An Introduction to Semiconductors part 2

In the last talk I covered diodes and rectifiers. In this talk I'll be covering transistors in their many guises and introducing integrated circuits. I won't be going into any great detail as to how transistors are made or exactly how they work. When I was at university one of the things I remember was one lecturer going through a long series of equations, the result of which was showing exactly how a transistor worked. Although I understood it at the time it's just a memory now! I'll also mention some other commonly used semiconductor devices.

During the Second World War there were many advances in electronics technology including radar where, as frequencies increased, valves were found to be less suitable. In particular silicon diodes were found to be eminently suitable as detectors. After the war more research was done into semiconductors and in late 1947 John Bardeen, Walter Brattain and William Shockley working at AT&T's Bell Labs discovered that by placing two closely spaced gold electrodes onto a slice of germanium an amplifier could be made. For this discovery, the point contact transistor, they were jointly awarded the 1956 Nobel prize for physics.

Several manufacturers, including Mullard in the UK, began manufacturing point contact transistors but they were difficult to make, nevertheless several types were available in the early 1950s. The ratings meant they were only useable in small signal stages. Figure 1 shows the construction of an RCA transistor and figure 2 shows a Mullard OC50 in which the case is actually the base connection.

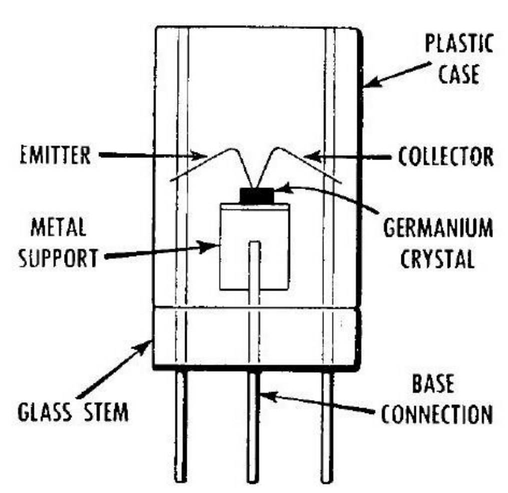


Figure 1- Point Contact Transistor (RCA)

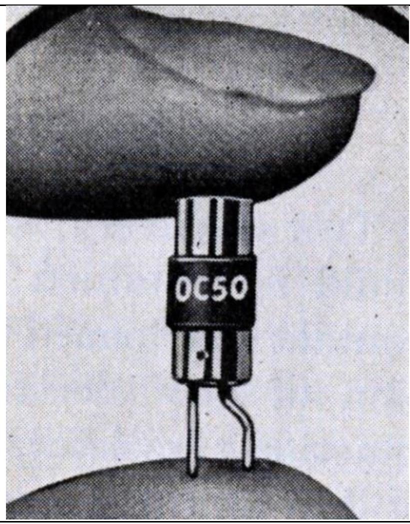


Figure 2 - Mullard OC50

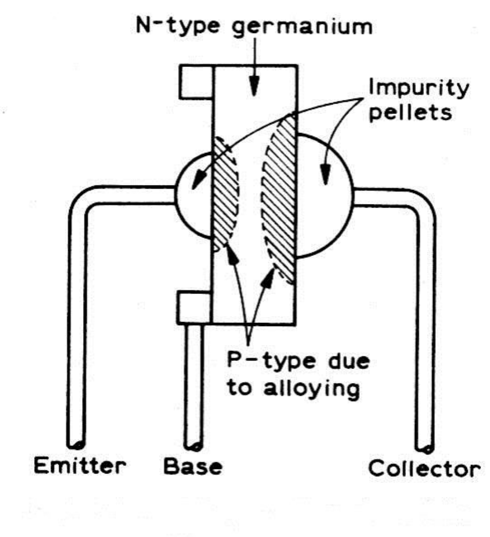


Figure 3 - Junction Transistor Construction

Within three years of the discovery of the point contact transistor a new type of transistor was invented, the bipolar junction transistor which proved easier to manufacture and quickly became the dominant type.

As with the point contact type the base is a crystal of N type germanium but instead of having emitter and collector electrodes touching the base they are formed by alloying specific impurities, typically indium. The alloying of the electrodes produces P type regions for the emitter and collector as shown in figure 3.

The distance between the two P type regions forms the base region and this distance affects the high frequency performance of the transistor.

The most well known transistors of this type in the UK are the Mullard OC series with the OC71 being aimed at audio applications and the OC44 and OC45 being aimed at RF applications.

The majority of the junction transistors were germanium PNP types although there were a few NPN types produced. Some of the early silicon transistors, such as the OC200 types, were PNP junction types using a similar manufacturing process as for the germanium junction transistors.

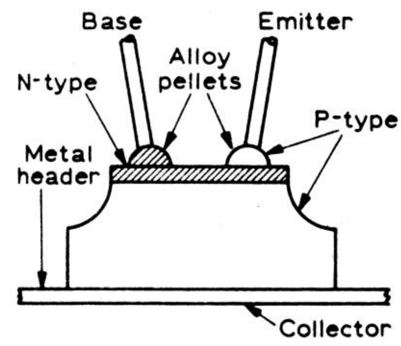


Figure 4 - Mesa Transistor Construction

Once the junction transistor had been introduced work continued to improve the performance of the transistors. One development was the MESA transistor where the base region is diffused into the collector region and the emitter is alloyed to the base region as shown in figure 4. The name comes from its resemblance to a flat topped mountain or mesa. The collector is bonded to a metal header which can form part of the encapsulation.

This form of construction can be used for both germanium and silicon transistors but because the diffusion of N type germanium into P type germanium is not an easy process only PNP germanium transistors were made using this form of construction. There were no such problems with silicon so both PNP and NPN silicon transistors were available.

Mesa transistors had better high frequency performance than junction transistors and were used in higher frequency RF and IF stages.

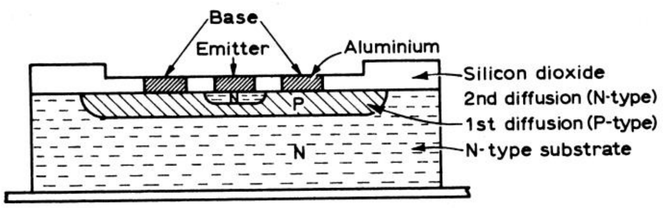


Figure 5 - Silicon Planar Transistor Construction

Further developments with silicon resulted in the planar transistor. Here silicon dioxide is grown on the surface of the collector region. The silicon dioxide is then selectively etched away over the base region then the base impurity is diffused into the collector region. Silicon dioxide is then re-grown over the base region which is then selectively etched over the emitter region and the emitter impurity is then diffused into the base region. Figure 5 shows the construction of an NPN transistor. For a PNP transistor the collector, base and emitter region diffusion types are reversed. Etching of the silicon dioxide used photographic masks to define the areas for the base and emitter regions meaning many transistors could be made simultaneously on a single slice of silicon. The transistors also had more consistent characteristics and the construction method reduced the cost of silicon transistors and paved the way for the manufacture of integrated circuits.

Once the manufacturing methods had been perfected, silicon transistors became the dominant type used for all applications.

All transistors have a leakage current, that is the current that flows between the emitter and collector when the base is unbiased. This current is temperature dependant, increasing as the temperature increases and limits the maximum operating temperature. If the temperature increases, the leakage current increases which can increase the temperature of the transistor which increases the leakage current, further increasing the temperature. This is known as thermal runaway and can destroy the transistor unless steps are taken to prevent it. Silicon transistors can operate at higher temperatures than germanium and have much lower leakage currents.

Power Transistors

The transistors described so far have all been small signal types with maximum collector currents of a few 10s of mA and maximum power dissipation of around 100mW to 300mW. The main problem with increasing the power dissipation was getting rid of the heat. Early low power transistors such as the OC72 were basically the same construction as the lower power OC71 but had a metal cover to help remove the heat. When higher collector currents were required a different form of construction was required where the transistor wafer is mounted directly onto a metal base which forms the collector connection and allows the heat to be removed to the outside world more easily. The metal case also allowed the transistor to be mounted onto a heatsink to improve the heat removal and keep the transistor temperature down. One of the most common case styles for power transistors was the TO3 and the similar, but smaller, TO66.

Bases and Connections

With only three connections you'd think that it would be a simple matter to create a "universal" pinout but with the number of manufacturers producing transistors there were almost as many variants of cases and pinouts as shown in Figure 6.

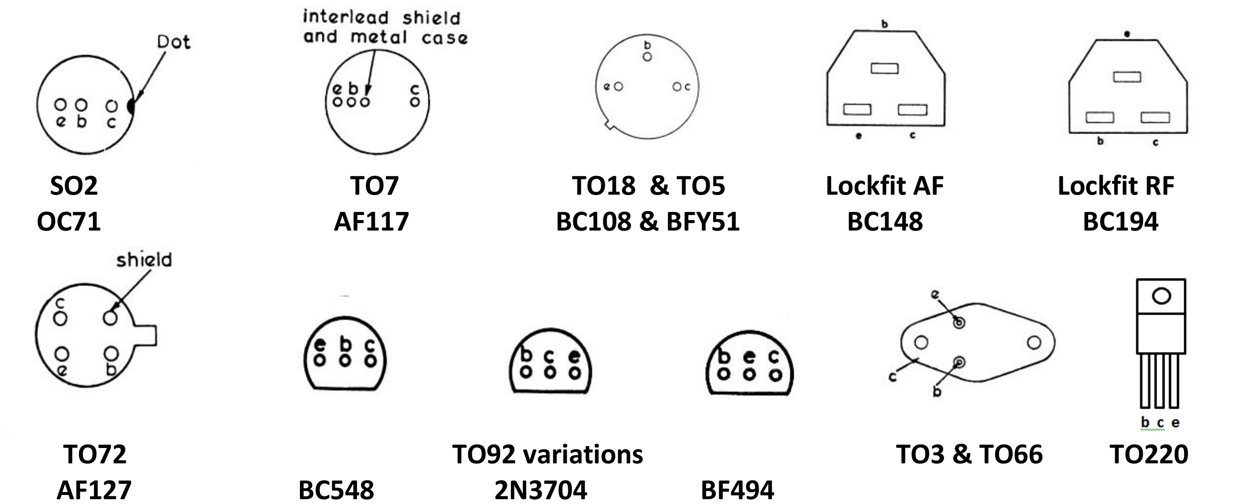


Figure 6 - Base connections with examples of transistors. Note - TO220 is shown from the front of the device, all others are shown from the underside.

