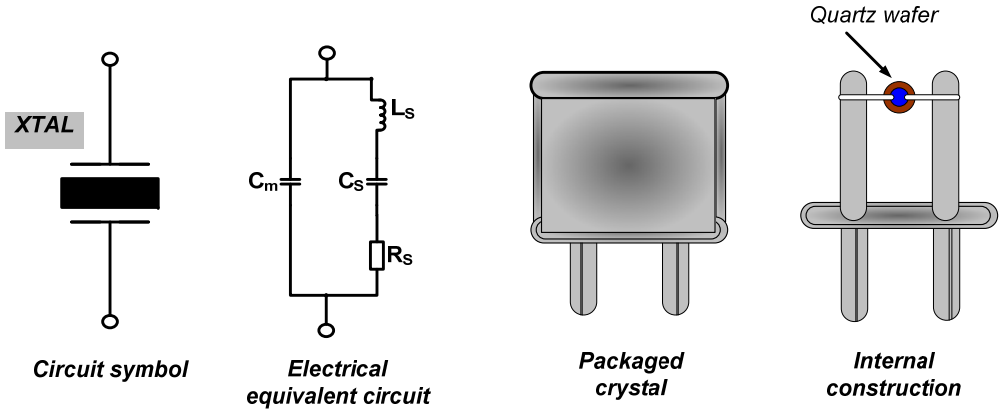


Figure 109 The Crystal Oscillator

Series resonance will occur in the crystal when the reactances in the series branch are equal.

Parallel resonance will occur at a higher frequency, when the inductive reactance of L_S equals the reactance of the parallel capacitor, C_m.

An oscillator using the crystal as a series resonant tank circuit is shown in figure 110.

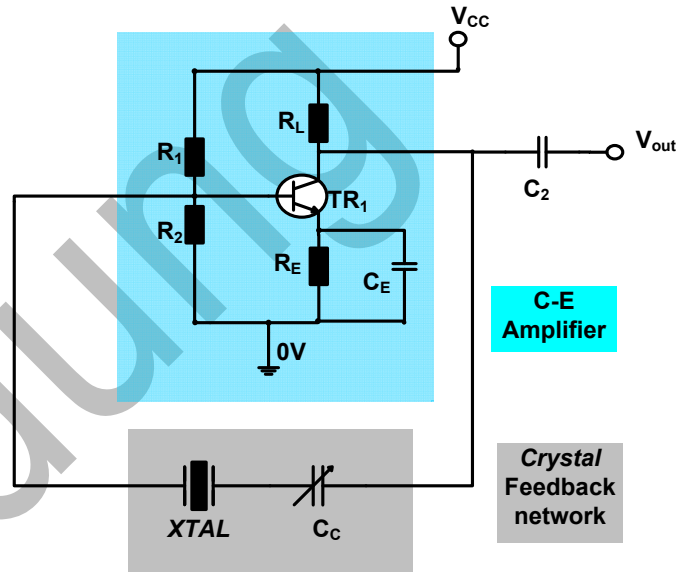


Figure 110 Series Resonant Crystal Oscillator Circuit

The impedance of crystal is at its minimum at the series resonance frequency, therefore providing maximum feedback.

The crystal-tuning capacitor, C_C, is used to 'fine-tune' the oscillator frequency by 'pulling' the resonant frequency of the crystal slightly up or down.

Other circuits exist to utilise the crystal oscillator in a parallel resonance combination.

1.3.14 Multivibrators

Multivibrators are a class of digital circuits using feedback to produce either two stable states, one stable state or no stable states, depending on the desired configuration.

There are three main types of Multi-vibrator:

- The Mono-stable Multi-vibrator (One-shot).
- The Bi-stable Multi-vibrator.
- The Astable Multi-vibrator.

Monostable Multi-Vibrator

A Mono-stable multivibrator has one stable state. It produces a single pulse in response to a triggered input. It is also commonly called a 'One-shot' multi-vibrator.

Once the trigger has been applied, the 'One-shot' will remain in the unstable state for a pre-determined length of time. After the pre-determined length of time has expired (time-out) the 'One-shot' will return to its stable state.

The period of time that the device remains in the unstable state will determine the pulse width at its output.

A 'One-shot' will reside in its stable state by default. It will change to its unstable (changing) state when an external trigger is applied. A Mono-stable multi-vibrator constructed from Bi-polar transistors is shown in figure 111.

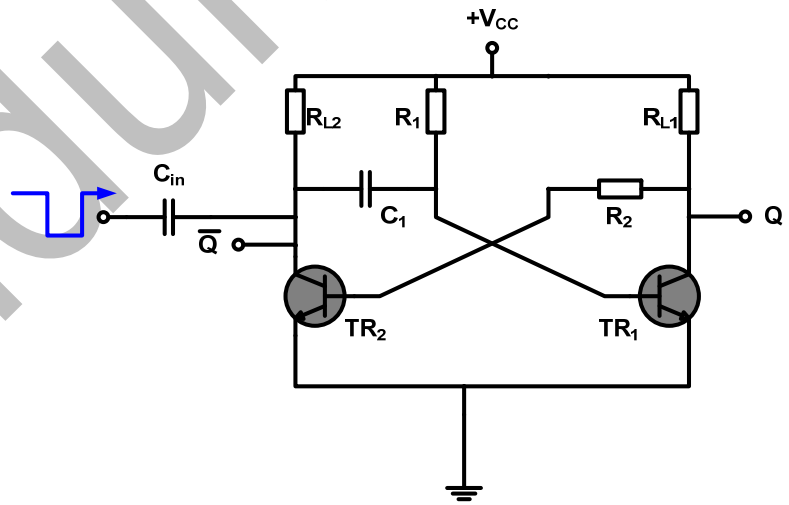


Figure 111 Mono-Stable Constructed From Transistors

The transistors are constructed using the Common Emitter structure. The components of the inverting amplifier are TR_1, R_{L1}, R_1 and C_1. The components of the inverter are R_2, R_{L2} and TR_2.

In the stable state, TR_1 will be 'on' while TR_2 is 'off'. In the unstable state, TR_1 will be 'off' while TR_2 is 'on'. The time period of the unstable state is determined by the RC time constant generated by R_1 and C_1. The Base and Collector waveforms of the respective transistors are shown in figure 112.

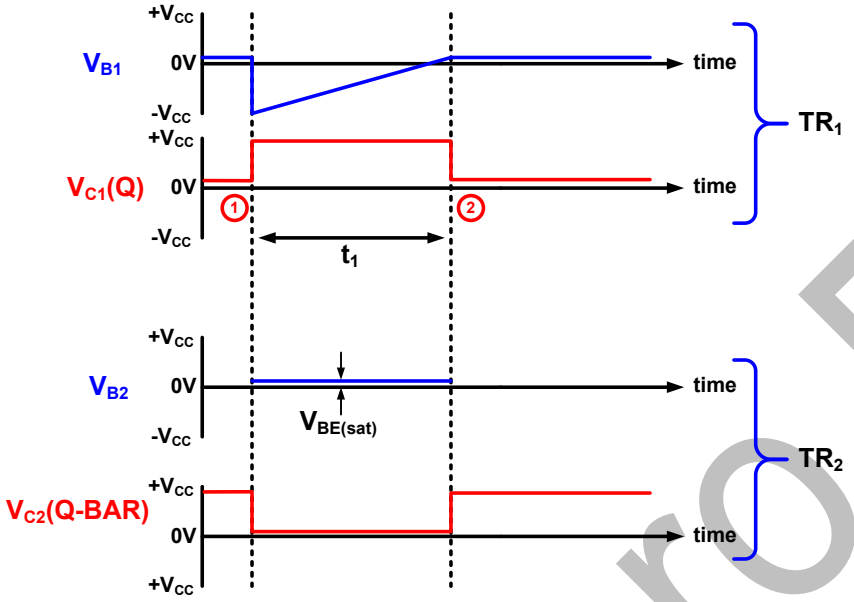


Figure 112 Mono-Stable Characteristic Waveforms

Stable State

In the stable state, no current flows through C_1, but current flow through R_1 is sufficient to keep TR_1 'on' (operating at Saturation).

The Base voltage V_{B1} equals $V_{BE(sat)}$ (0.7V). The *Q* output at V_{C1} is approximately $V_{CE(sat)}$ (≈ 0.2V) because the full V_{CC} is dropped across R_{L1}. This means that V_{B2} is approximately 0.2V, keeping TR_2 'off'. This means that Q-Bar (V_{C2}) is equal to V_{CC} as there is no voltage drop across R_{L2}.

Transition at Point 1 (Figure 113)

A short, negative-going pulse is applied to the input capacitor, C_{in}. Initially V_{C2} and V_{B1} are pulled down by the same magnitude as the pulse.

TR_1 is switched 'off' and Q (V_{C1}) rises toward V_{CC}. R_{L1} and R_2 now provide sufficient Base current to saturate TR_2. V_{C2} (Q-bar) is then held at $V_{CE(sat)}$.

C_1 ensures that V_{B1} initially drops by the same magnitude as V_{C2}. Therefore V_{B1} equals almost $-V_{CC}$.

Duration 't_1'

V_{B1} rises exponentially from almost $-V_{CC}$ towards $+V_{CC}$ as C_1 charges up through R_1. When V_{B1} reaches V_{BE}, TR_1 switches 'on' and the circuit switches back to its stable state.

Transition at Point 2 (Figure 112)

As V_{B1} switches TR_1 'on', V_{C1} drops to $V_{CE(sat)}$. This action drops V_{B2} to approximately 0.2V, switching TR_2 'off'. This causes V_{C2} (Q-Bar) to rise to $+V_{CC}$ with the time constant of $R_{L2} \times C_1$.

The duration of the unstable state is determined by the time taken for V_B to rise (exponentially towards $+V_{CC}$) from ≈$-V_{CC}$ to ≈0V. This is shown in figure 113.

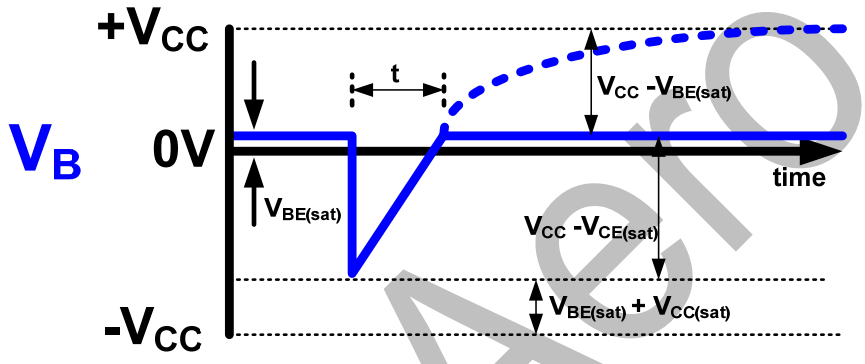


Figure 113 Duration of the Unstable State

The time period of the unstable state is shown in equation 10a. The rise toward $+V_{CC}$ is exponential causing the normal RC time constant to change slightly. The values of $V_{BE(sat)}$ and $V_{CE(sat)}$ have been ignored for simplicity.

$$V_{CC} = 2 \times V_{CC} \times (1 - e^{(\frac{-t}{C \times R})})$$

$$\frac{V_{CC} = 2 \times V_{CC} \times (1 - e^{(\frac{-t}{C \times R})})}{V_{CC}}$$

$$1 = 2 \times (1 - e^{(\frac{-t}{CR})})$$

$$\frac{1}{2} = 1 - e^{(\frac{-t}{CR})}$$

$$1 - \frac{1}{2} = e^{(\frac{-t}{CR})}$$

$$\frac{-t}{CR} = 1 \times \ln_{\frac{1}{2}}$$

$$\frac{t}{CR} = \ln_2$$

$$t = C \times R \times \ln_2$$

$$t = 0.693 \times C \times R$$

Therefore;

$$t = 0.693 \times C_1 \times R_1$$

Equation 10a: Duration of the Unstable State

This pulse is triggered by the leading edge of the shorter input pulse. Trigger pulses should not be re-applied until after each output pulse has been generated and the circuit has recovered from it (V_{C2} reaches $+V_{CC}$).

Mono-Stable Constructed From Logic Gates

Figure 114 shows a basic Mono-stable Multivibrator constructed from a logic gate and an inverter.

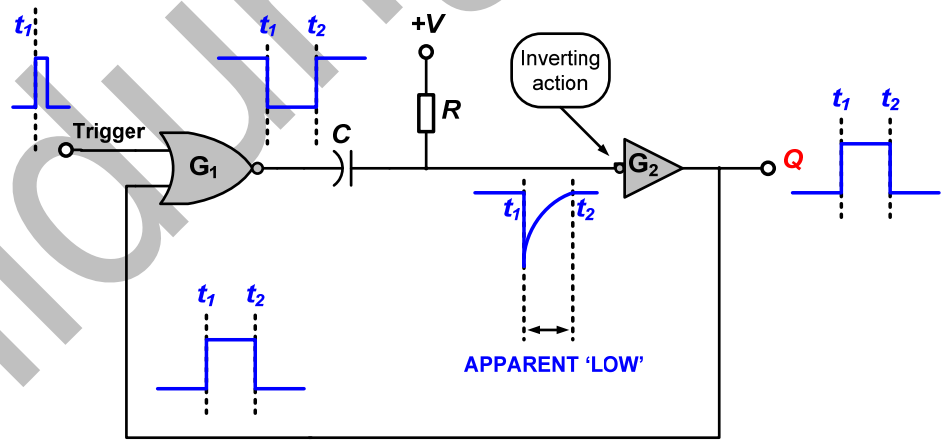


Figure 114 'One-Shot' Multivibrator Constructed from Logic Gates

When a pulse is applied to the trigger input, the output of the 'Nor' gate, G_1, goes low (logic '0').

The high (logic '1') to low transition is coupled to the inverter, G_2, through the capacitor, C.

The apparent 'low' signal at the input to the inverter causes its output to go high. The high output is connected back to G_1, keeping the output of G_1 in the low state.

Up to this point in time the trigger pulse has caused the output of the system, Q, to go high.

The capacitor then begins to charge to the high voltage level through the resistor R. The rate at which it charges is determined by the RC time constant.

When the capacitor charge reaches a certain level, it appears as the high state at the input to the inverter, G_2. The output, Q, then returns to the low state.

It can be seen that the output of the inverter, Q, goes high in response to the trigger input. It remains high for a time determined by the RC time constant. After this time period has expired, the output, Q returns to the low state.

A single narrow trigger pulse produces a single output pulse. The duration of this pulse is determined by the RC time constant.

There are two basic types of Mono-stable Multivibrator:

- Non-retriggerable.
- Re-triggerable.

Non-Retriggerable

A non-retriggerable 'One-shot' will not respond to any additional trigger pulses once it has been triggered.

This will apply between the time it reaches its unstable state and then returns to its stable state. It will ignore any additional trigger pulses that occur before it 'times out'. This is illustrated in figure 115.

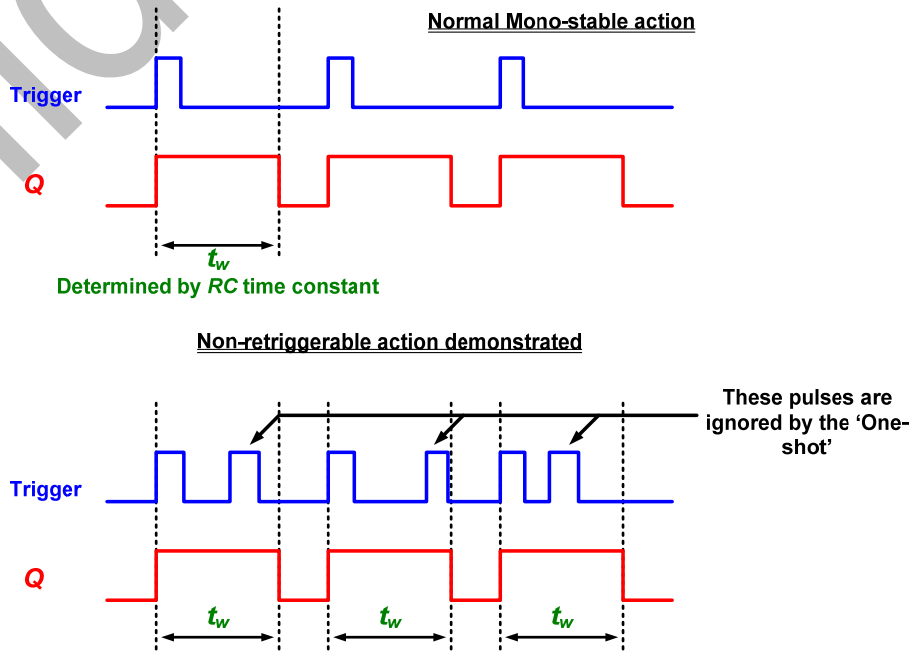


Figure 115 Non-Re-triggerable Mono-Stable Action

Re-Triggerable

A re-triggerable 'One-shot' can be triggered before it 'times out'. The result of triggering is an extension of the pulse width of the output, Q. This is illustrated in figure 116.

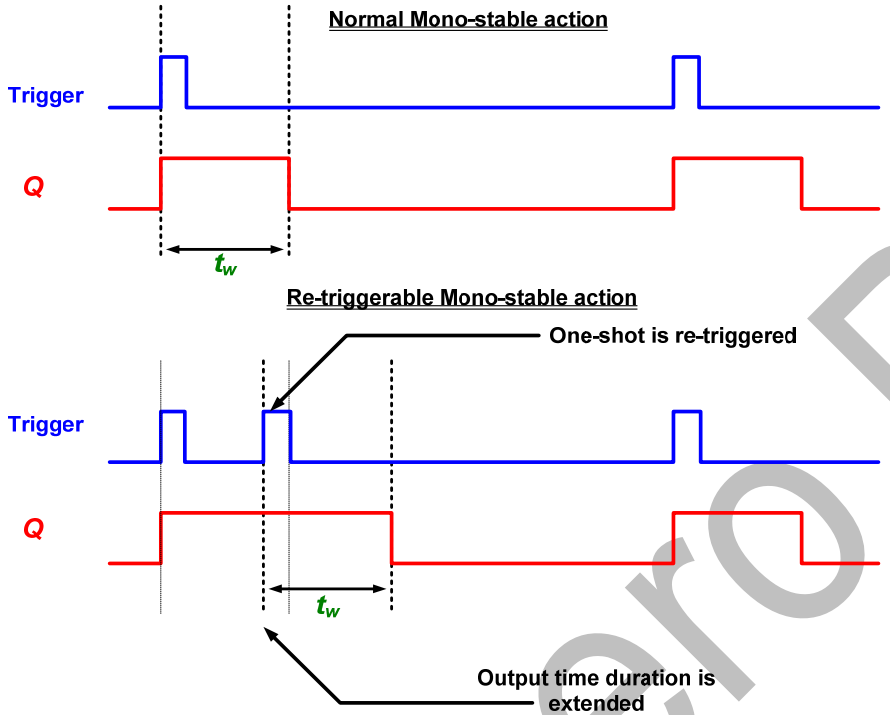


Figure 116 Re-Triggerable Mono-Stable Action

The 'Schmitt Trigger' Action

The Schmitt trigger action uses a characteristic called 'Hysteresis' to prevent erratic switching between states. Erratic switching can occur when a triggered output is slow to change or when a trigger voltage 'hovers' above and below the switching level. This is illustrated in figure117.

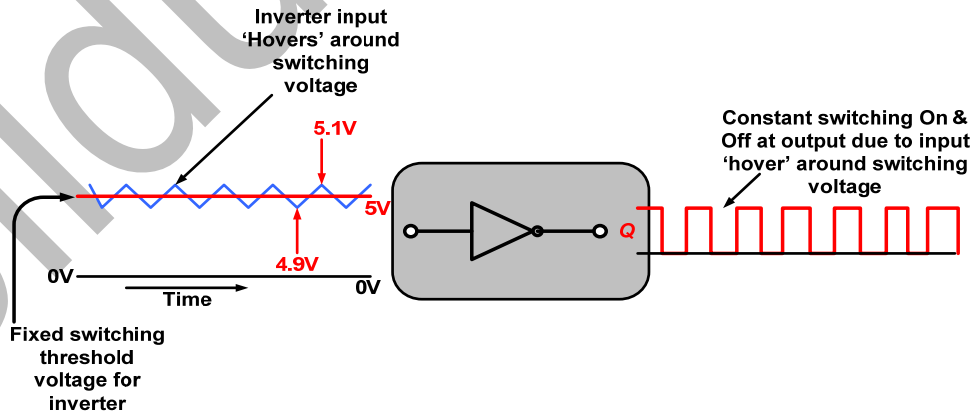
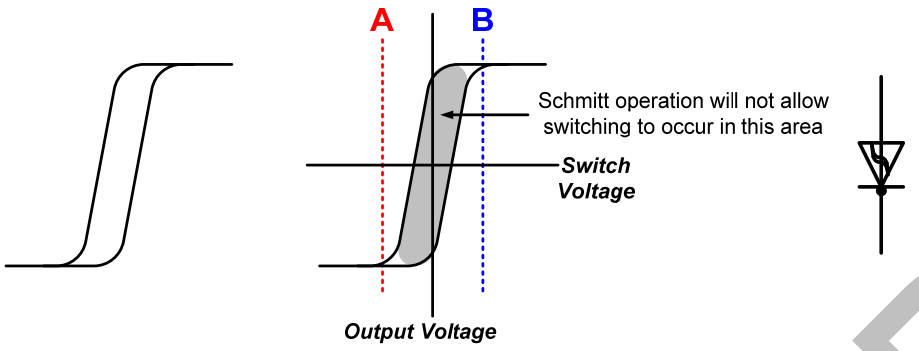


Figure 117 Erratic Switching

The Hysteresis characteristic ensures that state switching only occurs when it is desired. The symbol for the Schmitt action is shown in figure 118. Integrated circuits that show this symbol utilise the Hysteresis characteristic.



'Low' to 'high' transition will occur at point 'A'

'High' to 'low' transition will occur at point 'B'

The Schmitt condition will force the output into either the 'High' or 'Low' state

Figure 118 Schmitt Action Symbol

The logic symbols for a Mono-stable Multi-vibrator are shown in figure 119. The R/C inputs are used for the reference voltage from external components.

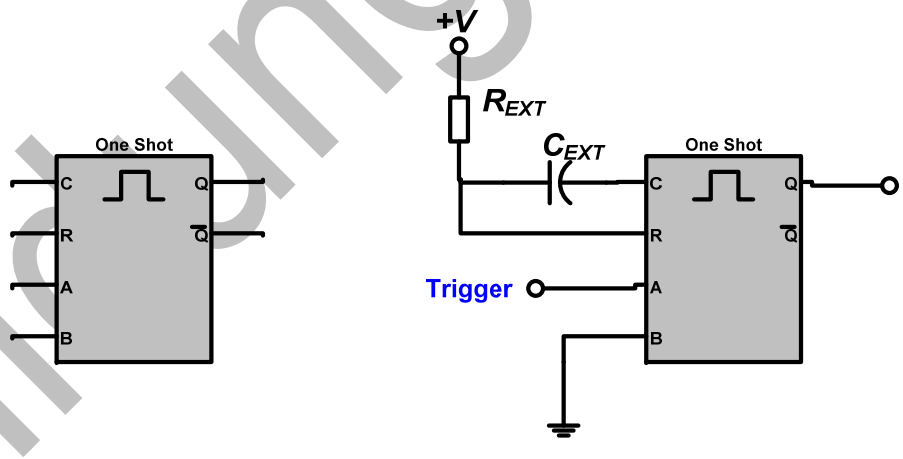
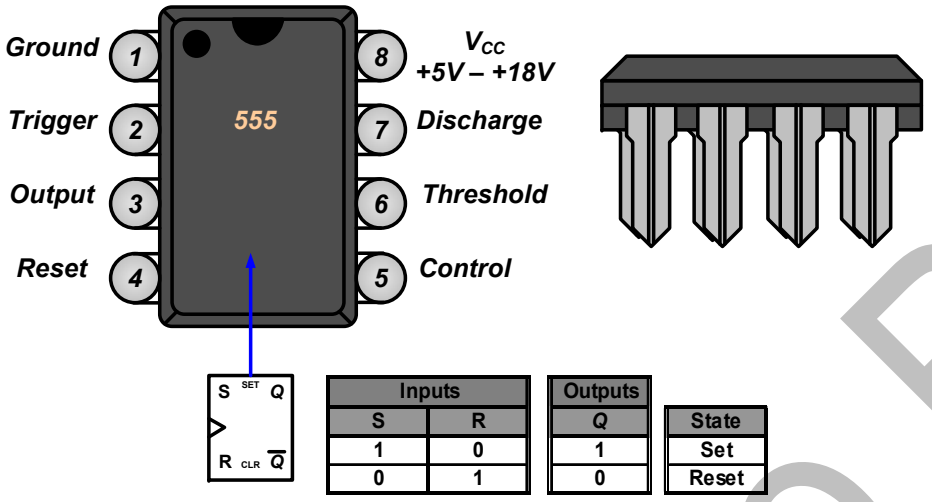


Figure 119 Mono-Stable Multi-Vibrator Logic Symbols

The '555 Timer' as a Mono-Stable Multi-Vibrator

The 555 timer is a versatile and widely used device. It can be configured as a Mono-stable Multi-vibrator. The integrated circuit construction of the 555 timer is shown in figure 120.



Inputs		Outputs	State
S	R	Q	
1	0	1	Set
0	1	0	Reset

Figure 120 555 Timer Integrated Circuit

Basic Operation of a 555 Timer

The block diagram of the 555 timer is shown in figure 121. The comparator is a device using an operational amplifier.

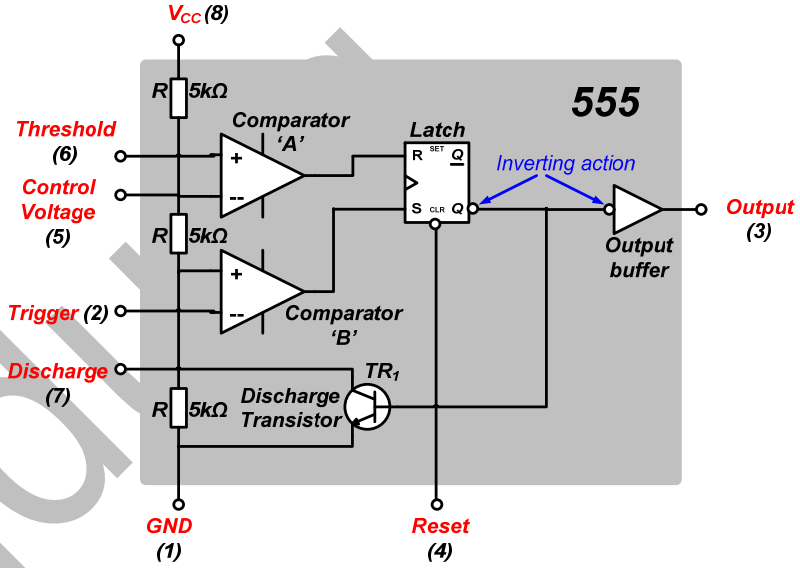


Figure 121 555 Timer Functional Circuit

A Comparator is a device that gives a 'high' output (logic '1') when the voltage on its positive input (+) is greater than the voltage on the negative (-) input.

A Comparator output goes low (logic '0') when the voltage on its negative input (-) is greater than the voltage on the positive (+) input.

The voltage divider consists of three 5kΩ resistors. This provides a trigger level of 1/3 V_{CC} and a threshold level of 2/3 V_{CC}.

The control voltage input (pin 5) can be used to externally adjust the trigger and threshold levels to different values if desired.

When the normally 'high' trigger input (pin 2) momentarily goes below 1/3 V_{CC}, the output of Comparator 'B' switches from 'low' to 'high'.

This action 'Sets' the SR latch. Normally the Q output of the SR latch after the 'Set' function would become 'high' (logic '1'). In this case, an inverting function has been added to the Q output to cause it to go 'low'.

This 'low' signal allows the discharge transistor, TR_1, to switch off. The Q output from the SR latch is then re-inverted and fed to a signal buffer to give the desired 'high' output (pin 3).

The output will remain 'high' until the normally 'low' threshold input (pin 6) goes above 2/3 V_{CC}. This causes the output of Comparator 'A' to switch from 'low' to 'high'.

This action will 'Reset' the latch, causing the output (pin 3) to become 'low'. The internal inverting action will also switch on the discharge transistor TR_1.

The external 'Reset' input can be used to reset the latch independent of the of the threshold circuit.

The Trigger and Threshold inputs are controlled by external components connected to produce either Mono-stable or Astable action.

Mono-Stable Operation of a 555 Timer

An external resistor and capacitor can be used to condition the 555 timer to behave as a non-retriggerable Mono-stable Multi-vibrator. This is shown in figure 122.

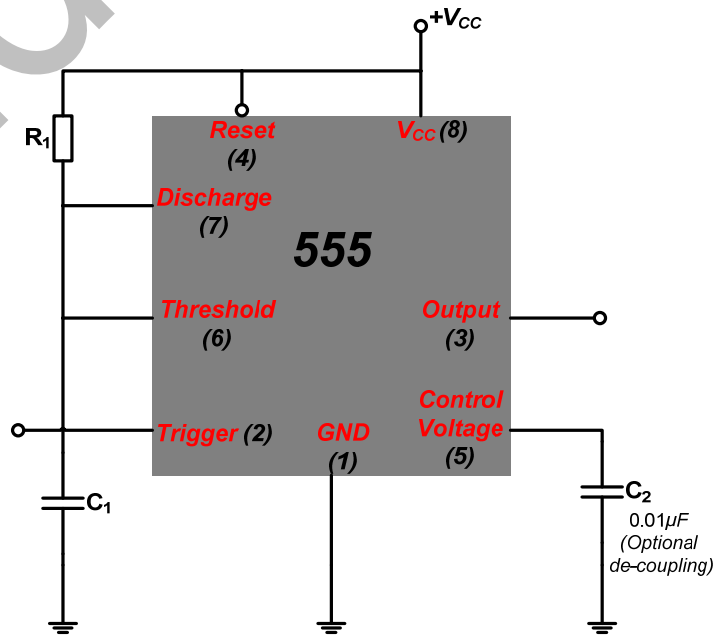


Figure 122 555 Timer Mono-Stable Operation

The pulse width of the output is determined by the time constant of R_1 & C_1. This is shown in equation 10b.

$$t_w = 1.1 \times R_1 \times C_1$$

Equation 10b: 555 Monostable Time Constant

The control voltage input (pin 5) is not used and is connected to a de-coupling capacitor, C_2, to prevent noise from affecting the trigger and threshold values.

Before a trigger pulse is applied, the output (pin 3) is 'low' and the discharge transistor; TR_1 is 'on', keeping C_1 discharged. This is shown in figure 123.

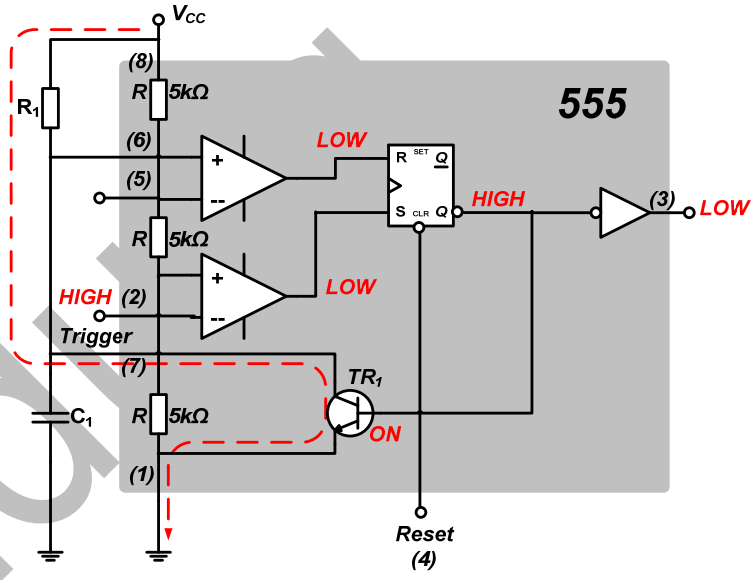


Figure 123 Prior to Triggering

When a negative-going (high to low transition) trigger pulse is applied, the output goes 'high' and the discharge transistor turns 'off', allowing C_1 to begin charging through R_1. This is shown in figure 124.

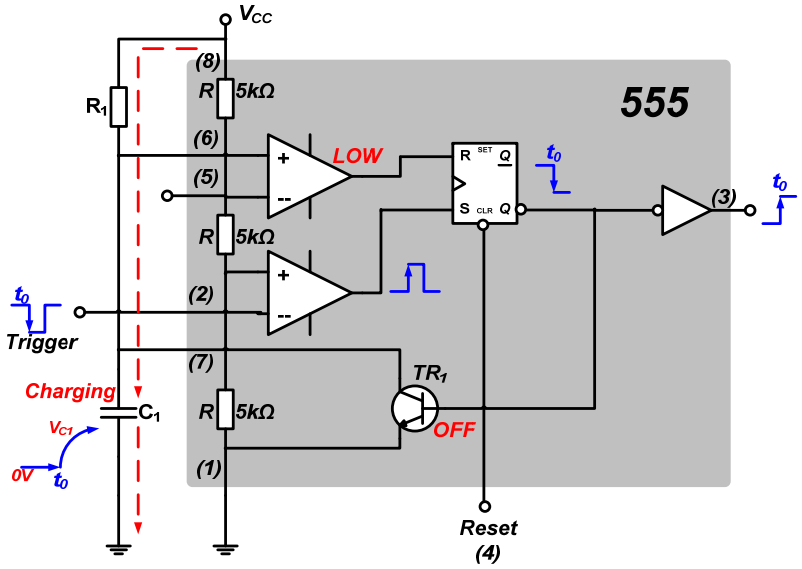


Figure 124 When Triggered

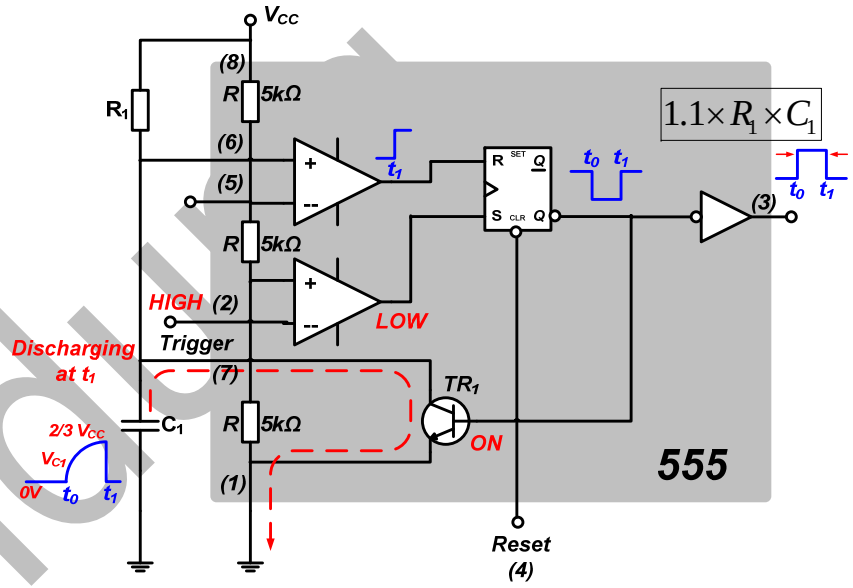


Figure 125 At the End of the Charging Interval

Once C_1 charges to 1/3 V_{CC}, the output returns to the 'low' state and switches TR_1 'on'. This action immediately causes C_1 to discharge. This is shown in figure 125. The rate of discharge will depend on the period of time the output (pin 3) was in the 'high' state.

Bi-Stable Multi-Vibrator

A Bi-stable Multi-vibrator (flip-flop) has two stable states. A Bi-stable device can applied in the form of a latch.

The Latch uses a feedback arrangement where the outputs are fed back to the opposite inputs. A Bi-stable circuit constructed from Bi-polar transistors is shown in figure 126.

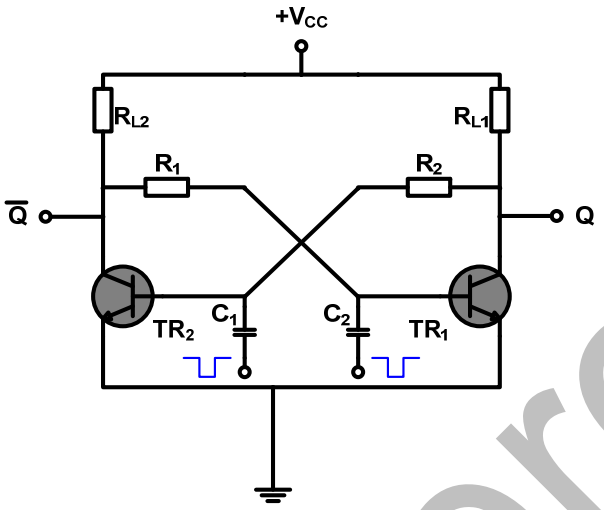


Figure 126 Bi-Stable Multi-Vibrator Constructed from Bi-Polar Transistors

Both transistors are arranged as inverters. The components of one inverter are R_1, R_{L1} and TR_1. The components of the other inverter are R_2, R_{L2} and TR_2.

In one stable state, TR_1 will be 'on' while TR_2 will be '*off*'. In the other stable state, TR_1 will be 'off' while TR_2 be will '*on*'.

Initial Conditions

The circuit will power up into either one of its two stable states.

Stable State 1

If TR_2 is initially assumed to be off, V_{C2} is at V_{CC}. Current through R_{L2} and R_1 is sufficient to keep TR_1 on. V_{B1} is equal to $V_{BE(sat)}$. V_{C1} (Q) is equal to $V_{CE(sat)}$ (≈0.2V). This means that V_{B2} is 0.2V causing TR_2 to remain off.

Transition 1

A short negative-going pulse is applied to the Base of TR_1 through the input capacitor, C_2. The negative voltage at VB_1 switches TR_1 off causing V_{C1} to rise to $+V_{CC}$.

This means that R_{L1} and R_2 now supply sufficient current to switch on and saturate TR_2. V_{C2} (Q-Bar) is then held at 0.2V ($V_{CE(sat)}$). This causes V_{B1} to drop to 0.2V which will hold TR_1 *off*, despite the removal of the input pulse.

Stable State 2 And Transition 2

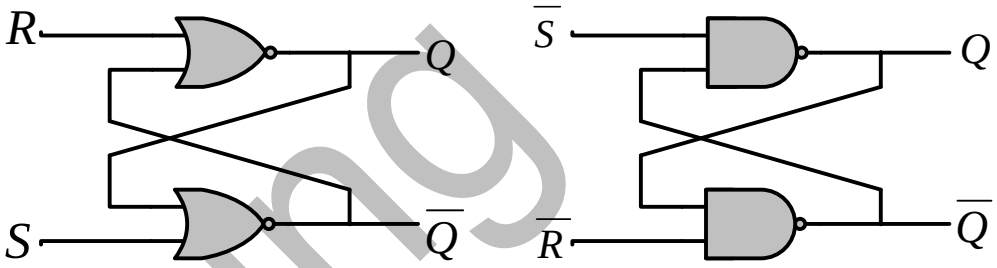
The circuit is symmetrical so the other stable state and transition correspond exactly to the previous ones. The transistor references can be transposed.

Timing

A change of state is initiated by the leading edge of the input trigger pulse. The new state is maintained until the other input pulse occurs.

The SR Latch as a Bi-Stable Multi-Vibrator

The SR (Set-Reset) latch can also be used to illustrate Bi-stable operation. An 'Active-High' input SR latch using two cross-coupled NOR gates is shown in figure 127.



'Active high' input *SR* NOR latch

'Active low' input *SR NAND* latch

Figure 127 Versions of the SR Latch

Astable Multi-Vibrator

An Astable Multi-vibrator has no stable states. It will oscillate between two semi-stable states. The Astable Multi-vibrator constructed from Bi-polar transistors is shown in figure 128.

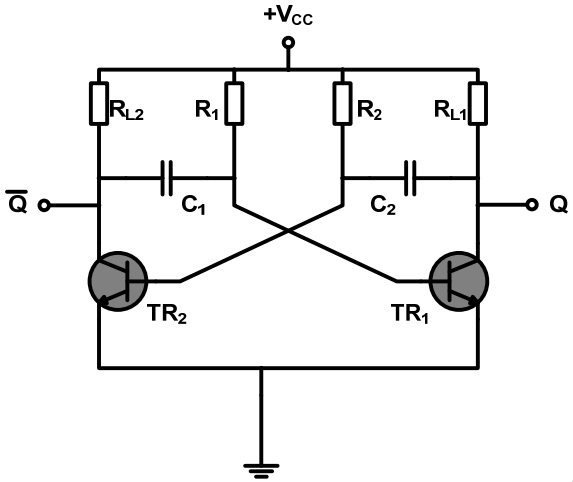


Figure 128 Astable Multi-Vibrator Constructed from Transistors

The Astable Multi-vibrator constructed with Bi-polar transistors is structured using two inverting amplifier arrangements.

The components of one inverting amplifier are R_1, R_{L1} and TR_1. The components of the other inverter are R_2, R_{L2} and TR_2. The inverters are AC coupled into a loop by C_1 and C_2.

In one unstable state, TR_1 will be 'on' while TR_2 will be 'off'. In the other unstable state, TR_2 will be 'on' while TR_1 will be 'off'.

The components C_1 and R_1 determine the time of one unstable state. The components C_2 and R_2 determine the time for the other unstable state.

The characteristic waveform of the Astable Multi-vibrator constructed from Bi-polar transistors is shown in figure 129.

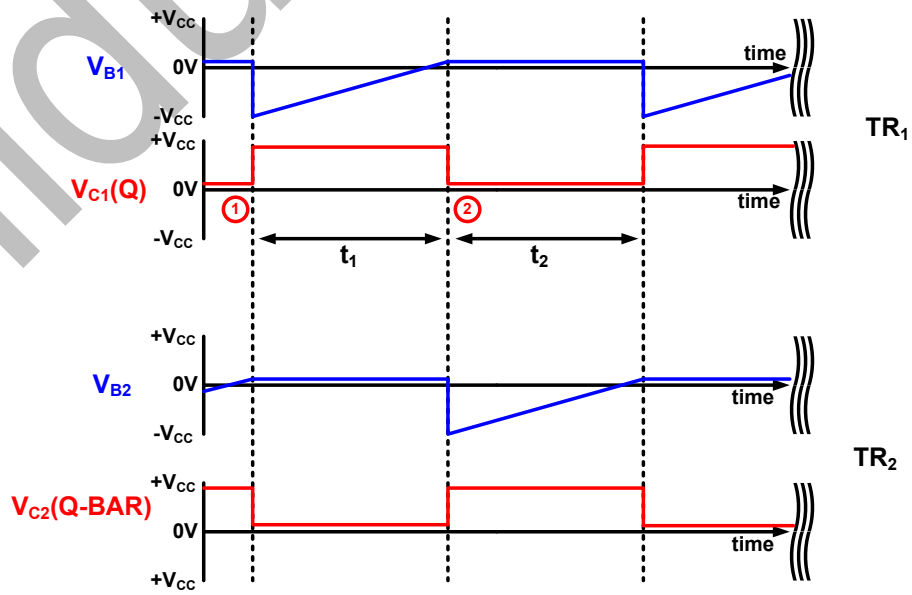


Figure 129 Astable Characteristic Waveform

The circuit has no stable state so an arbitrary entry must be made into the cycle. A useful time for analysis is immediately prior to 'point 1' on the graph (figure 129).

At this point, TR_1 is 'on', and TR_2 is 'off'. Therefore, V_{C1} is at $V_{CE(sat)}$ and V_{C2} is at $+V_{CC}$. V_{B1} is at $V_{BE(sat)}$ and V_{B2} is at ≈0.2V.

Transition at point 1

C_2 charges up via R_2 as V_{C1} is held at $V_{CE(sat)}$. V_{B2} reaches V_{BE} and turns T_{R2} 'on'. V_{C2} (Q-Bar) then drops from $+V_{CC}$ to $V_{CE(sat)}$.

Initially C_1 pulls down V_{B1} by the same magnitude. V_{B1} drops from just above 0V to $-V_{CC}$. TR_1 is now switched off. V_{C1} (Q) then rises to $+V_{CC}$ with the time constant $R_{L1} \times C_2$. TR_2 is then held 'on' by the Base current sustained by R_2.

Duration t1

V_{B2} remains at Saturation ($V_{BE(sat)}$). TR_2 is 'on', so V_{C2} is held at $V_{CE(sat)}$. V_{C1} remains at $+V_{CC}$ while V_{B1} remains at a negative voltage (TR_1 is off).

V_{B1} begins to rise exponentially from almost $-V_{CC}$ towards $+V_{CC}$ as C_1 charges up through R_1. When V_{B1} reaches V_{BE}, TR_1 switches on and the circuit changes states again.

Transition at point 2

C_1 charges up via R_1 as V_{C2} is held at $V_{CE(sat)}$. V_{B1} reaches V_{BE} and turns TR_1 'on'. V_{C1} (Q) then drops from $+V_{CC}$ to $V_{CE(sat)}$.

Initially C_2 pulls down V_{B2} by the same magnitude. V_{B2} drops from just above 0V to almost $-V_{CC}$. TR_2 is now switched 'off'. V_{C2} (Q-Bar) rises to $+V_{CC}$ and is determined by the time constant is $R_{L2} \times C_1$. TR_1 is then held 'on' by the Base current sustained by R_1.

Duration t2

V_{B1} remains at Saturation ($V_{BE(sat)}$). TR_1 is 'on', so V_{C1} is held a $V_{CE(sat)}$. V_{C2} remains at $+V_{CC}$ while V_{B2} remains at a negative voltage (TR_2 is off).

V_{B2} begins to rise exponentially from almost $-V_{CC}$ towards $+V_{CC}$ as C_2 charges up through R_2. When V_{B1} reaches V_{BE}, TR_2 switches 'on' and the circuit changes states.

The duration of each state is shown in equation 10c. As stated in equation 10a, the duration given as $t = 0.693 \times C \times R$.

$$t_1 = 0.693 \times C_1 \times R_1$$
$$t_2 = 0.693 \times C_2 \times R_2$$

Cycle time;

$$Cycletime = t_1 + t_2$$

Frequency;

$$f = \frac{1}{(t_1 + t_2)}$$

Equation 10c: Astable Time Duration

1.3.15 Field Effect Transistor

A Field-Effect Transistor (FET) is a type of transistor that uses an induced electric field to within its structure to control current.

The Junction Field-Effect Transistor (JFET) is a type of field-effect transistor (FET) that operates with a reverse-biased junction to control current in a channel. The JFET can also be referred to as a JUGFET in some texts.

A FET transistor behaves differently to a Bi-polar transistor. The Bi-polar junction transistor (BJT) is a current-controlled device (Base current controls the value of the collector current). The FET is a voltage-controlled device.

The voltage at the Gate terminal controls the value of current flow through the device. The FET also has very high input resistance when compared to the BJT device. This makes it superior in many applications.

The JFET operates with a reverse biased junction to control current in the channel. JFETS can operate in two categories depending on their construction:

- n-channel.
- p-channel.

The two types of JFET are illustrated in figure 130. The basic structure of the n-channel JFET consists of a wire lead connected to each end of the 'n' channel; the Drain is at the upper end, the Source is at the lower end.

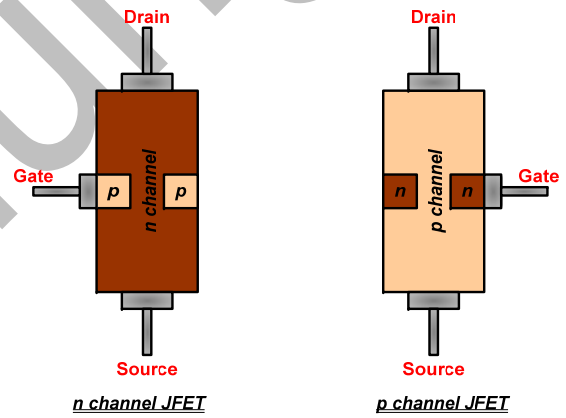


Figure 130 JFET Structure

Two P-Type regions are diffused into the N-type material to form an N-type channel. The P-type regions are connected to the Gate wire lead. For simplicity the Gate is only connected to one of the 'P' regions on the illustration.

JFET Basic Operation

An n-channel device has been used to illustrate the basic JFET operation. This is shown in figure 131.

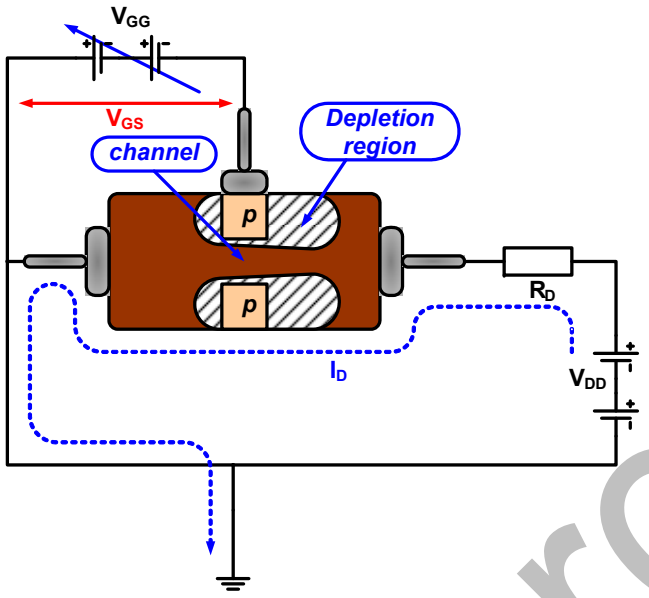


Figure 131 JFET Biased for Conduction

The V_{DD} signal provides the Drain-Source voltage and supplies the current from the Drain to the Source.

The V_{GG} signal sets the reverse biased voltage between the Gate and the Source.

The striped areas surrounding the P regions represent the 'depletion' region created by the effect of reverse bias.

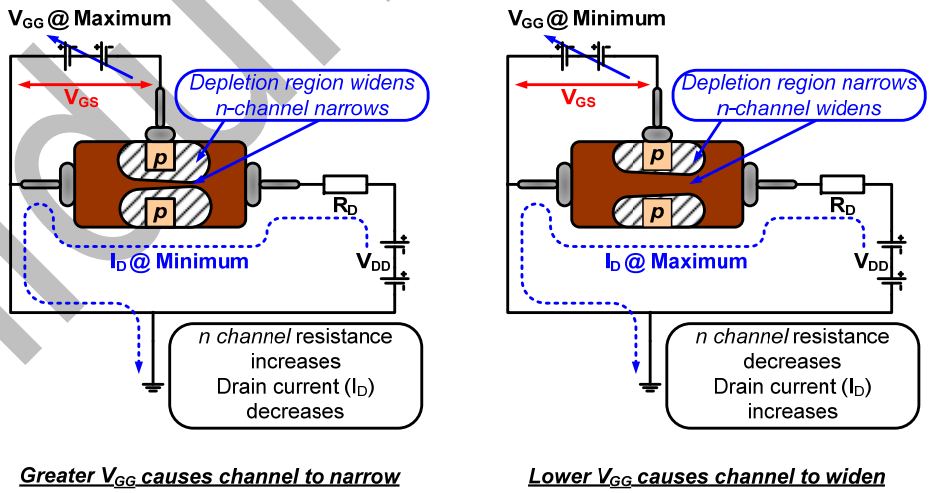


Figure 132 Effects of V_{GG} on Channel Width

The depletion region (normal P-N junction operation) is wider toward the Drain end of the channel because the reverse biased voltage between the Gate and the Drain is greater than that between the Gate and the Source.

The JFET must always be operated with the Gate-Source p-n junction reverse biased.

The effect of biasing the Gate-Source junction with a negative Gate voltage produces a depletion region in the n-channel and increases its resistance. This is illustrated in figure 132.

The channel width can be controlled by varying the Gate voltage. Therefore, the level of Drain current (I_D) can also be controlled.

The circuit symbols for the n-channel and p-channel JFETs are shown in figure 133. The arrow on the diagram points 'in' for the n-channel Gate. The arrow points 'out' on the p-channel Gate.

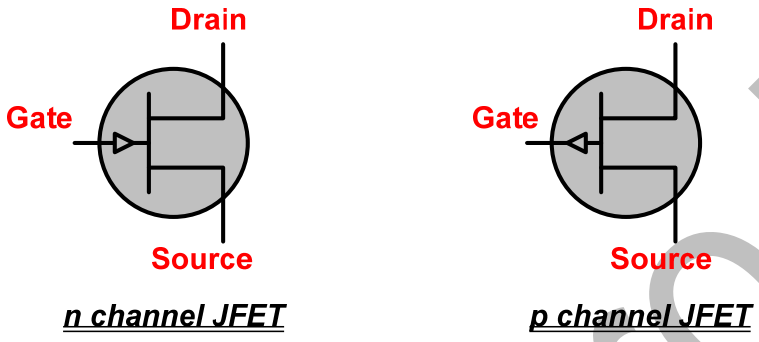


Figure 133 JFET Circuit Symbols

JFET Characteristics

Initially it should be considered that the Gate to Source voltage (V_{GS}) is 0V. This can be simulated by shorting the Gate to the Source. This is shown in figure 134 where both are grounded.

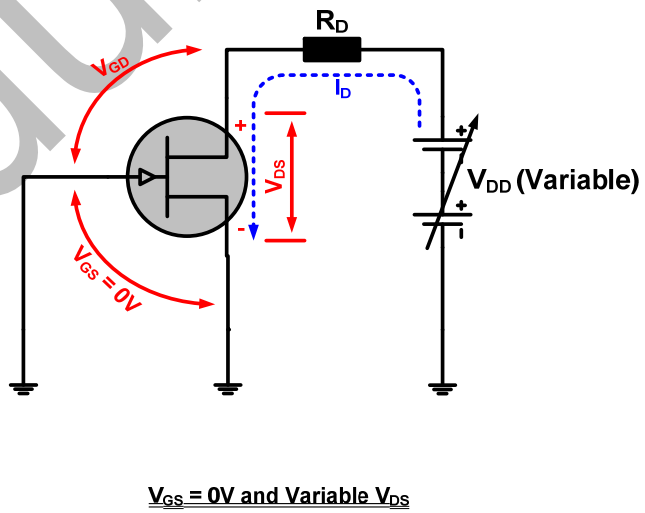


Figure 134 JFET Characteristics with V_{GS} @ 0v and Variable V_{DS} (V_{DD})

As V_{DD} and the Drain to Source voltage (V_{DS}) is increased from 0V, I_D will increase proportionally between point A and point B (as shown in figure 136). This is known as 'The Ohmic region'. In this region, the channel resistance is effectively constant because the depletion region is not large enough to have a

significant effect. In this region, V_{DS} and I_D are related by 'Ohm's law'.

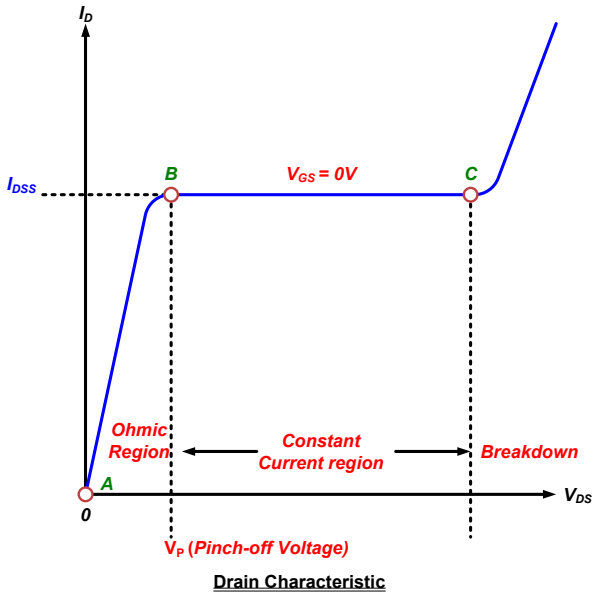


Figure 135 Drain Characteristic Curve

At point B, the curve levels off and I_D becomes constant. As V_{DS} increases from point B to point C, the reverse-biased voltage from the Gate to the Drain (V_{GD}) produces a depletion region large enough to offset the increase in V_{DS}. This keeps I_D relatively constant.

Pinch-off voltage

When V_{GS} = 0V. The value of V_{DS} that causes I_D to become constant (point B on the curve) is the Pinch-off voltage (V_P).

For any given JFET, the value of V_P is fixed. A continued increase in V_{DS} beyond the Pinch-off voltage produces an almost constant drain current. This value of drain current is known as I_{DSS} (Drain to Source with Gate Shorted) and is specified on JFET data sheets. The value of I_{DSS} is the maximum drain current that a specific JFET can produce regardless of the external circuit and must operate with V_{GS} at 0V.

As shown in figure 11f, breakdown occurs at point C when I_D begins to increase very rapidly with any further increase in V_{DS}. Breakdown can result in irreversible damage to the device.

