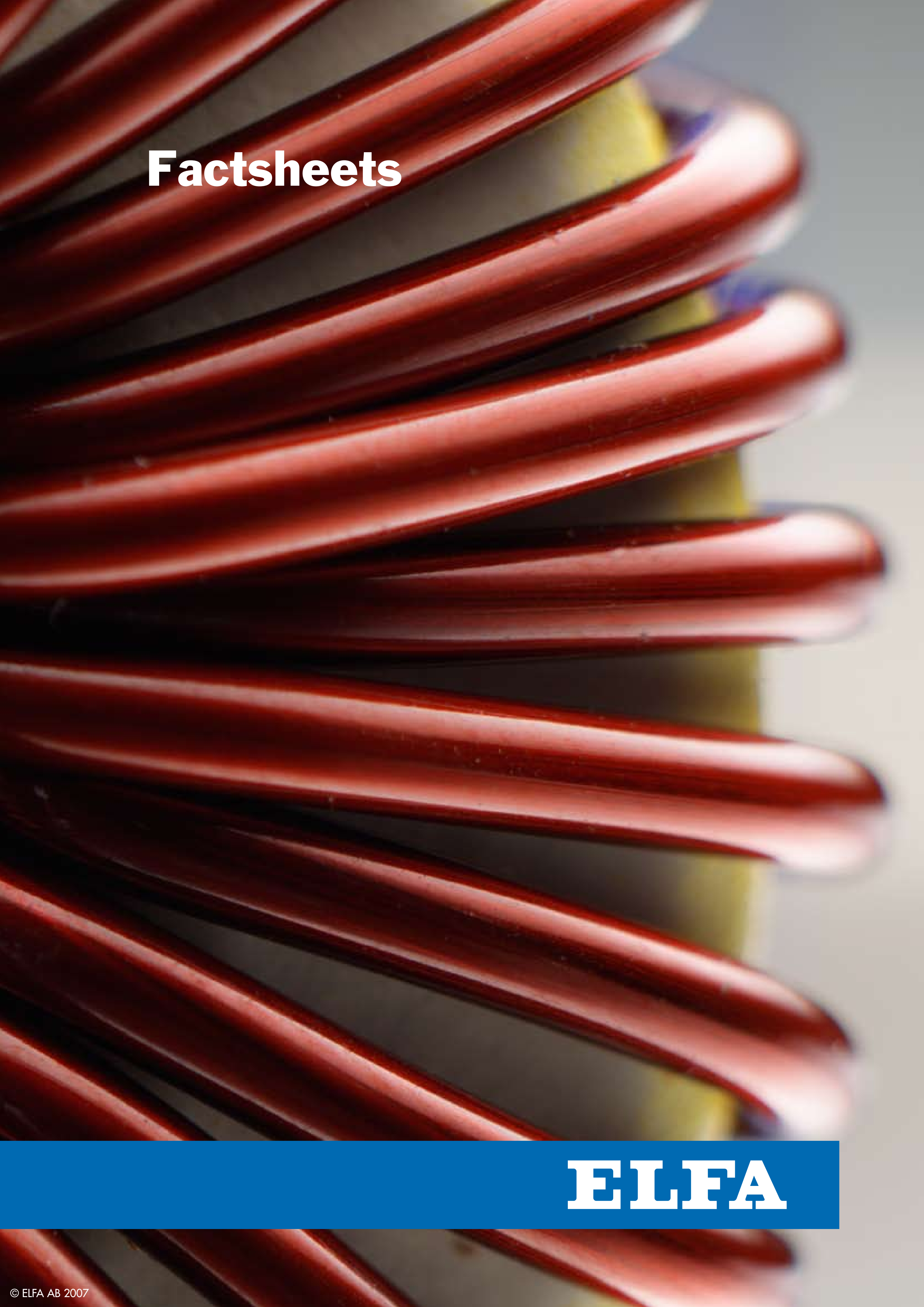
Factsheets



ELFA

Fuses

The fuse is a security and protection component that cuts the power. More commonly it means a component that senses the current consumption in a circuit and cuts the power if the consumption gets to large, like at short-circuits or overload.

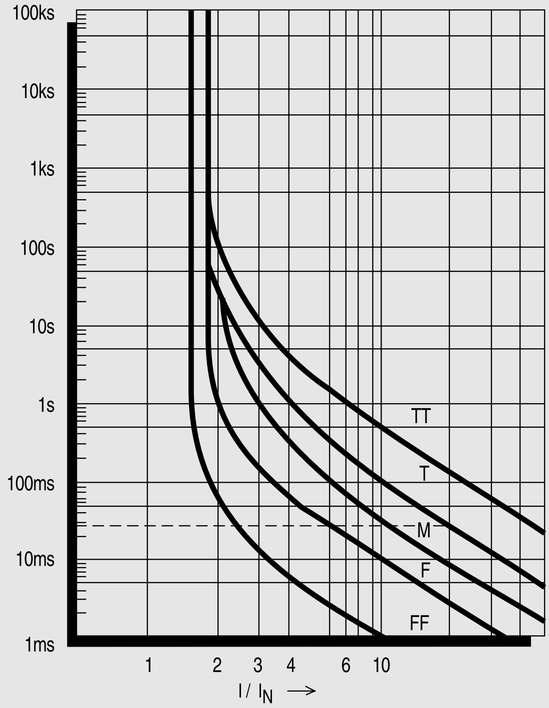
Data

The rated voltage is the greatest extended working voltage and the type of voltage (AC voltage or DC voltage) at which the fuse may be used.

The current rating is the working current at which the fuse is designed to be used. The current rating is somewhat lower than the current which is able to flow for an extended period without tripping the fuse. The difference between these two currents varies for different standards (e.g. CSA, IEC, Miti, UL).

The breaking characteristics describe the correlation between how quickly the fuse trips and how high the current is. The main groups are fast and delayed action fuses. Fast fuses are used in particular cases when a fuse which trips as quickly as possible is required, e.g. in an instrument input. Sometimes, these fuses are also necessary from the point of view of safety. It is necessary to use delayed action fuses when the load exhibits high current during startup, e.g. when a motor is turned on. Transformers also give increased inrush current, and this applies to a particularly great extent to toroidal core transformers.

There are standardised characteristics. For fuses to IEC, there is FF (very fast), F (fast), M (intermediate), T (delayed action), TT (very delayed action). For fuses to UL, there is e.g. T-D (delayed action) and D (delayed action). For circuit breakers there is B (fast), C (delayed action) and D (very delayed action).



This figure shows the pre-arching time for different types of fuses. I/In states the raltion between the load current and the rated current of the fuse.

The breaking capacity is the greatest current which the fuse is capable of breaking at a specific voltage without it short-circuiting or reconnect. The specification for the breaking capacity can include, for example, the magnitude of the breaking current, the magnitude and type (AC voltage or DC voltage) of the working voltage, and the cosφ of the load. The breaking capacity must be sufficient for all conditions. In the event of a short-circuit, for example, the entire current which the supply side is able to provide can flow.

Models

Thermofuses are available in many different models.

Glass tube fuses and *ceramic fuses* are the types most commonly used. In European equipment the fuses commonly used are 5 × 20 mm in size, while American equipment uses slightly larger fuses, 6.3 × 32 mm. The ceramic models have greater breaking capacity.

Many variants of special fuses of other sizes and with other properties are available. *Subminiature fuses* can be found, for example, at the input of certain measuring instruments in order to protect them against overloading. They are available in models for mounting in holders or for permanent mounting, both for hole mounting and surface mounting.

Automotive fuses are made in two models: as a ceramic bar 6 × 25 mm and with a metal strip on top which melts, or plastic-enclosed with two parallel blade terminals. The latter type is used in modern cars. The advantage is that the contact with the fuse holder is so much more secure than in the older ceramic fuses, where oxide often caused disruption after a few years of use.

Circuit breakers can be reset and do not need to be replaced after having been tripped. For most applications, the fuse must be designed such that automatic resetting is prevented as long as the overload state continues. Resetting is done manually.

Circuit breakers which operate thermally can be designed so that they have a long service life. They are manufactured for various tripping characteristics. Some fuses are designed with electromagnetic fast tripping for currents which exceed the current rating of the fuse to a high degree. Types without fast tripping are normally delayed action and are therefore suitable for use where there are high inrush currents.

Thermal fuses are influenced by the ambient temperature on account of the way in which they work. The rated value of the fuse is normally given at +20 °C. The manufacturer ETA gives the following conversion factors for its fuses for different ambient temperatures (rated value of the fuse = trigger current x conversion factor):

Amb temp (°C)	–20	0	20	30	40	50	60	70
Conversion factor	0.8	0.9	1.0	1.1	1.2	1.3	1.45	1.65

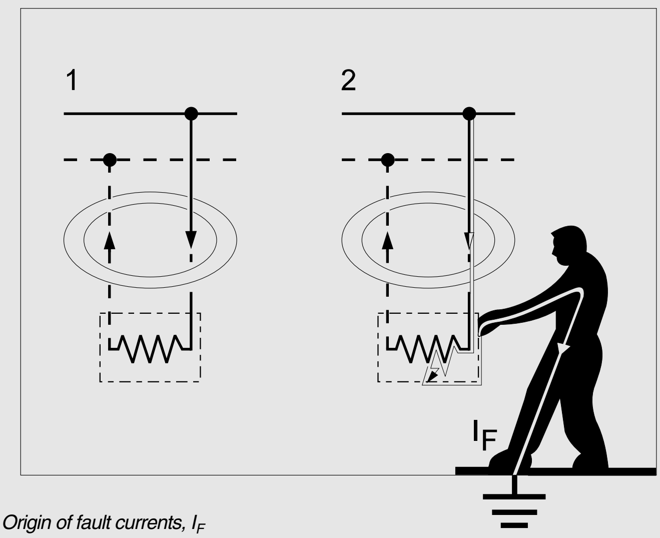
Auto-reset polymer fuses replaces normal glass tube fuses in most low current applications. After having tripped due to surge or excess temperature, they only need to cool down to return to low resistance level. This technology is also used for over-voltage protection. Polymer fuses are manufactured in several versions such as for hole mounting, surface mounting and also a foil version which is particularly suitable for battery packs. They are suitable for protection in motors, transformers, power supplies, loudspeakers, alarms, telephones, test instruments, PCBs etc.

Thermal fuses sense the ambient temperature and break the circuit if the temperature exceeds a certain limit. This makes them suitable for protecting most electrical and electronic equipment from overheating. Thermal fuses can be constructed with a melting body that cuts the power. An other construction uses a bimetal spring that bends from the heat and resets when cooling.

Residual current devices

Fault currents

Fault currents are currents which flow to the system's neutral point via the protective conductor or directly via the ground due to an insulation fault in electrical installations or devices.

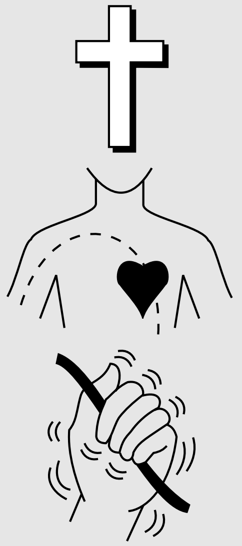


Origin of fault currents, I_F

1. Fault-free current circuit.
2. Faulty current circuit (defective device).

People or animals who touch defective or live parts and are thus traversed by a fault current are exposed to danger of the highest degree.

Fault currents in the form of creeping currents which flow to earth via damp girders, for example, can constitute a fire hazard or contribute to the breakdown of insulation material.



80 mA	*When exposed for more than 0.5 to 1 second – immediate death.*
30 mA	*Irregular heartbeat and increased blood pressure.*
15 mA	*Muscle cramps – if a hand is holding a live object, it cannot be opened.*
0,5 mA	*Not perceptible.*

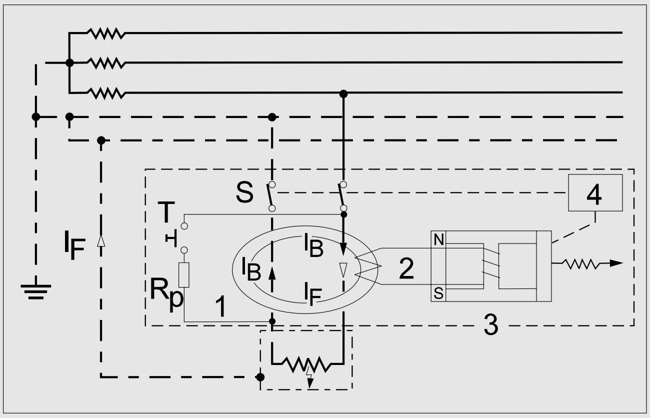
This is how people react to current.

Fault current breaking principle

Kirchhoff's laws form the basis of residual current device technology, according to which the sum of incoming currents is equal to the sum of outgoing currents. In residual current devices, these currents are measured and compared with one another. If the totals of the currents are not the same, i.e. if a fault current has occurred, the defective part is disconnected by means of a trigger device. This takes place quickly and even at small fault currents before people or animals are injured or property damaged.

Design of the residual current device

The essential components of a residual current device are the total current transformer, the magnetic trigger device and the contact system with the contact mechanism. All the phase conductors and the neutral conductor on the incoming side of the object to be protected are fed through the total current transformer (1). See diagram.



The principle of residual devices.

The current through each of these conductors induces a magnetic flux Φ in the total current transformer. In a fault-free device, the incoming and outgoing currents are the same. The magnetic fields formed by these currents (in the total current transformer) therefore cancel out one another.

If the currents in the conductors are different, a magnetic flux occurs in the transformer $(\Phi_B + \Phi_F) - \Phi_B = \Phi_F$, caused by the currents $I_B + I_F$ in the phase conductor and I_B in the neutral conductor respectively.

The magnetic flux Φ_F induces a voltage in the secondary winding (2) which actuates a secondary current through the winding of the magnetic trigger (3).

This current weakens the magnetic field in the magnetic trigger to such an extent that the armature is released and opens the main contacts via the trigger mechanism (4).

In the picture, you can also see the test button (T) which simulates a fault current via the resistor (Rp). This permits the continual testing of the functional safety of the residual current device.

Light and lamps

Our sources of electric lighting are designed such that they convert electrical current into radiation. The link between the luminous efficacy, service life and the electrical output of a light source is thus very important.

Marking of light sources

Like other technical products, these light sources are stamped with important information. The user can then find the right type of light source with the right voltage and also have the opportunity to choose the right lamp for the right occasion. **Filament lamps** are marked with the voltage and the wattage, or, in the case of small lamps, with the current (in milliamps). **Fluorescent lamps** and other **discharge lamps** are marked only with the wattage.

The link between voltage, output and current can be described by the following formulas.

$$U = R \times I \qquad \text{and} \qquad P = U \times I$$

where U is the voltage in volts, I is the current in amps, R is the resistance of the lamp in ohms, and P is the power in watts.

In the case of AC power circuits which do not have purely resistive loading, such as fittings for fluorescent lamps or motors, a further factor has to be taken into account, the phase shift factor cosφ. This is due to the phase shift which occurs between the voltage and current in a circuit of this type. The relationship then becomes:

$$P = U \times I \times \cos\varphi$$

Technical lighting quantities and units

The basic quantities and units in lights and lighting technology are as follows:

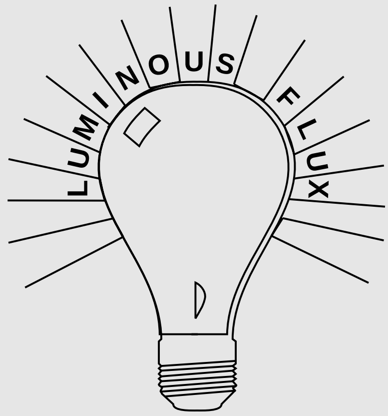
A few technical lighting quantities and units

Quantity	Quantity symbol	Unit
Luminous flux	Φ (fi)	lm (lumen)
Luminous intensity	I	cd (candela)
Illuminance	E	lx (lux)
Luminance	L	cd/m^2
Luminous efficiency	η (äta)	lm/W

These are used to indicate the lighting properties, light distribution, efficiency, etc. of the light sources and fittings. They are necessary for designing lighting, and the results which are achieved are always stated in these quantities and units.

Luminous flux (Φ) – lm

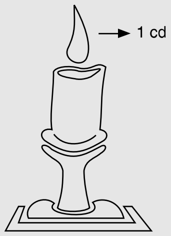
The luminous flux is expressed in lumen (abbreviated lm) and is the total light which radiates out from a light source. However, the flux of light is not the same in all directions.



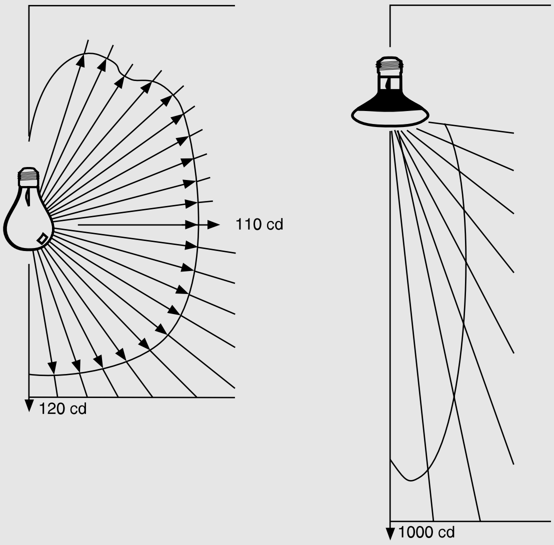
The luminous flux is the total light which radiates out from a light source.

Luminous intensity (I) – cd

The luminous intensity is expressed in candela (abbreviated cd) and is the light from a light source in *one* specific direction.



To gain some idea of how much 1 candela is, a candle with a diameter of 25 mm gives a light intensity of approx 1 cd.



Filament lamps do not give the same light intensity in all directions. A 100-watt, 1000-hour normal lamp gives a light intensity of approx 120 cd in the direction of the lamp shaft and approx 110 cd at right angles to the lamp shaft.

A 100-watt reflector lamp with a radiation angle of 35° radiates all light in almost the same direction with a light intensity of 1000 cd along the length of the lamp thanks to the reflector.

Illuminance (E) – lx

The illuminance is expressed in lux (abbreviated lx) and is a measure of the flux of light which hits a surface.

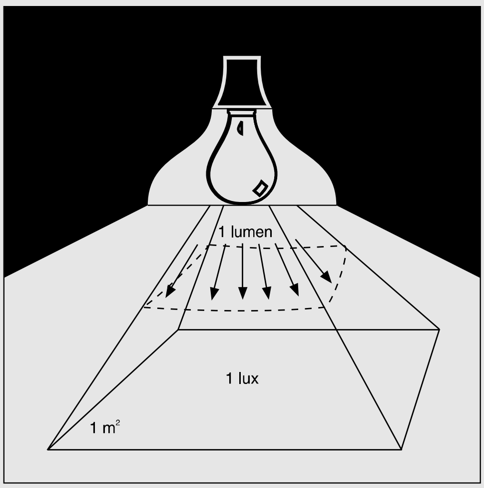
The illumination – E – is the relationship between the total flux of light – Φ – which hits a surface, and the size of this surface – A.

$$E = \Phi/A$$

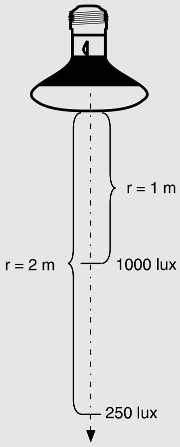
Example: When a surface – A – of 1 m × 1 m (=1 m^2) is evenly illuminated with a flux of light of 1 lumen, the illumination will be

$$E = \Phi/A = 1 \text{ lm} / 1 \text{ m}^2 = 1 \text{ lx}$$

The above is actually true only if we have the same flux of light on the entire surface. However, this is very rare, and therefore the result we get is the average illumination.



Ilustration of illuminance E.



The illumination depends on the distance.

The illumination is dependent on the distance to the light source and diminishes the greater this distance is, according to a specific law. As E = Φ/A och I = Φ/ω as we discovered earlier and A = ω × r^2 (r is the distance) then

$$E = I / r^2$$

where E is the illuminance expressed in lx, I is the luminance cd and r is the distance in m.

If we want to calculate the illumination – E – which can be achieved at a specific point, this can be done by finding out what light intensity – I – the light source is emitting in the direction of this point.

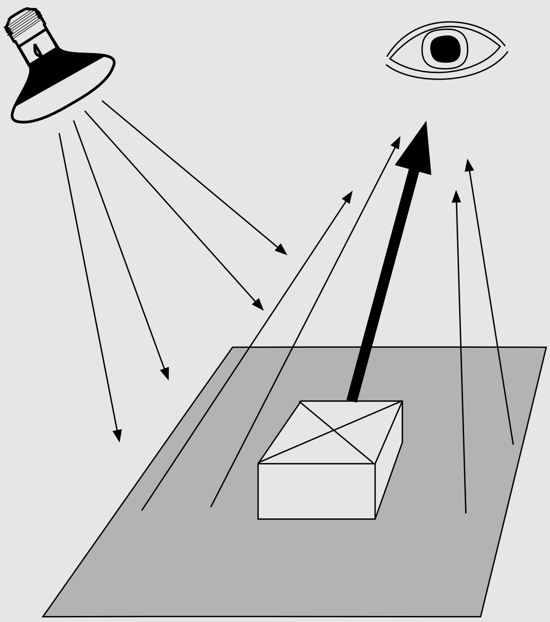
Example: The illumination for a light intensity of 1000 candela at 1 m distance: E = I / r^2 = 1000/1^2 lx = 1000 lx; at 2 m distance: E = I / r^2 = 1000/2^2 lx = 250 lx.

This relationship forms the basis for the design of floodlights, lighting which is suspended from a great height, spotlights and the like.

Luminance (L) – cd/m^2

The luminance is expressed in candela per m^2 (cd/m^2) or cm^2 (cd/cm^2), and is a measure of the light impression that the human eye gets from a shining surface; ”perceived brightness”. The luminance is defined as the light intensity in relation to the projection of the shining surface perpendicular to the direction of sight; in other words the light intensity in relation to the size which the eye perceives of the shining surface.

The luminance of a reflecting surface depends on the light striking it and on the **reflectiveness** of the surface in the direction of sight.



Luminance is the experienced light unit from a surface.

This concept is very important in the context of street lighting. A black road surface reflects very poorly, and thus its luminance is low. This is the opposite of a light road surface, which gives far better visibility thanks to its high luminance.

Good reflective properties also mean that less lighting can be installed under such conditions.

When it comes to dazzling, the concept of luminance is an important factor. If the greatest differences in luminance are in the field of vision, this can lead to dazzling and irritation.

If we were to look at a floodlight at night, it would dazzle very strongly, as opposed to if we were to look at the same floodlight on a sunny day. The floodlight has the same luminance by day or by night. However, at night the ambient luminance is very low. Thus the **contrast** is great, and the level of **dazzle** is high. During the day, the ambient luminance can generally be just as great as that of the floodlight. Thus the contrast is small and there is no dazzle.

Luminous efficiency (η) – lm/W

The luminous efficiency is expressed in lumen per watt (abbreviated lm/W) and is the measure of the effectiveness or efficiency of a light source. The light yield indicates how much flux of light a light source is giving out in relation to the electrical power used.

$$\eta = \Phi/P$$

The greater the light yield, the more efficient the light source, as a rule. However, in this context it is also necessary to take the service life of the light source into account.

Comparison of properties of different light sources.

Type	Power	Luminous flux lm	Luminous efficiency lm/W	Service life h
Filament lamp	60 W	730	12.1	1000
	100 W	1380	13.8	1000
Low-volt halogen	20 W	350	17.5	2000
Low-energy lamp	11 W	600	54.5	8000
Fluorescent lamp	36 W	3450	95.8	12000
Mercury fluore-scent lamp	80 W	4000	50	15000

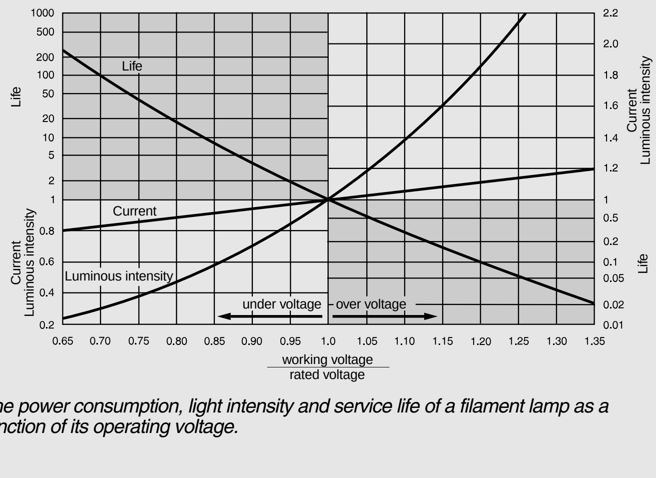
Lamps

Filament lamps are manufactured with sockets to an international standard. Lamps with a screw thread are designated E5.5, E10, E14 and E27, for example, where the numbers indicate the external diameter of the thread in mm. Lamps with bayonet sockets are designated BA7s, BA9s, BA15s, etc., for example. There are also miniature lamps with push-in sockets of "wedge" type and "telephone" type, as well as lamps with two terminals, "Bi Pin" type, which are mounted in sockets or soldered in. There are also lamps designed as glass tubes, along with different models of miniature lamp.

Filament lamps work, as we know, by the filament being heated up by a current passing through it to a temperature which is so great that the filament gives off a visible light. Tungsten is used as the filament material, which has a fusing point of 3 655 K. In an ordinary filament lamp, the filament reaches a temperature of between 1 800 and 2 500 K. The higher the temperature, the whiter the light (which is usually just expressed in K), but the higher temperature also reduces the service life.

Some applications demand lamps with a whiter light, e.g. lighting for photography and film. For this purpose, there are filament lamps which are able to withstand between 2 500 and 2 900 K due to the fact that they are filled with gas.

A filament lamp draws more current at the moment it is switched on, up to approx 12 times more, than when it has become so hot that it shines. The switch-on time is shortest for low-current lamps. Within 20 ms, the current is down to approx twice the nominal for a 0.1 A lamp. The current when the lamp is switched on can be limited by means of a preheating current which is permitted to flow through the lamp without switching it on.



The power consumption, light intensity and service life of a filament lamp as a function of its operating voltage.

If the working voltage of the filament lamp differs from the rated voltage, its properties are changed. In the diagram you can see how the service life is reduced to 0.05 times the nominal at a surge voltage of 25 %. On the other hand, the luminous efficacy is then 2.1 times greater. At the same time, the colour temperature increases (whiter light). In special applications, this may be desirable. In other cases, such as in indicator lamps for equipment where reliability is paramount, there may be just cause to reduce the voltage. However, in such instances it may be more appropriate to use LED lamps as indicator lamps.

The service life of a filament lamp is affected not only by the size of the voltage, but also by whether it is operated using DC voltage instead of AC voltage. This halves the service life. Knocks and vibration will also reduce the service life. Low-voltage lamps are best able to withstand knocks and vibration. An increased ambient temperature will also reduce the service life.

Halogen filament lamps have a filament just as in ordinary, evacuated lamps, but here there is a constant tungsten-halogen process. Tungsten-halogen is gaseous and transparent. It forms a deposit by means of heat circulation on the filament but not on the glass bulb. This means that the flux of light remains constant throughout the entire service life of the lamp.

The luminous efficacy is better than in ordinary filament lamps, and the colour temperature is higher, approx 3 000 K, which is useful in applications such as lighting for film and photography, in slide projectors, for illuminating works of art, etc. Another advantage is that they have a longer service life than ordinary filament lamps.

Fluorescent lamps give a very high luminous efficacy, often 100 lm/W or more. This should be compared with halogen lamps, for example, which give approx 12–25 lm/W, or ordinary tungsten filament lamps which give up to 18 lm/W at 2 500–2 900 K or 1–8 lm/W at 1 800–2 500 K. The service life is approx six times higher than for a filament lamp. The tubes are manufactured with colour temperatures of between 2 700 and 6 300 K and for UV radiation.

In series with the fluorescent lamps there must be a ballast which limits the current. This ballast also has another task: to provide sufficient ignition voltage. With the assistance of a starter, the ballast (like the two filaments of the tube) is traversed by current before the fluorescent lamp comes on. The energy stored in the ballast gives rise to a voltage pulse when the current ceases, upon which the fluorescent lamp comes on. The size of the ballast must be adjusted to suit the output of the lamp. Fluorescent lamps are made as a rule for 230 Vac Greater deviations demand a change of ballast.

Compact fluorescent lamps are available, with special sockets or an ordinary E27 thread. In the latter case, driver circuits, i.e. ballast and starter, are always included.

Fluorescent lamps have a phase angle of $\cos\varphi = 0.4–0.5$. Therefore, in permanent installations they should be phase compensated by a capacitor to approx $\cos\varphi = 0.9$.

LED lamps are completely different to ordinary filament lamps in that they have no filament. Instead, they have LEDs with semiconductive material which emit the light. When an electron falls into and is bound to a hole in the semiconductor, the hole is extinguished. Energy is thus released. This energy turns into heat in ordinary silicon semiconductors, but it is possible to get visible light in different colours or IR light by using other semiconductor materials, depending on the material or doping used. Red, orange and yellow are obtainable from gallium arsenide phosphide (GaAsP), while gallium phosphide (GaP) is used for green and blue.

LED lamps are usually adapted to a specific supply voltage. There also exist what are known as LED lamps which contain only one LED with a forward voltage drop of approx 2 V. These lamps *must be fitted with a series resistor* in accordance with the formula:

$$\text{Resistance (k}\Omega) = \frac{\text{Supply voltage (V)} - \text{Forward voltage (V)}}{\text{Forward current (mA)}}$$

Example: We want to operate the lamp using 10 mA and have a 5 V supply voltage. By the formula, this gives (5–2)/10 = 0.3 kΩ.

We have protected the LED with the resistor by limiting the current. However, we must *reverse the polarity of the LED correctly!*
A reverse-biased LED will not light up. Furthermore, it is destroyed immediately if the voltage is approx 5 V or higher.

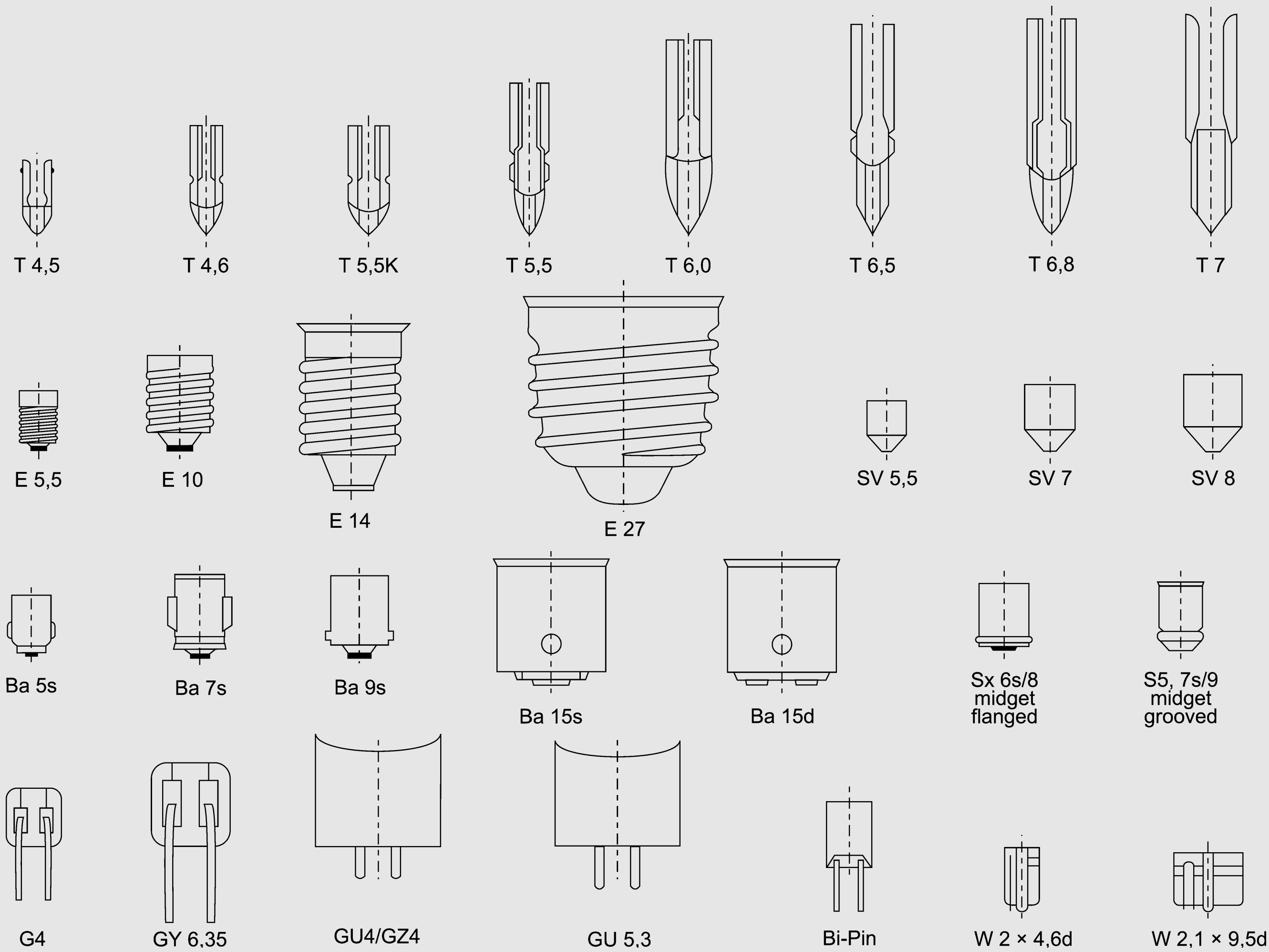
Neon lamps are a completely different type of lamp to filament lamps. They consist of two electrodes in a glass casing filled with an inert gas. When the voltage applied exceeds a certain value, the gas will become conductive by being ionised. In this instance, too, it is necessary to connect a series resistor in order to limit the current. The voltage over the lamp itself will then be constant. The voltage value depends on the gas pressure selected by the manufacturer. Usually, the operating voltage is between 60 and 150 V. However, the firing voltage is higher. The supply voltage must therefore be at least equal to the firing voltage. The series resistor is rated as follows:

$$\text{Resistance (k}\Omega\text{)} = \frac{\text{Supply voltage (V)} - \text{Operating voltage (V)}}{\text{Current (mA)}}$$

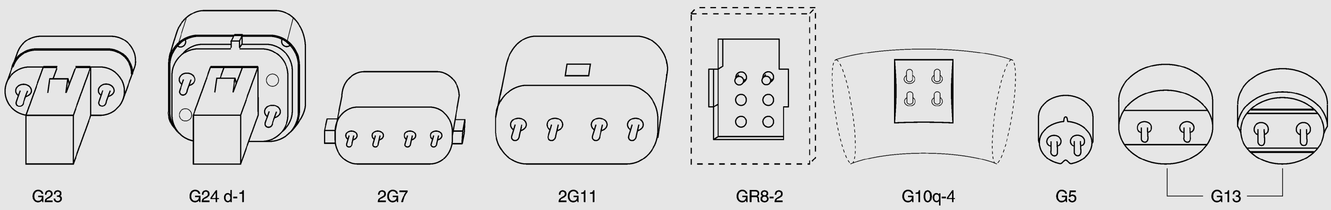
The neon lamp can be used as a voltage stabiliser due to the way in which it works. **The neon stabiliser tube** works in the same way. In fact, it is only the mechanical design which differs from that of the neon lamp, along with the fact that the operating voltage of the stabiliser tube is well specified.

Neon lamps with integral series resistors are also available. The supply voltage for these is stated.

Lamp bases. Scale 1:1



Fluorescent lamp bases. Scale 1:2



Switches and relays

Switches

The term switch is used to describe a wide range of components, most of them manually operated, that close or open an electrical circuit or switch it from one line to another.

The voltage being handled by the switch must be known in advance because higher voltages require better insulation. The strength of current is another important factor. Note that closing a switch can provoke a significant rush of current in many loads.

The current capacity depends on the design of the contact surfaces, the selected material, the dimensions and the pressure of the contact. It is also different for direct current or alternating current. This is due, for example, to the arcing which can occur when a switch is opened. To extinguish the arc, either the distance between the contact surfaces must become sufficiently large, or the current must be reduced. In alternating currents, the current is regularly reduced to zero, thereby making it easier to extinguish the arc.

Alternating current also prevents the migration of material from one pole to the other. Some manufacturers specify plus and minus poles for the connections, in which case one of the contacts is silver-plated while the other is made of solid silver. The polarity is selected so that the material migrates from the solid silver contact to the silver-plated contact.

The insulation material should be selected on the basis of the intended application of the switch. The insulation material causes losses that increase as the frequency increases. High-frequency signals require ceramic or PTFE insulation. In the case of very high frequencies, the conducting paths also have to be impedance-matched to keep losses and signal reflections to a minimum.

From the mechanical point of view, there are several types of switch:

Toggle switches usually require considerable mechanical force and actuator movement, but the positions are distinct.

Slide switches do not have such distinct switching positions. They are only used with low voltages, e.g. miniature switches in DIL packages.

Microgap switches are ideal when a small actuator movement is required together with very little operating force. Ingenious gearing means that the movement of the contacts can be magnified. Switches of this type contain curved split springs allowing the moving contact to move to either position. Between these positions, the switch is unstable. The result is precise and distinct positions, and fast changeover.

Reed switches are sensitive to magnetic fields. A glass tube contains a metal blade which, when acted on by a magnetic field, closes the electric circuit between the two electrodes in the switch. Reed switches are available either as standalone components or they form part of reed relays.

When using reed switches, remember:

- Sensitivity is reduced if the connecting pins are cut or bent.
- The glass may crack if the contacts are bent. To prevent this, grasp the contacts with pliers where they emerge from the glass. This does not apply to reed switches with flat glass tubes and flat contacts.

Functional descriptions

SP (Single Pole) = 1 pole. **DP** (Double Pole) = 2 poles.
ST (Single Throw) = two-position contact with an output for one position only (i.e. make or break). **DT** (Double Throw), **CO** (Change Over) = two-position contact with separate outputs for each of the positions (i.e. changeover).

Form A: Make contact. **Form B**: Break contact. **Form C**: Changeover contact, break-before-make. **Form D**: Changeover contact, make-before-break. If the name includes a number, this relates to the number of contacts.

Shorting, make-before-break: When the switch changes from one position to another, the contact to the new position is made before the previous position is broken. This means that the connections to the outputs for the old and new positions are short-circuited for an instant (unless the switch has completely separate contacts for each position).

Non-shorting, break-before-make: When the switch changes from one position to another, the contact to the old position is broken before the contact to the new position is made. There is no short circuit between the outputs.

Instantaneous, momentary, (on), (off), normally closed (NC), normally open (NO), opening, closing are all terms used to describe contacts with a starting position to which they return when the button is released. The term "spring-return", however, can also mean that just the button returns to its original position, and not the contacts. The terms "(on)" and "(off)" mean that the contacts return from an on/off position respectively. The term "normal" refers to the function in the starting position, and "opening" and "closing" indicate how the position of the contacts changes from the starting position.

Latching action means that the contact position changes the first time the button is pressed, and the position does not revert to the starting position until the button is pressed a second time.

Group actuation means that a number of switches are interconnected in a group in such a way that when one of the switches is actuated, any switches already actuated revert to their starting position.

Relays

Relays are switches that can be remote controlled, normally by passing electric current through a coil whose magnetic force actuates mechanical contacts.

The main differences between relays consist of their contact function and coil characteristics.

Contact function

Contacts can have a make function (form A), break function (form B) or change-over function (form C).

The contacts are designed with a specific maximum power in mind. This maximum power depends on the contact pressure and on the size and material of the contact surfaces. Complete relay data includes the maximum voltage, current and power.

Coil data

Relays are manufactured with actuating coils for direct current or alternating current, usually for voltages between 5–220 V. When selecting relays, you may need to take its own power consumption into account. In d.c. relays, consumption depends on the resistance of the relay coil, and the higher the resistance, the lower the power consumption of the relay.

You can use the coil voltage (U) and resistance (R) to calculate the power consumption (P) with the following formula:

$$P=U^2/R$$

Specific relay types

Latching relays remain in the switched position until a pulse of reverse polarity is applied. If there are two coils, one can be used for making and the other for breaking.

Solid state relays usually consist of one drive stage and one output stage. An isolating component is normally placed between the stages, for example an opto-isolator or reed relay. Depending on the type of current and the power, the output stage consists of a transistor, a triac or two opposing thyristors.

Reed relays consist of a reed switch (described in Switches above) actuated by a coil.

Contact protection

Switches and relays will only achieve the service life quoted by the manufacturer if the specified data relating to voltage, current and power are adhered to.

Contact data is normally quoted only for resistive loads. In the case of capacitive or inductive loads, the switching capacity of the relay is reduced. The manufacturer’s data sheet contains details of the inductive load capacity of a relay.

Capacitive loads

When a filament lamp, motor or capacitive load, e.g. fluorescent tube, is started, a rush of current occurs that exceeds the usual rated current by 10–15 times. This can be counteracted by connecting a choke, an NTC resistor or a PTC resistor in series with the load. For d.c. and a.c. applications.

Inductive loads

When an inductive load is switched off, e.g. a solenoid valve or motor, a break transient occurs that can exceed the supply voltage by several hundred volts. To prevent this transient occurring, various kinds of contact protection can be used, such as Comgap, varistors, double zener diodes, diodes and RC links.

Comgap. Comgap is a plasma type transient protection. When the voltage over the Comgap exceeds the rated voltage, the resistance falls from more than 10 MΩ to a few mΩ in less than 1 ϑs. For d.c. and a.c. applications.

Varistors. At a specific voltage, the resistance is changed from a very high to a very low value in an extremely short time. The varistor absorbs the energy from the transient, maintaining the voltage at an acceptable level. Unfortunately, the break time is impaired by the varistor.

With operating voltages between 24–28 V, the varistor should be connected over the load, and with operating voltages between 100–240 V it should be connected over the relay contacts. For direct and alternating current.

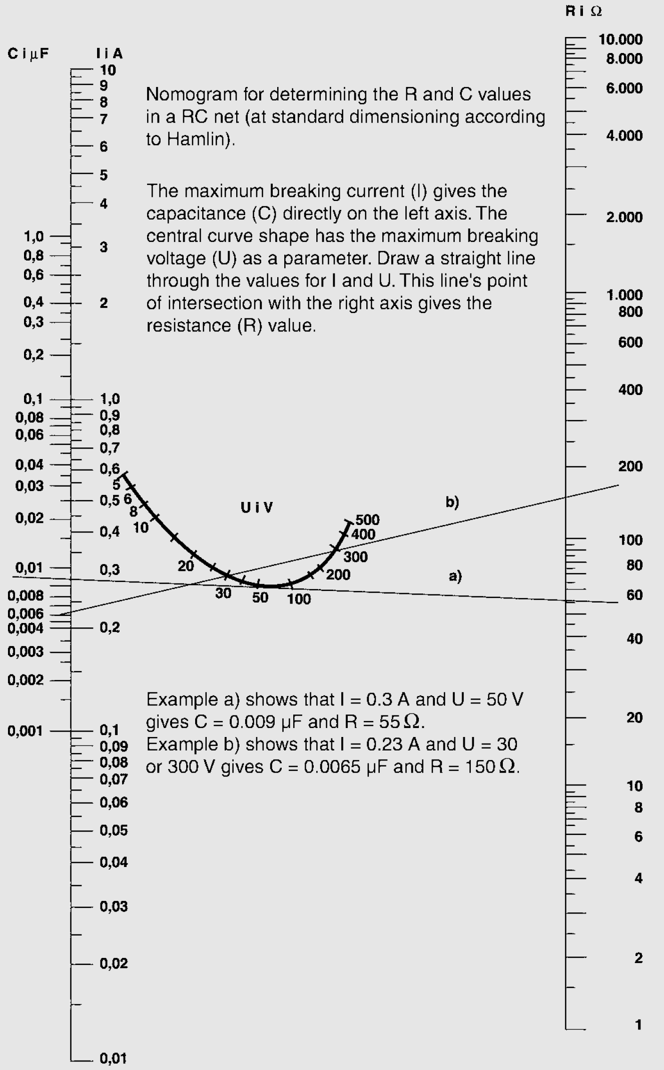
Double zener diodes. Two opposing zener diodes connected in series and fitted in parallel over the contact or load. They operate in a similar way to varistors. For direct and alternating current.

Diodes. A normal diode or transient voltage suppressor diode connected over the load. If a zener diode is connected in series with the diode, the break time is not affected as much as by a diode on its own. For direct current. Some types of transient voltage suppressor diodes can also be used with alternating current.

RC combination. An RC combination consists of a resistor and a capacitor connected in series. It can be connected in parallel over the contacts or the load. In some circumstances, it should also be used with purely resistive loads, for example when mercury relays are used.

The RC combination can also protect against break transients and prevent the occurrence of radio interference.

The RC combination should be connected over the load with supply voltages up to 24–28 V, and over the contacts when the voltage is between 100–240 V. For direct and alternating current, and can be combined with other methods of contact protection.



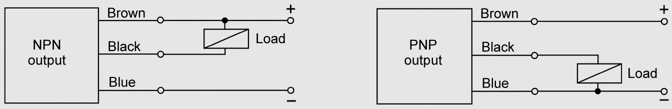
Dimensioning of a RC-link as contact protection.

Sensors

A sensor (or transducer) is a component that detects something and converts it into a signal.

Sensors are subdivided into two main groups: those generating a signal that is in one of just two different states (e.g. make/break) and those generating a (more or less) proportional signal.

The two-state sensors frequently have a transistor output instead of mechanical make/break operation. The transistor output can be either the 3-wire type or the 2-wire type. The 3-wire type normally comes in two different designs, PNP and NPN, with standardised connections and colour-coding of the wires. The output of a PNP sensor closes to the positive pole of the supply voltage when it is activated, whereas the output of an NPN sensor closes to the negative pole. 2-wire sensors are also called Namur sensors. The best way to describe them is as a variable resistor with high impedance when the sensor is activated and low impedance when it is not. Sensors with transistor outputs are usually protected against short circuits and incorrect polarity.



Connection of NPN resp. PNP-type sensors.

Some types of sensors:

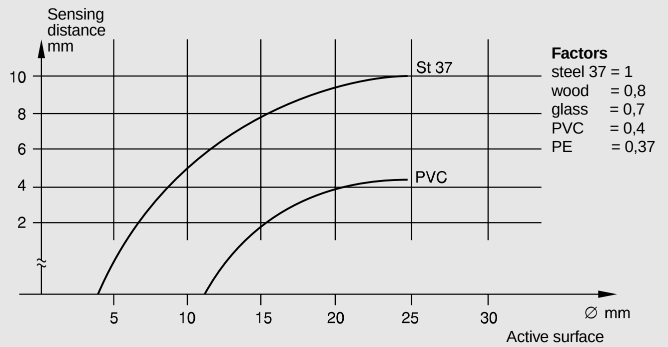
Level guards may contain a float with a magnet that can act upon a reed switch.

Pressure sensors carry out switching if a preset pressure is reached.

Tilt switches react to small changes in the angle of inclination, and they are used in position sensors.

Capacitive sensors output a resistance that varies with the capacitance when near a metal or a liquid, for example. The lower the dielectric constant, the closer the sensor has to be placed before it is actuated. Examples of dielectric constants: Air 1, polyamide 4–7, glass 5–15, metal 50–80 and water 80.

The sensor can detect water on the other side of a glass panel, for example - in other words, a higher dielectric constant can be detected through a material with a lower dielectric constant.



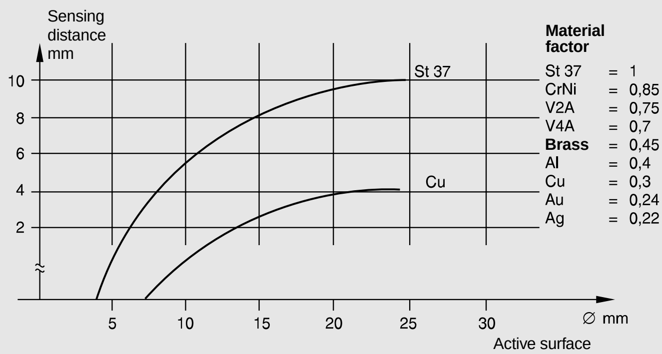
Capacitive sensors: Typical relationship (for different materials) between the contact distance and the size of the sensing surface.

Usually, the sensor has to be sensitive to a particular level, and should react at around that level. Some sensors are supplied complete with a switching function. They include an oscillator that is activated when a particular capacitance is reached near a medium. The sensors also contain amplifiers and transistor power amplifiers.

This type of capacitive sensor can be used to monitor the level of liquid or powder in containers, as an impulse counter to count parts, for sensing conveyor belts and V-belts, or for detecting the position of a product on the conveyor belt. Detection is entirely non-contacting. This type of sensor is maintenance free, is not subject to wear, generates distinct pulses and does not suffer from contact rebound or sparking. High frequencies are also tolerated.

Inductive sensors, at their most basic, consist simply of a coil that reacts to a change in magnetic field.

There are sensors that do not require an external magnetic field, reacting instead to a change in the magnetic field generated by the sensor itself. The field is affected by metal objects. The magnetic field is generated by an internal signal generator.



Inductive sensors: Typical relationship (for different materials) between the contact distance and the size of the sensing surface.

There are sensors with built-in trigger circuits ensuring a definite make or break under the specified influence of the magnetic field. The make or break action takes place with hysteresis so it is concise. In fact, these sensors should be classified as inductive switches. The output directly controls electronic circuits, relays or contactors.

The switching is non-contacting at a certain distance from the sensor. There are unshielded and shielded sensors - the shielded sensors have shorter contact spacing, but can be entirely enclosed in metal.

The usual applications for inductive sensors include use as non-contacting limit switches. They are particularly suitable in positioning applications or for counting objects.

Optical sensors contain a phototransistor, a photodiode or a photoresistor as the receiving sensor, see also the Optical Components factsheet. The sensor as a whole normally contains a transmitter and a receiver for modulated infra-red (IR) light and an IR sensitive receiver.

There are three main types:

- Combined transmitter/receiver working with a reflector. This type reacts when the light is interrupted.
- Combined transmitter/receiver, recording a light object which, as it passes within range of the sensor, produces a reflection of the signal, resulting in a reaction at the output.
- Separate transmitter and receiver. They can both be present in a combined sensor. The light is passed via optical fibre to the transmission and receiving points.

The sensors have semiconductor or relay outputs. The fact that the light is modulated means that the sensors are sensitive to IR light with an overlaid modulation frequency that distinguishes it from interference. Additional reliability is obtained in sensor systems using synchronised light, in which the receiver is only sensitive to modulated light in phase with the light from the transmitter.

Impulse counters and timers

Impulse counters

The traditional impulse counter is electromechanical, with the numbers being incremented mechanically. Fully electronic impulse counters have liquid crystal or LED display windows, and the counter settings have a battery backup or are stored in EEPROM.

Some counters have a preset at which upwards or downwards counting starts. Systems with built-in intelligence can also calculate the frequency, for example revolutions per minute or the periodic time, i.e. the time between two pulses.

The counting is based on incoming voltage pulses or by closing an input circuit, depending on the counter design.

Timers

Keeping track of the number of hours a piece of equipment runs is a valuable tool in improving service and maintenance functions. Operation timers work in one of three ways:

- The most straightforward principle uses the mains frequency, 50 Hz, as a reference. In principle, this type is a stepper motor connected to a mechanical counter.
- DC timers. This type contains an oscillator. The amplified clock signal acts on a stepper motor connected to a mechanical counter.
- Battery powered, with a built-in oscillator acting on a counter. The electronic circuits are created in CMOS, with a liquid crystal display to keep power consumption to a minimum.

Alarms

Alarms are usually created with sirens producing extremely high acoustic pressure. Many types of sirens consist of an electroacoustic converter and a piezoelectric type speaker with horn. Built-in driver circuits produce a constant, undulating or pulsating sound. Indoor sirens normally use high frequencies to create as much irritation as possible. Outdoor sirens should use lower frequencies so they can be heard over longer distances.

Alarms can also be created with light from xenon beacons.

Alarms are triggered by sensors of various kinds, including:

- Mechanically actuated pressure contacts.
- Magnetic contacts, for example fitted to windows and doors. One part is a magnet, the other is a reed switch.
- IR detectors reacting to heat and movement simultaneously, for example from humans.

The choice of sensor has to take account of the site of installation. Do you want it to be possible to remain in the room with the alarm triggered? In houses, the most common types are magnetic switches, possibly also with piezo sensors on the windows.

Fans

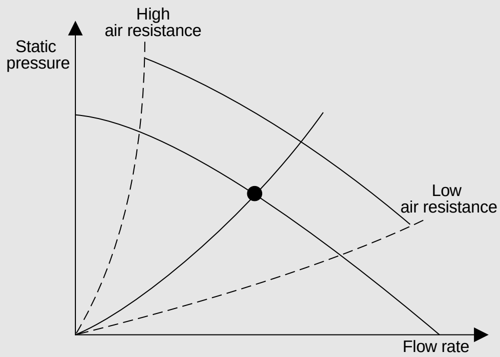
Compact circuits and greater component density mean that the slightest loss of energy by the individual components can cause overheating.
As the operating temperature increases, the life time of the components is reduced. That is why it is so important to remove excess heat. The most straightforward way of doing this is to use a fan. Fans can either introduce cool air or remove the warm air from the package. From the point of view of extending service life, the best approach is for the fan to introduce cool air. This also cools the fan itself and creates overpressure in the equipment.

The most common fan types are as follows:

Axial fans, the most widely-used in electronics applications. They are available in a range of sizes for various requirements in terms of air throughput, air pressure, sound levels, etc. They are manufactured using ball bearings or sleeve bearings. Fans with sleeve bearings are suitable for most applications, but fans with ball bearings are recommended if a long service life is required or if the ambient temperature is high. Fans with sleeve bearings should be installed with the shaft in a horizontal position. To obtain greater air throughput, two fans can be installed next to each other in the same panel. For greater air pressure, on the other hand, two fans can be installed one after the other in the same duct.

Radial fans are used in applications requiring greater air pressure at a given air quantity. Radial fans are usually noisier than comparable axial fans.

Tangential fans operate very quietly with a given air quantity. They have the disadvantage that they are physically very large and that the air pressure is low.



Pressure-flow rate curves for selecting the operating range of the fan. The fan achieves optimum efficiency and noise levels between the curves drawn with dotted lines. The static pressure (and air resistance) should be neither too high nor too low.

Selecting fans

You can use the following simplified formula as a guide to the air throughput you need to remove a given heat loss:

$$V = 3.0 \times P/(T_2 - T_1)$$

where

V = Air throughput in m^3/h
P= Heat loss in W
T_1 = Ambient temperature in °C
T_2 = Max permitted temp in °C inside the package.

To make sure the selected fan is up to the task, you should carry out a practical test in the intended application, for example by measuring the temperature increase inside the package. If the desired result is not obtained, you can replace the fan with another more powerful fan, or alternatively install more fans in parallel.

Heatsinks

The amount of heat given off by semiconductors like power transistors and diodes is too great for the component itself to transfer to the surrounding air satisfactory. To prevent temperatures rising to unacceptable levels, the components have to be helped to get rid of the excess heat.

One way of doing this is to install a heatsink, which transfers the heat generated by the transistor into the surrounding air through conduction and radiation.

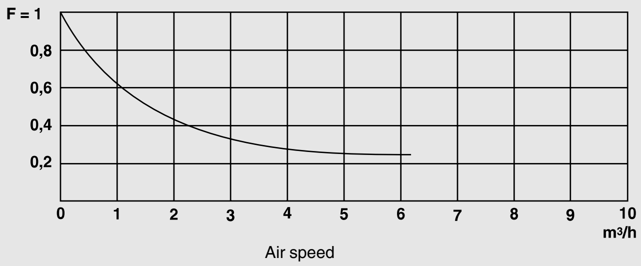
A flat metal panel is the simplest form of heatsink, but not the most effective. A more complex structure usually makes more sense in terms of cost, size and weight.

In a semiconductor, heat is generated in the barrier layer. The heat is then generally transfered from there to the case, and then to the surrounding air through the heatsink. This method of heat transmission is similar to the flow of current through electrical conductors. Accordingly, the thermal resistance (K in°C/W) corresponds to electrical resistance (R in V/A).

The following simple formula can be used to calculate the heatsink:

$$T_j - T_{amb} = P \times (K_{j\text{-}m} + K_{m\text{-}h} + K_h)$$

T_j = temperature in barrier layer.
T_{amb} = temperature of surrounding air.
P = heat generated in the semiconductor.
$K_{j\text{-}m}$ = thermal resistance between the barrier layer and the case. You can find this value in the manufacturer's data sheet.
$K_{m\text{-}h}$ = thermal resistance between the case and the heatsink. This value depends on the size and structure of the contacting surface. You can find the value in the data sheet.
K_h = thermal resistance of the heatsink, i.e. the thermal resistance between the contacting surface and the surrounding air.



Forced cooling using a fan. The thermal resistance of the cooling flange is multiplied by the factor F to obtain the reduced value for various air speeds.

The *thermal resistance* between the semiconductor and the heatsink should be kept as low as possible by using a large, flat and well machined contacting surface. Screws should be tightened to the recommended torque to ensure good heat conductivity, but without reducing mechanical strength. Any air pockets can be filled with silicone grease between the semiconductor and heatsink, but do not apply any more than is necessary. Thick layers reduce heat transmission. The thermal resistance $K_{m\text{-}h}$ varies between 0.14 and 0.05°C/W.

It is sometimes desirable to insulate the semiconductor from the heatsink using an insert. Different types have different thermal resistance. The thermal resistance of a 0.05 mm thick mica insert is about 1°C/W, of an 0.4 to 0.06 mm thick *mica insert*, silver-plated on both sides, about 0.5 °C/W and of a 3 mm thick aluminium insert with insulating aluminium oxide, about 0.3°C/W. Inserts made of kapton, silicone rubber and beryllium oxide are also available.

Beryllium oxide makes the best inserts, and is very useful in power amplifiers. However, the material is not for sale in Sweden because it is extremely poisonous. If you are servicing radio equipment manufactured outside Sweden, you should bear in mind that it may contain beryllium oxide in the insulating inserts. If an insert is broken it is potentially **life threatening** to breathe in the dust. The resulting chronic beryllium poisoning can cause problems with asthma. Long-term exposure can cause cancer. The thermal grease may also contain beryllium oxide.

The thermal resistance of many heatsinks is quoted for black, vertical surfaces. If the heatsink is installed so that the cooling surface is horizontal, thermal resistance increases by about 20%, and if the surface is shiny instead of matt black anodized, the thermal resistance increases by 15%. Note, however, that heatsinks available in different colours will have the quoted thermal resistance!

To increase the cooling effect, you can turn to forced air cooling using a fan. Another method is to use a *peltier element* with the cold side against the semiconductor and the warm side facing the surrounding air or a heatsink. Some heatsinks even have ducts for water or freon cooling.

If circuits generate large amounts of heat in short bursts, the thermal impedance is another significant factor. This is a time-dependent value, corresponding to the inertia, or mass, in the system. The thermal resistance within the semiconductor case is of crucial important in the case of very short lived thermal surges.

Electromagnets and motors

Pull and push type magnets

Pull and push type magnets should be selected on the basis of whether they are subjected to continuous connection, 100%, or reduced connection, e.g. 25%. For example, a pull type magnet might be connected for 20 seconds, then released for 60 seconds.

The force of attraction/repulsion varies with distance, but the variation is not linear. The position in which they are installed should also be considered. The forces usally quoted refer to horizontal installation. When magnets are installed vertically, you need to take account of the anchor weight, which either counteracts or reinforces the electromagnet depending on whether it is a push type electromagnet acting downwards or a pull type electromagnet acting upwards.

Small motors

Small motors are designed according to various different principles. Here is a short description of a few types:

Permanent magnet motors are the most common d.c. motors. They have excellent starting torque. The speed falls in a linear relationship with the current, and the current rises in a linear relationship with the torque.

Air-core d.c. motors are so called because the rotor core does not contain iron, only the copper winding. Iron is responsible for considerable losses when the magnetic poles are reversed frequently. This affects small d.c. motors, which often work at very high speeds. That is why it is advantageous to remove the iron from the rotor. The iron in the rotor is replaced with a stationary iron cylinder.

Air-core d.c. motors have a low moment of inertia, giving them a low mechanical time constant. This type of motor is suitable for use as a servo.

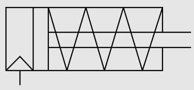
Stepper motors have a permanent magnet as the rotor, and a stator with two or four windings. At every phase change in the windings, the rotor moves step-by-step at an angle determined by the number of poles in the rotor and the number of phases. The mechanical step angle is 360/(n×p), where p is the number of poles and n is the number of phases in the motor.

Their operating properties make stepper motors particularly suitable for positioning applications, for example on an X-Y recorder. Stepper motors are best driven by a special driver circuit, which is driven in turn by a microprocessor.

Pneumatics

Cylinders

Cylinders convert pneumatic energy into mechanical energy (linear movement).

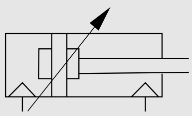


Single action

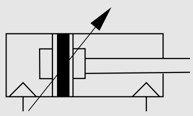
Cylinders in which the air pressure acts in one direction only. The return stroke uses external forces and/ or built-in springs.

Double action

Cylinders in which the pressure acts in both directions.



Cylinders with adjustable damping at both ends.



Cylinders with magnetic pistons for non-contacting position indication.

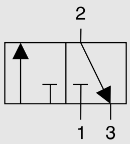
Directional valves

Directional valves are used to direct the air flow. The position can be changed manually, electrically or using compressed air. Each of the possible valve positions is represented in a separate box. The drawings show the position of the valves in the idle state. The air flow is illustrated with arrows and lines. The port types are indicated by the labels next to the connections - the letters that used to be used are due to be replaced by numbers in accordance with the preliminary CETOP recommendation RP68P.

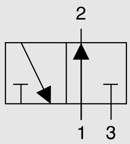
Numbers in RP68P:	Port type:
1	Inlet port
2, 4, 6	Outlet port
3, 5, 7	Exhaust port
10, 12, 14	Control port

3/2-valve

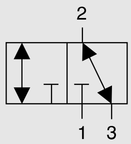
3 ports/2 positions. Usual application: supplying and removing a single volume, e.g. single action cylinders.



Supply valve



Exhaust valve



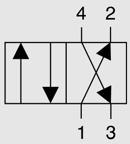
Supply/exhaust valve

In a supply valve in the idle state, outlet port two is connected to exhaust valve 3, and inlet port 1 is blocked. When the valve is actuated, the outlet port is connected to the inlet port, and the exhaust port is blocked.

In an exhaust valve in the idle state, inlet port 1 is connected to outlet port 2, and exhaust valve 3 is blocked. When the valve is actuated, the outlet port is connected to the exhaust port, and the inlet port is blocked. The function of a 3/2-way valve can also be provided by a 4/2-way or 5/2-way valve if either of outlet ports 4 or 2 is blocked (see below).

4/2-way valve

4 ports/2 positions. Alternating supply and exhaust operation on two volumes, e.g. double action cylinders.

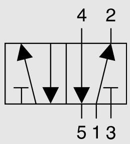


In 4/2-way valves, the two outlet ports 4 and 2 are alternated between inlet port 1 and exhaust port 3.

Both the outlet ports share exhaust port 3. This means it is not possible to restrict the exhaust flow as a way of regulating the speed in each direction separately in an attached double action cylinder.

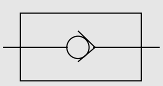
5/2-way valve

5 ports/2 positions. Mainly used as a 4 port, 2-way valve and in special applications.

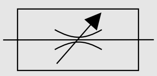


In 5/2-way valves, the two outlet ports 4 and 2 are alternated between inlet port 1 and exhaust ports 5 and 3. Outlet ports 4 and 2 use exhaust ports 5 and 3 respectively. This makes it possible to restrict the exhaust flow as a way of regulating the speed in each direction separately in an attached double action cylinder.

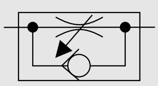
Non-return and throttle valves



Non-return valve
Allows air to flow in one direction only. It opens when the inlet pressure exceeds the outlet pressure.



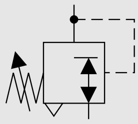
Adjustable throttle valve
Adjustable restriction of air flow in both directions.



Adjustable non-return throttle valve
Adjustable restriction of air flow in one direction. Air flow in the other direction is unrestricted.

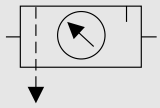
Pressure regulating valves

Devices that react when pressure at the inlet or outlet moves away from a predefined setting (mechanical, pneumatic or electric operation). Illustrated by a single box in which the broken line indicates the air line acting on the valve when there is a change of pressure. The arrow indicates the direction of flow, and the spring symbol represents the pressure setting.

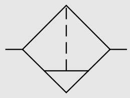


Pressure regulating valve (reducing valve)
A fluctuating inlet pressure is reduced to a constant pressure at the output.

Air treatment

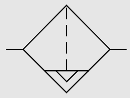


Air handling device
Simplified illustration.

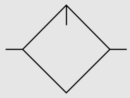


Manual drainage.

Filter with water separator
Separates solid impurities and water in mist form (condensate).

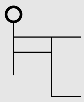


Automatic drainage.

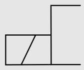


Oil mist lubrication
Adds oil mist to the compressed air in order to lubricate the equipment receiving the air.

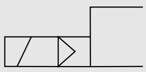
Control



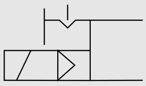
Lever



Solenoid



Solenoid and air pilot

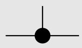


Solenoid and air pilot with manual operation

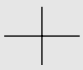
Power transmission

Air line for main flow

Air line for control flow

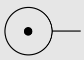


Crossing in which air lines are connected to each other

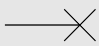


Crossing in which air lines are not connected to each other

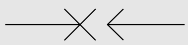
Flexible air line



Compressor

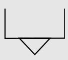


Plugged air line

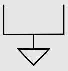


Connection with air line

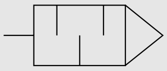
Electrical conductor



Exhaust without pipe connection



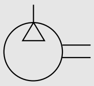
Exhaust with pipe connection



Silencer

Shut-off valve (simplified symbol)

Others



Compressor
Device for compressing gases, in this context turning air into compressed air. It converts mechanical energy into pneumatic energy.

Manometer
Measures pressure.

Connectors

Connectors constitute a link which can easily be broken. This provides flexibility in a system.