Figures 7 and 8 show several small signal transistors, both germanium and silicon, showing the variety of case styles from the glass cases of the OC71 and OCP71 to the metal cases of the red spot, MAT 121, GET872, OC170, BC107 and BF258, the plastic cases of the BC149, BC182 and ZTX653 and the ceramic/epoxy case of the TSC2000. The OC84 has the same glass case as the OC71 but with a metal cover to help with heat dissipation.

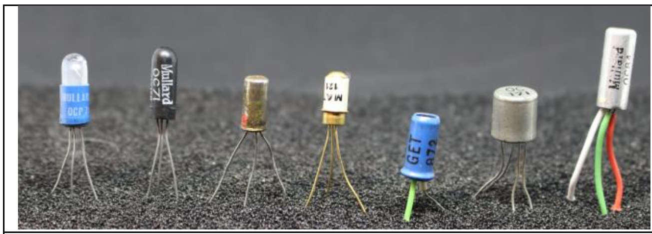


Figure 7 - Small Signal Germanium Transistors
Left to right – OCP71, OC71, Red Spot, MAT121, GET872, OC170, OC84

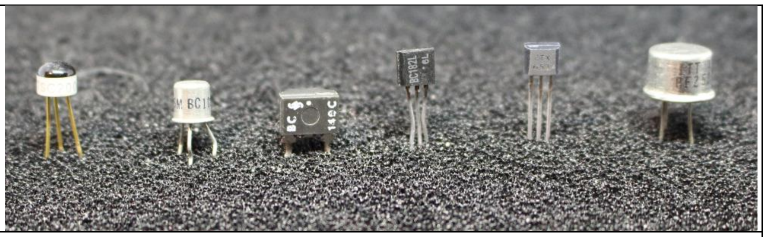


Figure 8 - Small Signal Silicon Transistors
Left to right – TSC2000, BC108, BC149, BC182, ZTX653, BF258

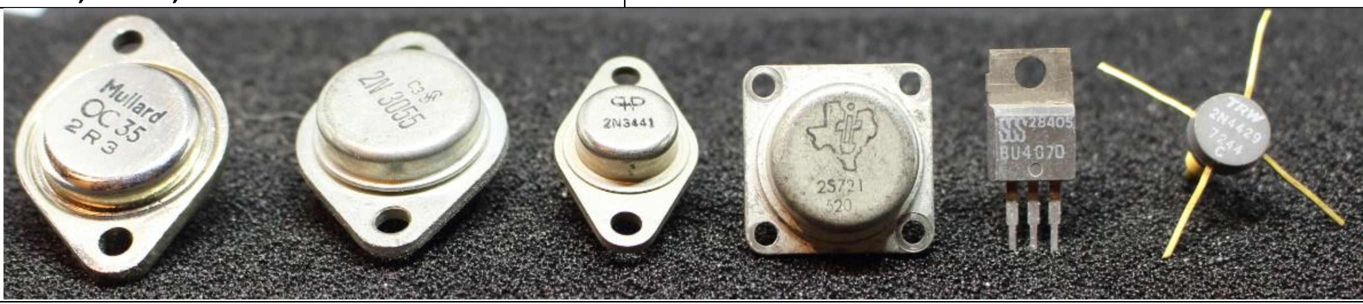


Figure 9 - Germanium and Silicon Power Transistors
Left to right – OC35, 2N3055, 2N3441, 2S721, BU407D, 2N4429

Figure 9 shows several power transistors, the OC35 and 2N3055 are in TO3 cases, the 2N3441 is a TO66 case, the 2S721 is a TO53 case, the BU407 is a TO220 case and the 2N4429 is an RF power transistor in a TO117A case.

The majority of the cases styles have been standardised as JEDEC (Joint Electron Device Engineering Council) standards with a TO (Transistor Outline) number. There have been many styles and many, especially metal cased power transistors, are now obsolete. The early cases such as the Mullard OC series were glass but later metal cases, such as the TO18, soon took over as they were easier to manufacture especially for silicon planar transistors. The TO92 plastic encapsulated transistor has largely replaced the TO18 types.

In the late 60s the plastic encapsulated Lockfit transistor was introduced as an alternative to the TO18 type but was not universally accepted as the plastic TO92 case had been introduced, meaning it was limited to a few transistor types. Several transistors were produced as TO18, Lockfit and TO92, the BC108, BC148 & BC548 being one example. The three parts are electrically identical but have the TO18, Lockfit and TO92 cases respectively.

In the early days of transistor manufacturing, the processes were not as precise as they were later and there were wide variations in the characteristics. The transistors were tested and, dependant on the measured characteristics, were placed in different “bins” for gain, frequency response, low noise etc. An example of this is the OC44 and OC45 where the OC44 could operate at higher frequencies and was therefore more suited to the frequency changer stage of a radio whereas the OC45 was more suited to the IF stages. Because of the testing there were large numbers of transistors which, although they worked, fell outside the required characteristics. These were sold off on the surplus market as red spot (AF) and white spot (RF) at a fraction of the price of the full spec transistors. One well known seller of surplus transistors was Clive Sinclair who bought surplus transistors, tested them and sold them in his products and as individual transistors. The MAT121 shown in figure 7 is one of his transistors.

Figure 7 also shows an OCP71. This has similar characteristics to an OC71 but it has a transparent case making it light sensitive. In the diode talk I showed how light on a diode can cause a current to flow giving a voltage across a load resistor. This also happens in a transistor, the light causing a current flow into the base emitter junction. This is amplified by the transistor action to give a higher current in the collector. In the standard OC71 the black case prevents this action but in the OCP71 it is encouraged making it useful as a detector in beam type sensors and optical audio links. The OCP71 used to sell for about 3 to 4 times the price of the OC71. However people soon found that if the paint was scraped off an OC71 it became photo sensitive and was much cheaper than an OCP71. The story goes that Mullard then started filling the OC series of glass transistors with an opaque material to prevent them being used as cheap photo transistors.

Transistor Characteristics

We’ll now look at the characteristics of a transistor using the curve tracer. Each plot shows the collector current on the Y axis versus the emitter collector voltage on the X axis. For each plot the X and Y axis settings and the base current are shown below the plot.

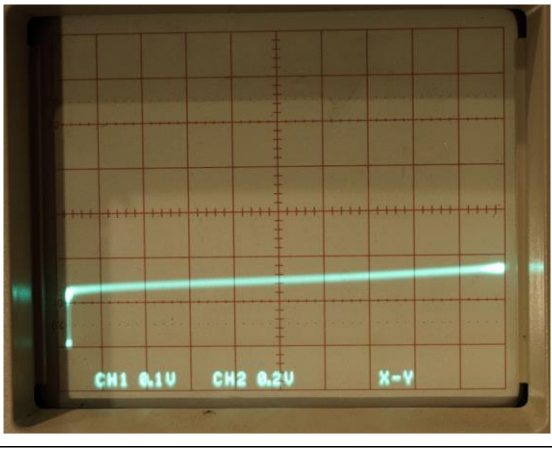


Figure 10 - OC71
Ib = 20µA, X = 2V/div, Y = 1mA/div

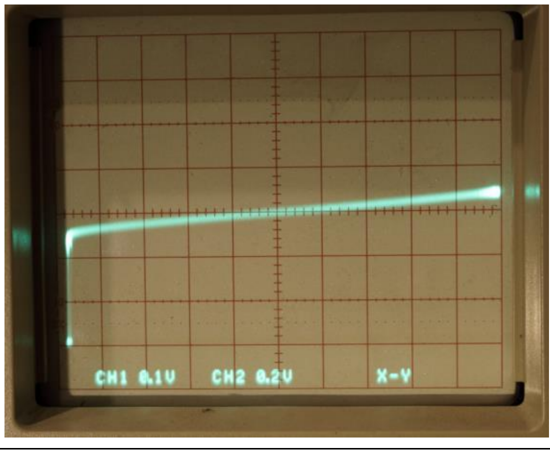


Figure 11 - OC71
Ib = 40µA, X = 2V/div, Y = 1mA/div

If you remember from the diode talk the forward voltage was dependant on the semiconductor material. The same goes for transistors. The base emitter junction is forward biased and the voltage across the junction is dependent on the semiconductor, around 0.3V for a germanium transistor and around 0.6V for a silicon transistor. This means the base voltage has to be about 0.3V or 0.6V relative to the emitter before the transistor will start passing any current.

Figure 10 shows an OC71 with a base current of 20µA and figure 11 shows the same transistor with a base current of 40µA. You can see that the current gain is approximately 60 at a low emitter collector voltage rising to about 90 at 20V.