JFETS are always operated below breakdown and within the constant current region (between points B & C).

V_{GS} Controls I_D

A bias voltage, V_{GG}, can be connected from the Gate to the Source. As V_{GG} is increased, V_{GS} becomes increasingly negative (with respect to the Gate). This creates a family of characteristic curves. This is shown in figure 136.

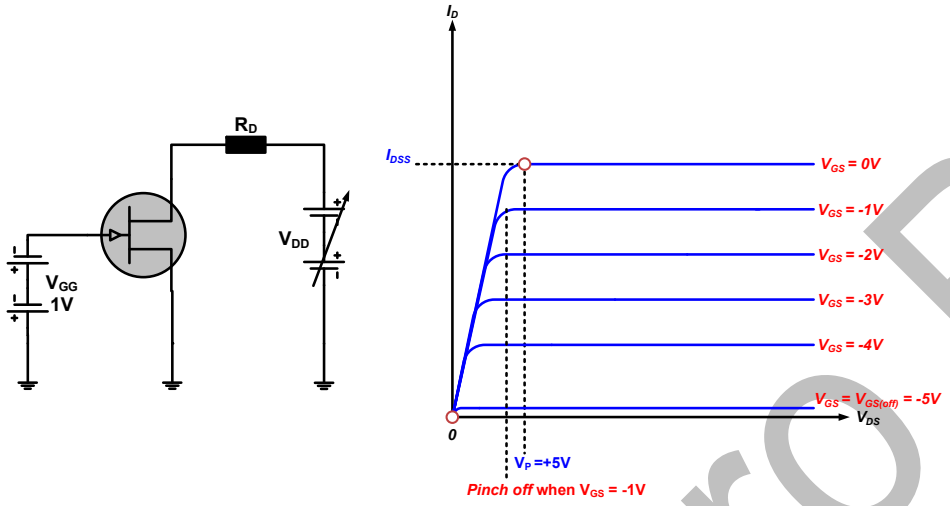


Figure 136 Characteristic Curves due to Increasing V_{GG}

The drain current, I_D, decreases as the negative magnitude of V_{GS} increases. It should be noted that for each increase in V_{GS}, the JFET reaches 'Pinch-off' at values of V_{DS} less than the specified value of V_P. This means that the drain current is also controlled by V_{GS}.

Cut-off Voltage

The value of V_{GS} that makes I_D approximately zero is the cut-off voltage, $V_{GS(off)}$. The JFET must be operated between V_{GS} = 0V and $V_{GS(off)}$.

For this range of Gate-Source voltages, I_D will vary from a maximum of I_{DSS} to a minimum of almost zero.

As shown earlier, for the n-channel JFET, an increasing negative V_{GS} causes a decreasing I_D in the constant current region.

When V_{GS} has a large enough negative value, I_D is reduced to zero.

This Cut-off effect is caused by widening the depletion region to a point where it completely closes the channel. This is shown in figure 137.

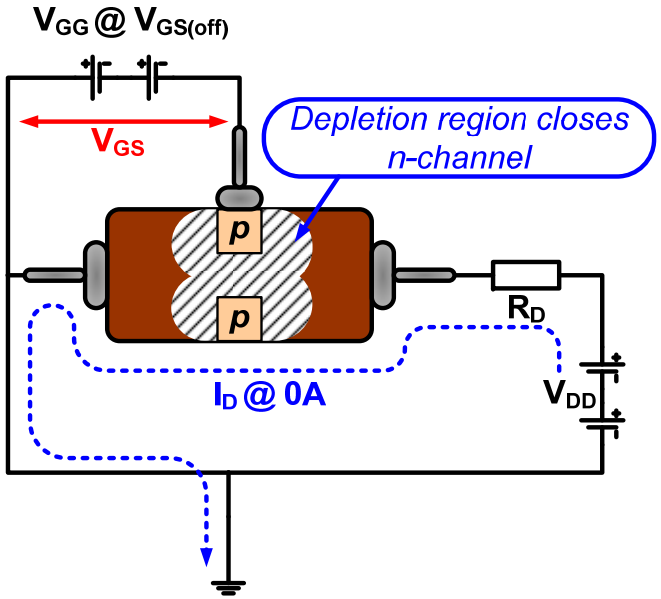


Figure 137 JFET at Cut-off

The basic operation of the p-channel JFET is the same as for an n channel JFET, except that it requires a negative V_{DD} and a positive V_{GS}.

Comparison of Pinch-off and Cut-off

There is a connection between the Pinch-off and Cut-off states. V_P is the value of V_{DS} when the drain current becomes constant and is measured at V_{GS} = 0V.

However, Pinch-off occurs for V_{DS} values less than V_P when V_{GS} is non-zero. This shows that although V_P is constant, the minimum value of V_{DS} at which I_D becomes constant varies with V_{GS}.

The Cut-off voltage, $V_{GS(off)}$, and V_P are always equal in magnitude but opposite in polarity.

A data sheet will usually give either $V_{GS(off)}$ or V_P, but never both. However, if one value is known, the other can be easily determined.

JFET Input Resistance and Capacitance

A JFET operates with its Gate-Source junction reverse-biased. Therefore the input resistance at the Gate is very high. This high input resistance is one advantage of the JFET over the Bi-polar Transistor which operates with a forward biased Base-Emitter junction.

The JFET data sheets will often specify the input resistance by giving a value of the Gate-reverse-current, I_{GSS}, at a certain Gate-Source voltage. Equation 11a shows the method used to calculate the input resistance. It should be noted that polarity is obsolete in a resistance value. This has been indicated using vertical lines surrounding the equation.

$$R_{IN} = \left| \frac{V_{GS}}{I_{GSS}} \right|$$

Equation 11a: Input Resistance

Example:

The 2N3970 data sheets shows a maximum I_{GSS} of 250 pico-Amps (10^{-12}) when V_{GS} = -20V at 25^{O}C.

I_{GSS} increases with temperature so the input resistance decreases.

The input capacitance, C_{ISS}, of a JFET is considerably greater than that of a Bi-polar Transistor because the JFET operates with a reverse-biased p-n junction.

As stated in the previous section, a reverse-biased p-n junction behaves as a capacitor. The capacitance depends on the value of the reverse voltage, which changes the width of the depletion region.

Self-Biasing a JFET

FET biasing is used to select a proper DC Gate-Source voltage to establish a desired value of Drain current.

As demonstrated earlier, the Gate-Source junction must be reverse-biased. This condition requires a negative V_{GS} for an n-channel JFET and a positive V_{GS} for a p-channel JFET. This can be achieved using the 'Self-Bias' arrangement shown in figure 138.

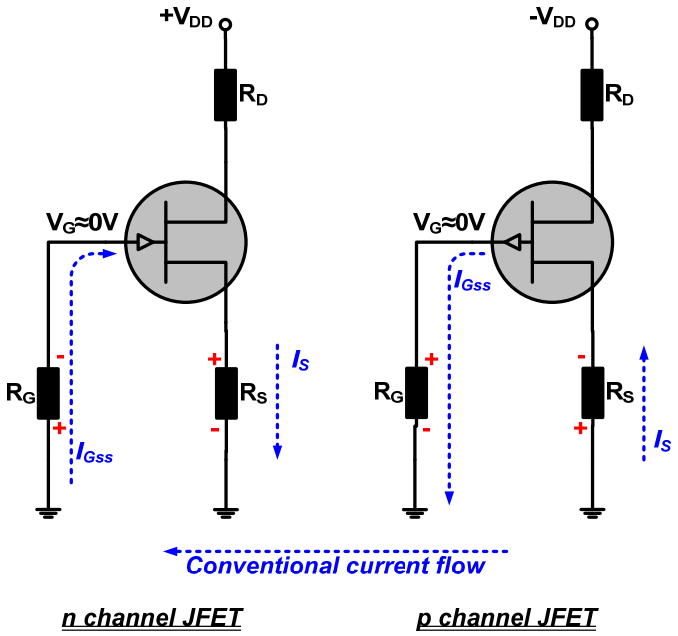


Figure 138 Self Biasing a JFET

The Gate is biased at approximately 0V by resistor, R_G, connected to ground.

The reverse leakage current, I_{GSS}, produces a negligible voltage across R_G, but can be neglected in most cases. It should e assumed that R_G has no voltage drop across it.

For the n-channel JFET, I_S produces a voltage drop across R_S and makes the Source positive with respect to ground. The following should be considered:

- $I_S = I_D$.
- $V_G = 0V$.
- $V_S = I_D \times R_S$.

The Gate-Source voltage can be calculated. This is shown in equation 11b.

$$V_{GS} = V_G - V_S = 0V - I_D \times R_S$$

Thus;

$$V_{GS} = -I_D \times R_S$$

Equation 11b: n-channel Gate-Source Voltage

For the p-channel JFET, the current through R_S produces a negative voltage at the Source, making the Gate positive with respect to the Source. Since $I_S = I_D$, the Gate-Source voltage is shown in equation 11c.

$$V_{GS} = +I_D \times R_S$$

Equation 11c: p-channel Gate-Source Voltage

For the following equations, the n-channel JFET is used for demonstration. The application of the equation is the same for the p-channel JFET, except for the opposite polarity of the signals. The Drain voltage has been determined in equation 11d.

$$V_D = V^{DD} - (I_D \times R_D)$$

Since $V_S = I_D \times R_S$, the Drain-Source voltage is;

$$V_{DS} = V_D - V_S$$

$$V_{DS} = V_{DD} - I_D(R_D + R_S)$$

Equation 11d: Drain Voltage

1.3.16 MOSFETS

The other type of FET device is a MOSFET. The MOSFET is a 'Metal-Oxide-Semi-Conductor' 'Field Effect Transistor.

The MOSFET is different to the JFET because it has no p-n junction structure; instead, the gate of the MOSFET is insulated from the channel using a layer of silicon dioxide (SiO_2). There are two basic types of MOSFET:

- Depletion MOSFET.
- Enhancement MOSFET.

Depletion MOSFET (D-MOSFET)

The Drain and Source of the Depletion MOSFET are diffused into the substrate material and then connected by a narrow channel that is adjacent to the insulated gate. The n-channel and p-channel devices are shown in figure 139.

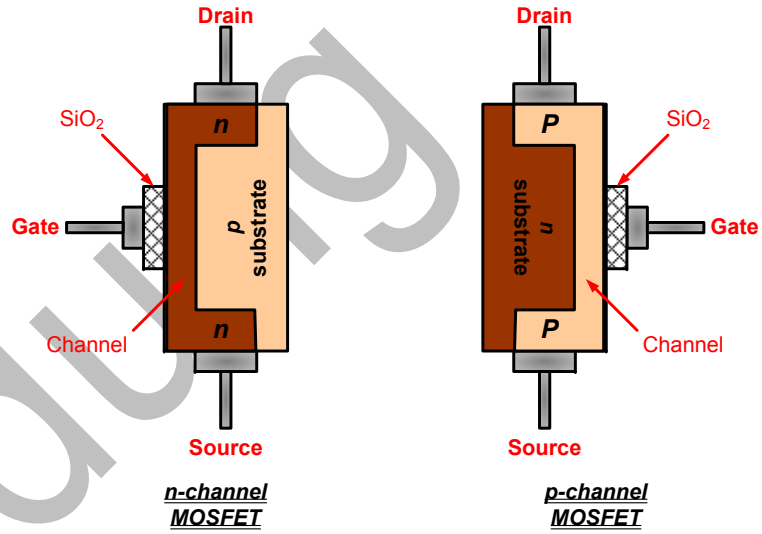


Figure 139 Basic Structure of the D-MOSFET

The n-channel device will be used to demonstrate the basic operation. The p-channel operation is the same, except that the voltage polarities are opposite those of the *n-channel* device.

The D-MOSFET can be operated in two modes:

- Depletion mode.
- Enhancement mode.

Since the Gate is insulated from the channel, either a positive or negative Gate voltage can be applied. The n-channel

MOSFET operates in the 'Depletion mode' when a negative Gate-Source voltage is applied. The n-channel MOSFET operates in the 'Enhancement' mode when a positive Gate-Source voltage is applied.

Depletion Mode

The Gate can be assumed to behave as one plate on a parallel plate capacitor. The channel will behave as the other plate. The silicon dioxide insulating material will behave as he dielectric.

With a negative Gate voltage, the negative charge on the Gate will repel 'conduction electrons' away from the channel, leaving a positive charge (holes) in their place. This means the n-channel is depleted of some of its electrons, so the channel conductivity is decreased.

As the negative value at the Gate increases, the n-channel becomes depleted of more electrons. At a sufficiently high value of negative Gate voltage, $V_{GS(off)}$, the channel becomes completely depleted, causing the Drain current to be zero. The Depletion mode is illustrated n figure 140.

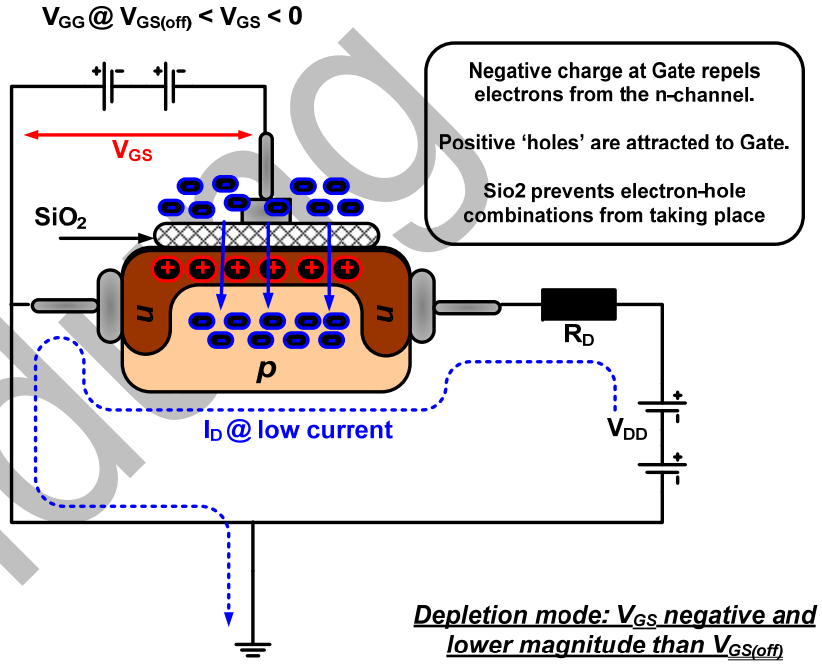


Figure 140 Depletion Mode D-MOSFET Operation

Much like the n-channel JFET, the n-channel D-MOSFET conducts drain current for Gate-Source voltages between $V_{GS(off)}$ and 0V.

Enhancement Mode

Unlike the JFET, the D-MOSFET conducts for V_{GS} above 0V. The Enhancement mode MOSFET is illustrated in figure 141.

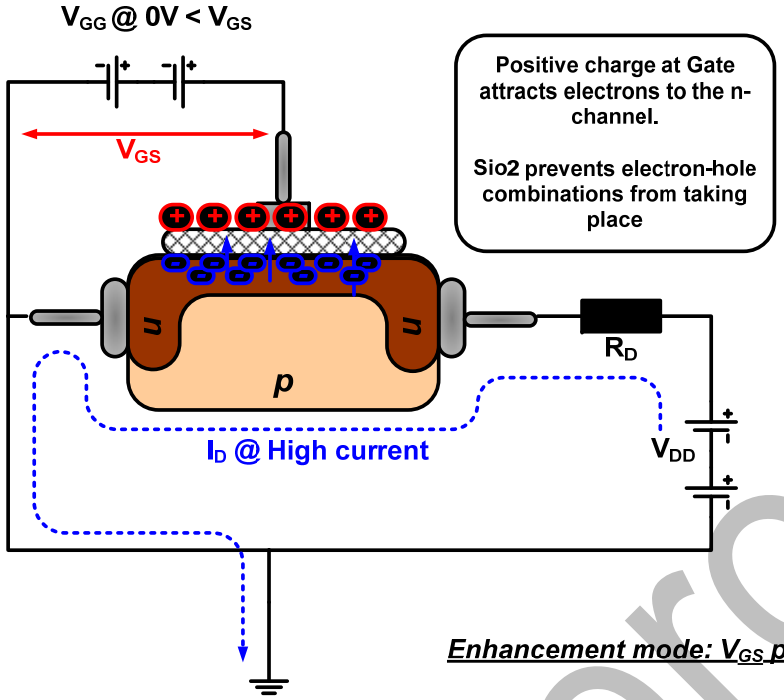


Figure 141 Enhancement Mode D-MOSFET Operation

With a positive Gate voltage, more conduction electrons are attracted into the channel, thus increasing (enhancing) the channel conductivity.

The schematic symbols for both the n-channel and p-channel depletion/enhancement MOSFETS are shown in figure 142. The substrate, indicated by the arrow is usually connected internally to the Source. An inward substrate arrow is for the n-channel MOSFET. An outward arrow is for the p-channel MOSFET.

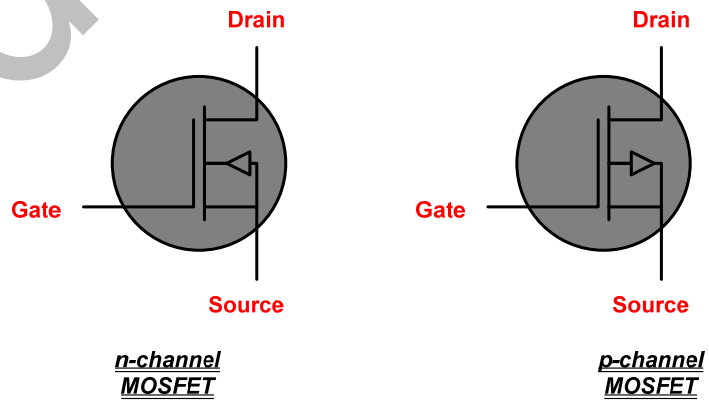


Figure 142 D-MOSFET Symbols

D-MOSFET Biasing

A simple method of D-MOSFET self biasing is to set V_{GS} = 0V to allow an AC signal at the Gate to vary the Gate-Source voltage above and below 0V. This is shown in figure 143.

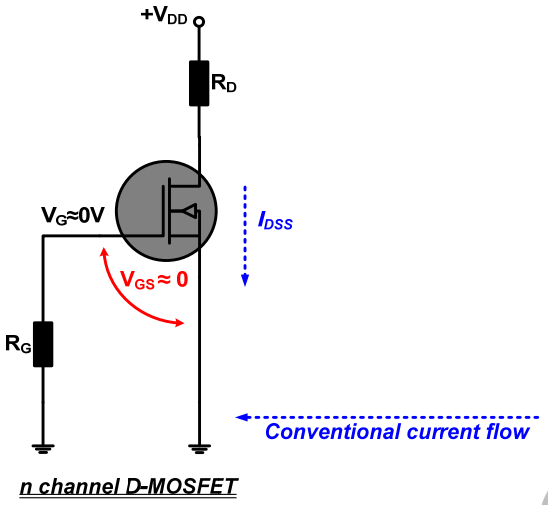


Figure 143 Self-Biasing a D-MOSFET

Since V_{GS} ≈ 0V, then I_D = I_{DSS}. The Drain-Source voltage is expressed in equation 12a.

$$V_{DS} = V_{DD} - (I_{DSS} \times R_D)$$

Equation 12a: D-MOSFET Drain-Source Voltage

Enhancement MOSFET (E-MOSFET)

This type of MOSFET operates only in the Enhancement mode. There is no Depletion mode available.

The E-MOSFET differs in construction from the D-MOSFET because it does not have a physical structural channel. The basic substrate extends completely to the SiO_2 layer. This is shown in figure 144.

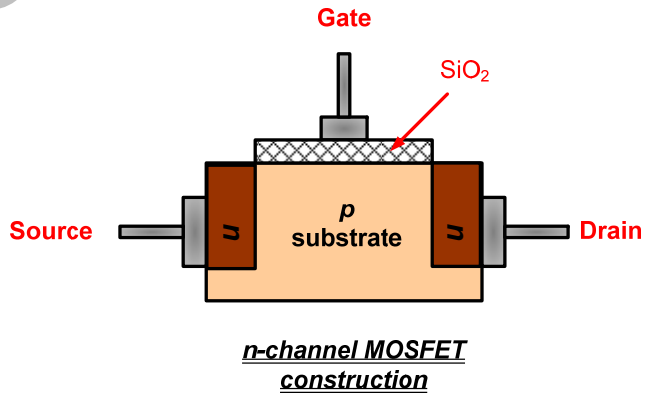


Figure 144 E-MOSFET Construction of an n-channel Device

For an n-channel device, a positive Gate voltage above a threshold value ($V_{GS(th)}$) induces a channel by creating a thin layer of negative charge in the substrate region adjacent to the SiO_2 layer.

The positive Gate voltage repels holes away from the insulating SiO_2 layer and attracts electrons from the Source.

As detailed earlier in the course, an actual current flow is negative to positive. (Conventional current is positive to negative).

The electrons cannot reach the Gate through the SiO_2 layer so they continue through the channel to the Drain (positively charged supply). This is illustrated in figure 145.

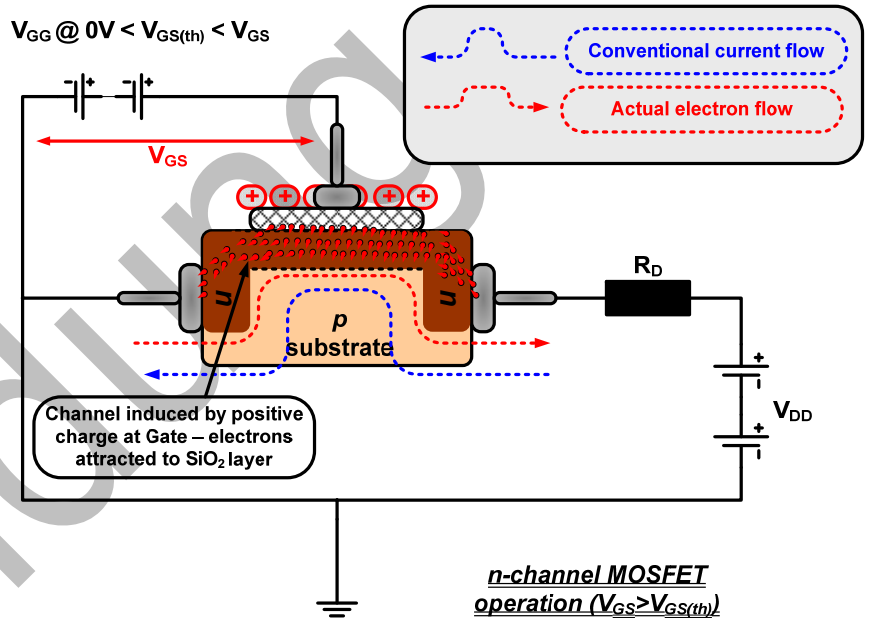


Figure 145 Induced Channel

The conductivity of the channel is enhanced by increasing the Gate-Source voltage, thus pulling more electrons into the channel. If the Gate voltage is below the threshold, there is no channel.

The schematic symbols for the n-channel and p-channel MOSFET are shown in figure 146. The broken lines indicate the absence of a structural channel.

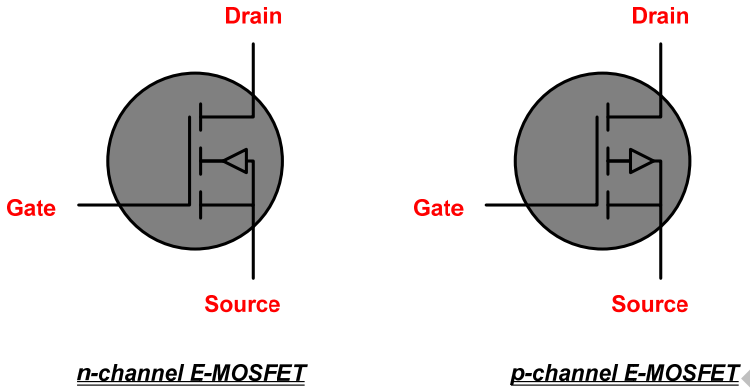


Figure 146 E-MOSFET Schematic Symbols

E-MOSFET Basing

The E-MOSFET must be biased so that the applied V_{GS} is greater than the required threshold $V_{GS(th)}$.

This can be achieved in a similar fashion to the BJT voltage divider bias arrangement when V_B must exceed V_{BE} (0.7V). Figure 147 shows a 'Drain feedback' bias arrangement and a 'Voltage Divider bias arrangement.

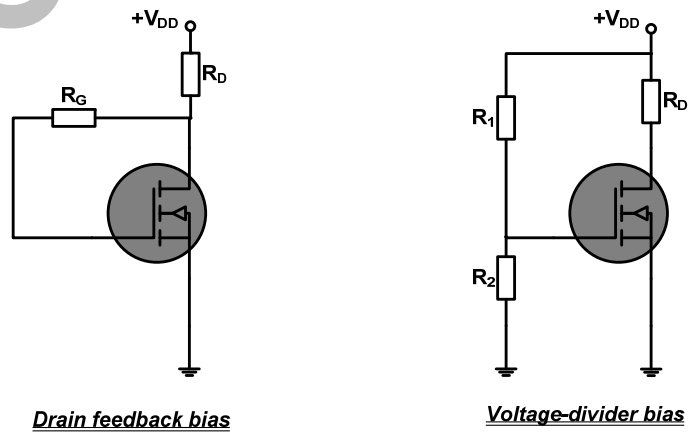


Figure 147 Self-Biasing an E-MOSFET

In the Drain-feedback bias arrangement there is negligible Gate current. Therefore there is no voltage drop across R_G. As a result, the Gate-Source voltage, V_{GS}, is equal to the Drain-Source voltage, V_{DS} (V_{GS}=V_{DS}).

In the voltage divider bias arrangement, the Gate-Source voltage is shown in equation 12b.

$$V_{GS} = \left(\frac{R_2}{R_1 + R_2}\right) \times V_{DD}$$

Equation 12b: E-MOSFET Voltage Divider Bias

The Drain-Source voltage can be calculated using the supply, V_{DD} and the voltage drop across R_D ($V_{RD} = I_D \times R_D$). This is shown in equation 12c.

$$V_{DS} = V_{DD} - (I_D \times R_D)$$

Equation 12c: E-MOSFET Drain-Source Voltage

Considerations

Due to the fact that the Gate is insulated from the channel, the input resistance is extremely high (ideally infinite). The Gate leakage current, IGSS, for a typical MOSFET is measured in the pico-Amp range (10-12). The Gate reverse current for a typical JFET is in the nano-Amp range (10-9).

The input capacitance results from the insulated Gate structure. Excess static charge can accumulate because the input capacitance combines with the very high input resistance and can result in damage to the device due to electro-static discharge (ESD). Normal ESD precautions should be taken when handling MOSFETS.

1.4 INTEGRATED CIRCUITS

1.4.1 Operational Amplifiers

The earliest operational amplifiers were used to conduct basic mathematical operations, such as addition, subtraction, integration and differentiation. They were previously constructed to operate with high voltages.

Current Op-amps are linear integrated circuits (IC's) and use significantly lower dc (direct current) supplies.

Symbol and Terminals

The standard Op-amp has two input terminals, the inverting input (labelled '**-**') and the non-inverting input (**+**). It also has one output terminal. The symbol for an operational amplifier is shown in figure 148.

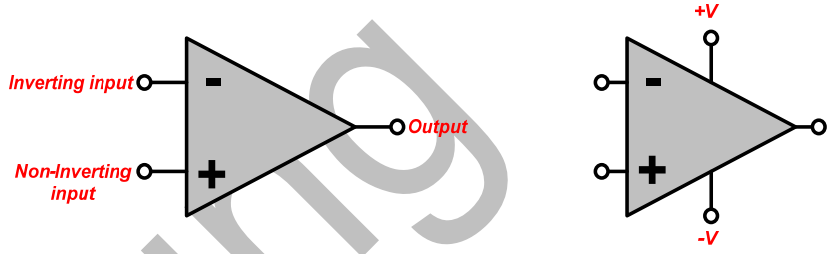


Figure 148 Basic Operational Amplifier Symbols

The typical Op-amp requires two DC supply voltages. One is positive and the other is negative. Usually, the DC supply lines are omitted off schematic symbols for simplicity, but it is always understood that they are present.

The Ideal Op-Amp

The appropriately illustrate the Op-amp, the ideal characteristics must be considered. A practical Op-amp will never meet those standards, but for simplicity it is easier to understand from the 'perfect scenario' perspective.

The ideal Op-amp has infinite voltage gain, infinite bandwidth and infinite input impedance (like an open circuit). Under these conditions it will not load the input driving source. This is illustrated in figure 149.

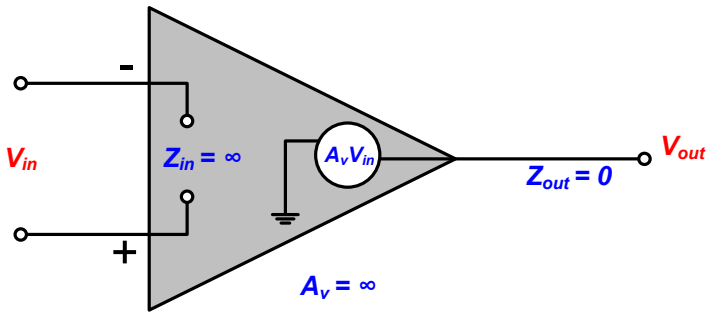


Figure 149 Ideal Op-Amp

The input voltage, V_{in}, appears between the two input terminals. The output voltage is $A_v * V_{in}$. This is indicated by the symbol for the internal voltage source.

The concept of infinite input impedance is a particularly valuable analysis tool for the various Op-amp configurations.

The Practical Op-Amp

Many IC op-amps approach parameter values that can be treated as ideal. However, the perfect ideal device can never be made. The practical Op-amp is shown in figure 150.

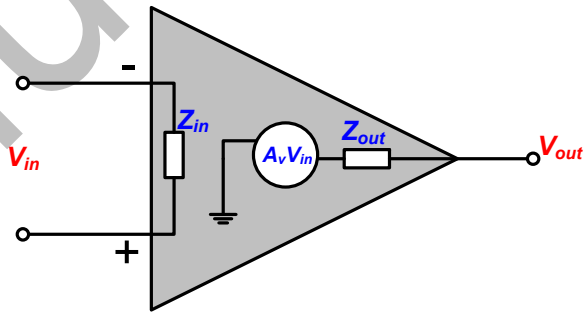


Figure 150 Practical Op-Amp

Peak-peak output voltage is usually limited to just below the two supply voltages. Output current is also limited by internal restrictions such as power dissipation and component ratings. A comparison table for the ideal and practical Op-amp characteristics are shown in figure 151.

OP-AMP COMPARISON TABLE		
	IDEAL	**PRACTICAL**
1	Input resistance = ∞	Very high input resistance (MΩ)
2	Output resistance = 0	Very low output resistance (Ω)
3	Voltage gain = ∞	Very high voltage gain
4	Bandwidth = ∞	Very high bandwidth (KHz/MHz)
5	Output voltage = 0 when V_{IN}+=V_{IN}-, independent of the magnitude of V_{IN}+	Output voltage almost 0 when V_{IN}+=V_{IN}-, increases with larger values of V_{IN}+
6	Characteristics are independent of temperature	Impdeances change with variance in temperature, altering the output currents and voltage gains

Figure 151 Ideal/Practical Comparison Table

The characteristics of a practical op-amp are high voltage gain, low output impedance and wide bandwidth. The basic requirements of an Op-amp are shown in figure 152.

BASIC REQUIREMENTS OF AN OP-AMP	
1	High differential voltage gain
2	DC Coupled
3	Differential input
4	High input impedance (ideally infinite)
5	Low output impedance (ideally zero)
6	High rejection of common-mode input signals and effects including temperature and supply voltage variations
7	Wide output voltage swing (ideally approaching the supply line voltages
8	Wide common-mode input voltage range
9	Should have no tendancy to oscillate under normal feedback situations

Figure 152 Basic Requirements of an Op-Amp

The Differential Amplifier

In its basic form, the Op-amp typically consists of two or more differential amplifier stages. The differential amplifier (diff amp) is fundamental to the Op-amp's internal operation.

The Diff-amp stages that make up part of the Op-amp provide high voltage gain and common-mode rejection (this will be defined later). The differential amplifier has two outputs (compared to the standard Op-amp with one output). This is shown in figure 153.

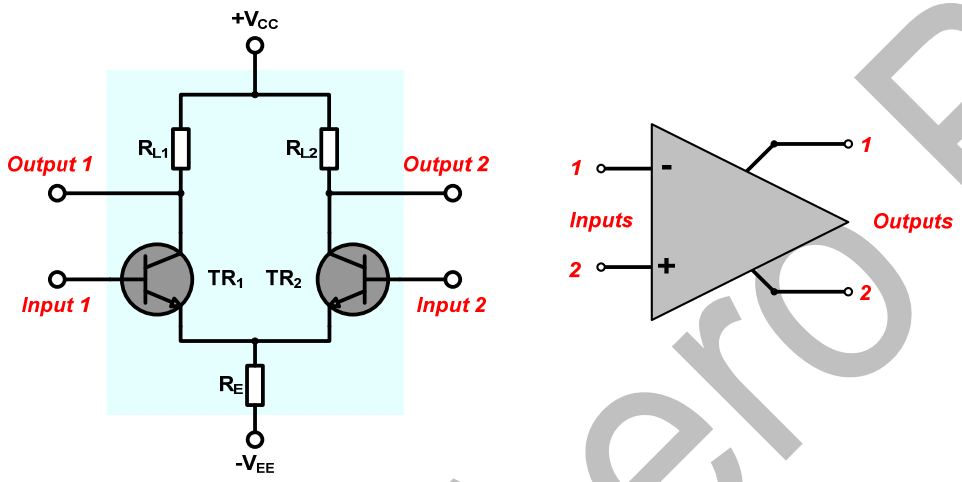


Figure 153 The Basic Differential Amplifier

Basic Operation

Although an Op-amp typically has more than one Diff amp stage, a single Diff amp can be used to demonstrate its operation. The following will discuss the basic DC analysis of the Diff amp operation and the understanding of basic voltage amplifier operation will be applied.

When both inputs are initially grounded (0V) the transistor Emitters of TR_1 and TR_2 are held at -0.7V. This is illustrated in figure 154.

It can be assumed that the transistors are identically matched by careful process control throughout manufacturing so that their DC Emitter currents are the same when there is no input signal.

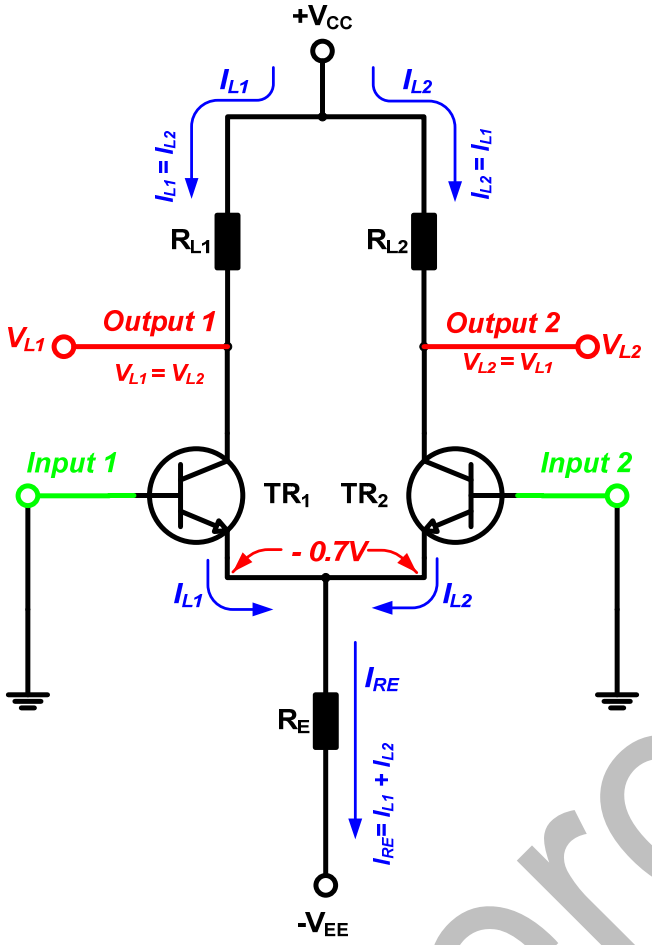


Figure 154 Diff Amp with both Inputs Grounded

The following assumption can then be made;

$$I_{E1} = I_{E2}$$

Both Emitter currents combine (Kirchoff's current law) through R_E, so that;

$$I_{E1} = I_{E2} = \frac{I_{RE}}{2}$$

The I_{RE} term can be determined with ohm's law to be;

$$I_{RE} = \frac{V_E - V_{EE}}{R_E}$$

Based on the approximation that $I_C \approx I_E$, it can be assumed that;

$$I_{L1} = I_{L2} \approx \frac{I_{RE}}{2}$$

Both Collector currents and both collector resistors (R_{L1} & R_{L2}) are equal when the input voltage signal is zero. This is shown in equation 1a.

$$V_{L1} = V_{L2} = V_{CC} - (I_{L1} \times R_{L1})$$

Equation 1a Collector Voltages With 0v Input Signal

It can now be assumed that input 2 remains grounded while a positive bias voltage is applied to input 1.

The positive voltage on the base of TR_1 increases I_{L1} and raises the Emitter voltage. This action reduces the forward bias (V_{BE}) of TR_2 because its Base is held at 0V (ground). This causes I_{L2} to decrease. This is illustrated in figure 156.

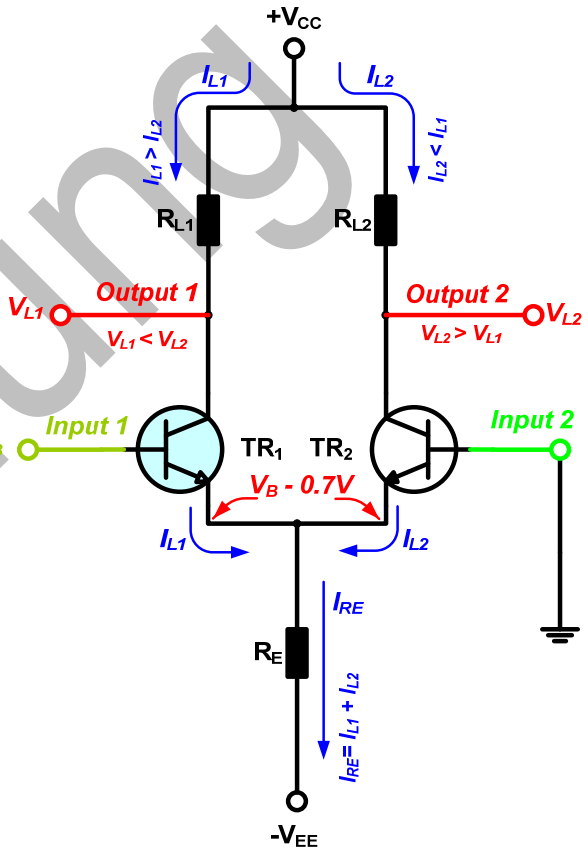


Figure 155 Bias Applied to Input 1, Input 2 Held at Ground

The increase in I_{L1} causes decrease in V_{L1}. The decrease in I_{L2} causes an increase in V_{L2}.

It can now be assumed that input 1 is grounded and a positive bias voltage is applied to input 2. The positive bias causes TR_2 to conduct more and increases I_{L2}, raising the Emitter voltage.

This reduces the forward bias of TR_1 because its Base is held at ground and causes I_{L1} to decrease. This is shown in figure 157.

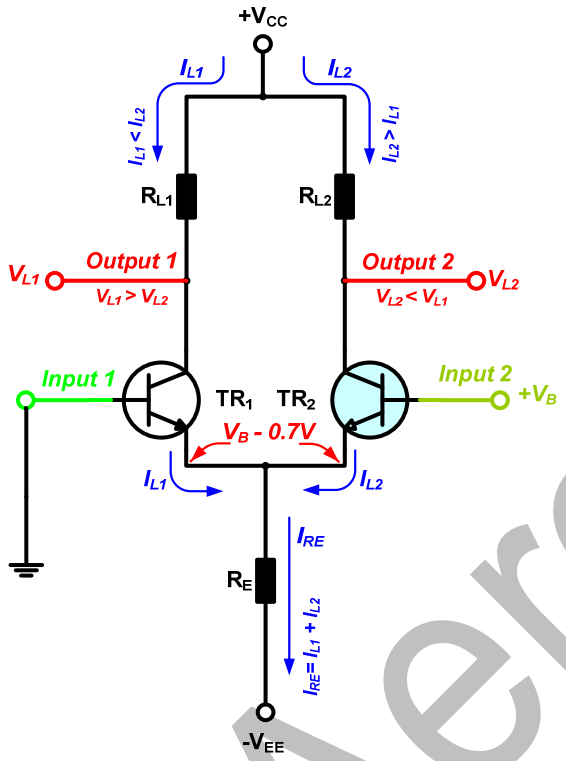


Figure 156 Bias Applied to Input 2, Input 1 is Held at Ground

The increase in I_{L2} causes V_{L2} to decrease. The decrease in I_{L1} causes V_{L1} to increase. The effect is identical to when input 2 is grounded with a bias applied to input 1.

Modes of Signal Operation

There are three main types of input mode;

- Single-ended input.
- Differential input.
- Common-mode input.

Single-ended Input

When a Diff-amp is operated in this mode, one input is grounded and the signal voltage is applied only to the other input.

When the signal voltage is applied to input 1, an inverted, amplified signal voltage appears at output 1.

In addition, a signal voltage appears in-phase at the Emitter of TR_1. Since the Emitters of TR_1 and TR_2 are common, this Emitter signal behaves as an input to TR_2. This will function as a Common-Base amplifier.

The signal is then amplified by TR_2 and appears non-inverted at output 2. This is illustrated in figure 158.

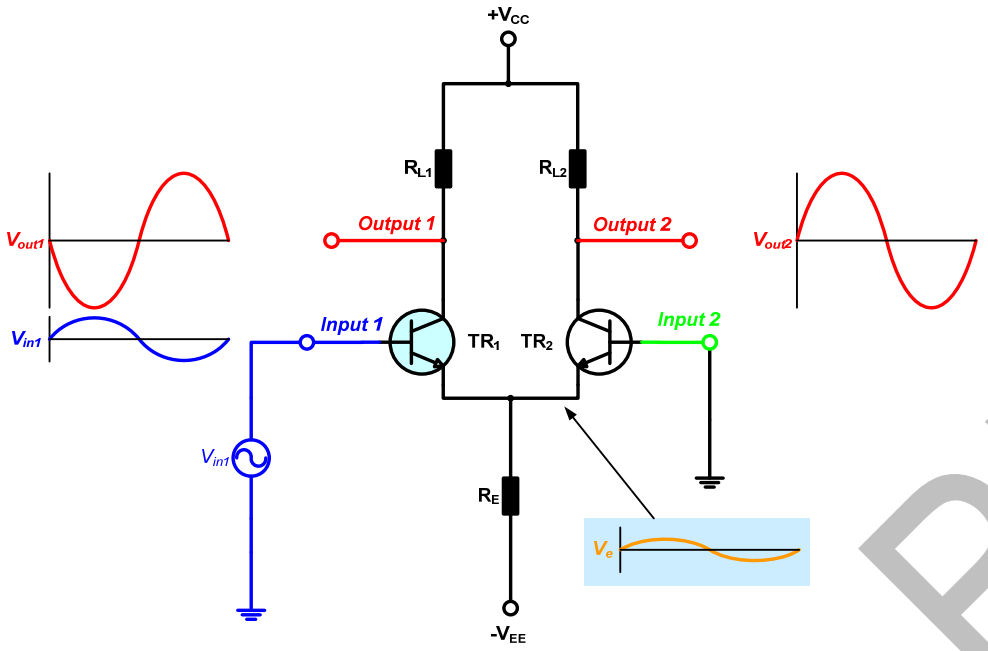


Figure 157 Single-Ended Mode to Input 1

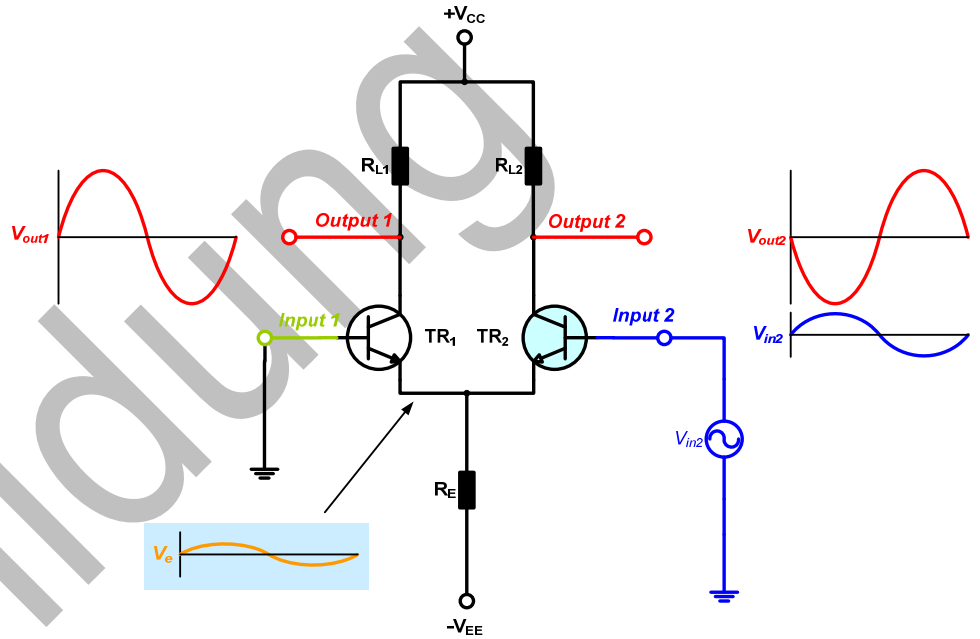


Figure 158 Single-Ended Mode to Input 2

When the signal is applied to input 2 with input 1 grounded, an inverted, amplified signal appears at output 2.

Under this condition, TR_1 acts as the Common-Base amplifier, causing a non-inverted, amplified signal to appear at output 1. This is shown in figure 159.

Differential Input

In this mode, two signals of opposite polarity (anti-phase) are applied to the inputs. This type of operation is known as 'Double-ended'. Each input affects both outputs. This is illustrated in figure 160. It can be illustrated by initially describing each input acting individually.

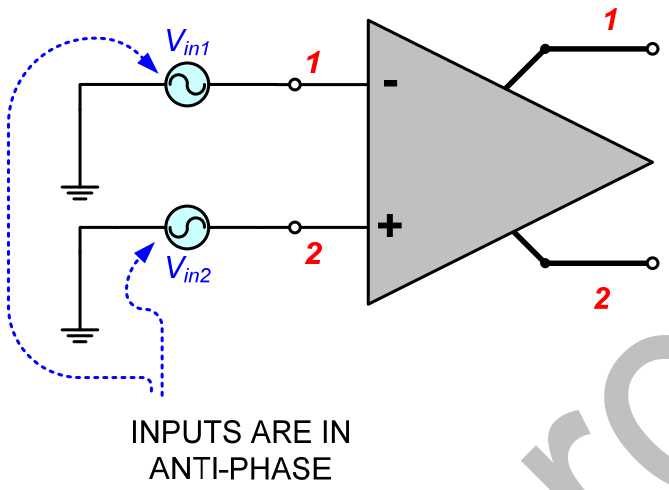


Figure 159 Differential Inputs

The output due to input 1 acting alone as a single-ended amplifier is shown in figure 161. The output due to input 2 acting alone is also shown in figure 161.

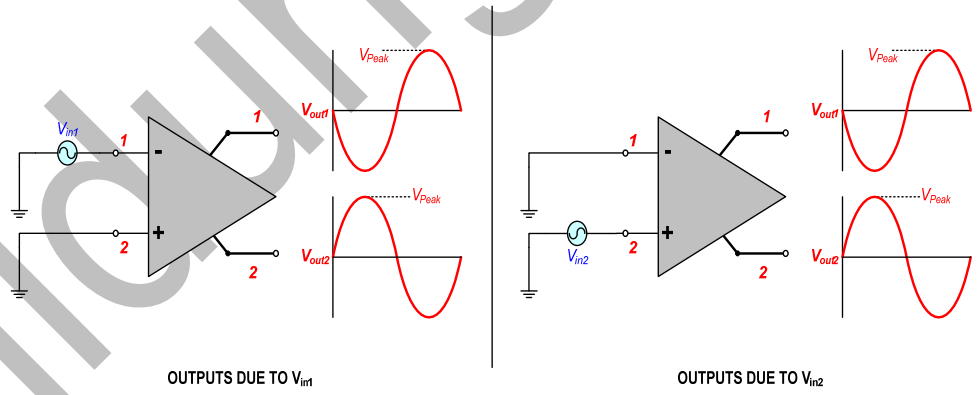


Figure 160 Differential-Mode Inputs Considered Individually

It should be noted in this case that the polarity of Output 1 is identical in both cases. The same is also true for Output 2.

If both Output 1 signals and both Output 2 signals are superimposed, the total differential outputs can be obtained. This is shown in figure 162.

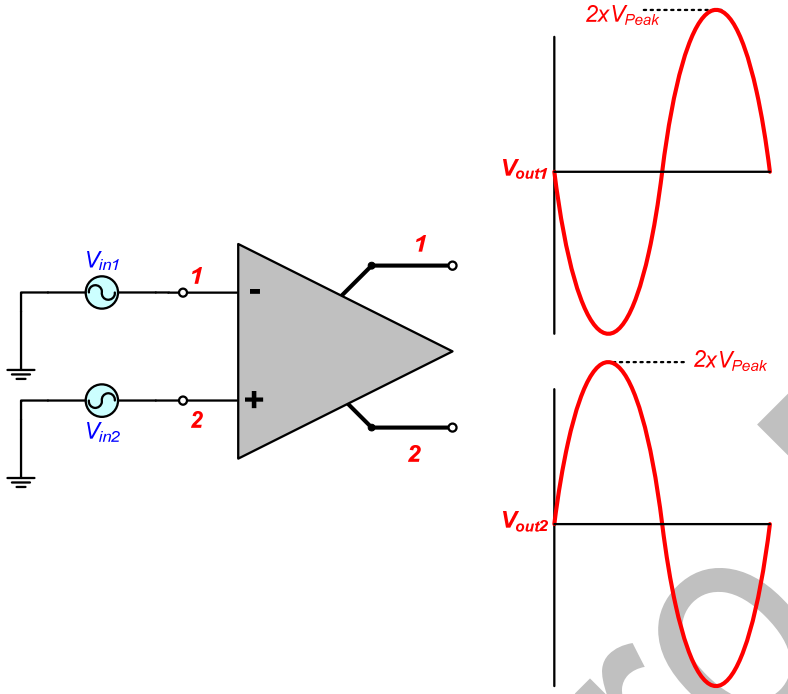


Figure 161 Total Outputs due to Differential Inputs

The new output signal at Output 1 is the sum of Input 1 and Input 2 due to the inverted contribution from Input 1 and the in-phase contribution from Input 2.

The new output signal at Output 2 is the sum of Input 1 and Input 2 due to the In-phase contribution from Input 1 and the inverted contribution from Input 2.

Common-Mode Input

When operating in the common-input mode, the inputs are applied at the same phase, frequency and amplitude. This has been illustrated in figure 164.

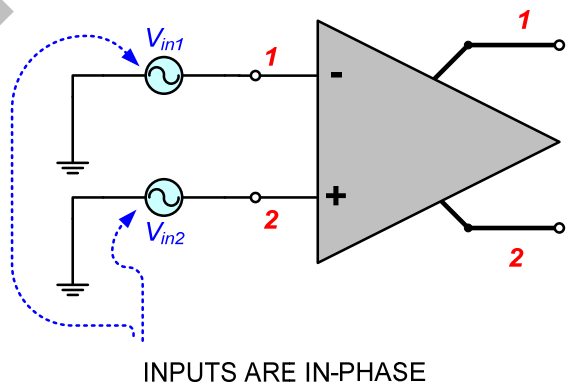


Figure 162 Common-Mode Operation

By considering each input signal applied individually, the operation can be understood. The output signals acting due to the Input 1 applied signal only can be found in figure 164. The output signals acting only due to the Input 2 applied signal can also be found in figure 164.

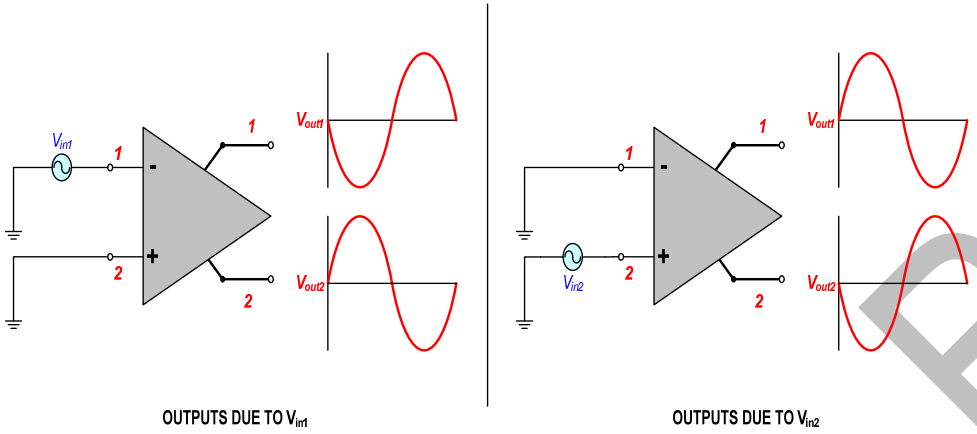


Figure 163 Common-Mode Inputs Considered Individually

It should be noted that the resultant signals on Output 1 are opposite in polarity. This is also true for the signals on Output 2.

When the input signals are applied to both inputs, the outputs can be superimposed and they are found to cancel out. The result of this is a signal on both outputs that is almost zero. This is shown in figure 165.

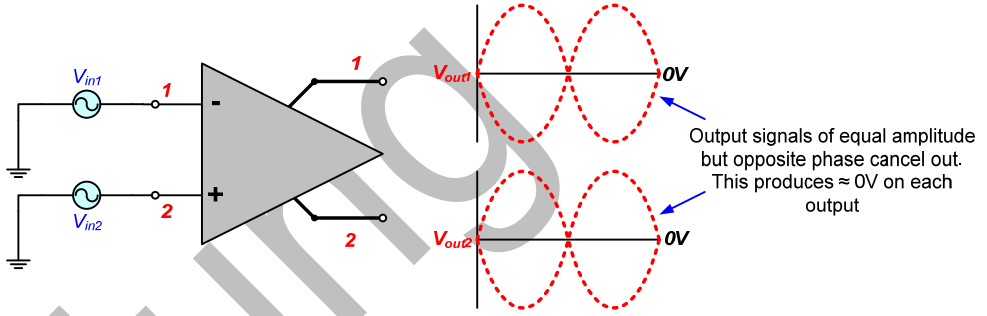


Figure 164 Outputs Cancel when Common-Mode Inputs are Applied

This action is called Common-mode rejection. It is important when unwanted signals appear commonly on both Diff-amp inputs.

Common-mode rejection means that this unwanted signal will not appear on the outputs to distort the desired signal.

Common-mode signals (noise) are generally the result of radiated energy on the input lines from adjacent lines or even the 50Hz power supply.

Common-Mode Rejection Ratio

Desired signals appear on one input only or in anti-phase on both input lines. These desired signals are amplified and appear on the outputs (as previously discussed).

Unwanted signals (noise) appearing with the same polarity on both input lines effectively cancel out and should not appear at the outputs.

The measure of amplifier's ability to reject Common-mode signals is a parameter known as the 'Common-mode rejection ratio' (CMRR).

Under ideal conditions, a Diff-amp provides a very high gain for desired signals (Single-ended or Differential) and a zero-gain for Common-mode signals.

Practical Diff-amps do exhibit a very small Common-mode gain (Usually less than 1), while providing a high Differential voltage gain (Usually several thousand).

The higher the Differential-mode gain (with respect to the Common-mode gain), the better the performance of the Diff-amp in terms of rejection of Common-mode signals.

A good measure of the Diff-amp's performance in rejecting unwanted Common-mode signals is the ratio of the Differential-mode gain ($A_{v[d]}$), to the Common-mode gain ($A_{v[cm]}$). This is the Common-mode rejection ratio (CMRR) and is shown in equation 1b. It should be noted that there are no measurement units because this is a ratio.

$$CMRR = \frac{A_{v[d]}}{A_{v[cm]}}$$

Eq 1b: Common-Mode Rejection Ratio (CMRR)

A higher value of CMRR is desirable. A very high value of CMRR means that the Differential-mode gain, $A_{v[d]}$, is high and the Common-mode gain, $A_{v[cm]}$, is low. The CMRR is often expressed in decibels (*dB*) as shown in equation 1c.

$$CMRR = 20\log_{10}\left(\frac{A_{v[d]}}{A_{v[cm]}}\right)$$

Equation 1c: CMRR (dB)

Example:

A Diff-amp has a differential-mode voltage gain of 2000 and a Common-mode gain of 0.2.