Connectors are generally made in accordance with some standard or specification, such as BS (British Standard), CCTU (French standard), DIN (German standard), IEC (European standard), MIL (U.S. military standard), etc. This standardisation is extremely important as regards the potential to switch between different manufacture of the same connector, so that the connector fulfils the same environmental requirements, lifetime, etc.

When selecting connectors, it is important to have a clear idea of what demands are to be placed on the connector in terms of current, voltage, lifetime and environment. There is no connector which can be used universally. The ideal would be a contact which has zero resistance when closed and infinite resistance when open.

The choice of contact material, plating thickness and the quality of the plating are decisive for the lifetime of the contact element. Different areas of application naturally have an impact on the requirements placed on the connector.

In connector pins and sockets, brass is an extremely common and low-cost material. There are also different quality levels as regards elastic deflection properties and hardness. Phosphor bronze, which has excellent elastic deflection properties, is a much better material. Beryllium copper is normally used in socket connectors and springs in extremely good quality connectors.

The contact elements are normally coated and *plated* with various materials to reduce transition resistance. These coatings can comprise gold, silver, rhodium, palladium, tin, nickel, copper, etc., either on their own or in various combinations. Of these combinations, gold and nickel have proven to be an extremely good combination as regards transition resistance, mechanical stresses and long-term stability. Hard alloys provide good resistance to wear, but can also produce high transition resistance in the event of low current. The oxide layer that is formed can create a diode effect in combination with the metal, resulting in distortion. It is therefore advantageous for audio connectors to be gold-plated. In connectors which transfer high currents, gold plating is not suitable due to its low melting point. In this case, silver-plating is more appropriate as it has the best conductivity, but you have to take care not to break a large current over the contact, as otherwise an arc could cause the silver to melt.

The connector manufacturer normally declares either the plating thickness and/or the number of operations, i.e. the number of times the connector can be pushed in and pulled out.

For e.g. Eurocard connectors, there are three different performance classes in accordance with DIN:

Class I:	500 operations
Class II:	400 operations
Class III:	50 operations

In basic connectors, Bakelite, Makrolon, nylon, ceramic, PVC, etc., are used as insulation materials. Better insulation materials include silicon rubber, DAP, PTFE, nylon 66 and Delrin which have good high frequency and temperature properties. Brass, ABS, steel, stainless steel, rubber, aluminium, etc., are used in contact bodies, casings, covers and strain reliefs.

Termination methods

Below is a brief description of the five most common types of connection.

Soldering

Soldering is a method which is relatively easy to carry out. It does not require expensive equipment and the conductor dimension is not critical. Its weaknesses include uneven quality, the fact that some component connections have poor solderability, and that pollutants can occur in the contact point. Soldering is also time-consuming. Quality is affected e.g. by the skill of the person carrying out the soldering, the choice of solder and flux, and the quality of the tool. See also the Soldering Factsheet.

Crimping

Crimping can be carried out quickly and easily. The result is an even connection quality, and it is possible to achieve a safe and gas-proof connection. However, the method requires special tools and in addition places demands on the choice of conductor area.

IDC

IDC, or Insulation Displacement Connection, is used when connecting multi pole contacts to ribbon cables. All conductors can be connected in a few seconds. The cable is clamped and stripped at the same time. The connection is gas-proof and safe. However, the conductors are fine, and as a result current and voltage capacity are restricted. This connection method requires special tools and only certain types of ribbon cable can be used. Ribbon cables with this connection method are used in PCs for example for internal connection of hard disks etc.

Press-fit

Press-fit technique has also been developed for multi-pole PCB connections. All types of Euro connectors, for example, are available today in Press-fit design. This technique is based on connector pins having a resilient connection that fit through-plated PCB holes. As the connector pins are pressed into the board, solder is scraped off the sides of the plated hole and a new, entirely oxide-free and gas-tight connection is created. One of the advantages of this technique is that one avoids the significant heating that is otherwise produced by soldering connectors with numerous pins. Furthermore, no subsequent cleaning of the connections is required and the connector pins can easily be made extra long to allow complementing wire wrap connection.

Wire wrapping

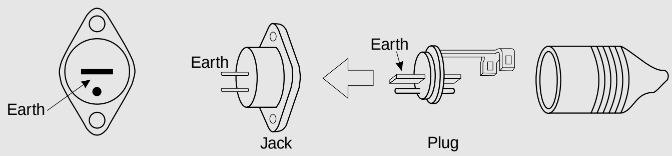
Wire wrapping is a method which is well suited for building prototypes. Properly conducted wire wrapping produces a secure and gas-tight contact. It is easy to make changes in a wire-wrapped construction by unwrapping and re-wrapping the wires. One of the disadvantages is that the method takes up a great deal of space. The pin has a rectangular profile. It has to have a particular height to give room for a number of wires. A wire is normally wrapped 5–7 turns. Wire wrapping requires special tools. See also the Wire Wrapping factsheet.

Fixed connections

Connectors are always a weak point in a system, and in some cases only connectors which satisfy military specifications are good enough. Error intensity rises with the number of poles. In some electronic equipment, however, the demands on high MTBF (Meantime Between Failure) are so stringent that connectors must be entirely or partially replaced by *fixed connections.* This can apply e.g. in space applications, where shaking, vibrations, temperature changes and possibly gases or fluids can impair contact function.

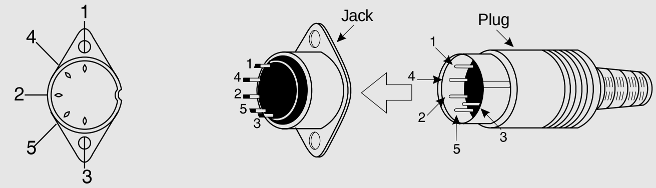
Applications

2-pole DIN connectors are used for loudspeaker outputs in the event of moderate output power. The wide pin is always connected to earth. In the event of higher outputs, pole screws are used.



2-pole DIN connectors, loudspeaker connectors.

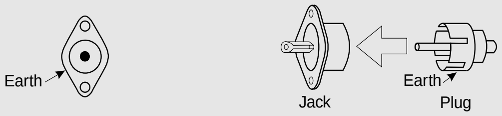
5-pole DIN connectors are used extensively in European HiFi apparatus, both for inputs and outputs. The pin connection for various types of apparatus are standardised in accordance with that set out below:



5-pole DIN connector

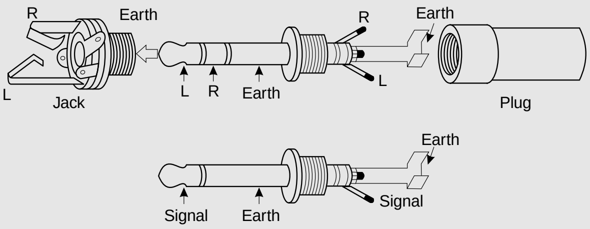
Type of apparatus	Connectors for	Input V	Input H	Output V	Output H	Earth
Amplifier	Pickup, tuner	3	5			2
"	Tape recorder	3	5	1	4	2
Tuner	Amplifier			3	5	2
"	Tape recorder			1	4	2
Disc player	Amplifier			3	5	2
Tape recorder	Amplifier	1	4	3	5	2
"	Microphone	1	4			2

Phono connectors are preferred in Japanese and American apparatus, but are also fairly widely used in European apparatus, in the latter case along with or connected in parallel with DIN connectors. In order to connect a stereo signal with phono connectors, two shielded cables are required, each with a phono connector.



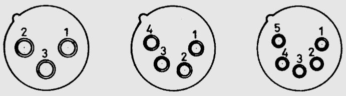
Phono connectors.

Tele plugs and tele jacks are primarily used for outputs to headphones and inputs for microphones. 3-pole connectors are used for stereo, with left and right channel in the same connector. 3-pole connectors are also used in blanced systems. Hot is then connected to "L" and cold to "R".



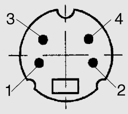
Tele plugs and tele jacks.

XLR-connectors are the most common connectors for microphones and sound systems. In balanced systems pin 1 is connected to earth, pin 2 to the positive conductor (hot, send), and pin 3 to the negative conductor (cold, return). In unbalanced systems, pin 3 or 2 can also be earthed. In some American equipment (microphones, mixing desks), pins 2 and 3 may be swapped.



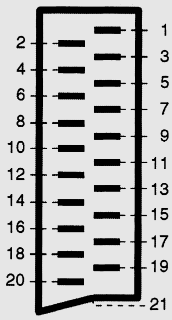
Contact configuration for XLR connectors, view from the solder side of the male connector.

The S-video connector is a 4-pole mini-DIN-connector. It is used for the 2 video signals from video cameras. The signals are luminance Y (signal level: 1 V incl. sync; impedance: 75 ohm) and chrominance C (signal level: 0,3 V burst; impedance: 75 ohm).



Contact configuration for S-video connectors view from the solder side of the female connector. Pin 1: Earth for luminance Y. Pin 2: Earth for chrominance C. Pin 3: Luminance Y. Pin 4: Chrominance C.

The SCART connector is used for video and audio signals in Television contexts, see diagram.



Scart connector.

Pin	Signal name	Description	Signal level	Impedance
1	AOR	Audio OUT Right	0.5 V rms	<1 kΩ
2	AIR	Audio IN Right	0.5 V rms	>10 kΩ
3	AOL	Audio OUT Left + Mono	0.5 V rms	<1 kΩ
4	AGND	Audio Earth		
5	B GND	RGB Blue Earth		
6	AIL	Audio IN Left + Mono	0.5 V rms	>10 kΩ
7	B	RGB Blue IN	0.7 V	75 Ω
8	SWTCH	Audio/RGB switch / 16:9		
9	G GND	RGB Green Earth		
10	CLKOUT	Data 2: Clockpulse Out		
11	G	RGB Green IN	0.7 V	75 Ω
12	DATA	Data 1: Data Out		
13	R GND	RGB Red Earth		
14	DATAGND	Data Ground		
15	R	RGB Red IN / Chroma	0.7 V 0.3 V burst 1)	75 Ω
16	BLNK	Blanking Signal	1–3 V 2) 0–0,4 V 3)	75 Ω
17	VGND	Composite Video Earth		
18	BLNKGND	Blanking Signal Earth		
19	VOUT	Composite Video OUT	1.0 V	75 Ω
20	VIN	Compos. Video IN / Luminance	1.0 V	75 Ω
21	SHIELD	Earth/Chassis		

1) Chroma. 2) RGB. 3) Composite.

Fibre optic conductors

There are many reasons for using fibre optic cables instead of copper wire. Optical transmission means that the link is not sensitive to electromagnetic interference, an important consideration in the industrial context. Another advantage of optical transmission is its potential for broadband use, making it suitable for telephony, computer and TV signals in digital form.

All optical fibres share the same structure, in which a core is surrounded by a cladding. The refractive index of the core is slightly higher than that of the cladding. The beams of light remain in the core because they are reflected when they graze the core-to-cladding interface. A protective coating usually encloses the inner components.

Fibres are available in two basic types: single-mode and multimode. Multimode fibres can in turn be divided into two types: *stepped-index* and *graded-index.*

In stepped-index fibres, the core is thick enough, at around 50 µm, to cause the light to be transmitted in many modes of propagation. If the core is made small enough (down to about 5–10 µm with a light wavelength of 1.3 µm), the optical cable carries only one mode. This is a single-mode fibre. Single-mode fibres have a high transmission capacity, but the thin core makes it difficult to splice the cables.

Graded-index fibres are a compromise between single-mode fibres and stepped-index fibres. The refractive index of the core of graded-index fibres smoothly tapers from the core centre to the cladding. A beam of light travelling at an angle from the centre of the cable is constantly refracted back towards the centre of the cable. The core of a graded-index fibre is thick enough to allow the transmission of different modes of propagation.

The principal factors limiting the distances in light transmission are dispersion and attenuation. The effect of dispersion is that the beams of light travel through the cable at different speeds. The dispersion of the light pulse places a limit on the maximum pulse repetition frequency, thereby limiting the transmission bandwidth. Different modes travel at different speeds in a multimode fibre, placing a further limit on bandwidth. The problem is avoided in single-mode fibres. In single-mode as well as multimode fibres, dispersion occurs in the material itself, because the refractive index of the glass varies depending on the wavelength and manufacturing flaws.

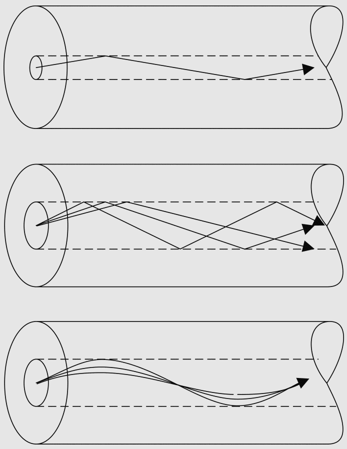
Dispersion depends on the speed which in turn varies depending on the refractive index following

$$v = c_0/n$$

where c_0 is the speed of light in a vacuum and n is the refractive index.

Attenuation and dispersion vary according to the transmission wavelength. The first fibres to appear in 1970 had an attenuation of 20 dB/km. Since then, manufacturers have learnt how to reduce attenuation in optical fibre, and transmission wavelengths have also been increased, which also has the effect of reducing attenuation. The first generation of optical fibre systems used a wavelength of 0.85 µm, the second 1.3 µm and the third 1.55 µm. The third-generation wavelength produces the least attenuation, a theoretical 0.16 dB/km, while dispersion is at its minimum at around 1.3 µm.

Splicing and connecting optical fibre cable is a difficult process. This is especially true of single-mode fibre in which the thin cores in each segment of cable have to line up with each other exactly. A phenomenon called Fresnel reflections also occur at the interface, with a minimum theoretical limit (about 4%). Contact attenuation varies between 0.2 and 2 dB depending on the type.



Three types of optical fibre: 1 – A single-mode fibre carries just one mode of propagation. This means that all beams of light are reflected from the cladding interface at the same angle. The result is that all beams are the same length and travel at the same speed, and there is no dispersion. 2 – In a thicker fibre, the beams are reflected at many different angles, causing dispersion. 3 – In graded-index fibres, the beams are steadily refracted.

PCB production

There are a number of different methods available for the construction of PCBs. The two methods briefly described below are:

1. Direct transfer of design to board.
2. Photographic transfer of design to board.

These methods are suitable for production of single boards or small volumes.

Direct transfer

In direct transfer, the symbols and paths are taped, rubbed or drawn directly onto a well-cleaned copper clad board. This is then etched in a solution consisting of sodium peroxide sulphate and water. Iron chloride can also be used, but this is not recommended for health reasons. Note that transfer symbols must be of a special type, designed for direct etching and not provided with carrier film.

Photographic transfer

In photographic transfer, use is made of a copper clad board coated with a layer of lacquer, i.e. *photoresist*, which is sensitive to ultra-violet radiation (UV light). This photoresist can be positive or negative (as in photography), but for the sake of simplicity only positive resist is considered here. Although the resist is not especially sensitive to ordinary light, it should not be exposed to light for long periods.

The procedure for production of a PCB with the photographic method is as follows:

Production of the design original (layout)

A design original is produced with curving tape, transfer symbols and a transparent plastic fim. You simply tape and rub off the design onto the plastic film. Symbols and characters are positioned very quickly and accurately thanks to the transparency of the film. It is easy to get the characters on the transfer strip to adhere. It is sufficient just to press or rub with one finger. Rubbing afterwards with the silicon treated protective sheet on the back of the strip causes the symbols to stick even better. The symbols have a definition of ±0.05 mm, are very thin and have a high resistance to mechanical agitation. To apply a line from the transfer, place the line against the underlay and cut it off through the plastic film to the desired length. To correct the symbols, scrape them away with a knife, lift them off with tape or use a special eraser. When the design original is ready, you can use it directly or alternatively make a ''working copy'' on positive/positive film.

Exposure

The original or possible copy is laid directly (scale 1:1) onto the emulsion side of the resist-coated board and the board is exposed with the help of a UV lamp or a more professional UV light box. The UV light must have a wavelength in the range 350-370 nm. Quartz lamps or sun lamps, which have a maximum wavelength of 256 nm, are not recommended. When carrying out an exposure, it is important to position the film on the board as precisely as possible and to be careful that dust and other foreign particles do not affect the result. The exposure time, especially if using a UV lamp, depends on the height of the lamp above the illuminated board, any possible glass sheet between the lamp and film laminate, etc. The following can however be used as guideline values:

With a UV lamp of 300 W, a glass sheet and a distance of 40–50 cm: 5–7 min.
With a UV light box: ca 2–5 min.

NOTE: USE SAFETY GLASSES when working with UV light.

Developing

Exposure is followed by developing. This is done in a bath of sodium hydroxide and water. A 1.5% solution is used. The developing time varies between 30 s and 4 min depending on the type of resist, and must be adjusted accordingly. After developing, the design should appear clearly. When positive resist is used, those parts that have not been exposed to light (i.e. are covered by the tape) will be protected during etching, thus leaving the design in place. After developing, the board is washed thoroughly with running water BEFORE etching.

Etching

It is an advantage here to use sodium peroxide sulphate for both direct transfer and photo transfer. For the etching bath to be effective, the powder is dissolved in boiling hot water, giving an etching solution temperature of about +50°C, which should be maintained for best results. Remember to protect hands and eyes from the chemicals. Use gloves and safety glasses! Suitable vessels for developing and etching are ordinary photographic developing trays. It is best, however, to use some sort of etching tank with air pump and ideally an immersion heater.

Quality assurance

If the exposure time of the photoresist is too short or if the wrong lamp is used, this will be apparent during developing and/or subsequent etching. The wrong type of lamp, too short exposure time or developer that is too old or weak cause parts of the photoresist to be left on the copper surface. These pieces, which are often very hard to see because they can be very thin, will appear as non-etchable pads during etching. To avoid this, the following should be observed:

1. The original must be flat black. No UV light is allowed to get through the symbols that make up the transferred design. The exception to this is the direct positive film that gives deep red UV-blocking lines. A UV-blocking original simplifies and enables full exposure of the photoresist. This gives the right conditions for a fully developed photoresist. It is advisable to over-expose by 1–2 minutes if there is uncertainty about the effectiveness of the lamp.
2. Use fresh developer. A used and spent developer can give poor results, especially if it has been used many times.
3. Check during developing that all exposed photoresist really has disappeared. If exposure has been too short, developing time must be extended up to 10 minutes. Stirring hastens the process.
4. During etching of the board, it is necessary to stir the etching solution the whole time one way or another to ensure that active parts are always in contact with the copper surface. At temperatures below +30°C, much of the etching capability is lost, which is why some form of heating arrangement is preferable.

NOTE: USE PROTECTIVE GLOVES and SAFETY GLASSES when working with chemicals, especially during etching and developing.

Dimensioning of foil conductors

Resistance of foil conductors

The resistance R of a copper foil conductor can be calculated from the formula:

$R = \rho_{Cu} \times l/(b \times t) = (\rho_{Cu}/t) \times (l/b)$

where ρ_{Cu} = resistivity of Cu, l = conductor length, b = conductor width and t = foil thickness.

ρ_{Cu}/t is for 70 µm foil $0.25 \times 10^{-3}\ \Omega$, for 35µm foil $0.5 \times 10^{-3}\ \Omega$ and for 17.5 µm foil $1.0 \times 10^{-3}\ \Omega$.

Example: The resistance of a 0.35 µm thick copper foil conductor of length 10 cm and width 1 mm is thus
$(\rho_{Cu}/t) \times (l/b) = 0.5 \times 10^{-3}\ \Omega \times (10 \times 10^{-2}\ m / 1 \times 10^{-3}\ m) = 0.05\ \Omega$

Maximum current and minimum spacing of foil conductors

The maximum allowed current depends on the foil thickness, the conductor width and how high a temperature the conductor can be allowed to withstand. If there is place on the PCB, it is appropriate to use foil conductors of width 1.57 mm (0.062") or 1.27 mm (0.05"). The minimum conductor width that can be produced in photographic exposure of the conductor design is 0.3 mm.

Table. Max current in A through foil conductors on PCB with 17.5 µm foil.

Conductor width (mm)	Allowable temperature in conductor 10°C	20°C	30°C	60°C	75°C	100°C
0.5	0.6	1	1.2	1.7	2	2.3
1.0	1.1	1.5	2	3	3.2	3.7
1.5	2	2.6	3.4	4.3	5	6
2	2.3	3.2	4	5	6	7
4	4	5	7	9	10	11
6	5	7	9	12	13	14
8	6	9	11	14	16	18
10	7	10	13	16	19	21

Table. Max current in A through foil conductors on PCB with 35 µm foil.

Conductor width (mm)	Allowable temperature in conductor 10°C	20°C	30°C	60°C	75°C	100°C
0.5	1.3	2	2.3	3	3.5	4
1.0	2	2.8	3.1	4	5	6
1.5	2.6	3.7	4.4	6	7	8
2	3.2	5	6	8	9	10
4	5.5	8	10	11	15	16.5
6	8	11	13	18	21	23
8	9.5	13	16	22	24	26
10	11	16	20	27	29	33

Table. Max current in A through foil conductors on PCB with 70 µm foil.

Conductor width (mm)	Allowable temperature in conductor 10°C	20°C	30°C	60°C	75°C	100°C
0.5	2.4	3.2	4	5	6	7
1.0	3.3	4.5	6	8	9	10
1.5	4.3	6	8	10	12	13
2	5	8	10	13	14	15
4	9	13	15	21	23	25
6	12	16	22	30	32	35

Allowable minimum distance between conductors, the insulation distance, is dependent on production method and maximum voltage; see table.

Table. Allowable minimum distance between foil conductors by photographic exposure of the conductor pattern.

Voltage (V) between conductors:	50	150	300	500
Min distance (mm) between conductors:	0.3	0.6	1.2	1.8

PCBs manufactured by sub-contractor

PCBs designed by you in a CAD-programme can be delivered by ELFA thanks to a contract of cooperation with the PCB manufacturer Elprint. It is possible for you to order individual boards for prototypes and smaller series, see *Customer adapted PCBs* on page 424.

Enclosures

Boxes

Enclosure boxes and cases are probably the last thing one considers when doing installations and constructions, but the issue is so important that one really should think about enclosures at a very early stage.

Several factors are important when adapting a design to a specific environment, such as durability, moisture resistance, flame resistance and screening.

The price is also an important factor for certain customer circles, as are appearance and practicality. The cost must also include adaption of both design and enclosure. It seldom pays to produce a completely unique mechanical design. A factory-manufactured case, box or rack system saves a lot of time, and thus expense.

The **material** constitutes the basis of the enclosure and its properties. Steel is very stable, but heavy, and it corrodes unless the surface treatment is very effective. Aluminium alloys are much lighter. Extruded sections and plates allow very flexible case systems. Plastic is not one, but many materials, with widely varying properties in terms of mechanical durability, processability, temperature resistance and flammability.

Most light-coloured plastic materials that cases are made of are poorly resistant to sunlight (UV radiation). They tend to yellow over time. Some plastics are manufactured with added UV inhibitors, which significantly improves their resistance to sunlight. Read more about this in the section about *Plastics*.

Flame resistance is specified in the UL 94 standard, which cases from e.g. OKW conform to. The materials are classified as follows:

94V-0 means that the test object is extinguished within five seconds on average. None of the test objects burn for more than 10 seconds. None of the test objects emit particles when burned. One example of such a material is flame-resistant ABS.

94V-1 means that the material is extinguished within 25 seconds on average, that the test objects never burn for more than 60 seconds, and that they never emit particles when burned.

94V-2 is the same as 94V-1, except that particles are emitted during burning. Examples of plastics that meet 94V-2 include flame-resistant polystyrene and polycarbonate.

If the test object should burn for more than 25 seconds, it falls under 94HB. Such materials include polystyrene, ASA and ABS plastics.

Screening has become more and more important since rapid rise times in logic circuits generate a wide spectrum of overtones. In some cases, sensitive circuits should also be screened against incoming fields.

Interference radiation can often be sufficiently deflected by filtering the incoming and outgoing cables. At frequencies above 1 MHz, however, a screening cage according to the principle *Faraday's cage*, i.e. an electrically conductive shell, is required. An aluminium alloy is often suitable in such cases, because the material is lightweight and easy to process. Zinc alloys are used for somewhat better screening properties.

Purely magnetic, extremely low-frequency fields require magnetic materials, such as iron. One excellent magnetic material is the metal alloy Mumetal®, which is used to seal LF transformers for example. However, aluminium is more conductive than iron, which often makes it a better choice.

High demands for damping also place demands on the composition of the case: the metal parts must be in contact with each other, the distance between contact points cannot be too great (critical for how high frequencies the case can damp). Watch out for anodised or oxidised aluminium with a very high surface resistance of tens of MΩ.

However, demands for damping are often quite moderate, and you can use plastic cases with a conductive layer of foil, evaporated aluminium, nickel lacquer, sprayed-on carbon powder etc. According to the manufacturer PacTec, these materials vary greatly. For example, at 5 MHz a 0.5 mm thin layer of copper in epoxy can attenuate 60 dB, nickel in acrylic 50 dB, silver in acrylic 45 dB, silver spray plating 35 dB and graphite lacquer in acrylic 15 dB.

Heat abstraction is often important. Remember that decreasing the temperature by just a few degrees can lengthen the interval between errors by thousands of hours. Heat can be abstracted or dissipated through air cooling. See section about *fans*.

19" enclosure systems for electronics

Before mechanical devices for electronics were bound by any norms, Schroff adopted the 19" measurement for front plates (482.6 ± 0.4 mm) from the United States. This turned out to be a decisive move. Since then, Schroff has been a pioneer of this area and contributed significantly to the development of the 19" norm with the introduction of the Europac system.

The system's measurements meet current international norms: DIN 41 494, IEC 297, BS 5954, EIA RS 310-C.

Quite simply, the system is designed for mechanical construction of industrial electronics. The 19" width of the front plate and the division into U height units (1 U = 44.45 mm) was the first norm established that allowed modular construction.

Single and double Eurocards are well-known terms for most card formats. They fit card frames of 3 and 6 U module height.

The internal width is also divided up in steps of 5.08 mm (= 1 HP) and allow up to 84 HP within the 19" width.

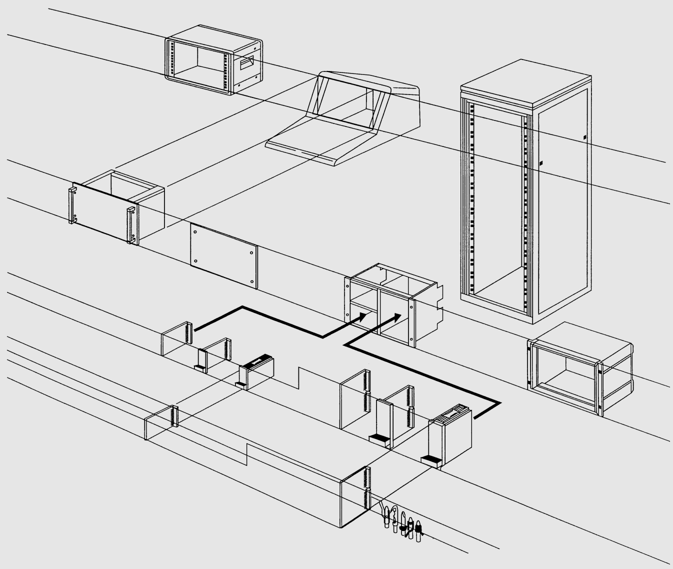
For standardised microcomputer systems, the module width of 4 HP (20.32 mm) has proven to be ideal. This division is simply called 1 *slot*.

The 19" system offers the user a complete construction system for all possible construction dimensions and electrical or environmental demands. The use of standard parts allows a variety of construction variations at a very low cost. The system also offers a complete range of accessories for many applications.

DIN divides the norm into the following levels (see figure):

- The *component level* includes PCBs and connectors.
- *Plug-in units* such as card cassettes or simple PCB modules.
- *Front plates and rack systems*
- *Cases* in many formats directly compatible with rack systems, like desktop cases or cases (with 19" brackets) for installation in 19" cabinets.
- *19" cases, cabinets and stands*

When rack systems are built into a case or cabinet, certain electrical and mechanical factors must be considered, both for screening and cooling the PCBs, which are often installed very close to each other.



19" mounting system for electronics

Ingress protection for electric equipment – IP ratings

Brief summary. For further information it is possible to order SS IEC 529 from SIS.

Exemples of designations

IP 4 4

Type of norm — IP

First digit — 4

Second digit — 4

An enclosure with this designation is protected against penetration by fixed objects greater than 1.0 mm and against pouring water.

First digit states degree of mechanical protection. Brief description.

0 No particular protection.
1 Protection against solid objects larger than 50 mm. E.g. body part like a hand (but no protection against intentional intrusion).
2 Protection against solid objects larger than 12 mm. E.g. like fingers or simular, not exceeding a length of 80 mm.
3 Protection against solid objects with a diameter or thickness exceeding 2.5 mm. E.g. tools, wires, etc.
4 Protection agains solid objects larger than 1.0 mm. E.g. wires or strips.
5 Protected against dust. Dust intrusion not completely prevented, but dust can not enter in amount great enough to prevent normal operation of the equipment.
6 Dustproof. No dust intrusion.

Second digit states degree of water protection. Brief description.

0 No particular protection.
1 Protection against trickling water. Trickling water (vertically falling droplets) must not be damaging.
2 Protection against trickling water at an inclination of 15°. Vertically trickling water must not have a damaging effect when sealing is inclined max. 15° from its normal position.
3 Protection against sprinkling water. Sprinkling water with an angle of max. 60° from the vertical line must not have a damaging effect.
4 Protection against oversprinkling of water. Water sprinkled against the sealing from any angle must not have a damaging effect.
5 Protection agains water jets. Water Skydd mot vattenstrålar. Water jet coming through a nozzle in any angle against the sealing must not have a damaging effect.
6 Protection against heavy seas. Water from heavy seas or water beeing splashed in powerful jets must not enter the encapsulation in a damaging amount.
7 Protection against brief immersion in water. Intrusion of damaging amount of water must not be possible when encapsulation is immersed in water at a certain pressure and for a certain time period.
8 Protection against affection caused by sustained immersion in water. The equipment is suitable for sustained immersion in water under conditions determined by the manufacturer.

Symbols

First digit

5	6

Second digit

1	3	4	5	7	8
					.....m .....m

Cables

Conducting materials

The conductor in virtually all conducting materials is made of copper. In exceptional cases, and for special purposes, silver, aluminium, constantan or similar materials are used. The resistance is an important characteristic of a conductor, and is described with the formula:

$$R = \rho(L/A)$$

where R is the resistance, ρ is a constant called resistivity that depends on the material and the temperature, L is the length and A is the area.

The resistance is dependent on the temperature. In metals,For metals the temperature dependence of the resistivity is almost a linear relationship. The resistance can therefore be described by the formula

$$R_T = R_{Tref} + \alpha(T - T_{ref})R_{Tref}$$

where R_T is the resistance at temperature T, R_{Tref} is the resistance at the reference temperature, α is a constant called the temperature coefficient, T is the temperature of the conductor and T_{ref} is the reference temperature.

The resistivity ρ and the temperature coefficient α for some common metals.

Metal	Resistivity at 20°C (10^{-6} Ωm)	Temp coeff (10^{-3}/°C)
Aluminium	0.027	4.3
Gold	0.022	4.0
Iron	0.105	6.6
Copper	0.0172	3.9
Nickel	0.078	6.7
Silver	0.016	3.8
Constantan	0.50	±0.03
Brass (die)	0.065	1.5
Steel (0.85% C)	0.18	–

The resistance produces a power loss that increases the temperature of the conductor. The term current density (S) - current/conductor area - can be used to select appropriate conductor dimensions in terms of increases in temperature. In the case of normal copper wire, short or loose conductors can have a current density of 6–10 A/mm². In large transformers in electronics equipment, 2.5 A/mm² is a common value, and in small transformers 3–3.5 A/mm². Use the following formula if you want to work out the diameter that corresponds to a particular current density and current:

$$d = 1.13 \sqrt{(I/S)}$$

where d is the diameter, I is the current and S is the current density.

In the table *Data of copper wire* that follows, the current for different wire diametres stated when S = 3 A/mm².

Under high frequencies, the electrons generally move along the surfaces of the conductor (the skin effect). In VHF and UHF, therefore, the wires used often have better conduction properties along the periphery (e.g. silver-plated wires) or have a larger surface area in relation to the size (so-called litz wire consisting of a large number of individually insulated conductors).

Conductors must be *insulated* with a suitable material before they can be placed next to each other, near earthed metal objects or other live conductors. By far the most widely used insulation material is polyvinylchloride (PVC). Other common insulators include rubber or EP rubber and plastics like polyethylene (PE), polypropylene (PP), polyurethane (PUR), polyamide (nylon), polytetrafluorethene (PTFE) (the Du Pont trade name is Teflon), FEP (Teflon FEP), silicone rubber and neoprene.

In components like transformers, various kinds of chokes and relays, enamelled wire is used, which is available in different temperature classes. It is often convenient to use directly solderable wire, but in transformers and solenoids, which generate a large amount of heat, it is preferable to use a more heat resistant wire in which the enamel has to be scraped away.

In coaxial cables, there is a solid or foam polyethylene layer between the inner conductor and the shield, with the outer cover made of PVC. In miniature cables and special low-loss cables, PTFE is used between the inner and outer conductor.

A **coaxial cable** has a characteristic impedance, making it suitable for use with high frequencies. The shield protects against electromagnetic high frequency fields. In lower frequencies, it only provides electrostatic protection.

To prevent electromagnetic interference, **twisted pair** is the best choice. Special audio cables use twisted pair enclosed in a single shield. There is sometimes an extra foil inside the shield to provide additional interference protection. There are also multicore cables with conductors shielded in pairs.

OPTICAL FIBRE can transmit light generated by an LED or laser to a light detector. The principle behind fibre optic cable is that the incoming beam, which is at a small angle in relation to the direction of the cable, is fully reflected by the walls of the cable because the core has a higher refractive index than the surrounding cladding.

Glass fibre cable has extremely good attenuation properties, only losing a few dB per km. Attenuation in plastic fibre is significantly higher, but it is a cheap alternative suitable for shorter distances, < 100 m, for example in a factory. Plastic fibre is cheap and easy to install, unlike glass fibre which requires special contacts and has a complex installation procedure. Plastic fibre generally has a core 1 mm in diameter, whereas a glass fibre cable can have a core diameter as low as 5-10 µm. See the factsheet on Fibre optic conductors.

Cable codes

Many different systems are used to identify and code cables. Some of the most widely used are presented below.

CENELEC is a European organisation whose function is to facilitate cross-border trade between its member countries by eliminating, as far as possible, any technical obstacles resulting from differences in national electricity regulations and standards. A cable designed and tested in accordance with a harmonisation document, HD, must show the HAR mark and a mark of origin.

Swedish codes

Example:

Mains Cable
RDOE Oil resistant (chloroprene rubber) rubber cable
REV Rubber cable for indoor use
RKK Round cable with plastic insulation
SKX Oval cable with plastic insulation
Low-current cable
EKKX Single-wire PVC-insulated telephone cable
RKUB Extra multi-stranded connection-cable for vehicles

German codes

LiYCY is a very common cable complying to a German code standard.

J Installation cable
S Signal cable
Li Multistrand conductor
C Braided copper shield
(L) Aluminium foil shield
Insulation and casing material
Y PVC
2Y PE
5Y PTFE
11Y PUR
2G Silicone rubber
5G Chloroprene rubber

Example: LiYCY = Multistrand, PVC insulation, shielded, PVC casing.

Power cables and installation cables – codes acc. to CENELEC

Construction of the designation

Example:	H	05	V		V		–		F	3	G	1.5				
Table:	1	2	3	4	5	6	–	7	8	9	10	11	+	9	10	11

Table 1 – Standard type

Symbol	Meaning
H	Cable corresponding to CENELEC standards
A	National standard accepted by CENELEC supplementing standard determined by CENELEC.

Table 2 – Rated voltage

Symbol	Meaning
01	100/100 V
03	300/300 V
05	300/500 V
07	450/750 V

Table 3 – Insulation

Symbol	Meaning
B	Ethylene-propylene rubber (EPR) for 90°C operating temperature
G	Ethylene-vinyl-acetate (EVA)
J	Glass-fibre braid
M	Mineral material
N	Chloroprene rubber (CR) or similar material
N2	Special chloroprene compound for welding cables acc. to HD 22.6
N4	Chlorosulphonated polyethylene (CSM) or chlorinated polyethylene
N8	Especially water durable polychloroprene rubber
Q	Polyurethane (PUR)
Q4	Polyamide
R	Ethylene-propylene rubber (normally) or corresponding synthetic elastomer for 60°C operating temperature.
S	Silicone rubber
T	Textile braid, impregnated or unimpregnated
T6	Textile braid, impregnated or unimpregnated, over each part in multi-conductor cables
V	Polyvinyl chloride (PVC) (normal)
V2	PVC for max. 90°C operating temperature
V3	PVC for cables installed at low temperatures
V4	Cross-linked PVC
V5	Oil resistant PVC
Z	Polyolefin-based cross-linked compound with little emission of corrosive gases, suitable for cables with low smoke emission when burning
Z1	Polyolefin based thermoplastic-compound with little emission of corrosive gases, suitable for cables with low smoke emission when burning

Table 4 – Metal sheaths, concentric conductors, screens

Symbol	Meaning
C	Concentric copper conductor
C4	Copper screen, constructed as braid over cabled parts

Table 5 –Non-metallic sheaths

See contents of table 3.

Table 6 – Special constructions and special designs

Symbol	Meaning
D3	Construction with tensile strength made up of one or many components of textiles or metal placed in the center of a round cable or distributed in a flat cable
D5	Filling (no tensile strength) in the center of a cable (only for lift cables)
(none)	Cable with circular cross-section
H	Flat, divideable cablel, sheathed or unsheathed
H2	Flat, undivideable cable
H6	Flat cable with 3 or more conductors, acc to HD 359 or EN 50214
H7	Cable with dual layer of extruded insulation
H8	Coiled conductor

Table 7 – Conductor material

Symbol	Meaning
(none)	Copper
-A	Aluminium

Table 8 – Conductor shape

Symbol	Meaning
-D	Multi-strand conductor for use in welding cable acc. to HD 22 Part 6 (other flexibility than acc. to HD 383, Class 5)
-E	Fine-strand conductor for use in welding cable acc. to HD 22 Part 6 (other flexibility than acc. to HD 383, Class 6)
-F	Multi-strand conductor for connection cable (flexibility acc. to HD 383, Class 5)
-H	Multi-strand conductor for connection cable (flexibility acc. to HD 383, Class 6)
-K	Multi-strand conductor for cable for fixed installation (if nothing else is stated, flexibility acc. to HD 383, Class 5)
-R	Few-strand, round conductor
-U	Massive, round conductor
-Y	Spinning conductor

Table 9 – No. of conductors

Symbol	Meaning
(number)	No. of conductors

Table 10 – Protective conductor

Symbol	Meaning
X	Green/yellow protective conductor missing
G	Green/yellow protective conductor included

Table 11 – Conductor area

Symbol	Meaning
(number)	Nominal conductor area (in mm^2)
/	Limiting signs before a number stating the area (in mm^2) for concentric conductors
Y	Spinning conductor, conductor area not stated

Publ. w. permission of the Swedish Electrical Commission, SEK.

Power, control and installation cables–codes acc. to Swedish standard SS 424 17 01

First letter – Conductor

A Aluminium
B Aluminium alloy
E Copper, single stranded (class 1)
F Copper, coarse stranded (class 2)
J Steel wire
R Copper, fine stranded (class 5)
S Copper, extra-fine stranded (class 6)

Second letter – Insulation

B Flame protected thermoplastic polyolefin (halogen-free, low smoke emission)
C Impregnated paper
D Rubber + rubber sheath
E Ethylene-propylene rubber
H Silicone rubber
I Urethane plastic
K PVC
L Polyethylene (PE)
O Neoprene rubber
Q Flame protected thermoplastic polyolefin (halogen-free, low smoke emission)
T Fluoride plastic
V Rubber without outer sheat
X Crossed-linked polyethylene (PEX)
Z Flame protected cross-linked polyolefin (halogen-free, low smoke emission)

Third letter – Sheath or other construction detail

A Aluminium-foil screen
B Flame protected thermoplastic polyolefin (halogen-free, low smoke emission)
C Concentric copper wire
F Copper wire braid
I Urethane plastic sheath
J Steel wire reinforcement
K PVC
L Plastic-coated aluminium-band screen
O Chloroprene rubber
P Zinc-coated steel-band reinforcement
Q Flame protected thermoplastic polyolefin (halogen-free, low smoke emission)
R Plastic-coated aluminium-band reinforcement
T Steel wire reinforcement
U No outer sheath
V Ethylene-propylene rubber
X PVC, oval cross-section
Z Flame protected cross-linked polyolefin (halogen-free, low smoke emission)

Fourth letter – Construction detail or use

B Hook-up wire
E Reinforced design
F Braid of copper or steel wire
H Lift cable
J Laying in ground
K PVC
P Reinforcement of zic-coated steel bands
R Control cable
S Self-supporting
T Heavy connection cable
V Laying in water
Z Cable for neon-light systems

Fifth letter – Construction detail or use

E Reinforced design
K PVC
L PE

Published with permission of the Swedish Electrical Commission, SEK.

Telecommunication cables – Codes acc. to Swedish standard SS 424 16 75

First letter – Optical or electrical conductor

A Aluminium, unclad
B Aluminium alloy
C Bronze
D Glass/plastic, fibre
E Copper, single-strand
F Copper, stranded
G Glass/glass, fibre
H Fibre bundle
J Copper-clad steel wire
K Coaxial pair
L Conducting plastic
M Copper, multi-strand
P Plastic/plastic, fibre
R Copper, extra multi-strand
S Copper, fine strand
T Copper, extra fine strand
Z Conductor with a yarn kernel

Second letter – Conductor insulation or secondary protection

A Acrylic coated fibre band
C Combination of cell and homogenous polyolefin
I Thermoplastic polyurethane elastomer
J Fibre without secondary protection
K PVC
L Polyethylene
M PP
N PA
O Thermoplastic elastomer
P Paper, unimpregnated
Q Halogen-free, flame protected material
R Polyester
S Slotted core
T Fluorethane plastic, e.g. PTFE, FEP
U Cellular polyolefin

Third letter – Cover or other construction detail

A Screen of aluminium band
B Lead sheath
C Combination of cellular and homogenous polyolefin
D Cable made of dielectric material only
E Individually screened parts or twisted groups
F Metal strand braid, metal strand spun or metal strain relief
G Non-metal reinforcement of braid, wire cover or strain relief
H Parts put around a strain relief
I Thermoplastic polyurethane elastomer (TPU)
J Steel band reinforcement
K PVC
L Polyethylene (PE)
M Metal sheath, ungrooved
N PA
O Thermoplastic elastomer
P Zinc-coated steel band reinforcement
Q Halogen-free, flame protected material
R Polyester
S Slotted core
T Zinc-coated steel wire reinforcement
U Without cover
W Metal sheat, grooved
X Oval cross-section
Z Copper band screen

Fourth letter – Construction detail or property

A Screen of aluminium band
B Connection wire
C Cable with embedded suspending wire
D Cable made of dielectric material only
E Reinforced version or low-capacitance cable
F Metal strand braid, metal strand spun or metal strain relief
G Non-metal reinforcement of braid, wire cover or strain relief
H Parts put around a strain relief
I Thermoplastic polyurethane elastomer (TPU)
J Steel band reinforcement
K PVC
L Polyethylene (PE)
M Metal sheath, ungrooved
N PA
O Thermoplastic elastomer
P Zinc-coated steel band reinforcement
Q Halogen-free, flame protected material
R Signal cable
S Self-supporting
T Zinc-coated steel wire reinforcement
U Fire-retardant cable
W Metal sheath, grooved
X Non weather-proof cable
Y Weather-proof cable
Z Copper band screen

Fifth letter – Construction detail or property

B Halogen-free flame-protected cable
C Cable with embedded suspending wire
D Cable made of dielectric material only
E Reinforced version or low-capacitance cable
F Metal strand braid, metal strand spun or metal strain relief
G Non-metal reinforcement of braid, wire cover or strain relief
H Parts put around a strain relief
I Thermoplastic polyurethane elastomer (TPU)
J Steel band reinforcement
K PVC
L Polyethylene (PE)
N PA
P Zinc-coated steel band reinforcement
Q Halogen-free, flame protected material
R Signal cable
T Zinc-coated steel wire reinforcement
U Fire-retardant cable
V Water block
W Metal sheath, grooved
X Non weather-proof cable
Y Weather-proof cable

The following letters specify properties and are stated in alphabetical order.

B Halogen-free, flame-protected cable
C Cable with suspending wire embedded in sheath
D Cable made up of dielectric material only
E Reinforced design or low-capacitance cable
H Parts put around strain relief
S Self-supporting cable
U Fire-retardant cable
V Water block
X Not weather-proof cable
Y Weather-proof cable

Published with permission by the Swedish Electrical Commission, SEK.

Cable colour-coding and numbering

Swedish telephone cable standard, e.g. EKKX

Pair no.	Colours of conductor pair	Pair no.	Colours of conductor pair	Pair no.	Colours of conductor pair	Pair no.	Colours of conductor pair
1	White – Blue	6	Red – Blue	11	Black – Blue	16	Yellow – Blue
2	White – Orange	7	Red – Orange	12	Black – Orange	17	Yellow – Orange
3	White – Green	8	Red – Green	13	Black – Green	18	Yellow – Green
4	White – Brown	9	Red – Brown	14	Black – Brown	19	Yellow – Brun
5	White – Grey	10	Red – Grey	15	Black – Grey	20	Yellow – Grey

High-current cable 450/750 V

Cond	Colour code
2	Light blue, brown
3	Yellow/green, brown, light blue
4	Yellow/green, brown, light blue, black
5	Yellow/green, black, brown, light blue, grey

DEF STAN 61-12 (British defence standard)

Cables with up to 25 conductors

No.	Colour	No.	Colour	No.	Colour
1	Red	10	Pink	19	Yellow/blue
2	Blue	11	Turquoise	20	White/blue
3	Green	12	Grey	21	Blue/black
4	Yellow	13	Red/blue	22	Orange/blue
5	White	14	Green/red	23	Yellow/green
6	Black	15	Yellow/red	24	White/green
7	Brown	16	White/red	25	Orange/green
8	Violet	17	Red/black		
9	Orange	18	Red/brown		

Cables with 26 to 36 conductors

No.	Colour	No.	Colour	No.	Colour	No.	Colour
1	Red	10	Pink	19	Yellow/blue	28	Orange/green
2	Blue	11	Turquoise	20	White/blue	29	Grey/green
3	Green	12	Grey	21	Blue/black	30	Yellow/brown
4	Yellow	13	Red/blue	22	Orange/blue	31	White/brown
5	White	14	Green/red	23	Green/blue	32	Brown/black
6	Black	15	Yellow/red	24	Grey/blue	33	Grey/brown
7	Brown	16	White/red	25	Yellow/green	34	Yellow/violet
8	Violet	17	Red/black	26	White/green	35	Violet/black
9	Orange	18	Red/brown	27	Green/black	36	White/violet

DIN 47100 (twisted pair and multicore)

Pair no.	Core no.	Colour	Pair no.	Core no.	Colour	Pair no.	Core no.	Colour
1	1	White	9	17	White/grey	17	33	Green/red
1	2	Brown	9	18	Grey/brown	17	34	Yellow/red
2	3	Green	10	19	White/pink	18	35	Green/black
2	4	Yellow	10	20	Pink/brown	18	36	Yellow/black
3	5	Grey	11	21	White/blue	19	37	Grey/blue
3	6	Pink	11	22	Brown/blue	19	38	Pink/blue
4	7	Blue	12	23	White/red	20	39	Grey/red
4	8	Red	12	24	Brown/red	20	40	Pink/red
5	9	Black	13	25	White/black	21	41	Pink/red
5	10	Violet	13	26	Brown/black	21	42	Pink/black
6	11	Grey/pink	14	27	Grey/green	22	43	Blue/black
6	12	Blue/red	14	28	Yellow/grey	22	44	Red/black
7	13	White/green	15	29	Pink/green			
7	14	Brown/green	15	30	Yellow/pink			
8	15	White/yellow	16	31	Green/blue			
8	16	Yellow/brown	16	32	Yellow/blue			

AWG table

AWG stands for American Wire Gauge, an American system for denoting conductor thicknesses. With each consecutive AWG-size, the thickness is changed by a constant factor. The AWG system was devised in 1857 by J.R. Brown.

AWG dim.	Cond. structure ×AWG	Cond. structure ×diam (mm)	Conductor cross-s. (mm^2)	Cond diam uninsulated (mm)
1	1×1	1×7.35	42.4	7.35
1	259×25	259×0.45	42.1	9.50
1	817×30	817×0.25	41.4	9.70
2	1×2	1×6.54	33.6	6.54
2	133×23	133×0.57	34.4	8.60
2	665×30	665×0.25	33.8	8.60
3	1×3	1×5.83	26.7	5.83
3	133×24	133×0.51	27.2	7.60
4	1×4	1×5.19	21.1	5.19
4	133×25	133×0.45	21.6	6.95

(continued)

(Continued)

AWG dim.	Cond. structure ×AWG	Cond. structure ×diam (mm)	Conductor cross-s. (mm^2)	Cond diam uninsulated (mm)
5	1×5	1×4.62	16.8	4.62
6	1×6	1×4.11	13.2	4.11
6	133×27	133×0.36	13.6	5.51
7	1×7	1×3.66	10.5	3.67
8	1×3.66	1×3.26	8.37	3.26
8	133×29	133×0.29	8.61	4.38
9	1×9	1×2.91	6.83	2.91
10	1×10	1×2.59	5.26	2.59
10	105×30	105×0.25	5.32	2.85
11	1×11	1×2.30	4.17	2.30
12	1×12	1×2.05	3.31	2.05
12	18×25	18×0.45	3.09	2.24
12	37×28	37×0.32	2.99	2.31
13	1×13	1×1.83	2.70	1.83
14	1×14	1×1.63	2.08	1.63
14	18×27	18×0.36	1.94	1.76
14	41×30	41×0.25	2.08	1.83
15	1×15	1×1.45	1.65	1.45
16	1×16	1×1.29	1.31	1.29
16	18×29	18×0.29	1.23	1.40
16	28×30	28×0.25	1.32	1.47
17	1×17	1×1.15	1.04	1.15
18	1×18	1×1.03	0.824	1.03
18	7×26	7×0.40	0.897	1.03
18	19×30	19×0.25	0.963	1.02
19	1×19	1×0.91	0.653	0.91
20	1×20	1×0.81	0.519	0.81
20	7×28	7×0.32	0.563	1.01
20	10×30	10×0.25	0.507	0.97
21	1×21	1×0.72	0.412	0.72
22	1×22	1×0.64	0.325	0.64
22	7×30	7×0.25	0.355	0.80
22	19×34	19×0.16	0.382	0.78
23	1×23	1×0.57	0.259	0.57
24	1×24	1×0.51	0.205	0.51
24	7×32	7×0.20	0.227	0.64
24	19×36	19×0.13	0.241	0.62
25	1×25	1×0.45	0.163	0.45
26	1×26	1×0.40	0.128	0.40
26	7×34	7×0.16	0.140	0.50
26	19×38	19×0.10	0.154	0.50
27	1×27	1×0.36	0.102	0.36
28	1×28	1×0.32	0.080	0.32
28	7×36	7×0.13	0.089	0.40
28	19×40	19×0.08	0.092	0.39
29	1×29	1×0.29	0.065	0.29
30	1×30	1×0.25	0.051	0.25
30	7×38	7×0.10	0.057	0.33
30	19×42	19×0.06	0.057	0.36
31	1×31	1×0.23	0.040	0.23
32	1×32	1×0.20	0.032	0.20
32	7×40	7×0.08	0.034	0.26
33	1×33	1×0.18	0.025	0.18
34	1×34	1×0.16	0.020	0.16
34	7×42	7×0.06	0.022	0.21
35	1×35	1×0.14	0.016	0.14
36	1×36	1×0.13	0.013	0.13
37	1×37	1×0.11	0.010	0.11
38	1×38	1×0.10	0.009	0.10
39	1×39	1×0.09	0.006	0.09
40	1×40	1×0.08	0.005	0.08
41	1×41	1×0.07	0.004	0.07
42	1×42	1×0.06	0.003	0.06

Data on copper wire

Diam Bare mm	Diam Enamelled mm	Area mm^2	AWG no	Resistance at 20°C Ω/km	Current at 3 A/mm^2 mA	Length m/100 g	Weight 100 g/km
0.04	0.05	0.0013	46	13700	3.8	8200	0.12
0.05	0.06	0.0020	44	8750	6	5400	0.18
0.06	0.07	0.0028	42	6070	9	3800	0.22
0.07	0.08	0.0039	41	4460	12	2800	0.35
0.08	0.09	0.0050	40	3420	15	2100	0.47
0.09	0.11	0.0064	39	2700	19	1700	0.59
0.10	0.12	0.0078	38	2190	24	1400	0.71
0.11	0.13	0.0095	37	1810	28	1100	0.91
0.12	0.14	0.011		1520	33	950	1.00
0.13	0.15	0.013	36	1300	40	820	1.21
0.14	0.16	0.015	35	1120	45	710	1.40
0.15	0.17	0.018		970	54	620	1.60
0.16	0.18	0.020	34	844	60	560	1.80
0.17	0.19	0.023		757	68	490	2.05
0.18	0.20	0.026	33	676	75	440	2.25
0.19	0.21	0.028		605	85	390	2.55
0.20	0.22	0.031	32	547	93	360	2.77
0.25	0.27	0.049	30	351	147	230	4.35
0.30	0.33	0.071	29	243	212	160	6.25
0.35	0.38	0.096	27	178	288	120	8.35
0.40	0.43	0.13	26	137	378	90	11.15
0.45	0.48	0.16	25	108	477	70	14.10
0.50	0.53	0.20	24	87.5	588	57	17.50
0.55	0.58	0.24		72.3	715	47	21.01
0.60	0.64	0.28		60.7	850	40	25.0
0.65	0.69	0.33	22	51.7	1.0 A	34	29.4
0.70	0.74	0.39		44.6	1.16	29	34.5
0.75	0.79	0.44		38.9	1.32	25	40.0
0.80	0.84	0.50	20	34.1	1.51	22	45.5
0.85	0.89	0.57		30.2	1.70	20	50.0
0.90	0.94	0.64	19	26.9	1.91	18	55.5
0.95	0.99	0.71		24.3	2.12	16	62.5
1.00	1.05	0.78	18	21.9	2.36	14	71.5
1.10	1.15	0.95		18.1	2.85	12	83.5
1.20	1.25	1.1		15.2	3.38	10	100.0
1.30	1.35	1.3	16	13.0	3.97	8.5	118.0
1.40	1.45	1.5		11.2	4.60	7.5	140.0
1.50	1.56	1.8		9.70	5.30	6.4	155.0
1.60	1.66	2.0	14	8.54	6.0	5.5	179.0
1.70	1.76	2.3		7.57	6.7	5.0	200.0
1.80	1.86	2.6	13	6.76	7.6	4.5	225.0
1.90	1.96	2.8		6.05	8.5	4.0	250.0
2.00	2.06	3.1	12	5.47	9.40	3.5	285.5

Coils and chokes

Inductive components such as coils and chokes exist for frequency selective purposes. A coil which is mostly used for suppressing AC voltages is often known as a choke.

As a rule, coils and chokes consist of a number of turns of copper wire wound tightly together, with or without some form of core. They are made in various models, with inductances from a few nH up to tens of H (henry).

Inductance is the property of a coil which counteracts all changes in the current flowing through it. This is done via an opposite-direction voltage which occurs in the coil, known as counter EMF (electromotive force). A coil with an inductance of 1 H has a counter EMF of 1 V as the current through it changes at a rate of 1 A/s (1 H=1 Vs/A).

Application areas

The following are examples of areas of application for coils and chokes:

Tuned filters (oscillatory circuits). For selecting or blocking certain frequencies. Here, attempts are made to attain coils with a high Q factor and with good stability. The coils are often air-cored or have cores of iron powder or ferrite (if so, often with an air gap). Toroids and adjustable potcores with or without shielding are common.

RFI filters. For damping undesirable high-frequency signals (interference). The coil must have high impedance over a broad frequency range (low Q factor). Ferrite cores are appropriate for this. If the current is low, toroidal cores are often used which have a closed magnetic circuit and a small interference field. In the case of higher currents, an air gap is introduced, or a core is used with an open magnetic circuit, such as a ferrite rod.

*DC current filtration and storage of energy.*As a choke in switching power supplies, for example, in order to filter out high-frequency ripple, and as an energy storage choke in DC/DC converters. Here, it is important for the choke to be able to withstand high DC currents without saturation of the core material (high saturation flux density). Iron powder is the most common core material used for these applications.

Impedance of the coil

Coils have a frequency-dependent resistance known as reactance, and a DC voltage resistance, which is the resistance in the wire. The *inductive reactance* (X_L) is calculated on the basis of the formula

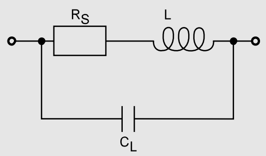
$$X_L = \omega L$$

where ω = angle frequency($2 \times \pi \times f$), in rad/s, f = frequency in Hz and L = inductance in Henry.

The impedance (Z) of the coil at a certain frequency is the combination of resistance and reactance:

$$Z = \sqrt{(X_L^2 + R^2)}$$

In order to understand more readily the coil as a component, we can use a simplified equivalent diagram:



The equivalent diagram of the coil. L = induktance, R_S = series resistance (resistance in the wire + other losses in the wire and core), C_L = self-capacitance in the coil, e.g. the capacitance between wire turns (also known as distributed capacitance, parasitic capacitance, stray capacitance).

Q factor (Q = Quality), or the factor of merit of the coil, is the ratio of the coil's reactance to series resistance. A lower resistance gives a higher Q factor and steeper filters.

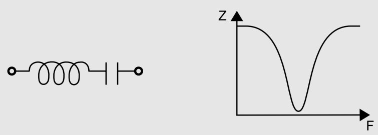
$$Q = X_L/R_S$$

Resonance

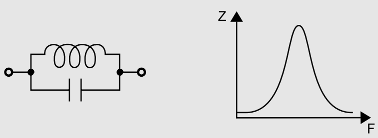
Together with a capacitor, a coil forms an oscillatory circuit. This circuit has a resonance frequency, i.e. a frequency at which the coil and the capacitor reactance are the same. At this frequency, the total impedance is at its lowest if they are connected in series, and at its highest if they are connected in parallel. The formula for the *resonance frequency* is:

$$f = 1/(2\pi \times \sqrt{(LC)})$$

The frequency is expressed in Hz if L is expressed in H and C is expressed in F. If L and C are expressed in µH and µF respectively, the frequency is expressed in MHz.



Serial resonance circuit.



Parallel resonance circuit.

Self-capacitance (C_L) in the coil forms an oscillatory circuit with the inductance. Its resonance frequency is known as *self resonance frequency* (SRF). Self-capacitance can cause problems at high frequencies if it is not taken into consideration. The test frequency for the Q factor should be at most a tenth of this frequency.

Calculations on coils without cores

The energy stored in a choke can be interesting. It can be calculated using the formula:

$$W = 1/2\ L \times I^2$$

where W = energy in Joules, L = coil inductance in henry and I = current in ampere through the coil.

Before we go into the calculation of a coil, we should state that even a straight wire gives rise to an inductance. This is something that one should bear in mind as regards HFs. It is a matter, then, of keeping cables as short as possible, e.g. in series with a decoupling capacitor, otherwise an oscillatory circuit could occur.

The formula for *the inductance in a wire* is:

$$L = 0{,}002\ s\ (\ln(4\ s/d) - x)$$

With the length (s) and diameter (d) of the wire expressed in cm, the inductance is expressed in µH. The factor x depends on the frequency and form. A straight wire and a high frequency will give x = 1, a low frequency will give x = 0.75. If the wire is bent, the inductance will be lower. A circle with one turn will give x = 2.45 at high frequencies and 2.20 at low frequencies, while a square will give 2.85 and 2.60 respectively.

If one wishes to increase the inductance, it is possible to surround the cable with a magnetic material such as a ferrite bead, or to wind several turns of the wire in a spiral. In the latter case, the wire is bent but the combined effect is great. The inductance of the coil increases in proportion with the square of the number of turns in the coil.

The inductance of a *single-layer, air-cored coil* can be calculated according to the formula:

$$L = (0{,}08d^2n^2) / (3d + 9s)$$

where the coil length (s) and diameter (d) are expressed in cm. n is the number of turns. The coil inductance is expressed in µH. The highest Q factor is obtained if the coil length is between 2 and 2.5 times greater than its diameter. The coil diameter should be greater than 5 times the diameter of the wire.

A *short, multi-layer, air-cored coil* can be calculated according to the formula:

$$L = (0{,}08d^2n^2) / (3d + 9s + 10a)$$

where d = average diameter and a = radial thickness of the winding, all in cm. The inductance is expressed in µH.

For *printed coils* which are etched on a PCB laminate with a sheet thickness of 35 µm, the inductance is calculated using the following formula:

$$L = nD_m\ (nK_1 + K_2)$$

where L = inductance in µH,
n = number of turns and
D_m = (average diameter of the coil in cm.

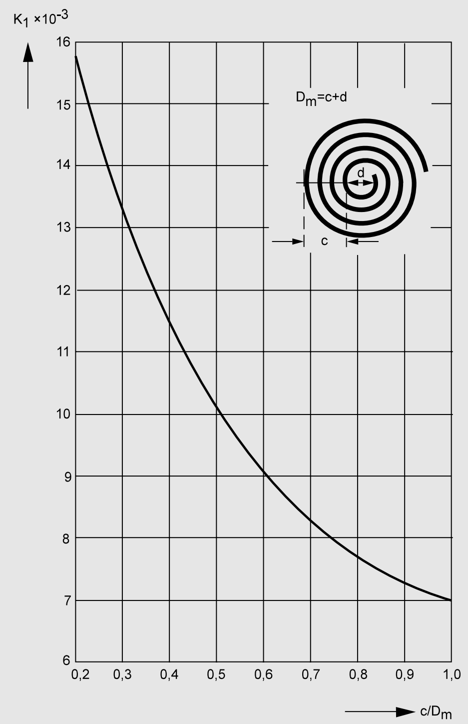
K_1 and K_2 are constants which depend on the shape of the coil. See diagram below for calculation of D_m, K_1 and K_2.

Example: Calculate the inductance of a printed spiralized coil with $d_1 = d_2 = 0.5$ mm, sheet thickness 35 µm, 14 turns and d = 10 mm.

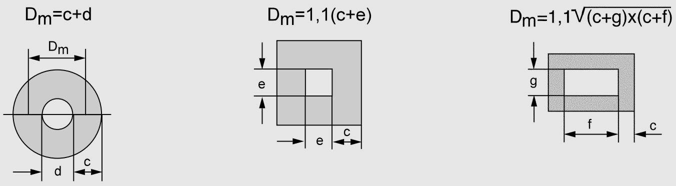
From the diagram below you can determine:
$c = n\ (d_1 + d_2) = 14\ (0.5 + 0.5)$ mm = 1.4 cm
$D_m = c + d = 1.4 + 1.0$ cm = 2.4 cm
$c/D_m = 1.4/2.4 \approx 0.58$ which gives $K_1 = 9.2 \times 10^{-3}$
$(d_1 + d_2) / d_1 = (0.5 + 0.5)/0.5 = 2$ which gives $K_2 = 3.5 \times 10^{-3}$

From this you get:
$L = nD_m\ (nK_1 + K_2) = 14 \times 2.4\ (14 \times 9.2 \times 10^{-3} + 3.5 \times 10^{-3})$ µH = 4.45 µH.

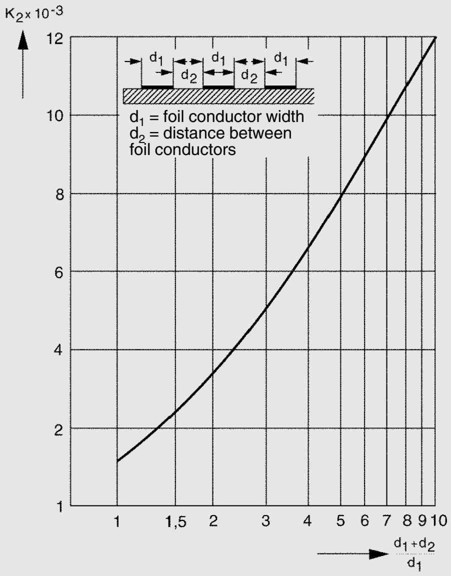
Calculation of K_1 for printed coils.



Calculation of D_m for printed coils.



Calculation of K_2 for printed coils.



Coils with cores

To increase inductance, as already noted, it is possible to provide a core and/or a case made of a ferromagnetic material. The most common materials are ferrites and iron powder. These are known as magnetically soft materials, i.e. they lose most of the magnetic flux when the field strength is taken away. The opposite, magnetic hardness, can be seen in permanent magnets.

Ferrite is a sintered ceramic, microcrystalline cubic material and consists of iron oxide (Fe_2O_3) and a combination of metals. The most common combinations are manganese/zinc (MnZn) and nickel/zinc (NiZn).

MnZn ferrites have the highest permeability (μ_i) and saturation flux density (B_s), while NiZn ferrites have higher resistivity (lower losses) and are best suited to frequencies greater than 1 MHz.

Ferrite offers advantages such as very high permeability (μ_i 100–10000), low losses and high frequency capability, but it has a low saturation flux density (B_S <0.5 T). This means that ferrite is saturated very easily and care must be taken where high DC currents are involved. One way of surmounting this problem is to use a ferrite core with an open magnetic circuit, e.g. a rod, or to introduce an air gap somewhere in the circuit.

Ferrite cores are used in HF inductances and RFI filters and in power transformers up to 1 MHz. They are manufactured as toroids, potcores (potcores, RM cores, etc.), C- and E cores (and variants of these), rods, threaded rods, beads, blocks, etc.

Iron powder cores, as the name suggests, consist of powdered iron in which the particles are isolated from one another by oxidising the surface, for instance. When a binding agent has been added, the material is compressed into the shape of the core and baked in a furnace.

The greatest advantage of iron powder cores compared with ferrite cores is that they are able to withstand high currents through the wire, the saturation flux density (B_S) is approx 1.5 T. They are also temperature-stable, give good Q factors and are able to withstand high frequencies. The major disadvantage is their low permeability ($\mu_i = 2–90$). This results from the fact that the large number of small air gaps between all the iron particles in total make a large air gap (distributed air gap).