The data for the OC71 shows a typical current gain of 41 at a collector current of 1mA and emitter collector voltage of 2V. The plots show how variable the current gain can be and the test conditions should always be stated when the gain is specified.

Figures 12 and 13 show the characteristics of an OCP71. The base current for both plots is the same but for figure 13 the transistor is illuminated using a torch. You can see that the light effectively increases the base current.

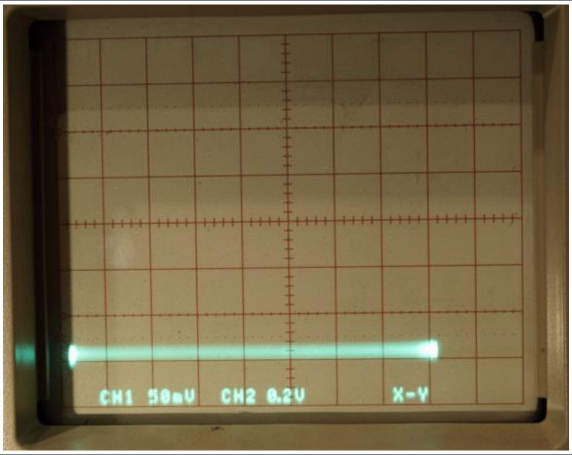


Figure 12 - OCP71 – dark
Ib = 20µA, X = 2V/div, Y = 0.5mA/div

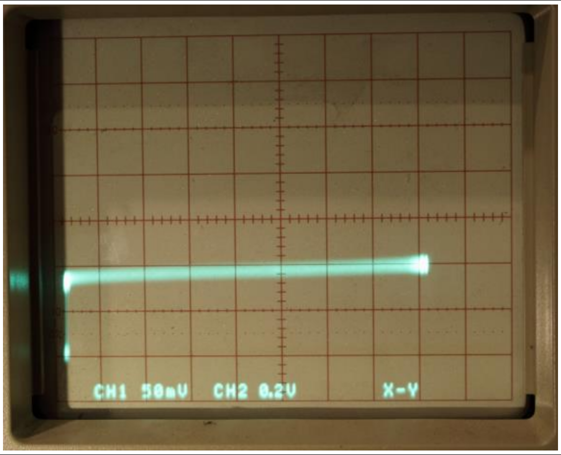


Figure 13 - OCP71– light
Ib = 20µA, X = 2V/div, Y = 0.5mA/div

A bipolar junction transistor, PNP or NPN, has two junctions with one being designated as the base emitter junction and one as the base collector junction. Now these two junctions are normally different because of the construction of the transistor, especially in the case of a silicon planar transistor, but can they be reversed. The answer is yes but the characteristics will be different. Figure 14 shows the normal characteristics of a BC108 and figure 15 shows the characteristics of the same transistor with the emitter and collector reversed. Both the breakdown voltage and current gain are lower but it still functions as a transistor.

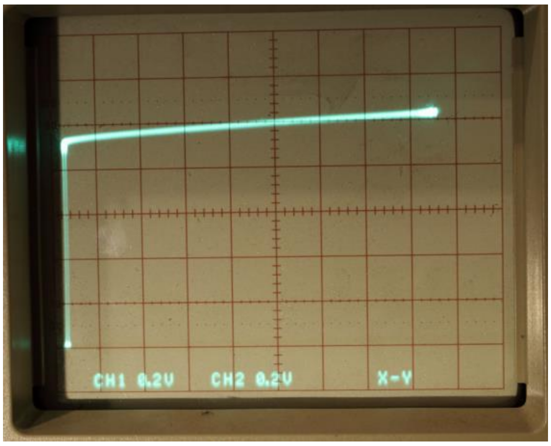


Figure 14 - BC108
Ib = 20µA, X = 2V/div, Y = 2mA/div

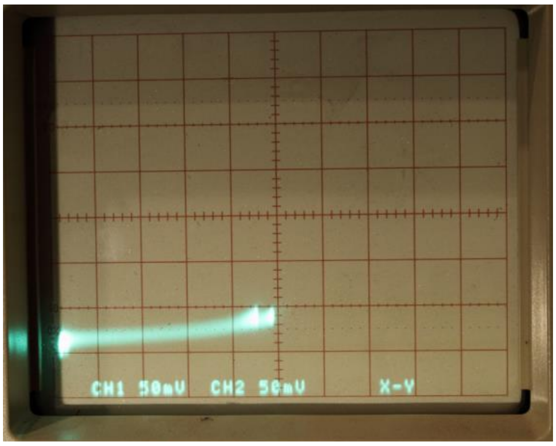


Figure 15 - BC108 EC swapped
Ib = 20µA, X = 0.5V/div, Y = 0.5mA/div

Some germanium transistors were made with interchangeable emitter and collector junctions, one example being the OC140 which was extensively used in early 625 line to 405 line TV standards converters.

Saturation

The collector current in a transistor is proportional to the base current, increase the base current and the collector current will increase but there is a limit to this increase. Figure 16 shows a test circuit with the theoretical collector current and power dissipation shown in figure 17. If we assume the transistor has a current gain of 100 and the supply voltage is 12V then increase the base current and the collector current will increase and the voltage across the 1k collector resistor will increase and hence the transistor collector voltage will decrease until it reaches zero. Increasing the base current beyond this point will not result in any change of collector voltage. The transistor is then said to be saturated. In practice the collector voltage does not reach zero but sits at a few 10s of mV. Figure 17 shows how the collector power dissipation varies, reaching a maximum when the voltage across the transistor is half the supply.

This arrangement is most commonly used to drive relays, lamps or LEDs. It is usual to over drive the base, or provide more current than is necessary, to take into account any variation of current gain and ensure the transistor saturates.

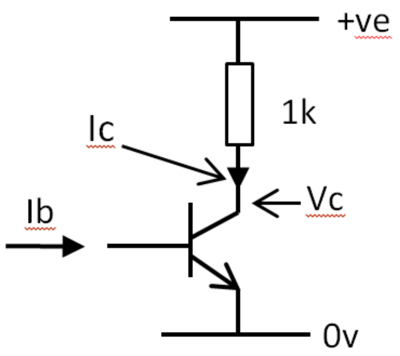


Figure 16 - Test circuit for saturation

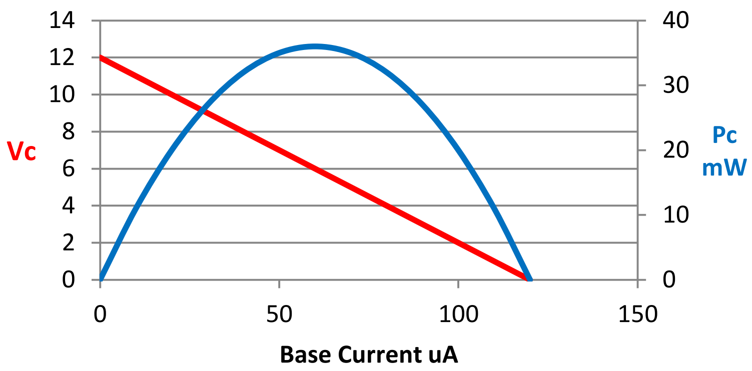


Figure 17 - Collector voltage and transistor power dissipation vs base current.

One application where a transistor is used as a switch which has to saturate is the Line Output Transistor in CRT TVs. In a line output stage, both valve and transistor, the active device is not used as a linear amplifier, as in the field output stage, but as a switch relying on the properties of the line output transformer to generate the line deflection coil current. However there are several considerations to be taken into account when using a transistor in the line output stage.

One is the time taken to turn off the transistor. A transistor base collector junction has capacitance and when the transistor is saturated, charge is accumulated in this capacitance. When the base drive current is removed this capacitance keeps the transistor turned “on” as it discharges. As it does so the transistor comes out of saturation and the collector voltage rises. If we look at the power dissipation plot in figure 17 we can see that when the base current is zero the transistor is “off” and the power dissipation is zero. When the transistor is turned on the collector voltage is low, virtually zero, so in spite of the high current the power dissipation is very low. However in between the power dissipation rises to a maximum when the voltage across the transistor is half the supply. In a line output transistor these conditions occur as the transistor is switched off. The longer it takes to turn off, the longer it is in the “in between stage” and the higher the power dissipation will be. To speed up the transistor turn off, the stored charge in the base collector has to be removed quickly. This can be achieved by applying a negative going spike to the base, assuming it’s an NPN transistor, and is often achieved by means of an inductor in series with the base. Some early transistor logic circuits achieved this speed up by connecting a resistor from the base to a negative supply.

Darlington Transistors

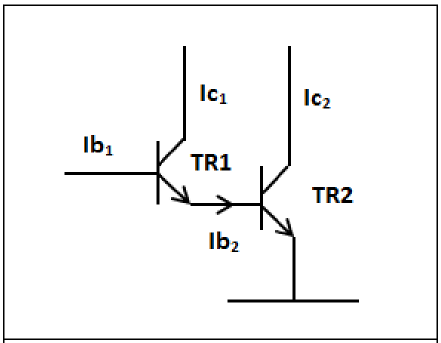


Figure 18 - Darlington Transistor

The collector current is the base current multiplied by the current gain, typically around 100 to 200. However in some applications a higher current gain is needed. This can be achieved by adding a second transistor as shown in figure 18. The collector current of TR2, Ic_2, is its base current, Ib_2, multiplied by its current gain. Similarly the collector current of TR1, Ic_1, is its base current, Ib_2, multiplied by its current gain. But Ib_2 is virtually the same as Ic_1 therefore the collector of TR2 is approximately the base current of TR1 multiplied by the current gain of TR1 and the current gain of TR2. If the current gain of TR1 and TR2 were both 100 then the combined current gain would be around 10,000.