Determine the CMRR and express it in decibels.

Solution;

$A_{v[d]}$ = 2000 and $A_{v[cm]}$ = 0.2, therefore;

$$CMRR = \frac{A_{v[d]}}{A_{v[cm]}} = \frac{2{,}000}{0.2} = 10{,}000$$

Expressed in decibels;

$$CMRR = 20\log_{10}(10{,}000) = 80dB$$

A CMRR of 10,000 means that he desired input signal (Differential) is amplified 10,000 times more than the unwanted noise (Common-mode).

If the amplitudes of the Differential input signal and the Common-mode noise are equal, the desired signal will appear on the output 10,000 greater in amplitude on the output than the noise signal. Therefore, the noise would have been essentially eliminated.

Example:

The Diff-amp shown below has a Differential voltage gain of 2500 and a CMRR of 30,000

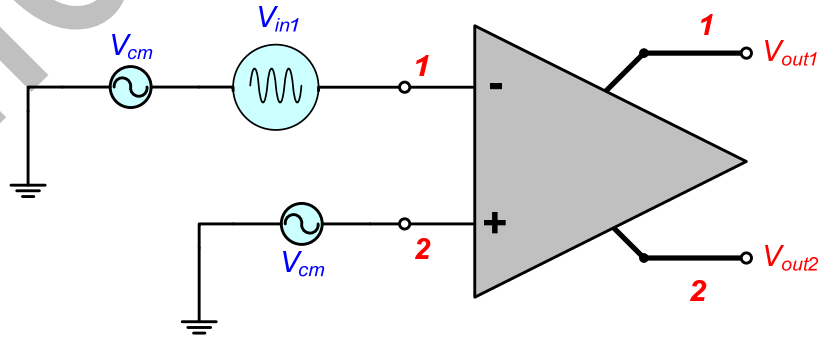


Figure 165 Part (a)

In part (a), a single-ended input signal of 500μV rms is applied. Simultaneously, a 1V, 60Hz Common-mode interference signal appears on both inputs as a result of radiated pick up from the United States' AC mains supply.

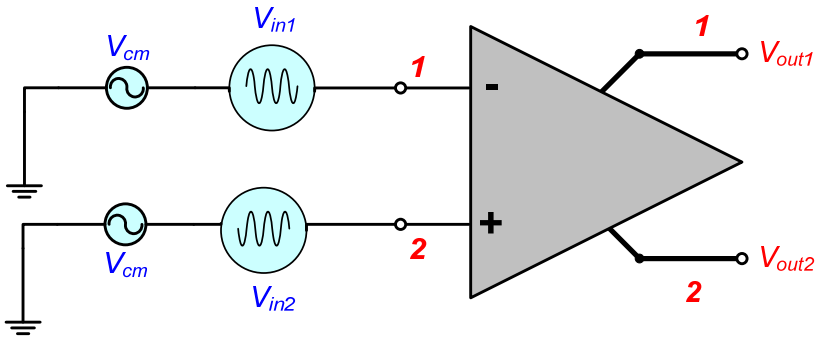


Figure 166 PART (b)

In part (b), Differential input signals of 500μV rms are applied to the inputs. The Common-mode interference is the same as part (a).

- Determine the Common-mode gain.
- Express the CMRR in dB.
- Determine the rms output signal for parts (a) and (b).
- Determine the rms interference voltage on the output.

Solution;

Common-mode gain

Note that;

$$CMRR = \frac{A_{v[d]}}{A_{v[cm]}}$$

Therefore, the Common-mode gain ($A_{v[cm]}$) can be transposed so that;

$$A_{v[cm]} = \frac{A_{v[d]}}{CMRR} = \frac{2500}{30{,}000}0.083$$

CMRR

$$CMRR = 20\log_{10}(30{,}000) = 89.5dB$$

Rms output signal for part (a)

From part (a), the Differential input voltage, $V_{in[d]}$, is the difference between the voltage on Input 1 and that on Input 2. Since Input 2 is grounded, that voltage is 0V. Therefore;

$$V_{in[d]} = V_{in1} - V_{in2} = 500\mu V - 0V = 500\mu V$$

The output signal voltage in this case is taken at Output 1.

$$V_{out1} = A_{v[d]} \times V_{in[d]} = (2500) \times 500\mu V = 1.25Vrms$$

Rms output signal for part (b)

From part (b), the Differential input voltage is the difference between the two anti-phase 500μV signals.

$$V_{in[d]} = V_{in1} - V_{in2} = 500\mu V - (-500\mu V) = 1000\mu V = 1mV$$

The output signal voltage is therefore;

$$V_{out1} = A_{v[d]} \times V_{in[d]} = (2500) \times (1mV) = 2.5Vrms$$

This shows that a differential input (two anti-phase input signals) results in a gain that is double that for a single-ended input.

Rms interference voltage on the output

The Common-mode input is 1V rms. The Common-mode gain, $A_{v[cm]}$, is 0.083. The interference voltage on the output is;

$$A_{v[cm]} = \frac{V_{out[cm]}}{V_{in[cm]}}$$

$$V_{out[cm]} = A_{v[cm]} \times V_{in[cm]} = (0.083) \times (1V) = 83mV$$

An Op-Amp Arrangement

Two Differential amplifiers can be cascaded and connected to an Emitter follower to form a simple Op-amp. This is shown in figure 167.

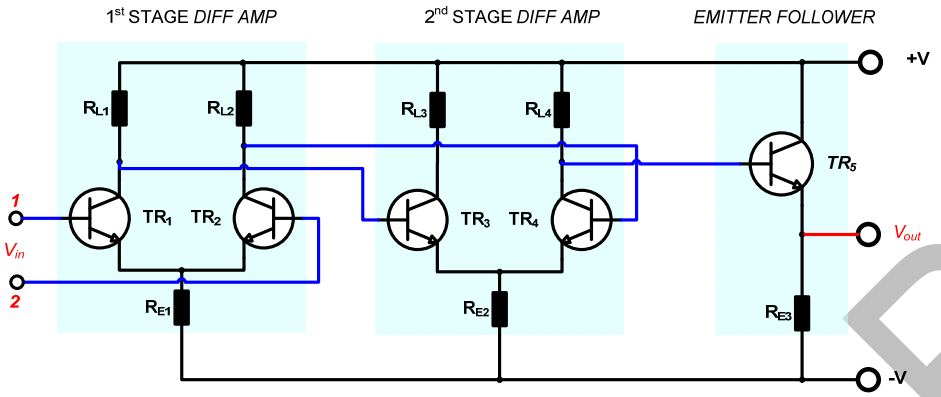


Figure 167 Op-Amp Circuit

The first stage can be applied with either a Single-ended or Differential input arrangement. The Differential outputs of the first stage feed to the Differential inputs of the second stage. The simplified amplifier circuit diagram is shown in figure 168.

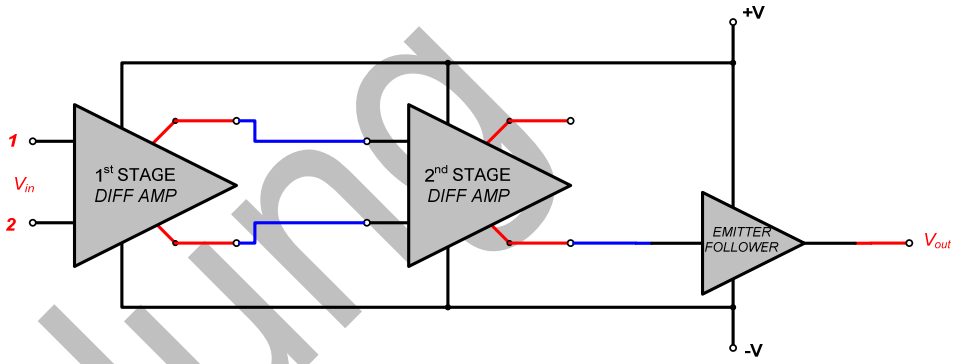


Figure 168 Amplifier Circuit Diagram

The output of the second stage is Single-ended to drive an Emitter-follower to achieve low output impedance. Both differential stages together provide high voltage gain and a high CMRR.

1.4.2 Op-Amp Parameters

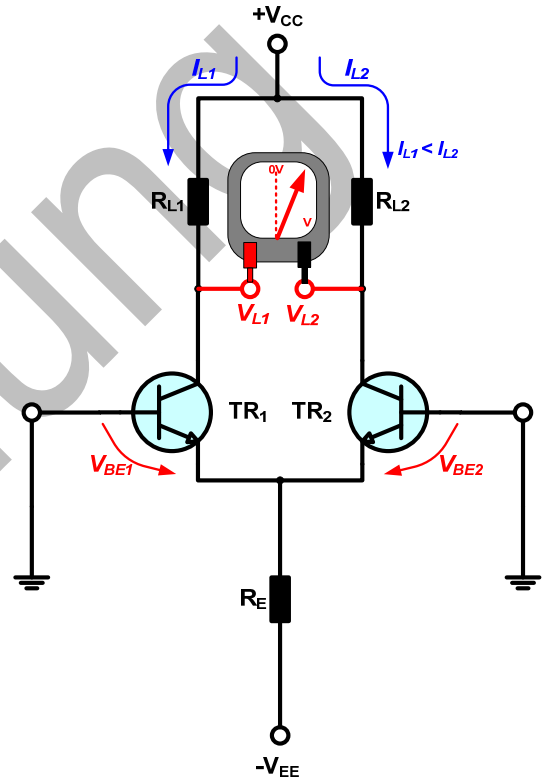
Several Op-amp parameters are defined in this section. These will be listed below;

- Input Offset Voltage.
- Input Bias Current.
- Input Impedance.
- Input Offset Current.
- Output Impedance.
- Common-Mode Rejection Ratio (CMRR).
- Slew rate.

Input Offset Voltage

The ideal Op-amp produces zero volts out for zero volts in. A practical Op-amp, however, even with a small dc voltage, $V_{OUT(error)}$, appears at the output when no Differential input voltage is applied.

Its primary cause is a slight mismatch of the Base-Emitter voltages of the Differential input stage. This is illustrated in figure 169.



A V_{BE} mismatch (V_{BE1} is different to V_{BE2}) causes a small output error voltage

Figure 169 Base-Emitter Mismatch

The output voltage of the Differential input stage is shown in equation 2a. A small difference in the Base-Emitter voltages of T_{R1} and T_{R2} causes a small difference in the Collector currents. This results in a non-zero value of V_{OUT}. It should be remembered that the Collector resistors are equal.

$$V_{OUT} = (I_{L2} \times R_L) - (I_{L1} \times R_L)$$

Equation 2a: Differential Input Stage Output Voltage

As specified on an Op-amp data sheet, the Input Offset Voltage, V_{OS}, is the Differential DC voltage required between the inputs to force the Differential output to zero volts. This is demonstrated on figure 170.

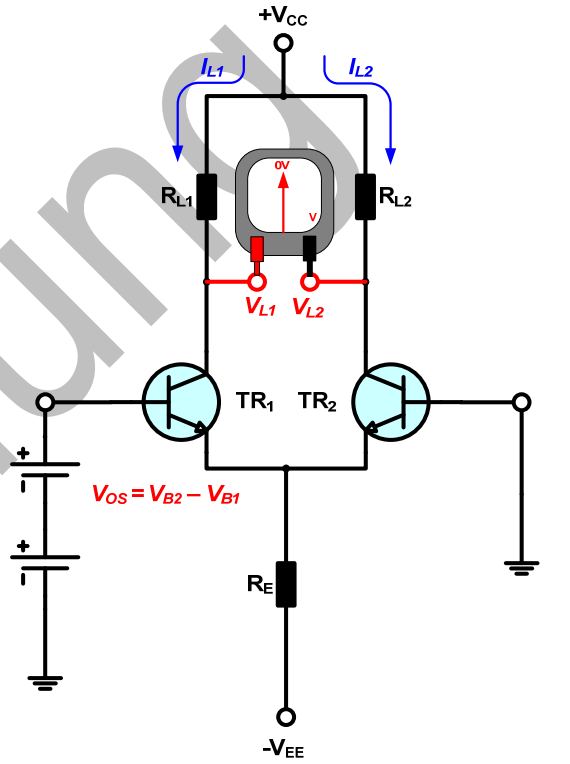


Figure 170 V_{os} Required to Force Vout to Zero Volts

Typical values of Input Offset Voltage are in the range of 2mV or less. The ideal value is 0V.

Input Offset Voltage Drift with temperature

The Input Offset Voltage Drift is a parameter related to V_{OS} that specifies how much change occurs in the Input Offset Voltage for each degree change in temperature.

Typical values range anywhere from about 5µV per degree Celsius. Usually Op-amps with a higher nominal value of Input Offset Voltage exhibit a higher drift.

Input Bias Current

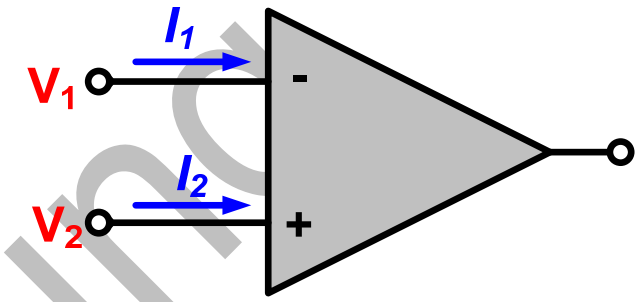
It has been seen that the input terminals of a Diff-amp are the transistor bases. Therefore, the input currents are the Base currents.

The Input bias current is the direct current required by the inputs of the amplifier to properly operate the first stage.

By definition, the Input bias current is the average of both input currents and is calculated in equation 2b. Input bias current is illustrated in figure 171.

$$I_{BIAS} = \frac{I_1 + I_2}{2}$$

Equation 2b: Input Bias Current



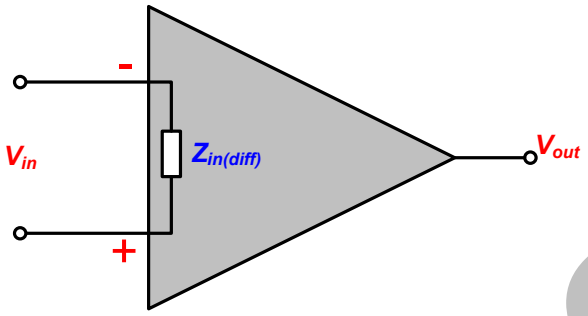
$$I_{BIAS} = \frac{I_1 + I_2}{2}$$

Figure 171 Input Bias Current of an Op-Amp

Input Impedance

There are two basic methods to specify the input impedance. These are specified using the Differential mode and the Common-mode.

The Differential input impedance is the total resistance between the inverting and the non-inverting inputs. This is shown in figure 172.

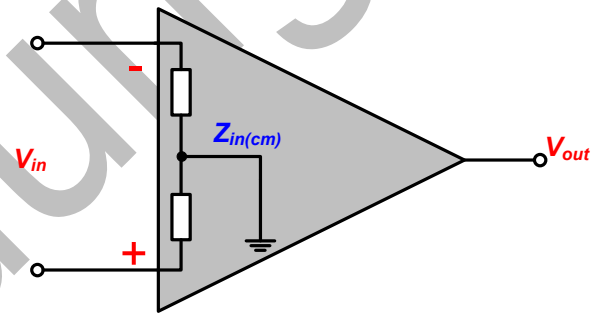


DIFFERENTIAL INPUT IMPEDANCE

Figure 172 Differential Input Impedance

Differential impedance is measured by determining the change in bias current for a given change in Differential input voltage.

The Common-mode input impedance is the resistance between each input and ground. It is measured by determining the change in bias current for a given change in Common-mode input voltage. This is illustrated in figure 173.



***COMMON-MODE* INPUT IMPEDANCE**

Figure 173 Common-Mode Input Impedance

Input Offset Current

Ideally the two input bias currents are equal, so their difference will be zero. In a practical Op-amp the input currents can never be exactly equal.

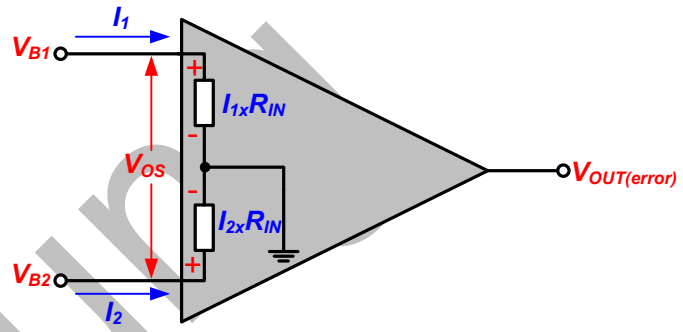
The Input offset current is the difference between the input bias currents. This has been expressed in equation 2c.

$$I_{OS} = |I_1 - I_2|$$

Equation 2c: Input Offset Current

The actual magnitudes of offset current are usually ten times less than the bias current. In many applications, the offset current can be ignored because it is negligible.

High gain, high impedance amplifiers should have tiny values of I_{OS} because the difference in currents through large input resistances will cause a substantial offset voltage. This is shown in figure 174.



EFFECT OF *INPUT OFFSET CURRENT*

Figure 174 Effect of Input Offset Current

The offset voltage developed by the Input offset current is shown in equation 2d.

$$V_{OS} = (I_1 \times R_{IN}) - (I_2 \times R_{IN}) = (I_1 - I_2) \times R_{IN}$$
$$V_{OS} = I_{OS} \times R_{IN}$$

Equation 2d: Offset Voltage Developed by Input Offset Current

The error caused by I_{OS} is amplified by the gain Av of the Op-amp and appears at the input as shown in equation 2e.

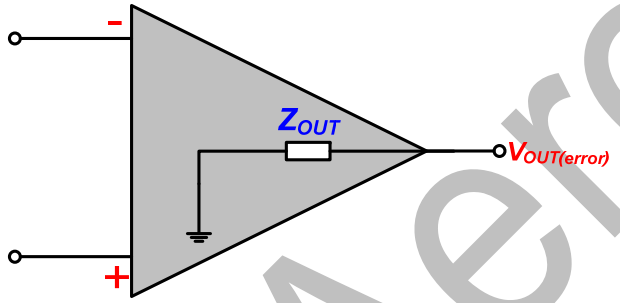
$$V_{OUT(error)} = A_v \times I_{OS} \times R_{IN}$$

Equation 2e: Error Caused By I_{os}

A change in the offset current with temperature also affects the error voltage. Values of temperature coefficient for the offset current in the range of 0.5nA/^{0}C are common.

Output Impedance

The Output impedance is the resistance viewed from the output terminal to of the Op-amp. This can be seen in figure 175.



OP-AMP *OUTPUT* IMPEDANCE

Figure 175 Output Impedance

Common-Mode Input Voltage Range

All Op-amps have limitations on the range of voltages over which they will operate. The Common-mode input voltage range is the range of input voltages which, when applied to both inputs, will not cause clipping or other output distortion.

Many Op-amps have Common-mode input voltage ranges of ±10V with DC supply voltages of ±15V.

Open-Loop Voltage Gain, A_{ol}

The Open-loop voltage gain of the Op-amp is the internal voltage gain of the device and represents the ratio of output voltage to input voltage when there are no external components.

The Open-loop voltage gain is set entirely by the internal design. Open-loop voltage gain can range up to 200,000. Data sheets often refer to the Open-loop voltage gain as the large signal voltage gain.

Common-Mode Rejection Ratio

The Common-mode rejection ratio (CMRR) has been previously discussed when discussing the Diff amp. As stated, it is a measure of an Op-amp's ability to reject Common-mode signals.

An infinite value of CMRR means that the output is zero when the same signal is applied to both inputs (Common-mode).

An infinite CMRR cannot be practicably achieved. A good Op-amp does have a very high value of CMRR. Common-mode signals are undesirable interference voltages (such as the 50Hz mains supply). A high CMRR enables the Op-amp to eliminate these interference signals from the output.

The accepted definition of CMRR from an Op-amp is the open-loop voltage gain, A_{ol}, divided by the Common-mode gain, A_{cm}. This is commonly expressed in decibels (*dB*). This is shown in equation 2f.

$$CMRR = \frac{A_{ol}}{A_{cm}}$$

$$CMRR(dB) = 20\log_{10}[\frac{A_{ol}}{A_{cm}}]$$

Equation 2f: CMRR in dB

Slew Rate

The maximum rate of change of the output voltage in response to s step input voltage is the Slew rate of the Op-amp. The Slew rate is dependant on the Frequency response of the amplifier stages within the Op-amp.

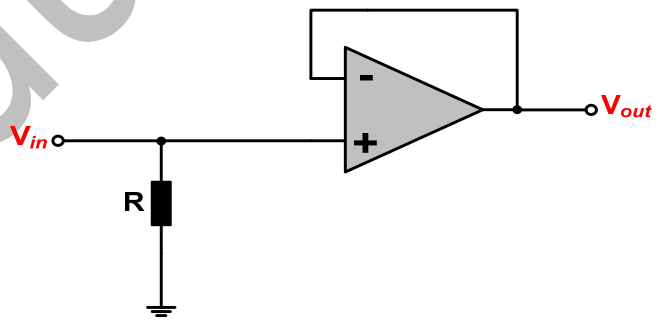
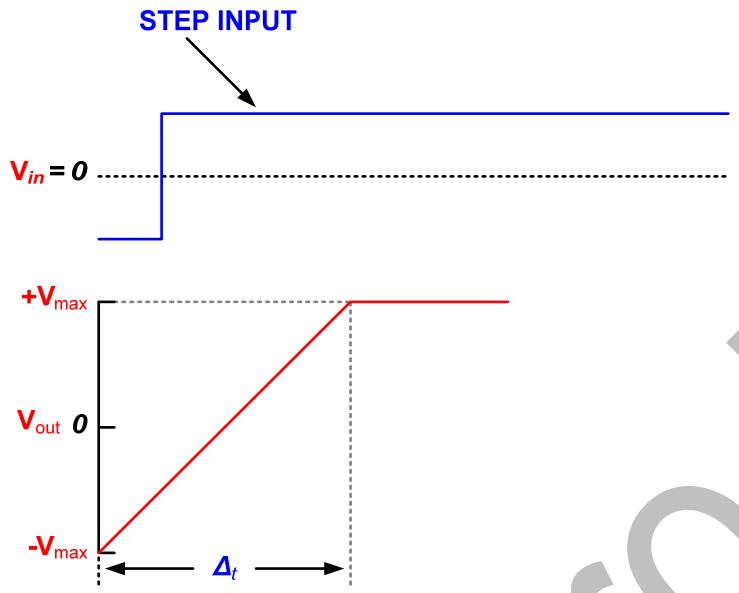


Figure 176 Measurement of Slew Rate

Slew rate is measured with an Op-amp connected as shown in figure 176. The connection to this Op-amp is designed for 'Unity' gain (Gain of 1). It is a Non-inverting configuration and will be discussed later. It depicts the slowest Slew rate (worst case scenario).

A pulse is applied to the input so that the ideal output voltage can be measured. The width of the input pulse must be sufficient to allow the output to 'Slew' from its lower limit towards its upper limit. This is shown in figure 177.



STEP INPUT VOLTAGE AND RESULTING OUTPUT VOLTAGE

Figure 177 Applied Step Input to Measure Slew Rate

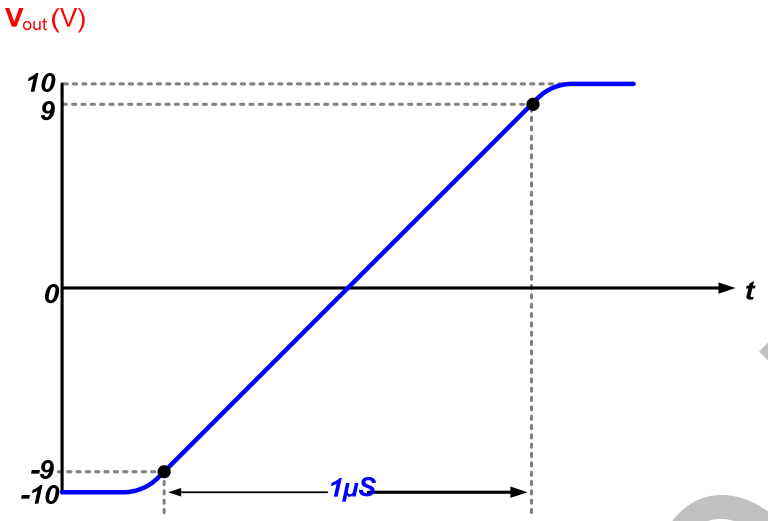
It can be seen that a time interval, Δ_t, is required for the output voltage to change from its lower limit, $-V_{max}$, to its upper limit, $+V_{max}$, when the step input is applied. The Slew rate is expressed as shown in equation 2g, where $\Delta V_{OUT}=+V_{max}-(-V_{max})$. The unit of the *Slew rate* is in Volts/micro-second (V/µS).

$$Slewrate = \frac{\Delta V_{OUT}}{\Delta_t}$$

Equation 2g: Slew Rate

Example:

The output voltage of a certain Op-amp appears as shown below in response to a step input;



Solution;

The output goes from the lower limit to the upper limit in 1μS. Since this response is not ideal, the limits can be taken at the 90% (as indicated). The upper limit becomes +9V and lower limit becomes -9V. The Slew rate is;

$$Slewrate = \frac{\Delta V_{OUT}}{\Delta_t} = \frac{+9V - (-9V)}{1\mu S} = 18V / \mu S$$

Frequency Response

The internal amplifier stages that make-up an Op-amp have voltage gains limited by junction capacitances. Although the Diff amps used in Op-amps are different from the basic amplifiers, the same principles apply. An Op-amp has no internal coupling capacitors so the lower frequency response will extend down to 0Hz.

Negative Feedback

Negative feedback is a very useful concept in Op-amp applications. A portion of the output voltage of an amplifier is returned to the input with a phase angle that opposes (or subtracts from) the input signal. Negative feedback is illustrated in figure 178.

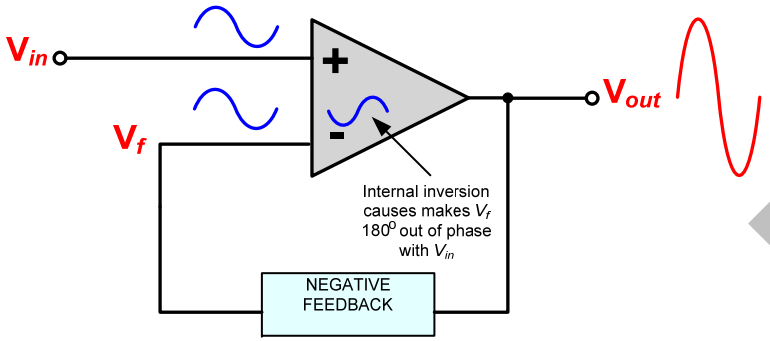


Figure 178 Negative Feedback

The inverting input (-) makes the feedback signal 180^0 out of phase with the input signal. The Op-amp has extremely high gain and amplifies the difference in the signals applied to the inverting and non-inverting inputs.

A tiny difference in these two signals is all the Op-amp requires to produce the output. When negative feedback is present, the non-inverting and inverting inputs are almost identical.

The functions of negative feedback and the reasons behind the identical signals will now be explored.

It should be assumed that a 1V input signal is applied to the non-inverting terminal. The open loop gain of the Op-amp is 100,000. The amplifier responds to the voltage at its non-inverting input terminal, causing the output to climb towards saturation.

Immediately, a fraction of this output is returned to the inverting terminal through the feedback connection.

If the feedback signal were to reach 1V there would be nothing left for the Op-amp to amplify. Therefore the feedback signal attempts to match the input signal (Although it never exactly matches it).

The gain is controlled by the amount of feedback used. When troubleshooting an Op-amp circuit with negative feedback, it should be noted that the two inputs will look identical on an oscilloscope, but are in fact slightly different.

It can now be assumed that something happens to reduce the internal gain of the Op-amp. This causes the output signal to reduce slightly. This will return a smaller signal to the inverting input through the feedback connection. The difference between the signals is larger than it was previously.

The output now increases to compensate for the original drop in gain. The net change in the output is so small that it can barely be measured. This illustrates that any variation in the amplifier is immediately compensated for by the negative feedback, resulting in a stable predictable output.

Why Use Negative Feedback?

As shown, the inherent open-loop gain of a typical Op-amp is very high (usually greater than 100,000). Therefore, a small difference in the two input voltages will drive the Op-amp into its saturated output states.

The Input offset voltage of the Op-amp can also drive it into saturation. For example, if V_{in} is assumed to be 1mV and A_{ol} = 100,000, then;

$$V_{in} \times A_{ol} = (1mV) \times (100{,}000) = 100V$$

Since the output level of an Op-amp cannot reach 100V, it is driven into saturation. The Op-amp is limited to its maximum output levels for both positive and negative input voltages of 1mV. This is shown in figure 4b. The use of the Op-amp in this configuration is restricted and is usually limited to comparator applications.

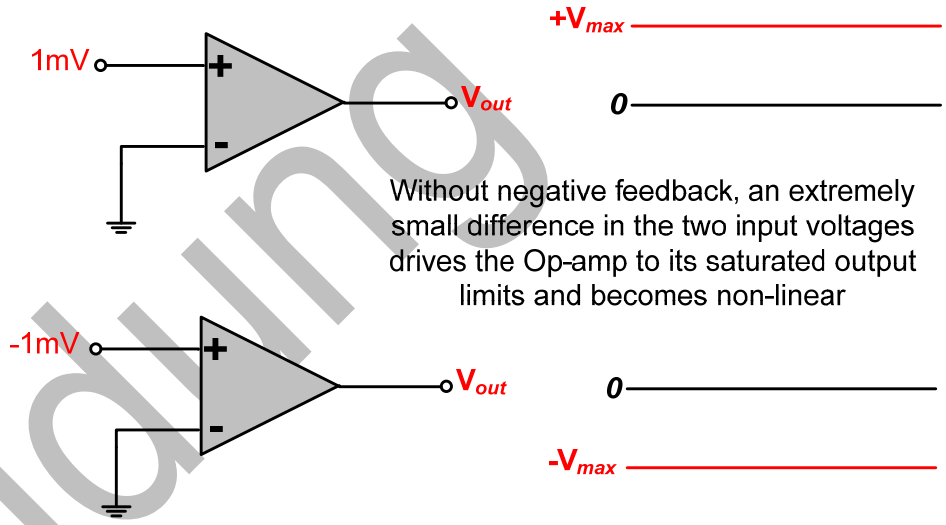


Figure 179 Op-Amp without Feedback

With negative feedback, the closed-loop voltage gain (A_{cl}) can be reduced and controlled so that the Op-amp can function as a linear amplifier.

In addition to providing a controlled, stable voltage gain, negative feedback also provides control of the input and output impedances. It can also be used to control bandwidth. The effects of negative feedback are shown in figure 180.

	VOLTAGE GAIN	INPUT IMPEDANCE (Z_{IN})	OUTPUT IMPEDANCE (Z_{OUT})	BANDWIDTH
WITHOUT NEGATIVE FEEDBACK	A_{ol} is too high for linear amplifier applications	Relatively high	Relatively low	Relatively narrow because the gain is so high
WITH NEGATIVE FEEDBACK	A_{cl} is set to desired value by the feedback network	Can be increased or reduced to a desired value depending on the type of circuit	Can be reduced to a desired value	Significantly wider

Figure 180 Effects of Applied Negative Feedback

Configurations with Negative Feedback

There are three basic ways that an Op-amp can be connected using negative feedback. It can be used to stabilize the gain and frequency response.

The extremely high open-loop gain of an Op-amp creates an unstable situation because a small noise voltage on the input can be amplified to a point where the output is driven out of its linear region.

Unwanted oscillations can also occur. The open-loop gain can also vary significantly from one device to another.

Negative feedback takes a portion of the output and applies it back out of phase with the input causing a reduction in gain. This closed-loop gain is usually much less than the open-loop gain and independent of it.

Closed Loop Voltage Gain, A_{cl}

The Closed-loop voltage gain is the voltage gain of an Op-amp with negative feedback. The amplifier configuration consists of the Op-amp and an external feedback network that connects the output to the inverting input. The Closed-loop voltage gain is then determined by the component values in the feedback network and can be precisely controlled by them.

1.4.3 Non-Inverting Amplfier

An Op-amp connected in a closed-loop configuration as a non-inverting amplifier is shown in figure 181. The input signal is applied to the non-inverting input (+).

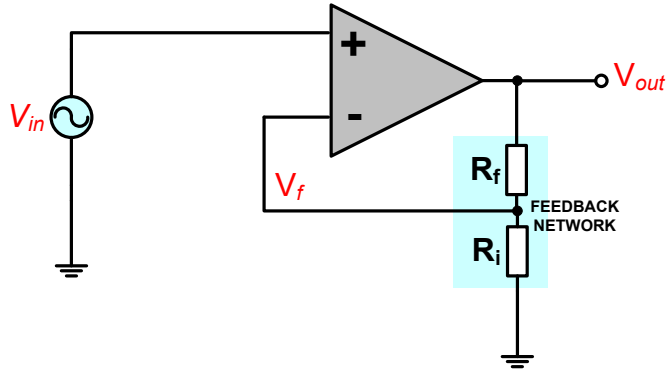


Figure 181 Non-Inverting Amplifier

A portion of the output is applied back to the inverting input (-) through the feedback network. This constitutes negative feedback.

The feedback fraction, B, is determined by R_f and R_i which provide a voltage divider network. The attenuation of the feedback network is the portion of the output returned to the inverting input. This determines the gain of the amplifier. The smaller feedback voltage, V_f, can be written;

$$V_f = \left(\frac{R_i}{R_i + R_f}\right) \times V_{out} = B \times V_{out}$$

The differential voltage, V_{diff}, between the Op-amp's input terminals is illustrated in figure 182. This voltage can be expressed as;

$$V_{diff} = V_{in} - V_f$$

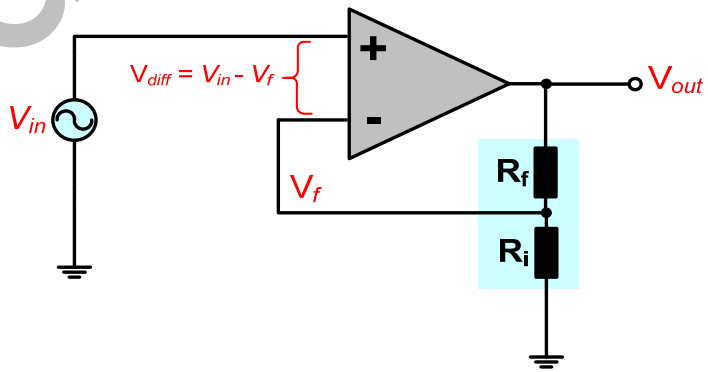


Figure 182 Differential Input, V_{in}-V_f

The input differential voltage is forced to be very small as a result of the negative feedback and the high open-loop gain, A_{ol}. Therefore a close approximation is;

$$V_{in} \cong V_f$$

By substitution;

$$V_{in} \cong B \times V_{out}$$

This can be transposed so that;

$$\frac{V_{out}}{V_{in}} \cong \frac{1}{B}$$

The ratio of the output voltage to the input voltage is the closed-loop gain. This result shows that the closed-loop gain for the non-inverting amplifier, $A_{cl(NI)}$, is approximately;

$$A_{cl(NI)} = \frac{V_{out}}{V_{in}} \cong \frac{1}{B}$$

The fraction of the output voltage, V_{out} that is returned to the inverting input is found by applying the voltage-divider rule to the feedback network.

$$V_{in} \cong B \times V_{out} \cong (\frac{R_i}{R_f + R_i}) \times V_{out}$$

This can be transposed so that;

$$\frac{V_{out}}{V_{in}} = (\frac{R_f + R_i}{R_i})$$

This can be expressed further as follows;

$$A_{cl(NI)} = \frac{R_f}{R_i} + 1$$

This shows that the closed-loop voltage gain, $A_{cl(NI)}$, of the non-inverting (NI) amplifier is not dependant on the Op-amp's open-loop gain but can be set by selecting values of R_i and R_f.

This equation is based on the assumption that the open-loop gain is very high compared to the ratio of the feedback resistors, causing the input differential voltage, V_{diff}, to be very small. When an equation is needed, the output voltage can be expressed as;

$$V_{out} = V_{in}\left(\frac{A_{ol}}{1+(A_{ol}\times B)}\right)$$

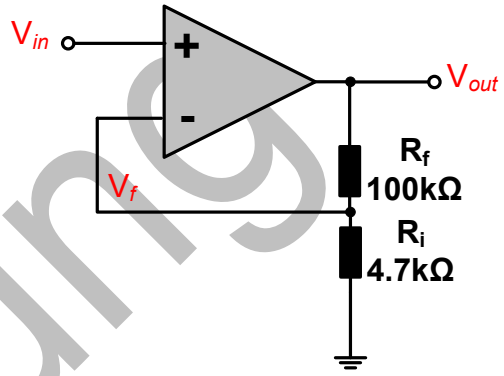
The exact solution is shown in equation 3a.

$$A_{cl(NI)} = \frac{V_{out}}{V_{in}} = \left(\frac{A_{ol}}{1+A_{ol}B}\right)$$

Equation 3a: Closed-Loop Gain of a Non-Inverting Amplifier

Example:

Determine the closed-loop voltage gain of the amplifier below, the open-loop voltage gain is 100,000;



Solution;

This is a non-inverting Op-amp configuration; therefore, the closed loop voltage gain is;

$$A_{cl(NI)} = \frac{R_f}{R_i} + 1 = \frac{100k\Omega}{4.7k\Omega} + 1 = 22.3$$

$$V_{out} = V_{in}\frac{R_f}{R_i} + 1$$

1.4.4 Voltage Follower

The voltage-follower configuration is a special case of the non-inverting amplifier whereby the output voltage is fed back to the inverting input (-) by a direct connection. This is shown in figure 183.

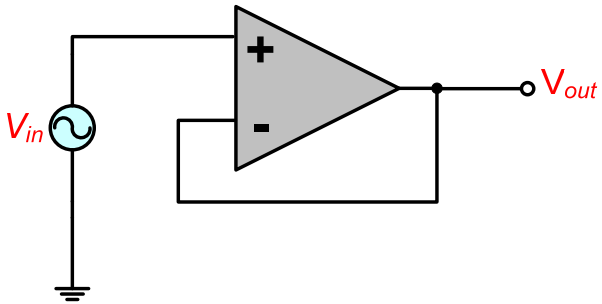


Figure 183 Op-Amp Voltage-Follower

The direct connection has a voltage gain of 1. The closed-loop voltage gain of a non-inverting amplifier is $\frac{1}{B}$ as previously derived. In this case, B = 1 so the closed-loop gain of the voltage-follower is;

$$A_{cl(VF)} = 1$$

The most important features of the voltage-follower are its very high input impedance and its low output impedance. These features make the voltage-follower an ideal buffer-amplifier for interfacing high impedance sources and low impedance loads.

.

1.4.5 Inverting Amplifier

An Op-amp connected as an inverting amplifier with a controlled amount of voltage gain is shown in figure 184. The input signal is applied through a series input resistor (R_i) to the inverting input (-). The output is fed back through R_f to the inverting input. The non-inverting input (+) is grounded.

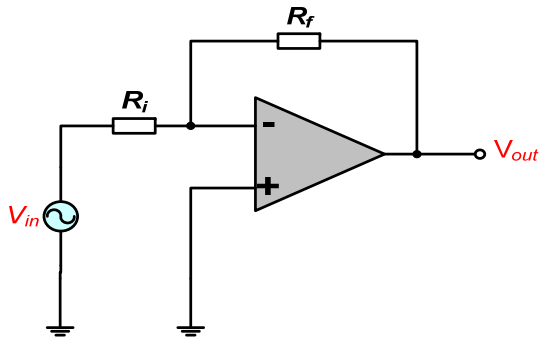


Figure 184 Inverting Amplifier

At this point, the ideal Op-amp parameters discussed earlier are useful in simplifying the analysis of the circuit. The concept of infinite input impedance is of great value.

Infinite input impedance implies that there is no conventional current into the inverting input. If there is no current through the input impedance, then there can be no voltage drop between the inverting and non-inverting inputs.

This means that the voltage at the inverting input (-) is zero because the non-inverting input (+) is grounded. This zero voltage at the inverting input terminal is described as virtual ground. This is illustrated in figure 185.

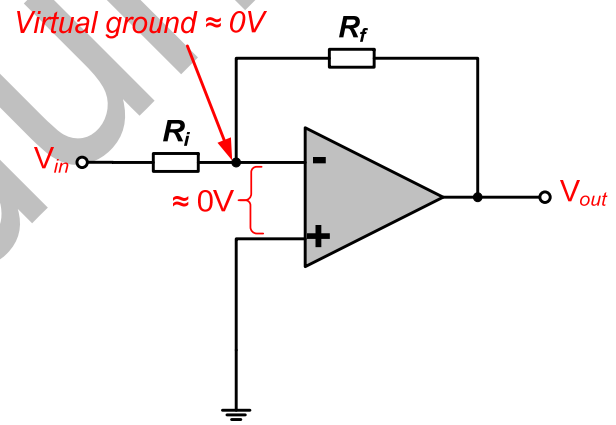


Figure 185 Virtual Ground

Since there is no current at the inverting input, the current through R_i and R_f are equal. This is illustrated in figure 186.

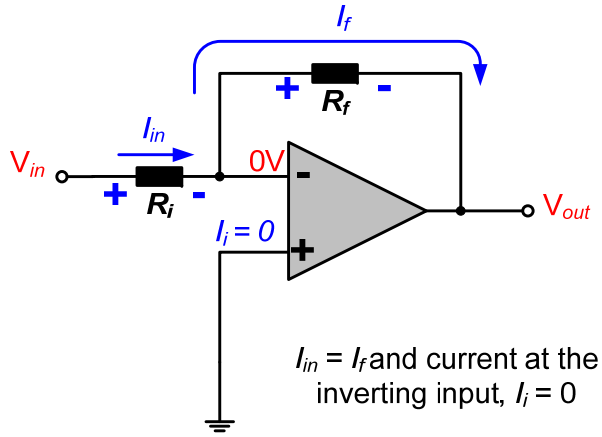


Figure 186 Current at the Inverting Input

It can be expressed that;

$$I_{in} = I_f$$

The voltage across R_i equals V_{in} because of virtual ground on the other side of the resistor. Therefore;

$$I_{in} = \frac{V_{in}}{R_i}$$

Also, the voltage drop across R_f equals $-V_{out}$ because of virtual ground. Therefore;

$$I_f = -\frac{V_{out}}{R_f}$$

Since $I_f = I_{in}$;

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

This can be transposed so that;

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

As expected, $\frac{V_{out}}{V_{in}}$ is the overall gain of the inverting amplifier. The closed-loop gain of the inverting amplifier is shown in equation 3b.

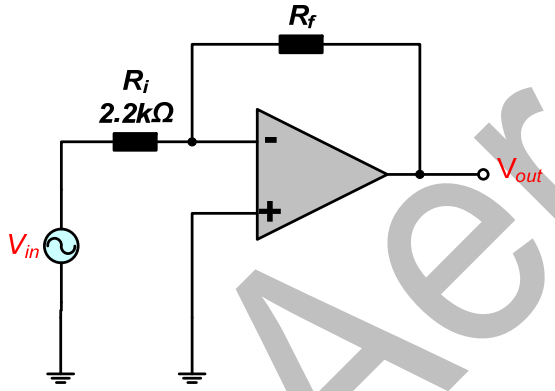
$$A_{cl(I)} = -\frac{R_f}{R_i}$$

Equation 3b: Closed Loop Gain of the Inverting Amplifier

The closed-loop gain of the inverting amplifier is the ratio of the feedback resistance to the input resistance. The closed-loop gain is independent of the Op-amp's internal open-loop gain. The negative feedback stabilizes the voltage gain. The '-' sign indicates that the output is inverted.

Example;

An Op-amp configuration is given below. Determine the value of R_f required to produce a closed-loop voltage gain of -100.



Solution;

Using the value of R_i at 2.2kΩ and $A_{cl(I)}$ = 100, R_f can be calculated as shown;

$$A_{cl(I)} = -\frac{R_f}{R_i}$$

This can be transposed so that;

$$R_f = -A_{cl(I)} \times R_i = -(-100) \times 2.2k\Omega = 220k\Omega$$

Op-Amp Impedances

Negative feedback connections affect the input and output impedances of the Op-amp. The effects on the inverting and non-inverting amplifiers can be examined.

Input and Output Impedances of a Non-Inverting Amplifier

The input impedance of a non-inverting amplifier configuration , $Z_{in(NI)}$, is shown in figure 187.

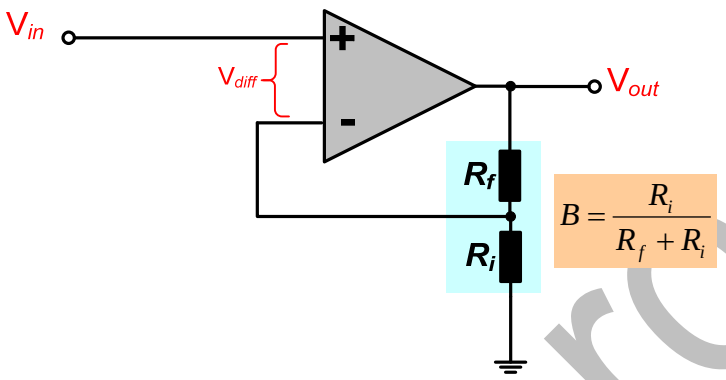


Figure 187 Input Impedance of a Non-Inverting Amp

$Z_{in(NI)}$ is greater than the internal input impedance of the Op-amp itself (without feedback) by a factor of $1 + A_{ol} \times B$. The input impedance is expressed in equation 4a.

$$Z_{in(NI)} = (1 + A_{ol} \times B) Z_{in}$$

Equation 4a: Input Impedance of a Non-Inverting Amp

The output impedance with the negative feedback, $Z_{out(NI)}$, is less than the Op-amp's output impedance by a factor of $\frac{1}{1 + A_{ol} \times B}$. The output impedance is expressed in equation 4b.

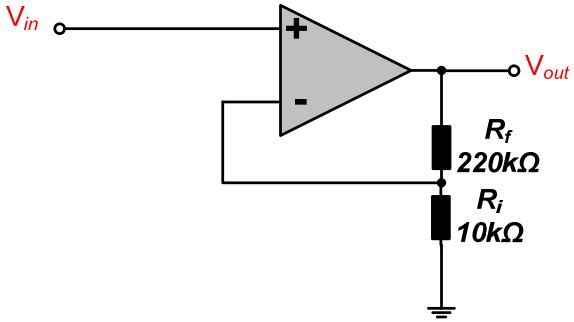
$$Z_{out(NI)} = \frac{Z_{out}}{1 + A_{ol} \times B}$$

Equation 4b: Output Impedance of a Non-Inverting Amp

The negative feedback in a non-inverting configuration increases the input impedance and decreases the output impedance.

Example:

- Determine the input and output impedances of the amplifier below. The input impedance, Z_{in}, is 2MΩ and the output impedance, Z_{out} is 75Ω. The Open-loop gain, A_{ol} is 200,000.
- Find the closed-loop voltage gain



Solution;

The attenuation, B, of the feedback circuit is;

$$B = \frac{R_i}{R_i + R_f} = \frac{10k\Omega}{230k\Omega} = 0.0435$$

$$Z_{in(NI)} = (1 + A_{ol} \times B) \times Zin = (1 + 200{,}000 \times 0.0435) \times 2M\Omega = 17.4G\Omega$$

$$Z_{out\,(NI)} = \frac{Z_{out}}{1 + A_{ol} \times B} = \frac{75\Omega}{1 + 8700} = 8.6m\Omega$$

The closed-loop gain is;

$$A_{cl(NI)} = \frac{1}{B} = \frac{1}{0.0435} = 23$$

Voltage-Follower Impedances

Since the voltage-follower is a special case of the non-inverting configuration, the same impedance formulae are used. The difference is that B = 1. The input and output impedances are shown in equations 4c and 4d.

$$Z_{in(VF)} = (1 + A_{ol}) * Z_{in}$$

Equation 4c: Input Impedance for a Voltage-Follower

$$Z_{out(VF)} = \left(\frac{Z_{out}}{1 + A_{ol}}\right)$$

Equation 4d: Output Impedance for a Voltage-Follower

As can be seen, the voltage-follower input impedance is greater for a given A_{ol} and Z_{in}, than for the non-inverting configuration with the voltage-divider feedback circuit.

The output impedance is much smaller because B is normally less than 1 for a non-inverting configuration

Input and Output Impedances of an Inverting Amplifier

For the inverting amplifier configuration, the input impedance, $Z_{in(I)}$, approximately equals the external input resistance, R_i. This is illustrated in figure 188. The inverting amplifier input impedance is shown in equation 4e.

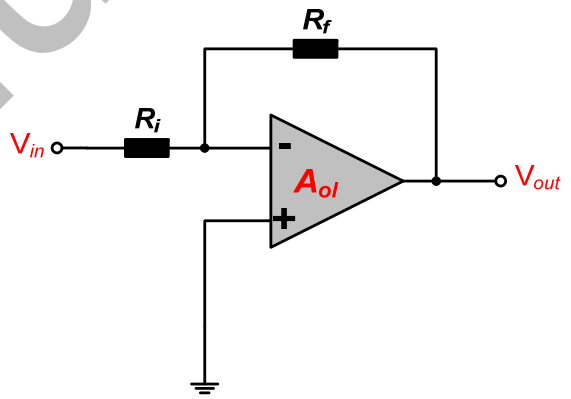


Figure 188 Inverting Amplifier

$$Z_{in(I)} \cong R_i$$

Equation 4e: Input Impedance of an Inverting Amplifier

The output impedance, $Z_{out(I)}$, approximately equals the internal output impedance of the Op-amp, Z_{out}. This is shown in equation 4f.

$$Z_{out(I)} \cong Z_{out}$$

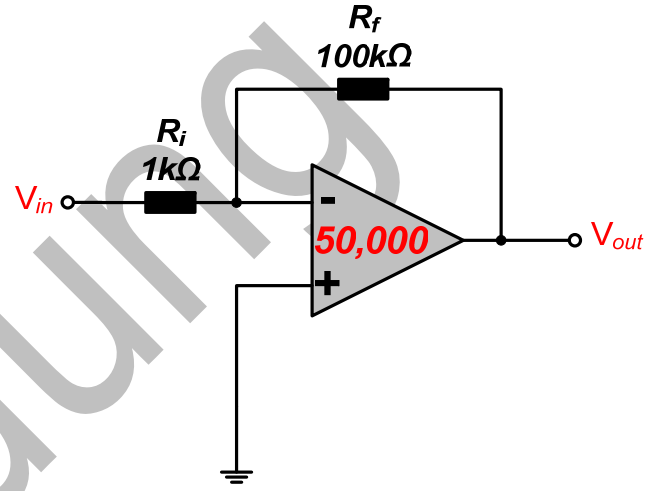
Equation 4f: Output Impedance of an Inverting Amplifier

Example:

The Op-amp shown below has the following parameters;

A_{ol} = 50,000
Z_{in} = 4MΩ
Z_{out} = 50Ω

- Find the values of the input and output impedances.
- Determine the closed-loop voltage gain.



Solution;

$$Z_{in(I)} \cong R_i = 1k\Omega$$

$$Z_{out(I)} \cong Z_{out} = 50\Omega$$

$$A_{cl(I)} = -\frac{R_f}{R_i} = -\frac{100k\Omega}{1k\Omega} = -100$$

1.4.6 Op-Amp Troubleshooting

Situations will occur whereby an Op-amp and its associated circuitry have become faulty. The Op-amp is a complex integrated circuit (IC) with multiple possible internal failures. The user cannot troubleshoot the Op-amp internally because it is an IC. It should be treated as a single device. If it fails, it should simply be replaced.

In the basic configurations, there are few external components that can fail. These consist of the feedback resistor, the input resistor and the Op-amp itself. A potentiometer may also be included to compensate for the Offset voltages. There may also be faulty connections on the PCB/contact pads.

The faults for the three basic configurations can be explored;

- The Non-inverting amplifier.
- The Voltage-follower.
- The Inverting amplifier.

Offset-Null Compensation

Most IC Op-amps provide a means of compensating for offset-null voltage. This can be performed by connecting an external potentiometer to designated pins on the IC package. This is shown in figure 189, using the 741 Op-amp as an example.

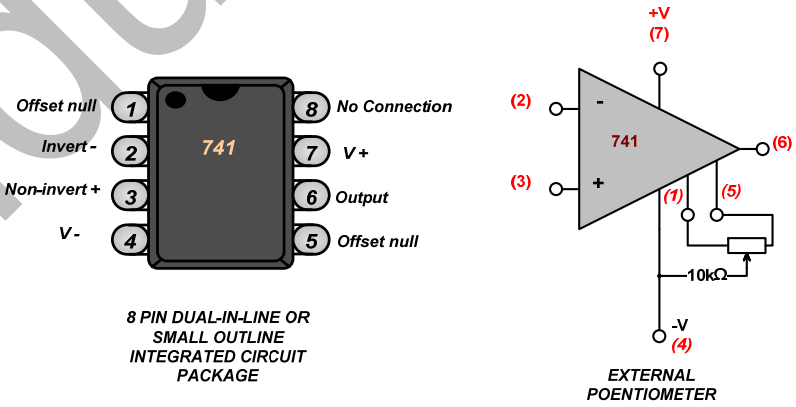


Figure 189 The 741 Op-Amp with Null Potentiometer Connection

The two terminals are labelled ‘Offset-null’. With no input, the potentiometer is simply adjusted until the output voltage reads 0V, as shown in figure 190.

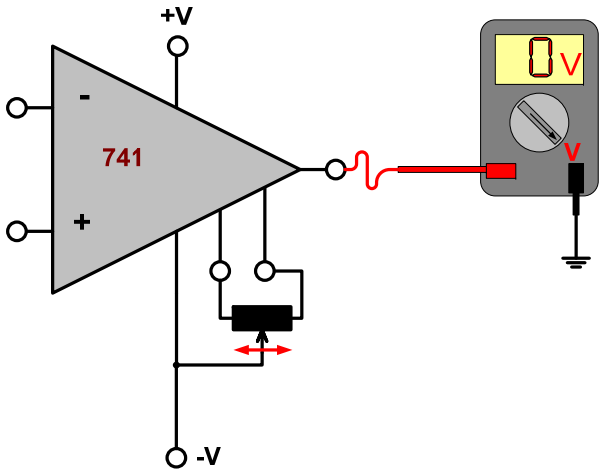


Figure 190 Setting the Null

Faults in a Non-Inverting Amplifier

The first step to be carried out with any suspected faulty circuit is to check for the appropriate supply voltages and ground connections. Then the following checks should be performed;

- Open feedback resistor.
- Open input resistor.
- Open/Incorrectly adjusted offset potentiometer.
- Faulty Op-amp.

Open Feedback Resistor

If the feedback resistor, R_f, is open, the Op-amp is operating with its very high open-loop gain. This causes the input signal to drive the device into non-linear (saturated) operation. This results in a severely clipped output signal. This is shown in figure 191.

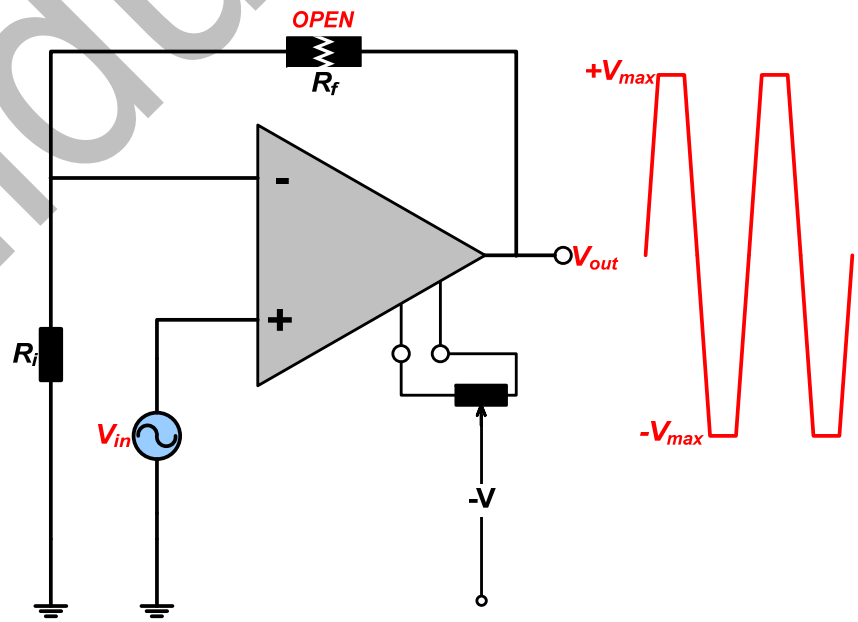


Figure 191 Open Feedback Resistor

Open Input Resistor

When the input resistor is open-circuit, the closed-loop configuration is still intact. Since R_i is open and effectively behaving as equal to infinity (∞), the closed-loop gain in this case is;

$$A_{cl(NI)} = \frac{R_f}{R_i} + 1 = \frac{R_f}{\infty} + 1 = 0 + 1 = 1$$

This shows that the amplifier behaves like a voltage follower. The output signal would become identical to the input as shown in figure 192.