Iron powder cores are used primarily in chokes for the filtration of DC voltage or low-frequency (50 Hz) AC voltage. They are also used as energy storage chokes in switch regulators, tuned filters and in impedance matches at high frequencies, for example. Iron powder cores are manufactured chiefly as toroids.

Iron cores are used almost exclusively for mains transformers, as the losses (induced eddy currents due to low resistivity), despite lamination of the core, are so great that frequencies in excess of 1 kHz are impractical to handle.

When designing coils with ferromagnetic cores, it is necessary to be familiar with magnetism to a certain extent. Therefore, we shall begin with a little basic theory on magnetism.

Magnetic field

When a current passes through a coil which is wound on a core, a *magnetomotive force* (mmf) occurs, which in turn gives rise to a magnetic flux (Φ) through the core. The magnitude of this flux is dependent on the core's reluctance (R_m). Reluctance can be regarded as "magnetic resistance" (by way of analogy with Ohm's law, E = I × R).

$$\text{mmf} = \Phi \times R_m$$

mmf is measured in ampere-turns (N × I), but is written A as turns are unitless. Sometimes you see this written At (ampere-turns), although this is incorrect, but simpler to understand. The magnetic flux is expressed in Weber (Wb).

If the mmf is viewed in relation to the effective magnetic path length (l_e) in metres, you get **magnetic field strength** (H) in A/m (or At/m).

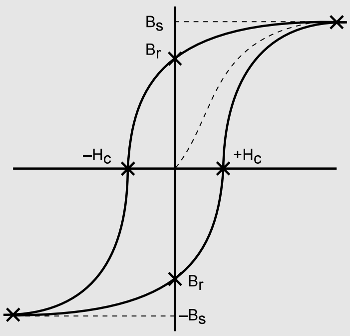
$$H = N \times I/l_e$$

Thus the field strength is the number of turns times the current divided by the magnetic path length. Note that the path length is not the same as the physical length of the core.

The flux density (B), or induction as it is also known, is the flux (Φ) divided by the effective magnetic area (A_e):

$$B = \Phi/A_e$$

The flux density (B) is expressed in tesla (T). 1 T = 1 Wb/m^2.



Hysteresis curve.

The hysteresis curve (B-H loop) is a method of showing the flux density (B) of a material in relation to the field strength (H). In a ferromagnetic material in a state of rest, there is a chaos of molecular magnets which point randomly in different directions. These magnets "cancel out" one another. When a magnetic field is applied, the molecular magnets will turn and point in the same direction as the magnetic flux, and more will do this the higher the field strength (H). When all the molecular magnets are pointing in the same direction, the material is saturated (B_s), and no higher a flux density can be achieved even if the field strength (H) is further increased. When the field strength is reduced, the curve does not follow its old path as a number of molecular magnets do not return entirely to their original position. When the field strength is zero (H = 0), there is a certain amount of flux left in the material. This flux density is known as remanence (B_r). An opposite-direction flux is required in order to return the flux to zero. The field strength required for this is known as coercive force (H_C) or coercivity.

Permeability

The flux density (B) can be viewed in relation to the field strength (H):

$$B = \mu \times H$$

where μ is the *permeability* and can be compared to "magnetic conductivity" (cf "reluctance" above). Viewed graphically, the permeability is the slope on the hysteresis curve. Permeability is a broad term, and is actually $\mu_o \times \mu_r$, where μ_o is the permeability in an absolute vacuum, and μ_r is the permeability of the material relative to μ_o. For example, $\mu_r = 100$ means that the permeability of the material is 100 times higher than the permeability in a vacuum. Thus the formula can be written as follows:

$$B = \mu_o \times \mu_r \times H$$

The permeability in a vacuum is $4\pi \times 10^{-7}$ (H/m)

In a closed magnetic circuit such as a toroid, μ_r is known as *initial permeability* μ_i (or toroidal permeability, μ_{tor}). This gives a correct value only with a small flux density (B<0,1 mT). μ_i is most often the permeability given by manufacturers in their material specifications.

In a magnetic circuit with an air gap, μ_r is known as *effective permeability*, μ_e. The relationship between μ_e and μ_i is described using the formula:

$$\mu_e = \mu_i / (1 + (G/l_e \times \mu_i))$$

where G = length of the air gap and l_e = magnetic path length.

Since the permeability of a material is not linear in relation to B and H (see the hysteresis curve), we also refer to other types of permeability.

Amplitude permeability (μ_a), which is the permeability when only an AC current passes through the coil. Even at just a few mT, there can be a great discrepancy from μ_i. Greatest is the around half saturation flux density (B_s): here it can be 2–3 times as great as μ_i. Thus the permeability varies depending on the field strength.

Reversible permeability or incremental permeability ($\mu\Delta$) is when there is an AC current superimposed on a DC current, e.g. a filter choke in a power supply. Here, the permeability varies depending on the magnitude of the field strength. An iron powder core keeps the permeability high up to 10 000s of A/m, while ferrite is saturated even at a few hundred A/m and loses all permeability.

Magnetic losses

When referring to *complex permeability*, we take into account the losses in the coil.

To account for the magnetic losses, we add a resistive term to the permeability.

$$\mu = \mu_s^{I} - j\mu_s^{II}$$

where $\mu_s^{I} = \mu_i$ and $\mu_s^{II} = \tan\delta \times \mu_i$. In manufacturers' datasheets, it is often possible to deduce μ_s^{I} och μ_s^{II} versus frequency directly from a diagram.

The magnetic losses ($\tan\delta_m$) can be divided into three parts: hysteresis loss ($\tan\delta_h$), which is dependent on the flux density (B), eddy-current loss ($\tan\delta_F$), which is frequency dependent, and a residual loss ($\tan\delta_r$), which is constant.

$$\tan\delta_m = \tan\delta_h + \tan\delta_F + \tan\delta_r$$

In the datasheets, a loss factor $\tan\delta/\mu_i$) can be deduced at a given frequency. $\tan\delta/\mu_i$ increases logarithmically in relation to the frequency. Here, we take into account eddy-current losses and residual losses ($\tan\delta_F + \tan\delta_r$), but not hysteresis losses ($\tan\delta_h$). To show the hysteresis loss, the hysteresis constant (η_B) is given. The hysteresis loss for a specific flux density can be calculated from this constant.

$$\tan\delta = \eta_B \times B \times \mu_i$$

For a core with an air gap, the magnetic losses ($\tan\delta\mu$) can be multiplied by the ratio μ_i/μ_e. Except for losses in the core, there are losses in the wire ($\tan\delta_w$). These wire losses can also be divided into three parts:resistive loss ($\tan\delta_R$), which is the resistance in the wire, eddy-current loss ($\tan\delta_C$), which is frequency-dependent, and dielectric losses in the insulation ($\tan\delta_d$), which can be viewed as a series resistance to the self-capacitance. The latter two are relatively small compared with the resistive loss (at moderate frequencies).

$$\tan\delta_w = \tan\delta_R + \tan\delta_C + \tan\delta_d$$

The skin effect

The resistive loss ($\tan\delta_R$) can be regarded as the DC resistance if the frequency is not in excess of 50 kHz. If the frequency is higher, what is known as the "skin effect", which increases the AC resistance, should be taken into account.

When a current passes through a cable, a magnetic field is formed, not only around, but also within the cable. This magnetic field inside the cable, which is at right angles to the current direction, in turn induces an eddy current lengthwise along the cable. The permeability of copper is low (($\mu_r \sim 1$), but its resistivity is also low, which means that the eddy currents at frequencies in excess of 50 kHz may be considerable. The longitudinal eddy currents travel against the current direction in the centre of the cable, and with the current direction along the outer edge of the cable. This gives a current concentration in the outer edge of the cable and thereby reduces the active area of the cable, which in turn increases the resistance.

The term **skin depth** means the depth at which the current density is decreased to 37 %(1/e). This depth is also the same as the wall thickness which a tube of the same length with a DC resistance corresponding to the AC resistance of the cable would have. This depth can be calculated using the formula

$$\delta = 1/\sqrt{(f\mu\pi\rho)}$$

where δ = skin depth in metres, f =frequency in Hz, μ = permeability $\mu_o \times \mu_r$ and ρ =conductivity in S/m. In the case of copper, $\mu_r = 1$ and $\rho = 5.8 \times 10^7$. The resistance can then be calculated using the formula:

$$R_{AC} = R_{DC} \times A / (2\pi \times r \times \delta) = R_{DC} \times r / (2 \times \delta)$$

where R_{AC} = AC resistance, R_{DC} = DC resistance, A = wire area, r = wire radius, and δ = skin depth.

Moreover, a wire in a coil induces eddy currents in adjacent wires, which further increases the AC resistance.

One way in which the effect of the eddy currents can be reduced is to use what is known as *litz wire* instead of a solid conductor. Litz wire consists of a number (3 to 400) of insulated strands bundled together which constantly change position within the bundle. The AC resistance of litz wire is the same as its DC resistance.

The resistance in a copper cable is approx 30 % higher at 100 °C than at 25 °C.

If there is a shield (copper can) or a component made of a ferromagnetic material (e.g. X7R or Z5U capacitor) in the vicinity, losses occur in it ($\tan\delta_s$). These losses are often regarded as negligible.

The total loss in a coil is:

$$\tan\delta = \tan\delta_m + \tan\delta_w + \tan\delta_s$$

As a rule, the best Q factor is obtained when the wire losses are as great as the core losses.

Calculations for coils with cores

To be able to calculate a core simply, you indicate in the datasheet the *effective magnetic dimensions*, known as the effective path length l_e, the effective area A_e and the effective volume V_e. If the core is not a toroid, the dimensions for a toroid with similar properties are given. The ratio $\Sigma l_e/A_e$ is known as the core factor. In datasheets from European manufacturers, l_e, A_e and V_e are often expressed in mm (mm^2, mm^3), but in cm (cm^2, cm^3) in American ones.

$1\ mm^2 = 10^{-6}\ m^2$	$1\ mm^{-1} = 10^3\ m^{-1}$
$1\ cm^2 = 10^{-4}\ m^2$	$1\ cm^{-1} = 10^2\ m^{-1}$
$1\ mm^3 = 10^{-9}\ m^3$	
$1\ cm^3 = 10^{-6}\ m^3$	

To calculate the inductance, the following formula is used:

$$L = \mu_o \times N^2 / ((1/\mu_r) \times (\Sigma\ l_e / A_e))$$

which can also be expressed as follows:

$$L = N^2 \times \mu_o \times \mu_r / (\Sigma\ l_e / A_e)$$

To simplify calculations, the permeabilities and core factor are often removed and an *inductance factor* A_L given.

$$A_L = \mu_o \times \mu_r / (\Sigma\ l_e / A_e)$$

If these two formulae are combined, we get:

$$L = N^2 \times A_L$$

The A_L value is most often expressed in nH/N^2.

Example: A 100 µH coil is required, and the core has an A_L value of 800 nH/N^2. If we remove N from the above formula, we get:

$$N = \sqrt{(L / A_L)}$$

$$N = \sqrt{(100000/800)} \approx 11 \text{ turns}$$

Remember to express L in nH if A_L is expressed in nH/N^2.

Ferrite rods are not only used as antenna cores, but often also as cores in HF and RFI chokes. They have an open magnetic circuit which means that it is possible to run high currents through the coil without saturating the core. The permeability (μ_{rod}) is, with the exception of the initial permeability (μ_i), dependent on the ratio of length to diameter. μ_{rod} can be deduced from the figure below.

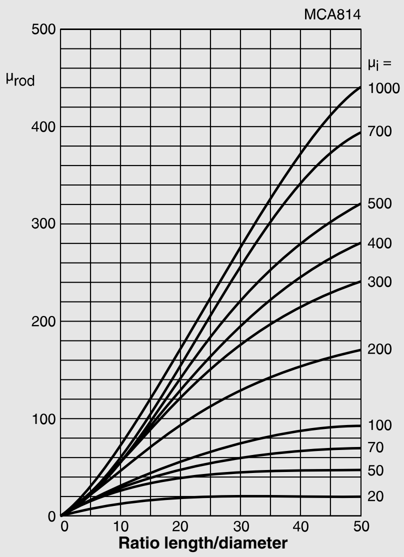


Diagram showing the permeability of the ferrite rod as a function of the ratio of its length to its diameter.

As the inductance is strongly dependent on the length of the winding and its location on the rod, it is difficult to give an A_L value. Instead, the inductance has to be calculated using the formula:

$$L = \mu_o \times \mu_{rod} \times N^2 \times A / l$$

where $\mu_o = 4\pi \times 10^{-7}$, μ_{rod} = permeability of the rod which can be seen from the diagram, N = number of turns, A = area of the rod and l = the length of the winding centred on the rod.

Flux density of the core

It is important to calculate the flux density (B) in the core in order to avoid saturation (B_S). A saturated core has a permeability of 1 ($\mu_r = 1$) and thereby an inductance as if it were air-cored. In addition, great losses occur, which results in heat, particularly at high frequencies. There are a number of methods which can be used to calculate the flux density, such as by first calculating the field strength (H) using the formula:

$$H = N \times I / l_e$$

and then calculating the flux density using the formula:

$$B = \mu_o \times \mu_r \times H$$

The flux density (B) is expressed in tesla (T) in the above case and in all cases below. If the current is purely direct, the following formula can be used:

$$B = L \times I / (N \times A_e)$$

where L = inductance, I = current, N = number of turns and A_e = effective area. For full-wave rectified, unfiltered DC voltage, the following applies:

$$B = U_{eff} / (19 \times N \times A_e \times f)$$

where U_{eff} = effective value of the ripple voltage and f = frequency. Often, there is a DC voltage with a certain ripple. Then it is necessary to calculate this, or, if one can content oneself with slightly too high an approximate value, to calculate the peak voltage as DC voltage.

If the voltage is an AC voltage, the following formula for sine waves can be applied:

$$B = \sqrt{2} \times U_{eff} / (\omega \times N \times A_e)$$

where U_{eff} = effective value of the voltage and ω = angular frequency ($2 \times \pi \times f$). For square waves, the formula is as follows:

$$B = 2{,}5 \times \hat{U} / (f \times N\, A_e)$$

$\hat{U}$ is the peak voltage.

Heat generation

In applications of over 100 kHz, the problem is seldom saturation, but most often heat generation. The wire in the coil is heated by both the DC current and the AC current, while the core is heated only by the AC current. In the table below the maximum flux density values (AC current) can serve as guidance for both ferrite and iron powder:

Guide values for maximum flux density in relation to frequency for avoiding high temperature in the core.

Frequency:	100 kHz	1 MHz	7 MHz	14 MHz	21 MHz	28 MHz
Flux density:	50 mT	15 mT	6 mT	4,5 mT	4 mT	3 mT

In DC current applications with ripple current, e.g. filter chokes in a power supply, there are negligible losses if the total flux density does not exceed 200 mT for most ferrites, and 500 mT for iron powder cores.

Temperature dependence

The permeability of a ferrite or iron powder core is strongly temperature dependent. In general, it increases up to a certain temperature (the Curie temperature, T_C, ϑ_C), where it steeply drops to 1. The temperature coefficient is denoted by α_F and indicates the change per K within a specified temperature range. The inductance change (Δ_L) in relation to the temperature change can be calculated using the formula:

$$\Delta L = \alpha_F \times \mu_i \times \Delta\vartheta \times L$$

where $\Delta\vartheta$ is the temperature change in K. If the core has an air gap, ϑ is multiplied by the ratio of μ_e / μ_i.

The higher the temperature, the higher the losses. When the permeability increases, there is a higher flux density and thus higher hysteresis losses (tan δ_h). In addition, the resistivity is reduced when the temperature is increased, which means that the eddy current losses (tan δ_f) are increased.

A spontaneous increase in the permeability occurs directly after demagnetisation on account of the fact that the material is subject to a slowly decreasing AC current field or that the Curie temperature has been exceeded. It returns to its normal value in accordance to a logarithmic function. This temporary instability is known as disaccommodation. It is described by the disaccommodation factor (D_F) which is related to the initial permeability (μ_i). The inductance change in relation to the time is calculated using the formula:

$$\Delta L = -D_F \times \mu_i \times \log(t_1 / t_2) \times L$$

where t_1 and t_2 are the two times after demagnetisation between which the inductance change is calculated. Here, too, ΔL is reduced, if the core has an air gap, by a ratio of μ_e / μ_i.

Magnetic units

In the USA, for instance, other units are often used.

Quantity	SI-unit	Other unit
Magnetomotive force (mmf):	A (ampere; ”ampere-turns”)	1 G (gilbert) = 1.257 A
Flux (Φ):	Wb (weber)	1 M (maxwell) = 10^{-8} Wb
Field strength (H)	A/m (”ampere-turns”/m)	1 Oe (oersted) = 79.6 A/m
Flux density (B):	T (tesla)	1 G (gauss) = 10^{-4} T

Resistors

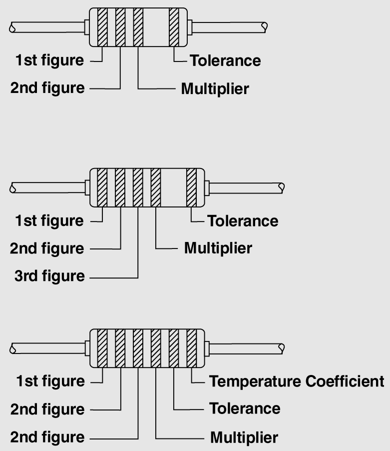
Resistors are the most commonly occurring component in electronic apparatus. They comprise a body that is normally insulated with connections. They contain a resistor element, made from a material with a known resistivity (ρ), in the form of a bar, a tube, film, surface layer or wire of a particular length (l) and area (A). This is described with the formula

$$R = \rho \times l / A$$

The unit for resistance (R) is the ohm (Ω). 1 ohm is the resistance which, at 1 Volt, allows through the charge quantity 1 Coulumb per second, i.e. 1 Ampere.

Resistors which are intended to have a resistance that is independent of current, voltage and external factors such as e.g. temperature and light, are called *linear resistors*, or simply resistors. If the resistor is to vary its resistance depending on current, voltage or some external factor, it is called a *non-linear resistor* or given a name which indicates what the resistance is dependent upon.

Resistor markings



Resistor markings

The resistance, tolerance and occasionally temperature coefficient of small resistors are often marked with 4 to 6 colour rings.

Colour codes for resistor markings.

Colour	No.	Multiplicator		Tolerance ± %	Temp coeff ± ppm/K
Black	0	10^0	1	20	200
Brown	1	10^1	10	1	100
Red	2	10^2	100	2	50
Orange	3	10^3	1000	3	15
Yellow	4	10^4	10000	0... +100	25
Green	5	10^5	100000	0.5	–
Blue	6	10^6	1000000	0.25	10
Violet	7	10^9	10000000	0.1	5
Grey	8	10^{-2}	0.01	–	1
White	9	10^{-1}	0.1	–	–
Gold	–	10^{-1}	0.1	5	–
Silver	–	10^{-2}	0.01	10	–

There are occasionally only three colour rings. In this case the tolerance, which is ±20%, is not marked. On rare occasions, other variants of colour markings occur, e.g. certain MIL-specified resistors, which have a final ring indicating failure rate. A pink final ring was previously used on extremely stable resistors.

Remember that chokes, capacitors, thermistors and fuses can be similar in appearance and can be colour marked in the same way.

Larger resistors are often marked with text. In this case, R or E (for ohm), k (for kilo-ohm) and M (for Mega-ohm) are written in front of the comma.

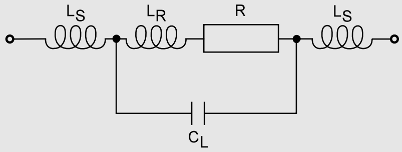
0R1 = 0,1 Ω
0E1 = 0,1 Ω
4k7 = 4,7 kΩ
22M = 22 MΩ

Three or four figure codes are occasionally used, where the first two or three figures are the significant figures and the last figure is the number of zeros.

100 = 10 Ω
101 = 100 Ω
103 = 10 kΩ
4754 = 4,75 MΩ

Frequency dependence

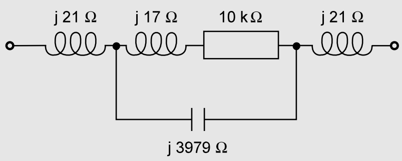
In order to understand the resistor's behaviour more easily, we can use a simplified comparison chart:



The comparison chart of the resistor.
R = resistance, C_L = internal capacitance (also called leak, parasitic and stray capacitance), L_R = inductance in the resistor element and L_S = inductance in the leads.

Here you can see that there are plenty of capacitive and inductive parts in a resistor. In AC applications (in particular HF), these give rise to reactances which, combined with the resistance, produce an impedance to which consideration must be given in some cases.

For example: what impedance will a 10 kΩ metal film resistor have at a frequency of 400 MHz? We estimate C_L at 0.1 pF. The connecting wires are 10 mm long and have a diameter of 0.6 mm. Using the formula for inductance in a straight wire (see the Coils and Chokes factsheet), we get an inductance (L_S) of 8.4 nH in each connecting wire. The inductance in the resistor element (L_R) can be calculated using the formula for a single-layered, air-cored coil. We estimate the body's diameter to be 2 mm, and the spiralisation's length to be 4 mm and 3 turns. The formula produces 6.9 nH. When converted to reactance, this gives 3979 Ω at C_L, 21 Ω at L_S and 17 Ω at L_R.



Example. Comparison chart for a 10 kΩ metal film resistor at 400 MHz frequency.

We can consider the inductive reactances to be negligible. The impedance (Z) at the parallel connection will be:

$$1/Z = \sqrt{(1/R)^2 + (1/X_{CL})^2}$$

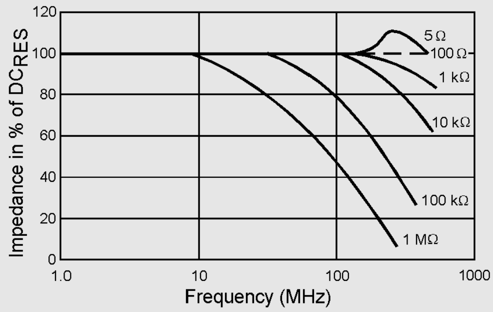
Which can also be written:

$$Z = R \times X_{CL} \times 1 / (\sqrt{R^2 + X_{CL}^2})$$

$$Z = 10\text{ k} \times 3979 \times 1 / (\sqrt{10\text{ k}^2 + 3979^2}) = 3697\ \Omega$$

The 10 kΩ resistor has an impedance of just 3.7 kΩ at 400 MHz.

Film resistors under 100 Ω can, as a rule, be considered to be inductive (the impedance increases with the frequency), 100 to 470 Ω which is almost ideal. Over 470 Ω, the resistors are capacitive (the impedance reduces with increased frequency). The higher the resistance value, the greater the capacitance. From the following diagram, impedance as a % of resistance can be read off as a function of resistance and frequency, for metal film resistors from a manufacturer.



The frequency dependence of impedance in metal film resistors

Wire wound resistors have both large inductance and capacitance, which means that they have a resonance frequency where the impedance is greatest. In the event of low frequencies they are inductive, and in the event of high frequencies they are capacitive.

Temperature dependence

A resistor which is passed by a current warms up. The amount of heat is dependent on output development (P), which is the same as the current (I) through the resistor multiplied by the voltage (U) which drives the current (P = U × I).

The output/heat development relationship is called *thermal resistance* (R_{th}). The temperature of the resistor can be calculated with the formula

$$T_{hs} = T_{amb} + P \times R_{th}$$

Ths = the "hot spot" temperature, i.e. the temperature at the warmest point on the surface. T_{amb} = ambient temperature. P = output in W and R_{th} = thermal resistance in K/W. The maximum T_{hs} is dependent on e.g. the insulation, sealing and resistance material and the thermal resistance (R_{th}) between the resistor element and the surface.

The output capacity specified in the data sheet (max continuous output) is the output at which the rise in temperature (P × R_{th}) and the specified ambient temperature (T_{amb}) combined have achieved the maximum temperature the resistor can handle without influencing parameters such as e.g. long-term stability and tolerance.

If the ambient temperature is higher than the temperature at which the output capacity is specified (as a rule 25, 40 or 70°C), the output capacity is reduced linearly to zero (derating) at the 'zero output temperature'. For epoxy lacquered resistors this temperature is approximately 150°C, for silicone-insulated and aluminium housed resistors approximately 200°C, and for glazed resistors approximately 350°C.

If you still exceed the maximum temperature (T_{hs}) for the resistor, you will shorten the resistor's expected *lifetime*. If you exceed it by a large margin, the actual lifetime can be just seconds or fractions of seconds.

There are various standards for how to test the output capacity, which the manufacturers follow. These standards differ from each other as regards e.g. installation procedure, pin length, air circulation (vertical or horizontal installation), ambient temperature, rise in temperature, surface temperature and expected lifetime. For this reason, a resistor which, according to one manufacturer, can handle 1 W, will only handle 1/10 W according to another, despite the fact that they are the same size. Experience shows that it is seldom practical to "remain at max output", in particular when the temperature of a solder connection should not exceed 100°C to prevent premature ageing.

Resistance tolerance is the maximum deviation of resistance, expressed in %. The resistance is measured in accordance with standards regarding type of measurement equipment, voltage, temperature, pin length, etc. On standard resistors, the tolerance is ±1–10%, but there are special types going down to ±0.005%.

All resistors are somewhat temperature dependent, and this is specified through a *temperature coefficient*. The unit is usually ppm/K (millionths per degree, 10^{-6}/K). The temperature coefficient varies in size in different types of resistor. Carbon resistors have a relatively large negative coefficient (–200 to –2000 ppm/K depending on resistance), while there are special metal film resistors with a coefficient below ±1 ppm/K.

Max working voltage is the maximum direct or alternating voltage which can be continuously applied to the resistor. This only applies to resistances above the 'critical' resistance, i.e. the resistance at which max voltage gives the maximum output development that the resistor can cope with. For resistances below the critical resistance, the maximum voltage is:

$$U = \sqrt{R \times P}$$

The *isolation voltage* is the voltage which the insulation around the resistor element can handle.

Noise

Noise occurs in all resistors. Both the 'thermal noise' which arises in everything which conducts current, due to the fact that all electrons do not always move in the current direction, and a current noise which depends on the resistor type. The thermal noise, which is independent of resistor type, can be calculated using the following formula:

$$U = \sqrt{(4kTRB)}$$

where U = noise voltage rms in volts, k = Boltzmann constant (1.38×10^{-23} J/K), T = absolute temperature in kelvin, R = resistance in Ω and B = bandwidth in hertz. The current noise, which is due e.g. to the type of resistor material, irregular area and impurities in the resistor material, is generally specified in the manufacturers' data sheets. The noise level is specified in µV/V or in dB. 0 dB is equivalent to 1 µV/V. The total noise is a combination of the thermal noise and the current noise.

$$\text{Total noise} = \sqrt{(\text{current noise}^2 + \text{thermal noise}^2)}$$

Voltage dependence

The resistance of all resistors is somewhat dependent on voltage, which normally lies between 10 and 1000 ppm/V. This gives rise to a distortion in the form of overtones, if you add an alternating current. This is often called non-linearity, and specifies a relationship between the signal's and the third overtone's voltages, in dB.

Design

The carbon composite resistor, or composite resistor, is an old type of resistor. It is built as a carbon rod or carbon tube with soldered connecting wires. The composition of the material in the carbon body determines the resistance value. These resistors have the advantage of having a low inductance. For this reason, they are suitable for use in pulse application such as in RC circuits for surge protection and in switching power supplies. Another advantage is that the resistors can handle temporary overloads without burning out. A major disadvantage is the large internal capacitance, approx. 0.2–1 pF depending on type and resistance value. The large internal capacitance, which is due to the structure of the carbon particles and binding agent, means that the carbon composite resistor is more or less unusable at frequencies above 5–10 MHz. It has a high temperature coefficient (–200 to –2000 ppm/K), large voltage dependence (200–500 ppm/V), high noise and poor long-term stability.

Carbon film resistors comprise a ceramic tube on which a resistive layer of carbon is vaporised. The film is spiralised up to approx. 10 turns with a diamond cutting point or laser in order to achieve the correct resistance value. The reactance of the inductance created through the spiralisation is small compared with the reactance of the internal capacitance of approx. 0.2 pF. They have a high temperature coefficient (–200 to –1000 ppm/K). The voltage dependence is below 100 ppm/V. The noise level is slightly high and the long-term stability is poor. However, carbon film resistors are extremely cheap to produce.

The metal film resistor differs from the carbon film resistor in that the carbon film is replaced by a metal film. The manufacturing process is largely the same. The resistors have good HF properties when the internal capacitance is small (below 0.2 pF). In the event of high resistance values and high frequencies, the reactance can still be important. The temperature coefficient is low (5–100 ppm/K), the voltage dependence is approx. 1 ppm/V, low noise and good long-term stability. However, the pulse capacity is low, even lower than for carbon film resistors. For this reason, care should be taken when replacing the carbon film with metal film in pulse applications.

The thick film resistor is occasionally called the metal glaze or cermet resistor. The film comprises a blend of metal oxides and glass or ceramics, screen printed on a ceramic base. They have good HF properties at low resistance values. Internal capacitance is approx. 0.1–0.3 pF. The voltage dependence is below 30 ppm/V. The long-term stability is very good. The resistors are pulse resistant, reliable and can handle high temperatures. The noise level is comparable with carbon film resistors. Surface mounted resistors are usually made from thick film.

Thin film resistors have an extremely thin film of metal, usually nickel-chrome, which is vaporised on a glass or ceramic base. The resistor is etched and laser-trimmed in order to achieve the correct resistance. The HF properties are seldom good. The temperature coefficient is extremely good, even below 1 ppm/K can be manufactured. The voltage coefficient is below 0.05 ppm/V. The long-term stability is extremely good. The noise is the lowest of all types of film resistor. Output and pulse capacity are low. The high stability means that thin film resistors are often used in precision applications, such as in extremely accurate voltage dividers.

Metal oxide resistors have a film made of metal oxide, often tin oxide, which has been spiralised. The HF properties are moderate as the internal capacitance is approx. 0.4 pF. The temperature coefficient is approx. ±200 ppm/K, the voltage dependence is below 10 ppm/V and the noise level is low. They are pulse resistant and can handle high temperatures, which makes them a good alternative to wire wound power resistors, particularly at high resistances.

Resistor networks are made of thick or thin film. Consequently they comprise a ceramic substrate with pressed-on resistors and cable paths. Two types of hole mounted casings are available, SIL casings (Single In Line) with one row with 4–14 pins and 2–24 resistors, and DIL casings (Dual In Line) with two rows with a total of 14–20 pins and 7–36 resistors. A number of different types of casing are made for surface mounting. Resistor networks are often produced specially to match particular applications. It is then possible to have different internal connections between the resistors, different resistance values for the resistors, as well as to supply the networks with other components such as capacitors and diodes.

The benefits of resistor networks include that they save room on the PCB, that the resistors' temperature drift is followed, that installation is simple and thereby saves time, which in turn results in a lower price for the installed component.

Wire wound resistors comprise a wire with high resistivity, normally made of nikrothal (CrNi), kanthal (CrAlFe) or konstantan (CuNi), wound around a base of ceramic, glass or fibre glass. They are insulated with plastic, silicone, enamel or encased in an aluminium housing. The latter to distribute the heat more easily in a cooling base. For precision purposes, they are manufactured from a high-quality and stable wire, and for power applications from a thick and hard-wearing wire. The HF properties are poor. High inductance (0.1–10 µH) and high capacitance (0.2–10 pF) depending on the number of wire turns and the dimensions of the base. In order to reduce the inductance, it is possible to wind the wire in various ways, e.g. bifilar winding, cross winding (Ayrton Perry winding) or section winding in various directions. In the precision type, the temperature coefficient is low (1–100 ppm/K). The voltage dependence is approx. 1 ppm/V. The noise is extremely low and the long-term stability is good. However, the output capacity is low. Output types have a temperature coefficient of –50 to +1000 ppm/K depending on the type of wire. Voltage dependence and noise as with the precision type. The long-term stability is greatly dependent on the surface temperature of the resistor (T_{hs}). When installing wire wound output resistors, it is important to remember that the surface temperature can be as high as 200–400°C. Temperatures this high can affect surrounding components, materials and soldering.

NTC resistors are non-linear resistors whose resistance is heavily dependent on the temperature of the resistor body. As the name implies (Negative Temperature Coefficient), they have a negative temperature coefficient, i.e. a resistance which decreases as the temperature increases. They are built up of polycrystalline semiconductors which comprise a mixture of chrome, manganese, iron, cobalt and nickel. These substances are sintered together with a plastic binding agent.

The resistance alters according to the formula

$$R = A \times e^{B/T}$$

where A and B are constants which are dependent on the material and T is the temperature. However, this is a simplified formula. Over large temperature ranges, B alters slightly with the temperature.
In order to calculate an approximate resistance (R_1) at a certain temperature (T_1), you can use the above formula if you know the resistance (R_2) at a reference temperature (T_2) and the B value.

$$R_1 = A \times e^{B/T1}$$

$$R_2 = A \times e^{B/T2}$$

If we divide these two, we get:

$$R_1/R_2 = A \times e^{B/T1} / (A \times e^{B/T2})$$

We can then eliminate A and move R_2, which gives us the 'Beta' formula:

$$R_1 = R_2 \times e^{(B/T1\ -\ B/T2)}$$

The Beta formula gives a relatively accurate value within the temperature range the B value is specified. $B_{25/85}$ specifies that the B value is fairly correct within the range 25 to 85°C.

The *power constant* (D) is the amount of power in Watts (or mW) which is required for the resistor's temperature to rise by 1 K above the ambient temperature.

The *time constant* (τ) is the time an NTC resistor needs to reach 63.2% ($1 - e^{-1}$) of the new resistance value in the event of a temperature change, without being heated by the current passing through it. This is a measure of how quickly it reacts, and is dependent on e.g. the resistor's mass.

The NTC resistor is used e.g. to: measure temperature, temperature regulation, temperature compensation, time delays, limiting starter currents and measuring flows.

PTC resistors have a positive temperature coefficient, i.e. a resistance which increases with increasing temperature. They are produced in a similar way to NTC resistors, but have a base of $BiTiO_3$ which is doped with various substances. By adding plenty of oxygen during the cooling period after sintering, a positive temperature coefficient is achieved. The resistance is slightly reduced at low temperatures, but when the material's *Curie temperature* (T_C) is reached, resistance increases strongly.

The *switch temperature* (T_{sw}) is the temperature at which the resistance is 2 × min resistance. PTC resistors are produced with switch temperatures between 25 and 160°C (up to 270°C if they are produced as heating elements).

Switch time (t_{sw}) is the time it takes for the PTC resistor to achieve the switch temperature through warming up by the current passing through, in the event of a constant voltage. At this time, the current is reduced by half. The switch time can be calculated using the following formula:

$$t_{sw} = h \times v \times (T_{sw} - T_{amb})/(I_t^2 \times R_{25} - D \times (T_{sw} - T_{amb}))$$

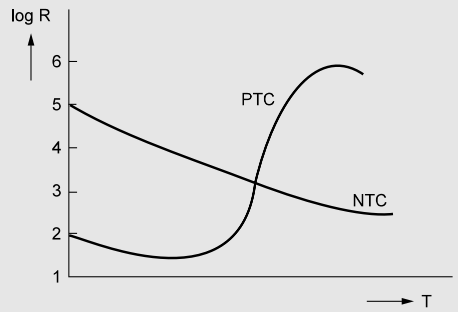
where

h = specific heat figure of the ceramic, e.g. 2.5×10^{-3} J × K^{-1} × mm^{-3},
v = volume of the ceramic in mm^3,
T_{sw} = switch temperature,
T_{amb} = ambient temperature,
I_t = current in A,
D = power constant in W/K.

By temperature coefficient we mean the PTC resistor's maximum temperature coefficient in the section where the curve is steepest.

It is important that max voltage is not exceeded. If it is exceeded, a short-circuit would probably occur and the resistor would be destroyed. Neither is it possible to connect several PTC resistors in series to achieve a higher voltage capacity. Most of the voltage will still end up over one of the resistors, which would then fail.

PTC resistors are used as overcurrent protectors for e.g. motors, self-regulating heating elements, demagnetising links in colour TVs, time delay circuits and for temperature indication.



The temperature dependence of the resistance value in PTC resp. NTC resistors.

The varistor, or VDR (Voltage Dependent Resistor), is a resistor whose resistance declines greatly with increased voltage. Today, varistors are normally made from granulated zinc oxide, doped with various substances such as Bi, Mn, Sb, etc., which are sintered to a tablet. The contact surfaces between the granules (millions), which work as a semiconductor transition with a voltage drop of approx. 3 V at 1 mA, form long chains. The total voltage drop depends on the granule size and the thickness of the varistor. Up to this voltage (varistor voltage), when the current through it is ≤1 mA, the varistor has high impedance. If the varistor voltage is exceeded, the current through the varistor increases logarithmically, i.e. the resistance falls. A varistor can switch from high impedance to low impedance in less than 20 ns. The diameter of the varistor determines the output capacity and lifetime. The structure of the granules means that the varistor has an internal capacitance of 50 – 20000 pF depending on voltage and size.

We can utilise the non-linearity to provide protection against voltage transients, which occur in the event of thunderstorms or the switching of inductive loads. A varistor can be used for both direct and alternating voltage. A ripple transient reduces the resistance in a varistor to 0.1 – 50 Ω dependent on transient voltage, varistor voltage and the diameter of the varistor.

Varistors are installed between phase and zero, and possibly at earth, in 230 Vac networks to damp incoming transients. On voltage feeds in an apparatus between + and –. Between the parts and to earth on signal cables. Above a contact which breaks a coil to prevent sparks. Above a triac to reduce radio interference.

Light dependent resistors, which are also referred to by the abbreviation LDR, are as the name implies resistors which vary their resistance depending on the amount of light (photo conductance). The stronger the light, the lower the resistance.

Light dependent resistors are primarily made from two different materials. Cadmium sulphide (CdS) which is sensitive to roughly the same spectrum of light as the human eye. Cadmium selenide (CdSe) whose sensitivity is more towards the infrared end of the scale. CdS has maximum sensitivity at 515 nm and CdSe at 730 nm, although by mixing the two materials, it is possible to achieve different curve forms with a maximum sensitivity between 515 and 730 nm.

In the dark, cadmium sulphide or selenide have no or few free electrons and the resistance is high. When energy in the form of light is added, valence electrons are released and move to the conduction band. The resistance will then be low.

The extent of the change in resistance is due, apart from the composition of the material and the type of production process, to the area and distance between the electrodes and the area which can be illuminated. Light dependent resistors have a relatively large temperature dependence, 0.1 to 2%/K. The response time varies from 1 ms to several seconds, depending on the strength of the light and the illuminated time or the time in darkness. The CdSe type is faster than the CdS type. Both possess a certain "memory effect". After long static light conditions, the resistance is temporarily displaced. The CdSe type has a more powerful memory effect than the CdS type.

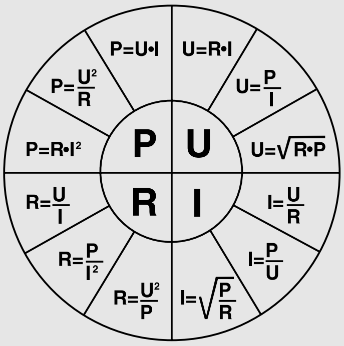
Standard series of values in a decade according to IEC-63

E192	E96	E48	E192	E96	E48	E192	E96	E48	E192	E96	E48
100	100	100	178	178	178	316	316	316	562	562	562
101			180			320			569		
102	102		182	182		324	324		576	576	
104			184			328			583		
105	105	105	187	187	187	332	332	332	590	590	590
106			189			336			597		
107	107		191	191		340	340		604	604	
109			193			344			612		
110	110	110	196	196	196	348	348	348	619	619	619
111			198			352			626		
113	113		200	200		357	357		634	634	
114			203			361			642		
115	115	115	205	205	205	365	365	365	649	649	649
117			208			370			657		
118	118		210	210		374	374		665	665	
120			213			379			673		
121	121	121	215	215	215	383	383	383	681	681	681
123			218			388			690		
124	124		221	221		392	392		698	698	
126			223			397			706		
127	127	127	226	226	226	402	402	402	715	715	715
129			229			407			723		
130	130		232	232		412	412		732	732	
132			234			417			741		
133	133	133	237	237	237	422	422	422	750	750	750
135			240			427			759		
137	137		243	243		432	432		768	768	
138			246			437			777		
140	140	140	249	249	249	442	442	442	787	787	787
142			252			448			796		
143	143		255	255		453	453		806	806	
145			258			459			816		
147	147	147	261	261	261	464	464	464	825	825	825
149			264			470			835		
150	150		267	267		475	475		845	845	
152			271			481			856		
154	154	154	274	274	274	487	487	487	866	866	866
156			277			493			876		
158	158		280	280		499	499		887	887	
160			284			505			898		
162	162	162	287	287	287	511	511	511	909	909	909
164			291			517			920		
165	165		294	294		523	523		931	931	
167			298			530			942		
169	169	169	301	301	301	536	536X	536	953	953	953
172			305			542			965		
174	174		309	309		549	549		976	976	
176			312			556			988		

E24	E12	E6	E3	E24	E12	E6	E3	E24	E12	E6	E3
10	10	10	10	22	22	22	22	47	47	47	47
11				24				51			
12	12			27	27			56	56		
13				30				62			
15	15	15		33	33	33		68	68	68	
16				36				75			
18	18			39	39			82	82		
20				43				91			

Ohm's Law

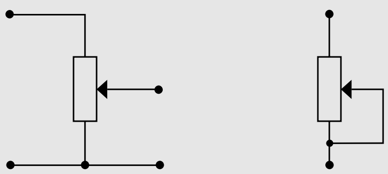
This is a simple guide for the magnitudes of voltage U (volt), current I (ampere), resistance R (ohm) and output P (watt). The formulas in the external sections are used to calculate the size of the internal sector's magnitude.



For example: When connecting an LED to 24 volts, an ignition resistor is required to restrict the current to e.g. 20 mA (0.02 ampere). We then look in sector R of the circle (resistance) and use the formula R = U/I. This produces: 24 volt/0.02 ampere = 1200 ohm. To know what output our ignition resistor must be able to handle, we take sector P (output) and use e.g. the formula P = U × I, i.e. 24 volts × 0.02 ampere, and this produces 0.48 watts. We choose e.g. a resistor of 1200 ohm and 1/2 watt.

Potentiometers

A potentiometer is a variable resistor which can be actuated mechanically. It has two connections coupled to either end of the resistor element, and a third coupled to a sliding contact (runner) which can be moved over the resistor track. The name potentiometer comes from its function of regulating potential, or as it is more commonly expressed, its function as a voltage divider. By using just the one end connection and the variable connection, the potentiometer can be used as an adjustable resistor (rheostat).



Voltage divider. *Rheostat.*

A large number of different versions of potentiometer can be produced, depending on the intended application.

Panel potentiometers are intended to be operated from a panel. They are installed in the panel with a bush (threaded neck around the spindle) or with screws. Occasionally they are installed e.g. in a bracket behind the panel and only the spindle passes through the panel. The potentiometer is designed as a rotary potentiometer with a circular resistance track and a spindle which moves the runner with a rotating movement, or as a slide potentiometer with a linear track. For more basic purposes, a resistor element made of carbon is used , which is inexpensive. However, for more demanding applications, cermet, conductive plastic or wire wound tracks are used.

The precision potentiometer is a type of panel potentiometer which is manufactured primarily in two versions. Multi-turn with wire wound track, in order to be set extremely precisely, or single turn with plastic or wire wound track, without mechanical stops at the limits. The latter has a high resolution and long lifetime in order to be used e.g. as an angle sensor. In order to measure a straight movement, there are linear position sensors, where the track is linear and the runner is operated with a rod.

Trimmer potentiometers are produced with carbon or cermet tracks, single or multi-turn and with or without a casing. As a rule, they are smaller than panel potentiometers as there is no spindle or bush, and because they have lower mechanical requirements. A trimmer often has a lifetime of just 200 rotations. This is true when the contact pressure on the runner is extremely high to produce a high level of stability. There are two types of multi-turn trimmer potentiometers, one with a linear track and geared spindle drive, where a long threaded spindle moves the runner, and one with a circular track with a runner which is turned with a worm gear unit.

Attenuator adapters comprise resistors in T- or π-links, which means that input and output impedances are constant when the degree of attenuation is altered. In variable attenuator adapters, these resistors comprise connected potentiometers which are ganged (operated with one and the same spindle). In many contexts, it is important to know exactly what the attenuation is. For this reason there are stepped attenuator adapters, where you can use switches to combine the desired attenuation.

A joystick is one or more potentiometers which are regulated with a straight protruding shaft (lever). The are used as control devices in one, two or three dimensions (X, Y and Z phases). The potentiometers in a joystick are specially produced for this application and have an angle of rotation of just 30–60°. Carbon tracks are used in the cheapest types, while it is most common to use plastic tracks in high quality types. Joysticks are often supplemented with microswitches and provided with special types of handle.

The resistance track in a potentiometer is made, as mentioned previously, from various materials in order to make use of the material's benefits for a particular application.

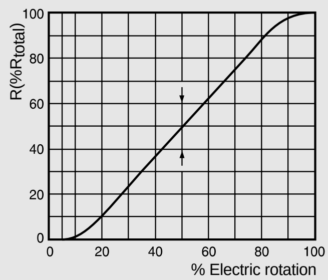
It is cheapest and simplest to make a **carbon track**. This is made from a carbon compound which is pressed onto a base of phenol card. Carbon track potentiometers can only handle low outputs. They have poor resolution and linearity, high noise and a short lifetime. However, they are extremely cheap to produce, which makes them suitable in many non-critical applications.

A variant of the carbon track is a **conductive plastic track**, where fine-grained carbon power is mixed with plastic and pressed onto a base. The benefits include infinite resolution and low noise, both when the runner is stationary (static noise) and when it is being moved (dynamic noise). The fact that it is possible to have extremely low contact pressure means that a long lifetime is possible. The disadvantages of the conductive plastic track are the low output capacity and the runner's poor current capacity, as well as the large temperature dependence of ±1000 ppm/°C. Plastic track potentiometers are used e.g. in industrial applications where stringent demands are placed on resolution and lifetime, as well as in audio equipment where the low noise is an advantage.

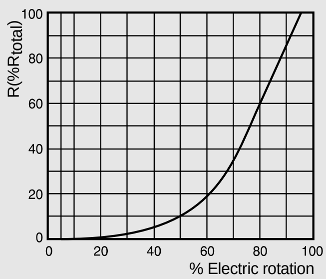
On the other hand, high output capacity is one of **the cermet track's** benefits. The cermet track comprises a mixture of metals and ceramics which are pressed onto a ceramic base. The track is stable as regards temperature, has output capacity, gives good resolution and has low static noise. As it can handle high contact pressure from the runner, the long-term stability is extremely good. For this reason, the cermet track is common in trimmer and panel potentiometers.

Wire wound tracks are used to achieve high output capacity or good temperature and long-term stability. Wire wound potentiometers are preferable when high current is being passed through the runner. In multi-turn precision potentiometers, a wire wound track is occasionally used which is overlaid with a layer of conductive plastic to increase resolution. Other areas of application for wire wound potentiometers are e.g. as adjustable series resistors (rheostats) to regulate current to different types of load.

In order to fit various applications, resistor tracks are made with various *curve forms*. A *linear* potentiometer has a resistance track with constant resistivity and area along the entire length of the track, and the resistance change is therefore the same along the entire track. In a *logarithmic* potentiometer, the resistance track is usually divided into three sections. Each section is linear but with different resistances. When the runner is at the beginning of the track, the resistivity is low and the resistance change small. At the end of the track the resistivity is high, and at this point resistivity changes much more rapidly than at the beginning of the track. Apart from the most common curve forms, linear and logarithmic, a large number of different curves are produced to suit particular purposes.



Linear potentiometer.



Logarithmic potentiometer.

Max power is the highest power which the potentiometer can handle. It is important to remember that the specified power applies to the entire track. If only part of the track is used, such as in a rheostat connection for example, the power capacity is reduced proportionally. The current through the runner must not exceed the current through the resistance track at max power. This current can be exceeded when you e.g. measure resistance between an end outlet and the runner with a normal multimeter, and move the runner towards the end outlet.

By *max working voltage* we are normally referring to isolation voltage. This is the highest voltage which may be connected to the potentiometer. Over the resistance track, max voltage is also restricted by max permitted power, which can be calculated with the formula

$$U = \sqrt{(R \times P)}$$

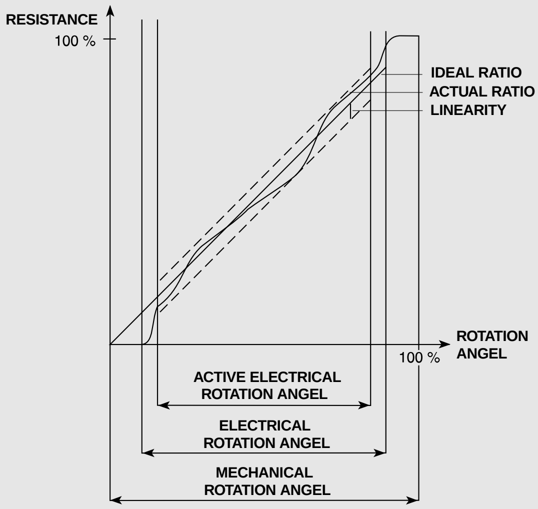
where U = voltage over the entire resistance track, R = resistance and P = power. *Test voltage* is a voltage between one of the connections and the outer casing of the potentiometer. This is often limited as regards time.

The *tolerance* of the potentiometer's resistance is seldom of importance. In a voltage divider, the decisive factor is the relationship between the resistances on either side of the runner. In a rheostat connection, the tolerance means that you can have different max resistances, but if the value is selected so that the runner ends up somewhere in the middle of the track, the tolerance is also unimportant here.

The *temperature range* can be described in two different ways: Dynamic temperature range, where the potentiometer satisfies all data when the runner is in movement. Static temperature range is with a stationary runner. They are primarily distinguished by e.g. rotating torque at low temperatures.

The *temperature coefficient* describes the change in resistance depending on temperature. This is specified in ppm/°C (millionths per degree). A voltage divider is stable as regards temperature as the resistance on either side the runner changes by the same amount.

In many contexts, e.g. during angle measurement, linear precision potentiometers are used. It is important that the actual curve form follows the theoretical one as closely as possible. This is called *linearity*, where the largest deviation in resistance is expressed as a percentage. The linearity is dependent on factors such as the purity of the raw material and variations in the area of the resistance track.



Linearity deviation of a potentiometer.

If you have a potentiometer coupled like a voltage divider, the load's resistance will be connected in parallel with part of the resistance track. This means that the linearity is negatively affected. A load resistance which is twice the potentiometer's resistance gives a linearity error of approx. 11%. To be counted as negligible, the load's resistance should be at least 100 times greater than that of the potentiometer.

The *electric angle of rotation* is the angle under which a change in resistance takes place. Active electrical angle of rotation is approximately 20° smaller. Approximately 10° at the beginning and end of the track can be influenced by the connections' fastening. Linearity is often measured only within the angle. *Mechanical angle of rotation* is approx. 30° greater than the electrical angle in order to achieve good contact with the end connection.

When the runner is at one limit, zero resistance is not achieved, but rather *limit resistance* or min resistance. This is due in part to the transition resistance between the runner and the track, resistance in the connections and mechanical tolerances which can mean that the runner does not reach the end of the track. The limit resistance is expressed as a percentage but with a minimum in ohms (e.g. "1% or 2 Ω") where the highest value applies.

The *contact resistance* which exists between the runner and the track, in particular when the runner is in motion, is largely dependent on current. Extremely low currents have difficulty bridging the diode effect brought about by a thin layer of oxide. The contact resistance varies considerably when the runner is in motion. This is called *CRV* (Contact Resistance Variation) and can be viewed as a noise. By the term *ENR* (Equivalent Noise Resistance), consideration is also being given to the resistance variations which exist in the track. A wire wound potentiometer has a high ENR as the resistance leaps every time the runner moves from one wire turn to another. CRV is expressed as a percentage of the total resistance and ENR in ohms.

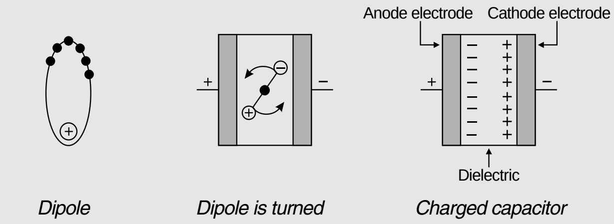
Capacitors

A capacitor consists of two plates (electrodes) separated by an insulating material (dielectric). The electrodes can then be charged without electrons jumping from the negative to the positive electrode. The **capacitance**, C, of the capacitor is its capacity to be charged with the charge Q, in coulombs per volt of the applied voltage U. This is expressed in the formula:

$$C = Q / U$$

The unit is coulomb/volt which is now called farad (F).

Capacitance increases with increased electrode area and reduced distance between the electrodes. To reduce the distance between the electrodes, air is not normally used as a dielectric but a material which can be made very thin, e.g. plastic, ceramics or an oxide layer. These materials often contain dipoles which contribute to an even higher capacitance. In a dipole, the electrons of the atoms can form extended oval paths around the nucleus of the atom creating a negative centre of gravity beside the positive nucleus. The dipoles can turn and adopt the same direction as the electric field when they are attracted by the charged electrodes. This causes the effect of the distance between the electrodes to be reduced and the capacitance increased. This effect on the capacitance is described by the property **permittivity**.



Dipole *Dipole is turned* *Charged capacitor*

The following formula applies:

$$C = \varepsilon \times A/d$$

where C = the capacitance in farad, A = the area in m^2, d = the distance between the electrodes in m and ε = the permittivity, which actually is $\varepsilon_o \times \varepsilon_r$ where ε_o is the permittivity in a vacuum, ε_r is a relative number which describes the permittivity of the dielectric in relation to the permittivity in a vacuum. ε_r is often called dielectric constant or capacitivity number.

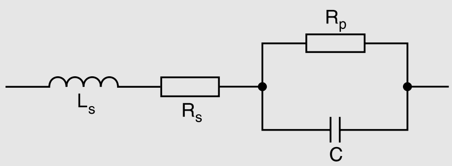
$$\varepsilon_o = 8{,}85 \times 10^{-12}\ \text{F/m}$$

As is shown in the table, the selection of dielectric determines the capacitance and size of the capacitor to a large extent. However, the materials have other properties and shortcomings which mean that the material with the highest constant cannot always be used.

ε_r for certain materials:

Air	1	Mica	4–8
Water	80	Aluminium oxide Al_2O_3	7
Glass	10	Tantalum oxide Ta_2O_5	11
Impregnated paper	3.5–6	Ceramics class 1	5–450
Pertinax	3.5–4.5	Ceramics class 2	200–15000
Polyester	3.3	Ceramics class 3	10000–50000
Polycarbonate	2.8	Ceramics NP0	60
Polypropylene	2.2	Ceramics X7R	1500
Polystyrene	2.6	Ceramics Z5U	5000

To understand the capacitor as a component we can use the following simplified equivalent diagram:



The capacitor's equivalent diagram.
R_s = the series resistance of the supply leads, electrodes and any electrolyte and the losses of the dielectric, L_s = the inductance of the supply leads and electrodes, C = the capacitance, R_p = the insulation resistance of the dielectric.

By *ESR* (equivalent series resistance) is meant all power losses of the capacitor which, apart from the series resistance (R_s) of the supply leads and electrodes, includes the dielectric losses (R_p) which occur when the dielectric is exposed to a change in field strength of the electric field. ESR is variable with frequency and temperature.

The power losses cause a rise in temperature which must be controlled if significant. To specify the dissipation resistances, a *dissipation factor* (tanδ) is indicated. It is described with the formula:

$$\tan\delta = ESR/X_c$$

Therefore, the dissipation factor is the quotient between ESR and reactance X_C. The power lost in the capacitor is calculated with the formula:

$$P = U^2 \times \omega \times C \times \tan\delta$$

If the frequency applied is the same as that at which ESR is specified, the formula can also be written as:

$$P = U^2 \times ESR / X_C^2$$

This formula is valid only if ESR is much smaller than the absolute value of $X_C - X_L$. (X_L see below.)

ESL (equivalent series inductance) is the inductance of the supply leads and electrodes L_s. The inductance of modern capacitors is usually between 10 and 100 nH.

The *impedance* of a capacitor is obtained with the formula:

$$Z = \sqrt{(ESR^2 + (X_C - X_L)^2)}$$

where Z = the impedance in Ω, X_C and X_L are the capacitive and inductive reactances at the particular frequency.

A capacitor also has a *self-resonance frequency* where X_C and X_L have the same value and outweigh each other. At this frequency, the impedance is equal to ESR.

The resistance of the insulating dielectric (R_p) is never infinite but always slightly conductive. This gives rise to a current through the dielectric called *leakage current* and causes the capacitor to have a certain self-discharge. This can be a critical factor in e.g. timing circuits.

Many properties of a capacitor are *variable with temperature* such as the dielectric constant, ESR and leakage current. The correct type of dielectric must therefore be selected according to the temperature range in which the capacitor will operate.

To describe the change in capacitance in relation to temperature, a *temperature coefficient* is specified. It is often given in ppm/°C (millionth parts per degree).

Moreover, many parameters are more or less dependent on frequency and voltage which must be taken into account in the selection of a dielectric.

Pulse rating is a way of describing with which speed a capacitor can be charged and discharged. The change in voltage gives rise to a current in electrodes and supply leads, the resistance of which produces a power loss. If the current density of the electrodes becomes great, the resistivity and hence the dissipation will increase. At a very high current, the electrodes can start to vaporise and overpressure can build up in the capacitor with disastrous consequences. Further, the change in voltage can cause dissipation in the dielectric which, in combination with the resistive losses, increase the temperature of the capacitor. The pulse rating is specified at an operating voltage equal to the rated voltage. If the operating voltage is lower, the pulse rating can be multiplied by the quotient between rated and operating voltage.

The pulse rating given on data sheets can be specified under very different conditions. The number of pulses, frequency, temperature rise etc. vary between different standardised testing methods.

The current created by a change of voltage can be calculated with the following formula:

$$I = C \times (\Delta V/\Delta t)$$

If the capacitance C and pulse rating ΔV/Δt are given in µF and V/µs respectively, current I in A is obtained.

The *max operating voltage* is dependent on several factors such as the voltage capacity and thickness of the dielectric, the distance between the connection wires and the case. The voltage capacity also varies with temperature and frequency. It is therefore important not to exceed the maximum voltage for the actual operating conditions. Even if a direct breakdown of the dielectric does not occur, too high a field strength can cause long term changes in the dielectric.

When a capacitor has been charged and the dipoles of the dielectric have been created and turned in line with the voltage field, not all dipoles return to the original position when the capacitor is discharged. These remaining dipoles cause a certain voltage to reappear in the discharged capacitor. This is called *dielectric absorption* and is more or less present in all capacitors. In certain applications such as sample and hold circuits and in audio equipment it is desirable that it is reduced as far as possible. The dielectric absorption is measured in percent of the original voltage after short circuiting for a certain time. However, several standards on measurement are available.

Application areas

As *coupling capacitors* blocking a d.c. voltage but conducting an a.c. voltage. As *decoupling capacitors* short circuiting an a.c. voltage superimposed on a d.c. voltage.

In *tuned filters* and *resonance circuits,* where the capacitor determines the frequency, often in combination with a resistor or a coil as e.g. in an oscillator or a distribution filter for a loudspeaker.

Power supply units contain capacitors to *store energy* which is used to filter (smooth) a d.c. voltage.

In *timing circuits* the charging and discharging curve of a capacitor is used to determine time as for example in an unstable multivibrator.

For *EMI suppression* the capacitor is used to absorb voltage transients, as in an RC circuit connected across e.g. a relay coil. Capacitors, e.g. X and Y capacitors, are also used to suppress high frequency interference (RFI).

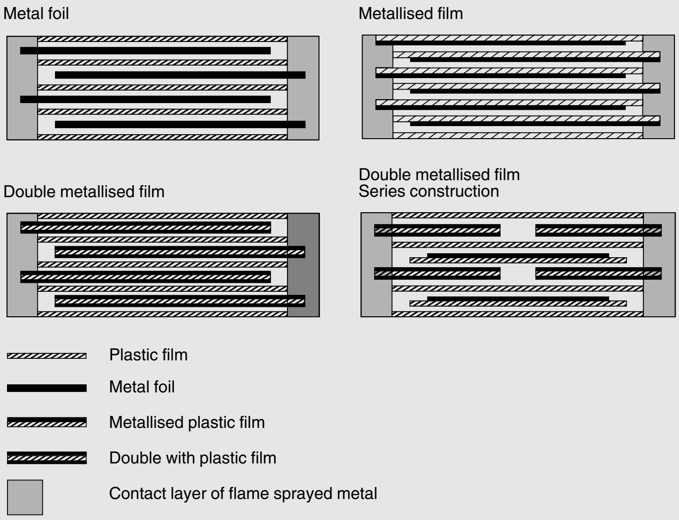
For high voltage a.c. current a *capacitive voltage divider* is often used to e.g. measure voltage. It has smaller power losses than a resistive voltage divider.

Capacitor types

Plastic film capacitors use a plastic film as dielectric. Dissipation from these capacitors is low due to the low resistance of the electrodes and their insulation resistance is high. Their manufacture has been automated and the price can therefore be kept low. They are non-polarised (either electrode can be positive or negative) and they have a very low leakage current.

Plastic film capacitors are used e.g. as coupling and decoupling capacitors in analogue and digital circuits, in timing circuits and in tuned filters. They are manufactured with capacitances from 10 pF to 100 µF

The electrodes consist of a metal foil or a metallisation. The latter is a thin vaporised metal layer which has the advantage, in case of a flash-over, that the metal coating will be vapourised around the point of the flash-over and any short-circuit is hence avoided. Many different types of winding can be used. A few of the most common are shown below:



Different constructions of plastic film capacitors.

As described in the diagram the windings may be arranged so that the capacitor is made up internally by two capacitors connected in deries. This increases the pulse durability of the capacitor.

In early constructions of plastic film capacitors, the supply lead was connected to one end of the winding. On a modern plastic film capacitor, the side of the rolled up foil is coated with a metal contact layer in a process called flame spraying. In this way, the complete long edge of the foil or film strip can be connected to the connection wire and the resistance and inductance of the capacitor can be reduced considerably.

Many different types of plastic are used in capacitors:

Polyester (PET, polyethylene terephthalate, mylar) can easily be made thin (approx. 1 µm is possible) and is easy to metallise. Small dimensions and low prices are therefore feasible. Polyester offers the worst performance of the modern plastic materials. Polyester capacitors with metal foil electrodes often have the designation KT and MKT if metallised. These capacitors are used in less critical applications such as decoupling.

Polycarbonate (PC) can also be made thin and is relatively easy to metallise. The material has a lower dielectric constant than polyester and the capacitors are therefore somewhat larger and more expensive. However, they do have a considerably lower dissipation and greater stability. These capacitors are designated KC and MKC if metallised. Polycarbonate capacitors are used in critical applications where their great stability is useful as e.g. in tuned filters and in oscillators.

Polypropylene (PP) is diffcult to make thin. It also requires pre-treatment before being metallised. Polypropylene capacitors are therefore both large and expensive compared to those of polyester or polycarbonate. Among its advantages are low dissipation, high stability and low dielectric absorption. Polypropylene capacitors with foil electrodes are called KP and MKP if metallised. Polypropylene capacitors are often used in pulse applications and where a low dielectric absorption is required e. g. in sample and hold circuits and in audio equipment.

Polystyrene (styrol, styroflex) is an old plastic material which is being increasingly replaced by polycarbonate and polypropylene. It can be metallised only with difficulty and its low dielectric strength (voltage capacity) makes the film considerably thicker than the other plastic materials. On the other hand, it has a very low dissipation, high stability and low dielectric absorption. Polystyrene is primarily used in critical filter applications.

Polyphenylene sulfide (PPS) is a new material, the most important properties of which are a high temperature resistance, good stability and very low power losses. However, the voltage capacity of the material is low and film thickness must therefore be increased.

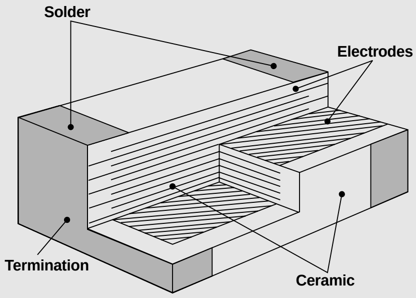
Comparison table for plastic materials (typical values).

	Polyester	Polycarbonate	Polypropylene	Polystyrene
Dielectric constant:	3.3	2.8	2.2	2.5
Tanδ at 1 kHz:	5×10^{-3}	1×10^{-3}	2×10^{-4}	2×10^{-4}
Tanδ at 100 kHz:	18×10^{-3}	10×10^{-3}	3×10^{-4}	3×10^{-4}
Max op temp °C:	125	125	100	70
Dielectr absorption %:	0.2-0.25	0.12-0.2	0.05-0.1	0.02-0.05
Temperature coeff ppm/°C:	+400	+150	-200	-150
Dielectr strength V/mm:	250	180	350	150

Paper capacitors have now been replaced by plastic film capacitors in most applications. In spite of a high dielectric constant, paper capacitors are both bigger and more expensive than plastic ones. The benefits of paper capacitors are their pulse rating and their low carbon content (approx. 3 % compared to 40–70 % for plastic) which provide excellent self-healing properties and negligible fire risks. Today they are used almost exclusively as EMI suppression capacitors (X and Y capacitors), where the advantages of paper as against plastic can be utilised.

Sometimes both plastic and paper foil is used in the capacitor winding. The dielectric is called mixed when the benefits of different materials are made use of.

Ceramic capacitors are made with one or more ceramic plates having imprinted metal electrodes. A ceramic capacitor with a dielectric with only one layer is called a single layer, single plate or plate capacitor. If a capacitor is constructed with a dielectric of several layers with intermediate electrodes, it is called a multilayer or monolithic capacitor. The range of available materials and designs is enormous. Ceramic capacitors are manufactured with capacitances from 0.5 pF up to several 100 µF. However, capacitors above 10 µF are rare for reasons of cost.



Section of a multilayer capacitor

The ceramic materials are divided into three groups:

Class 1 are materials with a low dielectric constant. They are the most stable not only with regard to temperature but also frequency, voltage and time. They have very low power losses even at high frequencies. Capacitors with one layer are manufactured with capacitances from 0.47 to 560 pF. Multilayer capacitors are made with an NP0 dielectric with values from 10 pF to 0.1 µF. They are used for example in HF applications and in temperature critical, frequency determining applications such as oscillators.