This combination is known as a darlington transistor. It can be made from a couple of discrete transistors but is also available as a single component. Figure 19 shows the internal configuration of a TIP121 darlington power transistor. Here the collectors of both transistors are connected together to form a convenient three pin package. This transistor has a minimum current gain of 1000 at a collector current of 3A whereas a TIP31 power transistor has a minimum current gain of 10 at a collector current of 3A.

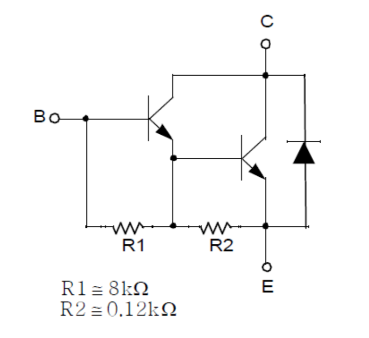


Figure 19 - TIP121 Darlington Transistor

You may think that adding extra transistors to form a triple or quadruple darlington to give an increased current gain would be an advantage but any leakage current in the first transistor would be multiplied by the current gains of all the following transistors. This limits the number of transistors that can be cascaded in this way. It would not be advisable to use multiple germanium transistors in this way because of their higher leakage current.

One disadvantage of the darlington transistor, such as the TIP121, is that the saturation voltage is higher than a standard transistor as the collector voltage has to be enough to provide sufficient drive to the base of the second transistor.

Field Effect Transistors

The Field Effect Transistor actually predates the junction transistor by a couple of decades being patented in the 1920s. However the practical implementation had to wait until suitable semiconductor manufacturing techniques had been developed. Although practical devices were developed in the 1950s it was in the 1960s that they became widely available, one popular device was the 2N3819 used in many projects in magazines in the late 1960s.

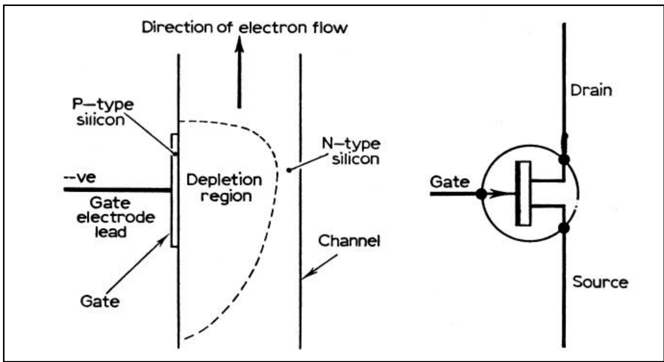


Figure 20 – N type depletion mode FET

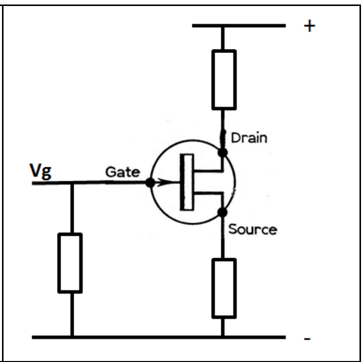


Figure 21 – FET biasing

The electrodes on a FET are the Source, Gate and Drain equivalent to the Emitter, Base and collector of a junction transistor or the Cathode, Grid and Anode of a valve respectively.

There are two types of FET, depletion and enhancement mode and both are available as N type and P type. Figure 20 shows an N type depletion mode FET. Here the source is connected to the negative supply and with zero volts on the gate a current will flow from the source to drain. If a negative voltage is applied to the gate a depletion region is formed which reduces the current flow. No current flows in the gate meaning the FET is very similar a valve. Figure 21 shows a typical biasing arrangement which is very similar to a valve and works in a similar manner.

An enhancement mode FET operates in the opposite manner to the depletion mode. With zero volts on the gate no current flows between the source and drain. If a positive voltage is applied to the gate of an N type enhancement mode FET a current will flow between the source and drain. This makes it suitable for switching circuits similar to a bipolar transistor but with a very low drive current.

P type depletion mode and enhancement mode FETs work in a similar manner but with the polarities reversed.

The FET described above is a junction type but there is another type with an insulating layer between the gate and the body of the FET. This is a MOSFET (Metal Oxide Semiconductor Field Effect Transistor). The most common use is as enhancement types in modern processor ICs which we'll come to later.

Most of the early FETs were low power signal types but in the 1970s FETs with higher current ratings were developed. These are now the preferred devices in applications such as switch mode power supplies as they are available with voltages and current ratings of up to 800V and 15A.

MOSFETs are also fairly easy to drive although care has to be taken when switching them off due to the gate capacitance. Some devices have significant gate capacitance which can hold enough charge to keep the FET turned on when the drive is removed. Ideally they need to be driven from a driver that can both source and sink current to charge and discharge the gate capacitance as quickly as possible to reduce the switching times and the power dissipation as shown in figure 22. For applications that require fast switching specialised driver ICs are available.

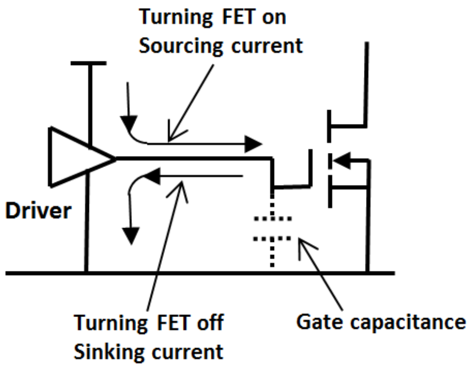


Figure 22 - Driving a MOSFET switch

Figures 23 to 25 show the characteristics of a P channel depletion mode FET and an N channel Enhancement mode power FET.

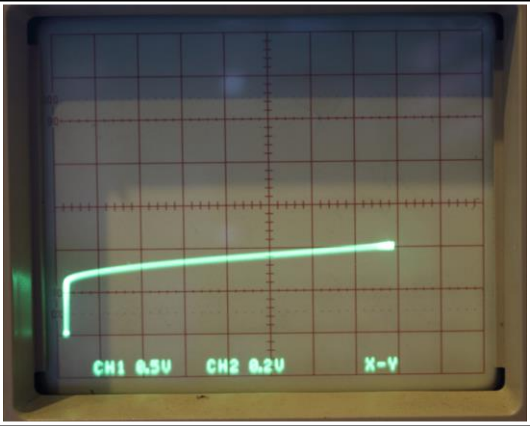


Figure 23 – J174, Vg = 3V

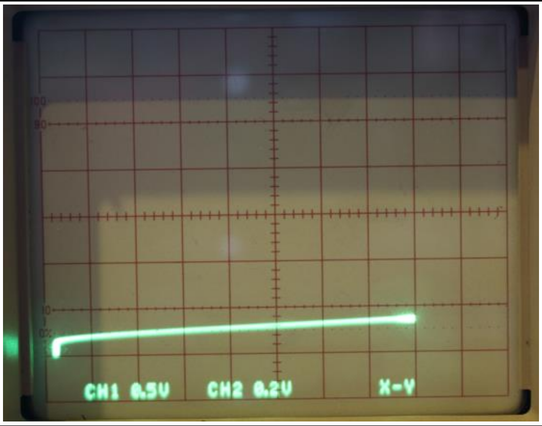


Figure 24 – J174, Vg = 4V

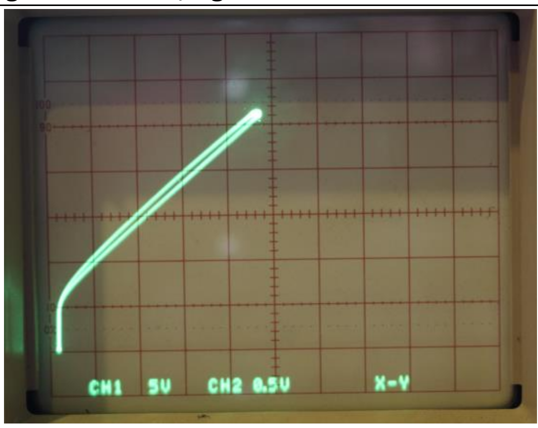


Figure 25 - IRF540, Vg = 3.4V

Figures 23 & 24 – J174 P channel depletion mode FET
X = 2V/div, Y = 5mA/div

Figure 25 - IRF540 N channel enhancement Mode power MOSFET
X = 5V/div, Y = 50mA/div

Note - The gate voltage polarities for these two FETs are the same but are appropriate as one is a P channel depletion mode and the other an N channel enhancement mode.

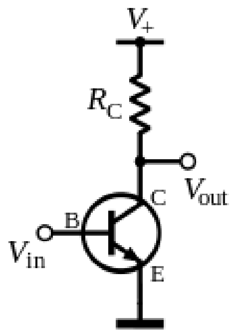
The power Mosfet has a lower "on" resistance as it is intended for use in switch mode power supplies meaning the currents are on the limit for the simple curve tracer used.

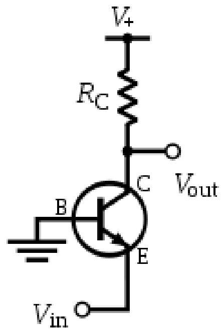
Circuit Configurations

There are three basic configurations for a transistor amplifier depending on which electrode is used for the signal input and which is used for the signal output. Figures 26 to 28 show the configurations for junction transistors and their basic characteristics. Note that no biasing components are shown.