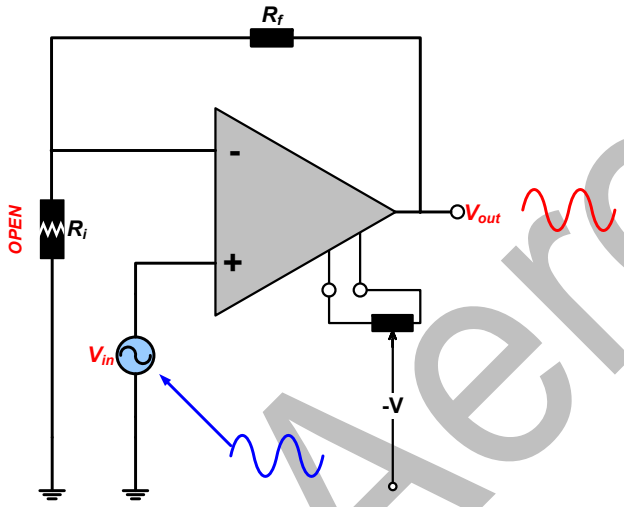


Figure 192 Open Input Resistor

Open/Incorrectly Adjusted Offset-Null Potentiometer

The output offset voltage will cause the output signal to begin clipping on only one peak as the input signal is increased to sufficient amplitude. This is shown in figure 193.

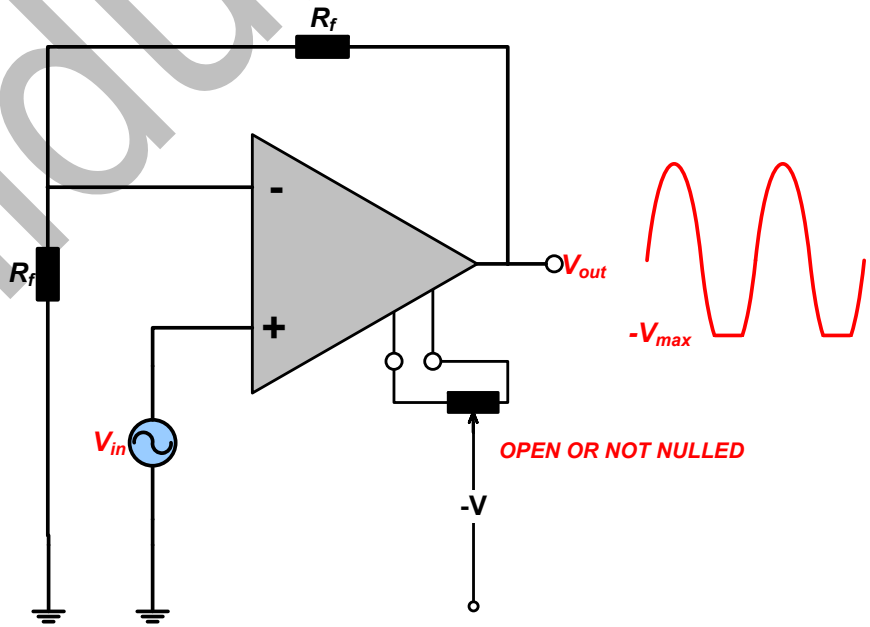


Figure 193 Open/Incorrectly Adjusted Offset-Null Potentiometer

Faulty Op-amp

An internal Op-amp failure will result in a loss or distortion of the output signal. Before replacing a faulty Op-amp it is best to ensure that there are no external failures or incorrect design conditions.

Faults in a Voltage-Follower

With the exception of a faulty Op-amp, faulty external connection or a problem with the Offset-null pot, the only issue that can occur with a Voltage-follower is an open feedback loop. This would have the same effect as an open feedback resistor.

Faults in an Inverting Amplifier

The following faults can occur in an inverting amplifier;

- Open feedback resistor.
- Open input resistor.

Open Feedback Resistor

If R_f is open, the input signal still feeds through the input resistor and is amplified by the high open-loop gain of the Op-amp. This is shown in figure 194.

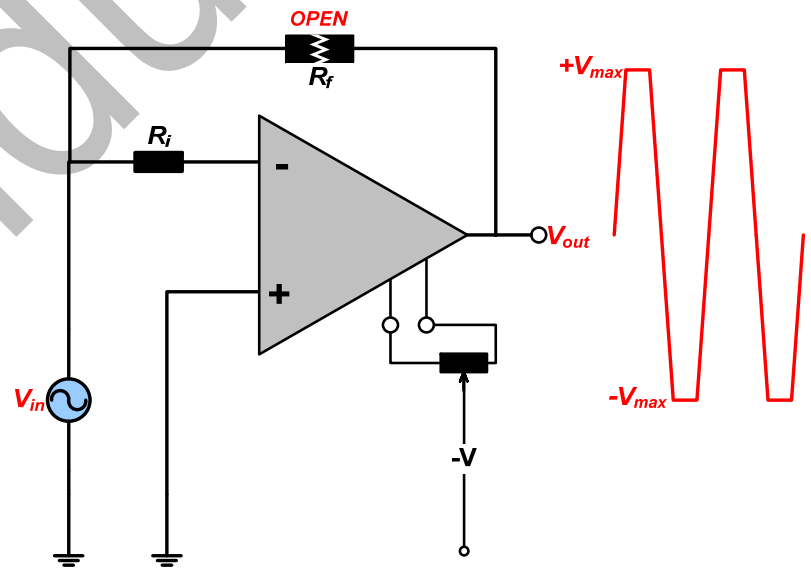


Figure 194 Open Feedback

This forces the device to be driven into saturated (non-linear) operation. This is the same result as the non-inverting configuration.

Open Input Resistor

This prevents the input signal from getting to the Op-amp input so there will be no output signal. This is shown in figure 195.

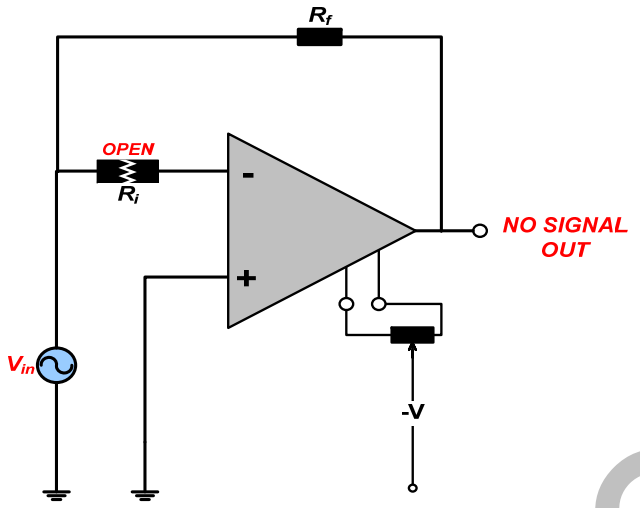


Figure 195 Open Input

Failures in the Op-amp (or the Offset-null pot) will have the same results as the non-inverting amplifier.

1.4.7 Comparators

Operational amplifiers are often used to compare the amplitude of one voltage to another. When applied as a Comparator, the Op-amp must be used in the Open-loop configuration. In this configuration, the input voltage is applied to one input, while a reference voltage is applied to the other.

Zero-Level Detection

One application of the Comparator is to determine when an input voltage exceeds a certain level. A zero-level detector is shown in figure 196.

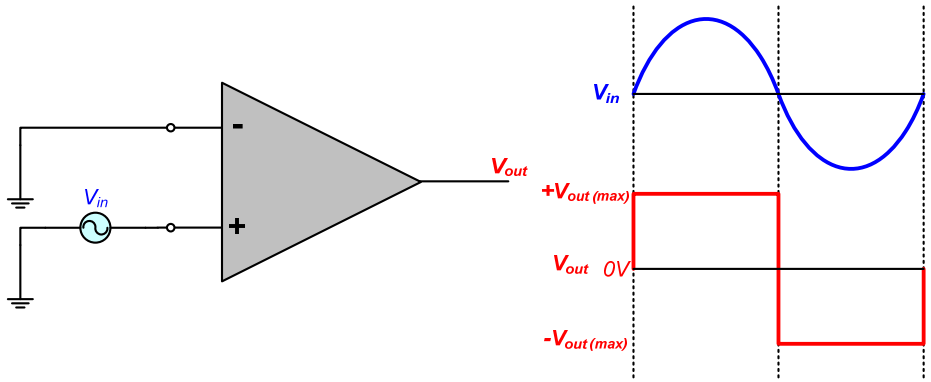


Figure 196 The Op-Amp as a Zero-Level Detector

The inverting input (-) is grounded while an input signal is applied to the non-inverting input (+). The high open-loop voltage gain causes the amplifier to be driven into saturation, even with a small difference in voltage between the inputs.

For example, consider an op-amp with an A_{ol} of 100,000. A voltage difference of 0.25mV (between the inputs) produces an output of 25V (0.25mV x 100,000), if the op-amp was capable of that voltage level.

Most Op-amps have output voltage limitations of less than ±15V. This means that the device would be driven into saturation by a 25V gain. For many comparison applications, special op-amp comparators are selected.

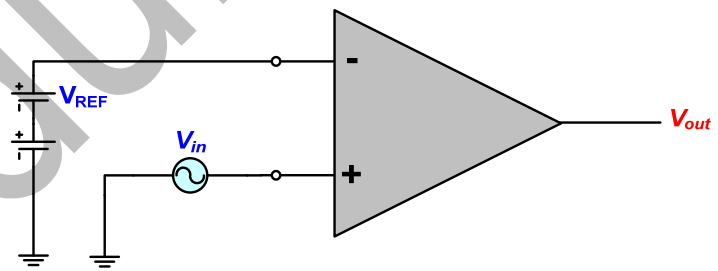
Figure 8a shows the result of a sinusoidal input voltage applied to the non-inverting input of the zero-level detector.

When the sine wave is negative, the output is at its maximum negative value. When the sine-wave crosses 0V, the amplifier is driven into its opposite state and the output goes to its maximum positive level.

The zero-level detector can be used as a squaring circuit to convert a sine-wave to a square-wave.

Non-Zero Level Detection

The Zero-level detector can be modified to detect voltages other than zero by connecting a fixed reference voltage to the inverting input (-). This is shown in figure 197.

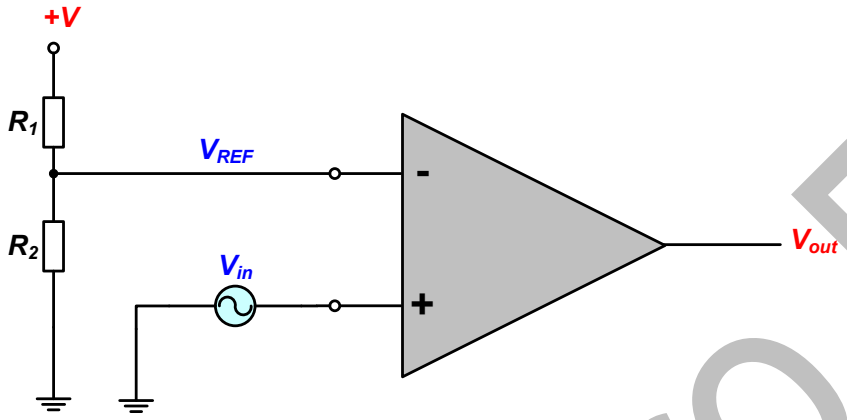


BATTERY REFERENCE

Figure 197 Non-Zero Level Detection

A more practical arrangement is shown in figure 198, using a voltage divider to set the reference voltage. The user should be familiar with the potential divider formula to calculate the reference voltage.

$$V_{REF} = \frac{R_2}{R_1 + R_2} \times (+V)$$



VOLTAGE DIVIDER REFERENCE

Figure 198 Practical Non-Zero Level Detection

Provided the input voltage (V_{in}) is less than V_{REF}, the output remains at the maximum negative level. When the input voltage exceeds the reference value, the output goes to its maximum positive state. This is shown in figure 199 with a sinusoidal input voltage.

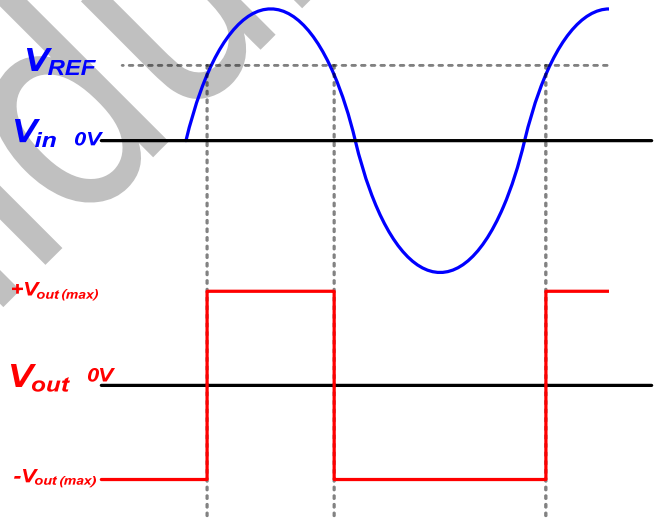
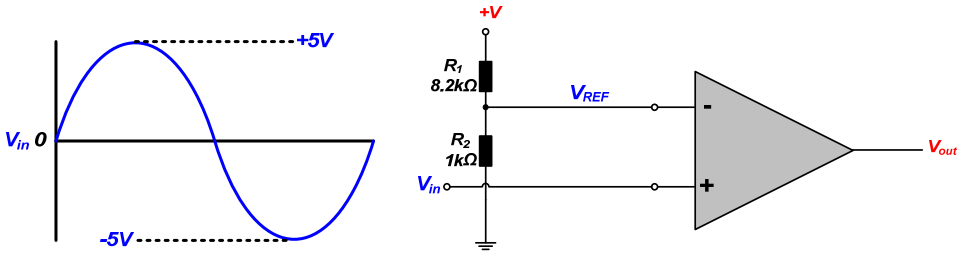


Figure 199 Waveforms

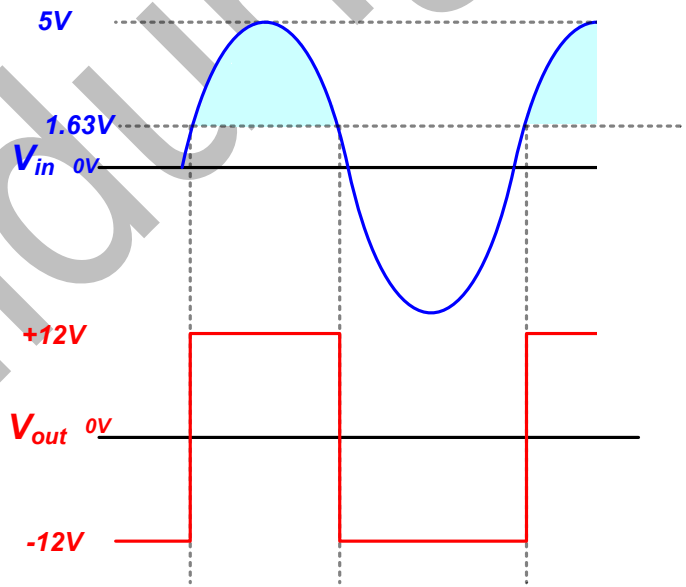
Example;

An input signal is applied to the comparator circuit shown below;



Show the output signal with respect to the input signal. Assume that the maximum output levels of the op-amp are ±12V.

Each time the input exceeds +1.63V, the output voltage switches to its +12V level. Each time the input drops below +1.63V, the output switches back to its -12V level.



Solution;

The reference voltage is set by R_1 and R_2;

$$V_{REF} = \frac{R_2}{R_1 + R_2} \times (+V) = \frac{1k\Omega}{8.2k\Omega + 1k\Omega} \times (15V) = 1.63V$$

1.4.8 Summing Amplifier

The summing amplifier is a variation of the inverting op-amp configuration. The summing amplifier has two or more inputs. Its output voltage is proportional to the negative of the algebraic sum of its input voltages.

The summing amplifier can be covered in several elements, we will cover three of those in this section;

- The Two-input summing amplifier.
- The summing amplifier with gain greater than unity.
- The Digital-analogue (DA) converter.

The Two Input Summing Amplifier

A two input summing amplifier is shown in figure 200. The two voltages, V_{IN1} and V_{IN2}, are applied to the inputs and produce conventional currents, I_1 and I_2.

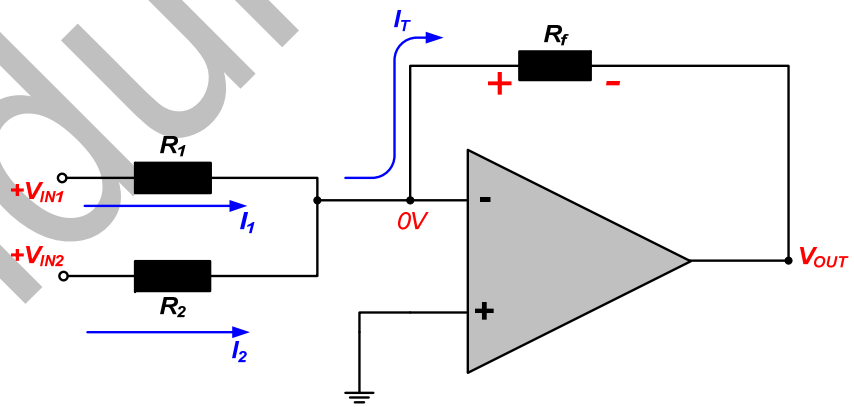


Figure 200 A Two Input Summing Amplifier

Using the ideal op-amp concepts of infinite impedance (No current from the input) and virtual ground, it is demonstrated that the inverting input (-) of the op-amp is 0V.

Therefore, the total current, which is the sum of I_1 and I_2, is forced through R_f.

$$I_T = I_1 + I_2$$

Since $V_{OUT} = -I_T \times R_f$;

$$V_{OUT} = (I_1 + I_2) \times R_f = -(\frac{V_{IN1}}{R_1} + \frac{V_{IN2}}{R_2}) \times R_f$$

If all three of the resistors are equal (R_1=R_2=R_f=R) then;

$$V_{OUT} = -(\frac{V_{IN1}}{R} + \frac{V_{IN2}}{R}) \times R$$

$$V_{OUT} = -(V_{IN1} + V_{IN2})$$

Equation 5a: Sum of the Input Voltages

Equation 5a shows the output voltage is the sum of the two input voltages. A general expression is given in equation 5b for a summing amplifier with 'n' inputs. This will apply when all the input resistors are of equal value, as shown in figure 201.

$$V_{OUT} = -(V_{IN1} + V_{IN2} + V_{IN3} + ...V_{IN\,n})$$

Equation 5b: Sum of Input Voltages with Equal Input Resistance Values

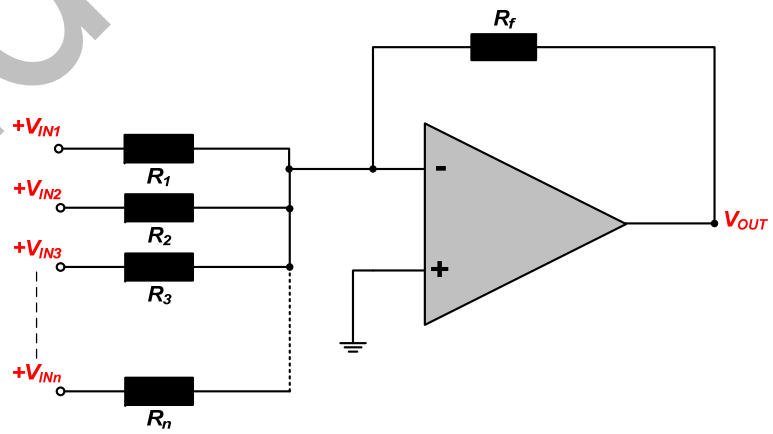
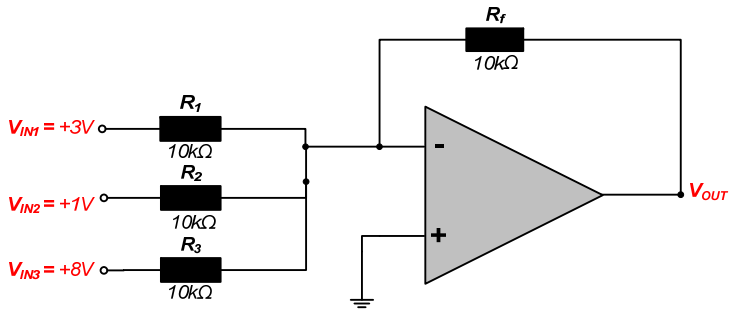


Figure 201 Summing Amplifier with 'N' Inputs

Example:

Determine the output voltage in the circuit below;



Solution;

$$V_{OUT} = -(V_{IN1} + V_{IN2} + V_{IN3})$$
$$= -(3V + 1V + 8V) = -12V$$

Summing Amplifier with Gain Greater than Unity

When R_f is larger than the input resistors, the amplifier has a gain of $\frac{R_f}{R}$, where R is the value of each input resistor. The general expression for the output is shown in equation 5c.

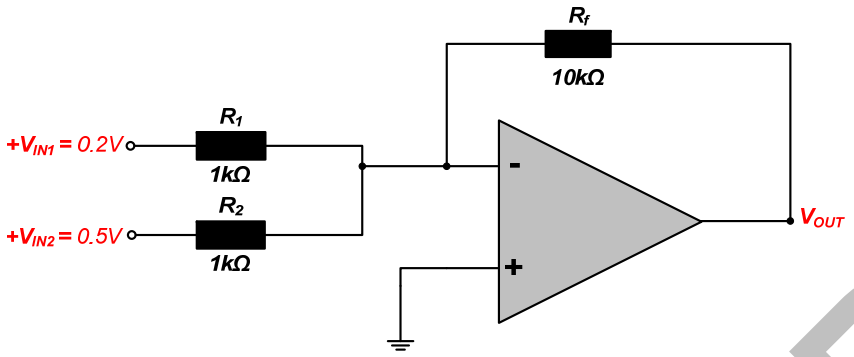
$$V_{OUT} = -\frac{R_f}{R} \times (V_{IN1} + V_{IN2} + ... + V_{INn})$$

Equation 5c: General Expression for Output of Summing Amplifier

It can be seen that the output is the sum of all the input voltages multiplied by a constant determined by the ratio of $\frac{R_f}{R}$.

Example:

Determine the output voltage for the summing amplifier shown below;



Solution;

R_f = 10kΩ and R = R_1 = R_2 = 1kΩ. Therefore;

$$V_{OUT} = -\frac{R_f}{R} \times (V_{IN1} + V_{IN2})$$
$$= -\frac{10k\Omega}{1k\Omega} \times (0.2V + 0.5V)$$
$$= -10 \times (0.7V)$$
$$= -7V$$

The Digital-Analogue (DA) Converter

Digital to Analogue (DA) conversion is a very important part of many systems. This example uses the op-amp summing amplifier conditions to produce a 'Binary weighted' DA converter.

A Binary-weighted DA converter uses resistor values to represent the binary weights of each of the input bits of the digital code. Figure 202 shows the Op-amp used as a DA converter.

$$V_{OUT} = I_f \times R_f$$

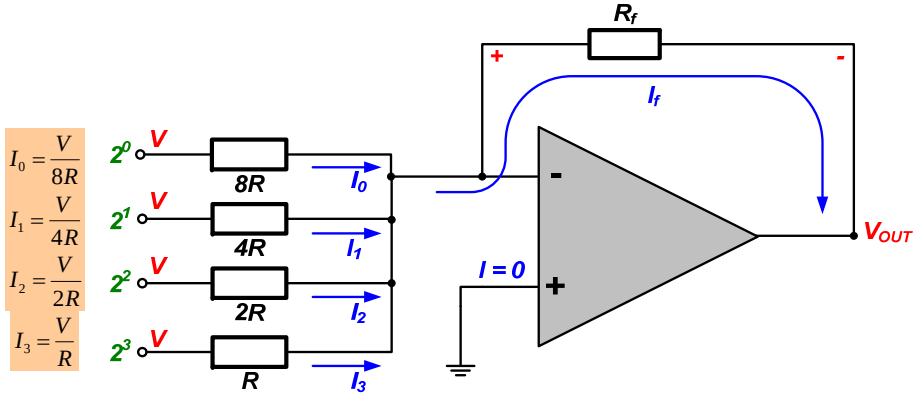


Figure 202 Four-Bit DA Converter with Binary Weighted Inputs

Each of the input resistors will either have a current flow or zero current flow, depending on the input voltage level.

If the input voltage is zero (Binary '0'), the current is also zero.

If the input voltage is high (binary '1'), current can flow.

The amount of current flow is dependant on the value of the input resistor. This is different for each input resistor.

Since there is ideally no current into the op-amp inverting input (-), each input will add together and go through R_f.

The inverting input (+) is held at 0V (virtual ground), so the voltage drop across R_f is equal to the output voltage. Therefore $V_{OUT} = I_f \times R_f$.

The values of the input resistors are chosen to be inversely proportional to the binary weights of the corresponding input bits.

The lowest value resistor (R) corresponds to the highest binary-weighted input (2^3). The other resistors are multiples of R (2R, 4R and 8R). The other multiples of R correspond to the other binary weights (2^2, 2^1 and 2^0).

The inputs currents are also proportional to the binary weights. Therefore, the output voltage is proportional to the sum of the binary weights because the sum of the input currents flows through R_f.

The lowest value resistor (R) will have the highest current and provide the highest current contribution to the summing amplifier when the input voltage is active. This input will represent the most-significant-bit (MSB).

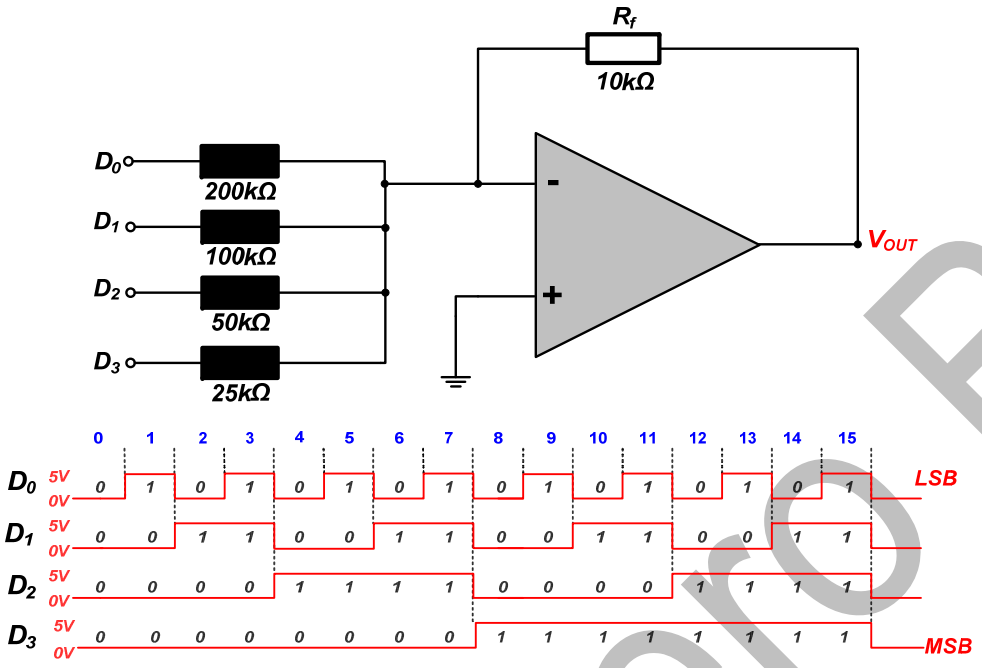
The highest value resistor (8R) will have the smallest current. The current will be eight times smaller than that through R. This will provide the smallest current contribution to the summing amplifier when the input voltage is active. This input will represent the least-significant bit (LSB).

The main disadvantage is of this type of DA converter is the number of different resistor values. For example, an 8 bit DA converter requires 8 resistors ranging from a value of R to 128R in binary-weighted steps.

These ranges of resistors require tolerances of 1 part in 255 (less than 0.5%) to accurately convert the input. This ensures that this type of DA converter is difficult to mass produce.

Example:

Determine the output of the DA converter (shown below) if the waveforms representing a sequence of 4-bit numbers are applied to the inputs. Input D_0 is the Least-significant bit (LSB).



Solution;

The current for each of the weighted inputs must be determined first. Since the inverting input of the Op-amp (-) is at 0V (Virtual ground), and a binary '1' corresponds to 5V, the current through any of the resistors is 5V divided by the resistance value.

$$I_0 = \frac{5V}{200k\Omega} = 0.025mA$$

$$I_1 = \frac{5V}{100k\Omega} = 0.05mA$$

$$I_2 = \frac{5V}{50k\Omega} = 0.1mA$$

$$I_3 = \frac{5V}{25k\Omega} = 0.2mA$$

Ideally, there is no current flow into the inverting amplifier of the Op-amp due to its own high impedance. Therefore, it can be assumed that all of the current flows through the feedback resistor, R_f.

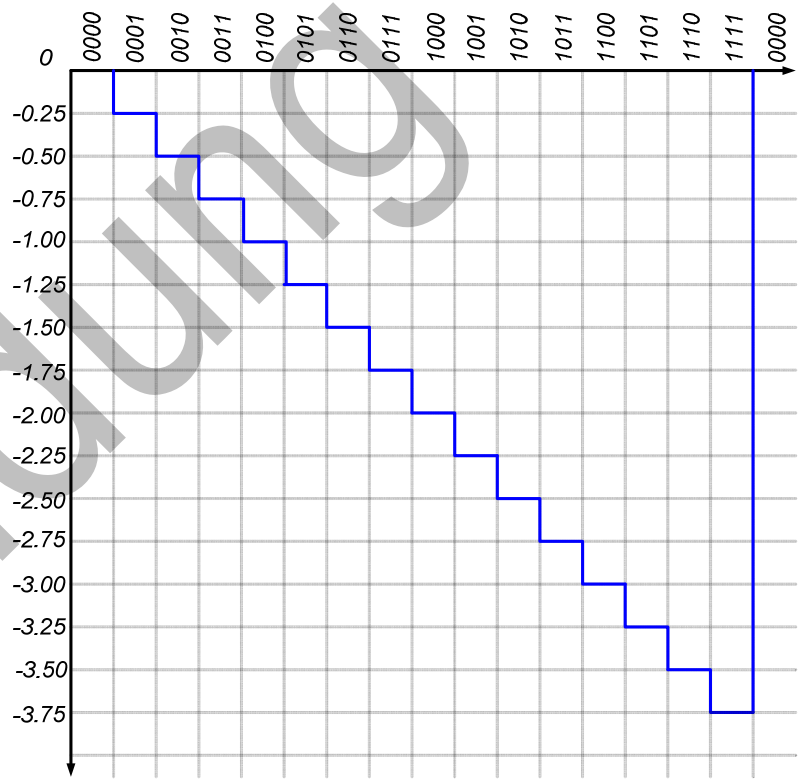
Since one end of R_f is at 0V, the voltage drop across R_f is equal to the output voltage, which is negative with respect to virtual ground.

$$V_{OUT(D0)} = (10k\Omega) \times (-0.025mA) = -0.25V$$

$$V_{OUT(D1)} = (10k\Omega) \times (-0.05mA) = -0.5V$$

$$V_{OUT(D2)} = (10k\Omega) \times (-0.1mA) = -1V$$

$$V_{OUT(D3)} = (10k\Omega) \times (-0.2mA) = -2V$$



From the diagram above, it can be seen that the first binary input code is '0000'. This produces an output voltage of 0V.

The next input code is '0001'. This produces an output voltage of -0.25V.

The next code is '0010'. This produces an output of -0.5V.

The next code is '0011'. This produces an output of -0.75V (-0.25V + -0.5V).

Each successive increase in the binary code increases the output voltage by -0.25V.

For this particular binary sequence on the inputs, the output is a staircase waveform ranging from 0V to -3.75V in 0.25V increments.

1.4.9 The Integrator

An Op-amp integrator simulates mathematical integration. This is basically a summing process that determines the total area under the curve of a function.

The integrators shown in this description are idealized to illustrate basic principles. Practical integrators include additional resistors in parallel with the feedback capacitor to prevent saturation.

An ideal integrator is shown in figure 203. The feedback element is a capacitor that forms an RC circuit with the input resistor.

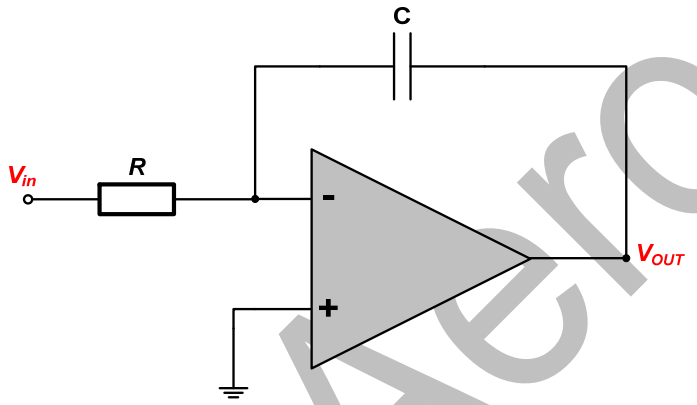


Figure 203 Ideal Op-Amp Integrator

Although a large value resistor is normally used in parallel with the capacitor to limit the gain, it does not affect the basic operation and is not shown for the purposes of this discussion.

Capacitor Charge

To understand the operation of the integrator, it is important to review how a capacitor charges. The charge, Q, on a capacitor is proportional to the charging current and the charge time.

$$Q = I_C \times t$$

In terms of voltage, the charge on a capacitor is:

$$Q = C \times V_C$$

From these two relationships, the capacitor voltage can be expressed as:

$$V_C = (\frac{I_C}{C}) \times t$$

This expression has the form of an equation for a straight line beginning at zero with a constant slope of I_C/C.

From mathematical algebra, the general formula for a straight line is y = mx + b. In this case, y = V_C, m = I_C/C, x = t and b = 0.

It should be recalled that the capacitor voltage in a simple RC circuit is exponential (not linear). This is because the charging current continuously decreases as the capacitor charges the rate of change of the voltage to continuously decrease.

The key point when using an Op-amp with an RC circuit to form an integrator is that the capacitor's charging current is held constant. This produces a linear (straight line) voltage rather than an exponential voltage.

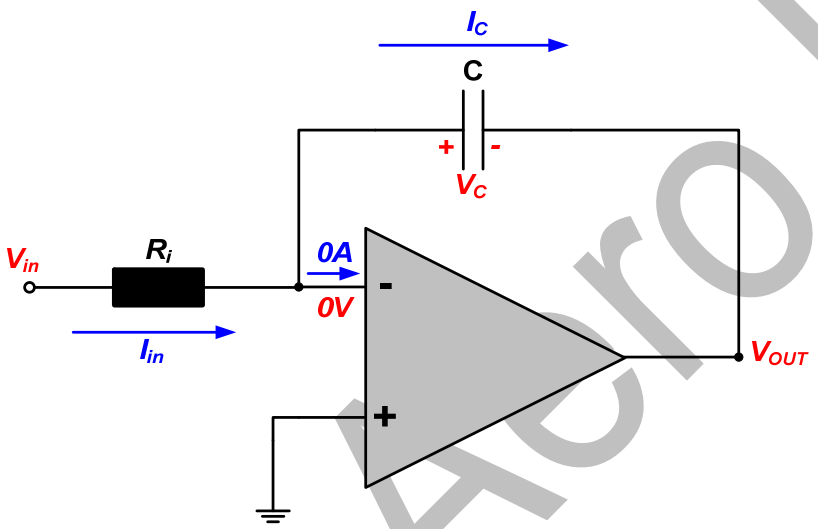


Figure 204 Currents in an Integrator

This is demonstrated in figure 204. The inverting input (-) of the Op-amp is held at virtual ground (0V) so the voltage across R_i equals V_{in}. Therefore, the input current is:

$$I_{in} = \frac{V_{in}}{R_i}$$

If V_{in} is a constant voltage, then I_{in} is also constant because the inverting input always remains at 0V. This means that there is a constant voltage drop across R_i.

The high impedance of the Op-amp causes negligible current from the inverting input. All of the input current is directed through the capacitor. The capacitor current is therefore:

$$I_C = I_{in}$$

Capacitor Voltage

Since I_{in} is constant, so is I_C. The constant I_C charges the capacitor linearly and produces a linear voltage across C.

The positive side of the capacitor is held at 0V by the virtual ground of the Op-amp. The voltage on the negative side of the capacitor decreases linearly from zero as the capacitor charges. This is shown in figure 205. This is known as a negative ramp.

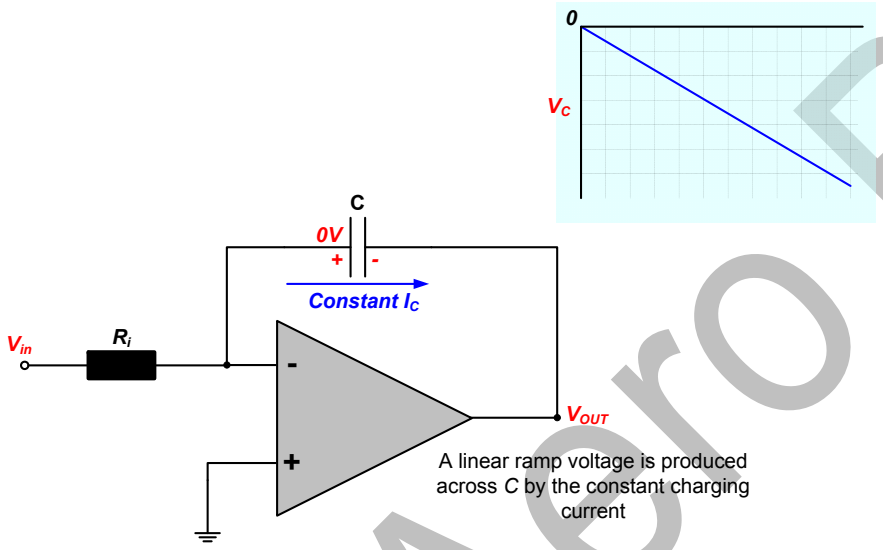


Figure 205 Negative Ramp Voltage

The Output Voltage

V_{OUT} is the same voltage as the negative side of the capacitor. When a constant input voltage in the form of a step or pulse (a pulse has constant amplitude when high) is applied, the output ramp decreases negatively until the Op-amp saturates at its maximum negative level. This is shown in figure 206.

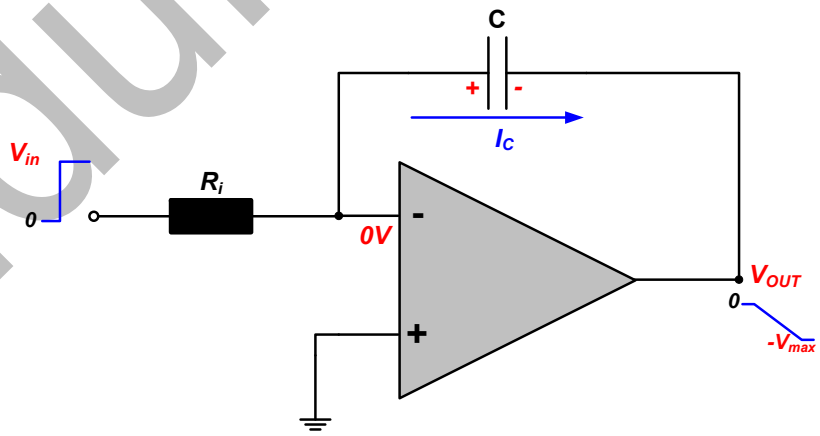


Figure 206 Step/Pulse Input Voltage Rate of Change of the Output

The rate the capacitor charge (and therefore the slope of the output ramp) is set by the ratio I_C/C.

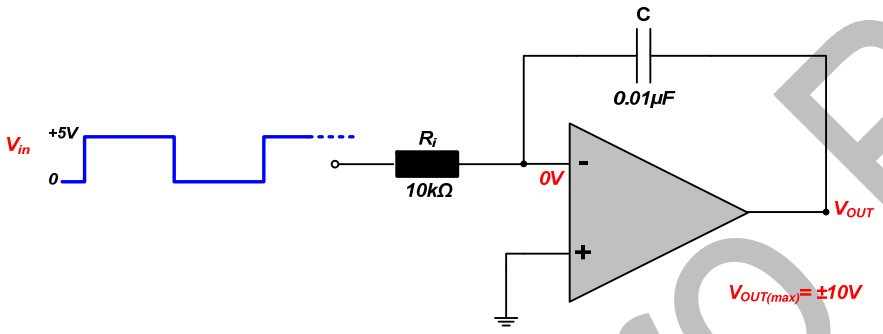
Since $I_C = V_{in}/R_i$, the rate of change (or slope) of the integrator's output voltage is $\Delta V_{OUT}/\Delta t$. This is shown in equation 7.

$$\frac{\Delta V_{OUT}}{\Delta t} = -\frac{V_{in}}{(R_i \times C)}$$

Equation 7: Rate of Change of Integrator Output Voltage

Example:

a) Determine the rate of change of the output voltage in response to the first input pulse in a pulse waveform, for the Integrator below. The output voltage is initially zero.



b) Describe the output after the first pulse. Draw the output waveform.

Solution;

a) The rate of change of the output voltage during the time that the input pulse is high is:

$$\frac{\Delta V_{OUT}}{\Delta t} = -\frac{V_{in}}{R_i \times C}$$
$$= -\frac{5V}{(10k\Omega) \times (0.01\mu F)}$$
$$= -\frac{50kV}{s}$$
$$= \frac{-50mV}{\mu S}$$

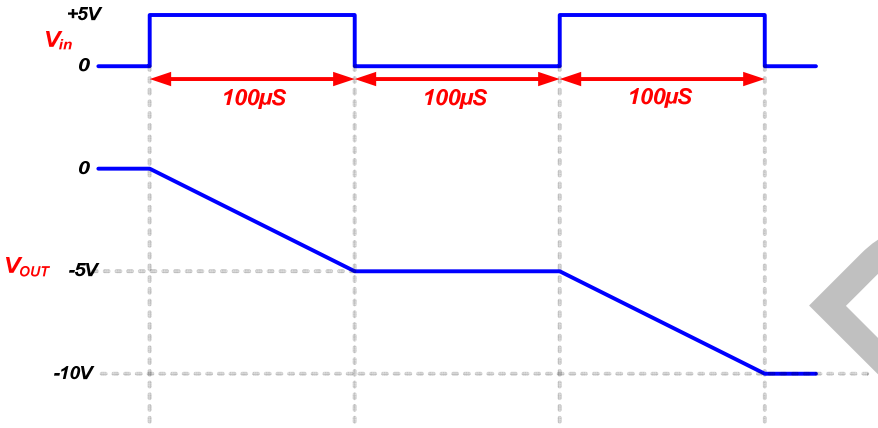
c) The rate of change was found to be -50mV/μS. When the input is (+) 5V, the output is a negative-going ramp.

When the input is 0V, the output is a constant level. In 100μS the voltage decreases.

$$\Delta V_{OUT} = -(50mV/\mu S) \times (100\mu S) = -5V$$

The negative-going ramp reaches -5V at the end of the input pulse. The output voltage then remains constant at -5V for the time that the input is zero.

When the next pulse occurs, the output again becomes a negative-going ramp which reaches -10V at the end of the second pulse.



The output will remain at -10V thereafter because it has reached its maximum negative limit.

1.4.10 The Differentiator

The Op-amp differentiator simulates mathematical differentiation. This is the process used to determine the instantaneous rate of change for a function.

The differentiators shown in this description are idealized to illustrate basic principles. Practical differentiators may include a series resistor to reduce high frequency noise.

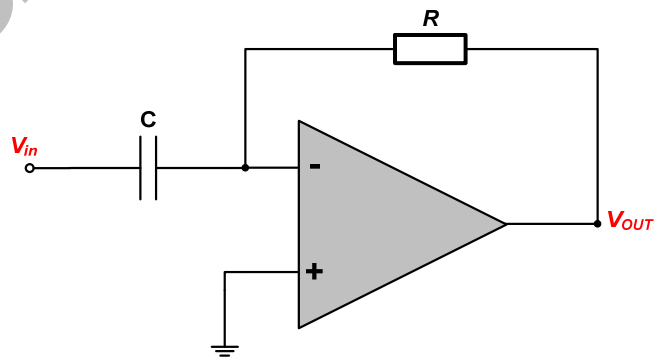


Figure 207 Ideal Differentiator

The ideal differentiator is shown in figure 207. The position of the resistor and capacitor is visibly different from the integrator. The capacitor is now used as the input element.

The differentiator produces an output that is proportional to the rate of change of the input voltage.

A small value resistor is usually used in series with the capacitor to limit the gain. This resistor does not affect the basic operation so it has not been shown for the purpose of this demonstration.

To illustrate the operation of the differentiator, a positive-going ramp voltage can be applied to the input. This is shown in figure 208.

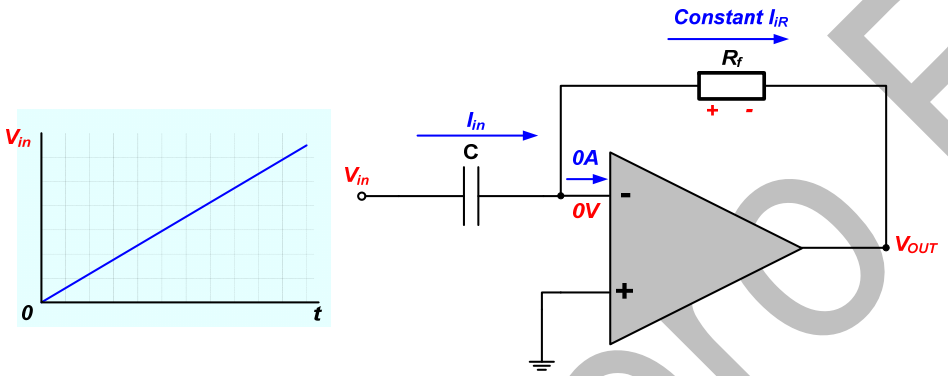


Figure 208 Ramp Input Applied to Differentiator

In this case $I_C = I_{in}$ and the voltage across the capacitor is equal to V_{in} at all times ($V_C = V_{in}$) due to the virtual ground on the inverting input.

Using the basic formula ($V_C = (\frac{I_C}{C}) \times t$), the capacitor current is ($I_C = (\frac{V_C}{t}) \times C$).

The current at the inverting input is negligible so $I_R = I_C$. Both currents are constant because the slope of the capacitor voltage ($\frac{V_C}{t}$) is constant.

The output voltage is also constant and equal to the voltage across Rf because one side of the feedback resistor is always 0V (virtual ground). This is expressed in equation 8.

$$V_{OUT} = I_R \times R_f = I_C \times R_f$$
$$V_{OUT} = -(\frac{V_C}{t}) \times R_f \times C$$

Equation 8: Differentiator V_{out}

The output is negative when the input is a positive-going ramp. This is shown in figure 208.

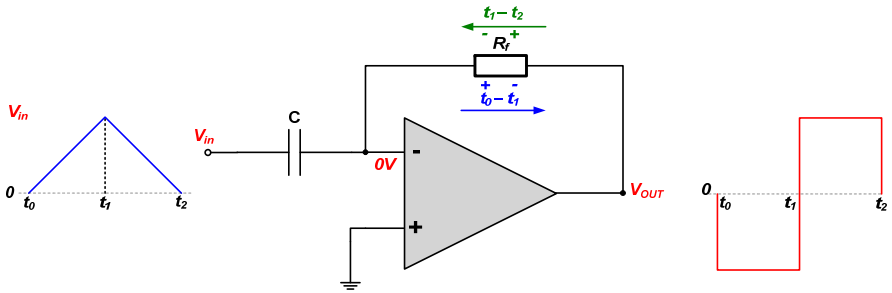


Figure 209 Differentiator Output

During the positive slope of the input, the capacitor is charging from the input source. The constant current through the feedback resistor is shown.

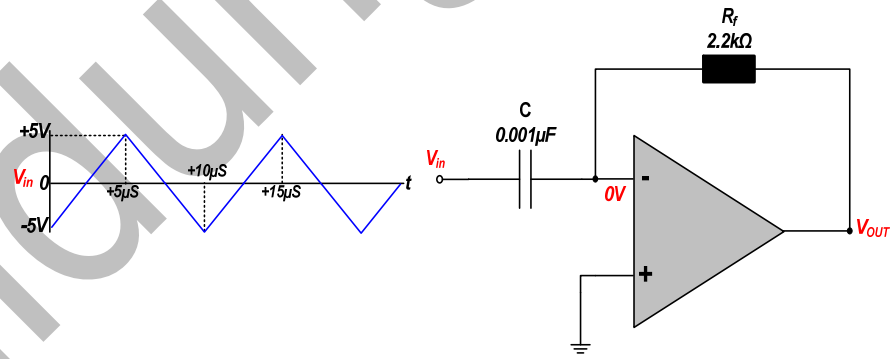
During the negative slope of the input, the current is in the opposite direction because the capacitor is discharging.

The term $\frac{V_C}{t}$ is the slope of the input. If the slope increases, V_{OUT} increases. If the slope decreases, V_{OUT} decreases.

This means the output voltage is proportional to the slope (which is the rate of change) of the input. The constant of proportionality is the RC time constant (R_f *x* C).

Example:

Determine the output voltage of the Op-amp differentiator for the tri-angular wave input shown below;



Solution;

Starting at t =0s, the input voltage is a positive-going ramp ranging from -5V to +5V (a variation of +10V in total) in 5μS.

The input then changes to a negative-going ramp ranging from +5V to -5V (A variation of -10V) in 5μS.

The time constant is;

$$R_f \times C = (2.2k\Omega) \times (0.001\mu F) = 2.2\mu S$$

Determine the slope (rate of change),($\frac{V_C}{t}$), of the positive-going ramp and calculate the output voltage:

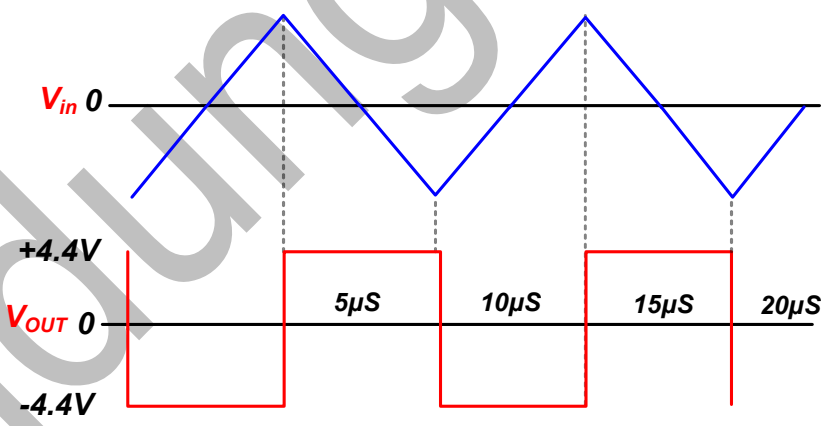
$$\frac{V_C}{t} = \frac{10V}{5\mu S} = 2V/\mu S$$

$$V_{OUT} = -(\frac{V_C}{t}) \times (R_f \times C)$$
$$= -(2V/\mu S) \times 2.2\mu S$$
$$= -4.4V$$

The slope of the negative-going ramp will also be -2V/µS. Calculate the output voltage:

$$V_{OUT} = -(\frac{V_C}{t}) \times (R_f \times C)$$
$$= -(-2V/\mu S) \times 2.2\mu S$$
$$= +4.4V$$

The output voltage waveform can now be shown below.



It has been shown above that a triangular-wave input will result in a square-wave output when applied to an Op-amp differentiator.

The Op-amp Differentiator can also be applied with Sinusoidal inputs. Standard rules of differentiation show that a Cosine wave is the result of the differentiation of a sine wave. This is shown in figure 210.

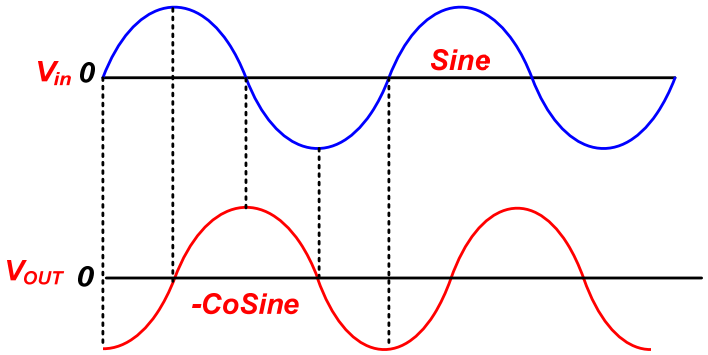


Figure 210 Differentiation of a Sine Wave

A square-wave input will result in a 'Tyre tread' output. This is shown in figure 211.

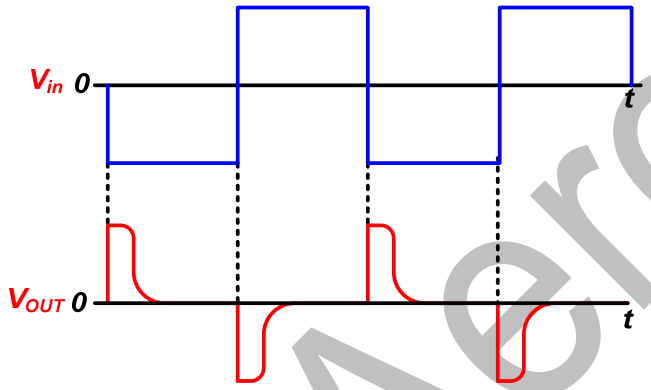


Figure 211 Differentiation of a Square-Wave

1.4.11 Schmitt-Trigger

A comparator with hysteresis

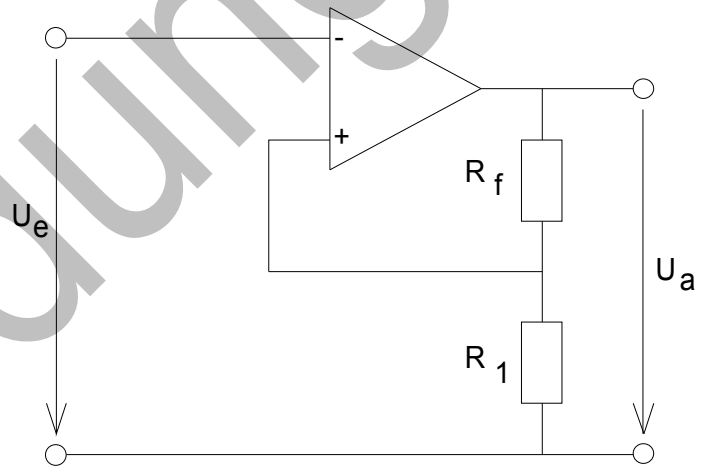


Figure 212 Schmitt-Trigger

In electronics, a **Schmitt** (or **Schmid**t) **trigger** is a special kind of comparator circuit.
A Schmitt trigger has one input voltage and one output voltage. The output can be either high or low. When the input is below a certain threshold, the output is high; when the input is above a certain (higher) threshold, the output is low; and when the input is between the two thresholds, the output retains its value.

The trigger is so named because the output retains its value until the input changes sufficiently to trigger a change. This delayed action is called *hysteresis*, and implies that the Schmitt trigger has some memory.

The benefit of a Schmitt trigger over a similar system with a single input threshold is that the Schmitt trigger is more stable. With only one input threshold, a noisy input signal near that threshold could rapidly switch back and forth, causing the output to switch back and forth from low to high. With the Schmitt trigger, a noisy input signal near one threshold could cause only one switch in output value, after which it would have to move to the other threshold in order to cause another switch.
The Schmitt trigger was invented by US scientist Otto H. Schmitt; today, they are typically built using operational amplifiers, and the reference voltage levels can be adjusted by controlling the resistances of R 1 and R 2

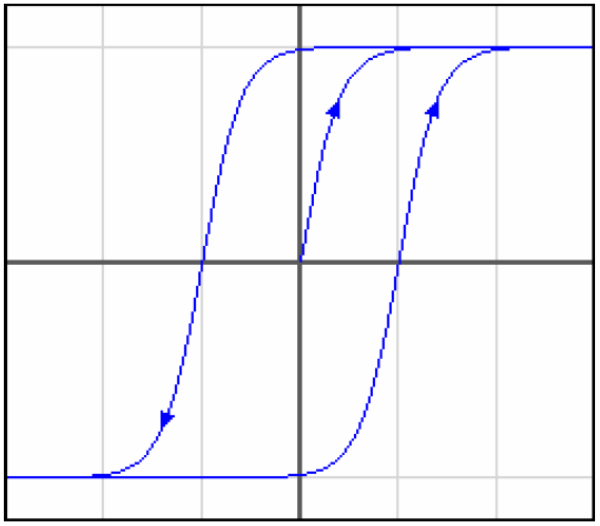


Figure 213 Characteristic curve of a Schmitt-trigger

1.4.12 Integrated Circuit Technologies

When working with Integrated circuit Op-amps or digital logic gates, the user should be familiar with their operational characteristics and parameters, as well as their logical and behavioural functions. This section will define those operational properties and discuss their effects on any analogue and digital outputs.