Class 1 dielectrics have an almost linear temperature coefficient. They are designated by a P or an N indicating whether the coefficient is positive or negative, as well as a figure giving the temperature coefficient.

Denomination and marking of class 1 capacitors.

Dielectric	Temp coeff ppm/°C	Colour code	EIA-designation
P100	+100 ±30	red/violet	M7G
NP0	0 ±30	black	C0G
N075	–75 ±30	red	U1G
N150	–150 ±30	orange	P2G
N220	–220 ±30	yellow	R2G
N330	–330 ±60	green	S2H
N470	–470 ±60	blue	T2H
N750	–750 ±120	violet	U2J
N1500	–1500 ±250	orange/orange	P3K

Class 2 consists of materials with a high dielectric constant. They have a non-linear dependence on temperature, frequency and voltage. The group covers many different types of dielectric with varying properties. Class 2 dielectrics have low losses at moderate frequencies. They age 1–5 % per log decade hour. The original values can be regained by heating above the Curie temperature of the ceramic which is approx. 150 °C.

They are manufactured as single layer capacitors with capacitances from 100 pF to 0.1 µF and as multilayer capacitors from 10 pF to 10 µF. They are used in uncritical applications such as coupling and decoupling.

Class 2 dielectrics are designated by a K and a figure indicating the dielectric constant or, according to EIA, by three characters where the first two indicate a temperature range and the third shows the change in capacitance within this temperature range.

EIA denomination of class 2 capacitors.

1:st character	Lower limit of temp. range	2:nd character	Upper limit of temp. range	3:rd character	Capacitance change
Z	+10 °C	2	+45 °C	A	±1,0 %
Y	–30 °C	4	+65 °C	B	±1,5 %
X	–55 °C	5	+85 °C	C	±2,2 %
		6	+105 °C	D	±3,3 %
		7	+125 °C	E	±4,7 %
		8	+150 °C	F	±7,5 %
		9	+200 °C	P	±10 %
				R	±15 %
				S	±22 %
				T	+22, –33 %
				U	+22, –56 %
				V	+22, –82 %

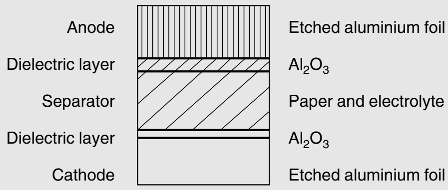
Z5U therefore means that the capacitance can vary between +22 and –56 % within the temperature range from +10 to +85 °C.

Class 3 is based on a semiconducting material which often has a sintered granulated internal structure, where the individual grains, having a low capacitance, between them create a high aggregate capacitance. The material offers largely the same or a somewhat worse performance than capacitors of class 2, but its voltage capacity is low. 16–50 V is often the maximum operating voltage. Due to the extremely high dielectric constant, great capacitances can be obtained with small dimensions and at a low price. The capacitors in the group are manufactured with capacitances from 1000 pF up to 1 F.

Mica capacitors are constructed like the ceramic multilayer capacitors, but as the capacitor element need not be baked at high temperature, silver can be used in the electrodes. Mica is actually the mineral muscovite which is extracted in Indian mines where the quality is particularly high. Muscovite is a tough and resistant mineral which splits into thin plates which can be provided with electrodes and stacked. Its electrical properties such as insulation resistance, power losses and stability are excellent och comparable to those of the best plastics and ceramic materials. However, mica capacitors are relatively large and expensive which is why they have been replaced by e.g. polypropylene capacitors to a large extent. Capacitors with mica dielectrics are often used in HF applications , where their low losses as well as high frequency and temperature stability are utilised. They are manufactured with capacitance values from 1 pF to 0.1 µF.

Electrolytic capacitors have aluminium or tantalum electrodes, where the surface of the anode electrode (positive terminal) has been oxidised and this very thin oxide layer is used as dielectric. To reduce the distance between the dielectric oxide layer and the cathode electrode (negative pole) an electrolyte with low resistivity is used.

Wet aluminium electrolytic capacitors have an electrolyte consisting of e.g. boric acid, glycol, salt and solvents. The electrodes are etched in acid baths in order to make their surfaces porous. In this way, the area of the electrodes can be increased up to 300 times. The dielectric oxide layer of the anode is formed (built up), in a bath of hydrated electrolyte, to a thickness of approx. 13 Å per volt, of the voltage it is designed for. Even the cathode is provided with a thin oxide layer (approx. 40 Å). To prevent contact between the electrode oxide layers which can then be damaged, a thin paper is placed as separator between the electrodes. As the electrolyte is negative, the capacitor case is connected to the negative terminal. However, the case cannot be used as connector.



Construction of a wet electrolytic capacitor.

The oxide layer acts as a diode and conducts the current in the reverse direction. The maximum reverse voltage is 1.5 V. If this voltage is exceeded the consequences can be disastrous.

The ESR of a wet aluminium electrolyte is relatively high due to the high resistivity of the electrolyte compared to e.g. aluminium or copper.

The *temperature dependence* is very great particularly at low temperatures. At the lower end of the temperature range, ESR can be 20 times higher than at room temperature. The change in capacitance due to temperature is ±20 % within the temperature range.

The *leakage current* of the dielectric is specified at the rated voltage. At a lower voltage, the leakage current is reduced. At half the voltage, the leakage current is only 20 % of that specified. The leakage current is increased with rise in temperature. At the upper end of the temperature range, the leakage current is 10 times as high.

Service life is a vague term. The service life of an electrolytic capacitor is its operating life until some parameter has reached a value outside the specified limit values. Many different standards are available for measurement of service life and comparisons are therefore difficult. The parameters which are measured are e.g. capacitance, dissipation factor and leakage current. It is primarily the electrolyte which is affected by ageing and changing in different ways. The electrolyte is broken down through chemical reactions and even the oxide layer can be destroyed. Modern electrolytic capacitors make use of very volatile solvents which evaporate in spite of efficient sealing and the capacitor dries out. A high temperature in the capacitor accelerates ageing markedly. A temperature reduction of 10 °C doubles its service life.

Wet aluminium electrolytic capacitors are manufactured with capacitances from 0.1µF to 0.5 F. Higher voltage capacities than approx. 500 V are not produced. The most common applications for aluminium electrolytic capacitors is as filter capacitors (reservoir capacitors) in power supplies. For a.c. voltage purposes, special aluminium electrolytic capacitors, so called **non-polarised aluminium electrolytic capacitors**, are manufactured in which each supply lead is connected to an anode electrode with oxide layer. Between the anodes, a cathode foil without supply lead is fitted.

The production of **dry aluminium electrolytic capacitors** started at the very beginning of the 20th century. However, those capacitors have little in common with the dry aluminium electrolytic capacitors of today. To separate the two groups, the modern types with manganese dioxide or organic semiconductors as electrolyte are often called **solid aluminium electrolytic capacitors** (SAL).

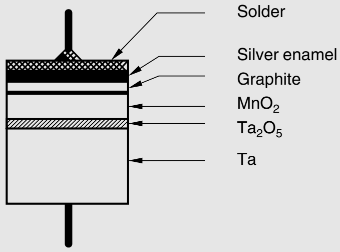
The manganese dioxide type has an electrolyte of manganese dioxide, which has a low resistance. The aluminium electrodes are etched and dipped in forming baths where an oxide layer is formed. A woven fibre glass layer, which also acts as a separator between the electrodes, is coated with manganese dioxide and is placed between the electrodes which are wound or folded into a compact capacitor element. The capacitor is then provided with a suitable case and supply leads.

This capacitor offers many advantages compared to other electrolytic capacitors such as a long service life, as the electrolyte cannot evaporate, a wide temperature range, from –55 to +175 °C and from –80 to +200 °C for certain types, a high temperature capacity, ability to continuously handle 30 % of rated voltage in the reverse direction and the fact that overheating does not lead to short circuit. Its service life is not so dependent on temperature as that of other electrolytes but it does vary with voltage. These capacitors are manufactured with capacitances from 0.1 to 2200 µF.

The second type has an organic semiconductor as electrolyte. It consists of a complex salt, called TCNQ, which has excellent electrical and thermal properties. This type too has etched electrodes with a separator between them. Moreover, it has an ESR comparable to that of ceramic and plastic capacitors and to obtain an equally low ESR with a wet aluminium electrolyte, its capacitance value would have to be increaased approx. 50 times.

This capacitor is suitable as e.g. filter capacitor in switching power supplies, where the high frequency makes the ESR value more important than the capacitance. It cannot handle such high temperatures as the manganese dioxide capacitor and 105 °C is the maximum permissible temperature. At low temperatures (down to –55 °C) it has, like the manganese dioxide type, a very small capacitance and ESR deviation. It can handle approx. 10 % of the rated voltage as reverse voltage. Its service life is more temperature dependent than that of ordinary wet electrolytic capacitors. 2000 h at 105 °C is increased to 20,000 h at 85 °C. An overvoltage can cause short circuiting, but if the current is less than 1 A, the temperature will be lower than 200 °C when the electrolyte is broken down and the capacitor is not damaged permanently. These capacitors are manufactured with values from 0.1 to 220 µF.

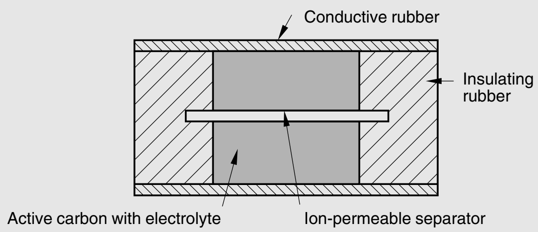
Tantalum electrolytic capacitors are manufactured with tantalum oxide as the dielectric. This material has excellent electrical properties. The anode of the capacitor is made from tantalum powder which is pressed and sintered into a porous cylinder or cube around a piece of tantalum wire. Approx. 50 % of its volume consists of air, which means that its internal surface is 100 times larger than the external surface. After it has been provided with an oxide layer in an acid bath, the capacitor element is dipped in a manganese nitrate solution which fills all its pores. The nitrate is converted with heat to manganese dioxide which becomes the dry electrolyte. To obtain contact with the cathode electrode, consisting of a conductive argent, the capacitor element is coated with a layer of carbon graphite. The older type of tantalum capacitor with a wet electrolyte and silver enslosure has been replaced by the dry type for several reasons including costs.



Construction of a tantalum electrolytic capacitor.

A tantalum capacitor has a low ESR due to the tantalum and the low resistivity of the manganese oxide. It is also considerably smaller than an aluminium electrolytic capacitor of equivalent value. Tantalum capacitors are used in applications such as coupling, decoupling, energy storage and in timing circuits where the low leakage current is of benefit. The greatest disadvantage with tantalum capacitors is that they tend to short circuit if the voltage or temperature is too high. This can cause fire in the capacitor. During the early development of the tantalum capacitor, a series resistance of 3 Ω per volt was recommended in order to limit the charging and discharging current, resulting in power loss and heating. For modern capacitors, a circuit impedance of 0.1 Ω per volt is recommended, which means that no series resistance at all is required most of the time, as the resistance of copper paths and conductors provide sufficient safety. Maximum reverse voltage is 15 % of the rated voltage at 25 °C, but is reduced at higher temperatures. At 85 °C it can only handle 5 % in the reverse direction. Tantalum electrolytic capacitors have a high temperature stability. They are manufactured with capacitances from 0.1 to 1000 µF.

The **double layer capacitor** (backup capacitors, supercap, goldcap etc.) is something in between a capacitor and a battery. In contrast to all other types of capacitor, this type has no dielectric and it operates according to the theory, published by Helmholtz in 1879, on electrical double layers, which is built on the property of electrical charges to attract each other and to form a positive and a negative layer on different sides of the contact surface between two media. The capacitor is constructed with several cells connected in series and consisting of two layers of active carbon particles impregnated with an electrolyte. Between the carbon layers there is an ion-permeable separator. The two layers and the separator are enclosed in vulcanised rubber. When a voltage is applied to the capacitor and the carbon particles of the anode layer are positively charged and the carbon particles of the cathode are negatively charged, the negative ions of the electrolyte pass through the separator and are collected around the positive carbon particles. In the same way, the positive ions are collected in the cathode layer. In this way, large charges can be stored in this type of capacitor. 1 gram of carbon powder can in theory provide a capacitance of 200 to 400 farad.



Construction of a double layer capacitor.

As the electrolyte of the cells contains water, the maximum voltage capacity will be 1.2 V per cell. At that voltage, water divides into oxygen and hydrogen. The capacitors have a high ESR of 1 to 300 Ω which strongly limits the discharge current. They can be fully charged in approx. 1 minute and have a service life of more than 10,000 charging/discharging cycles or 10 years with continuous charging. The leakage current (self-discharge) is approx. 1µA, which means that approx. 50 % of the voltage is left after a month. Temperature dependence is high. Within the temperature range of –25 to + 70 °C, the capacitance can change from –50 to +150 %. ESR is 3 times higher at –25 °C than at room temperature. They are non-polarised but the connection to the case should preferably be negative.

This capacitor is manufactured with capacitances from 10 mF up to 22 F and recent developments tend towards greater capacitances. It is used almost exclusively as voltage stand-by for e.g. memories and microprocessors. However, it is also used to store energy for short term requirements such as extra energy to start a motor, make a relay switch or generate an ignition pulse.

Capacitance table

Capacitance Conversion Table				
0.000001 µF	=	0.001 nF	=	1 pF
0.00001 µF	=	0.01 nF	=	10 pF
0.0001 µF	=	0.1 nF	=	100 pF
0.001 µF	=	1 nF	=	1000 pF
0.01 µF	=	10 nF	=	10000 pF
0.1 µF	=	100 nF	=	100000 pF
1 µF	=	1000 nF	=	1000000 pF
10 µF	=	10000 nF	=	10000000 pF
100 µF	=	100000 nF	=	100000000 pF

Diodes, transistors and thyristors

General information about Semiconductors

In the salad days of semiconductors, germanium (Ge) was the basic element. With a melting point 420 degrees lower than silicon, which does not melt until 1410 C, germanium was easier to manage. The first transistors were manufactured from germanium back in 1947. During the seven years following that event, about a hundred different transistors are introduced in the United States. In 1954, Gordon Teal of Texas Instruments demonstrates the very first samples of silicon transistors and already in 1955 the first commercial types are available on the market. Diodes, transistors and integrated circuits are made up of semiconductors.

The silicon transistor has better properties than germanium transistors, like higher breakdown voltage and higher power resistance. By using silicon, it also became possible to overcome many of the earlier temperature-related problems. When germanium is heated, the number of free electrodes will increase and, in turn, increase the current that flows through the transistor, also leading to an increase in temperature. A positive reconnection, resulting in a rush of current, will burn down the transistor in the end unless current is limited.

Semiconductors constitute the largest group of active components, comprising everything from simple diodes to highly advanced integrated circuits. The basis of these components' construction is formed by the P-N junction. A semiconductive material can be doped with various interfering materials, which will result in an excess, n-type, or a deficit, p-type of electrons. Typical doping materials are phosphate and boron. A term that lacks foundation in the purely physical sense but is still used to describe this deficit of electrons is the presence of ''holes''.

In the junction between a P-doped material and a N-doped material, a region is formed which will conduct current in one direction. In the diode, the simplest of the semiconductor components, this rectification is used to great advantage.

Diodes

The most important function for a diode is to act as a one-way vent for electrons. When the diode has forward-biased voltage, it will conduct current while it will block current when reverse-biased voltage is applied. A simple yet highly useful quality.

Diodes have been manufactured from both selenium and germanium, but the silicon diode reigns the market supremely these days. There are still special contexts, however, where germanium diodes can be of interest due to their low forward voltage drop, which is only 0.3 V, compared to 0.7 V for silicon diodes. Otherwise, the silicon diode has proven to be a reliable component that suits almost any application, ranging from rectifier diodes in mains parts to applications with radio frequency, voltage references and solar panels. Diodes are frequently used as serial switches to control signal paths in audio applications and as shunted components for connecting and disconnecting oscillators in RF applications.

Silicon diodes have a forward voltage drop of 0.7 V for small signal diodes, while power diodes can have a forward voltage drop of 1 V or more. If the reverse-biased voltage exceeds the specified value, the diode will be destroyed.

Avalanche diodes are special diodes unharmed in the event of an exceeded reverse-biased voltage. Surge voltage is absorbed by the diodes, making them suitable as transient and surge voltage suppressors.

Fast recovery is a fast type of diode, which means that it is ideal for switching. They are therefore often referred to as switch diodes and their recovery (switch-over) time lies somewhere between 1 and 500 ns.

Low-leakage diodes are another variety with a very low leakage current in the reverse direction.

Zener diodes work like regular diodes in the forward direction, but have a very specific reverse voltage. This means that the diode is used in the reverse direction to make use of the so-called Zener voltage, which is the voltage where breakdown occurs. A resistor or a current generator must be serially connected with a Zener diode to limit the current. A good Zener diode has a well-defined Zener voltage. The curve should demonstrate a sharp knee shape. The temperature drift should also be minimal. The best Zener diodes are in the 5.6–6.2 V range. Lower voltage leads to an increasingly negative temperature coefficient, while a higher voltage leads to a positive coefficient. For this reason, it can be a good idea to connect Zener diodes in series. A regular silicon diode is sometimes placed in series with a high-voltage Zener diode so that the temperature coefficients will counteract each other. The combined resistances of the diodes, however, will flatten the Zener peak somewhat.

There are also diodes in the range below 2 V. They are called **stabistors**. They work in the forward direction, separating them from Zener diodes.

Transient suppressor diodes are actually Zener or avalanche diodes that cut off voltage peaks. They are used to protect electronics components and systems. The cut-off is distinct and very rapid. This type of diode can temporarily tolerate high currents that occur at the cut-off.

Capacitance diodes, or varactor diodes, act as a voltage-controlled capacitance. Regular reverse-biased diodes also do, since their capacitance increases as voltage decreases, but capacitance diodes are optimised for this very task. The difference is the doping profile in the P-N junction. In general terms, we differentiate between gradual, abrupt and hyperabrupt junctions. The practical differences are in the change of capacitance as a function of the voltage change, where hyperabrupt diodes demonstrate the steepest sequences. Capacitance diodes replace rotating plate condensers in tuned circuits. They can also be used in stages for frequency duplication, for switching in narrow-band systems and in parametric amplifiers.

A **Diac** is a triac with no gate terminals. When the specified voltage is exceeded, a breakdown occurs and the diac becomes conductive until the holding current becomes too low. It is conductive in both directions and is used to control triacs.

Constant current diodes are really field effect transistors where the source and drain are connected.

Tunnel diodes have a breakdown already at a very low forward voltage, approx. 0.1 V. When high current flows through the diode, its forward voltage drop increases to a point where the current instead decreases as voltage increases, i.e. a negative resistance. If voltage is increased with about 0.3 V, the curve turns and shows positive resistance characteristics. Because of its negative resistance, the tunnel diode can be used as an active element in an oscillator. The negative resistance compensates the circuit's loss resistances and self-oscillation occurs.

PIN diodes are usually used as switches in high-frequency situations. They have low resistance in the forward direction and low capacitance when voltage is reversed. This provides low attenuation in the on position and high attenuation in blocked position. The defining characteristic of this diode is inertia during switch-over. This means that the diode does not continuously change its properties in proportion to the radio signal, which in turn means that it does not cause distortion. The diode works primarily as a resistor for high frequencies. The inertia, the recovery time from reverse voltage, τ, depends on the life span of the minority carriers. PIN diodes for the microwave range can have τ equal to a few ns, while there are PIN diodes that are useable all the way down to a few MHz with a τ of several ms. The lower limit frequency = $1/2\ \pi\ \tau$. Below this level, the diode works like a regular P-N junction.

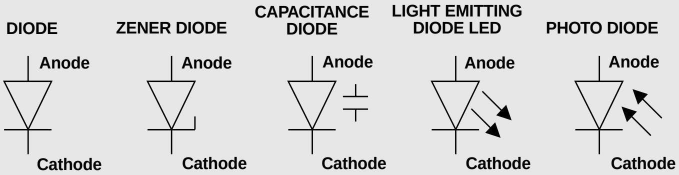
The PIN diode's resistance in the forward direction can be varied between 1 and 10,000 Ω by varying the current through it. This can be used in current-controlled attenuation kits. PIN diodes have an intrinsic (I= Intrinsic) layer of resistive material, which is placed between regions of highly doped P and N material.

Step recovery is a diode type that, like PIN diodes, has three layers. However, it differs in that the resistance change occurs abruptly at a small change of the charge between P and N. The abrupt voltage change can cause a brief transient, which is tantamount to several overtones to the input frequency. One common application area is frequency multipliers for high frequencies.

Gunn diodes, named after J.B. Gunn of IBM, are primarily used as oscillators in the microwave range. A Gunn diode is only a diode to the extent that it has two connections. It has no rectifying effect. In microwave applications, Impatt diodes (Impact Avalanche Transit Time) are also often used as amplifiers after Gunn diode oscillators.

Light-Emitting Diodes (LEDs) use the fact that photons are created if special crystal materials are used in the P-N junction. Read more about LEDs in the Optical Components Fact Sheet.

Solar cells are also a type of large diodes where one uses the photoelectric properties of semiconductors. When photons are generated close to the P-N junction, pairs of ''holes'' and electrons are formed. The voltage is approximately 0.5 V for each cell and maximum current is dependent on the surface of the cell, but often lies in the 1 to 2 A range. By employing connection in parallel or in series, it is possible to construct solar energy systems virtually unlimited in size.



Some common diode symbols.

Transistors

The transistor has the ability to work both as a current and as a voltage-controlled amplifier. Transistors usually have three connections. These connections are called Emitter, Base and Collector in bipolar transistors. In field effect transistors they are called Drain, Gate and Source. The term transistor derives from the words ''*trans*fer'' and ''res*istor*''. A transistor can be regarded as a conductor of an electrical charge and a variable resistor. The bipolar transistor functions as a current amplifier. A small current in the base enables a greater current between the collector and the emitter. The equivalent to the base in the field effect transistor is called a gate, but instead of current it is the voltage on the gate connection that enables current between the source and the drain.

There is a great number of specialised transistors for various application areas available today. Transistors appear as individual components, often as power stages or low noise amplifiers, but they are, above all, building blocks in integrated circuits. **Small signal transistors** can be optimised for low noise and/or high frequency. **Switching transistors** must be fast and have a low saturation voltage drop. The properties of **power transistors** should, as their name suggests, include a high power and/or voltage tolerance. Some transistors, like e.g. HF power transistors, have special design and structure in order to optimise the high frequency properties.

Field effect transistors are the most common discrete components today. They have taken over the role traditionally held by bipolar transistors. Integrated circuits, in combination with field effect transistors, provide, in most cases, unsurpassed performance in both low and high frequency applications as well as in power supply and switch applications.

The **bipolar transistor** can best be described as two diodes directed toward the base (PNP) or from the base (NPN).

Unipolar transistors can be divided into JFET (Junction FET) or field effect transistors of barrier layer type, and MOSFET (Metal Oxide Semiconductor FET) transistors. A JFET is based on a barrier layer whose width varies according to the voltage applied. It has extremely high input resistance and can be regarded as a voltage-controlled current generator. In a MOSFET, input resistance is even higher and the controlling electrode can be viewed as isolated. The input resistance is at least 100 MΩ. The input capacitance, however, makes the impedance decrease as frequency increases. Powerful power-MOSFETs may have a very high input capacitance, from a few hundred to several thousand pF, which is an important factor in design work, even where low frequency final stages are concerned.

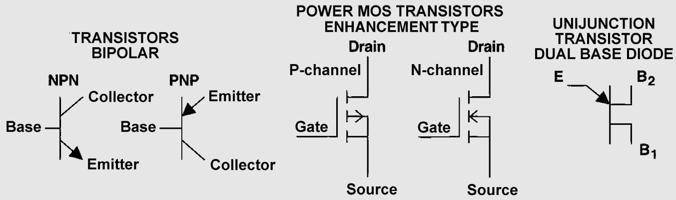
MOSFET transistors have a dominating position today as power switches with their excellent characteristics regarding switch time, power tolerance, large SOA (Safe Operating Area) as well as good dV/dT properties.

Field effect transistors have many advantages. One such important advantage is that the negative temperature coefficient for the transistor's output current is capable of preventing a thermal current surge in linear stages.

There are two kinds of field effect transistors: depletion mode and enhancement mode. The enhancement mode type does not draw any current until gate voltage is added. The depletion mode type, though, does draw current when the base voltage is zero. To throttle the transistor, the gate must be given a positive voltage if it is a P-type FET or a negative voltage if it is a N-type FET.

The **IGBT** (Insulated Gate Bipolar Transistor) is a component that combines the advantages of power-MOSFETs and bipolar power transistors in an excellent way. The IGBT has a low power loss in combination with the bipolar transistor's high tolerance for complex loads and the MOSFET's simple driving.

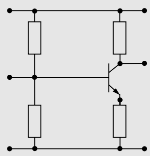
Double-base diodes are also known as **unijunction transistors**. In principle, they are constructed like a homogeneous N-doped bar. There is a P-doped region in the middle of this bar. This creates two diodes with inverse parallel connection with base terminals at either end of the bar and the emitter at the P-doped region. If a voltage is added to this bar, the potential will be proportional to the distance from one end. This means that the bar functions as a voltage divider. If the voltage between the emitter and the negative base terminal falls below the bar's potential at the emitter point, the P-N junction will be blocked. If the base-emitter voltage exceeds this potential, the resistance in the bar will decrease since the emitter attracts some of the electrons coming from the negative base. The base-emitter voltage will thereby decrease as the emitter current increases. The result will be a characteristic with negative resistance that can be used to create an oscillator. Double-base diodes are used in e.g. pulse oscillators and trigger circuits.



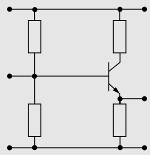
Some common transistor symbols. The MOS-transistor of depletion type is drawn like the enhancement type but with a continuous line between drain and source. Continuous = conducting when idle, broken = breaking when idle.

Basic circuits for transistors

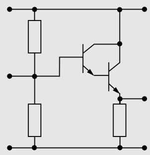
CE
Advantage: High amplification.
Drawback: Can be unstable in some cases.



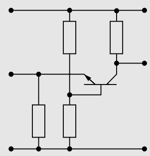
CC
Advantage: Low output impedance.
Drawback: Voltage amplification lower than 1. Instead, current amplification is used to generate an impedance transformation.



Darlington (CC)
Advantage: Low output impedance, high input impedance.
Drawback: Voltage amplifi-cation less than 1. Two emitter drop voltages limit the output swing at low feed voltage.



CB
Advantage: Stable against self-oscillation and often used in HF circuits.
Drawback: Lower amplification than CE circuits.

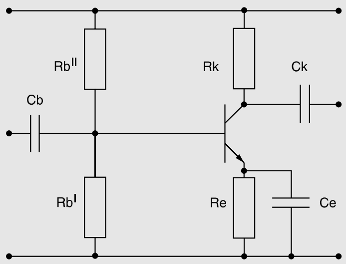


Tips on dimensioning of the CE-circuit

Let the potential drop be 1 V or more over Re for good temperature stability and so that the amplification is not affected so much by the spread in the transistors' current amplification factor. Voltages over Rb^I should be changed to 1 + 0.7= 1.7 V for silicon transistors as the base emitter voltage drop is approximately 0.7 V (somewhat less for small-signal transistors and higher for power transistors).

Re gives negative emitter feedback, which decreases the amplification. This stabilises against temperature drift and reduces the amplification spread in the connection due to the difference in current amplification factors in individual transistors.

At frequencies higher than zero (direct current), you do not want negative feedback, but the greatest possible amplification. This is why Ce exists as a short-circuit to earth. The value of Ce in relation to Re determines the lower limit frequency. The choice of Cb and Ck is also related to the lower limit frequency. Note that the input impedance is determined by parallel-connected values of Rb^I and Rb^{II}, even in parallel with the transistor's input impedance.



Dimensioning of the CE-cir-

Thyristors

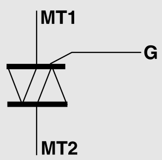
Thyristors are four-layer components (PNPN). The symbol is the same as for a diode, but with one additional connection, Gate or Knob. Thyristors can be characterised as two transistors connected to each other. Thyristors do not become conductive (ignite) until the knob is connected to a positive voltage and a control current I_{gt} is added. The thyristor will remain ignited irrespective of whether the control current is broken or whether a voltage with negative polarity is connected to the knob. It can be returned to blocking mode by:

– Reducing the anode current to below the holding current I_h (this is specified in data sheets).
–Breaking the anode current.

The **triac** can be described as two inverse parallel connected thyristors built into the same package and with a common knob. It is triggered by positive or negative pulses and switches off when the voltage above it is zero. The terminal closest to the knob is called MT1 (main terminal 1) and the other MT2. The trigger pulse is always referred to MT1.

A useful component in control circuits is the trigger diode, or **diac**. It can be characterised as a triac without a knob. It has a low ignition voltage, about 30 V. When that voltage is exceeded, the diac ignites and lets the trigger pulse through onto the triac.

TRIAC SYMBOL



Component designations for semiconductors

Several independent systems have evolved and are used today for semiconductor designations. The world's most widespread coordinating organisation in the electronics industry is Joint Electron Device Engineering Councils (JEDEC). All JEDEC-affiliated manufacturers produce components according to central specifications. The oldest European organisation for standardising and administering type numbers is Pro Electron. It was formed in Brussels in 1966. The system allows component grouping according to their areas of use as well as to materials.

The European Pro Electron System

Two or three letters followed by a 3- or 4-digit group of numbers provide a rough understanding of the component type as well as the power class.

The first letter indicates the material:
A Ge, germanium or a material with a band gap of 0.6-1 eV
B Si, silicon or another material with a band gap of 1-1.3 eV
C GaAs, gallium arsenide or another material with a band gap greater than 1.3 eV

The second letter indicates component type:

A	Diodes, signal, low-level
B	Capacitance diodes
C	Transistors, low frequency, low level
D	Transistors, low frequency, power
E	Tunnel diodes
F	Transistors, HF, low level
H	Diodes, Hall-effect components
L	Transistors, HF, power
N	Opto switches
P	e.g. Photo transistors
Q	e.g. LEDs, laser diodes
R	Thyristors, low level
S	Transistors, switch, low power
T	Thyristors, power
U	Transistors, power, switch
W	Surface wave components
X	Diodes, HF-multiplicator
Y	Rectifiers, booster
Z	Zener diodes, voltage reference

A third letter indicates that the component is designed for industrial or professional applications. This letter is usually W, X, Y or Z. After these letters comes a serial number with 3 to 4 digits and, in some cases, an additional letter that can indicate the amplification factor for example.

The American JEDEC System

The American system is not entirely without its ambiguities. Roughly speaking, a transistor that begins with 2N, like e.g. 2N2222, can be a bipolar transistor, while 2N3819 is a JFET. A designation that begins with 3N, like e.g. 3N128, indicates that it is a MOSFET. Some manufacturers also use letter designations like TIP34, MJE3055, etc.

The Japanese JIS System

First digit:
1 Two connections
2 Three connections
3 Four connections

The two following letters:

SA	PNP transistors and Darlington (HF)
SB	PNP transistors and Darlington (LF)
SC	NPN transistors and Darlington (HF)
SD	NPN transistors and Darlington (LF)
SE	Diodes
SF	Thyristors
SG	Gunn diodes
SH	Unijunction transistors
SJ	P-channel FET
SK	N-channel FET
SM	Triacs, bidirectional thyristors
SQ	LEDs
SR	Rectifier diodes
SS	Signal diodes
ST	Avalanche diodes
SV	Capacitance diodes, PIN diodes
SZ	Zenerdioder

The serial number consists of two to four digits within a number range between 10 - 9999. This is followed by a suffix that consists of one or several letters. The last letter indicates the area in which the semiconductor is used.

D Approved by the Japanese telecom authorities (NTT)
G The component is used for communication
M Approved by the Japanese Navy (DAMGS)
N Approved by the Japanese Broadcasting Corporation (NHK)
S Designed for industrial applications

The Japanese Industrial Standard (JIS) designation does not indicate whether a semiconductor is manufactured from silicon or germanium. The first two characters are frequently omitted on drawings as well as on the printed serial number on the components themselves. This means that a 2SC940 type transistor may very well be marked C940.

Thermionic valves

Thermionic valves, which are probably regarded by many as the predecessors of transistors, have not been replaced in all applications. There are a number of special cases in which these valves are still used, e.g. for transmitter power amplifiers for high frequencies. X-ray tubes and Geiger-Müller tubes are of special designs, as is the cathode ray tube. Further into the future, we will be needing valves in the form of spares.

One area in which these valves are really undergoing a revival is in power amplifiers for audio use. The distortion of these valves is different to that of bipolar transistors.

The even tones dominate in these valves, and the uneven harmonic components which are so unpleasant to the ear are weaker. The saturation properties in the output transformer contribute towards this "valve sound" sought by some. This is the case not least in guitar and bass amplifiers in which the signals are often clipped. The softer clipping means that it is possible to modulate the power amplifier to a higher average power without it sounding poor. This is the reason why a valve amplifier is able to sound a lot louder than a transistor amplifier, despite the fact that they both have the same measured output.

The newly awakened interest in these valves for audio use has led to the development of special audio valves. They are also available in matched pairs, selected according to spectrum analysis.

Certain properties can be deduced from the designations. In this respect, the USA and Europe have different standard designations:

European standard designations

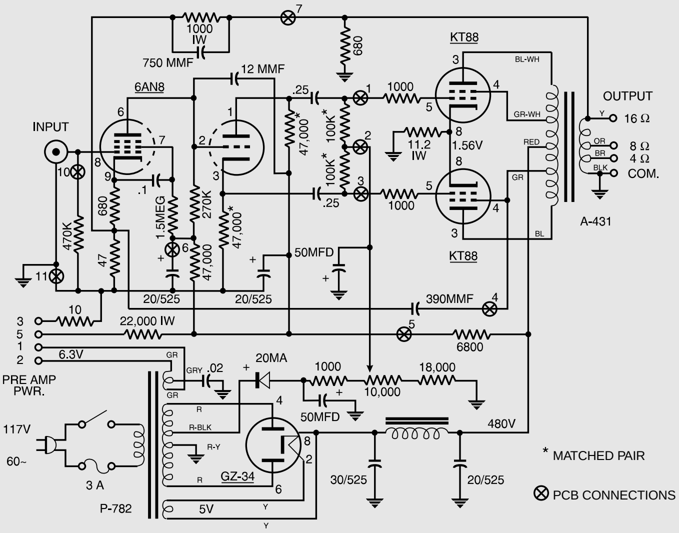
● The first letter indicates the filament voltage/filament current: A = 4 V, E = 6.3 V, D = 1.4 V battery voltage, G = 5 V, H = 150 mA series filament, K = 2 V battery voltage, P = 300 mA series filament, U = 100 mA series filament, V = 50 mA series filament. The first letter Q indicates that this is a tetrode for transmitter power amplifiers.

- The second letter specifies the type of valve: A = diode, B = double diode, C = triode, F = pentode for small signals, H = heptode, L = power pentode, M = "magic eye", Y = rectifier.
- The third letter states whether the valve has dual functions or more functions: ECC83, for instance, specifies a double triode with a 6.3 Vac filament.

Sometimes digits and letters are reversed in order to indicate that the valve concerned is of a special type. E83CC, for example, is the same as ECC83 but is a long-life model.
There are also letter-notated special valves which do not follow this system at all. The British valve KT66, for instance, has a 6.3 V filament voltage in spite of the misleading first letter K, which should have meant 2 V.

American standard designations

American valve designations generally start with a digit which indicates the filament voltage, but otherwise it is not possible to deduce anything from these designations. The American valve 12AX7 is the equivalent to the European valve ECC83, which may seem confusing. This is due to the fact that this double triode has two filaments which can be connected in parallel for 6.3 V or connected in series for 12.6 V.



One of the best valve power amplifiers made in the 1960s was the Mark III from Dynaco. Production of this valve power amplifier ceased a long time ago, but experts in analogue technology who have practical skills should be able to make a power amplifier themselves, possibly with certain modifications.

Optical components

LEDs

LEDs (Light-Emitting Diodes) emit light (photons) when current is transmitted in a forward direction from the P material to the N material. Light, which has a rather well-defined spectrum, is created by a recombination of charge carriers in the P-N junction. Materials in group III and V, but also from II and IV of the periodic table, are primarily used for semiconductor materials. They are therefore known as III-V or II-IV materials. The most widely used materials and their typical colours (wavelengths) are:

Gallium arsenide, GaAs, produces infrared to red light (650 nm).
Gallium arsenide phosphide, GaAsP, produces red to yellow light (630-590 nm).
Gallium phosphide, GaP, produces green to blue-green light (565 nm).
Gallium nitride, GaN, produces blue light (430 nm).
Indium gallium nitride, InGaN/YAG, produces white light.

Forward-biased voltage is applied to LEDs. This means that the current has to be limited by a series resistor. The forward voltage drop is approx. 1.4 V for GaAs, 2 V for GaAsP and 3 V for GaP.
For hole-mounted diodes, the cathode lead is usually shorter than the anode lead. When it comes to surface-mounted diodes, the cathode side is usually colour-marked.
LEDs come as independent components or as parts of segmented modules (displays) but can also be found in buttons with built-in illumination. Two-coloured LEDs use two diodes connected in parallel but turned in opposite directions for simple forms of indication or multi-colour combinations in displays and buttons.

Light detectors

Light detectors operate with or without an external voltage source. It is a collective term that covers several types of components.

Photo-diodes are actually regular diodes with reverse-biased voltage applied. When the P-N junction is lighted, leakage current will increase. Schottky diodes function in the same way, i.e. with a metal/semiconductor junction.

Photo-conductors, or **photo-resistors**, alter their resistance according to the light level. The highest sensitivity occurs at a certain wavelength determined by the selected semiconductor material and the interference degree. They have broad bandwidth and are light-sensitive, but have a long reset time.

PIN diodes use reverse-biased voltage. They have broad bandwidth and low noise and are very fast.

Photo-transistors function in the same manner as regular transistors, but the excess charge in the base is created by incoming light instead of current. Photo-transistors are somewhat slower than photo-diodes.

Avalanche photo-diodes are faster than photo-transistors. They also have higher gain.

Opto-isolators

Opto-isolators consist of a combined light transmitter/light detector. They can be used to transfer signals between units that are not galvanically connected with each other. The transmitter of an opto-isolator usually consists of a LED, and the receiver consists of a photo-resistor, photo-diode, photo-transistor or photo-triac. Opto-isolators often replace pulse transformers, like e.g. in primary switching power supplies. They can easily be automatically installed and have no lower limit frequency unlike pulse transformers. There are also opto-isolators with such linear characteristics that they can be used to transfer analogue signals.

Laser

Laser, Light Amplification by Stimulated Emission of Radiation, appears as:

Optically pumped solid-state lasers, e.g. ruby lasers.
Gas-discharge lasers, e.g. HeNe lasers.
Current pumped semiconductor lasers.

A laser generates light of a particular frequency and all outgoing radiation is in phase, so-called coherent radiation. The semiconductor laser is a P-N junction in which holes and electrons recombine, creating photons just like in a LED. The difference is that the LED does not have the advantage of gain through stimulated emission offered by the semiconductor laser, which produces a much more intense and coherent light. In general, GaAs with very high doping is used. The P-N junction is layered in a rectangular cross section with the end faces acting as mirrors to the laser cavity.

Displays

Displays, or display windows, can be made up of cathode ray tubes, like in TV receivers and monitors, LCDs (LCD = Liquid Crystal Display), electroluminescent screens (EL) and, when it comes to smaller displays, LED matrixes.

Liquid Crystal Displays (LCDs) consume very little power and are therefore suitable in battery-powered systems. A liquid is held between two substrates on which electrodes have been etched. When this liquid is exposed to an electric field, the crystals change direction, interrupting the light. The various types have very different characteristics. The early types, called Twist Nematic, had very poor contrast and the viewing angle was quite limited. This is a particular problem when it comes to larger displays. So-called Super Twist Nematic (STN) LCDs provide significantly enhanced contrast and the display can be viewed from angles of ±45 degrees. LCDs do not emit any light but are frequently backlit in appropriate colours. The illumination used for backlighting can be provided by LEDs, cold cathode tubes or electroluminescent (EL) sheets. Reflective displays reflect incoming light and transreflective displays have a reflecting background that do let light through and which can therefore be backlit.

LCDs -cSTN displays cSTNs (Colour Super Twist Nematic) are passive colour displays with low power consumption suitable for battery-powered systems, especially if they are of reflective or transreflective type.

LCDs -TFT displays TFTs (Thin Film Transistor) provide good contrast, 40:1, and high speed that allows moving images. Increase of the contrast is achieved by each pixel having its own transistor manufactured in amorphous silicon directly on the glass screen. The transistor provides, by means of its amplification, a more powerful drive of the liquid crystal. The light transparency is only about 3 % which means that the backlight consumes much energy. TFT displays in colour use the same technique as cathode ray tubes for colour. The individual pixels are oriented in a RGB (Red-Green-Blue) format.

LCDs -LTPS-TFT displays LTPS-TFTs (Low Temperature Poly Silicon-Thin Film Transistor) resemble regular TFT displays, but by using crystallised silicon on the glass substrates one can achieve a higher degree of integration and a larger amount of driver electronics can be constructed directly on the glass substrate which enables low-power displays for battery supply.

EL displays have good luminance, c. 100 cd/m^2, and relatively good contrast, c. 20:1. The colour is yellow. A supply voltage with a minimum of 80 V and a minimum of 60 Hz is connected to a layer of zinc and phosphorus. The voltage causes a migration of electrons in the phosphorous material which emits light.

Plasma displays (Gas discharge displays) have excellent contrast, up to 150:1, but they must be supplied with a high voltage. There are plasma screens for television use with high luminance, c. 400 cd/m^2. The cells in a colour plasma display are able to function because ionised gas emits ultraviolet radiation which, in turn, provides energy to phosphorous points in the same way as the electron beam in a CRT for TV.

Vacuum fluorescent displays have a high light intensity with a luminance that is about 45 times higher than that of EL displays. The colour is often green, but can also be white, orange or blue.

Cathode ray tubes (CRTs) are still the type of indicators that provide the highest luminance, up to 700 cd/m^2, as well as high contrast. The operating system around a CRT is elaborate: a video amplifier for brightness control, a complicated deflection system, convergence correction for colour tubes and often circuits to counteract the image distortion caused by e.g. a flat and square screen.

Display modules mean that the display itself, whether it is of LED, LCD, TFT, VF or EL type, is supplemented with driver electronics that multiplex the segments or that contain a decoder and usually a microprocessor to be operated directly with an ASCII code or a video signal.

The **speed**, in other words the transition from white to black or vice versa, varies distinctly between different indicator types. A LED display can change within 10 ns, while the transition time may be less than 0.1 ms for a CRT, about 1 ms for a plasma display, 0.1 ms to 1 s for an EL display and 10 ms to 1 s for a LCD/TFT. The transition time for liquid crystals increases rapidly as the temperature drops and their function is often completely gone at temperatures below –20 degrees.

Operational amplifiers

The operational amplifier first appeared in the 1960’s and has developed in many respects since then. In rough terms, it can be described as an amplifier with one inverted (–) and one non-inverted (+) input. The voltage difference between these is amplified and normally the operational amplifier has a very high gain. With feed-back the desired gain can be obtained. The feed-back increases the bandwidth and improves the linearity. Most types of operational amplifiers can be fed back down to a gain of one, without any stability problems. Some types cannot manage this, and must then be compensated with an external RC circuitry.

Which *maximum swing* you can obtain on the output depends on the voltage supplies that are used. Traditionally ±15 V has been dominant, but nowadays there are many amplifiers belonging to several families with a number of different application areas and supply voltages. A single supply voltage means that it is not as easy to work with balanced signals.

Some applications require low *offset voltage*, i.e. low voltage deviation, on the input and low *temperature voltage drift*. To meet these requirements, *the chopper amplifier* was developed. The input voltage is chopped with a high frequency in an analogue gate and a capacitor stores between the sampling occasions. This chopper technique makes it possible to obtain a fault voltage of only ±1 µV. The voltage drift can be as low as 0.05 µV/°C. Voltage drift can be as low as 0.05 µV/°C. The chopper technique is primarily used for static or very low frequency signals.

When both of the operational amplifier’s inputs are used as a single balanced input, it is important that the common mode signals are balanced out. The data sheets specify the attenuation in dB, so-called CMRR or *Common Mode Rejection Ratio*.

Speed is usually specified in *slew rate*. This means the maximum voltage derivative or simply how many volts the signal can raise in a single µs. High voltage derivative corresponds to wide bandwidth.

The noise is determined by the device’s *noise factor*. It is usually specified as nV/√Hz. This means that the noise voltage increases with the square-root of the bandwidth used.

A high input impedance is required in many applications. In these cases, it is suitable to use an operational amplifier with FET or MOSFET at the input. With BIFET technique, one can mix FET and bipolar circuits on the same chip. MOSFET amplifiers provide an even higher input resistance, since the inputs are purely capacitive in principle, but in practice the input resistance is the same as for FET. This is due to the fact that MOSFETs must be protected with protective diodes and that the leak current inside them decreases the input resistance.

Two special cases of operational amplifiers are the Norton amplifier and the transconductance amplifier.

The *Norton amplifier* has a very low input resistance and it is completely current-controlled. In principle, both inverted inputs can be seen as a diode. It could also be called a ”current differential amplifier”.

The **transconductance amplifier** has a high ohm differential input. Characteristically, there is a third input from which you can control the current gain.

Comparators are in fact constructed in the same way as operational amplifiers, but are optimised to quickly switch from full positive to full negative output voltage, and vice versa, upon a small change of the input voltage. A certain degree of positive feed-back is sometimes used to provide hysteresis for the switching levels. It provides more reliable switching and reduces the risk of oscillation if the input signal changes slowly.

An **instrumentation amplifier** is a development of the operational amplifier with integrated resistors that provide a fixed or programmable gain. High suppression of common mode signals (CMMR) is another requirement since the differential input is often used in measurement applications. The instrumentation amplifier often constitutes a high ohm load for signal sources with very low output voltage.

A **unity gain amplifier** is an operational amplifier where the input is connected to the operational amplifier’s output. This type of operational amplifier provides a gain of 1, hence its name. It is used to increase the drive ability, like an emitter follower, and can also be used within a feed-back loop, e.g. after an operational amplifier.

A **low power amplifier** is specially adapted to draw as little power as possible and have the lowest supply voltage possible. There are amplifiers that draw less than 1 uA standing current and low voltage variants can perform well in many portable measurement applications that contain only two battery cells acting as power supply.

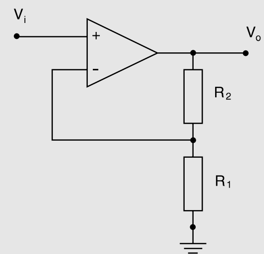
A **video amplifier** has been optimised to amplify video signals. These operational amplifiers normally have bandwidths over 100 MHz. They have also been adapted to have low noise and good phase properties. Additionally, many types have such a high operational ability that they can be loaded with 75 Ω.

A **low noise amplifier** designed for special measurement, audio and video applications offers noise properties that enable the design and construction of advanced systems for sound and video production with professional performance.

An **insulation amplifier** performs a linear transfer of signals where the inputs and outputs have galvanically separated earth potentials. This transfer can be optical, inductive or capacitive. This kind of amplifier can handle many thousands of volts between the inputs and outputs and insulation resistance can be higher than 10 MΩ. This amplifier type is also suitable when it is desirable to suppress common mode signals of more than 100 dB. This means that you can take care of small signals that are attached on a strongly varied potential. Application examples can be found within medical technology where it is vital to monitor patients and have a high insulation resistance between patients and equipment. Collection of measurement values in environments with a high noise field is another application.

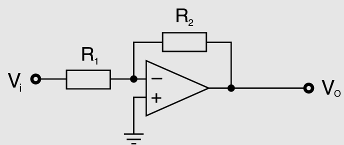
Non-inverting amplifier

$$V_o = \left(1 + \frac{R_2}{R_1}\right) V_i$$



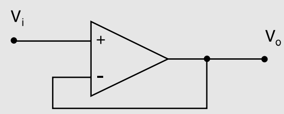
Inverting amplifier

$$V_o = -\frac{R_2}{R_1} V_i$$



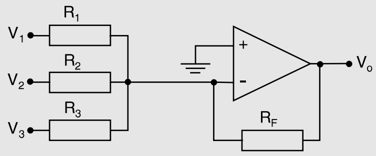
Voltage follower

$V_0 = V_i$
High input impedance and low output impedance.



Sum amplifier

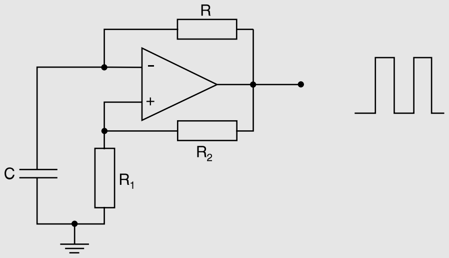
If R_F, R_1, R_2 and R_3 are made equally large, the output voltage becomes:
$V_0 = -(V_1 + V_2 + V_3)$



Square wave generator

$$f_o \approx \frac{1}{2\,RC\,ln\left(\frac{2\,R_1}{R_2}+1\right)}$$

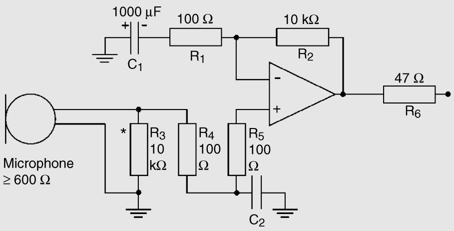
R_1 and R_2 are selected so that R_1 is approx. 1/3 of R and R_2 is 2 to 10 times R_1.



Microphone amplifier

Gain = 40 dB
For 20 dB gain $R_2 = 1\ k\Omega$

* R_3 should be 10 × the microphone impedance



A/D and D/A converters

Circuits for converting analogue signals to digital, and vice versa, are finding an increasing number of applications. There are many reasons for this. Digital circuits and microprocessors are cheap to mass-produce. Purely analogue circuits are difficult to produce as consideration to be given to analogue parameters such as noise, voltage drift, voltage discrepancies, frequency response, etc.

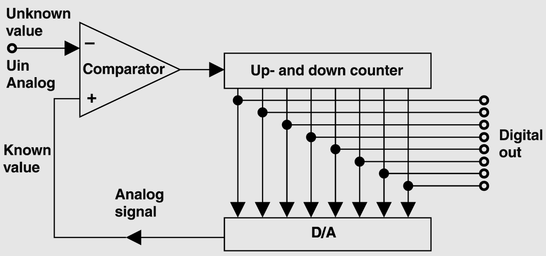
Through digital signal handling of what were originally analogue signals, it is possible to gain better control of the system's parameters and thereby reduce the need for fine tuning in production and subsequent control measurements and adjustments during service. Traditional analogue circuits in e.g. communication radios are increasingly being replaced with signal processors. The software in these deals with the algorithms which produce the impact of e.g. a filter (IIR or FIR), a detector or a modulator.

A/D converters

A common area of application for the circuits is in computers, e.g. for collecting measurements. The analogue measurement values are converted into digital words in an analogue/digital converter. This A/D converter is usually preceded by a multiplexer, which ensures that a single converter can handle the measurement readings from different sensors in the right order. There are A/D converters with built-in multiplexers and with matching interfaces directly to a microprocessor, which simplifies connection and saves circuits. Occasionally, a sample-and-hold-circuit on the A/D converter's input is used to freeze an analogue value during the conversion period.

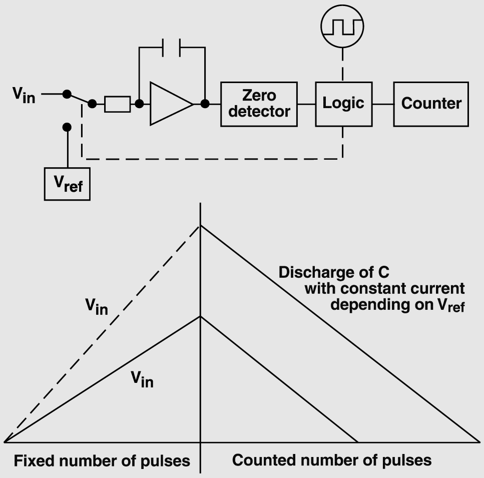
The conversion time varies greatly depending on which principle is used. A/D converters follow three main principles: Ramp, successive approximation and flash.

A/D converters with **successive approximation** are the most common. They try to convert the largest bit (the most significant) first, followed by the second-largest, etc., and continue until the digital value corresponds with the analogue value at the input.



A/D-converter using successive approximation.

Multimeters generally have a **ramp converter**, unless they are precision instruments which use successive approximation. The ramp converter takes a fairly long time to perform a conversion, but it is cheap to produce. The are a great many variants of this principle. In digital multimeters, dual ramp converters are most common. For a fixed period of time, e.g. for 1,000 clock pulses, a voltage is built up over a capacitor. The voltage over the capacitor is proportional to the measurement voltage (input voltage). The input signal is then disconnected. The capacitor is discharged over a number of clock pulses. The counter setting of these corresponds directly to the input voltage.



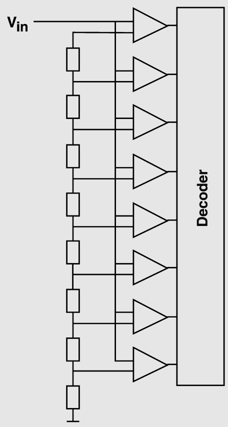
A/D-converter, ramp converter.

Some applications require extremely fast A/D converters, e.g. digital oscilloscopes and digital spectrum analysers. In this case, the **flash converter** is unparalleled as regards speed. Instrument manufacturers produce converters for their own use which can handle 1 GHz or more. There are standard types on the open market which can handle several hundred MHz. The flash converter basically comprises a ladder of comparators. These convert at the same time and immediately produce a digital value.

A variant on this theme, the **"half-flash"**, performs the conversion in two steps. As a result, the method is half as fast as full flash, but it does produce higher resolution for a given cost. As flash converters have a large number of comparators at the input, they have a low and widely varying input impedance. They should therefore be preceded by a driver with good driving capacity to ensure that the impedance variations do not produce linearity errors.

By using mean value formation, it is possible to increase the number of bits' resolution over the number employed. An 8 bit converter can therefore produce e.g. 10 bit resolution. The conversion requires a number of sets of words to give a mean value, and the conversion time is therefore extended drastically.

A special type of mean value formation is the **sigma/delta converter**. These are also called delta/sigma or bit stream converters. This is basically a 1 bit converter (!) which uses mean value formation to produce up to 20 bits, although with extremely low bandwidth. This technology is now used in CD-players. Bit stream converters are cheap to produce, give good linearity, and the problem of voltage peaks for the largest bit transitions is eliminated. They are also cheaper to produce because most of the circuit is made up of digital functions.



A/D-converter, flash converter.

D/A converters

D/A converters can be constructed with weighted resistance (1, 2, 4, 8, 16 Ω, etc.), or with ladder networks supplied with current or voltage. Other variants are also available in monolithic circuits.

The specifications of a D/A converter include information about resolution (number of bits, accuracy of output signal), settling time and "slew rate", max change coefficient of the output signal. Audio equipment, e.g. CD-players, places extremely high demands on the performance of D/A converters.

Logic circuits

A simple way of classifying today's logic families is to examine their supply voltage. The traditional TTL (Transistor-Transistor-Logic) family is supplied with 5 V. Today, there are also families for 3.3 V and 2.5 V designed to meet new requirements of memory components as well as new, extremely fast bus architectures.

Another classification is based on technology. The usual Bipolar and CMOS technologies have been complemented with BiCMOS so that three common technologies exist today.

Bipolar families

Bipolar families normally operate with a supply voltage that is relatively critical. For the 74 family (5 V supply voltage) it should lie within the 4.75 to 5.25 V range. Voltage should be disconnected from certain dispersed parts of the construction since the load varies according to the signals and current spikes arise when the transistors of the totem pole outputs can conduct current simultaneously on a short term basis. The connection wires of the by-pass capacitors should be as short as possible. In order to obtain a sufficient interference margin, even when circuits are handling fast signals, the groundplane should be stable. The level limits for a TTL output are a maximum of 0.4 V for ''0'' and a minimum of 2.4 V and up to the supply voltage for ''1''. The fact that the ''1'' level is not higher is due to the TTL output's structure with a voltage drop across a 130 Ω resistor, a transistor and a diode section. For an input, a maximum of 0.8 V and a minimum of 2.0 V apply. In the worst case, one should therefore count on an interference margin of 0.4 V for ''1'' and 0.4 V for ''0''. One usually counts on 0.7 V for ''1'' however.

74 Standard-TTL is the original TTL family. More modern variants are used today to great advantage. The delay is approx. 10 ns per gate and power development is 10 mW.

74S Schottky-TTL was the first TTL family of fast circuits. An included Schottky diode prevents a transistor from bottoming out. Today, the faster AS is preferred. The delay is approx. 3 ns per gate for S-TTL. The power is 20 mW per gate. The Schottky diode consists of a metal N junction instead of a P-N junction as in a normal diode. The Schottky diode has low capacitance, a lower forward voltage drop than the silicon diode and it is also easy to integrate.

74AS Advanced Schottky-TTL has a delay of approx. 1.5 ns per gate. The power consumption amounts to 22.5 mW per gate.

74LS Low Power Schottky-TTL is used today as a replacement for Standard-TTL. The delay is approx. 9 ns per gate, i.e. somewhat faster than the Standard-TTL. Besides, power consumption is only 2 mW per gate.

74ALS Advanced Low Power Schottky-TTL combines speed and low power consumption. This means a step delay of 4 ns and 1 mW per gate.

74F FAST-TTL is extremely fast with a delay of 3 ns and a power consumption of 4 mW per gate.

CMOS families

One can discern two main groups among these circuits. Those that operate with CMOS levels and those that operate with TTL levels. The latter group can be used with bipolar TTL circuits if certain design rules are observed. The circuits' outputs consist of complementary MOS transistors, hence the name. The standby power consumption is very low, c. 10 nW per gate. It rises with increased working frequency however and at a few MHz it is roughly the same as in ALS-TTL. The interference margins are much higher for the CMOS circuits. They can be increased further by raising the supply voltage. Thanks to this facility, CMOS has taken over the role as high level logic from the earlier bipolar high level families. This is valuable in industrial environments where the working frequencies are moderate. One should note that the combination of a high operating voltage and a high clock frequency may result in too great a power development in the circuits.

The **4000 family**, the earliest among the CMOS families, appeared towards the end of the 1960's. It is slow compared to the TTL families. The delay is approx. 20 ns per gate. The pin configuration also differs from TTL. The supply voltage can be between 3 and 15 V (18 V in some cases). The family also includes buffered versions, 4000B. Compared to unbuffered versions, they have longer throughput delay but a better interference margin, constant output impedance, higher gain and lower input capacitance. There is a risk of oscillation if buffered versions are supplied with a slowly rising edge.

74C is a variant of the 4000 series with pin connections in accordance with TTL, but with CMOS levels.

74HC and **74HC4000** are replacements for the 74C and 4000 families. The pin configuration is unaltered but they are considerably faster. The delay is approx. 8 ns per gate. The supply voltage is allowed to lie between 2 and 6 V. The interference margins are 1.4 V for both high and low level.

74HCT constitutes a variant of HC, adjusted for TTL levels. It has the same speed as HC. The supply voltage is allowed to lie between 4.5 and 5.5 V. The interference margins are 0.7 V for low and 2.4 V for high level provided that HCT is connected to HCT. If HCT is connected to LS-TTL, 0.47 and 2.4 V apply, and from LS-TTL to HCT 0.4 and 0.7 V apply, i.e. the same values as for LS-TTL to LS-TTL.

ACL with its variants **AC** (CMOS levels) and **ACT** (TTL levels) appeared in 1985. It is considerably faster than HC. The gate delay is less than 3 ns. The advantages also include a high and symmetrical drive capacity, able to drive as well as sink 24 mA. Some types drive ±48 mA or ±64 mA. The output can drive transmission cables directly. These can consist of coax cables, cables with twisted pairs or microstripline. The receiving end should then be loaded with a resistor, e.g. 300 Ω, in order to avoid reflections from the extremely high impedance input.

FCT is constructed in CMOS, but can also be connected to TTL inputs and outputs. The output is able to sink 64 mA and drive 15 mA. One variant, the FCT-T, gives 3.3 V in the high mode which means a nominally high TTL level, while the complementary CMOS transistors in FCT act as resistors to power supply or earth. FCT-T is about as fast as F while FCT is faster.

AHC Advanced High-Speed CMOS can drive 8 mA at 5 V and has a typical gate delay of 5.2 ns.

BiCMOS families

These families comprise bipolar transistors and complementary MOS transistors. The bipolar transistors are used for the outputs in order to provide a high drive capacity, while the MOS transistors are used to obtain high ohm inputs and current-saving, internal circuits.

BCT is a BiCMOS family that primarily comprises bus driver circuits. The outputs can drive 25 ohm transmission cables which means that they are able to sink 188 mA temporarily. Inputs and outputs become high ohm when supply voltage is disconnected. The input, constructed with CMOS, has a threshold voltage of 1.5 V and is therefore adjusted for TTL circuits.

ABT is a very fast BiCMOS family where the bipolar transistors have as much as 13 GHz f_T. The gate delay is 4.6 ns. The circuits are suitable for bus adjustment which requires high speed and a good drive capacity. The outputs sink 64 mA and drive 32 mA. Delay times that are independent of temperature is one of the advantages of this family. The static power consumption is very low and at high frequencies it is lower than that of CMOS.

Low-Voltage families

ALB Advanced Low-Voltage BiCMOS is specially designed for 3.3 V, providing a drive capacity of 25 mA with a gate delay of 2.2 ns. The inputs have clamping diodes to eliminate overshoot and undershoot.

ALVC Advanced Low-Voltage CMOS is a 3.3 V CMOS family with a step delay of 2 ns and a drive capacity of 24 mA. The family is specially adjusted for constructing advanced memory systems with e.g. SDRAM.

AVC Advanced Very Low-Voltage CMOS is a family that can operate with a supply voltage as low as 1.8 V with a step delay of only 3.2 ns.

ALVT Advanced Low-Voltage BiCMOS is a family adjusted for 2.5 and 3.3 V for high-speed bus systems. It has a delay of 2.5 ns with a drive capacity of 64 mA. It is also able to handle hot-swap systems with live voltage during extraction and insertion of circuit boards, so-called live insertion.

The **LVC Low-Voltage CMOS** family is a further development of 74HC, where performance characteristics like high speed and drive capacity have been maintained in spite of the fact that the supply voltage has been decreased to 3.3 V nominal. Lower voltage means lower power consumption as well as fewer battery cells in battery-powered equipment. 74LVC is pin compatible with 74HC and has a supply voltage range of 1.0 to 3.6 V. This family comprises the bulk of the circuits in 74HC and are only manufactured as surface-mounted components. When junctioned with 5 V logic, 74LV can be driven from bipolar TTL but not from 74HC(T). It is self capable of driving TTL and 74HCT. Where 74HC is concerned, it falls outside the specification which means that the driven circuit may draw more current than usual.

LVT Low-Voltage BiCMOS is 5 V-tolerant, 3.3 V family with a step delay of 3.5 ns and a drive capacity of up to 64 mA for advanced, high-speed backplane solutions.

Special logic families

ABTE/ETL Advanced BiCMOS Technology / Enhanced Transceiver Logic provides a drive capacity of up to 90 mA and is adjusted for the VME64-ETL specification.

BTL/FB+ Backplane Transceiver Logic for the IEEE 1194.1 and IEEE 896 (Futurebus+) standards. It has a delay of less than 5 ns and a drive capacity of 100 mA.

GTL/GTLP Gunning Transceiver Logic och Gunning Transceiver Logic Plus are adjusted for high-performance backplane solutions with a clock frequency of 80 MHz or more. The drive capacity is variable up to 100 mA and the outputs have rise/fall time control in order to reduce reflections and EMI.

Earlier logic families

ECL Emitter-Coupled Logic was mostly used for gaining extreme speed. The levels are typical (the MECL 10000 family) -0.9 V for ”1” and -1.75 V for ”0”, which means that the swing is 0.85 V. Internally, the operations of the circuits are based on a differential amplifier where current is directed towards one or the other of the collector outputs. This direction of current prevents the circuits from entering saturation and guarantees a high speed. There are ECL circuits that can manage a clock frequency of quite many GHz. They require relatively much power.

RTL Resistor Transistor Logic belongs to the oldest type of semiconductor-coupled logic. It is resistance-coupled and contains relatively few transistors. The outputs consist of a number of resistors to one transistor input. One disadvantage is crosstalk between the various inputs and another is that the circuit becomes slow. This type of coupling never saw the light of day as an IC. **DCTL** did however. The DCTL is a variant of the resistance-coupled logic where each input has a transistor in order to prevent crosstalk. Step delay is long however, 50 to 100 ns. Its interference margin is low though, only around 0.2 V. The sole importance of these circuits today is as spare parts to older apparatus.

DTL Diode Transistor Logic was the first large logic family. It was common in the mid-1960’s but was soon replaced by the much faster TTL family. At that time, it was not unusual to mix DTL and TTL circuits in the same design since the levels were not that far apart. In DTL, many of the resistors from RTL have been replaced by diodes which occupy a smaller silicon area. The logic tasks are performed by diodes and a transistor which resets the signal level. The output, with a transistor and a collector resistor, offered insufficient drive capacity to the positive voltage and a better drive capacity would have been needed due to the leak current in the input diodes. TTL offered a solution with its totem pole output and its input with multi-emitter transistors. Today, the sole importance of DTL is as a spare part.

DTLZ, HLL, HNIL are examples of earlier, bipolar circuit families classified as high level circuits. They are connected to 12 or 15 V supply voltage. The circuits are slow but have a very good interference margin. Capacitors are sometimes connected to make the circuits even slower to increase their interference tolerance.

Outputs

The totem pole output is the most common output type within TTL logic. It is not suitable in some contexts, however, so a couple of alternative outputs have been developed.