In the Common Emitter configuration the signal input is to the base with the output taken from the collector. This is the most common arrangement for a transistor amplifier. In the Common Base configuration the signal is applied to the emitter with the output from the collector. This is most often seen in RF amplifiers. The final configuration is the Common Collector, or Emitter Follower. Here the input is to the base with the output taken from the emitter. This is most commonly used to provide a low impedance output for an amplifier.

The equivalent configurations for FETs are Common Source, Common Gate and Common Drain. These and the junction transistor configurations are equivalent to the valve Common Cathode, Common Grid and Cathode Follower respectively.





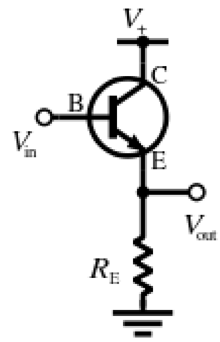


Figure 26 - Common Emitter	Figure 27 - Common Base	Figure 28 - Common Collector
Input impedance - Medium	Input impedance - Low	Input impedance - High
Output impedance- Medium	Output impedance - High	Output impedance - Low
Voltage gain- Medium	Voltage gain - High	Voltage gain -1
Current gain- Medium	Current gain - 1	Current gain - High

Unijunction Transistor

The unijunction transistor was a by-product of research into germanium tetrode transistors. It comprises a bar of lightly doped N type silicon with contacts at each end (Base 1 and Base 2) and a heavily doped P type junction part way along the bar as shown in figure 29.

The B1 to B2 connection is resistive and symmetrical meaning it reads the same value when measured with a multi-meter whichever way round the meter leads are connected.

Figure 30 shows the emitter voltage vs current characteristic. This has a negative resistance section making it suitable as an oscillator. The typical oscillator circuit, showing the values used on the demo board, is shown in figure 31. This was used in many applications in the 1960s and 70s, as it uses only five components.

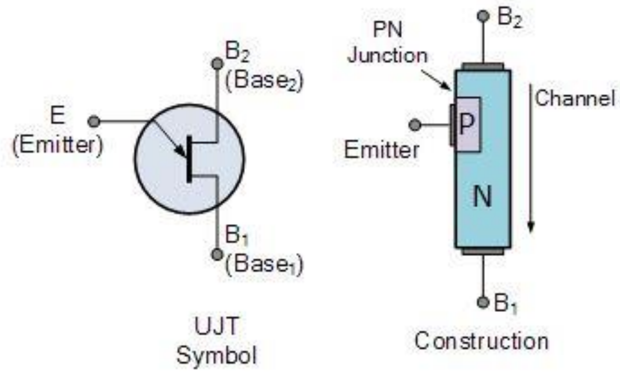
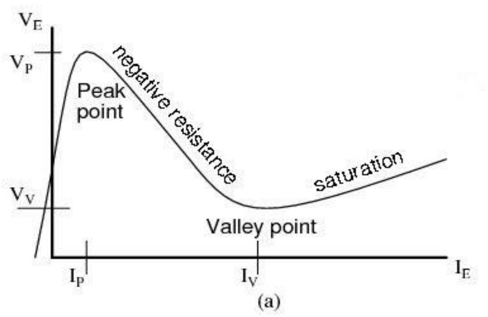


Figure 29 - Unijunction transistor symbol and construction

Figures 32 and 33 show the waveforms for the demo board oscillator.



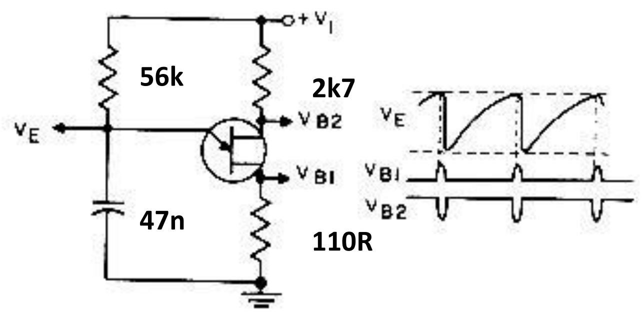
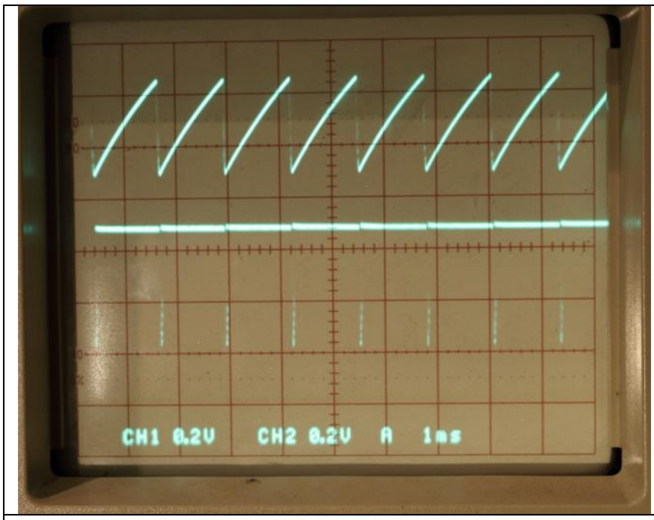


Figure 30 - Unijunction Emitter voltage vs Emitter current	Figure 31 - Circuit of Demo Board



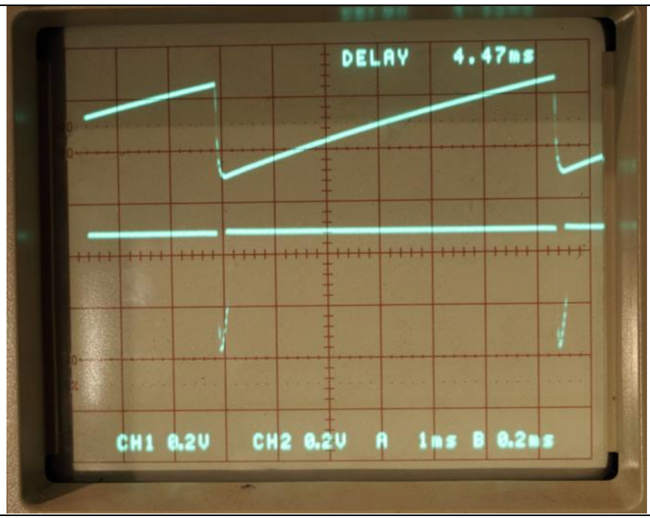


Figure 32 – Emitter and B2 Waveforms	Figure 33 – Close up of Emitter and B2 Waveform

Thyristor or Silicon Controlled Rectifier (SCR)

This is the semiconductor equivalent of the Thyratron. It comprises of four alternating semiconductor layers of P and N type silicon with three PN junctions equivalent to a PNP and an NPN transistor as shown in figure 34.

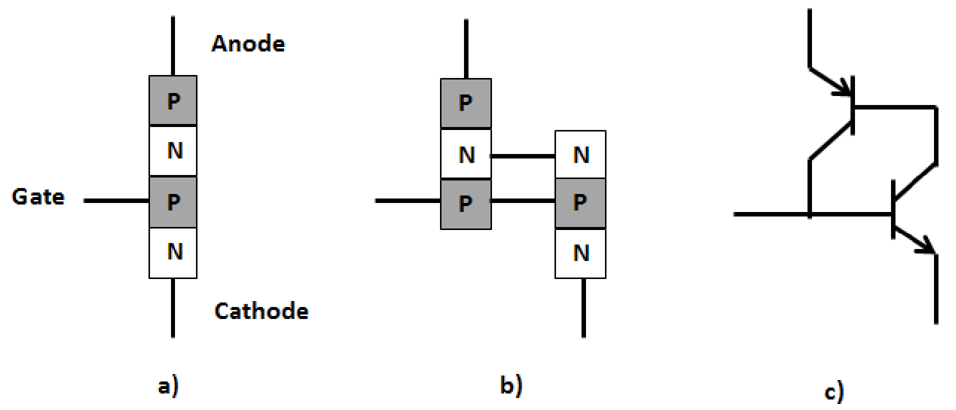


Figure 34 - Thyristor Construction (a), Equivalent Structure (b) and Equivalent Circuit (c)

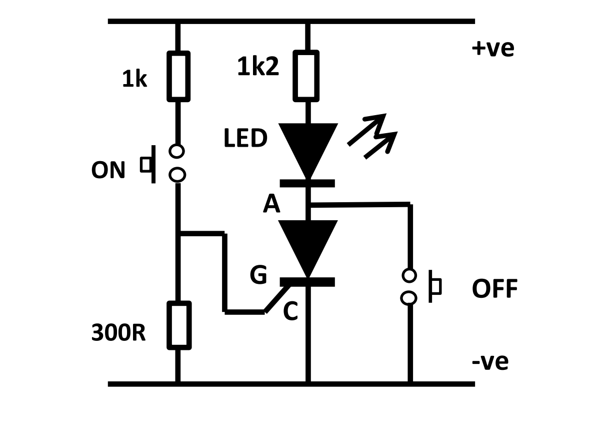


Figure 35 - Thyristor Demo Board

Figure 35 shows the demo board circuit. The anode load is an LED, to indicate when the thyristor is “on”. The gate is connected to ground by the 300R resistor and can be connected to the +ve supply via the ON pushbutton and the 1k resistor.

Looking at the equivalent circuit in figure 34 when power is applied there is no base current into the lower NPN transistor from the gate connection, therefore there is no collector current and no base current into the upper PNP transistor.