The creation of logic gates at component level (Transistors, resistors etc) will be discussed for the main technologies in use. The inclusive technologies under discussion are:

- Transistor-Transistor Logic (TTL).
- Complementary Metal-Oxide Semi-conductor (CMOS).

Other logic technologies do exist, although they are primarily evolutions of the two main technologies above.

The main operational parameters to be considered when designing circuits and selecting integrated circuits for their application are:

- DC supply voltages.
- Noise immunity.
- Power dissipation.
- Propagation delay and race hazards.
- Speed-Power product.
- Loading and fan out.

DC Supply Voltages

The standard DC supply voltage for **TTL** is +5V. **CMOS** devices are available in two different supply voltage categories. CMOS devices are available in +5V and +3.3V (this is known as low-voltage CMOS).

The DC supply voltages are usually omitted from logic diagrams for simplicity. The DC supply is usually connected to the V_{CC} pin of an IC package. The ground pin is usually connected directly to ground.

Both the supply and ground connections are distributed internally to all elements within the package. This is shown in figure 214 for a 16 pin package.

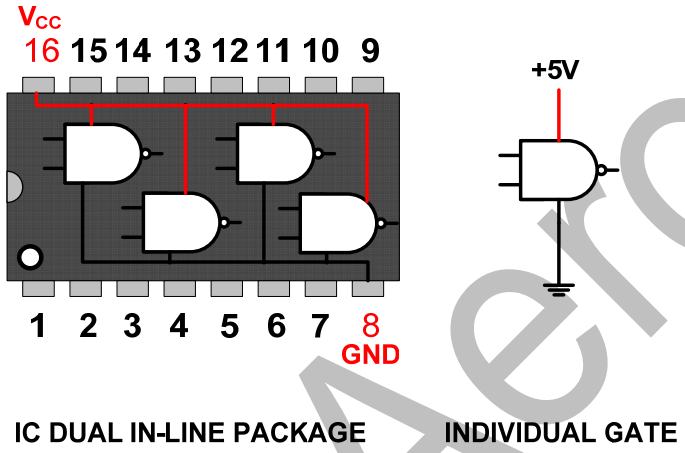


Figure 214 Example of V_{cc} and Ground Connections

CMOS logic levels

There are four different logic level specifications:

V_{IL} (Voltage-in LOW).
V_{IH} (Voltage-in HIGH).
V_{OL} (Voltage-out LOW).
V_{OH} (Voltage-out HIGH).

For CMOS circuits, the ranges of input voltages (V_{IL}) that can represent a valid LOW (logic *0*) are from 0V-1.5V for the +5V logic. These levels are 0V-0.8V for the +3.3V logic.

The ranges for the input voltage (V_{IH}) that can represent a valid HIGH (logic 1) are from 3.5V to 5V for the +5V logic. These levels are 2V-3.3V for the +3.3V logic.

The range of values between 1.5V and 3.5V on the +5V CMOS logic are regions of unpredictable and undesirable performance. The values in this range are disallowed.

The range of values between 0.8V and 2V on the +3.3V CMOS logic are regions of unpredictable and undesirable performance. The values in these ranges are disallowed.

When an input voltage is within this range, it can be interpreted as either a high or a low voltage by the logic circuit. Therefore, CMOS gates cannot be operated reliably when the input voltages are un-allowed ranges.

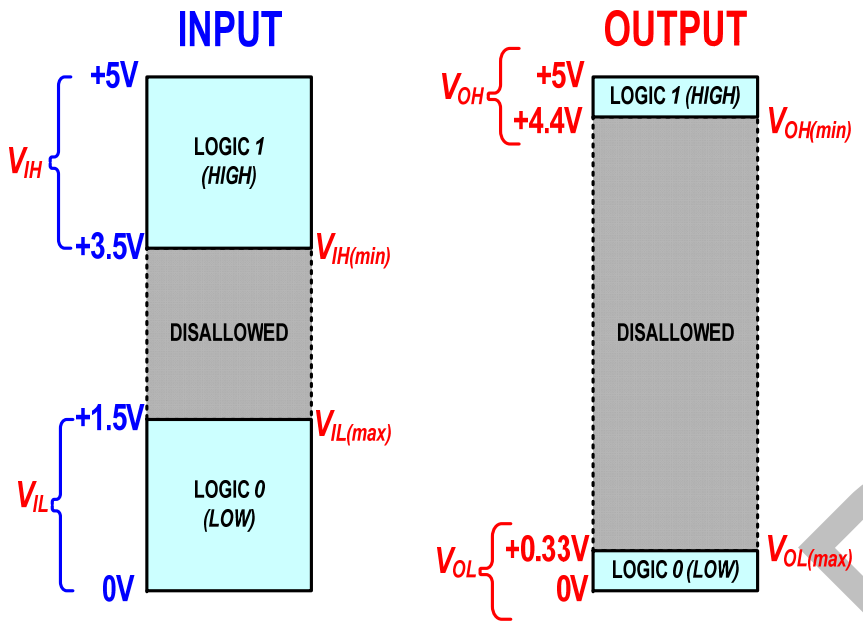


Figure 215 Valid Logic Levels for the +5v CMOS

The ranges of valid logic levels for the +5V CMOS are shown in figure 215. It should be noted that the minimum high output voltage, ($V_{OH(min)}$), is greater than the minimum high input voltage, ($V_{IH(min)}$).

The maximum low output voltage, ($V_{OL(max)}$), is less than the maximum low input voltage, ($V_{IL(max)}$).

TTL logic levels

The input and output logic levels for TTL are shown in figure 216. As demonstrated with CMOS, the four different logic levels are:

V_{IL} (Voltage-in LOW).
V_{IH} (Voltage-in HIGH).
V_{OL} (Voltage-out LOW).
V_{OH} (Voltage-out HIGH).

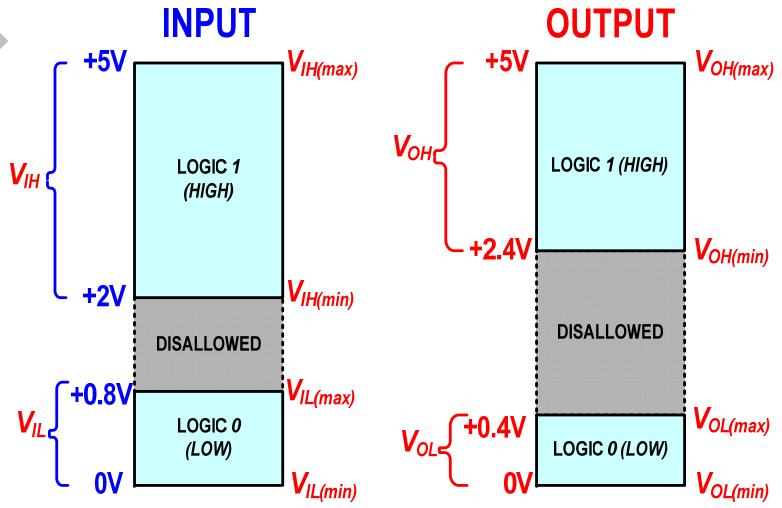


Figure 216 Valid Logic Levels for TTL

2 CIRCUIT BOARDS

2.1 TYPES OF CIRCUIT BOARDS

Correct mechanical and electrical constructions are a prerequisite for electronic circuits. In order to get a good mechanical stability electronic circuits are built on boards. Dependent on the purpose of the electronic circuit several kinds of boards are available, such as

⇒ Soldering strips

⇒ Raster boards

⇒ Rag raster boards

⇒ Veroboards

⇒ Printed circuit boards.

2.2 CONVETIONAL CIRCUIT BOARDS

2.2.1 Soldering Strips

Simple electronic circuits with low voltage supply or battery supply can be built very easily on soldering strips. Soldering strips are no longer used in industrial circuits. For test set -ups or in the field of hobby electronics, however, soldering strips have still been employed.

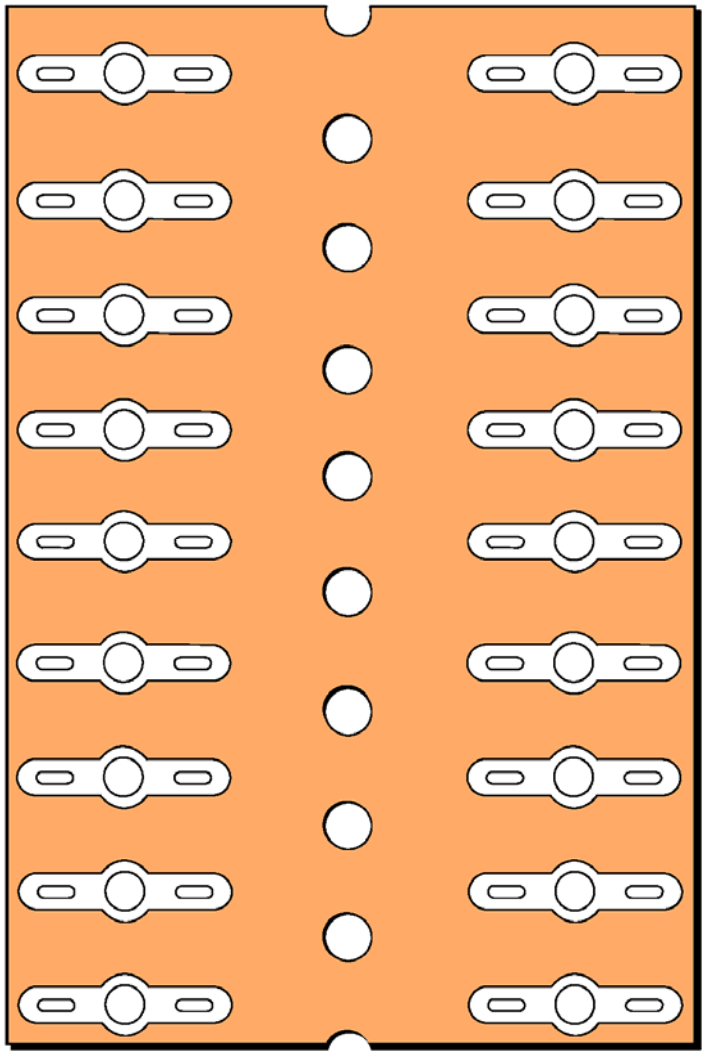


Figure 217 Soldering Strips

2.2.2 Raster Boards

Raster boards are used for simple electronic circuits. They consist of hard paper without copper coating. The board is perforated with holes 1.3 mm in diameter and in a raster of either 2.54 mm or 5.08 mm. The components are fitted to one side, designated 'component side'. The connections of the components are led through the holes and soldered together on the back side of the component side. The leads of the components or silver -coated wire may be used for the connection of the components.

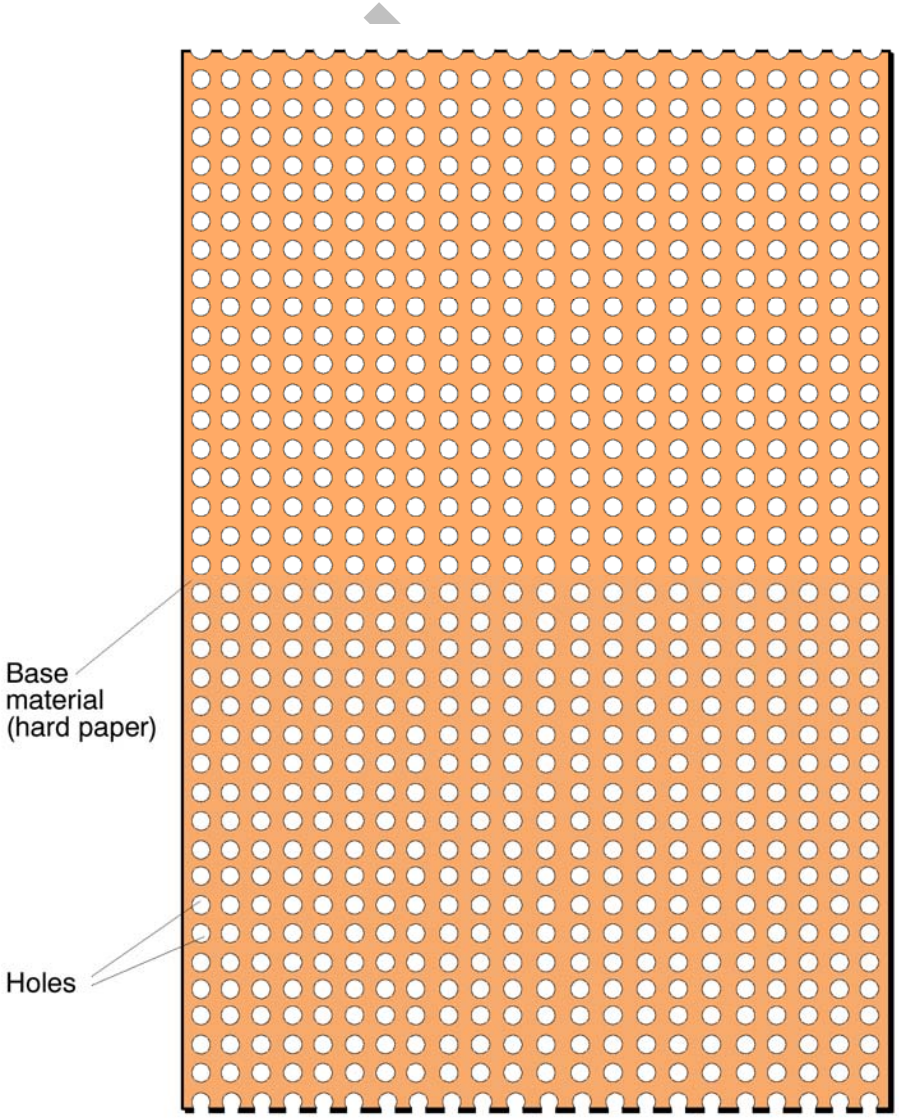


Figure 218 Raster Board

2.2.3 Tag Raster Boards

Raster boards with copper coating around each hole (tag) are designated 'tag raster boards'.

The copper coating has a thickness of 35 μm. The diameter of the hole is 1.0 mm to 1.3 mm with a tag diameter of 3 mm. The base material may be hard paper or epoxy resin paper in 1.5 mm thickness.

The components can be fitted to each side of the tag raster board. The leads of the components, silver -coated wire or solder bridges, may be used for the connection. Raster boards are available with different tag rasters, e.g. 2.54 mm or 5.08 mm, and as an additional feature with an integrated plug -in junction.

The advantages of tag raster boards are:

⇒ Test set -ups can be built without preparation in a very short time.

⇒ The circuit and circuit layout can be changed very easily.

⇒ Tag raster boards are reusable.

Mechanical instability and low adhesion of the tags are disadvantages of tag raster boards. Therefore great care must be taken when components are unsoldered from tag raster boards. The components and connections must be removed without excessive force, otherwise the tags may be disconnected from the board.

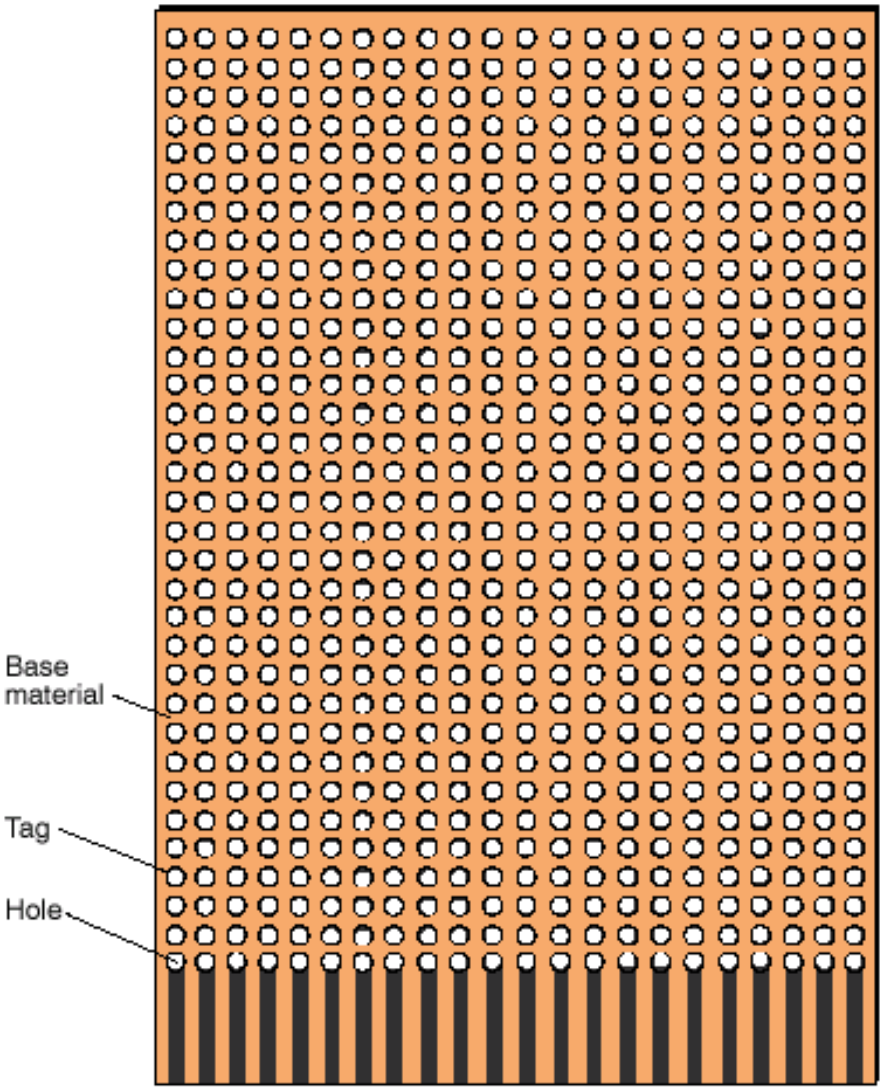


Figure 219 Tag Raster Board

2.2.4 Veroboards

Prototypes and test set -ups of electronic circuitry can be built very easily on Veroboards. Veroboards consist of high -quality hard paper as base material. One side of the base material is coated with straight copper conductor tracks arranged in parallel. The copper tracks are perforated in a defined raster whereby 2.54 mm or 5.08 mm are commonly used. Veroboards are available in different forms.

Advantages of veroboards are:

⇒ Test set-ups can be built without preparation in a very short time.

⇒ The circuit can be changed very easily.

⇒ Veroboards are reusable.

Mechanical instability caused by the perforation of the board and low adhesion of the copper conductors are disadvantages of veroboards. Therefore great care must be taken when the components of the veroboard are unsoldered. The components and connections must be removed without exerting excessive force.

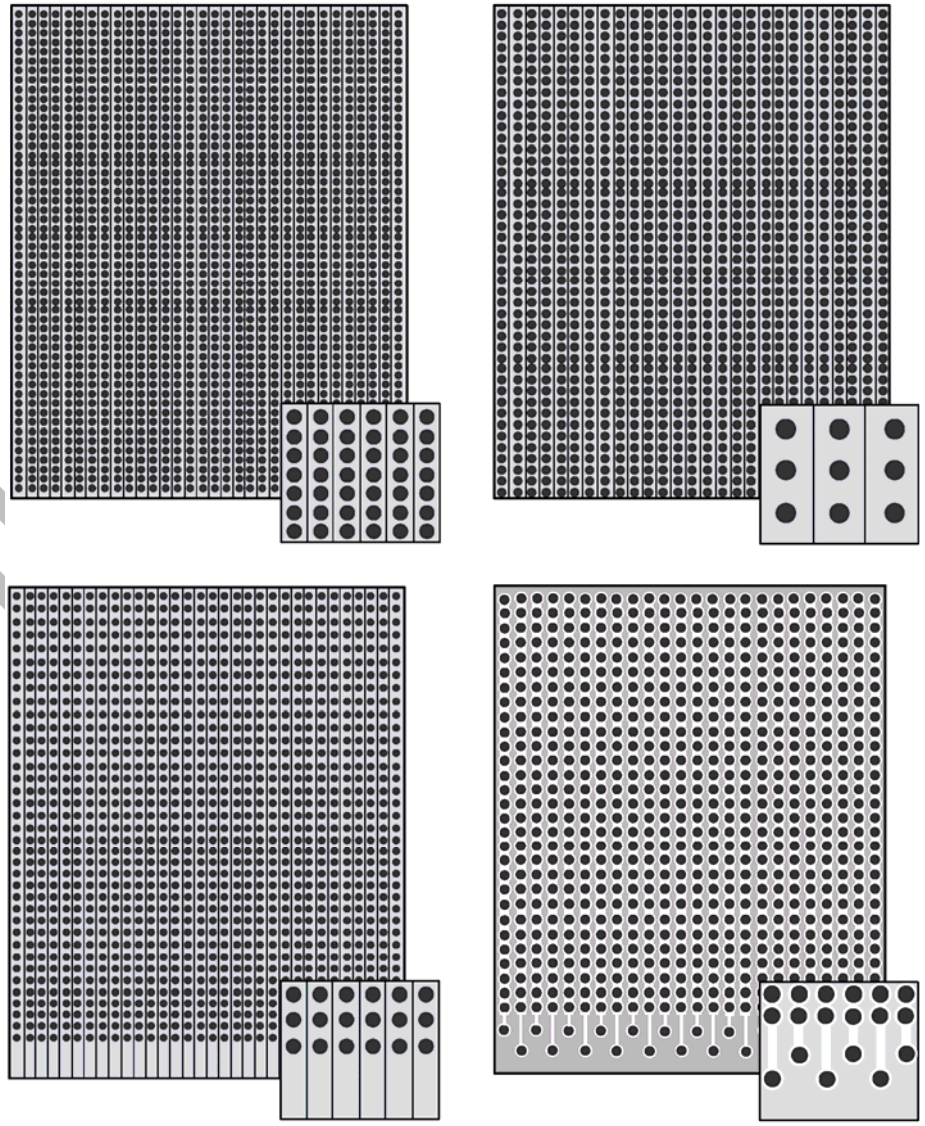


Figure 220 Veroboards

2.3 PRINTED CIRCUIT BOARDS

Today, there is an increasing diversity of types of electronic modular assemblies. Many of them have been built with new kinds of basic materials and manufacturing processes and techniques.

A firm understanding of these elements of construction will enable you to fully analyse the repair problem at hand and accomplish the task without further damage to the circuit board.

Modules and Assemblies

Most modern electronic systems are produced with building blocks commonly known as modules. The general concept of modular construction has become the standard method for constructing the most complex of systems. Primarily, the modular concept permits us to build up complex Systems with assemblies that are readily interchanged.

Modular Construction

The basic module is a printed circuit board assembly whose end function terminates at the connector device. It may take many forms ranging from the simple, single sided, un-coated circuit board with standard sized components mounted in a low density manner, to a double-sided or multi-layered circuit board packaged with mechanical hardware and multi-leaded sub-modules or micro-miniature components that are coated with or encapsulated within a variety of coating materials. Variations in termination and mounting techniques also tend to complicate matters.

Sub-Modules

Sub-modules are separate assemblies of components that have been connected or mounted to a circuit board assembly. They may be found in a variety of sizes and shapes and may have two or more connecting leads. The leads may be of various spacing, dimensions and materials, such as copper, nickel, kovar or (cobalt, iron alloy). The sub-module assembly may be conformally encapsulated with a potting material. These contribute to the repair problem. Internally, parts of the sub-module may be interconnected by either soldering or welding.

Figure 221 Example of Sub-modular Design

Circuit Board Elements

Printed circuit boards are constructed of two main elements. The first is the base material, sometimes called the substrate. The second is the circuitry made of a conductive foil bonded to the base material.

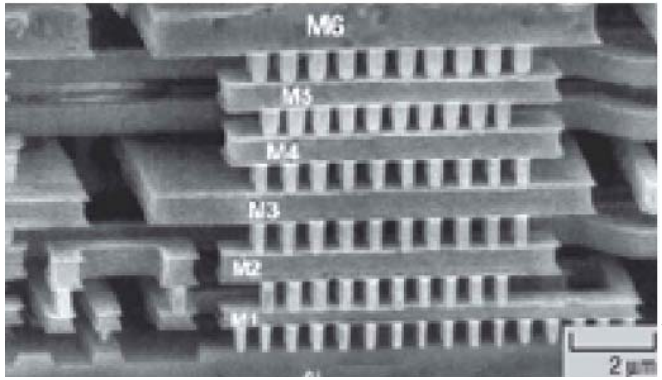


Figure 222 Example of Integrated metal Substrate

Base Material or Substrate

The base materials, listed in Figure 223, include most of those which have been used over the years.

Material	Colour
Phenolic Paper	Tan
Phenolic Cotton	Tan with weave
Phenolic Glass	Tan with weave
Silicone Glass	White to light tan
Epoxy Glass	Green with weave
Epoxy Composite	Green

Figure 223 Base materials used for Substrates

One of the earliest materials used was a laminate impregnated with phenol resins. This material is primarily for moderate temperature applications. Its colour is either natural or a shade from tan to brown. It is usually opaque and has no discernible weave since the paper laminating material does not have a weave.

Another material, used for higher temperature applications, is laminated fibreglass cloth impregnated with epoxy resin. Its colour can be natural, tan, green or others. It will usually be translucent and a cloth texture can be noted just below the surface. This material has a mechanical strength that is two to three times greater than that of phenolic laminates and it also has greater bond strength to the laminated copper. Other base materials used for special applications include ceramic, steel, Mylar, polyamides and others.

Most circuit boards are rigid and will keep their shape. However, some boards such as Dupont, Mylar and Kapton, are purposely made of thin and flexible material so they can be formed to fit the available space, or can be bent around corners to replace bulky wire harnesses.

Whatever the specific material used, its qualities must include high insulation resistance, low moisture absorption, good temperature stability, resistance to shock and where important, resistance to fungus growth.

Recognising the type of base material used helps determine the bond and mechanical strength factors for the specific module being repaired.

Circuitry

The circuitry is the second main element of a printed circuit board. It is made of a conductive foil bonded to the base material.

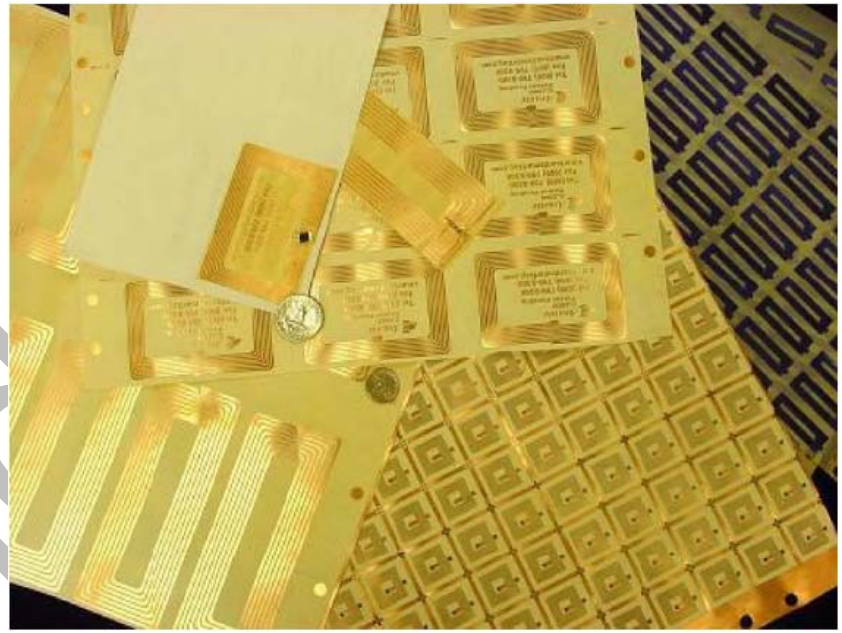


Figure 224 Etched Conductive Foil Pattern

The type of foil most used is copper, which is commonly found in two main thicknesses: 0.0014 inches (.036mm) and 0.0028 inches (.071mm). The thinner (0.0014 inches) foil is known as one-ounce copper, since one square foot of it weighs one ounce. The thicker (0.0028 inches) foil is known as two ounce copper. A third type, commonly used for internal circuits on multi-layer boards, is known as half-ounce copper and has a thickness of 0.0007 inches (.018mm).

Figure 225 Typical Modular Circuit

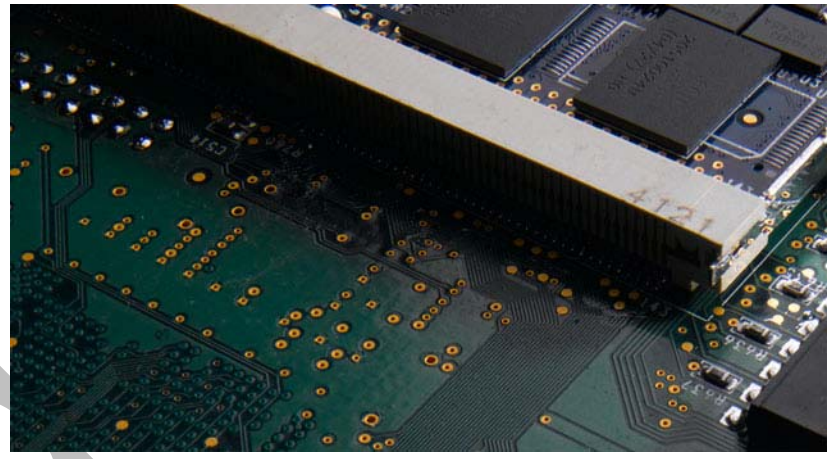


Figure 226 Typical Circuit Breakdown

The circuit consists of the conductors (also called runs), edge connectors, ground planes, heat sinks (because their large mass helps to dissipate heat), and pads or terminals as shown in figure 225.

The copper is extremely thin in comparison to the circuit board, this is a very important relationship to comprehend because the improper application of heat to circuitry, during repair, can overheat the circuitry and cause delamination of the copper as well as heat damage to the board itself.

Production of Printed Circuit Boards

Industrial electronics as well as consumer electronics face the following problems regarding their products:

⇒ The products must be small and of good quality.

⇒ The costs must be as low as possible.

Printed circuit boards (PCB) meet these requirements. In combination with semiconductors and highly integrated circuits electronic circuits can be realized in very small sizes and at very low costs.

The Principal Production Procedures

The printed circuit board can combine several tasks, from simple replacement of hand -wired circuits up to the combination of isolation purposes, carrier purposes and circuit layer.

The process of generating printed circuit boards is not uniform and differs in the employed technique. The mostly used techniques are the additive technique and the subtractive technique.

Principle of the Subtractive Technique

Great plates of base material with either one side or both sides coated with a copper layer are used for the subtractive technique. These plates are split up into boards.

In the next step the positive picture of the copper conductor layer is applied to the boards either via photo printing, screen-printing or offset printing. The lacquer applied is resistant to etching.

The following etching process removes all the copper not covered with the etching -resistant lacquer. As a result only the copper covered with lacquer remains on the boards.

Thereafter, the boards are cleaned, i.e. the lacquer is removed from the boards. The holes necessary to fit the components are either drilled or punched into the boards.

Before the components can be fitted, the boards are covered with a special lacquer as surface protection. In the last step the components can be fitted to the boards.

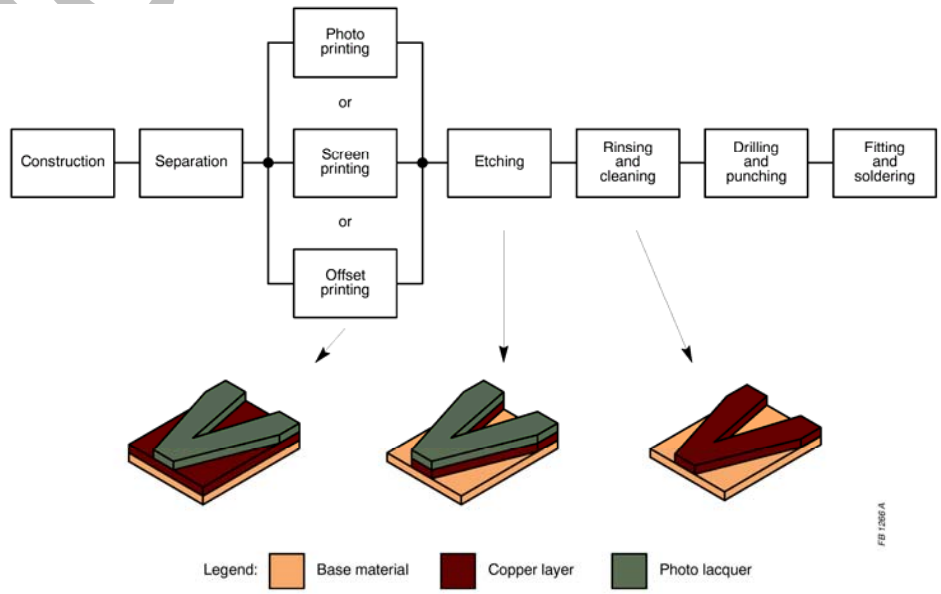


Figure 227 Principle of subtractive Technique

Principle of the Additive Technique

Base material without copper layer is used for the additive technique. As it is known from the subtractive technique, the great plates are divided into the necessary boards. In the next step holes for fitting the components are either drilled or punched into the boards.

The boards are pretreated by a chemical process for the copper to be fitted in the following step. Before the copper can be applied to the surface of the boards either via photo printing, screen printing or offset printing. The lacquer applied is again etching-resistant.

The result is that the copper can only adhere to those surfaces of the board which are not covered with lacquer. Thus the copper layer is applied to the base material.

Subsequently the board is cleaned and covered with surface protection lacquer. In the last step the components can be fitted to the board

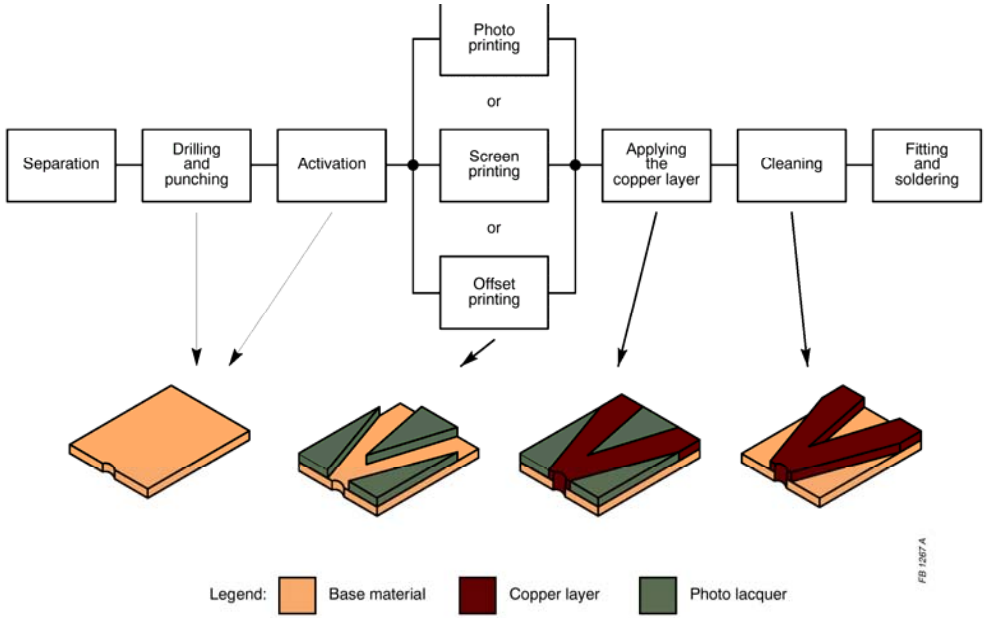


Figure 228 Principle of the Additive Technique

Through-Holes in Circuit Boards

In addition to its base material and conductive, foil, the usual circuit board also has a number of holes drilled through it to allow the mounting of components. These holes are of two kinds: unsupported and supported plated-through. This is shown in figure 229.

Unsupported holes are simply holes in the base material, drilled through from one side to the other, but providing no electrical connection between the two sides of the board.

Plated-through holes (supported), on the other hand, have their walls lined with copper so that they form a continuous electrical path, called a through connection, running through the board, from one side to the other. Plated-through holes are also known as reinforced holes. In addition to the use of plating, reinforcement may be through the use of copper eyelets or funnelets that have been inserted into the hole to create continuity from one side of the board to the other (while retaining compatibility with the copper foil).

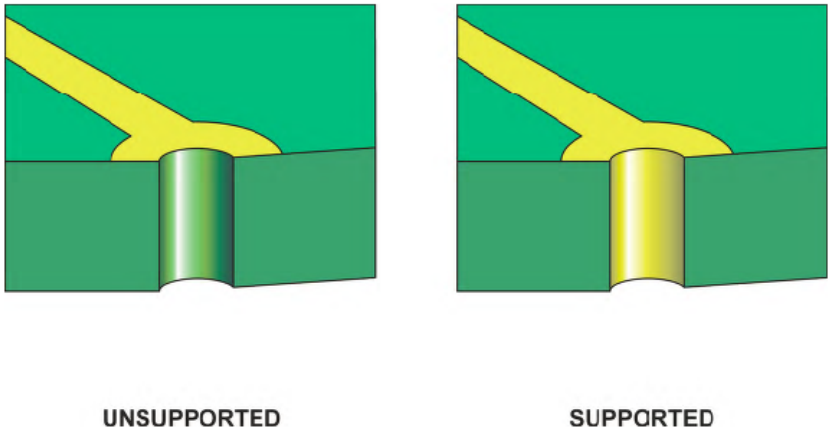


Figure 229 Through Holes in Circuit Boards

Eyelets

Eyelets are usually made of pure copper. The copper material in these devices must have a softness factor (Ductility Level) that will prevent splitting when they are set in a circuit board. Most eyelets are plated with tin-lead or pure tin.

Circuit Board Configurations

If circuits exist on only one side of a board, it is a single-sided board; if on both sides, it is a double-sided board. A third type is the multi-layered board, which is a sandwich of a number of layers, each one with its own circuitry. The following descriptions summarise the construction of each of these types of board.

Single-sided Boards

The single-sided board has combinations of conductors, terminals (pads), and ground planes bonded to one side of the base material. The usual board also has a number of holes drilled through it to allow the mounting of components. This is shown in figure 230.

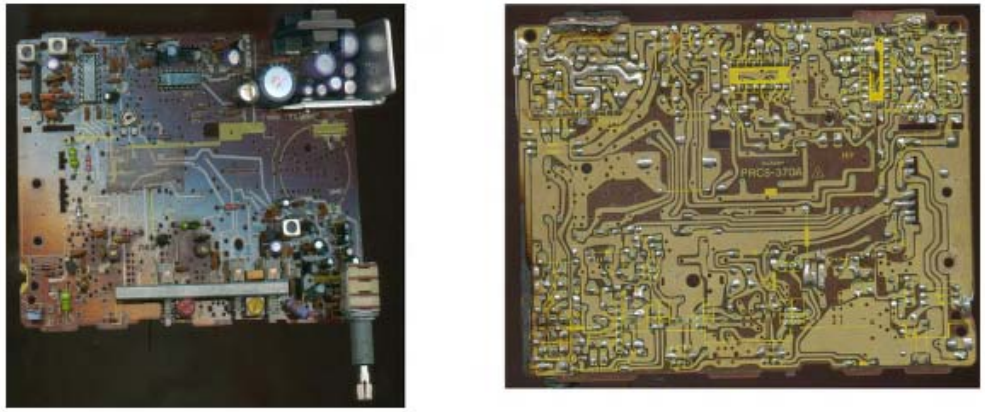


Figure 230 Single-sided Circuit Board

In a single-sided board, the pads are normally unsupported; there is no through-hole copper to support the pad to the board. The components are normally mounted on the back (non-circuit) side of the board, and the component leads are fed through the holes and soldered to the pads.

Double-sided Boards

The double-sided board has combinations of conductors, terminals and ground planes located on both sides of the board. In most applications, there is a continuous conductive path, or interconnection, between circuitry on both sides. This is usually via a plated-through hole, in which copper is plated to the hole walls and also to the pads on either side of the board. This is known as a supported hole configuration. In some boards, this through-hole support can also come from eyelets or funnelets that have been inserted into the hole and fused to the pads on either side.

In construction, the components are normally mounted on one side of the board, with the leads fed through the hole. The leads are soldered from the circuit side, but the solder flows through the hole, providing a continuous solder joint through the entire hole, with fillets located on both sides for additional lead support and solder joint integrity. This is shown in figure 231.

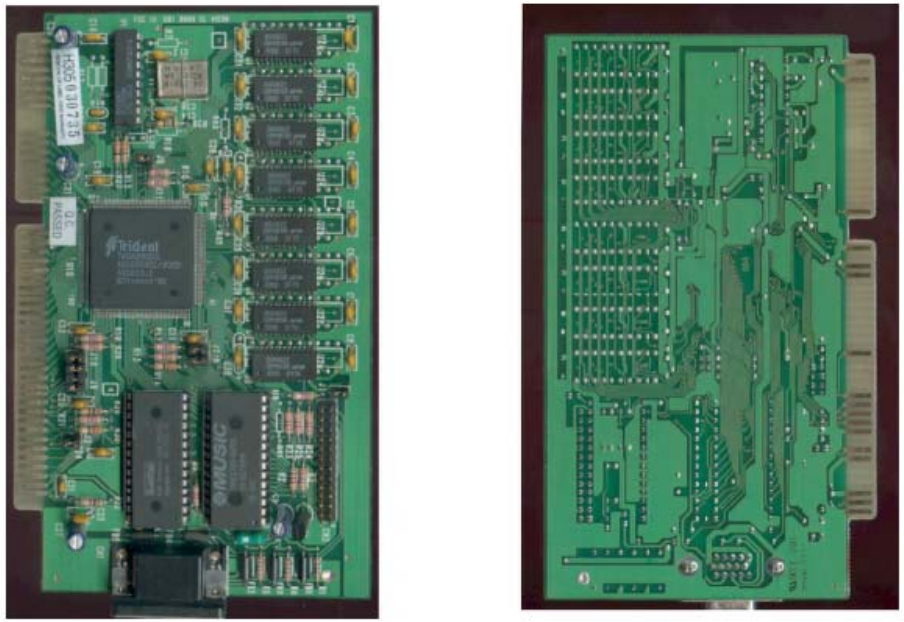


Figure 231 Double-sided Circuit Configuration

Multi-Layer Boards

Multi-layer boards, as shown in figure 232, are similar in construction to the double-sided type of board, wherein a plated-through hole is utilised to provide a continuous electrical connection between the circuitry on both outside surfaces of the board. The multi-layer board, however, utilises one or more **'internal'** conductor planes that are sandwiched in layers within the board (up to 48 layers). These are connected, in many locations, to the plated-through holes in the circuit board.

It is very important to consider this internal construction for component removal, since the internal planes can act as heat sinks absorbing the heat applied to the surface solder and pad, and preventing it from completely melting the joint.

Excessive heat at the hole/internal plane interface can cause damage and loss of continuity with the internal plane, often with no outward sign of damage. Since all plated through holes in a multi-layer may not interconnect with all (or any) of the internal planes, each plated through hole and its associated solder and lead must be considered independently with respects to thermal mass, heat transfer and dwell time on the joint.

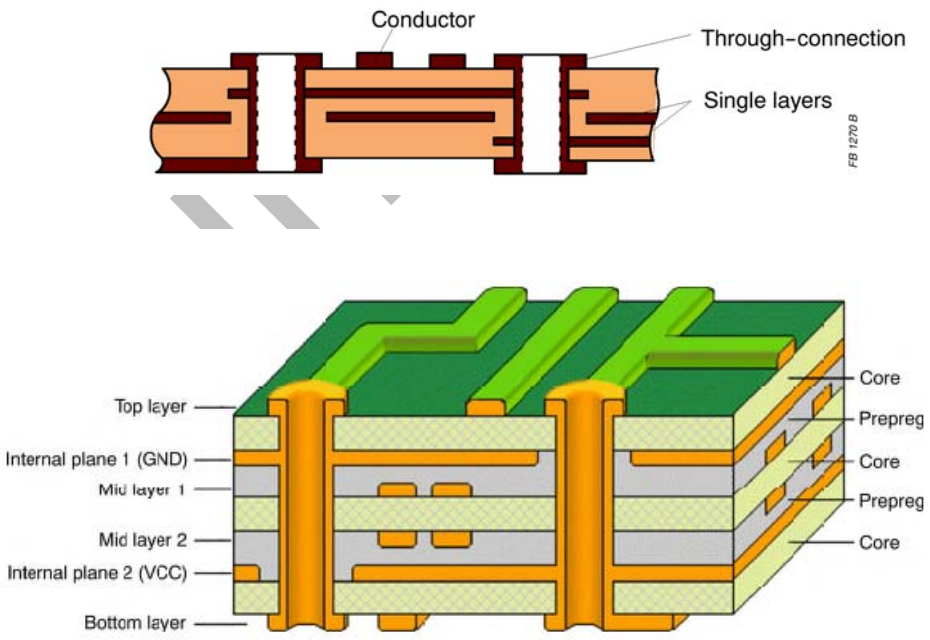


Figure 232 Multi-layer Board

Flexible Circuits

Though not as common as the rigid substrate type of printed circuit board, flexible printed wiring assemblies and flexible flat cables as shown in figure 233 are finding increased use in a variety of industries including aerospace.

Flexible printed wiring consists of etched conductors of electrolytic or rolled copper laminated to one or more layers of a flexible insulating base of Mylar or Kapton.

This type of circuitry is generally custom designed for a particular application, such as controlling positions on antenna and often used as an interconnection system within a component package.

Flat cable, on the other hand, is most often used like conventional harnesses: to interconnect separate component packages, to go around corners in high density packages, or to reduce the effects of vibration fatigue on the conductors.

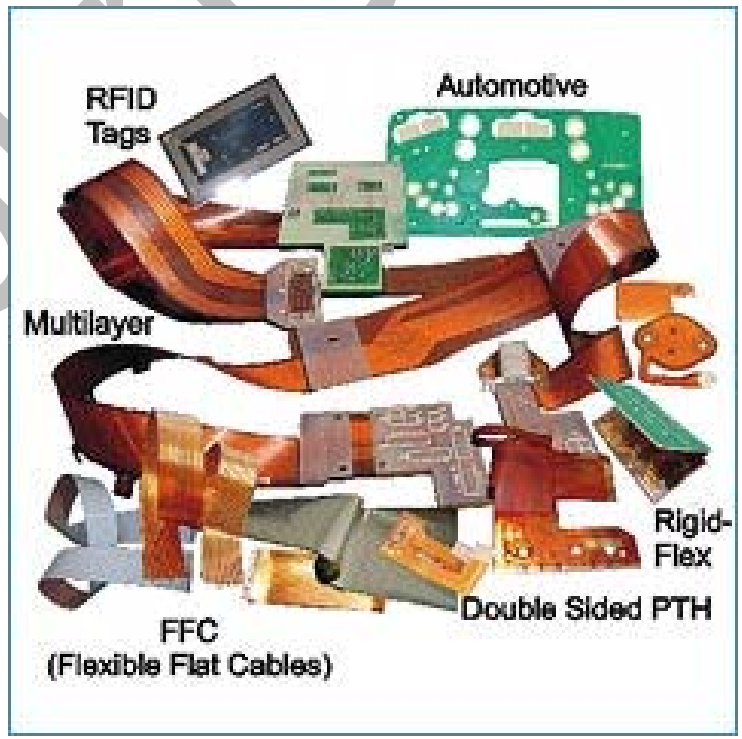


Figure 233 Examples of Flexible Circuit Boards

Component Types and Mounting Techniques

Common components used in electronics are usually referred to as the discrete components or piece parts of a modular assembly. They can be separated into categories by function, such as resistors, capacitors, transistors, diodes, etc. In each category there are several sub-categories: for example, resistors may be carbon composition, vitreous enamel, wire wound or ceramic and may be fixed, variable or adjustable. The same may be true for other discrete components.

In addition, a component may be classified as axial leaded, such as the resistor or multi-leaded, such as the transistor, dual inline pack or LSI micro-processor style package.

Each type of component has its own peculiarities and may be mounted in one of many different ways, depending on weight, shape, size, mechanical configuration, heat sensitivities and package size restrictions. Additionally, the component may be subject to electrical damage and require special precautions during handling.

In most instances, discrete components are mounted flush to the board surface to aid in heat dissipation and may be mechanically supported (depending on weight), conformably coated or otherwise bonded to the work piece surface. In other instances, the components may be vertically mounted.

Figure 234 shows the six most common methods of terminating soldered component leads to printed circuit boards. These are:

The Straight-Through (Unclinched) Lead

This type of termination is the easiest to remove and replace because there are no bends in the leads after they pass through the holes.

The Clinched Lead

This method is usually used with unsupported holes, but may be used with supported holes. The lead is inserted into the hole and then the lead end is being 90 degrees onto the conductor. This helps stabilise the component on the board so that it will not move during soldering.

The Swaged (Spaded) Lead

The lead ends extending through the printed circuit board are swaged or flattened to increase their width beyond that of the hole, thus retaining the components in their proper position during the handling and soldering. This method required special procedures for removing the component.

Surface Mounting

The leads of flat packs and other planar-mounted components are mounted and soldered onto the flat surfaces of the pads on the board, rather than passing through holes in the board.

Off-set Pad Mounting

This is a variation of the Surface Mounting method. The lead is passed through a hole in the board and is bent so that it can be surface-mounted to a pad on the opposite side. In some instances, the pad to which the lead is joined is some distance away from the hole.

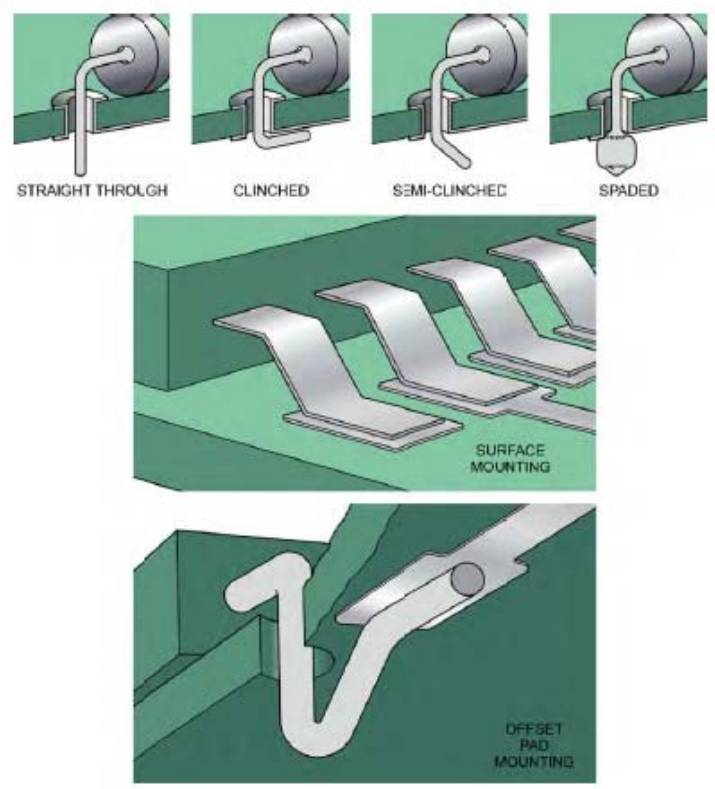


Figure 234 Soldered Component Termination Methods

Circuit Board Packaging

On some boards, the components are mounted widely spaced apart; this is known as low-density packaging. Boards with high-density packaging have many components mounted very close together. With low-density packaging, you have more space to work with and it is easier to get at the components in order to remove them. But with high-density packaging, circuit runs are much narrower and more closely spaced and components may be mounted so close together that some may be totally inaccessible. In this case, you may need to partially disassemble the board to get at them. This is where your knowledge of how the board was put together is useful; you have to 'de-manufacture' it to get at the parts, then **'re-manufacture'** it when you replace them.

Coatings

A major problem in the repair of some printed circuit board assemblies is the removal of conformal coatings that cover circuitry and components. These coatings are used to a variety or reasons:

- To provide electrical insulation.
- To cushion against mechanical shock and vibration.
- To prevent abrasion.
- To help support and bond the components.
- To function as a heat sink, or to protect against humidity or fungus.

In most cases, these must be partially removed to gain access to the solder joints during component test or removal.

The general characteristics of coatings, for purposes of removal, are hardness, degree of transparency, thickness, solubility and thermal properties.

Here are some of the main types of coating materials you will encounter in electronics, along with a few of their identifying characteristics.

Varnishes are hard, applied in thin coats, translucent or opaque and non-soluble in mild solvents.

Acrylic Lacquer is hard, usually thin, transparent and soluble in lacquer thinner or xylene.

Epoxies are hard, may be applied thickly or thinly, may be transparent or opaque and are non-soluble in mild solvents.

Silicone Varnish is soft, may be applied thickly or thinly, is transparent and is soluble in most mild solvents.

Polyurethane may be hard or soft, may be thickly or thinly coated, is always transparent and is non-soluble in mild solvents.

RTV rubber (or silicone rubber) is opaque, soft and usually very thick. The colour is normally white, red-brown or black.

Parylene is uniformly thin, non-soluble, transparent and very tough.

Coatings are applied in a number of different ways:

- Dipping.
- Spraying.
- Brushing.
- Vacuum condensation.

Spraying is used for thin applications, while the dipping method is used for thicker coating.

Parylene is applied by vaporising the required quantity of material in a vacuum changer; the coating then condenses uniformly on the entire assembly. Masking is used over areas not requiring coating and is removed after the coating process.

Once they are applied, coatings must be dried or cured. This is done in a number of ways, depending on the type of coating used. Air drying and heat curing are two of the most commonly used methods

3 SERVOMECHANISMS

3.1 INTRODUCTION

Servo mechanisms, also called SERVO SYSTEMS or SERVOS for short, have countless applications in the operation of electrical and electronic equipment. In working with radar and antennas, directors, computing devices, ship's communications, aircraft control, and many other types of equipment, it is often necessary to operate a mechanical load that is remote from its source of control. To obtain smooth, continuous, and accurate operation, these loads are normally best controlled by synchros.

As you may already know, the big problem here is that synchros are not powerful enough to do any great amount of work. This is where servos come into use. A servo system uses a weak control signal to move large loads to a desired position with great accuracy. The key words in this definition are move and great accuracy. Servos may be found in such varied applications as moving the rudder and elevators of a model airplane in radio-controlled flight, to controlling the diving planes and rudders of nuclear submarines. Servos are powerful. They can move heavy loads and be remotely controlled with great precision by synchro devices.

They take many forms. Servo systems are electro-mechanical, electro-hydraulic, hydraulic, or pneumatic. Whatever the form, a relatively weak signal that represents a desired movement of the load is generated, controlled, amplified, and fed to a servo motor that does the work of moving the heavy load.

3.1.1 Control Systems

The main purpose of a control system is to keep a physical quantity at a desired value. A control system is designed to level out external influences that act upon the system to be controlled.

In general terms, a control loop, as shown in figure 235, consists of the following elements:

⇒ controlled system (process)

⇒ measuring transducer (sensor)

⇒ controlling system (controller)

⇒ correcting element (actuator).

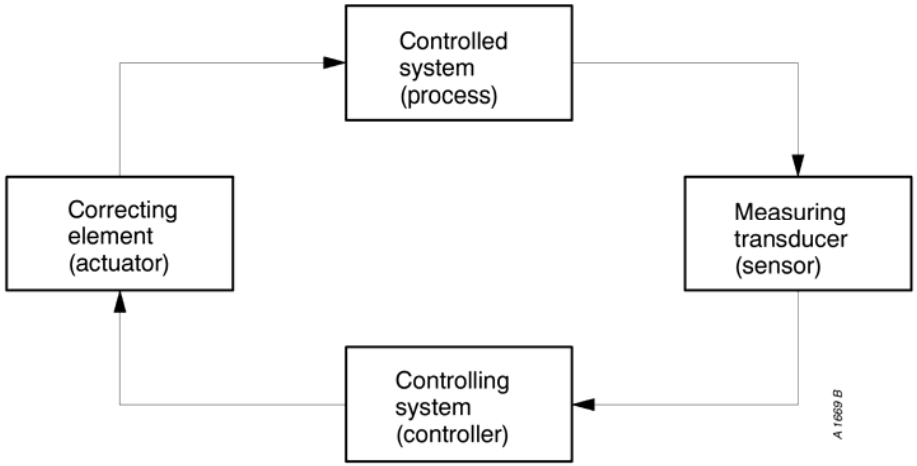


Figure 235 Basic elements of Closed-Loop Control

The controlled system (process) is the processing plant range which is to be operated to keep the controlled physical quantity at the desired value.

Because two different systems must 'communicate', additional devices are used as 'interpreters'. These are the measuring transducer (sensor) and the correcting element, consisting of a motive element (actuator) and an ability to provide an influence on the process.