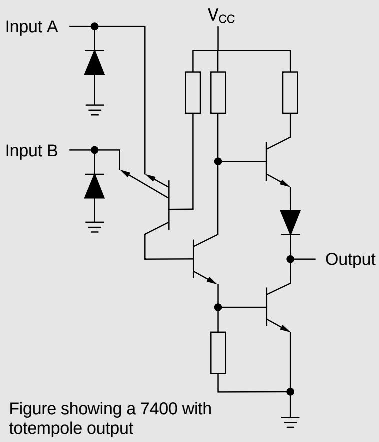
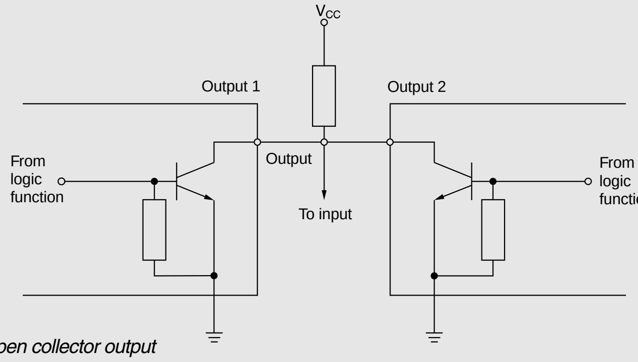


Figure showing a 7400 with totempole output

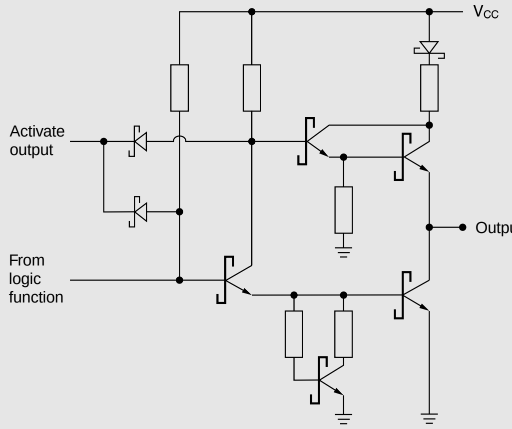
Totem pole output

The open collector output is used whenever it is desirable to activate an input from several connected outputs (wire-OR function), or as a driver for loads which are supplied with a high voltage and/or a high current. To be able to connect the output further to other logic circuits, a collector resistor is connected. An open collector output consists only of an NPN transistor’s collector connection. The output either leads to earth (on position) or is entirely open (off position).



Open collector output

The tri-state output is used to connect several outputs to a single input, e.g. in a computer bus in computers. A special connection is used to activate the output so that the transistor (one of the output transistors) will be able to conduct.



Tri-state output

Some design rules

Fan-in and fan-out must be taken into account by the designer. A gate in Standard-TTL, type 7400, has fan-out = 10. This means that it can be loaded with 10 inputs. This, in turn, means a 0.4 mA supply from a high output or 16 mA to earth for a low output. It is also possible to mix TTL-adjusted CMOS with TTL. A HCT output can e.g. be loaded with 2 pcs. of Standard TTL, 2 pcs. of S-TTL, 2 pcs. of AS-TTL, 10 pcs. of LS-TTL, 20 pcs. of ALS-TTL or 6 pcs. of F-TTL.

Power consumption can be very low by using CMOS. In quiescent mode, i.e. statically, the logic draws very little current. Dynamically, current increases with an increased working frequency. This is due to the outputs being loaded with capacitances from conductive patterns and the inputs of other logic circuits. At extremely high frequencies, there is therefore no difference in power consumption between bipolar circuits and CMOS.

Interference safety is something that has to be designed from the very beginning. In general, CMOS is superior to TTL and one should not select faster circuits than necessary. Preferably, fast circuits should only be selected where it is necessary in the design. Lower voltage to the CMOS circuits reduces the degree of generated interference, but interference margins are decreased at the same time. Use buffer circuits, transmission cables and terminations where high speed signals are to be transmitted over a longer distance. The circuits should have short connections to the transmission cable. Keep signals and earth connections joined. Allow them to accompany each other on the circuit board or position signal conductors on one side and a groundplane on the other. Watch out for earth loops that might pick up or emit interference.

Programmable logic circuits

Programmable logic circuits are increasingly replacing the traditional logic families in new designs. They can be characterised as circuits with configurable blocks of logic and flip-flops. These blocks can be connected rather freely and programmed by means of memory cells to create complex logic designs. Different kinds of architecture and a great number of manufacturers have produced a multitude of programmable logic circuits on today’s market.

Some common types of programmable logic:

SPLD Simple Programmable Logic Device
CPLD Complex Programmable Logic Device
FPGA Field Programmable Gate Array
FPIC Field Programmable InterConnect

SPLD (also known as PAL, GAL, PLA or PLD depending on the manufacturer) is the smallest and least expensive form of programmable logic. A SPLD is typically comprised of a few up to around ten or twenty macrocells. Each of these macrocells normally contain a few type 7400 serial circuits and can be connected with the other cells on the circuit. Programming is usually performed with EEPROM or FLASH technology.

CPLD (also known as EPLD, PEEL, EEPLD or MAX depending on the manufacturer) is rather similar to SPLD but has a significantly higher capacity. A typical CPLD is often 2 to 100 times larger than a SPLD and may contain from tens to a few hundred macrocells. A group of 8 to 16 macrocells is typically grouped together into a larger function block. The macrocells within such a function block are normally fully connected but not fully connected between the function blocks.

FPGA (also known as LCA, pASIC, FLEX, APEX, ACT, ORCA or Virtex depending on the manufacturer) differs from SPLD and CPLD by offering the highest integration of logic capacity. A FPGA is typically comprised of an array of logic blocks surrounded by programmable I/O blocks connected by programmable interconnect. A typical FPGA contains from under a hundred to tens of thousands of logic blocks and an even greater amount of flip-flops. Most FPGAs do not provide 100% interconnect between logic blocks. Instead, sophisticated software like routers connect the blocks in a highly effective manner. Even memory blocks and complex functions like processor cores are often integrated in a FPGA. The variation is great though between different manufacturers and FPGA families.

A **FPIC** is really not a logic device but rather a programmable ''wiring'' device. Through programming, it is able to connect a great number of I/O pins on the circuit.

VHDL

To simplify construction and programming of programmable logic circuits a common language has been developed that is a standard today in the IEEE. This is called VHDL (Very High Speed Integrated Circuit Hardware Description Language). It is a language that describes the structure and behaviour of the circuit, like which inputs and outputs the circuit has and the logic connection between them. Development tools for this purpose are also available that also can simulate the function before the circuit is programmed. For programming purposes a standard from JEDEC has been developed that provides the appearance of the file that the programming equipment uses as input data.

Microprocessors

The early history of microcomputers can be traced to ENIAC (the Electronic Numerical Integrator And Calculator), a predecessor to today's microprocessors and PC industry. ENIAC was able to perform 5000 additions and 300 multiplications per second when it was first introduced in November 1945. Its cost was half a million dollars, it consumed 150 kW and its weight, with its 19 000 electron tubes, was around 30 tons.

It was Alan Turing who introduced the idea that programme and data can co-exist in the memory of a computer. John von Neuman formulated these notions in a scientific essay in 1945, laying the foundation to ''the von Neuman architecture''. Data and programme memory co-exist, side by side with the control unit, the arithmetic unit and the I/O unit.

The key to the first microprocessor became the tests conducted by Jack Kilby (the Nobel Prize laureate for Physics in the year 2000) in October 1958 as he switched on the voltage to the very first integrated circuit at Texas Instruments. The ensuing development in the 1960's provided Frederico Faggin of Intel with the essentials for designing the MCS4004 processor in 1971. A 4-bit CPU with 46 instructions and an instruction cycle time of 10 microseconds. This processor was soon followed by the 8008, the 8080, the 68 family from Motorola and, somewhat later, the Z80 from Zilog.

There are two main groups of microprocessors. Processors for multi-circuit solutions are called **microprocessors** and single chip or single package computers are called **microcontrollers** or **microcomputers**.

The **microcomputer** is self-sufficient and requires no extra components in order to function. It contains data and programme memory integrated on the chip and I/O functions like e.g. A/D- and D/A converters as well as digital in-/outputs. There is a great selection from a great number of manufacturers available. All of these are, in one way or another, adapted in order to minimise the number of components in applications where microcomputers are used, like in everything from microwave ovens, blood analysing equipment, scales and remote controls to bank cards and musical greeting cards.

Microcomputers, even the most basic ones, are usually programmed in high level languages, even though assembler programming still exists. The most common kinds are C programming or some form of object-oriented programming (OOP) type C++.

Programme storage for microcomputers come in many different forms. Traditionally, many volume applications have used ROM (Read Only Memory), where the circuits are programmed already at the manufacturing stage. This has produced the lowest cost, at the same time requiring very large volumes to be profitable. It is more common today to use one of the field-programmable memory types like FLASH, EPROM or EEPROM, which all offer their respective, quite special advantages.

FLASH-based microcomputers offer the advantages of simple programming and updating programmes without removing the circuit. Comparisons can be made to the programme updates in e.g. BIOS for PC motherboards. With FLASH technology, it is also possible to get extremely large programme memories integrated in the microcomputer.

EPROM-based microcomputers are frequently of the OTP (One Time Programable) type. This offers the advantage of programmes that cannot be changed or updated. There also used to be an advantage in price compared to FLASH-based microcomputers who were more expensive at the time. In circuits that contain a window on the package, it is possible to erase programmes with UV light. These circuits are relatively expensive and are often only used for programme development.

EEPROM-based microcomputers offer simple programming with a virtually unlimited number of reprogrammings. FLASH and EPROM both contain limitations in that regard. It is relatively hard to make large EEPROM memories though, so the memory size is seldom larger than a few kilobyte for less expensive types.

The **microprocessors** are being developed in two directions, **CISC** or **RISC**. CISC (Complex Instruction Set Computer) has previously been the dominant processor type. RISC (Reduced Instruction Set Computer) has mostly been used for fast workstations as well as image handling and signal analysis. Modern CISC types have taken over quite many of the functions previously only found in RISC processors, like pipelining, a parallel execution which means that instructions and data are divided between separate buses (Harvard architecture), as well as cache memories, etc.

However, selecting a processor is not only a question of choosing the best performance. In some equipment the cost aspect may be dominant, in another the power consumption, etc. The development environment is also crucial, that there are emulators and software for the microprocessor or microcomputer.

The **development environment**, for programme development as well as debugging, is often crucial for the success of a microprocessor or microcomputer project. Emulation boards and systems for ICE (In Circuit Emulation) are important and time-saving tools in combination with a well-functioning high level language and debugging software.

High level language saves time. It is estimated that it takes the same amount of time to programme and debug one line in a high level language as one line in Assembler code. This makes it 10 to 100 times more efficient to write even small programmes in a high level language. It can save a lot of money, especially where smaller projects are concerned, if money is put into a more expensive component, containing a larger programme memory, instead of spending one's time developing a programme.

The **memory** in microcomputers consists of 2 kinds, data memory and programme memory. The data memory on the chip usually ranges from 1 kbyte up to 32 kbyte (256 kbit) and the programme memory from 16 kbyte up to 1 Mbyte (8 Mbit). Programme memories of several hundred kbyte offer great possibilities for effective programming in a high level language of simple as well as highly advanced systems.

Common functions frequently integrated in microcomputers are e.g. timers, watchdog, serial interface, A/D and D/A converters and display drivers.

Timers come in a multitude of versions. There are, for example, simple counters with 8-, 16- or 32-bit resolution. The oscillator or clock of the microcomputer are often used count or measure time or pulses, either directly or via a frequency divider. But there are also timers that function as advanced supplementary systems constructed with registers and which can be programmed to create advanced pulse trains or control sequences for e.g. stepper motor control or signals with pulse width modulation.

A **watchdog** (monitoring circuit) is required by most systems in order to handle blockings if the processor/microcomputer should stop. Blockings can arise as a result of interference in the supply voltage or faulty programmes. The blocking can be detected with a watchdog and the microcomputer is restarted in a preconceived way.

Serial interface. Serial communication come in many different designs. Earlier, traditional asynchronous protocols like RS232 need a UART (Universal Asynchronous Receive Transmit) or its synchronous counterpart USART. A system developed by Philips supported by many circuits is the I^2C (Inter-IC) bus. Even USB (Universal Serial Bus) and Ethernet are supported by peripheral systems today.

A/D and D/A converters have been developed with increased resolution year by year. There are often integrated analogue multiplexers for a number of channels, making it possible to monitor several analogue signals simultaneously.

Display drivers. In many cases, the only output signal from a microcomputer system is the display. With integrated display drivers, no external components are required for directly handling smaller displays.

Microprocessors and **microcomputers** can easily be divided into the word lengths (bus widths) they work with. In general, it can be said that a larger word length comes with a higher calculation performance. More bits in the word means that more digits are processed simultaneously. 8 or 16 bits are the most common word length for microcomputers, but there are also 4 bits in simple, often older microcomputers. 32 and 64 bits are standard for microprocessors and a wordlength of 128 and 256 bits are common for special graphics processors.

The **digital signal processor (DSP)** is a special type of processor designed for extremely fast processing of numerical algorithms. Many of the functions traditionally ascribed to the DSP are also included today in simpler microcomputers. One example is the MAC instruction (Multiply Add and aCcumulate) that can make most routines for signal processing more effective. In combination with integrated A/D converters, DSP instructions can enable microcomputers to handle e.g. voice recognition or routines for recognising handwriting on a LCD with press-sensitive film.

Memory circuits

All kinds of computers, whether they are consumer products like video cameras or professional products like control systems for space probes, have semiconductor memories. Memory capacity has evolved from 1024 bits per circuit back in 1971 to 1 073 741 824 bits per circuit some thirty years later. An increase that means that the memory capacity is doubled every 18th month (Moore's Law). This development can be illustrated with the example that a vintage 1970 semiconductor memory was able to store 2 lines of text. Thirty years later, some 200 standard books can be stored in a 1 Gbit FLASH memory.

The most common semiconductor memories on the market can be divided into volatile and non-volatile memories.

Volatile memories

Volatile memories are memories that lose the information when the power supply is cut off.

DRAM (**Dynamic Random Access Memory** means that information can be read without minding the sequence or the exact location of the stored information. Each memory position/cell is addressed individually/at random. The information is stored as a charge in cells, small capacitors really, whose charge is upheld by refresh logic on the memory circuit.

FPM, (Fast Page Mode) a DRAM with a technology no longer used for new design. This memory type is common in earlier PCs from the mid-nineties.

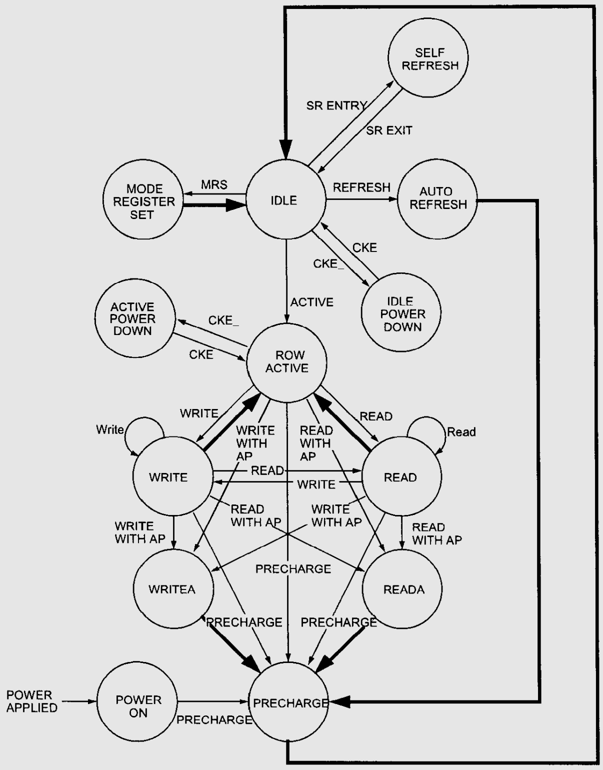
EDO (Extended Data Output) is the successor to FPM and is usually regarded as having a better access time (shorter) than FPM, requiring approx. 25 % fewer memory cycles in order to read and write an equal amount of data. This is possible as the addressing of the next cell can start at the same time as the content of the present cell is presented on the bus.

SDRAM (Synchronous DRAM) uses a separate clock to synchronise input signals to the memory. PC66, PC100 or PC133 for 66, 100 or 133 MHz clock speed are the most frequent types. The memory contains logic functions that enable data to be read in blocks, without providing the memory with new addresses. Reading keeps pace with the memory clock. For a PC100 memory, this means that memory bandwidths of several hundred megabytes/second can be obtained.

DDR-SDRAM (Double Data Rate Synchronous DRAM). By using each edge of the clock signal, the rising as well as the falling edge, it is able to transfer double the amount of data compared to a standard SDRAM.

D-RDRAM (Direct Rambus DRAM). Rambus is a company that has developed a special addressing methodology that enables the data bandwidth to be maintained at 1.6 GByte per second. In their RDRAM they have also abandoned the multiplexed address bus used by all other DRAMs. RDRAM is encapsulated in a BGA (Ball Grid Array) or CSP (Chip Scale Package) package with around 100 connections per package. These package types can easily handle many connections with maintained low capacitance and inductance, providing excellent high frequency properties.

SRAM (Static Random Access Memory) Compared to a DRAM, the A SRAM is constructed so the information does not need to be rewritten in order to be upheld in the memory. Furthermore, information is not stored as a charge in a capacitor but in a flip-flop that consists of a number of cross-connected transistors. SRAM has lower power consumption and is generally regarded as being faster than DRAMs. The most common SRAM applications are as memories in battery-fed systems and as high-speed cache memories like L2 cache for PCs.



The complex STATE diagram for a typical, modern SDRAM.

Non-volatile memories

Non-volatile memories are memories that does not lose the information when the power supply is cut off.

FLASH EPROM a electrically eraseable and programmeable read memory, has received its name due to the mode in which it is erased. The memory is organised in segments and each segment can be erased with a simple operation dubbed a "flash". The erasing functions in accordance with what is called the Fowler-Nordheim tunnel effect, where electrons pass through an extremely thin dielectric layer and remove the charge from a floating gate that is contained in each memory cell. FLASH memories are the most common kind of semiconductor memories for non-volatile storage of data. In regular FLASH memories, one bit of information is stored in each cell. The charge level of this cell determines if it is a "1" or a "0" that is being stored. Modern types of large FLASH memories store information on several levels. With four charge levels, each memory cell is able to store two bits of information. This multiple level technology is used in FLASH memories with 512 Mb or more. FLASH memories are e.g. used in BIOS circuits for PCs, cellular phones and digital cameras. FLASH is also used as storage media in different types of memory cards as well as programme storage for a vast number of microcomputer-based systems.

EEPROM (Electrically Erasable Programmable Read Only Memory), an electrically eraseable and programmeable read memory. The function of an EEPROM corresponds with that of a FLASH memory, with an additional possibility for individual programming of separate memory cells. EEPROMs cannot rival FLASH memories when it comes to size, but an EEPROM allows about 10 to 100 times more write operations than a FLASH memory. The fact that information is rewritable so many times makes EEPROMs suitable for several applications where data has to be continuously updated. Some common application areas for EEPROMs include e.g. serial memories for storing phone numbers in cellular phones, as programme memories for small, reprogrammable microcomputers or smart cards like GSM SIM cards.

EPROM (Electrically Programmable Read Only Memory), an optically eraseable and electrically programmeable read memory. EPROMs are optically erased with ultraviolet light and was the first type of non-volatile but electrically programmable read only memory to appear. This memory type already existed in the early 1970's and has dominated read only memories for 25 years, but has now been replaced by FLASH memories to a great extent. No new designs are made today with EPROMs.

PROM, a non-eraseable and electrically programmable read memory, is a predecessor to EPROM. This memory type still exists only to a much smaller extent. Programming is performed by burning built-in fuses in the circuit which are made of a nickel/chromium alloy. Therefore, the circuit cannot be erased or reprogrammed.

ROM (Read Only Memory), a non-eraseable mask programmed read memory. For a long time, read only memories with mask programming have been highly cost-effective alternatives for storing large amounts of data on a memory chip. The memory's information is added already at the manufacturing stage by modifying one of the masks which then forms the basis of the memory construction. The purchaser of the memory pays a -usually significant- fee for this mask to the semiconductor manufacturer. ROMs are usually only used in systems or apparatus that have been manufactured in extremely large series.

Transformers

A transformer consists, in its simplest form, of an iron core with two windings. If the current in the primary winding is sine shaped, the flow in the core will also change according to a sine function. This flow change induces a sine shaped voltage in the secondary winding. If the flow was constant, on the other hand, no voltage would be induced in the secondary winding. This means that the transformer does not carry direct current.

From this simple description, we can deduce that the transformer has two tasks:

- To carry an alternating voltage from the primary to the secondary side and, if so desired, to obtain a galvanic isolation between the primary and secondary sides at the same time.
- To transform (= carry and convert) one alternating voltage to another. This can be described with a simple formula:

$$U_p / U_s = n_p / n_s$$

where U_p = primary voltage
U_s = secondary voltage
n_p = turns of primary winding
n_s = turns of secondary winding

A **mains transformer** is an example of a transformation of one alternating voltage to another, e.g. 230 V to 11 V. Input power = output power – loss effects. This means that if one extracts e.g. 1 A on the secondary side, at least 0.05 A will flow on the primary side.

The transformer is dimensioned for a certain maximum power which must not be exceeded. This means that the winding resistances should be low enough not to produce too large voltage drops. It also means that the transformer's core must be big enough not to be saturated. Its size is determined not only by the transmitted power but also by the frequency. The general principle is: the lower the frequency, the bigger the core required.

The core is not solid since that would cause eddy currents to arise, which would result in considerable loss. Instead, transformer sheets stacked into a pack of insulated leaves are used. These are often cut into the shape of the letters E and I. They form an EI core together where the coil lies in the centre so that as much as possible of the magnetic field is gathered around it.

The leak flow is critical in certain applications. This applies to e.g. Hi-Fi amplifiers and measuring equipment where the field induces a hum. In these cases, toroidal core transformers are usually a better choice since they have a very small leak flow. One property of the toroidal core is that its initial current is more powerful than that of an EI core transformer. It is also able to carry interference on the network to a greater extent. Toroidal cores are seldom used for outputs exceeding 500 VA.

A transformer with separate primary and secondary windings is called a **full transformer**. This transformer type provides galvanic separation between input and output.

The **auto transformer** has a common primary and secondary winding. Therefore, this transformer type does not provide any galvanic isolation between inputs and outputs, but can be used to transform voltages both upwards and downwards. Due to the ''tight'' connection between the windings and the fact that the winding occupies less space, this transformer type is slightly smaller than one with two windings.

The **variable transformer** is usually a variant of the auto transformer in which the connection of the secondary winding is moved in such a way that the secondary voltage can be varied. It is suitable for use in laboratories where it is desirable to study the behaviour of certain apparatus under varying mains voltage. Variable transformers are also manufactured as full transformers.

The **isolating transformer** is a full transformer used to provide a power supply that is separated from the mains network. It is used in measurement laboratories, for example, where the earthed mains outlets cannot be used since acquired earth loops are then likely to influence the test results.
As we know, at least one of the power supply system's poles carries voltage to earth. The secondary winding of the isolating transformer can be left without earth connection, which means that it does not produce any voltage to earth (the secondary voltage is ''floating''). This floating voltage significantly reduces the risk for those working in the laboratory.
The transformer can also be provided with a shield between the primary and secondary side to prevent interference from being carried capacitatively.

The **safety isolating transformer** and the **isolation transformer for safety purposes** should be used to limit the risk of electric shock in certain installations and equipment. Specific cases are described in Swedish regulations. The transformers should have adequate isolation between the primary and secondary side, as well as a limited output voltage of e.g. 12, 24, 42 or 115 V depending on the application area. A safety isolating transformer provides a safety low voltage of 50 V maximum, while an isolation transformer is used for safety purposes and provides a voltage between 50 –125 V.

A so-called **toy transformer** should be used for toys as it provides a safety low voltage of 24 V maximum and is very safe even if handled carelessly.

The **bell transformer** is a safety transformer designed for doorbells and similar products. It may have a maximum short-circuit current of 10 A in order to avoid possible damage to the bell wire and only needs to handle short-term loads.

These and other safety transformer types such as **razor transformers** and **hand lamp transformers** should conform to requirements specified in international standards.

The **low frequency transformer** has a special purpose that differs quite distinctly from that of mains transformers. This transformer type is not primarily used to transform one alternating voltage to another, but to transform one impedance to another. This is used for adjustment between e.g. two amplifier stages or adjustment between a preamplifier and a loudspeaker.

The impedance conversion constitutes the square of the winding ratio (voltage ratio), i.e. a transformer with a winding ratio of 10:1 has an impedance ratio of 100:1.

Low frequency transformers for Hi-Fi use should be able to carry the entire tone frequency range 20 Hz to 20 kHz without variations in attenuation and without major phase distortion. In practice, this means that they should have an even greater frequency range. It is therefore considerably more difficult to design and construct a low frequency transformer than a mains transformer which only needs to function well on a single frequency.

The **output transformer** is a quite critical component. It has become relevant once again in Hi-Fi amplifiers as well as guitar amplifiers constructed with thermionic valves. The valves should be loaded with an optimum impedance which is derived from valve curves during the design work. This involves a significant number of kΩ which are then adapted to the low speaker impedance with a transformer. The high impedance entails many winding turns which give capacitance between the turns. In order to avoid a resonance in the vicinity of the tone frequency range, one attempts to keep this capacitance down by winding the transformer in sections where the primary and secondary windings are mixed. This also increases the coupling factor between the windings. Special alloys are sometimes used to keep core losses down.

The **small low frequency transformer** is used between e.g. a microphone or a pick-up with a moving coil and the amplifier input. The requirement for a wide bandwidth also applies here. It is especially important for a transformer for low level signals that it is well shielded against hum fields. Mu metal provides very efficient shielding.

The **modem transformer** provides galvanic isolation between the modem and the telecom network. It is designed to conform to the requirements imposed by the telecom authorities. Note that these requirements may vary quite significantly between different countries. In Sweden, for example, a test voltage of 2.5 kV is sufficient, while countries like Britain or Germany require a test voltage of 4 kV.

The **intermediate frequency transformer** consists of two connected resonance circuits. It is designed for a specific working frequency, e.g. 455 kHz (AM) or 10.7 MHz (FM), which can be trimmed using the coils' trim cores. For AM, SSB and CW, one usually wants the smallest bandwidth possible, i.e. the highest Q factor possible, while transformers for FM broadcasting should have approx. 250 kHz in order not to produce distortion. Hi-Fi tuners usually require a greater bandwidth since one is primarily seeking low distortion, while higher distortion can be tolerated in a car radio in order to gain higher sensitivity instead.

The **current transformer** is used for magnetic measurement of current through a conductor. This means that the current path does not need to be broken in order to perform the measurement. This transformer type is e.g. used in connection with residual current devices.

The **switch transformer** is often used in power supplies and DC/DC converters instead of conventional transformers. The frequency of a switching power supply is significantly higher than the network frequency, often as high as a couple of hundred kilohertz or even a few megahertz.

Power supplies

230 V alternating voltage is excellent for distributing electrical energy to different consumers within a building, but it must then be converted into a suitable voltage at the site of consumption. There are a number of methods for performing this conversion which are more or less suitable depending on the area of application.

The simplest power supply comprises a transformer, rectifier bridge and smoothing filter. The losses in these and charging and discharging of the capacitor(s) due to the pulsating direct current, however, give a relatively high output impedance which means that the output voltage varies sharply with the load. This type of mains part occurs almost solely in regulated adaptors for applications where a constant voltage is not required.

To avoid voltage variations with varying load, we must adjust the voltage in some way, e.g. by supplementing the simple power supply as above with a linear regulator circuit. The simplest form of **parallel regulator** consists of a zener diode. This can be supplemented with an emitter tracker to allow a higher current. However, the emitter tracker's output resistance increases the output resistance. This principle is found in regulated adaptors in a simple block unit.

It is better to build a **series regulator** with a servo loop. One can be built so that it emits very little ripple and with low output resistance. All regulated current, however, must pass through a series transistor in which high power loss develops if the current is large. This type of power supply therefore has a relatively low conversion factor. This principle is excellent for laboratory units.

In a **secondary switching** unit you chop (switch) the current. By varying the pulse width you can control the transferred energy. In this way a servo loop can regulate the output voltage. The switch technique means that the conversion factor can be very high. However, it is still quite a big unit due to the large transformer. Its size is determined not only by the transmitted power but also by $d\varphi/dt$.

By increasing the frequency you can reduce the size of the transformer. The frequency is usually 20–50 kHz, but up to 2 MHz occurs. In this case you chop the voltage on the primary side, i.e. we have a **primary switched** unit.

Switching in itself emits disturbances and it is important that the unit is provided with efficient filters on the inputs and outputs. Resonant conversion, an old principle which is increasingly being used, means that you work with nearly a sine instead of a square wave. This imposes lower requirements for filtering and shielding and such a unit often gives smaller disturbances.

DC/DC converters, as the name specifies, convert one direct voltage to another. The incoming direct voltage is chopped, transformed into another voltage and stabilised via reconnection to the chopper or with linear regulation. Such DC/DC converters are available with very small dimensions for mounting on a PCB. Some DC/DC converters have galvanically separate inputs and outputs.

Interference

There are a number of different types of electrical interference which can cause problems for sensitive electronics. The fact that disturbances come from outside –, i.e. from lightning, switching in power stations, connection and disconnection of phase compensation and switching in transformer stations – is normally quite well known. But a large number of disturbances also arise indoors. These disturbances usually come from lifts, fluorescent lamps, copiers, cooling installations, compressors, etc. A coffee percolator also emits interference on our main network. It is chiefly when switching these on and off that the interference arises. The disturbances that arise can be transients, voltage spikes, voltage variations, frequency variations and distortion.

A number of different types of protection are available to protect sensitive electronics from these different interference types.

Filter

Filters are the simplest form of protection against transients and voltage peaks. The filters have a damping effect only at frequencies above around 50 kHz. They do not protect against voltage variations. The filters normally contain varistors or "Comgap". These remove to a certain extent transients that can disturb electronics. But there is a risk that the "cut" disturbance is still sufficient to damage sensitive electronics. These filters also have an isolation voltage normally of between 600–1400 V which, if connected to the mains which has 4 kV, causes the isolation voltage to fall at exactly the outlet that it is connected to and you thereby run the risk that the disturbances will be "drawn" to exactly this outlet.

Interference suppression transformer

The interference suppression transformer suppresses interference from approx 100 Hz upwards, i.e. intermediate frequency interference, so-called ringing. The interference suppression transformer is primarily suitable for protecting computer equipment and sensitive electronics from transients, voltage spikes and earth disturbances. The interference suppression transformer is built up of a transformer with integral shields for picking up and diverting interference, these normally also have an unbroken protective earth between the input and output. There is a new "computer earth" on the output. Personal safety in the event of any isolation fault in the connected equipment is maintained through a built-in automatic fuse. Like the mains supply itself, the efficient interference suppression transformer has an isolation voltage of 4 kV.

Magnet stabiliser

The magnet stabiliser is a special type of transformer based on the ferro-resonance principle. The main task of the magnet stabiliser is to stabilise the voltage. Computers normally have built-in switching power supplies that cope with voltage variations of approx ±10–15 %. Magnet stabilisers manage to regulate the voltage within a very wide range, a stabilizer for 230 V regulates the voltage to this level from approx 135 V. In the event of overvoltage it always protects the connected load. It also has a filtering effect. The magnet stabiliser also usually has a separate "data earth" and an isolation voltage of 4 kV.

Uninterruptible power supply

A UPS is used to protect computers and other sensitive electronic equipment from sudden voltage variations, transients and power failures that can cause destructive consequences. These systems contain power converters, a battery and monitoring circuits.

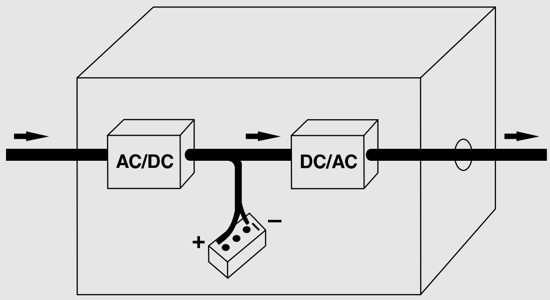
The units are usually based on two different principles:

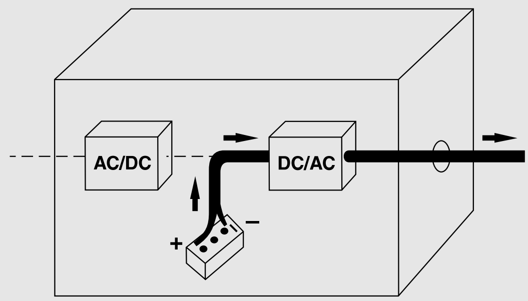
On-line system (UPS – Uninterruptable Power Supply)

The mains voltage is converted from 230 Vac to direct voltage; this is performed through a combined charger and rectifier. The converted voltage trickle charges the normally built-in, maintenance free lead-acid battery, the UPS charger also feeds the built-in DC to AC converter which in turn converts the voltage to 230 Vac. In the event of a power failure, the built-in battery provides the connected equipment with current for a number of minutes, normally 10–20 minutes. The connection time is completely uninterrupted. These types of unit also protect the connected equipment from transients, voltage spikes, voltage variations and frequency variations. To further increase operational security, the on-line systems can be provided with an internal bypass which steps in upon a major overload and in the event of a possible fault on the DC to AC converter. In bypass mode the connected equipment is supplied from the mains.

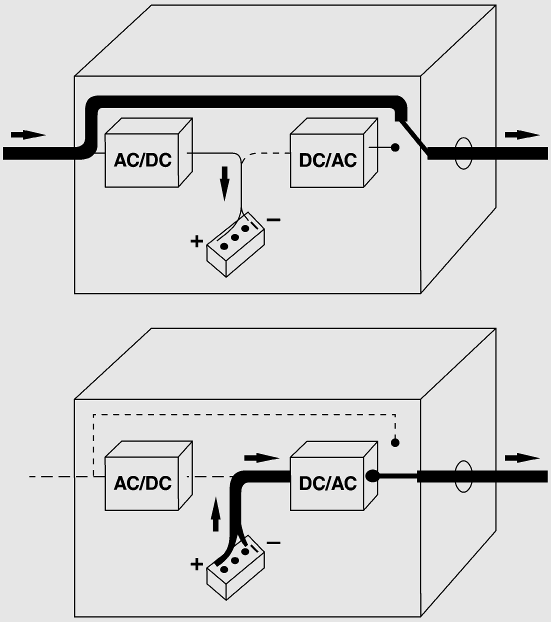
Off-line system (SPS – Standby Power Supply)

During normal operations the mains voltage passes unregulated to the output and at the same time the normally built-in maintenance-free lead-acid batteries are trickle charged. When the incoming mains voltage falls below a certain level (typically 197 V) a sensor circuit connects the DC to AC converter and the battery supply. This connection causes a voltage failure of between 2–10 msec. An off-line system has a certain transient suppressing function via an integral mains filter. The reserve time for an off-line system is normally between 10–20 minutes. The off-line systems' output voltage curve form during battery operation can be an angle wave or a sine wave.





On-line system. The battery serves as continuous power source and is trickle charged by supplying mains. At mains power loss the battery is already connected and operation continues uninterrupted.



Off-line system. The mains powers the system at the same time as the battery is trickle charged. At an interruption a qiuck connection of the battery and inverted rectifier takes place. The changeover is not instantaneous but there is a break of 2–10 ms.

Batteries

A battery is an item of equipment which converts chemical energy into electricity. Batteries are usually divided into two groups: *primary batteries* and *secondary batteries*. The designations are old and came about because in the past we often charged a secondary battery from a primary battery.

Primary batteries are used once and then discarded. The chemical reaction that creates the electrical energy in the battery is not reversible.

Secondary batteries, on the other hand, can be recharged and used again. The chemical reaction in them can be converted by applying a current instead of loading the battery with one. These batteries are used to store energy and are called rechargeable batteries.

Primary batteries

This group includes e.g. zinc-carbon batteries, alkaline batteries, magnesium batteries, mercury batteries, silver oxide batteries and lithium batteries.

The zinc-carbon battery has been our most common battery. Its plus pole comprises a carbon rod around which there is pulverised manganese dioxide (brown stone). The minus pole comprises zinc. It is designed physically as a zinc cup. There is an acid electrolyte of ammonium chloride and zinc chloride between the poles. Outside the zinc casing the batteries are usually provided with a sealed casing to prevent leakage. If the acid electrolyte comes out, it can corrode battery holders, control cards and components.

The formerly most common standard type, also called a transistor battery, is gradually being replaced by the motor battery which not only withstands higher power outlets but also has a higher capacity as purer materials are used.

A new battery produces 1.5 V but the voltage falls in pace with the capacity being extracted. The capacity is sharply limited at temperatures below 0 °C.

Alkaline batteries have an electrolyte which is alkaline, consisting of potassium hydroxide. The electrodes consist of zinc oxide as a minus pole and manganese dioxide as a plus pole, i.e. the same materials as in the zinc carbon battery. The capacity is higher than in motor batteries and the alkaline battery withstands higher power output. The differences in capacity between transistor and motor batteries and alkaline batteries are greatest with a big load. Therefore the alkaline battery is suitable for use in e.g. small tape recorders of the "freestyle" type, in flash units, etc. The alkaline battery works efficiently within the temperature range –30 °C to
+70 °C.

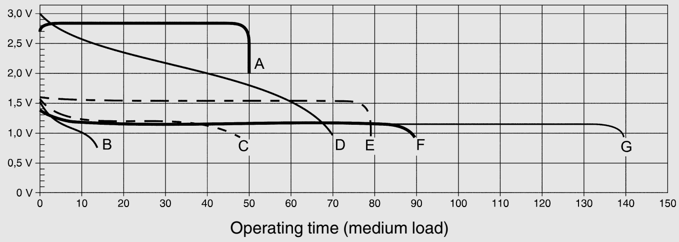
The silver oxide battery has a minus pole of zinc and a plus pole of silver oxide. The electrolyte is alkaline. The biggest advantage is that the output voltage remains relatively constant at 1.5 V and then falls abruptly. It is primarily used in cameras, calculators and clocks. Alkaline button cells are available as a cheaper alternative, but their voltage falls with capacity extraction and they therefore cannot be used in voltage sensitive equipment.

The mercury battery has a minus pole of zinc, a plus pole of mercury and an electrolyte of potassium hydroxide. It gives 1.35 V (1.4 V also occurs) during its consumption period, after which the voltage drops sharply. The area of application is the same as for silver oxide batteries.

Lithium batteries are today found in a number of different designs and with varying materials in their cathode and electrolyte. The most common areas of application are memory backup, clocks, cameras, calculators and safety apparatus which imposes strict requirements for capacity and reliability, and in equipment exposed to severe environmental stresses, due to the lithium battery's ability to work even under extreme temperature conditions.

The nominal cell voltage is 3 V, except Li/Thionyl chloride which holds 3.6 V.

Rechargeable lithium cells are also available on the market today.



Comparison between primary battery types (according to Duracell).
A – Lithium/SO_2 . B – Zink-Carbon. C – Alkaline. D – Lithium/MnO_2 .
E – Silver Oxide. F – Mercury. G – Zink/Air.

Zinc/air is the third type of environmentally friendly primary battery. It nominally gives 1.4 V. In the battery, zinc is oxidized catalytically by the acid in the surrounding air. The cell is sealed in the factory and in this condition it can be stored for up to 4 years. When the seal is broken, it can be used for three to four months, after which the content is carbonised. The output voltage stays at around 1.3 to 1.2 V throughout the discharge phase. The energy density is very high or more than twice as high as for the lithium battery.

The zinc/air cell works within –20 °C to +60 °C, but the current withdrawal declines with reduced temperature. The capacity is also affected by the relative air humidity and the carbon dioxide content in the air. Another disadvantage is that the zinc-air cell can emit a relatively limited current. This can produce function disturbances in some push-pull connected hearing aids, but in most cases the zinc-air cell can replace the mercury cell in hearing aid applications. The zinc-air cell is also suitable in e.g. pagers and telemetry equipment.

Secondary batteries

- Low weight
- Long life
- High capacity
- Easy charging
- Big current withdrawal
- Environmentally friendly
- Small temperature dependence

These are certainly acknowledged to be desirable properties for mobile equipment. We all come into contact with equipment containing some form of rechargeable batteries. Everyone increasingly wants freedom of movement without leads. But there is a constantly growing range of rechargeable batteries which all have different properties.

Here we will present the most common types of rechargeable batteries, their different properties and how they should be treated to function best to last as long as possible. We will concentrate on the three types that have achieved the greatest success with consumers: lead-acid, nickel cadmium and the new nickel metal hydride rechargeable battery.

Lead acid batteries

Secondary batteries have existed since 1860 when Raymond Gaston Planté invented the lead acid battery. Lead acid batteries represent approx 60 % of all rechargeable batteries sold. Lead-acid batteries are usually the most economic alternative as the cost per Ah, especially in larger sizes, is clearly the lowest for this type of rechargeable battery. An excellent feature for this rechargeable battery type is that it withstands tough requirements for treatment both in purely physical terms and with regard to charging and discharging. The lead-acid design is excellent as a starter battery and a backup power battery. Unfortunately the material in the electrodes is lead, which in itself entails advantages for charging and discharging, but also means that it is heavy.

The market was previously dominated by the open lead-acid batteries, but the most commonly occurring type of lead-acid batteries today are the valve controlled or sealed type, primarily in the industrial market sector. In the text below we have thus chosen to concentrate on this type of lead-acid batteries.

In this context it should also be mentioned that there are different types of valve controlled lead-acid batteries. For example, there are special types of lead-acid batteries where the electrodes are spirally wound with a thin separator between them in a cylindrical encapsulation. These types have a very low internal resistance which allows a very large current withdrawal over a short period.

Charging

The lead-acid battery is charged with a constant voltage. Electrodes in lead and a sulphuric acid electrolyte generate a cell voltage of 2 V. These cells are usually built together into a pack containing 3 or 6 cells. If the rechargeable battery is used in cyclical operation, i.e. it is charged and then used, in order to then be charged again, the charge voltage will be 2.40–2.50 V/cell, i.e. 14.4–15.0 V for a 12 V rechargeable battery. Many people use lead-acid batteries for maintenance operations or constant charging, as the rechargeable battery is not normally used but is constantly charged in order to be fully charged when required, e.g. in UPSs or for alarm installations. The charge voltage will then be 2.25–2.30 V/cell or 13.5–13.8 V for a 12 V rechargeable battery. The charger should also be adjusted so that the charge current is normally approx 10 % of the rechargeable battery's capacity, 0.1 C, or approx 5 % for constant charging. The maximum charge current should never exceed one third of the rechargeable battery's capacity.

When the battery is charged, the pole voltage will increase markedly and this rise is so marked that it can be used as a measurement of the charge status. Therefore a mains supply with a constant (stabilised) output voltage can also be used as an automatic charger. The required peak voltage for the battery is adjusted on the power supply without a battery. When the battery reaches this voltage, the charge current will fall to a value which only compensates for the self-discharging of the battery. If charging continues even though the battery is fully charged, the current will only be used to form oxyhydrogen gas by the water in the electrolyte. The cell voltage is then 2.4 V. A regulated power supply has low ripple on the output voltage which is an advantage if the battery will be charged while it is connected to sensitive apparatus. A fuse should be incorporated in the charge to prevent uncontrollable large currents from the battery if a short-circuit should arise.

Discharging

With regard to discharging, this design undoubtedly has its biggest advantages for discharging with high currents over a short period. A valve controlled lead-acid battery can normally be loaded in the short-term (< 5 sec) with a current corresponding to 15 times the rechargeable battery's capacity. The maximum continuous current withdrawal should not exceed 3 times the capacity.

Life

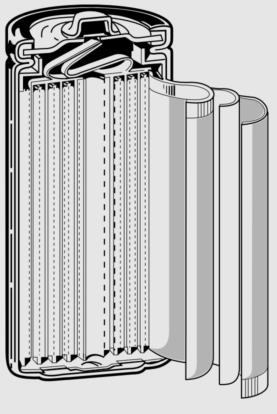
The most common lead-acid battery has a normal life of 3–5 years, but there are types with an estimated life of over 10 years. These are chiefly used as UPS units within telecommunications and for alarm purposes. A better measurement of life is often to specify how many cycles a rechargeable battery copes with before the capacity falls to 60 % of the original value. This figure is significantly affected by how much capacity is used for every discharge (depth of the discharge). A standard value is approx 500 cycles when using 50 % of the capacity for every discharge.

Summary regarding lead-acid batteries

Weight is a clear minus for the lead design. **Life** varies greatly between different types but is very good when compared with other rechargeable battery types. **Capacity** is often set in relation to weight, which does not give the lead design any advantage. However, it is easier and cheaper to manufacture lead-acid batteries in greater capacities than other battery types. **Charging** is a clear plus for these rechargeable batteries as it is very easy and does not require any special monitoring circuits. Unfortunately, with the best will in the world the lead-acid battery cannot be called environmentally friendly as it contains considerable quantities of the **environmentally hazardous** metal, lead. This rechargeable battery type is not especially **temperature dependent** during discharging (the capacity can be affected negatively at low temperatures), but charging should be performed at room temperature, otherwise the charge voltage must be adjusted for it to become fully charged.

Nickel cadmium rechargeable batteries

The first alkaline rechargeable battery, NiFe (Nickel ferrous rechargeable battery), was invented in 1899 by a Swede called Jungner. It was not until 1932 that the alkaline rechargeable battery was given electrodes of nickel and cadmium, and it was first in the 1960s that it really achieved wide commercial use. Today the NiCd rechargeable battery is the most common rechargeable battery in small consumer applications



Cross-section of a NiCd cell.

A major factor in the success of the wave of wireless equipment sweeping over us now lies in the improvements in performance and price which the NiCd rechargeable battery has undergone over recent years.

Nickel cadmium batteries offer a high energy density (high energy content in relation to weight), great possibilities for high current loads and a long life in number of cycles. NiCd batteries are normally used most in sizes from a few mAh to approx 10 Ah. These rechargeable batteries were formerly only available in one design which had to cover the whole area of application, but they are now available in a number of specialised designs to create the best possible performance for the intended types of use. For example, some will have the highest possible capacity, others can be charged very quickly or can work under high ambient temperatures.

The cells are designed with one negative cadmium electrode and one positive nickel electrode. The electrolyte, consisting of potassium hydroxide in water, acts as an ion conductor. To prevent short-circuiting, the two electrodes are kept electrically insulated from each other using a porous separator, usually consisting of plastic materials. In the cylindrical cells, the electrodes are spirally wound to create a large a surface as possible (high capacity) with as thin a separator as possible (low internal resistance = high discharge current). The electrochemistry in the rechargeable battery is designed so that all gases that form during charging (oxygen is formed through electrolysis of the water) are recombined and removed from the gaseous phase. Naturally, all cells are provided with a safety valve that prevents overpressure from forming in the cell during very large overcharging.

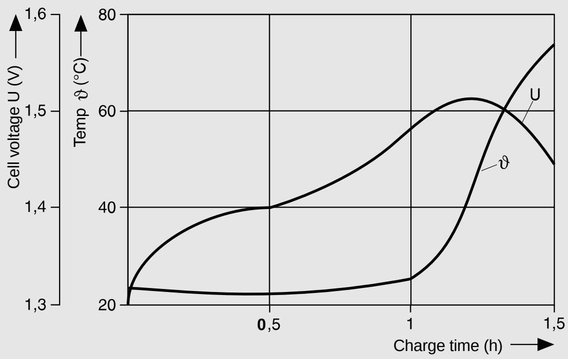
Charging

The nickel cadmium battery is charged with a constant current. Electrodes in nickel and cadmium and an electrolyte of potassium hydroxide give a cell voltage of 1.2 V. You have to supply more energy during charging than you withdraw during the discharge phase. You can usually reckon that you must supply 140 % of withdrawn capacity, i.e. you obtain a charge factor of 1.4. NiCd batteries are normally charged with 0.1 C over 14–16 h. Charging can be determined using the following formula:

$I = C \times 1.4 / t$

I = Charge current in A
C = Capacity in Ah
1.4 = Charge factor
t = Charge time in hours

The cell voltage will increase constantly during the charging phase to reach approx 1.45–1.5 V at the end of the charging cycle. For charge currents under 0.2 C the charge does not need to be monitored. Charging should be performed at room temperature. Check the polarity on connection. A NiCd-battery will be destroyed by incorrect polarity on the charger.



Fast charging of a NiCd cell. The curves show cell voltage and cell temperature when charging a fast charging NiCd battery with 1.0 C over a period of approx 90 min. The curves show that the rechargeable battery is fully charged when just over 70 minutes have passed and the rechargeable battery has reached approx 45°C external temperature.

Fast charging (0.5–1.5C)

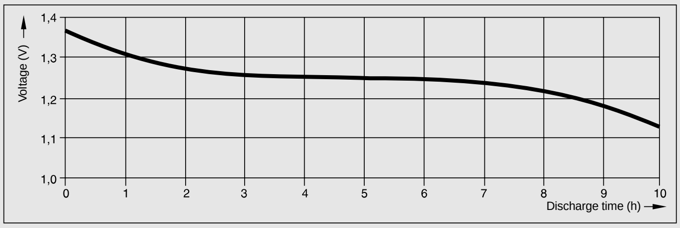
The nickel cadmium battery has very good properties with regard to its ability to take a high charge over a limited period. The shorter the time you want to charge for, the more careful you need to be in monitoring the charge. The cell voltage in the NiCd gradually increases during charging and then finally decreases slightly when the cell is fully charged. At the same time the cell temperature increases sharply. Modern fast chargers use the $-\Delta V$ method, i.e. they sense this voltage reduction and then cancel the charge (see figure). You should try in the first instance to avoid the cell temperature rising too much as this gives the cell a significantly shorter life. It is therefore recommended that a (auto-reset) thermal fuse is used as extra security. The surface temperature for a fully charged cell that has been fast charged is approx 45 °C. The thermal fuse is connected in series with the charge and on the outside of the battery pack. This breaks the charge when the temperature exceeds 45 °C. Fast charging <1 C can also be performed only with timer monitoring according to the formula stated above, but this should also be combined with a thermal fuse to avoid overheating of the cells. Fast charging is not suitable for high temperature cells or button cells.

Trickle charging

This charging method is most common for high temperature cells and button cells. This means that the rechargeable battery is constantly being charged so that it is then available in the event of a loss of voltage, e.g. as backup in a computer. For cylindrical NiCd cells, these should be trickle charged with 0.03–0.05 C while button cells should be trickle charged with 0.01 C. A cylindrical cell of 800 mAh should thus be trickle charged with 24–40 mA.

Discharging

The NiCd cell has unique load properties. You can charge a cell right up to 100 C, but only over a very short period. The maximum current withdrawal should constantly not exceed 8–10 C for approx 4–5 min. The NiCd cell also has the advantage that the cell voltage is very constant (1.2 V) throughout the discharge sequence. The final voltage (the pole voltage when the cell no longer has any capacity left to supply) is generally defined as approx 1.0 V. Unfortunately the NiCd batteries have the disadvantage that self-discharge is quite high, approx 1 % per day, which gives poor conversion efficiency during trickle charging which must then compensate for this fact.



The discharge voltage's time sequence at a load of 0.1 C.

Life

A number which most people have heard is that NiCd batteries manage 1000 charges and discharges. But you should also observe that the number of cycles that a rechargeable battery of this type manages depends greatly on how it is treated. As mentioned previously, in the event of overcharging the internal temperature ion the cell increases, which accelerates degeneration of the materials in the cell. The same thing happens for strong discharging. When a battery pack consisting of several cells is discharged, there are differences in the remaining capacity, so that some cells reach the final voltage before others. These will then become over-discharged and will have a negative effect on the life for the whole pack. In the event of major over-discharging, when the cell voltage falls to 0.2 V, the cell can reverse its polarity. NiCd cells always work best when fully cycled, i.e. they are discharged to 1.0 V before commencing charging. In this way you avoid differences in residual capacity and achieve the best total function for the battery pack.

Summary regarding nickel cadmium batteries

Weight is a clear advantage with NiCd, especially as this is small in relation to **capacity. The life** or chiefly cyclability is very good for this rechargeable battery type. **The charge** requires accuracy if you want to fast charge with a very high current without reducing the rechargeable battery's life, but is otherwise quite easy to manage. The NiCd cell is **temperature dependent** as the internal resistance increases with falling temperature. Today, types of NiCd batteries are available which are specially designed to function under a high ambient temperature. They are used e.g. in emergency light fittings.

The NiCd battery contains **the very environmentally hazardous material cadmium**, which must be limited in the environment. At present there is no satisfactory alternative to this rechargeable battery type. We must ensure that all NiCd batteries are returned to suppliers who provide equipment which includes such rechargeable batteries.

Nickel metal hydride batteries

NiMH batteries have been available since the middle of the 1970s. It is only now when public opinion has started to demand a more environmentally friendly alternative to NiCd that the major manufacturers have chosen to develop the system for the consumer stage. The environmental debate that has flared up recently has often discussed the environmentally hazardous NiCd batteries and their possible successors, Nickel metal hydride batteries. In fact, these rechargeable batteries have some advantages in comparison with NiCd, but naturally also disadvantages. In much of today's normal consumer equipment it will be possible to replace the environmentally hazardous NiCd batteries, while many other major areas of application with special requirements for rechargeable batteries will have to wait a while longer.

In this presentation we will mainly compare the system with NiCd to demonstrate both similarities and differences between the systems and to explain more clearly how these new rechargeable batteries should be treated to provide optimum energy for as long as possible.

NiMH is the secondary battery system which has the highest energy density of existing systems on the market. It is also the system which has the highest capacity in relation to the rechargeable battery's volume. This is the absolute biggest advantage with NiMH compared with NiCd. NiMH will be used in all normal areas of application, e.g. mobile phones, freestyles, video camera equipment, etc, where you can benefit from the higher capacity in the form of longer operating time. However, the price is also still significantly higher. In future, the price will fall when the materials are cheaper for this design.

The battery system is based on storing hydrogen in a metal alloy (the battery system seen previously called a nickel-hydrogen battery). A sintered nickel plate forms the positive electrode and the negative electrode comprises a special alloy consisting of precious earth metals, nickel, manganese, magnesium, aluminium and cobalt. None of today's manufacturers will specify the percentage for a respective ingredient as this alloy determines the rechargeable battery's properties. The separator consists of polyamide or polyolefine and the electrolyte is alkaline. When charging and discharging, you move hydrogen between the different electrodes. The ability to bind hydrogen in the metal alloy determines the rechargeable battery's capacity. The biggest problem being tackled now is that the higher the capacity you achieve the less willing the system is to react, which then limits the discharge current and the charge time. Like NiCd, the NiMH battery is provided with a safety valve which prevents too great an overpressure from forming in the cell.

Charging

NiMH has a higher capacity/volume than NiCd. This means that more active material is compressed into the same casing. This in turn has the result that the materials have a smaller volume in the casing to expand in, which makes the system less willing to react. Consequently, NiMH must be charged more slowly than NiCd and must also be monitored more carefully in order to avoid overcharging. Both rechargeable battery systems have a cell voltage of 1.2 V. Normal charging is performed in the same way, i.e. with a charge current of 0.1 C for 14–16 hours. This means that the charge factor is the same for both systems, 1.4. In the same way, the cell voltage will also increase in amount at the end of the charging phase to 1.45–1.5 V. No charge control apart from a timer is necessary as the charge current is <0.2 C.

Fast charging

If a NiCd battery can be fast charged in approx 15 min, the corresponding minimum charge time for NiMH is approx 1 hour. The temperature increase that occurs in the cell when it approaches a fully charged state occurs much more quickly for NiMH. The voltage reduction that occurs simultaneously is also significantly smaller, so that the accuracy for the circuits that will sense it must be higher. It is recommended that when you fast charge NiMH batteries you should always use at least two of the available safety systems ($-\Delta V$, surface temperature >45 °C, timer). In this context it should also be mentioned that the NiMH battery's life is more sensitive to overheating of the cell than NiCd. However, it has not been possible to establish any "memory effect" for the NiMH cells. This is a phenomenon which sometimes occurs for NiCd cells where only a small part of the cell's capacity is used. When this occurs repeatedly, the cell "learns" so that the available capacity after charging decreases. The phenomenon can usually be rectified by cycling the battery fully a couple of times, i.e. by discharging it fully before charging and repeating this 3–4 times.

Trickle charging

This type of charging can only be recommended for NiMH batteries of the button cell type. For cylindrical rechargeable batteries, a constant charge current always entails a reduced life. For button cell variants, however, there are no differences here from NiCd.

Discharging

As mentioned above, the active materials in the NiMH cell have less space to expand in inside the cell. This reduces the tendency to react. It is thus natural that the maximum discharge current is also lower than for the NiCd cell. Normally, discharge currents over 3–5 C cannot be recommended. However, there is absolutely no difference between the systems' final voltage, approx 1.0 V. Unfortunately, however, NiMH has a higher self-discharge, approx 1.5 % per day compared with 1.0 % per day for NiCd. This means that the storage time for a fully charged rechargeable battery that you want to have available for fast use will be shorter than for a corresponding rechargeable battery of the NiCd type.

Life

Because NiMH is a relatively new rechargeable battery system, we have no practical long-term studies to establish its life. According to the information from manufacturers marketing their rechargeable batteries in Sweden, the life should not be shorter than for NiCd, i.e. approx 1000 cycles. You should note that this figure applies for ideal conditions, e.g. charging with 0.1 C over 14 hours at room temperature for every charge. It also does not take into account any overcharging that can arise and have a negative effect on its life. A realistic cycling length under normal circumstances should be approx 500–800 cycles.

Summary regarding nickel metal hydride batteries

NiMH is the only rechargeable battery type that does not contain environmentally hazardous heavy metals and is therefore **more environmentally friendly** than the other rechargeable batteries. **Weight**, in relation to capacity, is the second advantage with this rechargeable battery type. It is the secondary rechargeable battery system with the highest **energy density. Life** is good with regard to cyclical use, but less good with regard to maintenance use. This does not apply, however, for button cell types, which have the same properties as corresponding NiCd types. **The charge** is more sensitive and must be monitored more carefully than for the other rechargeable battery types discussed here. Like NiCd, the NiMH cell is also **temperature dependent** and the recommended working temperature should be complied with.

Throughout this compilation we have adhered to general values for the different rechargeable battery systems. Because such big differences prevail between different types of rechargeable batteries, we recommend that you always check information on charging and discharging with a respective manufacturer or representative before use.

Solar cells

Solar cells use light to produce an electric current. The solar cell can be manufactured from many different materials, but silicon is commonly used.

Here we are talking of single (monocrystalline) or multi (polycrystalline) crystal cells and thin film (amorphous). The difference between single and multi crystal cells is not so great, they are actually only different ways of manufacturing the basic material for the cell. Thanks to more homogenous materials, the single crystal cell has a slightly higher conversion factor, i.e. it converts more energy per surface unit than the polycrystalline cell. However, the difference is quite small, 12–15 % for a single crystal and 10–14 % for a multi-crystal.

A normal solar cell of crystalline silicon is usually approx 10 × 10 cm and has a nominal voltage of approx 0.5 V. By connecting solar cells in series, you obtain solar cell panels. There are panels with a different number of cells depending on the area of application and the quality of the individual cell. A solar cell panel that will be used to charge a lead-acid battery at our latitudes needs to have at least 30 cells if it is of the single crystal type and 32 if the cells are of a multi-crystal type. With rising temperature, the voltage from the cell falls, which means that you may need a panel with even more cells if it is very hot where it will be installed.

A normal panel with 30–32 cells usually has a peak output of 40–45 W. You achieve other sizes by either adding more cells or by dividing the cells into smaller parts. However, this is quite expensive because it requires further steps in the manufacturing.

Thin film technology, however, offers several advantages from a manufacturing viewpoint as you can determine the characteristics exactly by positioning the cable pattern in a special way. A thin film panel is manufactured by adding a thin layer of active material directly on a treated sheet of glass. Using a laser you can then cut cells into the sizes and quantity you want. Unfortunately the conversion factor for this type of cell is significantly lower than for cells of a crystalline type, but for small applications like e.g. compact calculators, this type of cell has become extremely common. A standard panel in thin film for charging a battery normally has an output of approx 10 W.

Solar panels are normally used to charge batteries or to operate some type of consumer product directly, such as e.g. a water pump, fan, etc. A battery charging system is built up of one or more panels, a charge controller to give the battery a maximum charge and protection against overloads and damaging deep discharges, and a battery. The battery can be of different types, there is no special "solar battery". A normal car battery is not suitable because it is designed to give a lot of energy over a limited period and not to give less energy over a longer period, which is the normal operating condition in solar installations.

The solar panel will be mounted so that it is exposed to as much light as possible. The output power is directly proportional to how much light radiates in. You should choose a site which lies between SE and SW which is as shade-free as possible. Crystalline panels are especially sensitive to shading, and even if only one cell in the panel is in shade, you lose most of the energy. Diffuse shade is not as hazardous as more distinct shade. The angle to the sun is also important, during the winter months perpendicular mounting is preferable, while during the summer months an angle of 30–45° is most appropriate. The solar panel produces energy even if the sun is not shining, but naturally the energy produced is dependent on the radiated light. On a sunny summer day, the radiation in Sweden can be up to 1000 W/m^2 and then a correctly mounted panel charges a maximum of 3 A if the battery was not fully charged before. On a cloudy summer day the radiation may perhaps be 200 W/m^2 and then the current is not greater than approx 0.5 A.

Personal computers

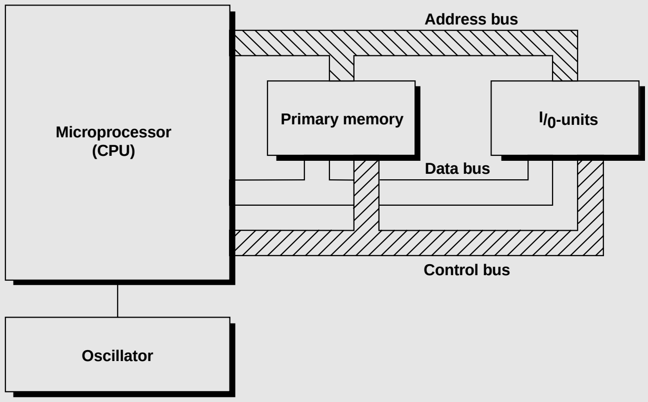
History

The PC, as we know it today, was born with the introduction of IBM's Personal Computer in 1981. A press release dated August 13 of that year informs us that ''IBM introduces the IBM 5150 PC Personal Computer in New York''. It continues: ''The PC has a 4.77 MHz Intel 8088 CPU with a 64 kB RAM which can be extended to 256 kB, a 40 kB ROM (BIOS), a 5.25" disk drive with a capacity of 160 kB, and comes with PC-DOS 1.0 from Microsoft. A complete computer with colour graphics costs US$ 6000''.

Several versions of ''microcomputers'' had already appeared. In January of the following year, 1982, Gregg Williams wrote in BYTE magazine: ''Which computer has colour graphics like Apple II and an 80-character screen like TRS-80 Model II, a redefinable character map like the Atari 800, a 16-bit microprocessor like Texas Instruments' TI 99/4 and a complete keyboard with both lower and upper case letters? The answer is IBM's PC.''

Many have tried to compare the development that the PC has experienced over the last twenty years to that of cars, boats or even the sizes of parking lots. Well, if parking lots had experienced the kind of development rate that computers have, the entire parking lots of cities like Stockholm, Gothenburg and Malmö would fit into a regular size living room. Intel's IA-64 Itanium, for example, uses 42 million transistors while the 8088 had 25000. The 8088 was able to address 1 MByte of data storage. A Pentium II addresses 64 GByte, or c. 64000 times as much.

CPU, the heart of the computer



Vital parts of a microcomputer.

The microprocessor's Central Processing Unit (CPU) is where the processing of data takes place through simple binary, logic operations executed by the CPU. The processor contains a number of storage areas called registers. The oscillator, or ''clock'' as it is usually referred to in everyday language, acts as a ''pacer'' for most actions in the PC. The oscillator gives the system clock signals, e.g. 300 MHz, 1.8 GHz or more, which control the work in the processor and the computer's various buses.

RISC and CISC

A definition of the terms RISC and CISC might be in order. To somewhat balance the kind of marketing where several companies have tried to outdo eachother when it comes to RISC (Reduced Instruction Set Computer) performance, an account from a historical point of view of what RISC architecture tried to achieve is provided below.

Early computers had ''accumulator architecture'' that was based on executing **operations** on **data** in an **accumulator** which were **stored** in **registers** that had, in turn, been **loaded** with data from the computer's **memory**.

Memory to memory architecture followed, providing the computer with an opportunity to work with registers that could contain both addresses and data, general registers. This made it possible to allow data to directly influence the execution of programmes.

Stack architecture offered an easy way to handle complete sets of registers, to store things like machine status and to switch tasks. But memory to memory architecture was still relatively slow. Loading registers from memory was a time-consuming operation. The development towards an increased number of internal registers was only natural, since the internal operations in the CPU were fast. The number of operations directly in memory were now reduced and the work with and between registers was optimised. Berkeley gave this architecture the name RISC. CISC (Complex Instruction Set Computer) was born simultaneously in order to provide a name for the older, ”traditional” type of computers.

RISC The fact that all instructions are contained in single words is often used to characterise RISC. These instructions are of equal length and use only a single CPU cycle to execute them. The computer uses an instruction pipeline, executing instructions with or between registers. The only memory operations are Load and Store and the RISC computer uses no microcode. It also has a great number of registers, usually 64 or more.

CISC can, as a design philosophy, be said to represent the ultimate goal that each instruction in a high level language (C++, Perl, Basic) should be able to be represented by a simple machine instruction. The drawback is a higher complexity, with larger chips that make it difficult to increase clock speed and performance.

A comparison of some well-known processors’ architecture and word length.

Architecture	4 bits	8 bits	16 bits	32 bits	64 bits
CISC			Z8000	Pentium-	
CISC	4040	Z-80	8086	Pro	PowerPC
CISC		8080	68000	Z80000	
CISC	4004	8051	TMS9900	Pentium	
CISC		2650		80486	
CISC	Am2901	6800	80C166	68020	IA-64
CISC		650x		68040	Itanium
CISC		F8		29000	
CISC		1808		SH	Alpha
RISC				ARM	

A brief overview of the most common microprocessors on motherboards, from the PC’s infancy in 1981 up to present times.