If we now supply base current to the NPN transistor by pressing the “ON” button what happens. Current flows in the collector of the NPN transistor causing current to flow in the base of the PNP transistor which causes current to flow into the base of the NPN transistor. Both transistors saturate and the voltage across the pair is reduced to a low level. This situation is self-sustaining so the external gate current can be removed and the device is turned “ON”. There is a minimum current to sustain this condition called the holding current.

This self-sustaining situation will continue indefinitely unless the current in both transistors is reduced to the point where the current is insufficient to maintain both “transistors” in the saturated state. This can be achieved in one of three ways

1 Reduce the voltage across the device until the current through the device is lower than the holding current.

2 Short circuit the device which reduces the current to zero.

3 Reverse the polarity of the supply which will cause the transistors to turn off.

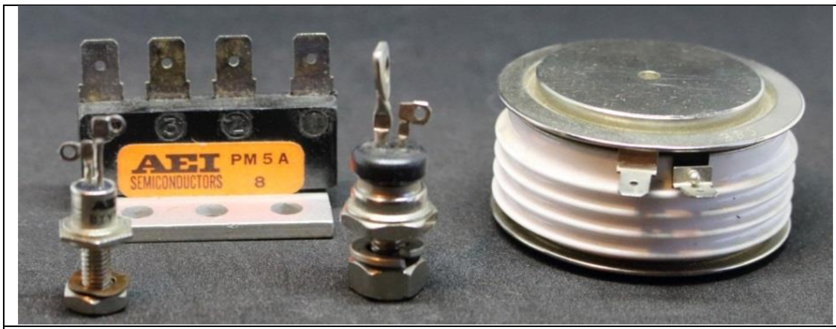


Figure 36 - A Selection of SCRs

One of the main uses for thyristors is to control an AC supply. In this application the thyristor will be able to conduct on the positive half cycle and will be turned off during the negative half cycle. By controlling the point in the positive half cycle at which the thyristor is triggered the power can be controlled. This was used in the regulated power supplies of many solid state TVs.

Triac

The triac is effectively two thyristors connected in parallel so that one will conduct on the positive half cycle with the other conducting on the negative half cycle. This allows better control of power to the load and by conducting on both half cycles of the mains waveform does not put a dc component onto the mains supply. There is a short time around the zero crossing of the mains waveform where the triac is off as there is insufficient current through the device to turn it on.

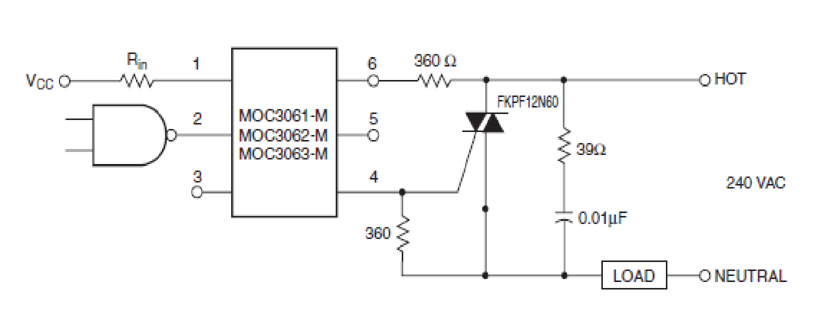


Figure 37 - Opto Isolated Triac Application Circuit

Triacs intended for switching the mains can be obtained with zero crossing detector which ensures the triac turns on and off around the zero crossing of the mains waveform. This can reduce the amount of RF interference generated due to the triac switching on at a peak of the mains waveform. Often these are combined with an opto-isolator allowing a simple logic level to switch an AC supply on and off. Figure 37 shows a typical application circuit taken from the MOC3061 data sheet. This uses the triac in the opto-isolator to trigger another, higher current, triac to switch power to the load.

Integrated Circuits

In September 1958 Jack Kilby, working for Texas Instruments, successfully demonstrated a working integrated circuit. In the patent application he described his new device as "a body of semiconductor material wherein all the components of the electronic circuit are completely integrated" shown in figure 38. His original circuit was made from germanium but a later, more practical, implementation was made from silicon by Robert Noyce of Fairchild Semiconductor. This incorporated the principle of p-n junction isolation which allowed the transistors to operate independently in spite of being on the same piece of silicon.

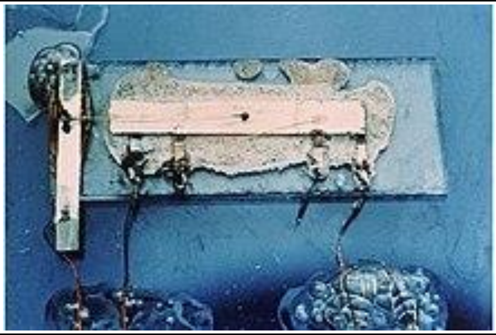


Figure 38 – The First IC

Over the next few years the technology improved, allowing more transistors to be fitted on an integrated circuit. This improvement became known as Moore's Law, after Gordon Moore of Fairchild Semiconductor. In a paper of 1965 he stated that the number of devices on an integrated circuit would double every year although this doubling has been revised to anywhere between 18 months and 2 years.

The basic manufacturing process of an IC is to use photographic masks to introduce the requisite N and P type impurities into a silicon die to create the transistors in a similar way to the silicon planar transistor. Looking at the equivalent circuit of an IC you will see very few resistors as these can take up more space than a transistor. Transistors with modified characteristics are often used instead of resistors.

Analogue ICs

One of the earliest type of IC was the analogue operational amplifier. The operational amplifier was used extensively in analogue computers with the early types using valves and later transistors. These often needed specialised techniques to remain stable. Integrating all the components onto one piece of silicon allowed a smaller and more stable amplifier to be made. Fairchild Semiconductor introduced the first op amp, the µA702, in 1963. Although it required odd supply voltages and had its limitations it did point the way for future op amps. The µA709 followed, then the LM101, each being an improvement on previous ICs. These op amps required external components for frequency compensation which kept the IC stable, the later ICs requiring fewer compensation components.

In 1968 probably the most famous, or should it be infamous, op amp was introduced, the 741. This had internal compensation requiring no external components to keep it stable. This became one of the most commonly used op amps for many years. Improvements were made to the design of op amps were made over the following years making op amps more suitable for low noise, higher frequency etc. applications. Dual and quad op amps were also introduced.

Many other types of IC were introduced to perform specific functions such as audio power amplifiers, radio IF amplifiers, FM stereo decoders. In the late 60s ICs stared to appear in TVs starting with the intercarrier sound IF and demodulator then the PAL decoder. Ultimately many functions were integrated into a single IC making it possible to make a colour TV with only a few ICs on a small PCB.

Voltage regulators were introduced, one of the most well known is the 723 which includes the voltage reference, error amplifier, current limiting all in one IC. It can be used stand alone as a 150mA regulator but with an external output transistor or transistors it can provide a regulated output with a maximum current of several amps.

With the introduction of digital ICs fixed voltage regulators to provide the necessary supply voltages were introduced. The 7805 supplied 5V at 1A with other variants providing 12V, 15V, 24V.

Not all the ICs introduced were successful and many became obsolete as better ICs were developed or the demand from equipment manufacturers dropped. This can cause problems when equipment using these ICs fails and spares are not available.

Figures 39 to 42 show a selection of analogue ICs in a variety of cases. Note DIL is Dual In Line where there is a row of pins either side of the IC package.