In the controlling system the desired value is compared with the measured value. Depending on the result, the controller generates a signal which is fed to the correcting element.

To obtain the optimum performance of the whole system, the controlling system may carry out very complex functions.

Control Variables

Control units operate as electronic, pneumatic, hydraulic or mechanic devices. Process inputs and outputs can be all kinds of physical values, for example temperature, flow, pressure or weight.

These physical values are variables. They are processed in the controlling unit.

There are generally five variables used to describe the functions of a complete closed-loop control system, as shown in figure 236.

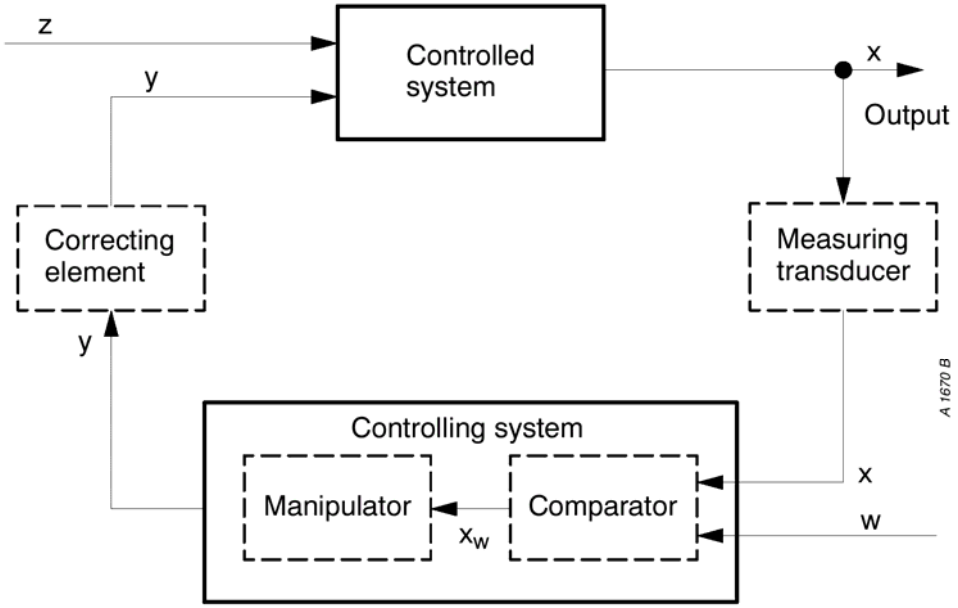
The controlled variable 'x' is the output of the controlled system which is fed back as the input to the controlling system. Here it is sensed and used to generate the controlling signal 'y'.

The command variable 'w' of a system is an input signal which is externally fed into the control circuit, as the intended value of the controlled variable 'x'.

The deviation 'x_w ' is an error signal resulting from the difference between the command signal 'w' and the controlled signal 'x'. Deviation may have a positive or negative sign, depending on the momentary magnitudes of the command signal and the controlled variable.

The correcting variable 'y' is the output signal of the controlling system and the input signal to the controlled system. It transmits the correcting action to the controlled system.

The disturbance variable 'z' represents all disturbing influences. These tend to affect the intended control function adversely.



x = Controlled variable (output feedback signal)

w = Command variable (command signal)

x_w = Deviation (error signal)

y = Correcting variable (correction signal)

z = Disturbance variable (system input signal)

Figure 236 Closed-Loop Control Variables

3.2 TYPES OF SERVOMECHANISM

Servomechanism may be classified according to two main categories; these are the **open loop** and **closed loop** systems. An input transducer will generate an electrical equivalent of the demanded angular position, the demand voltage; this voltage is amplified to a power level sufficient to enable the motor to drive the load. The output shaft velocity in a rate servo, or the final output shaft position in a remote position control (RPC) servo, will depend on the following factors.

- Variation of load conditions.
- Frictional forces within the motor and its load, and the mechanical interconnection (gear trains, clutches, linkages, drives etc).
- Variations of power supply.
- Value of the demand voltage.
- Variations of amplifier gain.

3.2.1 Open Loop Systems

Figure 234 shows such a system. The open loop system suffers from the major disadvantages that the above-mentioned factors do not remain constant. For example, for a given demand, the value of output voltage will vary as the gain of the amplifier alters with time and temperature. The frictional forces within the motor and load will change with velocity, temperature and load. Variation of supply voltage and frequency will cause variation of the speed of the motor and ultimately of the load speed and position, even if the load torque remains constant. There is no means of precisely controlling these factors and, therefore, the open loop system is not good enough for close tolerance control.

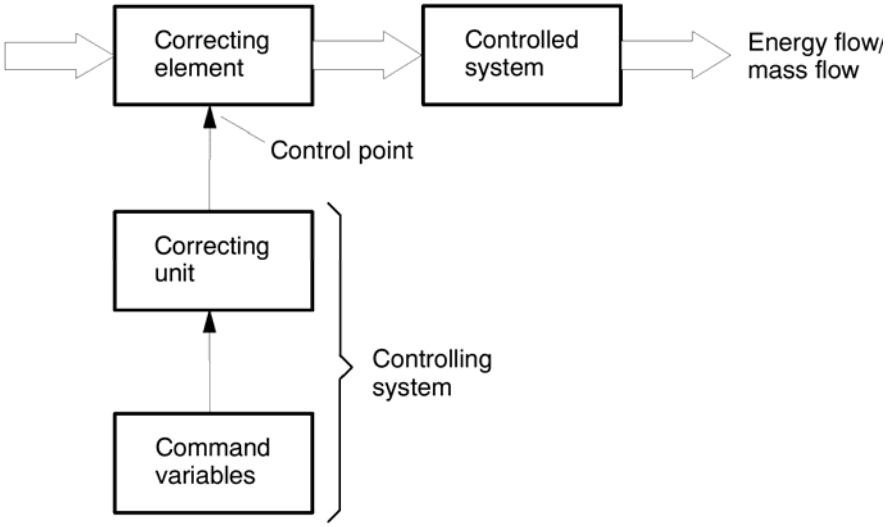


Figure 237 Open Loop System Block Diagram

To remove the variables and uncertainties present in open loop, it is necessary to obtain information about the behaviour of the output shaft and to compare this with the input demand. A control system that does this is said to be operating under closed loop conditions.

3.2.2 Closed Loop Control System

A closed-loop control system is another name for a servo system. To be classified as a servo, a control system must be capable of the following:

- Accepting an order that defines the desired result.
- Determining the present conditions by some method of feedback.
- Comparing the desired result with the present conditions and obtaining a difference or an error signal.
- Issuing a correcting order (the error signal) that will properly change the existing conditions to the desired result.
- Obeying the correcting order.

We have discussed the open- and closed-loop control systems and defined a servo system as a closed-loop control system. Although not technically accurate by definition, open-loop control systems are also often referred to in many publications as servo systems even though they lack one of the five basic requirements, that of feedback.

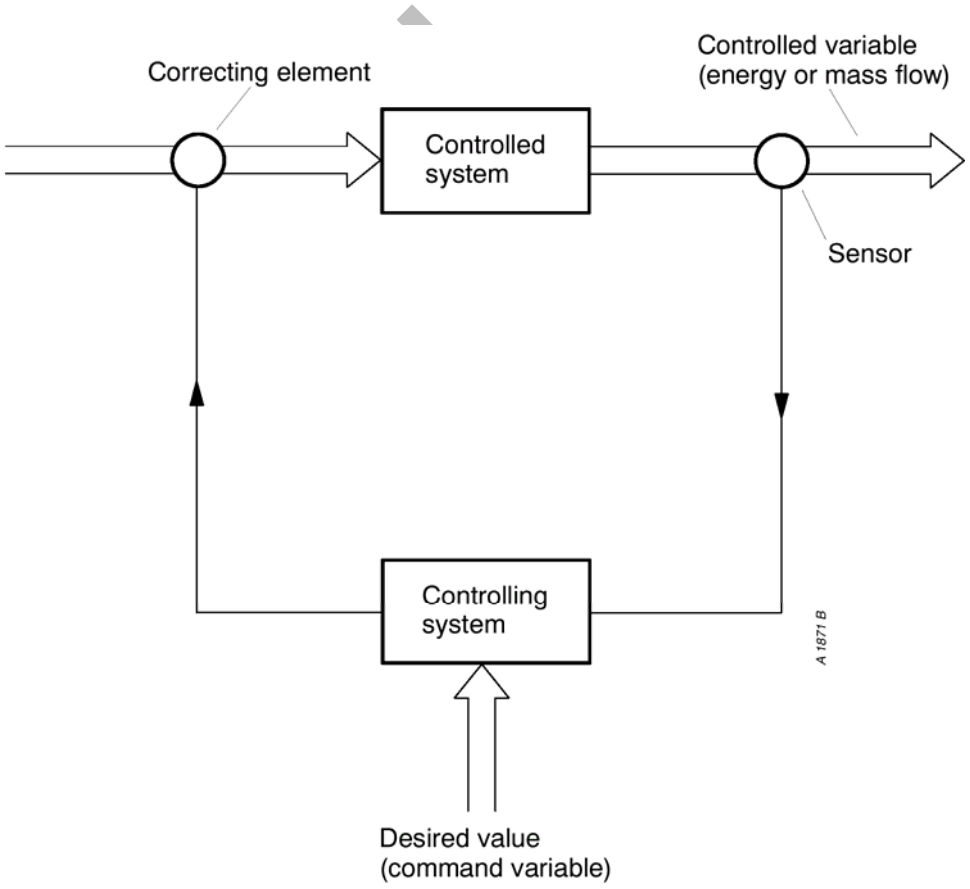


Figure 238 Basic Closed-Loop Control

For the following discussion of a servo system, refer to figures 239-242. This closed-loop servo system is one of the most common types in use today. It is normally made up of electro-mechanical parts and consists basically of a synchro-control system, servo amplifier, servo motor, and some form of feedback (response).

The synchro-control system provides a means of controlling the movement of the load, which may be located in a remote space. The servo amplifier and servo motor are the parts of the system in which power is actually developed (to move the load). In figures 239-242, assume that the control signal will be initiated by a handcrank input connected to the synchro transmitter (CX). The dials located on the CX and the CT indicates the positions of the synchro's rotors, while the dial on the load indicates the position of the load.

In figure 239, the dials of both the CX and the load indicate that the load is in the desired position. Because the load is where it should be, there will be no error signal present at the servo amplifier and no power to the servo motor.

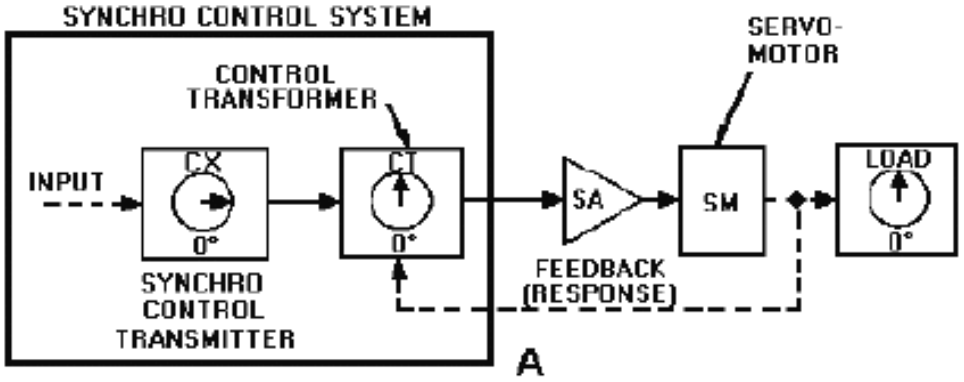


Figure 239 Basic Closed Loop Servo System

In figure 240, the rotor of the CX has been moved by the hand crank to 90º. (This indicates that it is ordered to move the load by 90º). Notice that the rotor of the CT is still at 0º. The CT now develops a signal, called the ERROR SIGNAL, which is proportional in amplitude to the amount the CT rotor is out of correspondence with the CX rotor. The phase of the error signal indicates the direction the CT rotor must move to reduce the error signal to zero or to "null out." The error signal is sent to the servo amplifier.

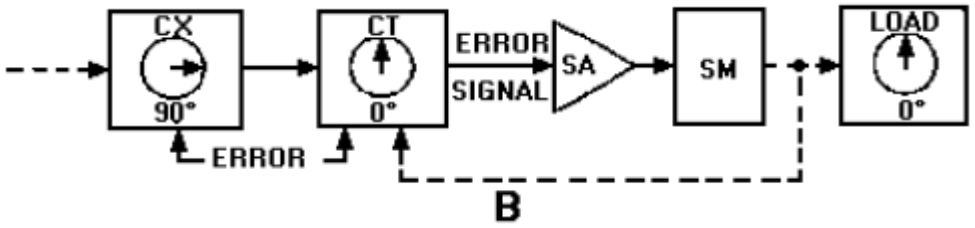


Figure 240 Closed Loop Servo System

In figure 241, the error signal has been amplified by the servo amplifier and sent on to the servo motor. The motor starts to drive in the direction that will reduce the error signal and bring the CT rotor back to the point of correspondence. In this case the motor is turning clockwise. The mechanical linkage attached to the servo motor also moves the rotor of the CT. This feedback causes the amplitude of the error signal to decrease, slowing the speed at which the load is moving.

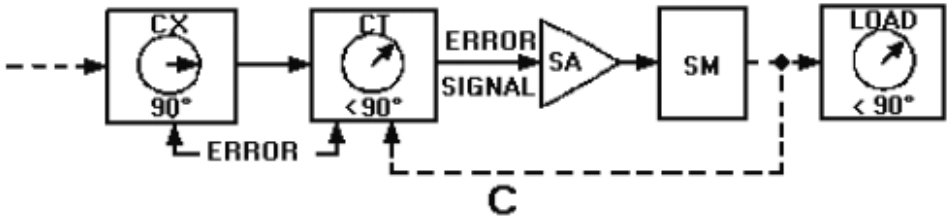


Figure 241 Closed Loop Servo System

In view D, the servo motor has driven both the load and the rotor of the CT, so that the CT is now in correspondence with the CX rotor. As a result, the error signal is reduced to zero (nulled). The load stops at its new position.

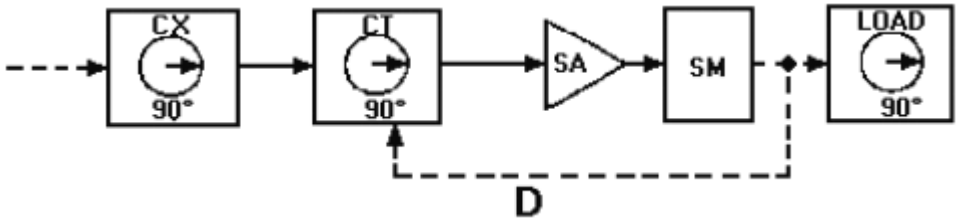


Figure 242 Closed Loop Servo System

Note that in this servo system, we moved a heavy load to a predetermined position through the simple turning of a hand crank. In responding to the hand crank, the servo system performed a basic positioning function.

Two key points for you to remember, thus far, about the operation of the closed-loop servo system are:

- The original error (movement of the CX rotor) was "detected" by the CT. For this reason the CT is called an ERROR DETECTOR.
- The servo motor, in addition to moving the load, also provides mechanical feedback to the CT to reduce the error signal. For this reason the servo motor is also called an ERROR REDUCER.

Classification

Closed loop servo mechanisms may be dc, ac or pulse operated and can be classified by:

- Application.
- Method of damping.
- Degree of damping.

Types of Closed Loop Servo Mechanisms

- Positional Servo: Control of position.
- Rate Servo: Control of rate and direction.
- Computing Servo: computes output functions from input information.

Basic Requirements of Closed Loop Servo Mechanisms

- Accept an order.
- Evaluate the position of the input and output.
- Evaluate the magnitude, direction and rate of change of any error.
- Carry out correcting orders.

Remote Position Servo System

DC servo systems are not often used in aircraft systems on modern aircraft due to practical problems of:

- Reliability of potentiometers and input/output transducers.
- Temperature stability of dc amplifiers.
- Maintenance of brush gear of dc motors.

The system does two basic things:

- Determines difference between input and output positions.
- Corrects difference between input and output positions.

Position Feedback

A Voltage (Vo) proportional to the output shaft angular position is applied in opposition to the input voltage (Vi) at a signal summation point. This position feedback signal must always reduce the input to the amplifier (error signal), from the summation point, towards zero.

Output Torque

The torque available for positioning the output load is limited only by the amplifier and dc motor. Due to inertia of the load there is a tendency for the load to overshoot the required position.

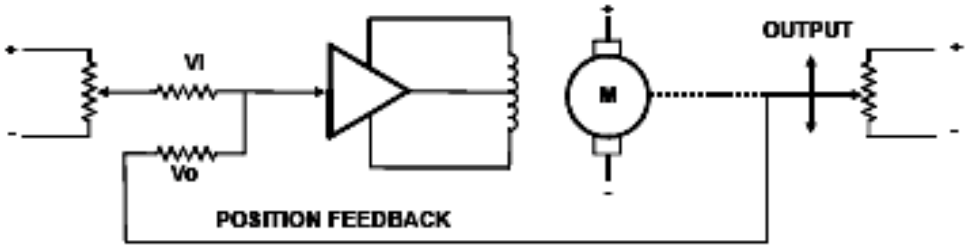


Figure 243 DC Remote Position Servo System

Servo systems using control synchros, ac amplifiers and ac servo motors are more commonly used on modern aircraft.

Error Detection

The relative positions of the input shaft angle (CT stator field) and the output shaft angle (CT rotor position) can be compared in the Control Transformer (CT).

Position Feedback

The CT rotor position is directly controlled by the output shaft which is in turn positioned by the ac servo motor.

Error Reduction

The servo motor must always position the CT rotor to reduce the error and position the load to the required angle. The Control Transmitter (CX) is isolated from the load and can be turned freely at all times, with the load accurately following its movement.

However, if no damping is provided, as before, the inertia of the load in a simple servo system will give a tendency to 'overshoot' with consequent effect on response time and stability.

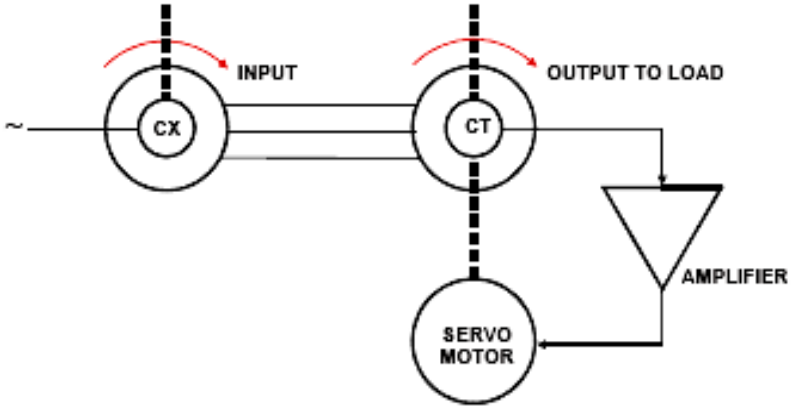


Figure 244 AC Remote Position Servo System

3.2.3 Control Response

Damping

Simple servo systems have a tendency to overshoot the required position due to the inertia of the load and the effect of the I/P suddenly causing a large error signal (step I/P). As the O/P shaft overshoots the required position a reverse error signal is applied. As the O/P shaft returns, it will again overshoot the required position. Neglecting friction, these oscillations will continue unless some form of 'damping' is applied to improve the control response. There are two basic types of input function:

- Step input function (change of input angle).
- Ramp input function (rate of change of input angle, or input shaft velocity).

The requirement is to reduce the servo motor speed before it reaches the required position. Damping is introduced to reduce this hunting to a minimum. Damping acts against the error signal to bring the system to rest sooner.

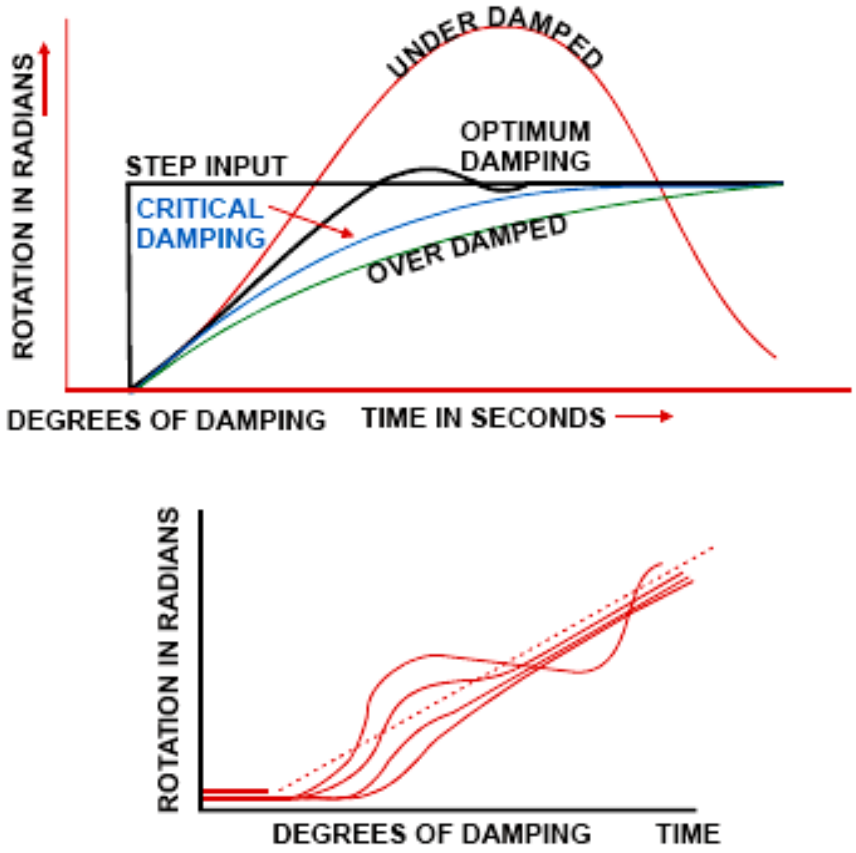


Figure 245 Degrees of Damping

System Damping

The simplest form of damping is the use of external friction. The two basic types are coulomb friction and viscous friction.

Coulomb Friction

This is a constant force independent of speed, eg, two plates rubbing together. In practice coulomb friction is not used as it causes an error. Coulomb friction is always present but is kept to a minimum.

Viscous Friction

This is a force proportional to velocity, eg, copper disc rotating between the poles of a permanent magnet. Removes the positional error of coulomb friction and number of oscillations remain the same independent of the size of the step I/P.

The disadvantage of mechanical friction is that it generates heat, wastes power and reduces accuracy of the system.

Angular Lag

If the input angle of a system is varied with time, then the output is required to follow this input. The output angle will however lag the input angle, the error being referred to as ANGULAR LAG. This must be so because the system is “error actuated”, ie, there must be an error signal in order to drive the servo motor. System damping in general is dependent upon:

- Applied motor torque.
- Inertia of the system and load.
- Viscous friction of the system.

System damping is required to reduce servo motor speed before it reaches the required position. It is always applied against the error signal and aids in limiting overshoot, improving stability and reducing response time.

Instead of mechanical friction, electrical friction may be used. Electrical damping is applied, as in friction damping, against the error signal. This is done by generating a signal proportional to the speed of the O/P shaft and applying it against the error signal. This is called Velocity Feedback Damping or Negative Feedback.

Velocity Feedback Damping

Velocity feedback is provided by including a tacho generator in the system. The tacho generator output signal is added algebraically to the control signal to the amplifier. The phase is arranged to oppose the control signal (negative feedback) when the system is approaching line up. When the system overshoots the line up position, the generator signal phase remains the same, but the control signal phase now inverts. The tacho generator signal and control signal (in phase) are summated and amplified and will provide a strong braking effect on the motor and so reduce any tendency to oscillate.

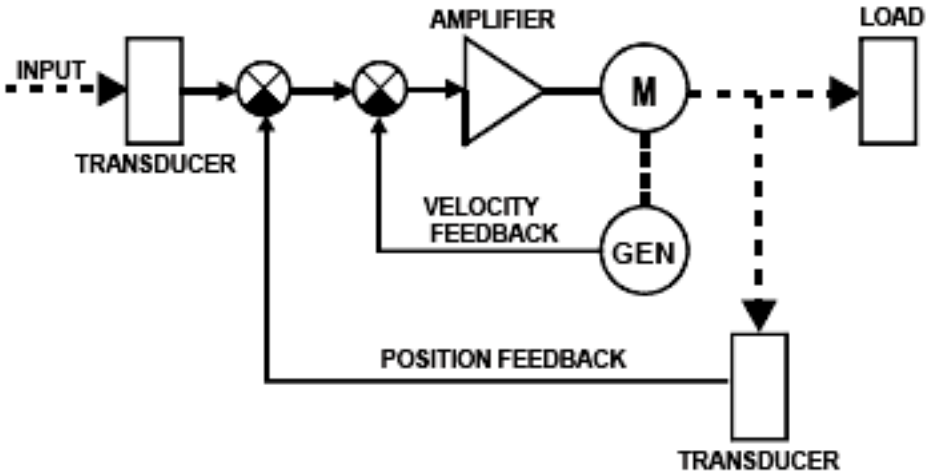


Figure 246 Velocity Feedback Damping

The disadvantages of velocity feedback damping are that the:

- Input to the amplifier is reduced when the system is following a moving input.
- Tacho generator adds to the viscous friction of the system.
- To overcome these problems and their effect on angular or velocity lag, transient velocity feedback damping may be used.

Transient Velocity Feedback Damping

During steady state conditions, the constant output voltage from the tacho generator, charges up C, this then blocks further signals to the summation point. During transient conditions, the charge on C must change, the current required to charge or discharge C being algebraically added the control signal at the summation point. Thus, feedback is only available from the tacho generator during changes in servo motor velocity and the input to the amplifier is not reduced when the servo motor is rotating at a constant velocity.

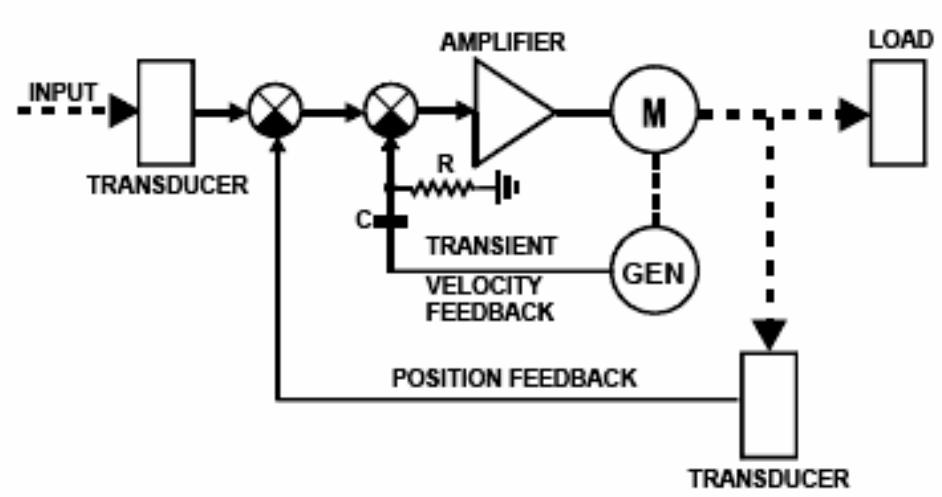


Figure 247 Transient Velocity Feedback Damping

Position servo Damping summary

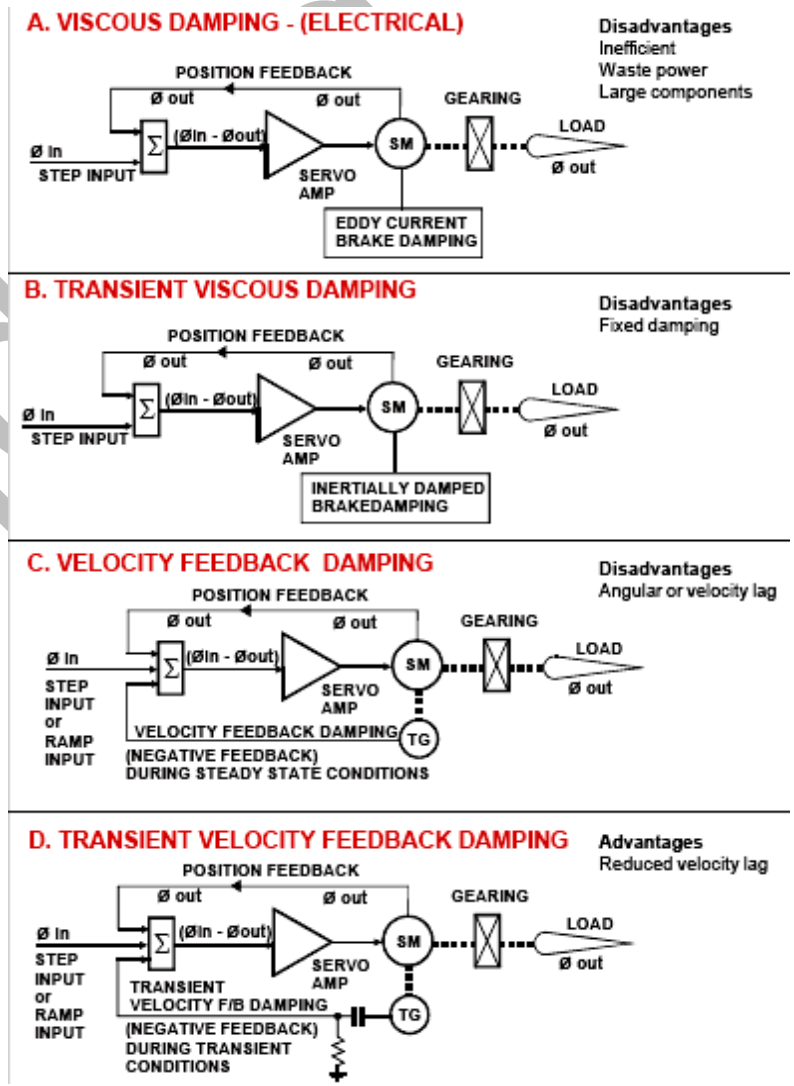


Figure 248 Position Servos Damping Summary

3.3 AC SERVO MOTORS

The servo motor must have good starting torque, be easily reversed and stop quickly when required.

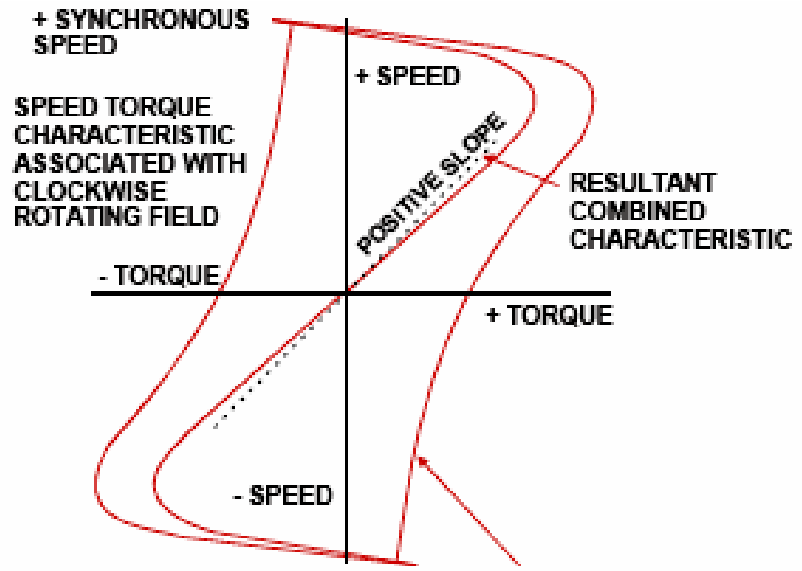


Figure 249 Conventional 2 Phase Induction Motor

A conventional two phase induction motor will continue rotating on a single phase, due to its 'positive' characteristic makes it unsuitable as a servo motor.

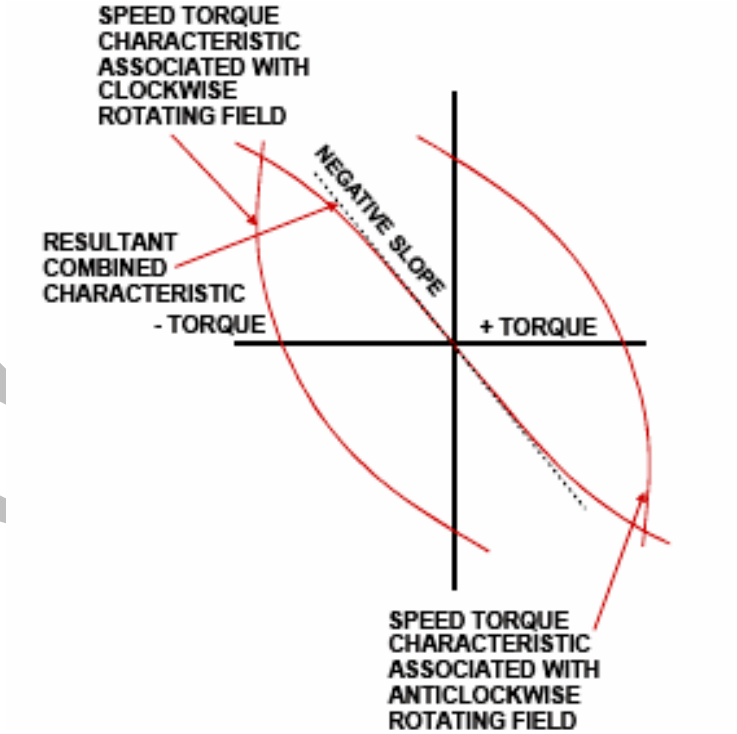


Figure 250 2 Phase Servo Motor Combined Characteristic

By the choice of materials and construction of the rotor (relatively high resistance) a motor can be designed to have a 'negative' combined characteristic. This servo motor characteristic provides a form of damping which is similar to viscous friction. When the control voltage is reduced, or becomes zero, a positive speed will produce a negative (braking) torque.

Construction

The servo motor shown in figure 251, has a two phase winding accommodated in slots in the stator which is comprised of iron laminations cemented together. The squirrel cage rotor slots are 'skewed' to minimise slot effect (or cogging), which is the reluctance of the rotor to move from one of a number of positions where the slots in the rotor and stator are aligned. Skewing the rotor slots ensures a smoother torque output. The squirrel cage rotor consists of iron laminations with the conductors and end rings cast in position. Aluminium conductors are generally used to ensure that the required speed/torque characteristic for the servo motor, which differs significantly from a conventional two phase ac motor, are obtained. Servo motors, in general, are subject to standardisation of frame sizes.

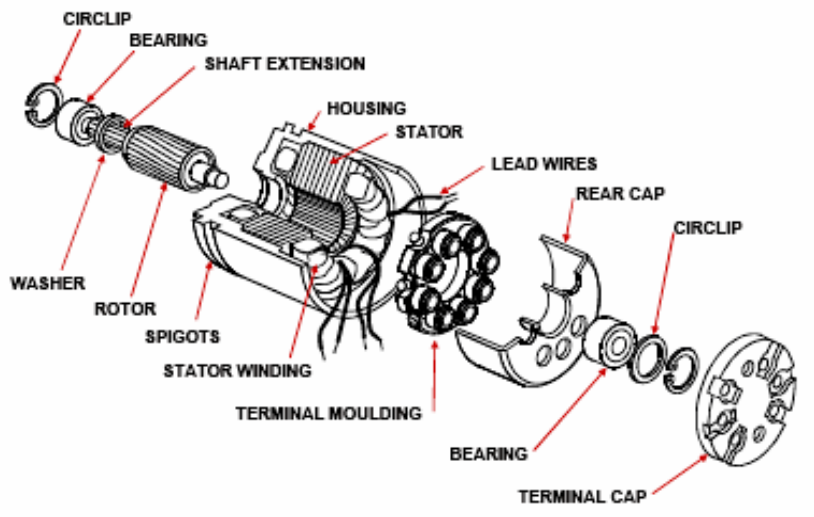


Figure 251 AC Servo Motor Construction

The two phase ac induction servo motor shown in figure 252, has an inertial wheel assembly attached to the rotor shaft. The inertial wheel provides transient viscous or velocity damping by mechanical means, for control of the servo motor response. One of the disadvantages of this method is that the degree of damping is pre-set. At uniform speeds the inertia wheel rotates with the shaft. When accelerating, there will be relative movement between the wheel and the motor shaft, the motor speed will lag behind the control voltage. When the system is approaching the line up position after misalignment there will be a large overshoot initially, but the oscillations will die away quickly.

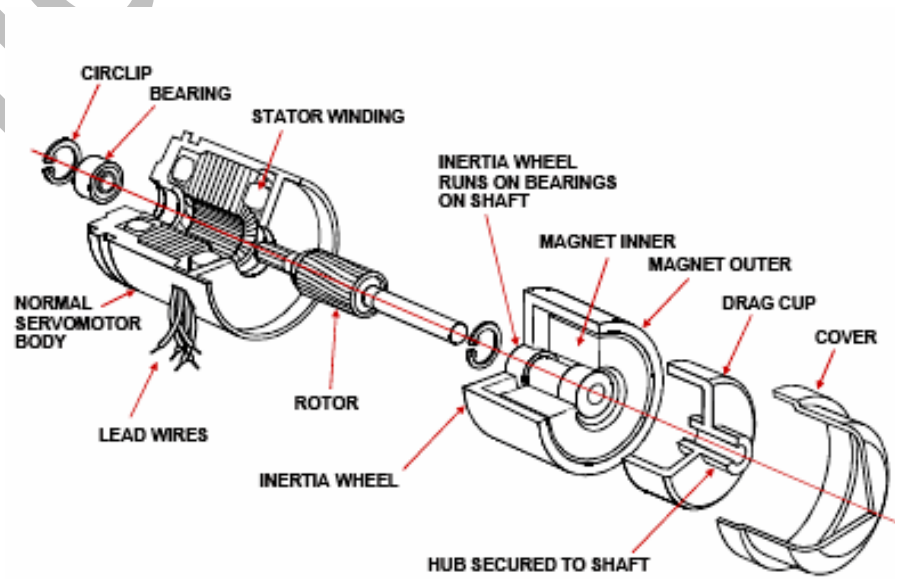


Figure 252 Inertially Damped Servo Motor

Tacho generators are used to provide an ac voltage which is proportional to servo motor shaft speed. Typically the output voltage can be from 0.6 to 3 volts per 1000 rpm, depending upon its construction. The tacho stator has two windings at 90°; one winding is the reference winding, which has a constant ac voltage applied to it. The other winding is the output winding, in which is developed on ac voltage proportional to speed. The rotor consists of a thin hollow aluminium or copper cylinder. Ideally there is no magnetic coupling between rotor and stator when the rotor is stationary.

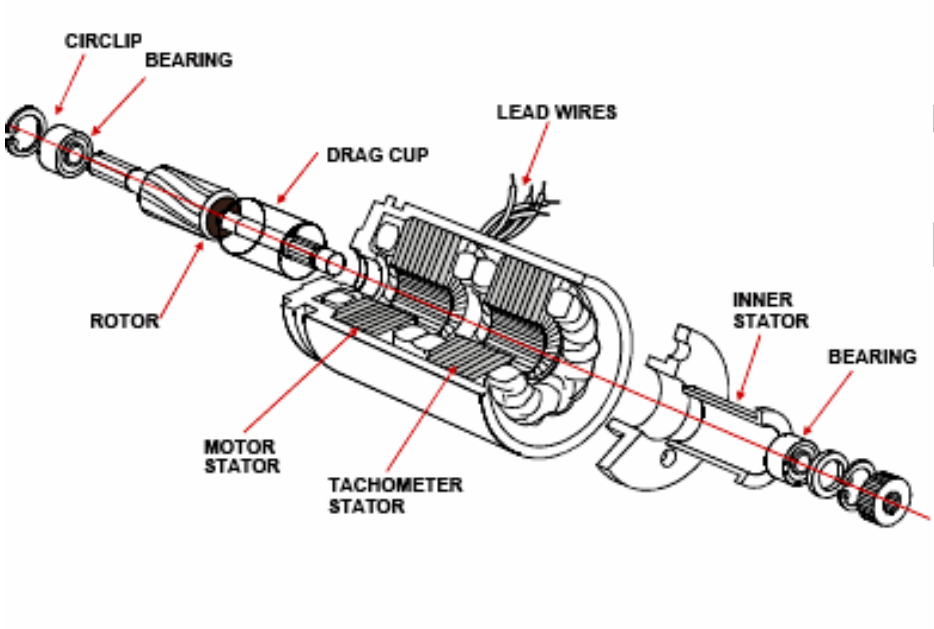


Figure 253 Servo Motor - Tacho Generator Assembly

With the Rotor Stationary

The flux patch from the energised reference winding will cut the cylindrical rotor, inducing an emf and circulating eddy currents. There will, however, be only minimum coupling with the output winding.

Output voltage will be minimum (zero).

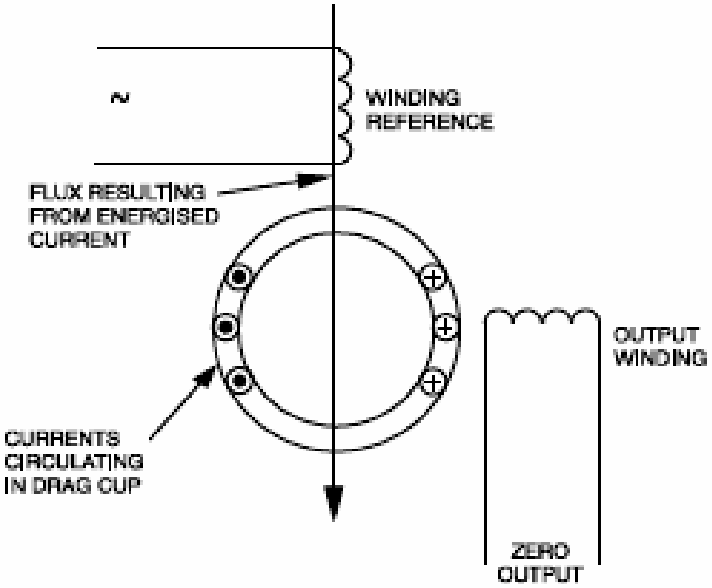


Figure 254 Tacho Generator – Operation

With the Rotor Rotating

The induced eddy currents will provide a coupling between reference winding and output winding, via the cylindrical rotor. A voltage will be induced in the output winding.

Output Voltage Amplitude

Will be independent upon the degree of coupling (ie, rotor speed).

Output Voltage Phase

Will be dependent upon the direction of rotation.

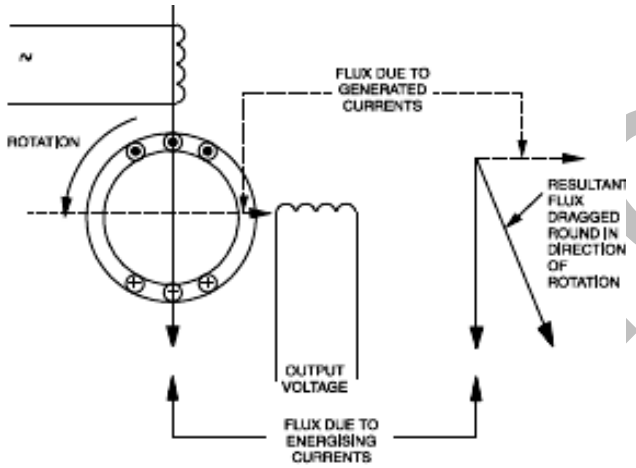


Figure 255 Tacho Generator – Operation

The accuracy of pointer displays could be greatly improved by gearing up the synchros, while employing a direct system to prevent ambiguity.

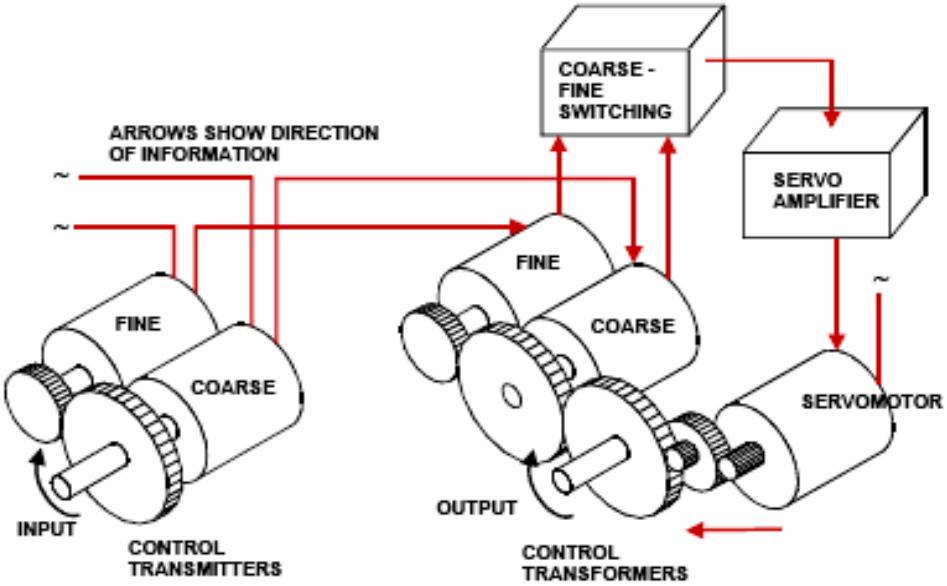


Figure 256 Coarse Fine Servo Systems

When the system is misaligned by a large amount, the ‘coarse’ synchro control signal is fed into the servo amplifier. When the servo motor has driven the output shaft sufficiently near to line up, the ‘fine’ control signal is switched to the amplifier. Thereafter, the system is controlled by the fine synchros.

3.3.2 Rate Servos

If a motor is required to drive at speed proportional to some incoming voltage, a rate, or velocity, servo may be used. The basic system is shown although in practice many refinements may be added.

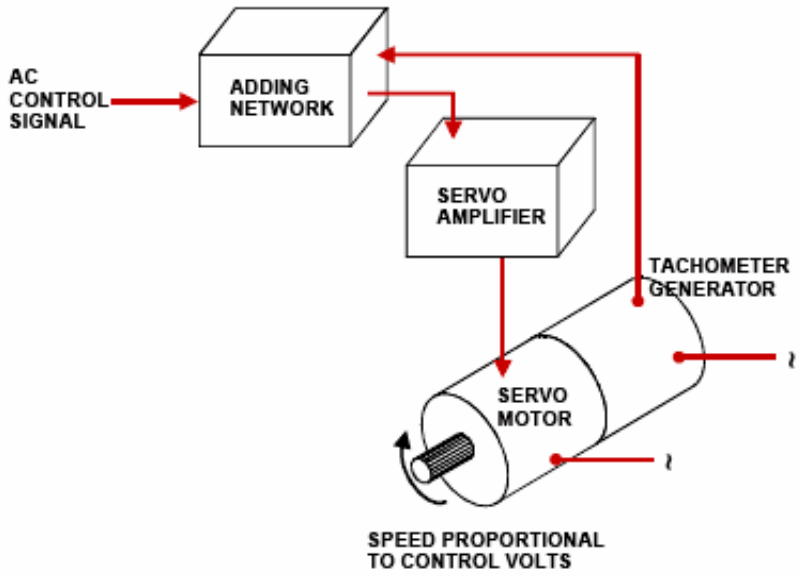


Figure 257 Rate Servos

The speed is measured by the tachometer generator which develops an equivalent voltage.

This is subtracted from the incoming voltage and the difference suitably amplified, drives the servo motor. Theoretically, the speed is almost independent of friction and load torque. In practice, the accuracy is dependent on linearity and freedom from temperature effects of the generator and the gain of the amplifier. It is usual, in rate servos, to employ generators with temperature compensation. To improve the speed regulation, the gain can be increased. The maximum gain setting is limited, as the system can easily become unstable with severe speed or positional oscillations.

SHORT TERM signals are applied to tacho generator field excitation supply. Only applicable above servo motor speeds as determined by THRESHOLD DETECTOR.

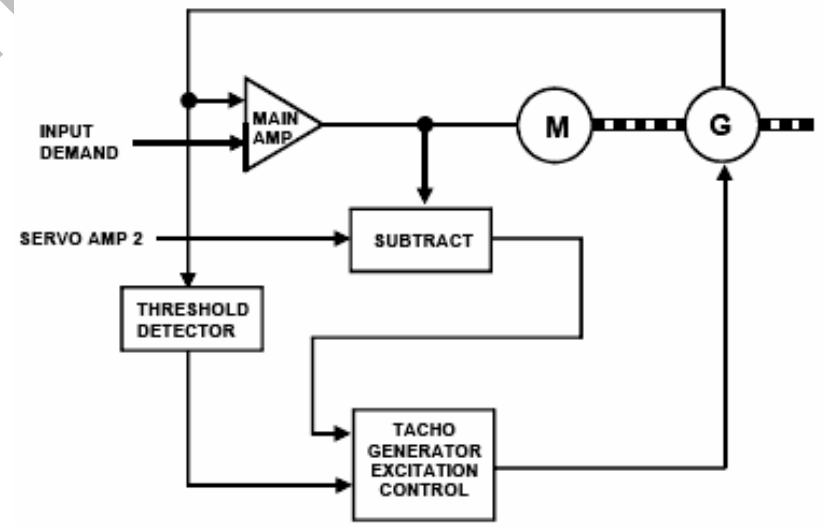


Figure 258 Rate Servos

Integrating Servos

In control systems it is often necessary to obtain a voltage or shaft movement representing the integral of some input function.

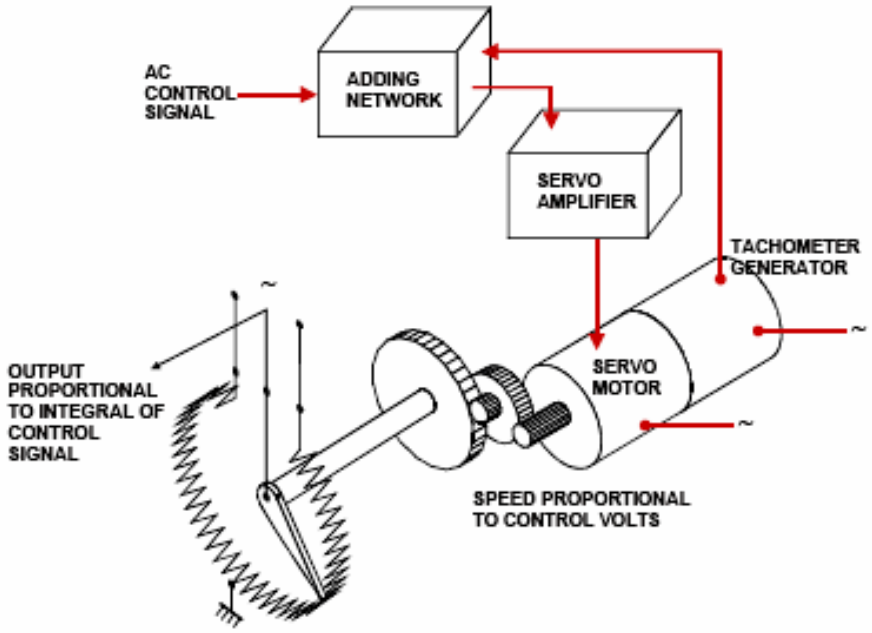


Figure 259 Integrating Servos

An integrating servo consists of a rate servo, in which some form of output is taken from load shaft, or from a transducer geared to it. In a perfect rate servo the number of revolutions of the shaft in a given time represents the integral of the input voltage level. Special attention must be paid to the linearity, residual voltages and degree of temperature compensation of the tachometer generator intended for use in an integrating servo. Integrating servos have an important field of application in correcting small errors which would lead to cumulative effects if allowed to persist. For example, if the automatic pilot in an aircraft fails to set the aircraft in the correct altitude, the aircraft may gradually gain or lose height. If the errors in height are integrated a signal will accumulate which can be used to correct the trim of the aircraft.

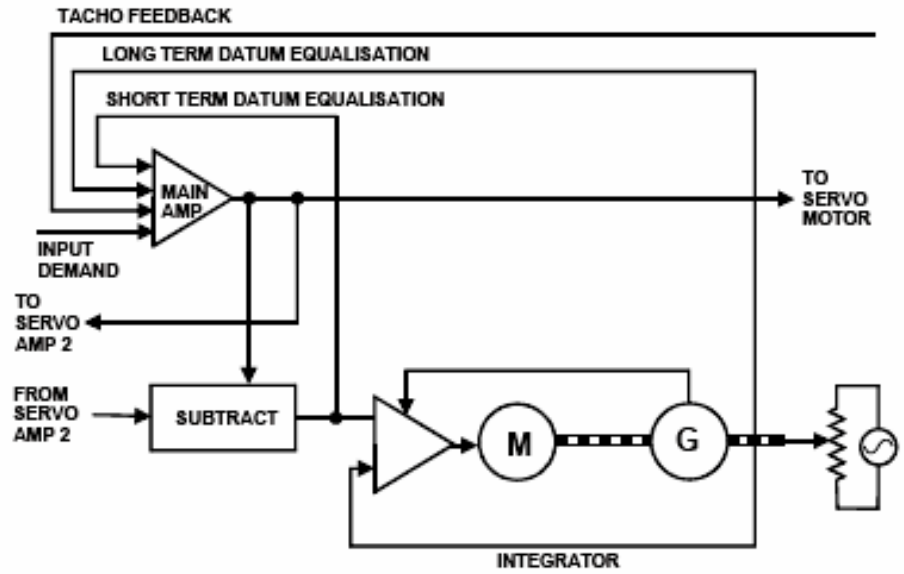


Figure 260 Integrating Servos

3.4 REMOTE INDICATING SYNCHRONOUS SYSTEMS

Remote indicating system consists of two distinct items, a transmitter and an indicator. When flying was in its infancy, the aircraft were quite small which enabled the pilot to see all around the aircraft. The pilot could see all the control surfaces just by looking from the cockpit, looking left and right he could see the ailerons, and by looking over his shoulder he could see the rudder and the elevators. The same could be said of engine parameters, many of the instruments then available operated on mechanical principles, for example engine speed, oil pressure and temperature, and displayed to the pilot in the cockpit by direct reading instruments. With the advent of larger and multi engine aircraft the pilot could no longer see the control surfaces from the cockpit, and mechanically operated instruments became severely limited in application. Therefore, an improved means of providing reliable and accurate indication of engine parameters and control surface position to the pilot became a requirement. This requirement was provided by the use of electrically remote reading indicating systems in which a sensor detects changes in a measured quantity or position and transmits this information electrically to a remote indicator in the cockpit for the pilot to view. The system introduced was a synchronous system consisting of a transmitter at the source (the medium to be measured) and transmitted electrically and synchronously to the indicator remote from the transmitter and the medium being measured, for example surface positions or engine parameters. There are several systems of this type used on aircraft and require either DC or AC power for operation.

3.5 DC SYNCRO SYSTEMS

The earliest of these systems is the Desynn system and requires DC direct current to power the system, which could be either 12 or 24 volts DC depending on the design requirement. There are three types of "Desynn" transmitters:

- Toroidal: resistor: rotary motion for the indication of position.
- Micro Desynn: linear motion for the indication of pressure.
- Slab Desynn: also used for pressure measurement.

Advantages of the Desynn system include:

- It removed the need for mechanical coupling between the medium to be measured and the indication.
- Hazardous fluids, pressures, and surface positions could be measured at source; also the system is not subject to small variations in the power supply.

Disadvantages:

- Friction caused by the surface resistance of wiper arms, which in turn causes wear on the Toroidal resistor, dirt or dust around the Toroidal resistor, which in turn causes a change in the resistance value.
- The Toroidal resistor transmitter is the basic system; the others are derivative of it.

Shown are the three Desynn systems diagrams:

- Basic- Desynn.
- Micro- Desynn.
- Slab- Desynn.