Processor spec.	Manu-fact-urer	Processor speed MHz	Bus MHz	Notes
8088	Intel	4.77	4.77	IBM’s first PC
V20	NEC	8 – 10	8 – 10	
8086	Intel	8	8	
V30	NEC	10 – 16	10 – 16	
80286	Intel	6 – 12	6 – 12	
386DX	Intel	16 – 33	16 – 33	
486DX	Intel	25 – 50	25 – 50	
486DX2	Intel	50 – 66	25 – 33	
486DX4 100	Intel	100	33	
486DX4 100	AMD	100	33	
586P75	AMD	133	33	
Pentium	Intel	60 – 150	60	
Pentium	Intel	66 – 200	66	
K5 Pxx	AMD	75/920/100/120	60	
686 Pxxx	Cyrix	120/133/150	66	
Pentium MMX	Intel	200/233/266	66	
Pentium Pro	Intel	150 – 200	66	
K6	AMD	166 – 300	66	
M II Pxxx	Cyrix	166/188	75	
M II	Cyrix	300 – 433	75	
Pentium II	Intel	233 – 450	66 – 100	Slot 1
Celeron	Intel	266 – 733	66	Slot 1/Socket 370/ FCPGA
K6-2	AMD	266 – 550	100	
K6-3	AMD	400/450	100	
Pentium III	Intel	450 – 1 GHz	100 – 133	Slot 1/Socket 370/ FCPGA
K7 Athlon	AMD	500 – 1.2 GHz	200	Slot A/Socket A
K7 Duron	AMD	600 – 800	200	Slot A/Socket A
K7 Thunderbird	AMD	800 – 1.1 GHz	200	Socket A
IA64 Itanium	Intel	1.5 GHz	400	

The bus structure, the highways of the computer

A **bus** is a system of wires that has a special function in the computer. They are divided into data buses, address buses and control buses.

The **data bus** handles the transfer of data between different units in the computer. A data bus can have different widths (number of parallel conductors): 8, 16, 32, 64 bits, etc. The wider the data bus, the more data that can be transmitted simultaneously on it. A wider bus provides a higher bus bandwidth which, in general, provides a faster computer. Generally, it is the CPU that determines when and how a transfer takes place on the data bus. Other units may also act as controllers, so-called **Bus Masters**. This feature is utilised in e.g. **DMA** (Direct Memory Access), transmission in which usually the system memory and a peripheral unit (like a hard disk that acts as a bus master) exchange information without data being guided or controlled in minute detail by the CPU.

The **address bus** carries information about the source and destination of a data transfer. The more wires (lines) there are in the address bus, the more addresses are available. For example, a processor with 32 lines can address 4 Gigabytes (2^{32} = 4 GB).

$1Ki=2^{10} = 1024$
$1Mi = 2^{10}\ Ki = 1048576$
$1\ Gi = 2^{10}\ Mi = 1073741824$

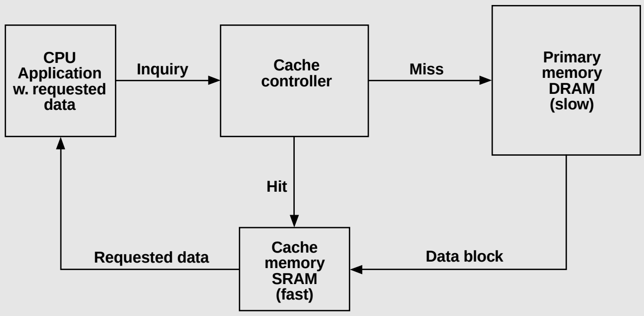
The **control bus** contains a number of lines that carry control signals. It determines, for example, the exact time lapses for writing to memory. This is usually called ”timing”. Communication between the computer’s units as well as error signals are also managed by signals on the control bus.

Primary memory

The computer has different memory types. The main memory, out of which programmes are run, and the secondary memory, which can store information even when the computer is turned off.

Main memory is the working memory in the computer required by the processor. The computer’s operating system as well as all programmes run by the computer are stored in this working memory (often some type of **DRAM**). In a PC, the operating system is stored on disk, in secondary memory (compare **DOS**, Disk Operating System), but is read into the working system when the computer is started or booted (”boot” is taken from the expression ”pull oneself up by the bootstraps”, compare **boot** and **BIOS**). The speed of the working memory is one of the great bottlenecks in a modern PC and an issue worth considering when judging the performance of a computer.

Cache memory means that a small section of high-speed memory, often **SRAM**, is used as an extra, small storage device next to the processor. When reading is performed, requested data is often already stored in this memory, which means that acquisition can be executed extremely fast. Cache memory is usually divided into L1 (Level 1) or internal cache, and L2 (Level 2) or external cache. An AMD K7, for example, has 128 kB L1 and up to 8 MB L2 cache.



Cache memory. A copy of the most frequently used information is stored in the cache memory. The processor generally writes to / reads from the fast cache memory and need not wait for the slower system memory.

Secondary memory

Secondary memory differs from the main memory and stores programmes and data that are not actively used at the time. We recognise secondary memories like hard disk drives and diskettes or floppy units. Memories which store a great amount of information that is retained even when the computer is turned off. The rapid development of the capacity of hard disk drives means that we have hard disk drives with up to 100 GBytes as a standard today.

Storage techniques for mass storage

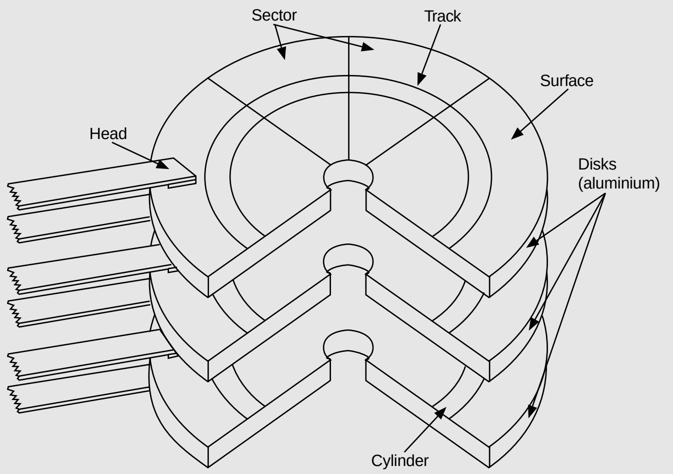
Information is stored magnetically on tapes, diskettes or hard disk drives and optically on CD-ROM or DVD.There are also hybrid techniques although they are not as dominant. Electrical mass storage occurs in various types of Flash memories, e.g. CF (Compact Flash), MMC (Multi Media Card), Smart Media or Sony Memory Stick. Flash with AND technique can reach storage capacities of several hundred MBytes.

Diskettes, or **”floppies”**, are thin, flexible plastic disks. The plastic bedding is coated with a binder mixed with a magnetic material and packaged in a case. Data is transferred from, or to, the disk with a read/write head that can be moved between the different tracks on the disk. The diskette has, ordinarily, a capacity of 1.44 MByte and today it is often replaced with a CD-ROM to transport software.

Hard disks (disk memories or fixed disks) are the most common type of secondary memory. These days, a hard disk is able to hold up to 100 Gbytes. A hard disk usually consists of one or several aluminium platters. These are precision polished and then coated with a very thin layer of magnetic material. The hard disk is able to pack information considerably more densely than a diskette.The disk rotates at about 10 000 rpm or more. One or several read heads float on an air cushion closely above the magnetic coating. With a sophisticated aerodynamic design, the read head can ”fly” very low (a few hundred nm, which is less than one hundredth of a human hair!) and the magnetic track width can be made extremely narrow.

Common standards for hard disk controllers are ST-506 (very old), ESDI (old), IDE and SCSI. IDE (Integrated Drive Electronics) is a standard introduced in 1986-87 by Compaq and Western Digital.

IDE is characterised by the fact that nearly the entire controller of the hard disk is located on the chassis of the disk and not on a separate control board. This gives the manufacturers ample opportunity to design control boards and disk systems relatively freely within this standard. Today,this technology has been developed further and is called Enhanced IDE (E-IDE) or Advanced Technology Attachment (ATA). E-IDE is an excellent alternative to SCSI for computers in the lower price ranges.



The construction of a hard disk.

SCSI (Small Computer System Interface) has a long tradition within the computer industry. It has long been used for disk systems to Macintosh computers and is nowadays also common in the rest of the PC world, acting as a standard for server systems. SCSI is a general standard suitable for all kinds of peripheral units. It is frequently employed for connection of back-up units (tape stations) and different CD-ROM and DVD units.

The SCSI standard comes in different versions. These differ in e.g. bus width, signal specifications, transfer type and bus speed as well as the possible intelligence of the different units. At present, the development of this standard has provided us with seven generations of SCSI and 640 MBytes/second has been defined as a standard for systems that require additional bandwidth. SCSI started out as a narrow bus with 50 connections that could transfer 1 byte of data per time unit. The more recent SCSI-2 and SCSI-3 are able to transfer 2 or 3 bytes simultaneously.

Summary, SCSI-standards.

Generic name	SCSI-1	SCSI-2	SCSI-3	Fast-20 Ultra SCSI SE	Ultra2 SCSI	Ultra3 or Ultra160 SCSI	Ultra320 SCSI
Bus width	8-bit data bus		Wide 16- and 8-bit data bus			Wide 16-bit databus	
Bus speed Mbyte/s	5	10	20	40	80	160	320
Signal standard		Single ended and HVD (High Voltage Differential) SCSI			LVD (Low Voltage Differential) SCSI		

Band/Tape devices for mass storage of data are used today for backup of computers.There are many competing standards. The most common are QIC, DAT and DLT. QIC (Quarter-Inch Cartridge) is available with various storage capacities ranging from under 100 MB to over 10 GB. QIC is normally used for backup of individual computers in the same manner as for more expensive and robust 4 and 8 mm DAT-based units. For network systems and data storage requiring transmission speeds of more than 2 MB per second, DLT (Digital Linear Tape) is used, which can store 100 GB and reach transmission speeds of over 6 MB per second. DLT has the advantage of direct reading after writing. A technique that enables direct data control when writing data to the tape. In this way, defects in the data tape can be handled in order to improve the safety.

CD-ROM/CD-R/CD-RW/DVD. You can normally store 650 or 700 MB on CDs. On a CD-ROM, a pattern of small depressions is etched onto the underside and then read by a laser. CD-ROMs are inexpensive to produce and are used as distribution media for both software and information. CD-ROM units come in various types. Transmission speed varies from the original 150 kB/s 1X (single speed) to 40X or more. CD-R and CD-RW offer the additional possibility of writing data as well. CD-R (Recordable) can be written on once whilte CD-RW (ReWritable) can be written over many times. The transmission speed for writing may be over 2 MB per second. DVD (claimed by many to stand for Digital Video Disc or Digital Versatile Disc, although neither explanation is official) is a 6 inch disk with a capacity of 4.7/8.5/9.4 or 17 GB of data. DVD normally uses MPEG-2 video compression and Dolby Digital or DTS (Digital Surround) for storing sound, but is also used for pure data storage of large amounts of data, like for example in dictionaries.

ROM-BIOS/FLASH-BIOS

ROM stands for Read Only Memory. Several important functions are located in a PC's BIOS. It contains the programme that makes it possible to start the computer. This so-called boot programme contains e.g. a small test programme that quickly checks the PC's different parts during start-up and then reads an operating system from a disk to the main memory. The BIOS is often stored in a Flash memory and can then be updated, e.g. in order to allow new types of peripheral units to be used.

Flash-ROM

Flash ROM is a type of electrically erasable memory. Like other types of ROM, it retains information after power has been switched off. The contents of a Flash memory can easily be changed with special software, and has been put to great use as a programme memory in many different types of peripheral units like CD units, laser printers, storage media for digital cameras etc. as well as a BIOS memory in PCs.

In- and out-devices

A computer would be totally unusable if you could not communicate with it. In-devices are used for feeding information and thereby controlling the computer. Out-devices make it possible to see or store the result of the computer's processing of information. In- and out-devices are connected via an interface (adaptor board) to the expansion bus. Secondary memories can also be viewed as in-out-devices.

Expansion buses

PCs have one or several expansion buses, board slots, where extra equipment like graphic and network boards can be inserted. There are several different standards for expansion buses today:

The **ISA bus**, which used to be the most common one but has now been replaced by PCI, was a further development of IBM's first PC bus from around 1981. ISA stands for Industry Standard Architecture and was originally called AT bus after IBM AT in which it was first used in 1984. Over time, the expansion bus has become more and more of a limitation on performance and new buses have therefore been developed like e.g. PCI and AGP.

MCA (Micro Channel Architecture) is one of IBM's expansion buses launched in 1987.

EISA (Extended Industry Standard Architecture) was an enhanced version of ISA, launched in 1989.

Nubus is the name of Apple's old 32-bit bus.

VL bus is a predecessor to the PCI bus. Used to be called ''Local Bus''.

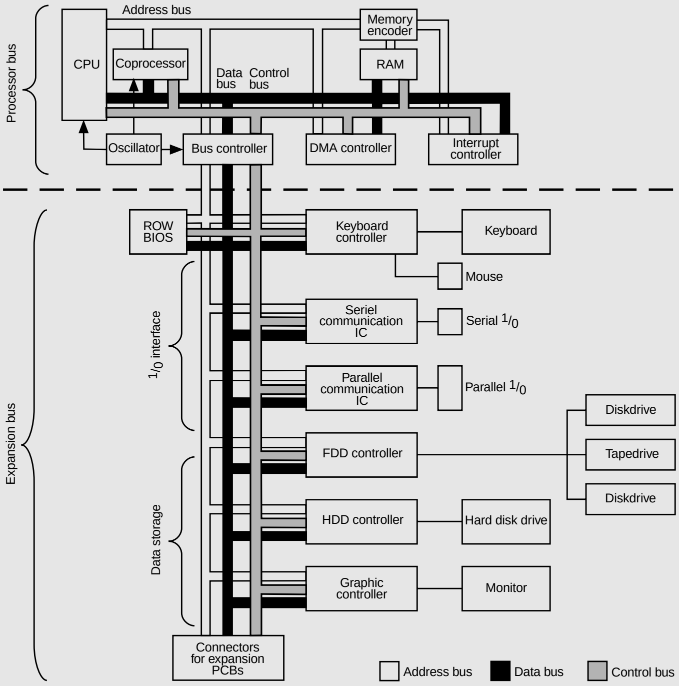
PCI (Peripheral Component Interconnect) is Intel's local bus which can manage up to 264 MB/s and can also be used in a 3.3 V system. It is currently a standard feature on all PCs.

AGP (Accelerated Graphics Port) is a screen interface for direct memory access that can easily handle an image resolution of e.g. 1024×768 pixels with 30 images per second. It was introduced in 1996 by Intel and is used with a separate card connector for the graphic card. 1*AGP (AGP) transfers 264 MBytes/s. 2*AGP (AGP 2x) transfers data on both edges of the clock signal, transmitting 528 MBytes/s. 4*AGP (AGP 4x) is able to transfer 1017 MBytes/s.

PCMCIA is an adjustment and further development of ISA. It therefore has the same bus width, i.e. 8 or 16 bits. PCMCIA has been developed for use in portable computers and subsequently it has a very small bus connector.

Input- and output-devices
Separate processor bus and expansions bus

In today's microcomputers, the processor bus is completely separated from the expansion bus (see figure). The processor bus is more powerful than the expansion bus due to a higher clock frequency and a wider bus. The performance of the expansion bus is limited by the standard (see earlier section about expansion buses), while the performance of the processor bus is determined by the choice of processor, amongst other things. Hence by separating the buses, it is possible to exploit the maximum performance of the processor bus while still maintaining a ''compatible'' speed on the expansion bus. With the buses separated, simultaneous transfers on both buses are also possible. Translation of bus width and clock frequency is performed in the so-called **bus controller**. Furthermore, in order to obtain an enhanced transfer performance between the two buses, the signals of the data and address buses are buffered.



A PC's bus-system with separate processor- and expansion bus

Input devices

The most common input device is the **keyboard**. A mouse **(pointer device)** is used to execute commands and to work with the marking of text and images on the screen. With the means of a **scanner** or **digital camera**, you can convert images into digital format for PC processing. A **bar code reader** of the same type as the ones used in shopping malls can be designed as a **pen reader** for PCs. It reads information in a bar code, and the information is then stored in a format that can be read and processed by a computer. Bar coding with an EAN code (European Article Numbering) is standard today on almost all kinds of merchandise. There are **joysticks** and **wheels** with or without tactile (mechanical) feedback in order to enhance the sensation when playing games. A joystick is usually connected via the computer's game port.

Output devices

The **display screen**, or **monitor**, is perhaps the most important part of a computer. A monitor should comply with the TCO 95 requirements that are intended to guarantee good ergonomics, user friendliness and environmental consideration. The equipment should also be prepared for recycling. Naturally, the computer and monitor should have a standby function that switches off the units after a certain amount of time. Such aspects will be satisfied by choosing equipment that has been approved in accordance with the EnergyStar/NUTEK demands for energy efficiency. There are several display technologies to choose between. When it comes to stationary models, displays with cathode ray tubes (CRTs) are still common although different types of TFT or plasma displays are rapidly making their entry, almost entirely lacking magnetic and electric interference fields. As for portable computers, TFT displays are completely dominant. In general, there is also some kind of **printer** connected to the PC, either directly or via a network.

A historical review of graphics

MDA was a standard for monochrome graphics, without colour.

Hercules was a monochrome graphic standard from an independent company that was widely used in the early 80 's. A low price and a good resolution were the main characteristics of Hercules graphics.

CGA was the first graphic standard for PCs. The maximum number of pixels that could be displayed, the so-called resolution,was 640 × 200

EGA, introduced in 1984, allows a resolution of 640 × 350 pixels, while VGA allows 640 × 480 pixels. Even today,VGA is still the most common graphic standard for PCs. **XGA** is an IBM standard dating back to 1987, which has a maximum resolution of 1024 × 768 pixels.

High resolution graphics

Pretty soon, both EGA and VGA were developed further into different "super" versions of the standards with higher resolution and more colours. Nowadays, there are a great variety of graphic cards and displays which support 1024 × 768 pixels and 256 colours or more. Common features for these cards are built-in, extended ROM, their own RAM with 1 MByte or more and perhaps their own graphic processor, which takes over the calculations from the PC's main processor.

Summary, graphics standards.

Graphics	Resolution	No. of colours
MDA	720 × 350 (text only)	monochrome
Hercules	720 × 384	monochrome
CGA	640 × 200	16
EGA	640 × 350	16
VGA	640 × 480	16
High resolution	800 × 600 or higher	16 or more

VGA cable assemblies

Connection of cale assembly for VGA monitor. Connectors: HD D-Sub 15-pin male connector to both PC and monitor.

Conn. no. PC, monitor	Signal name	Description	Dir. PC - monitor
1	RED	Red Video (75 Ω, 0.7 V_{p-p})	>
2	GREEN	Green Video (75 Ω, 0.7 V_{p-p})	>
3	BLUE	Blue Video (75 Ω, 0.7 V_{p-p})	>
4	ID 2	Monitor ID Bit 2	<
5	GND	Earth	—
6	RGND	Red Earth	—
7	GGND	Green Earth	—
8	BGND	Blue Earth	—
9	KEY	Key (no connection)	—
10	SGND	Sync Earth	—
11	ID 0	Monitor ID Bit 0	<
12	ID 1 or SDA	Monitor ID Bit 1	<
13	HSYNC or CSYNC	Horizontal Sync (Composite Sync)	>
14	VSYNC	Vertical Sync	>
15	ID 3 or SCL	Monitor ID Bit 3	<

Ports

External devices are connected via ports. The **parallel port** or **printer port** allows data transfer over 8 parallel wires simultaneously. There are also a number of wires for control and error messages from the printer. Besides being used for the printer, a parallel port can also be used for connecting networks, external diskette units, tape stations, etc. The 8 signal pins of the parallel port could originally only be used for output. It was, in other words, a **unidirectional** parallel port. Only some of the wires for the special printer functions could be used for input. Modern PCs always have a **bidirectional** parallel port. Further improvement in parallel port performance has been achieved with **EPP** (Enhanced Parallel Port). This is a hardware standard supported by the machine's BIOS.

With a **serial port**, only one bit at a time is transmitted over a wire. A serial port is therefore considerably slower than a parallel port, and is used in circumstances where the demands for transmission speed are not so high. The serial port is usually used when connecting a **modem** for Internet connection via the telecom net. The most common standard for serial connection on a PC is called RS-232. A serial port is controlled by a UART (Universal Asynchronous Receiver/Transmitter). This circuit has been available in a number of versions. The development stages have in turn been called 8250, 16450 and 16550. In today's PC, the UART is not a special circuit but a normal part of the PC's chipset. More advanced UARTs are buffered to allow high-speed communication.

Networks

Local Area Networks (LAN) are described in detail in the Fact Sheet's section "Data Communication" .

Computer glossary

10Base-2. BNC-connected thin coax cable RG58 for Ethernet.
10Base-5. AUI-connected thick coax cable RG8 for Ethernet.
10Base-F. Fibre-optic Ethernet.
10Base-T. RJ45-connected cable with twisted pairs for Ethernet.
3270. Synchronous, page-based terminal protocol for networks that use a central computer.
AC-3. The American Dolby Digital audio standard for DVDs.
ADPCM, Adaptive Delta Pulse Code Modulation. Compression primarily for digital sound. ITU standard for encoding voice transmissions.
ADSL, Asymmetrical Digital Subscriber Line. Technology for asymmetrical transmission via a regular copper wire with twisted pairs.
AGP, Accelerated Graphics Port. A PC page standard used to reduce the bottleneck between the PC's memory and the graphics board.
AMD, Advanced Micro Devices. Manufacturer of processors.
ANSI, American National Standards Institute. American authority of standardisation.
APM. Advanced Power Management. Cooperative project between Microsoft and Intel for power management, primarily in portable computers.
ARP. Address Resolution Protocol. Used in order to translate an IP address into a MAC address and vice versa.

ARPANET. Advanced Research Projects Agency. The predecessor to today's Internet. Was first developed in 1957 by the U.S. military as a computer network with high redundancy.
ASCII, American Standard Code of Information Interchange. An international standard for all characters, numbers and punctuations in the computer world.
ASP, Active Server Page. A code executed in the web server before it is sent to the user.
AT commands. Commands in accordance with the Hayes Standard which are used to control a modem, either directly or via a communication programme.
ATX. Standard for motherboards with the dimensions 305×244 mm.
Baby AT. Reduced version of the earlier AT standard for motherboards. Kept the same dimensions as the earlier standard, 8.5×11 inches.
BASIC, Beginner's All-purpose Symbolic Instruction Code. Programming language.
baud. Modulation rate unit, representing the number of signal changes each second. A measure of data transmission speed. Is often incorrectly said to represent the number of bits per second, bps.
BBS, Bulletin Board System. A computer that can be dialled via a modem, acting as a bulletin board.
BIOS. Basic Input/Output System, a part of the operating system stored in a ROM or Flash memory.
bit, **bi**nary digi**t**. A single bit can hold only one of two values: ”0” or ”1”. The smallest unit of information in a computer.
boot. Derived from the expression ”to pull oneself up by the bootstraps”. Computer world lingo for the process of loading the first piece of software that starts a computer.
BRI, Basic Rate ISDN or Basic Rate Interface. An ISDN interface that consists of two B-channels (bearer-channels) and one D-channel for transmitting the control signals.
BSA. Business Software Alliance. An alliance of software manufacturers for combatting illegal software copying.
BSD-Unix. The Unix version Berkeley Software Distribution.
bug. An error or defect in programmes. The expression allegedly derives from an incident in the infant days of computers, when a cockroach got stuck in one of the classic ENIAC's circuits, causing a programme failure.
bus. A collection of logic signals. Often a data or control bus.
byte. A data word that consists of 8 bits.
BYTE. A classic PC magazine (now defunct) that appeared in the 1970s.
C. Programming language developed in the early 70s by Dennis Ritchie on a Digital PDP-11.
C++. An extended version of C with OOP, object-oriented programming.
cache, cache memory. Fast local memory.
CAD, Computer-Aided Design.
CAE, Computer-Aided Engineering.
CAM, Computer-Aided Manufacturing.
CC, Carbon Copy, an e-mail copy.
CCD, Charge-Coupled Device. Common as electronic ”film”, i.e. light-sensitive arrays in digital cameras.
CCITT, Comité Consultatif International de Télégraphie et Téléphonie. European standardisation authority for telecommunications. Nowadays called ITU-T.
CDMA, Code-Division Multiple Access.
CD-R, Compact Disc-Recordable. Recordable CD.
CD-ROM, Compact Disc-Read Only Memory. Able to store about 700 MB.
CD-RW, Compact Disc-ReWritable. Rewritable CD.
CORBA, Common Object Request Broker Architecture. A standard that enables pieces of programmes, called objects, to communicate with one another irrespective of what programming language they were written in or what operating system they are currently running on.
CPU. Central Processing Unit. Really the actual processor in a computer. The term is frequently used nowadays for the entire computer's central unit with motherboard, internal memory, hard disk drive and disk station.
CRC, Cyclic Redundancy Check. An algorithm for detecting data transmission errors.
DES, Data Encryption Standard.
DHCP, Dynamic Host Configuration Protocol. Protocol used for automatic and dynamic assignment of IP addresses to workstations in a network.
DIN. Deutsche Industrie Norm. German industrial standard.
DMA, Direct Memory Access. Method for handling direct data transfers between memory and a peripheral device.
DNS, Domain Name Server. Translates domain names into numeric IP addresses and vice versa.
DOS, Disk Operating System.
DRAM. Dynamic Random Access Memory. Read and write memory.
DSP, Digital Signal Processor. Specially designed computer architecture for rapid, and often numeric, data processing.
EAN, European Article Numbering. Common PIN code system which can be read by pen readers and cash registers.
ECC, Error Checking and Correction. A technique for correcting errors.
ECDL. European Computer Driving Licence.
EDO, Extended Data Out. A type of DRAM.
EEPROM, Electrically Erasable Programmable Read Only Memory.
EPP, Enhanced Parallel Port.
EPS. Encapsulated PostScript. A format for importing and exporting PostScript graphics files.
Ethernet. A common standard for data transmission in networks, originally developed by Xerox back in 1976.
Firewall. Protection that can usually be implemented in both hardware and software, designed to prevent unauthorised users from accessing private networks, either through the Internet or through other networks.**FireWire**, standard for connection of external devices to the computer. The term IEEE 1394 is, however, more frequently used.
FORTRAN, FORmula TRANslation. A programming language.
free BSD. A free version of Unix, similar to Linux.
FTP, File Transfer Protocol. Programme and protocol used on the Internet for sending and copying files.
GIF, Graphics Interchange Format or Graphic Image File. Image format used on e.g. home pages on the Internet.
giga, G. One gigabyte is actually 1.073.741.824 bytes.
GNOME, GNU Network Object Model Environment. A graphic interface for Linux.
HAL 9000. The spaceship's computer in the movie ”2001: A Space Odyssey”.
hub. A central connection point in a network.
ICMP, Internet Control Message Protocol. An extension to the Internet Protocol (IP).
IrDA, Infrared Data Association.
ISDN, Integrated Services Digital Network.
ISO, International Standards Organisation.
Java. A programming language developed by Sun, designed to deliver executable code over networks.
JPEG, Joint Photographic Experts Group. A standard for image compression. JPEG is a lossy compression technique, which means that some amount of data is lost from the image.
KDE. Kool Desktop Environment. Open source graphic interface for Linux/X11.
LAN, Local Area Network. Computer network.
Linux. Perhaps the fastest growing Unix dialect.
MAC address. A hardware address in an Ethernet and the same thing as the Ethernet address that is unique to each adaptor.
modem. Acronym for **mo**dulator/**dem**odulator.
Moore's Law. A principle first introduced in 1965 by Gordon Moore, one of the co-founders of Intel. He predicted that the number of transistors per square inch on integrated circuits would double every 18 months.
MPEG, Moving Picture Experts Group.
nerd, usually computer nerd or geek. A person whose computer interest and skills are considerably greater than his, or her, social skills when it comes to other people. Excluding, perhaps, interactions with other like-minded persons.
open source. Indicating an open source code.
PERL, Practical Extradiction and Report Language. Often used for writing script for Internet applications. Originally developed for finding information in text files. Resembles C.
PGA. Pin Grid Array. Type of package in which the connecting pins are located on the bottom in concentric squares.
SCSI, Small Computer System Interface. Standard for data transmission between units in a computer.
SIMM, Single In line Memory Module.
TCP, Transmission Control Protocol. Handles the streams of data in TCP/IP communication. Guarantees delivery of data between the transmitting and receiving ends and handles error and flow control. Requires an affirmation for each data packet.
TCP/IP, Transmission Control Protocol/Internet Protocol. The protocols that the Internet is based on.
USB, Universal Serial Bus. A system that is similar to Apple's ADB (Apple Desktop Bus) but is intended for PCs.
VT100. Terminal type that many terminal emulation programmes can emulate or ”look like”.

Data communication

A brief overview of communication

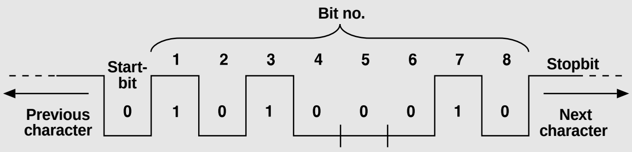
In the world of data and telecommunications, electric signals are transmitted over distances of just a few millimetres inside a semiconductor chip to tens of metres inside a data network, but also thousands of kilometres in cables at the bottom of the ocean or millions of kilometres up in space. All these cases are examples of data communication, but naturally the requirements are very different. When it comes to short distances, rise and fall times as well as voltage levels on the signals are the most vital issues. Where long distances are concerned, interference like noise, distortion and signal attenuation become most vital.

A communication link is the ”road” on which information travels. It might be a physical connection in the form of a copper wire, but it can just as well be an electromagnetic radio signal between two points or a logic connection inside a network. Even though we tend to imagine one transmitter and one receiver most of the time, the fact is that most of today's communication is what we call ''multicast'', i.e. one transmitter and several receivers (radio and television can act as examples).

The communication between transmitter and receiver can be of three different kinds: **Simplex.**, one-way communication (radio, TV, GP).**Half duplex** two-way communication with one at a time (Walkie-Talkie, Ethernet). **Full duplex**, communication in both directions (telephone, GSM). In reality this is often two parallel counter-directional simplex channels.

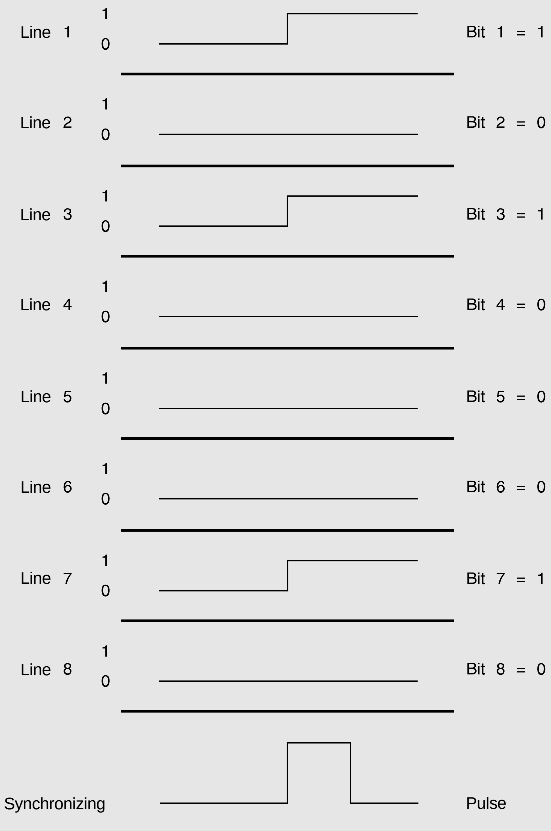
Transmission rate and baud rate

The speed of communications is, historically, often measured in bauds, but nowadays we usually talk about bits per second (bps). The difference being that baud represents signal changes per second ir modulation rate. If one needs to specify 1 signal change per bit, baud and bits/second amount to the same value. If it comes to 2 signal changes per bit however (this is the case for NRZ, Non-Return-to-Zero encoding), the bit speed becomes only half of the baud rate. In the same way, the baud rate for a V.32 modem is 2400, while the bit speed is 9600 since each ''bit space'' represents 4 bits of information by means of a combination of phase and amplitude modulation of the signal.



Serial transmission of ASCII ''E''. 8 bits. No parity.

An other way to increase the transmission rate is to use parallel conductors, parallel transmission. In the sketch below 8 conductors are used and the rate is 8 times the symbol rate.



Parallel transmission of ASCII ''E''.

Asynchronous and synchronous transmission

Serial information must be complemented in some way with information about the beginning and the end of a transmission. Just like the phone starts to ring when we are starting a phone call and we usually say ''good bye'' when we are ending it and hang up the phone. In the world of computers, this is corresponded by a start bit and a stop bit (see picture) for asynchronous communication and a special synchronising sequence for synchronous communication. Corresponding information is not necessary for parallel transmission, as each bit is identified on a special conductor or wire.

Modem

The word ''MODEM'' is a contraction of the words MODulator/DEModulator. The modem takes care of adjustment to the desired transmission standard for transmissions via the public telephone network or a private cable, from town to town or from country to country. The modem converts the computer's signals into tones that the network can handle. The connected equipment is referred to as DTE (Data Terminal Equipment, denoting the terminal or computer) and DCE (Data Communication Equipment, denoting the modem).

There are many types of modems, like e.g. short range modems or telephone modems. In Europe, speed and transmission type of telephone modems are defined with CCITT standards. The so-called V series defines the standards for the public telephone network. The most common ones are specified in the table below.

Modem standards for the common telephone network.

ITU/CCITT-standard	Rate reception/transmission bit per second	Comment
V.21	300	
V.22	1200	
V.22bis	2400	
V.23	1200/75, 75/122, 1200 half duplex	
V.24		Standard for connection between terminal and system
V.26	2400	
V.26bis	2400 and 1200	
V.27	4800	
V.29	9600	4-wire connection
V.32	Up to 9600	
V.32bis	Up to 14400	
V.32terbo	Up to 19200	Lucent replaced by V.34
V.32terbo+	Up to 21600	3Com (USR) replaced by V.34
V.33	14400	4-wire connection
V.34	Up to 33600	
V.90	Up to 56000/33600	Replaces X2, K56Flex
V.92	Up to 56000/47000	

For transmission between two computers, one often uses a null modem, i.e. a cable that is connected in such a way that the computers understand each other as though they were in fact communicating with a modem. Common connections for null modem cable assemblies fitting for the serial COM port on PCs (both 9-pin and 25-pin D-Sub connectors) are shown in table below.

Connection of null modem cable assemblies. Connectors: D-Sub 25- or 9-pole male connectors to computer.

Computer 1 Conn. no. 25-p (9-p)	Name of signal	Dir. of signal	**Computer 2** Conn. no. 25-p (9-p)	Name of signal
3 (2)	Received Data	<	2 (3)	Transmitted Data
2 (3)	Transmitted Data	>	3 (2)	Received Data
20 (4)	Data Terminal Ready	>	6+8 (6+1)	Data Set Ready + Carrier Detect
7 (5)	Signal Ground	–	7 (5)	Signal Ground
6+8 (6+1)	Data Set Ready + Carrier Detect	<	20 (4)	Data Terminal Ready
4 (7)	Request To Send	>	5 (8)	Clear To Send
5 (8)	Clear To Send	<	4 (7)	Request To Send

Connection of modem cable assemblies. Connector: D-Sub 25- or 9-pole male connectors to computer, D-Sub 25-pol female connector to modem.

Computer Conn. no. 25-p (9-p)	Name of signal	Dir. of signal	**Modem** Conn. no. 25-p	Name of signal
1 (–)	Shield	–	1	Shield
2 (3)	Transmitted Data	>	2	Transmitted Data
3 (2)	Received Data	<	3	Received Data
4 (7)	Request To Send	>	4	Request To Send
5 (8)	Clear To Send	<	5	Clear To Send
6 (6)	Data Set Ready	<	6	Data Set Ready
7 (5)	Signal Ground	–	7	System Ground
8 (1)	Carrier Detect	<	8	Carrier Detect
20 (4)	Data Terminal Ready	>	20	Data Terminal Ready
22 (9)	Ring Indicator	<	22	Ring Indicator

Modem control for dial-up modems

There are different ways to control a modem. The two most common ways are Hayes AT commands and CCITT V.25bis. Hayes is the most widespread of these two. Hayes has constructed the commands around the term AT, taken from the word ATTENTION. All commands start with AT. Here is an example: ATDT12345, which is interpreted as ATtention Dial **Tone** 12345.

Error correction and compression

One tries to reduce transmission errors that may occur in the telephone network by means of sophisticated protocols. CCITT's standard V.42 is one example of such a protocol. When large amounts of data are to be transmitted simultaneously, protocols that compress data is preferably used in order to reduce transmission times.

Short range modems

Short range modems, or baseband modems, are also known as Line Boosters. They are used to obtain a secure data transmission over long distances between equipment with serial communication. The most frequently used interface by far is the V.24/RS232, which can only be used up to about 15 metres with safety maintained. By using short range modems, the distance between equipment can be extended with several kilometres. Transmission is often performed over 4-wire (twisted pairs), but coax cable and fibre-optic cable can also be used. By using fibre-optic communication, it is possible to reach hundreds of kilometres with transmission rates of many Gigabits/second.

Local Area Networks (LAN)

There are several types of communication in local networks and for connection to the Internet. The most common type for local networks is Ethernet (IEEE 802.3) which is a CSMA/CD, or CS=Carrier Sense, MA=Multiple Access and CD=Collision Detect. In other words, it detects if it is an open channel, allows several users of the channel and detects if two users are trying to use the channel simultaneously. Ethernet comes in many forms with bandwidths of 10, 100 or 1000 Mbit per second. Transmission can be performed over distances up to about 100 metres.

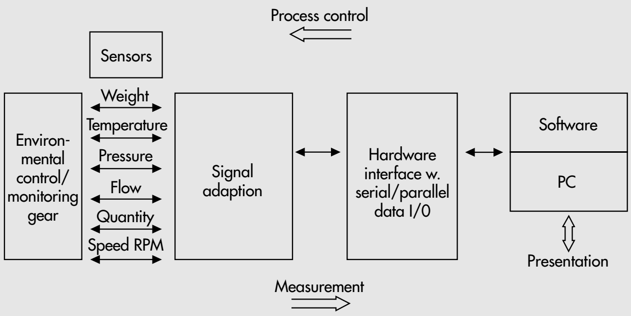
ISDN

Integrated Services Digital Network is often used for new installations of telecommunication services and allows transmission rates of up to 64 kbit/second for each channel. Each BRI-ISDN (Basic Rate Interface) connection contains 2 B channels with 64 kbit and one control channel.

xDSL

The new technology with Digital Subscriber Lines comes in several forms, the most common one being ADSL (Asymmetric DSL). It is offered to customers living within a radius of about 3 km from a telephone station and provides the user with a bandwidth of up to several Mbit/second. The asymmetry consists of a difference in speed between transmitted and received data.

Control systems



Block diagram of measurement and control.

Measurement and control can in principle be divided into the following parts:

- Measurement environment
- Transducers
- Signal conditioning
- Hardware for measurement/control/analysis
- PC
- Connections between computer and measurement/control equipment
- Software

Measurement environment

A number of different physical phenomena can be measured. Different types of transducer are used for this; see the following section.

An industrial environment often puts high demands on measuring instruments. Heavy-duty electrical machines can cause variations in the mains voltage that affect the equipment. Electromagnetic interference is a common occurrence. Moreover, the physical environment itself subjects measuring instruments and computers to severe stresses. Examples are extreme temperatures or temperature changes, humidity, dust, impurities and vibration. Such demanding environments need special arrangements, e.g.dust filters and special vibration-damping mountings.

Transducers

There are two main groups of transducers; those that give digital output signals and those that give analogue output signals. Examples of transducers with *digital output signals* are thermostats, level indicators of different types and optical transmitters. These transducers have only two states, I or 0, and can be connected directly to digital equipment. Transducers for temperature, flow, pressure, weight, speed, etc are *analogue transducers*. The output signal from an analogue transducer changes more or less linearly with the measured variable. Analogue signals may need to be conditioned before the measurement itself can take place see the section called Signal Conditioning.

Signal conditioning

The output signal from a transducer must usually be conditioned before it can be analysed and displayed by the hardware/software. Examples of conditioning are amplification, attenuation, filtering, isolation and linearisation. *Amplification* of the signal is usually carried out as close to the transducer as possible. This allows longer leads between transducer and measuring electronics because the output signal is less susceptible to interference. *Attenuation* may be necessary if the signal is much too strong. *Filtering* is used to remove undesirable components from the signal. In an industrial environment with high voltages, interference or earthing problems often require *isolation* of the signal from the other electronic equipment. Opto isolators and other devices are used for this purpose. *Linearisation* is used to compensate for non-linearity in, for example, temperature transducers.

Signal conditioning is also needed for *control* of processes. Closing of a valve, for example, may need 5 A at 220 Vac, while the output signal from a digital output device or analogue-digital converter (ADC) is 5 Vdc and 10s of mA.

Hardware for measuring, control and analysis

There are many different types of board with one or more of the following main built-in functions:

- ADC (analogue inputs)
- Signal conditioner, e.g.amplifier
- DAC (analogue outputs)
- Digital inputs and outputs
- Relays and contactors
- Counters and/or timers
- Hardware for data analysis

ADCs

An ADC (analogue-digital converter) produces a digital output signal that is directly proportional to the input signal. The higher the *resolution* of the converter, the more accurately the analogue signal can be represented. An 8-bit converter can, for example, give $2^8 = 256$ different analogue output levels. The so-called bit error is thus 1/256, i.e. less than 0.5% of the REF value, the largest input signal value. 12-bit converters are the most common ADCs. Too high a resolution is expensive and gives an even longer conversion time.

Measuring boards often have several independent *channels* on which measurements can be taken. The measurements (samplings)are distributed by multiplexing between the different channels. The *sampling frequency* is then reduced accordingly. If for example a board has a sampling frquency of 32,000 samples/second (32 kHz) and 8 channels are in use simultaneously, every channel is sampled at 4 kHz. *Nyquist's Sampling Theorem* says that sampling must occur at a frequency of at least double that of the highest frequency component to be measured. If, for example,
you are measuring a 10 kHz a.c. voltage, the sampling rate must be at least 20 kHz.

The input can be ”single-ended” or differential. Differential inputs are less sensitive to interference, and are therefore used in environments with high interference, long cable lengths between transducer and ADC or low input signals. Errors such as *linearity errors* or *amplification errors* can be eliminated with the help of either hardware or software.

Many boards contain circuits for adjustable *amplification, attenuation* or *filtering.* This makes the board useful in several different ways. Many boards allow selective amplification on respective channels, which means that the resolution of the ADC can be utilised better.

DACs

Digital to analogue conversion involves the transformation of a digital input signal to a corresponding analogue voltage or current. The specifications for a DAC include information about *resolution* (number of bits, accuracy of the output signal), *settling time* and *”slew rate”*, the maximum coefficient of change of the output signal.

Digital inputs and outputs

Digital inputs are used to read the status of digital transducers, e.g. level indicators and thermostats. Digital outputs are used for control of, for example, valves and relays. The specifications include *number of channels, max driving current* and the *speed* at which data can be received or transmitted.

Relays and contactors

Relays and contactors are used to directly control external equipment. More capacity can be controlled this way than directly from a digital output. Relays can be of the electromagnetic type or semiconductor type. Relay boards often consist of many relays on the same board.

Counters and/or timers

Counters and timers are used to register the occurrence of a digital event, to measure the interval between pulses or to generate square topped pulses. Important properties are the *number of bits*, which directly determines the size of the number that can be counted, and the clock frequency (time base). Some counters have several *channels*.

Hardware for data analysis

In general most of the analysis work is done by the CPU of today's PC, see below. In some high-performance applications, however, the central CPU is unable to process the signals quickly enough. Also, the transmission between the measuring equipment and the computer can be a bottleneck. Certain devices therefore have their own hardware for analysis and processing of data. Generally a high-performance DSP, Digital Signal Processor, is used. Double buffering is employed to collect and process items of data independently of each other.

PCs

In measurement and control systems managed by a PC, it is the computer that determines the general speed of the process. Also, if the measurement collection board has a very high performance level, then the PC must also be able to receive, analyse and display the input data sufficiently quickly.The software used can place large demands on the PC. For an application where simple measurements are made a couple of times a second, a cheap and simple PC is almost certainly enough, whereas a system with real time measurement and computation of high-frequency signals probably needs a 32-bit processor (or higher), a co-processor for floating point calculations, good memory architecture and a fast
disk system.

Industrial PCs

Industry often makes use of bespoke computers constructed of rack-mounted modular systems. These systems are easy to service. Special modifications are included to protect against dust, dirt, vibration and interference, etc. It is however becoming increasingly difficult for the developers of industrial PCs to keep up with the extremely fast improvements in performance that are constantly taking place with the ''standard'' PC.

Connections between computer and measurement and control equipment

There are three different main methods of connecting a PC to measurement or control equipment:

- **Via the PC's serial or parallel port** (this is still the most common method)
- **With measurement and control adaptor cards,** inserted into the computer's expansion bus
- **Via special busses,** developed for measurement and control purposes

Connection via the serial or parallel port

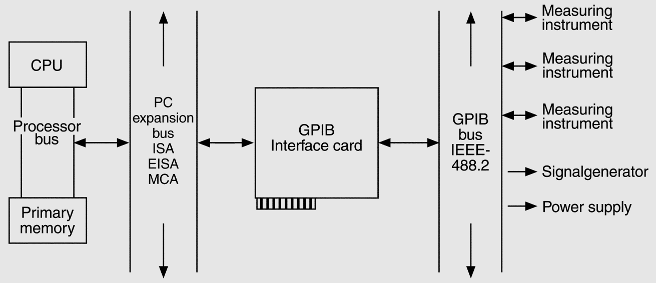
Less sophisticated equipment can be connected directly to the PC's ports. Generally the standard RS232 serial port is used. It is often a question of connecting equipment for measuring or controlling just one quantity, e.g. flow. Special expansion cards with many serial ports are also available, so-called multiport boards. The serial port has limited transfer capacity and can therefore not be used for measurement or control of very fast processes. Nor is it suitable for transmission over long distances. The parallel port is faster. See table. For measurement/control of slow processes however, e.g. temperature variations or flow, the serial port works well and is a cost-effective solution.

Connection with measurement and control adaptor cards

Here data collection or control takes place via a special expansion card on the computer's expansion bus. On the PC, the bus is limited to the normal ISA-bus maximum transfer speed. Computers with an MCA or EISA bus are faster, as are the busses on workstations such as the Sun Sparcstation, for example. Adaptor cards on the computer's expansion bus are considerably faster than the serial port. Measurement and control cards generally have several built-in functions and are therefore very flexible.

Connection via special buses for measuring technology

Specialised and advanced measuring instruments are often connected to a computer via a special instrument bus.



PC-GPIB block diagram.

GPIB bus = General Purpose Interface Bus. Also called HP-IB, IEEE-488 or IEC 625. Hewlett-Packard developed this communications bus for measurement devices in the middle of the 1960s. Several instruments can be connected to a PC over the GPIB bus by means of a GPIB interface card inserted into the PC. On the bus there are one or more talkers, listeners and controllers. The PC, as the controlling device, can play all three rolls with the help of suitable software. GPIB is an 8-bit bus and is relatively slow, ca 1 MB/s. GPIB is nevertheless very common in measurement and control technology.

VXI bus. Launched in 1987 as an extension of the VME and GPIB busses. VXI is a 32-bit bus and allows transfer rates greater than 10 MB/s. VXI is common in applications using the industrial PC.

MXI bus. Multisystem Extension Interface Bus, introduced 1989 by National Instruments. Supported today by HP and others. MXI is a 32-bit bus with support for several bus masters and a max transfer rate of 20 MB/s.

Transfer rates on different channels and buses that are used in measurement and control technology. Note that the rates are given in megabytes/second.

	Max transfer rate
PC ports	
Serial port	2-12 kB/s
Parallel port	1 MB/s
Expansion buses	
ISA bus	1-5 MB/s
EISA bus	33 MB/s
MCA bus	20 MB/s (MCA-2 40 MB/s)
Special buses for measurement	
GPIB	1 MB/s
VXI	c. 10 MB/s
MXI	20 MB/s
Network	
Network (Ethernet)	1, 10, 100, 1000 MB/s

Software

The user has, by means of software, ultimate control over the measurement and control processes.

Software performs one or more of the following functions.

Data collection

The software handles control of the measuring devices' inputs. *Data conditioning* can be done entirely with software, e.g. correction of linearity or other errors.

Control

Control of instruments and processes can be performed directly by software, which shows the results of changes in graphic form with pictures, symbols and figures.

Analysis

The analysis is a form of digital signal processing. The software works in conjunction with the hardware in the computer and the measurement and control devices to convert and analyse the data. The software can also produce *statistics*.

Presentation

The final link in the chain is the presentation of data for the operator. It is important that this is done in a simple and intuitive way, which at the same time ensures that all necessary information is disclosed. The presentation can be displayed on a monitor or on printers of different types.

Data storage

Data storage is another important part of measurement. Data must be stored in the appropriate form, in the appropriate place on, e.g., hard disk or tape for later inspection or analysis.

Measuring in general

Metrology

Putting it simple you can say that metrology is the science about measuring. Everything that has to do with measurement results, in development, measuring or analysis of a test, is covered by the laws and rules of metrology. This field covers everything from the abstract like e.g. interpretation of statistics or practical matters like which scale on a ruler to choose.

Calibration

Calibration means comparing a measurement device with a better standard. A standard in measurement contexts is a reference that is considered to be the correct one at a comparison. You calibrate in order to find out the test object's deviation in relation to the reference.

Adjustment is not the same as calibration. Adjustment entails that if an instrument does not comply with its specifications at calibration and if the quality routines allow it, you may adjust it. It is recommended that you save measurement data before adjusting (as-found) and after adjustment is finished (as-left).

Calibration can be performed on different levels depending on equipment and what it is used for. The most common form of calibration is traceable calibration which in short means that the reference normals that are used are traceable in relation to national and international standards. At *traceable calibration* methods and execution are guaranteed by a quality system like e.g. ISO 9001:2000. Then there is *accredited calibration*, which does not mean greater measurement inaccuracy but guarantees that the accredited laboratory has the necessary competence and routines for carrying out certain types of calibration and measurement jobs (e.g. calibration of reference normals). In this case the national accrediting agency (in Sweden Swedac) has accredited the organisation and a person responsible for ensuring methods and execution.

Traceability

The process of performing a measurment is only a part of the calibration process. During measurement all data concerning the test object's measurement values should be saved, manually or automatically on a computer (often it could be approved or disapproved instead of measurment values). After measuring all information regarding reference and object with measurment data should be compiled. This information reveals the traceability of the test. The information should include calibration date of the reference, and test date. Many laboratories save more data than that in order to conform to requirements of different authorities. Traceability is an unbroken chain of national and international standards. Putting it simply you can say that 1 volt in the USA coresponds to 1 volt in Europe, Sweden, etc.

Calibration sites

You can say that calibrations are carried out everywhere. But most of the metrology work performed in laboratories where surrounding factors like temperature, humidity, vibrations and electrical influence are closely monitored and controlled. There are five types of metrology laboratories with different metrologic functions.

Primary laboratories - here the highest level in metrology is performed and where research into methods for more precise and accurate measurments. They also calibrate according to primary and secondary standards.

Secondary laboratories - here primarily secondary and working standards are calibrated. Calibration with lower degree of accuracy but that requires special equipment and methods, is also performed here. Mobile calibration units are based on secondary laboratories.

Research laboratories - the metrology requirements in research laboratories differ from other laboratories as they depend on focus and aim of the research. A research laboratory can e.g. need the most accurate reference for current measurement on a single electron.

Calibration laboratories - Calibration laboratories are aimed at volume calibration with references that are calibrated by primary or secondary laboratories. The aim of a calibration laboratory is to calibrate as many instruments as possible in the shortest time possible in order to reduce the amount of time the final user is lacking his instrument, without waiving quality standards. The largest calibration laboratories calibrate tens of thousands instruments per year.

Mobile laboratories - Sometimes it is most practical end economic to bring the calibration equipment to the test object than the other way around, for e.g. military calibration, or in large corporations where the process is to expensive to allow it to stop by sending away the instrument. Mobile calibration puts great demands on the logistics as you often have many instruments to calibrate and a short time to do it in.

Measurement terms

Inaccuracy. A value that states how close a measured value is to a "true value". Metrologists prefer to use measurement instability factor as term instead of inaccuracy.

Measurement deviation. The difference between measured value and "true value" on the object. The real deviation is impossible to know exactly, you can only make an estimate.

Measurement instability. An estimate of the greatest fault in a measurement. Measurement instability factor is often stated as a probability that a value is within specifications.

Tolerance. In metrology it means limit values (measurement instability factor) within which an instrument is supposed to comply with specifications.

Verifcation. It refers to the job of making sure that the instrument conforms to specifed performance. If it does not, a decision to adjust, repair, downgrade specifications or simply to discard the instrument, is taken.

Calibration lable. A sticker put on the instrument showing its calibration status. It should display the certificate number of the instrument, who carried out the latest calibration and when the calibration was performed.

Measuring instruments

The Multimeter

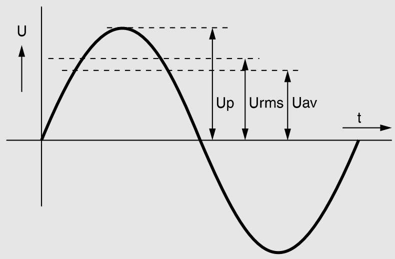
The multimeter, or the universal instrument, is needed in almost all electronic contexts. This combination instrument measures resistance, direct and alternating voltage and direct and alternating current. Increasing numbers of functions are being added, e.g. frequency, capacitance, inductance, transistor testing, etc.

For alternating current and voltage, you must understand how the instrument copes with rectification and how it presents the results. Peak rectification is usually used while the value is presented (with indicators or digits) in the form of RMS. This works well as long as we are recording a pure sine wave.

If we measure a distorted curve shape, we need a true RMS displaying *(TRMS)* multimeter. Either the instrument calculates the true RMS using an integrated circuit (does not manage such difficult signals, high crest factor) or using a converter where the input signal is converted to heat (energy) in a resistor which directly corresponds to the true RMS. On some instruments the direct voltage component can also be switched on or off.

The sine wave's three most important values:

U_p = *peak value*
U_{rms} = *RMS*
U_{av} = *(rectified) mean value*



It is important to define exactly what will be measured. The definition of the TRMS value (True Root Mean Square) for an alternating voltage is the value that in a resistive load produces the same quantity of heat (energy) as a direct voltage with the same value. For example, an RMS alternating voltage of 230 V makes a lamp light up as strongly as a direct voltage of 230 V. This means that the crest value for an alternating voltage is always higher than the RMS value, except in the case of a square wave where the RMS value is equal to the crest value.

The crest factor, or peak factor, is a measurement of the relationship between peak value and RMS. For a pure sine wave, it is 1.414:1 (i.e.$\sqrt{2}$), and for a square wave it is equal to 1. The higher the crest factor that the instrument manages the more difficult the signals that it manages to show correctly.

Some instruments manage to measure the peak value for a signal and using this you can calculate the crest factor.

Form factor is defined as the relationship between RMS and mean value. It is used internally in mean value sensing instruments which show the RMS. They are then calibrated for pure sine signals which have a form factor of 1.11.

When purchasing a multimeter you should be aware that there are occasions where it is not enough to have a mean value measuring instrument and there are occasions where a TRMS displaying instrument is necessary.

However, it is important always to have as good an understanding as possible of the signal's appearance and to take this into account when drawing conclusions regarding the measurements.

	Rectified mean value	RMS	Crestfactor	Formfactor	Error with a non TRMS multimeter (%)	Correction factor
Sine wave (V_p)	$\frac{2}{\pi} V_p$ ($0.637V_p$)	$\frac{1}{\sqrt{2}} V_p$ ($0.707V_p$)	$\sqrt{2}$ (1.414)	$\frac{\pi}{2\sqrt{2}}$ (1.111)	0	1
Square wave (V_p)	V_p	V_p	1	1	+11	0.900
Triangle wave (V_p)	$\frac{V_p}{2}$	$\frac{V_p}{\sqrt{3}}$	$\sqrt{3}$ (1.732)	$\frac{2}{\sqrt{3}}$ (1.155)	−4	1.039
Puls wave (V_p, T, t, $n = \frac{t}{T}$ = Duty cycle)	nV_p	$\sqrt{n}\, V_p$	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	$100\times\left(\frac{n\pi}{2\sqrt{2n}}-1\right)$	$\frac{2\sqrt{2n}}{n\pi}$

Fault display and correction factor at multimeter use. The table refers to average value sensing, TRMS instrument calibrated for sine signals. The correction factor for each curve form is the same as the relation between the form factor of the curve and the form factor for sine = 1.11. Besides fault dispaly according to the table the accuracy of the multimeter must be taken into consideration.

Accuracy

You should also be aware of the multimeter's accuracy and not rely blindly on what the figures show. The accuracy is often specified as a percentage plus a number of digits' deviation, e.g. *0.5 % ±2 digits*. This means that if the multimeter shows 225.5 V, this can really be 225.5 +0.5 % = 226.6 +2 digits – thus *226.8 V* or at the other extremity 225.5–0.5 % = 224.4 –2 digits – thus *224.2 V*. This applies for an instrument with a scale length of min 2999 (sometimes called 3-2/3 digits). Had it been an instrument with a scale length of 1999 (sometimes called 3-1/2 digits), the result would have been 225 V ±0.5 % ±2 digits = 222–228 V.

The conclusion is: Find out the instrument's scale length and accuracy in both percentages and in digits.

Clamp instruments

A clamp ammeter is a very useful instrument for measuring current (A), especially large currents. The major advantage is that you do not need to break up the power circuit but quite simply grip around one conductor on the circuit and read off the value directly on the instrument.

There are clamp instruments for both *alternating current* and *direct current*. AC clamps are the most common and are usually easier to manufacture (cheaper) than DC clamps. Often clamp instruments are combined with the measurement ranges for voltage and the resistance ranges.

There are also more advanced current clamp meters, which as well as the above, measure real and apparent power, and clamps especially for leak current fault detection. When choosing DC clamps, it should be noted that the instrument should have low remanence (remaining magnetism). The reverse leads to poorer accuracy over time. What is stated above regarding RMS and accuracy should also be taken into consideration for clamp instruments.

Problems with mains voltage overtones

Symptoms of overtones usually appear in electrical power supply systems where many non-linear single phase and three-phase power consumers are connected.

Every component in the electrical installation contributes in its way by emitting overtones or being affected by overtones. In total, this entails poorer performance and, in the worst cases, damage.

Odd overtones become a "ghost current" which overheats the zero conductor

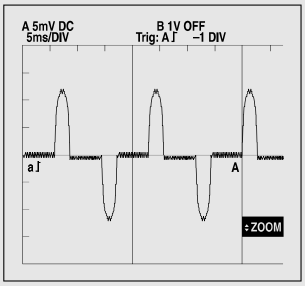
In a three-phase system, the zero conductor can carry a "ghost current" caused by non-linear loads connected to 230 V group cables.

In normal cases, with evenly distributed loads on the phases, the phase currents with a 50 Hz fundamental tone cancel each other out in the zero conductor. If, despite everything, a current flows in the zero conductor, this is probably the result of triple overtones, no. 3, 9, 15, etc. They do not equalise each other but are added to each other.

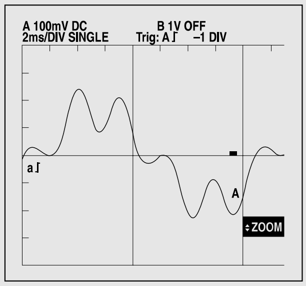
In an installation with many non-linear loads, the zero current can even become greater than the phase current! There is a high risk of overheating because the zero conductor, as distinct from the phase conductors, is not fused (it *must* not be fused). The zero conductor also often has a smaller area than the phases because it will normally carry a significantly lower current there.

A high current in the zero conductor also entails a higher voltage drop than normal between the zero and protective earth.

Problems of this type can be remedied by installing a 5-conductor system.



Single phase non-linear load current.



Three- phase non-linear load current.

Safety switches

Thermomagnetic safety switches are tripped by a bimetal being heated up by the passing current. The bimetal senses the current's true RMS (TRMS). Switches of this type provide better protection against overloading by overtone current than standard fuses and overload relays.

A peak value sensing electronic safety switch reacts to the current's 50 Hz peak value. However, it does not always react correctly for overtone currents. Because the peak value for the overtone current can be higher than the normal 50 Hz peak value, the switch is tripped too early for low nominal current. If, on the other hand, the peak value is lower than normal, the breaker will perhaps not trip when it really should.

Zero rails and connector clips

Zero rails and connector clips are dimensioned for the nominal phase current. They can therefore become overloaded when the zero conductors are overloaded with the sum of the overtones.

Distribution boxes

Overtones in distribution boxes can cause noise. The casing on a distribution box for a 50 Hz current can come into mechanical resonance due to the magnetic field formed by overtone currents. The casing then emits a noise.

Telecommunications

Problems with overtones often occur in telecommunications systems. To keep induced interference from phase current at as low a level as possible, telecommunications cables should be laid as close to the electrical installation's zero conductor as possible. However, this may mean that problems increase because any triple overtones in the zero conductor are transferred inductively and can be heard on the telephone.

Overtones' phase mode and effect.
Every overtone has a name, a frequency and a specific phase mode in relation to the fundamental tone (F). In an induction motor, this means that an overtone current with a positive phase mode radiates a rotating magnetic field with the same direction as the fundamental tone's magnetic field. An overtone current with a negative phase mode gives a magnetic field with a reverse rotation direction.
The first 9 overtones and their phase modes are shown in the adjacent table.

Name	F	2nd*	3rd	4th*	5th	6th*	7th	8th*	9th
Frequency	50	100	150	200	250	300	350	400	450
Phase more	+	–	0	+	–	0	+	–	0

* Even overtones disappear when the waveform is symmetrical (applies in a normal circuit).

Phase mode	**Rotation**	**Effect**
Positive	Forwards	Heating of conductor, safety switches, etc
Negative	Back	Heating (see above) + motor problems
Zero**	None	Heating + increased zero current in 3-phase 4 conductor systems

** Overtones with zero phase mode (odd multiples of 3rd overtone) are called triple overtones (3rd, 9th, 15th, 21st, etc).

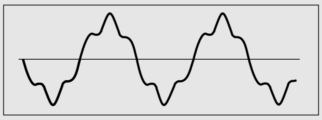
Fault detection for leak currents and mains interference

Problems with interference and leak current to earth are increasing. Clamp ammeters especially adjusted for measuring small currents are an excellent tool. For *leak current measurement*, the clamp is usually placed directly across all three phases and the zero or around phase and zero simultaneously. The result should be zero, otherwise you have a leak current.

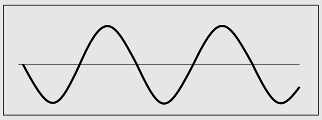
An ordinary clamp ammeter does not take into account whether the signal is distorted or not. This also means that the mean value shown by the instrument does not indicate whether the signal is distorted. Using a *frequency converter* overtone problems on the line can be localised. Using a band pass filter for 50 Hz, all overtones can be filtered out and only the fundamental tone is maintained during measurement.

When testing you first measure the current in "wide" mode and then in "narrow" mode. The difference between the two measurement values is the answer for the signal's overtone content. This function facilitates fault detection and you can quickly and easily trace the equipment generating interference on the network.

This function is not available in all clamp instruments.



Waveform in "wide" mode.



Waveform in "narrow" mode.

Waveforms for different settings of the instrument with frequency switch.

Insulation testers

Tests the insulation and conductive capacity in an electrical installation, machine, motor or apparatus by measuring the insulation resistance using a high voltage which immediately reveals initial sparking, etc. Insulation measurement can be performed using different test voltages, often from 50 V up to 10 kV (the 50 V range is used in an ESD context, etc.), and the result is read off in MΩ, sometimes GΩ. The higher the measurement voltage, the higher the resistance that can be measured.

Another important factor to take into consideration is poor contact at a contact point. If a large current flows through it, heat is formed at that point, which can lead to fire. It is therefore important that the resistance is as low as possible. As a result of this, an insulation tester should also be provided with a low ohmmeter with a relatively high test current (approx 200 mA).

Installation testers

Test instrument for electric installations. Helps to measure and document so that the installation complies with current demands and norms.

EN61557, european standard for CE marking of electric installations.

The installation contractor is responsible for:

- That the instructions of the material supplier are followed when installing.
- That the installation is carried out correctly according to current norm.
- CE marking the connection equipment.
- Documenting the build-up of the equipment, design, who has installed it and give it to the client.
- Checking and beginning use of the installation according to norm.
- That documentation is kept for 10 years.

When requested by the contractor when a new service is installed or added, the mains supplier is obliged to supply information about pre-impedance (Z) at the supply point, immediately before the electricity meter. If (Z) is known the contractor is able to use the correct dimensioned series coupling protection in reference to the supply mains short circuit power and the following installation.

The contractor is obliged to check the cut-out times at shorts or ground connection so that they do not surpass limit values. With information about pre-impedance (Z) this can be calculated but it is much easier to use an installation tester. These values should be documented.

The contractor is responsible for the control of the earth fault breaker, as well as function, tripping time, tripping current, touch voltage and document the data.

The contractor is obliged before an installation is passed on to the client or to the owner of the installation to document the installation. Simpler installations require only a catalogue, otherwise tables and diagrams must show the kind and structure of circuits, supply points, no. of conductors, conductor area, kind of conductor, length of cables and explain the disconnecting devices and their location.

The best choice is *always* to document testing of earth impedance and earth fault breakers. If something should happen with the installation, it is better to have it on record that everything was done right from the beginning, than to prove after the event that nothing has been done wrong.

The text above is freely interpreted from current norms. Complete information can be supplied by the electric security board.

Combination multimeters/oscilloscopes

This is a type of instrument which, as well as measuring signals and presenting the value in numbers, can also show the curve form for the signal. These instruments often combine an advanced multimeter with a digital oscilloscope. Further more they are small and easy to carry around. They also usually have some form of memory so that you can take the curve form "home" and analyse it in peace and quiet. The aim of the manufacturers is that the instruments should be easy to use so that multimeter users will be able to use these instruments.

Safety

Use of handheld measuring instruments in environments with high voltages and high power entails risks for the user.

This basically means that you should follow some simple routines, although the choice of instrument is also important.

The most common fault measurements are:

- Attempt to measure voltage when the fixed test leads are still in the power outlet.
- Attempt to measure voltage for the resistance range.
- Very high transient voltages on the measurement object.
- Exceeding max input voltage.

What makes a multimeter a safe multimeter?

There is no clear answer to this question because there are so many levels and specifications. But if you follow the specifications below you will have come a very long way.