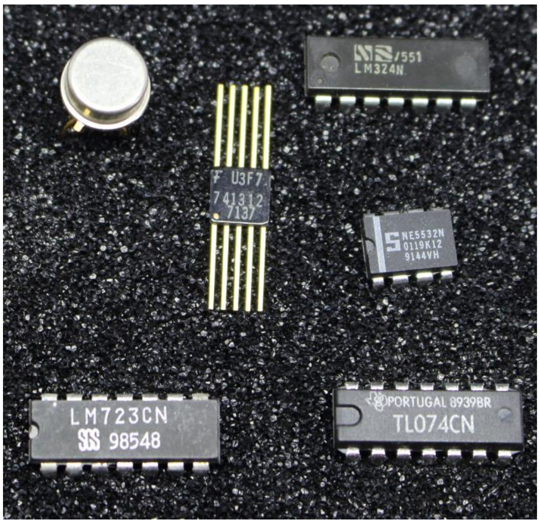


Figure 39 - A Selection of Analogue ICs

LM101 op amp in TO100 case

741 op amp in Flat pack

LM324 quad op amp in 14 pin DIL case

NE5532 dual op amp in 8 pin DIL case

LM723 PSU regulator

TL074 quad op amp in 14 pin DIL case

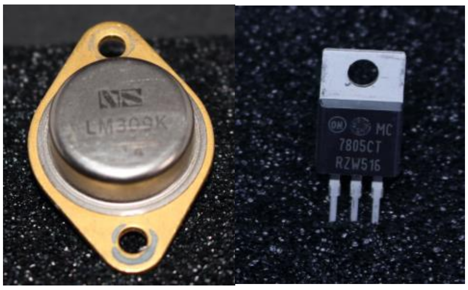


Figure 40 - Two 5V 1A Regulators

LM309 in TO3 case dating from the early 1970s

MC7805 in TO220 case dating from 2000s

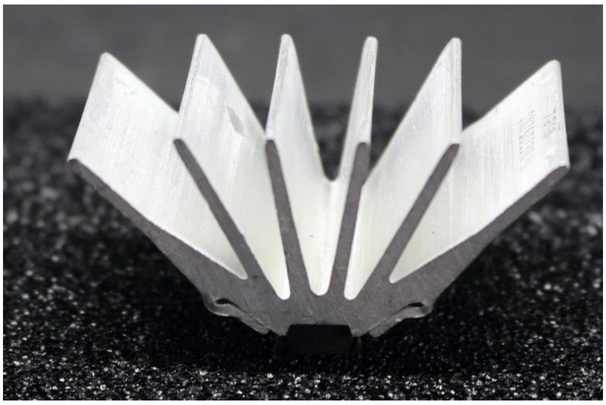


Figure 41 - Texas Instruments SN76013 audio amplifier

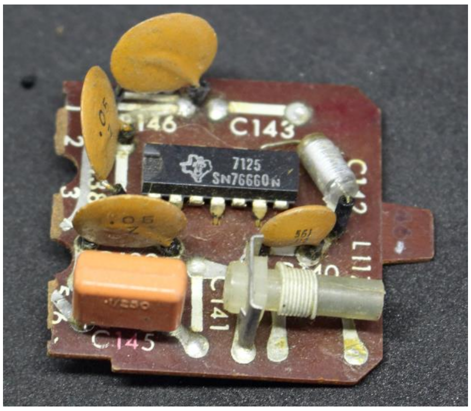
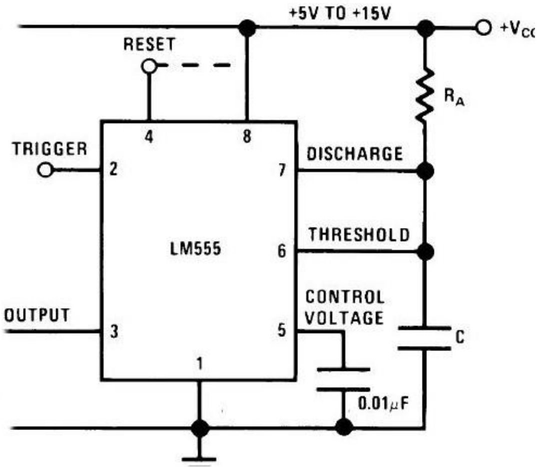


Figure 42 - SN76660 TV intercarrier sound IC

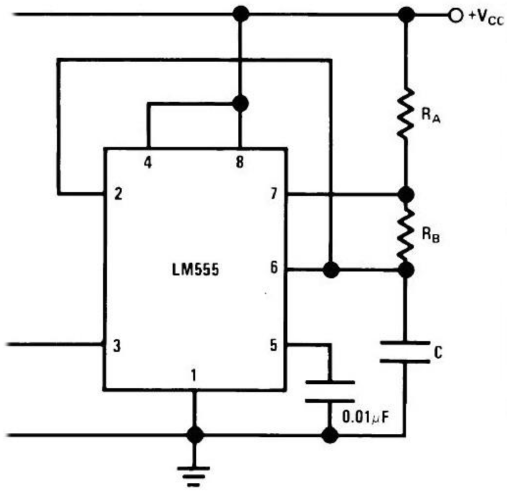
Some ICs, however, become classics. One of these is the 555 timer, introduced in 1972. This versatile IC can be used as an oscillator or a monostable. It has the advantage that the timings are independent of the supply voltage, which can be between 5V and 15V, and are predictable using very simple formulae and the output pin can sink or source current. Figures 43 and 44 show the two modes for the 555.

However there is design flaw in the 555. When the output pin changes state both the pull up and pull down output transistors are on simultaneously for a short time which causes a current spike from the supply. This can cause problems to any sensitive circuits on the same supply. The solution is to fit a decoupling capacitor across the supply close to the IC. Later a CMOS version was introduced which, apart from having lower power consumption, did not have this current spike problem.

There is a dual version of the 555, the 556, available in both bipolar and CMOS versions but it is important to decouple the supply, especially for the bipolar, version to prevent the current spike from one timer triggering the other timer.



$T = 1.1xR_AxC$



F = $\frac{1.44}{(R_A + 2R_B)xC}$

Figure 43 - 555 Monostable | **Figure 44 - 555 Astable**

Digital ICs

Concurrent with the development of analogue ICs digital ICs were developed. Early digital computers used valves which were soon replaced once reliable transistors became available. The advantages were smaller size and significantly reduced power consumption.

Once ICs had been invented it was inevitable that the technology would be applied to create digital ICs. Initially they were copies of the transistor circuits and were known a RTL (Resistor Transistor Logic) as shown in figure 45.

However there were disadvantages of the circuit. It has limited output drive capabilities and when the transistor is "on" the power dissipation is high. With the circuit shown there is a limit of three to the number of inputs but it is possible with a different input configuration to increase this to 8.

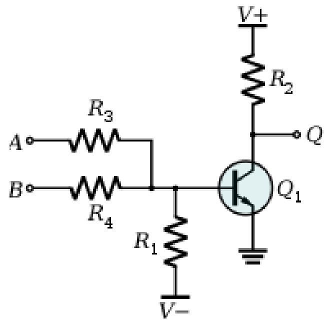


Figure 45 - RTL NOR gate

RTL gates were used extensively in the Apollo guidance computer.

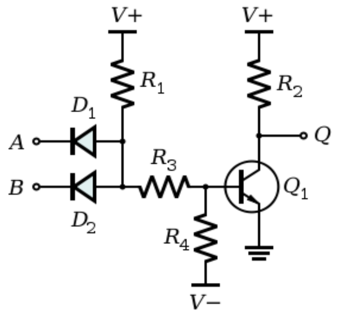
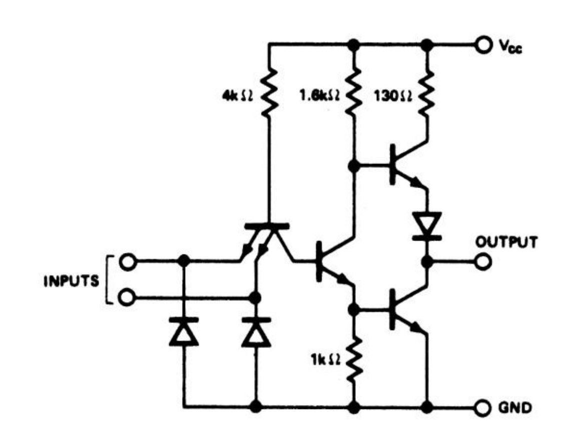


Figure 46 - DTL NAND gate

Following on from RTL the next development was Diode Transistor logic (DTL). This has a similar configuration to RTL but with the input resistors replaced by diodes with a pullup resistor as shown in figure 46. This allows more inputs than RTL and has a higher immunity to noise. The resistor R3 can be replaced by two diodes in series which removes the need for the negative supply. This also makes for a smaller IC as diodes take up less space than resistors in an IC.

However the speed of the DTL gate is very similar to the RTL gate.

The next development was to incorporate the input diodes into one device to create one of the most popular logic families, Transistor Transistor Logic (TTL). Figure 47 shows the circuit of a 7400 NAND gate. The multi emitter input transistor configuration takes up less area on the IC than the equivalent input of a DTL gate. The output stage also changed to one that can both actively sink and source current by replacing the collector resistor (R2 in figure 46) by a transistor to create the Totem pole output. All these changes significantly increased the speed resulting in TTL becoming the dominate logic family.



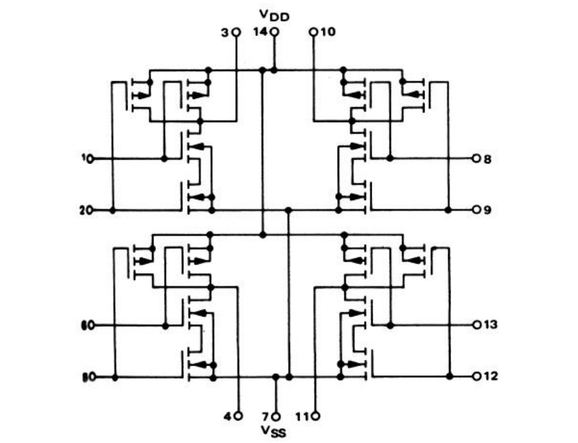


Figure 47 - Equivalent circuit of one gate in TTL 7400 quad two input NAND gate	Figure 48 - Equivalent circuit of CMOS MC14011 quad two input NAND gate

The logic families we have seen so far use bipolar transistors and consume significant current. The more complex the IC and the faster it runs the more current it consumes and consequently the more power it dissipates. In the late 1960s RCA introduced the CMOS logic family that used complimentary MOSFETs. Although it was slower than any of the other current logic families it had the advantage that it consumed very little current because either the upper transistors are “on” or the lower transistors are “on”. Only when the transistor switches does it consume any current. It could also run from a supply voltage between 3 and 15V compared to 5V for TTL. Figure 48 shows the circuit of a CMOS gate.

The input impedance of a CMOS gate was also very high which can cause problems if an input is left open circuit. With TTL if an input is left open circuit it assumes a logic 1 state although it is usually recommended to fit a pull up resistor to the positive supply. With an open circuit CMOS input, because of the high input impedance, any static charge can cause it to change state. Sometimes waving a hand near an unconnected input can induce enough static charge to cause a change of state. This often caused problems in CMOS circuits so it is recommended to use either tie unused inputs to ground or to the positive supply via a pull up resistor.

The standard family of TTL was the 74 series from Texas Instruments. Other pin compatible families were introduced, the 74L series which had lower power consumption but lower speed, the 74S series which incorporated Schottky diodes to give a higher speed but at increased power consumption and the 74LS which had the schottky diodes and the low power but had speeds similar to ordinary 74 series logic but with lower power consumption.