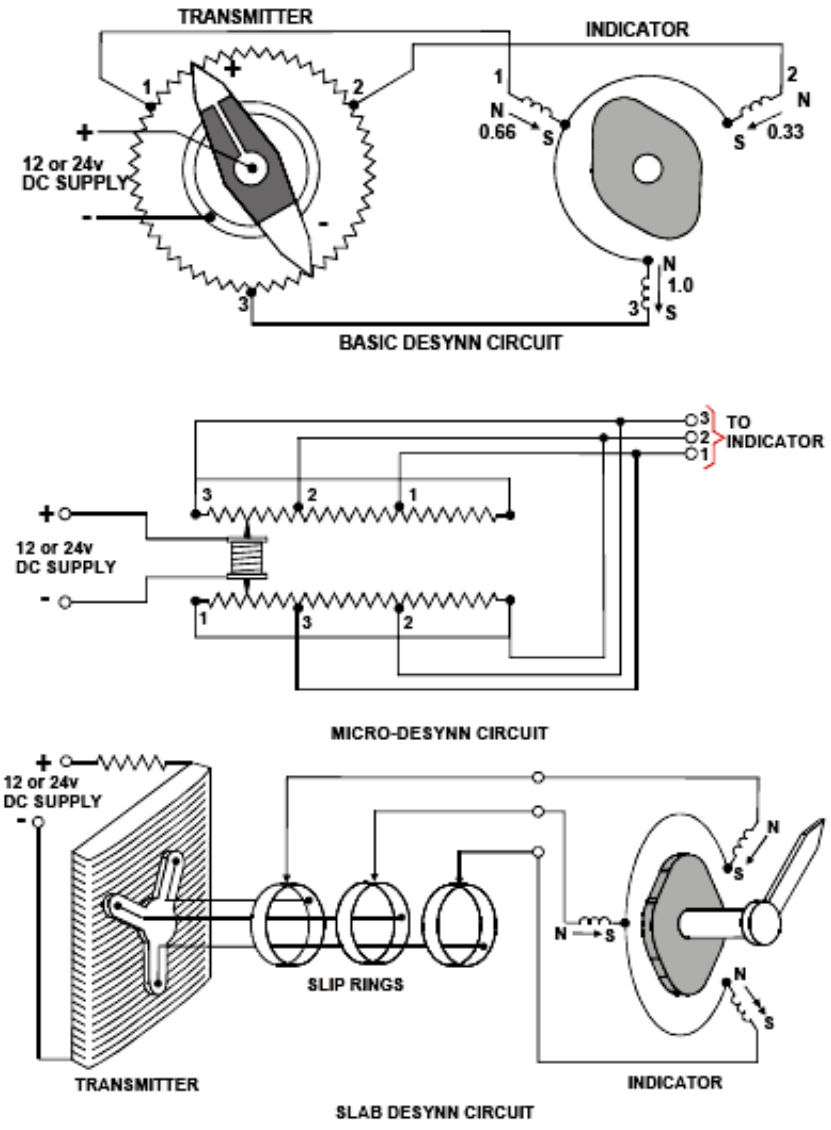


Figure 261 Desynn Systems

3.5.1 Desynn Systems

Testing Standard (Desynn)

Testing of Desynn indicators and transmitters is carried out in two separate parts.

Indicators are checked against a master potentiometer, which is calibrated in mechanical degrees. Transmitters are checked against a master Desynn indicator, with the use of a jig or fixture to hold the transmitter. All testing must be carried out in accordance with the manufactures test specification.

Desynn Transmission Systems

The Desynn system of DC data transmission is a relatively simple low torque system which lends itself satisfactorily to such aircraft applications as the remote indication of flap, rudder and elevator position, oil and fuel pressures, etc. The simple pointer and scale arrangement is generally acceptable for such indications and since arrangements of this kind can be operated satisfactorily with very small applied torques the low torque characteristics of the Desynn system does not present any undue difficulties when the system is employed to given indications of this type. The accuracy of the Desynn system is of the order of ±2½%; such accuracy is not high by modern standards, but it is reasonably realistic for the applications which have been mentioned. The DESYNN system of DC data transmission is a relatively simple low torque system designed for continuous rating on the 28v DC supply. (The name DESYNN implies a DC synchronous system.) In its simplest form, the transmitter consists of an endless resistance would on a ring former (toroidal resistor), having three fixed tappings at equidistant points which connect to the indicator (receiver). The control spindle carries two wiper arms which bear on the torodial resistor at points diametrically opposite. The arms are insulated from each other at the spindle and are connected to the supply via sliding contacts. The indicator consists of a 2 pole permanent magnet rotor pivoted to rotate freely within a three coil soft iron stator, with the pointer attached directly to the rotor. The stator coils are each connected to a tapping on the transmitter torodial resistor.

At the transmitter, the position of the wiper spindle is controlled by a linkage from the appropriate aircraft component, the position of which is to be indicated. When the supply is switched on, the voltages at the 3 equi-spaced tappings will be dependent upon the position of the wiper contacts. These voltages cause a magnetic field to be set up by the stator coils in the indicator and the permanent magnet rotor will align itself with this field and indicate accordingly. Any subsequent displacement of the transmitter spindle will change the stator voltages thus causing the indicator stator field to move in synchronism. The indicator rotor will therefore follow the stator field movement to continuously indicate the position of the transmitter spindle. A weak circular magnet, which does not affect the normal operation of the indicator, will return the pointer to an OFF SCALE position if the supply fails or is switched off.

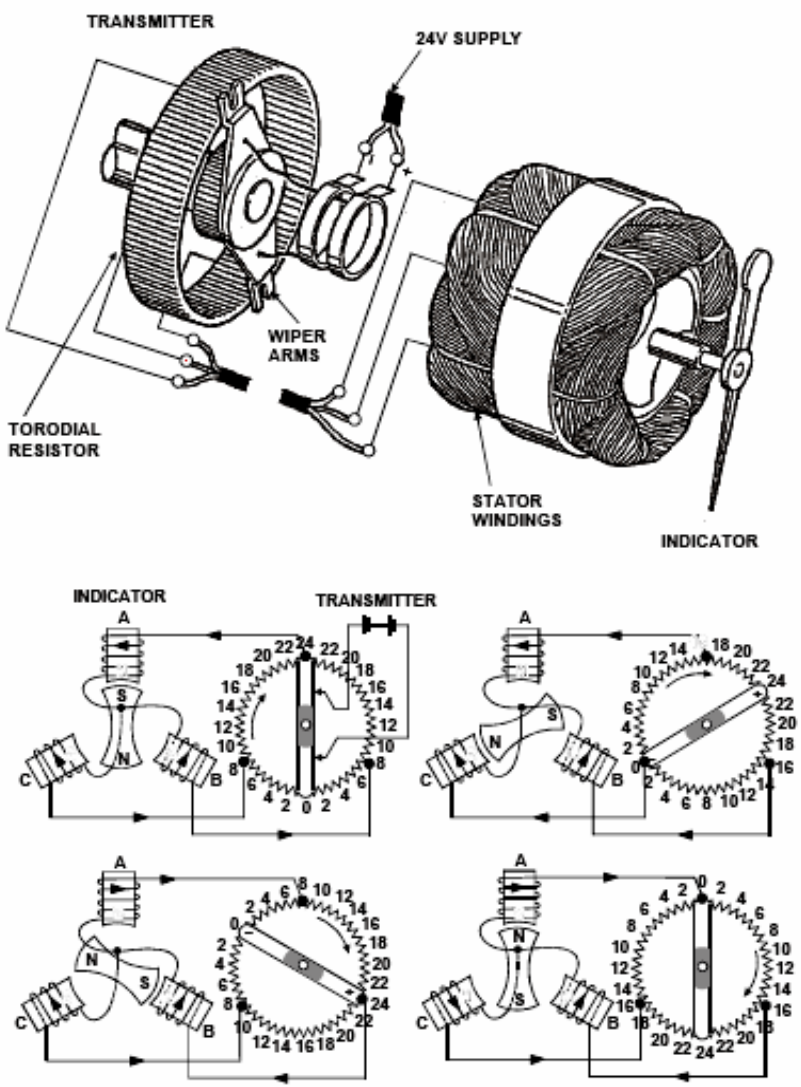


Figure 262 Desynn Transmission Systems

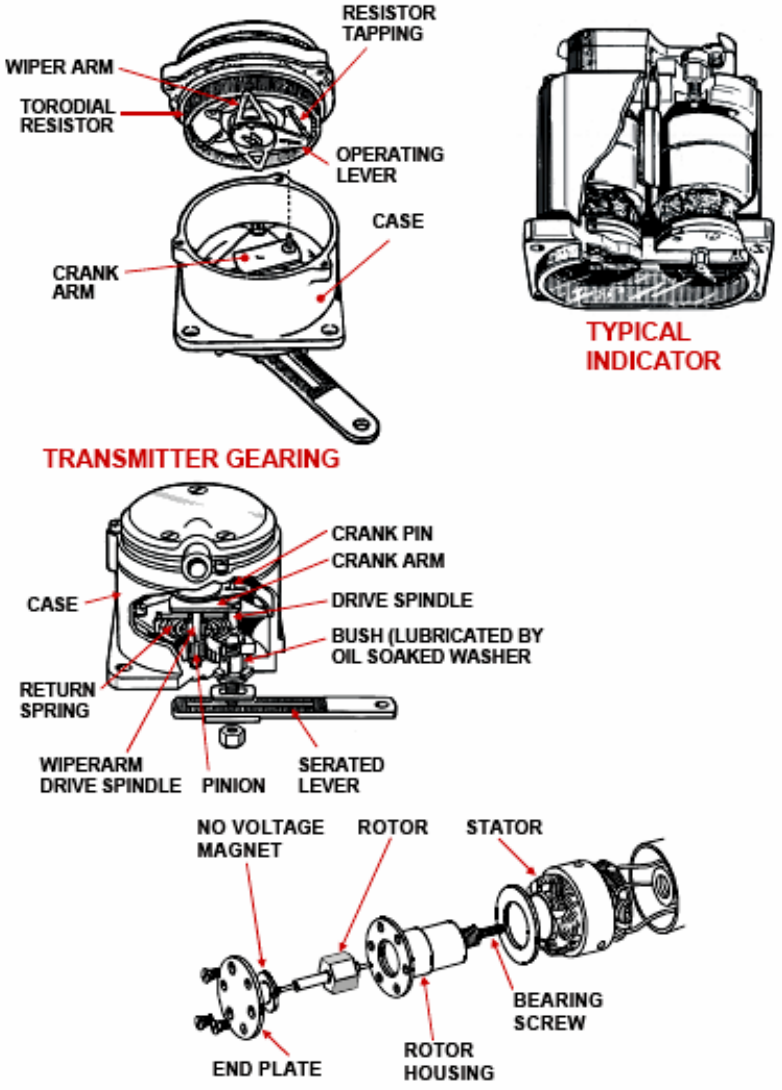


Figure 263 Desynn Transmission Systems

Typical Desynn Faults

SYMPTOM	CAUSE	REMEDY
Pointer “OFF SCALE”.	Power supply failure/”OFF” switch “ON”.	Check appropriate fuse.
Pointer displaced 180°. Rotation correct.	Power supply reversed.	Check transmitter for correct wiring.
Pointer rotation reversed.	2 connections crossed.	Check wiring connections between transmitter and indicator.
Pointer displaced 120° or 240° clockwise. Rotation correct.	3 connections crossed.	Check wiring connections between transmitter and indicator.
Pointer displaced and moves in 180° steps.	Open circuit or short circuit.	Carry out insulation resistance check and continuity checks.

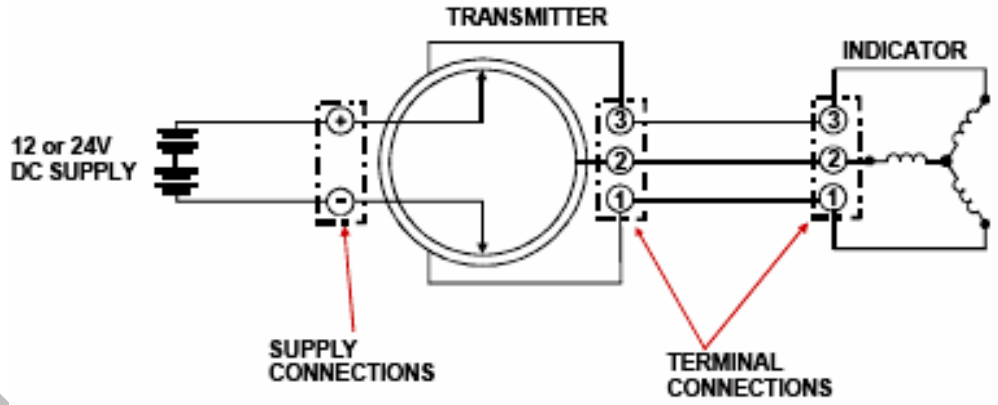


Figure 264 Desynn Transmission Systems

Micro Desynn

In applications where the movement of a prime mover is small and linear, the use of a basic system transmitting element is strictly limited. The micro Desynn transmitter was therefore developed to permit the magnification of such small movements and to produce, by linear movement of contacts, the same electrical results as the complete rotation of the contact arms of the basic transmitter.

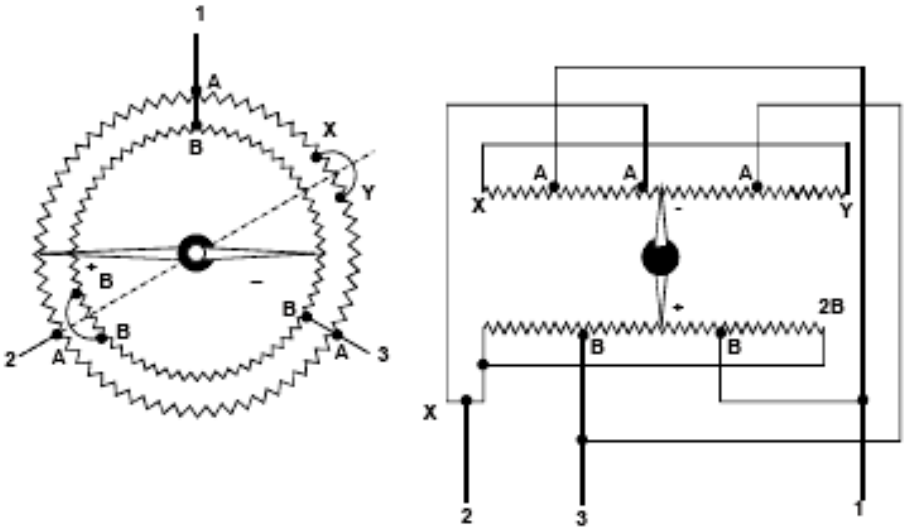


Figure 265 Micro Desynn Transmitter

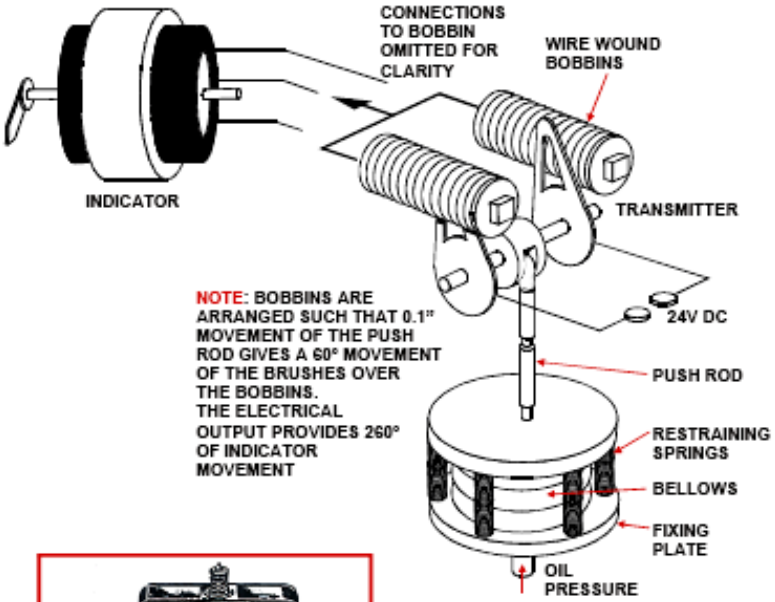


Figure 266 Desynn Transmission Systems – Micro Desynn

Each resistance bobbin is secured in place and accurately positioned so as to provide the necessary tapping points. The contact arms are mounted on a rocker shaft supported between the vertical parts of a U shaped bracket and movement of the transmitter's mechanical element is transmitted to the arms via a spring loaded operating pin and crank arm connected to the rocker shaft. Two beryllium copper hairsprings conduct current to the contact arms and also act together to return the rocker shaft and contact arms to their starting position.

Slab Desynn

The purpose of the Slab Desynn System is to provide more accurate indication than can be given with the standard Desynn system. The transmitter construction lessens the friction between the resistor and pick off brushes and provides a sine wave output to the indicator, which has high resistance coils. The Slab Desynn system is therefore not interchangeable with standard Desynn transmitters and indicators, nor can it be tested using the standard test set. The Slab Desynn transmitter assembly is supplied with 28 volts dc and consists of a resistor over which a contact plate assembly carrying three contacts 120º apart rotates on a pivot. The shaft of the contact plate assembly carries a pinion, hairspring and hairspring guard. The whole is supported in a frame assembly to which a brush holder with three brushes fitted.

The brushes make contact with slip rings on the contact plate assembly. Backlash between the sector and pinion is taken up by the hairspring, the free end of which is anchored to the Desynn frame assembly. A clamping ring holds the Desynn transmitter assembly in position so that the pinion meshes with the toothed sector.

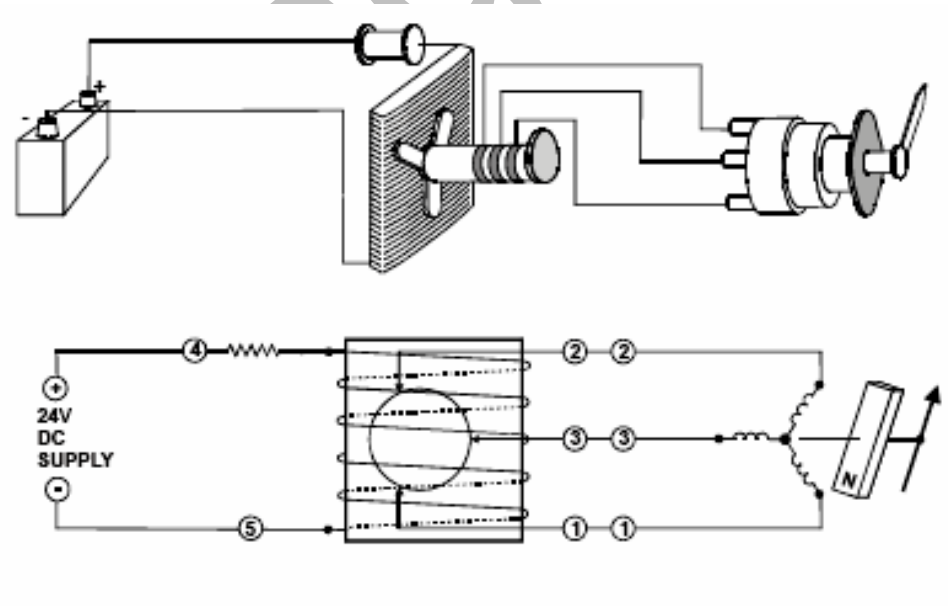


Figure 267 Slab Desynn Transmission

3.6 ALTERNATING - CURRENT "SYNCHRONOUS" SYSTEMS

These systems are generally referred to as synchro systems and operate in a similar way to the Desynn system, as a remote reading and indicating system. Synchro- systems are powered by AC power and operate on the transformer principle, the primary windings being the rotor and the secondary windings the stator.

Basically the system consists of a two pole single phase rotor, and a three phase stator for both the transmitter and the indicator. Using the transformer principle, when power is applied to the transmitter and receiver rotors, an equal and opposite voltage will be produced between the connecting lines of both stators. If both rotors are in the same angular positions, and then no current will flow in the stator coils, the synchros could be said to be at electrical (null).

When the rotor of the transmitter is rotated, an Electro motive force (emf) is set up between the rotor and the stator causing a unbalance between the stator coil voltages, this in turn will cause current to flow in the stator coils. The effect of this will produce a magnetic field in the stator of the indicator (receiver). This in turn will induce a torque on the rotor, which will then rotate to the same position as the transmitter. When the voltage and current unbalance become equal to the transmitter no further torque will be applied to the rotor of the indicator where once again the synchros will be at null.

The advantages of the A-C synchro, is the size. They are much smaller in comparison to the Desynn system, therefore, are easily installed into multiple instrument displays. They are much more accurate, and not prone to were and tear, consume far less power, and when in multiple combinations can be used to add, subtract, or multiply, for example computing air data computations for airspeed and altitude within an air data computer.

Disadvantages - none.

A-C SYNCHROS are the most common and widely used means of synchronous data transmission. Testing is carried out with a test fixture, where the stator is clamped stationary and the rotor is rotated through 360 degrees and compared to a master “Indicator” for accuracy.

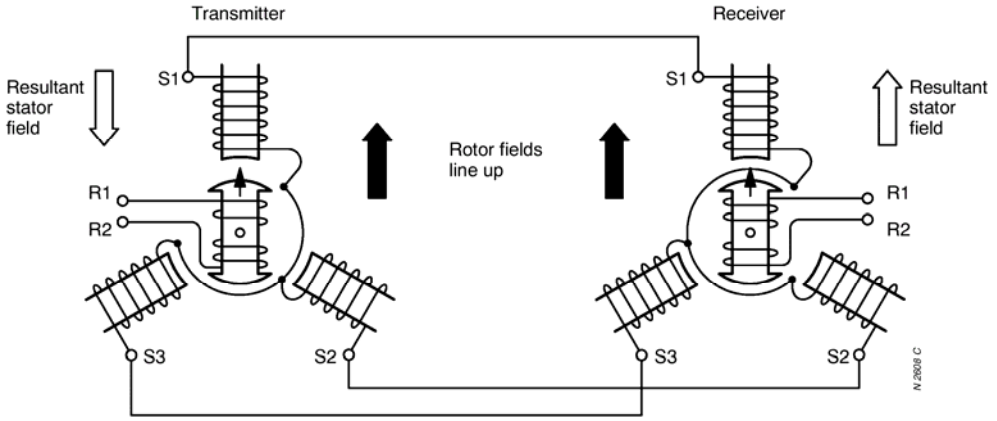


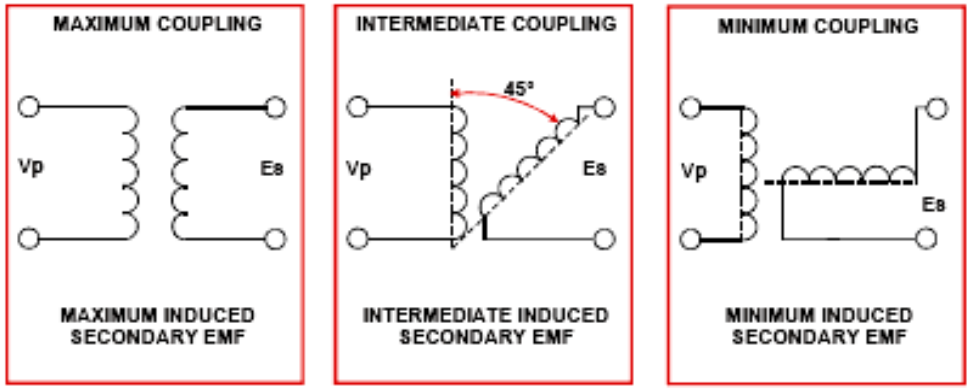
Figure 268 Basic AC Synchronous System

3.6.1 Principle of Operation

The components of an ac synchronous transmission system are an application of the principle of the transformer to an arrangement in which either the primary or secondary windings are wound on the rotor, which can be turned with respect to the fixed or stator windings. The amplitude of voltage induced into the secondary depends upon the relative angular position of the primary/secondary windings. These induced voltages must always be either IN PHASE or 180º OUT OF PHASE with each other and with the supply.

Transformer Principle

The amplitude of the emf induced in the secondary winding depends on the degree of 'magnetic coupling' between primary and secondary winding.



Es = Vp Cos θ° = Vp

Es = Vp Cos 45° = 0.707 Vp

Es = Vp Cos 90° = 0

Note: Assuming a turns ratio $\frac{Ns}{Np}$ of Unity

Figure 269 Transformer Principle

3.6.2 Torque Synchros

Torque synchros are particularly useful in pointer indication systems. Torque systems incorporating differentials may be used to add or subtract angular data. Apart from the necessary alternating supply, the torque system is completely self contained and needs no auxiliary equipment. The error in positioning is typically one degree when unloaded and can often be substantially reduced by gearing up the system. It must be remembered that the error in positioning increases with any increase in the torque being transmitted and the torque required to position the load has to be applied to the torque transmitter and loads the associated equipment, as there is no torque amplification.

Torque Synchro Construction

The stator body is made up of internally slotted laminations, in the slots of which are fitted three sets of windings S1, S2 and S3 spaced 120° apart. Despite similarity these must not be confused with normal three phase ac windings, where the voltages are equal in magnitude and 120° apart in phase relationship - in a synchro unit, the voltages are NOT equal in magnitude - as previously stated, the amplitude will depend upon the relative angle of the rotor with each stator winding.

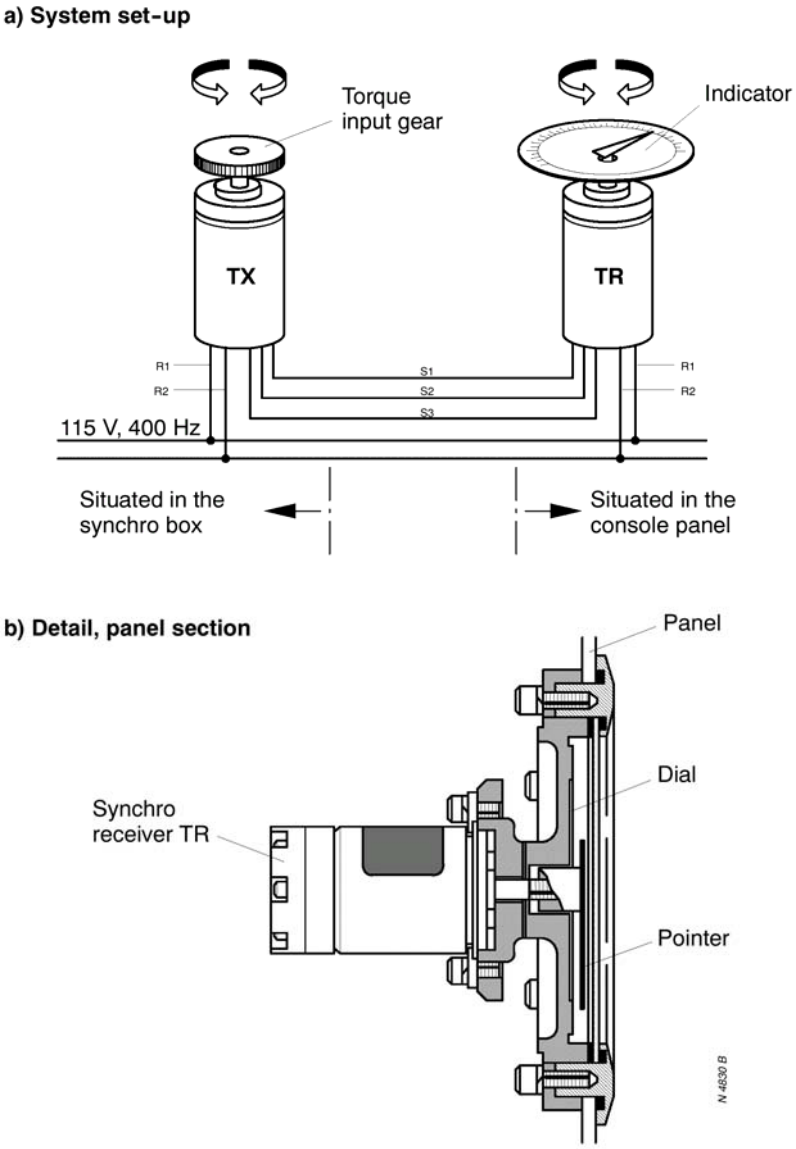


Figure 270 Torque Synchro Construction

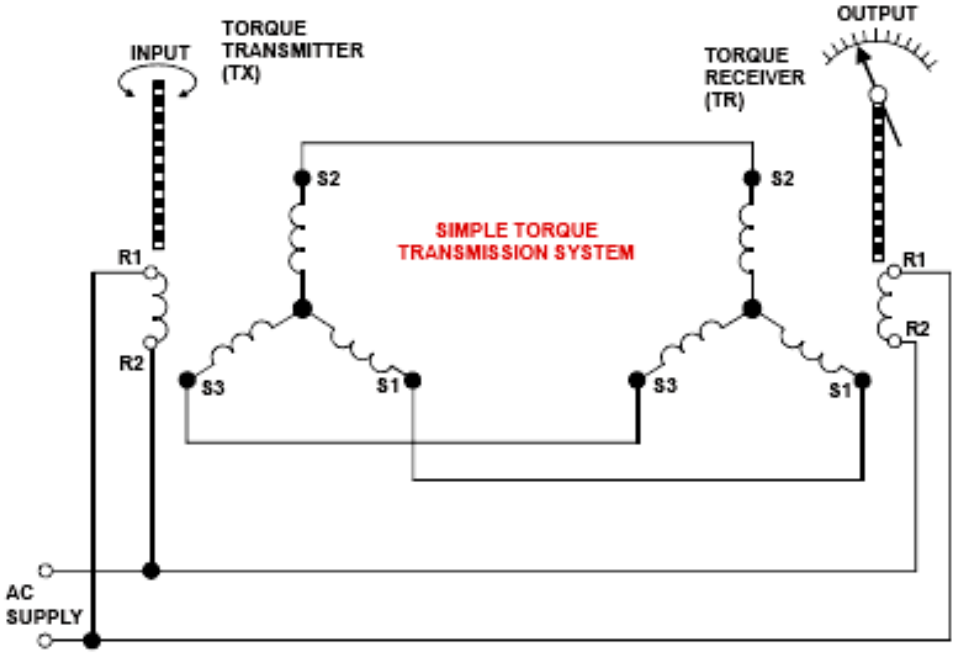


Figure 271 Torque Synchro Construction – Connection Diagram

Torque Synchro Faults

SYMPTOM	CAUSE	REMEDY
No movement of receiver	Power failure/OFF	Check appropriate fuse. Switch 'ON'
Receiver displaced 180° from transmitter Rotation correct	One pair of Rotor Connections reversed	Check rotors for correct wiring
Reverse rotation of receiver	2 Connections between stators reversed	Check stators for correct wiring
Receiver displaced 120° or 240° clockwise from transmitter Rotation correct	3 Connections between stators crossed	Check stators for correct wiring
Receiver displaced 0°, 60°, 120°, 180°, 240°, 300° and moves in 180° steps	Short circuit between 2 stator lines	Carry out insulation resistance check between stator lines
Receiver oscillates between 2 points approximately 75° apart	Open circuit on one stator line	Carry out continuity check on stator lines

Symbols Torque Syncro System

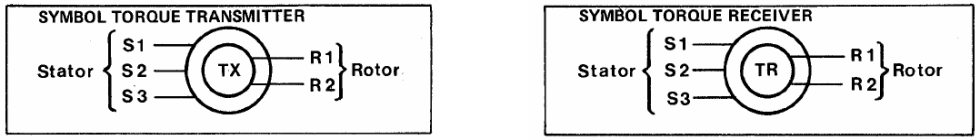


Figure 272 Symbol Torque Transmitter and Torque Receiver

Connections in Torque Synchro Systems

Only symmetrical connections between TX and TR and between TX, TDX and TR have been considered. Re-arrangement of the rotor and stator connections produces different results. The receiver rotor still moves synchronously with the transmitter rotor but it can do so from a different reference position or in the reverse direction.

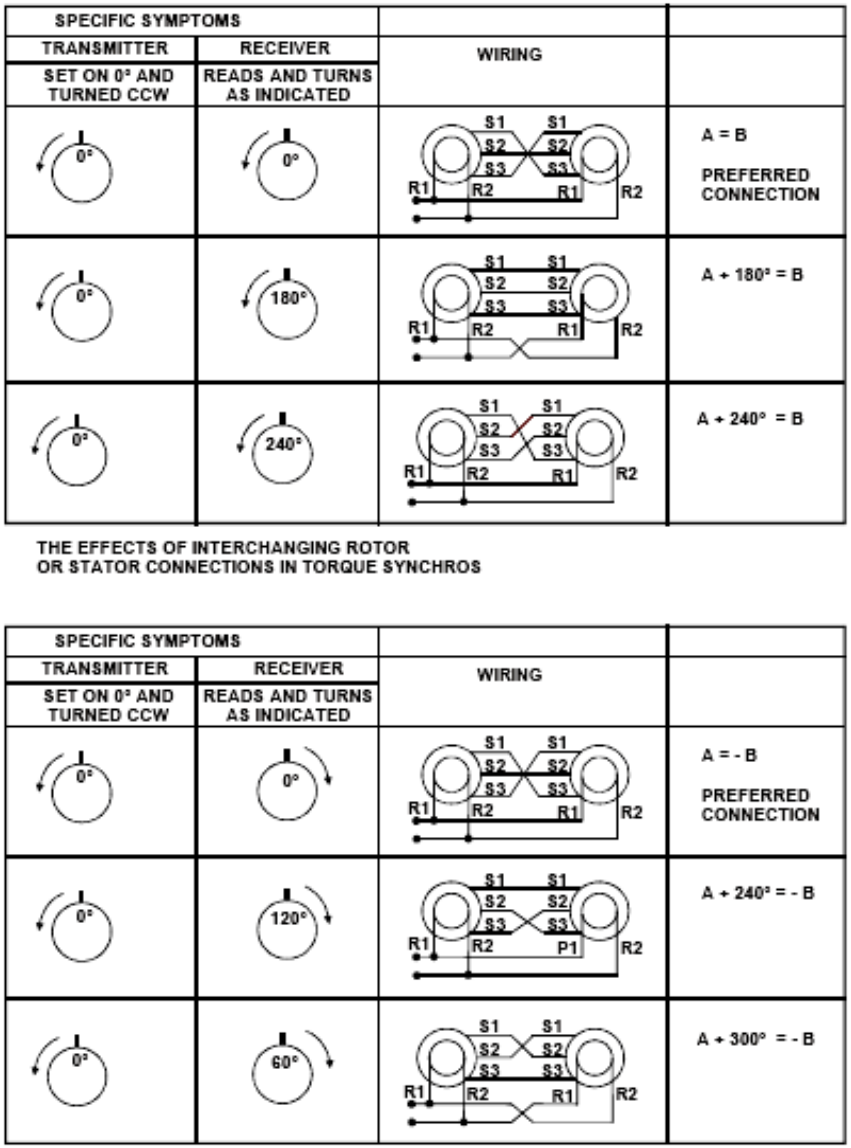


Figure 273 Connections in Torque Systems

The torque system, just mentioned, is used to transmit and indicate the angular position of a single input shaft. Under certain conditions it is necessary to transmit two angular positions, the synchro receiver indicating the sum or difference of the two angles. One method of applying a second angular input to a basic torque system is to rotate the transmitter stator (TX) through one angle and the rotor through another angle. There are two types of differential units - differential transmitters and differential receivers. The differential transmitter (TDX) accepts one electrical input and one mechanical input and produces one electrical output. The differential receiver (TDR) accepts two electrical inputs and produces one mechanical output see Figure 274.

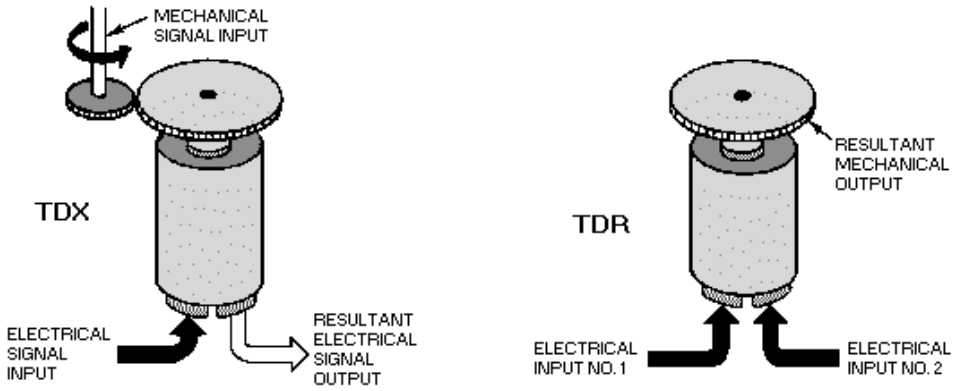


Figure 274 Differential Action in Torque Synchro System

In the torque differential transmitter, BOTH the rotor and stator windings consist of three Y-connected coils, as illustrated in figure 275. The stator is normally the primary, and receives its input signal from a synchro transmitter. The voltages appearing across the differential's rotor terminals (R1, R2, and R3) are determined by the magnetic field produced by the stator currents, the physical positioning of the rotor, and the step-up turns ratio between the stator and the rotor. The magnetic field, created by the stator currents, assumes an angle corresponding to that of the magnetic field in the transmitter supplying the signal. The position of the rotor controls the amount of magnetic coupling that takes place between the stator magnetic field and the rotor, and therefore, the amount of voltage induced into the rotor windings. If the rotor position changes in response to a mechanical input, then the voltages induced into its windings also change. Therefore, the output voltage of the TDX varies as a result of either a change in the input stator voltage or a change in the mechanical input to the rotor. This electrical output of the TDX may be either the SUM or the DIFFERENCE of the two inputs depending upon how the three units (the TX, the TDX, and the TR) are connected.

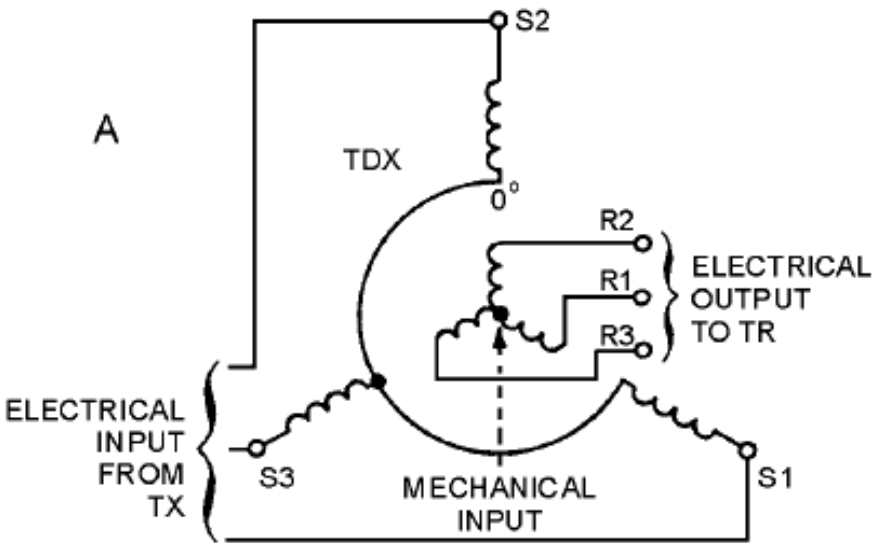


Figure 275 Torque Differential System TDX

The torque differential transmitter (TDX) and the torque differential receiver (TDR) are ELECTRICALLY IDENTICAL. The only difference in their construction is that the receiver (TDR) has a damper, which serves the same purposes as the damper in the TR — it prevents the rotor from oscillating. The real difference in the receiver lies in its application. It provides the mechanical output for a differential synchro system usually as the sum or difference of two electrical inputs from synchro transmitters. As in the case with the TDX, the TDR addition or subtraction function depends upon how the units in the system are connected. Basically, the torque differential receiver operates like electromagnets. In Figure 276, the rotor and stator of the torque differential receiver receive energizing currents from two torque transmitters. These currents produce two resultant magnetic fields, one in the rotor and the other in the stator. Each magnetic field assumes an angle corresponding to that of the magnetic field in the transmitter supplying the signal. It is the interaction of these two resultant magnetic fields that causes the rotor in the TDR to turn.

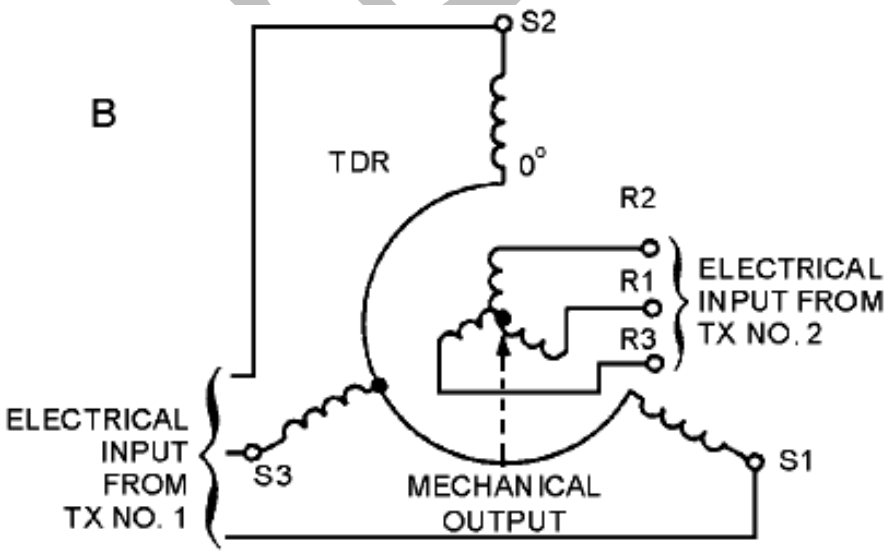


Figure 276 Torque Differential System TDR

Function Principle TDX

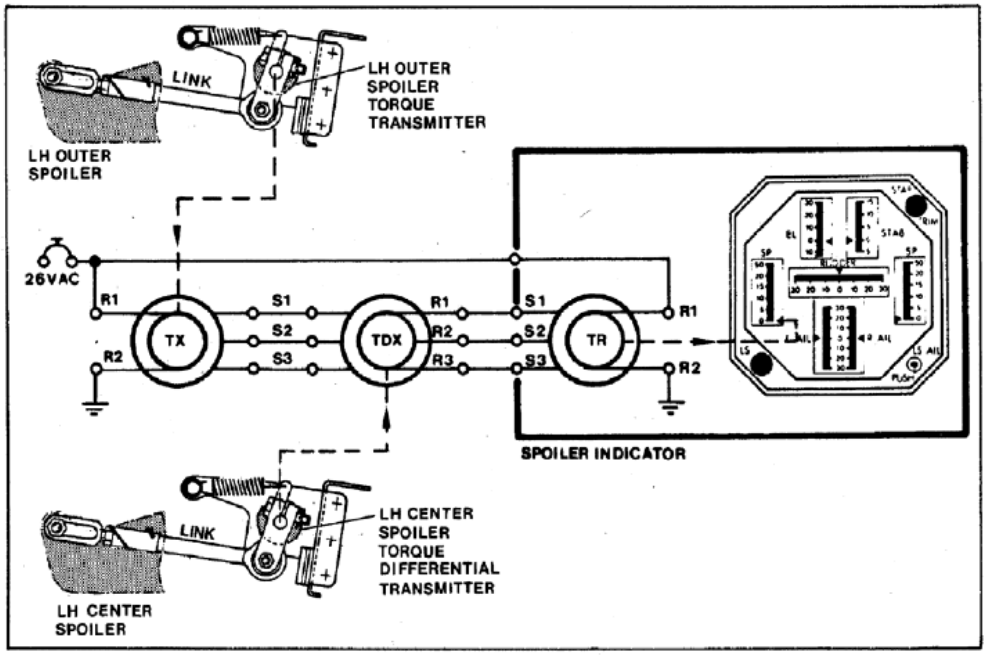


Figure 277 Example 1 Function TDX

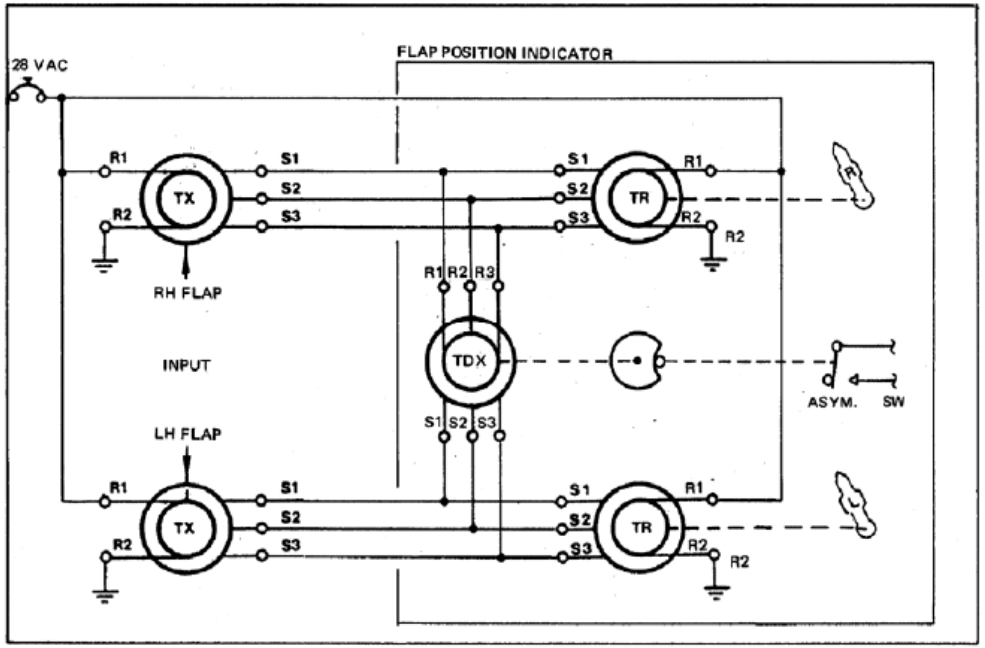


Figure 278 Example 2 Funktion TDX

3.7 CONTROL SYNCHROS

For many applications it may be better to transmit movements by means of control synchros used in conjunction with the servo systems. The accuracy with which data can be conveyed is closely related to the electrical errors of the synchros, usually a few minutes of arc. No appreciable mechanical restraint is imposed on the associated equipment at the driving end. In a system where a single transmitter operates a number of servo mechanisms, failure of one will not normally affect the others. Control synchros are primarily voltage devices and operate with relatively low temperature rises. No mechanical load is carried by them, so that the instruments may be smaller than in a comparable torque system.

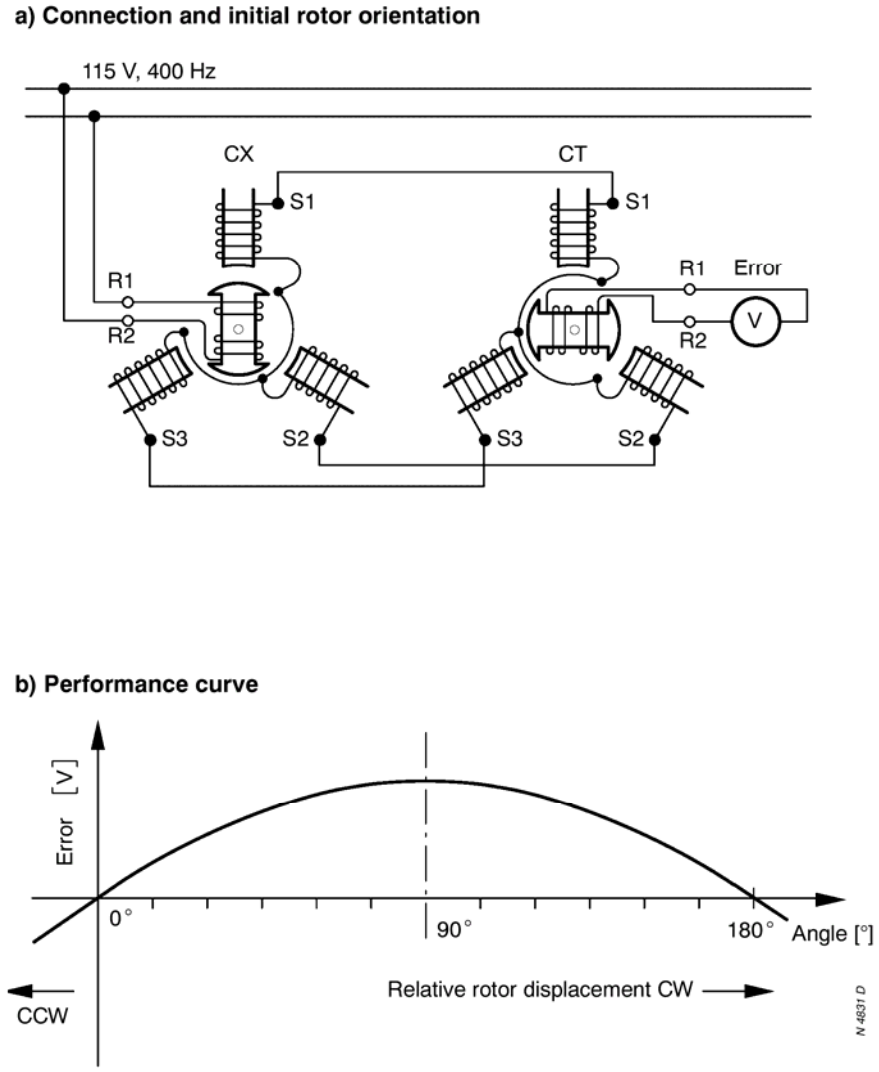


Figure 279 Typical Control Synchro

Indicating mechanisms now in production have large moments of inertia and require much greater torques to operate them than are available from Torque Transmission Systems. Comparatively large torqueses become available by the addition of a servo mechanism to the torque system. Control Transmission Systems are so called because the displacement signal from the transmitter is used to control a servo motor at the receiver. Any signal in the receiver rotor is amplified and the amplified output used to drive a servomotor. The servomotor is geared to the indicator mechanism and to the receiver rotor, which is turned until no signal is induced across it (the null position). There is a 90° difference between the relative rotor/stator positions of the transmitter and the control transformer when there is no servo signal (the transformer rotor is at its null point). Current still flows in the stator windings of control synchros even when at rest.

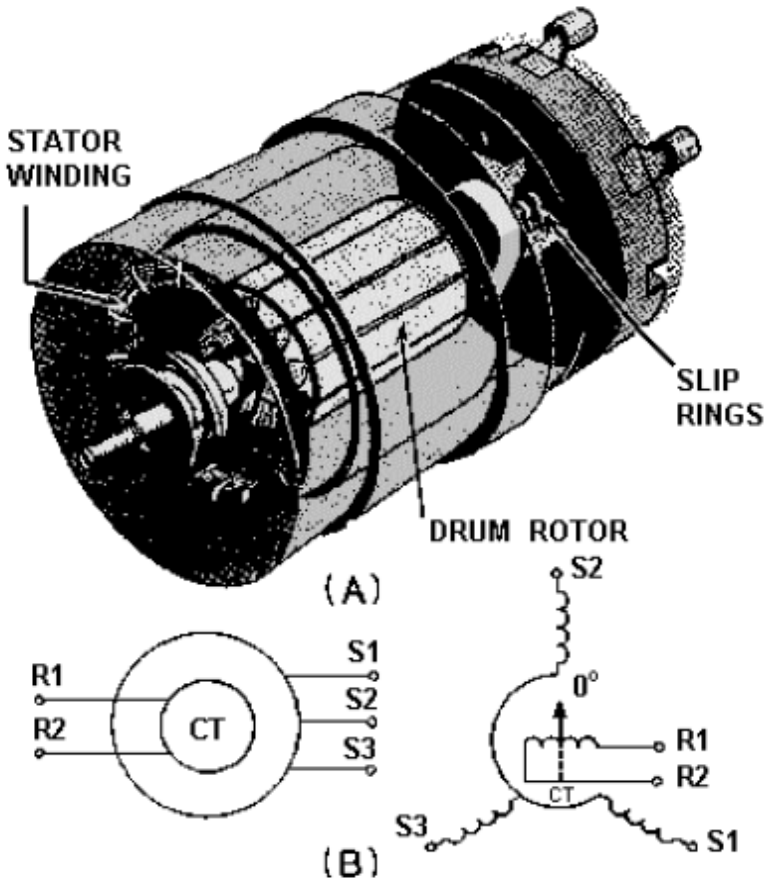


Figure 280 Control Transmission System (a) Phantom View (b) Schematic View

In the same way that differential synchros can be used in torque transmission systems so they can be used in control synchro systems. The sum or difference of two angles will be produced as an error signal.

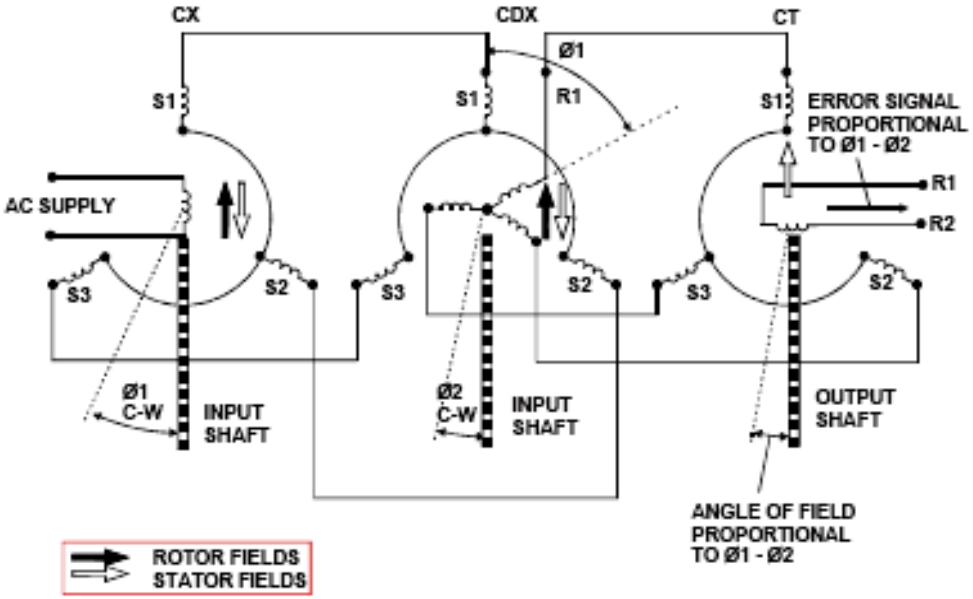


Figure 281 Control Differential Synchro

3.8 RESOLVER SYNCHRO

The general application of analogue computers to Air Data and Flight Path Computation has made it more desirable to have position information from a transducer transmitted in CARTESIAN co-ordinates, rather than POLAR co-ordinates.

A **polar** co-ordinate; indicates the bearing of a point P. Such information can be transmitted by standard synchros (ie, voltage proportional to shaft angle q).

Cartesian co-ordinates; indicate the same bearing by the distances X and Y measured along two axes; OX and OY at right angles to one another. Electrically, this can be achieved in the form of two signal voltages, whose amplitude is relative to the sine and cosine of the angle q where:

- X = OP Sin q.
- Y = OP Cos q.

Such information can be transmitted by a **resolver** synchro and converted back into a **polar** co-ordinate by a **resolver** synchro.

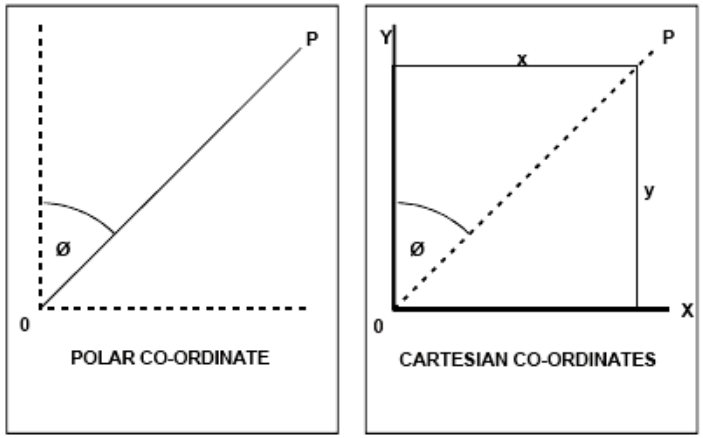


Figure 282 Trigonometric Function of a Resolver Synchro

The system is basically the same as the Synchro Control Transmission System, except for the electrical construction of the Synchro units. The resolver control transmitter consists of a two phase stator and a two phase rotor free to turn within the field of the stator. In each case, the two windings are physically at right angles to one another. In the application under consideration, only one rotor winding in the transmission is used and it can be seen the unused winding is short circuited. The resolver control transformer is electrically identical with the transmitter and the stator windings are interconnected as shown in the diagram. The transmitter rotor in use is energised from a 26v/115v 400 cps single phase AC supply.

Considering the system when the respective angular positions of the rotor/stator are as shown in the diagram, the maximum voltage will be induced across the stator coil Cos winding which is aligned with the rotor. No voltage will appear across the stator coil sine winding, which is at right angles to the rotor flux. The voltage at S1 and S2 will energise, via the connecting leads, the Cos winding of the control transformer whilst no current will flow in the sine winding. Unless the transformer rotor coil in use is at right angles to the energised cos winding, a voltage will, by transformer action, be induced across the rotor, the output of which controls the servo. Angular displacement of the transmitter rotor through 360° from a datum position will induce voltages across the two stator coils which will vary sinusoidally.