- Fused current inputs.
- Use of high power fuses (600 V or more) which cope with breaking/surge current.
- High voltage protection in resistance mode (500 V or more).
- Transient protection (6 kV or more).
- Safety test leads with slip protection and insulated banana jacks.
- The equipment must be tested and approved by an independent test body (UL, VDE, etc).
- It must be manufactured according to the IEC1010 standard.

Symbol	Meaning
○	OFF (power) SWITCH POSITION
|	ON (power) SWITCH POSITION
~	ALTERNATING CURRENT OR VOLTAGE
⎓	DIRECT CURRENT OR VOLTAGE
≂	EITHER ~ OR —
ϟ	DANGEROUS
⏚	GROUND
⚠	SEE EXPLANATION IN MANUAL
⧈	DOUBLE INSULATION (Protection Class II)
⊟	FUSE
UL	Underwriters Laboratories Inc., U.S.A.
FM	Factory Mutual Research Corp., U.S.A.
SA	Canadian Standards Association, Canada
TÜV RHEINLAND	Technischer Überwachungs-verein Rheinland
VDE	Verband Deutscher Electro-techniker (VDE) Germany
UR	UL recognition mark

Some common symbols on equipment. Even if the texts are in English they are self-explanatory. At the bottom are the symbols and names of some internationally accredited laboratories that carry out safety tests of different kinds.

The oscilloscope

The oscilloscope is the other common universal instrument. It enables you to see waveforms and sequences, superposed voltage, interference in the form of spikes, etc.

Choose an oscilloscope with sufficient bandwidth. Remember that the specification in MHz refers to sine waves. If we divide a square wave into its frequency components (according to Fourier), we find that the oscilloscope must have a bandwidth at least 10 times as high as the clock frequency in order that we can see rise times, jumps, etc. The oscilloscope's *rise time* is therefore a better parameter to use than its specified *bandwidth* in MHz.

Digital oscilloscopes convert the analogue input signal into a binary numerical value which can then be processed by digital circuits. Conversion is performed in an A/D converter, usually with 6 to 8 bits resolution. The highest frequency that can be registered corresponds to (maximum) half of the *sampling frequency* according to the *Nykvist theorem.* An anti-folding filter must be used which prevents mirroring of frequencies over the sampling frequency. This filter further limits the bandwidth because it cannot be done ideally.

The digital oscilloscope has the advantage of being able to memorise a waveform which is simply stored in a memory. The value can then be presented on the screen at any time, or fed to an external printer or a computer. Its greatest advantage lies in being able to memorise single sequences, but the limitations of the digital oscilloscope must also be understood. It can miss narrow spikes which can lie between two sampling points. It the narrow spike is recurrent, it can be recorded using an analogue oscilloscope, but this imposes greater requirements for the light intensity which is low if the spike occurs at a low frequency.

The oscilloscope should have a high acceleration voltage in order to give good light and good sharpness.

Because analogue oscilloscopes have some advantages and digital oscilloscopes have different ones, it may be advantageous to choose a combination oscilloscope that combines the best properties of each. There is also a new type of oscilloscope called *DRO (Digital Real time Oscilloscope).* Here, using a very high sampling speed you have a digital oscilloscope which behaves likes an analogue oscilloscope. A typical feature of these is that the sampling speed is 4-5 times greater than the analogue bandwidth. The advantages are also that a single sequence can be captured up to the analogue bandwidth and that problems with aliasing disappear.

The trigger facilities in the digital oscilloscope are more advanced than in the analogue one because you can often pre- and post-trigger or trigger on windows and logical conditions.

Probes for oscilloscopes should be chosen and used correctly. A probe which does not suppress the signal (1:1) has a capacitance which is in parallel with the oscilloscope's input. This capacitance, and the oscilloscope's input resistance load the measurement object. If you do not need the oscilloscope's maximum sensitivity, it is better to use a probe with suppression, e.g. 10 times. This "relieves" the oscilloscope's input impedance from 1 MΩ and e.g. 40 pF so that the measurement object is instead loaded with e.g. 10 MΩ and 15 pF. The probe includes a trimming condenser which must always be adjusted the first time the probe is connected to the oscilloscope. Connect the probe to the oscilloscope's trigger outlet. The square voltage is set as optimally as possible so that the square wave cannot be either a jump or a rounded front edge.

To make optimum use of the oscilloscope, it is important to choose a probe with a short rise time. This is added to the oscilloscope's own rise time. A good rule of thumb is that the probe should have a bandwidth twice as high as the oscilloscope. Some rated probes do not specify the probe's bandwidth but for which oscilloscope bandwidth the probe is suitable.

At high frequencies and/or high-resistance measurement points the described probes may not be adequate. Sometimes you can use 50 ohm impedance adapted input to avoid getting faulty results due to reflexes. An other good alternative is an *active probe* that has an amplifier at the tip. This gives a high-resistance probe with short connection conductors (no standing waves) and thereby best possible measurement. The backdrop besides the price is that it becomes large, can be experienced as clumsy and therefore be difficult for reaching certain measurement points.

The probe must withstand the voltage being measured. The peak voltage in e.g. a 230 V wall outlet is 325 V.

Some oscilloscope terms

A/D converter (ADC). Analogue to digital converter. An important part of a digital oscilloscope. Retrieves the signal and converts it into numbers stored in a memory. There should be a separate one for every channel, otherwise it reduces the total sampling speed for 2-channel operation. Cf. D/A converters below.

Aliasing (folding distortion). When a signal is sampled at a speed which is less than 2 times per period of the highest frequency component in the signal, an effect called aliasing appears. The result is a waveform which resembles the correct one but with a lower frequency, of the signal only contains one frequency component. Otherwise the frequencies that transcede half the sampling frequency will be transformed and the resulting signal will be distorted and hard to recognize.

Alternate mode. Several channels display simultaneously by the oscilloscope switching between the channels between the ''sweep'', so that a whole screen is drawn for one channel at a time. Used for short sweep times. Compare Chop mode.

Analogue bandwidth. Analogue bandwidth concerns the input amplifier for both analogue and digital oscilloscopes. This gives the highest frequency (sine wave) which can be reproduced without major changes in the curve's form and amplitude (–3 dB point gives a curve with 30% lower amplitude). Curve forms other than a sine require a higher bandwidth in order to reproduce the same frequency without changing the curve's waveform. If you want to study a square wave frequency, it is more appropriate to choose an oscilloscope according to the rise time.

ART Analogue Real Time Oscilloscope. Freely translated: analogue oscilloscope.

Auto-setup. Automatically gives a setting for best display for the curve. Not to be confused with automatic range adjustment which always follows and resets the settings as the input signal changes.

Averaging (mean value formation). "Averaging" is a waveform process technique which works for multiple waveforms. A number of measurements are taken at every measuring point and the mean value for the point is calculated. One advantage with "averaging" is that it reduces the noise in the signal.

Bandwidth. See "Analogue bandwidth".

Chop mode. Several channels display simultaneously by the oscilloscope switching between the channels during the same ''sweep'', so that only a short section of the curve for a channel is drawn before switching channel. (The curves are ''chopped up''). Used for long sweep times. Compare Alternate mode.

CRT (Cathode-Ray Tube). The picture tube in the oscilloscope. Should have a high acceleration voltage to give good light intensity.

Cursor. See "Cursors".

D/A converters (DAC). Digital to analogue converters in digital oscilloscopes. Converts the numerical value from the oscilloscope's memory and displays it as a curve on the screen. Cf. A/D converters above.

Digital bandwidth. See "Sampling" and "Single shot bandwidth".

DRO (Digital Real time Oscilloscope). A digital oscilloscope with a significantly higher (4–5 times) sampling frequency than the analogue bandwidth. Gives the digital oscilloscope an analogue feeling up to the maximum bandwidth and prevents aliasing.

DSO (Digital Storage Oscilloscope). Freely translated: Digital storage oscilloscope.

Dual-sweep (double time base). An oscilloscope which can display a signal with two independent time settings. You can then magnify a small part while at the same time viewing the whole curve. Also used for delayed sweep.

Delayed sweep. See "Dual sweep".

Glitch Capture (Peak Detect). Function for digital oscilloscopes which helps to capture short "spikes" irrespective of which time base is set. You can capture spikes between the sampling points.

Holdoff. Trigger Holdoff is a function that prevents the oscilloscope from triggering for a set adjustable time interval. The function is used for complex waveforms so that the oscilloscope will only trigger on the first point in the wave form which fulfils the trigger condition.

Cursor. Function with e.g. two crosses which can be moved along the curve on the screen to measure time, frequency and voltage. The answer is given as numbers on the screen. There are oscilloscopes with cursors that follow the curve automatically, as well as instruments which can measure voltage, time and frequency directly without using a cursor.

Memory depth (Memory length). Specifies how many measuring points a recorded curve contains (for example 1 Ki = 1024 points). The more points, the longer the time or the better the results on the screen.

Word length. Specifies the resolution in a vertical axis (for example 8 bits gives 256 points). The more bits, the better the resolution and the better the results on the screen.

Pre-trigger. Digital oscilloscopes have a facility for recording before the oscilloscope triggers.

Read-out. This function shows certain settings in plain text on the screen, e.g. 0.5 V/div, 20 ms/div. The trigger point for digital oscilloscopes is also often marked. This is also the name for the function for probes which then "tell" the oscilloscope whether it is a 1:1 or a 1:10 probe. If it is a 1:10 probe, you can see this on the probe's BNC connector which is then provided with a small pin.

Sampling. A digital oscilloscope samples (measures) the voltage of the signal at regular intervals, e.g. up to 10 million ops/sec, 10 MS/s. However, short spikes which occur between two samplings cannot be seen on the screen. Functions like "Glitch Capture" are then required. Approx 10 samplings per period are required in order to reproduce a signal. The sampling speed thus determines the bandwidth in digital mode, on condition that the sampling speed is not higher than the analogue bandwidth. See "DRO".

Sampling, real-time. Single shot recording (recording of a single sequence) of the signal means that all data points are taken during one period of the signal. The bandwidth is limited by the sampling speed.

Sampling, equivalent time. For repetitive signals, several points are taken for every period in order to build up the signal gradually. The bandwidth is the same as the analogue bandwidth. Here you should observe that "aliasing", and short interferences which do not repeat frequently in the same place, can be missed using this method.

Single shot bandwidth (Digital bandwidth). The single shot bandwidth or bandwidth is the highest frequency that the oscilloscope can display on the screen. THis is also known as real-time bandwidth. For analogue oscilloscopes (also for DRO) it is the input amplifiers that are decisive. For digital oscilloscopes, it is the sampling frequency that is decisive, but not for DRO. It is then real-time sampling that is referred to.

Rise time. The time it takes for the signal to rise from 10 to 90 % of its original value on the screen.

Trigger. Signal event is used to start the sweep on the oscilloscope.

Frequency counters

Frequency counters contain a counter which is connected for a certain period. An oscillator attends to this time window and it is primarily the oscillator's accuracy which decides the accuracy of the instrument. Multi-digit accuracy requires a long connection period at low frequency. Some frequency counters are therefore reciprocal, which means that it is instead the input signal's period that governs the number of count pulses from the oscillator. The counter value is inverted and then presented. A reciprocal counter can therefore have fast updating even at low frequencies. The trigger levels can be fixed or adjustable.

Universal counters also often have the capacity to measure parameters such as time interval, period time, ratio between two frequencies and rotational speed.

Measurement converters

The measurement converter adjusts signals from sensors and protects electronic equipment like PLC, electronic control and industrial computers. It gives galvanic insulation between sensors and other electronics, which eliminates earth currents and electrostatic interference at the same time as it filters out electrical interference.

The output signal from a measurement converter is independent of the size of the load (within certain limits). Its advantages can be summarised as follows:

- Transfer of measurement values over relatively large distances.
- Several measuring or recording instruments can be connected simultaneously to the same converter within the limit for permitted load. Special trimming measures are not necessary.
- Adjustment of cable resistance is not necessary for connected instruments.
- Cable installation is easy and cheap.
- Individual instruments or other measuring and recording devices can be disconnected from a circuit after short-circuiting of corresponding connecting cables without disturbing other equipment.
- For easy adjustment to panel instruments.

Generating instruments

Generating instruments are e.g. signal generators for high frequencies and tone generators and function generators for low frequencies.

From a **tone generator** one primarily requires low distortion and good amplitude accuracy throughout the frequency range. It usually has an output with a square wave.

The function generator is more universal. As well as sine waves it also gives edge waves and triangle waves, and sometimes even bursts. It can often give linear or logarithmic frequency sweeps which means that it can be used in automated measuring systems. A relatively high distortion, however, means that it is not suitable for distortion measurements.

For a **signal generator** for high frequency a number of requirements are imposed, depending on area of application. In general, it should be well shielded and have a good attenuator so that the amplitude accuracy is high.

Synthesis generators give good frequency accuracy through the frequency being referred to one, or more, crystal oscillators with low temperature operation. Unsuitable or simple connections can give high phase noise which makes the signal generator unsuitable for selectivity measurements (outside band measurements). The phase noise is not critical for inside band measurements if a high signal/noise ratio is not required.

System instruments

System instruments are those instruments that can be connected together so that they are controlled centrally from a computer. An instrument in such a constellation can also act as a "master" and the others as slaves. The most accepted standard is GPIB (General Purpose Instrument Bus), also called IEEE-488, HPIB and IEC625. GPIB is often used in automated measuring systems (ATE, Automatic Test Equipment).

A simpler and cheaper solution is RS232C with which, however, the speed is restricted.

Temperature measurement

Temperature measurement is the most common measurement in industry. Temperature is measured in the most varying contexts. Some examples: soldering, plastic moulding, the food provisions industry, fast charging of batteries, finding overloaded high voltage parts, various forms of process industry, etc.

When choosing a measurement method, the purpose of the measurement should be considered.

- Within what temperature range do you want to measure?
- How should the sensor be designed?
- How accurate a result is required?
- How fast should the measurement be?

Which is the most important of the following points:

- An exact numerical value.
- The difference between several measuring points.
- The difference between recurrent measurements for the same measuring point.

Some of the most common temperature measurement methods in industry will be described below.

Resistance sensors

The sensor comprises an encapsulated metal wire whose resistance increases as the temperature rises. These wires can consist of platinum, copper or nickel. The resistance is often 100 Ω at 0 °C, but 10, 500 and 1000 Ω also occur. The **Pt 100** is the most common of these. This designation stands for platinum, 100 Ω at 0 °C. Pt sensors can be manufactured for measurement from –250 to +800 °C.

The resistance change for changed temperature is relatively small, around 0.4 Ω/°C for Pt 100. In order to avoid too great a measurement error, you must compensate for the resistance in the cable between sensor and instrument. Very high accuracy can be achieved with a 4-wire connection.

The resistance in a Pt100 sensor, according to EN60751 (ITS90), should conform to the following formulas:
for –100°C < t < 0°C :

$$R_t = R_0\,(1 + At + Bt^2 + C(t - 100)\,t^3)$$

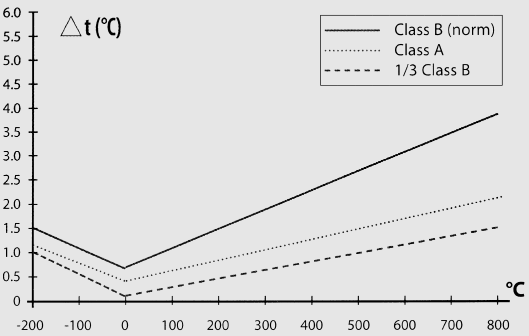
for 0°C < t < 850°C :

$$R_t = R_0\,(1 + At + Bt^2)$$

where
R_t is the resistance for the temperature t
R_0 the resistance at 0 °C
$A = 3{,}9083 \times 10^{-3}$ / °C
$B = -5{,}775 \times 10^{-7}$ / °C^2
$C = -4{,}183 \times 10^{-12}$ / °C^4

Some resistance and temperature values according to this formula are given in the table "Output signal from Pt-100 sensors and thermo elements type K" further on.

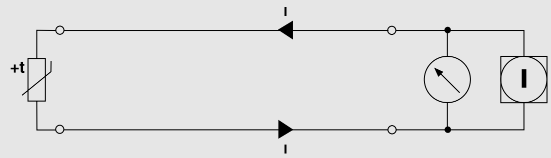
EN60751 defines three classes, A, B and ⅓B, for how much the sensor may deviate from the standardised values. See figure below.



Accuracy for Pt100 sensors. Δt= The allowed error, in degrees, for class A, B and ⅓B according to EN 60751.

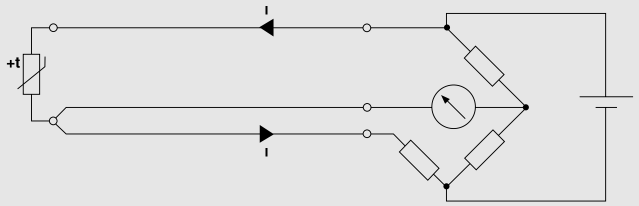
Connection of resistance sensors

One example of **2-wire connection**. The resistance in each of the cable's conductors amounts to 0.5 Ω. In total, the instrument senses the sensor's resistance plus 1 Ω. For Pt 100, 1 Ω corresponds to a temperature change of approx. 3 °C. The instrument will show approx. 3°C too high a temperature. 2-wire connection should only be used in situations where the sensor is situated close to the instrument or where high accuracy is not required.



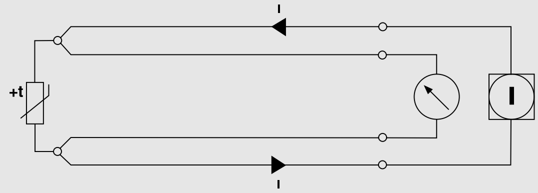
2-wire connection. The measuring current is carried in the same conductors as the voltage across the resistance is measured with. A voltage drop in the conductors causes incorrect indication if the cables are too long.

3-wire connection. In slightly simplified terms, one could say that the third wire is used to measure the cable's resistance so that the instrument will be able to compensate for this. The three conductors should have equal resistance. (2-conductors with a shield as a third conductor is not recommended). If the signal is to be transmitted over long distances or in a disturbing environment (close to heavy current cables or large electrical machines), shielded cable is recommended. Pt 100 with a 3-wire connection is often used in industry.



3-wire connection. A Wheatstone bridge can compensate for the cable resistance in this way. The voltage is measured with high impedance.

The most accurate version is the **4-wire connection**. This connection can also compensate for any difference in resistance between the four conductors (see 3-wire connection). It gives high accuracy and is primarily used for accurate laboratory measurement and calibration.



4-wire connection. The measuring current is carried in two conductors and the voltage is measured with high impedance with the other two conductors. High measurement accuracy can be achieved.

Thermoelements

This measurement method is based on the fact that different metals emit or take up different electron quantities at the same temperature. If two metallic conductors are connected in series and you measure the voltage between them, you achieve a voltage which varies with the temperature at the point where the two different materials meet. This voltage is called thermoelectric voltage. The voltage is small, approx. 40 µV/°C (for type K). Thermoelements are used in most industrial temperature measurements.

Thermoelements can be manufactured from many different metal combinations with different qualities, e.g. when measuring extremely high temperatures. Some types have been standardised for the sake of simplicity. A common standard is called **type K**. There are many instruments and sensors for type K. They provide an accuracy which is usually high enough (type K is frequently used in industry).

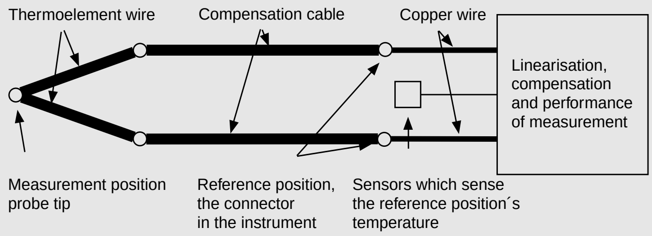
Output signal from Pt100-sensors and thermoelements type K.

Temperature °C	Resistance Pt100 Ω	Voltage Thermoelement type K µV
–50	80.31	–1889
–40	84.27	–1527
–30	88.22	–1156
–20	92.16	–778
–10	96.09	–392
0	100	0
10	103.90	397
20	107.79	798
30	111.67	1203
40	115.54	1612
50	119.40	2023
60	123.24	2436
70	127.08	2851
80	130.90	3267
90	134.71	3682
100	138.51	4096
110	142.29	4508
120	146.07	4920
130	149.83	5328
140	153.58	5735
150	157.33	6138

Connection of thermoelements

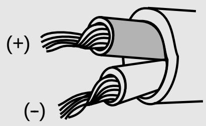
Because this principle is based on two different metals being connected in series in the sensor, the cable between the sensor and the instrument must consist of the same two metals as the sensor, or metals with the same thermoelectrical properties.

This type of cable is called **compensation cable**. Connectors, etc., should also be manufactured from the same metals. Otherwise a number of thermoelements will be connected in series, one at every point where different metals meet. This will cause measurement of temperature at several points, producing a significant measurement error. You must also observe the polarity of the sensor, compensation cable and connectors.



Outline diagram for connection of thermoelements

The figure below shows colour marking according to DIN IEC 584 and DIN 43714 for compensation cables. You should remember that 200 °C is the maximal temperature, even if the insulation material allows higher values. This is because the thermoelectrical properties are only guaranteed up to 200 °C. At higher temperatures thermoelement wire or thermal wire must be used. A mixture of different makes can entail measurement errors as exactly the same alloys are never used by the different manufacturers.
When splicing compensation cables, the cables should be in direct contact with each other. Twine the cables together and clamp them under the same screw. Then protect the join from oxidation.



DIN IEC 584 *White minus conductor, plus conductor as below*		**DIN 43714** *Red plus conductor, minus conductor as below*	
T Brown	*Cu - CuNi*	*U Brown*	*Cu - CuNi*
E Purple	*NiCr - CuNi*		
J Black	*Fe - CuNi*	*L Blue*	*Fe - CuNi*
N Pink	*NiCrSi - NiSi*		
B Grey	*Pt30Rh - Pt6Rh*		
K Green	*NiCr - NiAl*	*K Green*	*NiCr - NiAl*
R Orange	*Pt13Rh - Pt*	*S White*	*Pt13Rh - Pt*
S Orange	*Pt10Rh - Pt*	*S White*	*Pt10rh - Pt*

Colour marking for compensation cables according to DIN IEC 584 and DIN 43714. Do not mix type J (Fe-CuNi) with type L (Fe-CuNi). They have different temperature coefficients. The same applies for type T(Cu-CuNi) and U (Cu-CuNi). Note that specified materials refer to the thermoelements, while those in the compensation cable can vary.

Reference position, or **cold solder position**, refers to the point where the compensation cable changes to normal copper wire, usually inside the instrument. The sensor or the probe is called a measurement position or hot solder position.

If the reference position and the measurement position have the same temperature, e.g. +20 °C, the electrical voltage that the instrument senses is equal to zero. But the instrument should not display zero if the probe senses the temperature +20 °C. Therefore you must compensate for the temperature of the reference position, "compensation for cold solder position".
Subsequently, there is a temperature sensor next to the connector in every instrument for thermoelements. The accuracy of this sensor can cause measurement errors if the instrument is too hot or too cold. For most instruments, the highest measurement accuracy is obtained when the instrument is at normal room temperature.

Some standard thermoelements can be used from –200 °C and some measure over +1500 °C.

Thermistors

Thermistors are used as sensors for some instruments. There are two types of thermistors: PTC (= positive temperature coefficient, i.e. the resistance increases as the temperature rises) and NTC, negative temperature coefficient.

It is relatively easy to design the electronics required in the instrument to linearise the signal from the sensor. This makes the instrument cheap to manufacture.

It is not as accurate, but that can be improved by calibrating and trimming the instrument with a sensor. You can also trim the instrument in order to achieve greater accuracy within a limited temperature range. Thermistors are usually used within the range –50 °C to +150 °C, with a maximum of a few hundred degrees. Examples of application areas are "indoor/outdoor" thermometers and fever thermometers.

Semiconductor sensors

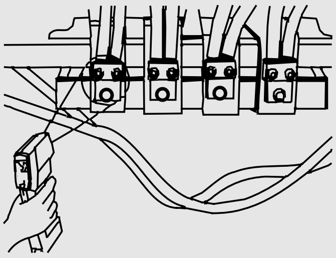
Semiconductor sensors are available in various versions, more or less intelligent. Some have mV output, others have computer- adjusted output in the component. They have a low price but a narrow temperature range, up to approx. +150 °C. The user personally designs the surrounding electronics required.

Temperature-sensitive indicators

Temperature-sensitive indicators look like tape strips, with one or more fields with the maximal temperature specified. When the specified temperature is exceeded, the field changes colour. The change remains so that you can later inspect it to see whether the object was exposed to too high a temperature.

IR-measurement, pyrometers

All objects hotter than the absolute zero point (approx. –273 °C) emit heat radiation in the form of infrared light, IR. This radiation increases with increased temperature. A pyrometer "sees" the radiation and presents the result as temperature.



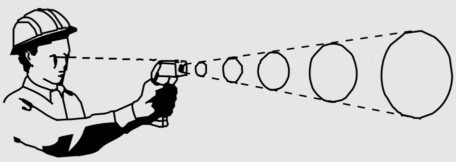
Contact-free temperature measurement can be used for fault detection in high voltage systems during operation. Overheated contact points can be a sign of overloading or poor contact.

During measurement, one must consider the surface material of the measurement object. Different surface materials have differing abilities to emit radiation at the same temperature. **The emission factor** ε describes this property. On many instruments, you can set this factor in order to measure different surfaces correctly. Other instruments have this permanently set to 0.9–1.0. The instrument then displays a slightly too low temperature when measuring on e.g. bare metal surfaces. Some surfaces can also act as a mirror and reflect heat radiation from hot objects in their vicinity. In these cases, one can paint the surface with matt black paint and/or draw up a calibration table.

This measurement principle has high repeatability, i.e. you obtain little difference between the measurement results every time you measure this way. Because you measure without making contact, you can measure on objects which cannot be measured using traditional measurement methods, e.g. high walls, very hot objects, rotating and other moving objects, and objects with high live voltage. Measurement is also very fast, since there is no probe with a mass that must be heated up by the measurement object.

Note, however, that the instrument will display a mean value if the sensor "sees" several zones with different temperatures. In this case, zones refer to both time and surface.

Pyrometry is the only measurement method able to measure temperature above 2000 °C.



For contact-free temperature measurement, the measurement surface becomes greater at longer distances.

Calibration of temperature instruments

You can carry out a simple calibration yourself of a temperature instrument. The probe or sensor is stirred in a mixture of ice and water. The result should be close to 0 °C. The probe or sensor is then put in steam from boiling water, or in the actual boiling water without touching the bottom of the boiling vessel. The instrument should then show close to 100 °C at normal air pressure.

The error indication is entered in a table and used when one needs exact measurement results. The instrument may need to be adjusted.

Radio wave dispersion

Radio waves are dispersed, like light in a straight line between two points, in an electromagnetic wave movement. The dispersion requires a clear view. On the whole, this limits the communication to the distance of the horizon. The range can be increased with a higher aerial.

However there are several factors which make it possible to receive broadcasts from remote places, far beyond the horizon.

Here are some examples of how the waves are dispersed:

- The so called **surface wave** is caused by a interaction with inducing currents in the ground. This wave follows the bend of the earth's surface and can therefore reach far beyond the horizon. It has farthest reach across water and is noticeable for frequencies of up to a few MHz, that is mostly long wave, but also medium-wave.
- The radio signals can bounce against the different **ionised strata** that appear in the atmosphere at different heights above the earth's surface. The lowest stratum is activated during daytime. This stratum attenuates primarily frequencies below 3 MHz, and is not reflective. The higher frequencies pass to the higher placed layers. These are also more ionised and act like mirrors. Therefore a short-wave connection is efficient during the day. Above a certain frequency the signals are no longer reflected and pass out into space. The highest usable frequency is called **MUF** (Maximum Usable Frequency). It varies not only during the 24 hours of the day, but also in different directions and at different times of the year. Above all it is affected by the sunspot number, which has a peak every 11th year. That is when MUF is at its highest. Sporadically, so-called E-strata can sometimes provide long range transmissions on VHF. These occur primarily during the summer.
- **Scatter**, or forward dispersion means that the signal is dispersed due to unevenness in a stratum, and a part of the signal goes towards the earth and can be received, even with a very weak intensity. Troposphere scatter is a common dispersion for signals on VHF and UHF, which allows long-distance connections. This requires high transmission power and directional aerials with high aerial gain.
- **Meteor trace connection** can be established by means of the strata that occur when meteor showers come into the atmosphere, burn and thereby ionise the air. These ionised strata form a good reflector and attenuate the reflected signals relatively little. However this kind of connection can only be established for a few seconds, which is compensated by transmitting telegraphy with high speed, e.g. 1.000 characters/s. The technology is used by broadcast-amateurs and the military.
- **Northern lights, Aurora Borealis**, provides strong ionised layers. A radio amateur who lives in the middle of Sweden can direct the aerial towards north and the ionised area, and in that way obtain connection with stations in the south. The reflected signals will change their frequencies through the doppler-effect because the strata are moving. The signal becomes strongly modulated with low frequent noise. The unmodulated telegraphy signals produce a buzzing tone and speech modulated signals, e.g. SSB will be very distorted. The radio signals received are often strong,despite moderate transmitting power from the station. Since the ionised area can be large, it is an advantage to have an aerial with a relatively wide reception angle, i.e. an aerial with low aerial gain.
- **Moon bounce** means as the name indicates, that the moon's surface is used as a reflector. The technology requires aerials with very high aerial gain, high transmit power and a very sensitive receiver (with low noise factor and narrow bandwidth).
- **Satellite**. There are some satellites for amateur-radio in operation. Long-distance connection with low power can be obtained by using these. They are normally on the VHF/UHF or SHF frequency range.

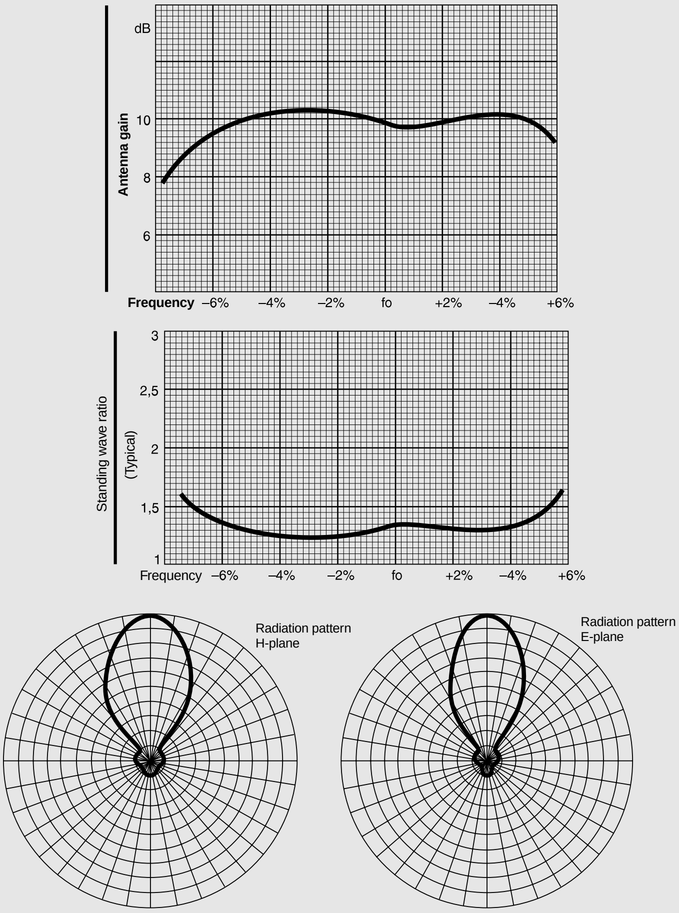
Aerials

An aerial can be omnidirectional or directional. an omnidirectional aerial emitts the same power in all directions but a directional aerial emitts more power in a certain direction, the so called direction of the main beam. This larger value can be seen as a gain in relation to an omnidirectional aerial and is designated as *antenna gain*. An aerial can be used for both transmission and reception. The aerial's so-called gain is the same in both cases. The aerial is completely passive and its gain is obtained through directional effect. The incoming or outgoing energy is concentrated in a narrow lobe. The narrower it is the higher the gain. Since the gain is not active, it is called 'aerial gain' . The aerial gain is usually specified in dB against a single dipole aerial.

If the aerial is optimised to obtain as narrow a head lobe as possible i.e. to obtain highest possible aerial gain, side lobes in the backward direction often occur. This is of minor importance for a transmitting aerial. As a rule, the highest possible aerial gain for reception on the VHF and UHF is aimed at, and side lobes do not have much effect. However, it is important to obtain the highest possible ratio between desired and disturbing signals, when receiving short-wave, where the stations are densely packed in the frequencies. In this case it could be an advantage to use an aerial that has less efficiency in the main direction and instead is optimised for lowest possible side lobes.

More elements and a longer boom contributes to more gain for an aerial of the Yagi type, e.g. a TV-aerial. This reduces the aerial's lobe angle, not only horizontally but also vertically. With radio communication it is often desirable to move the lobe around the horizon using a rotor. If the aerial has very high aerial gain (narrow lobes), it can also be necessary to change the angle vertically, i.e. tilt the aerial.

To obtain the best output power from the transmitter and signal/noise ratio from the receiver, the down-lead and aerial must be impedance adjusted to each other, so as much energy as possible is transmitted. Maladjustment creates standing waves.



Example of aerial gain, standing wave ratio and radiation diagram for a Yagi-aerial with 8 elements.

Wide-band aerials

Receiver aerials for 0.3–3 MHz can be made of wire in an L- or T-shape, or of wire diagonally laid towards a high point. The wire can be from a few metres up to 30–40 m long. It is also important that there is good earth connection to the receiver.

For receiver aerials for 3–30 MHz a 5 to 10 m long wire can be used. A considerably longer wire can be used without the impractical disadvantage of standing waves that reduces the aerial's efficiency at certain frequencies.

An active aerial could be a solution where you do not have the space for a longer aerial wire. The active aerial consists of a short feeler that has very high and capacitive impedance. The feeler is connected to an active circuit, which converts the impedance to 50 Ω and may also amplify the signal.

Tuned aerials

aerials for amateur radio are tuned for one or more amateur bands, to provide the best adjustment for transmitters and receivers.

A half-wave dipole is an excellent aerial. Its length is calculated using the formula:

$$L = v_c \times 0.95/(2 \times f)$$

Where L is the length of the aerial (in m), v_c is the speed of light in a vacuum (300 × 10^6 m/s) and f is the frequency (in Hz = s^{-1}). The constant 0.95 is required because the dispersion speed in a copper cable is lower than the speed of light.

A half-wave dipole has approx. 70 Ω impedance and therefore a 75 Ω coaxial cable should be chosen as a down-lead . The dipole has a radiating diagram in a figure of eight, i.e. it has two wide lobes across the aerial direction with tapering in the aerial wire's direction.

On short-wave it is desirable to have better directional effect, to be able to avoid superfluous noises. For long range communication on VHF and UHF it is important that there is a high aerial gain. Multiple Yagi aerials can be stacked sideways or upright to increase the aerial gain. Each doubling of the number of aerials provides an extra aerial gain of 3 dB.

Aerials for private radio are exclusively omni-directional and vertically mounted. By raising the aerial the range can be increased.

The down-lead from the aerials is usually 50 Ω coaxial cable. Note that the losses increase with longer length and higher frequency. See data for coaxial cables.

If the attenuation in the cable is e.g. 3 dB, the transmission power is reduced by half from e.g.100 to 50 W. This also means that there is less reception sensitivity, since the reception's noise factor also increases with 3 dB. For each deterioration, attenuation of 3 dB, the power reduces by half. In the example above the power is reduced to 25 W at 6 dB attenuation, to 12.5 W at 9 dB attenuation and so on.

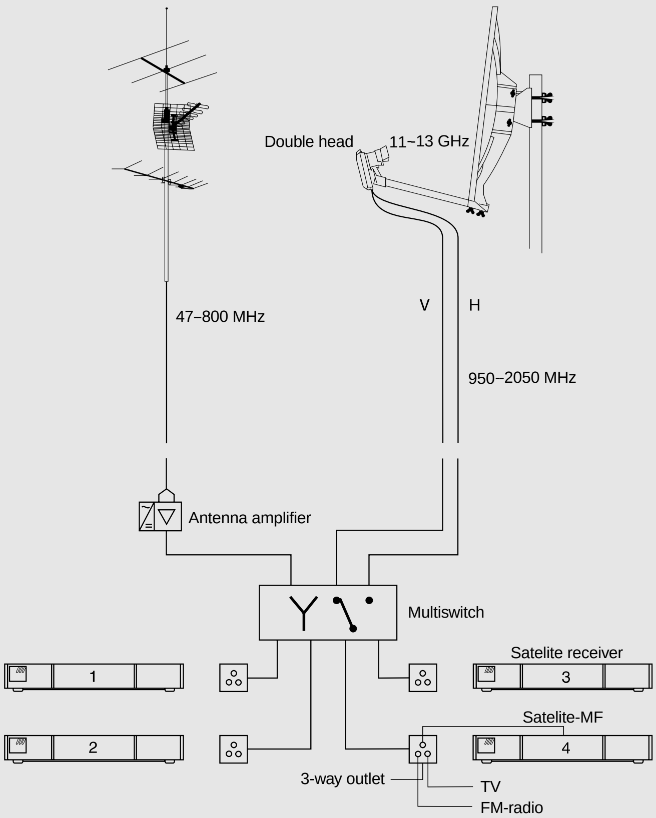
Lightning protection is necessary. The aerial mast should be earthed with a thick copper conductor to a good earth point, e.g. an earth spit pushed down 1.5–2 m into the ground. A transient protection should be connected between the receiver and the aerial.

Warning. Do not use silicone as protection against corrosion on the aerials, since it has an insulating effect.

TV and radio broadcast bands in Sweden

The TV and radio broadcast band is divided into 5 frequency bands which in turn are divided into a number of channels.

For technical reasons it is difficult to manufacture an aerial that operates efficiently over the whole of this frequency spectrum . However, there are so-called combination aerials, but they are a compromise and only operate satisfactorily if you live close to the TV transmitter. Even in these cases problems with shadow pictures especially on TV1 can occur.



System for terrestrial radio and television and satellite reception with distribution to several households.

Stacking of aerials

When mounting two or more aerials (for example UHF and VHF aerials) with a small distance between them, the respective radiation diagram will be affected. The appropriate stacking distance between two aerials of the same type depends on the boom length and the opening angle for the particular aerial model.

An easy way to find out the approximate distance is to set it to 2/3 of the boom length. A more accurate method, but still approximate for longaerials, is obtained using the formula:

$$\text{Distance} = \frac{\text{wavelength}}{2 \times \sin\left(\frac{\text{beamwidth}}{2}\right)} \times 0{,}8$$

where 0.8 equals the form factor for the aerial. The form factor can be obtained from the aerial manufacturer.

The aerial amplifier with common sense

If you have bad reception, don't assume that an aerial amplifier is the only and best solution. An aerial amplifier, no matter how good, cannot convert a bad signal to a good one. It can , however, compensate for the cable losses that occur between the aerial and the TV receiver, especially when many TV outlets have been installed.

If the reception is bad, start by checking your aerials and make sure they are not broken and rusty. If you have a so-called combination aerial, replace it with separate TV1 and TV2/4 aerials. If an aerial amplifier is still required, an outdoor amplifier should primarily be used and it should be mounted as close to the aerial as possible.

You must know which channel numbers are valid for the place where you live, in order to be able to order the correct TV-aerial (NOTE: Do not mix the channel numbers with the use the program companies' channel names). Below, you will find some tips. You can call Swedish Teracom for exact information.

Band plan for radio (FM) and television transmitters.

Band I	TV-channel 2–4	47–68 MHz
Band II	FM broadcasting	87.5–108 MHz
Band III	TV-channel 5–12	174–230 MHz
Band IV/V	TV-channel 21–69	470–854 MHz

The TV1 transmitters are normally within bands I and III.
TV2 and TV4 are always within bands IV/V.
There are also local slave transmitters that broadcast both TV1 and TV2/4 on the bands IV/V.

Radio communication

Radio communication is used in more and more applications: radio and TV from earth-bound or satellite distributing transmitters, communication radio for different applications, point-to-point communication for speech or data, navigation etc.

Modulation

The different techniques in *carrier wave modulation* (carrier wave modulation means that the signal is frequency translated to a higher frequency range) can be divided into a few main areas. You make a difference between *analogue* resp. *digital modulation*, and between *amplitude* and *frequency modulation.* (Alternative designations of the latter are linear resp. argument modulation.) The analogue mobile telephone systems use frequency modulation, FM, just like many communication radio systems. The digital modulation systems use commonly a combination of *ASK* (Amplitude Shift Keying), where the amplitude varies in one or several levels, and *PSK* (Phase Shift Keying), where the phase position is varied in certain positions.

Similar for AM, amplitude modulation, and FM is that they let the message, that is an analogue signal, affect a carrier wave. If we imagine the unmodulated carrier wave's voltage characterised as

$$s(t) = A \times \sin(\omega t + \Phi)$$

it means that at AM the amplitude, A, may vary with the message and at FM w may vary. At AM the same information will exist on "both sides of" the carrier wave's frequency in the frequency plan. You are then available to choose to transmit both (dual side band, *DSB*) or just one (single side band, *SSB*). You can choose at DSB also to transmit the carrier wave, as a reference signal. The only advantage with the latter is that detection can be carried out very easily, using an envelope detector. (An envelope detector is made up of a half-wave rectifier, a diode, followed by a low-pass filter, a RC circuit.) SSB needs a more narrow bandwidth than DSB.

The advantage with FM transmissions compared to AM is that a better signal-noise relation, S/N, is obtained after detection, which can be obtained at the expense of a wider bandwidth. Usually 200 kHz is used for transmission of speech and music.

An other type of argument modulation besides FM is *phase modulation* (PM). At PM f is varied. In practice it is often a combination of FM and PM because of measures for improved S/N, so called ''preemphasis'' and ''deemphasis''.

The alternative to carrier wave modulation is baseband modulation which means that the signal only is "adapted" to the transmission channel but still exists in the low frequency range, the baseband. An example of digital baseband modulation is PCM, pulse code modulation, which is used in e.g. digital telephone communication.

Listening

In Sweden, you can possess a radio receiver which can receive any frequencies. In some countries, however, you are only allowed to cover the frequency band dedicated for round radio transmissions for the public. See table. For bands other than the radio broadcasting bands in Sweden, you can listen to, but not transmit messages.

DX-ing, that is, listening to remote stations, is going through a renaissance. Many people want for example to be able to follow foreign newscasts, in which reports in many cases might differ significantly from the Swedish ones. The receiver should have large channel dispersion, good selectivity and good noise signal features. A good aerial is essential for a successful result. For more information, see section about *aerials.* Since the receivers have become smaller you can easily take them with you when you travel, and on holiday.

Radio broadcasting bands within long wave (LW), medium wave (MW) and short wave (SW).

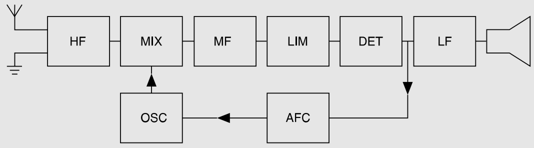
LW	148.5–283.5	kHz	25 m	11.65–12.05	MHz
MW	526.5–1606.5	kHz	22 m	13.6–13.8	MHz
74 m	3.95–4.0	MHz	19 m	15.1–15.6	MHz
49 m	5.95–6.2	MHz	16 m	17.55–17.9	MHz
41 m	7.1–7.3	MHz	13 m	21.45–21.85	MHz
31 m	9.5–9.9	MHz			

Scanner-listening, to listen to radio traffic on an automatic scanning receiver, can be both fun and exciting. This is allowed as long as nothing you have heard is forwarded. You can listen to local broadcasters such as police, fire brigade, taxi, but also long-range communication, for example, from aeroplanes.

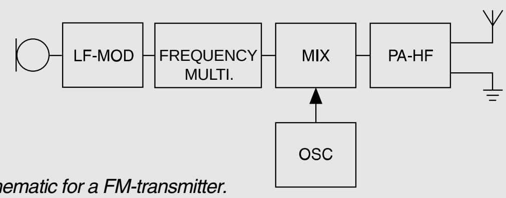
The receiver should have as wide a frequency spectrum as possible and a high scanning speed, however the large signal features normally become worse if the receiver has a very wide frequency spectrum. It should have two scanning modes, channel scanning and frequency scanning. Channel scanning means that it scans through the user programmed channels. Frequency scanning is continuously scanning all the channels between two specified frequencies.

Some interesting frequencies to listen to, e.g. for scanning.

1,5–2 MHz	Illegal cordless phones
46.5–47 MHz	Illegal cordless phones
79 MHz	Police, fire brigade, custom etc.
118–137 MHz	Air band (AM)
140–174 MHz	Miscellaneous communication
144–146 MHz	Amateur radio
155–162 MHz	Marine VHF
400–430 MHz	Miscellaneous communication
410–412 MHz	City police
432–438 MHz	Amateur radio
438–470 MHz	Miscellaneous communication
862–868 MHz	Cordless telephones
914–915 MHz	Cordless telephones



Block schematic for a FM (superheterodyne) receiver.
HF = high-frequency amplifier with filter, MIX = audio mixer, MF = medium-frequency amplifier, LIM = limiter, DET = detector, LF = low-frequency(audio) amplifier, AFC = feedback gor automatic frequency control, OSC = oscillator with frequency that changes for setting of desired reception frequency (station).



Block schematic for a FM-transmitter.

Advanced communication

As a private individual you can obtain a licence to possess and use your own transmitter under certain conditions.

Amateur radio certificate/license gives you permission to operate non-professional radio traffic for exercise purposes, communication and technical investigations if it is performed by personal interest and without any profit interests.

Amateur radio is an exciting hobby. Many radio-amateurs purchase a radio to talk to other broadcasting amateurs around the world, but many are also technically interested and are constantly trying to improve their equipment and skills. Amateur-radio stations have often technically higher performance than professional equipment.

There are different ways to contact other radio-amateurs and transmit messages to them.

CW Telegraphy. Covers a long distance with a low transmitting power, not noise-sensitive and also has a narrow bandwidth.

SSB Talk. Means Single Side Band and is the most efficient way to cover a long distance with speech transmission with low transmission power. SSB has a narrow bandwidth.

Packet radio. With the amateur packet radio system, data is sent in packages according to an error correcting data protocol called AX-25. Connection can be made to a station or to a mail-box, where information can be left or collected. To cover even longer distances, you can connect to one or many so-called digipeaters. In this way you can reach mail-boxes all over the world. **Packet cluster** is a so-called conference mode, where many people communicate simultaneously about real-time DX on the short-wave band.

RTTY, radioteletype, is also called remote print. It is a method used to transmit text. The method is old and well known. It is used for both two-way broadcasts between radio-amateurs and one-way broadcasts. e.g. news bulletins. Most larger news agencies broadcast the news on the KV-band, which can easily be listened to.

AMTOR is the radio-amateurs' name for error correcting RTTY. Other names that are used commercially are ARQ, FEC, Navtex and others.

Facsimile, fax pictures. Broadcast on a short-wave and via satellite, which include weather maps, press photos and marine information amongst others. These pictures can be received with the correct equipment.

SSTV, Slow Scan Television. TV-broadcasting with short-wave is a narrow but interesting branch of the amateur-radio hobby. The method builds upon broadcasting the picture line by line to the station.

Amateur radio bands in Sweden.

160 m	1.81–1.850;	MHz	10 m	28.0–29.7	MHz
	1.93–2.00	MHz	70 cm	432–438	MHz
80 m	3.5–3.8	MHz	23 cm	1240–1300	MHz
40 m	7.0–7.1	MHz	13 cm	2.3–2.45	GHz
30 m	10.1–10.15	MHz	6 cm	5.65–5.85	GHz
20 m	14.0–14.35	MHz	3 cm	10–10.5	GHz
17 m	18.068–18.168	MHz			
15 m	21.0–21.45	MHz	and different bands		
12 m	24.89–24.99	MHz	up to	250	GHz

The private radio band on 27 MHz is also called the citizen band (CB). The private radio band is used, for example, for communication between leisure-boats, in hunting teams, between cars etc. The band contains 40 channels. Today the stations operate with frequency modulation (FM). The station should be approved and T-marked. No licence is required.

Older stations that have AM can no longer be used.

Marine VHF consists of 60 channels within the frequency range 155.5–157.4 MHz. Max allowed output power from the transmitter is 25 W. The station can be used internationally by the licence holder. Here you can also call a coastal radio station that can transfer the connection to the public telephone network.

The connections have generally less noise than the private radio band, partly due to reduced risk of noise from foreign stations and partly due to better traffic discipline on the band. To use the marine VHF you must have at least a limited radio telegraph operator's certificate, so-called D-certificate. The Post & Telestyrelse can provide information about the test and its contents. The station must be approved and T-marked.

Companies and institutions can also obtain a certificate to use **communication radio** for point-to-point connections within for example the bands around 150 and 430 MHz.

The application is made to the Post & Telestyrelsen, which, after frequency planning, distributes certificates and frequencies.

Air band 118–135 MHz used by both private aeroplanes and for domestic and international flights, primarily for take off and landing directions.

Data transmission via radio occurs commercially as well as amongst broadcasting amateurs.

In the last few years there has been a tremendous increase in the usage of **mobile telephones**. This took place approxiamtely at the same time as the digital systems were introduced.

The original analogue system, NMT (Nordisk MobilTelefoni), is still accessible in Sweden but is not often used nowadays. Initially NMT operated at 450 MHz, but when the demand for higher capacity emerged they started to use the 900 MHz frequency band, where it shared space with the digital system GSM.

Groupe Spécial Mobil was the original meaning of GSM, though later the decision was made to keep the initials but to change meaning to Global System for Mobile communication. It was a European Union project, while USA used D-AMPS (Digital Advanced Mobile Phone System).

A several different systems has evolved with different geographical spread. The idea of a common system resulted in, what we call it today, the third-generation mobile phone technology. Its official term is UMTS, Universal Mobile Telephony System. Unfortunately, it has resulted in some different technical variations, where ETSI (European Telecommunications Standards Institute) has accepted two standards, partly TD-CDMA which is a time multiplexed system, partly a modulation/access-method called WCDMA, (Wideband Code Division Multiple Access). WCDMA has also been accepted as a standard in Japan and, meaning technically, that conversations are 'spread out' over a larger frequency range. This frequency range is shared by all ongoing conversations. A digital code is attached to each conversation and used at both spreading and detection. It has not, as with GSM, an absolute highest number of simultaneous conversations limit but in practice it means that the more ongoing conversations the more they interference each other, which results in a gradually lower quality (and/or lower transmission rate). Each conversation, or rather communication channel, can vary its transmission rate as needed.

The new systems uses frequencies around 2 GHz. This means that the range is shorter than for GSM, which can be an advantage in the cities, where the cells are small and the base stations must stand tight because of conversation density. On the other hand, it is a disadvantage out at the countryside as it requires so many base stations. There are more advantages with the 450 system MHz for sparsely populated areas.

There is a lot that distinguish the third generation system from the previous ones. Among other things there is a highly advanced power regulation system which changes the strength of the transmission up to 1600 times per second (for WCDMA) with the intention to not to transmit unnecessary powerful signals. It has advantages such as longer battery life and unnecessary radiation reduction. It is also necessary that all phones transmit about the same intensity so as not to interfere with the other users. Further, there is a larger transmission capacity available for 'phones', up to 2 Mbps, which makes it possible to transfer moving pictures using MMS (MultiMedia Services), connect to the Internet etc. You can use your 'phone' as a handheld computer, or your handheld computer/laptop as a phone. The concept of 'phone' begins to feel incomplete; instead phone should be translated to 'remote sound'. More like a mobile communication unit.

Interference suppression

In some cases noises occur after the transmitter has been installed. This happens for amateur radio broadcasting as well as private radio and mobile phones. There are two reasons for the noise occurring:

- The transmitter sends overtones or false frequencies. In this case the transmitter must be treated, e.g. be provided with a filter and possibly be shielded. The impedance adjustment between the transmitter and aerial, and also the earthing are important. The adjustment between the asymmetrical coaxial cable and a symmetrical feed aerial e.g. dipole, must be done with a balun transformer. Allow the coaxial cable to pass through some large ferrite cores or wind the cable several times in a ring. This reduces the risk of radiation from the aerial's down-lead.
- The radiated field has such a strength that it can be detected in e.g. an amplifier or tape-recorder, in a TV, a video-player or a hearing-aid. The fault must be remedied in the equipment which is being interrupted, not in the transmitter.

In the first case the fault is treated by the transmitter owner. In the second case, where the equipment being interrupted must be adapted, both of the parties must co-operate. If the noise occurs due to amateur-radio broadcasting, SSA can contribute with help from its noise official. The Post & Telestyrelsen investigates upon reports of radio and TV disturbance.

Tools and production aids

ELFA's aim is to have a range of tools and production aids that are functional, rational and ergonomically designed.

Tweezers are manufactured from chrome-plated steel as standard. They are also available in stainless steel (non-magnetic). In some sensitive environments such as clean-rooms, you can only use ceramic tweezers, which do not cause any metal-chips. They are chemically resistant and do not corrode. A further advantage is that the material is a good insulator, which together with other treatments can provide ESD-protection. See section *Electrostatic discharge*.

Pliers should be easily gripped, have insulated shanks and good precision in bearing and jaws. Cutters for larger wires should have bevelled edges. This provides however, a shock effect that might destroy components such as reed elements in glass tubes. For this purpose, non-bevelled pliers should be used. Make sure that you turn the flat side towards the sensitive component.

The non-bevelled oblique cutter with a spacer is a special tool, which leaves a bit of the wire for soldering.

A **microscope** makes it possible to accurately inspect small circuits. Heat-sensitive components are lighted up with cold-light, which is delivered to the object via a fibre conductor. Ring-lighting prevents shadows. The lens should have zoom. For some microscopes, one ocular can be screened off and the picture can be transferred to a TV-monitor via a video camera. It is also possible to connect a video printer for documentation. Another alternative is to connect a still picture camera.

Screwdrivers

Screwdrivers are available in different designs to suit different kinds of applications and requirements, e.g.:

- *High voltage resistance*, with insulated blade, to handle e.g. 1000 V.
- *Blades for different types of screws*, e.g. flat blade, hexagon, Phillips recessed, Pozidrive recessed and Torx. Choose your recessed driver carefully, since the angle on the driver differs from type to type. Torx is able to handle the highest rotating torque at a specific diameter. Torx is therefore becoming more common.
- *Loose blades*, also called "bits", gives you a very flexible system which allows you to replace a handle with a motor drive. As far as hand tools are concerned, the fixed ones are normally more stable and easier to use, and therefore preferable for the sizes you use often.
- A *screw holder* can be useful when assembling in narrow places. Sometimes a magnetic screwdriver can be helpful, but at other times it might be disastrous (e.g. when servicing tape-recorders or memory components). However, you can use a combined magnetiser/unmagnetiser.
- *Trimming drivers* for high frequency circuits must be non-magnetic. Therefore they are normally manufactured from plastic. For UHF and higher frequencies, the material must have a low dielectricity constant to prevent it interfering in the HF-fields.
  Ceramic drivers are very hard compared to plastic drivers, but also considerably more expensive. It is very important to chose a driver that fits exactly. An iron powder core can otherwise easily be damaged.
- The *ergonomics* are important. This means that the handle must be comfortable to hold and designed so that you get a good grip with your hand. Miniature screwdriver should preferably have rotating tops. A crank handle for screwdriver blades can simplify and speed up the assembly work.

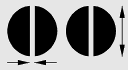
Choice of screwdriver

To facilitate the choice of a screwdriver we have created a table for the most common screw dimensions.

This is what you should do:
Measure the dimensions of the screw and then look up the suitable blade width and slot width for the screwdriver.

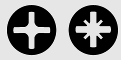
Example. The screw is 3.0 mm (M3) in diameter. You should then use a slotted driver with dim 4.0 × 0.80 mm.

Slotted screws



Blade width × Slot width mm	Screw size	Max. tightening torque in Nm Manual force	Machine
0.8 × 0.16	(Special)	–	–
1.0 × 0.18	(Special)	–	–
1.5 × 0.25	M1	–	–
1.8 × 0.30	M1.2	–	–
2.0 × 0.40	M1.6	0.40	0.42
2.5 × 0.40	M1.8	0.40	0.42
3.0 × 0.50	M2	0.70	0.80
3.5 × 0.60	M2.5	1.3	1.4
4.0 × 0.80	M3	2.6	2.9
5.5 × 1.0	M3.5	5.5	6.2
6.5 × 1.2	M4	9.4	10.5
8.0 × 1.2	M5	11.5	12.9
10.0 × 1.6	M6	25.6	28.7
12.0 × 2.0	M8	48.0	53.8

Cross recess
Phillips (PH) and Pozidriv (PZ)



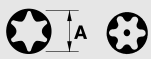
PH PZ

Size PH + PZ	Screw size	Max. tightening torque in Nm Manual force	Machine
00	(Special)	–	–
0	M1.6-M1.8	1	2
1	M2-M3	4	5
2	M3.5-M5	10	14
3	M6	20	42
4	M8	30	60

Hexagon

Dim. A mm	Screw size	Max. tightening torque in Nm Manual force
0.7	(Special)	0.08
0.9	(Special)	0.18
1.3	M1.4	0.53
1.5	M1.6-M2	0.82
2.0	2.5	1.9
2.5	M3	3.8
3.0	M4	6.6
4.0	M5	16
5.0	M6	30
6.0	M8	52
7.0	M10	78
8.0	M10	120
10.0	M12	220

Torx



Key no.	Dim. A mm	Screw size	Max. tightening torque in Nm Manual force	Machine
T5	1.17	(Special)	–	–
T4	1.28	(Special)	–	–
T5	1.42	M1.6	0.43	0.5
T6	1.70	M2	0.75	0.9
T7	1.99	M2.5	1.4	1.7
T8	2.31	M2.5	2.2	2.6
T9	2.50	M3	2.8	3.4
T10	2.74	**M3**-M3.5	3.7	4.5
T15	3.27	M3.5-M4	6.4	7.7
T20	3.86	**M4**-M5	10.5	12.7
T25	4.43	M4.5-**M5**	15.9	19.0
T27	4.99	M4.5-M6	22.5	26.9
T30	5.52	**M6**-M7	31.1	37.4
T40	6.65	M7-**M8**	54.1	65.1
T45	7.82	M8-M10	86.2	104
T50	8.83	**M10**	132	159
T55	11.22	**M12**	252	257
T60	13.25	M14	437	446

Bold text indicates the most common sizes.

All calculations are for metric screws.

Crimping

Crimping is a method used to create an electrical connection by permanently pressing a socket terminal around a conductor. Using a special tool the socket terminal is crimped into a pemanent deformation so that good electrical and mechanical functioning is achieved. Crimping tool and terminal must fit the conductor. Today's crimping tools are provided with a lock system, which makes the crimping process complete, and they are provided with gearing for low handle force

Comparison with other technologies

The method arose as an alternative primarily to soldering. The crimping technique is widely used within industry. Its popularity is based on the fact that the quality of the work is dependent not on the operator, but on the quality of the crimping tool. The features of the crimped connection very much depend on how much reduction of the sleeve/collar the crimping tool provides. This makes extensive demands on the crimping tool and its precision.

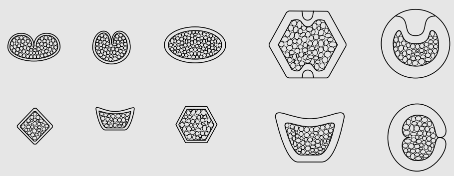
However there are alternatives within different fields of application: Thermal methods like *soldering* and *welding*, or mechanical methods like *terminal connection, wire wrapping, IDC connection* etc. Most of these methods have restrictions, which makes them less suitable for a wide range of applications. Crimping has become predominant especially for power cables.

Crimping has many advantages:

- speed
- reliability
- simplicity and easy accessibility
- low piece cost
- no heat
- no chemicals
- established feature standards
- easy control
- very large application areas

Different types of crimping

Depending on the conductor material, connection construction and usage requirements, many different forms and designs of crimping tools are available. A large amount of work has been done on testing designs and components. Make sure that you take advantage of this.



Example of cross sections for different types of crimping.

Coaxial connections

Crimping within coaxial cables is basically controlled by standardisation e.g. MIL, which specifies the measurements of the cable and connectors. Because of this, there are clear instructions about the dimensions for the crimping tool sockets. However within this area, it is particularly important that the crimping is performed correctly. Because the use of coaxial means working with low voltages and currents, even a relatively small mistake during the crimping could result in high transition resistance, which leads to faults in the coax-system. For this reason, make sure that you only use quality tools.

ESD

ESD stands for Electro Static Discharge. As early as the sixties it was discovered that the MOS-transistor was sensitive to ESD. Since then several different types of semiconductors have been developed, with even thinner conductors and shorter insulation distance between the conductors. This has resulted in an increasing sensitivity to ESD damage.

The types of damage caused by ESD can be divided into two groups. Either the components stop operating directly upon discharging or a latent fault occurs. The former type of damage is relatively easy to locate, whilst a latent fault means that the component's conductor-tracks are damaged without the fucntion stopping immediately. The result can be that the component acquires unwanted properties when functioning or operates intermittently. ESD-damaged components are costly both in terms of production and service.

The component's sensitivity

In the table below the most common components with their sensitivity level are specified. The values are general but still provide good guidelines for the respective component types.

Static sensitivity.

IC type	Level in V
MOS-FET	100 – 200
J-FET	140 – 10.000
C-MOS	250 – 2.000
Schottky-TTL	300 – 2.000
Bipolar transistors	380 – 1.500
ECL, board mounted	500 –
SCR	600 – 1.000

Electro static charge in everyday situations

Situation	Level in V, <20 % RH	Level in V, >65 % RH
Person walking on carpet	35.000	1.500
Person walking on PVC floor	12.000	250
Person at work-bench	6.000	100
Plastic folder for work order (PVC)	7.000	600
DIP in plastic box	12.000	3.500

Electro static charge

Static electricity occurs via contact, friction or separation of material. Static electricity always exists in our surroundings, in working areas, floors, chairs, clothes, packing material, paper or plastic folders. A person crossing a floor or working at a bench can build up a static charge of many thousand volts. The values in the table below show that in daily situations the voltages can reach levels which are a high risk factor for the components.

Measures to prevent static electricity

There is a basic rule for preventing ESD damage: **Avoid charging!** It runs like a red thread within ESD protection. By following the three guidelines below, you can obtain effective protection against ESD damage.

- Handle all ESD-sensitive products in an ESD-proof area.
- Transport all ESD-sensitive components in shielded boxes or packages.
- Check and test all ESD-protection to ensure correct function and quality.

ESD-safe area

An ESD-protected working place can be designed as follows. On the floor a *conductive floor mat* is placed, which has a connection to earth. As soon as a person approaches the working place and steps upon the mat, the static electricity is discharged via the mat. This means that products within the working area are protected against electrostatic discharges from temporary visitors. The work bench is provided with an *anti-static bench-top* which connects to the floor mat. The person who works at the bench is connected to the work place earth system via a *wrist band* to avoid any risks to the components. If insulated material that is not discharging cannot be avoided in the working place, an *ionising fan* should be used, which blows ionised air over the working area. This neutralises the stored charge within the insulated material and the risk of static discharge is eliminated.

ESD-protective packages

It is important here to distinguish between metallic, conductive and anti-static material. The materials have different electrical features and thereby different application areas.

Metallic and shielded materials have a metal layer that forms a Faraday's cage and prevents charges and electrical fields from penetrating. The metal layer in the bags is normally made of nickel or aluminium and provides very good protection for the components.

Conductive materials are conductive. Normally carbon is mixed with the plastic material to obtain desired features. The material is ageing-resistant and its conductive features are constant even when air humidity is low. The material's conductive features make it suitable for the manufacture of boxes where both mechanical and shielded protection is desired. The material's shielding features depend on its thickness. Therefore, bags made of conductive material have a limited shielding capacity compared to the metallic bags.

Low charge material (also called anti tribo electrical) is normally a chemically treated plastic. These materials do not have any shielding features, but the treatment gives the material a limited charging capacity against itself and other materials. The treatment makes the bags age and bags of low charge material should be considered as perishables. The material is only recommended as packing material for components that are not ESD-sensitive.

Control and maintenance

To obtain high quality ESD protection, all ESD material should be checked to ensure correct function and quality. Test instruments are available for checking wrist bands, benches, shoes and floors, and a measuring instrument for static electricity is also available. The equipment should be checked regularly. Wrist bands should be tested every day, whilst other work place equipment should be tested each month.

Handling chemicals

There are various laws aimed at protecting people and the environment against different chemicals. Some of these laws include:

- The Environment Act.
- The Act on Transporting Hazardous Goods.
- The Working Environment Act, where chemical risks are one of several risk areas.

The Environment Act

The Enviroment Act aims at promoting a durable line of progress ensuring that current and coming generations will be able to enjoy a healthy and good environment. This act lays down the basic rules regarding the import and handling of chemical substances and preparations.

Anyone who imports or manufactures a chemical product is responsible for carrying out internal investigations with regard to what health and environmental problems the product can cause. Products should be classified as to how dangerous they are, using the following categories:

Designation	Category	Symbol no.
T+	Extremely toxic	1
T	Toxic	1
C	Corrosive	2
Xn	Harmful to health	3
Xi	Irritating	3
N	Moderately harmful	4
F+	Extremely flammable	5
F	Very flammable	5
E	Explosive	6
O	Oxidizing	7

The Act on Transporting Hazardous Goods

This law contains provisions in which terms are defined such as:

- Transport
- Hazardous goods
- Means of transport

In the regulations regarding the transport of hazardous goods, chemicals are divided into various classes.

Class 1	Explosive substances and objects.
Class 2	Gases that are compressed, condensed or dissolved under pressure.
Class 3	Flammable fluids.
Class 4.1	Flammable solid substances.
Class 4.2	Self-igniting substances.
Class 4.3	Substances which generate flammable gas on contact with water.
Class 5.1	Oxidising substances.
Class 5.2	Organic peroxides.
Class 6.1	Toxic substances.
Class 6.2	Noxious substances and substances with a tendency to cause infections.
Class 7	Radioactive substances.
Class 8	Corrosive substances.
Class 9	Other hazardous substances and objects.

Hazardous goods is a concept comprising substances and products with properties that may be harmful to humans, the environment, property and other goods if they are not handled correctly during transport.

There are a number of regulations controlling the handling, marking and transport of hazardous goods.