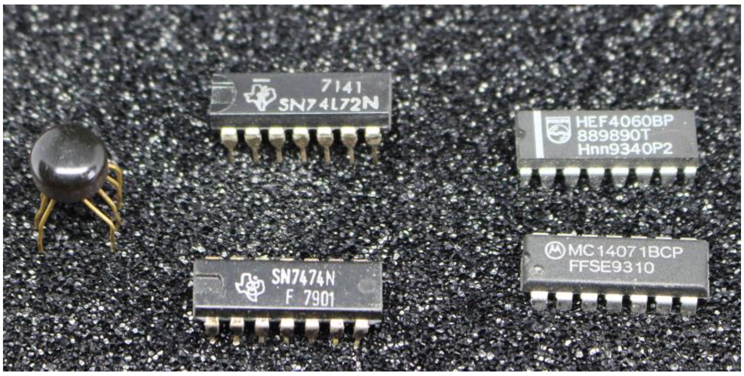


Figure 49 - Logic ICs, left to right – RTL flip-flop, TTL 74L72 & 7474, CMOS HEF4060 & MC14701

CMOS also had variant families as other manufacturers started producing their own versions. Mullard had the HEF series, then faster CMOS families, such as HC and HCT, were introduced which were functionally and pin compatible with TTL. Further developments brought about faster CMOS devices which became the favoured parts to use. Many of the early TTL devices are now obsolete.

In the logic families so far described the transistors are saturated which affects the speed at which they can switch because of the charge stored in the collector base junction. There is one family of logic where the transistors are never saturated. This is Emitter Coupled Logic (ECL). This was an early entrant into the logic families. Because the transistors never saturated they could switch at a higher speed than the saturating logic families. However this speed was achieved at the expense of higher power consumption.

The table below gives a comparison between some of the logic families.

Logic Family	Typical Speed	Typical Power per Gate @ 1MHz	Approximate Date of Introduction
RTL	4MHz	10mW	1963
DTL	4MHz	10mW	1962
74 series TTL	25MHz	10mW	1964
74L series TTL	3MHz	1mW	1964
74S series TTL	100MHz	19mW	1969
74LS series TTL	40MHz	2mW	1976
4000 series CMOS	5MHz	1.2mW*	1970
74HC series CMOS	50MHz	0.5mW	1982
ECL (III)	500MHz	60mW	1968

*When used for slow speed logic, CMOS consumes significantly less power.

Microprocessors and Beyond.

In the late 1960s a Japanese calculator company, Busicom, approached a small American company, that was producing memory ICs, to design a calculator IC for them. The American company decided that rather than design a specific IC they would produce an IC that could be programmed to perform the required function. That company was Intel and the IC they developed turned into the world's first microprocessor the 4 bit 4004 introduced in 1971.

The 4004 was the CPU (Central Processing Unit) and required additional support ICs to provide memory and Input and Output functions but it could be programmed to perform virtually any function.

Intel then introduced an 8 bit processor, the 8008, which led to the 8080, introduced in 1974. Other manufacturers began to introduce their own 8 bit processors, Motorola with its 6800, Zilog with its Z80, MOS Technology with its 6502 and RCA with its CMOS 1802. Because the 1802 used CMOS technology it was very low power and a radiation hardened version was used in various space probes.

Throughout the 70s and 80s faster processors and the peripheral ICs to go with them were introduced. Memory devices included RAM (Random Access Memory) to store temporary or changing data and EPROMs (Electrically Programmable Read Only Memory) to store the program data.

In the late 70s and early 80s personal computers began to appear. Many different types were introduced and were generally incompatible with one another. In the mid 80s IBM decided they would like to get into this growing market and designed their own personal computer. They needed an operating system to run it and approached a small company who wrote an operating system for them. That company was Microsoft.

Since then IBM compatible PCs have become faster and more powerful. Who can remember the quote by Bill Gates that 640k of RAM is enough? The PC I'm writing this on has a mere 8Gb of RAM (8 times the capacity of my first PCs hard drive!).

The processors used developed from 16 bits used in the early PCs through 32 bits to 64 bit processors now being required for modern operating systems. The speeds have also increased from around 8MHz for early systems to several GHz for the latest system. This doesn't mean there is a clock running at these GHz speeds being fed into the processor but there is an external clock of a much lower speed with an internal clock generator deriving the higher clock speeds internally to the processor.

Early processors had standard dual in line packages with typically 40 pins. This was enough for the data, address and control lines but modern processors have several hundred pins. Figure 50 shows an Intel Pentium 3 dating from around 2000 with 451 pins. The bluish area on the topside is the actual silicon IC fixed to a fibre glass PCB. The underside not only incorporates the pins but also has several decoupling capacitors.

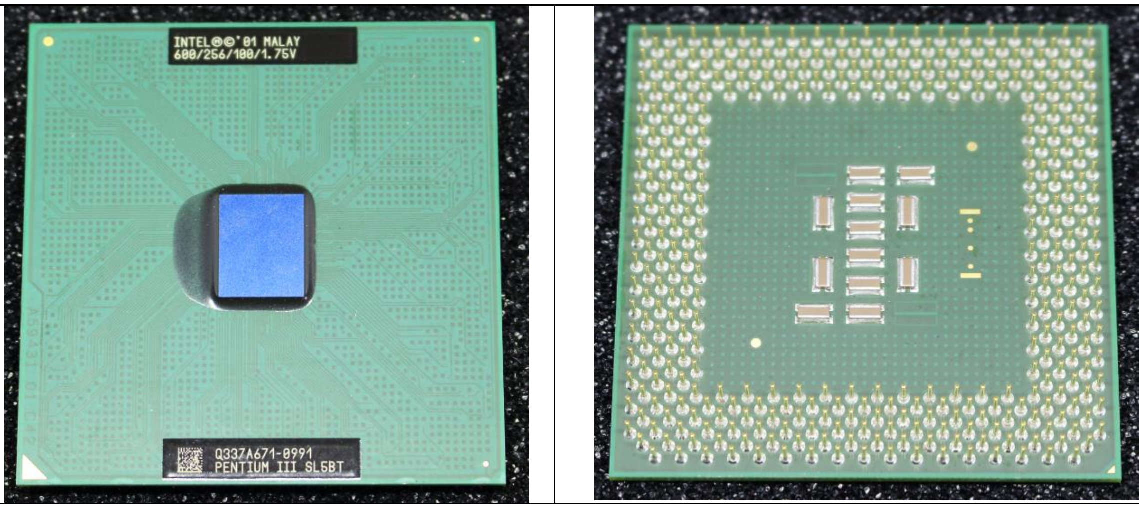


Figure 50 - Intel Pentium 3 Processor Top side view showing the actual processor chip and Underside view showing its 451 pins and decoupling capacitors

The early microprocessors had about 3000 transistors whereas the latest processors have several million. These transistors are MOSFETs as these allow a lower power consumption than bipolar types. To increase the speed of the processor the devices have to be made smaller. However this decreases the breakdown voltage of the devices so the supply voltage to the core of the processor is usually run at around 1V. The numbers of transistors used mean that the supply current is high. The P3 processor in figure 50 has a supply current for the processor core of up to 12A at 1.75V but some processors having a supply current of up to 30A for the processor core.

The program for a typical microprocessor system, not a PC, is normally held in a ROM (Read Only Memory). This would often be a mask programmed memory once the program has been proven and fully tested. During the program development the program would be programmed into an EPROM which could be erased by exposing it to UV light and then reprogrammed. Figure 51 shows a couple of EPROMs with the actual memory IC visible through the transparent cover. Some applications used an EPROM for the final program as creating a mask programmed ROM could be very expensive.

Microcontrollers

A microcontroller is effectively a microprocessor system on a single IC. By incorporating all the peripherals on a single IC, microprocessors to be used for many more applications. Early microcontrollers often had EPROM program memory, as shown in Figure 51 and 52, which required erasing with UV light before reprogramming could take place. Later microcontrollers used Flash memory which can be erased electrically before reprogramming. Figure 52 shows a selection of PIC microcontrollers from Microchip.

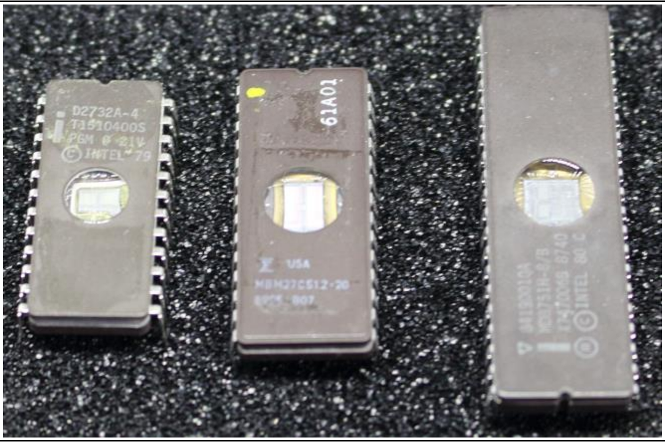


Figure 51 - UV erasable EPROMS and microcontroller

Figure 52 - A selection of PIC microcontrollers

These days there is hardly a device that doesn't have some form of microcontroller, mobile phones, digital cameras and many kids toys would not be possible without them.

References

https://en.wikipedia.org/wiki/Integrated_circuit

Data sheets for Intel P3 processor, 555 timer,MOC3061.