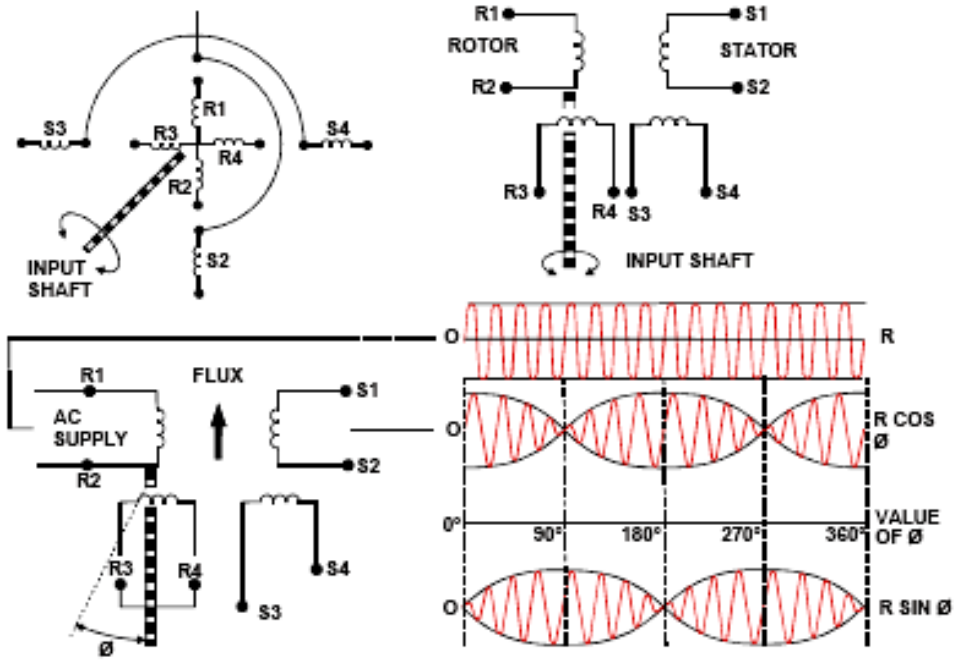


Figure 283 Resolver Synchro Control System

3.9 STANDARDISATION OF SYNCHROS

The basic idea of synchros is not a new one, but in the past many manufacturers working independently have produced individual systems such as Aysynn, Telesyn, Selsynn and Autosynn. While all these work on the same principle they are in no way standardised and are therefore not interchangeable. In order to distinguish the different types of synchros a code system is employed. A synchro is designated by a series of letters and numbers.

Standard Synchro Types and Symbols

The type of synchro is indicated by a code of two or three letters immediately after the frame size figure in the type designation.

The first letter classifies the instrument as control or torque:

- C Control.
- T Torque.

Succeeding letters identify the function of the instrument:

- D Differential.
- R Receiver.
- T Transformer.
- X Transmitter.

Types and Codes:

- CX Control Transmitter.
- CT Control Transformer.
- CDX Control Differential Transmitter.
- TX Torque Transmitter.
- TR Torque Receiver.
- TRX Torque Receiver Transmitter
- TDX Torque Differential Transmitter
- TDR Torque Differential Receiver
- The addition of the letter B indicates that the stator is rotatable.

3.10 SYNCHROTELS

Synchrotels can be used as synchro transmitters or transformers and have the advantages of having no brushes or slip rings and very low inertia. They are used for the output devices of capsule operated transducers.

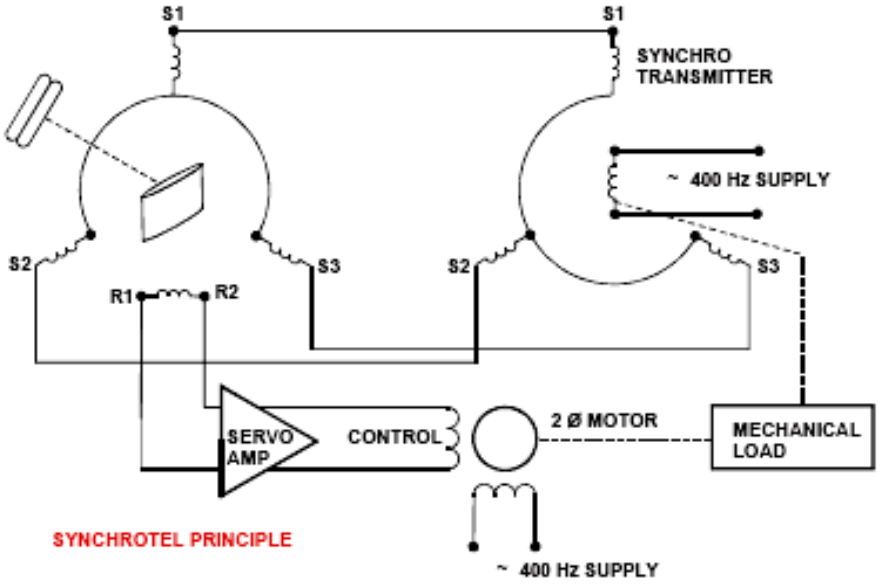


Figure 284 Synchrotel Principle

Construction

The stationary rotor coil fits into the high permeable case over the central core. The stator, which is a conventional poly-phase synchro type, fits over the same core. The rotor is an aluminium oblique section of a hollow cylinder which rotates in the air gap between the stator teeth and core.

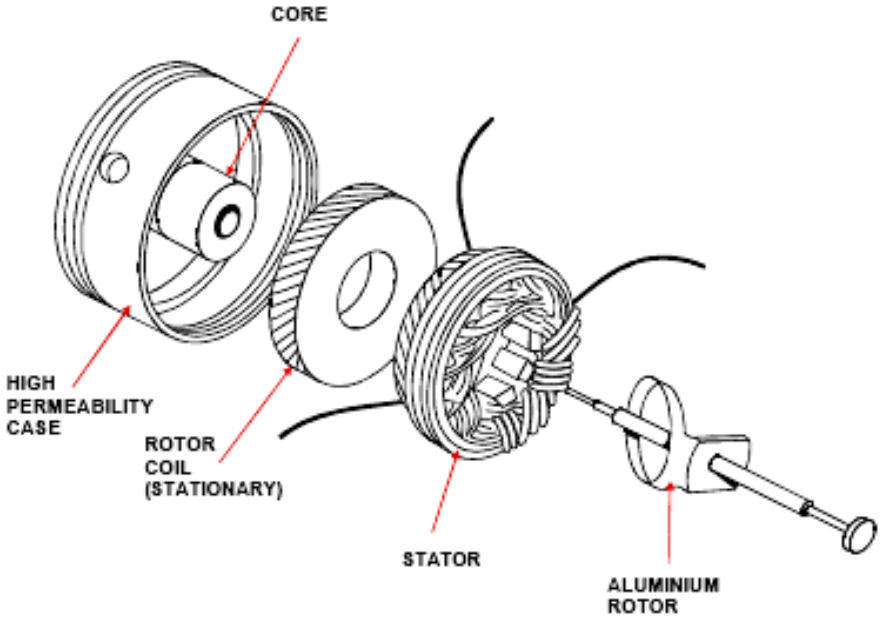


Figure 285 Synchrotel Construction

Operation

If the stator coils are excited with an ac voltage they will create a radial flux. Any of this flux which cuts the rotor loop will cause a current to circulate in the loop which can be likened to a single coil of wire. The flux generated will be at an angle to the core centre line and can be split into vector quantities; one radial opposing the flux causing the induced current (Lenz's Law) and the other axial along the centre core. The axial component will induce an emf in the fixed rotor coil whose amplitude is a sinusoidal function of the rotor position with respect to the stator flux. When used as a transmitter the rotor winding is energised producing an axial flux along the core which will produce a current flow in the rotor. The radial flux component will induce emf's in the stator coils in the same way as the rotor of a control synchro transmitter.

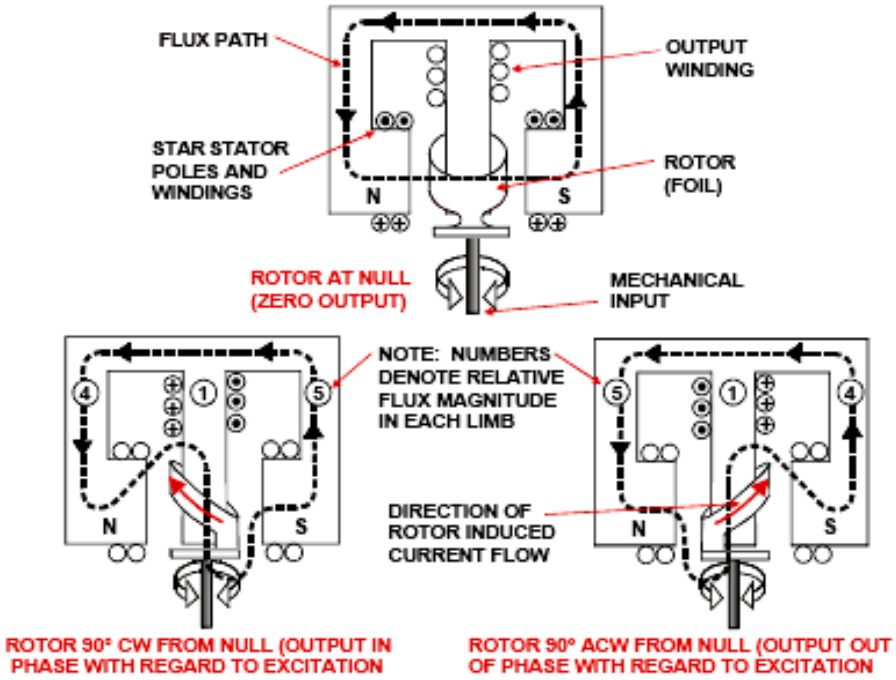


Figure 286 Synchrotel Operation

3.11 LINVARS

Linvars or Linear Variometers are a special form of resolver designed to give an alternating output directly proportional to shaft angle. This type give + 85 degrees of linear output about the null position which is where the rotor and stator slots are at right angles. As the rotor rotates in a uniform flux the proportion of flux linked with the rotor coil varies directly with rotor displacement.

There will be two zero positions.

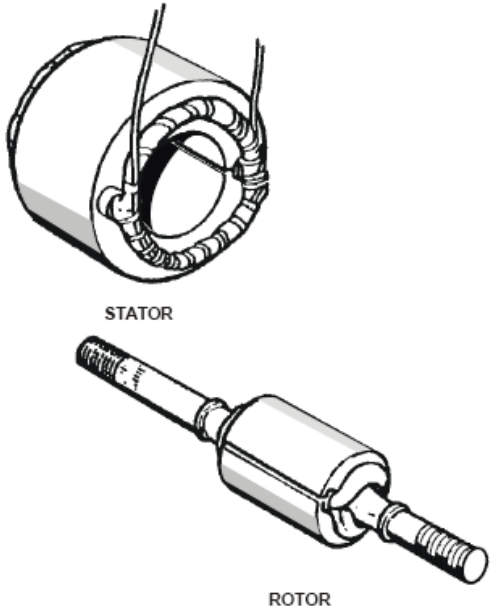


Figure 287 Linear Variometers

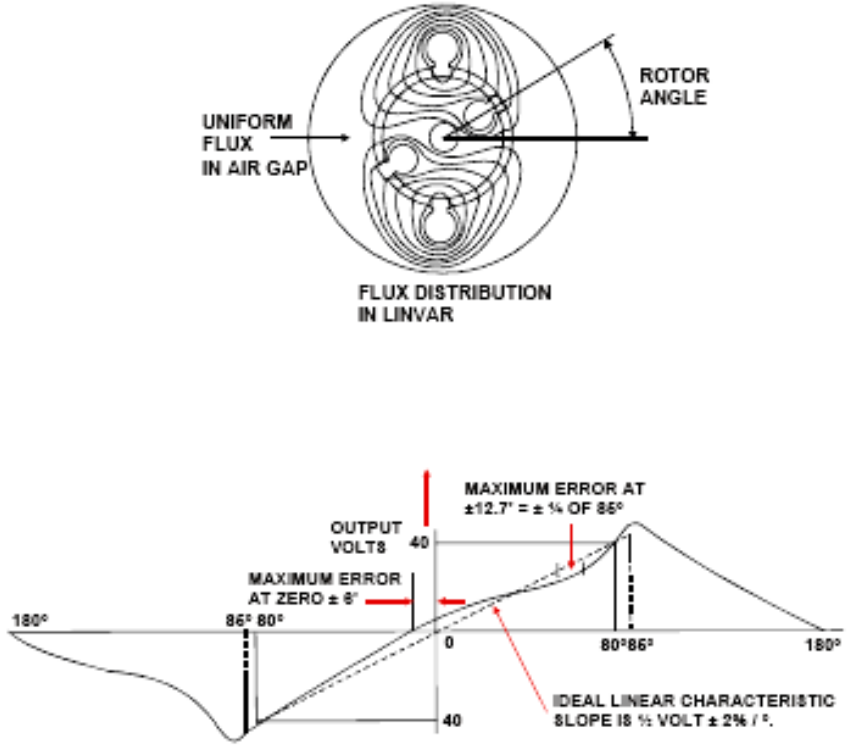


Figure 288 Linear Variometers

3.12 PROXIMITY SWITCH SYSTEM

The proximity switch is an inductive device that operates in conjunction with a steel target, the purpose of the target is to change the inductance of the inductive device.

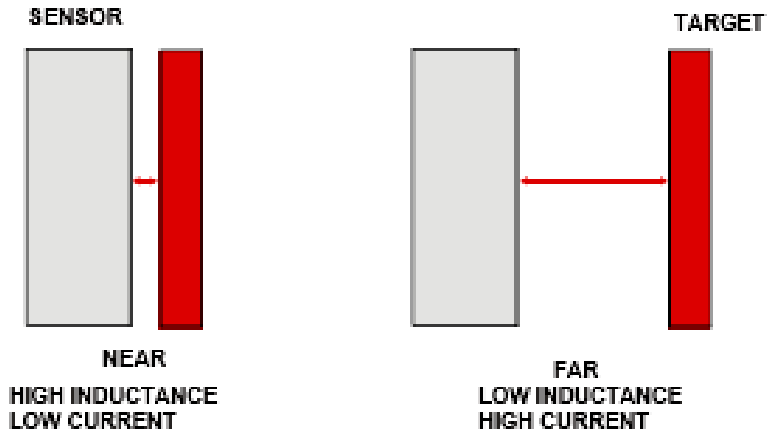


Figure 289 Proximity Switch System

The inductance of the device is changed, by placing a steel target next to the device. When the target is placed next to the device, the inductance will change to high inductance and current will change to low. When the target is moved away from the inductive device, the inductance will change to low and the current to high. Inductive Reactance XL = 2πf L ohms. Placing a target near, a proximity sensor will increase its inductance. An increase in inductance will increase the Inductive Reactance. An increase in Inductive Reactance will decrease the current. Such a device could be used to show a change of state, therefore, as the target is moved near or far the output state will change from high to low. This state is converted in the proximity switch electronic unit to a voltage. A typical output voltage for such a device is target near “0.3 volts” and target far “13 volts” this change in output voltage low to high could be viewed as a logic state, low as “logic zero” and high as “logic one”. The changing state of the sensor can be termed “in proximity” or “out of proximity” as the target is moved near or far. A typical diagram is shown for target near and target far. When near; the sensor/target inductance is five to eight millihenrys, and when far several millihenrys less.

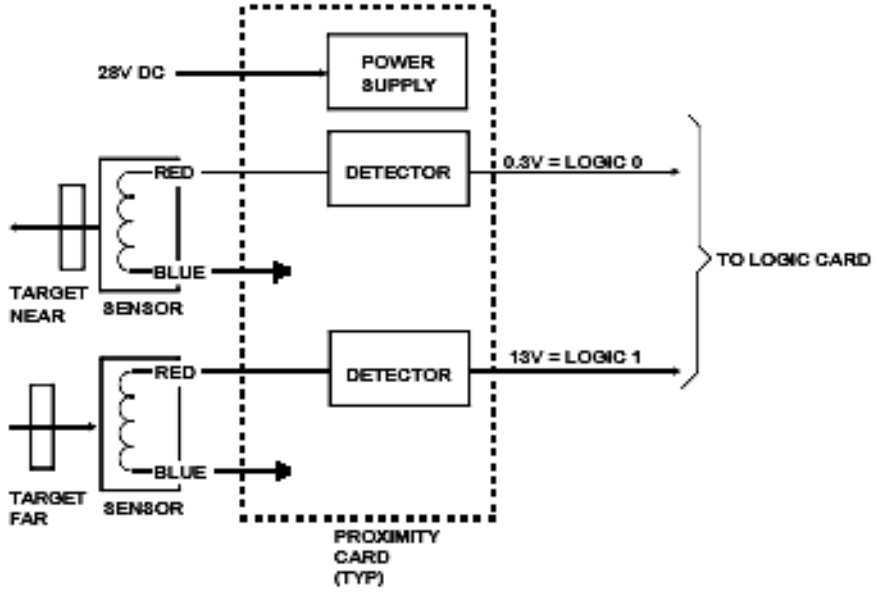


Figure 290 Proximity Switch System

Typical sensors are rectangular or cylindrical and are designated - Flange mount or Thread mount respectively. Targets are made from corrosion resistant steel and can be part of the aircraft structure or bolted to the aircraft structure in a fixed position or on movable linkage for example landing gear linkage.

Shown are typical sensors.

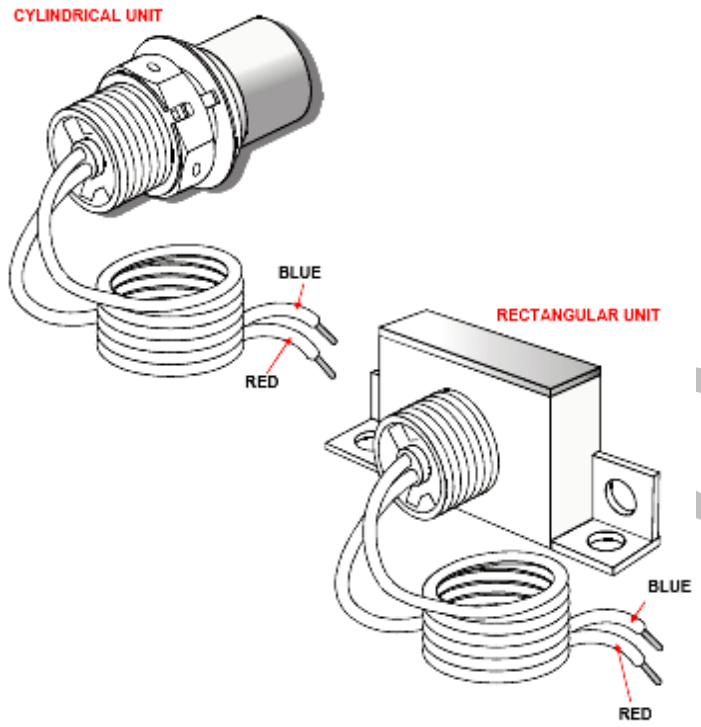


Figure 291 Proximity Switch System

3.12.1 Application

Proximity sensors are used extensively as signalling devices on all modern aircraft, for example passenger, cargo and equipment centre doors. These doors are indicated open or closed by proximity sensors. They are also used to indicate the position of landing gears, thrust reversers, leading edge slats, air ground sensing, valve position and failure protection. A typical example is landing gear, the landing gear is signalled in position both “up” and “down” and indicated to the flight crew. The flight crew will see three green lights in the flight deck indicating nose and main landing gears are in position “gears down and locked”, or “not locked” as the case may be. These systems normally have two systems, system one and system two. System one and system two indicate in position by illuminating the three green lights in the flight deck. If one of the sensors were to indicate out of position, caused by a “damaged or open circuit sensor” the other sensor will still indicate in position and indicate the landing gear is “down and locked” by illuminating the green light. As there are two sensors each sensor will illuminate its own light bulb, therefore, for the landing gear to be “not down and locked” both sensors must indicated out of position. Shown is a typical sensor arrangement for a nose landing gear.

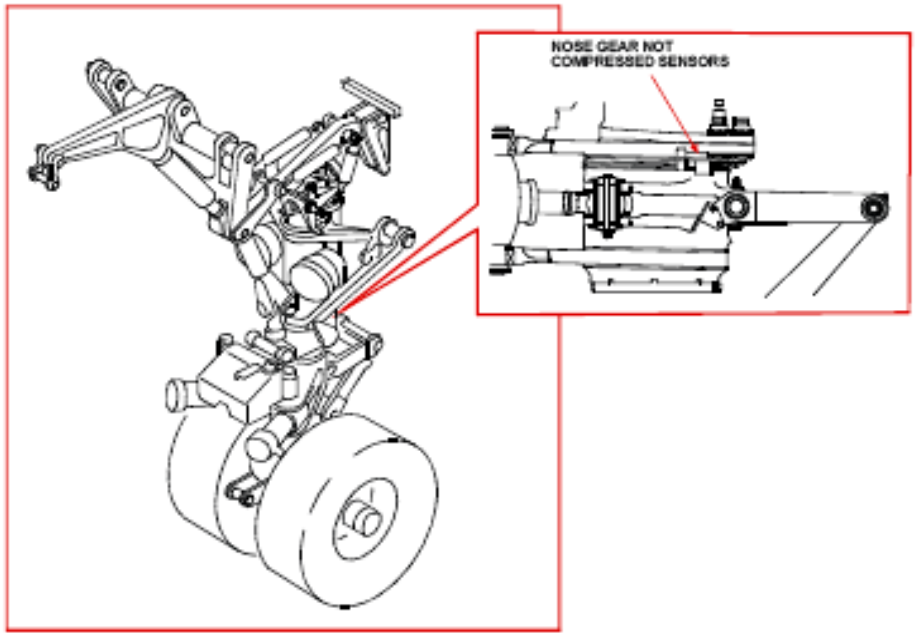


Figure 292 Proximity Switch System

Truck tilt sensors indicate that the aircraft is on the ground. As the aircraft lands the truck un-tilts, the sensors move into proximity and in doing so indicate aircraft on ground. The sensors through logic and driver cards in a box called Proximity Switching Electronic Unit (PSEU) will cause air ground relays to energise and indicate to the aircraft systems concerned that the aircraft is in the “on ground mode” not “in the air mode”. The reason for the differences between air and ground modes is that on the ground certain systems become live, while other systems become inactive for example cargo doors are only operative on the ground, whereas the ram air turbine is designed only to automatically deploy in the air. Shown is a typical sensor arrangement for truck tilt.

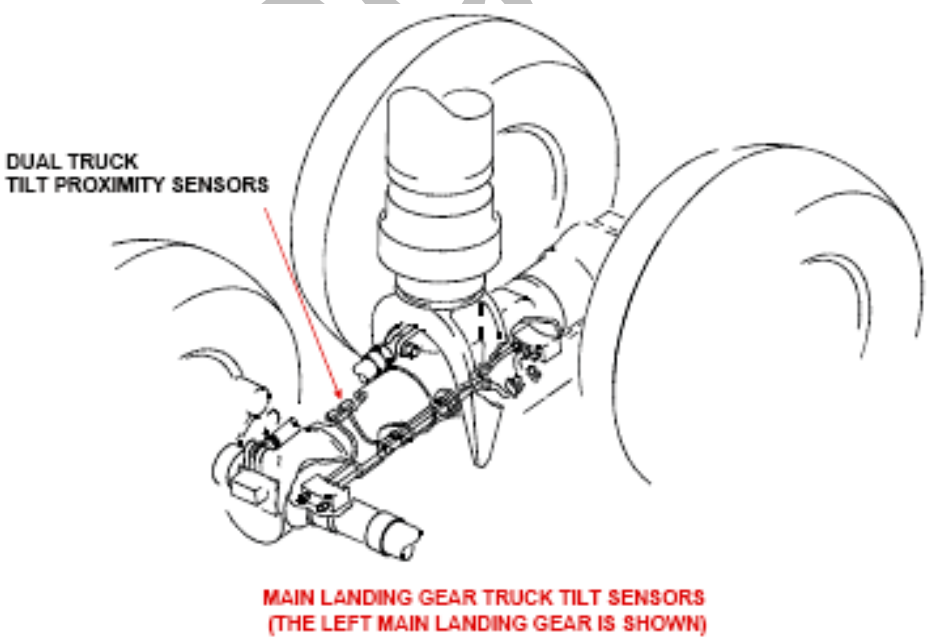


Figure 293 Proximity Switch System

A typical fault would be a failure to go into proximity; this may be caused by a sensor going open circuit, a target missing or an incorrectly rigged sensor. All sensors are installed into their position and rigged, this means the gap between the sensor and the target is adjusted for correct distance. If the gap is too wide, or the target is missing the sensor will be out of proximity. A fault like this would be detected by the (PSEU) and indicated to the flight crew either by a light illuminating or a message displayed on a status and maintenance screen. The fault will also be stored in the memory of the (PSEU) this is to enable maintenance engineers to fault find, locate and replace the faulty sensor or target. Testing and confirmation of a serviceable proximity switch system is accomplished by testing procedures performed by the Proximity Switch Electronics Unit

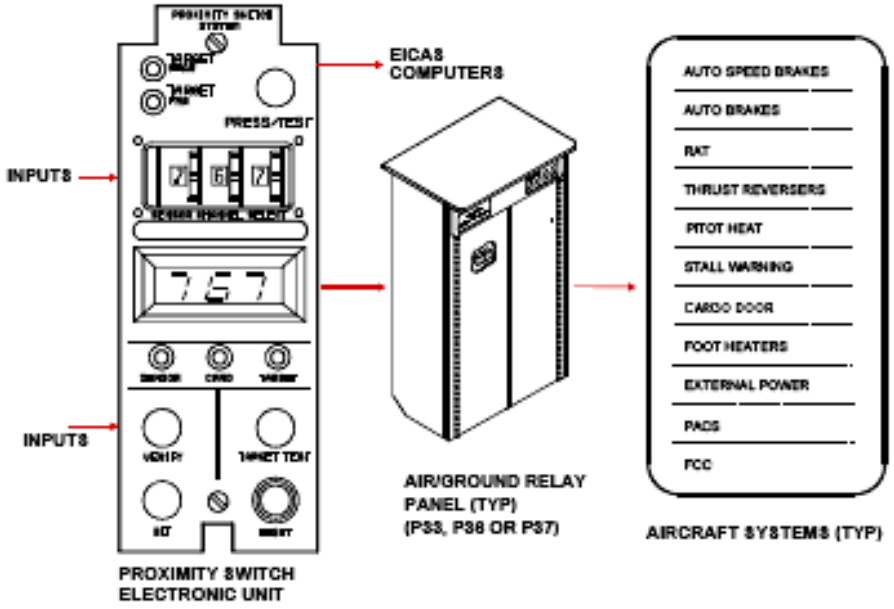


Figure 294 Failure Protection and Fault Detection

3.13 LINEAR VARIABLE DIFFERENTIAL TRANSFORMER (LVDT)

Depending on the direction of the primary winding - and the secondary side of a transformer the primary and secondary voltage can be either in phase or phase shifted 180°. The points on the windings indicate the phase relationship

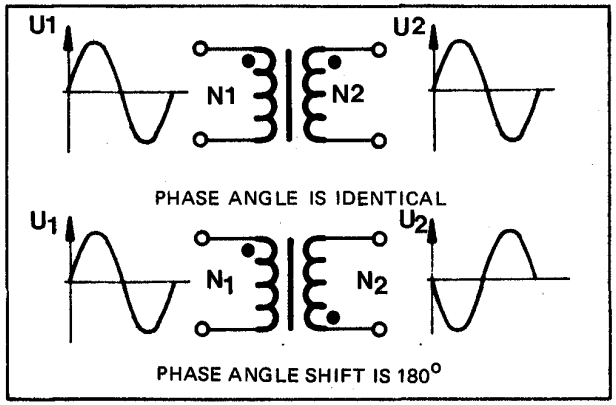


Figure 295 Groundfunction LVDT

Construction of a LVDT:

It consists of the primary winding N1 and the two oppositely in series connected secondary windings N2 and N3.

Depending on the position of the iron core, the outputvoltage changed in amplitude and phase position. Thus, the LVDT converts a mechanical position into an electrical signal. So the amplitude and phase of the mechanical position can be transmitted remotely (eg, flaps and rudder positions).

Function:

The movable iron core is in the middle position between N2 and N3. N1 induced in N2 and N3 two equal voltages of opposite phase:

Output voltage U2 = 0 V

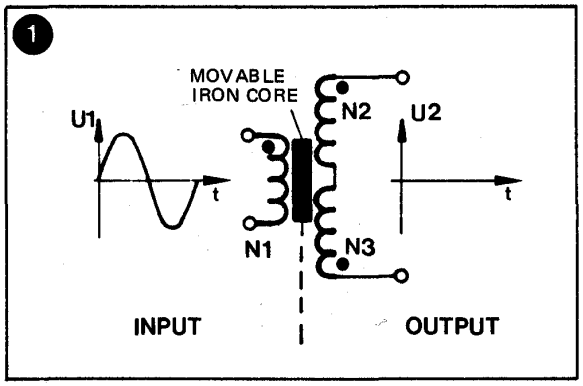


Figure 296 LVDT in centre Position

The movable iron core provides the magnetic couple between N1 and N2, so that the output of the LVDT produces a voltage U2 which is in phase to the Input voltage U1.

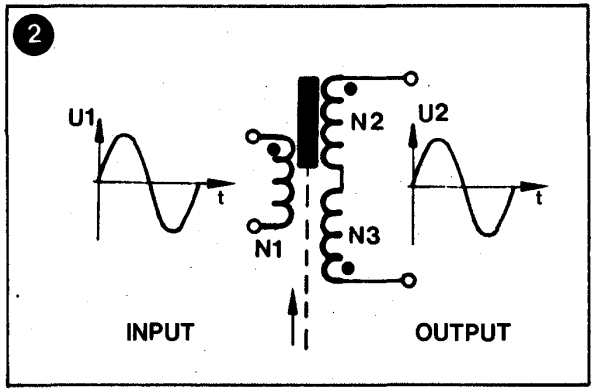


Figure 297 LVDT in phase Position

The movable iron core provides the magnetic couple between N1 and N3, so that the output of the LVDT produces a voltage U2 which is 180° out of phase to the Input voltage U1.

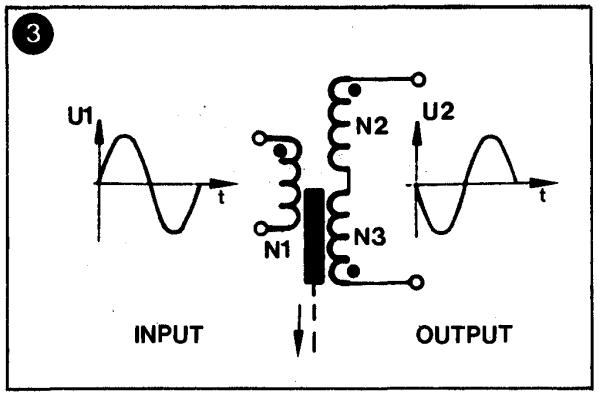


Figure 298 LVDT in out of phase Position

3.13.1 Rotary Variable Differential Transformer (RVDT)

As a new design, there is still the Rotary Variable Differential Transformer (RVDT) in which the movable iron core is able to rotate around an axis.
The function is identical to that of the LVDT's.

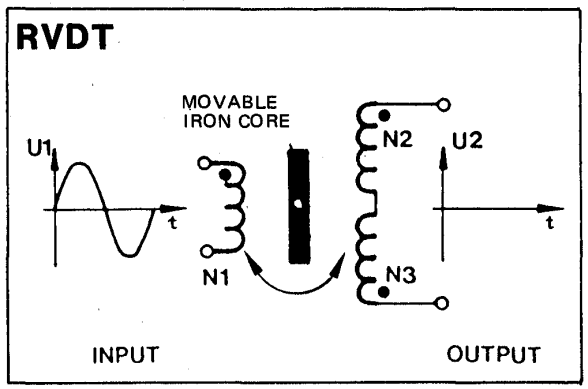


Figure 299 RVDT

3.14 DC AND AC RATIOMETERS

Introduction

Another type of remote reading indicator is the ratiometer, which can be one of two versions, DC or AC powered. They operate on a different principle to the synchronous systems. The first of these two systems to be described is the DC Ratiometer, it is used for sensing temperature and changes of temperature, for example the measurement of engine oil, fuel, and outside air temperature, again these hazardous temperatures can be measured at source and displayed on a remote indicator. The indicator is essentially a moving coil device, with a dial calibrated in degrees centigrade to indicate temperature. The transmitter in this case a temperature bulb (variable resistance) is located at the source, (the medium) to be measured. The indicator has two coils wound in opposition to each other on a former, which is pivoted with a pointer in a magnetic field and is free to rotate within that field. One coil (Coil B) is connected in series with a fixed value resistor this is the control coil, and the other coil, (Coil A) is connected in series with a temperature sensitive variable resistance which is the deflection coil. As can be seen in the diagram, they are connected together to form a parallel circuit which operates on the ratiometer principle. The ratiometer is designed to measure the ratio of currents flowing in two coils and, therefore, are not affected by variations in supply voltage. Any variation in supply voltage will affect both coils equally the tendency for one coil to move is counteracted by an equal and opposite tendency in the

other coil, therefore, both coils remain in the position determined by the resistance of the temperature bulb (variable resistance).

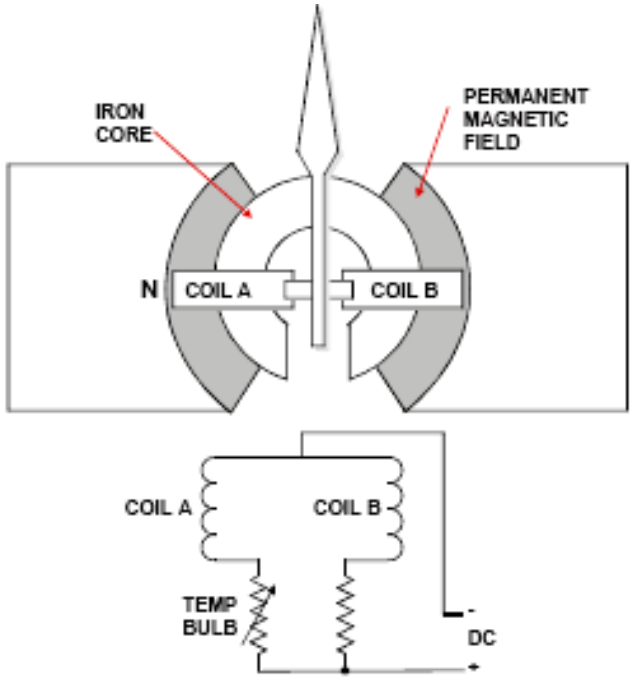


Figure 300 DC and AC Ratiometer

The fixed value resistor to the control coil calibrates the ratiometer so when at a certain temperature, the resistance is equal in the circuit, therefore, equal current will flow in both coils and the system will be balanced. This is usually at the centre of the scale, therefore, able to indicate an increase or decrease in temperature.

Operation

Whenever a change of temperature occurs at the variable resistor, the resultant change in resistance will unbalance the current ratio in the coils and cause them and the pointer to move to a new position, where the effect of the current and flux densities are alike. Where once again the balance is restored. This will register as either an increase or decrease in temperature. For example if the temperature of the variable resistance increases, the resistance will increase and in consequence the current will decrease in the deflection coil. (Coil - A). (Coil - B), the control coil will now be more dominant and will be driven into the larger air gap of the none-uniform magnetic field. Coil A the deflection coil, will be carried into the smaller air gap of the non-uniform magnetic field. The smaller the (air gap) the greater the effect of the permanent magnet field on the coil assembly and vice versa. When the effect of the (stronger coil) moving into a weaker field, is balanced by the (weaker coil) moving into a stronger field no further movement will take place. The coil with least current is compensated for by an increase in flux density.

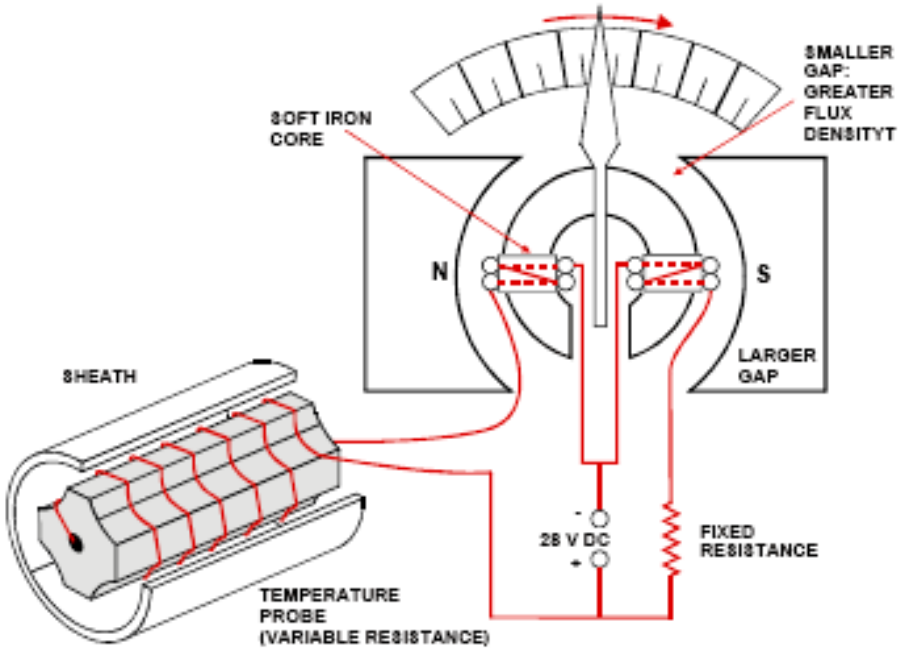


Figure 301 DC and AC Ratiometers

Note:

The current never changes direction; it will either increase or decrease in Coil A. Whereas in Coil B, the current flow is constant. With any moving coil indicator, the rotation of the measuring element is a product of current and magnetic field strength; the direction of rotation will depend on the direction of current relative to the magnetic field. As the ratiometer employs two coils there are two current flows, and it's the ratio of these currents associated with the ratio of flux densities, that will determine the position of the coils and pointer assembly. When power is removed from the ratiometer, the pointer is returned to an off scale position, this is achieved by a week spring to pull the pointer off scale. The supply voltage is applied to the ratiometer coils via the hairspring. Provided the supply voltage stays within limits (26 to 29 volts), the torque exerted by the hairspring will have no effect on the pointer movement. There are two types of temperature bulbs used in conjunction with the ratiometer, either platinum or nickel resistance bulbs. A ratiometer temperature gauge for use with platinum resistance bulb will be marked (PT LAW) on the rear of the instrument. Those for use with nickel resistance bulbs have no marking on the rear of the instrument, or they may be marked (NICKEL LAW).

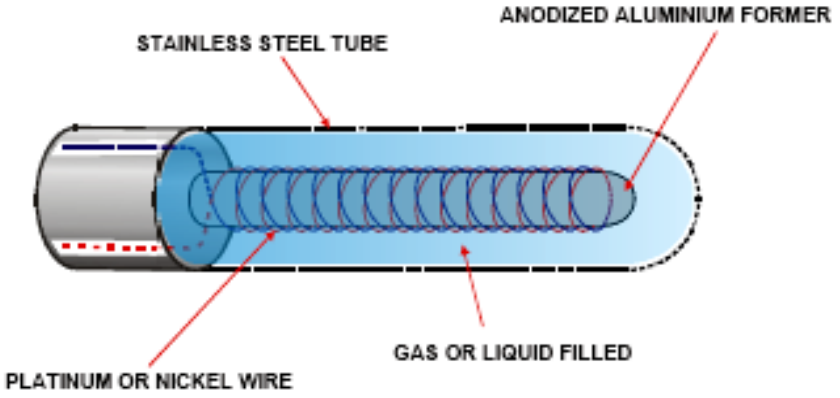


Figure 302 DC and AC Ratiometers

Advantages of a DC Ratiometer

- Provided the supply variations are within limits, (26 to 29 volts) there will be no change in indication.
- A remote reading system, without the need of mechanical coupling.
- Hazardous temperatures can be measured at source.

Typical Ratiometer Faults

- Pointer off scale; check power supply.
- Pointer shows full-scale deflection, check the temperature bulb, which may be open circuit, if so, all the current will go through the control coil and none to the deflection coil.
- Pointer shows minimum temperature indication, check the temperature bulb, which may be short circuit, if so, the resistance in the deflection coil will be zero. Therefore, there will be greater current in the deflection coil.

Testing the Ratiometer

Testing the ratiometer is carried out by connecting a decade-resistance box in place of the temperature bulb (variable resistance) the resistance values are set by the manufacture for calibration.

3.14.2 AC Ratiometer

Introduction

Another type of remote reading indicating system, is the AC Ratiometer, it is as its name states an AC operated instrument. The power supply to the ratiometer is 26 volts AC, with a frequency of 400 hertz. The operation of the AC Ratiometer in principle is the same as the DC Ratiometer, as it uses the current ratio of two coils by altering the resistance in the circuit. However this is an AC circuit and, therefore, the resistance is known as impedance, this will include inductance and capacitance.

The AC Ratiometer is used in conjunction with an inductor pressure transmitter, which is used to measure fluid pressures, for example engine oil pressure. When pressure is applied to the transmitter, the applied pressure produces an inductive change this output is then applied to the ratiometer that registers that change on a scale calibrated to indicate pressure

PSI (Pounds per Square Inch). Before the AC Ratiometer system is expanded upon, it should be noted however, the effects obtained by applying AC power to components having resistance, inductance or capacitance. There is no such thing as pure resistance, pure inductance, or pure capacitance. A wire-wound resistor, since it is wound in a form of a coil, has inductance and resistance. A coil has resistance as well as inductance. A capacitor has resistance as well as capacitance. However for the purpose of this explanation, it shall be assumed that the components are pure, which will enable their effects in a circuit that is supplied with AC power to be explained.

The AC Ratiometer consists of two coils mounted on laminated iron cores, the coils are arranged so as to produce an alternating flux across the air gaps of the cores. With a change in pressure at the transmitter, there will be a change of impedance in the transmitter coils, this change; will govern the currents flowing in the indicator coils. As the supply is AC and is alternating, the currents produced by the coils will produce an alternating flux across the air gaps in which the cam-shaped discs are situated. As this in itself will not produce movement of the cams, what is required is TORQUE to move the cams. “SHADING RINGS,” on the laminated iron core produce this. Shading rings, as can be seen in the diagram, are really LINEAR SHADED POLE MOTORS; current flowing through the coil wound around the laminated iron core produces a magnetic field “FLUX” across the air gap. This induces a magnetic field “FLUX” within the copper shading rings, which is out of phase with the “FLUX” in the laminated iron core. The two out of phase FLUXES acting across the air gap, produce EDDY CURRENTS in the aluminium cam, which in turn produce a magnetic field.

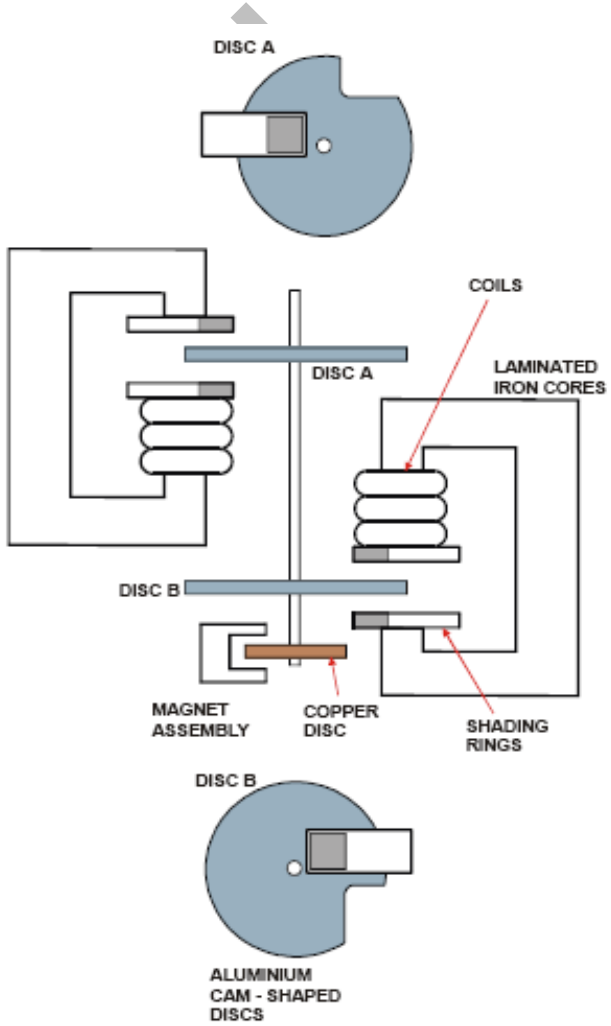


Figure 303 AC Ratiometer

The AC Ratiometer differs from the DC Ratiometer in that there is no fixed control coil or fixed deflection coil. The current ratio of the coils depend on the transmitter and the transmitter depends on the pressure either increasing or decreasing to alter the ratio of currents in the stator coils, this ratio of currents in the stator coils, will then be reflected by the ratiometer coils. As with the DC Ratiometer, with a power supply failure, the pointer will be returned to an off scale position by the hairspring.

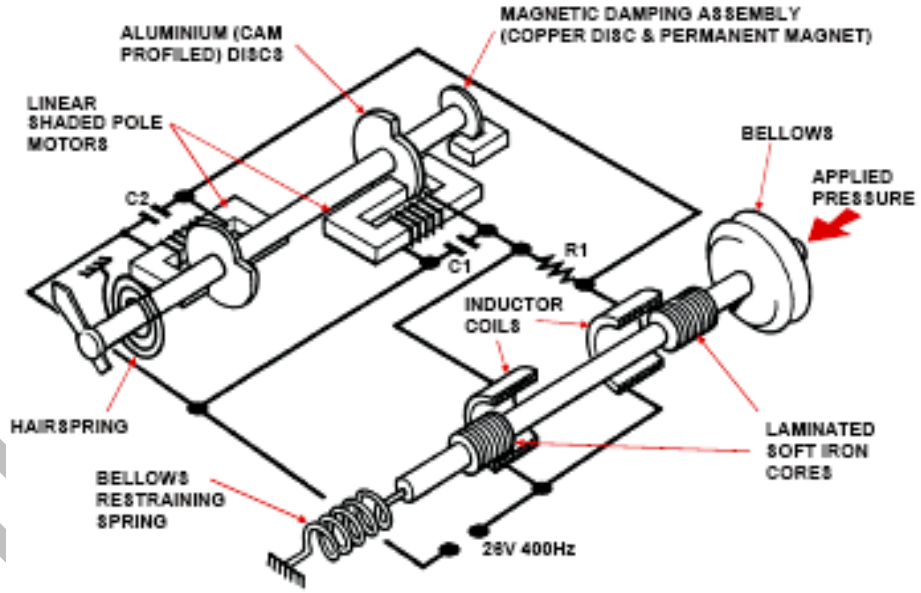


Figure 304 AC Ratiometer Testing

Testing

Testing of the indicator, is carried out in conjunction with a test set comprising of standard inductors, when powered, the indicators reading is compared with the values for the particular inductors used. Testing of the transmitter is carried out in conjunction with a Ratiometer test set and a Dead weight tester. All testing must be carried out in accordance with the manufactures test specification.

Faults with the System

Faults with the AC Ratiometer system are the same as the DC Ratiometer, an open circuit inductor coil in the transmitter or indicator, will either give full scale deflection or zero depending on which coil went open circuit and the position of the armature core in the transmitter.

3.14.3 Pure Resistance

In a purely resistive circuit connected across an AC supply, the current depends on the voltage applied to the circuit and on the resistance of the circuit. In other words, Ohms law applies at every instant, the current and the voltage rise and fall in phase with each other. As this is a purely resistive circuit, the current and voltage are in phase, power is developed in the resistance and heat is produced.

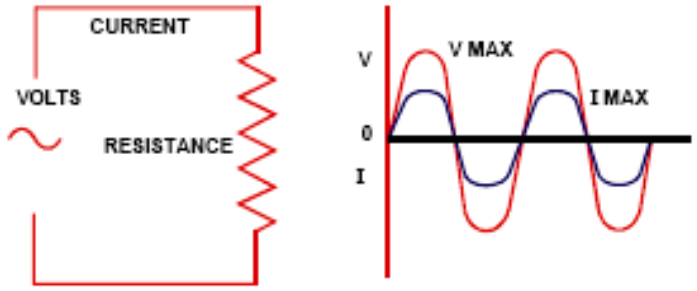


Figure 305 Pure Resistance

3.14.4 Pure Inductance

In a purely inductive circuit, the inductance opposes any change in the value of current in the circuit and its effect is such that with an AC power supply, the current through the inductance always lags behind the applied voltage. With a sine wave the angle of lag is 90 degrees. Since an inductance opposes any change in value of current at all times, and because the supply is alternating, the amount of opposition depends on the rate at which the current is changing, this is related to the supply frequency and the value of the inductance. As this is an AC circuit instead of the resistance of an inductance it is now called reactance, because the opposition or reactance of a coil depends on frequency, reactance, like resistance, is measured in Ohms. (In other words, AC Resistance.)

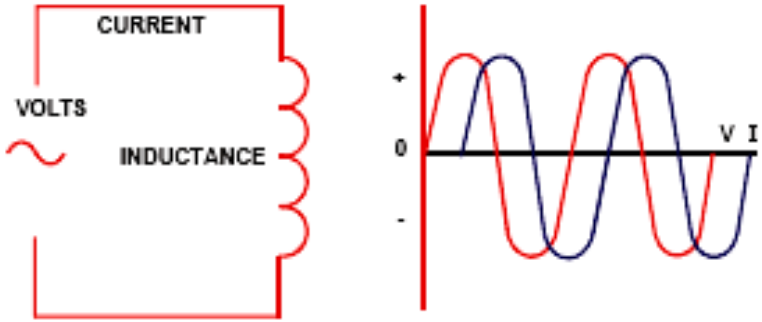


Figure 306 Pure Inductance

3.14.5 Pure Capacitance

In a purely capacitive circuit, the phase relationship between current and voltage is opposite to that in a pure inductance. Voltage exists across the plates of a capacitor only after current has flowed to charge the plates. The current leads the voltage in a capacitive circuit and with a sine wave the lead is 90 degrees. A capacitance opposes any change in the value of voltage applied to it and presents an opposition to current at all times with an alternating supply. Opposition to current is capacitive reactance and it is measured in Ohms just like a resistive circuit.

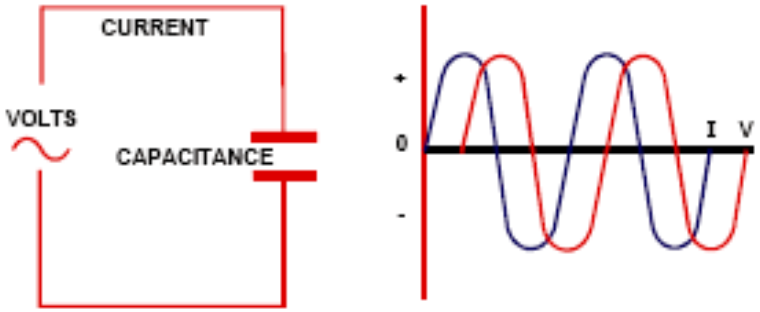


Figure 307 Pure Capacitance

3.14.6 Inductor Pressure Gauge Circuit

The AC Ratiometer inductor pressure gauge system has all three components that were just explained, and, therefore, it must be taken into account the effects of variations in the supply voltage, variations in supply frequency and variations in temperature have on the indicator itself. First the effect of supply voltage. As the supply variation will be common to both coils and the indicator being a ratiometer, the effect of a voltage supply change will be equal to both coils. Therefore, what happens to coil A, also happens to coil B, the ratio of current will remain the same, therefore, no change in indication. As the power supply is alternating AC, the effects of variations in the frequency must be taken into account and compensation provided for this effect. This is provided for by the two frequency compensation capacitors.

With an increase in frequency, coils A and B would oppose the current changes produced by the transmitter, therefore, the coil's reactance will increase, but this reactance change is cancelled by the capacitors because the frequency change on capacitors is exactly opposite to those produced by the coils. This way the effect of frequency variations is cancelled. An increase of temperature will reduce the ratio of impedance in the coils, which will in turn reduce the scale range of the indicator. This same increase in temperature will increase the resistance of the sensitive resistor across the coils, which will increase the range of the indicator. The two reactions are opposite, and the effects of temperature are therefore, cancelled

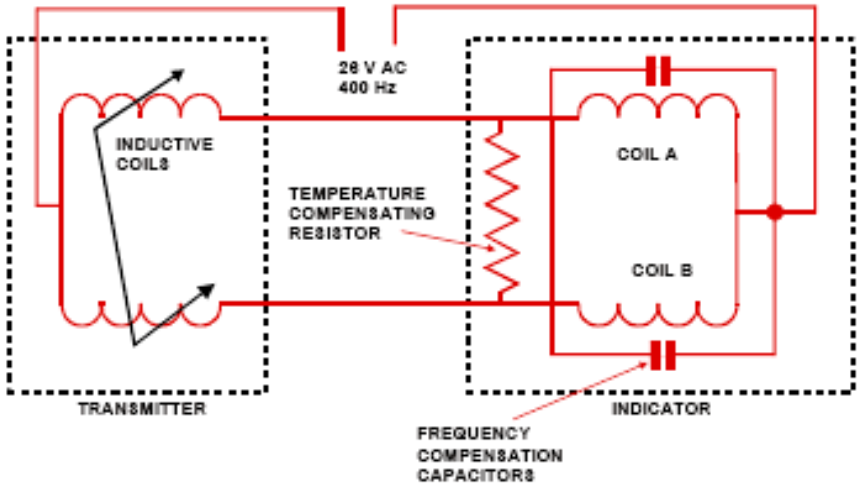


Figure 308 Inductor Pressure Gauge Circuit

3.14.7 Inductor Pressure Transmitter

Pressure Indicating System – Inductor Type

Present day systems used to measure engine oil and fuel pressures make use of an inductor type pressure transmitter mounted at the required locations on the engine. Electrical signals from the transmitters are passed to AC RATIOMETER type indicators that are calibrated in the appropriate pressure units.

Inductor Pressure Transmitters

A typical transmitter is shown in sectioned view and consists of a body containing two main subassemblies; a base, bellows and spindle assembly and stator and housing assembly. The base, bellows and spindle assembly consists of a flexible brass bellows soft soldered to a base at one end and a centre spindle at the other end. **Two armatures are located on the centre spindle with their poles at 90 degrees to each other**. A cup for locating a main spring is also attached to the centre spindle which in turn is supported in to guides sliding in two guide bushes.

The stator is constructed so that the two stator windings are positioned at 90 degrees to each other to reduce electrical coupling between the windings. The main spring seated between the inner face of the spring adjuster and the spring cup controls the movement of the bellows and spindle and, consequently, the armatures.

The fluid under pressure enters the bellows through a 4BA tapped hole in the base plate causing the bellows to expand and move the armatures in their respective stator windings. The armatures are normally displaced on the spindle such that the movement will result in entering farther into its coil to increase its impedance while the other is moving out of its coil and reducing its impedance. The stator windings are normally supplied at a constant 26v 400 Hz and the variation of impedance will cause the current in the windings to vary in proportion. The pressure being measured at any time will be

proportional to the RATIO of the currents in the two stator windings. It is usually arranged that the mid range point represents equal winding impedance and therefore equal currents (ie, a ratio of 1).

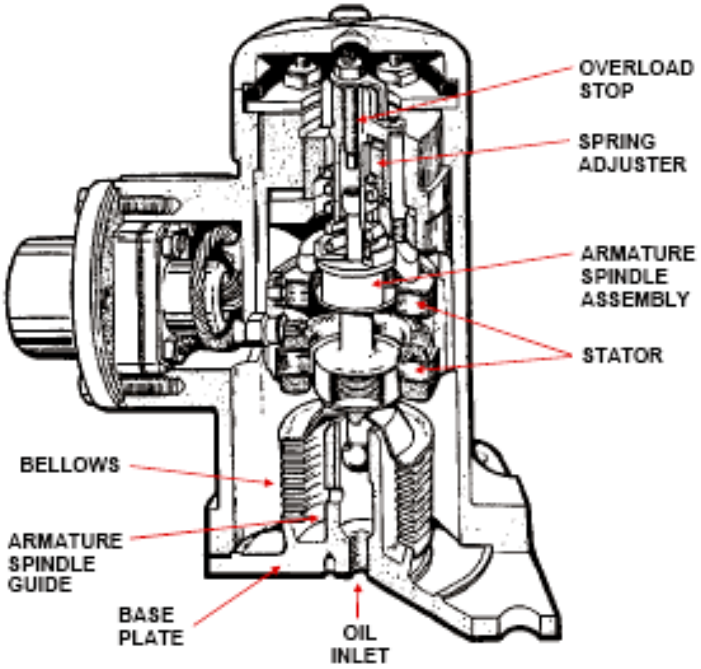


Figure 309 Inductor Pressure transmitter

Operation

Oil under pressure is admitted to the bellows, so causing them to expand and move the armature spindle towards the overload stop; displacement of the spindle alters the position of the armatures and therefore the inductance of the stators. This results in a decrease in current in the circuit of the stator nearest the bellows and an increase in current in the circuit of the stator farthest from the bellows. The current ratio so produced is denoted by the indicator.

Calibration

Calibration is carried out in the workshop by means of the spring adjuster, using a dead weight calibrator and suitable adapter. No adjustment can be made in situ on the aircraft. An overload stop is also suitably positioned to prevent over travel of the spindle and permanent damage to the bellows in the event of an excess pressure being applied.

4 LIST OF FIGURES