Transport information

Hazardous goods is marked in the catalogue with this symbol:

Please contact ELFA or your local reseller for information about transport times, transportation costs and possible smallest order quantity. We can not deliver hazardous products to countries where ELFA is not represented by a subsidiary or reseller.

The Working Environment Act

Basic rules for protecting against health risks and other harmful effects in working life. This law applies to all commercial handling of substances which are flammable, explosive, corrosive, toxic or in some other way harmful to health.

Bonding

Bonding as a method of joining materials has a wide range of application; we bond everything from toys to advanced constructions in the aerospace industry.

Adhesives can be divided into 3 main groups:

- Curing adhesive, e.g. 2-component epoxy and cyanoacrylate.
- Drying adhesive, where the solvent or water evaporates, e.g. contact adhesive.
- Melt adhesive, e.g. hotmelt adhesive.

Curing adhesive

Epoxy adhesive of the 2-component type comprises a base and a hardener. It is extremely important that the correct mixture ratio is obtained. An incorrect mixture ratio produces an inferior joint, and in the worst case scenario curing will not take place at all.

These adhesives can have widely varying curing times, and may also require other conditions such as curing temperature.

Cyanoacrylate adhesive cures as a result of the influence of moisture. Cures extremely quickly, and is often called second adhesive. Produces hard joints, but usually poor values with regard to mechanical stresses. Cyanoacrylate adhesive with added rubber is available which has significantly better values.

Silicone adhesive produces soft, elastic joints. Available as both 1 and 2-component adhesive. 1-component adhesive cures with the aid of moisture, which means that the adhesive cannot be used in closed spaces or in joints thicker than 5–6 mm.

2-component adhesive is available with several different curing systems. The most suitable system for electronics is the oxime type. It produces no corrosive products and hardly shrinks at all. In order to achieve good adhesion to glass, acetic acid-curing adhesives should be used. These are not suitable for electronics.

Drying adhesives

Drying adhesives comprise plastics or elastomers dissolved in a solvent or water. During evaporation, the adhesive becomes increasingly viscous, eventually becoming a solid material.

When using contact adhesive, part of the solvent should be allowed to evaporate before bringing the different parts together.

Melt adhesives

Melt adhesives are normally thermoplastics which are melted and then allowed tcool off. Melt adhesives solidify rapidly, which can be an advantage e.g. in production.

General

In order to achieve good adhesion and a strong joint, it is important to have a suitable surface against which to bond.

Different adhesives have different requirements as to the surface to be bonded, although in general the surface should be clean and dry.

Some materials may need to be etched or activated. This applies primarily to ’oily’ plastics, e.g. PTFE or polyethylene.

Preparation

Before starting bonding, you should think through what stresses and strains the bonded joint will have to handle. For example:

- temperatures
- mechanical stresses
- water or other solvents
- will the joint be painted?

Soldering

Soldering is a method of joining that has been used for thousands of years for the production of, for example, ornaments and weapons. It is however over the last 50 years that soldering has become a meaningful method of attachment for the electronics industry. The development of soldering gained speed during the second world war when new methods were sought that could be suitable for mass production. In recent years, the theoretical understanding of soldering has increased quickly and we have been able to map out the way binding occurs between solder and base metal. As a result, it has been possible to develop new solder for different purposes, which has increased the areas of application within industry. We mainly describe here the processes that occur during soft soldering and also go through the basic knowledge needed to ensure a good result.

Soldering involves the joining together of metals of the same or different type with a metallic binder, solder, which has a lower melting point than the metal parts that are to be joined. The solder spreads out between the metal parts that are being joined through capillary action.

Soldering and brazing

Usually the term soldering is used when the solder temperature is lower than 450 °C and the term brazing is used when you work with an open flame generated by, for example, a gas burner or welding equipment. The gas burner uses butane gas and the flame reaches a temperature of around 1300 – 1500 °C. With welding equipment the temperature reaches approximately 2700 °C. The welding equipment consists of separate gas and oxygen bottles. The solder has a high melting point compared with soft soldering.

From a metallurgical viewpoint there is no difference in principle between soldering and brazing. When soldering cables, components and semiconductors, the objective always is to obtain a good joint with good metallic contact with low electrical resistance. There should be no or little mechanical strain in the joint.

The solder

Spelter solder is available in many different forms, e.g. wire, rod, film and paste.

For soldering you use solder containing an alloy of tin and lead in different proportions and with different melting points. An alloy with 63% tin and 37% lead is called a eutectic alloy and has a melting point of 183°C. The advantages of a eutectic alloy are that it has a low melting point and short melting range. When you heat the solder, it first enters a plastic half liquid state and then changes into liquid. The temperature difference between the solid and liquid states is called the melting range. When in the melting range, you must not expose the solder joint to vibration, as you then obtain a brittle joint of low strength and poor electrical conductivity.

Sometimes it is necessary to change the properties of the solder and this is done by alloying tin and lead with other substances.

- Copper, increases the life of soldering tips if approximately 2% is included in the solder alloy.
- Silver, used for soldering silver plated components so as to prevent the silver from being leached out of the silver coating. 2% is usually added to achieve this effect.
- Bismuth, cadmium and indium are used to lower the melting point of the solder.

Environmental care

Because lead is an undesirable metal when it comes to environmental care, we should of course try to avoid it. There is an alternative lead-free soldering wire on the market today. The disadvantage with it is the somewhat higher melting point of 217–227°C. Otherwise it has about the same strength, see below.

Solder strength

With soldering the aim is to achieve good electrical contact, but you also want a soldered joint with a certain durability. Maximum strength is obtained when the gap width is between 0.05–0.25 mm. The tension that exists in the surrounding area between solder and base metal is then the most favourable. The reduced strength with gap widths below 0.05 mm is because irregularities in the surface prevent the gap from being completely filled. Strength diminishes with time, and the soldered joint reaches its final strength, approximately 75% of its original strength, after about a year. Because spelter solder has a relatively low working temperature, the strength deteriorates rapidly with rising temperature. It also decreases quickly with the time under load, because spelter solder has a tendency to flow if it is put under load for a long period.

Flux

Because most metals oxidise quickly, it is necessary to add "flux", which removes existing oxidation and also prevents re-oxidation.

The main purpose of the flux is to make the soldering process itself possible and by means of its special properties to improve the solderability of the components. Flux must therefore fulfill the following requirements:

- Loosen the existing oxide layer and prevent new oxidation during the soldering process.
- Withstand heating up to the soldering temperature without vaporising.
- Let itself be displaced by the melted solder without leaving any dross or gas pockets.
- Not have an unfavourable effect on the metal or degrade the electrical properties of the joint.

There have been many theories that try to explain the technology behind the way that the flux works and some of these have been useful in developing new fluxes. The most common view is that the flux removes the oxide film from the metal and the solder, and dissolves or loosens the film and lets it disperse into the body of the flux. The melted flux also builds a protective shell around the metal, which prevents a new oxide film from forming.

The flux can be solid or liquid and can be applied in different ways. The most common method in soft soldering is to use solder wire containing cores of flux.

Usually flux is divided into different groups depending upon the admixture of activating agent.

Non-active flux

Resin dissolved in alcohol without addition of activating agent.

Weakly active flux

Resin dissolved in alcohol with small additions of amines/amides or halogens, usually chlorine. The amount of activating agent must not exceed 0.5%. The residue from weakly active flux does not cause corrosion and is not electrically conductive.

Strongly active flux

Resin dissolved in alcohol activated with more than 0.5% of halogens or other oxide loosening substances that give the same effect. The amount of halogens added is described in a number of standards specified by the manufacturer, e.g. BS 441 type 1 D.T.D 599 A etc. A common characteristic of strongly active fluxes is that they attract moisture from the atmosphere and therefore become conductive causing small insulation faults. They are also weakly electrically conductive and must be thoroughly dispersed by washing. Strongly active flux may not be used for soft soldering of defence materials.

Because wood rosin flux (flux with additives of wood rosin) gives rise to formaldehyde fumes, which can cause allergy problems, there is a synthetically produced flux that can be used instead.

It is becoming more common to use flux with lower solids content. Instead of the traditional wood-rosin solids content of up to 20%, you can obtain flux with a solids content of around 2–3 %. This gives insignificant flux residue after soldering.

The metal surface

If you look at a polished metal surface with enough magnification, you will discover that it resembles a rocky landscape with bumps and cavities. The outermost atoms of the metal surface have a capacity for attracting oxygen atoms from the surrounding air and reacting with them to form an oxide film over the metal. Because the air has a certain degree of humidity, water vapour is formed on the oxide layer, and there is almost always, on this skin, a layer of fat and dirt. This increases the surface tension on the metal, which leads to inferior solderability.

Wetting

Wetting is an expression that often occurs in soft soldering. Wetting depends on the surface tension of the metal surface that is to be joined. When soldering leaves an even permanent layer on the metal surface, it means that the solder has wetted well. Without wetting, there is no soldering effect. To obtain wetting, there must be a stronger attraction between the atoms in the solder and metal than between the atoms in the solder itself. This requires low surface tension and good flux effect. Experience shows that a joint that displays good wetting is a good soldering. Solderings with good wetting are therefore easy to check and give rise to low inspection costs.

Capillary action

Capillary action plays an important part in soldering. Every correctly executed soldering is based on the principle that the melted solder must be forced into the gap between the metal surfaces. If you place two plates with a gap between them into a liquid, then the latter is drawn up along the edges of the plates. This is because of the liquid's capacity to wet the metal. Good wetting causes the liquid to rise between the plates, while bad wetting can cause it to sink. The better the capillary action, the better the soldering gap is filled with solder.

Soldering work

Preparation and tin coating

In order to achieve satisfactory soldering it is important that the joint is dressed and thoroughly cleaned. The surface dressing, which usually consists of tin, silver or gold, must have good solderability characteristics. Solderability deteriorates with time, and a joint that is oxidised or contaminated in some other way must be cleaned before soldering. When tin coating with a soldering iron, the solder must be applied to the conductor and not the soldering tip so that the flux is not vaporised immediately.

Soldering and mounting

Before soldering, ensure that the soldering tip is free from impurities and that it has molten solder on the surface, which causes more rapid transfer of the heat Hold the solder on the surface of the soldering tip against the part of the joint that has the largest mass. This causes the whole of the joint to be heated up properly and minimises the risk of cold soldering. Then apply the solder to the heated joint and not the soldering tip. If you put the solder directly onto the heated tip, the flux is vaporised before reaching the joint. The amount of solder applied should not be more than is required to cover the surface of the joint with a thin layer. When the solder flows, stop the heating immediately. This prevents the solder from flowing outside of the joint.

The soldered components must not be placed under mechanical load or displaced out of position until the solder has safely hardened. Otherwise the soldering takes on a crystalline grainy and gray appearance, which is the sign of a deficient joint. Components and conductors should if possible not be held in place with pliers or tweezers as these can easily magnify the shaking of the hand.

When soldering into solder cups, a somewhat different technique is used. Multipole connectors having solder cups with sealed bottoms are soldered by half-filling each terminal with solder. The terminal is heated up until the solder is molten, the conductor is inserted into the terminal until it butts up against the bottom and is then held in place until the solder hardens. This method requires a well-tinned and dried soldering tip.

Soldering of PCBs

When soldering on a PCB it is important to choose:

- The correct temperature.
- A soldering iron with the right power rating.
- The right size of solder wire.

Repeated soldering on single-sided boards requires a soldering iron with a power rating of at least 40 W. The temperature varies between 300 and 350 °C depending on the skill of the operator. A higher tip temperature puts greater demands on the operator, but gives a shorter soldering time and less spreading of the heat.

When soldering through-plated double-sided boards, the same applies as for single-sided boards except that the temperature should not fall below 350 °C. This is because PCBs that are plated have a larger mass to warm up. You have to quickly heat up the joint to be soldered to the right temperature and then immediately apply the right amount of solder. If the solder is applied over too long a period or with too cautious an action, the flux vaporises and an inferior soldering results.

Soldering of a multi-layer board uses the same method of approach as for soldering of a through-plated double-sided board. However, the soldering time is somewhat longer because of the greater mass to be heated up.

An important consideration when soldering on plated boards is to ensure that the solder wets around the component pins on both sides of the board. This prevents oxides from coming in between the solder and the joint. When soldering on a PCB, it can also be a good idea to ensure that the component pins are clipped to the right length before the soldering work begins. If this is done after soldering, the soldered joint is exposed to the mechanical shock effects that occur when the component pins are clipped with diagonal cutting pliers.

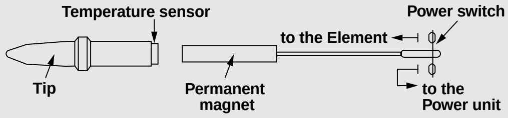
Soldering tools

Soldering station or soldering iron

The choice of soldering station (soldering pencil with transformer) or soldering iron (for connection directly to the mains) depends on the nature of the soldering work. The soldering station is preferred tool for industrial soldering on production contexts. A mains connected solder tool is more suitable from a handling point of view for service technicians working in the field but also for the hobbyist. Soldering stations and the more advanced mains connected soldering tools have automatic regulation of tip temperature. There are different techniques for regulation of the tip temperature, e.g. those we call mechanical or electronic.

Mechanical temperature regulation

The soldering iron consists of a permanent magnet that controls a sheet metal shield. At the end of the soldering tip sits a bit made of an alloy that at a certain temperature, the *Curie point*, becomes demagnetised. The magnet drops out and the electric circuit to the heating element is broken. When the temperature falls, the metal becomes magnetic again and the circuit is closed. The temperature at which this occurs varies depending on the composition of the bit.



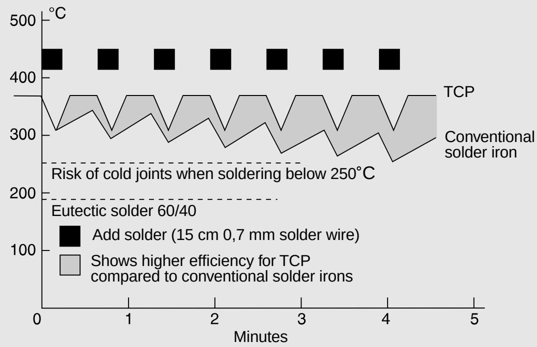
Soldering iron with mechanical temperature regulation.

Electronic temperature regulation

The tip temperature is varied electronically. In the soldering iron there is a sensor with inbuilt NTC or PTC resistance. The sensor butts up against the soldering tip and senses the temperature. Stepless regulation of the temperature is achieved by means of a potentiometer in the power supply unit.

The advantage of electronic temperature regulation versus mechanical is that, if there is a need for different tip temperatures during soldering, the adjustment is quicker and simpler with an electronic regulator.

The advantage of mechanical regulation is that it is not possible to change the temperature without changing the soldering tip. This can be an asset during, for example, the soldering of mass-produced products, when you will want to make it difficult for the operator to change the temperature.



Variations in temperature with/without temperature regulation.

Soldering iron tips

There are two different manufacturing methods for soldering tips – non-plated and plated. Basically, both types of tip are made of copper, which has good thermal conduction properties. The plated tips are surface treated in different processes in order to attain a high quality finish. The non-plated variety is not treated in this way.

Non-plated tips oxidise quickly, change shape quickly (the copper is 'eaten away'), have a short life and good thermal conductivity.

Plated tips have a long life, are easy to keep clean and have relatively good thermal conductivity.

The choice of soldering tip depends on the type and accessibility of the joint to be soldered.

Solder fumes

The fumes emitted during soldering contain formaldehyde, which can give rise to allergical difficulties when inhaled and should therefore be evacuated. However, the extraction should not be so strong as to have too much affect on the tip temperature. Extraction can be carried out in several different ways. One solution is to suck up the solder fumes directly at source, i.e. at the soldering tip. The fumes are transported through a tube and pipe system to a unit where the dangerous particles are filtered out.

Desoldering

Desoldering of components can take place in several different ways:

- Manual desoldering tool used with soldering pencil.
- Specially manufactured tips mounted directly on soldering pencil.
- Soldering braid, which is laid onto the soldering point and then applied warm via soldering pencil.
- Desoldering station with built-in vacuum. The station has a desoldering pencil to heat up the solder with. The vacuum is activated and the solder is sucked up into a container. Desoldering stations have different desoldering tips. The choice of desoldering tip depends on the diameter of the component's terminals and the solder.

Finishing

Cleaning after soldering

Cleaning of the PCB after soldering is often done for reasons of appearance but is also necessary when there is a risk of corrosion, e.g. during long-term storage, or when the board is likely to be subjected to extreme environmental conditions.

Cleaning may also be necessary to fulfil predetermined standards.

Because there are so many different fluxes on the market, it is also advisable to consult the flux manufacturer about suitable cleaning substances. Account must be taken of fire prevention regulations, chemical discharge regulations, health considerations, etc.

Protective coatings

The board is coated with lacquer to protect it from environmental effects.The protective lacquer is applied to the finished board. Certain parts may need to be masked before the lacquer is applied, e.g. connectors and potentiometers. This is done with solder masking latex or tape.

It is important to plan up-front for protective lacquering so that the design of the board is suitable for laquering.

All lacquers contain a solvent of some sorts and it is important that this solvent is compatible with the components and is acceptable from a health viewpoint.

Surface mounting

Surface mounting requires special tools, e.g. vacuum-driven picking pliers, rotating table for handling loose components, mounting station and hot plate for surface soldering or for hardening adhesive, etc.

The advantages of surface mount vs hole mount:

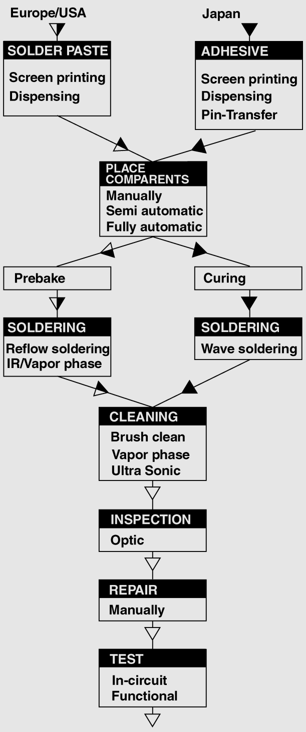
- Miniaturisation with up to 70% saving on PCB surface.
- Lower weight despite more components and functions in the same construction.
- Very good electrical properties at high frequencies.
- Improved quality and reliability.
- Lower component costs with better economy in large production runs.

Production of PCBs with surface-mount components requires a certain amount of new investment in production equipment for the different operations. What is needed is determined by the availability of components for the type of PCB to be manufactured and the choice of mounting and soldering technique involved.

Essentially there are two different manufacturing processes for pure surface mounting:

- Solder paste method
- Adhesive method

The solder paste method is used for pure surface mounting, while the adhesive method is much more useful for mixed mounting, i.e. for boards with both surface-mount and hole-mount components. As well as these two basic methods, there are also manufacturing processes with mixed mounting on both sides of the board, i.e. surface-mount mixed with hole-mount components. This naturally involves a significantly more complicated mounting procedure with hardening and soldering taking place repeatedly.



Surface mounting methods.

Transition to lead-free solder

Lead-free solder has a higher melting temperature than solder containing lead. It is therefor necessary for PCBs and components to have better temperature durability.

Lead-free solder is sensitive regarding polution from e.g. lead. Beware of component connections and solder residue that contain lead on PCBs.

Alloys in lead-free solder have less wettability than alloys in solder containing lead. The demands on the flux therefor are higher.

The solder joint has a duller appearance for lead-free solder and can therefor be mistaken for a cold solder joint.

For hand soldering you can avoid all to high soldering temperatures by selecting soldering tools with high temperature stability, sufficient power and good temperature heat conductivity. Shorter and thicker soldering tips have in general better heat conductivity properties. Keep the tip clean and check it regularly as it is more prone to wear when using lead-free solder.

For automated production the entire process should be controlled and tuned. Further information can be obtained e.g. at *http://www.ittf.no*: *”Why go lead-free?”* and *http://leadfree.ipc.org*.

Wire Wrapping

Wire wrapping as a method of connection was discovered by Bell Telephone Laboratories, USA, at the beginning of the 1950s. It was primarily developed for use in telephone system equipment.

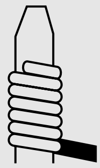
Wire wrapping involves the stretching of a single strand of wire around a square terminal pin with a special tool. The wire is stretched so tight that a gas-tight metallic connection is achieved that withstands temperature changes, corrosive environments, humidity and vibration.

Wire wrap bits and sleeves

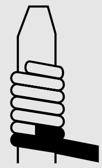
The choice of bit and sleeve is dependent on a number of different parameters, e.g. the wire diameter, the insulation diameter, the wire-wrap pin diagonal, whether normal or modified wrapping, and the length of the wire wrap pin.

Normal wrapping means that no insulation is wrapped round the pin. Modified wrapping means that 1.5 turns of insulation are wound round the pin for extra strain relief.

Normal wrapping



Modified wrapping



It is very important that the correct bit, sleeve and wire are used. Otherwise a satisfactory wrap is not achieved and contact problems result . It is also important that the wire has the correct stripped length to maintain maximum surface contact (too long a stripped length serves no purpose other than to take up space on the pin). On a 0.25 mm wire, the stripped length should be 25.4 mm, which gives 7 wraps on the pin.

Stripping must not damage the wire. It is therefore important to use a stripping tool designed especially for the job.

There are also special "Cut, Strip and Wrap" (CSW) bits and sleeves that do not need any pre-preparation of the wire. The bit and sleeve are constructed in such a way that they strip, cut and wrap the wire in an instant. This method does however put large demands on the wire insulation, which must be specially manufactured for the purpose.

Wire wrap tools

There are a number of different categories of wire wrap tools:

Simple hand tool. Usually a combined tool for wrapping, unwrapping and stripping.

Manual tool. Combined tool for wrapping and unwrapping. Also supplied with bit and sleeve. Designed for industrial use.

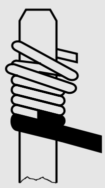
Battery powered tool. Suitable bit and sleeve are available. Can also be used for unwrapping. Designed for service, prototypes and short runs.

Mains powered tool. Suitable bit and sleeve are available. Designed for prototypes and production.

Pneumatic tool. Intended for production.

Automatic machine. Large production volumes.

The battery powered, mains powered and pneumatic tools can be supplied with a back force spring to counteract "over-wrapping", i.e. where the wire wraps round too many times above itself. If the tool is not fitted with this device then the operator must carefully "watch" the tool during wrapping.



”Overwrap”

Unwrapping tools are available in different forms depending on the wire diameter . The tool is provided with a "hook", which catches hold of the wire and unties it from the pin. A number of tools are provided with a sleeve that is "threaded" on the inside so as to lift off the wire. With this type of tool you avoid short circuiting the pins when unwrapping with the equipment still in operation.

Electrical safety

In order to allow products and appliances to be sold these must be safe for the user, pets, etc. Before Sweden joined the EU, according to Swedish law then the mandatory Semko-certificate was used for electrical appliances (also called a third-party certificate). There was also a so called appliance test according to current regulations. Swedish as well as international electrical safety standards were used to show that the device could be considered as safe.

Inside the EU the law that regulates electrical safety is the *Low-voltage directive* 73/23EEG, 93/68EEG.

When Sweden joined the EU European law was introduced in Swedish law. The demand for mandatory certification (Semko-approval) was replaced by a *CE-approval.*

The CE-approval is a demand by the authorities for enabling free trade with goods between the member countries. In short the CE-approval means that:

The CE-approval mark.

- The manufacturer (if inside the EU) or the importer (if the manufacturer is outside the EU) guarantees using the CE-approval that the product conforms to the safety standards that apply inside the EU. Supporting this statement the relevant harmonised EN-standards (European standards) are used.
- For each product documentation must be available. The simplest document is a so called declaration of conformity where the product is given an overall description and reference to standards are made, added are names and addresses of manufacturers or importers. This document must be readily available so that the authority can receive a copy within 3 working days. Among other documents the Technical File can be mentioned. It is a thorough documentation of the product including traceable measurement protocols.

A closely related direcitve is the *EMC-directive* (89/336EEG, 92/31EEG, 93/68EEG). This directive states how electric appliances must conform to set demands regarding electromagnetic radiation, both regarding emission and immunity. Most commonly a device must conform to the demands stated in the low-voltage directive and the EMC-directive before receiving a CE-approval.

In order for the system with declaration of conformity to work, the supervisory authorities (In Sweden the Electrical Safety Board) carry out a control of the market inside the EU. This means that the authority acquires products for control. If it should be the case that the product is not safe at such a control, the authority can choose between a number of measures. The easiest is a remark and the most powerful is a fine or imprisonment.

For further information on the subject visit the website of the Electrical Safety Board *www.elsak.se.*

Practical advice from SEMKO for the do-it-yourselfer

1. Make sure that you have **the necessary tools**. It is especially important that you have a good soldering iron and are able to solder correctly. If you have never soldered before, ask someone to teach you. Use non-acid flux, or solder which contains this, for all soldering. For more information, see section *soldering.*
2. Be careful when **soldering** so that the components and insulation are not damaged by the hot soldering iron.
3. Resistors that develop **heat** must be mounted away from the circuit board, conductors and other flammable parts.
4. Do not leave out any **fuses or protection components**.
5. Make sure that you are using components that are **of the correct size or rating**. For example, do not use a 0.5 W resistor, if the instructions say that you should use a 1 W resistor.
6. Make sure that you have plenty of **insulation space** around live parts. This is especially important between uninsulated live parts and touchable metal parts (the chassis and the secondary circuit).
7. Make sure that components, conductors etc. are safely fixed, so that there is no **risk of short circuiting** or heat damaging.
8. Do not bundle **mains voltage cables** with others.
9. Do not attempt to make (etch) **circuit boards** if they are designed for components that will conduct **mains voltage**.
10. Do not construct **the mains transformer** yourself, but choose a correctly dimensioned, ready made one.
11. Check that the plug, mains cable, switch and fuses are **Semko-approved**.
12. Check that the **fuses** have correct rating, so that you obtain the desired level of protection.
13. Never use **fuses** for a higher current than the building instructions specify. This can result in fire, and you could destroy valuable components.
14. If a fuse is tripped, you have most likely connected it incorrectly. You must **find the fault** and correct it before replacing the fuse.
15. Do not make **any temporary connections**. Instead, it is wise to do a proper job right from the very beginning.
16. Use properly **insulated copper wire**, especially for circuits that conduct mains voltage.
17. Make sure that the connection wires' **insulation** cannot be **damaged** by sharp metal edges, moving parts or hot components.
18. Only use speakers that are **matched** to the amplifier's impedance and output power.
19. Make sure that the device cover fits and provides **sufficient touch protection**, and that the cover has the necessary ventilation openings. It should not be possible to remove the cover without using tools.
20. Be careful when you **test** the device. Bear in mind that it is highly dangerous to touch live parts. Make sure there is someone else around to break the current if you get stuck.

Inspection and maintenance

Make sure that the DIY kit is **correctly connected** before any battery or mains voltage is connected.

Have you used **the correct soldering iron? NOTE!:** Transistors and capacitors can break if too high a temperature is used.

The **warranty** for parts is only valid if the kit has been assembled carefully and correctly.

The warranty for components is valid for **1 year** after purchase, but not for the assembled kit.

Check the current regulations with the appropriate **electrical safety authorities** in your country.

WEEE and RoHS

Producer responsibility for electrical and electronics products (WEEE)

Producer responsibility for electrical and electronics products puts the responsibility on manufacturers and distributors (so called producers) within the EU of handling products that are put out on the market in an environmentally friendly way. The purpose of letting the producers take the responsibility is that those who sell or use a product also should take responsibility for the costs that occur, the so called ”polluter pays-principle”.

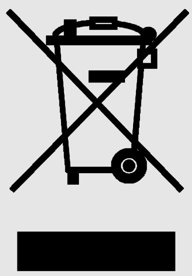
The purpose of the legislation is that electrical and electronic products should be designed and manufactured in such a way that the amount of waste is minimised and that the waste that is produced must not harm the environment. During 2005 producer responsibility for these products was introduced throughout the EU, the WEEE-directive.

The directive describes which commitments each member state must meet in order to introduce legislation.

The area of responsibility covers that the products are clearly marked, that collection is carried out and that the collected products are recycled in an environmentally friendly way. This in turn requires that systems for collection is set up and that there are clear requirements for recyclers. There are also requirements that producers inform all concerned parties, like consumers, recyclers, counties and authorities.

The product hierarchies concerned are:

1. Large appliances
2. Small appliances
3. IT-, telecommunication and office equipment
4. Home equipment (TV-, audio and video equipment)
5. Lighting equipment
6. Electric and electronic tools
7. Toys, leisure time and sports equipment
8. Medical technical equipment
9. Surveillance and monitoring equipment
10. Vending machines



Appropriate product marking.

Ban on use of certain health-hazardous substances (RoHS)

Electrical or electronics products may contain substances that are harmful to humans and the environment. Some of these substances are not harmful when the product is being used, but will be harmful as waste.
Handling these substances are both hazardous and expensive and it is therefor important to avoid these substances in the products.

Through the RoHS-directive certain health-hazardous substances in electrical and electronics products are banned.

As of July 1st 2006 newly manufactured electrical and electronics products must not contain:

Metals
Mercury, Hg (0,1 %)
Cadmium, Cd (0.01 %)
Bly, Pb (0,1 %)
Hexavalent chrome, Cr 6+ (0.1 %)

Flame retardants
Polybromine biphenyles, PBB (0.1 %)
Polybromine diphenylethers, PBDE (0.1 %)

There exist no technical solutions yet for exclusion of these substances in all the products in the product hierarchies included in the WEEE-directive, these means that the EU technical committee (TAC), has decided on exceptions.

For all EU member states the product hierachies 8 and 9 are not included, as well as servicing and repairs of those products manufactured before July 1st 2006. Batteries are not included in this directive. They are included in the battery directive.

Apart from these exceptions certain specific products are also not included and there is a continuous review of individual products in order to reach the goal to completely avoid the substances included in the RoHS-directive.

Of all products banned by the RoHS-directive lead is without doubt the most common substance in electrical and electronics products, and it is also the substance that is most difficult and expensive to replace for the electronics industry. Almost all solder points contain lead. As of July 1st 2006 all production must be using lead-free solder and this change affects the manufacturing process in many ways. The greatest change is the increase in process temperature which in turn exposes the components to higher temperatures. See also page 0000, *''Transition to lead-free solder''*.

PBB and PBDE are two types of flame retardants that exist in plastics. Hexavalent chrome has been used as anti-corrosion protection. Mercury has primarily been used in relays, sensors and fluorescent lamps. There exists replacements för mercury since many years and many countries have banned mercury since long. The use of cadmium as pigment and plating has also been restricted for many years.

ELFA's commitment

Some of the products sold by ELFA are included directly in both directives but the majority of the range comprises components that are used by our customers for production or repairs of products. The components used must comply with the RoHS-directive if they are to be used in finished products that are included in the WEEE-directive.

In order to keep the information concerning RoHS-compatibility of components and products current please see our website *www.elfa.se* or contact our personnel in the technical information department. We have also put together links on our website to the EU directives and to national legislature.

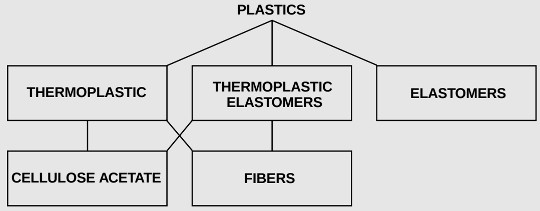
Plastics

The most important component in a plastic is the **polymer**. A polymer is a compound whose molecules are made up of a large number of identical groups of atoms. Plastics can be made of one polymer or several different polymers. Various additives are also used to modify the properties of the polymer. To summarise:

Polymer + Additive = Plastic

The use of plastics offers significant advantages in many applications, although you should also take into account the weaknesses and negative aspects of the material. One important property is the ability of plastic to resist particular environmental effects.

During manufacture and processing, the polymers create **chains of molecules** of various types. Depending on the shape of the chain of molecules, the plastics are classified as thermoplastics, thermosets and elastomers.



The relation between different types of plastics.

Thermoplastics have a linear or branched molecular structure. The molecules are held together by weak links. Materials with this structure are called thermoplastics. They have the characteristic feature that they become softer and more pliable as the temperature increases. As a result, they do not have a clearly-defined melting point. A thermoplastic can be heated and remoulded repeatedly without changing the properties of the material.

Thermosets are plastics in which the molecules form a network. Thermosets share the feature that during initial processing, they are given their final form as one, two or more components react with each other, creating a product with a network structure. The material then hardens. The material retains its shape with the application of heat, until the temperature increases to the point that the material starts to burn. Thermosets are usually rigid and brittle, and are characterised by low moisture absorption, dimensional stability and much lower pliability than most thermoplastics.

Elastomers have a similar structure to thermosets. The links and the chains of molecules are arranged more sparsely, which means they are easily moulded.

Reinforced plastics involve the use of strong fibres as a filler or in the filler, to reinforce (strengthen) the material.

Foamed plastics can either be thermosets or thermoplastics. Both types share the feature that the plastic forms a foam after the addition of a gas-generating agent or blowing agent. Foamed plastics are primarily used as heat insulation.

Additives

The properties of the different plastics can be altered by adding various ingredients to the polymer. Here are some examples of additives:

Plasticisers are added to make thermoplastics, in particular, softer and more resilient. Without a plasticiser, polymers are too hard and brittle for practical use. In thermoplastics, the effect of the plasticiser is to reduce the forces binding the polymer chains together, allowing the chains to slide over each other.

Stabilisers help polymers counter the effects of aging. Various kinds of stabilisers are used.

Fillers can improve the properties of the plastic, but their main benefit is to make the end product less expensive. The excessive use of filler indeed produces a cheaper product, but the result is frequently poor. Widely-used fillers include rock dust, chalk, clay, wood dust and cellulose.

Fire retardants are yet another type of additive. When it is set alight, a plastic passes through three stages: heating, physical degradation and ignition. When a polymer breaks down, gases are released. Some of these gases are highly flammable, some of them have a corrosive effect on surrounding metals, while others, for example carbon dioxide, smother the flames and are called self-extinguishing. The temperature of a polymer is also particularly dependent on the breakdown reactions it undergoes. The use of various additives, fire retardants, to inhibit these reactions means that any fires can be reduced in intensity or prevented altogether.

Examples of polymer types.

Thermoplastics

PVC	= Polyvinyl chloride	
LDPE	= Polyethylene (low density)	
HDPE	= Polyethylene (high density)	
PP	= Polypropylene	
PA	= Polyamide	
FEP	= Fluorinated ethene and propene	Poly-fluorinated carbons
PTFE	= Polyetrafluorethylene	
ETFE	= Ethenetetrafluorethene	
PMMA	= Polymethyl metacrylate = Plexi glass	
PS	= Polystyrene, standard styrene	
SAN	= Polystyrene, heat-resistant styrene	
ABS	= Polystyrene, acrylonitrile-butadiene-styrene	
PC	= Polycarbonate	
PETP, PET	= Polyester	
PUR	= Polyurethane (upholstery)	

Thermosets

UF = Urea-formaldehyde polymers = urea resins
MF = Melamine-formaldehyde polymers = Melamine
PF = Phenol formaldehyde = phenolic resins, Bakelite
EP = Epoxy resin = Araldite
UP = Unsaturated polyesters= Glass-fibre reinforced polyesters (a type of thermoplastic)
PUR = Polysiloxanes (for surface treatments)

Properties of plastics

The properties of the various plastics are highly dependent on the additives used. This is particularly relevant with regard to fire, since the additives can create different types of risk. The addition of **chlorine** or **bromine** to a plastic can result in the formation of dioxins in the event of fire.

PVC is a widely used plastic, for example in cables and packaging. PVC is manufactured in many different forms, offering a wide range of properties. A grey smoke is formed when rigid PVC is burnt, and black smoke when soft PVC is burnt.

Hydrochloric acid is produced when the plastic is burnt, but it is often neutralised by the additives used.

PVC is normally self-extinguishing, but it is flammable if the fire is sustained by other materials.

The chlorine content of PVC makes it undesirable from the environmental point of view, so it should be replaced with other plastics whenever possible.

The surface of PVC should be roughened before adhesion to ensure a good substrate. Adhesion with the use of PVC dissolved in an appropriate solvent. When bonding with another material, contact adhesive, polyurethane adhesive or two-component epoxy adhesive can be used.

PE is available as LD polyethylene (low density) and HD polyethylene (high density). LDPE is a soft material, used to make film for vapour barriers, plastic bags, cases, buckets, kitchen bowls and toys. HDPE is more rigid, and is used to make crates, pipes, containers, etc. It is relatively highly flammable and is not self-extinguishing. When the plastic is burnt it releases water and carbon dioxide. It gives off thin, white smoke that smells of candles burning, it sustains combustion and is easily flammable. It is recyclable. Ethylenes are lighter than water. PE plastics have very good electrical properties, (good insulation), as well as extremely low water permeability.

Polyethylene is very difficult to bond.

PP, polypropylene, is similar to HDPE but with greater surface hardness. It is highly heat resistant, tolerating temperatures up to +120°C, so it can be sterilised. It is used to make medical equipment, soft drink containers, components in domestic irons, toasters, fridges and freezers and certain car components. PP is the lightest of the widely used thermoplastics. Its electrical properties (insulation) are better than polyethylene, which makes PP very suitable as a raw material in telephone and high-frequency cables. Other positive properties include its high water impermeability and resistance to fissuring under chemical and physical influences. It is less resistant to cold than polyethylene.

Polypropylene is easily flammable and is not self-extinguishing. No hazardous compounds are released when it is burnt. Flame-retardant ingredients have been added to some products. Polypropylene is difficult to bond and has to be pre-treated. The material has to be ground and then bonded using a suitable cyanoacrylate adhesive.

PA, polyamide, is known to most of us as Nylon, the trade name. PA is a resilient and strong material suitable for use in textiles, V-belts, knife handles, cogs and motor casings. It is not highly flammable, nor is it self-extinguishing. When it is burnt, no materials are given off that are hazardous to humans. Burning PA produces nitric oxides, which can contribute to acidification. Thin white smoke is given off. Polyamide is difficult but not impossible to bond, using a solvent with a Nylon additive.

Two-component epoxy adhesive should be used to bond PA with other materials.

FEP, PTFE, ETFE can be grouped together under the term polyfluorinated carbons. The plastics are resilient, highly scratch resistant and can tolerate most chemicals. Their electric and dielectric properties are very good, too. Fluorinated carbons can withstand very low and very high temperatures, –190°C to +260°C. They have good sealing properties. In general, it is very difficult to bond these plastics.

FEP is self-extinguishing. When heated to +400°C, extremely aggressive and poisonous gases, hydrogen fluorides, are given off. The plastic is sold under the name Teflon (Du Pont trade mark).

Fluorinated carbons are in general very difficult to bond, having to be pretreated by grinding or etching.

PMMA is the most important plastic in the group of acrylic polymers. PMMA is characterised by high transparency, very good weather resistance and high surface hardness. PMMA is environmentally friendly. It is flammable and is not self-extinguishing. When the plastic is burnt, only carbon, hydrogen and oxygen are released. The trade name of PMMA is Plexiglas. Acrylic polymers are used in paints and sealing compounds.

PMMA can be bonded with various types of solvent, but the bond is weakened by UV radiation. Contact adhesive is used to bond PMMA to other materials.

PS stands for polystyrene resin. It has good hardness, rigidity and dimensional stability properties, and is relatively brittle. It is easy to process into finished products, and is cheap. It cannot withstand high temperatures. PS is not light-resistant so it is not suitable for outdoor use. It has excellent electrical properties (insulation capacity).

Polystyrene resin is highly flammable, is not self-extinguishing and produces large amounts of residues when burnt. Some polystyrene products include additives of fire-retardant ingredients.

Polystyrene can be bonded using solvents such as acetone or methylchloride. Contact adhesive is used to bond it to other materials.

Foamed polystyrene consists of many closed cells. The result is a rigid material with very low heat conductivity, 0.035 W/(m × K), and minimal water moisture absorbency, less than 3%. Foamed polystyrene is marketed as Frigolite and Styrolite.

SAN has better properties than polystyrene alone in terms of hardness, rigidity, and tensile strength. It is also more resistant to heat and chemicals. SAN is transparent, with a slight yellow colour, but it is usually coloured a very pale blue. SAN is used to make instrument panels, covers for office machines, fridge components and household items. SAN contains nitrogen, and produces acidifying nitric oxides when burnt.

ABS has the advantage of better impact resistance and resistance to chemicals and aging than polystyrene. ABS has low shrinkage and is non-transparent (opaque). By changing the proportions of the ingredient monomers, the properties of ABS can be altered within relatively wide margins. This means it can be tailored to specific applications. ABS is used to make covers for equipment like telephones, radio receivers, cameras, projectors, office machines, instrument panels in vehicles, helmets, crates and toys like Lego. Like SAN, ABS contains nitrogen, which produces acidifying nitric oxides when the plastic is burnt.

PC, polycarbonate, has very good mechanical properties, excellent dimensional stability, and outstanding impact resistance, as well as high rigidity. PC can withstand high temperatures, up to +110°C over long periods. Its high impact resistance is maintained when it is cold, and the material can be used at temperatures down to –100°C. It is moderately resistant to chemicals, but it is attacked or degraded by a number of organic solvents, and is broken down by alkalis. It is highly weather resistant, but the surface discolours to yellow by the action of UV light. The electrical properties (insulation) are very good for most applications. PC is used to make equipment covers and other components, electric and electronic components, toolboxes, handles for power tools, multipin connectors, moisture barriers for relays, etc, food containers, helmets and bullet-proof panels. PC is transparent. It is self-extinguishing, with an ignition point of at least +500°C. Carbon dioxide is the only product of combustion.

Small surfaces can be bonded using adhesives like methylene chloride. Two-component epoxy adhesive is used for bonding large surfaces and bonding to other materials.

PET, a kind of polyester, is produced as a thermoplastic or a thermoset. It is an unreinforced polyester of the normal thermoset type. It is hard, rigid and quite brittle, with good electrical properties (insulation) and moderate resistance to chemicals. It cannot withstand strong acids and bases, nor nonpolar solvents. It has good water and weather resisting properties. The plastic can be made self-extinguishing if acids containing chlorine are used at the prepolymer stage.

Linear polyester belongs to the thermoplastics group. It is used to make high-quality films and textile fibres. Terylene, Dacron and Tergal are trade names for this type of polyester.

Glass-fibre reinforced polyester has a better strength/mass ratio than many metals. Unsaturated polyester is used to make glass-fibre reinforced plastics, and as paint, bonding agent and putty. Glass-fibre reinforced polyester is used to make boats of various sizes, car bodies, electrical equipment, helmets, flag poles, masts, fishing rods and skis.

PUR, polyurethane, is produce as a thermoset or a thermoplastic. As a thermoset, it is being used more and more in industry, largely in vehicle manufacturing. It can be turned into anything from elastic and soft, to hard and wood-like, and its uses include bonding agents in paints and fillers, mattresses, upholstery, soles of shoes, fridge fittings, foamed plastic panels or heat and sound insulating foam. PUR is used less widely as an electrical material. The plastic is opaque, non-transparent. Rigid PUR foam can withstand dilute acids and bases, but swells by the action of ethanol, acetone and carbon tetrachloride. Semi-rigid and soft foamed PUR is less resistant to chemicals than the rigid version. It swells by the action of benzene and turpentine. It is not particularly weather resistant. The material yellows with age and in the presence of heat, it can take on a certain moisture content, making the material more brittle. Heating can cause isocyanates to be formed, so this plastic should not be burnt. Burning also produces acidifying nitric oxides. The blowing agent CFS, which can damage the ozone layer, has been used in the past to produce polyurethanes.

EP, epoxy resin, is a relatively expensive plastic. Non-reinforced EP has good impact resistance. It has good electric breakdown, resistance and radiation properties. The material can be used in a wide range of temperatures. EP is extremely resistant to the effects of chemicals. Applications for epoxy resin include laminates and reinforced plastics, paints, adhesives, casting resin and bonding agents. Epoxy resin laminates with glass-fibre reinforcement are widely used in the manufacture of PCBs. Epoxy adhesives provide excellent bonding with most materials. Enamels and paints based on epoxy resin have good adhesion, chemical resistance and durability, for example baking enamels. Casting resins, with or without filler, are used for casting and protecting sensitive electrical components.

PF, phenolic resins, have good mechanical properties, depending on the filler used. Phenolic resins have excellent dimensional stability, low shrinkage and high rigidity. Impact resistance is relatively low. The material can withstand extremes of heat, up to a maximum of +150°C. Chemical resistance is quite good, and phenolic resins tolerate water very well. Weather resistance is quite poor. The electrical properties are moderately good (good insulation), but the plastic should not be used in humid environments because of water absorption. The surface of phenolic resins carbonates when the material is burnt. Some types are self-extinguishing. Phenolic resins are used as a base in bonding agents in the manufacture of sandpaper and brake linings, and as an adhesive in watertight grades of plywood and particle board. Phenolic laminates are used to make radio components, switches and printed circuits.

UF, MF stand for urea-formaldehyde (UF) and melamine-formaldehyde (MF) polymers. The group as a whole is called the amino resins. The plastics have good mechanical properties. They are very hard and have extremely good resistance to wear. The surface properties are regarded as being the best of all plastics. The resistance to heat is good among urea-formaldehyde polymers and excellent among melamine-formaldehyde polymers. The amino resins are all resistant to chemicals. They can withstand dilute bases and acids, oils, greases and most organic solvents. The melamine version withstands boiling water, but the urea version do not. Weather resistance is poor, and the amino resins should not be used outdoors. They have good electrical characteristics, with very high tracking resistance. Amino resins cannot be electrostatically charged so they do not attract dust. Amino resins are self-extinguishing. They are used to make moulded items and laminates, and in bonding agents, adhesives and paints. Baking enamel, i.e. enamel that hardens quickly at high temperatures, is based on amino resin.

8-bit ASCII table for PCs

Dec value		0	16	32	48	64	80	96	112	128	144	160	176	192	208	224	240
▼	Hex value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	0	NULL	DLE ►	SP	0	@	P	`	p	Ç	É	á	░	└	╨	α	≡
1	1	SOH ☺	DC1 ◄	!	1	A	Q	a	q	ü	æ	í	▒	┴	╤	β	±
2	2	STX ☻	DC2 ↕	"	2	B	R	b	r	é	Æ	ó	▓	┬	╥	Γ	≥
3	3	ETX ♥	DC3 ‼	#	3	C	S	c	s	â	ô	ú	│	├	╙	π	≤
4	4	EOT ♦	DC4 ¶	$	4	D	T	d	t	ä	ö	ñ	┤	─	╘	Σ	⌠
5	5	ENQ ♣	NAK §	%	5	E	U	e	u	à	ò	Ñ	╡	┼	╒	σ	⌡
6	6	ACK ♠	SYN ▬	&	6	F	V	f	v	å	û	ª	╢	╞	╓	µ	÷
7	7	BEL •	ETB ↨	´	7	G	W	g	w	ç	ù	º	╖	╟	╫	τ	≈
8	8	BS ◘	CAN ↑	(	8	H	X	h	x	ê	ÿ	¿	╕	╚	╪	Φ	°
9	9	HT ○	EM ↓	)	9	I	Y	i	y	ë	Ö	⌐	╣	╔	┘	θ	•
10	A	LF ◙	SUB →	*	:	J	Z	j	z	è	Ü	¬	║	╩	┌	Ω	·
11	B	VT ♂	ESC ←	+	;	K	[	k	{	ï	¢	½	╗	╦	█	δ	√
12	C	FF ♀	FS ∟	,	<	L	\	l	¦	î	£	¼	╝	╠	▄	∞	η
13	D	CR ♪	GS ↔	−	=	M	]	m	}	ì	¥	¡	╜	═	▌	φ	²
14	E	SO ♫	RS ▲	.	>	N	^	n	~	Ä	Pt	«	╛	╬	▐	ε	■
15	F	SI ☼	US ▼	/	?	O	_	o	⌂	Å	∫	»	┐	╧	▀	∩	

Text description of control codes.

NUL = Null
SOH = Start of Heading
STX = Start of Text
ETX = End of Text
EOT = End of Transmission
ENQ = Enquiry
ACK = Acknowledge
BEL = Bell
BS = Backspace
HT = Horizontal Tabulation
LF = Line Feed
VT = Vertical Tabulation
FF = Form Feed
CR = Carriage Return
SO = Shift Out
SI = Shift In
DLE = Data Link Escape
DC1 = Device Control 1 (XON)
DC2 = Device Control 2
DC3 = Device Control 3 (XOFF)
DC4 = Device Control 4
NAK = Negative Acknowledgement
SYN = Syncronous Idle
ETB = End Of Transmission Block
CAN = Cancel Line
EM = End Of Medium
SUB = Substitute
(Also used as EOF=End Of File)
ESC = Escape
FS = File Separator
GS = Group Separator
RS = Record Separator
US = Unit Separator
SP = Space

Constants and units

Physical constants

Standard acceleration of gravity	$g_n = 9{,}806\,65\ m/s^2$
Speed of light in vacuum	$c_o \approx 2{,}99793 \times 10^8\ m/s$
Magnetic constant, permeability in vacuum	$\mu_o = 4\,\pi \times 10^{-7}\ H/m$ $\approx 1{,}257 \times 10^{-6}\ H/m$
Permittivity of vacuum	$\varepsilon_o \approx 8{,}854 \times 10^{-12}\ F/m$
Elementary charge	$e \approx 1{,}6021 \times 10^{-19}\ C$
Faraday constant	$F \approx 9{,}6487 \times 10^4\ C/mol$
Boltzmann constant	$k \approx 1{,}38 \times 10^{-23}\ J/K$

Temperature equivalents

0°C corresponds to 273.15 K
32.0°F corresponds to 273.15 K

Temp in °C corresponds to ((Temp in °F) – 32) / 1.8.
Temp in °F corresponds to (Temp in °C) ×1.8 + 32.

SI-units

Quantity	Name	Symbol	Expressed in other SI units	Expressed in SI base units
Base units				
Length (l)	meter	m		
Mass (m)	kilogram	kg		
Time (t)	second	s		
Electric current (I)	ampère	A		
Temperature (T)	kelvin	K		
Luminous intensity (I)	candela	cd		
Amount of substance (n)	mole	mol		
Derived units				
Frekvens (f)	hertz	Hz		s^{-1}
Force (F)	newton	N		$m \times kg \times s^{-2}$
Pressure, stress (p)	pascal	Pa	N/m^2	$m^{-1} \times kg \times s^{-2}$
Energy, work * (W)	joule	J	Nm, Ws	$m^2 \times kg \times s^{-2}$
Power (P)	watt	W	J/s	$m^2 \times kg \times s^{-3}$
Electric charge, quantity (Q)	coulomb	C	As	$s \times A$
Electric pot. diff. (V), electromotive force (U)	volt	V	W/A	$m^2 \times kg \times s^{-3} \times A^{-1}$
Capacitance (C)	farad	F	C/V	m^{-2}
Resistance (R)	ohm	Ω	V/A	$m^2 \times kg \times s^{-3} \times A^{-2}$
Conductance (G)	siemens	S	A/V	$m^{-2} \times kg^{-1} \times s^3 \times A^2$
Magnetic flux density (B)	tesla	T	Wb/m^2	$kg \times s^{-2} \times A^{-1}$
Magnetic flux (Φ)	weber	Wb	Vs	$m^2 \times kg \times s^{-2} \times A^{-1}$
Inductance (L)	henry	H	Wb/A	$m^2 \times kg \times s^{-2} \times A^{-2}$
Luminous flux (Φ)	lumen	lm	cd × sr	
Illuminance (E)	lux	lx	lm/m^2	$cd \times sr \times m^{-2}$

* mechanical as well as electric and thermal

SI prefixes

Factor	Name	Symbol
10^{24}	yotta	Y
10^{21}	zetta	Z
10^{18}	exa	E
10^{15}	peta	P
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	kilo	k
10^2	hecto	h
10^1	deca	da

Factor	Name	Symbol
10^{-1}	deci	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	mikro	µ
10^{-9}	nano	n
10^{-12}	piko	p
10^{-15}	femto	f
10^{-18}	atto	a
10^{-21}	zepto	z
10^{-24}	yokto	y

Example: 1 MΩ = 1000 kΩ   1 pF = 10^{-12} F = 10^{-6} µF   1 µm = 10^{-3} mm

IEC prefixes for binary multiples

Value	Name	Symbol
2^{60}	exbi	Ei
2^{50}	pebi	Pi
2^{40}	tebi	Ti
2^{30}	gibi	Gi
2^{20}	mebi	Mi
2^{10}	kibi	Ki

Example: 1 Kibit = 1 kibibit = 1 ”kilo binary”-bit = 1 × 2^{10} bit = 1024 bit

Conversion table for units of measurement

Length. SI unit metre (m).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 fermi	fm	1 femtometer	10^{-15} m
1 x unit	Xu	1.00208 mÅ	1.00208×10^{-13} m
1 atomic mass unit	amu	1 bohr	5.29177×10^{-11} m
1 angstrom	Å	10 nm	10^{-10} m
1 micron	µ	1⁄1000 mm	10^{-6} m
1 mil		0.001 in	0.0254 mm
1 inch (tum)	in, ”	1000 mil	2.54 cm
1 foot	ft, ’	12 in	30.48 cm
1 yard	yd	3 ft	0.9144 m
1 mile (statute mile)	mi	5280 ft	1609.344 m
1 nautical mile	nmi, NM	6076 ft	1852 m
1 astronomical unit	AU, ua		1.495978706×10^{11}m
1 Light year	Ly	6.32×10^4 AU	9.46053×10^{15} m
1 parsec	pc	2.06265×10^5 AU	3.0857×10^{16} m

Area. SI unit square metre (m^2).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 barn		100 fm^2	10^{-28} m^2
1 circular mil	CM	0.7854 mil^2	5.067×10^{-10} m^2
1 square inch	in^2	1.273×10^6 CM	6.4516 cm^2
1 square foot	ft^2	144 in^2	0.09290304 m^2
1 square yard	yd^2	9 ft^2	0.83612736 m^2
1 are	a		100 m^2
1 acre	ac, A	4840 yd^2	4046.86 m^2
1 hectare	ha	100 are	10 000 m^2
1 square mile	$mile^2$	640 acre	2 589 988 m^2

Volume. SI unit cubic metre (m^3).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 minim (UK)	min	59.2 µl	5.9194×10^{-8} m^3
1 minim (US)	min	61.6 µl	6.1612×10^{-8} m^3
1 cubic inch	cu in, in^3	1.64 cl	1.6387×10^{-5} m^3
1 UK fluid ounce	UK fl oz	2.84 cl	2.8413×10^{-5} m^3
1 US fluid ounce	US fl oz	2.96 cl	2.9574×10^{-5} m^3
1 US liquid pint	US lq pt	16 US fl oz, 0.473 l	4.7317×10^{-4} m^3
1 US dry pint	US dry pt	0.551 l	5.5061×10^{-4} m^3
1 UK pint	UK pt	20 UK fl oz, 0.568 l	5.5683×10^{-4} m^3
1 US (liquid) quart	US (lq) qt	2 US lq pt, 0.946 l	9.4635×10^{-4} m^3
1 litre	l	1 dm^3	10^{-3} m^3
1 US dry quart	US dry qt	2 US dry pt, 1.101 l	$1{,}1012\times10^{-3}$ m^3
1 UK quart	UK qt	2 UK pt, 1.137 l	1.1365×10^{-3} m^3
1 US liquid gallon	US gal	8 US dry pt, 3.785 l	3.785×10^{-3} m^3
1 UK gallon	UK gal	8 UK pt, 4.546 l	4.546×10^{-3} m^3
1 cubic foot	cu ft, ft^3	1728 in^3, 28.3 l	2.8317×10^{-2} m^3
1 US dry barrel	dbl	7056 in^3, 116 l	1.1563×10^{-1} m^3
1 US liquid barrel	bl	31.5 US gal, 119 l	1.1924×10^{-1} m^3
1 US petroleum barrel	bo	42 US gal, 159 l	1.5899×10^{-1} m^3
1 UK barrel	bl	36 UK gal, 164 l	1.6365×10^{-1} m^3
1 cubic yard	cu yd, yd^3	765 l	7.6455×10^{-1} m^3

Mass. SI unit kilogram (kg).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 atom mass unit	amu	1⁄12 of an atom C12	1.6605402×10^{-27} kg
1 point	pt	1⁄100 ct	2 mg
1 grain	gr	1⁄7000 lb	64.79891 mg
1 carat (metric)	ct		0.2 g
1 dram	dr	1⁄16 oz	1.7718 g
1 ounce	oz	1⁄16 lb	28.3495 g
1 pound (avoirdupois)	lb		0.45359237 kg
1 stone	st	14 lb	6.3503 kg
1 US quarter	qtr, qr	25 lb	11.34 kg
1 UK quarter	qtr, qr	28 lb, 2 st	12.70 kg
1 short hundredweight	sh cwt	100 lb	45.36 kg
1 long hundredweight	cwt	112 lb, 8 st	50.80 kg
1 short ton	sh t	2000 lb	907.18 kg
1 metric ton	t		1000 kg
1 (long) ton	t	20 cwt	1016.05 kg

Speed. SI unit metre per second (m/s).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 kilometre/hour	km/h	5⁄18 m/s	0.2778 m/s
1 foot/second	ft/s	1.097 km/h	0.3048 m/s
1 mile/hour	mph	1.609 km/h	0.4470 m/s
1 knot	kn	1.852 km/h	0.5144 m/s
1 mach	M	speed of sound	aprox. 340 m/s

Time. SI unit second (s).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 minute	min	1⁄1440 d	60 s
1 beat		1⁄1000 d	86.4 s
1 hour	h	60 min	3600 s
1 day	d	24 h	86 400 s
1 week		7 d	604 800 s
1 calendar year		365 d	31 536 000 s
1 tropical year (solar year)	a	365.242 d	31 556 926 s
1 sidereal year (stellar year)		365.256 d	31 558 153 s
1 anomalistic year		365.260 d	31 558 432 s
1 leap year		366 d	31 622 400 s

Angle. SI unit radian (rad).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 second	”	1⁄3600 °	4.4841368×10^{-6} rad
1 minute	’	1⁄60 °	2.9088821×10^{-4} rad
1 grad (gon)	g	1⁄400 rev., 0,9 °	1.5707963×10^{-2} rad
1 degree	°	1⁄360 rev.	1.7453286×10^{-2} rad
1 radian	rad	180/π °	57.2958 °

Force. SI unit newton (N).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 dyne			10^{-5} N
1 gram force	gf	1⁄1000 kgf	9.80665×10^{-3} N
1 pond	p	1 gf	9.80665×10^{-3} N
1 poundal	pdl	1 lb ft/s^2	1.38255×10^{-1} N
1 newton	N	1 kg m/s^2	
1 pound force	lbf	0.45359 kp	4.44822 N
1 kilogram force	kgf	1 kp	9.80665 N
1 kilopond	kp	1 kgf	9.80665 N

Pressure. SI unit pascal (Pa).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 newton/square metre	N/m^2		1 Pa
1 mm of water	mm H_2O		9.80665 Pa
1 mm of mercury	mmHg		133 Pa
1 torr		1 mmHg @ 0 °C	133.322 Pa
1 pound-force/square inch	psi, lbf/in^2	51.72 torr	6.8948×10^3 Pa
1 kilopond/square centimetre	kp/cm^2	1 at	9.80665×10^3 Pa
1 technical atmosphere	at	1 kp/cm^2	9.80665×10^3 Pa
1 bar	b	750.1 torr	10^5 Pa
1 atmosphere	atm	760 torr	1.01325×10^5 Pa

Torque. SI unit newton metre (Nm).

Unit	Symbol	Corresponds to	Corresponds to in SI units
1 pound inch	lbf in	1⁄12 lbf ft	0.112985 Nm
1 pound foot	lbf ft		1.35582 Nm
1 kilopond metre	kpm	1 kgfm	9.80665 Nm

Conversion table, inches/mm

Conversion table inches-metric

Inch	mm	Inch	mm	Inch	mm	Inch	mm	Inch	mm	Inch	mm	Inch	mm
1/64	0,40	11/64	4,37	21/64	8,33	31/64	12,30	41/64	16,27	51/64	20,24	61/64	24,21
1/32	0,79	3/16	4,76	11/32	8,73	1/2	12,70	21/32	16,67	13/16	20,64	31/32	24,61
3/64	1,19	13/64	5,16	23/64	9,13	33/64	13,10	43/64	17,07	53/64	21,03	63/64	25,00
1/16	1,59	7/32	5,56	3/8	9,53	17/32	13,49	11/16	17,46	27/32	21,43	1	25,40
5/64	1,98	15/64	5,95	25/64	9,92	35/64	13,89	45/64	17,86	55/64	21,83		
3/32	2,38	1/4	6,35	13/32	10,32	9/16	14,29	23/32	18,26	7/8	22,23		
7/64	2,78	17/64	6,75	27/64	10,72	37/64	14,68	47/64	18,65	57/64	22,62		
1/8	3,17	9/32	7,14	7/16	11,11	19/32	15,08	3/4	19,05	29/32	23,01		
9/64	3,57	19/64	7,54	29/64	11,51	39/64	15,48	49/64	19,45	59/64	23,41		
5/32	3,97	5/16	7,94	15/32	11,91	5/8	15,88	25/32	19,84	15/16	23,81		

Inch	0	1/16	1/8	3/16	1/4	5/16	3/8	7/16	1/2	9/16	5/8	11/16	3/4	13/16	7/8	15/16
0	0,0	1,6	3,2	4,8	6,4	7,9	9,5	11,1	12,7	14,3	15,9	17,5	19,1	20,6	22,2	23,8
1	25,4	27,0	28,6	30,2	31,8	33,3	34,9	36,5	38,1	39,7	41,3	42,9	44,5	46,0	47,6	49,2
2	50,8	52,4	54,0	55,6	57,2	58,7	60,3	61,9	63,5	65,1	66,7	68,3	69,9	71,4	73,0	74,6
3	76,2	77,8	79,4	81,0	82,6	84,1	85,7	87,3	88,9	90,5	92,1	93,7	95,3	96,8	98,4	100,0
4	101,6	103,2	104,8	106,4	108,0	109,5	111,1	112,7	114,3	115,9	117,5	119,1	120,7	122,2	123,8	125,4
5	127,0	128,6	130,2	131,8	133,4	134,9	136,5	138,1	139,7	141,3	142,9	144,5	146,1	147,6	149,2	150,8
6	152,4	154,0	155,6	157,2	158,8	160,3	161,9	163,5	165,1	166,7	168,3	169,9	171,5	173,0	174,6	176,2
7	177,8	179,4	181,0	182,6	184,2	185,7	187,3	188,9	190,5	192,1	193,7	195,3	196,9	198,4	200,0	201,6
8	203,2	204,8	206,4	208,0	209,6	211,1	212,7	214,3	215,9	217,5	219,1	220,7	222,3	223,8	225,4	227,0
9	228,6	230,2	231,8	233,4	235,0	236,5	238,1	239,7	241,3	242,9	244,5	246,1	247,7	249,2	250,8	252,4
10	254,0	255,6	257,2	258,8	260,4	261,9	263,5	265,1	266,7	268,3	269,9	271,5	273,1	274,6	276,2	277,8
11	279,4	281,0	282,6	284,2	285,8	287,3	288,9	290,5	292,1	293,7	295,3	296,9	298,5	300,0	301,6	303,2
12	304,8	306,4	308,0	309,6	311,2	312,7	314,3	315,9	317,5	319,1	320,7	322,3	323,9	325,4	327,0	328,6
13	330,2	331,8	333,4	335,0	336,6	338,1	339,7	341,3	342,9	344,5	346,1	347,7	349,3	350,8	352,4	354,0
14	355,6	357,2	358,8	360,4	362,0	363,5	365,1	366,7	368,3	369,9	371,5	373,1	374,7	376,2	377,8	379,4
15	381,0	382,6	384,2	385,8	387,4	388,9	390,5	392,1	393,7	395,3	396,9	398,5	400,1	401,6	403,2	404,8
16	406,4	408,0	409,6	411,2	412,8	414,3	415,9	417,5	419,1	420,7	422,3	423,9	425,5	427,0	428,6	430,2
17	431,8	433,4	435,0	436,6	438,2	439,7	441,3	442,9	444,5	446,1	447,7	449,3	450,9	452,4	454,0	455,6
18	457,2	458,8	460,4	462,0	463,6	465,1	466,7	468,3	469,9	471,5	473,1	474,7	476,3	477,8	479,4	481,0
19	482,6	484,2	485,8	487,4	489,0	490,5	492,1	493,7	495,3	496,9	498,5	500,1	501,7	503,2	504,8	506,4
20	508,0	509,6	511,2	512,8	514,4	515,9	517,5	519,1	520,7	522,3	523,9	525,5	527,1	528,6	530,2	531,8
21	533,4	535,0	536,6	538,2	539,8	541,3	542,9	544,5	546,1	547,7	549,3	550,9	552,5	554,0	555,6	557,2
22	558,8	560,4	562,0	563.6	565,2	566,7	568,3	569,9	571,5	573,1	574,7	576,3	577,9	579,4	581,0	582,6
23	584,2	585,8	587,4	589,0	590,6	592,1	593,7	595,3	596,9	598,5	600,1	601,7	603,3	604,8	606,4	608,0

Fraction to decimal

1/64 = .015625	11/64 = .171875	21/64 = .328125	31/64 = .484375	41/64 = .640625	51/64 = .796875	61/64 = .953125
1/32 = .03125	3/16 = .1875	11/32 = .34375	1/2 = .5	21/32 = .65625	13/16 = .8125	31/32 = .96875
3/64 = .046875	13/64 = .203125	23/64 = .359375	33/64 = .515625	43/64 = .671875	53/64 = .828125	63/64 = .984375
1/16 = .0625	7/32 = .21875	3/8 = .375	17/32 = .53125	11/16 = . 6875	27/32 = .84375	1 = 1.
5/64 = .078125	15/64 = .234375	25/64 = .390625	35/64 = .546875	45/64 = .703125	55/64 = .859375	
3/32 = .09375	1/4 = .25	13/32 = .40625	9/16 = .5625	23/32 = .71875	7/8 = .875	
7/64 = .109375	17/64 = .265625	27/64 = .421875	37/64 = .578125	47/64 = .734375	57/64 = .890625	
1/8 = .125	9/32 = .28125	7/16 = .4375	19/32 = .59375	3/4 = .75	29/32 = .90625	
9/64 = .140625	19/64 = .296875	29/64 = .453125	39/64 = .609375	49/64 = .765625	59/64 = .921875	
5/32 = .15625	5/16 = .3125	15/32 = .46875	5/8 = .625	25/32 = .78125	15/16 = .9375	

Electromagnetic radiation